

ISSN 0016-6758

Том 61, Номер 1

Январь 2025



# ГЕНЕТИКА



НАУКА

— 1727 —

# СОДЕРЖАНИЕ

---

## Том 61, номер 1, 2025

---

### Обзорные и теоретические статьи

Роль длинных некодирующих РНК в растениях

*А. Ю. Пронозин, Д. А. Афонников*

3

Взаимосвязь микроРНК с транспозонами в развитии остеоартрита

*Р. Н. Мустафин, Э. К. Хуснутдинова*

24

---

### Общая генетика

Механизмы модулирующего действия тимохинона (компоненты черного тмина, *Nigella sativa*), влияющие на активность некоторых ядерных и митохондриальных генов в тканях мышей после воздействия рентгеновского излучения

*С. А. Абдуллаев, Д. В. Фомина, Н. Ф. Раева, М. А. Попов, Т. Н. Максимова, Г. Д. Засухина*

38

---

### Генетика растений

Генетический полиморфизм лиственницы сибирской (*Larix sibirica* Ledeb.) в контрастных экотопах Республики Хакасия

*Н. В. Орешкова, А. В. Пименов, Т. С. Седельникова, С. П. Ефремов*

45

---

### Генетика животных

Генетическая оценка голштинского скота по микросателлитным маркерам ДНК

*Л. А. Калашникова, Т. Б. Ганченкова, Н. В. Рыжова, Я. А. Хабибрахманова,*

*И. Е. Багаль, И. Ю. Павлова, А. Е. Калашников*

55

Полиморфизм гена *NIPAL1* rs135173498 как маркер эмбриональной летальности у аулиекольского скота

*Е. В. Белая, Е. А. Климанова, В. М. Норкина, И. С. Бейшова*

67

Идентификация *Bos taurus* и *Bos grunniens* на основании SNP

*В. Н. Кипень, Ж. Т. Исакова, М. М. Патрин, К. Б. Чекиров, К. А. Айтбаев,*

*А. Р. Карыпова, М. И. Ирсалиев*

74

Структура STR-аллелофонда популяции быков холмогорской породы банка криоконсервированного семени, сохраненного в Республике Коми

*В. С. Матюков, А. С. Лейченко, Я. А. Жариков, С. В. Николаев*

82

---

### **Генетика человека**

Alu-полиморфизм генов-регуляторов аутофагии и апоптоза  
как фактор продолжительности жизни человека

*В. В. Эрдман, Д. Д. Каримов, И. А. Туктарова, А. А. Петинцева,  
Я. Р. Тимашева, Т. Р. Насибуллин*

91

---

### **Краткие сообщения**

Ассоциация VNTR-полиморфизма в гене *MIR137* с когнитивными функциями  
у больных шизофренией и здоровых лиц

*М. В. Алфимова, Г. И. Коровайцева, В. В. Плакунова, В. Е. Голимбет*

103

## РОЛЬ ДЛИННЫХ НЕКОДИРУЮЩИХ РНК В РАСТЕНИЯХ

© 2025 г. А. Ю. Пронозин<sup>1, 2, \*</sup>, Д. А. Афонников<sup>1, 2, 3</sup>

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

<sup>2</sup>Курчатовский геномный центр, Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, 630090 Россия

<sup>3</sup>Новосибирский национальный исследовательский государственный университет, Новосибирск, 630090 Россия

\*e-mail: pronozinartem95@gmail.com

Поступила в редакцию 19.04.2024 г.

После доработки 29.07.2024 г.

Принята к публикации 15.08.2024 г.

Длинные некодирующие РНК (днРНК) представляют собой класс линейных или кольцевых молекул РНК длиной более 200 нуклеотидов без открытых рамок считывания. Экспериментальные исследования показали участие днРНК в регуляции устойчивости к холодовому, солевому, тепловому стрессу, в развитии плодов, корня и листьев. Однако экспериментальные методы являются трудоемкими и дорогостоящими подходами и пока еще не могут быть использованы для массового изучения днРНК в масштабах всего генома. Для этих целей применяют биоинформационные подходы, которые направлены на широкомасштабное распознавание последовательностей днРНК в геномах и транскриптомах. Несмотря на растущее число исследований, посвященных структурно-функциональному анализу днРНК, данный тип молекул по-прежнему остается малоизученным. Это связано со множеством факторов, которые нужно учитывать при идентификации днРНК. Применение пан-геномов и пан-транскриптомов позволит повысить эффективность исследования и общее количество предсказанных днРНК по сравнению с использованием генома одного представителя вида. Настоящий обзор посвящен описанию молекулярных и биологических функций днРНК, экспериментальных и биоинформационных методов идентификации, а также закономерностям эволюции, выявлению и анализу днРНК в масштабах пан-геномов и пан-транскриптомов.

**Ключевые слова:** длинные некодирующие РНК, микроРНК, регуляция транскрипции, пан-геном.

**DOI:** 10.31857/S0016675825010016 **EDN:** VFMYSM

Некодирующие РНК (нкРНК) выполняют в геномах растений ряд важнейших функций, связанных с регуляцией экспрессии генов, гомеостазом физиологических параметров растений. Знание функциональных особенностей нкРНК позволяет целенаправленно улучшать различные свойства растений [1–3], что обусловливает интерес к изучению этих молекул. Некодирующие РНК можно разделить на две группы: структурные некодирующие (рибосомные РНК, транспортные РНК, малые ядерные РНК, малые ядрышковые РНК) и регуляторные некодирующие (малые некодирующие РНК и длинные некодирующие РНК) [4]. Наименее изученными из всех перечисленных классов молекул являются длинные нкРНК [5, 6].

Длинные некодирующие РНК (днРНК) являются важным классом некодирующих РНК в живых организмах и представляют собой линейные или кольцевые молекулы РНК, содержащие более

200 нуклеотидов без открытых рамок считывания [4, 6]. Функции и структура днРНК слабо изучены в сравнении с белок-кодирующими генами и генами малых некодирующих РНК. Однако участие днРНК выявлено в регуляции экспрессии генов [7], формировании структуры макромолекулярных комплексов [8], во взаимодействии с белками [9], в патогенезе [10]. Наиболее активно днРНК изучаются экспериментально и биоинформационически у человека и других модельных видов (мыши, крысы, дрозофилы и т. д.). Эти исследования позволяют идентифицировать последовательности днРНК в геномах и транскриптомах, расшифровать важные особенности их структуры, регуляции экспрессии и функционирования, классифицировать их по структуре и функциям. На сегодняшний день эти методы позволили идентифицировать большое количество последовательностей днРНК у животных. Так, согласно базе данных Gencode

[11], в геноме мыши описано около 13000 генов днРНК. В геноме человека насчитывается более 18000 генов днРНК [12]. Изучение днРНК у животных позволило классифицировать их по молекулярному механизму функционирования (промоторные, энхансерные, теломерные, сигнальные и др.) [13–15], положению в геноме (межгенные, инtronные, антисмыловые и др.) [7, 16–18]. Эти исследования приводят к созданию специализированных баз данных, посвященных днРНК и содержащих последовательности десятков тысяч днРНК и их структурно-функциональную аннотацию [19, 20]. Вследствие отсутствия у первичной структуры днРНК четко выраженных функциональных сегментов (как, например, домены у белковых последовательностей) на сегодняшний день предсказание функции днРНК на основе первичной структуры является сложной задачей [15, 21, 22].

Следует отметить, что получение экспериментальной информации о структуре и функции днРНК является трудоемким процессом и требует проведения дорогостоящих экспериментов. Поэтому для изучения этого класса РНК активно используются методы биоинформатики. В частности, они направлены на широкомасштабное распознавание последовательностей днРНК в геномах и транскриптомах [23–25], предсказание их взаимодействий с белками [26], с мРНК и другими некодирующими РНК [27], на регуляцию экспрессии [7], эволюцию [28, 29] и предсказание их структуры и функций [30].

днРНК также известны и у растений. Интересно отметить, что, несмотря на ряд специфических особенностей, функциональные и структурные характеристики днРНК животных и растений имеют много общего [31]. Известно, что у растений днРНК принимают участие в регуляции устойчивости к холодовому, солевому, тепловому стрессу [1–3], оказывают влияние на устойчивость к гипоксии [32], принимают участие в развитии плодов, корня и листьев [17, 33]. Тем не менее до сих пор наши знания о днРНК растений являются менее полными, чем о днРНК животных. Исследования [34, 35] показывают, что последовательности днРНК в процессе эволюции быстро накапливают замены, так что даже с последовательностями из близкородственных видов организмов они имеют слабую гомологию. Применение концепции пан-генома, которая подразумевает охват последовательностей, подверженных структурной вариации и, возможно, отсутствующих в референсной последовательности каждого представителя вида, позволит повысить эффективность исследования и общее количество предсказанных днРНК [36–40].

Настоящий обзор посвящен днРНК растений и включает описание их молекулярных и биологических функций, экспериментальных и биоинформационных методов идентификации,

закономерностям эволюции, выявлению и анализу днРНК в масштабах пан-геномов и пан-транскриптомов.

## КЛАССИФИКАЦИЯ днРНК ПО ПОЛОЖЕНИЮ В ГЕНОМЕ

днРНК классифицируют по многим параметрам, однако наиболее часто используют классификацию по положению в геноме относительно известных белок-кодирующих генов. Согласно этой классификации, днРНК делятся на четыре крупных класса [35], рис. 1:

- 1) инtronные — перекрываются с инtronом гена;
- 2) антисмыловые — ориентированы против направления транскрипции гена, кодирующего белок;
- 3) межгенные — расположены между двумя локусами генов;
- 4) смысловые — ориентированы в направлении транскрипции гена, кодирующего белок.

Инициация транскрипции инtronных днРНК происходит в направлении транскрипции гена мишени, терминация происходит в области перекрытия с экзоном гена мишени [41].

Транскрипция антисмыловых днРНК начинается на 3'-конце гена, кодирующего белок. Эти днРНК разделяются на две группы в зависимости от цис- и транс-геометрической изомерии. Транскрипция цис-изомеров антисмыловой днРНК (цис-днРНК) происходит с противоположного конца локуса белок-кодирующего гена, с которым данный транскрипт перекрываеться. За счет этого днРНК имеет высокий уровень комплементарности с данным геном. Транскрипция транс-изомеров антисмыловой днРНК происходит с локуса соседнего белок-кодирующего гена, с которым перекрываеться данный транскрипт. Таким образом, подобные транскрипты могут перекрываться с несколькими генами, но при этом имеют низкий уровень комплементарности с ними [42].

Длина межгенных днРНК превышает 200 нуклеотидов, данные транскрипты не перекрываются с кодирующими белки генами. Особенности транскриптов данного класса: высокая консервативность, выраженная тканеспецифичность, высокая стабильность. В работе [43] биоинформационными методами было показано взаимодействие межгенных днРНК с мобильными элементами. Проведенное исследование предоставляет систематическую оценку вклада мобильных элементов в состав, эволюционное происхождение и регуляцию днРНК.

Смысловые днРНК транскрибируются в направлении транскрипции кодирующих белок генов. Это делает их выделение в отдельный класс днРНК или вообще идентификацию как днРНК

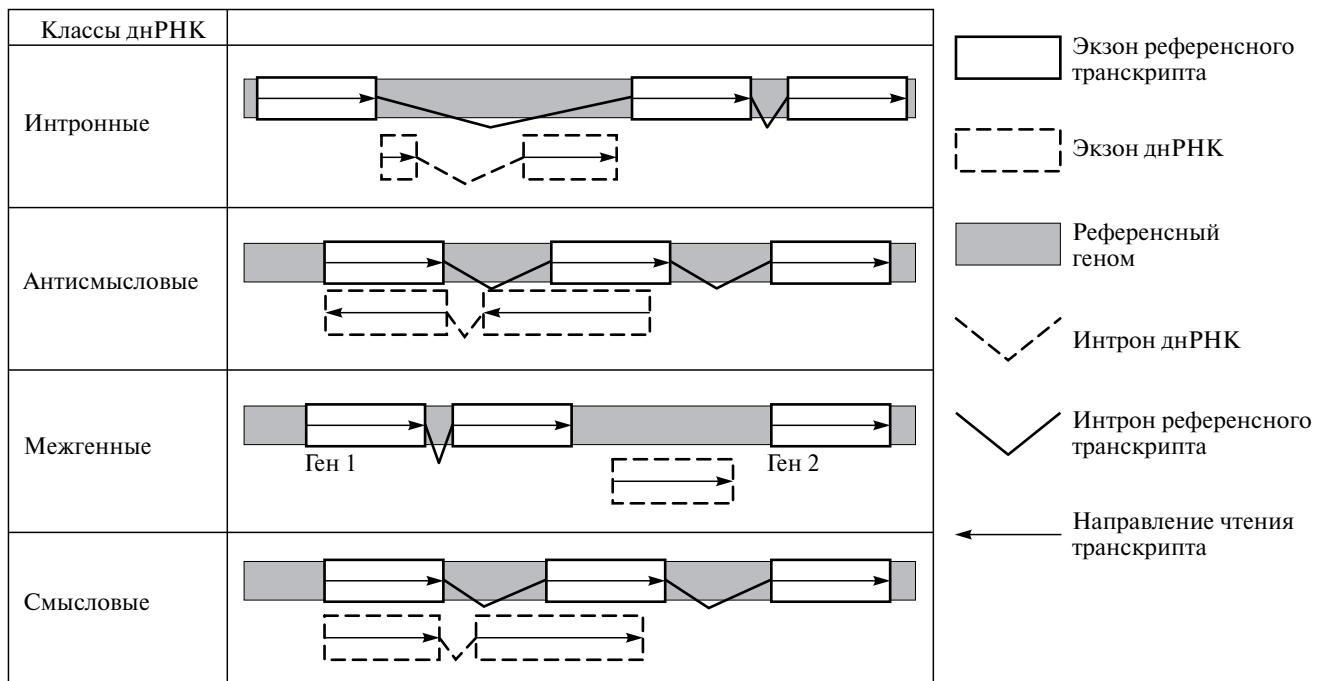


Рис. 1. Классификация днРНК на основе расположения и ориентации относительно ближайших или перекрывающихся генов, кодирующих белки.

достаточно противоречивым. Структурные характеристики таких РНК (длина, количество экзонов и инtronов) могут совпадать со структурой гена, кодирующего белок. Например, в работе [17] предполагается, что транскрипты данного класса днРНК могут полностью перекрываться с генами, кодирующими белок. Однако в работе [41] описываются примеры днРНК, возникших за счет альтернативного сплайсинга или укорачивания первого либо последнего экзона генов. Авторы предполагают, что эти гены могут представлять изоформы белков, кодируемых этими генами, а не днРНК. Например, в работе [34, 41] РНК, перекрывающиеся с экзоном гена, кодирующего белок, не считаются днРНК. Для растений найдено мало данных транскриптов, и ни для одного из них нет функциональной аннотации [41].

### МОЛЕКУЛЯРНЫЕ ФУНКЦИИ днРНК

днРНК имеют несколько общих характеристик с генами, кодирующими белки, что затрудняет их идентификацию. Так, большинство днРНК транскрибируется с помощью РНК-полимеразой II. Они имеют сходные эпигенетические профили, наличие сигналов сплайсинга и полигидрирования, а также сходные размеры экзонов и инtronов [44, 45]. Кроме того, некоторые днРНК имеют длину, превышающую 1000 нуклеотидов. Таким образом, они могут содержать открытые рамки считывания (ОРС). Размеры подобных ОРС могут варьироваться, некоторые из них могут

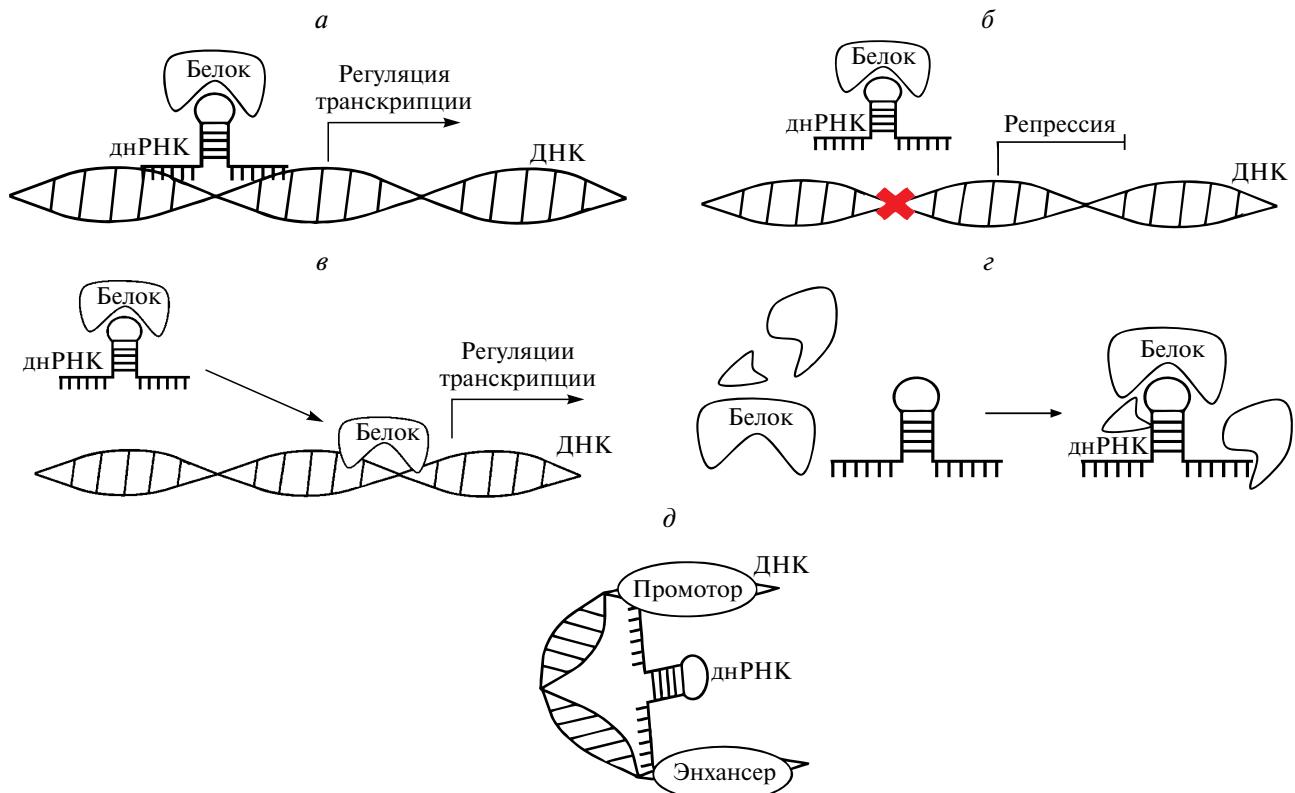
быть достаточно короткими. Например, в работе [46] упоминается, что днРНК в среднем могут содержать несколько транслируемых ОРС, имеющих длину 50 кодонов.

Отличительной особенностью днРНК является способность транскрибироваться РНК-полимеразой IV и V, в особенности у растений [47]. Большинство днРНК локализуются как в клеточном ядре, так и в цитоплазме [48]. Также днРНК демонстрируют низкий уровень экспрессии, при этом экспрессируются более специфично для определенных тканей [49]. Подобное наблюдается и при анализе консервативности структуры днРНК. Сравнение днРНК различных видов растений показало, что доля сходных последовательностей у разных видов мала. Однако днРНК одного вида демонстрируют достаточно высокий уровень консервативности первичной, вторичной и третичной структуры [15].

В работах [14, 15, 50] указывается, что на функционирование днРНК влияет и их внутриклеточная локализация. В работах [4, 15, 50, 51] выделяют пять основных функций днРНК, представленных на рис. 2.

### Сигнальные днРНК

днРНК чаще всего экспрессируются в определенных тканях. В некоторых работах отмечается, что днРНК экспрессируется специфически в ответ на различные стимулы [4, 52–54]. В связи с этим



**Рис. 2.** Молекулярные функции днРНК: *а* – сигнальные, днРНК получают сигнал для взаимодействия с ферментами, модифицирующими хроматин, для регуляции транскрипции; *б* – ловушки, связывают белки и препятствуют их присоединению к генам-мишеням; *в* – направляющие, направляют транскрипционные факторы к местам их функционирования, содействуют необходимому состоянию хроматина; *г* – каркасы, объединяют белки в комплексы и инициируют разные биологические процессы; *д* – энхансеры, могут транскрипционно активировать соседние гены.

можно предположить, что экспрессия некоторых днРНК находится под контролем транскрипционных факторов [55], осуществляется в определенный момент и в определенной ткани и может служить сигналом для запуска регуляторных процессов в клетке [55]. Такая специфическая экспрессия характерна для днРНК, которые называют сигнальными (рис. 2, *а*). Сигнальные днРНК могут выступать в качестве элементов регуляции. Известно, например, что сигнальные днРНК взаимодействуют с модифицирующими хроматин ферментами, такими как гистоновые метилтрансферазы, чтобы подавить экспрессию их генов-мишеней посредством образования гетерохроматина [56]. Такие взаимодействия являются частью регуляторных и сигнальных функций данного типа днРНК.

#### *днРНК-приманки или ловушки*

днРНК-приманки связывают белки и тем самым препятствуют их взаимодействию с регуляторными районами генов-мишеней [55]. Эти днРНК косвенным образом репрессируют транскрипцию, блокируя микроРНК, модифицирующие комплексы

(субъединицы хроматина), каталитические белки и факторы транскрипции [55, 57]. Также днРНК выступают в качестве ловушки для микроРНК, данный механизм взаимодействия подробно описан в подразделе “Механизмы взаимодействия днРНК и микроРНК”.

#### *Направляющие днРНК*

днРНК также выполняют функцию группировки/локализации факторов транскрипции в определенных геномных локусах. Такие днРНК называют направляющими, (рис. 2, *в*). днРНК направляют модификаторы хроматина и белковые комплексы или факторы транскрипции к определенным генам-мишеням. Они способствуют локализации регуляторных молекул в определенных локусах для осуществления транскрипционной регуляции [56]. днРНК имеют цис-, транс-геометрическую изомерию. Цис-изомеры направляют изменения хроматина, выступая в качестве комплементарных мишеней для малых РНК (таких как микроРНК, малые интерферирующие РНК) [55]. Транс-изомеры

направляют ДНК-мишень посредством РНК–ДНК взаимодействий [58].

#### днРНК-каркасы

днРНК-каркасы (скаффолды) играют важную роль при сборке макромолекулярных белковых комплексов и инициации различных биологических процессов [59], (рис. 2, *г*). Подобные днРНК способны соединять несколько белков в один рибонуклеиновый комплекс, который позволяет давлять или активировать транскрипцию в зависимости от наличия и природы вовлеченных днРНК и белков [55].

#### днРНК-энхансеры

днРНК, транскрипция которых осуществляется с энхансерных локусов, называются днРНК-энхансеры (рис. 2, *д*). Как и все днРНК, их последовательности считываются в одном направлении, имеют поли(A)-хвост и сплайсируются [60]. Их следует отличать от энхансерных РНК (эРНК), которые являются продуктом двунаправленной транскрипции, не сплайсированы и не имеют поли(A)-хвоста. Энхансерные днРНК на сегодняшний день довольно плохо изучены. Однако, согласно гипотезе [61], синтез энхансерных днРНК может служить способом поддержания открытого состояния хроматина в области энхансеров для облегчения их взаимодействия с промоторами посредством образования петель. Некоторые исследования показывают вовлеченность энхансерных днРНК в процесс активации экспрессии генов мишени [62, 63]. Также в нескольких работах показано понижение экспрессии генов мишени при нокдауне ближайшей к ним энхансерной днРНК [63, 64].

#### Механизмы взаимодействия днРНК и микроРНК

Особое внимание стоит уделить взаимодействию днРНК и микроРНК. В работе [65] указывается на участие данных взаимодействий в

регуляции устойчивости к болезням, холодовому, тепловому, биотическому и абиотическому стрессу у растений.

МикроРНК — это малые некодирующие РНК длиной 18–25 нуклеотидов, транскрибируемые из инtronов, некодирующих участков гена или межгенных спейсеров (участков ДНК между генами) [66]. На сегодняшний день выделяют четыре типа взаимодействия днРНК и микроРНК [67], (рис. 3). Более подробно разберем данные механизмы взаимодействия.

#### Деградация днРНК при встраивании микроРНК

МикроРНК функционируют в комплексе с белками семейства аргонавтов (AGO). Комплекс микроРНК–AGO встраивается в РНК-индуцируемый комплекс подавления экспрессии гена, либо блокируя синтез белка с мРНК, либо приводя к деградации мРНК [68]. Подобный механизм функционирования наблюдается и для днРНК. МикроРНК за счет РНК-связывающих белков встраивается в комплементарный участок днРНК, вызывая деградацию данных молекул [69, 70]. В этом случае днРНК выступает как конкурент мРНК и тем самым адсорбирует часть молекул микроРНК, снижая их воздействие на молекулы потенциальных мишени (мРНК). Важно, что при этом происходит деградация и самой днРНК (рис. 3, *а*).

#### днРНК-мишени для микроРНК

Впервые днРНК-мишени с эффектом мимикии, или конкурирующие эндогенные днРНК (competitive endogenous RNAs, ceRNAs), были обнаружены в геномах растений [71], (рис. 3, *б*). Механизм мимикии (“эффект губки”) заключается в блокировании взаимодействия микроРНК с мРНК-мишенью за счет связывания данной микроРНК с днРНК-приманкой через частично комплементарные последовательности [72].

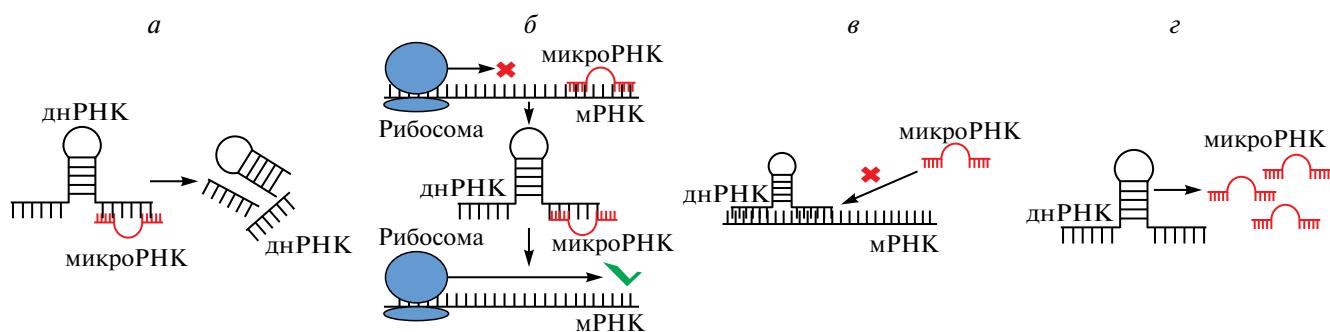


Рис. 3. Механизмы взаимодействия днРНК – микроРНК: *а* – деградация днРНК при встраивании микроРНК; *б* – днРНК выступает в качестве мишени (губки) для микроРНК; *в* – днРНК конкурирует с микроРНК для регуляции уровня экспрессии мРНК; *г* – днРНК выступает предшественником микроРНК (сплайсируется в микроРНК).

На сегодняшний день экспериментально подтверждено влияние эндогенных днРНК на регуляцию синтеза фосфатов в *Arabidopsis thaliana* [72], *Zea mays* [73] и *Medicago truncatula* [74] за счет связывания с микроРНК.

### Конкуренция днРНК и микроРНК

Механизмы взаимодействия микроРНК с эндогенными и конкурирующими днРНК схожи. днРНК выполняет роль мишени и снижает концентрацию микроРНК, (рис. 3, в). Однако днРНК может конкурировать и с микроРНК. Этот механизм функционирования днРНК возникает в случае, когда за счет комплементарных оснований они связываются с мРНК, не давая возможности в этих районах мРНК взаимодействовать с микроРНК. Подобный механизм позволяет предотвратить деградацию мРНК, с которой связываются днРНК [75–77].

### днРНК-предшественники микроРНК

Некоторые днРНК как первичные транскрипты могут служить предшественниками специфических микроРНК, (рис. 3, г). Малые РНК могут производиться за счет механизма сплайсинга и приводить к усилению посттранскрипционной регуляции соответствующих мРНК-мишеней [35]. Такой механизм функционирования днРНК был экспериментально подтвержден у человека [78]. Показано, что они отвечают за регуляцию развития и прогрессирования опухоли за счет контроля транскрипции микроРНК-супрессоров [79, 80]. Для растений днРНК-предшественники обнаружены биоинформационическими методами. В работах [81, 82] большинство выявленных микроРНК выравниваются на концевые участки последовательности днРНК, указывая на то, что эти днРНК являются их предшественниками.

## ЭКСПРЕССИЯ днРНК

На сегодняшний день появляется все больше работ по анализу днРНК у различных организмов, в которых применяются методы исследования экспрессии для данных молекул. Общепринятым считается, что днРНК демонстрируют более низкий уровень экспрессии по сравнению с мРНК, при этом экспрессируются более специфично для определенных тканей [34]. Однако в последнее время начинают появляться работы, демонстрирующие днРНК с уровнями экспрессии, равными таковым для генов, кодирующих белки [16, 83]. Подобные случаи являются единичными, и их результаты требуют экспериментального подтверждения.

Выделяют три направления исследования экспрессии днРНК: анализ экспрессии днРНК в тканях и в ответ на определенный раздражитель;

анализ дифференциальной экспрессии и ко-экспрессии с генами, кодирующими белки; анализ тканеспецифичности днРНК. Данные подходы применяются как для масштабного транскриптомного анализа, так и для анализа небольших выборок последовательностей.

Анализ экспрессии днРНК в тканях является наиболее распространенным направлением исследования. Большинство работ начинают свое исследование с данного направления, постепенно углубляясь в дифференциальную или ко-экспрессию.

Анализ тканеспецифичности днРНК является наименее исследованным, так как нет точного понимания, какой метод определения специфичности экспрессии гена в тканях наиболее точный [84]. Во многих работах специфичным считается ген, имеющий экспрессию только в одной ткани [16, 85, 86]. Одним из первых подходов является подсчет количества тканей, в которых данный ген экспрессируется [87]. Однако данный метод позволяет исследовать только гены с высокой экспрессией, что затрудняет выбор порога фильтрации последовательностей. Также существует проблема близкородственных тканей, таких как части стебля или мозга человека.

С развитием понимания о структуре и функциях днРНК стало появляться все больше подходов к решению данной задачи. Существующие методы делятся на две группы. В первой группе показателем специфичности является число, определяющее, насколько специфично экспрессируется исследуемый ген: 1 – если ген имеет специфическую экспрессию, 0 – если экспрессируется в нескольких тканях – Tau [88], Gini [89], tissue specificity index (TSI) [90]. Во второй группе специфичность экспрессии гена рассчитывается отдельно для каждой ткани – z-score, SPM [91], expression enrichment (EE) [92], при этом уровень специфичности варьируется между разными метриками.

В основном все метрики расчета специфичности экспрессии гена (Tau, EE score, TSI, Gini coefficient) похожи: для расчета используется уровень экспрессии гена в определенной ткани и количество экспериментов, которым данная ткань принадлежит. Например, формула для расчета Tau:

$$\tau = \frac{\sum_{i=1}^n (1 - \hat{x}_i)}{n-1}; \quad (1)$$

$$\hat{x}_i = \frac{x_i}{\max_{1 \leq i \leq n} (x_i)}, \quad (2)$$

где  $x_i$  – экспрессия гена в ткани;  $n$  – количество экспериментов с тканью.

Встречаются работы, использующие, помимо перечисленных выше, метрики, основанные на энтропии Шеннона, Jensen-Shannon (JS) [93]:

$$Hg = \sum_{1 \leq t \leq N} -p_{t|g} \log_2 (p_{t|g}); \quad (3)$$

$$p_{t|g} = \frac{\omega_{g,t}}{\sum_{1 \leq t \leq N} \omega_{g,t}}, \quad (4)$$

где  $p_{t|g}$  – распределение экспрессии гена в ткани и  $\omega_{g,t}$  – экспрессия гена. Таким образом,  $p_{t|g}$  показывает отношение экспрессии гена в определенной ткани на сумму экспрессий этого гена во всех тканях.

$$Q_{g|t} = Hg - \log_2 (p_{t|g}), \quad (5)$$

где  $Hg$  – энтропия, показывает, насколько ген тканеспецифичный. Если значение энтропии близко или равно 0, то данный ген специфичен для определенной ткани.  $Q_{g|t}$  показывает энтропию экспрессии исследуемого гена в одной ткани. Чем выше значение энтропии, тем выше специфичность экспрессии гена в определенной ткани.

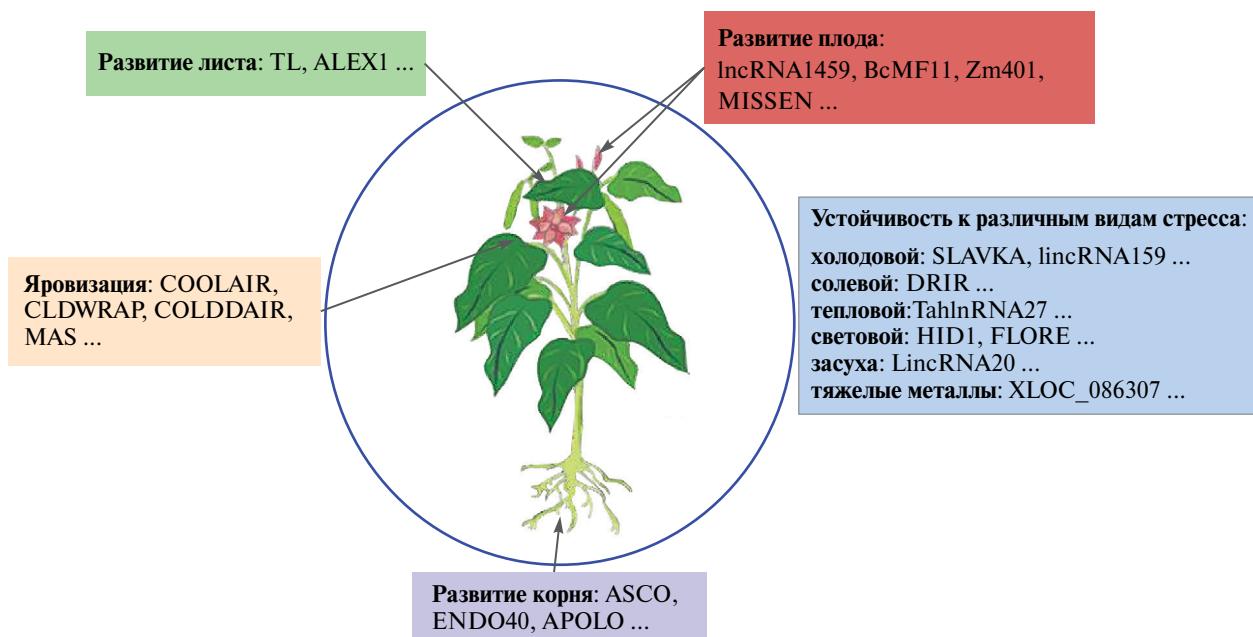
### БИОЛОГИЧЕСКАЯ РОЛЬ днРНК У РАСТЕНИЙ

Предыдущие разделы показали разнообразие и важность молекулярных функций днРНК у

растений. днРНК могут выступать как генетические и эпигенетические регуляторы экспрессии генов. В зависимости от геометрической изомерии они могут функционировать как цис-регуляторные элементы, взаимодействуя с близко расположеными генами и воздействуя непосредственно на гены в одной цепи, или транс-регуляторные элементы, действуя вдали от места синтеза. днРНК могут препятствовать связыванию транскрипционных факторов с промоторной областью, выступая в качестве репрессоров транскрипции. В результате этих молекулярных процессов контролируются функции растений на более высоком уровне организации: тканей, органов и всего организма. Примеры процессов, которые затрагивает регуляция при помощи днРНК у растений, приведены на рис. 4. К ним относятся: устойчивость растений к различным видам стресса, развитие органов растений, яровизация, развитие плода.

#### днРНК в процессе яровизации

Цветение является одним из важнейших адаптивных признаков, обеспечивающих переход репродуктивного роста и развития, происходящий в благоприятных условиях в течение жизненного цикла растения. Яровизация является важным механизмом, контролирующим цветение у некоторых видов растений, которые растут в вегетативном состоянии в холодные зимние сезоны и начинают цвети весной [94]. Также яровизация



**Рис. 4.** Биологическая роль днРНК у растений. днРНК принимают участие в регуляции генов, отвечающих за устойчивость растений к различным видам стресса (синий прямоугольник), также влияют на плодородие растений (красный прямоугольник), участвуют в развитии листа, корня (зеленый, фиолетовый прямоугольник) и в процессах яровизации (оранжевый прямоугольник).

является наиболее изученным регуляторным процессом у растений, в котором участвуют днРНК. Они также принимают участие в регуляции гена FLOWERING LOCUS C (*FLC*), отвечающего за устойчивость *A. thaliana* к холодовому стрессу [3]. На сегодняшний день в локусе *FLC* обнаружено два класса днРНК: COOLAIR – днРНК транскрибируется в антисмысловом направлении по отношению к промотору гена *FLC*; COLDAIR – днРНК транскрибируется с первого интрана гена *FLC* в смысловом направлении. днРНК COOLAIR связываются с локусом *FLC* в цис-положении и образуют R-петлю между промотором и первым интраном гена *FLC*, что приводит к подавлению экспрессии данного гена мишени. Данный процесс проходит как на автономном, так и в яровизированном пути [94].

#### днРНК и влияние на урожайность растений

Развитие нормального пыльцевого зерна имеет решающее значение для оплодотворения и развития семян. Пыльник действует как родительская ткань пыльцевого зерна и регулирует его развитие. Аномальное пыльцевое зерно может привести к развитию растений с мужской стерильностью [95]. Экспериментальные исследования [96, 97] показали вовлеченность днРНК в процесс контроля урожайности. Например, днРНК (LDMAR) риса (*Oryza sativa japonica*) вызывает запрограммированную гибель клеток пыльников в условиях длинного дня (LD, long day) за счет сплайсинга в микроРНК osa-smR5864. Однонуклеотидная мутация микроРНК osa-smR5864 приводит к мужской стерильности растения [96]. Снижение экспрессии днРНК (BcMF11) в *Brassica campestris* препятствует деградации тапетума и приводит к прерыванию образования пыльцевых зерен [97]. Также методами биоинформатики в работе [33] для пяти стадий развития пыльника *Brassica rapa* было выявлено 14 днРНК, ко-экспрессирующихся с 10 известными генами, кодирующими белки, которые играют решающую роль в развитии пыльцы. Как утверждают авторы, выявленные днРНК представляют интерес для дальнейших экспериментальных исследований вовлеченности днРНК в репродуктивные процессы растений.

#### днРНК в процессе фотоморфогенеза

Молекулярный сигнальный механизм фотоморфогенеза широко изучался у различных видов растений. Во время этого явления ряд белков ассоциируются в качестве первичных и вторичных сигнальных молекул. Среди них выделяют семейство транскрипционных факторов со структурой ДНК-связывающего домена спираль–петля–спираль [helix–loop–helix (bHLH)]. Среди членов семейства выделяют факторы взаимодействия с

фитохромом (PIF), которые выполняют функцию подавления фотоморфогенеза проростков в темное время суток [98]. В работе [77] обнаружена днРНК (HID1), которая функционирует как транс-регуляторный элемент, взаимодействуя с промотором гена *PIF3* (ключевая сигнальная молекула фотоморфогенеза, которая негативно регулирует реакции на красный свет), и ингибитирует его транскрипцию. В результате наблюдается удлинение зародышевого стебля (гипокотиля) у саженцев *A. thaliana*.

#### днРНК в процессах устойчивости к различным видам стресса

В природе растения подвергаются действию различных стрессовых факторов, таких как соль, засуха, холод, тепло, а также различные вирусные инфекции. Это приводит к проблемам в росте и плодородии у растений. Чтобы адаптироваться и выжить в таких неблагоприятных условиях, растения используют множество механизмов регуляции генов для восстановления клеточного гомеостаза. Современные исследования демонстрируют немаловажную роль днРНК в регуляции экспрессии генов в ответ на стрессовые состояния [99].

Вовлеченность днРНК в процессы устойчивости к абиотическим видам стресса демонстрируется у *A. thaliana* [81], *B. rapa* [100], винограда (Cabernet Sauvignon) [101], *O. sativa* [102] и многих других растений. Например, у днРНК *A. thaliana* выделяют следующие механизмы регуляции устойчивости к холодовому стрессу: модуляция / реконструкция хроматина, например днРНК COOLAIR, описанная выше; альтернативный сплайсинг днРНК TAS1 приводит к образованию варианта транскрипта с нерасщепленным интраном, более устойчивого к холодовому стрессу, чем сплайсированный вариант данной днРНК [103]; транскрипционная регуляция генов, например днРНК SVALKA подавляет транскрипцию гена *CBF1* (ответственного за запуск каскада реакций, направленных на увеличение устойчивости растений к низким температурам) за счет транскрипционной коллизии [104].

днРНК IncRNA973 у *A. thaliana* регулирует устойчивость растений к солевому стрессу путем модуляции экспрессии генов, ответственных за реактивные формы кислорода (SOD, CAT и POD), транскрипционные факторы (MYB5, WRKY46, NAC29 и ERF62) и ионный баланс (NHX7) [105].

Сверхэкспрессия днРНК At5NC056820 у *A. thaliana* повышает устойчивость растения к засухе [106]. днРНК TCONS\_00021861 у *O. sativa* за счет эффекта губки взаимодействует с miR528-3p и предотвращает ее взаимодействие с геном-мишенью YUCCA-7 (который участвует в биосинтезе триптофан-зависимых ауксинов и повышает устойчивость к засухе), что активирует синтез

индолил-3-уксусной кислоты (ИУК), повышающей устойчивость к засухе [107].

Подобный механизм наблюдается для днРНК TCONS\_00048391, которая является “губкой” для микроРНК *bra-miR164a*, подавляющей экспрессию гена-мишени NAC1 (ответственного за устойчивость к тепловому стрессу у *B. rapa*) [100].

Таким образом, даже внутри одного организма днРНК имеют различные механизмы регуляции устойчивости к абиотическим видам стресса.

### ЭВОЛЮЦИЯ днРНК

Эволюционные процессы приводят к изменчивости структуры и функций днРНК. В области изучения эволюции днРНК существует несколько направлений активных исследований. Прежде всего это вопросы, связанные с возникновением новых днРНК в геномах организмов. Одним из таких механизмов является дупликация генов. Появление в геноме дополнительной копии гена открывает свободу для эволюции одной из копий гена. Мутации, возникающие в одной из двух копий и ослабляющие исходную функцию гена, не будут отсеиваться отбором, потому что вторая копия может сохранить прежнюю функциональность. Таким образом, за счет дупликации возникают “псевдогены” (неактивные участки), изоформы и транскрипты днРНК [28].

Существует несколько гипотез возникновения днРНК за счет дупликации. Первая основывается на потере одной из изоформ гена, кодирующего белок (в процессе дупликации), способности кодировать белок. Это объясняется наличием у днРНК регуляторных элементов, сайтов сплайсинга, последовательностей экзонов и последовательностей, имеющих поли(А)-хвост. Наиболее известный пример днРНК, возникшей за счет дупликации гена, кодирующего белок, – человеческая днРНК *Xist* [29]. Вторая гипотеза основана на дупликации некодирующих РНК. Однако подобные случаи встречаются не так часто, как дупликация генов, кодирующих белки, вследствие высокой эволюционной лабильности днРНК [108].

Также одна из гипотез возникновения днРНК связана со встраиванием мобильных элементов (МЭ) в последовательность транскрипта. Мобильные элементы могут способствовать регуляции экспрессии генов при развитии и адаптации, а также служить основными источниками генетических вариаций в эволюции геномов. МЭ присутствуют в экзонах, сайтах начала транскрипции, сайтах полиаденилирования (поли(А)-хвосты) или в некоторых комбинациях этих частей. Исследование, проведенное на восьми растениях (включаящих *A. thaliana*, кукурузу, томат, тополь) [109], показало, что 59.7% днРНК, найденных во всех восьми организмах, имеют по крайней мере один

экзон с частично встроенным МЭ. В исследовании [43] показано участие МЭ не только в появлении днРНК, но и в регуляции и вариабельности днРНК.

Для оценки консервативности днРНК между геномами различных организмов сравнивают их последовательности [110]. Например, при выравнивании 67 тыс. днРНК риса на 39 тыс. днРНК кукурузы доля совпадающих позиций гомологов днРНК равна 20% [111]. При сравнении днРНК для эволюционно более далеких видов, например *A. thaliana* и *O. sativa*, уровень сходства значительно падает вплоть до отсутствия гомологии [109]. Только небольшое количество последовательностей днРНК демонстрирует уровень консервативности, сравнимый с консервативностью генов, кодирующих белки [112]. Одной из таких днРНК является IPS, участвующая в гомеостазе фосфора (Р<sub>и</sub>) у *A. thaliana*, *O. sativa*, *S. lycopersicum* и *S. moellendorffii*. IPS содержит консервативный мотив из 24 нуклеотидов в перечисленных видах, который служит мишенью для miR399. IPS первоначально появилась у *S. moellendorffii*, представителя древних споровых высших растений. За время эволюционного развития функция IPS не изменилась, что может указывать на роль данной днРНК в адаптации к суше в процессе миграции из водной среды.

Однако большинство днРНК показывают высокую вариабельность первичной структуры, что подтверждается в работе [113], посвященной анализу консервативности днРНК между 25 цветковыми растениями. Анализ показал, что только 6.79% днРНК консервативны у всех исследуемых растений. В работе [114] был проведен анализ консервативности днРНК (межгенных) риса относительно семи видов растений. В результате была показана низкая консервативность днРНК в сравнении с мРНК, однако более высокая в сравнении с мРНК, содержащими МЭ. Анализ консервативности структурных элементов днРНК показал высокую консервативность экзонов днРНК в сравнении с инtronами. В данных исследованиях показана быстрая эволюция днРНК, которая приводит к низкой консервативности последовательностей. Также отмечается, что процент консервативных днРНК выше среди родственных видов (в пределах одного рода), при этом консервативность резко убывает по мере увеличения эволюционного расстояния [113].

Несмотря на существующие теории возникновения и эволюции днРНК, к ним следует относиться с соответствующим скептицизмом вследствие отсутствия экспериментального подтверждения.

## днРНК В ПАН-ГЕНОМАХ И ПАН-ТРАНСКРИПТОМАХ

В природе существует высокое геномное и морфологическое разнообразие у видов, вызванное структурными вариациями. Они включают в себя инсерцию, дупликацию, инверсию, транслокацию, а также вариации копийности генов и их наличия или отсутствия. Данные вариации способны влиять на сегменты ДНК длиной в несколько мегабаз и в некоторых случаях изменять количество хромосом. Вследствие этого один референсный геном не способен охватить все содержимое гена одного вида. Например, в референсных геномах растений отсутствует ряд важных для сельскохозяйственных исследований генов. Для решения данной проблемы используется пан-геном.

Концепция пан-генома подразумевает охват последовательностей, подверженных структурной вариации и, возможно, отсутствующих в референсной последовательности каждого представителя вида. На сегодняшний день существует несколько формулировок термина “пан-геном”, которые представлены в работах [36–40].

В данном обзоре под пан-геномом понимается полный неизбыточный набор генов, представленных в геноме хотя бы одной особи, относящейся к данной таксономической единице [115]. Отметим, что это понятие можно распространить на таксономические единицы любого порядка [116]. Так, была проведена попытка создать общий пан-геном бактерий на основе 573 секвенированных ранее геномов бактерий [117]. Однако в настоящее время понятие пан-геном чаще всего употребляют именно в отношении вида.

Гены “пан-генома” могут быть разделены на две группы.

1. Консервативные гены – присутствуют в геномах всех представителей таксономической единицы.

2. Вариабельные гены – присутствуют в геномах не всех представителей вида. В свою очередь делятся на две категории: гены, представленные в двух и более геномах, и уникальные гены, представленные только в одном индивидуальном геноме.

При увеличении количества вариабельных генов, включенных в какой-либо конкретный пан-геном, общее количество генов в этом пан-геноме будет увеличиваться, в то время как количество консервативных генов будет постепенно уменьшаться. Причем каждая из этих выборок рано или поздно выйдет на плато, после которого она не будет значимо изменяться с добавлением новых геномов. При этом отмечается, что вариабельные гены в среднем эволюционно более молодые, чем консервативные [36]. Этот результат кажется логичным в свете того, что большую часть консервативных составляют гены домашнего

хозяйства. Функции консервативных и вариабельных генов различны. Функции консервативных связаны с важнейшими клеточными процессами, а вариабельных – с окружающей средой и защитными реакциями, рецепторной и антиоксидантной активностью, регуляцией генов и передачей сигнала.

При помощи консервативных и вариабельных генов можно подразделять пан-геномы на две категории – открытые и закрытые. Закрытыми принято считать пан-геномы, которые быстро выходят на плато по количеству генов, т. е. секвенирование геномов новых представителей исследуемого вида не привносит новых генов. Открытые, наоборот, увеличиваются в размере каждый раз при добавлении новых геномов в их состав.

Еще одним из методов анализа генного состава у нескольких представителей какого-либо таксона является анализ его транскриптомов. Нуклеотидные последовательности транскриптов позволяют оценить присутствие генов в геноме только в том случае, если они экспрессируются в какой-либо ткани или органе растения. Таким образом, определение пан-транскриптома можно сформулировать так: пан-транскриптом отражает совокупность всех молекул РНК, присутствующих в данном виде или особи, и состоит из консервативных и вариабельных транскриптов (уникальных для одного вида или особи).

Помимо одноклостидных полиморфизмов (SNP) и множественных структурных вариаций немаловажную роль играют днРНК. Например, полное секвенирование транскриптома показало, что транскрибируется чрезвычайно большая часть генов высших эукариот. Однако более 90% всех транскриптов не транслируются в белки [4, 44, 78, 118] и являются некодирующими последовательностями. Как было описано выше, днРНК выполняют важную роль во многих молекулярных и биологических процессах у растений. Однако, несмотря на растущее число посвященных структурно-функциональному анализу днРНК исследований, данный тип молекул по-прежнему остается малоизученным. Это связано со множеством факторов, которые нужно учитывать при идентификации днРНК. Размер днРНК схож с размером некоторых генов, кодирующих белки, что приводит к ошибкам перепредсказания. Также исследования показывают, что последовательности днРНК перерпевают быструю эволюцию, закономерности которой пока непонятны. Как правило, в процессе эволюции эти последовательности очень быстро накапливают замены, так что даже с последовательностями из близкородственных видов организмов они имеют слабую гомологию [113]. Таким образом, применение пан-геномов и пан-транскриптомов повысит эффективность исследования и общее количество предсказанных днРНК

по сравнению с использованием генома одного представителя вида. Также применение пан-генома позволит повысить точность и полноту исследуемого набора данных, так как пан-геном является совокупностью всех кодирующих и некодирующих последовательностей всех представителей вида. Более того, предсказанные для разных представителей одного вида днРНК могут быть выровнены на один пан-геном данного вида, это позволит произвести дальнейшее исследование вариабельных генов, поможет определить точное количество, тип и функции днРНК, внесенных каждым представителем исследуемого вида. Данный подход также позволит избежать необходимости выравнивать, сопоставлять и объединять результаты для геномов разных представителей одного вида, что существенно сократит время, трудозатраты и вычислительные ресурсы, необходимые для данного исследования. Например, исследование пан-транскриптома люцерны (*Medicago sativa* L.) [119] в тканях листьев, стебля и корней выявило 11677 новых днРНК. Данное исследование позволило выявить 2105 высоко экспрессионных и 1894 низко экспрессионных днРНК, а анализ генов онтологии показал вовлеченность данных днРНК в процессы регуляции устойчивости люцерны к засухе и солевому стрессу. В работе [120] при анализе пан-транскриптома кукурузы (*Zea mays*) было выявлено 644 новых днРНК. Однако на сегодняшний день работы, посвященные исследованию пан-геномов и пан-транскриптомов, в основном направлены на выявление и исследование новых генов, кодирующих белки. Тогда как работ, посвященных исследованию днРНК в масштабах пан-геномов и пан-транскриптомов, не так много, в особенности для растений.

### ЭКСПЕРИМЕНТАЛЬНЫЕ МЕТОДЫ ИДЕНТИФИКАЦИИ днРНК

До недавнего времени идентификация днРНК считалась побочным результатом исследований, направленных на выявление генов, кодирующих белки. Однако за последнее десятилетие стало очевидно, как важна роль днРНК в молекулярных и биологических процессах у различных организмов. Это привело к развитию экспериментальных технологий, специально направленных на идентификацию днРНК. На сегодняшний день для выявления и аннотации днРНК широко используют следующие экспериментальные методы [121]:

- 1) анализ РНК-секвенирования – выявление новых транскриптов днРНК;
- 2) микрочипы, разработанные для днРНК;
- 3) SAGE, CAGE – идентификация новых днРНК;
- 4) PARS, SHAPE – предсказание структуры днРНК.

Однако перед началом эксперимента необходимо подготовить днРНК. Для этого применяются различные методы очистки РНК, такие как денатурация путем нагревания, что может негативно повлиять на структуру днРНК [122]. Таким образом, разработаны новые методы, позволяющие избежать денатурации РНК. В большинстве этих подходов используются аффинные метки, которые участвуют в иммобилизации целевых РНК и рибозимов, что обеспечивает более высокую специфичность во время элюции [123]. Метод, предложенный [124], позволяет выявить днРНК с помощью ультрафильтрации и гель-фильтрации. Этот протокол очистки позволяет анализировать днРНК с сохраненными паттернами ко-транскрипционной укладки, а также поддерживает потенциально функциональные структурные элементы [124].

Анализ РНК-секвенирования (RNA-seq) является наиболее распространенным методом идентификации и аннотирования новых днРНК [121]. РНК-секвенирование (1) не требует предварительного знания изучаемых последовательностей, а значит, может быть использовано для изучения транскриптомов немодельных организмов, (2) дает более точные оценки уровня экспрессии и (3) требует меньшего количества РНК для проведения эксперимента [125]. Модификации анализа РНК-секвенирования, такие как цепь-специфичное РНК-секвенирование, позволяют разделять смысловые и антисмысловые классы днРНК [95].

Микрочипы были разработаны для анализа транскриптома, а также количественного и сравнительного анализа экспрессии мРНК между одним или двумя образцами различного происхождения. Стандартные платформы для микрочипов не содержат аннотации днРНК и не подходят для их выявления и анализа. Однако благодаря новым достижениям в технологии микрочипов теперь стало возможным создавать индивидуальные олигонуклеотидные микрочипы [126]. Микрочипы, разработанные для днРНК, используются для исследования профиля экспрессии днРНК в клетках и тканях [121]. Например, в работе [127] применяли микрочип с содержанием 3478 межгенных и инtronных последовательностей в геномах человека, мыши и крысы для поиска днРНК. Исследователи обнаружили 55 новых днРНК с высокой экспрессией, из которых экспрессия восьми была подтверждена в тканях мышей при помощи Нозерн-блот-анализа.

Оба молекулярно-генетических метода – серийный анализ генной экспрессии (SAGE) и кэп-анализ экспрессии генов (CAGE) – основаны на получении и прочтении коротких (обычно длиной 27 нуклеотидов) участков последовательности РНК. Это позволяет одновременно количественно и качественно охарактеризовать экспрессию тысяч различных генов путем анализа их транскриптов

[128]. В отличие от аналогичного метода серийного анализа экспрессии генов (SAGE, superSAGE), в котором фрагменты происходят из других частей транскриптов, SAGE в основном используется для определения точного местоположения сайтов начала транскрипции в геноме. Эти знания в свою очередь позволяют исследователю определять структуру промотора, необходимую для экспрессии гена [128]. Например, метод SAGE был применен в работе [96], посвященной изучению молекулярного патогенеза и биомаркеров предраковых поражений полости рта. В работе были идентифицированы 10 дифференциально экспрессирующихся днРНК, среди которых наибольшую экспрессию имеет днРНК NEAT1, принимающая участие в метастазировании рака.

Методы SHAPE и PARS [129, 130] позволяют получать данные о способе действия и взаимодействиях РНК с другими регуляторными молекулами [121]. В основе метода SHAPE (выборочное ацилирование 2'-гидроксильной группы, анализируемое методом удлинения праймера) лежит использование ангидрида N-метилизатиновой кислоты (NMIA) и его производных для исследования гибких регионов во вторичной структуре РНК. Метод SHAPE был применен для исследования вторичной структуры таких днРНК, как COOLAIR [131] и Braveheart [132]. В методе PARS применяются сразу две нуклеазы: S1 и V1. Первая нуклеаза (S1) расщепляет одноцепочечную ДНК или РНК, вторая (V1) – двухцепочечную. Таким образом, при проведении двух экспериментов параллельно можно получить информацию сразу и о неспаренных, и о спаренных нуклеотидах [130].

Несмотря на перечисленное разнообразие экспериментальных методов идентификации днРНК, данный тип молекул по-прежнему остается малоизученным. В первую очередь это связано с наличием общих характеристик днРНК с белок-кодирующими генами, описанных в разделе “Молекулярные функции днРНК”. Поэтому применение только экспериментальных методов для идентификации днРНК является неэффективным подходом для их исследования. Для массового их изучения в масштабах всего генома лучше применять биоинформационные методы в комбинации с экспериментальными.

## БИОИНФОРМАТИЧЕСКИЕ МЕТОДЫ ИДЕНТИФИКАЦИИ днРНК

Биоинформационные подходы основываются на решении следующих задач: предобработка данных, предсказание днРНК и структурно-функциональная аннотация [16, 24, 133]. Однако на сегодняшний день за счет применения технологий секвенирования нового поколения стремительно увеличивается количество экспериментов

по тотальному секвенированию РНК (RNA-seq). Чтобы повысить эффективность анализа транскриптомов для большого количества обрабатываемых данных, активно разрабатываются вычислительные конвейеры, которые позволяют выполнять стандартизованные этапы обработки данных в режиме, не требующем контроля пользователем вплоть до получения конечного результата. Главными достоинствами существующих конвейеров являются возможность сборки транскриптома для исследуемого организма и анализ дифференциальной экспрессии днРНК. Более подробное сравнение функционала биоинформационных конвейеров представлено в таблице П1.

Биоинформационные конвейеры позволяют автоматизировать и ускорить этапы анализа днРНК, перечисленные выше. Таким образом, предобработка данных включает контроль качества прочтений с помощью стандартных программ и выравнивание на геном.

Предсказания днРНК осуществляются программами CPC2 [134], CPAT [135], PLEK [133]. Перечисленные программы основаны на использовании методов машинного обучения (метод опорных векторов) для определения кодирующего потенциала исследуемого транскрипта. Основными критериями, по которым происходит оценка кодирующего потенциала, являются признаки открытых рамок считывания (ОРС). Встречаются случаи сходства признаков ОРС днРНК с ОРС генов, кодирующих белки, однако покрытие или целостность ОРС обычно у днРНК хуже.

Аннотация днРНК включает анализ структурных особенностей (экзон-инtronная структура, положение на хромосоме), экспрессии, ко-экспрессии, гомологичных последовательностей, консервативности. Анализ структурных особенностей днРНК в основном направлен на длину транскрипта, количество экзонов/инtronов и расположение на хромосоме. Анализ консервативности днРНК в работах можно разделить на две группы: консервативность внутри одного организма и консервативность между разными организмами. Консервативными считаются днРНК с уровнем гомологии выше 50% с днРНК других организмов, неконсервативными – с отсутствием гомологии с другими организмами [134].

Однако несмотря на достоинства конвейеров, они имеют свои недостатки. Большинство конвейеров разработаны на основе транскриптомных данных модельных организмов (человека, мыши, мухи). Вследствие этого для исследования другого организма нужно или использовать уже имеющиеся модели, что может внести ошибки в предсказание днРНК, или дополнительно обучать модель предсказания для данного организма, а затем

вносить эти данные в код конвейера, что может привести к неправильной работе программы.

Отметим, что, несмотря на общие черты, между структурой и функциями днРНК растений и животных существуют и отличия [31]. Например, в работе [136], посвященной анализу программ предсказания днРНК, было показано, что программы, разработанные для человека, плохо предсказывают днРНК растений и наоборот, что свидетельствует о различии структурных особенностей днРНК между человеком и растениями. Однако экспериментальных работ, направленных на сравнение структуры днРНК растений и животных, пока нет.

Поэтому для более точного выявления днРНК необходимы дополнительные фильтры последовательностей, не все конвейеры их предоставляют. На сегодняшний день современные биоинформационные конвейеры решают перечисленные выше недостатки. Так, ICAnnoLncRNA [137] и lncRNA-screen [138] позволяют обучать модель предсказания для любого исследуемого вида. Также ICAnnoLncRNA предоставляет несколько фильтров последовательностей (выравнивание на референсный геном, удаление мобильных элементов, фильтрация коротких изоформ и классификации), что позволяет получить наиболее точное количество выявленных днРНК.

При биоинформационном анализе днРНК производится сравнение вновь секвенированных транскриптов с уже известными последовательностями днРНК. Источником информации об этих последовательностях служат базы данных. Для их пополнения используются различные источники: результаты биоинформационической обработки транскриптомных экспериментов и предсказания днРНК, экспериментальные результаты, полученные при анализе литературных источников, совместное использование как биоинформационических методов, так и литературных источников. Базы данных, которые описывают структуру и функции днРНК растений, приведены в таблице П2.

## ЗАКЛЮЧЕНИЕ

На сегодняшний день различные исследования демонстрируют важность и незаменимость днРНК во многих клеточных процессах, таких как активация и подавление экспрессии генов, импринтинг и деметилирование ДНК, РНК-интерференция, ремоделирование хроматина и др. [139]. У растений днРНК участвуют в регуляции устойчивости к различным видам стресса [1–3], оказывают влияние на устойчивость к гипоксии [32], принимают участие в развитии плодов, корня и листьев. Таким образом, все больше повышается интерес к выявлению днРНК в различных видах. Для этого используют методы широкомасштабного распознавания последовательностей в геномах и транскриптомах.

Особый интерес представляет выявление днРНК в пан-транскриптомах и пан-геномах. Во многих работах [119, 120, 140], посвященных исследованию пан-геномов, отмечается более полная и точная информация о выявленных последовательностях генов. Однако, несмотря на растущее количество исследований, посвященных днРНК, их выявление по-прежнему остается проблематичной задачей. Для решения данных задач разрабатывают различные экспериментальные и биоинформационические подходы.

Работа выполнена при поддержке бюджетного проекта FWNR-2022-0020.

Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Настоящая статья не содержит каких-либо исследований с использованием в качестве объекта людей и 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. *Deng P., Liu S., Nie X., Wu L.* Conservation analysis of long non-coding RNAs in plants // *Sci. China Life Sci.* 2018. V. 61. P. 190–198.
2. *Wu H.-J., Wang Z.-M., Wang M., Wang X.-J.* Widespread long noncoding RNAs as endogenous target mimics for microRNAs in plants // *Plant Physiol.* 2013. V. 161. № 4. P. 1875–1884.
3. *Zhu Q.-H., Wang M.-B.* Molecular functions of long non-coding RNAs in plants // *Genes.* 2012. V. 3. № 1. P. 176–190.
4. *Назипова Н.Н.* Разнообразие некодирующих РНК в геномах эукариот // *Матем. биол. и биоинформат.* 2021. Т. 16. № 2. С. 256–298.
5. *Joshi A., Romanowska J.* Recent advances in computational-based approaches in epigenetics studies // *Epigenetics Methods.* 2020. P. 569–590.
6. *Kim E.-D., Sung S.* Long noncoding RNA: Unveiling hidden layer of gene regulatory networks // *Trends Plant Sci.* 2012. V. 17. № 1. P. 16–21.
7. *Karlik E., Ari S., Gozukirmizi N.* LncRNAs: Genetic and epigenetic effects in plants // *Biotechnol. Biotechnol.* 2019. V. 33. № 1. P. 429–439. <https://doi.org/10.1080/13102818.2019.1581085>
8. *Tsai M.-C., Manor O., Wan Y. et al.* Long noncoding RNA as modular scaffold of histone modification complexes // *Science.* 2010. V. 329. № 5992. P. 689–693. <https://doi.org/10.1126/science.1192002>
9. *Sousa C., Johansson C., Charonet C. et al.* Translational and structural requirements of the early nodulin gene *enod40*, a short-open reading frame-containing RNA,

- for elicitation of a cell-specific growth response in the alfalfa root cortex // *Mol. Cell. Biol.* 2001. V. 21. № 1. P. 354–366.  
<https://doi.org/10.1128/MCB.21.1.354-366.2001>
10. *Medvedeva Y.A., Lennartsson A., Ehsani R. et al.* EpiFactors: A comprehensive database of human epigenetic factors and complexes // *Database*. 2015. V. 2015. P. bav067.
11. *Frankish A., Diekhans M., Jungreis I. et al.* GENCODE 2021 // *Nucl. Acids Res.* 2021. V. 49. № D1. P. D916–D923.
12. *Гордиюк В.В.* Длинные некодирующие РНК – камертон в регуляции клеточных процессов // *Ukr. Biochem. J.* 2014. V. 86. № 2. С. 5–15.
13. *Zhao X., Li J., Lian B. et al.* Global identification of *Arabidopsis* lncRNAs reveals the regulation of MAF4 by a natural antisense RNA // *Nat. Commun.* 2018. V. 9. № 1. P. 5056.
14. *Li X., Wu Z., Fu X. et al.* lncRNAs: Insights into their function and mechanics in underlying disorders // *Mutat. Res. Mutat. Res.* 2014. V. 762. P. 1–21.
15. *Ahmad P., Bensaoud C., Mekki I. et al.* Long non-coding RNAs and their potential roles in the vector–host–pathogen triad // *Life*. MDPI. 2021. V. 11. № 1. P. 56.
16. *De Quattro C., Pè M.E., Bertolini E.* Long noncoding RNAs in the model species *Brachypodium distachyon* // *Sci. Rep.* 2017. V. 7. № 1. P. 11252.
17. *Ma L., Bajic V.B., Zhang Z.* On the classification of long non-coding RNAs // *RNA Biol.* 2013. V. 10. № 6. P. 924–933.  
<https://doi.org/10.4161/rna.24604>
18. *Chen L., Zhu Q.-H., Kaufmann K.* Long non-coding RNAs in plants: Emerging modulators of gene activity in development and stress responses // *Planta*. 2020. V. 252. № 5. P. 92.  
<https://doi.org/10.1007/s00425-020-03480-5>
19. *Griffiths-Jones S.* miRBase: The microRNA sequence database // *MicroRNA Protocols*. 2006. V. 342. P. 129–138.
20. *Amaral P.P., Mattick J.S.* Noncoding RNA in development // *Mamm. Genome*. 2008. V. 19. P. 454–492.
21. *Бейлерли О.А., Гареев И.Ф.* Длинные некодирующие РНК: какие перспективы? // Профилактическая медицина. 2020. Т. 23. № 2. С. 124–128.
22. *Blythe A.J., Fox A.H., Bond C.S.* The ins and outs of lncRNA structure: How, why and what comes next? // *Biochim. Biophys. Acta BBA-Gene Regul. Mech.* 2016. V. 1859. № 1. P. 46–58.
23. *Bryzghalov O., Makałowska I., Szcześniak M.W.* lncEvo: Automated identification and conservation study of long noncoding RNAs // *BMC Bioinformatics*. 2021. V. 22. № 1. P. 59.  
<https://doi.org/10.1186/s12859-021-03991-2>
24. *Zhao Q., Sun Y., Wang D. et al.* LncPipe: A nextflow-based pipeline for identification and analysis of long non-coding RNAs from RNA-Seq data // *J. Genet. Genomics*. 2018. V. 45.  
<https://doi.org/10.1016/j.jgg.2018.06.005>
25. *Talyan S., Filipów S., Ignarski M. et al.* CALINCA – a novel pipeline for the identification of lncRNAs in podocyte disease // *Cells*. MDPI. 2021. V. 10. № 3. P. 692.
26. *Campalans A., Kondorosi A., Crespi M.* Enod40, a short open reading frame – containing mRNA, induces cytoplasmic localization of a nuclear RNA binding protein in *Medicago truncatula* // *Plant Cell*. 2004. V. 16. № 4. P. 1047–1059.
27. *Unver T., Tombuloglu H.* Barley long non-coding RNAs (lncRNA) responsive to excess boron // *Genomics*. 2020. V. 112. № 2. P. 1947–1955.
28. *Khorkova O., Hsiao J., Wahlestedt C.* Basic biology and therapeutic implications of lncRNA // *Adv. Drug Deliv. Rev.* 2015. V. 87. P. 15–24.
29. *Duret L., Chureau C., Samain S. et al.* The *Xist* RNA Gene evolved in eutherians by pseudogenization of a protein-coding gene // *Science*. 2006. V. 312. № 5780. P. 1653–1655.  
<https://doi.org/10.1126/science.1126316>
30. *Graf J., Kretz M.* From structure to function: Route to understanding lncRNA mechanism // *BioEssays*. 2020. V. 42. № 12.  
<https://doi.org/10.1002/bies.202000027>
31. *Golicz A.A., Singh M.B., Bhalla P.L.* The long intergenic noncoding RNA (lncRNA) landscape of the soybean genome // *Plant Physiol.* 2018. V. 176. № 3. P. 2133–2147.
32. *Cheng F., Wu J., Fang L., Wang X.* Syntenic gene analysis between *Brassica rapa* and other Brassicaceae species // *Front. Plant Sci. Frontiers*. 2012. V. 3. P. 30895.
33. *Huang L., Dong H., Zhou D. et al.* Systematic identification of long non-coding RNAs during pollen development and fertilization in *Brassica rapa* // *Plant J.* 2018. V. 96. № 1. P. 203–222.  
<https://doi.org/10.1111/tpj.14016>
34. *Derrien T., Johnson R., Bussotti G. et al.* The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression // *Genome Res.* 2012. V. 22. № 9. P. 1775–1789.
35. *Ponting C.P., Oliver P.L., Reik W.* Evolution and functions of long noncoding RNAs // *Cell*. 2009. V. 136. № 4. P. 629–641.
36. *Golicz A.A., Bayer P.E., Barker G.C. et al.* The pangenome of an agriculturally important crop plant *Brassica oleracea* // *Nat. Commun.* 2016. V. 7. № 1. P. 13390.
37. *Meile L., Croll D., Brunner P.C. et al.* A fungal avirulence factor encoded in a highly plastic genomic

- region triggers partial resistance to *Septoria tritici* blotch // *New Phytol.* 2018. V. 219. № 3. P. 1048–1061. <https://doi.org/10.1111/nph.15180>
38. *Alcaraz L.D., Moreno-Hagelsieb G., Eguiarte L.E. et al.* Understanding the evolutionary relationships and major traits of *Bacillus* through comparative genomics // *BMC Genomics.* 2010. V. 11. № 1. P. 332. <https://doi.org/10.1186/1471-2164-11-332>
39. *Rasko D.A., Rosovitz M.J., Myers G.S.A. et al.* The pangenome structure of *Escherichia coli*: Comparative genomic analysis of *E. coli* commensal and pathogenic isolates // *J. Bacteriol.* 2008. V. 190. № 20. P. 6881–6893. <https://doi.org/10.1128/JB.00619-08>
40. *Merot-Lanthoene V., Tournebize R., Darracq O. et al.* Development and evaluation of a genome – wide Coffee 8.5K SNP array and its application for high-density genetic mapping and for investigating the origin of *Coffea arabica* L. // *Plant Biotechnol. J.* 2019. V. 17. № 7. P. 1418–1430. <https://doi.org/10.1111/pbi.13066>
41. *Budak H., Kaya S.B., Cagirici H.B.* Long non-coding RNA in plants in the era of reference sequences // *Front. Plant Sci. Frontiers.* 2020. V. 11. P. 441273.
42. *Britto-Kido S. de A., Neto J.R.C.F., Pandolfi V. et al.* Natural antisense transcripts in plants: A review and identification in soybean infected with Phakopsora pachyrhizi SuperSAGE Library // *Sci. World J.* 2013. V. 2013.
43. *Kapusta A., Kronenberg Z., Lynch V.J. et al.* Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs // *PLoS Genet.* 2013. V. 9. № 4. P. e1003470.
44. *Лукина С.С., Бурденный А.М., Филиппова Е.А. et al.* Роль длинных некодирующих РНК и ДНК-метилирования в патогенезе рака яичников // *Патол. физиол. и эксперим. терапия.* 2022. Т. 66. № 4. С. 143–156.
45. *Ulitsky I., Bartel D.P.* lincRNAs: Genomics, evolution, and mechanisms // *Cell.* 2013. V. 154. № 1. P. 26–46.
46. *Dinger M.E., Pang K.C., Mercer T.R. et al.* Differentiating protein-coding and noncoding RNA: Challenges and ambiguities // *PLoS Comput. Biol.* 2008. V. 4. № 11. P. e1000176.
47. *Xie C., Yuan J., Li H. et al.* NONCODEv4: Exploring the world of long non-coding RNA genes // *Nucleic Acids Res.* 2014. V. 42. № D1. P. D98–D103.
48. *Pal D., Rao M.R.S.* Long noncoding RNAs in pluripotency of stem cells and cell fate specification // *Long Non Coding RNA Biology* / Ed. Rao M.R.S. Singapore: Springer, 2017. V. 1008. P. 223–252.
49. *Татосян К.А., Зиневич Л.С., Демин Д.Э., Шварц А.М.* Функциональные особенности длинных некодирующих РНК, содержащих последовательности мобильных генетических элементов // *Мол. биол.* 2020. Т. 54. № 5. P. 718–724.
50. *Zhao Z., Zang S., Zou W. et al.* Long non-coding RNAs: New players in plants // *Int. J. Mol. Sci. MDPI.* 2022. V. 23. № 16. P. 9301.
51. *Kopp F., Mendell J.T.* Functional classification and experimental dissection of long noncoding RNAs // *Cell.* Elsevier. 2018. V. 172. № 3. P. 393–407.
52. *Huarte M., Rinn J.L.* Large non-coding RNAs: missing links in cancer? // *Hum. Mol. Genet.* 2010. V. 19. № R2. P. R152–R161.
53. *Hung T., Wang Y., Lin M.F. et al.* Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters // *Nat. Genet.* 2011. V. 43. № 7. P. 621–629.
54. *Loewer S., Cabilio M.N., Guttman M. et al.* Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells // *Nat. Genet.* 2010. V. 42. № 12. P. 1113–1117.
55. *Wang K.C., Chang H.Y.* Molecular mechanisms of long noncoding RNAs // *Mol. Cell.* 2011. V. 43. № 6. P. 904–914.
56. *Chen J., Wang H., Yao Y.* Experimental study of nonlinear ultrasonic behavior of soil materials during the compaction // *Ultrasonics.* 2016. V. 69. P. 19–24.
57. *Kotake Y., Nakagawa T., Kitagawa K. et al.* Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15INK4B tumor suppressor gene // *Oncogene.* 2011. V. 30. № 16. P. 1956–1962.
58. *Ye X., Wang S., Zhao X. et al.* Role of lncRNAs in *cis*- and *trans*-regulatory responses to salt in *Populus trichocarpa* // *Plant J.* 2022. V. 110. № 4. P. 978–993. <https://doi.org/10.1111/tpj.15714>
59. *Yang L., Froberg J.E., Lee J.T.* Long noncoding RNAs: Fresh perspectives into the RNA world // *Trends Biochem. Sci.* 2014. V. 39. № 1. P. 35–43.
60. *Lam M.T.Y., Cho H., Lesch H.P. et al.* Rev-Erbα repress macrophage gene expression by inhibiting enhancer-directed transcription // *Nature.* 2013. V. 498. № 7455. P. 511–515.
61. *Kim T.-K., Hemberg M., Gray J.M.* Enhancer RNAs: A class of long noncoding RNAs synthesized at enhancers // *Cold Spring Harb. Perspect. Biol.* 2015. V. 7. № 1. P. a018622.
62. *Kim T.-K., Hemberg M., Gray J.M. et al.* Widespread transcription at neuronal activity-regulated enhancers // *Nature.* 2010. V. 465. № 7295. P. 182–187.
63. *Lai F., Orom U.A., Cesaroni M. et al.* Activating RNAs associate with Mediator to enhance chromatin architecture and transcription // *Nature.* 2013. V. 494. № 7438. P. 497–501.
64. *Melo C.A., Drost J., Wijchers P.J. et al.* eRNAs are required for p53-dependent enhancer activity and

- gene transcription // *Mol. Cell.* 2013. V. 49. № 3. P. 524–535.
65. *Zhang P., Meng J., Luan Y. et al.* Plant miRNA–lncRNA interaction prediction with the ensemble of CNN and IndRNN // *Interdiscip. Sci. Comput. Life Sci.* 2020. V. 12. P. 82–89.
66. *Chen K., Rajewsky N.* The evolution of gene regulation by transcription factors and microRNAs // *Nat. Rev. Genet.* 2007. V. 8. № 2. P. 93–103.
67. *Jin Q., Zhao Z., Zhao Q. et al.* Long noncoding RNAs: Emerging roles in pulmonary hypertension // *Heart Fail. Rev.* 2020. V. 25. P. 795–815.
68. *Huntzinger E., Izaurralde E.* Gene silencing by microRNAs: Contributions of translational repression and mRNA decay // *Nat. Rev. Genet.* 2011. V. 12. № 2. P. 99–110.
69. *Yoon J.-H., Abdelmohsen K., Kim J. et al.* Scaffold function of long non-coding RNA HOTAIR in protein ubiquitination // *Nat. Commun.* 2013. V. 4. № 1. P. 2939.
70. *Mukherjee N., Corcoran D.L., Nusbaum J.D. et al.* Integrative regulatory mapping indicates that the RNA-binding protein HuR couples pre-mRNA processing and mRNA stability // *Mol. Cell.* 2011. V. 43. № 3. P. 327–339.
71. *Thomson D.W., Dinger M.E.* Endogenous microRNA sponges: Evidence and controversy // *Nat. Rev. Genet.* 2016. V. 17. № 5. P. 272–283.
72. *Franco-Zorrilla J.M., Valli A., Todesco M. et al.* Target mimicry provides a new mechanism for regulation of microRNA activity // *Nat. Genet.* 2007. V. 39. № 8. P. 1033–1037.
73. *Du Q., Wang K., Zou C. et al.* The PILNCR1-miR399 regulatory module is important for low phosphate tolerance in maize // *Plant Physiol.* 2018. V. 177. № 4. P. 1743–1753.
74. *Wang T., Zhao M., Zhang X. et al.* Novel phosphate deficiency-responsive long non-coding RNAs in the legume model plant *Medicago truncatula* // *J. Exp. Bot.* 2017. V. 68. № 21–22. P. 5937–5948.
75. *Faghihi M.A., Zhang M., Huang J. et al.* Evidence for natural antisense transcript-mediated inhibition of microRNA function // *Genome Biol.* 2010. V. 11. № 5. <https://doi.org/10.1186/gb-2010-11-5-r56>
76. *Kimura T., Jiang S., Nishizawa M. et al.* Stabilization of human interferon- $\alpha$ 1 mRNA by its antisense RNA // *Cell. Mol. Life Sci.* 2013. V. 70. № 8. P. 1451–1467. <https://doi.org/10.1007/s00018-012-1216-x>
77. *Wang Y., Pang W.J., Wei N. et al.* Identification, stability and expression of Sirt1 antisense long non-coding RNA // *Gene.* 2014. V. 539. № 1. P. 117–124.
78. *Cai X., Cullen B.R.* The imprinted H19 noncoding RNA is a primary microRNA precursor // *RNA.* 2007. V. 13. № 3. P. 313–316.
79. *Augoff K., McCue B., Plow E.F. et al.* miR-31 and its host gene lncRNA LOC554202 are regulated by promoter hypermethylation in triple-negative breast cancer // *Mol. Cancer.* 2012. V. 11. № 1. <https://doi.org/10.1186/1476-4598-11-5>
80. *Kallen A.N., Zhou X.B., Xu J. et al.* The imprinted H19 lncRNA antagonizes let-7 microRNAs // *Mol. Cell.* 2013. V. 52. № 1. P. 101–112.
81. *Amor B.B., Wirth S., Merchan F. et al.* Novel long non-protein coding RNAs involved in *Arabidopsis* differentiation and stress responses // *Genome Res.* 2009. V. 19. № 1. P. 57–69.
82. *Hirsch J., Lefort V., Vankerschaver M. et al.* Characterization of 43 non-protein-coding mRNA genes in *Arabidopsis*, including the MIR162a-derived transcripts // *Plant Physiol.* 2006. V. 140. № 4. P. 1192–1204.
83. *Lamin-Samu A.T., Zhuo S., Ali M., Lu G.* Long non-coding RNA transcriptome landscape of anthers at different developmental stages in response to drought stress in tomato // *Genomics.* 2022. V. 114. № 4. P. 110383.
84. *Kryuchkova-Mostacci N., Robinson-Rechavi M.* A benchmark of gene expression tissue-specificity metrics // *Brief. Bioinform.* 2017. V. 18. № 2. P. 205–214.
85. *Li L., Eichten S.R., Shimizu R. et al.* Genome-wide discovery and characterization of maize long non-coding RNAs // *Genome Biol.* 2014. V. 15. № 2. P. <https://doi.org/10.1186/gb-2014-15-2-r40>.
86. *Han L., Mu Z., Luo Z. et al.* New lncRNA annotation reveals extensive functional divergence of the transcriptome in maize // *J. Integr. Plant Biol.* 2019. V. 61. № 4. P. 394–405. <https://doi.org/10.1111/jipb.12708>
87. *Subramanian S., Kumar S.* Gene expression intensity shapes evolutionary rates of the proteins encoded by the vertebrate genome // *Genetics.* 2004. V. 168. № 1. P. 373–381.
88. *Yanai I., Benjamin H., Shmoish M. et al.* Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification // *Bioinformatics.* 2005. V. 21. № 5. P. 650–659.
89. *Ceriani L., Verme P.* The origins of the Gini index: Extracts from Variabilità e Mutabilità (1912) by Corrado Gini // *J. Econ. Inequal.* 2012. V. 10. P. 421–443.
90. *Julien P., Brawand D., Soumillon M. et al.* Mechanisms and evolutionary patterns of mammalian and avian dosage compensation // *PLoS Biol.* 2012. V. 10. № 5. P. e1001328.
91. *Xiao S.-J., Zhang C., Zou Q., Ji Z.L.* TiSGeD: A database for tissue-specific genes // *Bioinformatics.* 2010. V. 26. № 9. P. 1273–1275.

92. *Yu X., Lin J., Zack D.J., Qian J.* Computational analysis of tissue-specific combinatorial gene regulation: Predicting interaction between transcription factors in human tissues // *Nucl. Acids Res.* 2006. V. 34. № 17. P. 4925–4936.
93. *Huang X., Li S.Z., Wang Y.* Jensen-Shannon boosting learning for object recognition // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). IEEE. 2005. V. 2. P. 144–149.
94. *Marquardt S., Raitskin O., Wu Z. et al.* Functional consequences of splicing of the antisense transcript COOLAIR on FLC transcription // *Mol. Cell.* 2014. V. 54. № 1. P. 156–165.
95. *Liu X., Hao L., Li D. et al.* Long non-coding RNAs and their biological roles in plants // *Genomics Proteomics Bioinformatics.* 2015. V. 13. № 3. P. 137–147.
96. *Ding J., Lu Q., Ouyang Y. et al.* A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice // *Proc. Natl Acad. Sci. USA.* 2012. V. 109. № 7. P. 2654–2659. <https://doi.org/10.1073/pnas.1121374109>
97. *Song J.-H., Cao J.-S., Wang C.-G.* BcMF11, a novel non-coding RNA gene from *Brassica campestris*, is required for pollen development and male fertility // *Plant Cell Rep.* 2013. V. 32. P. 21–30.
98. *Kim J., Yi H., Choi G. et al.* Functional characterization of phytochrome interacting factor 3 in phytochrome-mediated light signal transduction // *Plant Cell.* 2003. V. 15. № 10. P. 2399–2407.
99. *Wang J., Meng X., Dobrovolskaya O.B. et al.* Non-coding RNAs and their roles in stress response in plants: 5 // *Genomics Proteomics Bioinformatics.* 2017. V. 15. № 5. P. 301–312. <https://doi.org/10.1016/j.gpb.2017.01.007>
100. *Wang A., Hu J., Gao C. et al.* Genome-wide analysis of long non-coding RNAs unveils the regulatory roles in the heat tolerance of Chinese cabbage (*Brassica rapa* ssp. *chinensis*) // *Sci. Rep.* 2019. V. 9. № 1. P. 5002
101. *Wang P., Dai L., Ai J. et al.* Identification and functional prediction of cold-related long non-coding RNA (lncRNA) in grapevine // *Sci. Rep.* 2019. V. 9. № 1. P. 6638.
102. *Chung P.J., Jung H., Jeong D.H. et al.* Transcriptome profiling of drought responsive noncoding RNAs and their target genes in rice // *BMC Genomics.* 2016. V. 17. № 1. P. 563. <https://doi.org/10.1186/s12864-016-2997-3>
103. *Calixto C.P., Tzioutziou N.A., James A.B. et al.* Cold-dependent expression and alternative splicing of *Arabidopsis* long non-coding RNAs // *Front. Plant Sci. Frontiers Media SA.* 2019. V. 10. P. 235.
104. *Jha U.C., Nayyar H., Jha R. et al.* Long non-coding RNAs: Emerging players regulating plant abiotic stress response and adaptation // *BMC Plant Biol.* 2020. V. 20. № 1. P. 466. <https://doi.org/10.1186/s12870-020-02595-x>
105. *Zhang X., Dong J., Deng F. et al.* The long non-coding RNA lncRNA973 is involved in cotton response to salt stress // *BMC Plant Biol.* 2019. V. 19. № 1. P. 459. <https://doi.org/10.1186/s12870-019-2088-0>
106. *Wang X., Fan H., Wang B., Yuan F.* Research progress on the roles of lncRNAs in plant development and stress responses // *Front. Plant Sci. Frontiers Media SA.* 2023. V. 14. P. 1138901.
107. *Chen J., Zhong Y., Qi X.* LncRNA TCONS\_00021861 is functionally associated with drought tolerance in rice (*Oryza sativa* L.) via competing endogenous RNA regulation // *BMC Plant Biol.* 2021. V. 21. № 1. P. 410. <https://doi.org/10.1186/s12870-021-03195-z>
108. *Kazemzadeh M., Safaralizadeh R., Orang A.V.* LncRNAs: Emerging players in gene regulation and disease pathogenesis // *J. Genet.* 2015. V. 94. P. 771–784.
109. *Zhu Y., Chen L., Hong X. et al.* Revealing the novel complexity of plant long non-coding RNA by strand-specific and whole transcriptome sequencing for evolutionarily representative plant species // *BMC Genomics.* 2022. V. 23. № S4. P. 381. <https://doi.org/10.1186/s12864-022-08602-9>
110. *Ulitsky I.* Evolution to the rescue: Using comparative genomics to understand long non-coding RNAs // *Nat. Rev. Genet.* 2016. V. 17. № 10. P. 601–614.
111. *Wang H., Niu Q.W., Wu H.W. et al.* Analysis of non-coding transcriptome in rice and maize uncovers roles of conserved lnc RNAs associated with agriculture traits // *Plant J.* 2015. V. 84. № 2. P. 404–416. <https://doi.org/10.1111/tpj.13018>
112. *Nitsche A., Stadler P.F.* Evolutionary clues in lncRNAs // *WIREs RNA.* 2017. V. 8. № 1. <https://doi.org/10.1002/wrna.1376>
113. *Sang S., Chen W., Zhang D. et al.* Data integration and evolutionary analysis of long non-coding RNAs in 25 flowering plants: 3 // *BMC Genomics.* 2021. V. 22. № 3. P. 739. <https://doi.org/10.1186/s12864-021-08047-6>
114. *Zhang Y.-C., Liao J.Y., Li Z.Y. et al.* Genome-wide screening and functional analysis identify a large number of long noncoding RNAs involved in the sexual reproduction of rice // *Genome Biol.* 2014. V. 15. № 12. P. 512. <https://doi.org/10.1186/s13059-014-0512-1>
115. *Pronzin A.Y., Bragina M.K., Salina E.A.* Crop pangenomes // *Vavilov. J. Genet. Breed.* 2021. V. 25. № 1. P. 57.
116. *Vernikos G., Medini D., Riley D.R., Tettelin H. et al.* Ten years of pan-genome analyses // *Curr. Opin. Microbiol.* 2015. V. 23. P. 148–154.

117. *Lapierre P., Gogarten J.P.* Estimating the size of the bacterial pan-genome // *Trends Genet.* 2009. V. 25. № 3. P. 107–110.
118. *Chekanova J.A., Gregory B.D., Reverdatto S.V. et al.* Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the *Arabidopsis* transcriptome // *Cell.* 2007. V. 131. № 7. P. 1340–1353.
119. *Medina C.A., Samac D.A., Yu L.-X.* Pan-transcriptome identifying master genes and regulation network in response to drought and salt stresses in Alfalfa (*Medicago sativa* L.) // *Sci. Rep.* 2021. V. 11. № 1. P. 17203.
120. *Jin M., Liu H., He C. et al.* Maize pan-transcriptome provides novel insights into genome complexity and quantitative trait variation // *Sci. Rep.* 2016. V. 6. № 1. P. 18936.
121. *Chowdhary A., Satagopam V., Schneider R.* Long non-coding RNAs: Mechanisms, experimental, and computational approaches in identification, characterization, and their biomarker potential in cancer // *Front. Genet.* 2021. V. 12. <https://doi.org/10.3389/fgene.2021.649619>
122. *Svergun D.I., Koch M.H.* Small-angle scattering studies of biological macromolecules in solution // *Rep. Prog. Phys.* 2003. V. 66. № 10. P. 1735.
123. *Schön P.* Atomic force microscopy of RNA: State of the art and recent advancements // *Seminars in Cell & Developmental Biology.* 2018. V. 73. P. 209–219.
124. *Chillón I., Marcia M., Legiewicz M. et al.* Native purification and analysis of long RNAs // *Methods in Enzymology.* 2015. V. 558. P. 3–37.
125. *Cheung F., Haas B.J., Goldberg S.M.D. et al.* Sequencing *Medicago truncatula* expressed sequenced tags using 454 Life Sciences technology // *BMC Genomics.* 2006. V. 7. № 1. <https://doi.org/10.1186/1471-2164-7-272>
126. *Au P.C.K., Zhu Q.-H.* Identification of lncRNAs using computational and experimental approaches // *Regulatory RNAs / Eds Mallick B., Ghosh Z.* Berlin; Heidelberg: Springer, 2012. P. 319–340.
127. *Babak T., Blencowe B.J., Hughes T.R.* A systematic search for new mammalian noncoding RNAs indicates little conserved intergenic transcription // *BMC Genomics.* 2005. V. 6. № 1. P. 104. <https://doi.org/10.1186/1471-2164-6-104>
128. *Shiraki T., Kondo S., Katayama S. et al.* Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage // *Proc. Natl Acad. Sci. USA.* 2003. V. 100. № 26. P. 15776–15781. <https://doi.org/10.1073/pnas.2136655100>
129. *Merino E.J., Wilkinson K.A., Coughlan J.L., Weeks K.M.* RNA structure analysis at single nucleotide resolution by Selective 2'-Hydroxyl Acylation and Primer Extension (SHAPE) // *J. Am. Chem. Soc.* 2005. V. 127. № 12. P. 4223–4231. <https://doi.org/10.1021/ja043822v>
130. *Kertesz M., Wan Y., Mazor E. et al.* Genome-wide measurement of RNA secondary structure in yeast // *Nature.* 2010. V. 467. № 7311. P. 103–107.
131. *Hawkes E.J., Hennelly S.P., Novikova I.V. et al.* COOLAIR antisense RNAs form evolutionarily conserved elaborate secondary structures // *Cell Rep.* 2016. V. 16. № 12. P. 3087–3096.
132. *Kim D.N., Thiel B.C., Mrozowich T. et al.* Zinc-finger protein CNBP alters the 3-D structure of lncRNA Braveheart in solution // *Nat. Commun.* 2020. V. 11. № 1. P. 148.
133. *Li A., Zhang J., Zhou Z.* PLEK: A tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme // *BMC Bioinformatics.* 2014. V. 15. № 1. P. 311. <https://doi.org/10.1186/1471-2105-15-311>
134. *Kang Y.-J., Yang D.C., Kong L. et al.* CPC2: A fast and accurate coding potential calculator based on sequence intrinsic features // *Nucl. Acids Res.* 2017. V. 45. № W1. P. W12–W16.
135. *Wang L., Park H.J., Dasari S. et al.* CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model // *Nucl. Acids Res.* 2013. V. 41. № 6. P. e74.
136. *Da Costa Negri T., Paschoal A.R., Alves W.A.L.* Comparison tools for lncRNA identification: Analysis among plants and humans // 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB). IEEE. 2020. P. 1–8.
137. *Pronozin A.Y., Afonnikov D.A.* ICAnnoLncRNA: A snakemake pipeline for a long non-coding-RNA search and annotation in transcriptomic sequences // *Genes.* MDPI. 2023. V. 14. № 7. P. 1331.
138. *Gong Y., Huang H.T., Liang Y. et al.* lncRNAscreen: An interactive platform for computationally screening long non-coding RNAs in large genomics datasets // *BMC Genomics.* 2017. V. 18. № 1. P. 434. <https://doi.org/10.1186/s12864-017-3817-0>
139. *Ким О.И., Кириченко Е.Ю., Кириченко Ю.Г. и др.* Длинные некодирующие РНК, ассоциированные с онкогенезом: биологическое значение и перспективы применения в диагностике // Клин. лаб. диагностика. 2016. Т. 61. № 1. P. 13–16.
140. *Gao L., Gonda I., Sun H. et al.* The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor // *Nat. Genet.* 2019. V. 51. № 6. P. 1044–1051.

**Table S1.** Comparison of lncRNA bioinformatic analysis pipelines

	LncPipe [1]	LncEvo [2]	CALINCA [3]	lncRNA Detector [4]	lncRNA- screen [5]	ICAnnoLncRNA [6]
Quality control and assembly of the transcriptome	+	+	+	—	+	—
Prediction model training for a non-model organism	—	—	—	—	+	+
Possibility of using data for plants	—	—	—	+	+	+
Prediction of lncRNA using specialized programs	+	+	—	+	+	+
Classification of lncRNAs according to their location relative to protein-coding genes	+	+	+	—	+	+
Removal of transcriptional noise and assembly redundancy	—	—	—	—	—	+
Deletion of sequences that are potential mobile elements	—	—	—	—	—	+
Differential expression analysis	+	—	+	—	+	—
Analysis of lncRNA conservation between different organisms	—	+	+	—	+	+
Analysis of lncRNA distribution in tissues based on its expression	+	—	+	—	+	+

**Table S2.** Databases for lncRNAs in plants

Title	Description	Data source	Reference
PLncDBv2.0	Plant database with more than 1,246,372 lncRNAs predicted for 80 organisms from chlorophytes to embryophytes. Presents information on expression, tissue specificity, mutations, and developmental stages of lncRNA transcripts in 13834 datasets from different organisms. Allows the user to study the relationship between lncRNAs and epigenetic markers	Methods for lncRNA prediction	[7]
PNRD	Includes from 166 plant species, over 25,739 lncRNAs of 16 types. Offers several analytical tools and includes a proprietary genomic search engine, protein-coding sequence prediction methods and a microRNA prediction method	Literature and databases	[8]
NONCODE v6	The database contains 16 animal species and 549,813 lncRNAs belonging to them, and 23 plant species and 94,697 lncRNA transcripts	Literature and databases	[9]
GREENC v2	Includes more than 495,000 lncRNA transcripts for 94 plant and algal species	Methods for lncRNA prediction	[10]
CANTATAdb v2	The database contains more than 239,631 lncRNAs from 36 plant species and 3 algal species	Methods for lncRNA prediction	[11]
EVLncRNAs v2	Includes 4010 lncRNAs for 124 species of organisms. Provides the user with a network of lncRNA interactions with microRNAs, proteins, genes, and other functional elements. Also contains several lncRNA analysis and prediction programs	Literature	[12]

## REFERENCES

1. *Zhao Q. et al.* LncPipe: A Nextflow-based pipeline for identification and analysis of long non-coding RNAs from RNA-Seq data // *J Genet Genomics*. 2018. Vol. 45, № 7. P. 399–401.
2. *Bryzgalov O., Makalowska I., Szcześniak M.W.* lncEvo: Automated identification and conservation study of long noncoding RNAs // *BMC Bioinformatics*. 2021. Vol. 22, № 1. P. 59.
3. *Talyan S. et al.* CALINCA—A Novel Pipeline for the Identification of lncRNAs in Podocyte Disease // *Cells*. MDPI, 2021. Vol. 10, № 3. P. 692.
4. *Shukla B. et al.* lncRNADetector: A bioinformatics pipeline for long non-coding RNA identification and MAPsInc: a repository of medicinal and aromatic plant lncRNAs // *RNA Biol*. 2021. Vol. 18, № 12. P. 2290–2295.
5. *Gong Y. et al.* lncRNA-screen: An interactive platform for computationally screening long non-coding RNAs in large genomics datasets // *BMC Genomics*. 2017. Vol. 18, № 1. P. 434.
6. *Pronozin A.Y., Afonnikov D.A.* ICAnnoLncRNA: A Snakemake Pipeline for a Long Non-Coding-RNA Search and Annotation in Transcriptomic Sequences // *Genes*. MDPI, 2023. Vol. 14, № 7. P. 1331.
7. *Jin J. et al.* PLncDB V2. 0: A comprehensive encyclopedia of plant long noncoding RNAs // *Nucleic Acids Res*. Oxford University Press, 2021. Vol. 49, № D1. P. D1489–D1495.
8. *Yi X. et al.* PNRD: A plant non-coding RNA database // *Nucleic Acids Res*. Oxford University Press, 2015. Vol. 43, № D1. P. D982–D989.
9. *Zhao Y. et al.* NONCODE 2016: An informative and valuable data source of long non-coding RNAs // *Nucleic Acids Res*. Oxford University Press, 2016. Vol. 44, № D1. P. D203–D208.
10. *Gallart A.P. et al.* GREENC: A Wiki-based database of plant lncRNAs // *Nucleic Acids Res*. Oxford University Press, 2016. Vol. 44, № Database issue. P. D1161.
11. *Szcześniak M.W. et al.* CANTATAb 2.0: Expanding the Collection of Plant Long Noncoding RNAs // *Plant Long Non-Coding RNAs* / ed. Chekanova J.A., Wang H.-L.V. New York, NY: Springer New York, 2019. Vol. 1933. P. 415–429.
12. *Zhou B. et al.* EVLncRNAs 2.0: An updated database of manually curated functional long non-coding RNAs validated by low-throughput experiments // *Nucleic Acids Res*. Oxford University Press, 2021. Vol. 49, № D1. P. D86–D91.

## The Role of Long Non-Coding RNAs in Plants

A. Y. Pronozin<sup>1, 2, \*</sup>, D. A. Afonnikov<sup>1, 2, 3</sup>

<sup>1</sup>*Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 630090 Russia*

<sup>2</sup>*Kurchatov Genomic Center, Institute of Cytology and Genetics, Siberian Branch,*

*Russian Academy of Sciences, Novosibirsk, 630090 Russia*

<sup>3</sup>*Novosibirsk National Research State University, Novosibirsk, 630090 Russia*

\*e-mail: pronozinartem95@gmail.com

Long non-coding RNAs (lncRNAs) are a class of linear or circular RNA molecules longer than 200 nucleotides without open reading frames. Experimental studies have shown the involvement of lncRNAs in the regulation of resistance to cold, salt, and heat stress, and in fruit, root, and leaf development. However, experimental methods are labor-intensive and costly approaches and cannot yet be used for genome-wide mass studies of lncRNAs. For this purpose, bioinformatic approaches that aim at large-scale recognition of lncRNA sequences in genomes and transcriptomes have been applied. However, despite the growing number of studies devoted to the structural and functional analysis of lncRNAs, this type of molecule remains poorly understood. This is due to the many factors that need to be considered when identifying lncRNAs. The use of pan-genomes and pan-transcriptomes will improve the efficiency of the study and the total number of predicted lncRNAs compared to using the genome of a single species representative. This review focuses on describing the molecular and biological functions of lncRNAs, experimental and bioinformatic methods of identification, patterns of evolution, detection and analysis of lncRNAs at the scale of pan-genomes and pan-transcriptomes.

**Keywords:** long non-coding RNAs, microRNAs, transcription regulation, pan-genome.

## ВЗАИМОСВЯЗЬ микроРНК С ТРАНСПОЗОНАМИ В РАЗВИТИИ ОСТЕОАРТРИТА

© 2025 г. Р. Н. Мустафин<sup>1</sup>, \*, Э. К. Хуснутдинова<sup>2</sup>

<sup>1</sup>Башкирский государственный медицинский университет, Уфа, 450008 Россия

<sup>2</sup>Институт биохимии и генетики – обособленное структурное подразделение Уфимского федерального исследовательского центра Российской академии наук, Уфа, 450054 Россия

\*e-mail: ruji79@mail.ru

Поступила в редакцию 12.05.2024 г.

После доработки 14.06.2024 г.

Принята к публикации 18.06.2024 г.

Согласно проведенным GWAS, остеоартрит ассоциирован с более 100 различными SNP, большинство из которых локализованы в инtronных и межгенных областях, где расположены гены транспозонов и произошедших от них некодирующих РНК. В ряде исследований определена также активация ретротранспозонов в тканях суставов и в периферической крови пациентов с остеоартритом. Сделано предположение о влиянии на этиопатогенез остеоартрита активированных транспозонов, вызывающих старение и связанное с ним воспаление. Для подтверждения данной гипотезы проведен поиск данных об изменении экспрессии специфических микроРНК, произошедших от мобильных генетических элементов при старении и остеоартрите. В результате найдено 23 таких микроРНК, участие которых в развитии болезни связано с воздействием на гены и сигнальные пути регуляции пролиферации и апоптоза клеток, воспалительные и метаболические процессы, механизмы деградации хряща. Изменение экспрессии данных микроРНК свидетельствует о том, что эпигенетические механизмы старения вовлечены в этиопатогенез остеоартрита вследствие патологической активации транспозонов, комплементарных последовательностям некодирующих РНК, произошедших от них в эволюции.

**Ключевые слова:** иммунная система, микроРНК, мобильные генетические элементы, ретроэлементы, транспозоны, остеоартрит.

**DOI:** 10.31857/S0016675825010023 **EDN:** VFLMDB

Остеоартрит (ОА) является самым распространенным типом артрита [1] и представляет собой гетерогенное многофакторное заболевание, характеризующееся ассоциацией более чем со 100 различными однонуклеотидными полиморфизмами (SNP – single nucleotide polymorphism), большинство из которых расположены в интранах генов и в межгенных областях [2, 3]. Болезнь характеризуется развитием воспаления в синовиальной оболочке сустава с деградацией хряща [4]. Помимо генетических факторов, на риск ОА влияют возраст, женский пол, семейный анамнез, курение, профессиональное воздействие избыточных нагрузок и ожирение, в большей степени остеоартриту подвержены женщины [5]. Глобальная распространенность ОА в мире, согласно статистическим данным за 2020-й год, составляет 7.6% всего населения, возрастая до 14.8% для людей старше 30 лет. Отмечена выраженная ассоциация ОА со старением [1]. Так, частота встречаемости ОА у людей старше 50 лет составляет уже 29.3% [6], а старше

70 лет – 40% [7]. Поскольку старение характеризуется развитием аутоиммунного асептического воспаления и гиперпродукцией интерферона в ответ на прогрессирующую гиперактивацию мобильных генетических элементов (МГЭ) [8, 9], можно сделать предположение о влиянии этих механизмов и на этиопатогенез ОА. Действительно, в тканях пораженных ОА суставов определяются транскрипты эндогенных ретровирусов HERV-E2 и HERV-WE1 [10], а в лейкоцитах крови больных ОА выявлено значительное снижение метилирования ретроэлементов (РЭ) LINE1 по сравнению со здоровым контролем, что свидетельствует об их активации [11].

МГЭ занимают не менее 45% генома человека и представляют собой генетические элементы, которые перемещаются внутри генома и подразделяются на РЭ (класс I) и ДНК-транспозоны (класс II). К РЭ относятся содержащие длинные концевые повторы (LTR – long terminal repeats) и не содержащие их элементы (включают автономные LINE и

неавтономные SINE и SVA) [12]. О вероятной роли МГЭ в развитии ОА свидетельствует расположение ассоциированных с болезнью SNP главным образом в инtronных и межгенных областях [2, 3], где находятся большинство МГЭ, а также произошедшие от них в эволюции гены микроРНК [13–15]. Кроме того, в опытах на мышах было показано, что при синовиальном воспалении суставов, пораженных ОА, выявляется снижение концентрации деацетилазы гистонов SIRT6. В результате индуцируется поляризация M1-макрофагов с высвобождением в них провоспалительных цитокинов [16]. Истощение SIRT6 наблюдается при старении и считается одним из эпигенетических драйверов данного процесса в связи со снижением сайленсинга МГЭ [17], продукты экспрессии которых стимулируют интерфероновый ответ [8]. Последний в свою очередь активирует M1-макрофаги, вызывая продукцию ими интерлейкинов IL-1 $\beta$ , IL-6, IL-12, фактора некроза опухоли- $\alpha$  (TNF- $\alpha$ ), активных форм кислорода и индуцильной синтетазы оксида азота (iNOS) [18]. В плазме крови и в синовиальной жидкости больных ОА определено также достоверное повышение концентрации CXCL10 (C-X-C motif chemokine ligand 10) – индуцильного интерфероном- $\gamma$  белка (Interferon gamma-induced protein 10, IP-10) по сравнению со здоровым контролем [19]. Перечисленные факты свидетельствуют о вероятной роли патологической экспрессии МГЭ при старении как драйверного процесса для развития ОА. Поскольку МГЭ являются регуляторами эпигенетических факторов [20], следует рассмотреть особенности их изменений при ОА.

Проведенный в 2021 г. GWAS с использованием образцов ДНК от 826 690 больных с различными типами ОА позволил определить независимую ассоциацию 100 различных SNP [2]. GWAS отдельных типов ОА также идентифицировали множество различных SNP, ассоциированных с болезнью. Например, с ОА тазобедренного сустава, согласно проведенному GWAS в 2023 г., ассоциированы 42 SNP [3]. Объяснить влияние такого количества генетических вариантов даже при помощи современных биоинформационных технологий очень сложно. В то же время результаты проведенных мета-анализов показывают достоверную ассоциацию ОА с аллельными вариантами лишь нескольких генов иммунной системы: *IL17A* [21], *IL1RN* [22], *IL6* [23] и компонента соединительной ткани *COL11A1* [24]. Данные ассоциации не могут объяснить сложную гетерогенную природу ОА. Однако расположение большинства ассоциированных с болезнью SNP в инtronах и некодирующих частях генома [2, 3] подтверждает предположение о роли МГЭ в этиопатогенезе ОА, поскольку МГЭ находятся главным образом в инtronах и межгенных областях [13–15]. На рис. 1 представлена схема

вероятных путей влияния МГЭ на развитие ассоциированного со старением ОА.

Представляют интерес результаты исследования нарушений экспрессии генов в тканях пораженных ОА суставов, поскольку это отражает влияние изменений эпигенетических факторов. Причиной таких изменений могут быть воздействия микроРНК, регулирующих экспрессию генов, кодирующих участвующие в воспалении факторы, стимуляция которых характерна также для старения под влиянием активированных МГЭ, вызывающих иммунный ответ. Согласно результатам ряда исследований, при ОА под влиянием микроРНК повышается экспрессия генов различных провоспалительных белков (табл. 1) [25–30], таргетное ингибирование которых с помощью регуляции эпигенетических факторов перспективно в лечении ОА [31]. Кроме того, согласно результатам анализа экспрессии генов в тканях пораженных ОА суставов, при болезни снижается экспрессия участвующих в иммунных реакциях генов *KLF2*, *KLF4* [32], *KLF9* [33], белковые продукты которых, крупель-подобные транскрипционные факторы, ингибируют воспаление, *JUN*, кодирующего транскрипционный фактор, который стимулирует апоптоз иммунокитов. При ОА определена также низкая экспрессия *MYC*, подавляющего пролиферацию клеток, стимулирующего их апоптоз и ингибирующего *IL-1 $\beta$* , TNF- $\alpha$ , IL-6, MMP-13. У больных ОА в тканях суставов выявлено снижение экспрессии генов: *NFKBIA*, ингибитора NFKB, предотвращающего образование комплексов NFKB/REL, связанных с воспалением [34]; *TFNAP3*, кодирующего индуцируемый фактором некроза опухолей белок цинковых пальцев, редактирующего убиквитин и участвующего в иммунных и воспалительных реакциях [35]; *MCL1*, регулятора апоптоза, необходимого для выживания фибробластов, макрофагов и лимфоцитов [36]; *CEACAM-1*, иммунного регулятора Т-лимфоцитов, подавляющего воспаление; *TNFRSF18*, кодирующего белок GITRL, индуцированного глюкокортикоидами лиганда рецептора TNF, регулирующего воспаление и оказывающего противовоспалительный эффект [26]. Причиной нарушения экспрессии этих генов может служить эпигенетическая дисрегуляция, обусловленная воздействием микроРНК вследствие патологической активации МГЭ.

#### Взаиморегуляция транспозонов и микроРНК

Патологическая активация МГЭ при старении, отражающаяся на развитии ОА, может быть обусловлена различными механизмами воздействия транспозонов на эпигенетическую регуляцию (рис. 2). Данные механизмы обусловлены наличием комплементарных последовательностей между МГЭ и микроРНК в связи с возникновением микроРНК от МГЭ в эволюции или

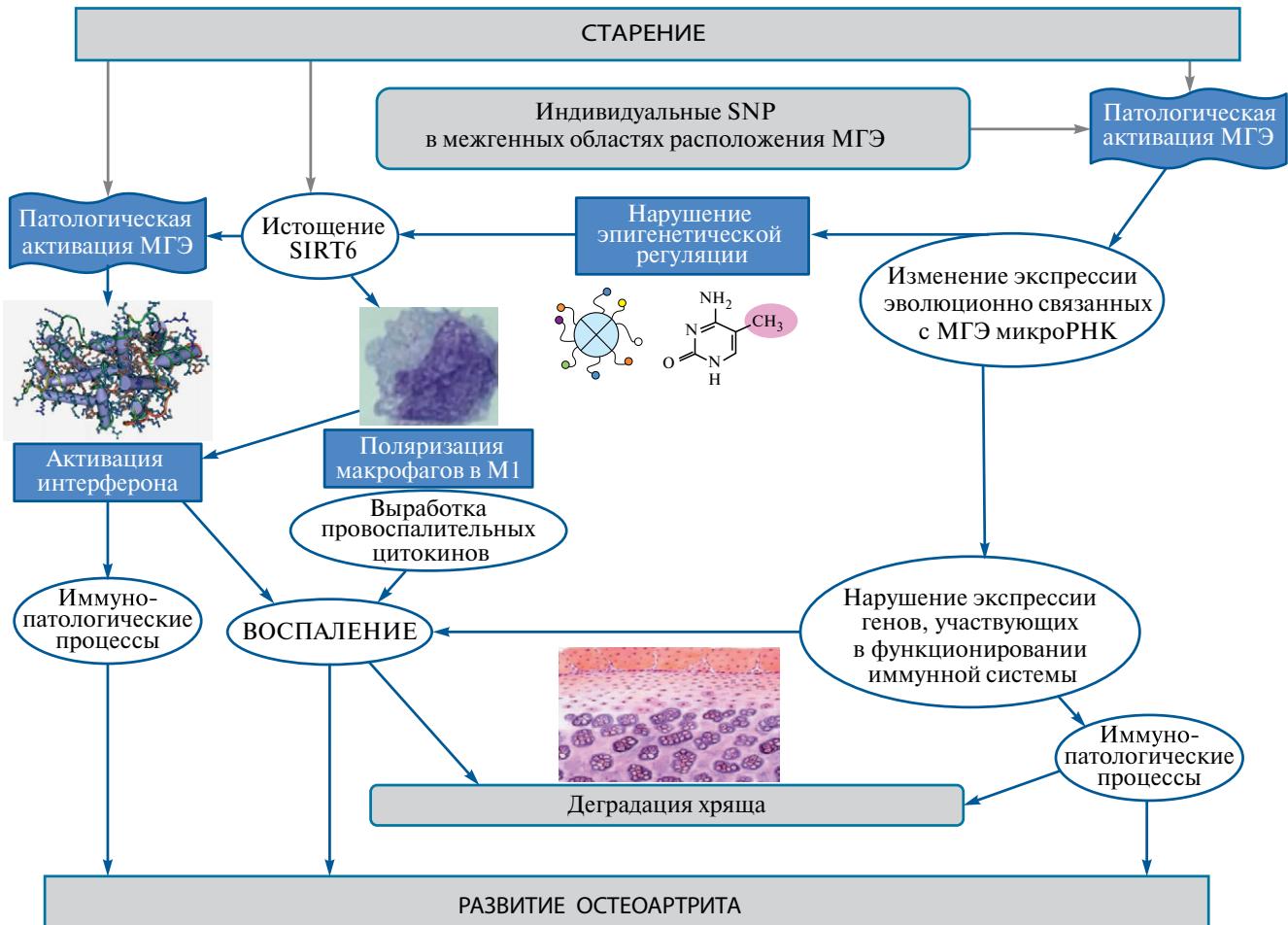


Рис. 1. Схема вероятных путей влияния активированных при старении мобильных генетических элементов (МГЭ) на эпигенетические механизмы развития остеоартрита.

непосредственным образованием микроРНК из транскриптов МГЭ [15]. Еще в 2016 г. G. Wei и соавт. была создана база данных о происхождении микроРНК от специфических МГЭ, названная MDTE DB (miRNAs derived from transposable elements database), в которую вошли 661 микроРНК человека [15]. Активированные МГЭ могут оказывать влияние на произошедшие от них микроРНК за счет связывания с ними продуктов транскрипции МГЭ, действующих как “губки” для микроРНК путем комплементарного связывания с нуклеотидными последовательностями по причине их эволюционного родства. Так блокируется влияние РНК-интерференции на мРНК генов-мишеней этих микроРНК [37]. Такой принцип регуляции определен не только у животных, но и у растений. Например, транскрипт LTR-содержащего ретровирусного элемента *MIKKI* (переводится с корейского языка как “приманка”), экспрессируемый в корнях риса, является имитатором для miR-171, которая дестабилизирует мРНК транскрипционных факторов корня, подобных SCARECROW. Процессированные транскрипты *MIKKI* действуют как ловушки

для miR-171, запуская их деградацию и обеспечивая накопление специфичных для корня мРНК транскрипционных факторов [38]. Транскрипты LTR-содержащих РЭ [39] и LINE1 функционируют в качестве молекул длинных некодирующих РНК, взаимодействуя со специфическими участками хроматина и регулируя экспрессию генов (в том числе контролируемых микроРНК) [40].

Некоторые микроРНК образуются непосредственно из генов МГЭ, которые являются основой для шпилечных структур пре-микроРНК. В результате формируются различные микроРНК, образующие регуляторную сеть управления экспрессией генов, меняющейся в ходе онтогенеза в тканях и органах человека. (Для анализа таких процессов было создано веб-приложение Brain miRTEexplorer [41]). Поэтому патологическая активация МГЭ приводит к образованию различных микроРНК из их транскриптов, влияющих на регуляторные сети других микроРНК в организме. МГЭ оказывают регуляторное воздействие на микроРНК за счет образования малых интерферирующих РНК (миРНК, siRNA – small interfering

**Таблица 1.** Гены, повышение экспрессии которых вовлечено в иммунные реакции при остеоартрите

Название гена	Название белка – продукта экспрессии гена	Функция белка [автор]
<i>C5AR1</i>	рецептор анафилотоксина C5a, экспрессируемый иммунными клетками	химический аттрактант и медиатор воспаления [25]
<i>CTLA4</i>	клеточный рецептор иммуноглобулинов	стимулирует иммунный ответ [26]
<i>EDNRB</i>	рецептор эндотелина типа B, связанный с G-белком	активирует фосфатидилинозитол-кальциевую систему [27]
<i>FSH</i>	фолликуло-стимулирующий гормон	стимулирует воспаление в суставе [26]
<i>HLA-DMB</i>	белки главного комплекса гистосовместимости класса II, DM beta	участвуют в иммунных реакциях [28]
<i>IL1B</i>	интерлейкин-1-бета	провоспалительный цитокин, вырабатываемый иммунными клетками [25]
<i>IL1R1</i>	рецептор интерлейкина 1	передача провоспалительных сигналов [27]
<i>IL4R</i>	рецептор интерлейкина 4	передача иммунных сигналов [29]
<i>IL6R</i>	рецептор интерлейкина 6	передача противовоспалительных сигналов [29]
<i>IL10</i>	интерлейкин 10	противовоспалительный цитокин, вырабатываемый иммунными клетками [25]
<i>IRAK3</i>	ассоциированная с рецептором интерлейкина-1 киназа	способствует передаче провоспалительных сигналов [30]
<i>RHOB</i>	малая везикулярная ГТФаза RhoB	активирует провоспалительные IL-1 $\beta$ , LPS, TNF $\alpha$ [30]
<i>SOX13</i>	транскрипционный фактор SRY-related HMG-box	автоиммунный антиген, модулирующий воспалительный ответ [30]
<i>TNFSF11</i>	член семейства факторов некроза опухолей	стимулирует активацию В- и Т-лимфоцитов и их инфильтрацию тканей сустава [27]

RNA) из транскриптов МГЭ. При этом миРНК являются конкурентными молекулами для связывания с мРНК-мишениями для микроРНК, нивелируя их воздействие на экспрессию генов. Этот эффект связан с защитными системами клеток-хозяев против активированных МГЭ в их геномах, запуская деградацию транскриптов МГЭ рибонуклеазами до миРНК. Последние оказывают посттранскрипционное ингибирование мРНК генов, которые не содержат фрагментов МГЭ в своем составе за счет частичной комплементарности последовательностей нуклеотидов [42].

Одним из путей взаимодействий микроРНК с МГЭ в регуляции активности генов является также подавление их экспрессии при связывании микроРНК со специфическими структурами ДНК, образованными благодаря встроенным в эти области МГЭ. В геноме человека Z-форма ДНК образуется

эндогенными ретровирусами, которые обеспечивают функциональные гены альтернативными промоторами [43]. Так, Z-форма ДНК, расположенная в промоторной области гена простагландинредуктазы (*PTGR1*), образована MER4 ретроэлементом. С последовательностями 5'-CACACACA-3' данной области связывается miR-6867-5p (содержащая комплементарные повторы 5'-GUGUGUG-3'), подавляя экспрессию гена *PTGR1* за счет ингибирования образования Z-формы (которая предполагает активацию экспрессии) [12]. Кроме того, у человека описан феномен РНК-направленного метилирования ДНК (RdDM – RNA-dependent DNA methylation), при помощи которого образованные из транскриптов МГЭ микроРНК [41] и миРНК [42] могут воздействовать на экспрессию произошедших от них в эволюции микроРНК за счет наличия комплементарных последовательностей в

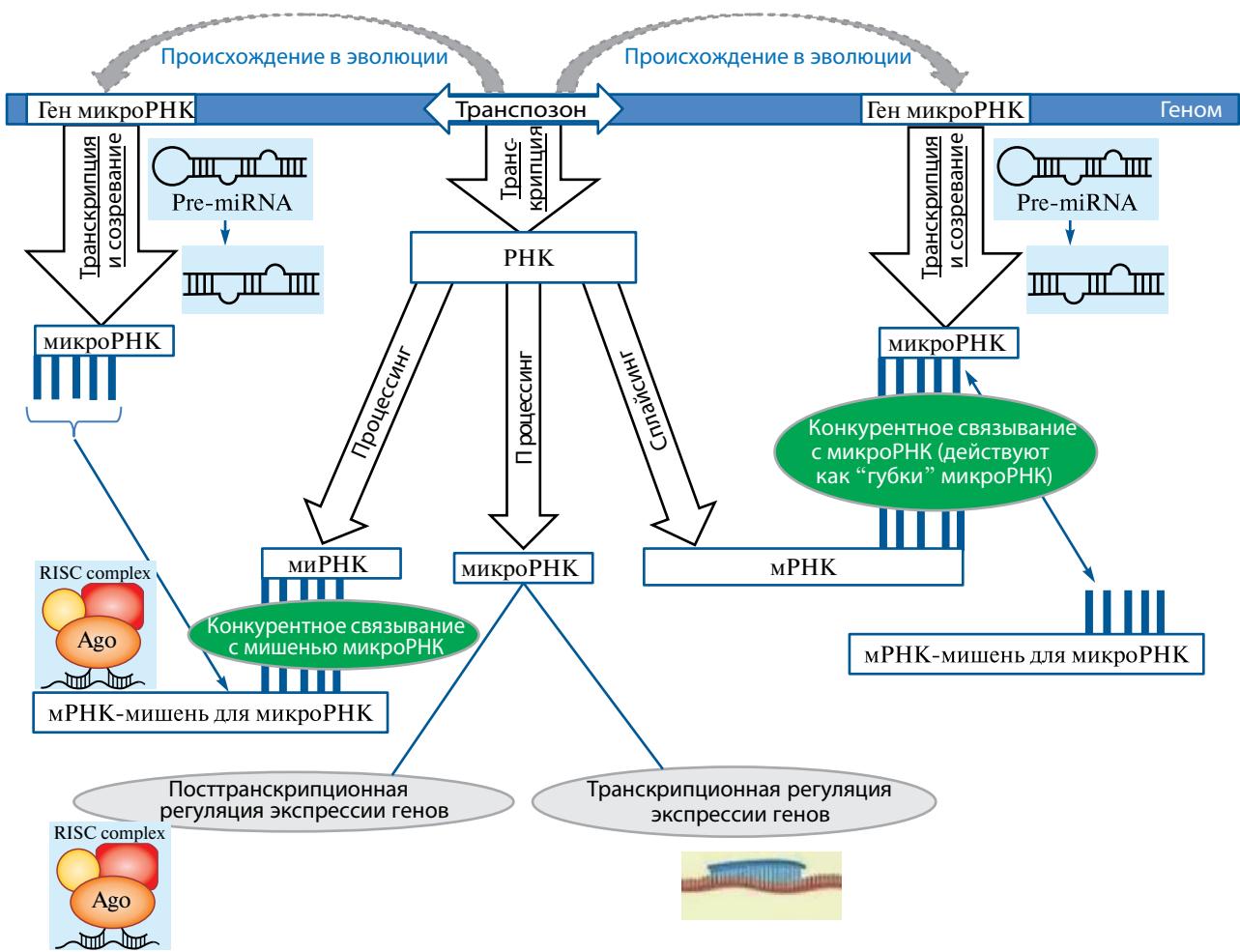


Рис. 2. Механизмы влияния транспозонов на эпигенетическую регуляцию микроРНК с участием малых интерферирующих РНК (миРНК).

структуре генома [44]. В то же время сами МГЭ являются мишениями для эпигенетической регуляции как произошедшими от них микроРНК [15], так и не имеющими с ними эволюционного родства за счет частичной комплементарности последовательностей. Например, микроРНК let-7 ингибирует экспрессию различных LINE1 за счет связывания с продуктом транскрипции ORF2p их генов, подавляя трансляцию на рибосомах [45].

*Влияние произошедших от транспозонов микроРНК, ассоциированных с механизмами старения, на развитие остеоартрита*

Вышеописанные механизмы воздействия активированных МГЭ на регуляторные эффекты произошедших от них микроРНК позволяют предположить, что дисрегуляция МГЭ при старении оказывает влияние на такие микроРНК, вовлеченные в патогенез ОА. Согласно данным научной литературы, 23 произошедшие от МГЭ микроРНК из базы данных MDTE DB [15] вовлечены в

механизмы старения и ОА (табл. 2). Действительно, у пациентов с ОА в макрофагах синовиальной жидкости пораженных суставов определена повышенная экспрессия miR-1246, произошедшей от LTR-ERVL [15]. Данная микроРНК подавляет экспрессию генов GSK3 $\beta$ , киназы гликогенсинтазы-3 $\beta$  и Axin2, Axis inhibition protein 2, белка ингибирования оси 2 у человека, способствуя активации путей Wnt/ $\beta$ -катенина и обусловленному этим воспалению [46]. Увеличение уровней miR-1246 выявлено также при старении фибробластов человека [47]. В суставах при ОА повышенная экспрессия miR-1271 [48], произошедшая от LINE2 [15], вызывает апоптоз хондроцитов за счет ингибирования митоген-активируемой протеинкиназы MAPK [49]. В экзосомах пациентов с ОА определена повышенная экспрессия miR-1290 [50], которая ингибирует ген CCNG2, кодирующий циклин, регулирующий клеточный цикл [51]. Высокий уровень miR-1271 и miR-1290 определен также при старении фибробластов человека [47].

Таблица 2. Особенности экспрессии микроРНК, произошедших от транспозонов, при старении и ОА

МикроРНК (транспозон-источник)	Механизм действия микроРНК при остеоартрите	Изменение экспрессии при старении (повышение – ↑; снижение – ↓) [автор]	Изменение экспрессии при остеоартрите (повышение – ↑; снижение – ↓) [автор]
miR-1246 (LTR-ERVL)	подавляет экспрессию GSK3 $\beta$ и Axin2, способствуя активации путей Wnt/β-катенина и воспалению [46]	↑ [47]	↑ [46]
miR-1271 (LINE2)	ингибитирует MAPK [49]	↑ [47]	↑ [48, 49]
miR-1290 (SINE/MIR)	подавляет CCNG2 [51]	↑ [47]	↑ [50]
miR-151a (LINE2)	ингибитирует экспрессию агматиназы, нарушая метаболизм полиаминов [49]	↓ [52]	↓ [5]
miR-192 (LINE2)	повреждает хондроциты в ответ на липополисахариды, вызывая воспаление [54]	↑ [55]	↑ [54]
miR-211 (LINE2)	подавляет экспрессию фибулина-4 и провоспалительных цитокинов [57]	↓ [56]	↓ [57]
miR-224 (MER-135)	ингибитирует экспрессию провоспалительного хемокина CCL1 [58]	↓ [60]	↓ [58]
miR-28 (LINE2)	ингибитирует экспрессию IL-34 [63]	↓ [62]	↓ [61]
miR-31 (LINE2)	подавляет мРНК генов <i>PAPOLG</i> , <i>SP1</i> , <i>SRC</i> , <i>ZC3H12C</i> [64]	↑ [65]	↑ [64]
miR-320b (LINE2)	участвует в генных сетях регуляции апоптоза с участием <i>YWHAZ</i> , <i>YWHAQ</i> , <i>YWHAH</i> , <i>YWAHE</i> , <i>YWHAB</i> , <i>YWHAG</i> , <i>SFN</i> [66]	↑ [67]	↑ [66]
miR-326 (hAT-Tip100)	ингибитирует SIRT1 с активацией воспаления и ангиогенеза [68]	↑ [70]	↑ [68]
miR-335 (SINE/MIR)	ингибитирует экспрессию генов <i>MMP13</i> , <i>VCAM1</i> [71], подавляет энхондральную оссификацию хряща [72]	↑ [73]	↑ [71, 72]
miR-340 (DNA-TE/TcMar)	воздействует на мРНК генов <i>YTHDF3</i> , <i>IGF2BP3</i> , ингибитирует сигнал ERK посредством подавления <i>FMOD</i> [74]	↓ [75]	↓ [74]
miR-374 (LINE2)	препятствует индуцированному липополисахаридами разрушению хряща за счет ингибиования Wnt5b [77]	↓ [76]	↓ [77]

Таблица 2. Окончание

МикроРНК (транспозон-источник)	Механизм действия микроРНК при остеоартрите	Изменение экспрессии при старении (повышение – ↑; снижение – ↓) [автор]	Изменение экспрессии при остеоартрите (повышение – ↑; снижение – ↓) [автор]
miR-378a (SINE/MIR)	ингибитирует мРНК генов <i>Sox6</i> и <i>Atg2a</i> [78]	↑ [79]	↑ [78]
miR-384 (LINE-Dong-R4)	подавляет экспрессию <i>S<sub>OX9</sub></i> и сигналинг NF-кБ, препятствуя пролиферации хрящевых клеток [80]	↑ [81]	↑ [80]
miR-421 (LINE2)	препятствует вызванному IL-1 $\beta$ апоптозу и воспалению [82]	↓ [83]	↓ [82]
miR-450b (LINE1)	регулирует экспрессию гена <i>SKAP2</i> в макрофагах [84]	↓ [85]	↓ [84]
miR-487b (SINE/MIR)	целевое ингибирирование Wnt5a-пути [92]	↑ [92]	↓ [93, 94]
miR-495 (ERV-L/MaLR)	ингибитирует мРНК гена <i>AKT1</i> , с подавлением p-S6, p-mTOR и пролиферации хондроцитов [86]	↑ [87]	↑ [86]
miR-576 (LINE1)	целевое ингибирирование Wnt5a-пути [97]	↑ [96]	↓ [95]
miR-708 (LINE2)	связывается с мРНК гена <i>SATB2</i> , ингибируя регенерацию хряща [88]	↑ [89]	↑ [88]
miR-885 (SINE/MIR)	подавляет экспрессию генов <i>IGF1R</i> , <i>CTNNB1</i> , <i>OXR1</i> [90]	↑ [91]	↑ [90]

Проведенный в 2021 г. комплексный анализ регуляторной сети развития ОА показал снижение экспрессии miR-151a [5], которая произошла от LINE2 [15]. Уровень данной микроРНК в сыворотке крови значительно выше у пожилых людей по сравнению с молодыми [52]. Прямой мишенью miR-151a является мРНК гена *AGMAT*, кодирующего агматиназу – ключевой фермент метаболизма агматина, который выполняет роль нейротрансмиттера. Соответственно, подавление экспрессии агматиназы под влиянием miR-151a может стать причиной нарушений иннервации тканей и органов при старении, в том числе и суставов при ОА [53]. Уровень miR-192, возникшей от LINE2 [15], повышается при ОА и оказывает повреждающее действие на хондроциты за счет ингибирирования экспрессии гена *GDF11*, кодирующего секретируемый лиганд суперсемейства трансформирующих факторов роста бета, рекрутирующего транскрипционные факторы SMAD, необходимые для

роста и размножения клеток [54]. При старении в тканях почек также повышается уровень miR-192 [55]. Низкая экспрессия miR-211, произошедшей от LINE2 [15], ассоциирована с короткой продолжительностью жизни, что свидетельствует о ее прогностическом действии в отношении старения [56]. При ОА уровень miR-211 также снижен. MiR-211 способствует дифференцировке хондроцитов путем подавления экспрессии гена *EFEMP2*, EGF containing fibulin extracellular matrix protein 2, который кодирует фибулин-4, препятствующий выработке провоспалительных цитокинов и разрушающих хрящ протеиназ [57].

MiR-224, произошедшая от ДНК-транспозона MER-135 [15], ингибитирует экспрессию провоспалительного хемокина CCL1, препятствуя также деградации хряща [58]. Были разработаны наночастицы с miR-224 в качестве перспективного метода лечения ОА [59]. MiR-224 ассоциирована со

старением головного мозга. Ее мишенью является ген *CHOP*, C/EBP homologous protein, который участвует в регуляции митохондриальных белков [60]. В сыворотке крови больных ОА определен сниженный уровень miR-28 [61]. При физиологическом старении также уменьшена экспрессия miR-28 [62]. Мишенью miR-28 является мРНК гена *IL-34*, интерлейкина-34. Соответственно низкий уровень miR-28 при старении и ОА способствует воспалению в связи с усилением выработки IL-34 [63].

При тяжелом ОА определена повышенная экспрессия miR-31, мишенью которой являются мРНК генов: *PAPOLG*, кодирующего полиадениловую полимеразу; *SP1*, specificity protein 1, кодирующего транскрипционный фактор цинкового пальца; *SRC*, кодирующегоprotoонкоген нерекцепторной тирозинкиназы; *ZC3H12C*, кодирующего эндорибонуклеазу. Продуктом гена *SP1* является транскрипционный фактор, играющий важную роль в дифференцировке костей и хондроцитов, регуляции роста клеток [64]. Повышенный уровень miR-31 определен при старении эндотелиальных клеток человека [65]. Произошедшая от LINE2 miR-320b [15] ассоциирована с быстрым прогрессированием ОА и была предложена в качестве прогностического биомаркера болезни [66]. Данная микроРНК вовлечена в генные сети с участием генов – регуляторов апоптоза. К ним относятся гены путей сигнальной трансдукции: *YWHAZ*, 14-3-3 protein zeta/delta; *YWHAQ*, 14-3-3 protein theta; *YWHAH*, 14-3-3 protein eta; *YWAHE*, 14-3-3 protein epsilon; *YWHAB*, 14-3-3 protein beta/alpha; *YWHAG*, 14-3-3 protein gamma и *SFN*, супрессора опухоли стратифина [66]. Повышенная экспрессия miR-320b ассоциирована со старением фибробластов человека [67]. Возникшая от ДНК-транспозона hAT-Tip100 miR-326 [15] способствует развитию ОА путем ингибирования экспрессии гена *SIRT1*, сиртуина-1, НАД-зависимой деацетилазы и стимуляции *VEGF*, сосудистого эндотелиального фактора роста с активацией воспаления и ангиогенеза [68]. Уровень miR-326 значительно повышен у больных ревматоидным артритом с положительным ревматоидным фактором [69]. Экспрессия miR-326 повышена в фибробластах кожи при старении [70].

При ОА определен повышенный уровень miR-335 [71], которая подавляет энхондральную оссификацию суставного хряща [72] путем ингибирования генов: *MMP13*, matrix metallopeptidase 13, кодирующего матриксную металлопептидазу, участвующую в деградации хряща; *ICAM1*, vascular cell adhesion molecule 1, кодирующего белок суперсемейства иммуноглобулинов, участвующий в регуляции адгезии лейкоцитов [71]. Высокий уровень miR-335 определен при старении в целом, а также при старении культуры астроцитов человека и гиппокампа головного мозга мышей по сравнению с

молодыми клетками и гиппокампом молодых мышей и ведет к нарушению консолидации памяти в гиппокампе головного мозга за счет ингибирования мРНК гена *PSD95*, который кодирует белок постсинаптической плотности [73]. При ОА снижение экспрессии произошедшей от ДНК-транспозона TcMar miR-340 [15] активирует гены *YTHDF3* (кодирует РНК-связывающий белок), *IGF2BP3* (ген мРНК-связывающего белка инсулиноподобного фактора роста), *FMOD* (ген интерстициального протеогликана фибромодулина) и сигнал ERK, extracellular signal-regulated kinase, способствующий пролиферации, подвижности и выживанию клеток [74]. При старении уровень miR-340 в сыворотке крови снижается [75].

Уровень miR-374, произошедшей от LINE2 [15], снижен при старении [76], а также в хрящевой ткани пораженных суставов при ОА [77]. MiR-374 препятствует индуцированному липополисахаридами разрушению хряща за счет ингибирования *Wnt5b*. Название гена образовано из слов Wingless и Int-1. Ген *Wnt5b* кодирует белки семейства сигнальных белков WNT, передающих сигналы в клетку через рецепторы клеточной поверхности, стимулируя пролиферацию и дифференцировку клеток, в том числе физиологическую регенерацию хондроцитов. Соответственно, ингибирование *Wnt5b* ведет к деградации хряща [77]. У пациентов с ОА в синовиальной оболочке пораженных суставов экспрессируется на высоком уровне miR-378, особенно на поздней стадии заболевания. Мишенью miR-378 являются мРНК генов *Sox6*, sex determining region Y-box 6, кодирующего регулятор хондрогенеза, и *Atg2a*, autophagy related 2A, кодирующего связанный с аутофагией белок [78]. На моделях мышей внутрисуставные инъекции анти-miR-378 лентивируса замедляли прогрессирование ОА, способствуя регенерации и подавлению патологической гипертрофии [78]. При старении тимуса также выявлено повышение уровней miR-378 [79]. MiR-384, возникшая в эволюции от LINE-DONG-R4 [15], подавляет экспрессию генов *Sox9*, SRY-box transcription factor 9, и сигнал NF-кВ, nuclear factor kappa B, препятствуя пролиферации хрящевых клеток. Белок *Sox9* регулирует транскрипцию гена антиミュллерова гормона во время дифференцировки хондроцитов. NF-кВ является регулятором транскрипции, стимулирующим экспрессию генов, участвующих в иммунных реакциях, в том числе регуляцию воспаления [80]. MiR-384 негативно регулирует возрастную остеогенную дифференцировку мезенхимальных стволовых клеток костного мозга, способствуя старению [81].

Экспрессия произошедшей от LINE2 miR-421 [15] в хондроцитах пораженных ОА суставов снижена. Данная микроРНК препятствует вызванному IL-1 $\beta$  апоптозу и воспалению [82]. Старение также

ассоциировано с уменьшением уровней miR-421 [83]. В фибробластоподобных клетках при OA снижена экспрессия miR-450b, мишенью которой является мРНК гена *SKAP2*, кодирующий с Src-киназой фосфопротеин 2, который играет важную роль в сигнальных путях киназы в макрофагах [84]. Поскольку активация макрофагов при OA [16] способствует прогрессированию заболевания за счет продукции ими интерлейкинов IL-1 $\beta$ , IL-6, IL-12, TNF- $\alpha$ , активных форм кислорода и iNOS [18], нарушение сигнальных путей киназы в данных клетках имеет значение в патогенезе OA [84]. В экспериментах снижение miR-450b выявлено при старении фибробластов мышей. Фибробlastы являются необходимыми клеточными компонентами суставов, поэтому их старение способствует дегенеративным процессам и прогрессированию OA [85]. У больных OA определен повышенный уровень miR-495 в хрящевой ткани пораженных суставов [86]. MiR-495 ингибирует мРНК гена *AKT1*, RAC-alpha serine/threonine-protein kinase, кодирующую протеинкиназу регуляции роста и пролиферации клеток с опосредованным подавлением *p-S6*, ribosomal protein S6, p-mTOR, phosphorylated mammalian target of rapamycin и пролиферации клеток [86]. Данная микроРНК произошла от ERV-L/MaLR [15]. MiR-495 блокирует клеточный цикл в S-фазе и способствует апоптозу клеток, индуцируя старение мезенхимальных стволовых клеток человека [87].

При остеоартрите повышение уровня miR-708 вызывает ингибирование гена *SATB2*, special AT-rich sequence-binding protein 2, белковый продукт которого способствует регенерации хряща при OA. Поэтому снижение экспрессии *SATB2* под влиянием miR-708 вызывает нарушение регенерации и как результат – деградацию хряща [88]. Усиленная экспрессия miR-708 ассоциирована также со старением [89]. Эволюционным источником гена miR-708 является LINE2 [15]. Повышенная экспрессия miR-885, произошедшей от SINE/MIR [15], ассоциирована с OA [90] и старением. Мишенью miR-885 является мРНК гена *IGF1R*, insulin-like growth factor 1 receptor, вовлеченного в клеточную интернализацию IGF-1 и активацию сигнальных каскадов PI3K/Akt/GSK-3 $\beta$ , phosphoinositide 3-kinases/ AKT serine/threonine kinase/ glycogen synthase kinase-3 beta. MiR-885 нацелена также на мРНК генов: *CTNNB1*, catenin beta-1, регулятор канонической передачи сигналов Wnt; *MAN1C1*, mannosidase alpha class 1C member 1, белковый продукт которого участвует в N-гликозилировании белков; *OXR1*, oxidation resistance 1, кодирующего белок регуляции чувствительности к окислительному стрессу [91].

Для некоторых произошедших от МГЭ микроРНК определено противоположное изменение экспрессии идентичных микроРНК при старении

и OA, что свидетельствует о том, что не все механизмы старения пересекаются с этиопатогенезом OA, но изменение МГЭ влияет на нарушение экспрессии микроРНК, содержащих идентичные последовательности. Так, произошедшая от SINE/MIR miR-487b [15], являющаяся прямой мишенью длинной нкРНК MAR1 (muscle anabolic regulator 1), оказывает целевое ингибирование мРНК гена *Wnt5a*, что ведет к подавлению путей регуляции миогенеза, способствуя старению мышечной ткани [92]. Активация путей *Wnt5a* при OA способствует прогрессированию заболевания в связи с низкой экспрессией miR-487b [93], которая способствует хондрогенной дифференцировке мезенхимальных стволовых клеток [94]. Сходные изменения определены в отношении экспрессии произошедшей от LINE1 [15] miR-576, уровень которой снижен в хондроцитах при OA по сравнению с нормой [95]. У пожилых людей определена ассоциация повышенной экспрессии miR-576 с гериатрическим синдромом дряхлости (прогрессирующее ухудшение физического здоровья) [96]. Мишенью miR-576 также является *Wnt5a* [97]. Таким образом, анализ научной литературы позволил выявить 23 микроРНК, произошедших от МГЭ и участвующих в патогенезе OA (см. табл. 2).

## ЗАКЛЮЧЕНИЕ

Предложена гипотеза, согласно которой патологическая активация МГЭ при старении способствует развитию OA при наличии индивидуальных особенностей SNP в межгенных и инtronных областях, где расположены гены МГЭ. Этим можно объяснить значительную распространенность OA, увеличивающуюся с возрастом, а также влияние факторов внешней среды на развитие заболевания, поскольку МГЭ являются высокочувствительными сенсорами генома к стрессовым воздействиям. При старении в организме происходят воспалительно-дегенеративные процессы, обусловленные активацией интерферонового ответа на продукты экспрессии МГЭ. Сходные механизмы описаны в патогенезе OA. Кроме того, произошедшие от МГЭ микроРНК, характеризующиеся взаиморегуляцией со своими эволюционными источниками за счет комплементарности нуклеотидных последовательностей, оказались вовлечеными в развитие OA. Было выявлено 23 таких микроРНК, участие которых в патогенезе OA обусловлено ингибированием экспрессии генов, вовлеченных в иммунные, воспалительные и дегенеративные процессы. В перспективе возможно использование таких микроРНК для таргетной терапии OA.

Исследование проведено без внешнего финансирования.

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

Авторы заявляют, что у них нет конфликта интересов.

### СПИСОК ЛИТЕРАТУРЫ

1. *GBD 2021 Osteoarthritis Collaborators.* Global, regional, and national burden of osteoarthritis, 1990–2020 and projections to 2050: A systematic analysis for the Global Burden of Disease Study 2021 // Lancet Rheumatol. 2023. V. 5. e508–e522. [https://doi.org/10.1016/S2665-9913\(23\)00163-7](https://doi.org/10.1016/S2665-9913(23)00163-7)
2. *Boer C.G., Hatzikotoulas K., Southam L. et al.* Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations // Cell. 2021. V. 184. P. 4784–4818.e17. <https://doi.org/10.1016/j.cell.2021.07.038>
3. *Faber B.G., Frysz M., Boer C.G. et al.* The identification of distinct protective and susceptibility mechanisms for hip osteoarthritis: Findings from a genome-wide association study meta-analysis of minimum joint space width and Mendelian randomisation cluster analyses // EBioMedicine. 2023. V. 95. <https://doi.org/10.1016/j.ebiom.2023.104759>
4. *Chen X., Wu Q., Cao X. et al.* Menthone inhibits type-I interferon signaling by promoting Tyk2 ubiquitination to relieve local inflammation of rheumatoid arthritis // Int. Immunopharmacol. 2022. V. 112. <https://doi.org/10.1016/j.intimp.2022.109228>
5. *Jiang Y., Shen Y., Ding L. et al.* Identification of transcription factors and construction of a novel miRNA regulatory network in primary osteoarthritis by integrated analysis // BMC Musculoskelet. Disord. 2021. V. 22. P. 1008. <https://doi.org/10.1186/s12891-021-04894-2>
6. *Allen K.D., Thoma L.M., Golightly Y.M.* Epidemiology of osteoarthritis // Osteoarthritis Cartilage. 2022. V. 30. P. 184–195. <https://doi.org/10.1016/j.joca.2021.04.020>
7. *Vos T., Flaxman A.D., Naghavi M. et al.* Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: A systematic analysis for the Global Burden of Disease Study 2010 // Lancet. 2012. V. 380. P. 2163–2196. [https://doi.org/10.1016/S0140-6736\(12\)61729-2](https://doi.org/10.1016/S0140-6736(12)61729-2)
8. *De Cecco M., Ito T., Petrushen A.P. et al.* L1 drives IFN in senescent cells and promotes age-associated inflammation // Nature. 2019. V. 566. P. 73–78.
9. *Gorbunova V., Seluanov A., Mita P. et al.* The role of retrotransposable elements in ageing and age-associated diseases // Nature. 2021. V. 596. P. 43–53. <https://doi.org/10.1038/s41586-021-03542-y>
10. *Bendiksen S., Martinez-Zubiavrra I., Tümmeler C. et al.* Human endogenous retrovirus W activity in cartilage of osteoarthritis patients // Biomed. Res. Int. 2014. V. 2014. <https://doi.org/10.1155/2014/698609>
11. *Teerawattanapong N., Udomsinprasert W., Ngarmukos S. et al.* Blood leukocyte LINE-1 hypomethylation and oxidative stress in knee osteoarthritis // Heliyon. 2019. V. 5. <https://doi.org/10.1016/j.heliyon.2019.e01774>
12. *Lee D.H., Bae W.H., Ha H. et al.* The human PTGR1 gene expression is controlled by TE-derived Z-DNA forming sequence cooperating with miR-6867-5p // Sci. Rep. 2024. V. 14. P. 4723. <https://doi.org/10.1038/s41598-024-55332-x>
13. *Conley A.B., Jordan I.K.* Cell type-specific termination of transcription by transposable element sequences // Mob. DNA. 2012. V. 3. P. 15. <https://doi.org/10.1186/1759-8753-3-15>
14. *Daniel C., Behm M., Öhman M.* The role of Alu elements in the cis-regulation of RNA processing // Cell. Mol. Life Sci. 2015. V. 72. P. 4063–4076. <https://doi.org/10.1007/s00018-015-1990-3>
15. *Wei G., Qin S., Li W. et al.* MDTE DB: A database for microRNAs derived from Transposable element // IEEE/ACM Trans. Comput. Biol. Bioinform. 2016. V. 13. P. 1155–1160. <https://doi.org/10.1109/TCBB.2015.2511767>
16. *Chen J., Chen S., Cai D. et al.* The role of Sirt6 in osteoarthritis and its effect on macrophage polarization // Bioengineered. 2022. V. 13. P. 9677–9689. <https://doi.org/10.1080/21655979.2022.2059610>
17. *Van Meter M., Kashyap M., Rezazadeh S. et al.* SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age // Nat. Commun. 2014. V. 5. P. 5011. <https://doi.org/10.1038/ncomms6011>
18. *Zhou F., Mei J., Han X. et al.* Kinsenoside attenuates osteoarthritis by repolarizing macrophages through inactivating NF- $\kappa$  B/MAPK signaling and protecting chondrocytes // Acta. Pharm. Sin. B. 2019. V. 9. P. 973–985. <https://doi.org/10.1016/j.apsb.2019.01.015>
19. *Saetan N., Honsawek S., Tanavalee S. et al.* Association of plasma and synovial fluid interferon- $\gamma$  inducible protein-10 with radiographic severity in knee osteoarthritis // Clin. Biochem. 2011. V. 44. P. 1218–1222. <https://doi.org/10.1016/j.clinbiochem.2011.07.010>
20. *Мустафин Р.Н., Хуснутдинова Э.К.* Некодирующие части генома как основа эпигенетической наследственности // Вавил. журн. генетики и селекции. 2017. V. 21. P. 742–749.
21. *Lu F., Liu P., Zhang Q. et al.* Association between the polymorphism of IL-17A and IL-17F gene with knee osteoarthritis risk: A meta-analysis based on case-control studies // J. Orthop. Surg. Res. 2019. V. 14. P. 445. <https://doi.org/10.1186/s13018-019-1495-0>

22. *Budhiparama N.C., Lumban-Gaol I., Sudoyo H.* Interleukin-1 genetic polymorphisms in knee osteoarthritis: What do we know? A meta-analysis and systematic review // *J. Orthop. Surg. (Hong Kong)*. 2022. V. 30. <https://doi.org/10.1177/23094990221076652>
23. *Deng X., Ye K., Tang J., Huang Y.* Association of *rs1800795* and *rs1800796* polymorphisms in *interleukin-6* gene and osteoarthritis risk: Evidence from a meta-analysis // *Nucleosides Nucleotides Nucleic Acids*. 2023. V. 42. P. 328–342. <https://doi.org/10.1080/15257770.2022.2147541>
24. *Rodriguez-Fontenla C., Calaza M., Evangelou E. et al.* Assessment of osteoarthritis candidate genes in a meta-analysis of nine genome-wide association studies // *Arthritis Rheumatol.* 2014. V. 66. P. 940–949. <https://doi.org/10.1002/art.38300>
25. *Liu Y., Lu T., Liu Z. et al.* Six macrophage-associated genes in synovium constitute a novel diagnostic signature for osteoarthritis // *Front. Immunol.* 2022. V. 13. <https://doi.org/10.3389/fimmu.2022.936606>
26. *Yang L., Chen Z., Guo H. et al.* Extensive cytokine analysis in synovial fluid of osteoarthritis patients // *Cytokine*. 2021. V. 143. <https://doi.org/10.1016/j.cyto.2021.155546>
27. *Pan L., Yang F., Cao X. et al.* Identification of five hub immune genes and characterization of two immune subtypes of osteoarthritis // *Front. Endocrinol (Lausanne)*. 2023. V. 14. <https://doi.org/10.3389/fendo.2023.1144258>
28. *Xu J., Chen K., Yu Y. et al.* Identification of immune-related risk genes in osteoarthritis based on bioinformatics analysis and machine learning // *J. Pers. Med.* 2023. V. 13. P. 367. <https://doi.org/10.3390/jpm13020367>
29. *Cheng P., Gong S., Guo C. et al.* Exploration of effective biomarkers and infiltrating Immune cells in Osteoarthritis based on bioinformatics analysis // *Artif. Cells. Nanomed. Biotechnol.* 2023. V. 51. P. 242–254. <https://doi.org/10.1080/21691401.2023.2185627>
30. *Li J., Wang G., Xv X. et al.* Identification of immune-associated genes in diagnosing osteoarthritis with metabolic syndrome by integrated bioinformatics analysis and machine learning // *Front. Immunol.* 2023. V. 14. <https://doi.org/10.3389/fimmu.2023.1134412>
31. *Grandi F.C., Bhutani N.* Epigenetic therapies for osteoarthritis // *Trends. Pharmacol. Sci.* 2020. V. 41. P. 557–569. <https://doi.org/10.1016/j.tips.2020.05.008>
32. *Knights A.J., Redding S.J., Maerz T.* Inflammation in osteoarthritis: The latest progress and ongoing challenges // *Curr. Opin. Rheumatol.* 2023. V. 35. P. 128–134.
33. *Zhang J., Zhang S., Zhou Y. et al.* KLF9 and EPYC acting as feature genes for osteoarthritis and their association with immune infiltration // *J. Orthop. Surg. Res.* 2022. V. 17. P. 365. <https://doi.org/10.1186/s13018-022-03247-6>
34. *Zhang Q., Sun C., Liu X. et al.* Mechanism of immune infiltration in synovial tissue of osteoarthritis: A gene expression-based study // *J. Orthop. Surg. Res.* 2023. V. 18. P. 58. <https://doi.org/10.1186/s13018-023-03541-x>
35. *Xia D., Wang J., Yang S. et al.* Identification of key genes and their correlation with immune infiltration in osteoarthritis using integrative bioinformatics approaches and machine-learning strategies // *Medicine (Baltimore)*. 2023. V. 102. <https://doi.org/10.1097/MD.00000000000035355>
36. *Xu L., Wang Z., Wang G.* Screening of biomarkers associated with osteoarthritis aging genes and immune correlation studies // *Int. J. Gen. Med.* 2024. V. 17. P. 205–224. <https://doi.org/10.2147/IJGM.S447035>
37. *Cornec A., Poirier E.Z.* Interplay between RNA interference and transposable elements in mammals // *Front. Immunol.* 2023. V. 14. <https://doi.org/10.3389/fimmu.2023.1212086>
38. *Cho J., Paszkowski J.* Regulation of rice root development by a retrotransposon acting as a microRNA sponge // *eLife*. 2017. V. 6. <https://doi.org/10.7554/eLife.30038>
39. *Lu X., Sachs F., Ramsay L. et al.* The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity // *Nat. Struct. Mol. Biol.* 2014. V. 21. P. 423–425. <https://doi.org/https://doi.org/10.1038/nsmb.2799>
40. *Honson D.D., Macfarlan T.S.* A lncRNA-like role for LINE1s in development // *Dev. Cell.* 2018. V. 46. P. 132–134. <https://doi.org/https://doi.org/10.1016/j.devcel.2018.06.022>
41. *Playfoot C.J., Sheppard S., Planet E., Trono D.* Transposable elements contribute to the spatiotemporal microRNA landscape in human brain development // *RNA*. 2022. V. 28. P. 1157–1171. <https://doi.org/10.1261/rna.079100.122>
42. *McCue A.D., Nuthikattu S., Slotkin R.K.* Genome-wide identification of genes regulated in trans by transposable element small interfering RNAs // *RNA Biol.* 2013. V. 10. P. 1379–1395. <https://doi.org/10.4161/rna.25555>
43. *Lee D.H., Bae W.H., Ha H. et al.* Z-DNA-containing long terminal repeats of human endogenous retrovirus families provide alternative promoters for human functional genes // *Mol. Cells.* 2022. V. 45. P. 522–530. <https://doi.org/10.14348/molcells.2022.0060>
44. *Chalertpet K., Pin-On P., Aporntewan C. et al.* Argonaute 4 as a effector protein in RNA-directed DNA methylation in human cells // *Front. Genet.*

2019. V. 10. P. 645.  
<https://doi.org/10.3389/fgene.2019.00645>
45. *Tristán-Ramos P., Rubio-Roldan A., Peris G. et al.* The tumor suppressor microRNA let-7 inhibits human LINE-1 retrotransposition // *Nat. Commun.* 2020. V. 11. P. 5712.  
<https://doi.org/10.1038/s41467-020-19430-4>
46. *Peng S., Yan Y., Li R. et al.* Extracellular vesicles from M1-polarized macrophages promote inflammation in the temporomandibular joint via miR-1246 activation of the Wnt/β-catenin pathway // *Ann. N. Y. Acad. Sci.* 2021. V. 1503. P. 48–59.  
<https://doi.org/10.1111/nyas.14590>
47. *Dhahbi J.M., Atamna H., Boffelli D. et al.* Deep sequencing reveals novel microRNAs and regulation of microRNA expression during cell senescence // *PLoS One.* 2011. V. 6.  
<https://doi.org/10.1371/journal.pone.0020509>
48. *Lu M.Y., Yang Y.H., Wu X. et al.* Effect of needle-knife on chondrocyte apoptosis of knee joint in rabbits with knee osteoarthritis based on CircSERPINE2-miR-1271-5P-ERG axis // *Zhongguo Zhen Jiu.* 2023. V. 43. P. 447–453.  
<https://doi.org/10.13703/j.0255-2930.20220411-k0001>
49. *Xie W.P., Ma T., Liang Y.C. et al.* Cangxi Tongbi Capsules promote chondrocyte autophagy by regulating circRNA\_0008365/miR-1271/p38 MAPK pathway to inhibit development of knee osteoarthritis // *Zhongguo Zhong Yao Za Zhi.* 2023. V. 48. P. 4843–4851.  
<https://doi.org/10.19540/j.cnki.cjcm.20230510.708>
50. *Ju C., Liu R., Zhang Y. et al.* Exosomes may be the potential new direction of research in osteoarthritis management // *Biomed. Res. Int.* 2019. V. 3.  
<https://doi.org/10.1155/2019/7695768>
51. *Qin W.J., Wang W.P., Wang X.B. et al.* MiR-1290 targets CCNG2 to promote the metastasis of oral squamous cell carcinoma // *Eur. Rev. Med. Pharmacol. Sci.* 2019. V. 23. P. 10332–10342.  
[https://doi.org/10.26355/eurrev\\_201912\\_19671](https://doi.org/10.26355/eurrev_201912_19671)
52. *Noren Hooten N., Fitzpatrick M., Wood W.H. 3<sup>rd</sup> et al.* Age-related changes in microRNA levels in serum // *Aging (Albany NY).* 2013. V. 5. P. 725–740.  
<https://doi.org/10.18632/aging.100603>
53. *Xie Y., Zhang Y., Liu X. et al.* miR-151-5p promotes the proliferation and metastasis of colorectal carcinoma cells by targeting AGMAT // *Oncol. Rep.* 2023. V. 49. P. 50.  
<https://doi.org/10.3892/or.2023.8487>
54. *Wang Y., Yu C., Zhang H.* Lipopolysaccharides-mediated injury to chondrogenic ATDC5 cells can be relieved by Sinomenine via downregulating microRNA-192 // *Phytother. Res.* 2019. V. 33. P. 1827–1836.  
<https://doi.org/10.1002/ptr.6372>
55. *Sataranatarajan K., Feliers D., Mariappan M.M. et al.* Molecular events in matrix protein metabolism in the aging kidney // *Aging Cell.* 2012. V. 11. P. 1065–1073.  
<https://doi.org/10.1111/acel.12008>
56. *Smith-Vikos T., Liu Z., Parsons C.* A serum miRNA profile of human longevity: Findings from the Baltimore Longitudinal Study of Aging (BLSA) // *Aging (Albany NY).* 2016. V. 8. P. 2971–2987.  
<https://doi.org/10.18632/aging.101106>
57. *Liu H., Luo J.* miR-211-5p contributes to chondrocyte differentiation by suppressing Fibulin-4 expression to play a role in osteoarthritis // *J. Biochem.* 2019. V. 166. P. 495–502.  
<https://doi.org/10.1093/jb/mvz065>
58. *Liu Y., Zhang Y.* Hsa\_circ\_0134111 promotes osteoarthritis progression by regulating miR-224-5p/CCL1 interaction // *Aging (Albany NY).* 2021. V. 13. P. 20383–20394.  
<https://doi.org/10.18632/aging.203420>
59. *Chen H., Chen F., Hu F. et al.* MicroRNA-224-5p nanoparticles balance homeostasis via inhibiting cartilage degeneration and synovial inflammation for synergistic alleviation of osteoarthritis // *Acta Biomater.* 2023. V. 167. P. 401–415.  
<https://doi.org/10.1016/j.actbio.2023.06.010>
60. *Francisco S., Martinho V., Ferreira M. et al.* The role of microRNAs in proteostasis decline and protein aggregation during brain and skeletal muscle aging // *Int. J. Mol. Sci.* 2022. V. 23. P. 3232.  
<https://doi.org/10.3390/ijms23063232>
61. *Beyer C., Zampetaki A., Lin N.Y. et al.* Signature of circulating microRNAs in osteoarthritis // *Ann. Rheum. Dis.* 2015. V. 74. e18.  
<https://doi.org/10.1136/annrheumdis-2013-204698>
62. *Morsiani C., Bacalini M.G., Collura S. et al.* Blood circulating miR-28-5p and let-7d-5p associate with premature ageing in Down syndrome // *Mech. Ageing Dev.* 2022. V. 206.  
<https://doi.org/10.1016/j.mad.2022.111691>
63. *Zhou S.L., Hu Z.Q., Zhou Z.J. et al.* miR-28-5p-IL-34-macrophage feedback loop modulates hepatocellular carcinoma metastasis // *Hepatology.* 2016. V. 63. P. 1560–1575.  
<https://doi.org/10.1002/hep.28445>
64. *Costa V., De Fine M., Carina V. et al.* How miR-31-5p and miR-33a-5p regulates SP1/CX43 expression in osteoarthritis disease: preliminary insights // *Int. J. Mol. Sci.* 2021. V. 22. P. 2471.  
<https://doi.org/10.3390/ijms22052471>
65. *Dellago H., Preschitz-Kammerhofer B., Terlecki-Zaniewicz L. et al.* High levels of oncomiR-21 contribute to the senescence-induced growth arrest in normal human cells and its knock-down increases the replicative lifespan // *Aging Cell.* 2013. V. 12. P. 446–458.  
<https://doi.org/10.1111/acel.12069>
66. *Ali S.A., Espin-Garcia O., Wong A.K. et al.* Circulating microRNAs differentiate fast-progressing from slow-progressing and non-progressing knee osteoarthritis

- in the Osteoarthritis Initiative cohort // *Ther. Adv. Musculoskelet. Dis.* 2022. V. 14. <https://doi.org/10.1177/1759720X221082917>
67. *Dalmasso B., Hatse S., Brouwers B. et al.* Age-related microRNAs in older breast cancer patients: biomarker potential and evolution during adjuvant chemotherapy // *BMC Cancer.* 2018. V. 18. P. 1014. <https://doi.org/10.1186/s12885-018-4920-6>
68. *Lin Z., Ma Y., Zhu X. et al.* Potential predictive and therapeutic applications of small extracellular vesicles-derived circPARD3B in osteoarthritis // *Front. Pharmacol.* 2022. V. 13. <https://doi.org/10.3389/fphar.2022.968776>
69. *Paradowska-Gorycka A., Wajda A., Rzeszotarska E. et al.* miR-10 and Its negative correlation with serum IL-35 concentration and positive correlation with STAT5a expression in patients with rheumatoid arthritis // *Int. J. Mol. Sci.* 2022. V. 23. P. 7925. <https://doi.org/10.3390/ijms23147925>
70. *Yang X., Tan J., Shen J. et al.* Endothelial cell-derived extracellular vesicles target TLR4 via miRNA-326-3p to regulate skin fibroblasts senescence // *J. Immunol. Res.* 2022. V. 2022. P. 3371982. <https://doi.org/10.1155/2022/3371982>
71. *Wilson T.G., Baghel M., Kaur N. et al.* Characterization of miR-335-5p and miR-335-3p in human osteoarthritic tissues // *Arthritis Res. Ther.* 2023. V. 25. P. 105. <https://doi.org/10.1186/s13075-023-03088-6>
72. *Xia S., Zhao J., Zhang D. et al.* MiR-335-5p inhibits endochondral ossification by directly targeting SP1 in TMJ OA // *Oral Dis.* 2023. V. 20. <https://doi.org/10.1111/odi.14736>
73. *Raihan O., Brishti A., Molla M.R. et al.* The age-dependent elevation of miR-335-3p leads to reduced cholesterol and impaired memory in brain // *Neuroscience.* 2018. V. 390. P. 160–173. <https://doi.org/10.1016/j.neuroscience.2018.08.003>
74. *Duan Y., Yu C., Yan M. et al.* m6A regulator-mediated RNA methylation modification patterns regulate the immune microenvironment in osteoarthritis // *Front. Genet.* 2022. V. 13. <https://doi.org/fgene.2022.921256>
75. *Zhang H., Yang H., Zhang C. et al.* Investigation of microRNA expression in human serum during the aging process // *J. Gerontol. A. Biol. Sci. Med. Sci.* 2015. V. 70. P. 102–109. <https://doi.org/10.1093/gerona/glu145>
76. *ElSharawy A., Keller A., Flachsbart F. et al.* Genome-wide miRNA signatures of human longevity // *Aging Cell.* 2012. V. 11. P. 607–616. <https://doi.org/10.1111/j.1474-9726.2012.00824.x>
77. *Shi F.L., Ren L.X.* Up-regulated miR-374a-3p relieves lipopolysaccharides induced injury in CHON-001 cells via regulating Wingless-type MMTV integration site family member 5B // *Mol. Cell. Probes.* 2020. V. 51. <https://doi.org/10.1016/j.mcp.2020.101541>
78. *Feng L., Yang Z., Li Y. et al.* MicroRNA-378 contributes to osteoarthritis by regulating chondrocyte autophagy and bone marrow mesenchymal stem cell chondrogenesis // *Mol. Ther. Nucleic Acids.* 2022. V. 28. P. 328–341. <https://doi.org/10.1016/j.omtn.2022.03.016>
79. *Guo D., Ye Y., Qi J. et al.* Age and sex differences in microRNAs expression during the process of thymus aging // *Acta Biochim. Biophys. Sin. (Shanghai).* 2017. V. 49. P. 409–419. <https://doi.org/10.1093/abbs/gmx029>
80. *Zhang W., Cheng P., Hu W. et al.* Inhibition of microRNA-384-5p alleviates osteoarthritis through its effects on inhibiting apoptosis of cartilage cells via the NF- $\kappa$ B signaling pathway by targeting SOX9 // *Cancer Gene Ther.* 2018. V. 25. P. 326–338. <https://doi.org/10.1038/s41417-018-0029-y>
81. *Li X., Wu J., Zhang K. et al.* MiR-384-5p targets Gli2 and negatively regulates age-related osteogenic differentiation of rat bone marrow mesenchymal stem cells // *Stem. Cells Dev.* 2019. V. 28. P. 791–798. <https://doi.org/10.1089/scd.2019.0044>
82. *Zhang H., Xiang X., Zhou B. et al.* Circular RNA SLTM as a miR-421-competing endogenous RNA to mediate HMGB2 expression stimulates apoptosis and inflammation in arthritic chondrocytes // *J. Biochem. Mol. Toxicol.* 2023. V. 37. <https://doi.org/10.1002/jbt.23306>
83. *Li G., Song H., Chen L. et al.* TUG1 promotes lens epithelial cell apoptosis by regulating miR-421/caspase-3 axis in age-related cataract // *Exp. Cell. Res.* 2017. V. 356. P. 20–27. <https://doi.org/10.1016/j.yexcr.2017.04.002>
84. *Chen Y.J., Chang W.A., Wu L.Y. et al.* Identification of novel genes in osteoarthritic fibroblast-like synoviocytes using next-generation sequencing and bioinformatics approaches // *Int. J. Med. Sci.* 2019. V. 16. P. 1057–1071. <https://doi.org/10.7150/ijms.35611>
85. *Nidadavolu L.S., Niedernhofer L.J., Khan S.A.* Identification of microRNAs dysregulated in cellular senescence driven by endogenous genotoxic stress // *Aging (Albany NY).* 2013. V. 5. P. 460–473. <https://doi.org/10.1863/aging.100571>
86. *Zhao X., Wang T., Cai B. et al.* MicroRNA-495 enhances chondrocyte apoptosis, senescence and promotes the progression of osteoarthritis by targeting AKT1 // *Am. J. Transl. Res.* 2019. V. 11. P. 2232–2244.
87. *Li X., Song Y., Liu D. et al.* MiR-495 promotes senescence of mesenchymal stem cells by targeting Bmi-1 // *Cell Physiol. Biochem.* 2017. V. 42. P. 780–796. <https://doi.org/10.1159/000478069>
88. *Wang Y., Su Q., Tang H. et al.* Microfracture technique combined with mesenchymal stem cells inducer represses miR-708-5p to target special at-rich sequence-binding protein 2 to drive cartilage repair and regeneration in rabbit knee osteoarthritis //

- Growth Factors. 2023. V. 41. P. 115–129.  
<https://doi.org/10.1080/08977194.2023.2227269>
99. Lee B.P., Buric I., George-Pandeth A. et al. MicroRNAs miR-203-3p, miR-664-3p and miR-708-5p are associated with median strain lifespan in mice // Sci. Rep. 2017. V. 7.  
<https://doi.org/10.1038/srep44620>
90. Kwak Y.H., Kwak D.K., Moon H.S. et al. Significant changes in serum microRNAs after high tibial osteotomy in medial compartmental knee osteoarthritis: potential prognostic biomarkers // Diagnostics (Basel.). 2021. V. 11. P. 258.  
<https://doi.org/10.3390/diagnostics11020258>
91. Behbahanipour M., Peymani M., Salari M. et al. Expression profiling of blood microRNAs 885, 361, and 17 in the Patients with the Parkinson's disease: Integrating interaction data to uncover the possible triggering age-related mechanisms // Sci. Rep. 2019. V. 9. P. 13759.  
<https://doi.org/10.1038/s41598-019-50256-3>
92. Zhang Z.K., Li J., Guan D. et al. A newly identified lncRNA MaR1 acts as a miR-487b sponge to promote skeletal muscle differentiation and regeneration // J. Cachexia Sarcopenia Muscle. 2018. V. 9. P. 613–626.  
<https://doi.org/10.1002/jcsm.12281>
93. Chang L., Yao H., Yao Z. et al. Comprehensive analysis of key genes, signaling pathways and miRNAs in human knee osteoarthritis: based on bioinformatics // Front. Pharmacol. 2021. V. 12.  
<https://doi.org/10.3389/fphar.2021.730587>
94. Alizadeh A.H., Lively S., Lepage S. et al. MicroRNAs as prognostic markers for chondrogenic differentiation potential of equine mesenchymal stromal cells // Stem Cells Dev. 2023. V. 32. P. 693–702.  
<https://doi.org/10.1089/scd.2022.0295>
95. Díaz-Prado S., Cicione C., Muñoz-López E. et al. Characterization of microRNA expression profiles in normal and osteoarthritic human chondrocytes // BMC Musculoskelet. Disord. 2012. V. 13. P. 144.  
<https://doi.org/10.1186/1471-2474-13-144>
96. Ipson B.R., Fletcher M.B., Espinoza S.E., Fisher A.L. Identifying exosome-derived microRNAs as candidate biomarkers of frailty // J. Frailty Aging. 2018. V. 7. P. 100–103.  
<https://doi.org/10.14283/jfa.2017.45>
97. Luo J., Liu L., Shen J. et al. MiR-576-5p promotes epithelial-to-mesenchymal transition in colorectal cancer by targeting the Wnt5a-mediated Wnt/β-catenin signaling pathway // Mol. Med. Rep. 2021. V. 23. P. 94.  
<https://doi.org/10.3892/mmr.2020.11733>

## Relationship of microRNAs with Transposons in Osteoarthritis Development

R. N. Mustafin<sup>1, \*</sup>, E. K. Khusnutdinova<sup>2</sup>

<sup>1</sup>Bashkir State Medical University, Ufa, 450008 Russia

<sup>2</sup>Institute of Biochemistry and Genetics, Ufa Federal Research Centre, Russian Academy of Sciences, Ufa, 450054 Russia

\*e-mail: ruji79@mail.ru

Conducted GWAS identified the association of osteoarthritis with more than 100 different SNPs, most of which are located in intronic and intergenic regions where genes encoding transposable elements and non-coding RNAs derived from them are located. A number of studies have also determined the activation of retroelements in joint tissues and in peripheral blood of patients with osteoarthritis. An assumption has been made that activated transposons, which cause aging and associated inflammation, influence the etiopathogenesis of osteoarthritis. To confirm this hypothesis, a search was conducted for data on changes in the expression of specific microRNAs derived from transposons during aging and osteoarthritis. As a result, 23 such microRNAs were found, the participation of which in the development of the disease is associated with an impact on genes and signaling pathways regulating cell proliferation and apoptosis, inflammatory and metabolic processes, and mechanisms of cartilage degradation. Changes in expression of these microRNAs indicate that the epigenetic mechanisms of aging are involved in osteoarthritis etiopathogenesis due to pathological activation of transposable elements complementary to the sequences of non-coding RNAs derived from them in evolution.

**Keywords:** immune system, microRNA, transposable elements, retroelements, osteoarthritis.

## МЕХАНИЗМЫ МОДУЛИРУЮЩЕГО ДЕЙСТВИЯ ТИМОХИНОНА (КОМПОНЕНТА ЧЕРНОГО ТМИНА, *Nigella sativa*), ВЛИЯЮЩИЕ НА АКТИВНОСТЬ НЕКОТОРЫХ ЯДЕРНЫХ И МИТОХОНДРИАЛЬНЫХ ГЕНОВ В ТКАНЯХ МЫШЕЙ ПОСЛЕ ВОЗДЕЙСТВИЯ РЕНТГЕНОВСКОГО ИЗЛУЧЕНИЯ

© 2025 г. С. А. Абдуллаев<sup>1, 2, \*</sup>, Д. В. Фомина<sup>1</sup>, Н. Ф. Раева<sup>1</sup>, М. А. Попов<sup>3</sup>,  
Т. Н. Максимова<sup>4</sup>, Г. Д. Засухина<sup>1, 5</sup>

<sup>1</sup>Федеральный медицинский биофизический центр им. А. И. Бурназяна, Москва, 123098 Россия

<sup>2</sup>Институт теоретической и экспериментальной биофизики Российской академии наук,  
Московская область, Пущино, 142290 Россия

<sup>3</sup>Московский областной научно-исследовательский клинический институт  
им. М. Ф. Владимировского, Москва, 129090 Россия

<sup>4</sup>Первый Московский государственный медицинский университет им. И. М. Сеченова, Москва, 119435 Россия

<sup>5</sup>Институт общей генетики им. Н. И. Вавилова Российской академии наук, Москва, 117971 Россия

\*e-mail: saabdullaev@gmail.com

Поступила в редакцию 05.06.2024 г.

После доработки 05.07.2024 г.

Принята к публикации 09.07.2024 г.

В работе обсуждается перспективный растительный препарат – тимохинон, компонент черного тмина (*Nigella sativa*), исследованный при многих патологиях на экспериментальных животных (мыши, крысы), характеризующийся положительным эффектом и отсутствием токсического действия. Препарат исследован в широком диапазоне доз при инъекционном и пероральном применении. Тимохинон обладает противомикробными, противовирусными, антивоспалительными, радиопротекторными свойствами. Основной повреждающей компонентой при действии ионизирующего излучения является оксидативный стресс. По этой причине в последнее время оценку радиопротекторов проводили по способности препарата снижать уровень оксидативного стресса. В качестве маркеров оксидативного стресса нами были использованы параметры изменения экспрессии генов ядерной и митохондриальной ДНК, осуществляющих важнейшие функции в клетке. Мышиам линии C57Bl/6 вводили тимохинон (10 мг/кг), через 30 мин проводили облучение (6 Гр). Через 6 и 24 ч исследовали экспрессию генов в клетках головного мозга и селезенки методом ПЦР в реальном времени. Было показано, что активность ядерных генов после воздействия радиации увеличивалась, но приходила в норму, если за 30 мин до облучения мышам вводили тимохинон. Митохондриальные гены также модифицировались в направлении активности контрольных клеток. Результаты анализов показывают, что тимохинон обладает защитными свойствами и может быть перспективен в качестве радиопротектора.

**Ключевые слова:** радиопротекторы, тимохинон, ядерные и митохондриальные гены.

**DOI:** 10.31857/S0016675825010033 **EDN:** VFHEAX

Необходимость изучения радиационного воздействия на человека связана с широким спектром его применения в медицине с диагностическими или лечебными целями, при контакте с различными дозами радиации в производственных условиях [1]. Предполагается увеличение продолжительности и длительности космических полетов, что может сопровождаться усилением воздействия радиоактивного фона, поэтому необходимы

дополнительные профилактические меры для сохранения здоровья космонавтов. Например, некоторые растительные препараты, обладающие широким спектром защитных эффектов, в том числе радиопротекторными свойствами, могут быть использованы в профилактических целях [2].

Особое внимание уделяется влиянию радиотерапии как одному из методов лечения онкозаболеваний, поскольку побочные эффекты действия

радиации в ряде случаев могут превышать положительные результаты лечения. Это связано с оксидативным стрессом, сопровождающим действие радиации, так как воздействие свободных радикалов играет немалую роль в патогенезе некоторых заболеваний и, как следствие, ведет к нарушению баланса между оксидативными и антиоксидантными параметрами. Так, оксидативный стресс может быть причиной развития радиационно-индуцированной пневмонии, которая может развиться даже через 6 мес. после радиотерапии рака легких. Радиопротекторы с антиоксидантными свойствами могут предохранять клеточные мембранные и ДНК организма от побочных эффектов проведения радиотерапии [3]. Осложнениями при радиотерапии рака простаты могут быть циститы, ректиты, а при раке груди – фиброзы, что также диктует необходимость применения защиты окружающих опухоли тканей. Одним из путей защиты нормальных тканей при применении радиотерапии может служить использование малых доз радиации, которые стабилизируют нормальные клетки, делая их более устойчивыми к высоким дозам радиации [4].

В последние годы в качестве радиопротекторов или соединений, ослабляющих токсическое действие излучения при радиотерапии, стали использовать натуральные продукты: растения и их компоненты [5]. Все больше исследований посвящено изучению профилактических и лечебных свойств тимохинона (TX) – компонента *Nigella sativa*, не обладающего токсическими свойствами и положительно влияющего на различные эндогенные и экзогенные патологии человека [6]. Черный тмин и его активный компонент – TX – проявляют различные фармакологические активности в отношении многих заболеваний (вирусных, бактериальных, а также при кардиопатологии и опухолеобразовании) благодаря антиоксидантным свойствам. Защита клеток от оксидативного стресса и воспаления регулируется через *Nrf2*- и *NFKB*-пути. Этот препарат характеризуется нейро-, нефро-, гастро- и антиканцерогенными эффектами [7]. TX ингибирует пролиферацию опухолевых клеток, что связано с повышением активности фосфатазы и гомолога тензина, что в конечном итоге приводит к подавлению *P13K* (АКТ-путь). Известно, что остановка клеточного цикла при действии TX сопровождается влиянием на активность генов *P53*, *STAT3*, а также на митохондриальный путь апоптоза. Некоторые авторы считают, что TX перспективен для профилактики и лечения ряда патологий [8]. TX обладает также радиопротекторными свойствами, так как способен снижать показатели оксидативного стресса в клетках головного мозга крыс, подвергшихся тотальному облучению. Это связано с повышением активности супероксиддисмутазы и перехватом свободных радикалов, образующихся при воздействии радиации. Такой механизм

защищает клеточные мембранные, что способствует предохранению клеток головного мозга от радиационно-индуцированных повреждений [9]. TX способствовал активации антиоксидантных ферментов (глутатионпероксидазы, глутатион S-трансферазы, супероксиддисмутазы и др.) у крыс после краинального облучения. TX предотвращал повышение уровня малонового диальдегида, вызванного воздействием радиации [10]. Как известно, повреждающее действие ионизирующего излучения на опухолевые клетки сопровождается также влиянием и на нормальные клетки организма [11]. У крыс, которым за 30 мин до облучения всего тела (6 Гр) вводили TX в концентрации 10 мг/кг, исследовали клетки печени, околоушные железы, мозг, яички по критерию содержания малондиальдегида, который является маркером оксидативного стресса, а также глутатионпероксидазы. В облученных клетках с предварительно введенным TX отмечались снижение уровня малондиальдегида и активация антиоксидантных ферментов. У крыс, получавших масло *Nigella sativa* при облучении головы, восстанавливалась гистологическая структура околоушных желез, уменьшался фиброз вокруг протоков и снижался уровень *TGF-β* [12]. Авторы считают, что TX является идеальным протектором, защищающим нормальные клетки при облучении опухолей. Таким образом, показано, что TX снижает эффект радиации, защищая нормальные клетки от повреждающего действия и повышая иммунный статус организма.

Авторы настоящей статьи, изучая радиопротекторную активность TX, использовали показатели оксидативного стресса. Исследованы также показатели влияния TX на активность ряда генов, в том числе митохондриальных, которые были ранее охарактеризованы ими как чувствительные к действию радиации [4, 13, 14]. Митохондриальные гены более чувствительны к повреждающим факторам. Так, при воздействии на клетку радиацией или химическими мутагенами в митохондриальных генах возникает в 3–50 раз больше повреждений, чем в ядерных генах [15]. По этой причине была исследована активность ряда генов ядерной и митохондриальной ДНК (я- и мтДНК) после действия радиации и введения TX.

## МАТЕРИАЛЫ И МЕТОДЫ

В исследовании использованы самцы мышей линии *C57Bl/6* двухмесячного возраста и массой 22–25 г, полученных из питомника экспериментальных животных Филиала Института биоорганической химии РАН (Пущино, Московская обл.). В ходе эксперимента животные находились в стандартных условиях вивария ФГБУ ГНЦ ФМБЦ им. А. И. Бурназяна ФМБА России. Животных содержали в поликарбонатных клетках на установке ИВК (индивидуальная вентиляция клеток) с

подачей стерильного воздуха (Фармбиолайн, Финляндия). Они были акклиматизированы в течение одной недели до начала экспериментов. Мышей кормили стандартным гранулированным кормом для лабораторных животных (ООО “Мест”, Москва) со свободным доступом к чистой питьевой воде. Они были размещены по пять особей в клетках при стандартном 12-часовом цикле свет/тень при температуре  $22 \pm 2^\circ\text{C}$  и при влажности  $45 \pm 5\%$ .

Облучение животных проводили на рентгеновской биологической установке РУСТ-М1 при напряжении 200 кВ, токе на трубке 2.5 мА, фильтре алюминиевом 1.5 мм. Мощность дозы рентгеновского облучения – 1 Гр/мин. Мышей в пластиковых контейнерах (по 5 особей вместе) подвергали облучению в дозе 6 Гр.

Тимохинон (2-изопропил-5-метил-1,4-бензодиоксанон) (Merck, Darmstadt, Германия) вводили животным внутрибрюшинно по 10 мг/кг массы тела. Вещество вводили за 30 мин до облучения.

Для изолирования тканей головного мозга и селезенки мышей умерщвляли путем обезглавливания через 6 и 24 ч после облучения. В качестве контроля использовали группы как необлученных, так и облученных мышей, не получавших ТХ. Головной мозг и селезенку изолировали и замораживали при  $-80^\circ\text{C}$  до проведения анализов.

Общая РНК из тканей селезенки и головного мозга была извлечена с помощью набора для выделения РНК (ExtractRNA, “Евроген”, Россия) в соответствии с инструкциями производителя. Концентрацию РНК определяли с использованием спектрофотометра NanoVuePlus (GEHealthcareTM, США) и доводили до 400 нг/мкл. 2 мкг РНК подвергали обратной транскрипции в комплементарную ДНК (кДНК) с использованием обратной транскриптазы MMLV (“Евроген”) в соответствии с протоколом производителя в общем объеме реакции 20 мкл. ПЦР в реальном времени проводили на приборе ДТ-прайм (ДНК-технология, Россия) с использованием набора qPCRmix-HS SYBR LowROX (Евроген) в соответствии с протоколом производителя. Реакционная смесь содержала 4 мкл кДНК, разбавленной в 20 раз, и 250 нМ каждого праймера. Последовательности праймеров и условия ПЦР к ядерным генам были использованы из ранее опубликованной нами работы [16], последовательности праймеров к митохондриальным генам представлены в работе [17]. Анализ кривых плавления был выполнен для всех генов, и специфичность, а также целостность продуктов ПЦР были подтверждены присутствием одного пика. Для анализа использовали метод 2-ДАСТ. ПЦР-тесты проводили в трех экземплярах для каждого образца. Уровни транскрипции генов-мишней были нормализованы по уровню эталонного

гена *GAPDH*. Нормализованные значения экспрессии генов анализировали с использованием программного обеспечения Prism GraphPad 7.0. Экспрессия гена у контрольных (интактных) мышей была принята за 100%. Статистическую обработку результатов осуществляли с помощью программы STATISTICA7.0. Для оценки значимости различий применяли непараметрический критерий Манна–Уитни. Различия считали значимыми при уровне показателя  $p < 0.05$ .

## РЕЗУЛЬТАТЫ

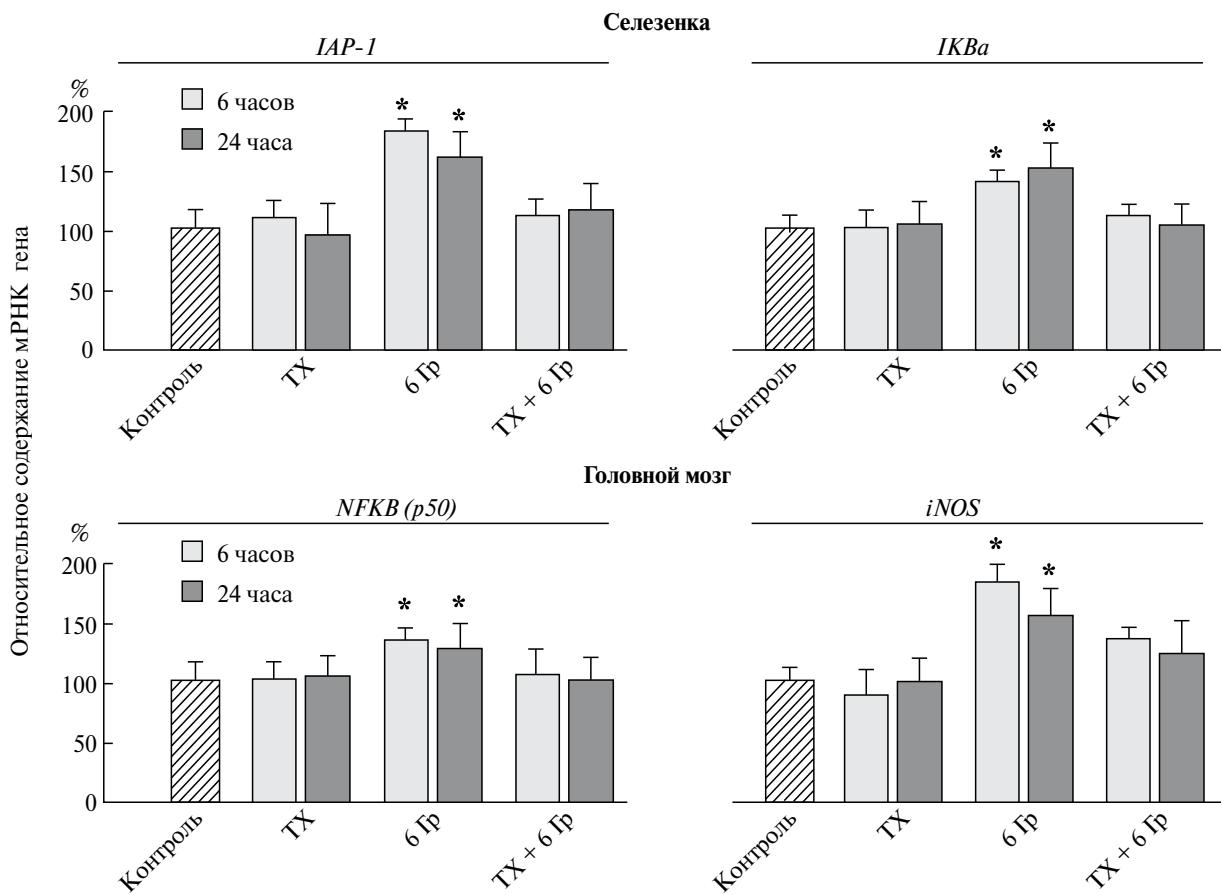
Исследована экспрессия ряда ядерных и митохондриальных генов ДНК в тканях селезенки и головного мозга мышей, которым до облучения (6 Гр) вводили ТХ в дозе 10 мг/кг.

На рис. 1 видно, что в ткани селезенки после облучения мышей отмечалось статистически значимое повышение экспрессии онкогенов *IAP-1* и *IKBa* по сравнению с животными, которым вводили только ТХ без облучения. В тканях головного мозга облученных мышей наблюдали повышение экспрессии гена *NFKB* (*p50*), контролирующего процессы пролиферации клеток и воспаления, и онкогена *iNOS* по сравнению с контрольными животными. Из данных рис. 1 видно, что введение ТХ мышам приводило к снижению экспрессии этих генов как в тканях селезенки, так и головного мозга мышей через 24 ч после облучения. Эти данные могут свидетельствовать о протекторном действии ТХ на исследованные ядерные гены по показателю модификации их активности.

На рис. 2 представлены результаты влияния ТХ на экспрессию генов мтДНК, участвующих в окислительном фосфорилировании и энергетическом балансе в тканях селезенки и головного мозга облученных мышей. Эти результаты показывают, что через 6 и 24 ч после облучения происходит снижение активности всех исследованных генов: вторая субъединица NADH-дегидрогеназы – компонент комплекса I, ген цитохрома *b* (*CYT-B*) – компонент комплекса III, шестая субъединица АТФ-сингтазы – компонент комплекса V. Снижение экспрессии этих генов регистрировали в тканях селезенки и головного мозга. Более выраженное снижение активности этих генов в пострадиационный период наблюдали в ткани селезенки по сравнению с показателями в ткани головного мозга. Однако у облученных животных, которым вводили ТХ, наблюдали повышение активности генов в обеих исследованных тканях, что также подтверждает способность ТХ к защите не только ядерной, но и митохондриальной ДНК.

## ОБСУЖДЕНИЕ

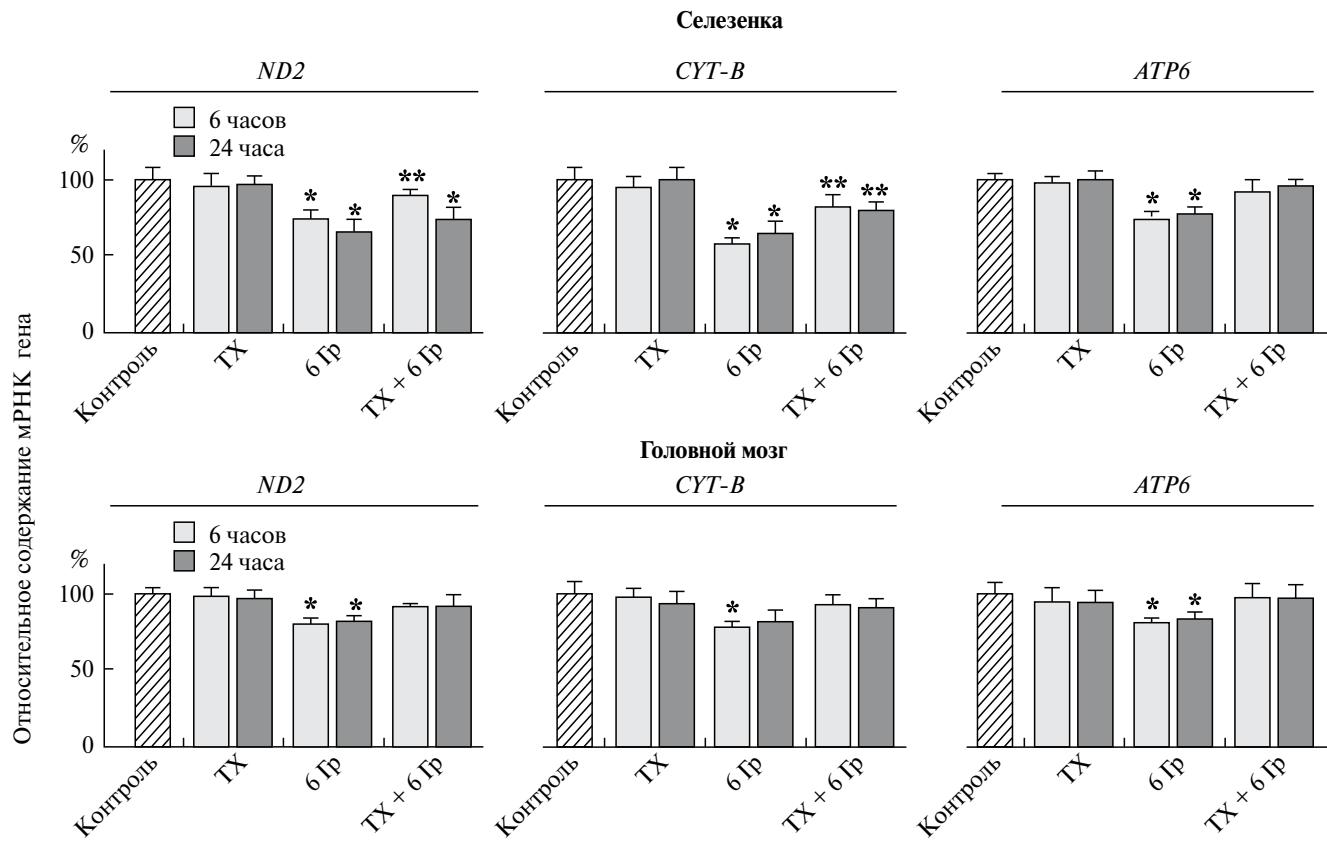
Одним из эффективных и распространенных методов лечения онкозаболеваний является



**Рис. 1.** Влияние TX на экспрессию онкогенов (*IAP-1*, *IKBa*, *NFKB (p50)*, *iNOS*) в тканях селезенки и головного мозга мышей через 6 и 24 ч после воздействия рентгеновского излучения в дозе 6 Гр. Уровень экспрессии генов у необлученных (контроль) мышей принят за 100%. Данные представлены как среднее  $\pm$  SEM из 5-6 независимых экспериментов. Отличия от контроля статистически значимы при  $p < 0.05$  (\*).

радиотерапия. Однако радиация влияет не только на опухолевые клетки, но и на нормальные за счет повышения уровня свободных радикалов. Ранее было показано, что такие изменения могут быть причиной повреждения легочной ткани и приводят к развитию через 6 мес. радиационно-индуцированной пневмонии, часто фатальной. На крысах при облучении (5 Гр) легких и введении TX за 30 мин до облучения (50 мг/кг) и в течение десяти дней после облучения было показано снижение всех показателей оксидативного стресса [3]. TX не обладает токсическими свойствами, но разнообразие его активности связано с концентрацией соединения и методом введения (внутривенное, внутримышечное или пероральное). Протекторный эффект TX перитонеально при исследовании на печени мышей был эффективен при дозах 4, 8, 12, 5, 25 и 50 мг/кг, однако увеличение дозы до 90.3 мг/кг могло привести к гибели животных. Субхроническое применение TX не сопровождалось какими-либо признаками токсичности с питьем для мышей в течение 90 дней [6]. Авторы доказывают,

что использование TX до радиотерапии предохраняло крыс от побочных эффектов. Так, было показано, что у 80% животных при радиотерапии головного мозга развивалась катаракта, однако ее развитие снижалось при использовании TX до или после облучения. Таким образом, авторы также полагают, что TX обладает протекторным эффектом. Измерение радиопротекторного эффекта рядом авторов осуществляется показателями индуцированного оксидативного стресса, который вызывает дисбаланс между повреждающим воздействием и защитными механизмами клетки. Защита клетки сопряжена с активностью генов и их регуляторов (некодирующих РНК). Активация онкогенов после воздействия радиации и снижение их экспрессии при действии протекторов могут, несомненно, служить маркером эффективности препарата. Авторами настоящей статьи было показано, что онкогены яДНК после воздействия радиации активировались, тогда как в опытах с радиацией и TX наблюдалось снижение их экспрессии. Гены мтДНК также модифицировались после радиационного



**Рис. 2.** Влияние TX на экспрессию генов mtДНК, участвующих в окислительном фосфорилировании (ND2, CYT-B, ATP6) в тканях селезенки и головного мозга мышей, через 6 и 24 ч после воздействия рентгеновского излучения в дозе 6 Гр. Уровень экспрессии генов у необлученных (контроль) мышей принят за 100%. Данные представлены как среднее  $\pm$  SEM из 5-6 независимых экспериментов. Отличия от контроля статистически значимы при  $p < 0.05$  (\*),  $p < 0.01$  (\*\*).

воздействия и активировались в опытах с TX. Следует подчеркнуть, что в последнее время mtДНК стала исследоваться наиболее интенсивно, так как ее связь с яДНК оказалась сопряженной и оказывающей влияние на ее активность. Так, ген *P53*, контролирующий клеточный гомеостаз и ряд других функций, при болезнях сердечно-сосудистой системы связан с митохондриальной дисфункцией и оказывает влияние на энергетический метаболизм, оксидативный стресс, митохондриальный апоптоз и аутофагию [18].

Обсуждая действие TX, нельзя не остановиться на особенностях его влияния на опухолевые клетки. Ряд авторов считают, что TX способен ингибировать пролиферацию и индуцировать апоптоз. При этом действуются некоторые эпигенетические механизмы, включая убиквитин-подобный растительный гомеодомен (*RHD*), ген RING fingerdomains (*UHRF1*), ДНК-метилтрансферазу 1 и гистондикацетилазу (*HDAC1*). TX обладает способностью связываться с *UHRF1* и *HDAC1*. Следует отметить, что TX формирует стабильную связь с цинком, находящимся в активном центре белка

*HDAC1*. Эти данные были получены на клеточных линиях MCF-7, HELA, Jurkat, где TX действовал дозозависимым образом, индуцируя апоптоз [19]. В эксперименте на крысах, которые получали перорально 50 мг/кг TX 3 раза в неделю, было показано ингибирование числа метастазов в печени и селезенке животных. В клеточных культурах TX снижал экспрессию ДНК-метилтрансферазы 1 (*DNMT1*). Авторы работы [20] считают, что TX обладает протекторными свойствами против канцерогенов, вызывающих рак грудной железы, через эпигенетические механизмы *DNMT1*.

Таким образом, TX является перспективным препаратом, характеризующимся многовекторностью действия на разные пути жизнедеятельности клетки, направленной на осуществление защитных систем организма. На животных были показаны положительные результаты действия TX, или черного тмина при многочисленных болезнях разной природы: антидиабетический, антимикробный, антипаразитарный, противовоспалительный, анти-гиперхолестеринемический и антигистаминный эффекты, эффекты на нервную,

кардиоваскулярную, респираторную системы [19, 21]. Вышеперечисленные авторы рекомендуют ТХ для профилактики и лечения многих патологий исходя из многочисленных данных, полученных на животных. Популяционные исследования, по-видимому, ожидаются. К профилактическим эффектам ТХ можно отнести радиопротекторное действие, которое выражается в модификации активности генов, осуществляющих восстановление баланса защитных систем энзиматической и неэнзиматической природы.

Работа выполнена по теме ФГБУ ГНЦ ФМБЦ им. А. И. Бурназяна “Технология-3” (госзадание № 123011300105-3).

Все эксперименты с животными проводились в соответствии с Европейской конвенцией о защите позвоночных животных, используемых в экспериментальных и других научных целях, директива 2010/63/EU.

Исследование одобрено Комитетом по биомедицинской этике ФГБУ ГНЦ ФМБЦ им. А. И. Бурназяна ФМБА России (протокол № 20 от 9 февраля 2024 г.).

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Ильин Л.А. Медицинские аспекты противодействия радиологическому и ядерному терроризму. М.: Наука, 2018. 392 с.
2. Ушаков И.Б. Космос, радиация, человек. М.: Научтехлитиздат, 2021. 352 с.
3. Dogru S., Taysi S., Yucel A. Effect soft thymoquinone in the lungs of rats against radiation-induced oxidative stress // Eur. Rev. Med. and Pharmacol. Sci. 2024. V. 28. № 1. P. 191–198. [https://doi.org/10.26355/eurrev\\_202401\\_34904](https://doi.org/10.26355/eurrev_202401_34904)
4. Михайлов В.Ф., Засухина Г.Д. Новый подход к стимуляции защитных систем организма малыми дозами радиации // Успехи соврем. биол. 2020. Т. 140. № 3. С. 244–252. <https://doi.org/10.31857/S0042132420030060>
5. Altay H., Demir E., Binici H. et al. Radioprotective effects of propolis and caffeic acid phenethyl ester on the tongue-tissues // Eur. J. Theor. 2020. V. 26. P. 202–207. <https://doi.org/10.5152/eurjther.2020.19047>
6. Taysi S., Algburi F., Mohammed Z. et al. Thymoquinone: A review on its pharmacological importance and its association with oxidative stress, COVID 19 and radiotherapy // Mini Rev. Med. Chem. 2022. V. 22. № 14. P. 1874–1875 <https://doi.org/10.2174/1389557522666220104151225>
7. Sadeghi E., Inenshahidi M., Hosseinzadeh H. Molecular mechanisms and signaling pathways of black cumin (*Nigella sativa*) and its active constituent, thymoquinone: A review // Mol. Biol. Rep. 2023. V. 50. P. 5439–5454. <https://doi.org/10.1007/s11033-023-08363-y>
8. Tiwari G., Cupta M., Devhare L., Tiwari R. Therapeutic and phytochemical properties of thymoquinone derived from *Nigella sativa* // Curr. Drug Res. Rev. 2023. <https://doi.org/10.2174/2589977515666230811092410>
9. Demir E., Taysi S., Ulusal H. et al. *Nigella sativa* oil and thymoquinone reduce oxidative stress in the brain tissue of rats exposed to total head irradiation // Int. J. Radiat. Biol. 2020. V. 96. № 2. P. 228–235. <https://doi.org/10.1080/09553002.2020.1683636>
10. Akyuz M., Taysi S., Baysal E. et al. Radioprotective effect of thymoquinone on salivary gland of rats exposed to total cranial irradiation // Head Neck. 2017. V. 39. № 10. P. 2027–2035. <https://doi.org/10.1002/hed.24861>
11. Koc M., Deniz C., Eryilmaz M. et al. Radioprotective effects of melatonin and thymoquinone on liver, hfrontid gland, brain, and testis of rats exposed to total body irradiation // Turk. J. Med. Sci. 2023. V. 53. P. 902–908. <https://doi.org/10.55730/1300-0144.5654>
12. Ahmed S., Bakz M. Will *Nigella sativa* oil protect parotid glands of rats against cranium gamma radiation? Histological and immunohistochemical evaluation // BMC Complement Med. Ther. 2024. V. 24. P. 111. <https://doi.org/10.1186/s12906-024-04410-8>
13. Abdullaev S., Gubina N., Bulanova T. et al. Assessment of nuclear and mitochondrial DNA, expression of mitochondria-related genes in different brain regions in rats after whole-body X-ray irradiation // Int. J. Mol. Sci. 2020. V. 21. <https://doi.org/10.3390/ijms21041196>
14. Abdullaev S.A., Glukhov S.I., Gaziev A.I. Radioprotective and radiomitigative effects of melatonin in tissues with different proliferative activity // Antioxidants (Basel). 2021. V. 10. <https://doi.org/10.3390/antiox10121885>
15. Газиев А.И. Пути сохранения целостности mitochondrialной ДНК и функций митохондрий в клетках, подвергшихся воздействию ионизирующей радиации // Радиац. биол. Радиоэкология. 2013. Т. 53. № 2. С. 117–136.
16. Михайлов В.Ф., Салеева Д.В., Рождественский Л.М. и др. Активность генов и некодирующих РНК как подход к определению ранних биомаркеров радиоиндуцированного опухолеобразования у мышей // Генетика. 2021. Т. 57. № 10. С. 1131–1140. <https://doi.org/10.31857/S0016675821100076>
17. Long G., Chen H., Wu M. et al. Antianemia drug roxadustat (FG-4592) protects against doxorubicin-induced cardiotoxicity by targeting antiapoptotic and antioxidant pathways // Front. Pharmacol. 2020. V. 11. <https://doi.org/10.3389/fphar.2020.01191>

18. Wang H., Yu W., Wang Y. et al. P53 contributes to cardiovascular diseases via mitochondria dysfunction: A new paradigm // Free Radic. Biol. Med. 2023. V. 208. P. 846–858.  
<https://doi.org/10.1016/j.freeradbiomed.2023.09.036>
19. Alsanosi S., Sheikh R., Sonbul S. et al. The potential role of *Nigella sativa* seed oil as epigenetic therapy of cancer // Molecules. 2022. V. 27.  
<https://doi.org/10.3390/molecules27092779>
20. Kaleem M., Kayali A., Sheikh R. et al. *In vitro* and *in vivo* preventive effects of thymoquinone against breast cancer-role of DNMT1 // Molecules. 2024. V. 29. P. 434.
21. Салеева Д.В., Раева Н.Ф., Абдуллаев С.А. и др. Профилактический и терапевтический потенциал тимохинона при ряде патологий человека на основе определения активации клеточных компонентов, осуществляющих защитные функции по активности генов и некодирующих РНК // Госпитальная медицина: наука и практика, 2023. Т. 6. № 2. С. 27–36.  
<https://doi.org/10.34852/GM3CVKG.2023.75.38.015>

## **Mechanisms of Modulating Action of Thymoquinone (Component of Black Cumin, *Nigella sativa*), Affecting the Activity of Some Nuclear and Mitochondrial Genes in Mice Tissue After Exposure to X-Ray Radiation**

**S. A. Abdullaev<sup>1, 2, \*</sup>, D. V. Fomina<sup>1</sup>, N. F. Raeva<sup>1</sup>, M. A. Popov<sup>3</sup>,  
T. N. Maksimova<sup>4</sup>, G. D. Zasukhina<sup>1, 5</sup>**

<sup>1</sup>*Burnazyan Federal Medical Biophysical Center of Russia, Moscow, 123098 Russia*

<sup>2</sup>*Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences, Moscow oblast, Pushchino, 142290 Russia*

<sup>3</sup>*Vladimirsky Moscow Regional Research Clinical Institute, Moscow, 129090 Russia*

<sup>4</sup>*Sechenov First Moscow State Medical University, Moscow, 119435 Russia*

<sup>5</sup>*Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, 117971 Russia*

\*e-mail: saabdullaev@gmail.com

The paper discusses a promising herbal preparation – thymoquinone, a component of black cumin (*Nigella sativa*), studied in experimental animals (mice, rats) in many pathologies, characterized by a positive effect and lack of toxic effect. The drug has been studied in a wide range of doses for injection and oral administration. Thymoquinone has antimicrobial, antiviral, anti-inflammatory, and radioprotective properties. The main damaging component of ionizing radiation is oxidative stress. For this reason, radioprotectors have recently been evaluated based on the drug's ability to reduce oxidative stress. As markers of oxidative stress, we used parameters of changes in the expression of nuclear and mitochondrial DNA genes that perform essential functions in the cell. *C57Bl/6* mice were administered thymoquinone (10 mg/kg) after 30 min. irradiation was performed (6 Gy). After 6 and 24 hours, gene expression in brain and spleen cells was studied using real-time PCR. It was shown that the activity of nuclear genes increased after exposure to radiation, but was normalized if thymoquinone was administered to mice 30 minutes before irradiation. Mitochondrial genes were also modified to target the activity of control cells. The test results show that thymoquinone has protective properties and may be promising as a radioprotector.

**Keywords:** radioprotectors, thymoquinone, nuclear and mitochondrial genes.

## ГЕНЕТИЧЕСКИЙ ПОЛИМОРФИЗМ ЛИСТВЕННИЦЫ СИБИРСКОЙ (*Larix sibirica* Ledeb.) В КОНТРАСТНЫХ ЭКОТОПАХ РЕСПУБЛИКИ ХАКАСИЯ

© 2025 г. Н. В. Орешкова<sup>1, 2, 3, \*</sup>, А. В. Пименов<sup>1</sup>, Т. С. Седельникова<sup>1</sup>, С. П. Ефремов<sup>1</sup>

<sup>1</sup>Институт леса им. В. Н. Сукачева – обособленное подразделение Федерального исследовательского центра “Красноярский научный центр Сибирского отделения Российской академии наук”, Красноярск, 660036 Россия

<sup>2</sup>Федеральный исследовательский центр “Красноярский научный центр Сибирского отделения Российской академии наук”, Красноярск, 660036 Россия

<sup>3</sup>Сибирский федеральный университет, Красноярск, 660041 Россия

\*e-mail: oreshkova@ksc.krasn.ru

Поступила в редакцию 07.06.2024 г.

После доработки 14.07.2024 г.

Принята к публикации 16.07.2024 г.

С помощью микросателлитных маркеров (SSR) проведена оценка генетического полиморфизма пяти ценопопуляций лиственницы сибирской (*L. sibirica*), произрастающих в Республике Хакасия. Наиболее высокие значения среднего числа аллелей на локус выявлены в ценопопуляциях *L. sibirica* с оステненных лугов долины озера Фыркал ( $3.9 \pm 0.458$ ) и облесенных суходольных склонов в долине озера Агасыр ( $3.9 \pm 0.348$ ). Для последней ценопопуляции характерны также максимальное значение эффективного числа аллелей на локус ( $2.302 \pm 0.283$ ) и самые большие показатели наблюдаемой и ожидаемой гетерозиготности ( $H_O = 0.487 \pm 0.074$ ,  $H_E = 0.492 \pm 0.070$ ). Анализ популяционной структуры свидетельствует о 0.8%-ном избытке гетерозиготных генотипов относительно популяции ( $F_{IS} = -0.008 \pm 0.031$ ) и 3.4%-ном дефиците гетерозиготных генотипов ( $F_{IT} = 0.034 \pm 0.034$ ) относительно вида. Дифференциация изученных ценопопуляций *L. sibirica* по SSR-маркерам составляет 4.3% ( $F_{ST} = 0.043$ ). Наименьшее генетическое расстояние (0.036) выявлено между экотопически максимально сходными (лесными) ценопопуляциями *L. sibirica*, а наибольшее генетическое расстояние (0.077), напротив, диагностировано между экотопически контрастными (болотной и оステнной) ценопопуляциями.

**Ключевые слова:** генетический полиморфизм, гетерозиготность, хвойные, микросателлитные маркеры, лиственница, *Larix*.

**DOI:** 10.31857/S0016675825010046 **EDN:** VEZVBB

Лиственница сибирская (*Larix sibirica* Ledeb.) относится к числу основных лесообразующих видов boreально-зональной зоны Евразии. На территории России леса с участием лиственницы (*Larix* Mill.) занимают около 264 млн га, или 38% всей лесопокрытой площади. В пределах области своего распространения лиственница сибирская образует широкий спектр морфологической изменчивости, проявляющейся в наличии внутривидовых форм и морфотипов деревьев, адаптированных к различным условиям произрастания [1]. При исследовании генетической изменчивости популяций *L. sibirica* севера Красноярского края, Урала, Алтая и Тывы установлено отсутствие тесной взаимосвязи между географическим положением выборок и степенью их генетической подразделенности при высоком уровне внутривидового генетического

разнообразия [2, 3]. Тем самым диагностирована важная особенность *L. sibirica*: основной пул генетической изменчивости вида сосредоточен на внутривидовом (ценопопуляционном) уровне, доля межпопуляционной (географической) компоненты значительно меньше. Оценка генетического разнообразия ряда видов хвойных, в том числе *L. sibirica*, на ценопопуляционном уровне позволила выявить адаптивную составляющую их внутривидовой дифференциации [4–7]. При выборе экспериментальных объектов для подобного рода исследований целесообразно фокусировать внимание на тех частях ареалов, где в целом оптимальные для вида условия произрастания характеризуются значительной почвенно-гидрологической и фитоценотической неоднородностью, наличием

орографических и фенологических межпопуляционных барьеров [8].

Таким критериям соответствует южносибирская часть ареала *L. sibirica* в пределах Республики Хакасия, где на относительно небольшой площади сосредоточены степные, лесные, луговые и болотные фитоценозы, сложно структурированные по высотным поясам гор, экспозициям невысоких кряжей, межгорным речным долинам и приозерным котловинам со свойственными для данных ландшафтов мозаичностью и резкими градиентами факторов среды [9]. Здесь широко представлены чистые и смешанные древостои *L. sibirica* горнотаежного, лесостепного и пойменно-болотного рядов развития, различающиеся по своей морфотипической, возрастной и пространственной структуре, по уровню и характеру антропогенной трансформации [10–12].

Цель настоящей работы – диагностика генетического разнообразия, структуры и внутривидовой дифференциации ценопопуляций лиственницы сибирской, произрастающих в Республике Хакасия.

## МАТЕРИАЛЫ И МЕТОДЫ

Объектами исследования послужили пять ценопопуляций *L. sibirica* из различных экотопов Республики Хакасия. Информация о местоположении пунктов сбора материала, об особенностях фитоценозов и морфологии деревьев представлена в табл. 1 и на рис. 1. Отбор экспериментальных материалов (образцов хвои) проводился в июле 2020 г. Общее количество отобранных деревьев составило 150 шт., количество образцов в каждой выборке – 30.

Для проведения генетического анализа в выборках лиственницы сибирской были использованы ядерные микросателлитные маркеры, разработанные ранее сотрудниками лаборатории лесной геномики СФУ для *L. sibirica*, а также японскими исследователями K. Isoda и A. Watanabe для лиственницы Кемпфера (*Larix kaempferi* (Lamb.) Carrière) [13, 14].

Отбор праймеров осуществляли путем подбора и оптимизации условий амплификации программы ПЦР. По результатам данного тестирования были отобраны полиморфные локусы, демонстрирующие хорошо интерпретируемые электрофорограммы. Отобранные полиморфные микросателлитные локусы для дальнейшего исследования выборок из популяций *L. sibirica* представлены в табл. 2.

Препараты тотальной ДНК были выделены модифицированным методом с применением цетил-триметиламмониумбромида (СТАВ) из образцов тканей хвои, высушенной при помощи силикагеля [15].

Для проведения ПЦР использовали готовые реакционные смеси для амплификации ДНК “GenePak PCR Core” производства ООО “Лаборатория Изоген”, содержащие ингибиованную для “горячего старта” Таq-ДНК-полимеразу, дезоксинуклеозидтрифосфаты и хлорид магния. Программа амплификации включала первичную денатурацию в течение 1 мин при 94°C, затем девять циклов “touchdown” с понижением на 1°C, каждый цикл: 30 с при 94°C, 30 с при 63°C, 1 мин при 72°C; далее 24 цикла без “touchdown”: 30 с при 94°C, 30 с при 53°C, 30 с при 72°C; финальная элонгация составляла 10 мин при 72°C. Продукты амплификации разделяли путем электрофореза в 6%-ном полиакриламидном геле с использованием Трис-EDTA-бортного электродного буфера в камерах для вертикального фореза. Окраску геля проводили в растворе бромистого этидия. В качестве маркера стандартных длин использовали ДНК плазмиды pBR322 *E. coli*, обработанную рестриктазой НрАII.

Расчет основных показателей для выборок пяти ценопопуляций лиственницы сибирской был проведен при помощи программы GenAIEx 6.51b2 [16]. Ошибки генотипирования, возникшие из-за *null-аллелей*, были идентифицированы и скорректированы с помощью программы MICRO-CHECKER [17]. Для построения филогенетического дерева использовались пакеты adegenet и poppr в R, построение производилось методом невзвешенных парногрупповых средних (UPGMA). Филогенетическое дерево было сгенерировано на основе стандартного генетического расстояния М. Няя [18]. Визуализация дерева осуществлялась с помощью онлайн-инструмента iTOL, версия 6 [19].

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В процессе исследования 10 ядерных микросателлитных локусов в пяти выборках лиственницы сибирской из Республики Хакасия было выявлено 45 аллельных вариантов, 29 (около 64 %) из которых оказались общими. У изученных нами ценопопуляций идентифицированные микросателлитные локусы частично различались по составу и частотам встречаемости выявленных аллелей. Наибольшее количество аллелей выявлено в следующих выборках: болотный лиственничник в долине оз. Агаскыр (AG-b) – 40 (из них 7 редких, частота встречаемости которых менее 5%); лиственничное редколесье в долине оз. Фыркал (FRK) – 39 (9 редких); лиственничное редколесье, остеиненное в пойме р. Карыш (KR) – 38 (8 редких). У ценопопуляции из долины р. Тунгужуль (TNZh) – наиболее возрастного и крупномерного насаждения, напротив, отмечается наименьшее число аллельных вариантов – 33 (из них 3 редких). Самый высокий уровень аллельного разнообразия в исследованных выборках лиственницы имеет

**Таблица 1.** Биоэкологическая характеристика ценопопуляций *L. sibirica* в Ширинском и Орджоникидзевском районах Республики Хакасия

Код ценопопуляционной выборки, географические координаты	Орографические и фитоценотические условия, таксационные характеристики насаждений
TNZh 54°16' с. ш., 89°38' в. д.	Лиственничное редколесье (13–15 экз./га) высокотравно-луговое полидоминантное (свыше 140 видов травянистых растений) на темно-серых и дерново-подзолистых почвах пологих склонов долины р. Тунгужуль (правый приток р. Белый Июс). 600–640 м над ур. моря. Возраст деревьев – 180–360 лет, высота – 18–34 м, диаметр – 58–92 см. Большое разнообразие крупномерных морфотипов деревьев: низкоросло-кряжистые крупносуковатые; с уплощенно-шатровыми, широко раскидистыми, куполообразными и даже сдвоенными формами крон; сильно сбежистыми, многосучковатыми и толстокорымыми стволами и др.
KR 54°24' с. ш., 89°59' в. д.	Лиственничное редколесье (6–17 экз./га) оstepненно-разнотравное на сухих примитивных почвах разрушающихся глыбистых останцев, их каменистых россыпей скального обрамления поймы р. Карыш (бассейн оз. Итколь). 500–520 м над ур. моря. Возраст деревьев – 90–140 лет, высота – 5–8 м, диаметр – 13–16 см. Одиночно стоящие деревья представлены различными формами: кривостволовыми, мелкосуковатыми, нередко суховершинными; в группах преобладают прямостволовые, сильно сбежистые, крупносуковатые, низкорослые деревья с этажно-асимметричными, либо уплощенными, флагообразными и иными формами крон
AG-s 54°58' с. ш., 89°15' в. д.	Редкостойный лиственничник (40–60 экз./га) на сухих и свежих минеральных почвах по южному склону в долине оз. Агасыр – бассейн р. Печище (левый приток р. Черный Июс). 540–555 м над ур. моря. Возраст деревьев – 130–185 лет, высота – 19–23 м, диаметр – 44–56 см. Большую часть крупномерных деревьев характеризуют низкие, раскидистые, суковатые кроны однобоко-шатровой формы, возможно, сформировавшиеся под влиянием системного ветрового воздействия
AG-b 54°58' с. ш., 89°16' в. д.	Болотный лиственничник с елью на пятнисто-мерзлотных, среднемощных (до 120 см), избыточно влажных торфяных почвах эутрофного ряда водно-минерального питания в долине оз. Агасыр – бассейн р. Печище (левый приток р. Черный Июс). 540 м над ур. моря. Выборка сформирована тонкомерным подростом (2–2.5 тыс. экз./га, возраст – 50–70 лет, высота – 4–6 м, диаметр – 8–11 см)
FRK 54°30' с. ш., 89°46' в. д.	Лиственничное редколесье паркового типа (8–9 экз./га), низкоросло-разнотравное с кустарниками на оstepненных лугах южной оконечности долины оз. Фыркал (правобережная пойма р. Белый Июс). 540 м над ур. моря. Возраст деревьев – 160–200 лет, высота – 12–14 м, диаметр – 50–60 см. Низкорослые, поодиночке рассеянные деревья с сильно сбежистыми стволами и крупными скелетными ветвями, формирующие флагообразные, зонтиковидные, конусообразные, асимметрично-шатровые, разветвленные, щитовидные и другие формы крон

локус *bcLK232*, в котором выявлено восемь аллелей. Локус *Ls\_2672894* в ценопопуляции FRK оказался мономорфным.

Анализ генетического разнообразия *L. sibirica* в исследованных насаждениях Республики Хакасия показал (табл. 3), что наиболее высокие значения среднего числа аллелей на локус были выявлены в выборках AG-s ( $3.9 \pm 0.348$ ) – облесенных

суходольных склонов в долине оз. Агасыр и FRK ( $3.9 \pm 0.458$ ) – оstepненных лугов долины оз. Фыркал, а эффективного числа аллелей на локус – в выборке AG-s ( $2.302 \pm 0.283$ ). Самые большие показатели наблюдаемой и ожидаемой гетерозиготности также выявлены в лесной выборке AG-s ( $H_O = 0.487 \pm 0.074$ ,  $H_E = 0.492 \pm 0.070$ ). В трех (AG-s, AG-b, FRK) из пяти выборок *L. sibirica* выявлен дефицит гетерозиготных генотипов. Высокое

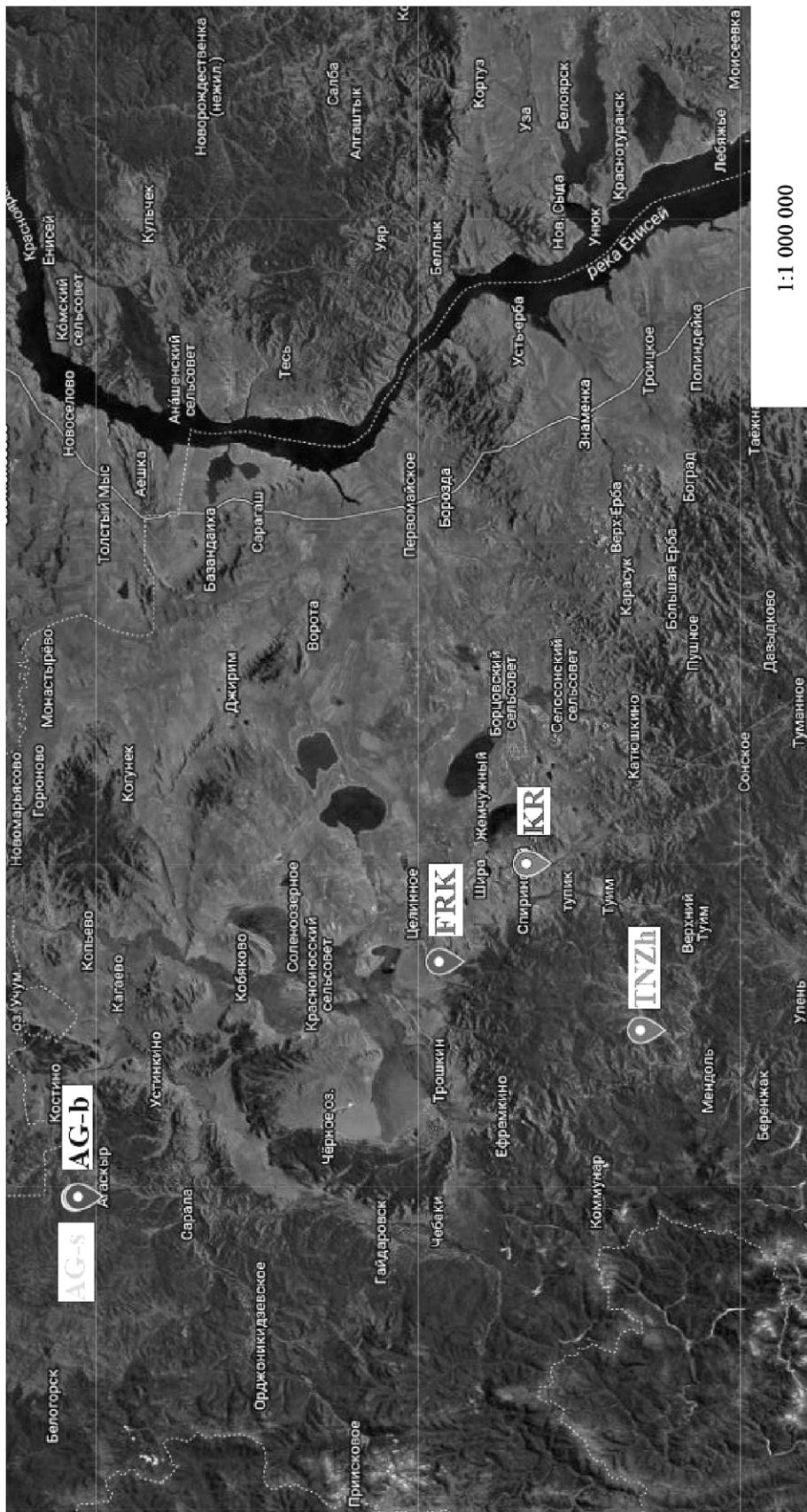


Рис. 1. Карта расположения изученных ценопопуляций *L. sibirica* в Ширинском и Орджоникидзевском районах Республики Хакасия (кодировка ценопопуляционных выборок соответствует табл. 1).

Таблица 2. Отобранные ядерные микросателлитные локусы для *L. sibirica*

Локус	Мотив	Последовательности праймеров	Длина фрагмента, п.н.	Источник литературы
<i>bcLK232</i>	$(AG)_{19}$	F: TGTTGCTGGGTGTTAGA R: GGGTAATAGTCCAGTCTTG	142–178	[14]
<i>bcLK224</i>	$(AG)_{17}$	F: GGAGAGGCCACTACTATTATTAC R: ATGCCTTCCTTCATTCCCTCT	152–168	
<i>bcLK066</i>	$(TG)_{12}$	F: GCAACCGACAATGATTACATAG R: CCTAAAACGTAACTTTGCTCAAT	155–172	
<i>Ls_954234</i>	$(ATT)_{15}$	F: TGGCGTTGGCTAAGTTGAA R: GGTTGATTATGTGTGTATGTGG	171–204	
<i>Ls_752897</i>	$(AAG)_{15}$	F: GCAGATGTTGATACAGTGGAGG R: CAGCTTCATTCGTGGCTAAT	216–264	
<i>Ls_417667</i>	$(AAT)_{16}$	F: CAGAGGATCTCATTCTGTTGA R: CTCGAAGGCCAATTAGGATAAA	207–243	
<i>Ls_2672894</i>	$(TTTG)_{11}$	F: CAAAGGATGGAATGTGTCTCAA R: GTTGGTATGGTTCCCAGAGTG	152–164	
<i>Ls_2552367</i>	$(CTAT)_{10}$	F: AAAGGTGCAATCACGTAAAGAC R: ATCGAAGCGGAAAATGTGTA	184–196	
<i>Ls_1008427</i>	$(ATAG)_{13}$	F: CACCCCTATCCCACAAATCTTA R: ATTTATCTTGGCCCTCATGC	152–174	
<i>Ls_305132</i>	$(GTCGGA)_{7}$	F: GCAGAGGCCGTTATTGATCTAT R: CCCTCGTTCCCTCTGACTA	210–240	

Таблица 3. Показатели генетической изменчивости, рассчитанные для пяти выборок *L. sibirica* по результатам SSR-анализа

Ценопопуляционные выборки	$N_A$	$N_E$	$H_O$	$H_E$	$F$
TNZh	$3.3 \pm 0.260$	$2.03 \pm 0.224$	$0.487 \pm 0.060$	$0.454 \pm 0.058$	$-0.083 \pm 0.053$
KR	$3.8 \pm 0.416$	$2.04 \pm 0.172$	$0.467 \pm 0.056$	$0.469 \pm 0.056$	$-0.008 \pm 0.055$
AG-s	$3.9 \pm 0.348$	$2.30 \pm 0.283$	$0.487 \pm 0.074$	$0.492 \pm 0.070$	$0.012 \pm 0.053$
AG-b	$3.6 \pm 0.340$	$1.87 \pm 0.174$	$0.390 \pm 0.057$	$0.419 \pm 0.059$	$0.054 \pm 0.043$
FRK	$3.9 \pm 0.458$	$2.21 \pm 0.248$	$0.473 \pm 0.073$	$0.483 \pm 0.068$	$0.030 \pm 0.047$
В среднем для всех исследованных ценопопуляций	$3.7 \pm 0.162$	$2.09 \pm 0.098$	$0.461 \pm 0.028$	$0.463 \pm 0.027$	$0.0004 \pm 0.023$

Примечание.  $N_A$  – среднее число аллелей на локус,  $N_E$  – эффективное число аллелей на локус,  $H_O$  – наблюдаемая гетерозиготность,  $H_E$  – ожидаемая гетерозиготность,  $F$  – индекс фиксации;  $\pm$  – стандартная ошибка.

значение индекса фиксации Райта найдено для популяции AG-b на торфянистых почвах заболоченной долины оз. Агаскыр ( $0.054 \pm 0.043$ ).

Анализ данных по генетическому полиморфизму исследованных выборок показал, что в целом значения наблюдаемой и ожидаемой гетерозиготности ( $H_O = 0.461$ ,  $H_E = 0.463$ ) сопоставимы

с аналогичными показателями, приведенными для *L. sibirica* из других местопроизрастаний [3, 20–22]. Однако выявленное пониженное аллельное разнообразие ( $N_A = 3.7$ ;  $N_E = 2.09$ ), возможно, связано с небольшими размерами и орографической изоляцией между собой хотя и территориально близких,

но растущих в контрастных экотопах ценопопуляций *L. sibirica*.

Стоит также отметить, что повышенные значения основных показателей генетической изменчивости, полученные другими авторами, связаны еще и с разными наборами микросателлитных локусов, используемых в исследованиях. Для анализа популяционно-генетической изменчивости лиственницы сибирской ранее в основном использовались наиболее высоко вариабельные динуклеотидные микросателлитные локусы [3, 20–22]. В настоящей работе мы также использовали динуклеотидные локусы, но основная доля приходилась на три-, тетра- и гексануклеотидные локусы (7 из 10 локусов).

В трех (AG-s, AG-b, FRK) из пяти выборок *L. sibirica* выявлен дефицит гетерозиготных генотипов, наибольший наблюдается в популяции из болотного лиственничника долины оз. Агаскыр ( $F = 0.054$ ). Это указывает на вероятное близкородственное скрещивание в популяции, сформированной на момент отбора образцов главным образом молодыми деревьями — потомством небольшого числа деревьев, сохранившихся после интенсивных рубок, которые проводились в лиственничниках Хакасии в середине XX в. Ближе всего к равновесному состоянию ( $F = -0.008$ ) находится популяция из долины р. Карыш — антропогенно наименее нарушенное насаждение лиственницы в данном исследовании. Дефицит гетерозиготных генотипов ранее также был выявлен при

изучении генетического разнообразия природных популяций лиственницы сибирской с Урала [3, 22].

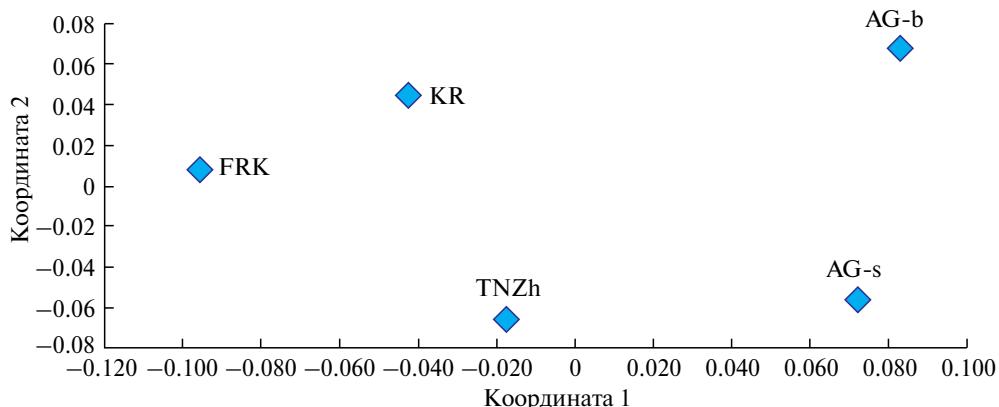
Анализ популяционной структуры и степени генетической подразделенности ценопопуляций лиственницы сибирской (табл. 4) [23] показал, что в изучаемых выборках наблюдается 0.8%-ный избыток гетерозиготных генотипов относительно популяции ( $F_{IS} = -0.008 \pm 0.031$ ) и 3.4%-ный дефицит гетерозиготных генотипов ( $F_{IT} = 0.034 \pm 0.034$ ) относительно вида. Полокусные значения показывают, что наиболее значимый дефицит гетерозигот наблюдается в локусах *Ls\_1008427* и *Ls\_954234*. Коэффициент инбридинга популяции относительно вида в целом ( $F_{ST}$ ), отражающий степень подразделенности популяций, варьирует от 0.013 (*Ls\_2672894*) до 0.091 (*bcLK232*), составляя в среднем  $0.043 \pm 0.007$ . Это свидетельствует о том, что только около 4.3% выявленной генетической изменчивости у изученных выборок лиственницы сибирской распределется между популяциями. Внутри популяций сосредоточено около 95.7% всего генетического разнообразия. Наибольший вклад в дифференциацию изученных популяций вносят локусы *bcLK232* и *bcLK224* (см. табл. 4).

Анализ гетерогенности аллельных частот с помощью критерия  $\chi^2$  показал, что у одного (*Ls\_2672894*) из десяти исследованных локусов наблюдаемые различия статистически недостоверны. У остальных локусов различия по частотам аллелей высоко достоверны ( $P < 0.001$ ) (см. табл. 4).

**Таблица 4.** Значение показателей *F*-статистик Райта для *L. sibirica*

Локус	<i>N</i>	$\chi^2$	$F_{IS}$	$F_{IT}$	$F_{ST}$
<i>bcLK232</i>	8	$63.751 (1.32 \cdot 10^{-4})^{***}$	0.020	0.109	0.091
<i>bcLK224</i>	4	$49.813 (5.125 \cdot 10^{-9})^{***}$	0.053	0.117	0.068
<i>bcLK066</i>	4	$58.437 (9.344 \cdot 10^{-11})^{***}$	-0.093	-0.048	0.041
<i>Ls_954234</i>	5	$157.324 (1.155 \cdot 10^{-28})^{***}$	0.092	0.118	0.028
<i>Ls_752897</i>	5	$50.711 (1.974 \cdot 10^{-7})^{***}$	-0.043	0.000	0.041
<i>Ls_417667</i>	3	$94.971 (1.873 \cdot 10^{-20})^{***}$	-0.191	-0.170	0.018
<i>Ls_2672894</i>	2	0.086 (0.770) <sup>ns</sup>	-0.037	-0.024	0.013
<i>Ls_2552367</i>	5	$150.415 (3.063 \cdot 10^{-8})^{***}$	-0.005	0.039	0.044
<i>Ls_1008427</i>	5	$56.307 (1.798 \cdot 10^{-8})^{***}$	0.156	0.201	0.053
<i>Ls_305132</i>	4	$63.819 (7.514 \cdot 10^{-12})^{***}$	-0.027	0.002	0.029
Среднее			-0.008 ±0.031	0.034 ±0.034	0.043 ±0.007

Примечание. *N* — число аллелей;  $\chi^2$  — тест на гетерогенность, уровень значимости ns — not significant, \*\*\* —  $P < 0.001$ ;  $F_{IS}$  — коэффициент инбридинга особи относительно популяции;  $F_{IT}$  — коэффициент инбридинга особи относительно вида;  $F_{ST}$  — коэффициент межпопуляционной дифференциации; ± — стандартная ошибка.



**Рис. 2.** Проекция изученных ценопопуляционных выборок *L. sibirica* на плоскости двух координат по данным PCA-анализа матрицы генетических расстояний (кодировка ценопопуляционных выборок соответствует табл. 1).

Результаты теста распределения генетической изменчивости (AMOVA) с учетом иерархических уровней показали, что наибольшее генетическое разнообразие было найдено внутри особей (94 из 100%). Генетическое разнообразие между особями, составляющее 2 и 4%, приходится на межпопуляционную составляющую.

Оценка генетической дифференциации между исследованными популяциями лиственницы сибирской была проведена с использованием стандартного генетического расстояния ( $D_N$ ) М. Нея [18] на основании частот аллелей десяти микросателлитных локусов. Наименьшее генетическое расстояние (0.036) было выявлено между экотипически максимально сходными (лесными) ценопопуляциями TNZh и AG-s (табл. 5). Наибольшее генетическое расстояние (0.077), напротив, было диагностировано между экотипически контрастными ценопопуляциями – болотной AG-b и оステпенной FRK.

Выявленный уровень генетической дифференциации в пяти выборках лиственницы сибирской наглядно показывает расположение ценопопуляций на плоскости двух координат (рис. 2).

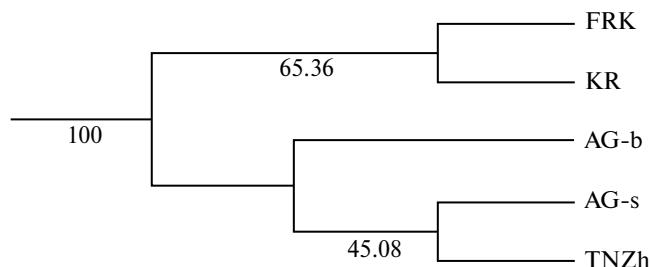
Результаты теста Мантела ( $R = 0.479$ ,  $P = 0.100$ ) свидетельствуют об отсутствии корреляции между матрицами генетического и географического расстояния в исследованном регионе [24].

**Таблица 5.** Попарные генетические расстояния Нея для ценопопуляционных выборок *L. sibirica*

Ценопопуляционные выборки	TNZh	KR	AG-s	AG-b	FRK
TNZh	0.000				
KR	0.038				
AG-s	0.036	0.060			
AG-b	0.061	0.052	0.041		
FRK	0.046	0.042	0.072	0.077	0.000

Филогенетическое дерево для пяти выборок лиственницы сибирской из разных экотопов Республики Хакасия, построенное методом невзвешенных парногрупповых средних (UPGMA), сгенерированное на основе стандартного генетического расстояния М. Нея [18] с оценкой бутстреп-поддержки узлов ветвления, иллюстрирует группировку ценопопуляций (рис. 3).

На дендрограмме четко выделяются две группы с бутстреп-поддержкой в 100%. В первую со значением бутстрепа 65.36 вошли выборки из долины р. Карыш (KR) и из долины оз. Фыркал (FRK) ( $F_{ST} = 0.019$ ,  $D_N = 0.042$ ). Данные ценопопуляции – ландшафтные изоляты на крайнем пределе распространения древесной растительности в Ширинской степи – максимально близки между собой по экологическим условиям и таксационным характеристикам относительно других выборок. Второй кластер объединил лесные выборки лиственницы из восточных отрогов Кузнецкого Алатау: долин р. Тунжугуль и оз. Агаскыр ( $F_{ST} = 0.024$ ,  $D_N = 0.046$ ). Образующие данные выборки ценопопуляции лиственницы (болотный и суходольный лиственничники) экотипически поливариантны по условиям водно-минерального питания, почвенным характеристикам и фитоценотической структуре. Весьма вероятно, что совокупность этих факторов определяет крайне низкие для данного кластера значения бутстреп-поддержки, а



**Рис. 3.** Филогенетическое дерево пяти ценопопуляций *L. sibirica* (кодировка ценопопуляционных выборок соответствует табл. 1).

следовательно, и сомнительную достоверность его топологии.

В целом на основании данного филогенетического дерева с низкими значениями бутстрепа внутри полученных кластеров – “степного” и “лесного” – можно сказать о том, что ценопопуляции лиственницы в пределах каждого из этих двух контрастных экотопов Республики Хакасия генетически неоднородны и слабо дифференцированы. Вероятными причинами этого являются внутрипопуляционная гетерогенность условий произрастания и слабая репродуктивная изоляция территориально смежных в пределах той или иной межгорной долины ценопопуляций лиственницы сибирской.

Авторы благодарят Яну Викторовну Стенину, выпускницу магистерской программы “Геномика и биоинформатика” кафедры геномики и биоинформатики ИФБ и БТ СФУ, за участие в лабораторных исследованиях и помочь в обработке полученных данных.

Работа выполнена в рамках Государственно-го задания ФИЦ КНЦ СО РАН по теме “Биоразнообразие лесов Сибири: эколого-динамический, генетико-селекционный, физико-химический и ресурсно-технологический аспекты” № FWES-2024-0028.

Эта статья не содержит каких-либо исследований с участием людей или животных.

Авторы заявляют об отсутствии конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Биоразнообразие лиственниц Азиатской России / Под ред. С. П. Ефремова, Л. И. Милитина, Новосибирск: ГЕО, 2010. 159 с.
2. Орешкова Н.В., Белоконь М.М. Генетическая дифференциация популяций лиственницы сибирской (*Larix sibirica* Ledeb.) по микросателлитным локусам // Вестник Красноярского гос. аграрного ун-та. 2012. № 2. С. 111–116.
3. Васильева Ю.С., Сбоева Я.В., Чертов Н.В., Жуланов А.А. Оценка состояния генофонда западной расы лиственницы сибирской (*Larix sibirica* Ledeb.) Урала на основании полиморфизма микросателлитных маркеров // Бюлл. науки и практики. 2019. Т. 5. № 12. С. 98–110. <https://doi.org/10.33619/2414-2948/49/11>
4. Шейкина О.В., Демаков Ю.П., Гладков Ю.Ф., Унженина О.В. Генетическая изменчивость и дифференциация суходольной и болотной ценопопуляций сосны обыкновенной в Республике Марий Эл // Науч. журн. КубГАУ. 2013. № 94 (10). С. 1–12.
5. Орешкова Н.В., Седельникова Т.С., Пименов А.В., Ефремов С.П. Генетическая структура и дифференциация болотных и суходольных популяций сосны кедровой сибирской (*Pinus sibirica* Du Tour) по ядерным микросателлитным локусам // Генетика. 2014. Т. 50. № 9. С. 1059–1066. <https://doi.org/10.7868/S0016675814090100>
6. Oreshkova N.V., Sedel'nikova T.S., Efremov S.P., Pimenov A.V. Genetic Polymorphism of Siberian Stone Pine (*Pinus sibirica* Du Tour) in Kuznetsk Alatau // Contemporary Problems Ecology. 2020. V. 13. № 6. P. 569–576. <https://doi.org/10.1134/S1995425520060116>.
7. Oreshkova N.V., Sedel'nikova T.S., Pimenov A.V. Genetic diversity of coenopopulations of *Larix sibirica* Ledeb. and *Picea obovata* Ledeb. in Taimyr // Contemporary Problems Ecology. 2023. V. 16. № 5. P. 612–619. <https://doi.org/10.1134/S1995425523050098>
8. Санников С.Н., Шавнин С.А., Санникова Н.С., Петрова И.В. Эколого-генетические принципы выделения и классификации лесных генетических резерватов // Экология. 2015. № 1. С. 3–8. <https://doi.org/10.7868/S0367059715010151>
9. Куминова А.В., Зверева Г.А., Маскаев Ю.М. и др. Растительный покров Хакасии. Новосибирск, 1976. 423 с.
10. Седельникова Т.С., Пименов А.В. Числа хромосом форм *Larix sibirica* (Pinaceae) в Ширинской степи Республики Хакасия // Бот. журн. 2017. Т. 102. № 5. С. 693–697. <https://doi.org/10.1134/S0006813617050118>
11. Седельникова Т.С., Аверьянов А.С., Пименов А.В. Особенности пыльцы внутривидовых форм лиственницы сибирской в контрастных экотопах Южной Сибири // Лесоведение. 2021. № 3. С. 265–277. <https://doi.org/10.31857/S0024114821030104>
12. Аверьянов А.С., Барченков А.П., Пименов А.В., Седельникова Т.С. Морфологические признаки шишечек лиственницы сибирской на юге Сибири // Сиб. лесной журн. 2024. № 1. С. 13–20. <https://doi.org/10.15372/SJFS20240102>
13. Орешкова Н.В., Бондар Е.И., Путинцева Ю.А. и др. Разработка ядерных микросателлитных маркеров с длинными (трех-, четырех-, пяти- и

- шестинуклеотидными) мотивами для трех видов лиственницы на основе полногеномного *de novo* секвенирования лиственницы сибирской (*Larix sibirica* Ledeb.) // Генетика. 2019. Т. 55. № 4. С. 418–425.  
<https://doi.org/10.1134/S001667581904009X>
14. Isoda K., Watanabe A. Isolation and characterization of microsatellite loci from *Larix kaempferi* // Mol. Ecol. Notes. 2006. V. 6. № 3. P. 664–666.  
<https://doi.org/10.1111/j.1471-8286.2006.01291.x>
15. Doyle J.J., Doyle J.L. Isolation of plant DNA from fresh tissue // Focus. 1990. V. 12. № 1. P. 13–15.
16. Peakall R.O.D., 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>
17. Van Oosterhout C., Hutchinson W.F., Wills D.P.M., Shipley P. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data // Mol. Ecol. Notes. 2004. V. 4. № 3. P. 535–538.  
<https://doi.org/10.1111/j.1471-8286.2004.00684.x>
18. Nei M. Genetic distance between populations // Am. Naturalist. 1972. V. 106. P. 283–291.  
<https://doi.org/10.1086/282771>
19. Letunic I., Bork P. Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation // Bioinformatics. 2007. V. 23. № 1. P. 127–128.  
<https://doi.org/10.1093/bioinformatics/btl529>
20. Орешкова Н.В., Белоконь М.М. Оценка генетической изменчивости лиственницы сибирской с использованием микросателлитного анализа // Вестник МГУЛ – Лесной вестник. 2012. № 1 (84). С. 118–122.
21. Орешкова Н.В., Белоконь М.М., Жамъянсурен С. Генетическое разнообразие, популяционная структура и дифференциация лиственниц сибирской, Гмелина и Каяндеры по данным SSR-маркеров // Генетика. 2013. Т. 49. № 2. С. 204–213.  
<https://doi.org/10.7868/S0016675812120090>
22. Chertov N., Vasilyeva Yu., Zhulanov A. et al. Genetic structure and geographical differentiation of *Larix sibirica* Ledeb. in the Urals // Forest. 2021. V. 12. № 1401.  
<https://doi.org/10.3390/f12101401>
23. Wright S. The interpretation of population structure by F-statistics with special regard to systems of mating // Evolution. 1965. V. 19. № 3. P. 395–420.  
<https://doi.org/10.1111/j.1558-5646.1965.tb01731.x>
24. Mantel N. The detection of disease clustering and a generalized regression approach // Cancer Research. 1967. V. 27. № 2. P. 209–220.

## Genetic Polymorphism of Siberian Larch (*Larix sibirica* Ledeb.) in Contrasting Ecotopes of the Republic of Khakassia

N. V. Oreshkova<sup>1, 2, 3, \*</sup>, A. V. Pimenov<sup>1</sup>, T. S. Sedel'nikova<sup>1</sup>, S. P. Efremov<sup>1</sup>

<sup>1</sup>Sukachev Institute of Forest – a separate division of the Federal Research Center “Krasnoyarsk Science Center of the Siberian Branch of the Russian Academy of Sciences”, Krasnoyarsk, 660036 Russia

<sup>2</sup>Federal Research Center “Krasnoyarsk Science Center of the Siberian Branch of the Russian Academy of Sciences”, Krasnoyarsk, 660036 Russia

<sup>3</sup>Siberian Federal University, Krasnoyarsk, 660041 Russia

\*e-mail: oreshkova@ksc.krasn.ru

Using microsatellite markers (SSR), the genetic polymorphism of five coenopopulations of Siberian larch (*L. sibirica*) growing in the Republic of Khakassia was assessed. The highest values of the average number of alleles per locus were found in coenopopulations of *L. sibirica* from the steppe meadows of the valley of Lake Fyrkal ( $3.9 \pm 0.458$ ) and forested upland slopes in the valley of Lake Agaskyr ( $3.9 \pm 0.348$ ). The latter cenopopulation is also characterized by the maximum value of the effective number of alleles per locus ( $2.302 \pm 0.283$ ) and the highest rates of observed and expected heterozygosity ( $H_O = 0.487 \pm 0.074$ ,  $H_E = 0.492 \pm 0.070$ ). Analysis of the population structure indicates a 0.8% excess of heterozygous genotypes relative to the population ( $F_{IS} = -0.008 \pm 0.031$ ) and a 3.4% deficiency of heterozygous genotypes ( $F_{IT} = 0.034 \pm 0.034$ ) relative to the species. The differentiation of the studied coenopopulations of *L. sibirica* by SSR markers is 4.3% ( $F_{ST} = 0.043$ ). The smallest genetic distance (0.036) was detected between ecologically most similar (forest) coenopopulations of *L. sibirica*, and the largest genetic distance (0.077), on the contrary, was identified between ecologically contrasting (swamp and steppe) coenopopulations.

**Keywords:** genetic polymorphism, heterozygosity, conifers, microsatellite markers, larch, *Larix*.

## ГЕНЕТИЧЕСКАЯ ОЦЕНКА ГОЛШТИНСКОГО СКОТА ПО МИКРОСАТЕЛЛИТНЫМ МАРКЕРАМ ДНК

© 2025 г. Л. А. Калашникова<sup>1,\*</sup>, Т. Б. Ганченкова<sup>1</sup>, Н. В. Рыжова<sup>1</sup>, Я. А. Хабибрахманова<sup>1</sup>,  
И. Е. Багаль<sup>1</sup>, И. Ю. Павлова<sup>1</sup>, А. Е. Калашников<sup>1</sup>

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

\*e-mail: lakashnikova@mail.ru

Поступила в редакцию 20.12.2023 г.

После доработки 24.05.2024 г.

Принята к публикации 02.07.2024 г.

Представлены результаты исследования полиморфизма 12 микросателлитных локусов крупного рогатого скота голштинской породы ряда регионов России и зарубежных стран. Среднее количество аллелей на локус составило  $5.43 \pm 0.19$  при колебаниях в интервале 4–13 аллелей, среднее число эффективных аллелей –  $3.26 \pm 0.11$ . Установлен перечень из 29 наиболее часто встречающихся аллелей. Выявлено 22 приватных аллеля, при этом частота приватных аллелей составила 0.004–0.033. Показано, что количество локально распространенных аллелей в отечественных стадах выше, чем у животных зарубежной селекции. Средний уровень наблюдаемой гетерозиготности по всем локусам достиг  $0.681 \pm 0.017$  и варьировал в диапазоне 0.65–0.78 при индексе фиксации  $-0.131 \pm 0.005$ . Генетические расстояния между стадами отечественной селекции составили  $< 0.074$ . Выявлено, что стада коров распадаются на два кластера. В один кластер вошли животные из трех областей России вместе с быками Германии и Нидерландов, а в другой кластер – животные двух регионов, близких быкам Канады, США и Англии. При этом быки Дании и Финляндии оказались в отдельном кластере. Цель настоящей работы – оценка аллелофонда голштинского скота отечественной селекции и определение генетического профиля породы по STR-маркерам.

**Ключевые слова:** генетическая дифференциация, голштинская порода, черно-пестрый скот, микросателлиты, маркеры, породная принадлежность.

**DOI:** 10.31857/S0016675825010057 **EDN:** VEUINC

Голштинская порода крупного рогатого скота относится к молочному типу продуктивности. Она отличается высокими производственными показателями и способностью к адаптации к различным географическим условиям. В России сформирован отечественный массив животных голштинской породы [1–3]. Актуальность работы заключается в том, что результаты индивидуальной генетической идентификации используются не только для научных, сколько для практических целей – определения достоверности происхождения племенного материала и для генетической характеристики пород животных [4]. Необходимо отметить, что впервые проводимая работа соответствует положению о порядке проведения молекулярно-генетической экспертизы племенной продукции требованиям ЕЭК. Статистический анализ результатов генотипирования по микросателлитным локусам ядерной ДНК (STR, short tandem microsatellite repeats) позволил оценить уровень родства, генетического сходства и

расстояний между группами животных. При этом формирование матриц родства является основой геномной оценки [5].

В настоящее время в дополнение к основному методу установления записей о родословных по данным племенной книги используются молекулярно-генетические методы. В планах селекционно-племенной работы для определения родства особей внутри пород и породной принадлежности групп животных применяются STR, имеющие в отношении фенотипических признаков нейтральный характер и высокую степень полиморфизма [6–8]. Критерием оценки общего уровня генетического разнообразия животных голштинской породы является полиморфизм участков STR, перечень и частота аллельных вариантов [9].

В мировой практике племенной работы в последние десятилетия для решения вопросов идентификации рекомендуется применять

молекулярно-генетические методы на основе STR- и SNP-маркеров [10, 11] (руководство по генетике ICAR) [12]. Одной из главных проблем определения генетической структуры породы является факт, что экспертиза может осуществляться только по одинаковой панели генетических маркеров при идентичном способе определения генетического разнообразия. При использовании различных закрытых коммерческих панелей STR-локусов проводить сравнительные проверки не представляется возможным [13–16].

Цель исследований – оценка аллелофонда голштинского скота отечественной селекции и определение генетического профиля породы по микросателлитным маркерам геномной ДНК. Настоящая работа проводилась впервые именно с племенными животными при выполнении планов селекционно-племенной деятельности, инвентаризации голштинской породы, а также приведении породы в соответствие с требованиями ЕЭК.

## МАТЕРИАЛЫ И МЕТОДЫ

Работа проведена в лаборатории ДНК технологий ФГБНУ ВНИИПлем Минсельхоза России. Для характеристики генофонда голштинской породы отечественной и зарубежной селекции определен полиморфизм 12 микросателлитных локусов ядерной ДНК (BM1824, SPS115, TGLA53, TGLA227, ETH3, BM2113, TGLA126, TGLA122, BM1818, INRA023, ETH10, ETH225) [12].

Затем определен спектр аллелей 722 голов скота племенных хозяйств Курской (KUR, 124 гол.), Волгоградской (VLG, 100 гол.), Белгородской

(BEL, 100 гол.), Пензенской (PNZ, 100 гол.) областей и Краснодарского края (KRD, 96 гол.), а также быков-производителей североамериканской селекции из США (USA, 100 гол.) и Канады (CAN, 35 гол.) и европейской селекции – Великобритании (GBR), Нидерландов (NDL), Германии (DEU), Дании (DNM), Финляндии (FIN), Испании (SPA) (всего 65 гол.).

Первичную обработку данных фрагментного анализа, полученных в мультиплексной ПЦР, осуществляли на капиллярном секвенаторе Applied Biosystems 3130 с помощью программного обеспечения (ПО) Genemapper (вер. 6). Статистическую обработку данных вычисляли в модуле GenAIEx (вер. 6.5). Для выявления межпопуляционных различий внутри породы сделан расчет генетических расстояний между исследованными группами в ПО Mega (вер. 6) [12].

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В настоящей работе проведены исследования полиморфизма 12 STR-локусов, представляющих панель оценки достоверности происхождения ISAG (Int. Society of Animal Genetics) [3, 8, 16–18] голштинской породы (в особенности племенных животных действующего генофонда) в целом, что формирует цель исследований как инвентаризацию породы и в значительной степени превышает объемы аналогичных филогенетических исследований. Локусы имели высокий полиморфизм (табл. 1), среднее количество аллелей на локус  $N_a$  в среднем по всем локусам и генетическим группам составило  $5.43 \pm 0.19$  (в Тюменском регионе

Таблица 1. Усредненные показатели генетического разнообразия

Генетическая группа	$N_a \pm SD$	$N_a, p \geq 5\% \pm SD$	$N_e (cp.) \pm SD$	$I \pm SD$
BEL	$7.91 \pm 0.57$	$4.25 \pm 0.47$	$3.65 \pm 0.42$	$1.44 \pm 0.10$
PNZ	$7.16 \pm 0.72$	$4.50 \pm 0.41$	$3.68 \pm 0.39$	$1.43 \pm 0.09$
VLG	$7.16 \pm 0.61$	$4.25 \pm 0.35$	$3.74 \pm 0.38$	$1.44 \pm 0.09$
GBR	$6.41 \pm 0.39$	$4.50 \pm 0.43$	$3.63 \pm 0.37$	$1.41 \pm 0.09$
KUR	$6.33 \pm 0.55$	$4.50 \pm 0.37$	$3.34 \pm 0.31$	$1.36 \pm 0.09$
KRD	$6.33 \pm 0.55$	$4.08 \pm 0.35$	$3.75 \pm 0.36$	$1.41 \pm 0.09$
CAN	$5.91 \pm 0.54$	$4.25 \pm 0.41$	$3.55 \pm 0.39$	$1.37 \pm 0.10$
USA	$5.66 \pm 0.48$	$3.83 \pm 0.38$	$3.34 \pm 0.36$	$1.30 \pm 0.09$
DEU	$4.75 \pm 0.49$	$4.75 \pm 0.49$	$3.30 \pm 0.33$	$1.28 \pm 0.10$
NDL	$4.08 \pm 0.35$	$4.08 \pm 0.35$	$2.97 \pm 0.24$	$1.17 \pm 0.09$
DNM	$4.00 \pm 0.52$	$4.00 \pm 0.52$	$2.87 \pm 0.45$	$1.11 \pm 0.14$
SPA	$2.66 \pm 0.25$	$2.66 \pm 0.25$	$2.46 \pm 0.26$	$0.87 \pm 0.11$
FIN	$2.25 \pm 0.32$	$2.25 \pm 0.32$	$2.02 \pm 0.33$	$0.69 \pm 0.13$

$4.6 \pm 0.52$  [19]), при колебаниях в диапазоне 4–13 аллелей, среднее число эффективных аллелей  $N_e$  –  $3.26 \pm 0.11$ , а средний информационный индекс Шеннона  $I$  достиг  $1.26 \pm 0.03$ , что также подтверждается ранее полученными данными у заводчиков Тюмени [19]. Полученные ранее данные по образцам черно-пестрой и холмогорской пород скота показали от 4 до 15 аллельных вариантов на локус (в 15-локусной панели) [20]. Для сравнения среднее число эффективных аллелей для локусов в популяции Хабаровского края составило 4.5 [21], а для уральской популяции Свердловской области – 4.02 [22].

Анализ отечественных стад голштинского скота в России показал высокий уровень полиморфизма: среднее количество наблюдаемых аллелей  $N_a$  превышало шесть аллелей на локус и варьировало в диапазоне 6.3–7.9.

Количество аллелей с частотой  $p > 5\%$  в отечественных стадах было в диапазоне 4.1–4.5, количество эффективных аллелей – в интервале 3.3–3.7, а информационный индекс  $I > 1.36$  (см. табл. 1).

При рассмотрении полиморфизма животных по локусам наибольшее в целом  $N_a$  (ср.) выявлено для TGLA53 ( $8.15 \pm 0.88$ ), число  $N_e$  (ср.) для того же локуса достигло  $5.22 \pm 0.40$ , наиболее высокое значение индекса  $I$   $1.74 \pm 0.11$  отмечено у TGLA53 (рис. 1). На втором месте по величине индекса  $I$  был локус TGLA227 ( $1.62 \pm 0.10$ ) с  $N_a$  (ср.)  $7.69 \pm 0.71$  и  $N_e$  (ср.)  $4.37 \pm 0.31$  (в Тюмени на первом месте с частотой 0.905 [19]). По аналогичным исследованиям, проведенным по 13-локусной панели, для черно-пестрой породы величина  $N_a$  (ср.) составляла  $6.57 \pm 0.32$ , а  $N_e$  (ср.)  $3.74 \pm 0.20$  [23]. Отдельно для голштинской породы величины параметров были около 1.8 и 2.1 с величиной индекса  $I$  около 1.5 соответственно [24]. В популяции Урала наблюдали величины  $N_a$  (ср.) и  $N_e$  (ср.) 10.9 и 4.03 соответственно [22].

Проведены сравнительные исследования частот встречаемости аллелей ( $p$ ) у групп скота отечественной и зарубежной селекции (см. рис. 1). Сравнительный анализ показал, что распределение аллелей внутри каждого локуса у животных, принадлежащих разным стадам и генетическим группам, имеет неравномерный и сходный характер [25, 26]. В каждом локусе были отобраны наиболее частые аллели и составлен перечень аллельных вариантов, имеющих среднее значение  $p > 0.5$  (табл. 2). При частоте  $p$  (ср.)  $> 0.6^1$  более 70% животных стада имеют такие аллели в составе гомо- или гетерозиготного генотипа. По результатам анализа распределения частот в каждом локусе выделены

аллельные варианты, которые встречаются у животных голштинской породы наиболее часто.

Для сравнения: в небольшой выборке коров голштинской породы Тюменского региона наибольшей частотой отличались аллели 117 (ETH3), 248 (SPS115) – 0.529–0.600, а в целом частота основных аллелей достигала 0.60 [19]. Примечательно, что в стадах родственной черно-пестрой породы скота Белоруссии наибольшая эффективность при контроле происхождения отмечена для локуса TGLA53 (0.848), а наименьшая – для локуса ETH3 (0.470) [27]. В популяции Хабаровского края частота встречаемости варьировала от 0.020 до 0.587, а с наибольшей частотой 0.538–0.587 встречались аллели: 117/ETH3, 117/TGLA126, 248/SPS115. При этом число эффективных аллелей колебалось от 2.4 в локусе SPS115 до 14.3 в локусе TGLA122 [21].

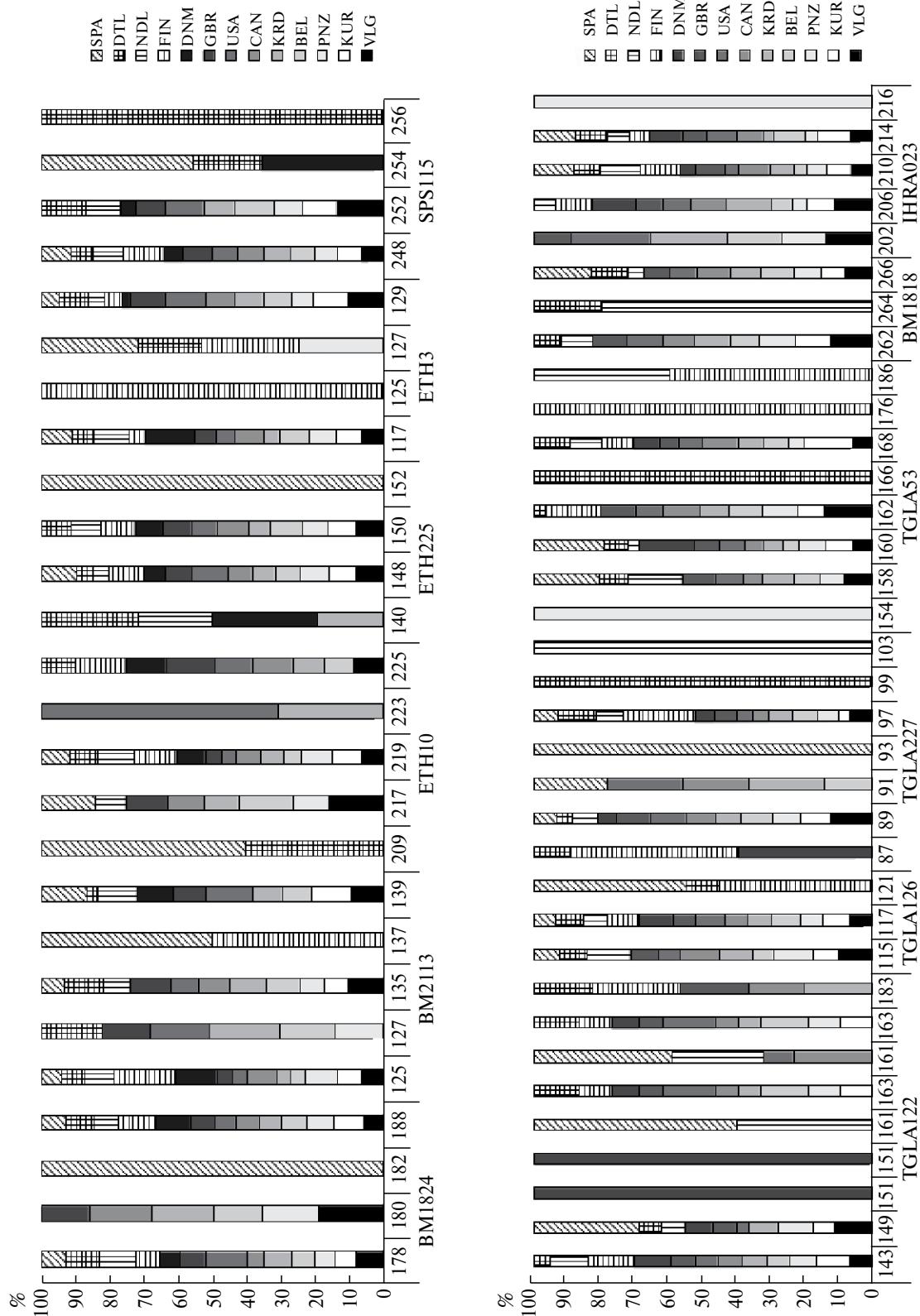
Приватные аллели были выявлены в шести из 13 изученных групп с величиной  $p$  0.5–3%: VLG (аллель 141 п.н./локус BM2113, 179/TGLA122); PNZ (121/BM2113, 95/TGLA227, 180/TGLA53, 216/INRA023 ( $p$  17.5%); BEL (186/BM1824, 123,147/BM2113, 138/ETH225, 113/113,109,125/TGLA126, 79/TGLA227, 156,178,182/TGLA53, 258/BM1818); KRD (159/TGLA122); CAN (145/ETH225); GBR (109/ETH3, 198/INRA023).

Таким образом, всего выявлено 22 приватных аллеля в 10 локусах из 12 (табл. 3, 4). Количество приватных аллелей на локус варьирует от одного до четырех (TGLA126), при этом аллель 216/INRA023 встречался более часто. В группе KRD найден только один аллель 159/TGLA122  $p$  0.5%, в VLG – два 141/BM2113 и 179/TGLA122  $p < 0.4\%$ , в PNZ – четыре 121/BM2113  $p$  0.5%, 95/TGLA227  $p$  1.5%, 180/TGLA53  $p$  3.3% и 216/INRA023  $p$  17.5%. В стаде BEL выявлено наибольшее количество аллелей с  $p < 2\%$ .

Вероятно, вышеуказанные аллели произошли от материнской породы, на основе которой было создано стадо животных с дальнейшим поглотительным скрещиванием с голштинскими быками. Поскольку приватные аллели в исследованных группах имеют низкие значения  $p < 0.4$ –3.3%, они не могут являться критерием породной принадлежности животных. В сравнении с полученными данными частоты условно приватных аллелей тюменской популяции составляли 0.014–0.043 [19].

При рассмотрении группы канадских быков североамериканской селекции CAN выявлен один приватный аллель 145/ETH225  $p$  1.4%, два аллеля 109/ETH3 и 198/INRA023 – у быков GBR  $p$  1.3%, в группах быков USA и DEU приватные аллели выявлены не были. Причина отсутствия приватных аллелей в группах быков могла быть связана с тем, что быки (или их ближайшие родственники) были использованы при формировании отечественных стад.

<sup>1</sup>выделены курсивом.



**Рис. 1.** Частоты аллелей микросателлитных локусов.

Таблица 2. Преобладающие аллели микросателлитных локусов

Локус	Аллель, п.н.	Генетические группы									
		VLG	KUR	PNZ	BEL	KRD	CAN	USA	GBR	NDL	DEU
BM1824	188	0.4	0.61	0.55	0.52	0.43	0.47	0.42	0.5	0.5	0.56
ETH10	219	0.39	0.52	0.56	0.3	0.4	0.43	0.27	0.29	0.64	0.5
ETH3	117	0.35	0.41	0.44	0.47	0.25	0.44	0.31	0.34	0.57	0.33
	129	0.47	0.48	0.29	0.4	0.37	0.4	0.56	0.48	0.21	0.39
SPS115	248	0.54	0.63	0.58	0.6	0.68	0.66	0.64	0.74	0.79	0.56
TGLA126	117	0.52	0.65	0.52	0.7	0.57	0.53	0.71	0.55	0.57	0.67
BM1818	266	0.46	0.42	0.48	0.59	0.54	0.59	0.48	0.44	0.29	0.64
INRA023	210	0.26	0.31	0.26	0.16	0.32	0.39	0.18	0.36	0.5	0.33

При сравнительном анализе распределения приватных аллелей оказалось, что в отечественных стадах количество аллелей (обнаруженных у < 25% генетических групп) было выше, чем у животных зарубежной селекции. Количество аллелей варьировало от 0.33 до 0.42.

Количество аллелей (обнаруженных у < 50% групп) у коров отечественной селекции также было выше, чем у животных зарубежной (размах 1.08–1.34).

Например, в выборках голштин в Сирийской Арабской Республике при среднем числе аллелей на локус, равном 6.18, число приватных аллелей было равно четырем [28]. В небольших нескольких стадах Казахстана приватных аллелей было два, а число редких аллелей в породе составило 14.9%, что меньше, чем в черно-пестрой породе, – 14 и 32.8% соответственно [29].

По результатам анализа, средний уровень наблюдаемой гетерозиготности  $H_o$  в группах был  $0.68 \pm 0.02$ , что соответствует несмещенному значению ожидаемой гетерозиготности  $H_e$   $0.68 \pm 0.01$ . Среднее значение индекса фиксации  $F$  отрицательное:  $-0.07 \pm 0.019$ , так как наблюдалось неродственное скрещивание [30].

Сравнительный анализ гетерозиготности групп по локусам выявил высокий уровень  $H_o$  у пяти локусов: BM2113 0.78, TGLA122 0.80, TGLA227 0.81, TGLA53 0.75 и INRA023 0.78, а уровень ожидаемой гетерозиготности  $H_e$  по этим локусам был 0.70, 0.74, 0.75, 0.77 и 0.70 соответственно.

Ранее в популяции Хабаровского края наибольшие уровни наблюдаемой и ожидаемой гетерозиготности (0.857 и 0.930 соответственно) были обнаружены в локусе TGLA122, а наименьшие (0.530 и 0.586 соответственно) – в локусе SPS115, при этом

Таблица 3. Перечень и частота приватных аллелей

Генетическая группа	Локус	Аллель, п.н.	Частота, $p$
VLG	BM2113	141	0
	TGLA122	179	0
PNZ	BM2113	121	0.01
	TGLA227	95	0.02
	TGLA53	180	0.03
	INRA023	216	0.18
	BM1824	186	0.01
BEL	BM2113	123	0.02
	BM2113	147	0.01
	ETH225	138	0.01
	ETH3	113	0.02
	TGLA126	109	0.02
	TGLA126	125	0.01
	TGLA227	79	0.01
	TGLA53	156	0.01
	TGLA53	178	0.01
	TGLA53	182	0.01
KRD	BM1818	258	0.01
	TGLA122	159	0.01
	ETH225	145	0.01
	ETH3	109	0.01
CAN	INRA023	198	0.01

средний уровень наблюдаемой и ожидаемой гетерозиготности составлял примерно 0.700 [21].

Для сравнения в выборке коров Тюменского региона наибольшим уровнем  $H_o$  характеризовался локус TGLA227 – 0.905, наименьшим – BM1824 – 0.667. В то же время  $H_e$  была наибольшей по локусу TGLA53 и составляла 0.941 [19].

В отдельных выборках голштинской и черно-пестрой пород популяций Казахстана уровень ожидаемой гетерозиготности варьировал от 0.596

в локусе SPS115 до 0.867 в TGLA227, от 0.603 в локусе BM1818 до 0.844 в BM2113 соответственно. Средний показатель ожидаемой гетерозиготности для этих пород составил 0.715 и 0.738 соответственно [29], что близко к средним показателям  $0.721 \pm 0.036$  и  $0.662 \pm 0.081$ , полученным позднее [29], где гетерогенность черно-пестрой породы была ниже.

В популяции Урала в голштинизированной черно-пестрой породе средние уровни  $H_o$  и  $H_e$

**Таблица 4.** Распространение приватных и локальных аллелей

Генетическая группа	Приватные аллели, % (ср.) ± SD	Локальные аллели, % $p \leq (ср.) 25\% \pm SD$	Локальные аллели, % $p \leq (ср.) 50\% \pm SD$
VLG	0.16 ± 0.11	0.42 ± 0.19	1.33 ± 0.33
KUR	–	0.33 ± 0.14	1.25 ± 0.33
PNZ	0.33 ± 0.14	0.33 ± 0.14	1.25 ± 0.33
BEL	1.00 ± 0.28	0.33 ± 0.19	1.08 ± 0.29
KRD	0.08 ± 0.08	0.33 ± 0.14	1.25 ± 0.33
CAN	0.08 ± 0.08	0.25 ± 0.13	1.08 ± 0.34
USA	–	0.17 ± 0.11	1.08 ± 0.36
GBR	0.17 ± 0.11	0.42 ± 0.19	1.33 ± 0.33
DNM	–	–	0.50 ± 0.23
FIN	–	–	0.08 ± 0.08
NDL	–	0.17 ± 0.11	0.58 ± 0.23
DEU	–	0.25 ± 0.13	0.92 ± 0.34
SPA	–	0.08 ± 0.08	0.33 ± 0.19

**Таблица 5.** Показатели гетерозиготности и значений индекса фиксации

Группа	$H_o$	$H_e$	$uH_e$	$F$
VLG	0.708 ± 0.025	0.706 ± 0.025	0.709 ± 0.025	–0.005 ± 0.016
KUR	0.694 ± 0.033	0.673 ± 0.029	0.676 ± 0.029	–0.031 ± 0.021
PNZ	0.652 ± 0.029	0.697 ± 0.027	0.700 ± 0.028	0.062 ± 0.029
BEL	0.774 ± 0.041	0.685 ± 0.035	0.689 ± 0.035	–0.131 ± 0.025
KRD	0.699 ± 0.032	0.702 ± 0.031	0.706 ± 0.031	0.005 ± 0.014
CAN	0.663 ± 0.045	0.684 ± 0.031	0.695 ± 0.031	0.039 ± 0.039
USA	0.667 ± 0.033	0.664 ± 0.033	0.668 ± 0.034	–0.006 ± 0.017
GBR	0.657 ± 0.038	0.691 ± 0.033	0.700 ± 0.033	0.048 ± 0.033
DNM	0.650 ± 0.093	0.572 ± 0.069	0.635 ± 0.077	–0.119 ± 0.075
FIN	0.542 ± 0.096	0.427 ± 0.071	0.569 ± 0.095	–0.293 ± 0.096
NDL	0.700 ± 0.054	0.632 ± 0.036	0.682 ± 0.040	–0.121 ± 0.070
DEU	0.776 ± 0.070	0.664 ± 0.032	0.704 ± 0.033	–0.152 ± 0.073
SPA	0.667 ± 0.112	0.531 ± 0.060	0.708 ± 0.080	–0.248 ± 0.156

составляли 0.73 и 0.72 соответственно при величине индекса  $F$  –0.004 [17].

Индекс  $F$  имел невысокие отрицательные значения: BM2113 –0.130, TGLA122 –0.091, TGLA227 –0.090, INRA023 –0.130, что так же, как на Урале, свидетельствовало о неродственном скрещивании. По наиболее полиморфному локусу TGLA53 отмечено положительное значение  $F$  0.080, что по сравнению с величиной  $H_e$  0.84 указывало на инбридинг (табл. 5). Дефицит гетерозигот популяций голштинского скота в других исследованиях наблюдался в диапазоне 0.01–0.091 [31, 32].

Анализ данных показал, что величина  $H_o$  (ср.) по всем локусам и группам национальной селекции варьировала от 0.652 в стаде PNZ и до 0.774 в группе BEL. Наблюдаемый уровень  $H_o$  в отечественных стадах в целом соответствует ожидаемым показателям гетерозиготности. Показатель  $H_e$  изменялся в интервале 0.68–0.71, а индекс  $F$  –0.131–0.005. Наибольшее отклонение значения  $F$  в отрицательную сторону свидетельствовало об аутбридинге в хозяйствах BEL (–0.139). В хозяйствах PNZ и KRD

ближкие к нулю положительные значения  $F$  (0.062 и 0.005) указывали на родственное скрещивание.

В группах быков североамериканской селекции уровень  $H_o$  составлял 0.66 (CAN) и 0.67 (USA) и соответствовал ожидаемым показателям 0.70 и 0.67 соответственно. Согласно величине  $F$  у быков CAN наблюдался инбридинг (0.039), а в группе USA – аутбридинг (–0.006).

В группах DNM и FIN более выражены отрицательные значения  $F$  (–0.119 и –0.293), среди них аутбридинг был выше, чем у быков NDL и DEU, где величина  $F$  достигла значений –0.151 и –0.152 соответственно.

Попарный анализ генетических расстояний по Нью [33, 34] между группами позволил оценить степень удаленности стад из разных регионов России от групп быков зарубежных стран (табл. 6).

Между отечественными стадами голштинской породы генетические расстояния ( $L$ ) изменились от 0.036 (VLG – KRD) до 0.074 (PNZ – KRD). Оказалось, что величины  $L$  ближе при сопоставлении животных из отечественных стад и ряда

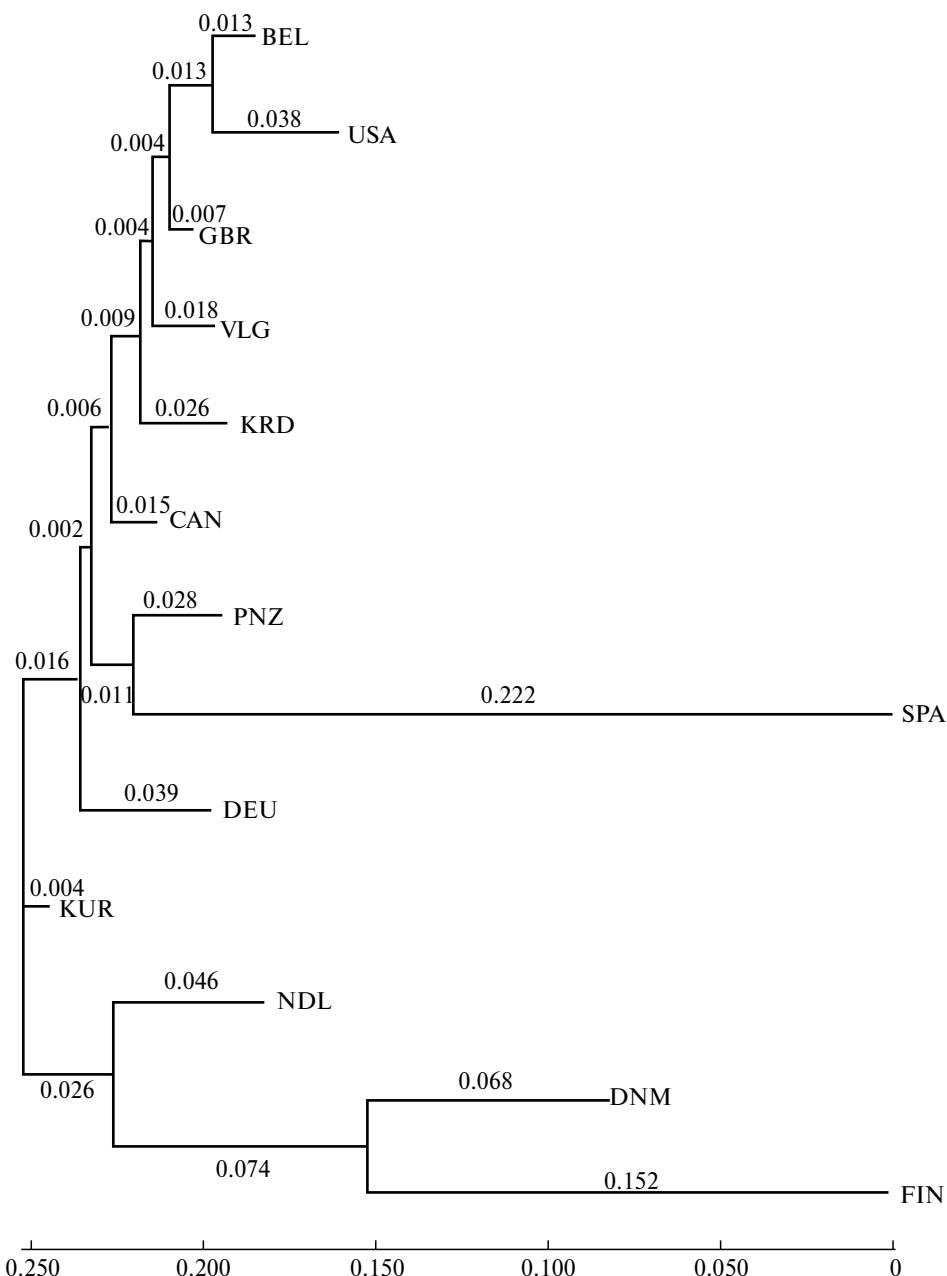
**Таблица 6.** Попарная матрица генетических расстояний, сходства и подразделения стад

Группы	VLG	KUR	PNZ	BEL	KRD	CAN	USA	GBR	DNM	FIN	NDL	DEU	SPA
VLG	–	0.96	0.96	0.96	0.97	0.98	0.94	0.99	0.77	0.76	0.95	0.98	0.97
KUR	0.05 0.01	–	0.95	0.96	0.94	0.98	0.95	0.97	0.84	0.84	0.99	0.99	0.96
PNZ	0.05 0.01	0.06 0.01	–	0.93	0.94	0.96	0.90	0.95	0.80	0.81	0.97	0.98	0.99
BEL	0.04 0.01	0.04 0.01	0.07 0.02	–	0.94	0.97	0.97	0.98	0.80	0.78	0.93	0.99	0.96
KRD	0.04 0.01	0.06 0.02	0.07 0.02	0.06 0.02	–	0.98	0.94	0.99	0.75	0.78	0.95	0.99	0.98
CAN	0.02 0.01	0.02 0.01	0.04 0.01	0.03 0.01	0.02 0.01	–	0.94	0.99	0.81	0.83	0.98	0.99	0.99
USA	0.06 0.02	0.06 0.02	0.11 0.02	0.04 0.01	0.06 0.02	0.07 0.02	–	0.97	0.76	0.75	0.90	0.96	0.92
GBR	0.01 0.01	0.03 0.01	0.06 0.01	0.03 0.01	0.01 0.01	0.01 0.01	0.03 0.01	–	0.78	0.81	0.98	0.99	0.99
DNM	0.26 0.10	0.18 0.10	0.22 0.10	0.22 0.09	0.29 0.10	0.22 0.10	0.27 0.10	0.25 0.10	–	0.83	0.86	0.83	0.81
FIN	0.28 0.14	0.18 0.12	0.22 0.13	0.25 0.14	0.24 0.13	0.19 0.13	0.29 0.14	0.21 0.13	0.19 0.15	–	0.89	0.86	0.78
NDL	0.05 0.03	0.01 0.02	0.03 0.03	0.07 0.04	0.05 0.03	0.02 0.02	0.11 0.04	0.03 0.03	0.16 0.10	0.12 0.12	–	0.99	0.99
DEU	0.02	0.01	0.02	0.01	0.01	0.01	0.05	0.01	0.19	0.15	0.01	–	0.99
SPA	0.03	0.04	0.01	0.04	0.02	0.01	0.09	0.01	0.21	0.25	0.01	0.01	–

зарубежных стран, чем при сопоставлении отечественных стад между собой. Значения  $L$  между VLG и быками CAN, DEU были 0.022 и 0.018 соответственно, в то время как между VLG и группами KUR, PNZ, BEL и KRD были от 0.036 до 0.046. Необходимо отметить существенные расстояния между аборигенными стадами и быками DNM и FIN, которые различались в 10 и более раз (0.175–0.287), так как, вероятно, быки NDL, DNM и FIN в меньшей степени использовались при формировании отечественных стад.

Анализ молекулярной вариансы (AMOVA, analysis of mol. variance,  $\sigma^2_G$ ) показал, что 96% вариансы представлено вариабельностью внутри отечественных стад и 4% дисперсии обеспечивают отличия между ними. При рассмотрении составляющих дисперсии по всей совокупности животных доля математического ожидания возрасала до 98%, а 2% определяли различия между группами разной селекции.

Анализ генетической идентичности по Нею [30, 33] ( $L_i$ ) показал, что уровень генетического



**Рис. 2.** Дендрограмма генетических расстояний.

сходства [34, 35] между группами отечественного скота составлял от 0.929 до 0.965. Сходство отечественного поголовья с животными зарубежной селекции подтверждается высокими значениями  $L_i > 0.9$ , за исключением групп быков DNM и FIN, величины  $L_i$  быков DNM и FIN составили 0.7–0.8 (см. табл. 6). По-видимому, направление селекции быков DNM и FIN, а также интенсивность их использования при формировании российского массива голштинского скота отличаются от быков североамериканской и немецкой селекции.

Генетические различия между группами и внутри них с уровнем инбридинга связаны по методике Фишера–Райта через показатели избытка или недостатка гетерозиготности. Коэффициент подразделения  $F_{st}$  внутри подгрупп по отношению к общей измеренной дифференциации между ними определяется как доля общей генетической изменчивости.

Коэффициент  $F_{st}$  был  $> 0$  и для отечественного скота находился в диапазоне 0.009–0.017. Величины  $F_{st}$  для быков USA были  $< 0.025$ , а для быков DEU  $< 0.030$ . Высокие значения  $F_{st}$  были у быков DNM (0.080–0.101) и FIN (0.121–0.156), что указывало на отличие этих групп от остального поголовья. Межпородное скрещивание черно-пестрой и голштинской пород способствует интродукции аллелей и повышению уровня генетического разнообразия [36] в противоположность чистопородному разведению, при этом наблюдалось снижение генетических различий между породами со значениями  $F_{st}$  с 0.058 до 0.026, а  $L$  – с 0.306 до 0.123 [36].

Для оценки генетической структуры породы с включением в анализ быков других стран использован метод оценки генетических отношений  $L$  (рис. 2) по принципиальным координатам РСоА, при этом между группами выявлены два отдельных кластера. В один кластер вошли животные VLG, KUR и PNZ вместе с быками DEU, в другой – группы BEL и KRD вместе с быками CAN, USA и GBR, отдельно оказались быки NDL, DNM и FIN. На основании анализа РСоА голштинский скот отечественной селекции подразделялся на две группы в зависимости от использования быков европейской или североамериканской селекции, имеющих наряду с общими признаками свои генетические особенности. Результаты филогенетического анализа по Нею [33, 34] подтвердили расположение популяций на основе показателей генетических расстояний, отражающих долю генетических вариаций (рис. 3). Величина рассчитанных генетических расстояний составляла от 0.002 до 0.222. Ранее выявляли генетические расстояния от 0.057 до 0.453 [31, 32]. Показано уменьшение генетических дистанций между анализируемыми породами за счет введения голштинской крови.

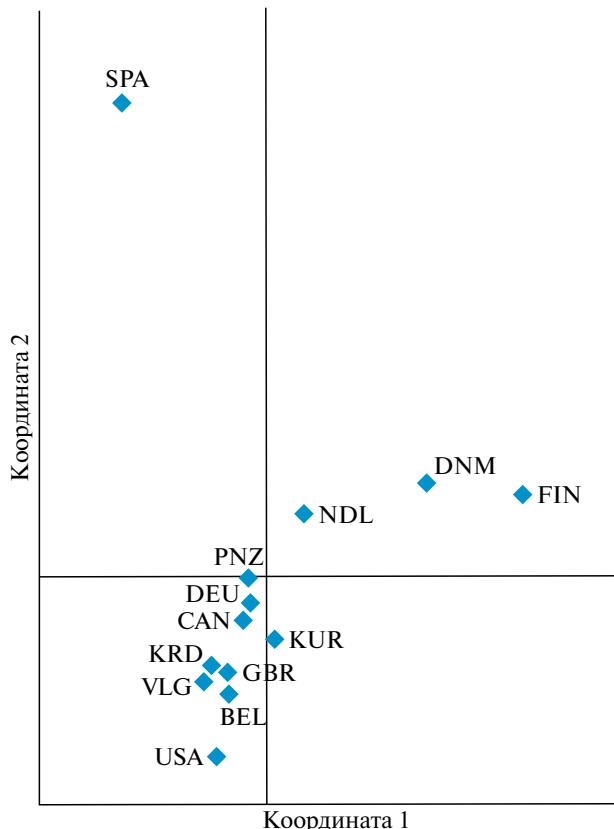


Рис. 3. Диаграмма распределения животных в системе принципиальных координат.

Доля генетических вариаций составила по первой оси 55%, по второй – 23%, третьей – 12%. Результаты размещения групп указывали на генетическую близость национальных групп VLG, KUR и PNZ с быками DEU, а животные групп BEL и KRD были генетически близки с быками североамериканской и европейской селекции – USA, CAN, DEU. Генетически отдаленными от отечественного голштинского скота оказались быки DNM и FIN (европейской селекции).

Таким образом, впервые проведено генотипирование и профилирование голштинской породы в целом на большом объеме поголовья, значительно превышающем аналогичные исследования. Подтверждено, что животные голштинской породы тесно связаны друг с другом независимо от страны происхождения, т. е. мировая популяция голштинского скота – это генетическая единица. В то же время расчеты показали на внутреннее подразделение голштейн в зависимости от происхождения и направления селекции. По результатам генотипирования установлен перечень аллелей, наиболее часто встречающихся в стадах и присутствующих у большинства животных. Наличие этих аллелей является типичным признаком голштейн и генетической характеристики породы. Отечественные

стада отличались от зарубежных большей степенью генетического разнообразия, наличием приватных аллелей, характеризующих исходное маточное поголовье, дифференцированы в зависимости от использования быков и примыкают к поголовью из Германии, США и Канады.

На основании филогенетического анализа скот национальной селекции подразделялся на два кластера в зависимости от интенсивности покрытия быками европейской или североамериканской селекции. Показано, что от отечественных стад более отдалены генетически быки Дании и Финляндии. Показатели генетических расстояний и генетической идентичности визуализированы при построении филогенетических деревьев и принципиальной системы координат, выявив отношения между национальными стадами и быками североамериканской и европейской селекции в зависимости от интенсивности их использования, подразделив быков Дании и Финляндии.

Работа поддержана Государственным заданием Минсельхоза России по теме “Проведение исследований по генетической характеристики голштинского скота отечественной селекции на основе STR-маркеров генома с целью разработки критериев определения породной принадлежности животных” № 082-00208-23-00 (НИОКР 123030600068-9) на 2023 г. (eLIBRARY ID: 54082343).

Исследование одобрено Этическим комитетом ФБУН Всероссийского научно-исследовательского института племенного дела Минсельхоза России, (02.03.2024, протокол № 1).

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

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. *Abblondi M., Sabbioni A., Stocco G. et al.* Genetic diversity in the italian holstein dairy cattle based on pedigree and snp data prior and after genomic selection // FRONT VET SCI. 2022. V. 8. <https://doi.org/10.3389/fvets.2021.773985>
2. *Meuwissen T., Hayes B., Goddard M.* Genomic selection: A paradigm shift in animal breeding // Animal Frontiers. 2016. V. 6. № 1. P. 6–14. <https://doi.org/10.2527/af.2016-0002>
3. *Мещеров Р.К., Мещеров Ш.Р., Ходыков В.П., Никулкин Н.С.* Породные и племенные ресурсы крупного рогатого скота голштинской породы черно-пестрой масти в Российской Федерации: реалии и перспективы // Агрозоотехника. 2023. Т. 6. № 2. С. 1–10. <https://doi.org/10.15838/alt.2023.6.2.6>
4. *Adamov N., Mickov L., Petkov V., Adamov M.* Microsatellite markers for pedigree verification in cattle // Macedonian J. Anim. Sci. 2011. V. 1. № 1. P. 9–15.
5. *MacHugh D.E., Loftus R.T., Cunningham P., Bradley D.G.* Genetic structure of seven European cattle breeds assessed using 20 microsatellite markers // Anim. Genetics. 1998. V. 29. № 5. P. 333–340. <https://doi.org/10.1046/j.1365-2052.1998.295330.x>
6. *Agung P.P., Saputra F., Zein M.S.A. et al.* Genetic diversity of Indonesian cattle breeds based on microsatellite markers // Asian-Australas J. Anim. Sci. 2019. V. 32. № 4. P. 467–476. <https://doi.org/10.5713/ajas.18.0283>
7. *Joshi P., Vyas P., Kashyap S.K.* Molecular characterization of Nagori cattle using microsatellite markers // J. Pharmacognosy and Phytochemistry. 2018. V. 7. № 2. P. 3250–3252.
8. *Gororo E., Chatiza F.P., Chidzwondo F., Makuza S.M.* Is neutral genetic diversity related to quantitative variation in semen traits in bulls? // Reprod. Domest. Anim. 2021. V. 56. № 10. P. 1293–1301. <https://doi.org/10.1111/rda.13991>
9. *Lenstra J.A., Groeneveld L.F., Eding H. et al.* Molecular tools and analytical approaches for the characterization of farm animal genetic diversity // Anim. Genetics. 2012. V. 43. № 5. P. 483–502. <https://doi.org/10.1111/j.1365-2052.2011.02309.x>
10. *Калашников А.Е., Ялуга В.Л.* Изменение встречаемости аллелей ЕАВ-локуса групп крови у скота холмогорской породы вследствие голштинизации // Изв. Кабардино-Балкарского гос. аграрного ун-та им. В. М. Кокова. 2022. № 3(37). С. 66–78. <https://doi.org/10.55196/2411-3492-2022-3-37-66-78>
11. *Калашников А.Е., Хрунова А.И., Калашников В.Е., Рыжова Н.Г.* Изменение гаплотипов локусов групп крови холмогорской породы при поглотительном скрещивании // Достижения и актуальные проблемы генетики, биотехнологии и селекции животных. Витебск, 2021. С. 27–29. EDN: IXBBNP
12. *Тяпугин С.Е., Калашникова Л.А., Новиков А.А. и др.* Генетическая идентификация сельскохозяйственных и диких видов животных // Методическое пособие. Лесные Поляны: ФГБНУ “ВНИИплем”, 2021. С. 1–98. EDN: DWIYJB
13. *Heslot N., Yang H.P., Sorrells M.E., Jannink J.L.* Genomic selection in plant breeding: A comparison of models // Crop Sci. 2012. V. 52. № 1. P. 146–160. <https://doi.org/10.2135/cropsci2011.06.0297>
14. *Kemp S.J., Brezinsky L., Teale A.J.* A panel of bovine, ovine and caprine polymorphic microsatellites // Anim. Genetics. 1993. V. 24. № 5. P. 363–365. <https://doi.org/10.1111/j.1365-2052.1993.tb00341.x>
15. *McKay S.D., Schnabel R.D., Murdoch B.M. et al.* An assessment of population structure in eight breeds of cattle using a whole genome SNP panel // BMC

- Genetics. 2008. V. 20. № 9. P. 37.  
<https://doi.org/10.1186/1471-2156-9-37>
16. Дунин И.М., Тяпугин С.Е., Калашникова Л.А. и др. Генофонд пород молочного скота в России: состояние, перспективы сохранения и использования // Зоотехния. 2019. № 5. С. 1–2.  
<https://doi.org/10.25708/ZT.2019.18.21.001>
17. 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/10.31857/S001667582104010X>
18. Филиппова Н.П., Павлова Н.И., Корякина Л.П. и др. Микросателлитный анализ якутского скота // Животноводство и кормопроизводство. 2018. Т. 101. № 4. С. 58–63. EDN: VQTBVK
19. Часовщикова М.А. Генетическая характеристика голштинской породы крупного рогатого скота с использованием микросателлитных ДНК-маркеров // Изв. ОГАУ. 2019. Т. 2. № 76. С. 191–193.
20. Калашникова Л.А., Хабибрахманова Я.А., Ганченкова Т.Б. и др. Генетическая характеристика крупного рогатого скота с использованием микросателлитов // Зоотехния. 2016. № 2. С. 9–11. EDN: VOJQUT
21. Шукюрова Е.Б., Лукашина А.А., Бузько А.Н. Генетическая характеристика голштинского крупного рогатого скота по ДНК-микросателлитам // Вестн. ДВО РАН. 2020. № 4(212). С. 47–52.  
<https://doi.org/10.37102/08697698.2020.212.4.008>
22. Зиновьевна Н.А., Гладырь Е.А. Генетическая экспертиза сельскохозяйственных животных: применение тест-систем на основе микросателлитов // Достижения науки и техники АПК. 2011. № 9. С. 19–20. EDN: OGBSUP
23. Гладырь Е.А., Горелов П.В., Маурчева В.Н. и др. Оценка результативности тест-системы на основе микросателлитов в проведении ДНК-экспертизы крупного рогатого скота // Достижения науки и техники АПК. 2011. № 8. С. 51–54. EDN: OBGJWR
24. Харзинова В.Р., Карпушкина Т.В. Генетическая экспертиза сельскохозяйственных животных на основе анализа микросателлитов // Проблемы и перспективы научно-инновационного обеспечения агропромышленного комплекса регионов: Сб. докл. Междунар. научно-практ. конф. Курск, 2019. С. 567–569. EDN: KRPNDO
25. Rege J.E.O., Lipner M.E. African animal genetic resources: Their characterisation, conservation and utilisation // Proc. Res. Planning Workshop. 1992. P. 19–21.
26. Spelman R.J. Utilisation of marker assisted selection in the New Zealand dairy industry. A thesis presented in partial fulfilment of the requirements for the degree of master of agricultural sci. in animal breeding and genetics at Massey University. Massey Univ. 1995. P. 1–95.
27. Танана Л.А., Епишко О.А., Глинская Н.А. STR-локусы в контроле происхождения крупного рогатого скота белорусской черно-пестрой породы // С.-х. журн. 2014. Т. 2. № 7. С. 204–207. EDN: TBISCV
28. Мохаммад А.А., Бакай А.В. Сравнительная характеристика генетической структуры крупного рогатого скота сирийской породы Шами с голштинской и абердин-ангусской породами // Главный зоотехник. 2019. № 12. С. 23–30. EDN: TYJTOQ
29. Карымсаков Т.Н., Гладырь Е.А., Нурбаев С.Д. и др. Сравнительная характеристика аллелофонда крупного рогатого скота трех родственных пород черно-пестрого корня, разводимых в Республике Казахстан // Молочное и мясное скотоводство. 2017. № 3. С. 11–14. EDN: YTDPGH
30. Кузнецов В.М. F-статистики Райта: оценка и интерпретация // Проблемы биологии продуктивных животных. 2014. Т. 4. С. 80–104. EDN: TFRDMN
31. Эрнест Л.К., Зиновьевна Н.А., Коновалова Е.Н. и др. Изучение влияния прилития крови голштинского скота на изменение генофонда крупного рогатого скота отечественных пород с использованием ДНК-микросателлитов // Зоотехния. 2007. № 12. С. 2–4. EDN: JWZSMD
32. Абдельманова А.С., Волкова В.В., Доцев А.В., Зиновьевна Н.А. Характеристика генетического разнообразия современной и архивной популяций крупного рогатого скота черно-пестрой породы с использованием микросателлитных маркеров // Достижения науки и техники АПК. 2020. № 2. С. 34–38.  
<https://doi.org/10.24411/0235-2451-2020-10207>
33. Pamilo P., Nei M. Relationships between gene trees and species trees // Mol. Biol. Evol. 1988. V. 5. № 5. P. 568–583.  
<https://doi.org/10.1093/oxfordjournals.molbev.a040517>
34. Rzhetsky A., Nei M. A simple method for estimating and testing minimum-evolution trees // Mol. Biol. Evol. 1992. V. 9. № 5. P. 945–967.
35. Nei M., Tajima F., Tateno Y. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data // J. Mol. Evol. 1983. V. 19. № 2. P. 153–170.  
<https://doi.org/10.1007/BF02300753>
36. Зиновьевна Н.А., Гладырь Е.А., Багиров В.А., Брем Г. Динамика биоразнообразия отечественного черно-пестрого скота под воздействием кроссбрининга // Вавил. журн. генетики и селекции. 2015. Т. 19. № 2. С. 72–75. EDN: UCRCFNH

## Genetic Evaluation of Holstein Cattle Makes Use of Microsatellite DNA Markers

**L. A. Kalashnikova<sup>1,\*</sup>, T. B. Ganchenkova<sup>1</sup>, N. V. Ryzhova<sup>1</sup>, Y. A. Khabibrakhmanova<sup>1</sup>,  
I. E. Bagal<sup>1</sup>, I. Y. Pavlova<sup>1</sup>, A. E. Kalashnikov<sup>1</sup>**

<sup>1</sup>*All-Russian Research Institute of Animal Breeding Ministry of Agriculture of Russian Federation,  
Moscow region, Lesnye Polyany, 141212 Russia*

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

The results of a research of polymorphism of 12 micro-satellite loci in Holstein cattle from an ordinal number of regions of Russia and external countries were presented. The average number of alleles per locus was  $5.43 \pm 0.19$ , with variation in the range of 4–13 alleles, the average number of effective alleles was  $3.26 \pm 0.11$ . A list of 29 most frequent alleles has been fixed. 22 private alleles were identified, and the frequency of private alleles was 0.004–0.033. It has been demonstrated that the amount of locally alleles in domestic herds is higher than in animals of external selection. The mean level of observed heterozygosity for all loci hold at  $0.681 \pm 0.017$  and varied in the range of 0.65–0.78 for a fixation index of  $-0.131 \pm 0.005$ . Genetic length between herds of domestic selection were  $< 0.074$ . It was revealed that groups of cow herds come down into two clusters. The first cluster included animals from three areas of Russia, associated with bulls from Germany and the Netherlands, and the second cluster included individuals from other two provinces closest to the males of Canada, the USA and GB. At once, the oxen of Denmark and Finland founds themselves in a separate cluster. The basis of this work was to evaluate the allele reservoir of Holstein cattle of domestic selection and determine the genetic profile of the breed by STR markers.

**Keywords:** genetic differentiation, Holstein breed, Holstein-Friesian (black and white), micro-satellites, markers, breed appurtenance.

## ПОЛИМОРФИЗМ ГЕНА *NIPAL1* rs135173498 КАК МАРКЕР ЭМБРИОНАЛЬНОЙ ЛЕТАЛЬНОСТИ У АУЛИЕКОЛЬСКОГО СКОТА

© 2025 г. Е. В. Белая<sup>1,\*</sup>, Е. А. Климанова<sup>2</sup>, В. М. Норкина<sup>2</sup>, И. С. Бейшова<sup>3</sup>

<sup>1</sup>Белорусский государственный педагогический университет им. Максима Танка, Минск, 220030  
Республика Беларусь

<sup>2</sup>Новосибирский государственный аграрный университет, Новосибирск, 630039 Россия

<sup>3</sup>Западно-Казахстанский аграрный технический университет им. Жангира хана, Уральск, 090009  
Республика Казахстан

\*e-mail: Kolyuchka005@rambler.ru

Поступила в редакцию 26.06.2024 г.

После доработки 09.09.2024 г.

Принята к публикации 12.09.2024 г.

Установлено, что у крупного рогатого скота аулиекольской породы SNP *G/A* rs135173498, ассоциированный с живой массой при рождении (*p*-value = 2,71E-06), характеризуется отсутствием живорожденных телят, гомозиготных по редкому генотипу при значительном объеме выборки в 497 животных. При этом телята с гетерозиготным генотипом не характеризовались снижением живой массы при рождении относительно значений этого показателя по выборке в целом. Проанализированы возможные механизмы развития фенотипического эффекта SNP rs135173498 гена *NIPAL1* как фактора эмбриональной летальности у аулиекольского крупного рогатого скота и, предположительно, у пород казахской белоголовой, абердин-ангусской и шароле.

**Ключевые слова:** GWAS, QTL, 3'-нетранслируемые участки, rs135173498/*NIPAL1*, генетический маркер, полногеномный поиск ассоциаций, эмбриональная летальность, крупный рогатый скот.

**DOI:** 10.31857/S0016675825010068 **EDN:** VEPCSW

Снижение плодовитости крупного рогатого скота является основной причиной выбраковки коров и фактором ухудшения экономической эффективности скотоводства. Тем не менее длительное время эмбриональная смертность не рассматривалась как причина снижения репродуктивной функции крупного рогатого скота. Только благодаря достижениям биотехнологии в области трансплантации эмбрионов было показано, что до 70% всех воспроизводительных потерь у крупного рогатого скота происходит за счет эмбриональной смертности, которая наступает в течение первого эмбрионального периода. Генетические исследования сельскохозяйственных животных свидетельствуют о присутствии в популяциях рецессивных аллелей, приводящих к абортируемости эмбрионов и снижающих жизнеспособность новорожденных. Поэтому в настоящее время выявление мутаций, приводящих к нарушению функций белкового продукта гена (LoF, loss-of-function) и снижающих в том числе фертильность крупного рогатого скота, имеет такое же важное значение, как поиск мутаций, ассоциированных с хозяйственными-полезными признаками.

К наиболее известным LoF-мутациям, обусловливающим наследственные аномалии и вызывающим эмбриональную и раннюю постэмбриональную смертность, в молочных породах крупного рогатого скота относятся LoF-мутации, вызывающие дефицит уридинмонофосфатсингтазы (DUMPS), дефицит лейкоцитарной адгезии (BLAD), комплексный порок позвоночника (CVM), брахиспинальный синдром (BY) у представителей голштинской породы, а также синдром Вивера у бурой швейцарской породы. Основная опасность отрицательных последствий таких генетических мутаций связана с использованием системы искусственного осеменения коров, при котором от одного быка – носителя летальных генов – могут быть получены десятки тысяч потомков, несущих в своем генотипе мутантный аллель, частота которого может резко увеличиться в течение малого числа поколений, что приведет к существенному повышению процента гибели племенного молодняка [1].

Новый подход, названный картированием гомозиготности и основанный на генотипировании десятков тысяч SNP (single nucleotide polymorphism) с помощью чипов средней и высокой плотности

(например, чип Bovine SNP50 v.2 BeadChip производства “Illumina”, США, позволяющий проводить анализ 54609 SNP), позволил выявить существенное количество LoF-мутаций, ассоциированных с эмбриональной летальностью [2].

Данный метод позволяет идентифицировать регионы, характеризующиеся отсутствием одного из гомозиготных генотипов. Подобные участки гена рассматриваются в качестве кандидатов для локализации генов, несущих LoF-мутации. Предложенный метод позволяет осуществлять идентификацию летальных генетических дефектов в течение очень короткого периода и при наличии ограниченного числа случаев их проявления. Выявленные посредством картирования гомозиготности дефекты получили название гаплотипов фертильности. При этом, если для диагностики генетических дефектов используется ДНК-чип, то результаты интерпретируются как идентификация соответствующих гаплотипов (ННД, ННВ, ННС, НН0), в то время как при прямом анализе причинной мутации речь идет об идентификации собственно генетических дефектов.

В настоящее время установлены гаплотипы, ассоциированные с фертильностью коров различных пород крупного рогатого скота: голштинской, джерсейской, бурой швицкой, монбельярдской, айрширской, симментальской, у красного нордического скота [3–8]. Так, в настоящее время у коров голштинской породы установлены гаплотипы, ассоциированные с фертильностью, такие как НН13 в гене *KIR2DS1* (нонсенс-вариант p.Gln159), НН21 в гене *NOTCH3* (делеция p.Cys44del), делеция в *RIOX1* (p.Ala133\_Glu142del), а также миссенс-вариант в гене *PCDH15* (p.Leu) [9]. У айрширской породы крупного рогатого скота с нарушением фертильности и эмбриональной смертностью плодов ассоциирован гаплотип АН1, обусловленный мутацией, нарушающей сплайсинг в гене *UBE3B* [10, 11]. У крупного рогатого скота красной датской породы в 2016 г. обнаружена LoF-мутация в первом экзоне гена холинергического рецептора никотиновой субъединицы бета 1 *CHRN1* (OMIA 002022-9913) на ВТА 19, представляющая собой делецию одной пары нуклеотидов (c.55delG), приводящую к образованию преждевременного стоп-кодона (p.Ala19Profs47) во втором экзоне гена и нарушению синтеза 96.0% белка. Соответствующий синдром получил название “множественный артогрипоз”. И по мнению Н. А. Зиновьевой, принимая во внимание экспоненциальный рост числа исследований в области структурной и функциональной геномики животных, в частности крупного рогатого скота, в ближайшее время следует ожидать открытия новых генетических дефектов [12].

Ранее нами был проведен полногеномный поиск ассоциаций с признаками живой массы на разных возрастах у крупного рогатого скота казахской

белоголовой и аулиекольской породы [13]. Среди прочих выявленных генов-кандидатов внимание привлекает расположенный на ВТА6 QLT-ассоциированный ген ENSBTAG00000009423 *NIPAL1* (NIPA like domain containing 1 – первый содержащий NIPA-подобный домен), содержащий SNP *G/A* rs135173498 пограничной значимости ассоциации с признаком живой массы при рождении (*p*-value = 2,71E-06). Особенностью данного гена-кандидата явилось отсутствие живорожденных телят, гомозиготных по редкому генотипу при значительном объеме выборки в 497 животных. При этом телята с гетерозиготным генотипом не характеризовались снижением живой массы при рождении относительно значений этого показателя по выборке в целом, а также практически не отличались от гомозигот дикого типа по данному показателю. Эти наблюдения наводят на мысль о возможной эмбриональной летальности у телят с гомозиготным генотипом. Цель настоящей работы – проведение анализа влияния rs135173498 на биологическую функцию белкового продукта гена *NIPAL1* для оценки возможности применения rs135173498/*NIPAL1* в качестве генетического маркера эмбриональной летальности у крупного рогатого скота аулиекольской и, возможно, других пород.

## МАТЕРИАЛЫ И МЕТОДЫ

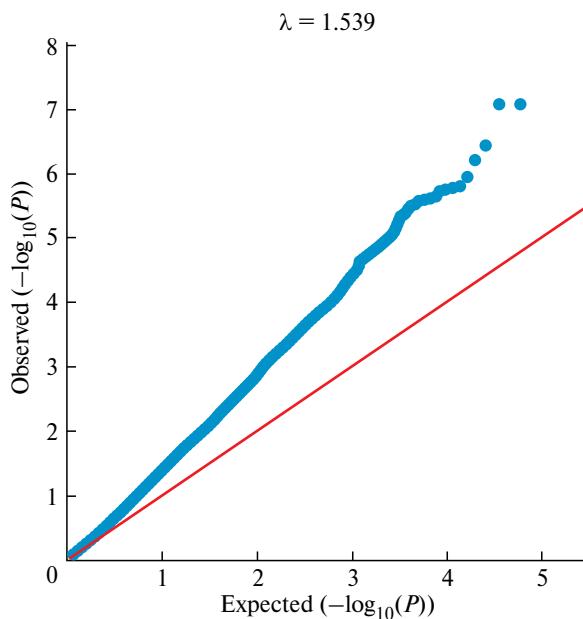
Объектом исследования явился полиморфизм rs135173498 гена *NIPAL1*, ассоциированный по данным SNP-чипирования с живой массой при рождении крупного рогатого скота аулиекольской породы. Предметом исследования явились возможные биологические механизмы развития эмбриональной летальности у гомозиготных носителей rs135173498. Информация об ассоциации rs135173498/*NIPAL1* с живой массой при рождении, а также частотах встречаемости аллелей и генотипов у телят аулиекольской породы получена в результате полногеномного поиска ассоциаций, детально описанного в наших более ранних работах [13–15]. В частности, выделение геномной ДНК и генотипирование проводили в Neogen Agrigenomics, Lincoln, NE, USA) в соответствии с протоколом производителя GeneSeek GGP Bovine 150 K, который содержит 150000 SNP (Neogen Corporation Company, Lincoln, NE, USA). Полученные результаты в GenomeStudio преобразовывались в формат Plink (.bed, .bim, .fam) [16].

Ассоциативный поиск был выполнен с использованием Plink – набора инструментов для анализа ассоциаций всего генома, а именно расчета линейной регрессионной зависимости, а также коэффициентов детерминации. Значимость коэффициентов регрессии полиморфных сайтов оценивали с использованием теста Вальда. Потенциальными значимыми считали сайты, *p*-значение которых

было выше частного от деления уровня значимости (0.05) на количество полиморфных сайтов, а также  $p$ -значение теста Вальда у них было меньше 0.05. Оценку проведенного анализа ассоциаций проводили с использованием визуальной оценки квартиль-квартиль графиков, а также коэффициента  $\lambda$  (рис. 1). Графики квантиль-квантиль (QQ plot), построенные для оценки достоверности модели при уровне доверия  $p \leq 0.005$ , демонстрируют значения  $\lambda = 1.539$ , что позволяет рассматривать доверительный уровень модели при  $p = 0.005$  как приемлемый.

Полученные полиморфные сайты были аннотированы идентификаторами rs с использованием базы данных SNPChimpV3, с помощью Ensembl, базы данных QTL крупного рогатого скота и Panther Classification System. Форматирование полученных файлов, а также визуализация производились с помощью интерпретируемого языка программирования Python (библиотеки pandas, numpy, scipy, matplotlib, seaborn).

Информация о распределении белка *NIPAL1* в клетке получена с помощью базы GeneCards. Данные по локализации в GeneCards формируются на основании высокопроизводительных скринингов на основе микроскопии, прогнозов первичной последовательности и автоматического интеллектуального анализа текста. Оценка достоверности локализации присваивается на основе типа свидетельства и источника и визуализируется как в таблице, так и в схематическом изображении ячейки. Шкала достоверности имеет цветовую кодировку:



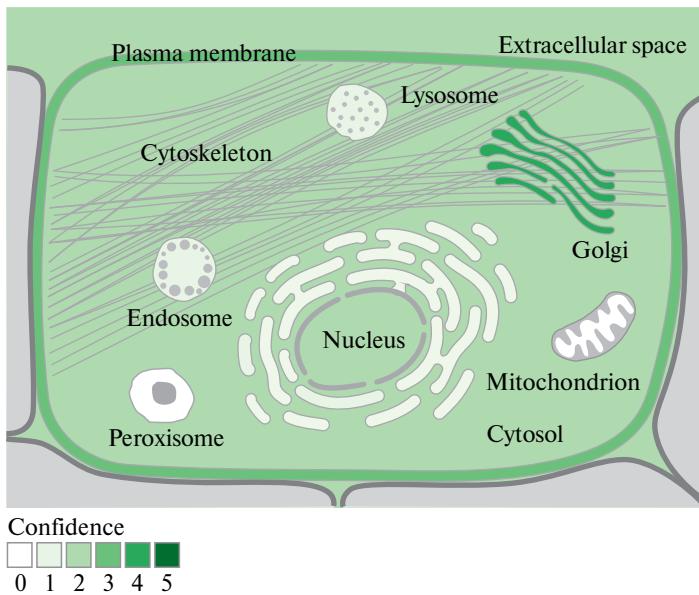
**Рис. 1.** Графики квантиль-квантиль (QQ plot). Пояснение к осям: Y — наблюдаемые значения  $p$  (преобразованы с помощью обратного десятичного логарифма  $p$ , [15]).

от светло-зеленого (1) для низкой достоверности до темно-зеленого (5) для высокой достоверности. Белый цвет (0) указывает на отсутствие признаков локализации. Анализ белок-белковых взаимодействий белкового продукта гена *NIPAL1* выполнен с помощью базы данных string-db.org.

## РЕЗУЛЬТАТЫ

Анализ биологических процессов, в которых участвует белковый продукт гена *NIPAL1*, позволил установить, что единственным биологическим процессом, контролируемым данным геном, является локализация. Полученные данные согласуются с результатами анализа распределения белка *NIPAL1* в клетке, полученными с помощью базы GeneCards (<https://www.genecards.org/>) и демонстрирующими аппарат Гольджи и плазматическую мембрану как основное место локализации белка *NIPAL1* (рис. 2), что согласуется также с литературными данными [17]. Подобно эндоплазматическому ретикулуму, аппарат Гольджи является местом созревания и сортировки белков, получаемых из эндоплазматического ретикулума (ЭПР). И белок *NIPAL1*, будучи клеточным компонентом мембран, участвующим в трансмембранных переносах неорганических катионов и, в частности, катионов магния, играет одну из ключевых ролей в реализации метаболических функций аппарата Гольджи, что подтверждается широким спектром его белок-белковых взаимодействий.

Карта белок-белковых взаимодействий белка *NIPAL1* дана на рис. 3. На нем отражаются прогнозируемые взаимодействия интересующего белка на основании экспериментальных данных и таких баз, как Uniprot, Ensembl и др. Выстраиваемые схемы взаимодействий имеют цветовую кодировку линий, соединяющих белки. Источники информации представлены желтыми (анализ научной литературы), голубыми (из баз данных) и розовыми (экспериментальные данные) линиями. Прогнозируемые взаимодействия отмечены красными (слияние генов, при котором две или более последовательностей ДНК, кодирующих разные белки, объединяются в одну новую последовательность, которая транскрибируется и транслируется в химерный белок) и синими линиями (совместная встречаемость связанных белков у вида — представление о встречаемости показывает наличие или отсутствие связанных белков у разных видов за счет вычисления предсказания белок-белковых взаимодействий, основанных на анализе частоты совместного появления двух белков в различных протеомах). Дополнительно черными линиями указаны ко-экспрессия (представление ко-экспрессии показывает гены, которые совместно экспрессируются у того же или у других видов), фиолетовыми — гомологичные белки, светло-зеленым — отображение списка важных групп взаимодействия



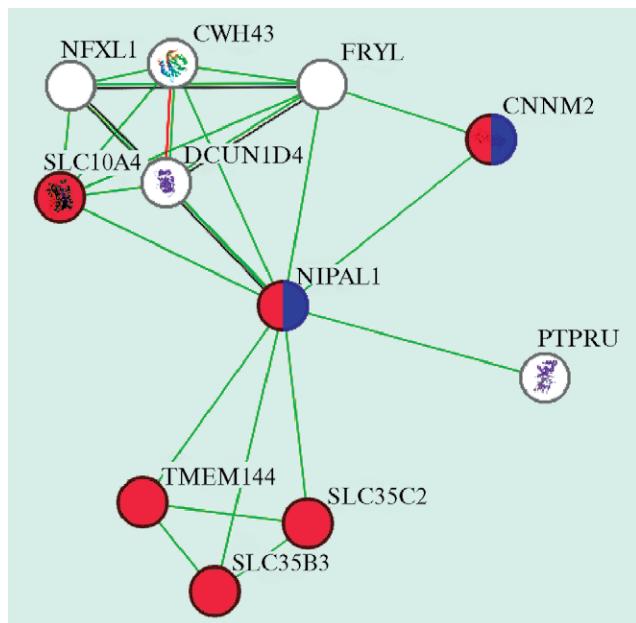
Compartment	Confidence
Golgi apparatus	4
Plasma membrane	3
Lysosome	1
Cytosol	1
Endosome	1
Endoplasmic reticulum	1
Nucleus	1
Mitochondrion	1
Cytoskeleton	1
Extracellular	1

**Рис. 2.** Внутриклеточная локализация NIPAL1 (GeneCards). Шкала достоверности имеет цветовую кодировку: от светло-зеленого (1) для низкой достоверности до темно-зеленого (5) для высокой достоверности; белый цвет (0) указывает на отсутствие признаков локализации.

белков на основании извлеченного из рефератов научной литературы.

Из схемы, приведенной на рис. 3, можно выделить десять наиболее вероятных участников белок–белкового взаимодействия протеина NIPAL1 у крупного рогатого скота по данным string-db.org. Из них шесть (SLC10A4, TMEM144, SLC35B3, SLC35C2, CNNM2 и NIPAL1) выполняют

функцию трансмембранных транспортных белков (выделены красным). NIPAL1 совместно с CNNM2 являются трансмембранными переносчиками ионов магния (отмечены фиолетовым цветом). Для белков NFXL1, CWH43, FRYL, DCUN1D4, PTPRU совместная с NIPAL1 функция не установлена (окрашены в белый). Однако по NFXL1, CWH43, FRYL, DCUN1D4 имеется публикация об ассоциации данных белков с массой тела у кур [18]. Прогнозируемое взаимодействие наблюдается у белков CWH43 и DCUN1D4 (красная линия). Для белков NFXL1, FRYL, DCUN1D4 и NIPAL1 указана совместная экспрессия (черные соединительные линии). Также все рассмотренные на рис. 3 белки связаны светло-зеленой линией, что свидетельствует о встречаемости данных белков в публикациях.



**Рис. 3.** Белок–белковые взаимодействия белкового продукта гена NIPAL1, построенные с помощью string-db.org (у крупного рогатого скота).

## ОБСУЖДЕНИЕ

Понимание фенотипических последствий полиморфизма rs135173498, приводящих к эмбриональной летальности, лежит не только в биохимических и цитологических характеристиках белка NIPAL1, но и в его локализации в 3'-области (3'-UTR) гена *NIPAL1*, которая представляет собой некодирующий участок мРНК, расположенный после кодирующей области, и является жизненно важным элементом, регулирующим экспрессию генов [19]. Наиболее известно, что 3'-нетранслируемые области (3'-UTR) информационных РНК (мРНК) регулируют процессы, “заявленные” на мРНК, такие как локализация мРНК, стабильность мРНК и трансляция. Кроме того, 3'-UTR могут устанавливать 3'-UTR-опосредованные белок–белковые

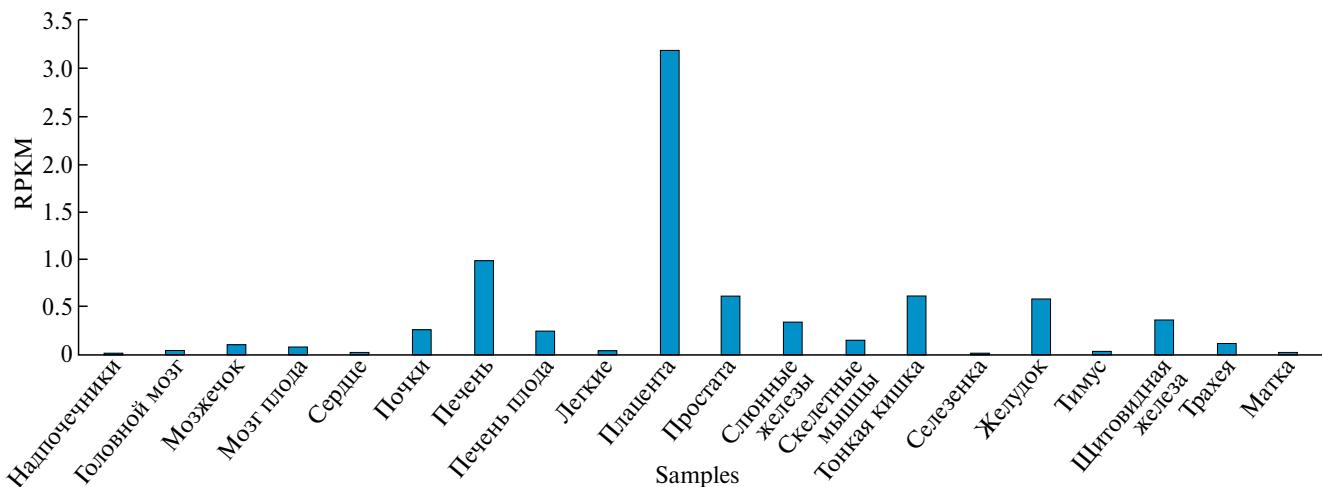


Рис. 4. Уровень экспрессии NIPAL1 в 20 различных тканях человека (цит. по [20]).

взаимодействия (ББВ) и таким образом могут передавать генетическую информацию, закодированную в 3'-UTR, белкам. В работе Mayg C. показано, что эта функция регулирует различные свойства белков, включая образование белковых комплексов или посттрансляционные модификации, но также предполагается, что они изменяют конформации белков. Следовательно, передача информации, опосредованная 3'-UTR, может регулировать свойства белка, которые не закодированы в аминокислотной последовательности [20].

По данным авторов, функции 3'-UTR реализуются на трех уровнях: 1) 3'-UTR регулируют процессы на уровне информационной РНК (мРНК), 2) 3'-UTR регулируют свойства белка, опосредуя 3'-UTR-зависимые белок–белковые взаимодействия (ББВ), 3) 3'-UTR регулируют различные свойства белка, опосредуя 3'-UTR-зависимые белок–белковые взаимодействия [20]. На уровне информационной РНК (мРНК) 3'-UTR регулируют посттрансляционные процессы путем связывания с различными эффекторными белками. РНК-связывающие белки связываются с 3'-UTR мРНК и привлекают различные эффекторные белки. Далее происходит привлечение экзосомы, которое в свою очередь приводит к дестабилизации мРНК, тогда как моторный белок приводит к регуляции локализации мРНК с помощью движения по микротрубочкам.

Регуляция свойств белка через 3'-UTR-зависимые ББВ позволяет альтернативным 3'-UTR определять альтернативные функции белков, несмотря на то, что кодируемые белки имеют идентичные аминокислотные последовательности. В этом случае 3'-UTR-зависимые ББВ опосредуются только длинной изоформой 3'-UTR, а не короткой изоформой 3'-UTR. Механизм 3'-UTR-регуляции различных свойств белка путем образования

3'-UTR-зависимого белкового комплекса заключается в 3'-UTR-зависимых посттрансляционных модификациях и в 3'-UTR-зависимом сворачивании белка [20]. В пользу гипотезы об ассоциации rs135173498/*NIPAL1* с эмбриональной летальностью свидетельствует также факт, что ген *NIPAL1* (по данным NCBI) экспрессируется в 20 типах тканей человека. При этом наибольший уровень экспрессии наблюдался в плаценте (рис. 4) [21].

Анализ распределения аллелей и генотипов rs135173498/*NIPAL1* в настоящем исследовании позволил установить, что относительная частота минорного аллеля *A* в исследованной популяции составляет  $0.058 \pm 0.0$ , что значительно ниже по сравнению с MAF<sub>E</sub>, предоставляемой базой данных Ensembl.org. В то же время относительная частота аллеля дикого типа *G* в изученной популяции аулиекольского скота достигла  $0.942 \pm 0.0$ . Наблюдаемое распределение животных с генотипами *AA/GA/GG* составило  $0/57/436$ , что практически соответствовало теоретически рассчитанному равновесному распределению  $2/54/437$  ( $\chi^2 = 1.86$ ). Однако, как было отмечено выше, ни одного живорожденного теленка с генотипом rs135173498/*NIPAL1/AA* обнаружено не было.

Нами установлено, что полиморфизм rs135173498 расположен в участке генома 6:64247306/68186013, породоспецифичном для казахской белоголовой породы, которая является одной из трех пород-прародителей аулиекольской (казахская белоголовая, шароле и абердин-ангусская) [22]. Следовательно, данная мутация может приводить к эмбриональной летальности также у телят этих трех пород, что обуславливает интерес к дальнейшему расширению исследования rs135173498/*NIPAL1* как возможного маркера эмбриональной летальности у аулиекольской, казахской белоголовой, шароле и абердин-ангусской пород крупного рогатого скота.

Будучи локализованным в участке 3'-UTR, полиморфизм rs135173498/*NIPAL1*, вероятно, влияет на посттрансляционные белок–белковые взаимодействия мРНК, снижая эффективность функции белка *NIPAL1* и приводя к нарушениям работы комплекса Гольджи в клетках плаценты; и, таким образом, способствует повышенной abortируемости гомозиготных зародышей с генотипом rs135173498/*NIPAL1/AA*, что, возможно, является причиной наблюдалемого распределения генотипов в исследованной выборке. Полученные данные позволяют рассматривать rs135173498/*NIPAL1* как потенциальный генетический маркер эмбриональной летальности у крупного рогатого скота аулиекольской породы. Тем не менее целесообразно дополнить исследование анализом генотипов abortируемого материала. С учетом происхождения аулиекольской породы от трех пород-прародительниц можно предположить наличие аналогичного фенотипического эффекта генотипа rs135173498/*NIPAL1/AA* также у казахской белоголовой, шароле и абердин-ангусской пород. Или у некоторых из них. В таком случае проведение скрининговых мероприятий, направленных на элиминацию животных носителей аллеля rs135173498/*NIPAL1/A* из поголовья и исключение их из селекционного процесса, может принести значительный экономический эффект за счет сокращения abortируемости эмбрионов.

Работа выполнена в рамках АР19680057 “Изучение хозяйствственно-полезных признаков и характеристика генофонда крупного рогатого скота казахской белоголовой породы методом ресеквенирования и транскриптомного анализа”, Республика Казахстан.

Исследование одобрено локальной Комиссией по биологической этике НАО “Западно-Казахстанский аграрный технический университет им. Жан-гир хана”, (4 марта 2024 г., протокол № 2).

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. Шайдуллин Р.Р., Фаизов Т.Х., Ганиев А.С. Характер распространения летальных генов у молочного скота // Уч. записки КГАВМ им. Н. Э. Баумана. 2015. № 2. С. 242–244.  
<https://doi.org/cyberleninka.ru/article/n/harakter-rasprostraneniya-letalnyh-genov-u-molochnogo-skota>
2. Charlier C., Coppieters W., Rollin F. et al. Highly effective SNP-based association mapping and management of recessive defects in livestock // Nat. Genet. 2008. V. 40. P. 449–454.  
<https://doi.org/10.1038/ng.96>
3. Van Raden P.M., Olson K.M., Null D.J. et al. Harmful recessive effects on fertility detected by absence of homozygous haplotypes // J. Dairy Sci. 2011. № 94. P. 6153–6161.  
<https://doi.org/10.3168/jds.2011-4624>
4. Fritz S., Capitan A., Djari A. et al. Detection of haplotypes associated with prenatal death in dairy cattle and identification of deleterious mutations in gart, shbg and slc37a2 // PLoS One. 2013. V. 8. № 6.  
<https://doi.org/10.1371/journal.pone.0065550>
5. Cooper T.A., Wiggans G.R., Van Raden P.M. et al. Genomic evaluation of ayrshire dairy cattle and new haplotypes affecting fertility and stillbirth in holstein, brown swiss and ayrshire breeds // <http://aipl.arsusda.gov>
6. Sonstegard T.S., Cole J.B., Van Raden P.M. et al. Identification of a nonsense mutation in cwc15 associated with decreased reproductive efficiency in jersey cattle // PLoS One. 2013. V. 8. № 1. P. 1–6.  
<https://doi.org/10.1371/journal.pone.0054872>
7. McClure M., Kim E., Bickhart D. et al. Fine mapping for weaver syndrome in brown swiss cattle and the identification of 41 concordant mutations across nrcam, pnpla8 and cttnbp2 // PLoS One. 2013. V. 8. № 3. P. 1–16.  
<https://doi.org/10.1371/journal.pone.0059251>
8. Kadri N.K., Sahana G., Charlier C. et al. A 660-Kb deletion with antagonistic effects on fertility and milk production segregates at high frequency in nordic red cattle: Additional evidence for the common occurrence of balancing selection in livestock // PLoS Genet. 2014. V. 10. № 1. P. 1–11.  
<https://doi.org/10.1371/journal.pgen.1004049>
9. Häfliger I.M., Spengeler M., Seefried F.R. et al. Four novel candidate causal variants for deficient homozygous haplotypes in holstein cattle // Sci. Rep. 2022. № 12. P. 1–13.  
<https://doi.org/10.1038/s41598-022-09403-6>
10. Cooper T.A., Wiggans G.R., Null D.J. et al. Genomic evaluation, breed identification, and discovery of a haplotype affecting fertility for ayrshire dairy cattle // J. Dairy Sci. 2014. V. 97. № 6. P. 3878–3882.  
<https://doi.org/10.3168/jds.2013-7427>
11. Гладырь Е.А., Терновская О.А., Костюнина О.В. Скрининг гаплотипа фертильности АН1 айрширской породы крупного рогатого скота Центрального и Северо-Западного регионов России // АгроЗооТехника. 2018. Т. 1. № 4. С. 1–13.  
<https://doi.org/10.15838/alt.2018.1.4.1>
12. Зиновьева Н.А. Гаплотипы фертильности голштинского скота // С.-х. биология. 2016. Т. 51. № 4. С. 423–435.  
<https://doi.org/10.15389/agrobiology.2016.4.423rus>
13. Белая Е.В., Наметов А.М., Шамшидин А.С. Полногеномный поиск ассоциаций с QTL мясной продуктивности у скота казахской белоголовой и аулиекольской

- пород // Главный зоотехник. 2022. № 7. С. 3–11.  
<https://doi.org/10.33920/sel-03-2207-01>
14. Белая Е.В. Биологические функции породоспецифичных SNP-маркеров мясной продуктивности у крупного рогатого скота казахской белоголовой и аулиекольской пород // Генетика и разведение животных. 2022. № 2. С. 33–39.  
<https://doi.org/10.31043/2410-2733-2022-2-33-39>
15. Белая Е.В., Бейшова И.С., Селионова М.И. и др. Полногеномный поиск QLT-ассоциированных SNP для прогнозирования наследственного потенциала продуктивности у казахского белоголового скота // Вестник АПК Ставрополья. 2022. № 3(47). С. 18–25.  
<https://doi.org/10.31279/2222-9345-2022-11-47-18-25>
16. Purcell S.M., Neale B., Todd-Brown K. et al. PLINK: A tool set for whole-genome association and population-based linkage analyses // Am. J. Hum. Genet. 2007. V. 81. P. 559–575.  
<https://doi.org/10.1086/519795>
17. Manialawy Y., Khan S.R., Bhattacharjee A. et al. The magnesium transporter nipal1 is a pancreatic islet-expressed protein that conditionally impacts insulin secretion // J. Biol. Chem. 2020. V. 295.
18. Dadousis C., Somavilla A., Ilska J.J. et al. A genome-wide association analysis for body weight at 35 days measured on 137,343 broiler chickens // Genet. Sel. Evol. 2021. V. 53(1). P. 1–14.  
<https://doi.org/10.1186/s12711-021-00663-w>
19. Wei W., Gao W., Li Q. et al. Comprehensive characterization of posttranscriptional impairment-related 3'-utr mutations in 2413 whole genomes of cancer patients // NPJ Genom. Med. 2022. V. 7. P. 1–12.  
<https://doi.org/10.1038/s41525-022-00305-0>
20. Mayr C. What are 3' utrs doing? // Cold. Spring. Harb. Perspect Biol. 2019. № 10. P. 1–15.  
<https://doi.org/10.1101/cshperspect.a034728>
21. Duff M.O., Olson S., Wei X. et al. Genome-wide identification of zero nucleotide recursive splicing in drosophila // Nature. 2015. №. 521(7552). P. 376–379.  
<https://doi.org/10.1038/nature14475>
22. Beishova I., Dossybayev K., Shamshidin A. et al. Distribution of homozygosity regions in the genome of Kazakh cattle breeds // Diversity. 2022. V. 14(4). P. 1–11.  
<https://doi.org/10.3390/d14040279>

## Polymorphism of the *NIPAL1* Gene rs135173498 as a Marker of Embryonic Lethality in Auliekol Cattle

A. V. Belya<sup>1, \*</sup>, E. A. Klimanova<sup>2</sup>, V. M. Norkina<sup>2</sup>, I. S. Beishova<sup>3</sup>

<sup>1</sup> Belarusian State Pedagogical University named after Maxim Tank, Minsk, 220030 Republic of Belarus

<sup>2</sup> Novosibirsk State Agrarian University, Novosibirsk, 630039 Russia

<sup>3</sup> Zhangir Khan West Kazakhstan Agrarian Technical University, Uralsk, 090009 Republic of Kazakhstan

\*e-mail: Kolyuchka005@rambler.ru

It was found that in Auliekol cattle, SNP G/A rs135173498 associated with live weight at birth (*p*-value = 2.71E-06) is characterized by the absence of live-born calves homozygous for the rare genotype with a significant sample size of 497 animals. At the same time, calves with a heterozygous genotype were not characterized by a decrease in live weight at birth relative to the values of this indicator for the sample as a whole. The article analyzes possible mechanisms for the development of the phenotypic effect of SNP rs135173498 of the *NIPAL1* gene as a factor of embryonic lethality in Auliekol cattle, and, presumably, in the Kazakh Whitehead, Aberdeen Angus and Charolais breeds.

**Keywords:** GWAS, QTL, 3' untranslated regions, rs135173498/*NIPAL1*, genetic marker, genome-wide association study, embryonic lethality, cattle.

## ИДЕНТИФИКАЦИЯ *Bos taurus* И *Bos grunniens* НА ОСНОВАНИИ SNP

© 2025 г. В. Н. Кипены<sup>1,\*</sup>, Ж. Т. Исакова<sup>2,4</sup>, М. М. Патрин<sup>3</sup>, К. Б. Чекиров<sup>4</sup>,  
К. А. Айтбаев<sup>2</sup>, А. Р. Карыпова<sup>4</sup>, М. И. Ирсалиев<sup>2</sup>

<sup>1</sup>Институт генетики и цитологии Национальной академии наук Беларусь, Минск, 220072 Республика Беларусь

<sup>2</sup>Научно-исследовательский институт молекулярной биологии и медицины, Бишкек, 720040

Кыргызская Республика <sup>3</sup>ООО “Максим Медикал”, Москва, 123423 Россия

<sup>4</sup>Кыргызско-Турецкий Университет “Манас”, Бишкек, 720038 Кыргызская Республика

\*e-mail: v.kipen@igc.by

Поступила в редакцию 02.07.2024 г.

После доработки 01.08.2024 г.

Принята к публикации 08.08.2024 г.

В статье исследованы образцы домашнего яка и трех пород крупного рогатого скота для оценки дифференцирующего потенциала полиморфных вариантов Chr4:68609356G>T (ген *JAZF1*), Chr14:35695388G>T (ген *SLCO5A1*) и Chr19:63181970C>G (ген *CEP112*). Подтверждена высокая точность (99.67%), специфичность (100%) и чувствительность (100%) предложенной тест-модели, состоящей из этих трех полиморфизмов, для идентификации домашних яков и крупного рогатого скота. Разработан быстрый и простой метод идентификации на основе данной модели с использованием технологии конкурентной аллель-специфической ПЦР (KASP), что позволяет существенно сократить временные и финансовые затраты на молекулярно-генетический анализ, а также снизить риск перекрестного загрязнения образцов.

**Ключевые слова:** *Bos taurus*, *Bos grunniens*, однонуклеотидный полиморфизм, идентификация, конкурентная аллель-специфическая ПЦР (KASP, Kompetitive allele specific PCR), генотипирование *in silico*.

**DOI:** 10.31857/S0016675825010079 **EDN:** VEOPTT

Домашний як (*Bos grunniens*) представляет собой доместицированную форму дикого яка (*Bos mutus*). Эти животные известны своей способностью выживать в экстремальных условиях, таких как низкие температуры, гипоксия и нехватка пищи, что делает их важным объектом для исследования молекулярно-генетических механизмов выживаемости и адаптации к высокогорным районам. Долгая и сложная история доместикации яка, а также физико-географические особенности его местообитания оказывают значительное влияние на генетическое разнообразие современных популяций этого животного. В настоящее время идет активное изучение *Bos grunniens* по всему миру, включая исследования различий в популяциях, генетического многообразия, геногеографии и филогенетических связей [1].

В горах Кыргызстана активно развивается якодомство, в то время как на более низких и средних высотах преобладает разведение крупного рогатого скота (КРС). В отличие от последнего, яки предпочитают использовать пастбищные корма низкого роста, что свидетельствует об их уникальных адаптационных возможностях. Домашний як имеет ограниченное географическое распространение

в высокогорных районах Центральной Азии, которое обусловлено его высокой адаптивностью к жизни в условиях гор и плато, подобно дикому родственнику – тибетскому яку.

Оценка генетического разнообразия яков и пород КРС, разводимых в Кыргызстане, в контексте возможной гибридизации между видами проводилась в нашем предыдущем исследовании [2]. Кроме того, в работе [3] были исследованы гибриды яка с КРС, по результатам которой в популяциях яка и гибридов первого поколения был выявлен видоспецифичный для яка паттерн из восьми ISSR-фрагментов. Также на основе микросателлитного анализа ранее был изучен аллелофонд яков и их гибридов с *Bos taurus*, который показал высокое генетическое разнообразие для гибридов первого поколения в сравнении с исходными видами [4].

Идентификация домашнего яка и КРС имеет важное значение для оценки возможной гибридизации между этими видами. Эта задача актуальна как для селекционеров, так и экспертов-криимиалистов, особенно в условиях ограниченной доступности референсных данных и высоких затрат на анализ STR-локусов. Оценка частот

распространения видоспецифичных SNP (Single Nucleotide Polymorphism) также представляет интерес с точки зрения динамики эволюционных изменений у домашних видов животных по сравнению с их дикими сородичами, что может быть связано с искусственным отбором, проводимым человеком.

Цель настоящего исследования – использование методов биоинформатики для выявления SNP с высоким потенциалом дифференциации, необходимых для идентификации принадлежности образцов как к биологическому виду *Bos taurus*, так и *Bos grinniens*. На основании проведенного анализа предлагается разработать тест-систему, включающую несколько SNP для идентификации домашнего яка и представителей КРС.

## МАТЕРИАЛЫ И МЕТОДЫ

**Биологические образцы.** Для молекулярно-генетического исследования были собраны образцы крови от 56 домашних яков (*Bos grinniens*) из высокогорного региона Калмак-Ашуу (Кочкорский р-н, Нарынская обл., Кыргызская Республика). Эта группа обозначена как YAK. Также были взяты образцы крови у 146 коров (*Bos taurus*) трех пород: абердин-ангусской ( $n = 45$ , группа ABR), голштинской ( $n = 51$ , группа HOL) и алатауской ( $n = 50$ , группаALA). Совокупность этих образцов составила выборку COW. Биологический материал был отобран сотрудниками опытных хозяйств и поступал в лабораторию для молекулярно-генетического анализа в пробирках типа Vacutainer (BD Vacutainer® Sodium Citrate Tubes), каждая из которых была снабжена биркой с описанием животного (видовая и породная принадлежность, возраст, пол).

**Выделение ДНК.** Образцы крови хранились при температуре  $-20^{\circ}\text{C}$  в течение 1–2 мес. либо при  $-80^{\circ}\text{C}$  при долговременном хранении. ДНК экстрагировали с использованием набора Blood-Animal-Plant DNA Preparation Kit (Jena Bioscience, Германия) в соответствии с инструкциями производителя. Количественная оценка выделенной ДНК проводилась с помощью NanoDrop 1000 (Thermo Fisher Scientific, США). Средняя концентрация ДНК составила  $72.4 \pm 23.3 \text{ нг/мкл}$ , при соотношении  $260/280 = 1.89 \pm 0.09$ .

**KASP.** Определение генотипа по SNP (версия генома *Bos\_taurus\_UMD\_3.1.1*, GCF\_000003055.6) Chr4:68609356G>T (*JAZF1*), Chr14:35695388G>T (*SLCO5A1*), Chr19:63181970C>G (*CEP112*) проводилось с применением технологии, основанной на конкурентной аллель-специфической ПЦР (KASP, Kompetitive allele specific PCR). Генотипирование проводили с использованием KASP Assay mix (KASP by Design, KBD) и KASP Master mix (LGC Biosearch Technologies, Великобритания; ООО “Максим Медикал”, РФ) в объеме 10 мкл

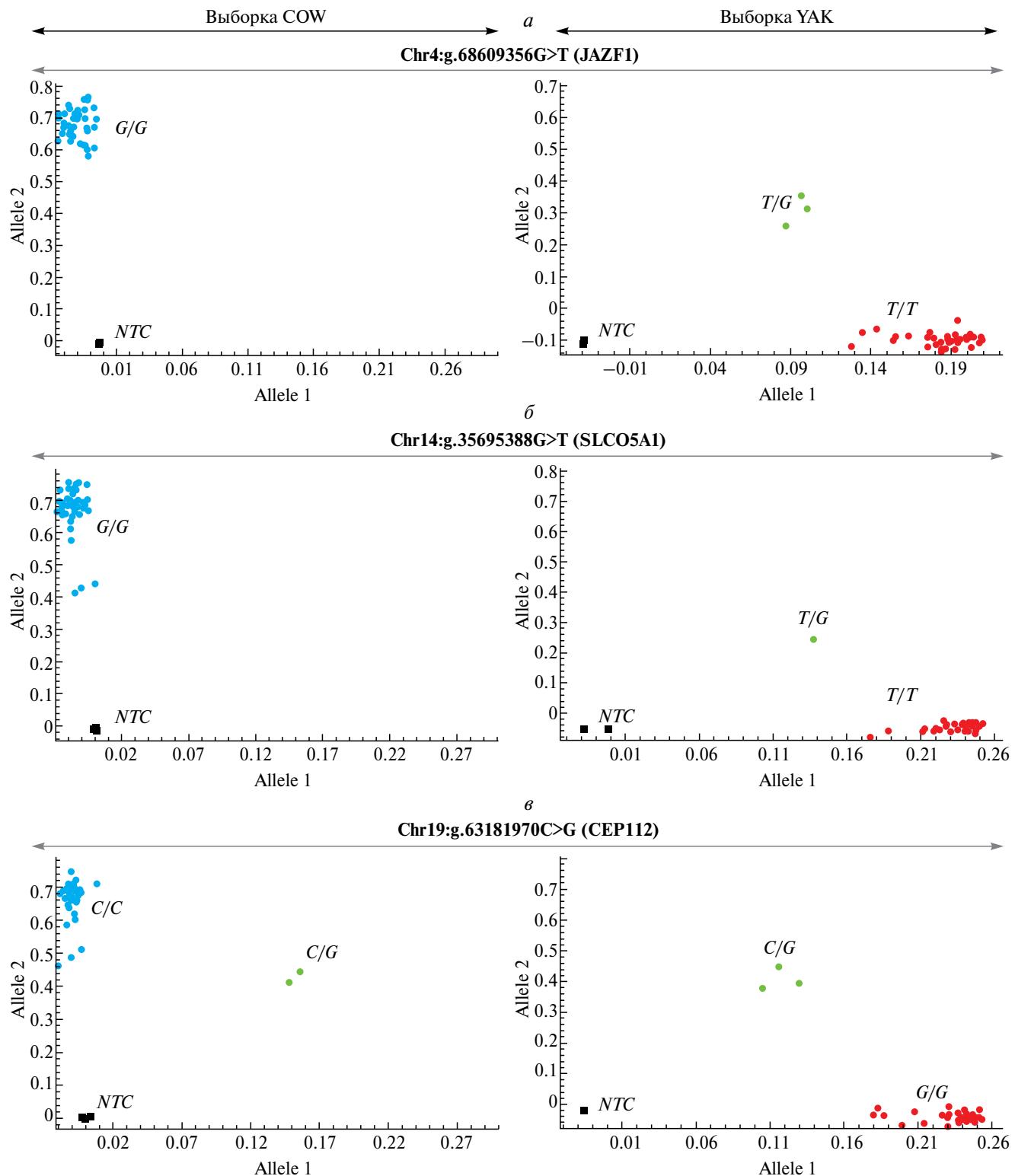
в термоцикльере QuantStudio™ 5 Real-Time PCR System (Thermo FS, США) в соответствии с рекомендациями производителя. На рис. 1 представлены 2D-графики аллельной дискриминации для SNP Chr4:68609356G>T (*JAZF1*), Chr14:35695388G>T (*SLCO5A1*), Chr19:63181970C>G (*CEP112*), использовавшихся в тестах для идентификации домашних яков и КРС.

**Определение генотипа *in silico*.** Для проведения генотипирования *in silico* использовались гены животных, доступные в базе данных NCBI, конвертированные в формат \*.fasta с помощью пакета SRA-Toolkit версии 2.11. Для определения генотипов использовалось оригинальное программное обеспечение GENIS, разработанное на языке Python версии 3.10. Подробности методологии описаны в статье [5].

В рамках исследования генотипирование было проведено для 316 особей домашнего яка и 385 особей КРС. Таким образом, в биоинформатическом анализе был задействован 701 отсеквенированный геном, файлы для которых расположены в базе Sequence Read Archive (SRA) [6]: PRJNA74739 (2012, Китайская Народная Республика – КНР), PRJNA217895 (2013, КНР), PRJNA285834 (2015, КНР), PRJEB18113 (2016, Швейцария), PRJNA508864 (2018, КНР), PRJNA431934 (2018, Австралия), PRJNA531398 (2019, КНР), PRJNA762180 (2021, Великобритания), PRJNA766811 (2021, КНР), PRJNA842787 (2022, КНР), PRJNA899924 (2022, КНР), PRJNA950586 (2023, КНР).

**Статистический анализ данных.** Для оценки потенциала SNP в качестве идентификационных маркеров использовался ROC-анализ (Receiver Operating Characteristic analysis) в программе SPSS версии 20.0. SNP полагался высокоеффективным маркером при условии, что нижняя граница асимптотического 95%-ного доверительного интервала (ДИ) для параметра площади под кривой (AUC) превышала 0.8.

Комплексную оценку дифференцирующего потенциала для совокупности SNP проводили с использованием программы MDR v.3.0.2 [7]. Вклад конкретного генотипа определялся величиной энтропии  $H$  (выраженной в %). При  $H = 100\%$  генотип способен однозначно дифференцировать, к какой группе относится образец. В программе MDR для определения оптимальной модели дифференциации использовались следующие высоко консервативные настройки: количество атрибутов (attribute count range) – от 1 до  $n$  (где  $n$  – количество переменных в модели); воспроизводимость модели (cross-validation count) – 100; анализ топ-моделей (track top models) – 1000; поиск конфигурации модели (search method configuration) – всесторонний (exhaustive); метод сравнения (ambiguous cell analysis) – точный тест Фишера (Fisher's exact test);



**Рис. 1.** График аллельной дискриминации результатов KASP для полиморфизмов: *a* – Chr4:68609356G>T (*JAZF1*); *b* – Chr14:35695388G>T (*SLCO5A1*); *c* – Chr19:63181970C>G (*CEP112*).

классификация ячеек (ambiguous cell assignment) – неклассифицированные (unclassified). Корректность модели оценивалась по значению сбалансированной точности (adj. Balanced Accuracy). Вероятность отнесения образца к одной из двух выборок – YAK или COW – рассчитывалась в SPSS v.20.0 с использованием логистической регрессии.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Отбор наиболее информативных SNP включал два этапа. На первом этапе были сформированы две тестовые группы, образцы в которые были включены случайным образом: BG-1 (домашний як,  $n = 56$ ) и BT-1 (КРС,  $n = 60$ ), для которых были определены генотипы по 947 SNP (перечень SNP представлен в файле “Дополнительные материалы”, табл. Д-1). В перечень потенциально информативных SNP вошли как ранее описанные SNP [8], так и фланкирующие их SNP из BovineHD BeadChip от Illumina® [9].

Для оценки информативности SNP был проведен ROC-анализ, который позволил выделить 64 SNP с наибольшими значениями параметра AUC (area under ROC curve) – площади, ограниченной ROC-кривой и осью доли ложных положительных классификаций. Эти маркеры были отобраны для дальнейших исследований. На втором, расширенном этапе генотипы этих 64 SNP были определены для дополнительных 260 животных *Bos grinniens* (группа BG-2) и 325 животных *Bos taurus* (группа BT-2). Список отобранных для генотипирования *in silico* животных представлен в табл. Д-2 (доп. мат.). Повторный ROC-анализ был проведен для уточнения SNP с наибольшим дифференцирующим потенциалом для различия домашнего яка и крупного рогатого скота. Полученные результаты использовались для построения ROC-кривых и оценки показателя AUC. По результатам биоинформационического анализа, значения AUC для отобранных SNP находились в диапазоне от 0.899 до 0.999, что указывает на их высокий дифференцирующий потенциал (доп. мат., табл. Д-3).

Для исследования взаимодействий SNP был использован метод многомерного сокращения размерности (MDR v.3.0.2). На основе данных, полученных в ходе биоинформационического анализа SRA, были отобраны три SNP с наибольшим дифференцирующим потенциалом: Chr4:68609356G>T (*JAZF1*), Chr14:35695388G>T (*SLCO5A1*) и Chr19:63181970C>G (*CEP112*). Значения энтропии  $H$  для этих SNP составили 78.19, 77.79 и 77.07% соответственно. Далее, с использованием технологии KASP, были определены генотипы для 202 образцов домашнего яка и крупного рогатого скота из Киргызстана. Повторный анализ с использованием метода MDR позволил рассчитать величину энтропии  $H$  для каждого SNP: для Chr4:68609356G>T (*JAZF1*)

$H = 85.16\%$ , для Chr14:35695388G>T (*SLCO5A1*)  $H = 85.16\%$ , для Chr19:63181970C>G (*CEP112*)  $H = 82.76\%$ . Результаты генотипирования представлены в табл. Д-4 (доп. мат.).

В результате ROC-анализа было установлено, что все три SNP обладают высочайшим дифференцирующим потенциалом для различия домашнего яка и крупного рогатого скота. Значения AUC для SNP составили: Chr4:68609356G>T (*JAZF1*) AUC = 1.0, 95%ДИ = [1.0–1.0],  $p = 4.17 \cdot 10^{-28}$ ; Chr14:35695388G>T (*SLCO5A1*) AUC = 1.0, 95%ДИ = [1.0–1.0],  $p = 4.17 \cdot 10^{-28}$ ; Chr19:63181970C>G (*CEP112*) AUC = 1.0, 95%ДИ = [0.999–1.0],  $p = 4.56 \cdot 10^{-28}$ . Частоты аллелей также подтвердили их значимость: аллель *G* для Chr4:68609356G>T в выборке COW составил 100%, аллель *G* для Chr14:35695388G>T – 100%, аллель *C* для Chr19:63181970C>G – 99.32%. В выборке YAK частоты аллелей были следующими: аллель *T* для Chr4:68609356G>T – 97.32%, аллель *T* для Chr14:35695388G>T – 99.11%, аллель *G* для Chr19:63181970C>G – 97.32%.

Таким образом, в тест-систему были включены три SNP, графическая интерпретация модели на основании анализа 202 образцов представлена на рис. 2, *a*. Согласно полученным результатам, сбалансированная точность идентификации, (adj. Balanced accuracy) домашнего яка и КРС при анализе 202 образцов по SNP Chr4:68609356G>T (*JAZF1*), Chr14:35695388G>T (*SLCO5A1*) и Chr19:63181970C>G (*CEP112*) составила 98.87% (специфичность модели – 100%, чувствительность – 100%). При объединении двух массивов данных – образцов из Киргызстана и образцов из NCBI-SRA, для которых имелись данные по трем SNP, с последующим анализом в MDR определено, что сбалансированная точность идентификации (adj. Balanced accuracy) домашнего яка и КРС составила 99.67% (специфичность модели – 100%, чувствительность – 100%). Графическая интерпретация модели на основании анализа 611 образцов представлена на рис. 2, *b*.

При оценке точности отнесения образца к одной из двух групп: домашний як или КРС, с использованием логистической регрессии, получены следующие данные (доп. мат., табл. Д-5):

при наличии генотипа *CC* (Chr19:63181970C>G) / *GG* (Chr4:68609356G>T) / *GG* (Chr14:35695388G>T) вероятность отнесения образца к выборке COW составляет 100% (в совокупности распространенность особей с данными генотипами в группе COW – 94.79%);

при наличии генотипа *CG* (Chr19:63181970C>G) / *GG* (Chr4:68609356G>T) / *GG* (Chr14:35695388G>T) или *CC* (Chr19:63181970C>G) / *GG* (Chr4:68609356G>T) / *TG* (Chr14:35695388G>T) вероятность отнесения образца к выборке COW

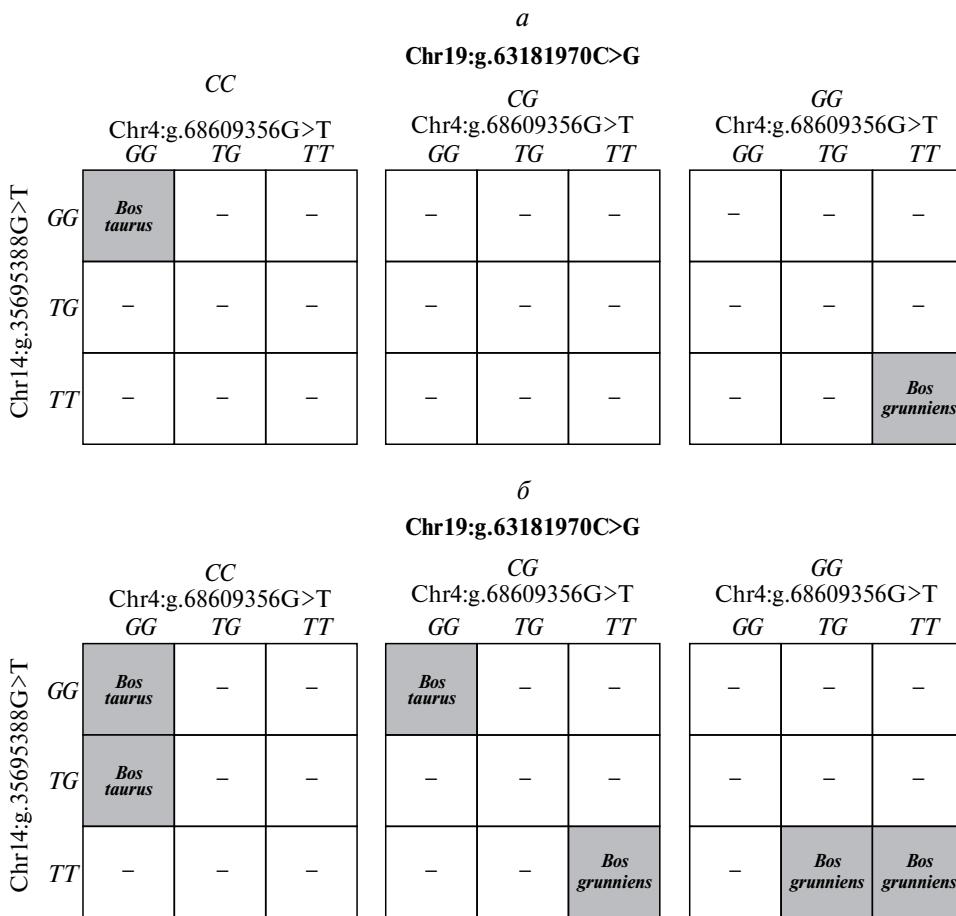


Рис. 2. Графическое представление модели из трех полиморфизмов для идентификации домашнего яка и КРС: *а* – 202 образца (образцы из Кыргызстана); *б* – 611 образцов (образцы из Кыргызстана и данные SRA).

составляет 100% (в совокупности распространенность особей с данными генотипами в группе COW – 4.22%);

при наличии генотипа *GG* (*Chr19:63181970C>G*) / *TT* (*Chr4:68609356G>T*) / *TT* (*Chr14:35695388G>T*) вероятность отнесения образца к выборке YAK составляет 100% (в совокупности распространенность особей с данными генотипами в группе COW – 90.87%);

при наличии генотипа *CG* или *GG* (*Chr19:63181970C>G*) / *TT* или *TG* (*Chr4:68609356G>T*) / *TT* (*Chr14:35695388G>T*) вероятность отнесения образца к выборке YAK составляет 100% (в совокупности распространенность особей с данными генотипами в группе COW – 7.21%);

в 1.31% случаев особь не удалось отнести ни к одному кластеру с заявленным в MDR уровнем точности в 99.0%.

Таким образом, в настоящей работе мы показали, что сбалансированная точность дифференциации *Bos taurus* и *Bos grunniens* составила более

99%, что вполне достаточно для решения большинства задач популяционной генетики. Однако в 1.31% случаев (согласно модели, рис. 2, *б*) образцы домашнего яка и КРС не могут быть корректно дифференцированы с уровнем значимости  $p < 0.01$ . Это обусловлено относительно небольшим количеством особей с редкими генотипами, такими как *CC* (*Chr19:63181970C>G*) / *TG* (*Chr4:68609356G>T*) / *GG* (*Chr14:35695388G>T*) – для КРС, или *GG* (*Chr19:63181970C>G*) / *TT* (*Chr4:68609356G>T*) / *TG* (*Chr14:35695388G>T*) – для домашнего яка. Если же использовать уровень значимости  $p < 0.05$  при MDR-анализе, то все 611 образцов дифференцируются корректно. Аналогичная ситуация показана для результатов дифференциации при использовании логистической регрессии.

В перспективе для увеличения точности модели, особенно для криминалистических приложений, может быть рекомендовано увеличение выборки исследуемых образцов за счет включения новых образцов с заведомо известной видовой принадлежностью. Увеличение числа редких генотипов при сохранении их относительной частоты

в выборках домашнего яка и КРС приведет к снижению процента образцов, которые не могут быть корректно дифференцированы.

#### Характеристика исследуемых генов

Ген *JAZF1* (JAZF zinc finger 1, NCBI Gene ID 616701) расположен на 4-й хромосоме (NC\_037331.1 (67992971..68321145), ARS-UCD2.0 (GCF\_002263795.3)). Ортолог данного гена у человека кодирует ядерный белок с тремя цинковыми пальцами типа C2H2 и действует как репрессор транскрипции. Этот ген участвует в липидном обмене, подавляя липогенез и усиливая липолиз, что приводит к уменьшению накопления липидов в жировой ткани. *JAZF1* также задействован в гомеостазе глюкозы, улучшая метаболизм глюкозы и чувствительность к инсулину. Белки *JAZF1* у КРС и человека имеют схожий размер. В исследовании Eusebi G.P. et al. изучались профили экспрессии генов в префронтальной коре агрессивных пород КРС, включая *JAZF1*, который показал пониженную экспрессию у испанской породы *Lidia*, известной своим агонистическим поведением [10]. Jin M. et al. обнаружили, что *JAZF1* участвует в адаптации овец *Ovis aries* к высокогорным условиям [11]. Zhao F. et al. выделили *JAZF1* как один из ключевых генов, влияющих на продуктивные характеристики КРС [12]. Исследователи установили, что полиморфизм *JAZF1* ассоциирован с ростом у людей [13].

Ген *SLCO5A1* (solute carrier organic anion transporter family member 5A1, NCBI Gene ID 535202) расположен на 14-й хромосоме (NC\_037341.1 (33483001..33770931), ARS-UCD2.0 (GCF\_002263795.3)). Ортолог данного гена у человека отвечает за активность трансмембранныго переносчика органических анионов, независимую от натрия. Белок *SLCO5A1* расположен во внутриклеточных мембраносвязанных органеллах и плазматической мембране. Размер белков у КРС и человека различается незначительно – 846 и 848 аминокислот соответственно. При ассоциативном анализе репродуктивных признаков у *Sus scrofa domesticus* и генных сетей на основании полногеномных исследований было показано, что ген *SLCO5A1* связан с транскрипционными факторами, участвующими в функционировании молочной железы и скелетных мышц [14]. В работе Gaddis K.L.P. et al. была обнаружена связь гена *SLCO5A1* с восприимчивостью к кетозу у джерсейской породы КРС [15]. В другом исследовании был определен регион на 14-й хромосоме генома *Bos taurus* (абердин-ангусы) с потенциально высокой ассоциацией с массой при рождении, где также расположен ген *SLCO5A1* [16].

Ген *CEP112* (centrosomal protein 112, NCBI Gene ID 617266) расположен на 19-й хромосоме (NC\_037346.1 (62280043..62600661), ARS-UCD2.0

(GCF\_002263795.3)). Ортолог этого гена у человека кодирует белок со спиральным доменом, который принадлежит к семейству эфекторных белков, контролирующих клеточное деление. *CEP112* был идентифицирован как компонент центросомы человека. Для этого гена характерен альтернативный сплайсинг, что приводит к образованию множества вариантов транскрипта. В исследовании Kim S. и соавт. установлено, что полиморфизм rs42699274 гена *CEP112* может служить генетическим маркером при оценке генетического разнообразия и дивергенции среди корейских пород КРС [17]. В другом исследовании, направленном на повышение точности геномного прогнозирования характеристик телосложения у корейского КРС голштинской породы, было показано, что полиморфизм гена *CEP112* является важным маркером для понимания генетических основ этих признаков и может служить надежной основой для прогнозов, основанных на геномных данных [18].

Известно, что прямое перенесение результатов, полученных для одного биологического вида, на другой вид некорректно и требует дополнительных исследований. Тем не менее полученные результаты могут быть полезны для дальнейшего изучения эволюции филогеографической структуры популяций домашнего яка и пород КРС в Кыргызстане. Они также могут способствовать сохранению генетических ресурсов для будущих селекционных исследований и расширить понимание того, как животные могут адаптироваться к изменениям климата. Все три SNP, включенные в тест-систему, расположены в инtronных областях генов и не должны оказывать влияние на функционирование экспрессируемых белков. Однако различные события одомашнивания, адаптация к различным климатическим зонам и дивергентный отбор по продуктивным признакам сформировали геномные различия между исследуемыми видами.

Ранее нами было показано, что при анализе пяти STR-локусов (*BM1818*, *BM1824*, *BM2113*, *CSSM66* и *ILSTS006*) точность классификации для *Bos grinniens* составила  $98.8 \pm 3.4\%$ , для *Bos taurus* –  $99.1 \pm 1.2\%$  [2]. При использовании же трех SNP (*Chr4:68609356G>T*, *Chr14:35695388G>T* и *Chr19:63181970C>G*) точность классификации составила 99.67% при максимально возможных значениях специфичности и чувствительности. При этом образцы *YAK\_107*, *YAK\_117* и *YAK\_131*, для которых по результатам STR-анализа точность отнесения к своему кластеру составила 70.3, 75.5 и 72.3% соответственно, по результатам SNP-анализа были однозначно (100% точность) классифицированы как домашние яки.

Итак, в настоящем исследовании нами предложена тест-система для дифференциации домашнего яка (*Bos grinniens*) и крупного рогатого скота (*Bos taurus*), основанная на анализе трех SNP

в генах *JAZF1*, *SLCO5A1* и *CEP112*. С использованием методов биоинформатики были определены SNP с высоким дифференцирующим потенциалом. Эти полиморфные варианты подтвердили свою эффективность на практике. Модель, основанная на трех SNP (Chr4:68609356G>T в гене *JAZF1*, Chr14:35695388G>T в гене *SLCO5A1* и Chr19:63181970C>G в гене *CEP112*), продемонстрировала высокую сбалансированную точность (не менее 99.67%) при анализе 611 образцов. Применение данного подхода позволило достичь 100% чувствительности и 100% специфичности в дифференциации двух видов. Метод KASP, используемый в молекулярно-генетическом анализе, отличается одностадийностью, что снижает риск кросс-контаминации и уменьшает трудовые затраты благодаря исключению этапов рестрикции и электрофореза. Дальнейшее расширение базы данных образцов с помощью генотипирования, особенно *in silico*, остается актуальной задачей для совершенствования и применения данной тест-системы.

Программное обеспечение GENIS разработано в рамках НИР “Биоинформационный подход к анализу данных полногеномного секвенирования для поиска однонуклеотидных замен, способных дифференцировать близкородственные биологические виды” (БРФФИ, 2023–2025 гг., Б23-060, рег. № 20231076).

Сбор биологических образцов домашнего яка и КРС выполнен в рамках правительственного задания Министерства образования и науки Кыргызской Республики (договор № 30-21 от 02.15.2021, № 129/1).

Молекулярно-генетические исследования выполнены при частичной финансовой поддержке ООО “МАКСИМ МЕДИКАЛ” в рамках частной научной инициативы ANALYSIS OF SRA DATA USING BIOINFORMATICS METHODS.

Исследование одобрено Этическим комитетом Научно-исследовательского института молекулярной биологии и медицины (Бишкек, Кыргызская Республика, 12.06.2024, протокол № 6).

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. Wang X., Pei J., Xiong L. et al. Genetic diversity, phylogeography, and maternal origin of yak (*Bos grunniens*) // BMC Genomics. 2024. V. 15. № 25(1). P. 481. P. 1–13.  
<https://doi.org/10.1186/s12864-024-10378-z>
2. Chekirov K.B., Isakova Zh.T., Kipen V.N. et al. Differentiation of *Bos grunniens* and *Bos taurus* based on STR locus polymorphism // Vavilov J. of Genet. and Breeding. 2023. V. 27(5). P. 488–494.  
<https://doi.org/10.18699/VJGB-23-59>
3. Stolpovsky Y.A., Kol N.V., Evsyukov A.N. et al. Comparative analysis of ISSR marker polymorphism in population of YAK (*Bos mutus*) and in F1 hybrids between yak and cattle in the Sayan-Altai region // Russ. J. Genet. 2014. V. 50. № 10. P. 1163–1176.  
<https://doi.org/10.7868/S0016675814100142>
4. Аль-Кейси Т.В. Сравнительное исследование аллелофонда яков и их гибридов с крупным рогатым скотом с использованием микросателлитов: дис. ... канд. биол. наук. М.: ВИЖ, 2011. 97 с.
5. Kipen V.N., Snytkov E.V. GENIS – methodological approach for *in silico* genotyping (validation on *Sus scrofa* sequencing) // Mathematical Biol. and Bioinf. 2024. V. 19(1). P. 36–51.  
<https://doi.org/10.17537/2024.19.36>
6. Sequence Read Archive (SRA) [электронный ресурс]. URL: <https://www.ncbi.nlm.nih.gov/sra>. Дата доступа: 20.06.2024.
7. Ritchie M.D., Hahn L.W., Roodi N. et al. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer // Am. J. Hum. Genet. 2001. V. 69(1). P. 138–147.  
<https://doi.org/10.1086/321276>
8. Kalbfleisch T., Petersen J.L., Tait Jr. R.G. et al. Using triallelic SNPs for determining parentage in North American yak (*Bos grunniens*) and estimating cattle (*B. taurus*) introgression // F1000Research. 2020. V. 9(1096). P. 1–26.  
<https://doi.org/10.12688/f1000research.25803.2>
9. BovineHD DNA Analysis Kit [электронный ресурс]. Режим доступа: <https://www.illumina.com/products/by-type/microarray-kits/bovinehd.html>. Дата доступа 21.06.2024.
10. Eusebi P.G., Sevane N., O'Rourke T. et al. Gene expression profiles underlying aggressive behavior in the prefrontal cortex of cattle // BMC Genomics. 2021. V. 22(245). P. 1–14.  
<https://doi.org/10.1186/s12864-021-07505-5>
11. Jin M., Wang H., Liu G. et al. Whole-genome resequencing of Chinese indigenous sheep provides insight into the genetic basis underlying climate adaptation // Genet. Sel. Evol. 2024. V. 2. № 56(1). P. 1–14.  
<https://doi.org/10.1186/s12711-024-00880-z>
12. Zhao F., McParland S., Kearney F. et al. Detection of selection signatures in dairy and beef cattle using high-density genomic information // Genet. Select. Evol. 2015. V. 47(1). № 49. P. 1–12.  
<https://doi.org/10.1186/s12711-015-0127-3>
13. Johansson A., Marroni F., Hayward C., Franklin C.S. et al. Common variants in the *JAZF1* gene associated with height identified by linkage and genome-wide association analysis // Hum. Mol. Genet. 2009.

- V. 18(2). P. 373–380.  
<https://doi.org/10.1093/hmg/ddn350>
14. *Lima Verardo L.L.* Gene networks from genome wide association studies for pig reproductive traits: Tese (Doutorado em Zootecnia). Viçosa: Universidade Federal de Viçosa, 2015. 170 p.  
<http://www.locus.ufv.br/handle/123456789/6773>
15. *Gaddis K.L.P., Megonigal J.H., Clay J.S., Wolfe C.W.* Genome-wide association study for ketosis in US Jerseys using producer-recorded data // *J. Dairy Sci.* 2018. V. 101(1). P. 413–424.  
<https://doi.org/10.3168/jds.2017-13383>
16. *Zeng X.* Angus cattle at high altitude: Pulmonary arterial pressure, estimated breeding value and genome-wide association study: Diss. ... PHD. Colorado: Department of Animal Sci., Colorado State Univ., 2016. 259 p.
17. *Kim S., Cheong H.S., Shin H.D. et al.* Genetic diversity and divergence among Korean cattle breeds assessed using a BovineHD single-nucleotide polymorphism chip // *Asian-Australas J. Anim. Sci.* 2018. V. 31(11). P. 1691–1699.  
<https://doi.org/10.5713/ajas.17.0419>
18. *Lee J., Mun H., Koo Y. et al.* Enhancing genomic prediction accuracy for body conformation traits in Korean Holstein Cattle // *Animals.* 2024. V. 14(7). P. 1–14.  
<https://doi.org/10.3390/ani14071052>

## Identification of *Bos taurus* and *Bos grinniens* Based on SNP

**V. N. Kipen<sup>1,\*</sup>, Zh. T. Isakova<sup>2,4</sup>, M. M. Patrin<sup>3</sup>, K. B. Chekirov<sup>4</sup>, K. A. Aitbaev<sup>2</sup>,  
 A. R. Karypova<sup>4</sup>, M. I. Irsaliev<sup>2</sup>**

<sup>1</sup>*Institute of Genetics and Cytology of the National Academy of Sciences of Belarus, Minsk,  
 220072 Republic of Belarus*

<sup>2</sup>*Research Institute of Molecular Biology and Medicine, Bishkek, 720040, Kyrgyz Republic*

<sup>3</sup>*Maxim Medical LLC, Moscow, 123423 Russia*

<sup>4</sup>*Kyrgyz-Turkish Manas University, Bishkek, 720038 Kyrgyz Republic*

\*e-mail: v.kipen@igc.by

The article examined samples of domestic yak and three breeds of cattle to assess the differentiating potential of the polymorphic variants Chr4:68609356G>T (*JAZF1* gene), Chr14:35695388G>T (*SLCO5A1* gene) and Chr19:63181970C>G (*CEP112* gene). The high accuracy (99.67%), specificity (100%) and sensitivity (100%) of the proposed test model consisting of these three polymorphisms for the identification of domestic yaks and cattle were confirmed. A fast and simple identification method has been developed based on this model using competitive allele-specific PCR (KASP) technology, which can significantly reduce the time and financial costs of molecular genetic analysis, as well as reduce the risk of cross-contamination of samples.

**Keywords:** *Bos taurus*, *Bos grinniens*, single nucleotide polymorphism, identification, kompetitive allele specific PCR (KASP), in silico genotyping.

## СТРУКТУРА STR-АЛЛЕЛОФОНДА ПОПУЛЯЦИИ БЫКОВ ХОЛМОГОРСКОЙ ПОРОДЫ БАНКА КРИОКОНСЕРВИРОВАННОГО СЕМЕНИ, СОХРАНЕННОГО В РЕСПУБЛИКЕ КОМИ

© 2025 г. В. С. Матюков<sup>1,\*</sup>, А. С. Лейченко<sup>1</sup>, Я. А. Жариков<sup>1</sup>, С. В. Николаев<sup>1</sup>

<sup>1</sup>Институт агробиотехнологий Федерального исследовательского центра Коми научного центра Уральского отделения Российской академии наук, Сыктывкар, 167023 Россия

\*e-mail: nipti38@mail.ru

Поступила в редакцию 03.07.2024 г.

После доработки 30.09.2024 г.

Принята к публикации 07.10.2024 г.

Исследовали полиморфизм микросателлитов у 162 чистопородных быков печорского типа холмогорской породы, холмогорской и голштинской пород, а также холмогоро-голштинских помесей. По 15 локусам микросателлитов выявили 132 аллеля, или 8.8 в среднем на локус. Из общего числа аллелей 78 (59.1%), или в среднем 5.2 аллеля на локус, встречались с частотами более 0.1 хотя бы в одной из групп быков разной генеалогии и породности. 21 аллель в 13 локусах встречался с частотами 0.15 и выше независимо от породы, генеалогии и породности группы. Наибольшее число аллелей, выявили в группах пучоро-холмогоро-голштинских и холмогоро-голштинских помесей. Максимальную генетическую дистанцию установили между кроссированными пучоро-холмогорскими быками и быками голштинской породы ( $D_N = 0.237$ ,  $F_{ST} = 0.045$ ). Высокую генетическую дифференциацию быков голштинской породы с кроссированными и чистолинейными производителями пучорского типа холмогорской породы подтвердил кластерный анализ. Генетическое различие классической холмогорской породы с голштинской было ниже. Кластерный анализ результатов генотипирования по микросателлитам быков холмогорской породы в одном массиве с выборками животных голштинской породы и холмогоро-голштинских помесей позволил получить дополнительную информацию для планирования мероприятий по поддержанию генетического разнообразия сохраненного поголовья чистопородного холмогорского скота.

**Ключевые слова:** холмогорская порода, пучорский тип, кластер, вероятность, генетическая дифференциация, микросателлиты.

**DOI:** 10.31857/S0016675825010081 **EDN:** VEHBQU

С 1970-х годов скотоводство Российской Федерации захватила повальная практика поглощения коммерческими породами отечественного скота. Этот процесс в XXI в. достиг своего апогея. В результате наша страна утратила многие отечественные генофонды. В Республике Коми скрещивание с импортными породами началось несколько позднее и менее интенсивно, чем во многих регионах Российской Федерации. Поэтому немногочисленное поголовье отечественного холмогорского скота до сих пор сохранилось в некоторых экономически слабых хозяйствах, расположенных в районах Крайнего Севера и Приполярья с неблагоприятными условиями кормления и содержания. Головное племпредприятие РГУСП “Коми” по искусственно осеменению и племенной работе сохранило банк глубокозамороженного семени чистопородных и помесных быков-производителей

холмогорской породы, уничтоженный в большинстве регионов.

Среди отечественных пород холмогорский скот на протяжении нескольких столетий считался одним из лучших по молочной продуктивности и наиболее приспособленным к разведению в условиях Российской Севера. В советский период он использовался для выведения некоторых пород и был распространен во многих регионах европейской части России, Сибири, на Чукотке, Камчатке и Сахалине. До нулевых годов XXI в. удельный вес чистопородного поголовья и низкокровных холмогоро-голштинских помесей все еще был значительным в холмогорской породе. По молочной продуктивности она конкурировала с чистопородным голштинским скотом. Так, по усредненным данным бонитировки молочного скота в хозяйствах всех категорий Российской Федерации за

2006–2009 гг. при уровне удоев около 5000 кг молока за 305 дней последней законченной лактации по удою и продукции молочного жира в расчете на один месяц жизни холмогорская порода уступила только айрширской. Разница с черно-пестрой и голштинской породами была незначительной [1], а с учетом дополнительно полученного приплода (в пересчете на молоко) превзошла голштинскую по продукции молочного жира + белка за месяц жизни на 0.23 кг, или 6.15% [2].

Родиной холмогорской породы считается пойма нижнего течения р. Северной Двины, заселенная в X–XIII вв. переселенцами из Новгорода Великого и с верховий Волги. Известно, что скот низовий Северной Двины, задолго до его скрещивания с западноевропейским, выделялся среди массива местного северного скота более крупными размерами и лучшей продуктивностью [3].

Восточнее ареала распространения холмогорского скота в бассейне р. Печоры разводили местный печорский северный комолый скот. С XVI в., а возможно и ранее, холмогорский скот по торговым путям из Холмогор и Беломорья начал проникать на среднюю и нижнюю Печору. Однако методичное поглощение и вытеснение северного комолого скота холмогорским на севере Республики Коми началось с завоза в 1931 г. в хозяйства Усть-Цилемского района Коми АССР 55 холмогорских быков. Затем завоз холмогорского скота стал постоянным в подсобные сельхозы ГУЛАГа, позднее повсеместно [4]. Между сторонниками и противниками поглощения местного скота холмогорским в 50–60-е годы прошлого столетия в Республике Коми велась острыя дискуссия. Компромиссным решением стало выведение и утверждение в качестве селекционного достижения зонального печорского типа холмогорской породы. Животные печорского типа по некоторым экстерьерным и хозяйственным признакам отличались от холмогорской породы ведущих племенных хозяйств других регионов. В условиях Приполярья они характеризовались более продолжительным сроком хозяйственного использования, реже выбывали по причине алиментарных заболеваний и заболеваний конечностей. В хозяйствах Республики Коми средний убой первотелок печорского типа холмогорской породы за 305 дней первой лактации на 165.7 кг молока, или 4.9 кг молочного жира, превышал аналогичные показатели сверстниц классических холмогорских линий. Обследование стад, неблагополучных по лейкозу, показало, что инфицированность ВЛКС коров печорского типа была на 4.8% ниже, чем сверстниц, происходивших от быков классических линий холмогорской породы ( $p \leq 0.05$ ) [5]. Многолетняя ветеринарная практика (к сожалению, не подтвержденная официальными документами) свидетельствовала о том, что по гематологической и посттубойной диагностике частота клинических

проявлений лейкоза крупного рогатого скота у печорского типа холмогорской породы была ниже, чем у животных, происходивших от производителей, завезенных из-за пределов Республики Коми.

В настоящее время печорский тип холмогорской породы, как и холмогорская порода в целом, находится на грани исчезновения и подлежат охране и сохранению в русле реализации мировой программы сохранения генетических ресурсов сельскохозяйственных животных [6]. Поэтому для проведения эффективных мероприятий по сохранению генофонда необходима всесторонняя оценка генетического разнообразия сохраненной популяции холмогорского скота. Эффективным инструментом и средством мониторинга генетического разнообразия, индивидуальной и популяционной идентичности являются молекулярные маркеры [7, 8]. Одними из наиболее доступных и эффективных маркеров для генетического анализа служат микросателлиты [9–11].

Цель исследования – с помощью генетико-статистического анализа и STR-маркеров изучить внутри- и межгрупповую генетическую изменчивость, а также кластерную структуру субпопуляций и групп чистопородного и помесного холмогорского скота, представленного банком криоконсервированного семени быков Республики Коми.

## МАТЕРИАЛЫ И МЕТОДЫ

Полиморфизм микросателлитов изучили у быков-производителей холмогорской породы, принадлежавших РГУСП “Коми” по племенной работе, биопродукция которых хранилась в криобанке головного племпредприятия. Молекулярно-генетические исследования полиморфизма микросателлитов на коммерческой основе провела лаборатория ДНК-технологий Всероссийского научно-исследовательского института племенного дела. ДНК для исследования выделяли из криоконсервированного семени быков-производителей. Анализ микросателлитного профиля производителей провели по 15 локусам: *SPS115*, *TGLA53*, *TGLA122*, *BM1818*, *ETH10*, *BM1824*, *INRA23*, *BM2113*, *TGLA227*, *CSRM60*, *TGLA126*, *ETH225*, *CSSM66*, *ETH3*, *ILSTS6*. Информацию о генотипировании быков по локусам микросателлитов заимствовали из базы данных программы “СЕЛЭКС”.

Общая численность выборки генотипированных быков составила 162 головы, из них чистопородных холмогорских 76 голов, в том числе печорского типа холмогорской породы (П) – 18, классической холмогорской породы (Х) – 38, кроссируемых печоро-холмогорских (ПХ) – 20 голов. Помесных холмогорских быков классических линий с кровностью по голштинской породе от 3 до 50% – 39 голов (ХГл<sup>1</sup>), с кровностью 50% и выше – 19 голов (ХГл<sup>2</sup>), пучоро-холмогоро-голштинских

помесей (ПХГл) – 18 голов, чистопородных голштинских (Гл) – 10 голов. Кровность помесных холмогоро-голштинских быков указана в процентах после аббревиатуры породы и породности. Под “классической” холмогорской породой “классическими” линиями в отличие от печорского типа холмогорской породы нами принят массив чистопородного холмогорского скота заводского разведения линий Наилучшего, Цветка, Тополя, Лимона, Хлопчатника и др. без кроссов с линиями печорского типа и без прилития крови других пород в течение не менее пяти поколений.

Результаты генотипирования быков по STR-локусам обработали с помощью программы GenAlEx 6.5<sup>1</sup> [12]. По группам разной породности рассчитали следующие параметры:  $N$  – число животных в группе;  $N_a$  – среднее число аллелей на локус;  $N_e$  – среднее эффективное число аллелей на локус;  $H_0$  – оценка средней наблюдаемой (фактической) гетерозиготности,  $H_E$  – оценка средней ожидаемой гетерозиготности на локус,  $uH_E$  – несмещенная оценка средней ожидаемой гетерозиготности на локус; индекс фиксации (инбридинга)  $F_{IS}$ . Парные генетические дистанции между группами быков рассчитали следующим образом:  $D_N$  – генетическая дистанция по Nei,  $F_{ST}$  – по Wright,  $uD_N$  – несмещенная генетическая дистанция по Nei и оценки  $F_{ST}$ ,  $G'_{ST}$  (Nei),  $G''_{ST}$  (Hed),  $G'''_{ST}$  (Hed),  $D_{ST}$ , полученные с использованием метода генерации случайных выборок в программе GenAlEx 6.5.

В программе Structure Version 2.3.4 [9] вычислили апостериорную вероятность ( $Q$ ) отнесения отдельных полилокусных STR-генотипов к разным кластерам без введения предварительной информации об их классификации по генеалогии. По каждой генеалогической и породной группе рассчитали средние вероятности членства особей в  $k$ -м кластере, начиная с  $k = 2$  до  $k = 4$ . Распределение средних вероятностей по кластерам свели в таблицы и визуализировали графически. Для детализации генетических различий особей, отнесенных к

различным кластерам при  $k = 3$  и  $k = 4$ , в пределах каждого кластера рассортировали особей по вероятностям членства и с порогом вероятности 0.75 сформировали в группы. Затем в программе GenAlEx 6.5 по сформированным группам рассчитали генетико-популяционные характеристики и генетические дистанции между ними. Статистическую обработку данных и графики выполнили в программе Excel.

## РЕЗУЛЬТАТЫ

В генотипированной по STR-локусам выборке быков-производителей выявили 132 аллеля, или 8.8 в среднем на локус. Из общего числа аллелей 78 (59.1%), или в среднем 5.2 аллеля на локус, встречались с частотами более 0.1 хотя бы в одной из породных групп быков. 21 аллель 13 локусов встречался с частотами 0.15 и выше во всех без исключения породных группах.

Наибольшее число STR-аллелей, которые не встречались в других выборках, выявили в ПХГл и ХГл. В сравнительно многочисленной группе чистопородных холмогорских быков выявили всего два таких аллеля, в группе быков печорского типа – два, у быков кроссированной печно-холмогорской группы – один.

Разные породные группы достоверно не различались по параметрам генетического разнообразия (табл. 1), хотя дифференциация по сочетаниям некоторых показателей между группами была заметной. Так, при меньшей численности выборки чистопородных голштинских быков по числу эффективных аллелей на локус они превзошли более многочисленную выборку печно-холмогорских быков и по этому показателю были фактически тождественны группам Х и П. Самая многочисленная выборка помесей ХГл<sup>1</sup> при высоких показателях абсолютного и эффективного числа аллелей на локус, исключая малочисленную группу Гл, уступила остальным группам по уровню наблюдаемой гетерозиготности. Кроме того, по этой группе

**Таблица 1.** Характеристика групп быков по генетико-популяционным показателям  $X \pm sx$

Выборки	$N$	$N_a$	$N_e$	$H_0$	$H_E$	$uH_E$	$F_{IS}$
Гл	10	$5.333 \pm 0.422$	$3.813 \pm 0.385$	$0.627 \pm 0.052$	$0.691 \pm 0.037$	$0.727 \pm 0.039$	$0.087 \pm 0.061$
П	18	$6.000 \pm 0.458$	$3.839 \pm 0.410$	$0.704 \pm 0.025$	$0.703 \pm 0.026$	$0.723 \pm 0.027$	$-0.012 \pm 0.037$
ПХГл	18	$6.467 \pm 0.533$	$4.138 \pm 0.405$	$0.721 \pm 0.039$	$0.726 \pm 0.027$	$0.749 \pm 0.028$	$0.014 \pm 0.027$
ПХ	20	$5.867 \pm 0.389$	$3.579 \pm 0.283$	$0.724 \pm 0.030$	$0.698 \pm 0.021$	$0.715 \pm 0.022$	$-0.040 \pm 0.034$
Х	38	$6.800 \pm 0.490$	$3.814 \pm 0.279$	$0.733 \pm 0.024$	$0.717 \pm 0.021$	$0.727 \pm 0.021$	$-0.024 \pm 0.026$
ХГл <sup>1</sup>	39	$7.133 \pm 0.601$	$4.112 \pm 0.386$	$0.701 \pm 0.023$	$0.729 \pm 0.023$	$0.738 \pm 0.023$	$0.035 \pm 0.023$
ХГл <sup>2</sup>	19	$6.400 \pm 0.349$	$3.972 \pm 0.299$	$0.688 \pm 0.033$	$0.725 \pm 0.023$	$0.745 \pm 0.024$	$0.054 \pm 0.027$

Примечание. ПХГл – печно-холмогоро-голштинские помеси со средней кровностью 27.7% по голштинской породе; ХГл<sup>1</sup> – холмогоро-голштинские помеси со средней кровностью 25.9% и ХГл<sup>2</sup> – 60.5% по голштинской породе.

выявлено достоверное нарушение генетического равновесия по генотипам пяти локусов из 15 (табл. 1).

Наиболее высокую генетическую дифференциацию по оценкам генетических дистанций установили между кроссированными пчоро-холмогорскими быками и голштинской породой. Максимальное различие между ними подтверждено разными методами (табл. 2). Несколько ниже по величине установлена генетическая дистанция между чистопородными группами быков Гл и Х. Независимо от способа расчета генетических дистанций корреляция между их оценками при попарном сравнении групп быков разной породности была высокой и колебалась в пределах от 0.894 до 1.0. Небольшое снижение значения коэффициентов корреляции установлено между показателями  $F_{ST}$  и  $F_{ST}'$  с  $G'_{ST}$  (Nei),  $G'_{ST}$  (Hed),  $G''_{ST}$ ,  $De_{ST}$ . Взаимосвязи между

оценками  $G'_{ST}$  (Nei),  $G'_{ST}$  (Hed) и  $G''_{ST}$ ,  $De_{ST}$  были близки к функциональным.

Деление всего массива на два кластера, соответствующих двум породам (голштинской и холмогорской), показало высокую генетическую дифференциацию быков голштинской породы с кроссированными и чистолинейными быками пчорского типа холмогорской породы и в меньшей степени с производителями холмогорской породы (рис. 1).

Средняя вероятность членства в первом кластере быков голштинской породы составила  $0.107 \pm 0.039$ , во втором –  $0.893 \pm 0.039$ , пчорского типа холмогорской породы соответственно  $0.828 \pm 0.052$  и  $0.172 \pm 0.052$ , пчоро-холмогорских –  $0.860 \pm 0.019$  и  $0.140 \pm 0.019$ , быков холмогорской породы –  $0.633 \pm 0.046$ ,  $0.367 \pm 0.046$ .

По сравнению с чистолинейными быками пчорского типа и кроссированными пчоро-холмогорскими производителями у холмогорских быков

**Таблица 2.** Генетические дистанции, рассчитанные разными методами при попарном сравнении групп быков разной генеалогии и породности

Выборки	$D_N$	$uD_N$	$F_{ST}$	$F_{ST}'$	$G'_{ST}$ (NEI)	$G'_{ST}$ (HED)	$G''_{ST}$	$De_{ST}$
Гл*П	0.180	0.082	0.033	<b>0.035</b>	<b>0.026</b>	<b>0.083</b>	<b>0.095</b>	<b>0.071</b>
Гл*ПХГл	0.132	0.028	0.024	0.025	0.007	0.022	0.026	0.019
П*ПХГл	0.074	0.000	0.014	0.014	-0.001	-0.002	-0.003	-0.002
Гл*ПХ	0.237	0.145	0.045	<b>0.045</b>	<b>0.049</b>	<b>0.152</b>	<b>0.173</b>	<b>0.130</b>
П*ПХ	0.049	0.000	0.010	0.010	-0.006	-0.018	-0.021	-0.015
ПХГл*ПХ	0.065	0.000	0.013	0.013	-0.001	-0.002	-0.002	-0.001
Гл*Х	0.185	0.106	0.034	<b>0.035</b>	<b>0.034</b>	<b>0.111</b>	<b>0.126</b>	<b>0.095</b>
П*Х	0.078	0.025	0.014	<b>0.015</b>	<b>0.009</b>	<b>0.029</b>	<b>0.033</b>	<b>0.024</b>
ПХГл*Х	0.058	0.000	0.010	0.011	0.000	0.001	0.001	0.001
ПХ*Х	0.063	0.016	0.013	0.013	0.007	0.020	0.024	0.017
Гл*ХГл <sup>1</sup>	0.141	0.061	0.028	<b>0.027</b>	<b>0.018</b>	<b>0.059</b>	<b>0.068</b>	<b>0.051</b>
П*ХГл <sup>1</sup>	0.072	0.019	0.014	0.014	0.006	0.020	0.023	0.017
ПХГл*ХГл <sup>1</sup>	0.058	0.000	0.011	0.010	-0.001	-0.002	-0.003	-0.002
ПХ*ХГл <sup>1</sup>	0.079	0.033	0.016	<b>0.016</b>	<b>0.012</b>	<b>0.038</b>	<b>0.044</b>	<b>0.032</b>
Х*ХГл <sup>1</sup>	0.058	0.023	0.010	<b>0.011</b>	<b>0.008</b>	<b>0.026</b>	<b>0.030</b>	<b>0.022</b>
Гл*ХГл <sup>2</sup>	0.121	0.021	0.023	0.023	0.003	0.011	0.012	0.009
П*ХГл <sup>2</sup>	0.110	0.038	0.020	<b>0.021</b>	<b>0.012</b>	<b>0.040</b>	<b>0.046</b>	<b>0.034</b>
ПХГл*ХГл <sup>2</sup>	0.096	0.018	0.016	0.017	0.005	0.016	0.019	0.014
ПХ*ХГл <sup>2</sup>	0.146	0.080	0.027	<b>0.028</b>	<b>0.028</b>	<b>0.090</b>	<b>0.102</b>	<b>0.077</b>
Х*ХГл <sup>2</sup>	0.130	0.076	0.022	<b>0.023</b>	<b>0.025</b>	<b>0.084</b>	<b>0.096</b>	<b>0.072</b>
ХГл <sup>1</sup> *ХГл <sup>2</sup>	0.065	0.011	0.012	0.012	0.002	<b>0.008</b>	0.010	0.007

Примечание. Полужирным шрифтом выделены достоверные значения генетических дистанций при уровне значимости от  $p \leq 0.05$ .

установлена более высокая вероятность членства во втором “голштинском” кластере (см. рис. 1). Распределение средних вероятностей членства быков классической холмогорской породы по отдельным кластерам было близким к распределению вероятностей в группе помесей ПХГл со средней кровностью по голштинской породе 27.7%. У холмогоро-голштинских помесей с повышением кровности по голштинам возрастала вероятность принадлежности ко второму кластеру.

При делении всего массива на три кластера во всех группах независимо от их генеалогии и породности в разных соотношениях присутствовали компоненты трех кластеров. В группе голштинских быков средние вероятности принадлежности особей к кластерам распределились следующим образом: первый –  $0.343 \pm 0.087$ , второй –  $0.564 \pm 0.095$ , третий –  $0.093 \pm 0.019$  (рис. 2).

У быков печорского типа самой высокой была вероятность принадлежности к третьему кластеру –  $0.666 \pm 0.102$  и минимальная – к первому ( $0.115 \pm 0.135$ ). Коэффициент корреляции между

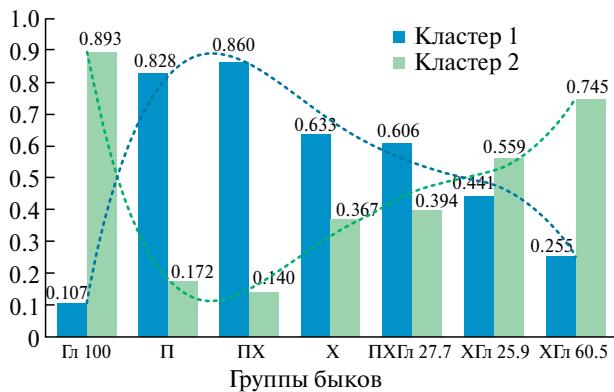


Рис. 1. Распределение средних оценок вероятностей принадлежности особей групп различной породности к кластерам при  $k = 2$ .

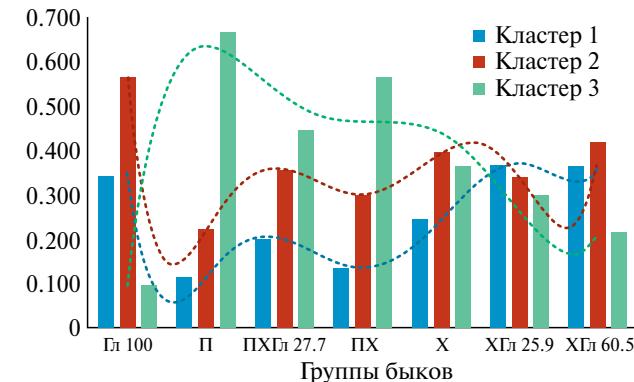


Рис. 2. Распределение средних оценок вероятностей принадлежности особей групп различной породности к кластерам при  $k = 3$ .

распределениями средних вероятностей членства особей в отдельных кластерах по этим двум группам был отрицательным ( $-0.785 \pm 0.122$ ). По сравнению с печорским типом у холмогорской породы средние вероятности членства по кластерам распределились более равномерно:  $0.245 \pm 0.087$ ,  $0.394 \pm 0.040$ ,  $0.361 \pm 0.044$  (рис. 2). При  $k = 3$  коэффициент корреляции распределения средних вероятностей членства быков классической холмогорской и голштинской пород по кластерам составил  $0.175 \pm 0.145$ , между холмогорской породой и печорским типом холмогорской породы –  $-0.473 \pm 0.120$ .

Таким образом, распределение средней вероятности членства особей в кластерах у холмогорской породы одновременно положительно коррелировало с распределениями вероятности по кластерам у печорского типа и голштинской породы.

Генетические дистанции  $F_{ST}$  между кластерами, оцененные по группам особей, сформированным с порогом вероятности членства в кластере 0.75 и выше, при  $k = 3$  следующие:  $1*2 - 0.055$ ,  $1*3 - 0.044$ ,  $2*3 - 0.027$ ; при  $k = 4$ :  $1*2 - 0.056$ ,  $1*3 - 0.080$ ,  $1*4 - 0.077$ ,  $2*3 - 0.042$ ,  $2*4 - 0.032$ ,  $3*4 - 0.053$ . Из этих данных видно, что некоторые оценки генетических дистанций между различными кластерами приближались или достигали величин межпородных различий [13]. Однако только генетическое различие между животными первых кластеров и массивом быков в целом достигало значений  $F_{ST}$  0.033 и 0.046. Для остальных кластеров оценки  $F_{ST}$  были значительно ниже: при  $k = 3 - 0.008-0.012$ , при  $k = 4 - 0.010-0.023$ .

Анализ индивидуальных профилей вероятности членства быков печорского типа холмогорской породы в кластерах показал, что при  $k = 2$  они распределились в пределах от 0.146 до 0.969. У 86.8% животных вероятность отнесения к первому кластеру достигла и превысила порог вероятности 0.75. Ко второму кластеру с вероятностью 0.75 и выше отнесен только один бык (2.6%). У животных классической холмогорской породы 36.8% особей с вероятностью 0.75 и выше отнесены к первому кластеру и 13.2% – ко второму. При  $k = 3$  у быков классической холмогорской породы индивидуальная величина  $Q$  в третьем кластере варьировала от 0.027 до 0.847, при этом только у 10.5% животных она достигла и превысила пороговое значение 0.75; 18.4% животных при пороге вероятности членства 0.75 отнесены ко второму и три животных (7.9%) к первому кластеру. У печорского типа холмогорской породы значения вероятности членства в третьем кластере распределились в пределах 0.125–0.924. У 50.0% особей был превышен порог вероятности 0.75 отнесения их к третьему кластеру, у 5.2% – ко второму кластеру и ни одного животного к первому.

## ОБСУЖДЕНИЕ

Считается, что микросателлиты вследствие их высокого полиморфизма позволяют получать надежные оценки генетической дифференциации между популяциями и внутрипопуляционными субъединицами [14]. Для этого в качестве мер генетической дифференциации широко используются генетические дистанции Wright –  $F_{ST}$  [13], Nei –  $D_N$  ( $G_{ST}$ ) и его модификации [15]. Однако на показатели  $F_{ST}$  Wright и  $G_{ST}$  Nei, полученные с помощью многоаллельных систем маркеров, влияет уровень гетерозиготности в пределах отдельных популяций. Поэтому R. W. Hedrick предложил рассчитывать стандартизированную величину  $G''_{ST}$  (Hed), а L. Jost разработал статистику  $D_{e_{ST}}$ , основанную на изменчивости эффективного числа аллелей [15]. В нашем исследовании коэффициенты корреляции оценок генетических дистанций, полученных разными методами, информативно были фактически равноценны.

Независимо от способов оценки генетических дистанций дифференциация холмогорской породы с голштинской была несколько выше, чем с печорским типом холмогорской породы, но ниже, чем между Гл и ПХ. Между группами быков П\*ПХ, ПХ-Гл\*ХГл<sup>1</sup>, ПХГл\*ПХ, П\*ПХГл генетические дистанции по большинству оценок были близки к нулю (см. табл. 2).

Редукция гетерозигот, которая прослеживалась в группах чистопородных голштинских быков и их помесей с холмогорской породой (см. табл. 1), пока не получила объяснения. Вероятнее всего, недостаток гетерозигот в этих группах являлся следствием инбредности использованных голштинских производителей, которая не компенсировалась повышением гетерозиготности, обусловленной разницей в генных частотах между скрещивающимися голштинской и холмогорской породами. В пользу этого предположения свидетельствует снижение до минимума дефицита гетерозигот в генетически наиболее гетерогенной группе ПХГл.

С. В. Николаев и В. Л. Ялуга [16] также установили возрастание дефицита гетерозигот у помесного холмогорского скота с увеличением кровности по голштинской породе. В группе помесей с кровностью по голштинской породе 3–24% они наблюдали генетическое равновесие между гетеро- и гомозиготами ( $F_{IS} 0.005 \pm 0.017$ ), а в группе с кровностью 75–91% редукция гетерозигот достигала 22.4% ( $F_{IS} 0.224 \pm 0.045$ ). Одновременно в этой группе наблюдаемая гетерозиготность по сравнению с другими группами достоверно снижалась. В. В. Волкова и соавт. [17] установили, что по STR-маркерам холмогорская порода по сравнению с голштинской характеризовалась более высоким уровнем аллельного разнообразия по среднему числу аллелей на локус, количеству

информационных аллелей и эффективному числу аллелей. Тенденция снижения аллельного разнообразия с повышением кровности по голштинской породе у холмогоро-голштинских помесей отмечена и в [16]. В нашем исследовании такая тенденция неочевидна и может быть распространена только на динамику числа эффективных аллелей. Тем не менее независимо от использования в [17, 18] разных панелей микросателлитов полученные в нашем исследовании оценки аллельного разнообразия, гетерозиготности и данные о более низком аллельном разнообразии у животных пучорского типа по сравнению с классической холмогорской породой хорошо согласовывались с [17, 18].

В результате геногеографических исследований полиморфизма mtДНК, гаплотипов Y-хромосомы и STR-маркеров получены новые данные о генезисе крупного рогатого скота Евразии и, в частности, холмогорского [19, 20]. При исследовании гаплотипов Y-хромосомы установлено сходство скота, обитающего на территориях севера Англии, прибалтийских низменностей, Скандинавии и Европейского Севера России. Эти данные позволили предположить, что холмогорский скот Архангельской области и территории Республики Коми, занятой пучорским типом холмогорской породы, активно колонизировался западноевропейскими породами черно-пестрого корня. Одновременно с этим исследование mtДНК показало сходство ярославского и холмогорского скота, включая пучорский тип, с северным местным скотом Финляндии [19].

В то же время по 30 локусам микросателлитов не удалось выявить потока генов в направлении от черно-пестрого скота к холмогорской породе [20]. На первый взгляд, результаты этих исследований могут показаться противоречивыми. Однако известно, что Y-хромосома наследуется только по мужской линии. Поэтому, если в популяции хотя бы однократно заменить местных самцов на мигрантов с последующим отбором для репродукции их мужских потомков от местных самок, то даже без повторных завозов в ходе воспроизводства популяции Y-хромосома самцов местной популяции будет вытеснена Y-хромосомой мигрантов. Ситуация по X-хромосоме может быть прямо противоположной, а по генам, локализованным в аутосомах, популяция может восстановить исходный аллелофонд по материнской породе. История холмогорского скотоводства, исследования полиморфизма ДНК подтвердили правильность такого предположения [2].

В настоящем исследовании были представлены голштинская и классическая холмогорская породы, пучорский тип холмогорской породы и группы холмогоро-голштинских помесей разной кровности, что позволило провести кластерный анализ каждой из перечисленных групп при  $k = 2 \dots 4$ .

В результате установили, что с увеличением числа кластеров величины и распределения средней вероятности членства особей в соответствующих кластерах менялись. Из сравнения диаграмм и линий тренда (см. рис. 1 и 2) видно, что с увеличением числа кластеров более детально проявляется сходство (различие) между вероятностными профилями групп разной породности, характеризующих особенности структуры их аллелофондов. При  $k = 2$  (см. рис. 1) имели место четкое различие профилей голштинской породы (Гл) с группами П и ПХ и по сравнению с ними значительный дрейф классической холмогорской породы в сторону голштинской породы. В то же время по генетическим дистанциям Гл с П, ПХ и Х такой четкой дифференциации не выявлено.

При подразделении групп на три кластера (см. рис. 2) максимальная дифференциация по вероятностному профилю установлена между Гл и П. В группе П средняя вероятность членства особей во втором кластере составила 0.220 (минимальная), у голштинской породы 0.564 (максимальная), в третьем кластере соответственно 0.666 и 0.093. Корреляция между распределениями средних вероятностей членства особей в кластерах была высоко отрицательной ( $-0.785$ ).

Использование кластерного анализа оказалось эффективным для исследования генетической дифференциации родственных между собой помесных групп. Так, в группах ПХГл и ХГл со сходной кровностью по голштинской породе средние вероятности членства особей соответственно в первом, втором и третьем кластерах распределились следующим образом: 0.200, 0.355, 0.445 и 0.369, 0.335, 0.296. Из этих данных видно, что средние вероятности членства во втором кластере (условно “голштинском”) в сравниваемых группах почти равны, а дифференциация между генетической структурой групп обусловлена разницей между вероятностями членства особей в третьем кластере (условно “печорского типа”) и первом (условно “переходном”). Из приведенных данных видно, что в группе ПХГл доля аллелофонда печорского типа выше, чем в группе ХГл. Это различие обусловлено более высоким генетическим (генеалогическим) вкладом (независимо от пола предков) печорского типа в ПХГл за счет снижения вклада холмогорской породы относительно ХГл.

Таким образом, кластерный анализ результатов генотипирования быков холмогорской породы по микросателлитам в одном массиве с выборками животных голштинской породы и холмогоро-голштинских помесей позволил получить дополнительную информацию для планирования и реализации мероприятий по сохранению и воспроизведству исчезающего генофонда. В частности, удалось: а) показать, что дифференциация печорского типа холмогорской породы с голштинской породой по

компонентам аллелофонда значительно выше, чем классического холмогорского скота; б) установить различие внутригрупповой генетической структуры между разными генеалогическими и породными группами, генетическую дифференциацию которых другими методами выявить не удалось; в) показать генетическую дифференциацию между группами особей, вероятность членства которых в отдельных кластерах была равна или превышала пороговое значение.

Практическая ценность полученной информации заключается в том, что она облегчит проведение индивидуального и группового поиска, отбора и подбора полилокусных STR-генотипов для размножения при планировании мероприятий по сохранению генетического разнообразия в поколениях воспроизводства (репродукции) исчезающей породы.

Работа выполнена при поддержке Минобрнауки РФ в рамках Государственного задания № FUUU-2023-0002 НИОКР 123033000033-4.

Исследование одобрено Этическим комитетом Института агробиотехнологий Федерального исследовательского центра Коми научного центра Уральского отделения Российской академии наук (01.07.2024, протокол № 2).

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. Матюков В.С., Тырина Ю.О., Кантанен Ю., Столповский Ю.А. Оценка селекционной ценности генофонда местного скота (на примере холмогорской породы) // С.-х. биология. 2013. № 2. С. 19–30.
2. Матюков В.С., Жариков Я.А., Зиновьева Н.А. Генетическая история и ценность генофонда исчезающей холмогорской породы // Молочное и мясное скотоводство. 2018. № 2. С. 2–8.
3. Резников Ф.И. Новые данные к истории холмогорского скота. Архангельск, 1949. 28 с.
4. Шубин П.Н., Котельников В.М. Породное преобразование крупного рогатого скота Коми АССР. Сыктывкар, 1965. 135 с.
5. Гагиев Г.И. Характеристика животных нового внутрипородного печорского типа – “ПХ-1” холмогорской породы // Научные основы молочного производства на Севере. Сыктывкар, 1998. С. 56–84.
6. Столповский Ю.А., Захаров-Гезехус И.А. Проблема сохранения генофондов доместицированных животных // Вавил. журн. генетики и

- селекции. 2017. Т. 21. № 4. С. 477–486.  
<https://doi.org/10.18699/VJ17.266>
7. *Demir E., Balcioğlu M.S.* Genetic diversity and population structure of four cattle breeds raised in Turkey using microsatellite markers // Czech. J. Animal Sci. 2019. V. 64. № 10. P. 411–419.  
<https://doi.org/10.17221/62/2019-CJAS>
8. *Saravanan K.A., Panigrahi M., Kumar H. et al.* Genome-wide assessment of genetic diversity, linkage disequilibrium and haplotype block structure in Tharparkar cattle breed of India // Animal Biotechnology. 2020. V. 33. № 1. P. 1–15.  
<https://doi.org/10.1080/10495398.2020.1796696>
9. *Pritchard J.K., Wen X., Falush D.* Documentation for structure software: Version 2.3. Software from <http://pritch.bsd.uchicago.edu/structure.html>
10. *Галинская Т.В., Щепетов Д.М., Лысенков С.Н.* Предубеждения о микросателлитных исследованиях и как им противостоять // Генетика. 2019. Т. 55. № 6. С. 617–632.  
<https://doi.org/10.1134/S0016675819060043>
11. *Матюков В.С., Жариков Я.А., Канева Л.А.* Анализ аллелофонда полутонкорунных овец печорской популяции с помощью STR-маркеров // Генетика. 2023. Т. 59. № 7. С. 843–849.  
<https://doi.org/10.31857/S0016675823060103>
12. *Peakall R., Smouse P.E.* GenAIEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-anupdate // Bioinformatics. 2012. V. 28. P. 2537–2539.
13. *Hall S.J.G.* Genetic differentiation among Livestock breeds-values for *Fst* // Animals. 2022. V. 12. № 9.  
<https://doi.org/10.3390/ani12091115>
14. *Kalinowski S.T.* Evolutionary and statistical properties of three genetic distances // Mol. Ecol. 2002. V. 11. № 8. P. 1263–1273.  
<https://doi.org/10.1046/j.1365-294X.2002.01520.x>
15. *Кузнецов В.М.* Сравнение методов оценки генетической дифференциации популяций по микросателлитным маркерам // Аграрная наука Евро-Северо-Востока. 2020. Т. 21. № 2. С. 169–182.  
<https://doi.org/10.30766/2072-9081.2020.21.2.169-182>
16. *Николаев С.В., Ялуга В.Л.* Сравнительная генетическая характеристика микросателлитного профиля голштинизированных и чистопородных холмогорских быков // Аграрная наука. 2023. № 7. С. 58–62.  
<https://doi.org/10.32634/0869-8155-2023-372-7-58-62>
17. *Волкова В.В., Романенкова О.С., Денискова Т.Е. и др.* Характеристика аллелофонда холмогорской породы крупного рогатого скота с использованием STR-маркеров // Молочное и мясное скотоводство. 2019. № 7. С. 3–7.
18. *Киселёва Т.Ю., Подоба Б.Е., Заблудский Е.Е. и др.* Анализ 30 микросателлитных маркеров у шести локальных популяций крупного рогатого скота // С.-хоз. биология. 2010. № 6. С. 20–25.
19. *Meng-Hua Li., Tapio I., Villkki J. et al.* The genetic structure of cattle populations (*Bos taurus*) in northern Eurasia and the neighbouring. Near Eastern regions: Implications for breeding strategies and conservation // Mol. Ecol. 2007. V. 16. № 18. P. 3839–3853.  
<https://doi.org/10.1111/j.1365-294X.2007.03437.x>
20. *Kantanen J., Edwards C.J., Bradley D. et al.* Maternal and paternal genealogy of Eurasian taurine cattle (*Bos taurus*) // Heredity. 2009. V. 103. № 5. P. 404–415.  
<https://doi.org/10.1038/hdy.2009.68>

## Structure of the STR Allele Pool of the Population of Kholmogoryk Breed Bulls, Saved in the Komi Republic of the Bank of Frosted Semen

V. S. Matyukov<sup>1, \*</sup>, A. S. Leichenko<sup>1</sup>, Ya. A. Zharikov<sup>1</sup>, S. V. Nikolaev<sup>1</sup>

<sup>1</sup>*Institute of Agrobiotechnologies, Federal Research Center Komi Scientific Center, Ural Branch of the Russian Academy of Sciences, Syktyvkar, 167023 Russia*

\*e-mail: nipti38mail.ru

Microsatellite polymorphism was studied in 162 purebred Pechora-type bulls of the Kholmogory, Kholmogory, and Holstein breeds, as well as Holstein-Kholmogory crossbreeds. For 15 microsatellite loci, 132 alleles were identified, or 8.8 on average per locus. Of the total number of alleles, 78 (59.1%), or 5.2 alleles per locus on average, occurred with frequencies greater than 0.1 in at least one of the groups of bulls of different genealogy and breed. Twenty-one alleles in 13 loci occurred with frequencies of 0.15 and higher, regardless of the breed, genealogy, and breed of the group. The largest number of alleles was found in the Pechora-Kholmogory-Holstein and Kholmogory-Holstein crossbreeding groups. The maximum genetic distance was established between the crossed Pechora-Kholmogory bulls and Holstein bulls ( $D_N$  0.237,  $F_{ST}$  0.045). High genetic differentiation of Holstein bulls with crossed and purebred Pechora-type sires of the Kholmogory breed was confirmed by cluster analysis. The genetic difference between the classical Kholmogory breed and the Holstein was lower. Cluster analysis of the results of genotyping by microsatellites of Kholmogory bulls in one array with samples of Holstein animals and Kholmogory-Holstein crosses made it possible to obtain additional information for planning measures to maintain the genetic diversity of the saved herd of purebred Kholmogory cattle.

**Keywords:** Kholmogory breed, Pechora-type bulls, cluster, possibility, genetic differentiation, microsatellites.

## ALU-ПОЛИМОРФИЗМ ГЕНОВ-РЕГУЛЯТОРОВ АУТОФАГИИ И АПОПТОЗА КАК ФАКТОР ПРОДОЛЖИТЕЛЬНОСТИ ЖИЗНИ ЧЕЛОВЕКА

© 2025 г. В. В. Эрдман<sup>1, 3, \*</sup>, Д. Д. Каримов<sup>1, 2</sup>, И. А. Туктарова<sup>1</sup>, А. А. Петинцева<sup>1</sup>,  
Я. Р. Тимашева<sup>1, 3</sup>, Т. Р. Насибуллин<sup>1</sup>

<sup>1</sup>Институт биохимии и генетики – обособленное структурное подразделение Уфимского федерального исследовательского центра Российской академии наук, Уфа, 450054 Россия

<sup>2</sup>Уфимский научно-исследовательский институт медицины труда и экологии человека, Уфа, 450106 Россия

<sup>3</sup>Башкирский государственный медицинский университет, Уфа, 450008 Россия

\*e-mail: danivera@mail.ru

Поступила в редакцию 07.06.2024 г.

После доработки 11.07.2024 г.

Принята к публикации 16.07.2024 г.

Для оценки вклада Alu-инсерций в генах-кандидатах старения и долголетия *ACE*, *PLAT*, *COL13A1*, *LAMA2*, *CDH4*, *SEMA6A*, *PKHD1L1*, *STK38L*, *HECW1*, *TEAD1* в выживаемость на фоне старческого физиологического и патологического фенотипа проведен анализ ассоциаций с продолжительностью жизни. Получены данные о причинах смерти 1382 человек преклонного возраста из выборки татар, проживающих в Республике Башкортостан (всего 1790 человек от 18 до 109 лет). Риск смерти оказался повышен у лиц с Alu-инсерцией в гене *STK38L* (Ya5ac2145\*II, *HR* = 2.07, *P* = 0.022). Alu-инсерционные варианты генов *HECW1* и *TEAD1* продемонстрировали протективный в отношении выживаемости эффект (Ya5NBC182\*II, *HR* = 0.71, *P* = 0.038 и Ya5ac2013\*II, *HR* = 0.74, *P* = 0.035 соответственно). С выживаемостью при различных клинических фенотипах ассоциированы Alu-полиморфные варианты генов *SEMA6A* (Yb8NBC597\*ID, *HR* = 0.54, *P* = 0.016 при цереброваскулярных заболеваниях), *TEAD1* (Ya5ac2013\*II, *HR* = 0.57, *P* = 0.016 при сердечно-сосудистых патологиях) и *LAMA2* (Ya5-MLS19\*ID, *HR* = 0.36, *P* = 0.03 при полиморбидности). Таким образом, с выживаемостью и достижением долголетия оказались ассоциированы гены, вовлеченные в регуляцию аутофагии и апоптоза.

**Ключевые слова:** старение, долголетие, Alu-полиморфизм, гены *TEAD1*, *HECW1*, *STK38L*, *LAMA2*, *SEMA6A*, анализ выживаемости.

**DOI:** 10.31857/S0016675825010092 **EDN:** VEFHND

Продолжительность жизни человека определяется комплексом средовых, поведенческих и наследственных факторов. При этом динамический характер эндогенных механизмов, регулирующих скорость старения организма, обусловливает специфические особенности поздних этапов онтогенеза. Абсолютное и относительное количество изменений в организме, вызванных повреждением молекул, клеток, органов и их систем, в совокупности приводит к снижению функциональной активности и нарушению гомеостаза как всего организма, так и его частей, без возможности полного выздоровления. Соответственно, процесс старения характеризуется повышением риска развития большого числа заболеваний и ростом вероятности смерти от всех причин [1, 2]. Важно отметить, что доля возрастных патологий, среди которых

ведущими являются болезни системы кровообращения, дыхания, онкологические заболевания, варьирует в разных возрастных группах [3].

Изучение роли полиморфизма генома в формировании возраст-ассоциированных заболеваний выступает одним из ключевых направлений молекулярно-генетических исследований детерминации продолжительности жизни. В частности, полиморфное состояние ДНК-сайтов, содержащих мобильные генетические элементы, ассоциировано с повышением нестабильности генома. Инсерции транспозонов могут вызывать различные мутации, хромосомные перестройки, влиять на эпигенетический ландшафт эукариотического генома, процессы регуляции транскрипции и экспрессию генов. Данные эффекты инсерционных событий в большей степени обнаруживаются на

поздних этапах онтогенеза [4]. Наиболее распространенным семейством транспозонов человека являются *Alu*-повторы. Локализованные в генах ключевых сигнальных путей *Alu*-полиморфные варианты вовлечены в физиологические и патологические внутриклеточные процессы [5, 6]. Ранее нами были установлены ассоциации с долголетием *Alu*-полиморфных локусов в генах, кодирующих протеинкиназу *STK38L*, протеинилазу *HECW1*, белок кальциевого канала *PKHD1L1*, клеточный рецептор *SEMA6A*, транскрипционный фактор *TEAD1*, ферменты плазмы *ACE* и *PLAT*, а также белки адгезии *CDH4* и внеклеточного матрикса *COL13A1* и *LAMA2* [7–9]. Участие данных генов в развитии возраст-зависимых заболеваний, в том числе сердечно-сосудистых, нейродегенеративных, онкологических, а также болезней обмена веществ, установлено в ряде исследований [10–20]. В то же время их вклад в выживаемость, в том числе при различных клинических фенотипах и, более того, в возрасте, превышающем среднестатистический показатель, ранее не был изучен. *Alu*-транспозоны являются одними из ключевых эндогенных факторов эволюционной адаптации и развития человека как вида [21, 22]. Тем самым данный вид полиморфизма генома можно рассматривать как важный молекулярно-генетический предиктор выживаемости в условиях старческого физиологического фона и, более того, патологического коморбидного состояния.

В настоящей работе мы проанализировали вклад *Alu*-полиморфных вариантов генов, кодирующих структурные компоненты клетки, межклеточных взаимодействий, а также вовлеченных в ключевые сигнальные пути клеточной активности, в выживаемость среди лиц преклонного возраста и долгожителей.

## МАТЕРИАЛЫ И МЕТОДЫ

Выборка из жителей Республики Башкортостан формировалась на протяжении 2001–2015 гг. и включала 1790 человек в возрасте от 18 до 109 лет, татар по этнической принадлежности. Критериями включения лиц среднего возраста (18–59 лет) в исследование являлись отсутствие в медицинском анамнезе сахарного диабета, перенесенного инфаркта и/или инсульта, аутоиммунных и онкологических заболеваний. В отношении людей старшей возрастной когорты (60–89 лет) допускалось наличие в анамнезе атеросклероза, кардиосклероза и церебросклероза. В группу долгожителей включались все лица, достигшие 90 лет.

Образцы ДНК получали методом фенольно-хлороформной экстракции из 8 мл цельной венозной крови. *Alu*-полиморфные локусы в генах *ACE*, *PLAT*, *COL13A1*, *LAMA2*, *CDH4*, *SEMA6A*, *PKHD1L1*, *STK38L*, *HECW1*, *TEAD1* были выбраны

в качестве генетических предикторов выживаемости на основании сведений об их функциональной значимости и ранее обнаруженной связи с возрастными заболеваниями и долголетием [7–20]. Генотипирование проводили методом ПЦР с последующим разделением амплифицированных фрагментов в 1%-ном агарозном геле. Условия идентификации *Alu*-полиморфных локусов приведены в табл. 1.

Информация о выживаемости и смертности всех лиц старше 60 лет и 115 человек старше 45 лет, вошедших ранее в группу исследования (всего 1382 человека), собиралась в период с 2022 по 2023 г. путем опроса их родственников. На момент завершения данного этапа работы (30 декабря 2023 г.) были получены сведения о статусе выживаемости 1069 человек (процент отклика составил 77.35%), из них 944 человека умерли, 125 были живы. Исследуемая выборка была дифференцирована по возрасту на контрольную группу, группу лиц старческого возраста и долгожителей. Верхняя граница для контрольной группы соответствовала показателю средней продолжительности жизни населения Республики Башкортостан, который был установлен на основании официальных данных Федеральной службы государственной статистики (<https://rosstat.gov.ru/>, дата обращения 25 января 2024 г.). Характеристика сформированных групп представлена в табл. 2.

Данные исследования обрабатывали на платформе IBM SPSS V22.0 (Чикаго, Иллинойс, США), а также в программной среде Python [23]. Частоты всех *Alu*-полиморфных вариантов генов были проверены на соответствие равновесию Харди–Вайнберга в контрольной группе. Изменение с возрастом распределения частот генотипов по каждому выбранному *Alu*-полиморфному локусу оценивали путем попарного сравнения возрастных групп с помощью  $\chi^2$ -теста Пирсона. Связь между *Alu*-полиморфными вариантами и смертностью от всех причин устанавливали с применением регрессии пропорциональных рисков Кокса. Кроме того, был проведен стратифицированный анализ выживаемости с учетом пола и по причинам смертности от различных патологий возраста. Кривые показателя относительного риска смерти (hazard ratio – *HR*) построены с использованием пакетов lifelines и matplotlib.

## РЕЗУЛЬТАТЫ

В этнической группе татар, проживающих в Республике Башкортостан, охарактеризовано распределение частот аллелей и генотипов по *Alu*-полиморфным локусам в интранах генов *ACE*, *COL13A1*, *LAMA2*, *TEAD1*, *PLAT*, *PKHD1L1*, *STK38L*, *CDH4*, *HECW1*, *SEMA6A* как в общей выборке, так и с учетом возраста. Наблюдаемое распределение частот

**Таблица 1.** Alu-полиморфные локусы, вошедшие в исследование, их локализация, условия амплификации и длина амплифицируемых фрагментов

Alu-элемент	Ген, локализация*	Последовательности праймеров	Температура отжига, °C	Аллели (длина фрагмента, п.н.)
Ya5ACE	<i>ACE</i> 17q23.3	F 5'-ctg gag acc act ccc atc ctt tct-3' R 5'-gat gtg gcc atc aca ttc gtc aga t-3'	68	<i>I</i> (490) <i>D</i> (190)
Ya5NBC182	<i>HECW1</i> 7p13	F 5'-gaa gga cta tgt agt tgc aga agc-3' R 5'-aac cca gtg gaa aca gaa gat g-3'	64	<i>I</i> (563) <i>D</i> (287)
Yb8NBC597	<i>SEMA6A</i> 5q23.1	F 5'-tga ggt gtt gca gac gat gt-3' R 5'-cgc atg ctt tag aga ata ccc-3'	63	<i>I</i> (429) <i>D</i> (108)
Yb8NBC516	<i>CDH4</i> 20q13.33	F 5'-ggg ctc agg gat act atg ctc-3' R 5'-gcc tag gcc tac cac tca ga-3'	60	<i>I</i> (445) <i>D</i> (124)
Ya5ac2145	<i>STK38L</i> 12p11.23	F 5'-tgt tct aat gac cat gcc tac tt-3' R 5'-tgc ctt tag gaa gct aca gat tta-3'	60	<i>I</i> (465) <i>D</i> (135)
Yb8AC702	<i>PKHD1L1</i> 8q23.2	F 5'-tgt ttg gaa ata agc caa aca at-3' R 5'-ggg tag caa cct ttt tca tct tt-3'	60	<i>I</i> (482) <i>D</i> (161)
Ya5ac2013	<i>TEAD1</i> 11p15.2	F 5'-tgg cag att ctg act ggc ta-3' R 5'-cac gta agg tga aaa ggg ga-3'	60	<i>I</i> (489) <i>D</i> (212)
TPA25	<i>PLAT</i> 8p11.21	F 5'-caa cca atg aaa acc act ga-3' R 5'-gtt ctc ctg aca tct tta ttg-3'	60	<i>I</i> (518) <i>D</i> (217)
Ya5ac1986	<i>COL13A1</i> 10q22.1	F 5'-tct agt ggg atg agg ata ac-3' R 5'-tgt gcc atg ggg taa gaa ac-3'	60	<i>I</i> (431) <i>D</i> (134)
Ya5-MLS19	<i>LAMA2</i> 6q22.33	F 5'-cta tga cgg agt aaa aag aag t-3' R 5'-gaa aga gtg cca acc ctg tcc-3'	63 (7 циклов) 60 (22 цикла)	<i>I</i> (401) <i>D</i> (106)

Примечание. F – прямой праймер, R – обратный праймер; п.н. – пар нуклеотидов; \* – данные базы UCSC.

генотипов всех изученных полиморфных маркеров в группе лиц среднего возраста соответствовало теоретически ожидаемому равновесию Харди–Вайнберга ( $P_{ХВ} > 0.05$ , табл. 3). Изменение спектра частот генотипов с возрастом оценивали путем попарного сравнения возрастных групп для каждого выбранного Alu-полиморфного локуса с помощью  $\chi^2$ -теста Пирсона (табл. 3). Наиболее выраженные различия в частотах генотипов между возрастными группами были отмечены для полиморфного локуса *HECW1*\*Ya5NBC182 ( $P < 0.001$ ). Для полиморфных локусов *LAMA2*\*Ya5-MLS19, *TEAD1*\*Ya5ac2013 и *SEMA6A*\*Yb8NBC597 статистически значимые изменения в распределении частот генотипов наблюдались для лиц, достигших долголетия ( $P < 0.05$ ). Среди лиц преклонного возраста (в группах лиц старческого возраста и долгожителей) наблюдалось отклонение в распределении частот генотипов по генетическому маркеру *STK38L*\*Ya5ac2145 ( $P < 0.01$ ). Кроме того, распределение частот генотипов в группе лиц старческого возраста по полиморфному локусу *COL13A1*\*Ya5ac1986 и

среди долгожителей по полиморфному локусу *PKHD1L1*\*Yb8AC702 отличалось от таковых в контрольной группе лиц среднего возраста ( $P < 0.05$ ).

Для оценки роли установленных Alu-полиморфных маркеров старения и долголетия в определении продолжительности жизни и достижения возраста долголетия был проведен анализ выживаемости. Согласно полученным результатам, относительный риск смерти от всех причин статистически значимо повышен у носителей инсерций в Alu-полиморфном локусе гена *STK38L*; у носителей инсерций в генах *TEAD1* и *HECW1* показано снижение относительного риска (табл. 4).

Детальный анализ полученных результатов показал, что носительство гомозиготного инсерционного генотипа гена *STK38L* в 2 раза повышает риск смерти от всех причин в общей группе исследованных лиц (Ya5ac2145\*II,  $HR = 2.07$ ,  $P = 0.02$ ; см. рис. 1, a). Alu-инсерции в генах *TEAD1* и *HECW1* ассоциированы со снижением риска смерти ( $HR = 0.74$ ,  $P = 0.035$  для генотипа Ya5ac2013\*II гена *TEAD1*;  $HR = 0.71$ ,  $P = 0.038$  и  $HR = 0.67$ ,  $P = 0.015$

Таблица 2. Характеристика группы исследования

Группы	<i>n</i> (%) всех случаев	Возрастной диапазон, лет	<i>M</i> ± <i>σ</i>
Всего	1790 (100)	18–109*	67.85 ± 21.22
мужчины	809 (45.2)		
женщины	981 (54.8)		
Возрастные группы			
Средний возраст	631	18–65 (для мужчин) 18–74 (для женщин)	43.93 ± 15.86
Старческий возраст	724	66–89 (для мужчин) 75–89 (для женщин)	84.11 ± 10.45
Долгожители	435	90–114	94.5 ± 3.61
Статус выживаемости			
Всего	1069	45–114	83.3 ± 10.8
живые	125 (11.7)	45–96	70.82 ± 11.8
умершие	944 (88.3)	45–114	85.59 ± 9.35
Причины смерти:			
старость	352 (37.29)	70–114	88.96 ± 6.21
ССЗ	241 (25.53)	46–104	84.67 ± 8.74
ЦВЗ	166 (17.59)	53–105	87.09 ± 8.63
ХОБЛ в анамнезе	26 (2.75)	53–98	83.52 ± 8.19
СД2Т в анамнезе	10 (1.06)	59–97	83.45 ± 8.19
рак	41 (4.34)	45–99	78.82 ± 11.83
другие причины	42 (4.45)	49–97	78.98 ± 13.35
полиморбидность	66 (6.99)	59–100	84.22 ± 8.54

Примечание. ССЗ – сердечно-сосудистые заболевания; ЦВЗ – цереброваскулярные заболевания; ХОБЛ – хроническая обструктивная болезнь легких; СД2Т – сахарный диабет второго типа; *n* – объем выборки (группы); *M* – средний возраст, *σ* – стандартное отклонение; \* – сведения о возрасте на момент сбора материала.

для генотипов Ya5NBC182\**II* и Ya5NBC182\**ID* гена *HECWI* соответственно; см. рис. 1, *б-г*).

Анализ выживаемости с учетом пола выявил схожую тенденцию в показателях относительного риска смерти отдельно для мужчин и женщин (табл. 4). При этом у мужчин для гомозиготного Alu-инсерционного генотипа гена *STK38L* показана более выраженная ассоциация с риском смерти от всех причин (Ya5ac2145\**II*, *HR* = 2.38, *P* = 0.041; см. рис. 2). Наличие Alu-инсерций в генах *TEAD1* и *HECWI*, хотя и снижает риск смерти, однако не достигает уровня статистической значимости (CI 0.47–1.31 и CI 0.61–1.49 для генотипов Ya5ac2013\**II* и Ya5ac2013\**ID* гена *TEAD1* соответственно; CI 0.35–1.14 и CI 0.33–1.11 для генотипов

Ya5NBC182\**II* и Ya5NBC182\**ID* гена *HECWI* соответственно; см. табл. 4). У женщин наличие гетерозиготного генотипа гена *HECWI* связано со снижением риска смерти от всех причин (Ya5NBC182\**ID*, *HR* = 0.65, *P* = 0.037; рис. 3).

Анализ выживаемости в группах индивидов, дифференцированных по причинам смерти, показал ассоциацию Alu-инсерций в генах *SEMA6A*, *TEAD1* и *LAMA2* со снижением риска смерти при различных патологических фенотипах. Наличие Alu-инсерции в гетерозиготном варианте в гене *SEMA6A* ассоциировано со снижением риска смерти от цереброваскулярных заболеваний (Yb8NBC597\**ID*, *HR* = 0.54, *P* = 0.016; рис. 4, *a*). Риск смертности от сердечно-сосудистых

Таблица 3. Распределение частот генотипов по Alu-полиморфизму генов в трех возрастных группах

Ген Alu-полиморфизм	Генотип	Средний возраст			Старческий возраст			Долгожители			
		<i>n</i>	<i>p</i> , %	$P_{XB}$	<i>n</i>	<i>p</i> , %	$P_{\chi^2*}$	<i>n</i>	<i>p</i> , %	$P_{\chi^2*}$	$P_{\chi^2**}$
<i>ACE</i> Ya5ACE	<i>II</i>	104	23.53	0.85	176	25.43	0.361	94	23.50	0.072	0.296
	<i>ID</i>	200	45.25		318	45.95		198	49.50		
	<i>DD</i>	138	31.22		198	28.61		108	27.00		
<i>HECW1</i> Ya5NBC182	<i>II</i>	189	42.19	0.47	250	45.29	0.041	125	45.29	< 0.001	< 0.001
	<i>ID</i>	203	45.31		219	39.67		128	46.38		
	<i>DD</i>	56	12.50		83	15.04		23	8.33		
<i>SEMA6A</i> Yb8NBC597	<i>II</i>	18	3.76	0.47	33	5.84	0.120	23	7.57	0.005	0.037
	<i>ID</i>	156	32.57		167	29.56		105	34.54		
	<i>DD</i>	305	63.67		365	64.60		176	57.89		
<i>CDH4</i> Yb8NBC516	<i>II</i>	145	39.94	0.19	191	39.63	0.074	100	35.09	0.483	0.216
	<i>ID</i>	144	39.67		216	44.81		124	43.51		
	<i>DD</i>	74	20.39		75	15.56		61	21.40		
<i>STK38L</i> Ya5ac2145	<i>II</i>	10	2.06	0.06	12	2.11	0.004	4	1.28	0.008	0.522
	<i>ID</i>	78	16.05		102	17.96		63	20.19		
	<i>DD</i>	398	81.89		454	79.93		245	78.53		
<i>PKHD1L1</i> Yb8AC702	<i>II</i>	119	23.20	0.54	133	20.75	0.421	72	20.06	0.025	0.160
	<i>ID</i>	284	55.36		346	53.98		186	51.81		
	<i>DD</i>	110	21.44		162	25.27		101	28.13		
<i>TEAD1</i> Ya5ac2013	<i>II</i>	137	27.45	0.05	157	27.79	0.966	97	29.39	0.024	0.024
	<i>ID</i>	224	44.89		250	44.25		161	48.79		
	<i>DD</i>	138	27.66		158	27.96		72	21.82		
<i>PLAT</i> TPA25	<i>II</i>	134	23.76	0.57	157	24.30	0.591	87	23.71	0.561	0.126
	<i>ID</i>	253	44.86		296	45.82		158	43.05		
	<i>DD</i>	177	31.38		193	29.88		122	33.24		
<i>COL13A1</i> Ya5ac1986	<i>II</i>	309	55.28	0.47	363	54.18	0.038	229	57.11	0.937	0.081
	<i>ID</i>	205	36.67		263	39.25		137	34.16		
	<i>DD</i>	45	8.05		44	6.57		35	8.73		
<i>LAMA2</i> Ya5-MLS19	<i>II</i>	128	21.84	0.13	129	19.03	0.095	60	14.81	< 0.001	< 0.001
	<i>ID</i>	262	44.71		310	45.72		224	55.31		
	<i>DD</i>	196	33.45		239	35.25		121	29.88		

Примечание. *n* – объем группы; *p* – частота генотипа;  $P_{XB}$  – *P*-значение критерия Харди–Вайнберга;  $P_{\chi^2*}$  – *P*-значение  $\chi^2$ -теста Пирсона; \* – результаты относительно группы лиц среднего возраста, \*\* – результаты относительно группы лиц старческого возраста.

Таблица 4. Ассоциация между Аল-полиморфизмом генов кандидатов старения и долголетия и смертностью от всех причин

Ген Ал-полиморфизм	Генотип	Общая группа		Мужчины		Женщины	
		HR (95% CI <sub>HR</sub> )	P	HR (95% CI <sub>HR</sub> )	P	HR (95% CI <sub>HR</sub> )	P
<i>ACE</i> Ya5ACE	<i>ID</i>	1.11 (0.88–1.4)	0.396	1.08 (0.72–1.63)	0.701	1.09 (0.8–1.48)	0.577
	<i>II</i>	0.98 (0.74–1.3)	0.911	0.99 (0.59–1.66)	0.974	0.88 (0.61–1.26)	0.475
<i>HECWT</i> Ya5NBC182	<i>ID</i>	0.67 (0.49–0.93)	0.015	0.61 (0.33–1.1)	0.102	0.65 (0.44–0.98)	0.037
	<i>II</i>	0.71 (0.52–0.98)	0.038	0.63 (0.35–1.14)	0.126	0.7 (0.47–1.04)	0.076
<i>SEMA6A</i> Yb8NBC597	<i>ID</i>	0.98 (0.79–1.22)	0.854	1.06 (0.72–1.56)	0.784	0.97 (0.73–1.28)	0.825
	<i>II</i>	0.63 (0.39–1.02)	0.058	0.62 (0.18–2.13)	0.447	0.7 (0.41–1.19)	0.184
<i>CDH4</i> Yb8NBC516	<i>ID</i>	1.11 (0.88–1.59)	0.276	0.78 (0.44–1.37)	0.380	1.28 (0.88–1.85)	0.198
	<i>II</i>	1.16 (0.86–1.57)	0.327	0.95 (0.54–1.66)	0.857	1.16 (0.8–1.69)	0.438
<i>STK38L</i> Ya5ac2145	<i>ID</i>	0.98 (0.76–1.28)	0.907	0.88 (0.57–1.36)	0.559	0.97 (0.68–1.39)	0.879
	<i>II</i>	2.07 (1.11–3.86)	0.022	2.38 (1.04–5.46)	0.041	1.55 (0.54–4.46)	0.415
<i>PKHD1LI</i> Yb8AC702	<i>ID</i>	1.07 (0.82–1.4)	0.632	1.46 (0.89–2.41)	0.138	1.02 (0.72–1.44)	0.919
	<i>II</i>	1.11 (0.82–1.51)	0.501	1.55 (0.88–2.73)	0.132	0.91 (0.62–1.35)	0.651
<i>TEAD1</i> Ya5ac2013	<i>ID</i>	0.91 (0.71–1.16)	0.443	0.95 (0.61–1.49)	0.835	0.83 (0.6–1.15)	0.261
	<i>II</i>	0.74 (0.57–0.98)	0.035	0.79 (0.47–1.31)	0.353	0.72 (0.51–1.02)	0.066
<i>PLAT</i> ТРА25	<i>ID</i>	1.16 (0.89–1.51)	0.323	0.76 (0.5–1.15)	0.199	1.36 (0.99–1.85)	0.056
	<i>II</i>	1.12 (0.89–1.42)	0.283	0.93 (0.59–1.46)	0.764	1.29 (0.91–1.83)	0.151
<i>COL13A1</i> Ya5ac1986	<i>ID</i>	1.17 (0.72–1.92)	0.53	1.55 (0.59–4.09)	0.379	1.01 (0.56–1.85)	0.962
	<i>II</i>	0.99 (0.61–1.62)	0.989	1.31 (0.49–3.52)	0.586	0.81 (0.45–1.44)	0.472
<i>LAMA2</i> Ya5-MLS19	<i>ID</i>	0.92 (0.73–1.15)	0.473	1.25 (0.84–1.86)	0.261	0.84 (0.62–1.13)	0.243
	<i>II</i>	1.08 (0.81–1.44)	0.609	1.41 (0.8–2.49)	0.238	0.92 (0.63–1.35)	0.672

Примечание. HR – hazard rate (уровень риска); CI – confidence interval (доверительный интервал); P – показатель уровня значимости.

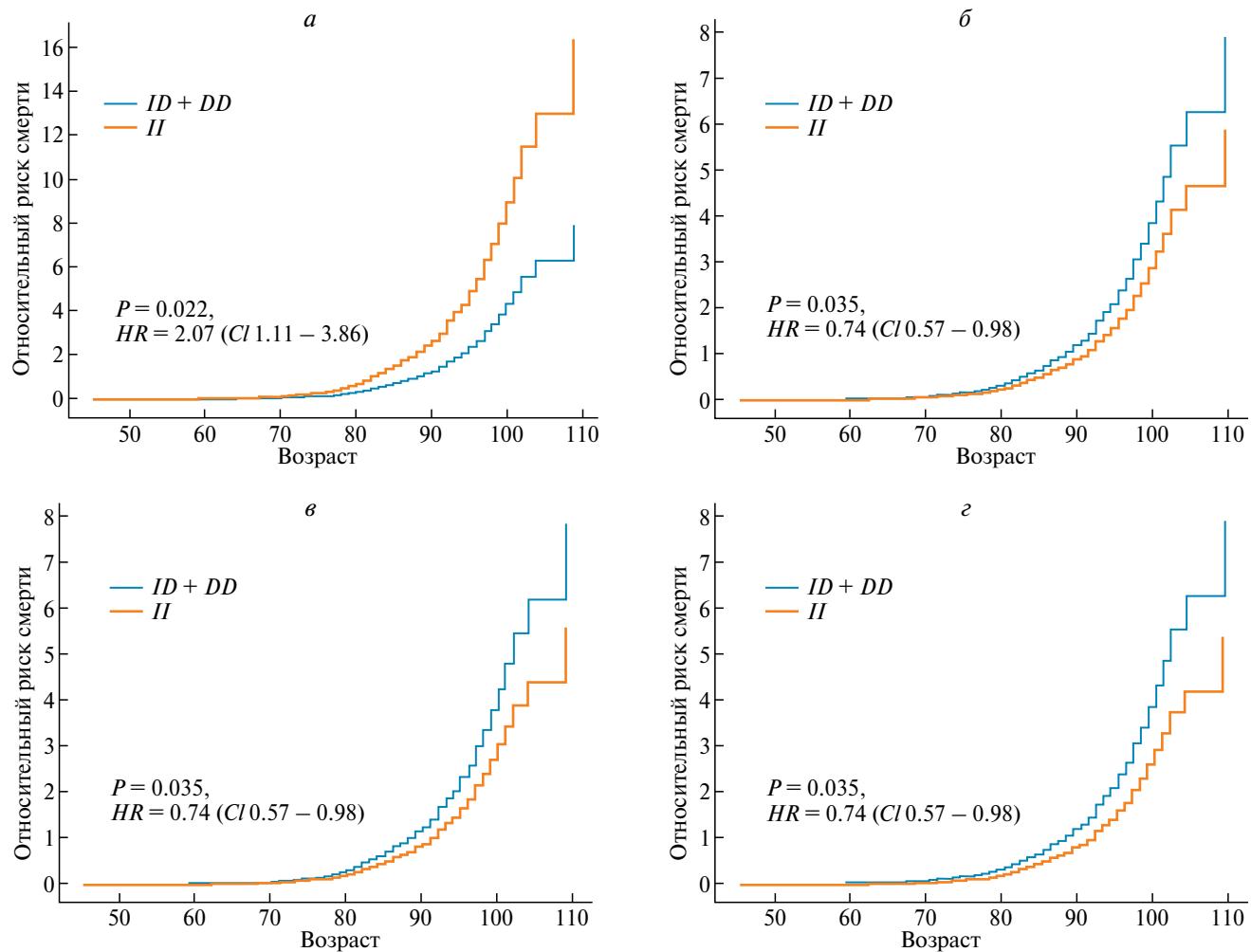


Рис. 1. Накопленный риск смертности от всех причин в общей группе, ассоциированный с Alu-инсерцией в генах *STK38L* (а), *TEAD1* (б) и *HECW1* (в и г).

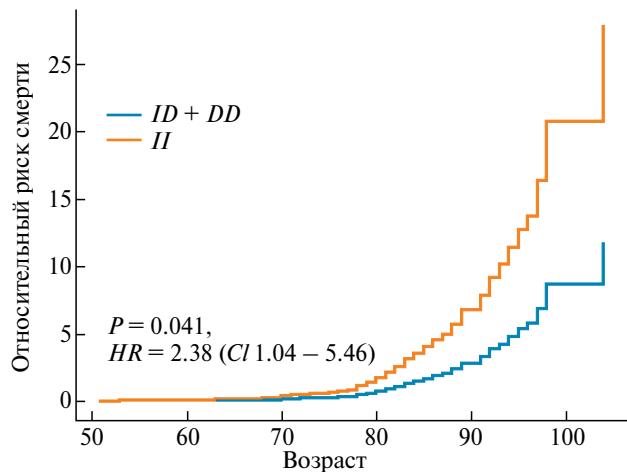
патологий снижен у носителей гомозиготного по Alu-инсерции в гене *TEAD1* генотипа (Ya5ac2013\*II,  $HR = 0.57$ ,  $P = 0.016$ ; рис. 4, б). Также снижение риска смертности ассоциировано с Alu-инсерцией в гене *LAMA2* среди индивидов с полиморбидностью (Ya5-MLS19\*ID,  $HR = 0.36$ ,  $P = 0.03$ ; см. рис. 4, в).

## ОБСУЖДЕНИЕ

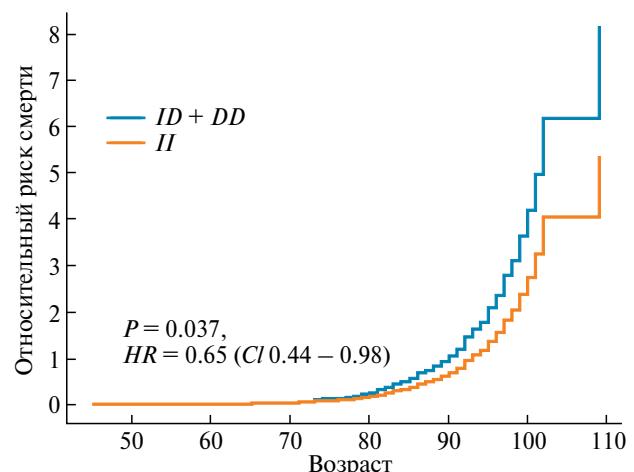
В рамках исследования старения и долголетия человека был проведен анализ выживаемости среди лиц, достигших возраста, превышающего среднепопуляционный показатель, по Alu-полиморфным вариантам генов ключевых структурных и функциональных белков в качестве предикторов. Согласно полученным результатам, Alu-инсерционный генотип Ya5ac2145 гена *STK38L* ассоциирован с риском смертности, в то время как для Alu-инсерционных генотипов Ya5NBC182 гена *HECW1* и Ya5ac2013 гена *TEAD1* продемонстрирован протективный в отношении выживаемости эффект. Кроме того, с

выживаемостью при различных клинических фенотипах оказались ассоциированы Alu-инсерционные локусы Yb8NBC597 гена *SEMA6A* и Ya5-MLS19 гена *LAMA2*. Ассоциированные с выживаемостью в условиях старческого фенотипа гены вовлечены в регуляторные пути контроля апоптоза и аутофагии.

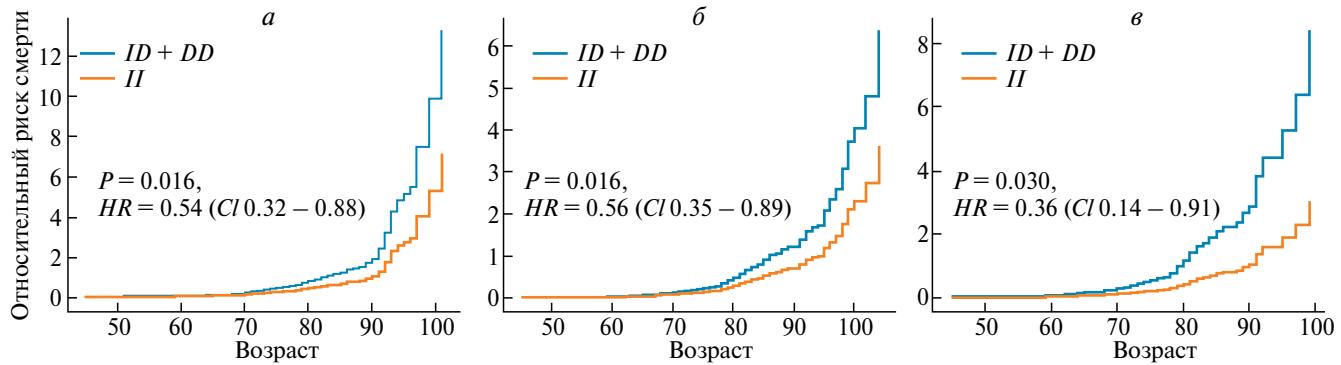
*HECW1* является C2 и WW домен-содержащим белком семейства убиквитинлигаз E3 и относится к семейству транскрипционных факторов (ТФ) NEDD4, регулирующих цитоплазматическую трансляцию, рибонуклеопротеиновый комплекс и биогенез рибосом, а также пути KEGG, включая Akt, p53, аутофагию и апоптоз [24]. Как и другие члены семейства NEDD4, *HECW1* связывается с LC3 – ключевым белком системы аутофагии, принимая тем самым участие в регуляции данного клеточного процесса. В раковых клетках продемонстрирован ингибирующий аутофагию эффект нокдауна NEDD4 [25]. Более того, *HECW1* усиливает проапоптотическую активность p53 независимо



**Рис. 2.** Накопленный риск смертности от всех причин в группе мужчин, ассоциированный с Alu-инсерцией в гене *STK38L*.



**Рис. 3.** Накопленный риск смертности от всех причин в группе женщин, ассоциированный с Alu-инсерцией в гене *HECW1*.



**Рис. 4.** Накопленный риск смертности в группах, дифференцированных по причинам смерти: от цереброваскулярных заболеваний, ассоциированный с Alu-инсерцией в гене *SEMA6A* (а); от сердечно-сосудистых заболеваний, ассоциированный с Alu-инсерцией в гене *TEAD1* (б); при полиморбидности, ассоциированный с Alu-инсерцией в гене *LAMA2* (в).

от его катализитической активности [26]. В целом контролируемая убиквитинлигазой E3 деградация белка играет фундаментальную роль в самообновлении, поддержании и дифференцировке раковых стволовых клеток [27]. Таким образом, активность HECW1 является принципиальным моментом для развития патологического фенотипа, формирующегося с возрастом. Белок HECW1 в изобилии присутствует в нейрональных тканях и благодаря участию в белковом гомеостазе является ключевым элементом в нормальном и патологическом развитии нервной системы [28]. Интересно, что существует обратная корреляция между онкологическими и нейродегенеративными заболеваниями, ключевой причиной которой может выступать локализация белка p53 в клетке, по-разному воздействующая на процессы апоптоза и аутофагии [29]. В данной работе установлена ассоциация выживаемости и долголетия с Alu-инсерцией Ya5NBC182 в гене *HECW1*. Можно предположить,

что Alu-транспозоны, влияя на активность генов, являются молекулярной основой для адаптационной пластичности тканей нервной системы. Однако тканеспецифичный характер вовлеченности HECW1 в сложную сеть регулирования апоптоза и аутофагии, особенно на поздних этапах жизни человека, требует дальнейшего всестороннего изучения.

ТФ TEAD1 и протеинкиназа STK38L являются членами консервативного пути Hippo, регулирующего размер органов и гомеостаз тканей [30]. Недавние исследования продемонстрировали роль сигнального каскада Hippo в стимулировании апоптоза и аутофагии [10]. Делекция генов, связанных с аутофагией и взаимодействующих с каскадами киназы Hippo, ассоциирована с нарастающей склонностью к спонтанному развитию различных заболеваний [31]. Белок TEAD1 (ТФ домена TEA) является одним из основных низкостоящих ядерных эффекторов передачи сигналов Hippo.

Он способен связываться с консенсусной последовательностью ДНК 5'-CATTCC-3', называемой MCAT-элементом [32]. Благодаря взаимодействию с различными ко-факторами, такими как YAP (yes-associated protein) и TAZ (ко-активатор транскрипции с мотивом связывания PDZ), TEAD связывается с MCAT-содержащими генами, которые регулируют рост клеток. Показано, что экспрессия белков TEAD усиливается при различных типах рака и коррелирует с плохой выживаемостью пациентов с онкологией [11]. Кроме того, TEAD регулирует экспрессию множества генов, участвующих в развитии сердечно-сосудистой системы, и, будучи основным молекулярным компонентом YAP/TAZ-сигнального пути, вовлечен в патофизиологические процессы, которые способствуют сердечно-сосудистым заболеваниям [33]. Можно предположить, что обусловленная Alu-вставкой Ya5ac2013 сниженная активность гена *TEAD1*, вовлеченного в контроль пролиферации и апоптоза, в условиях старческого фенотипа способствует выживаемости и защите от патологий сердца и сосудов.

Дополнительные ферменты пути Hippo включают семейство протеинкиназ NDR, в частности *NDR1/STK38* и *NDR2/STK38L* [34, 35]. Данные киназы регулируют широкий спектр возраст-чувствительных клеточных процессов, таких как контроль клеточного цикла, межклеточная коммуникация, апоптоз, аутофагия, гомеостаз питательных веществ и др. [36]. Экспериментально было продемонстрировано участие киназы *STK38* в системном метаболизме: на фоне диеты с повышенным содержанием жиров ее экспрессия значительно повышалась, что в свою очередь привело к развитию воспаления и инсулинерезистентности [37]. Установлено, что *STK38/STK38L* действует как основной фактор ответа на стрессовые воздействия и играет важную роль в аутофагии [10]. Экспрессия гена *STK38L* снижается под воздействием стрессовых факторов, при этом степень снижения коррелирует с хронологическим возрастом [36]. Выявленная нами ассоциация с риском смерти от всех причин в исследуемой группе Alu-инсерции Ya5ac2145 в гене *STK38L*, связанная со снижением активности гена, согласуется с представленными литературными данными.

В настоящей работе установлена ассоциация с выживаемостью на фоне полиморбидного статуса генотипа, гетерозиготного по Alu-инсерции в гене *LAMA2*, кодирующем основной компонент базальной мембраны ламинин. При изучении роли экспрессии гена ламинина в формировании различных старческих процессов были продемонстрированы противоречивые результаты, что можно объяснить тканеспецифичностью данного белка [12]. Интересно, что в мышцах с дефицитом  $\alpha 2$ -цепи ламинина повышается экспрессия генов

аутофагии [13]. С позиции адаптации и выживаемости в преклонном возрасте результаты данной работы можно рассматривать как взаимодействие различных компенсаторных механизмов.

Вовлеченность Alu-инсерционного полиморфизма гена *SEMA6A* в выживаемость при цереброваскулярных событиях напрямую подтверждает роль семафорина-6 в структурно-функциональной организации нервной системы [14]. Кроме того, многие исследования показали, что семафорины влияют на пролиферацию клеток, миграцию и апоптоз, воздействуя на компоненты сосудистой стенки и принимая тем самым участие во многих патологических процессах кровеносной системы [15].

Таким образом, Alu-полиморфные варианты, для которых установлены ассоциации с выживаемостью и смертностью, локализованы в генах, действованных в процессах апоптоза и аутофагии. Недавние достижения в понимании временных и пространственных последствий нарушения регуляции аутофагии для тканевого гомеостаза выявили сложную и многофакторную взаимосвязь между аутофагией и старением. Будучи высококонсервативным путем разрушения дефектных клеточных компонентов, аутофагия выступает важным эндогенным механизмом, обеспечивающим снятие клеточных стрессовых состояний, при этом хроническая активация аутофагии может привести к гибели клеток [38]. В целом полученные нами данные согласуются с концепцией о возраст-зависимом снижении количества белков, связанных с аутофагией и обеспечивающих транспорт к лизосомам, что свидетельствует о нарушениях аутофагии как одном из важных факторов старения организма [34]. Alu-ретротранспозоны способны влиять на функционирование генов множеством различных путей, в основном приводя к снижению уровня экспрессии генов. Это позволяет предположить, что данный тип генетического полиморфизма может быть ассоциирован с рядом патологических возраст-зависимых фенотипов и, соответственно, продолжительностью и качеством жизни человека.

Работа поддержана грантом РНФ № 24-25-00179.

Исследование одобрено Этическим комитетом ИБГ УФИЦ РАН (6.06.2024 г., протокол № 8).

Все процедуры, выполненные в исследовании с участием людей, соответствуют этическим стандартам институционального и/или национального комитета по исследовательской этике и Хельсинской 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 об отсутствии конфликта интересов.

### СПИСОК ЛИТЕРАТУРЫ

1. *Мушкамбаров Н.Н.* Геронтология in polemico. Монография. М.: “Мед. информ. агентство”, 2011. 464 с.
2. *Чупаха М.В., Белоусова О.Н., Сухатерина Е.В.* Характеристика биологического возраста и данных антропометрии при артериальной гипертензии на фоне метаболического синдрома у пациентов среднего и пожилого возраста // Соврем. проблемы здравоохранения и мед. статистики. 2024. № 1. С. 335–347.  
<https://doi.org/10.24412/2312-2935-2024-1-335-347>
3. *Maier H., Jeune B., Vaupel J.W.* Exceptional Lifespans. Springer Nature, 2021. 344 p.
4. *Le Breton A., Bettencourt M.P., Gendrel A.V.* Navigating the brain and aging: Exploring the impact of transposable elements from health to disease // Front. Cell. Dev. Biol. 2024. V. 12.  
<https://doi.org/10.3389/fcell.2024.1357576>
5. *Maxwell P.H.* What might retrotransposons teach us about aging? // Curr. Genet. 2016. V. 62. P. 277–282.  
<https://doi.org/10.1007/s00294-015-0538-2>
6. *Li M., Schifanella L., Larsen P.A.* Alu retrotransposons and COVID-19 susceptibility and morbidity // Hum. Genomics. 2021. V. 15. P. 2–11.  
<https://doi.org/10.1186/s40246-020-00299-9>
7. *Эрдман В.В., Каримов Д.Д., Насибуллин Т.Р. и др.* Роль Alu-полиморфизма генов *PLAT*, *PKHD1L1*, *STK38L* и *TEAD1* в формировании признака долгожительства // Успехи геронтологии. 2016. Т. 29. № 5. С. 709–716.
8. *Каримов Д.Д., Эрдман В.В., Насибуллин Т.Р. и др.* Alu-инсерционно-делеционный полиморфизм генов *COL13A1* и *LAMA2*: анализ ассоциаций с долгожительством // Генетика. 2016. Т. 52. №. 10. С. 1185–1193.  
<https://doi.org/10.7868/S0016675816100039>
9. *Erdman V.V., Karimov D.D., Tuktarova I.A. et al.* Alu deletions in *LAMA2* and *CDH4* genes are key components of polygenic predictors of longevity // Intern. J. of Mol. Sci. 2023. № 21.  
<https://doi.org/10.3390/ijms232113492>
10. *Wang D., He J., Huang B. et al.* Emerging role of the Hippo pathway in autophagy // Cell Death & Disease. 2020. V. 11. № 10. P. 880.  
<https://doi.org/10.1038/s41419-020-03069-6>
11. *Zhou Y.H., Huang T.T., Cheng A.S.L. et al.* The TEAD family and its oncogenic role in promoting tumorigenesis // Intern. J. Mol. Sci. 2016. V. 17. № 1. P. 138.  
<https://doi.org/10.3390/ijms17010138>
12. *Reed M.J., Damodarasamy M., Banks W.A.* The extracellular matrix of the blood-brain barrier: Structural and functional roles in health, aging, and Alzheimer’s disease // Tissue Barriers. 2019. V. 7. № 4.  
<https://doi.org/10.1080/21688370.2019.1651157>
13. *Carmignac V., Svensson M., Körner Z. et al.* Autophagy is increased in laminin  $\alpha 2$  chain-deficient muscle and its inhibition improves muscle morphology in a mouse model of MDC1A // Human Mol. Genet. V. 20. № 24. P. 4891–4902.  
<https://doi.org/10.1093/hmg/ddr427>
14. *Fard D., Tamagnone L.* Semaphorins in health and disease // Cytokine & Growth Factor Reviews. 2021. V. 57. P. 55–63.  
<https://doi.org/10.1016/j.cytogfr.2020.05.006>
15. *Zhang C., Hong C.D., Wang H.L. et al.* The role of semaphorins in small vessels of the eye and brain // Pharmacol. Research. 2020. V. 160.  
<https://doi.org/10.1016/j.phrs.2020.105044>
16. *Kaushik A., Parashar S., Ambasta R.K., Kumar P.* Ubiquitin E3 ligases assisted technologies in protein degradation: Sharing pathways in neurodegenerative disorders and cancer // Ageing Res. Reviews. 2024. V. 96. P. 102279.  
<https://doi.org/10.1016/j.arr.2024.102279>
17. *Le D., Brown L., Malik K., Murakami S.* Two opposing functions of angiotensin-converting enzyme (ACE) that links hypertension, dementia, and aging // Intern. J. Mol. Sci. 2021. V. 22. № 24.  
<https://doi.org/10.3390/ijms222413178>
18. *Loos R.J.F., Yeo G.S.H.* The genetics of obesity: from discovery to biology // Nature Rev. Genet. 2022. V. 23. № 2. P. 120–133.  
<https://doi.org/10.1038/s41576-021-00414-z>
19. *Yepes M.* The plasminogen activation system promotes neurorepair in the ischemic brain // Current Drug Targets. 2019. V. 20. № 9. P. 953–959.  
<https://doi.org/10.2174/138945012066181211144550>
20. *Resink T.J., Joshi M.B., Kyriakakis E.* Cadherins and cardiovascular disease // Swiss Med. Weekly. 2009. V. 139. № 0910. P. 122–134.
21. *Cordaux R., Batzer M.A.* The impact of retrotransposons on human genome evolution // Nature Rev. Genet. 2009. V. 10. № 10. P. 691–703.  
<https://doi.org/10.1038/nrg2640>
22. *Nefedova L.N., Kim A.I.* The role of retroelements in the evolution of animal genomes // Biol. Bul. Reviews. 2022. V. 12. № 1. P. 29–40.  
<https://doi.org/10.1134/S2079086422010042>
23. *Davidson-Pilon C.* Lifelines: Survival analysis in Python // J. Open Source Software. 2019. V. 4. № 40. P. 1317.  
<https://doi.org/10.21105/joss.01317>
24. *Cao L., Li H., Liu X. et al.* Expression and regulatory network of E3 ubiquitin ligase NEDD4 family in cancers // BMC Cancer. 2023. V. 23. № 1. P. 526.  
<https://doi.org/10.1186/s12885-023-11007-w>

25. *Li Y., Zhang L., Zhou J. et al.* Nedd4 E3 ubiquitin ligase promotes cell proliferation and autophagy // *Cell Proliferation*. 2015. V. 48. № 3. P. 338–347. <https://doi.org/10.1111/cpr.12184>
26. *Li Y., Ozaki T., Kikuchi H. et al.* A novel HECT-type E3 ubiquitin protein ligase NEDL1 enhances the p53-mediated apoptotic cell death in its catalytic activity-independent manner // *Oncogene*. 2008. V. 27. № 26. P. 3700–3709. <https://doi.org/10.1038/sj.onc.1211032>
27. *Quiroga M., Rodriguez-Alons A., Alfonsin G. et al.* Protein degradation by E3 ubiquitin ligases in cancer stem cells // *Cancers*. 2022. V. 14. <https://doi.org/10.3390/cancers14040990>
28. *Huang S.S., Hsu L.J., Chang N.S.* Functional role of WW domain-containing proteins in tumor biology and diseases: Insight into the role in ubiquitin-proteasome system // *FASEB Bioadv.* 2020. V. 2. P. 234–253. <https://doi.org/10.1096/fba.2019-00060>
29. Гомбоева Д.Е., Брагина Е.Ю., Назаренко М.С., Пузырев В.П. Обратная коморбидность между онкологическими заболеваниями и болезнью Гентингтона: обзор эпидемиологических и биологических доказательств // Генетика. 2020. Т. 56. № 3. С. 260–271. <https://doi.org/10.31857/S0016675820030054>
30. *Piccolo S., Dupont S., Cordenonsi M.* The biology of YAP/TAZ: Hippo signaling and beyond // *Physiol. Reviews*. 2014. V. 94. № 4. P. 1287–1312. <https://doi.org/10.1152/physrev.00005.2014>
31. *Ramaccini D., Pedriali G., Perrone M. et al.* Some insights into the regulation of cardiac physiology and pathology by the Hippo pathway // *Biomedicines*. 2022. V. 10. № 3. P. 726. <https://doi.org/10.3390/biomedicines10030726>
32. *Lin K.C., Park H.W., Guan K.L.* Regulation of the Hippo pathway transcription factor TEAD // *Trends Biochem. Sci.* 2017. V. 42. P. 862–872. <https://doi.org/10.1016/j.tibs.2017.09.003>
33. *Zhang Y., Ren Y., Li X. et al.* A review on decoding the roles of YAP/TAZ signaling pathway in cardiovascular diseases: Bridging molecular mechanisms to therapeutic insights // *Intern. J. Biol. Macromolecules*. 2024. <https://doi.org/10.1016/j.ijbiomac.2024.132473>
34. *Hergovich A.* The roles of NDR protein kinases in Hippo signalling // *Genes*. 2016. V. 7. № 5. P. 21. <https://doi.org/10.3390/genes7050021>
35. *Sharif A.A.D., Hergovich A.* The NDR/LATS protein kinases in immunology and cancer biology // *Seminars in Cancer Biology*. 2018. V. 48. P. 104–114.
36. *Jonischkies K., Del Angel M., Demiray Y.E. et al.* The NDR family of kinases: Essential regulators of aging // *Frontiers in Mol. Neurosci.* 2024. V. 17. <https://doi.org/10.3389/fnmol.2024.1371086>
37. *Rawat P., Thakur S., Dogra S. et al.* Diet-induced induction of hepatic serine/threonine kinase STK38 triggers proinflammation and hepatic lipid accumulation // *J. Biol. Chemistry*. 2023. V. 299. № 5. <https://doi.org/10.1016/j.jbc.2023.104678>
38. *Aman Y., Schmauck-Medina T., Hansen M. et al.* Autophagy in healthy aging and disease // *Nat. Aging*. 2021. V. 1. № 8. P. 634–650. <https://doi.org/10.1038/s43587-021-00098-4>

## Alu Polymorphisms of Autophagy and Apoptosis Regulatory Genes As Human Lifespan Factors

**V. V. Erdman<sup>1, 3, \*</sup>, D. D. Karimov<sup>1, 2</sup>, I. A. Tuktarova<sup>1</sup>, A. A. Petintseva<sup>1</sup>,  
Y. R. Timasheva<sup>1, 3</sup>, T. R. Nasibullin<sup>1</sup>**

<sup>1</sup>*Institute of Biochemistry and Genetics – Subdivision of the Ufa Federal Research Centre of the Russian Academy of Sciences, Ufa, 450054 Russia*

<sup>2</sup>*Ufa Research Institute of Occupational Medicine and Human Ecology, Ufa, 450106 Russia*

<sup>3</sup>*Bashkir State Medical University, Ufa, 450008 Russia*

\*e-mail: danivera@mail.ru

To assess the contribution to survival of Alu-insertions in the *ACE*, *PLAT*, *COL13A1*, *LAMA2*, *CDH4*, *SEMA6A*, *PKHD1L1*, *STK38L*, *HECW1*, *TEAD1* genes, which are candidates of aging and longevity, amid the senile physiological and pathological phenotype, was carried out the association analysis with life expectancy. Survival and mortality data were obtained for 1,382 elderly people, who were selected from the sample of Tatars residing in the Republic of Bashkortostan (total 1790 people from 18 to 109 years). Mortality risk was higher among carriers of the *STK38L* Alu-insertion genotype (Ya5ac2145\*II, *HR* = 2.07, *P* = 0.02). Alu-insertion in the *HECW1* and *TEAD1* genes has demonstrated a survival protection effect (Ya5NBC182\*II, *HR* = 0.71, *P* = 0.038 and Ya5ac2013\*II, *HR* = 0.74, *P* = 0.035 respectively). The survival amid the persons with various clinical phenotypes was associated with the Alu polymorphism of the *SEMA6A* (Yb8NBC597\*ID, *HR* = 0.54, *P* = 0.016 for the cerebrovascular diseases), *TEAD1* (Ya5ac2013\*II, *HR* = 0.57, *P* = 0.016 for the cardiovascular pathologies) and *LAMA2* (Ya5-MLS19\*ID, *HR* = 0.36, *P* = 0.03 for multimorbidity status) genes. Thus, the genes involved in the regulation of autophagy and apoptosis were associated with survival and longevity.

**Keywords:** aging, longevity, Alu-polymorphism, *TEAD1*, *HECW1*, *STK38L*, *LAMA2*, *SEMA6A* genes, survival analysis.

## АССОЦИАЦИЯ VNTR-ПОЛИМОРФИЗМА В ГЕНЕ *MIR137* С КОГНИТИВНЫМИ ФУНКЦИЯМИ У БОЛЬНЫХ ШИЗОФРЕНИЕЙ И ЗДОРОВЫХ ЛИЦ

© 2025 г. М. В. Алфимова<sup>1</sup> \*, Г. И. Коровайцева<sup>1</sup>, В. В. Плакунова<sup>1</sup>, В. Е. Голимбет<sup>1</sup>

<sup>1</sup>Научный центр психического здоровья, Москва, 115522 Россия

\*e-mail: m.alfimova@gmail.com

Поступила в редакцию 25.06.2024 г.

После доработки 30.07.2024 г.

Принята к публикации 02.08.2024 г.

Цель исследования состояла в анализе ассоциаций между когнитивными доменами, нарушенными при шизофрении, и полиморфизмом rs58335419, расположенным в GWAS-значимом локусе риска шизофрении – в гене *MIR137*, кодирующем miR-137. Больные с диагнозами шизофренического спектра ( $n = 787$ ) и здоровые добровольцы без наследственной отягощенности психозами ( $n = 622$ ) выполнили тесты на семантическую вербальную беглость (ВБ), внимание/рабочую память, вербальную эпизодическую память и управляющие функции. После коррекции на множественность сравнений нарушения ВБ были ассоциированы с гомозиготным носительством распространенного аллеля (с тремя повторами) у больных мужчин. Аналогичные тенденции имели место в объединенной выборке для внимания/рабочей памяти и общего индекса когнитивного функционирования. Аллель с четырьмя повторами не был связан с вариативностью когнитивных функций. Полученные результаты указывают на ассоциацию между гомозиготным носительством распространенного аллеля rs58335419 и более низкими показателями когнитивных функций, что противоположно действию данного полиморфизма на риск развития заболевания.

**Ключевые слова:** шизофрения, микроРНК, miR-137, VNTR, когнитивные функции, вербальная беглость.

**DOI:** 10.31857/S0016675825010103 **EDN:** VECUEW

Полногеномные исследования ассоциаций (GWAS), проведенные Консорциумом по психиатрической геномике, обнаружили высоко значимые связи риска шизофрении с полиморфными сайтами в гене *MIR137*, кодирующем микроРНК-137 (miR-137) [1]. Современные данные позволяют предположить существование в этом локусе нескольких автономных генетических вариантов, связанных с шизофренией [2]. Одним из них может быть расположенный внутри транскрипта pri-miR-137 вблизи последовательности pre-miR-137 полиморфизм rs58335419, представляющий собой вариабельное число tandemных повторов (VNTR) длиной 15 п.н. В популяциях обнаруживаются аллели с числом повторов от 3 до 14. Распространенным является аллель с тремя повторами ( $R3$ ). VNTR не сцеплен с однонуклеотидными маркерами, обнаруженными в GWAS шизофрении [2], и данные о направлении его влияния на риск развития болезни неоднозначны: имеются указания как на протективную роль мутантных аллелей (с

числом повторов больше трех,  $R>3$ ) [2], так и на их связь с возрастанием риска [3, 4].

Полиморфизм rs58335419 предположительно влияет на альтернативный сплайсинг pri-miR-137: аллели  $R > 3$  связаны с повышением числа транскриптов, конкурирующих с классическим вариантом miR-137; и таким образом их носительство может вести к усилению функционирования генов-мишеней miR-137. Один из потенциальных механизмов этого влияния – изменение длины CpG-островка, внутри которого находится *MIR137*, что сказывается на содержании динуклеотидов CpG и, возможно, на метилировании данного локуса [2].

Помимо шизофрении, генетические варианты в гене *MIR137* также ассоциированы с когнитивными способностями (по данным GWAS интеллекта [5]). Это позволяет предположить, что полиморфизм *MIR137* может влиять не только на риск заболевания, но и на вариативность сопутствующего заболеванию когнитивного дефицита. В ряде работ показана связь аллелей риска индексных

полиморфизмов GWAS шизофрении с нарушением скорости обработки информации, внимания, рабочей и эпизодической памяти [6–8], в других исследованиях связей с когнициями не найдено [9, 10]. Относительно ассоциаций с rs58335419 опубликовано две работы. В одной из них показано накопление *R4* в группе больных шизофренией с выраженным когнитивным дефицитом по сравнению с группами когнитивно сохранных больных и здорового контроля [4]. В другой здоровые выполняли тест на оценку управляющих функций (тест Струпа), и носители распространенного аллеля продемонстрировали более высокие значения одного из нестандартных показателей теста [11]. Таким образом, в обеих работах более высокие когнитивные показатели были связаны с распространенным аллелем. Однако эти данные получены на небольших выборках и для ограниченного числа когнитивных доменов и требуют подтверждения и расширения.

Цель исследования – оценка ассоциации rs58335419 с когнитивными функциями, нарушенными при шизофрении, у больных расстройствами шизофренического спектра и здоровых лиц в популяции русских. На основании описанных выше данных [3, 4, 11] мы предполагали, что носительство аллелей  $R > 3$  будет приводить к снижению когнитивных показателей в обеих группах. Дополнительно, основываясь на гипотезе о механизме действия VNTR на сплайсинг pri-miR-137 [2] и наших данных о связи единственного вариативно метилированного CpG (chr1:98511049, hg19) в CpG-островке *MIR137* с когнитивным дефицитом при шизофрении [14], мы оценили в подгруппах больных и здоровых ассоциацию между VNTR и метилированием этого CpG.

Выборка состояла из 787 больных психозами (средний возраст  $33.97 \pm 10.96$  лет; 56% женщин; рубрики МКБ-10: F20 – 687 человек, F23 – 45, F25 – 55; длительность заболевания  $8.38 \pm 8.63$  лет). Также обследовали 622 здоровых добровольца без наследственной отягощенности психозами ( $28.97 \pm 9.79$  лет; 61% женщин), прошедших когнитивное тестирование. Среди больных имели или получали высшее образование 63%, среди здоровых – 78%. Все участники были частью описанной ранее выборки, в которой показана ассоциация аллелей  $R > 3$  с риском шизофрении [3]. Критериями включения в данное исследование для больных и здоровых были: 1) возраст от 16 до 65 лет, 2) образование не менее девяти классов, 3) отсутствие дополнительных диагнозов, связанных с нарушением когнитивных функций. Когнитивное обследование больных проводили после стабилизации состояния. На момент обследования выраженность симптомов по Шкале позитивных и негативных синдромов (PANSS) составляла: позитивных

–  $26.99 \pm 8.42$  балла, негативных –  $21.17 \pm 7.84$ , общих психопатологических –  $33.76 \pm 12.31$ .

Когнитивное обследование включало стандартные методики для оценки различных доменов и занимало около 30 минут [12]. В настоящей работе были использованы результаты выполнения тестов на скорость обработки информации (семантическая вербальная беглость, ВБ), эпизодическую вербальную память (суммарное количество слов, воспроизведенных за пять повторений в Тесте слухоречевой памяти Рея, RAVLT), внимание и рабочую память (количество верно выполненных за одну минуту операций при обратном отсчитывании от 200 по 2 и 5), когнитивную гибкость (время выполнения Теста следования по маршруту, часть Б; TMT-В) и когнитивное торможение (коэффициент интерференции в тесте Струпа). Показатели были стандартизованы с помощью среднего и стандартного отклонения группы здоровых. Для лиц, которые полностью выполнили не менее трех тестов из пяти, также вычисляли общий когнитивный индекс, который представлял собой усредненную оценку по всем тестам. Генотипы в сайте rs58335419 определяли методом полимеразной цепной реакции, как подробно описано в предыдущей работе [3]. Использовали полученные ранее данные по метилированию для 43 больных и 38 здоровых [12]. Протокол исследования был одобрен этическим комитетом ФГБНУ НЦПЗ (№ 98 от 11.09.2007). Все участники подписали информированное согласие.

Связь между генотипом и когнитивными признаками оценивали с помощью ковариационного анализа (ANCOVA) с использованием программы JASP 0.18.3 [13]. Решение о применении данного параметрического метода принимали в соответствии с рекомендациями для больших выборок ( $n > 500$ ) на основании значений асимметрии и эксцесса распределения когнитивных показателей [14]. Гомозигот по аллелю  $R > 3$  объединили с гетерозиготами, в качестве ковариат вводили пол и возраст. Рассматривали главный эффект генотипа ( $R3/R3$  vs  $R > 3$ ) и эффекты его взаимодействия с диагностической группой (больные vs здоровые) и полом. Post-hoc – попарные сравнения – проводили методом Тьюки с использованием 1000 бутстреп-выборок. Учитывая, что мы анализировали пять когнитивных доменов, для основного анализа уровень значимости принимали равным  $p < 0.01$ . В остальных случаях использовали стандартный порог  $p < 0.05$ . Основываясь на данных о накоплении аллеля *R4* в группе больных с когнитивным дефектом [4], анализ повторяли для выявления эффекта данного аллеля.

Пациенты отличались от здоровых по возрасту ( $t = 8.91$ ,  $p < 0.001$ ) и уровню образования ( $\chi^2 = 35.23$ ,  $p < 0.001$ ). Все когнитивные показатели

пациентов были снижены относительно здоровых на 0.9 стандартного отклонения и более (для всех главных эффектов диагностической группы в ANCOVA  $p < 0.001$ ). Частоты генотипов и аллелей были такими же, как в выборке, частью которой являлась выборка настоящего исследования [3]. Больные:  $R3/R3 = 0.56$ ;  $R3/R > 3 = 0.38$ ;  $R > 3/R > 3 = 0.06$ . Здоровые:  $R3/R3 = 0.61$ ;  $R3/R > 3 = 0.34$ ;  $R > 3/R > 3 = 0.05$ . Для аллеля  $R3$  частоты составили 0.75 и 0.78, для  $R4 = 0.09$  и 0.08 у больных и здоровых соответственно. У больных наблюдалось повышение (относительно здоровых) частоты аллеля  $R > 3$  ( $\chi^2 = 4.22$ ,  $p = 0.040$ ) и тенденция к снижению количества гомозигот  $R3/R3$  ( $\chi^2 = 4.69$ ,  $p = 0.096$ ); по частоте  $R4$  различия между больными и здоровыми не были значимы ( $\chi^2 = 0.26$ ,  $p = 0.608$ ).

ANCOVA выявил главный эффект генотипа и эффект взаимодействия генотипа и пола на ВБ (табл. 1). Тест Тьюки подтвердил влияние генотипа на ВБ в группе больных мужчин ( $p = 0.002$ , Cohen's  $d = 0.43$ ). Пациенты-мужчины с генотипом  $R3/R3$  имели худшие показатели ВБ среди всех участников. Дальнейший анализ показал, что у больных различия по ВБ остаются значимыми при учете длительности болезни и выраженности трех синдромов PANSS (главный эффект генотипа:  $F(1, 562) = 8.30$ ,  $p = 0.004$ ,  $\eta^2_p = 0.02$ ; эффект взаимодействия пол x генотип:  $F(1, 562) = 4.04$ ,  $p = 0.045$ ;  $\eta^2_p = 0.01$ ). Общий когнитивный индекс отчасти повторял этот паттерн: самые низкие значения когнитивного индекса наблюдались у пациентов-мужчин с генотипом  $R3/R3$ , однако различия не были достоверны. Анализ для  $R4$  не выявил значимых эффектов генотипа на когнитивные показатели.

Ни главный эффект генотипа, ни эффекты его взаимодействия с группой и полом на метилирование CpG с координатой chr1:98511049 не были значимы. При учете выраженности симптомов и длительности болезни у больных имела место тенденция к взаимодействию генотипа с полом ( $F(1, 19) = 5.23$ ,  $p = 0.034$ ,  $\eta^2_p = 0.22$ ). Самые высокие показатели метилирования наблюдались у пациентов-мужчин с генотипом  $R3/R3$ . При этом у больных метилирование коррелировало с ВБ ( $r = -0.33$ ,  $p = 0.036$ ).

Таким образом, в настоящей работе мы впервые в большой когорте больных шизофренией и здоровых лиц проанализировали ассоциации rs58335419 в гене *MIR137* с когнитивными доменами, обычно нарушенными у пациентов, а также выявили у больных пол-специфичный эффект данного VNTR-полиморфизма на семантическую вербальную беглость, которая рассматривается в качестве показателя скорости обработки информации. Вопреки гипотезе, распространенный аллель

был связан с ухудшением ВБ у пациентов-мужчин. Аналогичные тенденции наблюдались для объединенной группы больных и здоровых: гомозиготы по общему аллелю демонстрировали тенденцию к снижению продуктивности внимания/рабочей памяти и индекса общего когнитивного функционирования. Отличие этого результата от данных двух предыдущих исследований – австралийского и колумбийского [4, 11] – может объясняться как этническими особенностями выборок, так и различиями в оценке когниций. Первое предположение представляется маловероятным в связи со сходством частот аллелей в нашей работе и в этих исследованиях. Относительно когнитивных показателей следует отметить, что Y. González-Giraldo с соавт. [11], как и мы, не нашли связи генотипа с главным показателем теста Струпа, отражающим когнитивное торможение, а при сравнении групп с выраженным когнитивным дефицитом и без такого в работе E. Mahmoudi с соавт. [4] на результаты могла повлиять общая тяжесть заболевания. На это указывает тенденция к накоплению аллеля  $R4$  не только у пациентов с выраженным дефицитом, но и, хотя и в меньшей степени, у когнитивно сохранных пациентов относительно контроля.

Расхождение в направлении влияния распространенного аллеля на риск заболевания и уровень когнитивного функционирования является неожиданным и требует дальнейшего изучения. Ранее, основываясь на том, что увеличение числа повторов с двух до трех и более является характеристикой линии человека и отличает его от других приматов, A. Pacheco с соавт. [2] предположили, что экспансия VNTR эволюционно выгодна для когнитивных функций из-за потенциальной негативной регуляции miR-137 со стороны длинных аллелей. Наши данные согласуются с этой гипотезой, однако содержательный анализ связи когниций с VNTR затруднен из-за того, что miR-137 принадлежит к числу микроРНК, которые могут оказывать как активирующее, так и ингибирующее воздействие на когнитивные процессы, в частности, связанные с обучением и долговременной памятью [15], а также из-за противоречивости данных о направлении изменений концентрации miR-137 при шизофрении [16, 17].

Для метилирования CpG с координатой chr1:98511049 мы не нашли значимого влияния со стороны VNTR. Однако имелась тенденция, указывающая на повышение метилирования у больных мужчин с генотипом  $R3/R3$ , что перекликается с нашими данными об отрицательной корреляции метилирования в этом сайте и когнитивным индексом у пациентов мужского пола [12]. Кроме того, выявленный нами пол-специфичный эффект VNTR в определенной степени соответствует доклиническим данным о половых различиях

**Таблица 1.** Значения когнитивных показателей и метилирования у больных и здоровых лиц в зависимости от пола и генотипа в сайте rs58335419

Признак	Пациенты		Контроль		Значимость эффекта
	гомозиготы <i>R3/R3</i>	носители аллеля <i>R &gt; 3</i>	гомозиготы <i>R3/R3</i>	носители аллеля <i>R &gt; 3</i>	
Вербальная беглость					
Мужчины	$-1.73 \pm 0.95$ <i>n</i> = 179	$-1.28 \pm 1.03$ <i>n</i> = 159	$-0.28 \pm 1.02$ <i>n</i> = 139	$0.01 \pm 0.94$ <i>n</i> = 93	Г: $F = 10.06$ , $p = 0.002^*$ ; ГхД: ns; ГхП: $F = 10.07$ , $p = 0.002^*$ ; ГхДхП: ns
Женщины	$-1.15 \pm 1.19$ <i>n</i> = 250	$-1.14 \pm 1.16$ <i>n</i> = 184	$0.11 \pm 0.97$ <i>n</i> = 231	$0.09 \pm 1.02$ <i>n</i> = 143	
Внимание/рабочая память					
Мужчины	$-1.01 \pm 1.05$ <i>n</i> = 130	$-0.88 \pm 1.01$ <i>n</i> = 121	$0.10 \pm 1.06$ <i>n</i> = 129	$0.18 \pm 1.11$ <i>n</i> = 91	Г: $F = 5.24$ , $p = 0.022$ ; ГхД: ns; ГхП: ns; ГхДхП: ns
Женщины	$-1.21 \pm 0.92$ <i>n</i> = 203	$-1.10 \pm 0.97$ <i>n</i> = 156	$-0.17 \pm 0.91$ <i>n</i> = 203	$0.05 \pm 0.96$ <i>n</i> = 119	
Эпизодическая вербальная память					
Мужчины	$-1.68 \pm 1.02$ <i>n</i> = 138	$-1.42 \pm 1.04$ <i>n</i> = 145	$-0.46 \pm 1.00$ <i>n</i> = 90	$-0.37 \pm 0.92$ <i>n</i> = 55	Г: ns; ГхД: ns; ГхП: ns; ГхДхП: ns
Женщины	$-1.13 \pm 1.02$ <i>n</i> = 207	$-1.01 \pm 1.20$ <i>n</i> = 167	$0.27 \pm 0.92$ <i>n</i> = 157	$0.30 \pm 0.91$ <i>n</i> = 79	
Когнитивная гибкость					
Мужчины	$-2.13 \pm 1.82$ <i>n</i> = 138	$-2.22 \pm 1.94$ <i>n</i> = 137	$-0.14 \pm 1.02$ <i>n</i> = 94	$0.13 \pm 0.80$ <i>n</i> = 60	Г: ns; ГхД: $F = 5.68$ , $p = 0.017$ ; ГхП: ns; ГхДхП: ns
Женщины	$-2.05 \pm 2.08$ <i>n</i> = 199	$-2.46 \pm 2.25$ <i>n</i> = 163	$-0.07 \pm 1.08$ <i>n</i> = 172	$0.18 \pm 0.93$ <i>n</i> = 90	
Когнитивное торможение					
Мужчины	$-0.68 \pm 0.92$ <i>n</i> = 127	$-0.85 \pm 0.91$ <i>n</i> = 125	$-0.004 \pm 1.07$ <i>n</i> = 87	$-0.02 \pm 0.96$ <i>n</i> = 56	Г: ns; ГхД: ns; ГхП: ns; ГхДхП: ns
Женщины	$-0.92 \pm 1.04$ <i>n</i> = 175	$-1.00 \pm 1.00$ <i>n</i> = 143	$-0.06 \pm 1.08$ <i>n</i> = 150	$0.13 \pm 0.84$ <i>n</i> = 88	
Общий когнитивный индекс					
Мужчины	$-1.48 \pm 0.81$ <i>n</i> = 146	$-1.37 \pm 0.84$ <i>n</i> = 145	$-0.16 \pm 0.67$ <i>n</i> = 93	$0.06 \pm 0.62$ <i>n</i> = 60	Г: $F = 4.95$ , $p = 0.026$ ; ГхД: ns; ГхП: ns; ГхДхП: ns
Женщины	$-1.30 \pm 0.90$ <i>n</i> = 212	$-1.38 \pm 1.00$ <i>n</i> = 171	$-0.03 \pm 0.62$ <i>n</i> = 173	$0.19 \pm 0.59$ <i>n</i> = 92	
Метилирование CpG на chr1:98511049. hg19					
Мужчины	$0.42 \pm 0.13$ <i>n</i> = 10	$0.34 \pm 0.18$ <i>n</i> = 9	$0.28 \pm 0.24$ <i>n</i> = 12	$0.35 \pm 0.20$ <i>n</i> = 4	Г: ns; ГхД: ns; ГхП: ns; ГхДхП: ns
Женщины	$0.21 \pm 0.19$ <i>n</i> = 15	$0.34 \pm 0.19$ <i>n</i> = 9	$0.29 \pm 0.18$ <i>n</i> = 17	$0.28 \pm 0.23$ <i>n</i> = 5	

Примечание.  $F$ -критерий и  $p$  представлены для номинально значимых эффектов. \* — эффекты, выдержавшие поправку на множественность сравнений. Г — генотип, Д — диагностическая группа, П — пол. Даны средние  $\pm$  стандартное отклонение.

уровней miR-137 в префронтальной коре [18], а также данным о негативной корреляции между периферическими уровнями эстрадиола и miR-137 у больных шизофренией женщин [17]. В целом полученные результаты указывают на ассоциацию между гомозиготным носительством распространенного аллеля rs58335419 и снижением скорости обработки информации при шизофрении в популяции русских, что противоположно действию данного полиморфизма на риск развития заболевания.

Работа выполнена в рамках государственного задания.

Исследование одобрено Этическим комитетом ФГБНУ НЦПЗ, (11.09.2007, № 98).

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 были старше 15 лет и подписали информированное согласие в соответствии с законом РФ “О психиатрической помощи и гарантиях прав граждан при ее оказании”.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. *Trubetskoy V., Pardiña A.F., Qi T. et al.* Mapping genomic loci implicates genes and synaptic biology in schizophrenia // *Nature*. 2022. V. 604. P. 502–508. <https://doi.org/10.1038/s41586-022-04434-5>
2. *Pacheco A., Berger R., Freedman R., Law A.J.* A VNTR regulates miR-137 expression through novel alternative splicing and contributes to risk for schizophrenia // *Sci. Rep.* 2019. V. 9. № 1. P. 11793–11804. <https://doi.org/10.1038/s41598-019-48141-0>
3. *Коровайцева Г.И., Олейчик И.В., Лежейко Т.В., Голимбет В.Е.* Изучение ассоциации VNTR-полиморфизма rs58335419 гена *MIR137* с риском развития шизофрении // Генетика 2024. Т. 60. № 2. С. 63–69. <https://doi.org/10.31857/S0016675824020065>.
4. *Mahmoudi E., Atkins J.R., Quidé Y. et al.* The *MIR137* VNTR rs58335419 is associated with cognitive impairment in schizophrenia and altered cortical morphology // *Schizophr. Bull.* 2021. V. 47. № 2. P. 495–504. <https://doi.org/10.1093/schbul/sbaa123>
5. *Savage J.E., Jansen P.R., Stringer S. et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence // *Nat. Genet.* 2018. V. 50. P. 912–919. <https://doi.org/10.1038/s41588-018-0152-6>
6. *Cummings E., Donohoe G., Hargreaves A. et al.* Mood congruent psychotic symptoms and specific cognitive deficits in carriers of the novel schizophrenia risk variant at MIR-137 // *Neurosci. Lett.* 2013. V. 532. P. 33–38. <https://doi.org/10.1016/j.neulet.2012.08.065>
7. *Kuswanto C.N., Sum M.Y., Qiu A. et al.* The impact of genome wide supported microRNA-137 (*MIR137*) risk variants on frontal and striatal white matter integrity, neurocognitive functioning, and negative symptoms in schizophrenia // *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* 2015. V. 168B. P. 317–326. <https://doi.org/10.1002/ajmg.b.32314>
8. *Cosgrove D., Harold D., Mothersill O. et al.* MiR-137-derived polygenic risk: effects on cognitive performance in patients with schizophrenia and controls // *Transl. Psychiatry*. 2017. V. 7. № 1. P. e1012. <https://doi.org/10.1038/tp.2016.286>
9. *Van Erp T.G.M., Guella I., Vawter M.P. et al.* Schizophrenia miR-137 locus risk genotype is associated with DLPFC hyperactivation // *Biol. Psychiatry*. 2014. V. 75. № 5. P. 398–405. <https://doi.org/10.1016/j.biopsych.2013.06.016>
10. *Kandratsenka H., Nestsiarovich A., Goloenko I. et al.* Association of *MIR137* with symptom severity and cognitive functioning in Belarusian schizophrenia patients // *Front. Psychiatry*. 2018. V. 9. <https://doi.org/10.3389/fpsyg.2018.00295>
11. *González-Giraldo Y., González-Reyes R.E., Forero D.A.* A functional variant in *MIR137*, a candidate gene for schizophrenia, affects Stroop test performance in young adults // *Psychiatry. Res.* 2016. V. 236. P. 202–205. <https://doi.org/10.1016/j.psychres.2016.01.006>
12. *Алфимова М.В., Кондратьев Н.В., Голов А.К., Голимбет В.Е.* Метилирование ДНК в локусе *MIR137HG*, ассоциированном с шизофренией и интеллектом, может быть связано с заболеванием и когнитивными функциями // Генетика. 2019. Т. 55. № 2. С. 207–213.
13. *JASP (version 0.18.3) [Computer software]*. 2024. <https://jasp-stats.org/>, дата обращения 01.03.2024.
14. *Kim H.Y.* Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis // *Restor. Dent. Endod.* 2013. V. 38. P. 52–54. <https://doi.org/10.5395/rde.2013.38.1.52>
15. *Grinkevich L.N.* The role of microRNAs in learning and long-term memory // *Vavilov J. Genet. and Breeding*. 2020. V. 24. № 8. <https://doi.org/10.18699/VJ20.687>
16. *Liu X., Dong L., Jiang Z. et al.* Identifying the differentially expressed peripheral blood microRNAs in psychiatric disorders: A systematic review and meta-analysis. // *Front. Psychiatry*. 2024. V. 15. <https://doi.org/10.3389/fpsyg.2024.1390366>

17. Peng Q., Dai Z., Yin J. et al. Schizophrenia plausible protective effect of microRNA-137 is potentially related to estrogen and prolactin in female patients // *Front. Psychiatry*. 2023. V. 14. <https://doi.org/10.3389/fpsyg.2023.1187111>
18. Garcia-Cerro S., Gómez-Garrido A., Garcia G. et al. Exploratory analysis of microRNA alterations in a neurodevelopmental mouse model for autism spectrum disorder and schizophrenia // *Int. J. Mol. Sci.* 2024. V. 25. № 5. <https://doi.org/10.3390/ijms25052786>

## Association of the VNTR Polymorphism in the *MIR137* Gene with Cognitive Functions in Schizophrenia Patients and Healthy Individuals

**M. V. Alfimova<sup>1,\*</sup>, G. I. Korovaitseva<sup>1</sup>, V. V. Plakunova<sup>1</sup>, V. E. Golimbet<sup>1</sup>**

<sup>1</sup>*Mental Health Research Center, Moscow, 115522 Russia*

\*e-mail: [m.alfimova@gmail.com](mailto:m.alfimova@gmail.com)

The study aimed to analyze the associations between cognitive domains impaired in schizophrenia and the rs58335419 polymorphism located within the schizophrenia GWAS-significant locus, in the gene *MIR137* encoding miR-137. Schizophrenia spectrum patients (n = 787) and healthy volunteers without a family history of psychosis (n = 622) completed tests of semantic verbal fluency (VF), attention/working memory, verbal episodic memory, and executive functions. After correction for multiple testing, VF abnormalities were associated with homozygosity for the common allele (with three repeats) in male patients. Similar trends occurred in the pooled sample for attention/working memory and the general index of cognitive functioning. The four-repeat allele was not associated with variance in cognitive performance. The results obtained indicate an association between homozygosity for the common allele at rs58335419 and lower levels of cognitive functions, which is opposite to the effect of this polymorphism on the risk of developing schizophrenia.

**Keywords:** schizophrenia, microRNA, miR-137, VNTR, cognition, verbal fluency.