СОДЕРЖАНИЕ

Том 57, номер 11, 2021

Обзорные и теоретические статьи

Метилирование генов в свободной циркулирующей ДНК плазмы крови как прогностический и предиктивный факторы при раке молочной железы
С. В. Рыков, Е. А. Филиппова, В. И. Логинов, Э. А. Брага 1225

Фенотипически нестабильные мутации – маркеры хромосомных перестроек с участием ДНК-транспозонов
Л. П. Захаренко 1241

Генетика растений

Изменчивость митохондриальной ДНК указывает на американское происхождение современных пихт
В. Л. Семериков, С. А. Семерикова, Ю. А. Путинцева 1246

Генетика животных

Изменчивость контрольного региона мтДНК восточноевропейской полевки *Microtus rossiaemeridionalis* (Rodentia, Arvicolini) из двух инвазивных популяций юга Дальнего Востока России
И. Н. Шереметьева, И. В. Картавцева, А. А. Емельянова, А. С. Лапин 1252

Генетика человека

Особенности геномного распределения регионов высокой гомозиготности у коренного населения Северной Евразии на индивидуальном и популяционном уровнях на основе анализа SNP высокой плотности
Н. А. Колесников, В. Н. Харьков, А. А. Зарубин, М. О. Раджабов, М. И. Воевода, М. А. Губина, Э. К. Хуснутдинова, С. С. Литвинов, Н. В. Екомасова, О. В. Штыгашева, Н. Р. Максимова, А. Л. Сухомясова, В. А. Степанов 1261

Биоинформационная аннотация генов предрасположенности к болезни Альцгеймера и ишемической болезни сердца
Н. Ю. Часовских, Е. Е. Чижик, А. А. Бобрышева 1276

Роль генов нейротрансмиттерной системы в развитии хронической обструктивной болезни легких
Г. Ф. Корытина, Л. З. Ахмадишина, О. В. Кочетова, Т. Р. Насибуллин, Ю. Г. Азнабаева, Ш. Р. Зулкарнеев, С. М. Измайлова, Н. Ш. Загидуллин, Т. В. Викторова 1286

Экспрессия гена β1-адренорецептора (*ADRB1*) в миокарде больных хронической сердечной недостаточностью
Э. Ф. Муслимова, Т. Ю. Реброва, Д. С. Кондратьева, Э. Л. Сондуев, Б. Н. Козлов, С. А. Афанасьев 1297

Тканевой гоносомный мозаицизм у пациентов с нарушением формирования пола, связанным с аномалиями дифференцировки гонад
Н. В. Опарина, Н. Ю. Райгородская, О. Ю. Латышев, Л. Н. Самсонова, Е. А. Володько, С. А. Суязова, В. Б. Черных 1306

Математические модели и методы

К анализу случайных процессов изонимии. II. Динамика дивергенции популяций
В. П. Пасеков 1318

Краткие сообщения

Конкордантность количественных признаков несахарного диабета и выживаемости единоутробных потомков у гибридов крыс линий *Brattleboro* и *WAG*
И. И. Хегай 1330

Идентификация патогенных CNV с помощью экзомного секвенирования при необъясненных задержках развития: исследование семейного трио
О. Ю. Наумова, П. В. Добрынин, Е. А. Гибитова, М. А. Жукова, С. Ю. Рычков, О. В. Жукова, Е. Л. Григоренко 1333

Разнообразие моногенных наследственных болезней среди русского населения Волго-Уральского региона
Р. А. Зинченко, Е. К. Гинтер, В. В. Кадышев, С. С. Амелина, Т. П. Васильева, О. Ю. Александрова, С. И. Куцев 1339

Contents

Vol. 57, No. 11, 2021

Reviews and Theoretical Articles

Genes Methylation in Cell-Free Circulating Plasma DNA as Prognostic and Predictive Factors in Breast Cancer

S. V. Rykov, E. A. Filippova, V. I. Loginov, and E. A. Braga 1225

Phenotypically Unstable Mutations as Markers of Chromosomal Rearrangements Involving DNA Transposons

L. P. Zakharenko 1241

Plant Genetics

Mitochondrial DNA Confirms the American Origin of Modern Firs

V. L. Semerikov, S. A. Semerikova, and Yu. A. Putintseva 1246

Animal Genetics

The mtDNA Control Region Variability of *Microtus rossiaemeridionalis* (Rodentia, Arvicolini) from Two Invasive Populations of the Russian Far East

I. N. Sheremetyeva, I. V. Kartavtseva, A. A. Emelyanova, and A. S. Lapin 1252

Human Genetics

Features of the Genomic Distribution of Runs of Homozygosity in the Indigenous Population of Northern Eurasia at the Individual and Population Levels Based on High Density SNP Analysis

N. A. Kolesnikov, V. N. Kharkov, A. A. Zarubin, M. O. Radzhabov, M. I. Voevoda, M. A. Gubina, E. K. Khusnutdinova, S. S. Litvinov, N. V. Ekomasova, O. V. Shtygasheva, N. R. Maksimova, A. L. Sukhomyasova, and V. A. Stepanov 1261

Bioinformatic Annotation of Genes for Alzheimer's Disease and Ischemic Heart Disease

N. Yu. Chasovskikh, E. E. Chizhik, and A. A. Bobrysheva 1276

Role of the Neurotransmitter System Genes in Chronic Obstructive Pulmonary Disease Development

G. F. Korytina, L. Z. Akhmadishina, O. V. Kochetova, T. R. Nasibullin, Yu. G. Aznabaeva, Sh. R. Zulkarneev, S. M. Izmailova, N. Sh. Zagidullin, and T. V. Victorova 1286

Expression of the β1-Adrenoreceptor Gene (*ADRB1*) in the Myocardium of Patients with Chronic Heart Failure

E. F. Muslimova, T. Yu. Rebrova, D. S. Kondratieva, E. L. Sonduev, B. N. Kozlov, and S. A. Afanasiev 1297

Inter-Tissue Gonosomal Mosaicism in Patients with Disorders Sex Development, Related to Gonadal Differentiation Abnormalities

N. V. Oparina, N. Yu. Raygorodskaya, O. Yu. Latyshev, L. N. Samsonova, E. A. Volodko, S. A. Suyazova, and V. B. Chernykh 1306

Mathematical Models and Methods

To Analysis of Random Processes of Isonymy. II. Dynamics of Population Divergence

V. P. Passekov 1318

Short Communications

Concordance of Quantitative Traits of Diabetes Insipidus and Litters Survival in Hybrids of *Brattleboro* and *WAG* Rat Lines
I. I. Khegai 1330

Identification of Pathogenic CNVs in Unexplained Developmental Disabilities Using Exome Sequencing: A Family Trio Study
O. Yu. Naumova, P. V. Dobrynin, E. A. Gibitova, M. A. Zhukova, S. Yu. Rychkov, O. V. Zhukova, and E. L. Grigorenko 1333

Diversity of Monogenic Hereditary Diseases among the Russian of the Volga-Ural Region
R. A. Zinchenko, E. K. Ginter, V. V. Kadyshev, S. S. Amelina, T. P. Vasilieva, O. Yu. Alexandrova, and S. I. Kutsev 1339

ГЕНЕТИКА, 2021, том 57, № 11, с. 1225–1240

ОБЗОРНЫЕ И ТЕОРЕТИЧЕСКИЕ СТАТЬИ

УДК 575:599.9

МЕТИЛИРОВАНИЕ ГЕНОВ В СВОБОДНОЙ ЦИРКУЛИРУЮЩЕЙ ДНК ПЛАЗМЫ КРОВИ КАК ПРОГНОСТИЧЕСКИЙ И ПРЕДИКТИВНЫЙ ФАКТОРЫ ПРИ РАКЕ МОЛОЧНОЙ ЖЕЛЕЗЫ

© 2021 г. С. В. Рыков[1, 2, *], Е. А. Филиппова[1], В. И. Логинов[1], Э. А. Брага[1, **]

[1]Научно-исследовательский институт общей патологии и патофизиологии, Москва, 125315 Россия

[2]Национальный исследовательский центр "Курчатовский институт" – ГосНИИгенетика, Курчатовский геномный центр, Москва, 117545 Россия

**e-mail: enterru@yandex.ru*

***e-mail: eleonora10_45@mail.ru*

Поступила в редакцию 13.01.2021 г.
После доработки 18.03.2021 г.
Принята к публикации 23.03.2021 г.

Несмотря на значительный прогресс в методах ранней диагностики и лечения рака молочной железы (РМЖ), сохраняется высокий шанс рецидива. Существующие методы выявления рецидивов не пригодны для раннего обнаружения, поскольку требуют инвазивных процедур и не способны оперативно сопровождать лечение. Многообещающим методом стала так называемая жидкая биопсия, основанная на анализе генетического материала из периферической крови, в частности фракции свободной циркулирующей ДНК (сцДНК) плазмы крови. Определение мутаций, копийности и метилирования отдельных генов в сцДНК позволяет отслеживать изменения как в опухолевом очаге, так и в патологическом процессе у пациента в целом, причем мало инвазивно и с хорошей точностью и оперативностью. В представленном обзоре обобщены данные литературы по прогностической и предиктивной ценности эпигенетических маркеров, основанных на анализе метилирования генов в сцДНК плазмы или сыворотки крови у больных РМЖ. Показано, что гены, гиперметилированные в сцДНК (*RASSF1A*, *RARB*, *SOX17*, *WNT5A* и др.), могут служить эффективными маркерами общей и опухоль-специфичной выживаемости и химиорезистентности. Важным преимуществом маркеров метилирования, в отличие от маркеров на основе однонуклеотидных замен, микросателлитов и вариаций числа копий, являются универсальность, раннее проявление и ясная связь с биологией патологического процесса. Анализ метилирования наиболее эффективных маркеров может решать задачи своевременного выявления метастазирования и рецидива заболевания, оперативно наблюдать за ответом на неоадъювантную и послеоперационную химиотерапию, что является основой для персонифицированного подхода к лечению больных РМЖ.

Ключевые слова: метилирование ДНК, гиперметилированные гены, свободная циркулирующая ДНК, рак молочной железы, плазма крови, предиктивный фактор, прогностический фактор.

DOI: 10.31857/S0016675821110126

Рак молочной железы (**РМЖ**) является самым распространенным злокачественным заболеванием у женщин в мире. В 2018 г. зарегистрировано более 2.1 млн новых случаев. Метастатический тип РМЖ диагностируется в зависимости от региона и этнической группы от 5–6 до 9% случаев первичного выявления [1]. РМЖ – гетерогенное заболевание, включающее различные типы опухолей, которые отличаются по морфологии, клиническому течению и чувствительности к терапии. Молекулярное типирование позволяет классифицировать опухоль, дать прогноз заболевания и персонализировать схему лечения, например определить требуемый объем хирургического вмешательства и средств терапии. В рутинной клинической практике применяется типирование по экспрессии генов рецепторов эстрогена, прогестерона и эпидермального фактора роста, которые делят РМЖ на четыре подтипа: люминальный А и Б, HER2+ позитивный и трижды негативный. Подтипы РМЖ отличаются между собой и по прогнозу течения, и ответом на медикаментозную терапию [2]. Однако детальное типирование вскрывает молекулярное разнообразие подтипов, что требует разработки и внедрения более точных и информативных групп маркеров [3].

Основной объект для молекулярного анализа – ткань опухоли, полученная с помощью биопсии. Биопсия является инвазивным методом и позволяет получить на анализ только небольшую часть

опухолевого очага, который вследствие гетерогенности злокачественного процесса не всегда отражает молекулярный портрет всей опухоли. Более доступным и полноценным объектом может служить периферическая кровь, которая содержит помимо форменных элементов циркулирующие опухолевые клетки, свободную циркулирующую ДНК и экзосомы. Так называемая "жидкая биопсия" позволяет изучить генетический материал опухолевого процесса в целом. Так, можно определить генотип опухоли и его изменения с течением болезни, что делает возможным обнаружение мутаций, которые могут быть пропущены при инвазивной биопсии [4]. Жидкая биопсия может использоваться для мониторинга ответа на терапию, оценки риска возникновения устойчивости к терапии и своевременно выявлять рецидив болезни [5, 6]. Данные молекулярного профиля опухолевого процесса, полученные анализом сцДНК, дополненные данными анализа циркулирующих опухолевых клеток, РНК и белковых маркеров, позволяют дополнять и предварять классическую биопсию и в итоге найти и оптимизировать комплексный подход к персонифицированному лечению больных РМЖ [7].

Вопросам диагностики РМЖ на основе метилирования генов в сцДНК плазмы или сыворотки периферической крови посвящен ряд обзоров [8, 9]. Однако получено также много экспериментальных данных по оценке метилирования генов в сцДНК как факторов выживаемости и химиорезистентности при РМЖ, которые не получили должного обсуждения в мировой литературе. В представленном обзоре суммированы и обобщены накопленные за последнюю декаду данные о прогностическом и предиктивном значении метилирования генов и их наборов в сцДНК плазмы или сыворотки крови у пациентов с РМЖ.

СВОБОДНАЯ ЦИРКУЛИРУЮЩАЯ ДНК КРОВИ В НОРМЕ И ПАТОЛОГИИ

Свободная циркулирующая ДНК (**сцДНК**, cell-free circulating DNA) в норме у здорового человека имеет эндогенное происхождение и появляется в кровотоке в результате гибели ядросодержащих клеточных элементов, созревания эритроцитов и тромбоцитов и в результате активной секреции нуклеиновых кислот во внеклеточное пространство [10]. Экзогенным источником сцДНК в кровотоке могут быть вирусные и бактериальные агенты при развитии инфекционных заболеваний и ДНК плода при беременности [11, 12].

Концентрация сцДНК в крови здоровых людей, измеренная различными методами, составляет: спектрофотометрией – 15–50 нг/мл, радиоиммуноанализом – 14 нг/мл, методами с флуоресцентными красителями – 7–57 нг/мл, количественной ПЦР – 16.57 нг/мл, и в норме не должна превышать 50–60 нг/мл [10]. Концентрация сцДНК является результатом динамического равновесия между скоростью поступления ДНК в кровоток и скоростью ее гидролиза сывороточными нуклеазами и выведением олигомеров из циркуляции. СцДНК выводится из кровотока в печени и почках. Среднее время полураспада сцДНК в крови составляет 10–15 мин [13]. С помощью анализа паттерна метилирования сцДНК плазмы крови здоровых доноров определено, что сцДНК в норме происходит из белых кровяных клеток (55%), предшественников эритроцитов (30%), эндотелиальных клеток сосудов (10%) и гепатоцитов (1%) [14].

СцДНК может состоять из фракций разного размера, что отражает отличие в способе выхода из клеток. Основная фракция сцДНК в норме выходит из клеток в результате апоптоза, при котором образуются фрагменты ДНК длиной 180–200 пн и кратные этой длине мультимеры. При некрозе в кровоток попадает высокомолекулярная ДНК, размеры фрагментов которой могут быть длиной более 10 тпн [15]. В кровяном потоке сцДНК представлены в основном в комплексе с нуклеосомами [16], а также могут находиться в комплексе с сывороточными белками, в составе апоптотических телец, быть адсорбированными на поверхности форменных элементов крови [17]. Концентрация сцДНК в кровотоке повышается при травматическом поражении организма [18], радиационном поражении организма [19], аутоиммунных заболеваниях [20] и при онкологических заболеваниях [21].

Свободная циркулирующая ДНК крови при канцерогенезе происходит из различных источников, включающих первичную опухоль, циркулирующие опухолевые клетки, отдаленные метастазы, а также из здоровых гемопоэтических и стромальных клеток. Быстрое накопление в крови сцДНК происходит в основном благодаря массивному выходу ДНК из клеток, подвергшихся апоптозу и некрозу. Свой вклад в увеличение концентрации сцДНК в крови также вносит снижение скорости удаления сцДНК из кровяного потока, что обусловлено множественной органной дисфункцией на фоне системного воспалительного процесса у пациентов с онкологией [22]. Концентрация сцДНК в крови варьирует у разных пациентов и зависит от типа, местоположения и стадии опухоли, в ряде случаев не выходя за пределы нормы [23, 24]. Фракция циркулирующей опухолевой ДНК (**цоДНК**), которая происходит из клеток опухоли, составляет от 0.01 до 90% всей сцДНК (рис. 1).

Многочисленные исследования продемонстрировали клиническую ценность анализа сцДНК на всех стадиях лечения: при диагностике, прогнозе, молекулярной классификации, при подборе стратегии лечения, определении ответа на терапию и при выявлении рецидива [25] (рис. 2).

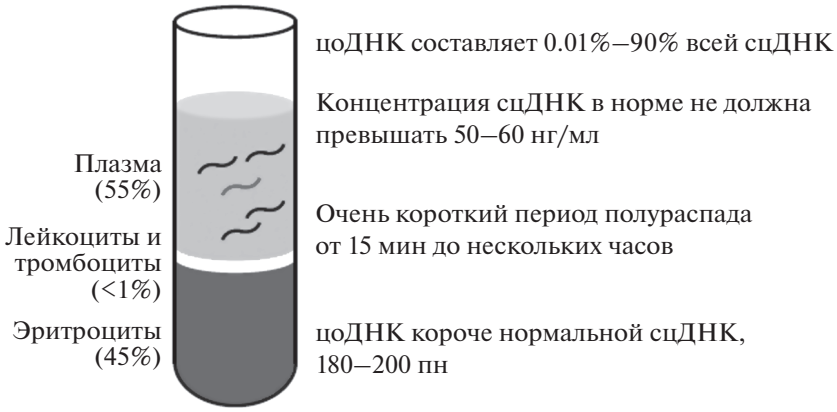

Рис. 1. Характеристика и место циркулирующей опухолевой ДНК в системе крови.

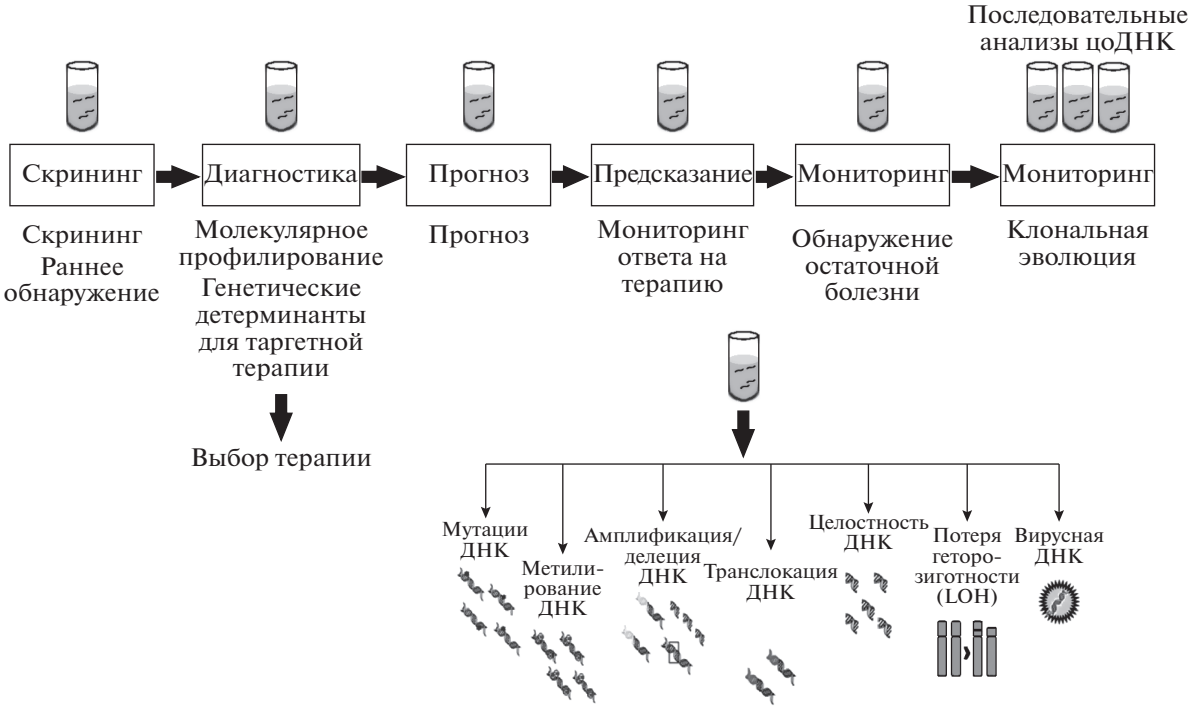

Рис. 2. Применение анализа сцДНК в диагностике и ведении онкологических заболеваний.

Паттерн метилирования сцДНК при онкопатологиях отличается от такового в норме [26]. Продемонстрирован высокий потенциал метилирования сцДНК не только при обнаружении злокачественного процесса, но и при определении места локализации первичной опухоли [27].

ЦИРКУЛИРУЮЩАЯ ДНК ПРИ РМЖ

Исследования регистрируют факт повышения концентрации сцДНК у пациентов с РМЖ [21]. Оценка концентрации сцДНК в плазме/сыворотке крови проведена различными способами. Так, с помощью ПЦР в реальном времени получены следующие значения: маркер *HBB* – 65 нг/мл (интервал 9–566) против 13 нг/мл (интервал 3–73) у здоровых доноров [28]; маркер *GAPDH* – 41600 гэ/мл (геном-эквивалентов/мл) (интервал 3644–192482) в сыворотке против 14764 гэ/мл (интервал 2198–87450) у здоровых [29]; маркер *GAPDH* – 471 нг/мл (предел отсечения при ROC-анализе) [30]; маркер *TERT* – 105.2 нг/мл против 77.06 нг/мл у здоровых [31]. Концентрация сцДНК существенно снижается после хирургического удаления опухоли – с 105.5 до 59 нг/мл. В то же время увеличение концентрации сцДНК после хирургического вмешательства является предиктивным фактором рецидива [31].

Похожую тенденцию показывает исследование с применением прямого флуоресцентного измерения концентрации сцДНК: 1010 ± 642 нг/мл – перед хирургическим вмешательством и 398 ± 162 нг/мл – после, что соответствует уровню контрольной здоровой группы – 395 ± 248 нг/мл. Уровень сцДНК коррелировал с размером опухоли, метастазами в региональных лимфоузлах и со стадией [32]. Методом флуориметрии показано, что уровень сцДНК в плазме повышается в процессе прогрессирования РМЖ и снижается после хирургического лечения. Медианная концентрация сцДНК в плазме составляет 0.5, 235, 422, 1280 и 0.5 нг/мл у пациентов с РМЖ, классифицированными по стадиям I, II, III, IV и после хирургического вмешательства соответственно [33].

Важная проблема в разработке методов диагностики РМЖ – поиск предела содержания сцДНК, после которого однозначно можно говорить о начале злокачественного процесса. Выявление высоких концентраций тотальной сцДНК не является опухоль-специфичным биомаркером. Продуктивным путем применения маркера концентрации сцДНК в онкологической практике является измерение в составе тотальной сцДНК доли цоДНК. Доля цоДНК, определенная с использованием маркеров на основе химерных генов и изменения копийности, варьировала от 1.4 до 47.9%. Выявление более 0.75% цоДНК позволило диагностировать злокачественный процесс с чувствительностью 90% и специфичностью 99% [34].

СОВРЕМЕННЫЕ МЕТОДЫ ОПРЕДЕЛЕНИЯ МЕТИЛИРОВАНИЯ ЦЕЛЕВЫХ ЛОКУСОВ В сцДНК ПЛАЗМЫ И СЫВОРОТКИ ПЕРИФЕРИЧЕСКОЙ КРОВИ ПРИ РМЖ

Маркеры на основе определения метилирования целевых локусов в сцДНК обладают характеристиками, привлекательными для практического клинического применения, – стабильностью, ранним проявлением, различием паттернов при различной локализации опухолевого очага [35–37].

Получение надежных данных с помощью маркеров метилирования требует тщательной верификации и оптимизации преаналитических процедур выделения фракции сцДНК. Ограничения на использование стандартных методов возникают в силу того, что сцДНК присутствуют в ограниченном количестве, сильно фрагментированы, и часто необходимо определять статус метилирования отдельных молекул, составляющих минорную часть пула всех внеклеточных ДНК [38]. Влияние на конечный результат может оказать не столько низкая концентрация сцДНК, сколько низкая доля цоДНК, а также загрязнение фракции свободной ДНК в результате лизиса клеточных компонентов крови. Неоптимизированная процедура обращения с кровью способствует выходу геномной ДНК из лейкоцитов, что существенно снижает возможность детекции опухолевых маркеров [39, 40].

Для определения метилирования сцДНК плазмы и сыворотки крови при РМЖ применяются как методы, основанные на бисульфитной конверсии ДНК, так и методы с использованием метил-чувствительных рестриктаз. Методы, основанные на использовании метил-чувствительных рестриктаз, в современной практике применяются редко и для повышения производительности сочетаются с высокопроизводительными методами анализа. Одним из таких методов является анализ MethDet56 (methylation detection), позволяющий исследовать с помощью чипа метилирование 56 генов [38, 41].

Методы, основанные на бисульфитной конверсии, являются золотым стандартом в эпигенетических исследованиях. Метод основан на способности гидросульфита взаимодействовать с цитозином в составе одноцепочечной ДНК с конверсией последнего в урацил; 5-метилцитозин при тех же условиях модификации не подвергается [42]. Основным недостатком метода бисульфитной конверсии является существенная деградация ДНК – до 90% от поступившего в реакцию генетического материала [43]. Требуется оптимизация стадии бисульфитной конверсии ДНК при работе с таким малым количеством генетического материала. Оптимизация проводилась по параметрам времени реакции, концентрации реагентов, температуры реакции и совмещения с процедурами очистки [44]. Еще одним способом оптимизации всей процедуры анализа метилирования в сцДНК стало исключение процедуры выделения ДНК из плазмы и сыворотки крови классическими методами (метод одношаговой МС-ПЦР (one-step MSP, OS-MSP)) [45].

Анализ бисульфит-конвертированной ДНК проводят в зависимости от поставленной задачи с помощью вариантов МС-ПЦР, таргетного или полногеномного секвенирования на NGS-секвенаторах. Метил-специфическая ПЦР основана на применении двух пар праймеров – для метилированной и неметилированной бисульфит-конвертированной ДНК [46]. Оптимизированная МС-ПЦР является высокочувствительным методом, который позволяет детектировать одиночный метилированный аллель, приходящийся на тысячу неметилированных [47]. Количественный вариант МС-ПЦР улучшает чувствительность, специфичность и воспроизводимость метода и позволяет детектировать метилированный аллель в присутствии 10000 неметилированных [48]. Разработано несколько вариантов количественного МС-ПЦР – MethyLight на основе TaqMan проб [48], метил-чувствительный анализ кривых плав-

ления с высоким разрешением (methylation-sensitive high-resolution melting analysis, MS-HRM) [49], капельная цифровая метил-специфичная ПЦР (droplet digital methylation-specific PCR, ddMSP) [50], мультиплексная количественная МС-ПЦР cMethDNA [51]. Для масштабного анализа метилирования ДНК (как геномной, так и внеклеточной) существуют два основных подхода: технологии с использованием микрочипов (microarray) и массового параллельного секвенирования. Наиболее информативным методом с полным покрытием генома является метод полногеномного бисульфитного секвенирования (Whole-Genome Bisulfite Sequencing, WGBS) [52]. Несмотря на информативность, метод WGBS не полностью отвечает критериям для практического клинического применения в силу высокой стоимости, низкого покрытия, низкой скорости получения и обработки данных [53]. Более практично применение таргетного секвенирования бисульфит-конвертированной ДНК [54, 55].

МЕТИЛИРОВАНИЕ ЦИРКУЛИРУЮЩЕЙ ДНК ПРИ РМЖ

Процесс злокачественной трансформации характеризуется глобальным гипометилированием и гиперметилированием отдельных локусов геномной ДНК. При новообразованиях молочной железы существенно меняется не только количество сцДНК, но и ее свойства, в частности метилирование CpG-динуклеотидов. Для фракции сцДНК плазмы крови наблюдается схожий эффект, что и в тканях злокачественных новообразований молочной железы, – существенное снижение глобального геномного метилирования 5метС в сравнении как с нормой, так и с доброкачественными новообразованиями. Уровень глобального гипометилирования сцДНК при РМЖ коррелирует со стадией заболевания и метастазированием в региональные лимфоузлы [56]. Схожие результаты получены при исследовании метилирования тотальной ДНК периферической крови, которая включает также ДНК форменных элементов крови [57, 58]. Однако другие исследования показывают, что уровень метилирования тотальной ДНК периферической крови, измеренный по метилированию LINE-1 регионов, может не отличаться от нормы [59]. Кроме того, обнаружено существенное различие уровня метилирования геномной ДНК лейкоцитов крови здоровых женщин и больных РМЖ; так, гипометилирование геномной ДНК лейкоцитов крови ассоциировано с риском возникновения РМЖ [60, 61].

Гиперметилирование CpG-островков промоторных регионов генов-супрессоров опухолевого роста обнаруживается с высокой частотой при РМЖ и гиперметилированные гены используются как эффективные маркеры РМЖ [62].

ГИПЕРМЕТИЛИРОВАНИЕ ОТДЕЛЬНЫХ ЛОКУСОВ В сцДНК ПЛАЗМЫ И СЫВОРОТКИ ПЕРИФЕРИЧЕСКОЙ КРОВИ КАК ПРОГНОСТИЧЕСКИЙ ФАКТОР ПРИ РМЖ

Прогностические маркеры позволяют осуществить подразделение пациентов на группы риска и использовать дифференцированные схемы лечения, что позитивно отражается на качестве жизни пациентов и эффективности терапии. В клинической практике маркеры на основе определения гиперметилирования отдельных генов, детектируемых во фракции сцДНК, могут являться независимыми прогностическими факторами общей и опухоль-специфической выживаемости при первичном РМЖ. Маркеры гиперметилирования отдельных генов могут быть также ассоциированы с основными прогностическими факторами: размером опухоли, метастазами в региональных лимфоузлах, отдаленными метастазами, экспрессией генов рецепторов гормонов, наличием циркулирующих опухолевых клеток.

Так, обнаружение метилирования *RASSF1A* и/или *APC* в плазме крови пациентов до применения химиотерапии независимо ассоциировано с негативным прогнозом общей выживаемости ($p = 0.001$) [63].

Гиперметилирование промотора гена *BRCA1* в сцДНК плазмы крови ассоциировано с высокой вероятностью рецидива, негативным прогнозом как опухоль-специфической ($p \le 0.0001$; медиана 14.2 мес.), так и общей выживаемости ($p = 0.0001$; медиана 24.3 мес.). Гиперметилирование *GSTP1* в сцДНК плазмы крови также ассоциировано с высокой вероятностью рецидива и негативным прогнозом опухоль-специфической выживаемости ($p = 0.03$, медиана 24.2 мес.) [64].

К маркерам худшего прогноза общей и опухоль-специфической выживаемости относятся гиперметилирование *PITX2* ($p = 0.001$ и $p = 0.023$ соответственно) и гиперметилирование *RASSF1A* ($p = 0.001$ и $p = 0.004$). Мультивариантный анализ выживаемости показывает, что маркеры метилирования *PITX2* и *RASSF1A* являются независимым от терапии прогностическим фактором общей выживаемости ($p = 0.021$, $p < 0.001$) [65]. Выявление метилирования минимум одного из маркеров *GSTP1*, *RASSF1A*, *RARB* в плазме крови пациентов с I и II клиническими стадиями указывает на худший прогноз общей выживаемости ($p = 0.002$).

Высокую концентрацию в плазме крови сцДНК в целом можно тоже использовать для негативного прогноза общей выживаемости ($p = 0.001$). Пациенты с высокой концентрацией сцДНК, а также с

метилированием ряда маркеров показывали худший прогноз общей выживаемости ($p = 0.001$) [66].

Метилирование *RARB* и *ESR2* наблюдали в 20 и 50% образцов сыворотки крови, что хорошо коррелировало с обнаружением в ткани опухоли. Гиперметилирование *RARb2* в сыворотке крови ассоциировано с плохим прогнозом опухоль-специфической ($p = 0.0001$; медиана – 22.1 мес.), и общей выживаемости ($p = 0.002$; медиана – 23.2 мес.) [67].

Метилирование *SOX17* обнаружено в 72.9% образцов опухолевой ткани и 58.1% сцДНК плазмы крови больных, а в ткани и плазме здоровых доноров метилирование *SOX17* не выявляется. Важно, что наблюдается существенная корреляция между обнаружением метилирования в опухолевой ткани и плазме крови того же пациента ($r_s = 0.688$, $p < 0.001$). Метилирование промотора *SOX17* в ткани и плазме ассоциировано со стадией болезни ($p < 0.001$ и $p = 0.001$ соответственно) и метастазами в лимфоузлах ($p < 0.001$ и $p = 0.001$ соответственно). Мультивариантный анализ демонстрирует, что обнаружение метилирования *SOX17* является независимым маркером плохого прогноза и общей ($p = 0.02$) и опухоль-специфической ($p = 0.001$) выживаемости [68].

Определены гены, метилирование которых в сцДНК может выступать эффективным прогностическим фактором общей и опухоль-специфической выживаемости при метастатическом РМЖ. Метилирование промоторной области гена *SEPTIN9* обнаруживали в сцДНК плазмы крови у 11% (9/82) пациентов с первичным РМЖ, у 52% (26/50) пациентов с метастатическим РМЖ и не выявляли у здоровых доноров (0/51) [69]. Количественные показатели метилирования панели из шести генов-маркеров (*AKR1B1*, *HOXB4*, *RASGRF2*, *RASSF1*, *HIT1H3C*, *TM6SF1*), суммированные в кумулятивном индексе метилирования, позволяют дать прогноз метастатического РМЖ. Так, медианы выживаемости без прогрессирования заболевания и общей выживаемости существенно короче у пациентов с высоким кумулятивным индексом метилирования (2.1 и 12.3 мес. соответственно), чем у пациентов с низким индексом (5.8 и 21.7 мес. соответственно) [70].

Выявление метилирования трех любых генов из набора маркеров (*KLK10*, *SOX17*, *WNT5A*, *MSH2*) у пациентов с метастатическим РМЖ с ранее выявленным метилированием *SOX17* или *WNT5A* ухудшает прогноз общей выживаемости ($p = 0.042$, $p = 0.043$, $p = 0.048$ соответственно). Особенно заметное снижение общей выживаемости ($p = 0.011$) и высокая вероятность летального исхода ($p = 0.017$) характерны для сочетания HER2/neu фенотипа и метилирования *SOX17*. Метилирование всех четырех генов указанного выше набора коррелирует с отсутствием ответа на химиотерапию ($p = 0.002$). В группе, получившей адъювантную терапию, обнаружение метилирования *KLK10* ассоциировано с рецидивом болезни ($p = 0.008$) и снижением интервала до выявления рецидива ($p = 0.013$) [71].

Обнаружение маркеров метилирования коррелирует с присутствием в кровотоке циркулирующих опухолевых клеток. Исследована связь между клиническими маркерами метастатического процесса: циркулирующими опухолевыми клетками и метилированием сцДНК плазмы крови. Гиперметилирование одного и более генов *APC*, *RASSF1A* и *ESR1* обнаружено в плазме 42 из 80 пациентов (53%), притом, что частота метилирования этих генов по отдельности составила 29, 35 и 20% соответственно, в то время как в контрольной группе метилирование этих генов найдено только у 3 из 20 (15%). Выявление маркеров метилирования в плазме коррелирует с обнаружением циркулирующих опухолевых клеток в крови ($p = 0.03$) пациентов с РМЖ. Данный эффект может быть объяснен как тем, что циркулирующие опухолевые клетки выступают источником метилированной сцДНК в плазме крови, так и тем, что присутствие клеток и метилированной сцДНК является фенотипическим проявлением агрессивной опухоли [72].

Установлена корреляция между обнаружением в крови циркулирующих опухолевых клеток и определением метилированных генов *APC* ($p = 0.012$) и *GSTP1* ($p = 0.001$) в сцДНК плазмы пациентов с РМЖ. Обнаружение метилированных генов *APC*, *RASSF1A* и циркулирующих опухолевых клеток коррелирует с агрессивными и прогрессирующими типами РМЖ ($p = 0.001$, $p = 0.031$). Частота встречаемости метилирования *APC* и *RASSF1A* различается между метастатическим РМЖ и РМЖ без метастазов; метилированный *GSTP1* чаще обнаруживается в плазме пациентов с большими опухолями ($p = 0.023$) и коррелирует с позитивным Her2/neu статусом ($p = 0.003$) [73].

Промотор гена *SOX17* гиперметилирован в 86% первичных опухолей. В ДНК циркулирующих опухолевых клеток *SOX17* метилирован в 34.5% ранних случаев РМЖ, 45.8% случаев метастатического РМЖ и в 1.6% здоровых доноров; в парных образцах сцДНК плазмы *SOX17* метилирован в 34.5, 40.7 и 1.6% случаев соответственно. Обнаружена существенная корреляция между уровнем метилирования *SOX17* в ДНК циркулирующих опухолевых клеток и сцДНК на ранних стадиях РМЖ, но не у пациентов с метастатическим РМЖ. Эти данные могут указывать на роль циркулирующих опухолевых клеток в формировании пула сцДНК плазмы периферической крови [74].

Эпигенетические механизмы вовлечены в подавление экспрессии гена эстрогенового рецептора при РМЖ, который является важным маркером

прогноза выживаемости и предиктивным маркером ответа на гормональную терапию тамоксифеном и ингибиторами ароматазы. На выборке из 100 пациентов с РМЖ установлена корреляция между гиперметилированием промоторов *ESR1* и *PGR* в плазме крови с ERα-негативным типом РМЖ ($p = 0.003$ и $p = 0.001$). Обнаружено также значительное гиперметилирование промоторов *ESR1* и *PGR* в сцДНК плазмы при тройном негативном РМЖ ($p = 0.03$ и $p = 0.05$) [75].

Выявлена значимая ассоциация между метилированием промотора гена *ESR1* в сцДНК плазмы крови и ER-статусом опухоли ($p = 0.0179$). Более высокие значения уровня метилирования *ESR1* показывают подтипы РМЖ с плохим прогнозом – тройной негативный и Her2+ (80 и 60% соответственно), чем подтипы с лучшим прогнозом – люминальный А и Б (28 и 5.9% соответственно) [76].

Метилирование участков промотора гена *ESR1* – *ER3*, *ER4*, *ER5* обнаружено с частотами 65, 26.7, 61.7% в опухолевой ткани и с очень близкими частотами 57.5, 21.7, 55.8% – в сыворотке крови. Обнаружена значимая связь между метилированием *ER3*, *ER4*, *ER5* в ткани и сыворотке и отрицательным статусом рецептора эстрогена ER (0.007, 0.001, 0.003 для сыворотки соответственно) [77].

Выявлена также связь между эндокринным фенотипом опухоли и метилированием других маркеров. Обнаружена корреляция между гиперметилированием *BRCA1* в тканях опухоли и плазме крови с негативным статусом ER, PR и HER2, а также тройным негативным РМЖ на выборке из 100 пациентов с РМЖ [64]. Отсутствие метилирования *WNT5A* коррелирует с фенотипом ER+, PR+, HER2– в группе адъювантной терапии и метастатической группе ($p = 0.040$ и $p = 0.016$ соответственно). Отсутствие метилирования *KLK10* статистически значимо для группы с фенотипом HER2–, PR+, получившей адъювантную терапию, и группы в устойчивой ремиссии ($p = 0.027$ и $p = 0.004$ соответственно). Присутствие метилирования *KLK10* при тройном негативном РМЖ ассоциировано с высокой вероятностью рецидива ($p = 0.014$) [71].

Анализ метилома сцДНК при химиотерапии ER+ РМЖ эксеместаном позволяет выявить группу устойчивых к лечению опухолей. Выявлено семь регионов генома, связанных с резистентностью к эксеместану, которые характеризуются различной плотностью метилирования. ROC-анализ маркеров метилирования перед гормональной терапией выявил, что высокая плотность метилирования Chr10 (46400000–46599999, *PTPN20*; AUC = 0.909, $p = 0.019$), Chr3 (67800000–67999999; *SUCLG2-AS1*; AUC = 0.932, $p = 0.013$), Chr3 (140200000–140399999; *CLSTN2*; *CLSTN2-AS1*; *TRIM42*; AUC = 0.864, $p = 0.037$) и Chr12 (101200000–101399999; *ANO4*; AUC = 0.886, $p = 0.026$) предсказывает устойчивую к терапии группу опухолей [78]. Данные по прогностическим факторам РМЖ сведены в табл. 1.

ГИПЕРМЕТИЛИРОВАНИЕ ОТДЕЛЬНЫХ ЛОКУСОВ В сцДНК ПЛАЗМЫ И СЫВОРОТКИ ПЕРИФЕРИЧЕСКОЙ КРОВИ КАК ПРЕДИКТИВНЫЙ ФАКТОР ПРИ РМЖ

В клинической практике маркеры на основе определения гиперметилирования отдельных генов, детектируемых во фракции сцДНК, могут применяться при контроле эффективности ответа на адъювантную и неоадъювантную химиотерапию и терапию метастатического РМЖ.

Неоадъювантная химиотерапия (НАХТ) проводится с целью уменьшения объема первичной опухоли, размеров и количества пораженных лимфатических узлов, элиминации отдаленных микрометастазов и повышения безрецидивной и общей выживаемости. Статистические оценки эффективности НАХТ разнятся в зависимости от иммуно-гистохимического типа РМЖ: для трижды негативного от 30 до 60% [79], для ER+ опухолей от 2–10% [80] до 69% [81], для HER2+ от 30–70% [82]. Выявление ответа на химиотерапию является важной клинической задачей.

Исследована динамика метилирования гена *RASSF1A* в плазме крови у группы из 21 пациента, прошедших НАХТ [83]. В четырех случаях, в которых достигнут полноценный ответ на терапию, метилирование *RASSF1A* в плазме снижается и не детектируется с 12-й нед. Напротив, у 17 пациентов с частичным или минимальным ответом на терапию метилирование *RASSF1A* в плазме детектируется дольше, либо вовсе сохраняется ($p = 0.02$).

Выявлено различие между реагирующей и резистентной группами до применения терапии по сниженной частоте метилирования маркеров *BRCA1* и *GSTP1* ($p = 0.02$ и $p = 0.003$ соответственно) у группы прошедших НАХТ пациентов с РМЖ [84]. Проанализирован уровень метилирования сцДНК плазмы крови пациентов в шести циклах химиотерапии по пяти маркерам: *BRCA1*, *MGMT*, *GSTP1*, *SFN* и *MDR1*. Выявлено, что метилирование *BRCA1* существенно коррелирует с ответом на терапию и различается между реагирующей и резистентной группами. Корреляция отмечена также между метилированием *BRCA1* и негативным статусом рецепторов гормонов.

Метилирование промотора *RASSF1A* в сцДНК плазмы крови обнаружено в 23% (20/87) образцов до НАХТ, что существенно выше, чем показатели обнаружения антигенов CEA (8.6%) и CA15-3 (7.4%) [85]. Отмечено снижение метилирования *RASSF1A* на выборке в целом, но степень измене-

Таблица 1. Прогностические маркеры на основе определения гиперметилирования генов во фракции сцДНК плазмы крови при РМЖ

Маркеры	Тип образца	Метод	Выборка пациентов и доноров	Клиническое значение	Ссылка
APC, RASSF1A, ESR1	Плазма и сыворотка	Количественная МС-ПЦР	80 пациентов, 20 здоровых доноров	Корреляция с обнаружением циркулирующих опухолевых клеток и прогноз общей и опухоль-специфической выживаемости	[72]
APC, GSTP1, RASSF1	Сыворотка	MethyLight	85 пациентов, 22 здоровых донора	То же	[73]
SOX17	Плазма	Количественная МС-ПЦР	114 пациентов с первичным и метастатическим РМЖ, 60 здоровых доноров	»	[74]
ESR1, PGR	Сыворотка	МС-ПЦР	100 пациентов	Корреляция с эндокринным фенотипом опухоли	[75]
ESR1	Плазма	Количественная МС-ПЦР	110 пациентов	То же	[76]
ESR1	Сыворотка	МС-ПЦР	120 пациентов	»	[77]
BRCA1, GSTP1	Сыворотка	МС-ПЦР	100 пациентов	»	[64]
SEPTIN9	Плазма	Количественный МСП-ПЦР	132 пациента, 51 здоровый донор	Прогноз общей и опухоль-специфической выживаемости	[69]
RASSF1A, APC	Сыворотка	MethyLight	86 пациентов, 10 здоровых доноров	То же	[63]
PITX2, RASSF1A	Плазма и сыворотка	MethyLight	428 пациентов	»	[65]
GSTP1, RASSF1A, RARB	Сыворотка	Одношаговая МС-ПЦР (OS-MSP)	336 пациентов	»	[66]
ESR2, RARB	Сыворотка	МС-ПЦР	100 пациентов, 30 здоровых доноров	»	[67]
SOX17	Плазма	МС-ПЦР	155 пациентов, 60 здоровых доноров	»	[68]
AKR1B1, *HOXB4*, *RASGRF2*, *RASSF1*, *HIT1H3C*, *TM6SF1*	Сыворотка	cMethDNA	141 пациент	»	[70]
KLK10, SOX17, WNT5A, MSH2	Плазма	Количественная МС-ПЦР	200 пациентов, 35 здоровых доноров	Прогноз общей и опухоль-специфической выживаемости, прогноз резистентности к химиотерапии	[71]
Chr3 (67800000–67999999), Chr3 (140200000–140399999;), Chr12 (101200000–101399999), Chr6 (32400000–32599999)	Плазма	Полногеномное секвенирование бисульфит-конвертированной ДНК	16 пациентов	Прогноз резистентности к химиотерапии	[78]

ния уровня метилирования зависит от ответа на терапию: наибольшее снижение наблюдается в группе пациентов с положительным откликом ($p = 0.006$) и в меньшей степени у резистентных пациентов. Сохранение метилирования после терапии коррелирует со степенью остаточной активности опухолевого процесса ($p = 0.008$). Прогностическое значение имеет сравнение уровня метилирования *RASSF1A* в сцДНК плазмы крови после года с момента хирургического лечения и после НАХТ. В группе пациентов, отрицательной по метилированию *RASSF1A* (37 пациентов), и в группе, в которой метилирования *RASSF1A* не обнаружено спустя год после лечения (10 пациентов), наблюдается устойчивая ремиссия. В то же время среди пациентов с метилированием *RASSF1A* до лечения (шесть пациентов) и с метилированием до лечения и спустя год после лечения (один пациент) – три пациента имели рецидив с метастазами.

На биопсийном материале опухолей РМЖ выявлены дифференциально метилированные регионы геномной ДНК, которые различаются между группой, отвечающей на НАХТ, и группой резистентной к терапии [86]. Проанализировано состояние метилирования промоторных областей генов *SLC9A3*, *C1QL2*, *DPYS*, *IRF4*, *ADCY8*, *KCNQ2*, *TERT*, *SYNDIG1*, *SKOR2* и *GRIK1* в материале опухоли и плазме крови для их использования в качестве предиктивных маркеров ответа на НАХТ у больных РМЖ. Наилучшие результаты показала диагностическая панель, состоящая всего из двух маркеров *IRF4* и *C1QL2*, имеющая диагностическую точность 0.75 с чувствительностью 75% и специфичностью 75% [86]. В другом исследовании на выборке из 36 пациентов с РМЖ выявлено статистически значимое различие в частотах метилирования промотора гена *ADCY8* в плазме крови между группами с различной степенью ответа на терапию – метилирование чаще наблюдалось в группе с плохим ответом на лечение [87].

Исследован способ определения и количественной оценки фракции сцДНК, происходящей из ткани РМЖ и нормальной ткани молочной железы, с помощью маркеров метилирования [88]. Использование маркеров на основе однонуклеотидных замен показывает низкие значения чувствительности и специфичности. Значение концентрации тотальной сцДНК плазмы крови не обладает способностью различать пациентов с РМЖ и здоровых доноров. Поэтому были использованы маркеры метилирования *KRT19*, *LMX1B*, *ZNF296*, специфические только для ткани молочной железы, что позволило получить высокие значения диагностического теста (чувствительность 80% и специфичность 97%). На 29 пациентах, которые получали НАХТ перед операцией, проследили динамику изменения концентрации цоДНК, специфичной для ткани молочной железы. Через 2–4 нед. после начала терапии наблюдается резкое уменьшение концентрации специфической сцДНК, которая достигает низшего значения сразу после операции, не отличаясь от здоровых показателей. Напротив, пациенты, у которых ответа на терапию не наблюдалось, перед операцией демонстрировали повышенный уровень специфической сцДНК. Уровень специфической сцДНК, измеренный в последний период неоадъювантной терапии, является предиктивным фактором остаточного заболевания. Уровень специфической сцДНК существенно ниже, по сравнению со значением перед началом терапии, у пациентов с полноценным ответом [88].

Адъювантная терапия с использованием гормональных препаратов, применяемая после хирургического удаления опухоли, способна увеличивать как безрецидивную, так и общую выживаемость больных. При гормонотерапии РМЖ применяют ингибиторы ароматазы: нестероидные (летрозол (Фемара), анастрозол (Аримидекс)) и стероидные (эксеместан), а также конкурентный ингибитор эстрогеновых рецепторов – тамоксифен.

Статус метилирования *RASSF1A* в сцДНК 148 пациентов был проанализирован перед хирургическим лечением и спустя один год после операции с дополнительной адъювантной терапией тамоксифеном [89]. Было обнаружено, что метилирование *RASSF1A* является независимым предиктивным маркером негативного прогноза (относительный риск рецидива и смерти составил 5.1 и 6.9 соответственно). Также подтверждено, что измерение уровня метилирования *RASSF1A* в сцДНК позволяет отслеживать эффективность терапии. Более поздние исследования подтвердили, что гиперметилирование *RASSF1A* снижает терапевтический эффект доцетаксела (Таутакс) при терапии РМЖ [90] и что его уровни положительно коррелируют с экспрессией рецепторов эстрогена при РМЖ [91].

Исследована динамика изменения метилирования промоторов генов *RARB*, *MSH2*, *MGMT*, *PGR* и *MDGI* в сцДНК плазмы периферической крови при хирургическом лечении с последующей терапией тамоксифеном на выборке из 20 пациентов ER+ РМЖ с использованием метода MethDet-56 [92]. Профиль метилирования сцДНК плазмы крови изменился: сразу после хирургического лечения выявлено снижение метилирования промоторов *RARB, MSH2* и *MGMT*; после применения тамоксифена возросло метилирование промотора *ESR1*; после оперативного лечения в комбинации с тамоксифеном выявлено снижение метилирования промоторов *PGR*, *MDGI*, *MGMT* и *RARB*.

Более 30% гормон-зависимых случаев РМЖ невосприимчивы к первичной гормонотерапии и около 40% из первоначально восприимчивых случаев приобретают устойчивость. Анализ метилома сцДНК при химиотерапии экземестаном

больных РМЖ с ER+ позволяет определить группу устойчивых к лечению и проконтролировать эффективность лечения на опухолях, чувствительных к терапии [78]. Выявлено семь регионов генома, связанных с резистентностью к эксеместану, которые характеризуются существенно различной плотностью метилирования. Кроме первичной устойчивости, изменения в метилировании генома обнаружены для приобретенной устойчивости. Изменение в плотности метилирования Chr6 (32400000–32599999; *LA-DRA; HLADRB5; HLADRB6; HLA-DRB1*) является предиктором приобретенной к эксеместану резистентности (AUC = 0.900, $p = 0.024$) и ассоциировано с коротким периодом до прогрессии заболевания (3.33 против 26.31 мес., $p = 0.000033$).

Метилирование маркера *ESR1*, обнаруженное в ДНК циркулирующих опухолевых клеток пациентов после химиотерапии препаратами эверолимус (Афинитор) или эксеместан, ассоциировано с отсутствием ответа на лечение ($p = 0.023$) [93].

Основным методом лечения метастатической формы РМЖ является химиотерапия. Единого стандарта лечения метастатического РМЖ не существует, что отражает комплексность заболевания и насущную необходимость в определении быстрых и надежных маркеров ответа на терапию.

Исследована связь между уровнем метилирования промоторов генов *ESR1* и *14-3-3-*σ в сцДНК плазмы крови и клиническими проявлениями у пациентов с метастазирующим РМЖ [94]. Уровень метилирования промотора *14-3-3-*σ существенно различался между группой метастазирующего РМЖ и контрольной здоровой группой ($p < 0.001$), а также между группой в ремиссии и группой с метастазами ($p < 0.001$). При проведении химиотерапии метастазирующего РМЖ уровень метилирования снижался у большинства пациентов, однако обособлялась группа, у которой уровень оставался неизменным. ROC-анализ позволяет различить группу не отвечающих на терапию пациентов с помощью величины отношения уровня метилирования перед первым циклом химиотерапии к уровню метилирования перед вторым циклом с чувствительностью 75% и специфичностью 66.7% при отсекающем значении в 2.39 [94].

Значимое снижение суммарного индекса метилирования генов-маркеров выявлялось у пациентов с ответом на терапию или без прогрессирования заболевания ($p = 0.010$). Выборка состояла из 29 пациентов с метастатическим РМЖ, у которых взяли образцы сыворотки перед терапией доцетакселом или капецитабином и по прошествии 18–49 дней. Данный эффект снижения индекса метилирования не наблюдался для пациентов с прогрессирующим течением заболевания. При прослеживании группы из 13 пациентов в процессе терапии у 10 наблюдали динамический характер изменения индекса метилирования в зависимости от течения болезни: снижение метилирования при первоначальном ответе на терапию и увеличение метилирования при последующем прогрессировании заболевания [51].

Методами массового параллельного секвенирования исследованы метиломы сцДНК плазмы крови пациентов с метастатическим РМЖ (отдаленные метастазы) после терапии, пациентов в устойчивой ремиссии и здоровых доноров [53]. Данные указали на различие в метилировании около 5.0×10^6 CpG-динуклеотидов между группой метастатического РМЖ и группами в ремиссии и здоровыми донорами. Между группой в ремиссии и здоровыми донорами наблюдается существенное сходство. Метилом метастатического РМЖ характеризуется глобальным гипометилированием и гиперметилированием отдельных локусов. Анализ выявил 21 новых гиперметилированных при метастатическом РМЖ CpG-островков в локусах: *BEND4*, *CDH4*, *C1QL3*, *ERG*, *GP5*, *GSC*, *HTR1B*, *LMX1B*, *MCF2L2*, *PAX5*, *PCDH10*, *PENK*, *REC8*, *RUNX3*, *SP8*, *SP9*, *STAC2*, *ULBP1*, *UNC13A*, *VIM*, *VWC2*. Целенаправленное секвенирование локусов *GP5*, *PCDH10*, *HRR1B* и *UNC13A* показало хорошее соответствие с данными полногеномного секвенирования и продемонстрировало статистически значимое различие в величине метилирования между метастатическим РМЖ и группой в ремиссии и здоровых доноров. Данные по предиктивным маркерам РМЖ сведены в табл. 2.

ЗАКЛЮЧЕНИЕ

Маркеры на основе метилирования отдельных генов сцДНК показывают значительный прогностический и предиктивный потенциал при РМЖ. Анализ сцДНК плазмы крови позволяет предсказать эффективность и проследить ответ на адъювантную и неоадъювантную химиотерапию. Важным преимуществом маркеров метилирования, в отличие от маркеров на основе однонуклеотидных замен, микросателлитов и вариаций числа копий, является универсальность, раннее проявление и ясная связь с биологией патологического процесса.

Вместе с обнадеживающими результатами, продемонстрированными на клиническом материале, необходимо отметить ряд объективных технических сложностей, возникающих на практике. Результаты анализа метилирования сцДНК зависят от преаналитической стадии выделения внеклеточной ДНК, которая может осуществляться различными способами и коммерческими наборами. Устоявшихся стандартов в данной области не существует. Исследователями сообщается о различной степени корреляции данных, полученных на биопсийном опухолевом материале

Таблица 2. Предиктивные маркеры на основе определения гиперметилирования генов во фракции сцДНК плазмы крови при РМЖ

Маркеры	Тип образца	Метод	Выборка пациентов и доноров	Клиническое значение	Ссылка
Chr3 (67800000–67999999), Chr3 (140200000–140399999), Chr12 (101200000–101399999), Chr6 (32400000–32599999)	Плазма	Полногеномное секвенирование бисульфит-конвертированной ДНК	16 пациентов	Контроль ответа на адъювантную гормональную химиотерапию РМЖ	[78]
RASSF1A	Сыворотка	MethyLight	148 пациентов	То же	[89]
RARB, MSH2, MGMT, PGR, MDGI	Плазма	MethDet-56	20 пациентов ER+, 20 здоровых доноров	»	[92]
ESR1	Плазма	Количественная МС-ПЦР	58 пациентов, 108 пациентов, 30 здоровых доноров	»	[93]
RASSF1A	Сыворотка	МС-ПЦР с анализом по метил-чувствительным кривым плавления (MS-HRM)	52 пациента	Контроль ответа на неоадъювантную химиотерапию РМЖ	[83]
BRCA1	Сыворотка	МС-ПЦР	30 пациентов, 30 здоровых доноров	То же	[84]
RASSF1A	Плазма	Одношаговая МС-ПЦР (OS-MSP)	87 пациентов	»	[85]
ADCY8	Плазма	МС-ПЦР	36 пациентов	»	[87]
KRT19, LMX1B, ZNF296	Плазма	Высокопроизводительное таргетное секвенирование ампликонов конвертированной ДНК	29 пациентов, 64 здоровых донора	»	[88]
AKR1B1, *HOXB4*, *RASGRF2*, *RASSF1*, *HIT1H3C*, *TM6SF1*	Сыворотка	cMethDNA	141 пациент	Контроль ответа на химиотерапию метастатического РМЖ	[70]
*14-3-3-*σ	Сыворотка	Количественная МС-ПЦР	77 пациентов без обнаруженных метастазов, 34 пациента с метастазами, 34 здоровых доноров	То же	[94]
AKR1B1, COL6A2, GPX7, HIST1H3, HOXB4, RASGRF2, TM6SF1, ARHGEF7, TMEFF2, RASSF1	Сыворотка	cMethDNA	29 пациентов	»	[51]
GP5, PCDH10, HRR1B, UNC13A	Плазма	Полногеномное секвенирование бисульфит-конвертированной ДНК, высокопроизводительное таргетное секвенирование ампликонов конвертированной ДНК	40 пациентов, 40 здоровых доноров, 40 пациентов в ремиссии	»	[53]

и материале плазмы крови. Это может быть как техническим следствием стадии анализа и интерпретирования, которая требует процедур калибровки и нормировки, так и следствием достижения предела обнаружения применяемых методов. Устоявшихся стандартов в данной области также не существует. Концептуальный характер различия данных биопсийного материала и данных сцДНК продемонстрирован для маркеров на основе определения мутаций геномной ДНК. Все это требует дальнейшей работы для внесения ясности в техническую сторону вопроса и широкого внедрения анализа метилирования сцДНК в клиническую практику.

Работа выполнена в рамках Государственного задания Минобрнауки РФ при поддержке ПНИ 0403-2020-0002 на 2020–2022 годы.

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

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

СПИСОК ЛИТЕРАТУРЫ

1. *Harbeck N., Penault-Llorca F., Cortes J. et al.* Breast cancer // Nat. Rev. Dis. Primers. 2019. V. 5. № 1. e66. https://doi.org/10.1038/s41572-019-0111-2
2. *Sørlie T., Perou C.M., Tibshirani R. et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications // Proc. Natl Acad. Sci. USA. 2001. V. 98. № 19. P. 10869–10874. https://doi.org/10.1073/pnas.191367098
3. *Lehmann B.D., Jovanović B., Chen X. et al.* Refinement of triple-negative breast cancer molecular subtypes: Implications for neoadjuvant chemotherapy selection // PLoS One. 2016. V. 11. № 6. e0157368. https://doi.org/10.1371/journal.pone.0157368
4. *Dawson S.-J., Tsui D.W.Y., Murtaza M. et al.* Analysis of circulating tumor DNA to monitor metastatic breast cancer // N. Engl. J. Med. 2013. V. 368. № 13. P. 1199–1209. https://doi.org/10.1056/NEJMoa1213261
5. *Siravegna G., Marsoni S., Siena S. et al.* Integrating liquid biopsies into the management of cancer // Nat. Rev. Clin. Oncol. 2017. V. 14. № 9. P. 531–548. https://doi.org/10.1038/nrclinonc.2017.14
6. *Nassar F.J., Chamandi G., Tfaily M.A. et al.* Peripheral blood-based biopsy for breast cancer risk prediction and early detection // Front Med. (Lausanne). 2020. V. 7. e28. https://doi.org/10.3389/fmed.2020.00028
7. *Wan J.C.M., Massie C., Garcia-Corbacho J. et al.* Liquid biopsies come of age: Towards implementation of circulating tumour DNA // Nat. Rev. Cancer. 2017. V. 17. № 4. P. 223–238. https://doi.org/10.1038/nrc.2017.7
8. *Tzanikou E., Lianidou E.* The potential of ctDNA analysis in breast cancer // Crit. Rev. Clin. Lab. Sci. 2020. V. 57. № 1. P. 54–72. https://doi.org/10.1080/10408363.2019.1670615
9. *Constâncio V., Nunes S.P., Henrique R. et al.* DNA Methylation-based testing in liquid biopsies as detection and prognostic biomarkers for the four major cancer types // Cells. 2020. V. 9. № 3. https://doi.org/10.3390/cells9030624
10. *Тамкович С.Н., Власов В.В., Лактионов П.П.* Циркулирующие ДНК крови и их использование в медицинской диагностике // Мол. биология. 2008. Т. 42. № 1. С. 12–23.
11. *Rafi I., Chitty L.* Cell-free fetal DNA and non-invasive prenatal diagnosis // Br. J. Gen. Pract. 2009. V. 59. № 562. e146–e148. https://doi.org/10.3399/bjgp09X420572
12. *Han D., Li R., Shi J. et al.* Liquid biopsy for infectious diseases: A focus on microbial cell-free DNA sequencing // Theranostics. 2020. V. 10. № 12. P. 5501–5513. https://doi.org/10.7150/thno.45554
13. *Schwarzenbach H., Pantel K.* Circulating DNA as biomarker in breast cancer // Breast Cancer Res. 2015. V. 17. № 1. e136. https://doi.org/10.1186/s13058-015-0645-5
14. *Moss J., Magenheim J., Neiman D. et al.* Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease // Nat. Commun. 2018. V. 9. № 1. P. 5068. https://doi.org/10.1038/s41467-018-07466-6
15. *Jahr S., Hentze H., Englisch S. et al.* DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells // Cancer Res. 2001. V. 61. № 4. P. 1659–1665. PMID: 11245480
16. *Roth C., Pantel K., Müller V. et al.* Apoptosis-related deregulation of proteolytic activities and high serum levels of circulating nucleosomes and DNA in blood correlate with breast cancer progression // BMC Cancer. 2011. V. 11. P. 4. https://doi.org/10.1186/1471-2407-11-4
17. *Rykova E.Y., Morozkin E.S., Ponomaryova A.A. et al.* Cell-free and cell-bound circulating nucleic acid complexes: mechanisms of generation, concentration and content // Expert Opin. Biol. Ther. 2012. V. 12. Suppl 1. P. S141–S153. https://doi.org/10.1517/14712598.2012.673577
18. *Rodrigues Filho E.M., Ikuta N., Simon D. et al.* Prognostic value of circulating DNA levels in critically ill and trauma patients // Rev. Bras. Ter. Intensiva. 2014. V. 26. № 3. P. 305–312. https://doi.org/10.5935/0103-507X.20140043
19. *Vasilyeva I.N.* Low-molecular-weight DNA in blood plasma as an index of the influence of ionizing radiation // Ann. N.Y. Acad. Sci. 2001. V. 945. P. 221–228. https://doi.org/10.1111/j.1749-6632.2001.tb03889.x
20. *Duvvuri B., Lood C.* Cell-free DNA as a biomarker in autoimmune rheumatic diseases // Front Immunol. 2019. V. 10. P. 502. https://doi.org/10.3389/fimmu.2019.00502
21. *Rohanizadegan M.* Analysis of circulating tumor DNA in breast cancer as a diagnostic and prognostic biomarker // Cancer Genet. 2018. V. 228-229. P. 159–168. https://doi.org/10.1016/j.cancergen.2018.02.002
22. *Delmonico L., Alves G., Bines J.* Cell free DNA biology and its involvement in breast carcinogenesis // Adv.

Clin. Chem. 2020. V. 97. P. 171–223.
https://doi.org/10.1016/bs.acc.2019.12.006

23. *Bettegowda C., Sausen M., Leary R.J. et al.* Detection of circulating tumor DNA in early- and late-stage human malignancies // Sci. Transl. Med. 2014. V. 6. № 224. e224ra24.
https://doi.org/10.1126/scitranslmed.3007094

24. *Diehl F., Schmidt K., Choti M.A. et al.* Circulating mutant DNA to assess tumor dynamics // Nat. Med. 2007. V. 14. № 9. P. 985–990.
https://doi.org/10.1038/nm.1789

25. *Stewart C.M., Kothari P.D., Mouliere F. et al.* The value of cell-free DNA for molecular pathology // J. Pathol. 2018. V. 244. № 5. P. 616–627.
https://doi.org/10.1002/path.5048

26. *Li L., Choi J.-Y., Lee K.-M. et al.* DNA methylation in peripheral blood: A potential biomarker for cancer molecular epidemiology // J. Epidemiol. 2012. V. 22. № 5. P. 384–394.
https://doi.org/10.2188/jea.JE20120003

27. *Kang S., Li Q., Chen Q. et al.* CancerLocator: Non-invasive cancer diagnosis and tissue-of-origin prediction using methylation profiles of cell-free DNA // Genome Biol. 2017. V. 18. № 1. P. 53.
https://doi.org/10.1186/s13059-017-1191-5

28. *Huang Z.H., Li L.H., Hua D.* Quantitative analysis of plasma circulating DNA at diagnosis and during follow-up of breast cancer patients // Cancer Lett. 2006. V. 243. № 1. P. 64–70.
https://doi.org/10.1016/j.canlet.2005.11.027

29. *Zanetti-Dällenbach R., Wight E., Fan A.X.-C. et al.* Positive correlation of cell-free DNA in plasma/serum in patients with malignant and benign breast disease // Anticancer Res. 2008. V. 28. 2A. P. 921–925.

30. *Gong B., Xue J., Yu J. et al.* Cell-free DNA in blood is a potential diagnostic biomarker of breast cancer // Oncol. Lett. 2012. V. 3. № 4. P. 897–900.
https://doi.org/10.3892/ol.2012.576

31. *Catarino R., Ferreira M.M., Rodrigues H. et al.* Quantification of free circulating tumor DNA as a diagnostic marker for breast cancer // DNA Cell Biol. 2008. V. 27. № 8. P. 415–421.
https://doi.org/10.1089/dna.2008.0744

32. *Agassi R., Czeiger D., Shaked G. et al.* Measurement of circulating cell-free DNA levels by a simple fluorescent test in patients with breast cancer // Am. J. Clin. Pathol. 2015. V. 143. № 1. P. 18–24.
https://doi.org/10.1309/AJCPI5YHG0OGFAHM

33. *Tangvarasittichai O., Jaiwang W., Tangvarasittichai S.* The plasma DNA concentration as a potential breast cancer screening marker // Indian J. Clin. Biochem. 2015. V. 30. № 1. P. 55–58.
https://doi.org/10.1007/s12291-013-0407-z

34. *Leary R.J., Sausen M., Kinde I. et al.* Detection of chromosomal alterations in the circulation of cancer patients with whole-genome sequencing // Sci. Transl. Med. 2012. V. 4. № 162. e162ra154.
https://doi.org/10.1126/scitranslmed.3004742

35. *Wang J., Han X., Sun Y.* DNA methylation signatures in circulating cell-free DNA as biomarkers for the early detection of cancer // Sci. China Life Sci. 2017. V. 60. № 4. P. 356–362.
https://doi.org/10.1007/s11427-016-0253-7

36. *Warton K., Samimi G.* Methylation of cell-free circulating DNA in the diagnosis of cancer // Front. Mol. Biosci. 2015. V. 2. P. 13.
https://doi.org/10.3389/fmolb.2015.00013

37. *Issa J.P.* DNA methylation as a clinical marker in oncology // J. Clin. Oncol. 2012. V. 30. № 20. P. 2566–2568.
https://doi.org/10.1200/jco.2012.42.1016

38. *Брызгунова О.Е., Лактионов П.П.* Современные методы исследования метилирования внеклеточных ДНК // Мол. биология. 2017. Т. 51. № 2. С. 195–214.
https://doi.org/10.7868/S0026898417010074

39. *Parpart-Li S., Bartlett B., Popoli M. et al.* The effect of preservative and temperature on the analysis of circulating tumor DNA // Clin. Cancer. Res. 2017. V. 23. № 10. P. 2471–2477.
https://doi.org/10.1158/1078-0432.ccr-16-1691

40. *Sherwood J.L., Corcoran C., Brown H. et al.* Optimised pre-analytical methods improve KRAS mutation detection in circulating tumour DNA (ctDNA) from patients with Non-Small Cell Lung Cancer (NSCLC) // PLoS One. 2016. V. 11. № 2. e0150197.
https://doi.org/10.1371/journal.pone.0150197

41. *Liggett T.E., Melnikov A., Yi Q. et al.* Distinctive DNA methylation patterns of cell-free plasma DNA in women with malignant ovarian tumors // Gynecol. Oncol. 2011. V. 120. № 1. P. 113–120.
https://doi.org/10.1016/j.ygyno.2010.09.019

42. *Darst R.P., Pardo C.E., Ai L. et al.* Bisulfite sequencing of DNA // Curr. Protoc. Mol. Biol. 2010. Chapter 7. Unit 7.9.1-17.
https://doi.org/10.1002/0471142727.mb0709s91

43. *Tanaka K., Okamoto A.* Degradation of DNA by bisulfite treatment // Bioorg. Med. Chem. Lett. 2007. V. 17. № 7. P. 1912–1915.
https://doi.org/10.1016/j.bmcl.2007.01.040

44. *Yi S., Long F., Cheng J. et al.* An optimized rapid bisulfite conversion method with high recovery of cell-free DNA // BMC Mol. Biol. 2017. V. 18. № 1. P. 24.
https://doi.org/10.1186/s12867-017-0101-4

45. *Yamamoto N., Nakayama T., Kajita M. et al.* Detection of aberrant promoter methylation of GSTP1, RASSF1A, and RARβ2 in serum DNA of patients with breast cancer by a newly established one-step methylation-specific PCR assay // Breast Cancer Res. Treat. 2012. V. 132. № 1. P. 165–173.
https://doi.org/10.1007/s10549-011-1575-2

46. *Herman J.G., Graff J.R., Myöhänen S. et al.* Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands // Proc. Natl Acad. Sci. USA. 1996. V. 93. № 18. P. 9821–9826.
https://doi.org/10.1073/pnas.93.18.9821

47. *Delpu Y., Cordelier P., Cho W.C. et al.* DNA methylation and cancer diagnosis // Int. J. Mol. Sci. 2013. V. 14. № 7. P. 15029–15058.
https://doi.org/10.3390/ijms140715029

48. *Eads C.A., Danenberg K.D., Kawakami K. et al.* MethyLight: A high-throughput assay to measure DNA methylation // Nucl. Acids Res. 2000. V. 28. № 8. e32. https://doi.org/10.1093/nar/28.8.e32

49. *Wojdacz T.K., Dobrovic A., Hansen L.L.* Methylation-sensitive high-resolution melting // Nat. Protoc. 2008. V. 3. № 12. P. 1903–1908. https://doi.org/10.1038/nprot.2008.191

50. *Yu M., Heinzerling T.J., Grady W.M.* DNA methylation analysis using droplet digital PCR // Methods Mol. Biol. 2018. V. 1768. P. 363–383. https://doi.org/10.1007/978-1-4939-7778-9_21

51. *Fackler M.J., Lopez Bujanda Z., Umbricht C. et al.* Novel methylated biomarkers and a robust assay to detect circulating tumor DNA in metastatic breast cancer // Cancer Res. 2014. V. 74. № 8. P. 2160–2170. https://doi.org/10.1158/0008-5472.CAN-13-3392

52. *Beck S., Rakyan V.K.* The methylome: Approaches for global DNA methylation profiling // Trends Genet. 2008. V. 24. № 5. P. 231–237. https://doi.org/10.1016/j.tig.2008.01.006

53. *Legendre C., Gooden G.C., Johnson K. et al.* Whole-genome bisulfite sequencing of cell-free DNA identifies signature associated with metastatic breast cancer // Clin. Epigenetics. 2015. V. 7. P. 100. https://doi.org/10.1186/s13148-015-0135-8

54. *Samorodnitsky E., Datta J., Jewell B.M. et al.* Comparison of custom capture for targeted next-generation DNA sequencing // J. Mol. Diagn. 2015. V. 17. № 1. P. 64–75. https://doi.org/10.1016/j.jmoldx.2014.09.009

55. *Widschwendter M., Evans I., Jones A. et al.* Methylation patterns in serum DNA for early identification of disseminated breast cancer // Genome Med. 2017. V. 9. № 1. P. 115. https://doi.org/10.1186/s13073-017-0499-9

56. *Pasha H.A., Rezk N.A., Riad M.A.* Circulating cell free nuclear DNA, mitochondrial DNA and global DNA methylation: Potential noninvasive biomarkers for breast cancer diagnosis // Cancer Invest. 2019. V. 37. № 9. P. 432–439. https://doi.org/10.1080/07357907.2019.1663864

57. *Severi G., Southey M.C., English D.R. et al.* Epigenome-wide methylation in DNA from peripheral blood as a marker of risk for breast cancer // Breast Cancer Res. Treat. 2014. V. 148. № 3. P. 665–673. https://doi.org/10.1007/s10549-014-3209-y

58. *van Veldhoven K., Polidoro S., Baglietto L. et al.* Epigenome-wide association study reveals decreased average methylation levels years before breast cancer diagnosis // Clin. Epigenetics. 2015. V. 7. № 1. P. 67. https://doi.org/10.1186/s13148-015-0104-2

59. *Wu H.-C., Delgado-Cruzata L., Flom J.D. et al.* Repetitive element DNA methylation levels in white blood cell DNA from sisters discordant for breast cancer from the New York site of the breast cancer family registry // Carcinogenesis. 2012. V. 33. № 10. P. 1946–1952. https://doi.org/10.1093/carcin/bgs201

60. *Choi J.-Y., James S.R., Link P.A. et al.* Association between global DNA hypomethylation in leukocytes and risk of breast cancer // Carcinogenesis. 2009. V. 30. № 11. P. 1889–1897. https://doi.org/10.1093/carcin/bgp143

61. *Kuchiba A., Iwasaki M., Ono H. et al.* Global methylation levels in peripheral blood leukocyte DNA by LUMA and breast cancer: A case-control study in Japanese women // Br. J. Cancer. 2014. V. 110. № 11. P. 2765–2771. https://doi.org/10.1038/bjc.2014.223

62. *Tang Q., Cheng J., Cao X. et al.* Blood-based DNA methylation as biomarker for breast cancer: a systematic review // Clin. Epigenetics. 2016. V. 8. e115. https://doi.org/10.1186/s13148-016-0282-6

63. *Müller H.M., Widschwendter A., Fiegl H. et al.* DNA methylation in serum of breast cancer patients: an independent prognostic marker // Cancer Res. 2003. V. 63. № 22. P. 7641–7645.

64. *Sharma G., Mirza S., Parshad R. et al.* Clinical significance of promoter hypermethylation of DNA repair genes in tumor and serum DNA in invasive ductal breast carcinoma patients // Life Sci. 2010. V. 87. № 3–4. P. 83–91. https://doi.org/10.1016/j.lfs.2010.05.001

65. *Göbel G., Auer D., Gaugg I. et al.* Prognostic significance of methylated RASSF1A and PITX2 genes in blood- and bone marrow plasma of breast cancer patients // Breast Cancer Res. Treat. 2011. V. 130. № 1. P. 109–117. https://doi.org/10.1007/s10549-010-1335-8

66. *Fujita N., Nakayama T., Yamamoto N. et al.* Methylated DNA and total DNA in serum detected by one-step methylation-specific PCR is predictive of poor prognosis for breast cancer patients // Oncology. 2012. V. 83. № 5. P. 273–282. https://doi.org/10.1159/000342083

67. *Mirza S., Sharma G., Parshad R. et al.* Clinical significance of promoter hypermethylation of ERβ and RARβ2 in tumor and serum DNA in Indian breast cancer patients // Ann. Surg. Oncol. 2012. V. 19. № 9. P. 3107–3115. https://doi.org/10.1245/s10434-012-2323-5

68. *Fu D., Ren C., Tan H. et al.* Sox17 promoter methylation in plasma DNA is associated with poor survival and can be used as a prognostic factor in breast cancer // Medicine (Baltimore). 2015. V. 94. № 11. e637. https://doi.org/10.1097/MD.0000000000000637

69. *Matsui S., Kagara N., Mishima C. et al.* Methylation of the SEPT9_v2 promoter as a novel marker for the detection of circulating tumor DNA in breast cancer patients // Oncol. Rep. 2016. V. 36. № 4. P. 2225–2235. https://doi.org/10.3892/or.2016.5004

70. *Visvanathan K., Fackler M.S., Zhang Z. et al.* Monitoring of serum DNA methylation as an early independent marker of response and survival in metastatic breast cancer: TBCRC 005 prospective biomarker study // J. Clin. Oncol. 2017. V. 35. № 7. P. 751–758. https://doi.org/10.1200/JCO.2015.66.2080

71. *Panagopoulou M., Karaglani M., Balgkouranidou I. et al.* Circulating cell-free DNA in breast cancer: size profiling, levels, and methylation patterns lead to prognostic and predictive classifiers // Oncogene. 2019. V. 38. № 18. P. 3387–3401. https://doi.org/10.1038/s41388-018-0660-y

72. *van der Auwera I., Elst H.J., van Laere S.J. et al.* The presence of circulating total DNA and methylated genes is associated with circulating tumour cells in blood from breast cancer patients // Br. J. Cancer. 2009. V. 100. № 8. P. 1277–1286. https://doi.org/10.1038/sj.bjc.6605013

73. *Matuschek C., Bölke E., Lammering G. et al.* Methylated APC and GSTP1 genes in serum DNA correlate with the presence of circulating blood tumor cells and are associated with a more aggressive and advanced breast cancer disease // Eur. J. Med. Res. 2010. V. 15. № 7. P. 277–286. https://doi.org/10.1186/2047-783x-15-7-277

74. *Chimonidou M., Strati A., Malamos N. et al.* SOX17 promoter methylation in circulating tumor cells and matched cell-free DNA isolated from plasma of patients with breast cancer // Clin. Chem. 2013. V. 59. № 1. P. 270–279. https://doi.org/10.1373/clinchem.2012.191551

75. *Mirza S., Sharma G., Parshad R. et al.* Clinical significance of Stratifin, ERalpha and PR promoter methylation in tumor and serum DNA in Indian breast cancer patients // Clin. Biochem. 2010. V. 43. № 4–5. P. 380–386. https://doi.org/10.1016/j.clinbiochem.2009.11.016

76. *Martínez-Galán J., Torres-Torres B., Núñez M.I. et al.* ESR1 gene promoter region methylation in free circulating DNA and its correlation with estrogen receptor protein expression in tumor tissue in breast cancer patients // BMC Cancer. 2014. V. 14. P. 59. https://doi.org/10.1186/1471-2407-14-59

77. *Hagrass H.A., Pasha H.F., Ali A.M.* Estrogen receptor alpha (ERα) promoter methylation status in tumor and serum DNA in Egyptian breast cancer patients // Gene. 2014. V. 552. № 1. P. 81–86. https://doi.org/10.1016/j.gene.2014.09.016

78. *Liu X.-R., Zhang R.-Y., Gong H. et al.* Methylome variation predicts exemestane resistance in advanced ER+ breast cancer // Technol. Cancer Res. Treat. 2020. V. 19. e1533033819896331. https://doi.org/10.1177/1533033819896331

79. *Garufi G., Palazzo A., Paris I. et al.* Neoadjuvant therapy for triple-negative breast cancer: Potential predictive biomarkers of activity and efficacy of platinum chemotherapy, PARP- and immune-checkpoint-inhibitors // Expert Opin. Pharmacother. 2020. V. 21. № 6. P. 687–699. https://doi.org/10.1080/14656566.2020.1724957

80. *Ades F., Zardavas D., Bozovic-Spasojevic I. et al.* Luminal B breast cancer: Molecular characterization, clinical management, and future perspectives // J. Clin. Oncol. 2014. V. 32. № 25. P. 2794–2803. https://doi.org/10.1200/JCO.2013.54.1870

81. *Early Breast Cancer Trialists' Collaborative Group (EBCTCG).* Long-term outcomes for neoadjuvant versus adjuvant chemotherapy in early breast cancer: meta-analysis of individual patient data from ten randomised trials // Lancet Oncol. 2018. V. 19. № 1. P. 27–39. https://doi.org/10.1016/S1470-2045(17)30777-5

82. *Caparica R., Lambertini M., Pondé N. et al.* Post-neoadjuvant treatment and the management of residual disease in breast cancer: state of the art and perspectives // Ther. Adv. Med. Oncol. 2019. V. 11. e1758835919827714. https://doi.org/10.1177/1758835919827714

83. *Avraham A., Uhlmann R., Shperber A. et al.* Serum DNA methylation for monitoring response to neoadjuvant chemotherapy in breast cancer patients // Int. J. Cancer. 2012. V. 131. № 7. e1166–e1172. https://doi.org/10.1002/ijc.27526

84. *Sharma G., Mirza S., Parshad R. et al.* DNA methylation of circulating DNA: A marker for monitoring efficacy of neoadjuvant chemotherapy in breast cancer patients // Tumour Biol. 2012. V. 33. № 6. P. 1837–1843. https://doi.org/10.1007/s13277-012-0443-y

85. *Takahashi H., Kagara N., Tanei T. et al.* Correlation of methylated circulating tumor DNA with response to neoadjuvant chemotherapy in breast cancer patients // Clin. Breast Cancer. 2017. V. 17. № 1. P. 61–69.e3. https://doi.org/10.1016/j.clbc.2016.06.006

86. *Sigin V.O., Kalinkin A.I., Kuznetsova E.B. et al.* DNA methylation markers panel can improve prediction of response to neoadjuvant chemotherapy in luminal B breast cancer // Sci. Rep. 2020. V. 10. № 1. P. 9239. https://doi.org/10.1038/s41598-020-66197-1

87. *Кекеева Т.В., Жинжило Т.А., Ненахова Ю.Н. и др.* Метилирование гена *ADCY8* в плазме крови как предиктивный маркер ответа на неоадъювантную химиотерапию рака молочной железы // Мед. генетика. 2018. Т. 17. № 10. С. 38–41.

88. *Moss J., Zick A., Grinshpun A. et al.* Circulating breast-derived DNA allows universal detection and monitoring of localized breast cancer // Ann. Oncol. 2020. V. 31. № 3. P. 395–403. https://doi.org/10.1016/j.annonc.2019.11.014

89. *Fiegl H., Millinger S., Mueller-Holzner E. et al.* Circulating tumor-specific DNA: A marker for monitoring efficacy of adjuvant therapy in cancer patients // Cancer Res. 2005. V. 65. № 4. P. 1141–1145. https://doi.org/10.1158/0008-5472.CAN-04-2438

90. *Gil E.Y., Jo U.H., Jeong H. et al.* Promoter methylation of RASSF1A modulates the effect of the microtubule-targeting agent docetaxel in breast cancer // Int. J. Oncol. 2012. V. 41. P. 611–620. https://doi.org/10.3892/ijo.2012.1470

91. *Kajabova V., Smolkova B., Zmetakova I. et al.* RASSF1A promoter methylation levels positively correlate with estrogen receptor expression in breast cancer patients // Transl. Oncol. 2013. V. 6. P. 297–304. https://doi.org/10.1593/tlo.13244

92. *Liggett T.E., Melnikov A.A., Marks J.R. et al.* Methylation patterns in cell-free plasma DNA reflect removal of the primary tumor and drug treatment of breast cancer patients // Int. J. Cancer. 2011. V. 128. № 2. P. 492–499. https://doi.org/10.1002/ijc.25363

93. *Mastoraki S., Strati A., Tzanikou E. et al.* ESR1 methylation: A liquid biopsy-based epigenetic assay for the follow-up of patients with metastatic breast cancer receiving endocrine treatment // Clin. Cancer Res. 2018. V. 24. № 6. P. 1500–1510. https://doi.org/10.1158/1078-0432.CCR-17-1181

94. *Zurita M., Lara P.C., Del Moral R. et al.* Hypermethylated 14-3-3-sigma and ESR1 gene promoters in serum as candidate biomarkers for the diagnosis and treatment efficacy of breast cancer metastasis // BMC Cancer. 2010. V. 10. P. 217. https://doi.org/10.1186/1471-2407-10-217

Genes Methylation in Cell-Free Circulating Plasma DNA as Prognostic and Predictive Factors in Breast Cancer

S. V. Rykov[a, b, *], E. A. Filippova[a], V. I. Loginov[a], and E. A. Braga[a, **]

[a] *Federal State Budgetary Scientific Institution "Institute of General Pathology and Pathophysiology", Moscow, 125315 Russia*

[b] *National Research Centre "Kurchatov Institute" – GosNIIgenetika, Kurchatov Genomic Center, Moscow, 117545 Russia*

**e-mail: enterru@yandex.ru*

***e-mail: eleonora10_45@mail.ru*

Despite significant progress in early diagnosis and treatment of breast cancer, a high chance for recurrence remains. The existing methods of detecting relapse are not suitable for early detection, since they require invasive procedures and are not able to promptly accompany treatment. A promising method has become a liquid biopsy based on the analysis of genetic material from peripheral blood, in particular the fraction of cell-free circulating DNA (cfDNA) of blood plasma. Determination of mutations, changes in copy number and methylation of individual genes in cfDNA makes it possible to track changes both in the tumor focus and in the pathological process in the patient as a whole, and it is little invasive and with good accuracy and promptness. This review summarizes the literature data on the prognostic and predictive value of epigenetic markers based on the analysis of gene methylation in cfDNA of plasma and serum in patients with breast cancer. It has been shown that genes hypermethylated in cfDNA (*RASSF1A*, *RARB*, *SOX17*, *WNT5A*, etc.) can serve as effective markers of overall and tumor-specific survival and chemoresistance. An important advantage of methylation markers, in contrast to markers based on single nucleotide substitutions, microsatellites and copy number variations, is their versatility, early manifestation, and a clear connection with the biology of the pathological process. Methylation analysis of the most effective markers can solve the problem of timely detection of metastasis and recurrence of the disease; promptly monitor the response to therapy during courses of neoadjuvant and adjuvant chemotherapy, which is the basis for personalized treatment of breast cancer patients.

Keywords: DNA methylation, hypermethylated genes, cell-free circulating DNA, breast cancer, blood plasma, predictive factor, prognostic factor.

ГЕНЕТИКА, 2021, том 57, № 11, с. 1241–1245

ОБЗОРНЫЕ И ТЕОРЕТИЧЕСКИЕ СТАТЬИ

УДК 575.21;575.22

ФЕНОТИПИЧЕСКИ НЕСТАБИЛЬНЫЕ МУТАЦИИ – МАРКЕРЫ ХРОМОСОМНЫХ ПЕРЕСТРОЕК С УЧАСТИЕМ ДНК-ТРАНСПОЗОНОВ

© 2021 г. Л. П. Захаренко*

Федеральный исследовательский центр, Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, 630090 Россия

**e-mail: zakharlp@bionet.nsc.ru*

Поступила в редакцию 01.02.2021 г.
После доработки 20.02.2021 г.
Принята к публикации 09.03.2021 г.

В обзоре анализировали правомочность использования нестабильных аллелей генов *Drosophila melanogaster* для оценки скорости перемещения мобильных генетических элементов по частоте смены фенотипа. На нескольких примерах показано, что нестабильность аллелей со встроенными в них транспозонами обусловлена в большей степени рекомбинациями между ними, чем истинным перемещением транспозонов с участием транспозазы.

Ключевые слова: нестабильные мутации, мобильные генетические элементы, рекомбинация, дрозофила.

DOI: 10.31857/S0016675821110151

Средняя скорость перемещения мобильных генетических элементов (**МГЭ**) в геноме *Drosophila melanogaster* незначительна и составляет $2.6 \times \times 10^{-4}$–$5.0 \times 10^{-4}$ для инсерций и 4.0×10^{-6} для эксцизий на копию за поколение [1, 2]. По другим данным, в отсутствии селекции скорость инсерций МГЭ намного ниже, но также превышает скорость эксцизий и составляет 2.1×10^{-9} и 1.4×10^{-10} на сайт за поколение соответственно [3]. Оценка скорости перемещения МГЭ из-за редкости событий связана с анализом большого количества данных. По этой причине данные полногеномного секвенирования все шире используются для оценки скорости перемещения МГЭ. Однако скорость перемещения МГЭ, измеренная с помощью цитологических методов [4, 5], может не совпадать по порядку величины с результатами полногеномного секвенирования [6] в силу существования “подводных камней”, приводящих к артефактам в интерпретации и связанных со сложностью анализа повторенных последовательностей [7–9], к которым относятся МГЭ. Оценка скорости перемещения МГЭ с помощью полногеномного секвенирования по некоторым данным может быть сильно завышена [10].

Очевидно, что мобильные элементы могут менять свое положение или число в геноме как за счет активности соответствующих транспозаз, так и за счет рекомбинаций между собой или между другими повторами. Мобильные генетические элементы обнаруживают в ряде случаев на концах перестроек, то есть они дают свой вклад в перестройки генома дрозофил разных видов на популяционном и на эволюционном уровне [11–13].

Некоторые авторы оценивают скорость перемещения МГЭ косвенно по смене фенотипа нестабильных мутаций, вызванных встройками МГЭ. При этом нестабильные мутации ревертируют к дикому типу или дают производные с другими морфологическими проявлениями. Подразумевается, что нестабильность таких генов обусловлена выщеплением МГЭ с участием транспозазы [14, 15]. Возникают вопросы, правомочно ли использовать смену фенотипа нестабильных аллелей для оценки истинной скорости перемещения МГЭ, обусловленной активностью транспозаз, и как перестройки влияют на нестабильность инсерционных аллелей?

Нестабильность аллелей по окраске зерен у кукурузы связана с хромосомными перестройками

Исследуя нестабильность аллелей по окраске зерен у кукурузы, Б. Макклинток обнаружила, что в некоторых линиях разрывы хромосом проходили рядом с локусом, который она назвала *Dissociation* (***Ds***). *Ds* мог перемещаться по геному в случайные места. Мобильность *Ds* зависит от присутствия другого локуса, названного *Activator* (***Ac***) [16, 17]. Более поздние молекулярно-генетические исследования показали, что *Ds* – это укороченный вариант *Ac*-последовательности, явля-

ющейся автономным мобильным элементом. *Ac* относится к семейству hAT транспозонов и кодирует всего один ген – транспозазу [18]. *Ac/Ds* элементы перемещаются, как все транспозоны hAT-семейства, посредством cut-and-paste механизма [19]. Полноразмерная копия *Ac* встречается редко и присутствует в геноме в небольшом количестве. В то же время вариантов *Ds* множество, однако к перемещению способны лишь те из них, что сохранили концевые последовательности [20].

Уровень нестабильности в разных аллелях кукурузы, обнаруженный Б. Макклинток, был различным. Часть мутантных аллелей была стабильна, другие аллели могли иметь нормальный фенотип, но были нестабильными и в следующих поколениях обнаруживали мутантные аллели с разной степенью экспрессии признака. Кроме этого, оказалось, что не все нестабильные аллели требовали присутствия активатора *Ac*, среди них были и автономные аллели, нестабильность которых была связана с перестройками, а не с активностью *Ac*. Позже выяснилось, что для образования перестроек необходимы две последовательности, расположенные относительно недалеко друг от друга. *Ac* более чем в 60% и в меньшей степени *Ds* (более 44%) обнаруживают недалеко от другой последовательности этого семейства – на расстоянии от нескольких нуклеотидов до 10 сантиморганид [21]. В зависимости от локализации и ориентации двух МГЭ после рекомбинации их друг с другом могли образоваться делеции, дупликации или рекомбинации, видимые при наличии морфологических маркеров даже на световом уровне на цитологических препаратах. В этом случае наличие концевых повторов необязательно, и рекомбинировать могут МГЭ в разных сочетаниях: *Ac/Ds*, *Ac/Ac* или *Ds/Ds* [21]. Перестройки, которые Б. Макклинток наблюдала на цитологическом уровне, были результатом альтернативной рекомбинации, которая проходила между соседними МГЭ на стадии сестринских хроматид [22, 23]. Перемещения МГЭ сами по себе перестроек не вызывали. Б. Макклинток удалось найти такие линии кукурузы, которые содержали удачную комбинацию вариантов *Ds* и *Ac*, и благодаря этому сделать свое выдающееся открытие без молекулярного анализа, наблюдая за перестройками хромосом на цитологическом уровне и за окраской зерен. Итак, в основе нестабильности аллелей кукурузы лежит не столько перемещение МГЭ, сколько рекомбинация между ними.

Ac/Ds-система кукурузы не уникальна, она работает в более чем 30 видах растений. Перестройки индуцируются сестринскими хроматидными обменами, гомологичной рекомбинацией и транспозицией МГЭ [24–27]. На базе МГЭ кукурузы созданы вектора, которые используются на других видах для манипуляции с хромосомами, чтобы индуцировать относительно большие хромосомные делеции [27, 28].

Мода на мутацию по гену yellow в Умани – результат распространения инверсии между двумя hobo-элементами, а вспышка мутабильности – результат рекомбинации между ними

Сходная ситуация была описана при анализе причины “моды на мутацию” и вспышки мутабильности по гену *yellow* в популяции *D. melanogaster* в Умани в 1980-е гг. В течение десятилетия в этой популяции мутанты по гену *yellow* встречались с повышенной (до 3%) частотой [29]. Часть аллелей была стабильна, другая часть проявляла нестабильность и давала каскад производных с разной частотой и с разным морфологическим проявлением. Анализ показал, что “мода на мутацию” была вызвана распространением в популяции инверсии регуляторной зоны гена *yellow* между двумя *hobo*-элементами [30–33], которые относятся к hAT-семейству транспозонов, как и *Ac/Ds*-транспозоны кукурузы [34].

В гомозиготном состоянии инверсия в регуляторной зоне гена *yellow* была относительно стабильна. Однако при скрещивании таких мутантных самцов с самками лабораторной линии со сцепленными Х-хромосомами, инверсия реинвертировала за счет рекомбинации между разнонаправленными *hobo*, окаймляющими инверсию. Тогда появлялись мухи с нормальным фенотипом, но в регуляторной зоне гена *yellow* оставалась встройка *hobo*-элемента, которая тем не менее не препятствовала нормальной работе гена. Такие псевдонормальные мутации были в свою очередь нестабильны и могли давать мутантные производные с разной степенью нестабильности. Видимая глазом фенотипическая нестабильность аллелей была лишь вершиной айсберга, поскольку инверсия не только реинвертировала, но и в ряде случаев мультиплицировалась с вовлечением в процесс части гена *yellow*. Разный уровень нестабильности был обусловлен разницей в относительном расположении рекомбинирующих МГЭ друг относительно друга и их ориентацией. Если рекомбинация шла между однонаправленными МГЭ и регуляторная зона гена *yellow* выпадала, то мутация была фенотипически стабильной, но на молекулярном уровне процессы могли продолжаться за счет рекомбинаций между мультиплицированными последовательностями гена *yellow* [31–33].

Рекомбинации между *hobo*-элементами у самцов в нестабильных линиях из Умани шли на премейотической стадии между сестринскими хроматидами, поскольку в фенотипически нормальных линиях, но с *hobo*-элементом в регуляторной зоне гена *yellow*, в некоторых семьях все потомки были мутантами по цвету тела. Доля мутантных

потомков в семье в такой линии зависела от стадии сперматогенеза, на которой проходила рекомбинация. Полногеномный сиквенс гибридных геномов подтверждает данные о том, что большая часть рекомбинаций идет на премейотической стадии [35]. В некоторых случаях получали каскад производных, где норма чередовалась с мутантным фенотипом, то есть рекомбинации в гене *yellow* приводили к чередованию в ряду поколений инверсий и реинверсий регуляторной зоны между соседними *hobo* [31–33]. Для сравнения: большая часть пертурбаций *P*-элемента также имеет премейотическое происхождение (32%), в то время как в мейозе наблюдается приблизительно 4% событий, связанных с *P*-элементом [36].

Нестабильность в гене singed связана с перестройками с участием транспозонов

Другой пример спонтанно возникшей нестабильности описан для гена *singed*^{49} *D. melanogaster*. В этом случае за нестабильность гена *sn* также были ответственны перестройки с вовлечением *hobo*-транспозона [37]. Кроме полноразмерной копии в производных нестабильного гена были обнаружены дефектные копии *hobo*. Инверсия между 7D и 2D с *hobo*-элементами на концах была обнаружена в одном из членов каскада производных (sn^s) как результат рекомбинации между разнонаправленными *hobo*. Хромосомные перестройки между транспозонами, связанные с делециями или инверсиями, часто обнаруживали и в других системах [12].

Описана также нестабильность в гене *singed D. melanogaster* с участием *P*-элемента. В нестабильных *sn* обнаруживают два *P*-элемента, расположенных голова к голове или тандемно в разной ориентации и в разной локализации [11, 38, 39]. Считается, что выпадение одного из них приводит к смене фенотипа. Однако выпадение *P*-элемента может быть как результатом выщепления за счет активности транспозазы, так и результатом рекомбинации между одинаково ориентированными МГЭ. Так, для одного из производных нестабильного *sn* описана инверсия между 17C и 7D1-2 с *P*-элементами на концах, сопровождающаяся сменой фенотипа [11]. Инверсии с *P*-элементами на концах способны реинвертировать с большой точностью с восстановлением исходного фенотипа [40].

В работах по гибридному дисгенезу у *D. melanogaster* при оценке активности *P*-элемента используют нестабильные аллели *singed* [4, 5]. Поскольку ген *singed* расположен в Х-хромосоме, самцов таких линий скрещивают с лабораторными самками со сцепленными Х-хромосомами. Скорость перемещения *P*-элемента косвенно оценивают по доле потомков с изменившимся по сравнению с родительским фенотипом. Однако даже если *P*-элемент переместился с участием транспозазы, а не за счет рекомбинаций, но событие произошло на ранних стадиях сперматогенеза, и соответствующий клон клеток размножился, то частота смены фенотипа будет завышена и не будет отражать истинную частоту перемещения МГЭ.

Таким образом, смена фенотипа в нестабильных аллелях и у кукурузы и у дрозофилы связана не столько с перемещением МГЭ с участием транспозазы, сколько с рекомбинацией между соседними МГЭ или другими повторами. В свете вышесказанного, истинную скорость перемещения МГЭ, обусловленную активностью транспозазы, а не рекомбинацией, нельзя однозначно оценивать по смене фенотипа нестабильных аллелей, особенно если нестабильность сопровождается каскадом производных. Экстраполировать частоту смены фенотипа нестабильного аллеля на активность МГЭ неправомочно, так как последовательность МГЭ в данной ситуации выступает как повтор, участвующий в рекомбинации, а не как источник транскрипционной активности, приводящий к перемещению МГЭ.

Работа поддержана бюджетным проектом 0259-2021-0016.

Автор благодарит И.К. Захарова за ценные замечания.

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

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

СПИСОК ЛИТЕРАТУРЫ

1. *Nuzhdin S.V., Pasyukova E.G., Mackay T.F.* Accumulation of transposable elements in laboratory lines of *Drosophila melanogaster* // Genetica. 1997. V. 100. Is. 1–3. P. 167–175.
2. *Maside X., Assimacopoulos S., Charlesworth B.* Rates of movement of transposable elements on the second chromosome of *Drosophila melanogaster* // Genet. Res. 2000. V. 75. Is. 3. P. 275–284. https://doi.org/10.1017/s0016672399004474
3. *Adrion J.R., Song M.J., Schrider D.R. et al.* Genome-wide estimates of transposable element insertion and deletion rates in *Drosophila melanogaster* // Genome Biol. Evol. 2017. V. 9. Is. 5. P. 1329–1340. https://doi.org/10.1093/gbe/evx050
4. *Eggleston W.B., Johnson-Schlitz D.M., Engels W.R.* P-M hybrid dysgenesis does not mobilize other transposable element families in *D. melanogaster* // Nature. 1988. V. 331. Is. 6154. P. 368–370. https://doi.org/10.1038/331368a0
5. *Eggleston W.B., Rim N.R., Lim J.K.* Molecular characterization of *hobo*-mediated inversions in *Drosophila melanogaster* // Genetics. 1996. V. 144. Is. 2. P. 647–656.

6. *Khurana J.S., Wang J., Xu J. et al.* Adaptation to *P* element transposon invasion in *Drosophila melanogaster* // Cell. 2011. V. 147. Is. 7. P. 1551–1563. https://doi.org/10.1016/j.cell.2011.11.042
7. *Novak P., Neumann P., Macas J.* Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data // BMC Bioinform. 2010. V. 11. P. 378. https://doi.org/10.1186/1471-2105-11-378
8. *Goubert C., Modolo L., Vieira C. et al. De novo* assembly and annotation of the Asian tiger mosquito (*Aedesal bopictus*) repeatome with dnaPipeTE from raw genomic reads and comparative analysis with the yellow fever mosquito (*Aedesa egypti*) // Genome Biol. Evol. 2015. V. 7. P. 1192–1205. https://doi.org/10.1093/gbe/evv050
9. *Mohamed M., Dang N.T., Ogyama Y. et al.* A transposon story: From TE content to TE dynamic invasion of *Drosophila* genomes using the single-molecule sequencing technology from Oxford nanopore // Cells. 2020. V. 9. Is. 8. E1776. https://doi.org/10.3390/cells9081776
10. *Treiber C.D., Waddell S.* Resolving the prevalence of somatic transposition in Drosophila // Elife. 2017. V. 6. e28297. https://doi.org/10.7554/eLife.28297
11. *Roiha H., Rubin G.M., O'Hare K. P* element insertions and rearrangements at the singed locus of *Drosophila melanogaster* // Genetics. 1988. V. 119. P. 73–83.
12. *Lim J.K., Simmons M.J.* Gross chromosome rearrangements mediated by transposable elements in *Drosophila melanogaster* // Bioessays. 1994. V. 16. Is. 4. P. 269–275. https://doi.org/10.1002/bies.950160410
13. *Dolgin E.S., Charlesworth B.* The effects of recombination rate on the distribution and abundance of transposable elements // Genetics. 2008. V. 178. № 4. P. 2169–2177. https://doi.org/10.1534/genetics.107.082743
14. *Paterson J., Simmons M.J., O'Hare K.* Transcription of the *singed-weak* mutation of *Drosophila melanogaster*: Elimination of *P*-element sequences by RNA splicing and repression of singed transcription in a *P* genetic background // Mol. Genet. Genomics. 2007. V. 278. P. 53–64. https://doi.org/10.1007/s00438-007-0227-z123
15. *Ota R., Kobayashi S. Myc* plays an important role in *Drosophila* P-M hybrid dysgenesis to eliminate germline cells with genetic damage // Commun. Biol. 2020. V. 3. P. 185. https://doi.org/10.1038/s42003-020-0923-3
16. *McClintock B.* The origin and behavior of mutable loci in maize // Proc. Natl Acad. Sci. USA. 1950. V. 36. Is. 6. P. 344–355. https://doi.org/10.1073/pnas.36.6.344
17. *McClintock B.* Chromosome organization and genic expression // Cold Spring Harb Symp. Quant. Biol. 1951. V. 16. P. 13–47.
18. *Kunze R., Weil C.F.* The hAT and CACTA superfamilies of plant transposons // Mobile DNA II / Eds Craig N.L., Craigie R., Gellert M. et al. Washington: ASM Press, 2002. P. 565–610.
19. *McBlane J.F., van Gent D.C., Ramsden D.A. et al.* Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps // Cell. 1995. V. 83. P. 387–395. https://doi.org/10.1016/0092-8674(95)90116-7
20. *Coupland G., Plum C., Chatterjee S. et al.* Sequences near the termini are required for transposition of the maize transposon *Ac* in transgenic tobacco plants // Proc. Natl Acad. Sci. USA. 1989. V. 86. P. 9385–9388. https://doi.org/10.1073/pnas.86.23.9385
21. *Vollbrecht E., Duvick J., Schares J.P. et al.* Genome-wide distribution of transposed dissociation elements in maize plant cell // The Plant Cell. 2010. V. 22. Is. 6. P. 1667–1685. https://doi.org/10.1105/tpc.109.073452
22. *Weil C.F., Wessler S.R.* Molecular evidence that chromosome breakage by *Ds* elements is caused by aberrant transposition // Plant Cell. 2010. V. 22. Is. 6. P. 1667–1685. https://doi.org/10.1105/tpc.109.073452
23. *Fujimoto S., Matsunaga S., Murata M.* Mapping of T-DNA and *Ac/Ds* by TAIL-PCR to analyze chromosomal rearrangements // Methods Mol. Biol. 2016. V. 1469. P. 207–216. https://doi.org/10.1007/978-1-4939-4931-1_17
24. *Bai L., Brutnell T.P.* The activator/dissociation transposable elements comprise a two-component gene regulatory switch that controls endogenous gene expression in maize // Genetics. 2011. V. 187. Is. 3. P. 749–759. https://doi.org/10.1534/genetics.110.124149
25. *Xuan Y.H., Peterson T., Han C.D.* Generation and analysis of transposon *Ac/Ds*-induced chromosomal rearrangements in rice plants // Methods Mol. Biol. 2016. V. 1469. P. 49–61. https://doi.org/10.1007/978-1-4939-4931-1_4
26. *Lazarow K., Doll M.L., Kunze R.* Molecular biology of maize *Ac/Ds* elements: An overview // Methods Mol. Biol. 2013. V. 1057. P. 59–82. https://doi.org/10.1007/978-1-62703-568-2_5
27. *Mielich K., Shtifman-Segal E., Golz J.C. et al.* Maize transposable lements *Ac/Ds* as insertion mutagenesis tools in *Candida albicans* // G3 (Bethesda). 2018. V. 8. Is. 4. P. 1139–1145. https://doi.org/10.1534/g3.117.300388
28. *Puchta H., Fauser F.* Gene targeting in plants: 25 years later // Int. J. Dev. Biol. 2013. V. 57. Is. 6–8. P. 629–637. https://doi.org/10.1387/ijdb.130194hp
29. *Захаров И.К., Иванников А.В., Скибицкий Е.С. и др.* Генетические свойства аллелей генов Х-хромосомы, выделенных из природной популяции *Drosophila melanogaster* в период вспышки мутаций // Докл. Акад. наук. 1995. Т. 341. № 1. С. 126–129.
30. *Грачева Е.М., Захаров И.К., Волошина М.А. и др.* Взрыв мутабильности по гену *yellow* в природной популяции *Drosophila melanogaster* связан с инсерцией *hobo*-транспозона // Генетика. 1998. Т. 34. № 4. С. 462–468.
31. *Захаренко Л.П., Захаров И.К., Романова О.А. и др.* Мода на мутацию в природной популяции *Drosophila melanogaster* Умани вызвана распростране-

нием *hobo* индуцированной инверсии в регуляторной зоне гена *yellow* // Генетика. 2000. Т. 36. № 6. С. 740–748.

32. *Захаренко Л.П., Захаров И.К., Волошина М.А. и др.* Причина сохранения высокой нестабильности по гену *yellow* в линиях *Drosophila melanogaster*, выделенных в период "моды на мутацию" в популяции Умани // Генетика. 2004. Т. 40. № 3. С. 316–321.
33. *Захаренко Л.П., Коваленко Л.В., Захаров И.К., Май С.* Персистентная локус-специфическая генетическая нестабильность *yellow2-717* и *NotchUc-1* у *Drosophila melanogaster* связана с мультипликацией *hobo* // Цитология. 2007. Т. 1. № 6. С. 497–502.
34. *Arensburger P., Hice R.H., Zhou L. et al.* Phylogenetic and functional characterization of the *hAT* transposon superfamily genetics // Genetics. 2011. V. 188. Is. 1. P. 45–57.
https://doi.org/10.1534/genetics.111.126813
35. *Hemmer L.W., Dias G.B., Smith B. et al.* Hybrid dysgenesis in *Drosophila virilis* results in clusters of mitotic recombination and loss-of-heterozygosity but leaves meiotic recombination unaltered // Mob. DNA. 2020. V. 11. P. 10.
https://doi.org/10.1186/s13100-020-0205-0
36. *Daniels S.B., Chovnick A. P* element transposition in *Drosophila melanogaster*: An analysis of sister-chromatid pairs and the formation of intragenic secondary insertions during meiosis // Genetics. 1993. V. 133. Is. 3. P. 623–636.
37. *O'Hare K., Tam J.L., Lim J.K. et al.* Rearrangements at a *hobo* element inserted into the first intron of the *singed* gene in the unstable sn^{49} system of *Drosophila melanogaster* // Mol. Gen. Genet. 1998. V. 257. Is. 4. P. 452–460.
https://doi.org/10.1007/s004380050669
38. *Nitasaka E., Yamazaki T.* Excision of one of two defective *P* elements as the cause of alternate mutational events (sn^+ and sn^e) of the *singed* bristle allele sn^w in *Drosophila melanogaster* // Japn. J. Genet. 1988. V. 63. Is. 4. P. 303–312.
https://doi.org/10.1266/jjg.63.303
39. *Ortori C.A., Chambers D., Brookfield J.F.* The molecular basis of instability of the *singed* (very weak) mutation in *Drosophila melanogaster* // Genet Res. 1994. V. 63. Is. 1. P. 19–26.
https://doi.org/10.1017/s0016672300032043
40. *Engels W.R., Preston C.R.* Formation of chromosome rearrangements by P factors in Drosophila // Genetics. 1984. V. 107. Is. 4. P. 657–678.

Phenotypically Unstable Mutations as Markers of Chromosomal Rearrangements Involving DNA Transposons

L. P. Zakharenko*

The Federal Research Center Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 630090 Russia

**e-mail: zakharlp@bionet.nsc.ru*

The suitability of using unstable alleles of *Drosophila melanogaster* to estimate the rate of movement of transposable elements (TEs) by the frequency of phenotype change was investigated. Several examples show that the instability of alleles with TE introduced in them is due more to recombinations between TEs than to the true movement of TEs by transposase.

Keywords: unstable alleles, transposable elements, recombination, *Drosophila*.

ГЕНЕТИКА, 2021, том 57, № 11, с. 1246–1251

ГЕНЕТИКА РАСТЕНИЙ

УДК 575.174.5+582.632.2

ИЗМЕНЧИВОСТЬ МИТОХОНДРИАЛЬНОЙ ДНК УКАЗЫВАЕТ НА АМЕРИКАНСКОЕ ПРОИСХОЖДЕНИЕ СОВРЕМЕННЫХ ПИХТ[1]

© 2021 г. В. Л. Семериков[1], *, С. А. Семерикова[1], Ю. А. Путинцева[2], **

[1]*Институт экологии растений и животных, Уральское отделение Российской академии наук, Екатеринбург, 620144 Россия*

[2]*Сибирский федеральный университет, Красноярск, 660041 Россия*

**e-mail: semerikov@ipae.uran.ru*

***e-mail: yputintseva@sfu-kras.ru*

Поступила в редакцию 20.12.2020 г.
После доработки 05.02.2021 г.
Принята к публикации 09.02.2021 г.

Приводятся результаты филогенетического анализа 15 видов пихт, представляющих все основные эволюционные линии рода *Abies*, и *Keteleeria davidiana*, использованной в качестве внешней группы. Данные включают нуклеотидные последовательности митохондриальной ДНК длиной около 28000 пн, полученные путем частичного ресеквенирования сборки митохондриального генома пихты сибирской *A. sibirica*. Установлено базальное положение линий мтДНК некоторых американских пихт, что подтверждает американское происхождение современных *Abies*. Митотипы большинства евроазиатских видов образуют кладу, дочернюю по отношению к американским пихтам, указывая на ее происхождение в результате одной миграции из Америки в Евразию. В то же время ранее полученные данные по ядерной и хлоропластной ДНК свидетельствуют о неоднократных миграциях пихт из Америки в Евразию. Этот конфликт митохондриальных и ядерных данных можно объяснить, предполагая гибридный захват митохондриальной ДНК аборигенных евроазиатских видов видами-мигрантами.

Ключевые слова: Abies, митохондриальная ДНК, интрогрессивная гибридизация, молекулярная филогения.

DOI: 10.31857/S0016675821100118

Миграции между Северной Америкой и Евразией оказали решающий вклад в формирование разнообразия многих таксономических групп умеренной и бореальной флор северного полушария (например [1, 2]). В ходе трансконтинентальных миграций переселенцы могут вытеснять родственные аборигенные виды, а при отсутствии репродуктивных барьеров – формировать зоны интрогрессивной гибридизации. В последнем случае распространяемый с пыльцой генетический поток со стороны более адаптивного вида "разбавляет" гены менее адаптивного вида, которым обычно является аборигенный вид, вплоть до их полной замены. В то же время менее подвижные гены, передающиеся по материнской линии и не распространяющиеся с пыльцой (например хлоропластные и митохондриальные у цветковых растений), наследуются гибридом от аборигенного вида, что проявляется в виде "гибридного захвата" цитоплазматических геномов. Древние события "гибридного захвата" проявляются в конфликте между филогенетическими деревьями, основанными на маркерах с разным типом наследования. В частности, такие конфликты характерны для родов сем. Pinaceae (сосны, ели) [3–7], у которых маркеры ядерного, хлоропластного (**хлДНК**) и митохондриального (**мтДНК**) геномов имеют двуродительское, отцовское и материнское наследование, соответственно [8]. В современную эпоху локальные процессы "гибридного захвата", хотя и не связанные с трансконтинентальными миграциями, так же широко распространены [9, 10].

Согласно последнему биогеографическому исследованию рода *Abies* [7], наиболее базальными и удаленными от остальных видов по ядерной и хлоропластной ДНК являются некоторые американские пихты, что указывает на американское происхождение современных *Abies*. В то же время евроазиатские виды входят в несколько клад, часть которых содержат также и американские виды, поддерживая предположение о множественных миграциях из Америки в Евразию. Однако дерево

[1] Дополнительная информация для этой статьи доступна по doi 10.31857/S0016675821100118 для авторизованных пользователей.

мтДНК [7] демонстрирует разделение гаплотипов на два кластера – кластер американских видов вместе с несколькими островными азиатскими видами и кластер большинства евроазиатских видов, что, в свою очередь, соответствует однократной миграции мтДНК. Одновременный анализ ядерных, хлоропластных и митохондриальных данных позволил обосновать гипотезу о нескольких волнах миграции пихт из Северной Америки в Евразию [7], которые сопровождались гибридным замещением ядерного и хлоропластного геномов аборигенных форм – продуктов предыдущих волн миграции, геномами новых мигрантов, что возможно ввиду относительной легкости межвидовых скрещиваний у пихт. При этом, благодаря гибридному захвату виды-мигранты приобретали митохондриальные геномы видов-аборигенов. Однако разрешение полученного мтДНК дерева [7] было невысоким. Не было установлено положение внешней группы и, как следствие, не были выявлены наиболее дивергентные клады и не были определены филогенетические отношения евроазиатской и американской групп: являются ли они сестринскими, или одна из групп является частью другой. Данное филогенетическое исследование направлено на разрешение этих неопределенностей. Для этого была использована выборка видов, относящихся ко всем основным группам, выявляемым согласно ядерному дереву видов *Abies* [7]. Нуклеотидные последовательности включали 27 регионов митохондриального генома.

МАТЕРИАЛЫ И МЕТОДЫ

В филогенетическое исследование были включены 15 видов *Abies*, представляющих все основные клады дерева видов *Abies*, выделенные ранее на основании изменчивости ядерной и хлоропластной ДНК [7], и все основные секции рода [11], а также *Keteleeria davidiana* в качестве внешней группы (табл. 1).

Для исследования филогении мтДНК рода *Abies* было использовано ресеквенирование сборки митохондриального генома пихты сибирской *A. sibirica*, полученной на основе полногеномного NGS-секвенирования [12]. Для этого в контигах данной сборки были случайно выбраны 42 фрагмента. Для их амплификации и секвенированияс помощью программы Primer3 [13] были разработаны 67 пар ПЦР-праймеров. Фрагменты имели длину 700–1500 пн. Условия ПЦР и секвенирования соответствовали описанным в работе [12]. Для дальнейшего использования были отобраны 27 пар праймеров, дававших амплификацию не менее чем у 13 видов *Abies* из 15.

Редактирование сиквенсных данных и выравнивание последовательностей осуществлялось с помощью программ CodonCodeAligner v. 3.7.1 (CodonCodeCorporation) и BioEdit v. 7.2.5 [14]. Для филогенетической реконструкции нуклеотидные последовательности отдельных фрагментов мтДНК были конкатенированы в объединенный алайнмент. Анализ проводился с использованием байесовского метода (BI) и метода максимальной экономии (MP). Все признаки были равновзвешены. Инделы и инверсии рассматривались как отдельные события и были закодированы как (1/0). BI выполняли с использованием MrBAYES v. 3.1.2 [15] с использованием модели нуклеотидных замен GTR + G + I. Были сгенерированы две цепи MCMC по 10 млн поколений каждая и с сохранением текущих данных каждые 1000 поколений. Для контроля сходимости цепей учитывали их стандартное отклонение. Когда последнее упало ниже 0.01, уже сгенерированные 10 млн деревьев были отброшены как “burn-in”, а последующие 10 млн деревьев были использованы для построения консенсусного дерева по принципу “majority rule”. MP-анализ был выполнен с помощью PAUP * v. 4.0b10 [16] с использованием эвристического поиска, случайного порядка выбора с алгоритмом TBR, опцией COLLAPSE, MaxTrees = 500 и опцией MulTrees. Поддержка ветвей была оценена путем “bootstap” с использованием 1000 повторов.

РЕЗУЛЬТАТЫ

В результате секвенирования 27 фрагментов мтДНК были получены выровненные последовательности (алайнмент) общей длиной 29968 пн (Приложение, табл. S1). Без делеций и отсутствующих данных длина алайнмента у *Abies* составляла 14015 пн, из которых изменчивыми являлись 153 признака. Нуклеотидное разнообразие $\pi = 0.002$.

Байесовское филогенетическое дерево имело существенно лучшее разрешение, чем дерево, построенное на основе метода максимальной парсимонии, при этом они не противоречили друг другу (рис. 1). Евроазиатские виды *A. sibirica*, *A. nephrolepis*, *A. koreana*, *A. fargesii*, *A. recurvata*, *A. firma*, *A. alba* и *A. pindrow* образовали поддержанную кладу (апостериорная вероятность 1). В нее не вошли митотипы восточноазиатских видов недавнего американского происхождения *A. sachalinensis* и *A. mariesii* [7]. Ближайшими родственниками пихты сибирской по мтДНК являются *A. fargesii* (Китай) и *A. koreana* (Корея). Среди евроазиатских пихт гималайский вид *A. pindrow* оказался генетически более обособленным, чем европейско-средиземноморский вид *A. alba*.

Митотипы американских видов не образовали монофилетичную кладу и разделились на несколько дивергентных линий. Клада евроазиатских видов находится внутри клады нескольких американских видов плюс *A. sachalinensis* (1/0.73 апостериорная вероятность/поддержка бутстрэп). *A. bracteata* и *A. amabilis* (западно-североамериканские виды) об-

Таблица 1. Использованный материал

Таксон	Характеристика образца	Естественное распространение	Секция
Abies sibirica Ledeb.	51°48′/87°15′ E Is24-2007	Сибирь, Восточная Европа	*Balsamea*
A. sachalinensis (F. Schmidt) Mast.	46°40′/141°50′ E Is41-2008	Хоккайдо, Сахалин и Курилы	*Balsamea*
A. nephrolepis (Trautv. ex Maxim.) Maxim.	49°00′/131°05′ E Is01-2008	Северо-Восточная Азия	*Balsamea*
A. alba Mill.	47°48′/13°02′ E Is6-2010	Европа	*Abies*
A. balsamea (L.) Mill.	MBG, voucher A0718	Восток Сев. Америки	*Balsamea*
A. recurvata Mast.	Kew, voucher 36850	Юго-Зап. Китай	*Momi*
A. mariesii Mast.	35°46′/137°48′ Is5n-2015	Япония	*Amabilis*
A. firma Siebold & Zucc.	36°13′/140°06′ E Is7n-2015	Япония	*Momi*
A. fargesii Franch.	M, v. K09-13-062Is28-2014	Центральный Китай	*Pseudopicea*
A. amabilis Douglas ex J. Forbes	M, Is19-2014	Запад Сев. Америки	*Amabilis*
A. bracteata (D. Don) A. Poit	NBG, Is12-2011	Запад Сев. Америки	*Bracteata*
A. koreana E.H. Wilson	MBG, voucher A048Is48-2013	Корея	*Balsamea*
A. pindrow (Royle ex D. Don) Royle	34°03′/74°24′ E Is31-2013	Зап. Гималаи	*Momi*
A. concolor (Gordon) Lindl. ex Hildebr.	37°45′/119°30′ W Is13-2012	Запад Сев. Америки, Центральная Америка	*Grandis*
A. magnifica A. Murray	37°45′/119°30′ W Is16-2012	Запад Сев. Америки	*Nobilis*
Keteleeria davidiana v. *formosana* (Hayata) Hayata	Kew, voucher 40615	Китай	

Примечание. Характеристика образца: указан источник согласно [7] и Приложения, табл. S1, координаты природных популяций (с.ш./долгота), MBG – Московский ботанический сад, M – ботанический сад г. Мустила, Финляндия, Kew – ботанический сад Кью, NBG – Никитский ботанический сад. Секции указаны согласно [11].

разовали кладу с поддержкой 1/0.84. Также весьма дивергентным оказался митотип азиатской пихты *A. mariesii* (о. Хонсю), который занял неопределенное филогенетическое положение в основании дерева.

ОБСУЖДЕНИЕ

Мы получили на настоящий момент наиболее поддержанное филогенетическое дерево мтДНК *Abies*. Дерево в основном не противоречит построенному ранее на основе трех фрагментов мтДНК [7]. Евроазиатские виды пихт разделились сходным образом: митотипы большинства из них формируют евроазиатскую кладу (рис. 1), при этом гаплотипы части островных восточноазиатских видов (в настоящем исследовании это *A. sachalinensis* и *A. mariesii*) не входят в данную кладу и более близки американским гаплотипам, что соответствует их положению в работе [7]. Поддержка евроазиатской клады (рис. 1), согласно полученным данным, оказалась существенно выше (100), чем в предыдущем исследовании (57/54) [7]. Впервые показано, что митотипы большинства евроазиатских видов образуют кладу, дочернюю по отношению к американским пихтам. Это поддерживает гипотезу об американском происхождении современных евроазиатских пихт и о происхождении мтДНК основной части евроазиатских пихт от одной, первоначальной, миграции из Америки в Евразию [7]. Вероятно, последующие волны миграции сопровождались вытеснением видов предыдущих миграции, но при этом митохондриальные геномы первой волны наследовались видами более поздних волн. Филогенетические деревья мтДНК в обоих ис-

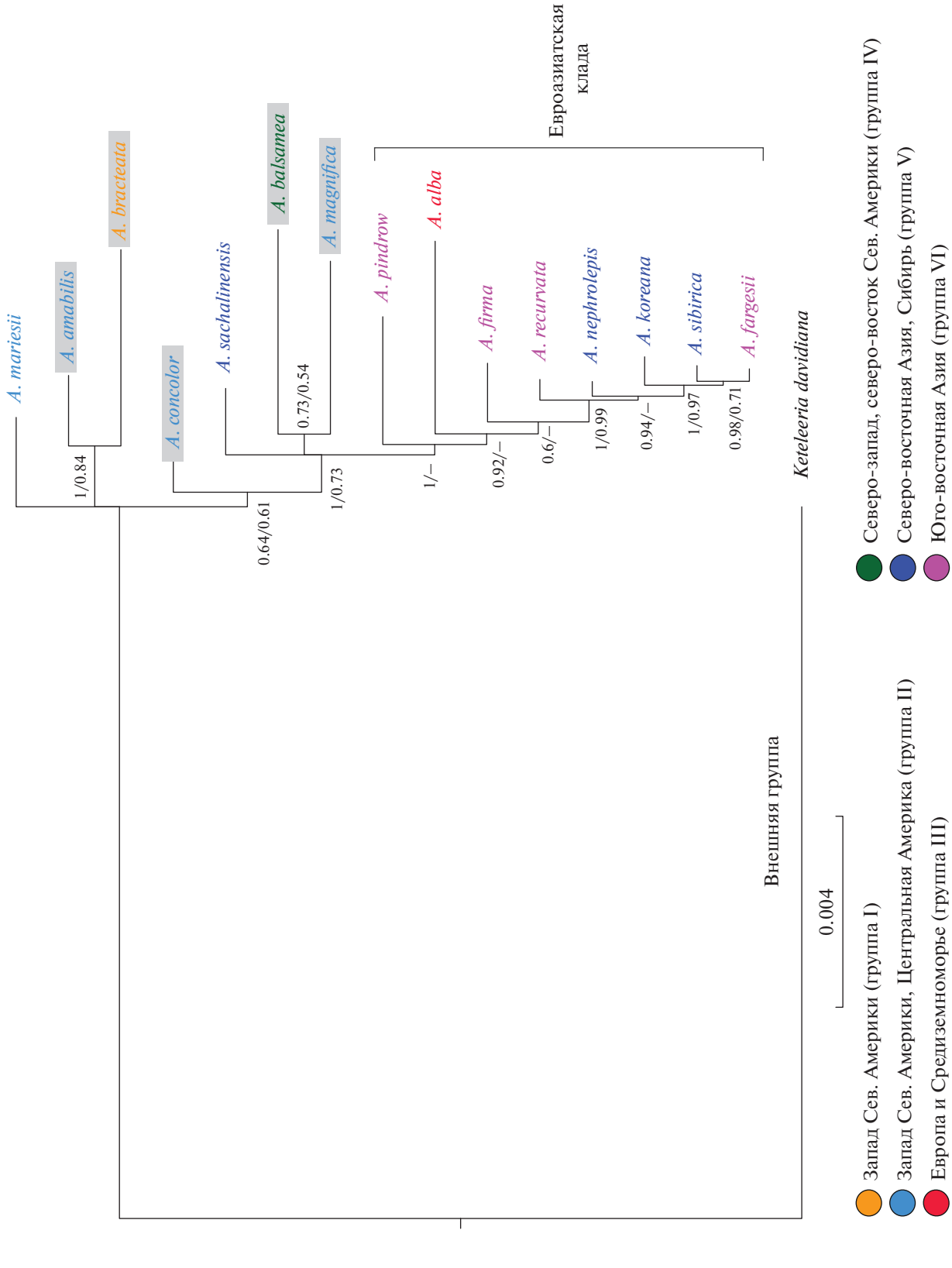

Рис. 1. Филогенетическое дерево 15 видов пихт, построенное на основе нуклеотидных последовательностей мтДНК. Статистическая поддержка полученных клад (BI апостериорная вероятность/MP бутстрэп) приведена в их узлах. Группы выделены на основе ядерной ДНК [7], их цветовое обозначение указано в соответствии с работой [7]. Названия таксонов из Северной Америки затенены.

следованиях существенно противоречили деревьям ядерной и хлДНК, в том числе в пределах Евразии. Например, митотипы родственных по ядерной ДНК *A. sibirica, A. koreana* и *A. nephrolepis* образуют кладу с митотипами *A. fargesii* и *A. recurvata*, хотя по ядерной ДНК они входят в две обособленные ветви – северо-восточноазиатскую и юго-восточноазиатскую. Несмотря на то, что по ядерной ДНК гималайский вид *A. pindrow* и является более близким родственником восточноазиатских видов, чем средиземноморский вид *A. alba*, но по мтДНК – наоборот, *A. alba* более близка к восточноазиатским видам. Такие противоречия могут быть результатом нескольких причин, включая неполный сортинг линий, гибридный захват, рекомбинацию мтДНК и взаимодействие этих факторов.

Несмотря на значительный объем данных секвенирования, многие узлы дерева оказались слабо поддержаны. Очевидно это может быть связано с низкой скоростью мутирования и, соответственно, с низкой изменчивостью мтДНК. Кроме того, филогенетический сигнал может значительно деградировать ввиду рекомбинации мтДНК. Благодаря этому при гибридном контакте филогенетически удаленных таксонов, возможно формирование химерных митотипов, как, например, у дальневосточных видов пихт [7]. В частности это может быть причиной значительного отклонения положения *A. mariesii* на дереве мтДНК от положения, реконструированного ранее, при использовании меньшего объема сиквенсных данных, когда митотип *A. mariesii* оказался родственным митотипам *A. balsamea* и *A. lasiocarpa* [7]. Мы ожидаем, что включение в анализ большей части видов *Abies* с использованием нескольких образцов на вид, увеличение количества и длины проанализированных фрагментов мтДНК заметно увеличат информативность филогенетического анализа для реконструкции истории пихт. Решение проблемы низкой изменчивости мтДНК видится в использовании NGS-секвенирования, которое позволяет получить значительно больший объем данных. Благодаря существенно большему числу копий на клетку мтДНК, по сравнению с ядерной ДНК, митохондриальные геномы могут быть собраны из прочтений (ридов) при относительно низкой глубине полногеномного секвенирования. Также может быть использовано обогащение геномных библиотек митохондриальной фракцией с помощью гибридизации с митохондриальными зондами “байтами” [17].

Работа выполнена в рамках государственного задания Института экологии растений и животных УрО РАН и при финансовой поддержке РФФИ, проект № 19-04-00795А.

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

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. *Tiffney B.H.* Perspectives on the origin of the floristic similarity between eastern Asia and eastern North America // J. Arnold Arboretum. 1985. V. 66. P. 73–94. https://doi.org/10.5962/bhl.part.13179
2. *Wen J., Nie Z.L., Ickert-Bond S.M.* Intercontinental disjunctions between eastern Asia and western North America in vascular plants highlight the biogeographic importance of the Bering land bridge from late Cretaceous to Neogene // J. Syst. Evol. 2016. V. 54. № 5. P. 469–490. https://doi.org/10.1111/jse.12222
3. *Tsutsui K., Suwa A., Sawada K. et al.* Incongruence among mitochondrial, chloroplast and nuclear gene trees in *Pinus* subgenus *Strobus* (Pinaceae) // J. Plant Res. 2009. V. 122. P. 509–521. https://doi.org/10.1007/s10265-009-0246-4
4. *Wang B.S., Wang X.R.* Mitochondrial DNA capture and divergence in *Pinus* provide new insights into the evolution of the genus // Mol. Phylogenet. Evol. 2014. V. 80. P. 20–30. https://doi.org/10.1016/j.ympev.2014.07.014
5. *Hao Z.Z., Liu Y.Y., Nazaire M. et al.* Molecular phylogenetics and evolutionary history of sect. Quinquefoliae (*Pinus*): Implications for Northern Hemisphere biogeography // Mol. Phylogenet. Evol. 2015. V. 87. P. 65–79. https://doi.org/10.1016/j.ympev.2015.03.013
6. *Ran J.H., Shen T.T., Liu W.J. et al.* Mitochondrial introgression and complex biogeographic history of the genus *Picea* // Mol. Phylogenet. Evol. 2015. V. 93. P. 63–76. https://doi.org/10.1016/j.ympev.2015.07.020
7. *Semerikova S.A., Khrunyk Y.Y., Lascoux M., Semerikov V.L.* From America to Eurasia: a multigenomes history of the genus *Abies* // Mol. Phylogenet. Evol. 2018. V. 125. P. 14–28. https://doi.org/10.1016/j.ympev.2018.03.009
8. *Neale D.B., Sederoff R.R.* Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine // Theor. Appl. Genet. 1989. V. 77. P. 212–216. https://doi.org/10.1007/BF00266189
9. *Watano Y., Kanai I., Tani N.* Genetic structure of hybrid zones between *Pinuspumila* and *P. parviflora* var. *pentaphylla* (Pinaceae) revealed by molecular hybrid index analysis // Am. J. Bot. 2004. V. 91. № 1. P. 65–72. https://doi.org/10.3732/ajb.91.1.65
10. *Semerikov V.L., Semerikova S.A., Polezhaeva M.A. et al.* Southern montane populations did not contribute to the recolonization of West Siberian Plain by Siberian larch (*Larixs ibirica*): A range-wide analysis of cytoplasmic markers // Mol. Ecol. 2013. V. 22. № 19. P. 4958–4971. https://doi.org/10.1111/mec.12433

11. *Farjon A., Rushforth K.D.*A classification of *Abies* Miller. (Pinaceae) // Not. Roy. Bot. Garden. Edinburg. 1989. V. 46. P. 59–79.
12. *Semerikov V.L., Semerikova S.A., Putintseva Y.A. et al.* Mitochondrial DNA in Siberian conifers indicates multiple postglacial colonization centers // Canad. J. For. Res. 2019. V. 49. № 8. P. 875–883. https://doi.org/10.1139/cjfr-2018-0498
13. *Untergasser A., Cutcutache I., Koressaar T. et al.* Primer3 – new capabilities and interfaces // Nucl. Ac. Res. 2012. V. 40. № 15. e115. https://doi.org/10.1093/nar/gks596
14. *Hall T.A.* BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT // Nucl. Ac. Symp. Series. 1999. V. 41. P. 95–98.
15. *Ronquist F., Huelsenbeck J.P.* MrBAYES 3: Bayesian phylogenetic inference under mixed models // Bioinformatics. 2003. V. 19. № 12. P. 1572–1574. https://doi.org/10.1093/bioinformatics/btg180
16. *Swofford D.L.* PAUP*. Phylogenetic analysis using parsimony (* and other methods). Version 4.0 beta10. Sunderland: Sinauer Associates, Massachusetts, 2002.
17. *Mamanova L., Coffey A.J., Scott C.E. et al.* Target-enrichment strategies for next-generation sequencing // Nat. Methods. 2010. V. 7. № 2. P. 111–118. https://doi.org/10.1038/nmeth.1419

Mitochondrial DNA Confirms the American Origin of Modern Firs

V. L. Semerikov[a, *], S. A. Semerikova[a], and Yu. A. Putintseva[b, **]

[a] *Institute of Plant and Animal Ecology, Ural Branch of Russian Academy of Science, Ekaterinburg, 620144 Russia*

[b] *Siberian Federal University, Krasnoyarsk, 660041 Russia*

**e-mail: semerikov@ipae.uran.ru*

***e-mail: yputintseva@sfu-kras.ru*

The results of phylogenetic analysis of 15 species of four, representing all the main groups of evolutionary lineages of the genus *Abies*, and *Keteleeria davidiana* as an outgroup, based on the nucleotide sequences of mitochondrial DNA with a size of about 28000 base pairs obtained by partial resequencing of the assembly of the mitochondrial genome of Siberian fir *A. sibirica* are presented. The basal position of mitochondrial DNA lineages of some American firs has been revealed, which confirms the American origin of modern *Abies*. The mitotypes of most Eurasian species form a daughter clade with respect to American firs, which indicates its origin as a result of one migration from America to Eurasia. At the same time, previously obtained data on nuclear (and chloroplast) DNA indicates repeated migrations of fir from America to Eurasia. This conflict between mitochondrial and nuclear data can be resolved by assuming the capture of the mitochondrial DNA of native species by the invading species during migrations.

Keywords: *Abies*, mitochondrial DNA, introgressive hybridization, molecular phylogeny.

ГЕНЕТИКА, 2021, том 57, № 11, с. 1252–1260

ГЕНЕТИКА ЖИВОТНЫХ

УДК 575.222:599.323.5

ИЗМЕНЧИВОСТЬ КОНТРОЛЬНОГО РЕГИОНА мтДНК ВОСТОЧНОЕВРОПЕЙСКОЙ ПОЛЕВКИ *Microtus rossiaemeridionalis* (Rodentia, Arvicolini) ИЗ ДВУХ ИНВАЗИВНЫХ ПОПУЛЯЦИЙ ЮГА ДАЛЬНЕГО ВОСТОКА РОССИИ

© 2021 г. И. Н. Шеремetьева[1, *], И. В. Картавцева[1], А. А. Емельянова[1], А. С. Лапин[2]

[1]*Федеральный научный центр биоразнообразия наземной биоты Восточной Азии Дальневосточного отделения Российской академии наук, Владивосток, 690022 Россия*

[2]*Хабаровская противочумная станция, Хабаровск, 680031 Россия*

**e-mail: sheremet76@yandex.ru*

Поступила в редакцию 02.02.2021 г.
После доработки 04.06.2021 г.
Принята к публикации 15.06.2021 г.

Изучена изменчивость контрольного региона мтДНК восточноевропейской полевки *Microtus rossiaemeridionalis* (Rodentia, Arvicolini) из двух инвазивных популяций юга Дальнего Востока России. Результаты продемонстрировали более низкое генетическое разнообразие в популяции восточноевропейской полевки из Советской Гавани по сравнению с популяцией из города Хабаровск. *F*-статистика Райта не выявила гетерогенности выборок из этих популяций. Высказано предположение, что обнаруженные гаплотипы относятся к субкладе EU европейской линии.

Ключевые слова: инвазия, изменчивость, контрольный регион мтДНК, восточноевропейская полевка, *Microtus rossiaemeridionalis.*

DOI: 10.31857/S0016675821110138

В настоящее время вторжение чужеродных видов в экосистемы рассматривается как одна из наиболее важных причин сокращения биологического разнообразия [1, 2], а прогнозирование инвазии является одной из основных задач охранной биологии [3, 4]. Еще в середине прошлого века Ч. Элтон [5] высказал предположение, что вторжения инвазивных видов происходят гораздо чаще, чем реально мы видим. Не каждая инвазия успешна, поскольку формируются популяции с очень низкой плотностью и/или малым размером. Подобные популяции нестабильны [6], в том числе из-за потери генетического разнообразия, инбридинга и дрейфа генов [7, 8]. Поэтому оценка уровня генетической изменчивости популяций и роли генетических процессов, происходящих в них, обеспечивающих успешность экспансии, являются ключевой задачей при изучении инвазивных видов.

Восточноевропейская полевка *Microtus rossiaemeridionalis* (=*M. mystacinus* De Filippi, 1865, *M. levis* Miller, 1908, *M. epiroticus* Ondrias, 1966, или *M. subarvalis* Meyer et al., 1969) – один из широкоареальных видов серых полевок группы “arvalis”. Основной ареал вида простирается от Южной Финляндии и Балтийского региона на восток – до Западной Сибири, и на юг – до Южного Кавказа и Северного Ирана, Турции и Греции. Для восточноевропейской полевки было показано наличие двух филогенетических линий мтДНК, выделенных согласно изменчивости гена цитохрома *b* [9, 10]. Эти линии имеют четкую географическую приуроченность: первая встречается на территории Северного Ирана, вторая – в Европе. При этом, европейская филогенетическая линия имеет три субклады: к TU подкладе относятся особи из Турции и Армении, GK – из Греции, EU – из большей части Европы, включая Украину и Румынию. Для восточноевропейской полевки, как и для других инвазивных млекопитающих, характерно относительно быстрое расширение ареала и встречаемость вида далеко за пределами нативного ареала [11]. В последние годы появилась информация о проникновении этой полевки с транспортом, в основном по Транссибирской магистрали [12–23], в отдельные районы Западной и Восточной Сибири и Дальнего Востока России. За счет значительной способности выживать и адаптироваться к новым условиям [24, 25] *M. rossiaemeridionalis* легко проникает в антропогенные биотопы, где достигает высокой численности [26, 27]. При этом способность приспосабливаться к антропогенной среде у этого вида превышает таковую

у полевой мыши *Apodemus agrarius* (Pallas, 1771), которая внесена в список "Самые опасные инвазивные виды России (ТОП-100)" [28]. В нативной части ареала восточноевропейская полевка является резервуаром возбудителей природно-очаговых инфекций (туляремии, чумы, клещевого энцефалита, геморрагической лихорадки с почечным синдромом и др.). Для инвазивной популяции из города Советская Гавань доказан факт циркуляции в ней хантавируса [29]. Таким образом, этот вид представляет потенциальную опасность для дальневосточного региона, поскольку одновременно происходит инвазия и новых для региона возбудителей инфекций, опасных для человека.

Высокий уровень генетического разнообразия, как правило, увеличивает устойчивость и успешность выживания популяций инвазивных видов [8, 30]. Цель настоящей работы – дать оценку генетического разнообразия двух инвазивных популяций восточноевропейской полевки юга Дальнего Востока для прогнозирования их устойчивости. В качестве молекулярно-генетического маркера был выбран некодирующий участок (контрольный регион) мтДНК, обладающий высокой скоростью накопления мутаций и позволяющий анализировать эволюционные процессы на популяционном и внутривидовом уровне [31]. Данный маркер часто используют для анализа внутривидовой и популяционной изменчивости как у серых полевок [32–36], так и у других млекопитающих [37–40]. Проведенное недавно исследование по сопоставлению изменчивости двух маркеров мтДНК у вида-двойника восточноевропейской полевки – *M. arvalis* s. str. формы obscurus показало, что несмотря на то, что число вариабельных сайтов в контрольном регионе меньше, чем в гене цитохрома *b*, число информативных сайтов больше [34]. Кроме этого было показано, что значения нуклеотидного разнообразия для последовательностей контрольного региона мтДНК выше в два раза, чем для гена цитохрома *b*.

МАТЕРИАЛЫ И МЕТОДЫ

В работе использовали 32 образца тканей восточноевропейских полевок, отловленных в двух инвазивных популяциях недавно обнаруженных на юге Дальнего Востока России (Хабаровский край) за пределами нативного ареала вида: г. Хабаровск (n = 10) и г. Советская Гавань (n = 22). Данные о численности полевок в двух популяций предоставлены Хабаровской противочумной станцией. Процент численности полевок рассчитывали как число попаданий особей на 100 ловушко-суток. Также в анализ включены гомологичные участки двух нуклеотидных последовательностей полной мтДНК восточноевропейской полевки, имеющиеся в наличии из основного ареала (точная локализация образца неизвестна) в GenBank/NCBI DQ015676 и NC_008064 [41]. Участки контрольного региона для этих нуклеотидных последовательностей были идентичны. Предварительно проведенное нами сравнение участка гена цитохрома *b* этих последовательностей с другими нуклеотидными последовательностями, имеющимися в GenBank/NCBI, показало, что данные образцы относятся к субкладе EU европейской линии. Еще одна нуклеотидная последовательность в GenBank/NCBI под номером DQ323955 в анализ не включена, поскольку содержала псевдоген [42].

Выделение ДНК из мышечных тканей, фиксированных 95% этиловом спиртом, проводили методом солевой экстракции [43] с добавлением протеиназы К. Для амплификации фрагмента контрольного региона мтДНК применяли праймеры Pro+ и Phe–, использованные ранее для других видов серых полевок [32, 44, 45]. Амплификацию проводили на наборах фирмы "СибЭнзим" (Новосибирск). Условия реакции подбирали эмпирически: начальная денатурация ДНК (94°C – 120 с), 40 циклов амплификации (94°C – 10 с, 52°C – 10 с, 72°C – 60 с) и достройка цепей (72°C – 420 с). Реакцию присоединения меченых нуклеотидов проводили при помощи набора Bright Dye Terminator ("Applied Biosystems", США) с использованием вышеуказанных праймеров.

Полученные на автоматическом секвенаторе ABI Prizm 3130 на базе Центра коллективного пользования ФНЦ Биоразнообразия ДВО РАН нуклеотидные последовательности редактировали и выравнивали вручную в программе BioEdit 7.0.9.0 [46]. Выравнивание проводили по нуклеотидной последовательности полного митохондриального генома из GenBank/NCBI под номером DQ015676 [41].

Сеть гаплотипов построена при помощи программы Network 5.0.0.3, где для расчета использован метод "median joining" [47]. Нуклеотидное и гаплотипическое разнообразие, число вариабельных сайтов, среднее число нуклеотидных различий оценивали с помощью DNAsp 6.10 [48]. Расчет статистик нейтральности (Tajima's D, Fu's F), F-статистику Райта и анализ демографической и пространственной экспансии проводили в программе Arlequin 3.1 [49].

РЕЗУЛЬТАТЫ

Для 32 образцов восточноевропейской полевки были получены нуклеотидные последовательности контрольного региона длиной 887 пн, расположенного с 15397-ой по 16283-ю пн мтДНК. В целом для двух популяций обнаружено 13 гаплотипов (табл. 1), различающихся между собой 14-ю нуклеотидными заменами (рис. 1). Частота вари-

Таблица 1. Показатели генетического разнообразия *Microtus rossiaemeridionalis* из двух инвазивных популяций юга Дальнего Востока

Исследуемые показатели	Популяции	
	Хабаровск	Советская Гавань
n	11	22
N	8	8
*V*s	25	8
*V*is	4	3
$k \pm S.E.$	2.11 ± 1.282	1.26 ± 0.826
$h \pm S.D.$	0.867 ± 0.099	0.654 ± 0.112
$\pi \pm S.D.$	0.0024 ± 0.0016	0.0014 ± 0.0004
Tajima's *D* (*P*)	−1.467(0.06)	−1.398(0.08)
Fu's *F*(*P*)	−10.253(0.000)	−28.447(0.000)
Показатели пространственной экспансии		
τ	1.55	0.14
θ	0.686	1.705
M	194.83	3942.16
SSD (*P*)	0.0075(0.70)	0.0187(0.38)
Показатели демографической экспансии		
τ	1.94	0.56
θ_0	0.3617	0.5625
θ_1	20.078	4257.205
SSD (*P*)	0.0072(0.75)	0.0186(0.24)

Примечание. *n* – объем выборки, *N* – число гаплотипов, *V*s – число вариабельных сайтов, *V*is – число сайтов, информативных с точки зрения максимальной экономии, *k* – среднее число попарных нуклеотидных различий, *h* – гаплотипическое разнообразие, π – нуклеотидное разнообразие, Tajima's *D* – коэффициент теста Таджимы; Fu's *F* – коэффициент теста Фу; τ – время экспансии в мутационных единицах, θ – мутационные параметры на начальном и конечном этапах роста численности, *M* – межпопуляционный обмен мигрантами, *SSD* – сумма квадратов отклонений между наблюдаемым и ожидаемым распределением попарных нуклеотидных различий, *S.D.* – стандартное отклонение *S.E.* – стандартная ошибка, *P* – уровень значимости.

абельных позиций составила 0.0158 от общей длины исследованного участка. Обнаружено 11 транзиций и три трансверсии. Все 13 гаплотипов выявлены для вида впервые. Все полученные гаплотипы внесены в GenBank/NCBI под номерами MW556643–MW556656. Только три гаплотипа H1, H3 и H2 были общими для семнадцати, трех и двух полевок соответственно. Самый распространенный гаплотип H1 (выявлен у каждой второй особи) отличался от гаплотипа из основной части ареала одной трансверсией.

Среднее число попарных нуклеотидных различий между гаплотипами в популяции из Хабаровска было несколько выше, чем в популяции полевок из г. Советская Гавань: 2.11 и 1.26 соответственно. Нуклеотидное и гаплотипическое разнообразие были минимальны в популяции из г. Советская Гавань, а максимальны в популяции из Хабаровска (табл. 1). Генетическая дистанция *p* между выборками полевок городов Советская Гавань и Хабаровск составила 0.0019 ± 0.0005. *F*-статистика Райта не выявила гетерогенности выборок восточноевропейской полевки из этих популяций $F_{st} = 0.0055$ ($P = 0.109$).

Построенная для определения филогенетических отношений медианная сеть гаплотипов имела структуру близкую к звездообразной (рис. 2). Центральный гаплотип H1 встречен с самой высокой частотой и был обнаружен во всех исследованных популяциях: Хабаровск – 0.4, Советская Гавань – 0.591. Гаплотипы не формировали кластеров в соответствии с их географической принадлежностью. Все гаплотипы отличалась друг от друга на 1–3 нуклеотидные замены. Это позволяет предположить, что гаплотипы H1–H13, так же как и гаплотип из основной части ареала, относятся к субкладе EU европейской линии.

Значение *D* Tajima для каждой популяции было отрицательным и не значимым. Результаты теста Фу (Fu's *F*s) на селективную нейтральность,

	2	5	3 0	4 2	5 4	5 9	6 6	7 6	4 6 4	6 2 5	6 7	7 0 6	7 3 9	7 6 9	8 2 3
				*									*	*	*
DQ 015676	C	G	T	T	A	T	C	A	T	A	C	T	A	A	A
H1	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.
H2	.	.	.	A	.	.	.	.	.	.	.	.	C	.	.
H3	.	.	.	A	.	.	.	.	.	.	T	.	.	.	.
H4	.	.	.	A	.	.	T	.	.	.	.	.	.	.	.
H5	.	.	.	A	.	.	.	.	.	.	.	.	.	.	T
H6	.	.	.	A	G	.	.	.	.	.	.	.	.	.	.
H7	.	A	.	A	G	.	.	.	.	.	.	.	.	C	.
H8	.	.	C	A	.	.	.	.	.	G	.	.	.	C	.
H9	.	.	C	A	.	.	.	.	.	.	.	.	C	C	.
H10	.	.	C	A	.	.	.	G	.	.	.	.	.	.	.
H11	.	.	C	A	.	C	.	.	.	.	T	.	.	.	.
H12	T	.	.	A	.	.	.	.	C	.	.	C	.	.	.
H13	T	.	.	A	.	.	.	.	.	.	.	.	.	.	.

Рис. 1. Гаплотипы контрольного региона восточноевропейской полевки из инвазивных популяций юга Дальнего Востока. Нуклеотидные замены представлены относительно последовательности варианта DQ015676 [30]. Сайты замен показаны от начала контрольного региона мтДНК. * – трансверсия.

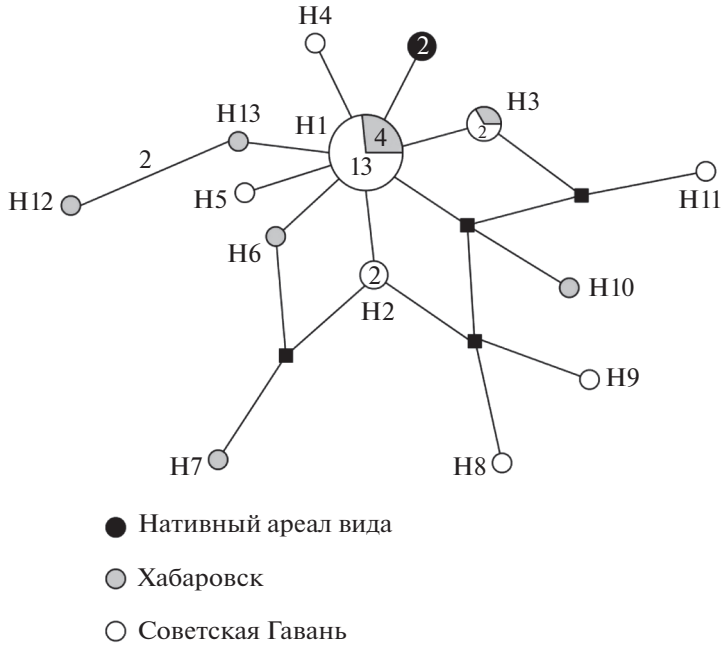

Рис. 2. Филогенетическая сеть гаплотипов мтДНК восточноевропейской полевки. Размеры кружков пропорциональны количеству образцов с данным гаплотипом. Сектором указана доля, а цифра в секторе – количество особей разных популяций. Цифры на ветвях сети соответствуют числу нуклеотидных замен больше двух.

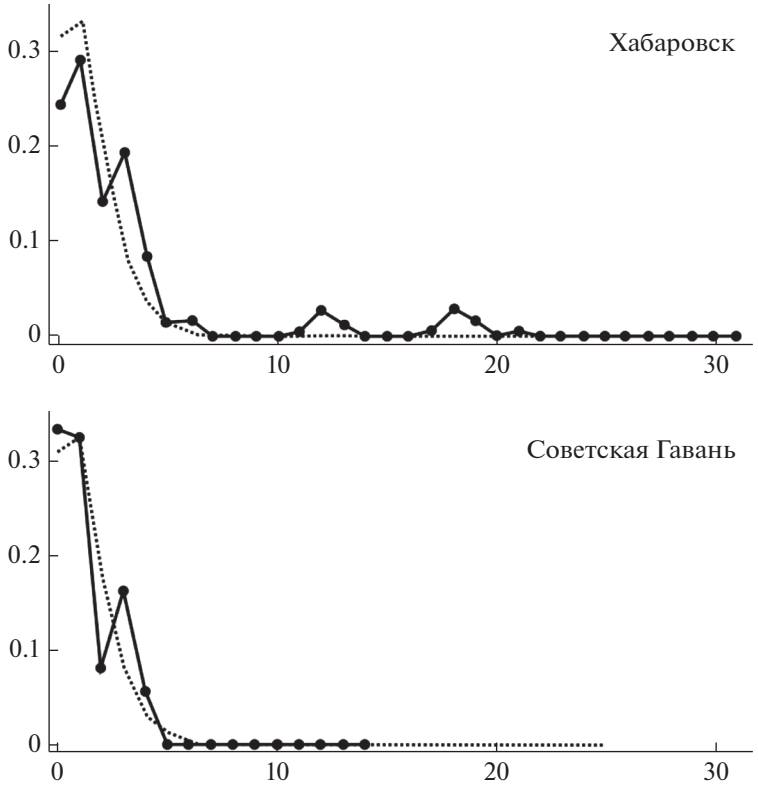

Рис. 3. Частота (ось ординат) попарных нуклеотидных различий (ось абсцисс) в популяциях восточноевропейской полевки с применением модели роста-уменьшения популяции. Наблюдаемая частота обозначена сплошной линией, ожидаемая – пунктирной.

были не только отрицательными, но и достоверно отличались от нуля (табл. 1). Распределение попарных нуклеотидных различий между гаплотипами полевок из популяции г. Советская Гавань имело унимодальную форму с модой в начале шкалы (рис. 3). Распределение попарных нуклеотидных различий между гаплотипами особей популяции из Хабаровска тоже можно считать одновершинным, но с некоторым допущением, поскольку в правой части распределения имеется небольшие дополнительные пики. Присутствие основного пика в области минимальных различий в популяциях из Советской Гавани и Хабаровска может являться следствием эффекта основателя и/или дрейфа генов.

Проверка выполняемости модели пространственной и демографической экспансии показала низкие и статистически недостоверные значения суммы квадратных отклонений *SSD* для двух популяций. Это, наряду с другими выше описанными параметрами, а также условными значениями возраста демографической и пространственной экспансии (табл. 1) свидетельствует в пользу гипотезы роста численности. Кроме того, условные значения возраста свидетельствуют в пользу большей молодости популяции г. Советская Гавань.

ОБСУЖДЕНИЕ

Генетические процессы, проходящие в инвазивных популяциях, аналогичны процессам, протекающим в малых изолированных популяциях. Эти популяции сталкиваются не только с явлением “эффект основателя” и с дрейфом генов, а также могут испытывать сильные колебания численности, что приводит к снижению генетического разнообразия и (как результат) снижает способность популяции сохраняться и адаптироваться к новой среде.

Полученные результаты продемонстрировали более низкое генетическое разнообразие в популяции восточноевропейской полевки из Советской Гавани по сравнению с популяцией из Хабаровска. При этом два гаплотипа (H1 и H3) были общими для особей обеих популяций, а остальные гаплотипы имели небольшое количество замен от гаплотипа H1. Учитывая тот факт, что проникновение вида в г. Советская Гавань проходило недавно, максимум 70 лет назад (время открытия регулярного движения поездов на участке железной дороги Комсомольск-на-Амуре–Ванино), можно предположить, что донором инвазивных полевок двух городов была одна популяция, находящаяся в пределах распространения субклады EU

европейской линии. Низкое генетическое разнообразие популяции г. Советская Гавань, может быть результатом "эффекта основателя" и стохастических процессов, проходящих в малых изолированных популяциях (дрейфа генов). Наблюдения за численностью этого вида в Советской Гавани показали резкие колебания численности в различные годы [29]. Так, с момента генетического подтверждения обитания восточноевропейской полевки в окрестностях и на территории Советской Гавани [18] численность вида, по нашим данным, в разные годы изменялась: 2012 – 2.3% попаданий на 100 ловушко-суток, 2013 – 1.3%, достигнув наибольшего значения в 2014 – 20.0% и снизившись в 2015 – 6.3%. Из материалов, документированных и хранящихся в архиве зоолого-паразитологической лаборатории Хабаровской противочумной станции можно проследить колебания численности *M. rossiaemeridionalis* и в более ранние годы. Если принимать во внимание, что до момента генетических исследований этот вид, несмотря на внешние морфологические отличия (более светлый окрас и меньшие размеры), рассматривали как дальневосточную полевку, то (согласно архивным данным) первая особь серой полевки в единственном экземпляре была добыта здесь в апреле 1991 г. Далее, с 2004 г., они стали появляться чаще. Ежегодные обследования г. Советская Гавань и его окрестностей, несмотря на постановку ловушек Геро в одних и тех же местах и в одно и то же время (конец сентября–начало октября), свидетельствовали о непостоянстве, периодичности попаданий зверьков. Так, заметное участие в отловах грызунов данный вид имел в 2004 (6.3% попаданий) и в 2008–2009 гг. (4.0–6.0%), тогда как в период 2005–2007 он не встречался.

Показатели генетического разнообразия для популяции восточноевропейской полевки Хабаровска достаточно высокие и незначительно снижены по сравнению с таковыми для аборигенных видов серых полевок (дальневосточная полевка и полевка Максимовича) этой территории [50]. Следует отметить, что филогеография этого вида на сегодняшний момент разработана недостаточно по сравнению с видом-двойником *M. arvalis* [51]. Наблюдения за численностью восточноевропейской полевки в Хабаровске показали более низкие значения, чем в г. Советская Гавань, при этом резких колебаний численности в этих популяциях отмечено не было. В период 2010–2012 гг. в отловах этот вид не встречался, тогда как в 2013 и 2015 гг. регистрировали менее 0.1%. Примерно средние значения наблюдали в 2014 (0.4%), 2016–2017 (0.1–0.5%) и 2019–2020 гг. (0.1–0.2%). Наибольший показатель отмечали в 2018 г. – 1.1%. При этом следует отметить, что начиная с 2017 г., вид начали регистрировать на большей территории города (ранее он в отловах встречался только на территориях прилежащих железнодорожной станции). Сделать заключение о состоянии популяции восточноевропейской полевки из Хабаровска в предшествующие годы (до 2010) сложно, поскольку, как было отмечено, до генетических исследований этот вид на данной территории принимали за дальневосточную полевку или полевку Максимовича. В Хабаровске оба эти аборигенные вида встречаются, при этом их молодых особей по внешним признакам сложно отличить от восточноевропейской полевки.

Принято считать, что малые изолированные популяции более подвержены быстрому падению численности и исчезновению за счет потери генетического разнообразия, особенно при отсутствии потока мигрантов из других популяций [7]. На территории юга Дальнего Востока восточноевропейская полевка в настоящее время обнаружена исключительно только в двух городах [18, 20], где может заселять парки, лесопарки, пустыри и обширные газоны, проникая в дома [52, 53]. При этом можно считать, что плотность популяций восточноевропейской полевки городов Хабаровск и Советская Гавань остается низкой, поскольку до сих пор эта полевка не встречена за пределами городов. В связи с этим можно предполагать, что обе данные популяции в настоящее время изолированы не только от популяций нативного ареала вида, но и друг от друга. Полученные в результате нашего исследования данные, указывающие на более низкое генетическое разнообразие популяции восточноевропейской полевки из Советской Гавани, а также отсутствие потока мигрантов, дают основание предположить, что эта популяция может быть подвержена риску исчезновения. Однако в литературе имеются сведения [54], что инвазивные виды, несмотря на негативные последствия снижения генетического разнообразия, часто способны все же успешно колонизировать новые территории и адаптироваться к новой среде.

Работа выполнена в рамках госзадания лаборатории эволюционной зоологии и генетики ФНЦ Биоразнообразия ДВО РАН № 0207-2021-0007 "Эволюционные аспекты формирования наземной биоты востока Азии".

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. *Hobbs R.J., Arico S., Aronson J. et al.* Novel ecosystems: Theoretical and management aspects of the new ecological world order // Global Ecol. and Biogeography. 2006. V. 15. P. 1–7. https://doi.org/10.1111/j.1466-822x.2006.00212.x

2. *Bellard C., Cassey P., Blackburn T.M.* Alien species as a driver of recent extinctions // Biol. Letters. 2016. V. 12. № 2. P. 20150623. https://doi.org/10.1098/rsbl.2015.0623
3. *Kareiva P., Marvier M.* What is conservation science? // BioScience. 2012. № 62. P. 962–969. https://doi.org/10.1525/bio.2012.62.11.5
4. *Bellard C., Thuiller W., Leroy B. et al.* Will climate change promote future invasions? // Global Change Biol. 2013. V. 19. № 12. P. 3740–3748. https://doi.org/10.1111/gcb.12344
5. *Elton C.S.* The Ecology of Invasions by Animals and Plants. Methuen, London, 1958. 159 p.
6. *Mayr E.* The Nature of Colonization in Birds // The Genetics of Colonizing Species / Eds Baker H.G., Stebbins G.L. N.Y.: Academic Press, 1965. P. 29–47.
7. *Примак Р.* Основы сохранения биоразнообразия. М.: Издательство Научного и учебно-методического центра, 2002. 256 с.
8. *Lee C.E.* Evolutionary genetics of invasive species // Trends Ecol. Evol. 2002. V. 17. P. 386–391. https://doi.org/10.1016/S0169-5347(02)02554-5
9. *Mahmoudi A., Darvish J., Aliabadian M. et al.* New insight into the cradle of the grey voles (subgenus *Microtus*) inferred from mitochondrial cytochrome b sequences // Mammalia. 2017. V. 81. № 6. P. 583–593. https://doi.org/10.1515/mammalia-2016-0001
10. *Holicová T., Sedláček F., Mácová A. et al.* New record of *Microtus mystacinus* in eastern Kazakhstan: Phylogeographical considerations // ZooKeys. 2018. V. 781. P. 67–80. https://doi.org/10.3897/zookeys.781.25359
11. *Малыгин В.М., Баскевич М.И., Хляп Л.А.* Инвазии видов-двойников обыкновенной полевки // Росс. журн. биол. инвазий. 2019. № 4. С. 71–93.
12. *Ковальская Ю.М., Малыгин В.М.* Восточноевропейская полевка (*Microtus rossiaemeridionalis*) в Сибири // Науч. докл. высшей шк. Биол. науки. 1985. № 1. С. 49–51.
13. *Мейер М.Н., Голенищев Ф.Н., Раджабли С.И., Саблина О.В.* Серые полевки фауны России и сопредельных территорий. СПб.: Изд-во ЗИН РАН, 1996. 319 с.
14. *Якименко Л.В., Крюков А.П.* Об изменчивости кариотипа восточноевропейской полевки *Microtus rossiaemeridionalis* (Rodentia, Cricetidae) // Зоол. журн. 1997. Т. 76. № 3. С. 375–378.
15. *Демидович А.П., Липин С.И.* Особенности биологии обыкновенных полевок в Иркутской области // Вестник Иркутской гос. с.-хоз. академии. 1997. № 3. С. 1–25.
16. *Гашев С.Н.* Восточноевропейская полевка – новый вид в списке млекопитающих Тюменской области // Ежегодник Тюменского областного краеведческого музея. Тюмень: ТОКМ, 1998. С. 161–165.
17. *Ильинских Н.Н., Москвитина Н.С., Сучкова Н.Г. и др.* Цитогенетическая нестабильность полевок и проблема сохранения природно-очаговых инфекций в городских местообитаниях Томска // Фунд. науки и практика. 2010. Т. 1. № 3. С. 53–54.
18. *Картавцева И.В., Тиунов М.П., Лапин А.С. и др.* Инвазия полевки *Microtus rossiaemeridionalis* на территорию Дальнего Востока России // Росс. журн. биол. инвазий. 2011. № 4. С. 17–24.
19. *Pavlova S.V., Tchabovsky A.V.* Presence of the 54-chromosome common vole (Mammalia) on Olkhon Island (Lake Baikal, East Siberia, Russia), and the occurrence of an unusual X-chromosome variant // Comp. Cytogenet. 2011. V. 5. № 5. P. 433–440. https://doi.org/10.3897/CompCytogen.v5i5.1720
20. *Лапин А.С.* Мелкие млекопитающие южной части Хабаровского края и Еврейской автономной области (фауна, экология, эпизоотологическое значение): Автореф. дис. ... канд. биол. наук. Хабаровск: Изд-во ДВГГУ, 2013. 24 с.
21. *Маркова Е.А., Стариков В.П., Ялковская Л.Э. и др.* Молекулярные и цитогенетические данные о находке восточноевропейской полевки *Microtus rossiaemeridionalis* (Arvicolinae, Rodentia) на севере Западной Сибири // Доклады Академии наук. 2014. Т. 455. № 5. С. 603–605.
22. *Моролдоев И.В., Шереметьева И.Н., Картавцева И.В.* Первая находка восточноевропейской полевки (*Microtus rossiaemeridionalis*) в Бурятии // Росс. журн. биол. инвазий. 2017. № 2. С. 88–94.
23. *Моролдоев И.В.* Новые сведения о распространении млекопитающих в Бурятии. Млекопитающие России: фаунистика и вопросы териогеографии. Материалы Международного совещания (Ростов-на-Дону 17–19 апреля 2019 г.). М.: Товарищество научных изданий КМК, 2019. С. 178–180.
24. *Fredga K., Jaarola M., Ims R.A., Steen H., Yoccoz N.G.* The 'common vole' in Svalbard identified as *Microtus epiroticus* by chromosomal analysis // Polar Research. 1990. V. 8. P. 283–290. https://doi.org/10.3402/polar.v8i2.6818
25. *Frafjord K.* Predation on an introduced vole *Microtus rossiaemeridionalis* by arctic fox *Alopex lagopus* on Svalbard // Wildlife Biol. 2002. V. 8. № 1. P. 41–47. https://doi.org/10.2981/wlb.2002.006
26. *Карасёва Е.В., Степанова Н.В., Телицына А.Ю. и др.* Экологические различия двух близких видов – обыкновенной и восточноевропейской полевок // Синантропия грызунов. М.: Наука, 1994. С. 60–76.
27. *Тихонова Г.Н., Тихонов И.А., Федорович Е.Ю., Давыдович Л.В.* Сравнительный анализ ориентировочно-исследовательского поведения видов-двойников *Microtus arvalis* и *M. rossiaemeridionalis* (Rodentia, Cricetidae) в связи с разной склонностью к синантропии // Зоол. журн. 2005. Т. 84. № 5. С. 618–627.
28. Самые опасные инвазионные виды России (ТОП-100) / Ред. Дгебуадзе Ю.Ю., Петросян В.Г., Хляп Л.А. М.: Т-во научных изданий КМК, 2018. 688 с.
29. *Лапин А.С., Высочина Н.П., Здановская Н.И. и др.* Экология и эпидемиологическое значение восточноевропейской полевки на юге Хабаровского края // Дальневост. журн. инф. патологии. 2015. № 27. С. 53–56.
30. *Ahlroth P., Alatalo R.V., Holopainen A. et al.* Founder population size and number of source populations enhance colonization success in waterstriders // Oecologia. 2003. V. 137. № 4. P. 617–620. https://doi.org/10.1007/s00442-003-1344-y

31. *Банникова А.А.* Молекулярные маркеры и современная филогенетика млекопитающих // Журн. общей биологии. 2004. Т. 65. № 4. С. 278–305.

32. *Haring E., Herzig Straschil B., Spitzenberger F.* Phylogenetic analysis of Alpine voles of the *Microtus* multiplex complex using the mitochondrial control region // J. Zool. Syst. and Evol. Res. 2000. V. 38. № 4. P. 231–238.

33. *Tougard C., Renvoise E., Petitjean A., Quere J.-P.* New insight into the colonization processes of common voles: Inferences from molecular and fossil evidence // PLoS One. 2008. V. 3. № 10. P. e3532. https://doi.org/10.1371/journal.pone.0003532

34. *Сибиряков П.А., Товпинец Н.Н., Дупал Т.А. и др.* Филогеография обыкновенной полевки *Microtus arvalis* (Rodentia, Arvicolinae) формы obscurus: новые данные по изменчивости митохондриальной ДНК // Генетика. 2018. Т. 54. № 10. С. 1162–1176.

35. *Li K., Kohn M.H., Zhang S. et al.* The colonization and divergence patterns of Brandt's vole (*Lasiopodomys brandtii*) populations reveal evidence of genetic surfing // BMC Evol. Biol. 2017. V. 17. № 145. P. 2–17. https://doi.org/10.1186/s12862-017-0995-y

36. *Gao J., Yue L., Jiang X. et al.* Phylogeographic patterns of *Microtus fortis* (Arvicolinae: Rodentia) in China Based on Mitochondrial DNA Sequences // Pakistan J. Zool. 2017. V. 49. P. 1185–1195. https://doi.org/10.17582/journal.pjz/2017.49.4.1185.1195

37. *Pergams O.R.W., Lacy R.C.* Rapid morphological and genetic change in Chicago-area *Peromyscus* // Mol. Ecol. 2008. V. 17. P. 450–463 https://doi.org/10.1111/j.1365-294X.2007.03517.x

38. *Звычайная Е.Ю., Данилкин А.А., Холодова М.В. и др.* Анализ изменчивости контрольного региона и гена цитохрома *b* мтДНК сибирской косули *Capreolus pygargus* Pall. // Изв. РАН. Серия биологическая. 2011. № 5. С. 511–517.

39. *Мальцев А.Н., Стахеев В.В., Богданов А.С. и др.* Филогенетические взаимоотношения внутривидовых форм домовой мыши *Mus musculus*: анализ изменчивости контрольного региона (D-петли) митохондриальной ДНК // Доклады Академии наук. 2015. Т. 465. № 3. С. 380–383.

40. *Баранова А.И., Холодова М.В., Давыдов А.В., Рожков Ю.И.* Полиморфизм контрольного региона мтДНК диких северных оленей европейской части России *Rangifer tarandus* (Mammalia: Artiodactyla) // Гентика. 2012. Т. 48. № 9. С. 1098–1104.

41. *Triant D.A., Dewoody J.A.* Accelerated molecular evolution in *Microtus* (Rodentia) as assessed via complete mitochondrial genome sequences // Genetica. 2006. № 128. P. 95–108. https://doi.org/10.1007/s10709-005-5538-6

42. *Triant D.A., DeWoody J.A.* Extensive mitochondrial DNA transfer in a rapidly evolving rodent has been mediated by independent insertion events and by duplications // Gene. 2007. № 401. P. 61–70. https://doi.org/10.1016/j.gene.2007.07.003

43. *Aljanabi S.M., Martinez I.* Universal and rapid salt extraction of high quality genomic DNA for PCR-based techniques // Nucl. Ac. Res. 1997. V. 25. № 22. P. 4692–4693. https://doi.org/10.1093/nar/25.22.4692

44. *Шереметьева И.Н., Картавцева И.В., Фрисман Л.В. и др.* Полиморфизм и генетическая структура полевки Максимовича *Microtus maximowiczii* (Schrenck 1858) (Rodentia, Cricetidae) Среднего Приамурья по данным секвенирования контрольного региона мтДНК // Генетика. 2015. Т. 51. № 10. С. 1154–1162.

45. *Haring E., Sheremetyeva I., Kryukov A.* Phylogeny of Palearctic vole species (genus Microtus, Rodentia) based on mitochondrial sequences // Mammalian Biol. 2011. V. 76. P. 258–267. https://doi.org/10.1016/j.mambio.2010.04.006

46. *Hall T.A.* BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98 // Nucl. Ac. Symp. Series. 1999. V. 41. № 41. P. 95–98. https://doi.org/10.1021/bk-1999-0734.ch008

47. *Bandelt H.J., Forster P., Röhl A.* Median-Joining networks for inferring intraspecific phylogenies // Mol. Biol. and Evol. 1999. V. 16. № 1. P. 37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

48. *Rozas J., Ferrer-Mata A., Sánchez-DelBarrio J.C. et al.* DnaSP 6: DNA sequence polymorphism analysis of large datasets // Mol. Biol. Evol. 2017. V. 34. P. 3299–3302. https://doi.org/10.1093/molbev/msx248

49. *Excoffier L., Laval G., Schneider S.* Arlequin ver. 3.0: An integrated software package for population genetics data analysis // Evol. Bioinformatics Online. 2005 V. 1. P. 47–50. https://doi.org/10.1177/117693430500100003

50. *Шереметьева И.Н.* Причины повышения генетического разнообразия в популяциях серых полевок рода *Alexandromys* на территории Среднего Приамурья // Совр. пробл. регион. развития [электронный ресурс]: Матер. VII Всеросс. науч. конф. (Биробиджан 9–11 октября 2018 г.) / Под ред. Фрисмана Е.Я. Биробиджан: ИКАРП ДВО РАН, 2018. С. 214–216. https://doi.org/10.31433/978-5-904121-22-8-2018-214-216

51. *Cetinturk D., Yigit N., Colak E. et al.* Inferring phylogenetic relationships in the common vole (*Microtus arvalis*) based on mitochondrial and nuclear sequence diversities // Turkish J. Zoology. 2021. V. 45. P. 117–130. https://doi.org/10.3906/zoo-2008-3

52. *Малыгин В.М.* Систематика обыкновенных полевок. М.: Наука, 1983. 208 с.

53. *Малыгин В.М., Рябов С.В.* Распространение и биология видов-двойников обыкновенной полевки в зоомедицинском аспекте // Дезинфекционное дело. 2013. № 4. С. 27–35.

54. *Uller T., Leimu R.* Founder events predict changes in genetic diversity during human-mediated range expansions // Glob. Change Biol. 2011. V. 17. P. 3478–3485. https://doi.org/10.1111/j.1365-2486.2011.02509.x

The mtDNA Control Region Variability of *Microtus rossiaemeridionalis* (Rodentia, Arvicolini) from Two Invasive Populations of the Russian Far East

I. N. Sheremetyeva[a], *, I. V. Kartavtseva[a], A. A. Emelyanova[a], and A. S. Lapin[b]

[a] *Federal Scientific Center of the East Asia Terrestrial Biodiversity Fareastern Branch, Russian Acalemy of Sciences, Vladivostok, 690022 Russia*

[b] *Khabarovsk Plague Control Station, Khabarovsk, 680031 Russia*

**e-mail: sheremet76@yandex.ru*

The mtDNA control region variability of *Microtus rossiaemeridionalis* (Rodentia, Arvicolini) from two invasive populations the Russian Far East was studied. The results demonstrated a lower genetic diversity in the Eastern European vole population from Sovetskaya Gavan than in the population from the city of Khabarovsk. Wright's *F* statistic did not reveal any heterogeneity in the samples from these populations. It was suggested that the discovered haplotypes belong to the EU subclade of the European lineage.

Keywords: invasion, variability, mtDNA control region, East European vole, *Microtus rossiaemeridionalis*.

ГЕНЕТИКА, 2021, том 57, № 11, с. 1261–1275

ГЕНЕТИКА ЧЕЛОВЕКА

УДК 575.22

ОСОБЕННОСТИ ГЕНОМНОГО РАСПРЕДЕЛЕНИЯ РЕГИОНОВ ВЫСОКОЙ ГОМОЗИГОТНОСТИ У КОРЕННОГО НАСЕЛЕНИЯ СЕВЕРНОЙ ЕВРАЗИИ НА ИНДИВИДУАЛЬНОМ И ПОПУЛЯЦИОННОМ УРОВНЯХ НА ОСНОВЕ АНАЛИЗА SNP ВЫСОКОЙ ПЛОТНОСТИ

© 2021 г. Н. А. Колесников[1, *, **], В. Н. Харьков[1], А. А. Зарубин[1], М. О. Раджабов[2], М. И. Воевода[3], М. А. Губина[4], Э. К. Хуснутдинова[5], С. С. Литвинов[5], Н. В. Екомасова[5], О. В. Штыгашева[6], Н. Р. Максимова[7], А. Л. Сухомясова[7], В. А. Степанов[1]

[1]Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, 634050 Россия

[2]Институт физики им. Х.И. Амирханова, Дагестанский федеральный исследовательский центр Российской академии наук, Махачкала, 367012 Россия

[3]Федеральный исследовательский центр фундаментальной и трансляционной медицины, Новосибирск, 630060 Россия

[4]Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, 630090 Россия

[5]Институт биохимии и генетики, Уфимский федеральный исследовательский центр Российской академии наук, Уфа, 450054 Россия

[6]Хакасский государственный университет им. Н.Ф. Катанова, Абакан, 655000 Россия

[7]Северо-Восточный федеральный университет им. М.К. Аммосова, Якутск, 677000 Россия

**e-mail: nikita.kolesnikov@medgenetics.ru*

***e-mail: nik.fleming@mail.ru*

Поступила в редакцию 29.01.2021 г.
После доработки 11.05.2021 г.
Принята к публикации 20.05.2021 г.

Анализ геномных данных по 886 889 аутосомных SNP позволил оценить уровень инбридинга на основе исследования регионов высокой гомозиготности (F_{ROH}) в выборке из 1836 человек, включающей 76 популяций коренного населения Восточной Европы, Кавказа, Волго-Уральского региона, Сибири и Средней Азии. Данные были получены с помощью биочипов Infinium Multi-Ethnic Global-8 Kit. Среди популяций Дагестана, внутри нахско-дагестанской языковой семьи, можно выделить представителей языковых групп с максимальным уровнем F_{ROH}: дидойская (0.0727) и андийская (0.0378) группы. Последний показатель близок к таковому в популяциях, говорящих на чукотско-камчатском и нивхском языках Сибири (0.0360). Для популяций Сибири характерно наибольшее значение суммарной длины коротких и средних регионов высокой гомозиготности на человека, по сравнению с другими исследованными популяциями. Популяции коренных народов Дагестана и Сибири в целом характеризуются большей степенью гомозиготизации генома и большей вариабельностью регионов с повышенной гомозиготностью по сравнению с другими популяциями Кавказа, Средней Азии, Европы и Волго-Уральского региона.

Ключевые слова: популяционная генетика, регионы высокой гомозиготности, популяции человека, инбридинг, Северная Евразия.

DOI: 10.31857/S0016675821110059

Существенная часть любого индивидуального генома человека представлена регионами высокой гомозиготности (англ. ROH – runs of homozygosity). Общая длина и распределение таких участков связаны с генетико-демографической историей популяции и отражают эффекты основателя, уровень инбридинга, а также могут быть следствием действия сильного направленного отбора [1, 2]. Современные методы анализа данных полных геномов или высокоплотных наборов SNP позволяют детально выявлять структуру и распределение ROH в индивидуальных геномах и популяциях с целью уточнения генетической истории популяций, оценки уровня инбридинга, поиска сигналов естественного отбора [1, 3]. Данные по распределению ROH в популяциях могут служить основой для идентификации участков генома, связанных с заболеваниями, в контексте картирования гомозиготности [4]. Общая длина и структура ROH характеризуются существенной межиндивиду-

альной и межпопуляционной вариабельностью. Как правило, сумма длин коротких и средних ROH у индивида возрастает от восточно-африканских к европейским и азиатским популяциям в силу накопления рекомбинационных событий. Напротив, общий суммарный размер длинных ROH показывает большие межиндивидуальные вариации, отражая, вероятно, эффекты недавнего инбридинга [5, 6]. Кроме того, показано, что частота распределения ROH в геноме коррелирует со скоростью рекомбинации, а также с сигналами недавнего положительного отбора [7].

Предыдущие исследования показали, что у неродственных людей широко распространены короткие ROH, измеряемые несколькими десятками тысяч пар нуклеотидов (тпн) [8]. Регионы высокой гомозиготности средних размеров – от сотен тысяч до нескольких миллионов пн – являются, как правило, результатом недавнего общего происхождения индивидов внутри популяции. Длинные ROH, измеряющиеся десятками миллионов пн, характерны для популяций с высоким уровнем близкородственных браков [5].

Среди многочисленных групп населения Северной Евразии, имеющих различную генетико-демографическую историю, особый интерес для исследования ROH представляют популяции коренных народов Северного Кавказа и Сибири, характеризующиеся, как правило, небольшой численностью и значительной географической изоляцией от других популяций. В частности, ранее для популяций Дагестана, говорящих на языках нахско-дагестанской семьи, были показаны высокие коэффициенты инбридинга, очень большое количество и большая суммарная длина гомозиготных блоков, по сравнению с окружающими популяционными группами Кавказа, Ближнего Востока, Европы, Центральной и Южной Азии [9]. Для популяций Сибири мы ранее на низкоплотных чипах обнаружили значительно большее число и суммарную длину ROH по сравнению с популяциями Волго-Уральского региона, Кавказа, Европы и Средней Азии [10].

Настоящая работа посвящена оценке уровня инбридинга, структуры и вариабельности ROH в популяциях Северной Евразии на основе анализа SNP высокой плотности.

МАТЕРИАЛЫ И МЕТОДЫ

Данные были получены с помощью биочипов “Infinium Multi-Ethnic Global-8 Kit”, содержащих 1779819 SNP, в выборке из 1836 человек. Популяции коренного населения Сибири ($N = 477$) представлены алтайцами (Б – с. Бешпельтир Чемальского муниципального района, $N = 24$ и К – с. Кулада Онгудайского района, $N = 25$), бурятами (А – п. Агинское Агинского района, $N = 23$ и К – с. Курумкан Курумканского района, $N = 28$), калмыками ($N = 29$), кетами ($N = 15$), коряками ($N = 20$), нивхами ($N = 13$), татарами (Т – г. Томск, $N = 20$), тувинцами ($N = 28$), удэгейцами ($N = 15$), хантами (К – с. Казым Белоярского района, $N = 30$ и Р – д. Русскинская Сургутского района, $N = 26$), хакасами (Т – сагайцами Таштыпского р-на, $N = 29$ и Ш – качинцами Ширинского р-на, $N = 26$), чукчами ($N = 25$), чулымцами ($N = 22$), эвенками (З – забайкальские (п. Чара Каларского района, с. Моклан и с. Тупик Тунгиро-Олёкминского района), $N = 25$ и Я – якутские, $N = 28$) и якутами ($N = 26$).

Коренные популяции Дагестана ($N = 518$) включают аварцев ($N = 24$), даргинцев ($N = 28$), лакцев ($N = 24$), табасаран ($N = 21$), лезгин ($N = 28$), агулов ($N = 24$), рутульцев ($N = 22$), цахуров ($N = 24$), арчинцев ($N = 24$), андийцев ($N = 17$), ахвахцев ($N = 24$), тиндинцев ($N = 18$), каратинцев ($N = 24$), багулалов ($N = 23$), ботлихцев ($N = 16$), хваршин ($N = 15$), цезов ($N = 24$), бежтинцев ($N = 22$), чамалинцев ($N = 24$), гунзибцев ($N = 17$), гинухцев ($N = 19$), кумыков ($N = 37$) и караногайцев ($N = 19$).

Волго-Уральский географический регион и популяции Европы ($N = 419$) представлены коми ($N = 30$), марийцами ($N = 30$), вепсами ($N = 30$), удмуртами (удмурты $N = 30$, Б – п. Балезино, $N = 28$ и Ш– с. Шаркан, $N = 18$),чувашами ($N = 26$), башкирами (Б – Бурзянский р-н, $N = 34$, П – пермские, $N = 15$ и С – Салаватский р-н, $N = 15$), бесермянами ($N = 16$), карелами ($N = 29$), мордвой (Э – эрзя, $N = 16$, М – мокша, $N = 30$, Ш – шокша, $N = 14$), татарами (К– казанские, $N = 33$) и русскими ($N = 33$).

Коренные популяции Кавказского региона ($N = 294$) включают абхазов ($N = 30$), адыгейцев ($N = 10$), балкарцев ($N = 50$), ингушей ($N = 30$), карачаевцев ($N = 22$), мегрелов ($N = 28$), ногайцев ($N = 20$), осетин (С – северные, $N = 30$, Ю – южные, $N = 17$), черкесов ($N = 30$) и чеченцев ($N = 27$). Популяции северных киргизов ($N = 28$), узбеков ($N = 22$), казахов (М – младший жуз, $N = 29$, С – старший жуз, $N = 26$) и дунган ($N = 23$) относятся к популяциям Средней Азии ($N = 128$).

При биоинформатической обработке полученных данных из анализа были исключены SNP с более чем 10% отсутствующих генотипов. Были исключены все позиции делеций и инсерций. Данные были предварительно отфильтрованы по минимальной частоте редкого аллеля (MAF, minor allele frequency > 0.01). В итоге после фильтрации исходных данных по 1779819 маркерам в финальный массив данных вошло 886889 аутосомных SNP. Кроме этого, были исключены образцы, имеющие более 5% пропущенных SNP, а также метисы первого, второго и третьего поколений. Для нахождения участков, которые с большей вероятностью представляют аутозиготность, перед поиском ROH в программе PLINK v1.9 [11] мы также удалили SNP с высоким неравновесием по сцеплению (LD) (англ. LD – linkage disequilibrium). Для этого использовали сканирование генома в рамке из 50 SNP со скользящим окном 5 SNP и

порогом попарного коэффициента сцепления (r^2 порог = 0.8). В результате контроль прошли 585715 SNP. Для анализа ROH в PLINK указали следующие параметры: –homozyg-snp 50, –homozyg-kb 300. Все остальные параметры использовались по умолчанию. Анализ был выполнен с учетом трех различных наборов ROH с учетом минимальной длины 0.5, 1.5 и 5 млн пн (малые, средние и большие ROH).

Для расчета коэффициента инбридинга F_{ROH} у каждого индивидуума размеры аутосомных областей ROH, превышающих 1.5 млн пн, суммировали и затем делили на общую длину аутосом, за исключением центромер (2708.613716 млн пн для сборки генома hg19). Как показано ранее, F_{ROH} для 1.5 млн пн лучше коррелирует с оценками инбридинга, полученными из анализа родословных [1].

Для оценки возраста ROH применяли программу GARLIC [2], использующую метод, основанный на вероятностной модели для ROH в классах с “коротким” (A, до 0.5 млн пн), “средним” (B, 0.5–1.5 млн пн) и “длинным” (C, более 1.5 млн пн) размерами. Короткие ROH отражают гомозиготность по древним гаплотипам, средние ROH являются результатом недавних демографических изменений в популяции (например, эффект “бутылочного горлышка”) и длинные ROH отражают гомозиготность возникшую в результате недавних кровнородственных браков. Для анализа ROH в GARLIC указали следующие параметры: *–build hg19*, *–error 0.001*, *–winsize 40*, *–auto-winsize*, *–auto-winsize-step 2*, *–kde-subsample 0*, *–ld-subsample 0*, *–resample 40*.

Для межпопуляционных сравнений были рассчитаны средние значения количества и совокупной длины ROH на геном для каждой популяции. Статистический анализ проводился в программной среде R. Статистическую значимость межпопуляционных различий оценивали при помощи теста Вилкоксона. Для коррекции уровня значимости по причине множественных сравнений использовалась поправка Бонферрони (уровень значимости = 0.05).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Коэффициент инбридинга F_{ROH}

Коэффициенты инбридинга F_{ROH} для всех индивидов из 76 популяционных выборок оценивали отдельно и группировали по географическим регионам и языковым группам для трех классов ROH (длина больше трех пороговых значений 0.5, 1.5 и 5 млн пн). Коэффициенты F_{ROH} для трех диапазонов длин ROH, среднее суммарное количество и средняя суммарная длина гомозиготных участков в исследованных популяциях представлены в табл. 1. Примечательно, что ROH > 1.5 млн пн не было выявлено лишь у четырех индивидов (кумыка, лезгина, эвенка и узбека).

Максимальные значения по $F_{ROH > 1.5}$ (0.0376) и $F_{ROH > 5}$ (0.0239) наблюдаются в популяциях Дагестана (табл. 1). Наибольшим коэффициентом геномного инбридинга для длин ROH > 1.5 млн пн обладают популяции гинухцев и гунзибцев (F_{ROH} = = 0.1039, F_{ROH} = 0.0887). Этот показатель для гинухцев является одним из самых высоких в мире и схож лишь с уровнем инбридинга в изолированной популяции индейцев каритиана из бразильской Амазонии с общей численностью в 320 человек ($F_{ROH > 1.5}$ = 0.1067) [3]. Также по уровню гомозиготизации генома значительно выделяется и ряд других малочисленных народов Дагестана – бежтинцы, ахвахцы и цезы (F_{ROH} 0.0689, 0.0678 и 0.0571 соответственно). Эти результаты хорошо соотносятся с предыдущими данными по популяциям Дагестана [9], где для народов нахско-дагестанской языковой семьи были выявлены более высокие коэффициенты инбридинга по сравнению с популяциями из других географических регионов и не нахско-дагестанскими популяциями Дагестана.

Внутри нахско-дагестанской языковой семьи по F_{ROH} этносы разделяются на три группы, в полном соответствии с их языковой классификацией. Максимальный средний уровень F_{ROH} выявлен для популяций дидойской (цезской) группы (0.0727), проживающих в высокогорных регионах в течение сотен поколений в относительной изоляции и имеющих также значительно более низкий уровень разнообразия по гаплогруппам Y-хромосомы, по сравнению с другими северокавказскими популяциями [12]. Для андийской группы показано гораздо меньшее значение (0.0378), находящееся на одном уровне с выборками чукчей, коряков и нивхов (0.0360). Наименьший коэффициент инбридинга выявлен для лезгинской группы (0.0193), занимающей в основном низменные территории и предгорья южной части Дагестана и отличающейся гораздо большей по сравнению с дидойской и андийской группами численностью представляющих ее этносов. Эти результаты полностью согласуются с данными об очень высокой частоте инбредных браков среди малочисленных и более изолированных популяций высокогорных районов по сравнению с равнинными [13].

Среди коренного населения Сибири максимальные значение $F_{ROH > 1.5}$ обнаружены у дальневосточных коряков, чукчей и нивхов (0.0422, 0.0409, 0.0363), что вполне закономерно в связи с их относительной малочисленностью и территориальной изолированностью. Высокое значение показано и для выборки хакасов-сагайцев из предгорного Таштыпского района (0.0318), вдвое превосходящее хакасов-качинцев равнинного Ширинского района. Минимальное значение характерно для томских татар (0.0071). Для остальных популяций наблюдается разброс $F_{ROH > 1.5}$ от 0.0112 до 0.0292. Несмотря на то что F_{ROH} в популяциях Дагестана значительно превышает по средним и длинным

Таблица 1. Значения коэффициентов F_{ROH} для трех диапазонов длин ROH, среднее суммарное количество (A, B, C) и средняя суммарная длина (a, b, c) для трех классов длин исследуемых популяций

Популяция	Регион	Группа	PLINK						GARLIC							
			F_{ROH}			F_{ROH}			NROH			SROH			Σ	
			0.5	1.5	5	0.5	1.5	5	A	B	C	a	b	c	NROH	SROH
Башкиры (Б)	В	1	0.0438	0.0122	0.0043	0.0388	0.0094	0.0035	544	189	14	95	95	31	746	221
Башкиры (П)	В	1				0.0450	0.0147	0.0065	561	188	17	95	95	45	765	235
Башкиры (С)	В	1				0.0368	0.0083	0.0034	439	157	10	69	74	24	606	167
Татары (К)	В	1				0.0334	0.0061	0.0023	575	192	9	100	97	20	776	216
Чуваши	В	1				0.0406	0.0103	0.0032	443	136	15	79	70	31	594	180
Бесермяне	В	2	0.0498	0.0169	0.0063	0.0509	0.0199	0.0105	678	202	16	118	110	55	896	283
Коми	В	2				0.0468	0.0164	0.0084	453	111	13	84	63	43	577	191
Удмурты	В	2				0.0504	0.0158	0.0047	476	143	19	89	81	46	637	216
Удмурты (В)	В	2				0.0494	0.0155	0.0037	576	197	24	107	107	52	796	266
Удмурты (Ш)	В	2				0.0514	0.0171	0.0039	574	205	28	98	105	57	806	259
Вепсы	В	3	0.0445	0.0107	0.0027	0.0458	0.0126	0.0031	432	145	18	76	75	38	594	188
Карелы	В	3				0.0432	0.0089	0.0024	453	153	16	80	78	31	622	189
Марийцы	В	4	0.0437	0.0117	0.0040	0.0496	0.0162	0.0044	450	159	23	81	83	51	632	215
Мордва (М)	В	4				0.0396	0.0082	0.0026	595	179	10	109	97	24	784	231
Мордва (Ш)	В	4				0.0460	0.0141	0.0067	635	189	13	109	98	39	837	246
Мордва (Э)	В	4				0.0396	0.0083	0.0026	596	204	16	100	99	30	816	229
Русские	В	5	0.0377	0.0061	0.0019	0.0377	0.0061	0.0019	607	189	8	111	102	19	804	231
Агулы	Д	6	0.0490	0.0193	0.0119	0.0464	0.0171	0.0104	617	163	11	114	93	46	790	253
Арчинцы	Д	6				0.0782	0.0459	0.0315	547	161	25	106	94	123	733	323
Лезгины	Д	6				0.0412	0.0124	0.0076	568	156	8	105	88	33	733	226
Рутульцы	Д	6				0.0433	0.0131	0.0072	445	131	10	78	70	34	586	182
Табасараны	Д	6				0.0350	0.0066	0.0028	573	196	9	101	99	20	778	220
Цахуры	Д	6				0.0499	0.0206	0.0120	463	132	14	81	70	54	609	206
Андийцы	Д	7	0.0696	0.0378	0.0240	0.0578	0.0279	0.0185	577	167	17	101	90	76	761	266
Ахвахцы	Д	7				0.1020	0.0678	0.0430	424	114	35	90	79	174	572	343
Багулалы	Д	7				0.0721	0.0400	0.0253	621	171	22	119	103	105	814	328
Ботлихцы	Д	7				0.0512	0.0196	0.0102	673	230	20	114	117	59	924	290
Каратинцы	Д	7				0.0775	0.0460	0.0321	547	164	24	101	93	122	735	316
Тиндинцы	Д	7				0.0569	0.0253	0.0139	536	164	16	95	91	66	715	253
Чамалалы	Д	7				0.0697	0.0377	0.0246	501	151	23	91	82	102	675	276

Таблица 1. Продолжение

Популяция	Регион	Группа	PLINK						GARLIC							
			F_{ROH}			F_{ROH}			NROH			SROH			Σ	
			0.5	1.5	5	0.5	1.5	5	A	B	C	a	b	c	NROH	SROH
Бежтинцы	Д	8	0.1126	0.0727	0.0460	0.1105	0.0689	0.0417	423	113	28	101	97	158	563	356
Гинухцы	Д	8				0.1437	0.1039	0.0733	259	71	27	80	85	226	357	391
Гунзибцы	Д	8				0.1305	0.0887	0.0549	318	98	33	85	95	204	449	383
Хваршины	Д	8				0.0773	0.0448	0.0309	594	183	24	108	102	121	801	331
Цезы	Д	8				0.1011	0.0571	0.0293	492	141	31	109	108	137	664	354
Караногайцы	Д	9	0.0318	0.0037	0.0010	0.0324	0.0039	0.0011	582	198	13	99	94	20	793	213
Кумыки	Д	9				0.0312	0.0036	0.0009	548	183	9	100	95	16	739	210
Даргинцы	Д	10	0.0696	0.0369	0.0243	0.0696	0.0369	0.0243	594	143	19	114	87	97	756	297
Лакцы	Д	11	0.0666	0.0347	0.0230	0.0666	0.0347	0.0230	587	179	20	107	100	93	786	300
Аварцы	Д	12	0.0761	0.0442	0.0320	0.0761	0.0442	0.0320	589	178	24	109	100	119	791	328
Абхазы	К	13	0.0335	0.0044	0.0006	0.0350	0.0047	0.0004	519	224	22	86	101	28	764	215
Адыгейцы	К	13				0.0339	0.0048	0.0004	552	193	14	91	91	21	759	204
Черкесы	К	13				0.0317	0.0037	0.0010	589	173	7	106	92	13	769	211
Балкарцы	К	14	0.0345	0.0069	0.0028	0.0357	0.0071	0.0027	514	151	8	100	85	21	674	205
Карачаевцы	К	14				0.0382	0.0104	0.0055	589	172	10	106	92	30	771	228
Ногайцы	К	14				0.0295	0.0030	0.0003	528	239	21	83	102	24	788	209
Ингуши	К	15	0.0396	0.0099	0.0030	0.0392	0.0094	0.0026	521	196	16	91	98	33	733	222
Чеченцы	К	15				0.0399	0.0105	0.0034	550	177	15	99	92	33	741	223
Осетины (С)	К	16	0.0363	0.0066	0.0017	0.0359	0.0063	0.0015	515	220	20	86	101	31	755	217
Осетины (Ю)	К	16				0.0367	0.0069	0.0018	570	212	20	92	97	31	801	220
Мегрелы	К	17	0.0369	0.0049	0.0006	0.0369	0.0049	0.0006	529	202	16	93	100	23	747	216
Алтайцы (Б)	С	18	0.0533	0.0171	0.0089	0.0519	0.0173	0.0086	517	163	14	93	86	46	695	224
Алтайцы (К)	С	18				0.0552	0.0168	0.0100	632	177	13	121	101	47	821	269
Татары (Т)	С	18				0.0338	0.0071	0.0037	483	127	5	86	68	17	615	171
Тувинцы	С	18				0.0591	0.0151	0.0061	731	242	19	139	136	46	992	322
Хакасы (Т)	С	18				0.0664	0.0318	0.0183	585	173	23	111	97	87	781	295
Хакасы (Ш)	С	18				0.0534	0.0146	0.0070	652	199	12	122	112	38	862	273

Таблица 1. Окончание

Популяция	Регион	Группа	PLINK						GARLIC							
			F_{ROH}			F_{ROH}			NROH			SROH			Σ	
			0.5	1.5	5	0.5	1.5	5	A	B	C	a	b	c	NROH	SROH
Буряты (А)	С	19а	0.0656	0.0181	0.0076	0.0599	0.0112	0.0027	691	317	31	116	152	50	1040	318
Буряты (К)	С	19а				0.0615	0.0166	0.0081	733	235	16	135	130	47	983	312
Калмыки	С	19а				0.0465	0.0058	0.0008	688	260	23	121	126	33	970	280
Эвенки (З)	С	19б				0.0632	0.0195	0.0080	608	195	19	111	110	55	822	276
Эвенки (Я)	С	19б				0.0786	0.0256	0.0105	705	221	23	138	133	69	949	339
Удэгейцы	С	19б				0.0774	0.0246	0.0134	764	219	14	152	136	59	997	347
Якуты	С	19в				0.0723	0.0233	0.0097	725	240	25	137	137	68	990	343
Кеты	С	20а	0.0707	0.0274	0.0127	0.0720	0.0259	0.0121	753	221	21	139	132	70	995	341
Ханты (К)	С	20б				0.0716	0.0280	0.0119	620	201	24	123	125	75	844	323
Ханты (Р)	С	20б				0.0716	0.0266	0.0102	633	212	24	120	126	70	869	316
Чулымцы	С	20в				0.0676	0.0292	0.0167	541	165	22	99	92	79	727	270
Нивхи	С	21а	0.1139	0.0398	0.0133	0.1073	0.0363	0.0119	699	239	28	147	161	84	965	393
Коряки	С	21б				0.1209	0.0422	0.0142	806	301	30	172	210	98	1136	480
Чукчи	С	21б				0.1136	0.0409	0.0137	703	257	32	149	174	100	992	424
Казахи (М)	С. А.	22а	0.0367	0.0055	0.0018	0.0360	0.0040	0.0006	575	198	13	100	98	19	786	217
Казахи (С)	С. А.	22а				0.0358	0.0039	0.0002	511	232	25	83	101	28	768	212
Киргизы	С. А.	22а				0.0414	0.0067	0.0022	589	194	8	106	103	19	791	227
Узбеки	С. А.	22а				0.0334	0.0073	0.0043	549	170	6	97	88	20	725	205
Дунгане	С. А.	22б	0.0562	0.0116	0.0070	0.0562	0.0116	0.0070	808	218	6	151	126	27	1032	304
	В		0.0438	0.0122	0.0043				535	173	16	94	90	37	723	221
	Д		0.0704	0.0377	0.0239				525	156	20	100	93	96	701	289
	К		0.0357	0.0065	0.0018				543	196	15	94	96	26	755	215
	С		0.0702	0.0229	0.0099				663	218	21	127	127	62	902	316
	С. А.		0.0406	0.0067	0.0028				606	202	12	107	103	23	820	233

Примечание 1. В – Волго-Уральский регион и Европа, К – Кавказ, Д – Дагестан, С – Сибирь, С. А. – Средняя Азия.

Примечание 2. 1 – алтайская/тюркская; 2 – пермская подгруппа; 3 – прибалтийско-финская подгруппа; 4 – волжско-финская подгруппа; 5 – восточнославянская группа; 6 – нахско-дагестанская/лезгинская ветвь; 7 – нахско-дагестанская/андийская группа; 8 – нахско-дагестанская/дидойская группа; 9 – алтайская/тюркская; 10 – нахско-дагестанская/даргинская ветвь; 11 – нахско-дагестанская/лакская ветвь; 12 – нахско-дагестанская/аваро-андийская группа; 13 – абхазо-адыгские языки; 14 – алтайская/тюркская; 15 – нахско-дагестанская/нахская ветвь; 16 – индоевропейская/иранская группа; 17 – картвельская семья; 18 – алтайская/тюркская; 19 – а) алтайская/монгольская, б) алтайская/тунгусо-маньчжурские языки, в) алтайская/тюркская; 20 – а) енисейская, б) уральская/обско-угорская, в) алтайская/тюркская; 21 – а) палеоазиатский язык, б) чукотско-камчатские языки; 22 – а) алтайская/тюркская, б) сино-тибетские языки/китайские языки.

ROH (значения $F_{ROH > 5}$, $F_{ROH > 1.5}$), для популяций Сибири по более коротким участкам гомозиготности ($F_{ROH > 0.5}$) наблюдаются схожие значения. Это может свидетельствовать о значительных эффектах инбридинга для этих популяций Сибири в прошлом с постепенным уменьшением доли кровнородственных браков к настоящему времени.

Популяции Волго-Уральского региона занимают промежуточное положение среди других регионов по коэффициенту инбридинга для всех трех групп гомозиготных регионов. Это можно объяснить как его территориальным положением на месте пересечения путей различных миграционных потоков, так и давней традицией заключения браков лишь с родственниками не ближе пятого–шестого колена, что характерно для башкир, казанских татар, чувашей, коми, марийцев и мордвы. Внутри Волго-Уральского региона для популяций с низкой эффективной численностью и проживающих на периферии наблюдаются более высокие значения уровня гомозиготности генома. Максимальное значение обнаружено у бесермян (0.0199), обособленных территориально и имеющих численность около 2 тыс. человек, минимальное значение – у казанских татар и русских (0.0061).

Популяции Кавказа и Средней Азии обладают минимальными значениями коэффициентов инбридинга, оцененных по ROH, – от 0.0030 до 0.0116 для $F_{ROH > 1.5}$. Среди популяций Средней Азии максимальные значения наблюдаются у дунган (0.0116) что, вероятно, объясняется их недавней миграцией из Китая с сохранением их языкового, культурного и демографического единства на фоне практически полного отсутствия межнациональных браков. Ранее было показано, что дунгане Узбекистана генетически намного ближе к популяциям Китая и Северо-Восточной Индии, чем к соседним с ними узбекским популяциям, и не имеют с ними недавнего обмена генами [14].

Оценка ROH с помощью PLINK

В отличие от модельного подхода для поиска ROH, этот анализ не зависит от оценки частоты популяционных аллелей. Общая длина и количество ROH для каждого индивидуального генома колеблется от 41.2 до 687.7 млн пн и от 61 до 394 соответственно. Медианное число ROH на индивидуальный геном в исследованных образцах составило 135.4, граница первого квартиля 121.8, граница третьего квартиля 151.5. Медианная длина ROH 128.4 млн пн, граница первого квартиля 103.7 млн пн, третьего 139.3 млн пн. Среднее значение совокупной длины ROH на индивидуума варьирует от 79.9 млн пн в выборке ногайцев до 289.1 млн пн у гинухцев. Минимальное среднее популяционное значение числа ROH на человека наблюдается у узбеков (103.7), максимальное – среди коряков (312).

Доля ROH различной длины значительно колеблется в зависимости от географического региона и языковой группы. Выборки из Дагестана и Сибири, особенно носители языков нахско-дагестанской группы, чукотско-камчатских и нивхского языков, разительно отличаются от таковых из других географических регионов. Среднее значение общей длины ROH на человека в нашей выборке, говорящей на языках дидойской подгруппы, 305.03 млн пн сравнимо со значением 308.6 млн пн у представителей чукотско-камчатских и нивхского языков. Это полностью согласуется с оценкой геномного инбридинга в этих популяциях.

По доле образцов с одним и более ROH длиной больше 5 млн пн по регионам выделяются три различные группы, которые явно связаны с эндогамией и изоляцией: большая часть дагестанских (78.1%) и сибирских (75.5%) образцов имеет длинные ROH, далее идут кавказский, Волго-Уральский регион и Европа (57.9 и 38% соответственно). Для Среднеазиатского региона наблюдается минимальная доля геномов, имеющих длинные ROH (15.6%). Это полностью соответствует полученным ранее данным о наибольшем генетическом разнообразии и многокомпонентности генофонда среднеазиатских популяций в сравнении с другими этносами Северной Евразии. При этом наблюдаются значительные различия между популяциями внутри каждого региона. Например, у казахов старшего жуза всего 7% образцов имеют ROH больше 5 млн пн. Для двух человек из эвенкийской (З) и казахской (С) популяций не было найдено ROH длиннее 1.5 млн пн.

Оценка ROH с помощью GARLIC

Общая длина и количество ROH, идентифицируемые с помощью пакета GARLIC, для каждого индивида колеблются от 124 до 788.6 млн пн и от 319 до 1237 соответственно. Медианное число ROH на индивидуальный геном в исследованных образцах составило 774.5, граница первого квартиля 754.1, граница третьего квартиля 798.5. Медианная длина ROH – 261.4 млн пн, граница первого квартиля 243.9 млн пн, третьего 282.8 млн пн. Среднее значение совокупной длины ROH на индивидуума варьирует от 167.3 млн пн в выборке башкир (С) до 480.2 млн пн у коряков. Минимальное среднее популяционное значение числа ROH на человека наблюдается у гинухцев (356.8) из-за большей длины ROH за счет сильного инбридинга, максимальное – среди коряков (1136).

Средние размеры ROH классов A и B одинаковы среди популяций одного географического региона, за исключением того, что популяции Дагестана и Сибири имеют большую изменчивость (табл. 2). Максимальные значения наблюдаются у дагестанских популяций (гинухцев, гунзибцев, ахвахцев, цезов).

Наибольшие средние значения класса C также наблюдаются у популяций Дагестана и Сибири,

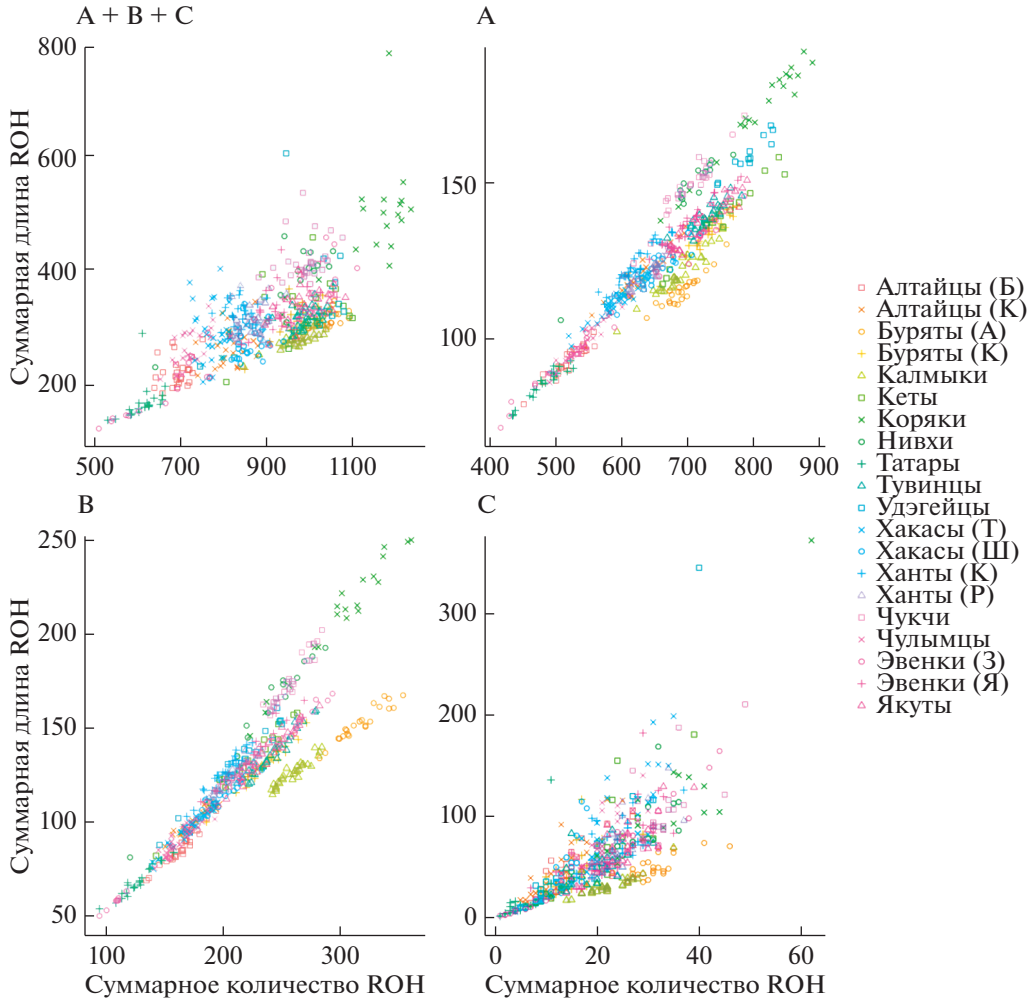

Рис. 1. Отношение суммарной длины ROH (млн пн) к суммарному количеству ROH в исследуемых популяциях Сибири.

далее у популяций Средней Азии, Волго-Уральского региона и Европы, минимальные средние значения – у кавказских популяций. У народов, имеющих смешанное происхождение, – казахов (М, С), бурятов (А), калмыков, караногайцев и кумыков средние размеры ROH аналогичны кавказским популяциям.

Для сибирских чукчей, коряков, нивхов для всех классов длин наблюдается значительное увеличение суммарной длины и количества ROH на человека по сравнению с другими популяциями, что свидетельствует о продолжительном инбридинге, а также эффекте “бутылочного горлышка” (рис. 1). Для трех этих этносов также наблюдается

Таблица 2. Суммарная статистика для значений границ классов длин (тпн) ROH по регионам

Статистика	Сибирь	Средняя Азия	Дагестан	Кавказ	Волго-Уральский регион и Европа	Все
Классы A и B						
Мин.	305.54	287.59	315.23	279.41	301.07	279.41
Макс.	417.37	369.87	668.22	370.51	370.79	668.22
Сред. знач.	366.96	333.15	390.07	321.02	338.75	358.55
SD	7.28	15.00	17.94	8.89	4.35	6.67
Классы B и C						
Мин.	942.65	777.099	981.114	787.342	1010.64	777.10
Макс.	1553.83	1402.82	2598.01	1200.04	1257.98	2598.01
Сред. знач.	1260.15	1115.65	1380.02	999.93	1112.28	1214.99
SD	38.33	120.00	77.19	44.52	20.96	31.39

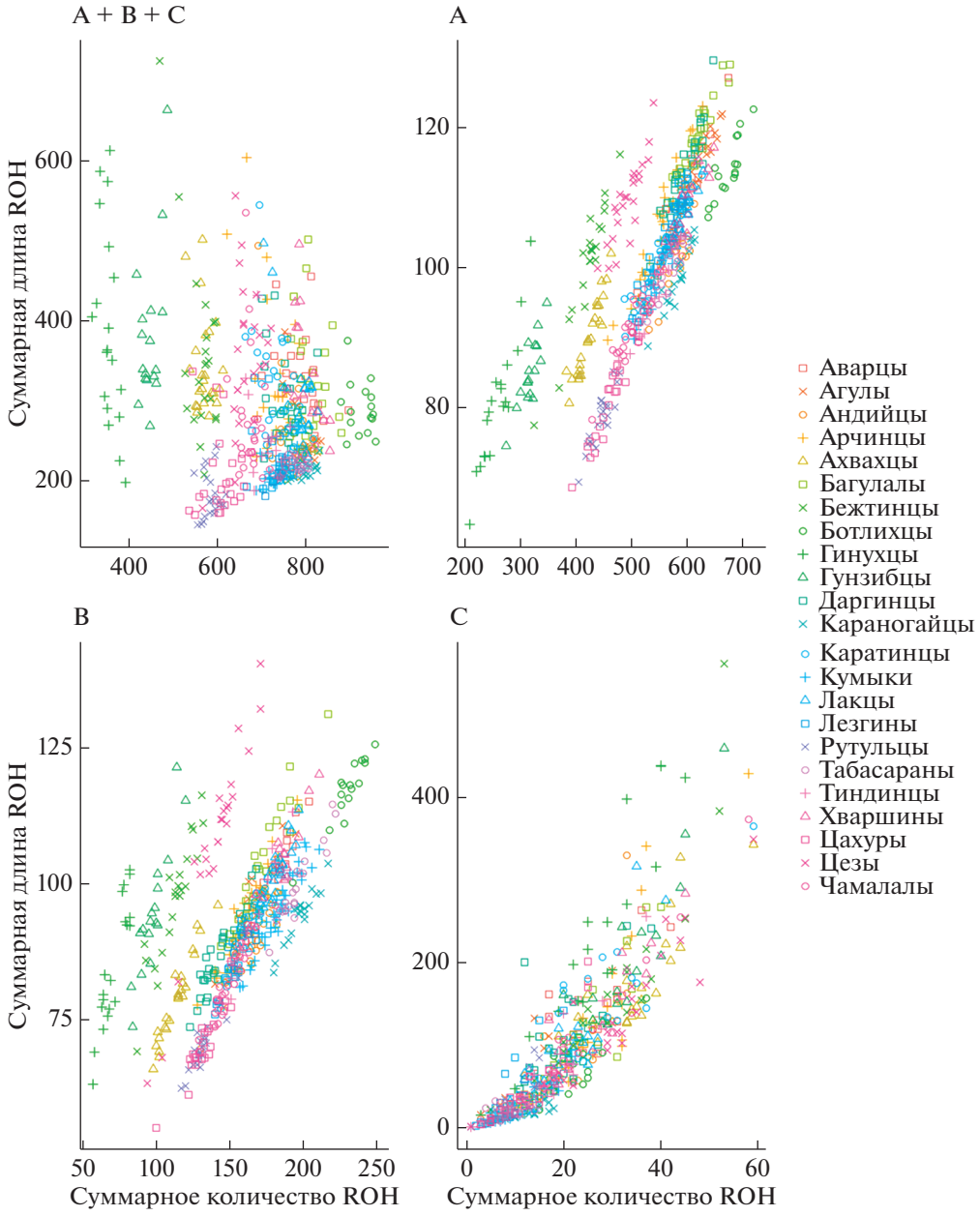

Рис. 2. Отношение суммарной длины ROH (млн пн) к суммарному количеству ROH в исследуемых популяциях Дагестана.

значительное увеличение суммарной длины среднего класса ROH на человека как по сравнению с остальными популяциями, так и по сравнению с классом коротких ROH внутри популяции, что подтверждает продолжительность инбридинга и небольшую эффективную численность популяций. Далее по коротким и средним длинам ROH выделяются кеты, якуты, эвенки (Я), удэгейцы, тувинцы и буряты (К). Для якутов и тувинцев, несмотря на большой эффективный размер популяции, большое количество и длина коротких и средних ROH вероятно связаны с сильным эффектом основателя, который выявляется на маркерах мтДНК и Y-хромосомы [15, 16]. В популяции эвенков (З) обнаружены образцы с признаками метисации, что может значительно занижать значения по длинам и количеству ROH. При анализе Y хромосомных линий у них наблюдаются европейские гаплогруппы, не характерные для коренных сибирских популяций. При анализе аутосомных блоков, идентичных по происхождению (IBD – англ. identical by descent), у эвенков также наблю-

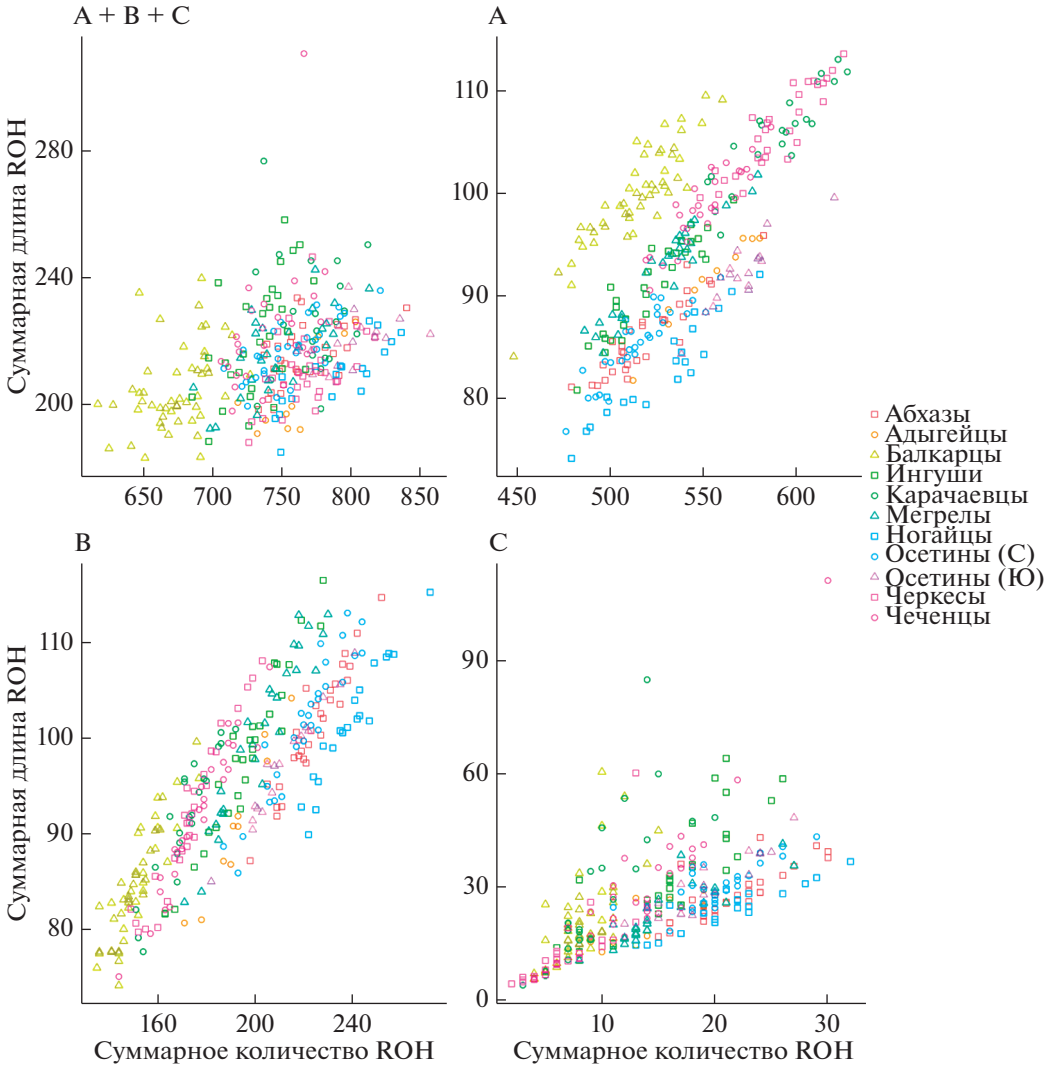

Рис. 3. Отношение суммарной длины ROH (млн пн) к суммарному количеству ROH в исследуемых популяциях Кавказа.

дается значительное количество длинных блоков (4 млн пн), общих с европейскими популяциями, что свидетельствует об относительно недавней метисации этой группы людей с пришлыми европейцами. Для бурятов (А) также характерно увеличение суммарной длины среднего класса ROH, но в большей степени не из-за длины фрагментов, а из-за их количества, что свидетельствует об эффекте “бутылочного горлышка” в прошлом (рис. 1). У бурятов (А) и калмыков наблюдается увеличение количества ROH, но без увеличения суммарной длины, наиболее ярко выражено для ROH средней длины, что свидетельствует о том, что они происходили из довольно небольших полуизолированных сообществ, где эндогамия не была редкостью. По средней суммарной длине и количеству для средних и коротких ROH максимальные значения наблюдаются у сибирских популяций (табл. 1).

Для популяций гинухцев, гунзибцев, бежтинцев, ахвахцев и цезов для всех классов ROH наблюдается распределение соотношения суммарной длины и количества, характерное для популяций, подверженных сильному инбридингу (рис. 2). Распределение ROH связано с размером популяции и их эффективной численностью, при этом относительно небольшие популяции Дагестана и Сибири имеют ROH большего количества и размера, чем большие по численности популяции этих регионов. Инбридинг добавляет небольшое количество очень длинных ROH для тех, кто является потомком кровнородственных браков, увеличивая дисперсию суммы ROH, видимую как сдвиг влево на графике длин ROH по сравнению с количеством ROH.

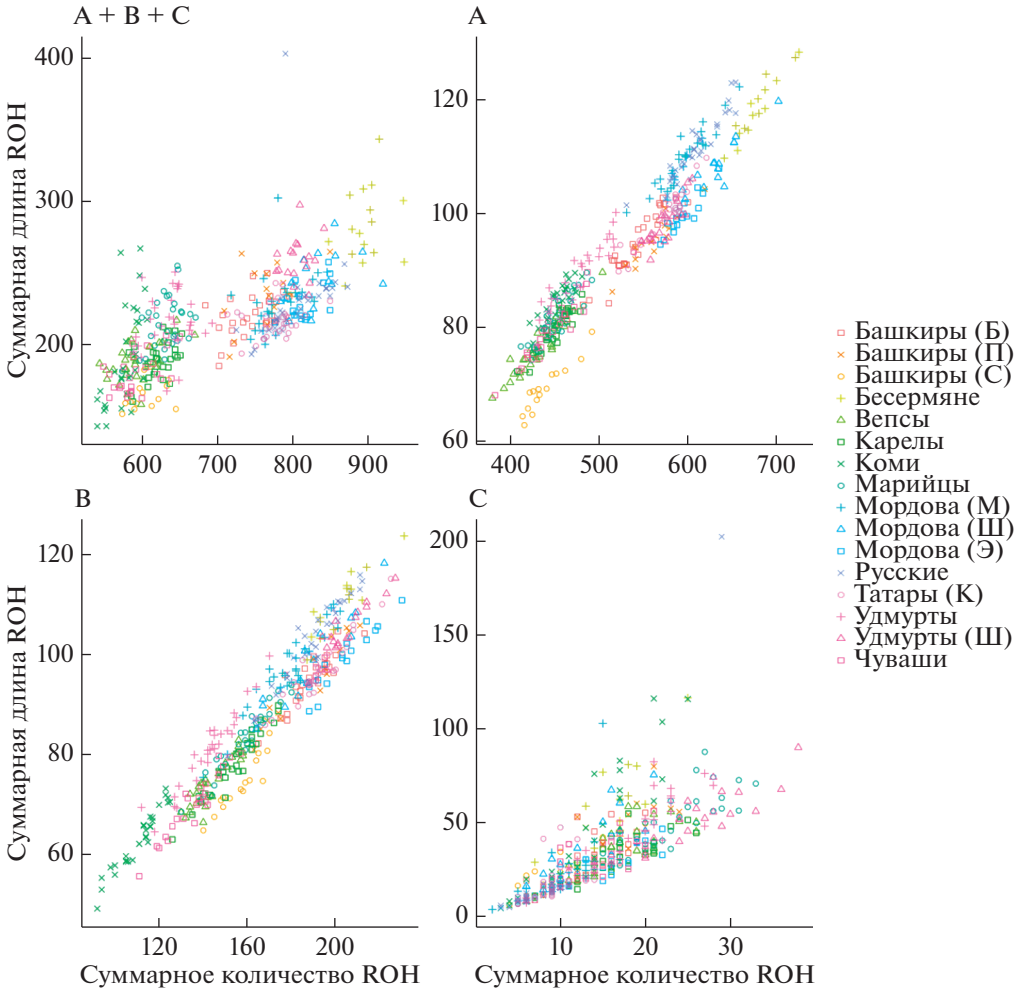

Рис. 4. Отношение суммарной длины ROH (млн пн) к суммарному количеству ROH в исследуемых популяциях Волго-Уральского региона и Европы.

Кавказский регион характеризуется меньшим разбросом по суммарной длине и количеству регионов высокой гомозиготности на индивида, чем в остальных географических регионах. Для длинных ROH наблюдается тенденция к увеличению количества, но не длины, что вероятно объясняется сильной метисацией населения. Среди кавказских популяций выделяется популяция балкарцев, являющаяся одним из самых высокогорных народов Европы и имеющая более длинные короткие ROH, чем у остальных популяций Кавказа, что может свидетельствовать о сильном инбридинге в прошлом (рис. 3).

В популяциях, смешанных по происхождению и характеризующихся высоким уровнем генетического разнообразия, объединяются различные гаплотипы, что обычно значительно снижает количество и длину ROH. Это обнаружено у томских и казанских татар, узбеков и башкир. Для Волго-Уральского региона и Европы наблюдается разделение на две группы по количеству ROH, где основной вклад вносят короткие и средние ROH. В первую группу входят марийцы, коми, вепсы, чуваши, карелы, удмурты и башкиры (С), во вторую – бесермяне, мордва (М, Ш, Э), удмурты (Ш, Б), татары (К), башкиры (Б, П) (рис. 4).

Эффект “бутылочного горлышка” в демографической истории конкретного этноса увеличивают количество ROH, которые обычно все еще относительно короткие, например популяции калмыков и бурятов (А). У некоторых популяций, имеющих как эффект “бутылочного горлышка”, так и распространенность близкородственных браков, наблюдаются большее количество и длина ROH с увеличением дисперсии суммы ROH (коряки, чукчи и нивхи).

Наибольшими отличиями среди популяций Средней Азии характеризуется популяция дунган, в ней наблюдаются высокие значения по количеству и суммарной длине для средних и коротких ROH и минимальные значения для длинных (рис. 5).

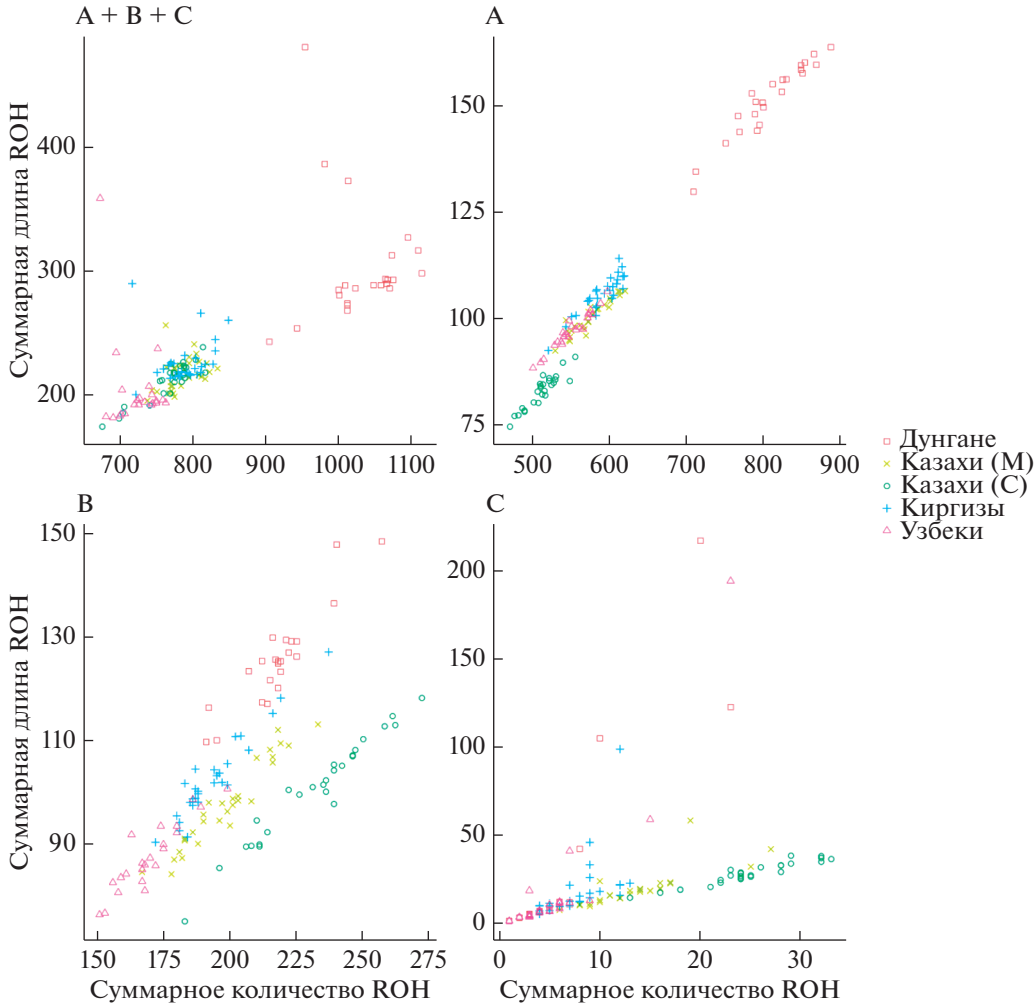

Рис. 5. Отношение суммарной длины ROH (млн пн) к суммарному количеству ROH в исследуемых популяциях Средней Азии.

Доля генома в ROH разной длины и разного количества сильно варьирует по регионам (рис. 6). Наши результаты согласуются с рядом недавних исследований, проведенных с использованием высокоплотных чипов. Уровень гомозиготности по всему геному показывает самый высокий уровень инбридинга в дагестанских и сибирских группах, говорящих на дидойских, чукотко-камчатских и нивхском языках, по сравнению с остальными северо-евразийскими популяционными выборками. Дагестанские и сибирские популяции обладают самыми длинными геномными участками гомозиготности для всех категорий длины ROH и имеют более высокую долю людей с длинными областями аутозиготности.

Высокий коэффициент инбридинга и наличие длинных аутозиготных сегментов согласуются с предыдущими выводами для коэффициентов инбридинга Дагестана [17–19], рассчитанными на основе данных родословных и высокоплотных чипов, а также с данными по Y-хромосомным маркерам указывающими на эффекты основателя и высокую подразделенность по языковым группам [12]. Аналогично для работ по популяционной генетике сибирских этносов с использованием нейтральных маркерных систем Y-хромосомной, митохондриальной ДНК, показано, что во многих популяциях наблюдается эффект основателя по отдельным гаплогруппам с низким генетическим разнообразием, которое свидетельствует о событиях экспансии численности в относительно недавнее время [15, 20, 21]. Интересно, что полученные нами в настоящей работе оценки существенно не отличаются от тех, которые были получены в предыдущем исследовании нашей группы [10] с использованием гораздо меньшего числа маркеров на меньших выборках сибирских популяций.

Для популяций Сибири характерно большее значение суммарной длины классов A и B на человека, что соответствует данным об увеличении

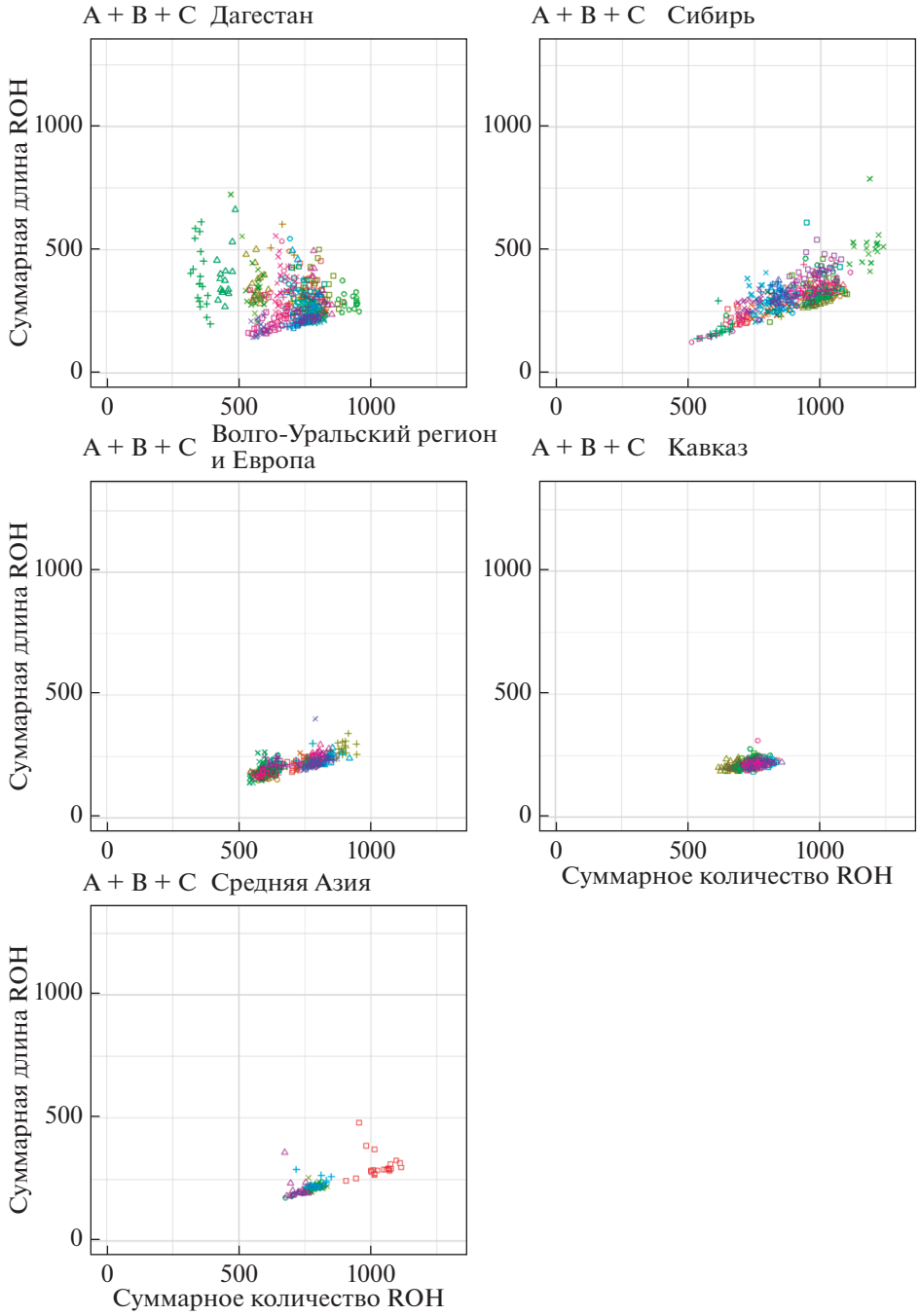

Рис. 6. Отношение суммарной длины ROH (млн пн) к суммарному количеству ROH в исследуемых популяциях для каждого географического региона.

длин классов A и B с удалением популяции от Африки. Эта тенденция похожа на наблюдаемое сокращение разнообразия гаплотипов с увеличением расстояния от Африки. Суммарная длина класса C ROH более изменчива и значительно больше в большинстве популяций Дагестана и Сибири по сравнению с популяциями Кавказа, Средней Азии, Европы и Волго-Уральского региона. Это также свидетельствует о том, что большая доля популяций из Дагестана и Сибири имеют более высокий уровень кровного родства, это аналогично данным, полученным по значениям коэффициента инбридинга (F_{ROH}).

Результаты проведенного анализа показывают, что генофонд населения Северной Евразии демонстрирует значительные различия в распределении суммы длин и количества по всем классам ROH как между исследуемыми географическими

регионами, этносами, так и внутри локальных популяций. Особенно это характерно для территориально удаленных популяций чукчей и коряков, а также культурно и географически изолированных горных популяций Дагестана. Коренные этносы высокогорных районов Дагестана показывают максимальные значения инбридинга, сохранившегося на высоком уровне до настоящего времени. Большее количество и длина коротких и средних ROH у сибирских популяций указывают на более древний и более продолжительный инбридинг. Получено подтверждение значительной генетической гетерогенности среднеазиатских популяций и относительно низком уровне инбридинга у большинства народов Кавказа и Волго-Уральского региона. Полученные результаты позволяют существенно уточнить имеющиеся данные об особенностях структуры генофонда всех исследованных этносов. Данные о аутозиготных блоках сцепления в конкретных популяциях также могут быть использованы при изучении их компонентного состава предрасположенности к различным наследственным заболеваниям.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-34-90101.

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

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. *McQuillan R., Leutenegger A.L., Abdel-Rahman R. et al.* Runs of homozygosity in European populations // Am. J. Hum. Genet. 2008. V. 83. № 3. P. 359–372. https://doi.org/10.1016/j.ajhg.2008.08.007
2. *Szpiech Z.A., Blant A., Pemberton T.J.* GARLIC: Genomic autozygosity regions likelihood-based inference and classification // Bioinformatics. 2017. V. 33. № 13. P. 2059–2062. https://doi.org/10.1093/bioinformatics/btx102
3. *Lemes R.B., Nunes K., Carnavalli J.E.P. et al.* Inbreeding estimates in human populations: Applying new approaches to an admixed Brazilian isolate // PLoS One. 2018. V. 13. № 4. P. e0196360. https://doi.org/10.1371/journal.pone.0196360
4. *Ropers H.* New perspectives for the elucidation of genetic disorders // Am. J. Hum. Genet. 2007. V. 81. № 2. P. 199–207. https://doi.org/10.1086/520679
5. *Kirin M., McQuillan R., Franklin C.S. et al.* Genomic runs of homozygosity record population history and consanguinity // PLoS One. 2010. V. 5. № 11. e13996. https://doi.org/10.1371/journal.pone.0013996
6. *Pemberton T.J., Absher D., Feldman M.W. et al.* Genomic patterns of homozygosity in worldwide human populations // Am. J. Hum. Genet. 2012. V. 91. № 2. P. 275–292. https://doi.org/10.1016/j.ajhg.2012.06.014
7. *Ceballos F.C., Joshi P.K., Clark D.W. et al.* Runs of homozygosity: Windows into population history and trait architecture // Nat. Rev. Genet. 2018. V. 19. № 4. P. 220–234. https://doi.org/10.1038/nrg.2017.109
8. *Frazer K.A., Ballinger D.G., Cox D.R. et al.* International HapMap Consortium A second generation human haplotype map of over 3.1 million SNPs // Nature. 2007. V. 449(7164). P. 851–861. https://doi.org/10.1038/nature06258
9. *Karafet T.M., Bulayeva K.B., Bulayev O.A. et al.* Extensive genome-wide autozygosity in the population isolates of Daghestan // Eur. J. Hum. Genet. 2015. V. 23. № 10.P. 1405–1412. https://doi.org/10.1038/ejhg.2014.299
10. *Колесников Н.А., Харьков В.Н., Зарубин А.А., Степанов В.А.* Особенности распределения регионов высокой гомозиготности в геномах представителей коренного населения Северной Евразии // Генетика. 2019. Т. 55. № 10. С. 1231–1236. https://doi.org/10.1134/S0016675819100072
11. *Purcell S., Neale B., Todd-Brown K. et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses // Am. J. Hum. Genet. 2007. V. 81. P. 559–575. https://doi.org/10.1086/519795
12. *Глазунова Е.О., Харьков В.Н., Раджабов М.О. и др.* Генофонд коренных народов Дагестана цезской группы по маркерам Y-хромосомы // Мед. генетика. 2016. Т. 15. № 4. С. 29–31. https://doi.org/10.1234/XXXX-XXXX-2016-4-29-31
13. *Гаджиев А.Г.* Антропология малых популяций Дагестана. Махачкала: Даг. филиал АН СССР, 1971. 368 с.
14. *Zhabagin M., Balanovska E., Sabitov Z. et al.* The connection of the genetic, cultural and geographic landscapes of transoxiana // Scientific Reports. 2017. V. 7. № 1. P. 3085. https://doi.org/10.1038/s41598-017-03176
15. *Харьков В.Н., Степанов В.А., Медведева О.Ф. и др.* Происхождение якутов: анализ гаплотипов Y-хромосомы // Мол. биология. 2008. Т. 42. № 2. С. 226–237.
16. *Харьков В.Н., Хамина К.В., Медведева О.Ф. и др.* Структура генофонда тувинцев по маркерам Y-хромосомы // Генетика. 2013. Т. 49. № 12. С. 1416–1425. https://doi.org/10.7868/S0016675813120035
17. *Колесников Н.А., Харьков В.Н., Раджабов М.О. и др.* Различия по уровню инбридинга в популяциях Да-

гестана: анализ регионов высокой гомозиготности // Мед. генетика. 2020. Т. 19. № 7(216). С. 21–22.

18. *Bulaeva K.B., Davudov O.M., Pavlova T.A. et al.* Genetic subdivision of Dagestan ethnic populations // Genetika. 2003. V. 39. № 1. P. 83–92.
19. *Bulayeva K.B.* Overview of genetic-epidemiological studies in ethnically and demographically diverse isolates of Dagestan, Northern Caucasus, Russia // Croat. Med. J. 2006. V. 47. № 4. P. 641–648.
20. *Деренко М.В., Малярчук Б.А., Возняк М. и др.* Разнообразие линий Y-хромосомы у коренного населения Южной Сибири // Докл. РАН. 2006. Т. 411. № 2. С. 273–277.
21. *Харьков В.Н., Хамина К.В., Медведева О.Ф. и др.* Генофонд бурят: клинальная изменчивость и территориальная подразделенность по маркерам Y-хромосомы // Генетика. 2014. Т. 50. № 2. С. 203–213. https://doi.org/10.7868/S0016675813110088

Features of the Genomic Distribution of Runs of Homozygosity in the Indigenous Population of Northern Eurasia at the Individual and Population Levels Based on High Density SNP Analysis

N. A. Kolesnikov[a, *, **], V. N. Kharkov[a], A. A. Zarubin[a], M. O. Radzhabov[b], M. I. Voevoda[c], M. A. Gubina[d], E. K. Khusnutdinova[e], S. S. Litvinov[e], N. V. Ekomasova[e], O. V. Shtygasheva[f], N. R. Maksimova[g], A. L. Sukhomyasova[g], and V. A. Stepanov[a]

[a] *Tomsk National Research Medical Center of the Russian Academy of Sciences, Research Institute of Medical Genetics, Tomsk, 634050 Russia*

[b] *Amirhanov Institut of Fysics, Dagestan Federal Research Center, Russian Academy of Sciences, Makhachkala, 367012 Russia*

[c] *Federal State Budget Scientific Institution "Federal Research Center of Fundamental and Translational Medicine", Novosibirsk, 630060 Russia*

[d] *Institute of Cytology and Genetics, Siberian Branch of Russian Academy of Sciences, Novosibirsk, 630060 Russia*

[e] *Institute of Biochemistry and Genetics, Ufa Federal Research Center of the Russian Academy of Sciences, Ufa, 450054 Russia*

[f] *Katanov Khakass State University, Abakan, 655000 Russia*

[g] *Ammosov North-Eastern Federal University, Yakutsk, 677000 Russia*

**e-mail: nikita.kolesnikov@medgenetics.ru*

***e-mail: nik.fleming@mail.ru*

The genomic data analysis (~886889 autosomal SNPs) allowed us to estimate the inbreeding level based on the analysis of runs of homozygosity (F_{ROH}) in a sample of 1806 people, including 76 indigenous populations of Siberia, Dagestan, the Caucasus, Volga-Ural region, Eastern Europe and Central Asia. Data was obtained using biochips Infinium Multi-Ethnic Global-8 Kit. In the populations of Dagestan, within the Nakh-Daghestanian language family, representatives of several language groups can be distinguished, with the maximum F_{ROH} level for the Dido group (0.0727) and the Andean group (0.0378), which is similar in values to the populations speaking Chukchi-Kamchatka and Nivkh languages of Siberia (0.0360). The Siberian populations are characterized by a greater value of the total length of short and medium regions of runs of homozygosity per person. The total length of long ROHs is more variable and much larger in the most of Dagestan and Siberia populations, compared to the populations of the Caucasus, Central Asia, Europe, and the Volga-Ural region.

Keywords: population genetics, ROH, runs of homozygosity, human populations, inbreeding, North Eurasia.

ГЕНЕТИКА, 2021, том 57, № 11, с. 1276–1285

ГЕНЕТИКА ЧЕЛОВЕКА

УДК 577.29

БИОИНФОРМАЦИОННАЯ АННОТАЦИЯ ГЕНОВ ПРЕДРАСПОЛОЖЕННОСТИ К БОЛЕЗНИ АЛЬЦГЕЙМЕРА И ИШЕМИЧЕСКОЙ БОЛЕЗНИ СЕРДЦА

© 2021 г. Н. Ю. Часовских[1], Е. Е. Чижик[1, *], А. А. Бобрышева[1]

[1]*Сибирский государственный медицинский университет, Томск, 634050 Россия*

**e-mail: evgenika06@gmail.com*

Поступила в редакцию 01.02.2021 г.
После доработки 12.04.2021 г.
Принята к публикации 20.04.2021 г.

Проведена функциональная аннотация для генов предрасположенности к болезни Альцгеймера (БА) и генов предрасположенности к ишемической болезни сердца (ИБС) при помощи Cytoscape v 3.6.0. Идентифицированные гены вовлечены в реализацию иммунного ответа, апоптоза, а также регулируют процессы нейрогенеза и ангиогенеза. По результатам проведенной функциональной аннотации гены предрасположенности к БА и гены предрасположенности к ИБС были отнесены к терминам в соответствии с генной онтологией и объединены в группы. Число общих групп функций, в которые были вовлечены гены предрасположенности к болезни Альцгеймера и ишемической болезни сердца, составило 106. Общие гены предрасположенности *APOE*, *APOA1*, *ABCA1*, вовлеченные в метаболизм жирных кислот, потенциально могут участвовать в механизмах ассоциации исследуемых заболеваний. Полученные результаты могут служить предпосылкой для дальнейших исследований вклада наследственных факторов в совместное проявление БА и ИБС.

Ключевые слова: ишемическая болезнь сердца, болезнь Альцгеймера, GWAS, функциональная аннотация, гены предрасположенности.

DOI: 10.31857/S0016675821110035

Болезнь Альцгеймера (**БА**) на сегодняшний день представляет острую проблему для здравоохранения всего мира. По оценкам Alzheimer's Disease International на 2019 г. данным заболеванием страдают более 50 млн человек по всему миру [1]. Многочисленные исследования генетических основ болезни Альцгеймера выявили гены, ответственные за разнообразные функции клеточных структур [2, 3]. Также результаты данных исследований показывают, что проявления БА могут сочетаться с другими патологиями, например с заболеваниями сердечно-сосудистой системы, в частности инфарктом миокарда.

В 2014 г. G. Liu с соавт. для выявления новых факторов риска развития болезни Альцгеймера интегрировали данные трех GWAS исследований, для чего использовали метаанализ на основе генов. Анализ путей проводили с использованием Киотской энциклопедии генов и геномов и базы данных генной онтологии [4]. Авторы впервые выявили участие путей, связанных с сердечно-сосудистыми заболеваниями, клеточными процессами и инфекционными заболеваниями, в развитии болезни Альцгеймера. Также в 2018 г. вышла статья W. Chen и соавт., которые рассматривали ген *ApoE4* в качестве мишени для лечения ишемической болезни сердца (**ИБС**) и болезни Альцгеймера. Было показано, что мутация гена аполипротеина Е ведет к нарушению обмена холестерина, что может привести к развитию ИБС и БА [5]; мутация *ABCA1*, а именно полиморфный вариант гена *ABCA1 – rs2230806*, приводит к тому же самому результату [6–8]. Вместе с тем имеющиеся на сегодняшний день данные не позволяют оценить совокупное участие различных наследственных факторов в механизмах сочетанного проявления БА и ИБС.

Таким образом, задача представленной работы – выявление общих функций генов предрасположенности к БА и генов предрасположенности к ИБС, характеризующих биологические процессы и вовлеченных в развитие данных заболеваний.

МАТЕРИАЛЫ И МЕТОДЫ

Гены предрасположенности к ИБС и гены предрасположенности к болезни Альцгеймера были извлечены из публичной базы данных DisGeNET [9]. Оценка функционального сходства генов предрасположенности к исследуемым забо-

леваниям осуществлялась с помощью алгоритма, реализованного в плагине ClueGO Cytoscape v3.6.0 [10]. ClueGO визуализирует связи между терминами GO и функциональными группами в биологических сетях. Для установления связи между генами предрасположенности к изучаемым заболеваниям использовались с 8-го по 15-й уровни иерархии на основе терминологии GO в категориях "биологический процесс" и "молекулярные функции". Для анализа использовался двусторонний гипергеометрический тест с $p < 0.05$.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Списки генов предрасположенности к БА и предрасположенности к ИБС были сформированы на основании анализа данных DisGeNET. В случае БА список составил 446 протеинкодирующих генов, а в случае ИБС – 324 гена.

По результатам проведенной функциональной аннотации гены предрасположенности к БА и гены предрасположенности к ИБС были отнесены к терминам в соответствии с генной онтологией (GO), при этом схожие функции генов были объединены в группы. 60 групп состояли только из генов предрасположенности к ИБС и 445 – из генов предрасположенности к БА. Кроме того, было выявлено 106 общих групп функций, условно разбитых нами на крупные блоки, соответствующие вовлеченности в процессы иммунной системы, функции транспортеров, сигналинг в клетке, липидный обмен и метаболизм, сердечно-сосудистую систему, нервную систему (табл. 1).

Помимо данных результатов выявлено 49 функций, которые нельзя было выделить в отдельную группу.

Функции иммунной системы, а именно, дифференцировка лейкоцитов (GO:0002521), дифференцировка миелоидных лейкоцитов (GO:0002573), регуляция секреции цитокинов (GO:0050707), дегрануляция секреторных гранул лейкоцитов (GO:0043299), хемотаксис гранулоцитов (GO:0071621), положительная регуляция миграции лейкоцитов (GO:0002687), сигнальный путь рецептора Т-клеток (GO:0050852) влияют на возникновение и поддержание воспаления при исследуемых заболеваниях. Действительно, имеются данные о том, что, например, в ряде случаев при наличии у пациента ИБС повышается концентрация IL-1β в крови из-за нарушения коронарного кровотока [11]. Кроме того, и при болезни Альцгеймера наблюдается повышенный уровень данного интерлейкина [12].

Дифференцировку лейкоцитов регулируют выявленные нами общие гены предрасположенности к болезни Альцгеймера и ишемической болезни сердца *AGER*, *CD40LG*, *EEF1A2*, *IL10*, *IL18*, *IL6*, *MMP9*, *MOK*, *PLCG2*, *PPARG*, *TGFB1*, *TLR4*, *TNF*, *VEGFA*. Интерлейкин-6 представляет собой цитокин с широким спектром биологических функций. Известно, что он играет важную роль в окончательной дифференцировке В-клеток в Ig-секретирующие клетки [13]. Интерлейкин-6 может выводиться в кровоток после сокращения мышц и способствовать расщеплению жиров и повышению резистентности к инсулину, вызывать рост миеломы и плазмоцитомы, а также дифференцировку нервных клеток [14, 15]. Продукт гена *IL18*, интерлейкин-18, относится к провоспалительным цитокинам, при связывании с IL18R1 и IL18RAP образует тройной сигнальный комплекс, который активирует NF-κB, запуская синтез медиаторов воспаления [16]. Ген *AGER* отвечает за образование рецептора конечных продуктов гликозилирования (RAGE), являющегося медиатором острого и хронического сосудистого воспаления при атеросклерозе. Взаимодействие RAGE с S100B после инфаркта миокарда может играть роль в апоптозе миоцитов путем активации передачи сигналов ERK1/2 и p53/TP53 [17]. Рецептор конечных продуктов гликозилирования также является рецептором амилоидного бета-пептида и способствует транслокации амилоид-бета-пептида (ABPP) через клеточную мембрану из внеклеточного во внутриклеточное пространство в корковых нейронах. Инициированная ABPP-передача сигналов RAGE, особенно стимуляция митоген-активируемой протеинкиназы p38 (MAPK), способна управлять транспортной системой, доставляющей ABPP в виде комплекса с RAGE в интранейрональное пространство [18, 19]. Частным случаем такого процесса является дифференцировка миелоидных лейкоцитов (GO:0002573), при которой неспецифическая миелоидная клетка-предшественник приобретает специфические признаки любой клетки линии миелоидных лейкоцитов.

В реализацию сигнального пути рецептора Т-клеток (GO:0050852) включены гены предрасположенности к ИБС и болезни Альцгеймера *NFKB1* и *PLCG2*. Ген *PLCG2* кодирует фермент 1-фосфатидилинозитол-4,5-бисфосфат-фосфодиэстеразу γ-2, который принимает участие в продукции молекул вторичного мессенджера диацилглицерина (DAG) и 1,4,5-трифосфата инозита (IP3) и является важным ферментом в трансмембранной передаче сигналов. *NFKB1* ответствен за образование ядерного фактора субъединицы NF-κB p105, являющегося конечной точкой ряда событий трансдукции сигнала, которые инициируются широким спектром стимулов, связанных с такими процессами, как воспаление, иммунитет, дифференцировка, клеточный рост, онкогенез и апоптоз [20].

Регуляцию секреции цитокинов могут осуществлять *IL10*, *IL1B* и *IL1A*. В хемотаксис вовлечен провоспалительный цитокин IL-1β, влияющий

Таблица 1. Список функций, ассоциированных с БА и ИБС одновременно

№	Наименование функциональной группы	Число генов, входящих в функциональную группу		Общие гены предрасположенности
		ИБС	БА	
Иммунная система				
1	Хемотаксис гранулоцитов (GO:0071621)	11	13	*CCL2*, *IL1B*, *IL1RN*, *PLA2G1B*, *PLA2G2A*
2	Дегрануляция лейкоцитов (GO:0043299)	23	46	*EEF1A2*, *HFE*, *HMOX1*, *IGF2R*, *LRP1*, *MMP9*, *MPO*, *NFKB1*, *OLR1*
3	Регуляция секреции цитокинов (GO:0050707)	17	19	*AGER*, *APOA1*, *CRP*, *IL10*, *IL1A*, *IL1B*, *LPL*, *LRP1*, *MOK*, *TLR4*, *TNF*
4	Сигнальный путь рецептора Т-клеток (GO:0050852)	5	19	*NFKB1*, *PLCG2*
5	Положительная регуляция миграции лейкоцитов (GO:0002687)	23	21	*AGER*, *IL1RN*, *IL6*, *IL6R*, *MOK*, *SERPINE1*, *TGFB1*, *TNF*, *VEGFA*
6	Дифференцировка лейкоцитов (GO:0002521)	30	54	*AGER*, *CD40LG*, *EEF1A2*, *IL10*, *IL18*, *IL6*, *MMP9*, *MOK*, *PLCG2*, *PPARG*, *TGFB1*, *TLR4*, *TNF*, *VEGFA*
7	Дифференцировка миелоидных лейкоцитов (GO:0002573)	15	29	*EEF1A2*, *MMP9*, *PPARG*, *TGFB1*, *TLR4*, *TNF*, *VEGFA*
Транспортеры				
8	АТФазная активность, связанная с трансмембранным движением веществ (GO:0042626)	8	9	*ABCA1*, *ABCB1*
9	Активность трансмембранного переносчика ионов переходного металла (GO:0046915)	1	7	–
10	Регулирование активности катионного канала (GO:2001257)	6	27	*MMP9*, *PLCG2*
11	Активность трансмембранного транспортера неорганического катиона (GO:0022890)	26	63	*CFH*, *MMP9*, *PLCG2*, *PTGS2*
12	Активность трансмембранного транспортера моновалентного неорганического катиона (GO:0015077)	17	28	*CFH*, *PTGS2*
Липидный обмен и метаболизм				
13	Окисление жирных кислот (GO:0019395)	5	12	*DECR1*, *IRS1*, *PPARG*
14	Метаболический процесс арахидоновой кислоты (GO:0019369)	8	8	*HPGDS*, *PTGS2*
15	Транспорт жирных кислот (GO:0015908)	11	14	*ACE*, *APOE*, *IL1B*, *PLA2G1B*, *PLA2G2A*, *PPARG*
16	Регуляция транспорта холестерола (GO:0032374)	16	14	*ABCA1*, *APOA1*, *APOE*, *CETP*, *LRP1*, *NFKB1*, *PON1*, *PPARG*
Сердечно-сосудистая система				
17	Регуляция развития сердечной мышечной ткани (GO:0055024)	10	12	*NR3C1*, *TGFB1*

Таблица 1. Продолжение

№	Наименование функциональной группы	Число генов, входящих в функциональную группу		Общие гены предрасположенности
		ИБС	БА	
18	Развитие кровеносных сосудов (GO:0001568)	63	53	*AGT*, *APOE*, *BCAS3*, *CCL2*, *HMOX1*, *IL10*, *IL18*, *IL1A*, *IL1B*, *IL6*, *IL6R*, *LDLR*, *LRP1*, *NOS3*, *NR3C1*, *PPARG*, *PTGS2*, *SERPINE1*, *TGFB1*, *VEGFA*
19	Регулирование вазоконстрикции (GO:0019229)	10	10	*ACE*, *AGT*, *PTGS2*
20	Негативная регуляция свертывания крови (GO:0030195)	13	6	*APOE*, *F2*, *NOS3*, *PLG*, *SERPINE1*
21	Развитие сосудистой сети клубочков (GO:0072012)	7	1	*IL6R*
	Сигналинг в клетке			
22	MAPK каскад (GO:0000165)	55	96	*AGER*, *AGT*, *APOE*, *CCL2*, *CCR5*, *CD40LG*, *IL18*, *IL1B*, *IL1RN*, *IL6*, *IRS1*, *LRP1*, *MOK*, *NFKB1*, *PLA2G1B*, *PLA2G2A*, *TGFB1*, *TLR4*, *TNF*, *VEGFA*
23	p38 MAPK каскад (GO:0038066)	6	7	*AGER*, *IL1B*, *IL6*, *MOK*, *VEGFA*
24	Регуляция активности протеинтирозинкиназы (GO:0061097)	7	18	*ACE*, *AGT*, *BDNF*, *LRP1*
25	Сигнальный путь внутриклеточного рецептора стероидного гормона (GO:0030518)	11	11	*ESR1*, *ESR2*, *NR3C1*
26	Активность фосфатидилинозитол-3-киназы (GO:0035004)	7	7	*IRS1*, *TGFB1*
27	Сигнальный путь рецептора фактора роста эндотелия сосудов (GO:0048010)	10	9	*IL1B*, *VEGFA*
28	Сигнальный путь β-рецептора трансформирующего фактора роста (GO:0007179)	17	11	*TGFB1*
29	Регуляция пептидил-тирозин фосфорилирования (GO:0050730)	24	45	*ACE*, *AGT*, *BDNF*, *EEF1A2*, *IL18*, *IL6*, *IL6R*, *LRP1*, *TGFB1*, *TNF*, *VEGFA*
30	Регуляция передачи сигнала белка Ras (GO:0046578)	17	15	*ABCA1*, *APOA1*, *APOE*
31	Активность циклин-зависимого белка серинтреонинкиназы (GO:0004693)	4	13	*AGER*, *MOK*
32	Активность протеинтирозинкиназы (GO:0004713)	13	33	*ACE*, *AGT*, *BDNF*, *CD40LG*, *IGF2R*, *LRP1*
33	Негативная регуляция апоптотического сигнального пути (GO:2001234)	20	35	*HMOX1*, *IL1A*, *IL1B*, *IL6*, *MMP9*, *NOS3*, *PTGS2*, *SERPINE1*, *TNF*
34	Регуляция внутреннего апоптического сигнального пути (GO:2001242)	5	30	*MMP9*, *PTGS2*
35	Секреция инсулина (GO:0030073)	15	20	*HMGCR*, *IL1B*, *IL1RN*, *IRS1*, *LRP1*, *TNF*

Таблица 1. Окончание

№	Наименование функциональной группы	Число генов, входящих в функциональную группу		Общие гены предрасположенности
		ИБС	БА	
36	Активность лиганда рецептора (GO:0048018)	37	41	*AGT*, *APOA1*, *BDNF*, *CCL2*, *CD40LG*, *IL10*, *IL18*, *IL1A*, *IL1B*, *IL1RN*, *IL6*, *IL6R*, *TGFB1*, *TNF*, *VEGFA*
37	Регуляция высвобождения цитохрома С из митохондрии (GO:0090199)	4	14	*IL6*, *MMP9*
	Нервная система			
38	Регуляция нейрогенеза (GO:0050767)	38	80	*AGER*, *AGT*, *APOE*, *BDNF*, *F2*, *IL1B*, *IL6*, *LDLR*, *LRP1*, *MOK*, *NR3C1*, *PPARG*, *TGFB1*, *TMEM106B*, *TNF*, *VEGFA*
39	Негативная регуляция синаптической передачи (GO:0050805)	6	16	*AGER*, *IL1B*, *MOK*, *PTGS2*
40	Регуляция синаптической передачи, глутаматергическая (GO:0051966)	5	12	*CCL2*, *PTGS2*, *TNF*
41	Дифференцировка дендритных клеток (GO:0097028)	6	8	*AGER*, *MOK*, *TGFB1*
42	Регенерация проекции нейрона (GO:0031102)	3	11	*APOA1*, *IL6*, *LRP1*
43	Регуляция нейрональной синаптической пластичности (GO:0048168)	2	15	*AGT*, *APOE*
44	Дифференцировка глиальных клеток (GO:0010001)	16	38	*AGER*, *F2*, *IL1B*, *IL6*, *LDLR*, *LRP1*, *MOK*, *PPARG*, *TGFB1*, *TLR4*, *TNF*
45	Регуляция синаптической пластичности (GO:0048167)	9	39	*AGER*, *AGT*, *APOE*, *MOK*, *PTGS2*
46	Миграция клеток конечного мозга (GO:0022029)	5	7	–
47	Развитие переднего мозга (GO:0030900)	15	35	*CRP*, *LRP1*
48	Регуляция развития дендритного отдела позвоночника (GO:0060998)	2	17	*APOE*
49	Регуляция глиогенеза (GO:0014013)	17	20	*AGER*, *F2*, *IL1B*, *IL6*, *LDLR*, *LRP1*, *MOK*, *PPARG*, *TGFB1*, *TNF*
50	Модуляция химической синаптической передачи (GO:0050804)	21	73	*AGER*, *AGT*, *APOE*, *BDNF*, *CCL2*, *IL1B*, *MOK*, *PTGS2*, *TNF*
51	Развитие нейронов (GO:0048666)	35	99	*AGER*, *AGT*, *APOA1*, *APOE*, *BDNF*, *IL6*, *LRP1*, *MOK*, *TMEM106B*, *UNC5C*, *VEGFA*
52	Закрытие нервной трубки (GO:0001843)	10	3	*MTHFR*, *TGFB1*
53	Регуляция аксоногенеза (GO:0050770)	10	22	*APOE*, *BDNF*, *LRP1*, *VEGFA*
54	Дифференцировка астроцитов (GO:0048708)	12	22	*AGER*, *F2*, *IL1B*, *IL6*, *LDLR*, *LRP1*, *MOK*, *TLR4*, *TNF*
55	Регуляция образования β-амилоида (GO:1902003)	2	15	*APOE*
56	Тау-протеинкиназная активность (GO:0050321)	1	8	*IL6*
57	Регуляция транспорта ионов кальция в цитозоль (GO:0010522)	9	20	*F2*, *PLA2G1B*, *PLA2G2A*, *TGFB1*

на процесс клеточной миграции. Помимо указанного влияния, IL-1β в центральной нервной системе влияет на выживаемость нейронов и может оказывать нейротоксическое действие [21].

Дегрануляция, т.е. процесс регулируемого экзоцитоза секреторных гранул лейкоцитов (GO:0043299), также связана с функцией общих генов предрасположенности к ИБС и болезни Альцгеймера: *EEF1A2*, *HFE*, *HMOX1*, *IGF2R*, *LRP1*, *MMP9*, *MPO*, *NFKB1*, *OLR1*. В частности, ген *IGF2R* вовлечен в реализацию транспорта фосфорилированных лизосомальных ферментов из комплекса Гольджи и клеточной поверхности в лизосомы. Лизосомные ферменты, несущие фосфоманнозильные остатки, специфически связываются с маннозо-6-фосфатными рецепторами в аппарате Гольджи, далее полученный комплекс рецептор–лиганд транспортируется в кислый прелизосомальный компартмент [22].

Хемотаксис гранулоцитов (GO:0071621) представляет собой движение гранулоцитов в ответ на внешний раздражитель. В реализацию данной функции включены гены *CCL2*, *IL1B*, *IL1RN*, *PLA2G1B* и *PLA2G2A*. Примечательно, что продукт гена *CCL2*, C-C мотив хемокина 2, является лигандом хемокинового рецептора CCR и посредством связывания с ним может индуцировать сильный хемотаксический ответ и мобилизацию внутриклеточных ионов кальция [23]. Считается, что данный белок играет главную роль в рекрутировании моноцитов в артериальную стенку на ранних стадиях развития атеросклероза [24].

За реализацию функции положительной регуляции миграции лейкоцитов (GO:0002687) ответственны следующие гены предрасположенности к ИБС и болезни Альцгеймера: *AGER*, *IL1RN*, *IL6*, *IL6R*, *MOK*, *SERPINE1*, *TGFB1*, *TNF*, *VEGFA*. Продукт гена *SERPINE1* ответствен за регуляцию клеточной миграции, независимо от его роли в качестве ингибитора протеазы [25]. Одной из функций эндотелиального фактора роста сосудов А, продукта экспрессии гена *VEGFA*, является участие в реализации клеточной миграции [26]. Сигнальный путь рецептора фактора роста эндотелия сосудов (GO:0048010) – это любая серия молекулярных сигналов, инициируемых связыванием внеклеточного лиганда с рецептором сосудистого эндотелиального фактора роста (VEGFR), расположенного на поверхности принимающей клетки и заканчивающегося регуляцией нижестоящего клеточного процесса, например транскрипции.

По полученным данным пять групп функций исследуемых генов вовлечены в процессы трансмембранного транспорта, которые также связаны с липидным обменом. Так, АТФазная активность, связанная с трансмембранным движением веществ (GO:0042626), является первичным активным переносчиком растворенного вещества через мембрану по реакции: АТФ + H_2O = АДФ + фосфат. Транспортный белок может быть временно фосфорилирован (транспортеры P-типа) или нет (транспортеры ABC-типа и другие семейства транспортеров). Первичный активный транспорт происходит вверх по градиенту концентрации растворенного вещества и управляется первичным источником энергии. За реализацию данной функции отвечают общие для ИБС и болезни Альцгеймера гены *ABCA1* и *ABCB1*. Установлено, что мутация гена *ABCA1* может привести как к развитию дефицита ЛПВП 1, так и к дефициту ЛПВП 2. Первый вариант дефицита характеризуется отсутствием холестерина ЛПВП в плазме, накоплением сложных эфиров холестерина и преждевременным развитием ИБС [27]. Дефицит ЛПВП 2 наследуется как аутосомно-доминантный признак и характеризуется умеренно низким уровнем холестерина ЛПВП, склонностью к преждевременному развитию ИБС и уменьшением оттока клеточного холестерина [28].

За счет функции активности трансмембранного транспортера неорганического катиона (GO:0022890), реализованной при помощи общих для исследуемых заболеваний генов *CFH*, *MMP9*, *PLCG2*, *PTGS2*, происходит передача неорганических катионов с одной стороны мембраны на другую. Частными случаями этого процесса являются активности трансмембранного переносчика ионов переходного металла (GO:0046915) и трансмембранного транспортера моновалентного неорганического катиона (GO:0015077). Управление же данной функцией происходит путем регулирования активности катионного канала (GO:2001257).

Выявленные функции липидного обмена и метаболизма, а именно метаболический процесс арахидоновой кислоты (GO:0019369) и окисление жирных кислот (GO:0019395) могут быть вовлечены в механизмы развития сосудистых нарушений при ИБС. Кроме того, имеются данные о том, что при БА наблюдается нарушение метаболизма ненасыщенных жирных кислот, в том числе и арахидоновой [29]. Окисление жирных кислот представляет собой процесс удаления из них одного или нескольких электронов. За реализацию данной функции отвечают гены предрасположенности к обоим заболеваниям: *DECR1*, *IRS1*, *PPARG*. Ген *DECR1* кодирует митохондриальный фермент 2,4-диеноил-КоА-редуктазу, который является вспомогательным в процессе β-окисления. Также он участвует в метаболизме ненасыщенных жирных эноил-КоА сложных эфиров, имеющих двойные связи в четных и нечетных позициях в митохондриях. Катализирует НАДФ-зависимое восстановление 2,4-диеноил-КоА с получением транс-3-еноил-КоА [30]. Продукт гена *IRS1*, субстрат инсулинового рецептора 1, может опосредовать контроль различных клеточных процессов с помощью инсулина [31].

Функции генов предрасположенности к ИБС и болезни Альцгеймера, *HPGDS* и *PTGS2*, связаны с метаболическим процессом арахидоновой кислоты (GO:0019369). Так, ген *PTGS2* отвечает за синтез фермента простагландин G/H синтазы 2, который за счет циклооксигеназной активности окисляет арахидонат до гидропероксиэндопероксида простагландина G2 (PGG2), а с помощью пероксидазной активности снижает PGG2 до гидроксиэндопероксида PGH2 [32, 33]. Примечательно, что ген *HPGDS* кодирует бифункциональный фермент, который катализирует как превращение PGH2 в PGD2 (простагландина, участвующего в сокращении/расслаблении гладких мышц и мощного ингибитора агрегации тромбоцитов), так и конъюгацию глутатиона с широким спектром арилгалогенидов и органических изотиоцианатов [34].

По полученным данным, исследуемые гены вовлечены в негативную регуляцию апоптотического сигнального пути (GO:2001234), в том числе внутреннего (GO:2001242). Также они регулируют процесс высвобождения цитохрома С из митохондрий (GO:0090199), при котором цитохром С способен перемещаться из межмембранного пространства митохондрий в цитозоль, что является ранним этапом апоптоза и приводит к активации каспазы. Высвобождение цитохрома С из митохондрий является центральным событием в сигнальной фазе апоптотического процесса, и его часто используют для мониторинга данного типа гибели клеток. Любое событие, которое вызывает апоптоз, в какой-то момент вызовет высвобождение цитохрома С из митохондрий [35].

В ходе исследования были выявлены группы функций, относящиеся к нервной и сердечно-сосудистой системам, включающие гены предрасположенности к БА и к ИБС.

Так, функция дифференцировки дендритных клеток (GO:0097028) регулируется генами *AGER*, *MOK* и *TGFB1*. Ген *TGFB1* участвует в качестве регулятора инициации воспаления, в том числе при участии дендритных клеток [36]. MAPK/MAK/MRK перекрывающаяся киназа, кодируемая геном *MOK*, способна фосфорилировать несколько экзогенных субстратов и подвергаться аутофосфорилированию [37]. Данная функция представляет собой частный случай дифференцировки лейкоцитов (GO:0002521).

Частным случаем нейрогенеза (GO:0050767) является регуляция глиогенеза (GO:0014013). Известно, что ген *LRP1*, участвующий в данном процессе, кодирует белок, связанный с рецептором липопротеинов низкой плотности (ЛПНП), который в свою очередь принимает участие в эндоцитозе и фагоцитозе апоптотических клеток, клеточном липидном гомеостазе, плазменном клиренсе остатков хиломикрона и активированного *LRPAP1* (α-2-макроглобулина), основного компонента амилоидных бляшек, обнаруженного у пациентов с болезнью Альцгеймера [38].

В регуляцию развития сердечной мышечной ткани (GO:0055024) вовлечены гены *NR3C1* и *TGFB1*. Продуктом гена *NR3C1* является глюкокортикоидный рецептор, изоформа-β которого обладает внутренней транскрипционной активностью, независимой от изоформы-α, при их совместной экспрессии [39], а также может играть роль в контроле метаболизма глюкозы, поддерживая чувствительность к инсулину.

Отдельно следует выделить такие гены предрасположенности к ИБС и болезни Альцгеймера, как *ABCA1* и *APOA1*, встречающиеся в функциях, описанных выше. Возникновение в данных генах мутаций может привести к семейному синдрому дефицита ЛПВП 1 типа и ЛПВП 2 типа, что влечет за собой нарушение соотношения липидных фракций и, как следствие, раннюю семейную ишемическую болезнь сердца [40, 41].

Важную роль в механизмах исследуемых нарушений может играть ген *SORT1*, поскольку сортилин способствует апоптозу нейронов, а находясь в адипоцитах участвует в формировании везикул для транспортера глюкозы SLC2A4/GLUT4, повышающего чувствительность к инсулину [42]. Полиморфизм гена *SORT1* изменяет уровень ЛПНП в плазме, что повышает предрасположенность носителя полиморфизма к инфаркту миокарда.

Интерлейкин-6, связывающийся с рецептором IL6R, вовлечен в развитие сосудистой сети клубочков и играет важную роль в дифференцировке В-клеток и нервных клеток, содействует расщеплению жиров и повышению резистентности к инсулину [43].

Таким образом, в результате проведенного исследования были сформированы уникальные перечни протеинкодирующих генов предрасположенности к болезни Альцгеймера и ишемической болезни сердца. Результаты функциональной аннотации демонстрируют, что общие функции генов предрасположенности к исследуемым заболеваниям связаны с развитием и функционированием сердечно-сосудистой и нервной систем. При этом процессы секреции цитокинов, миграции и хемотаксиса лейкоцитов, метаболизма жирных кислот, апоптоза также оказывают влияние на нарушение гомеостаза клеток нервной и сердечно-сосудистой систем при БА и ИБС.

Полученные результаты вносят вклад в развитие представлений о коморбидности сердечно-сосудистых заболеваний, в частности ИБС и болезни Альцгеймера. Биоинформационный подход позволяет провести комплексный анализ уже существующих исследований с точки зрения функциональных характеристик генов. Полученные теоретические знания могут служить основой для оценки

общих механизмов, их дальнейшего экспериментального изучения, а также для разработки новых методов воздействия на найденные мишени.

Финансовая поддержка при подготовке статьи не осуществлялась.

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

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. World Alzheimer Report 2019: Attitudes to Dementia. Alzheimer's Disease International, 2019. P. 166.
2. *Corrêa-Velloso J.C., Gonçalves M.C., Naaldijk Y. et al.* Pathophysiology in the comorbidity of Bipolar Disorder and Alzheimer's disease: pharmacological and stem cell approaches // Prog. Neuropsychopharmacol. Biol. Psychiatry. 2018. V. 80. Pt A. P. 34–53. https://doi.org/10.1016/j.pnpbp.2017.04.033
3. *Klimova B., Kuca K., Maresova P.* Global view on Alzheimer's disease and diabetes mellitus: threats, risks and treatment alzheimer's disease and diabetes mellitus // Curr. Alzheimer. Res. 2018. V. 15. № 14. P. 1277–1282. https://doi.org/10.2174/1567205015666180925110222
4. *Liu G., Yao L., Liu L. et al.* Cardiovascular disease contributes to Alzheimer's disease: Evidence from large-scale genome-wide association studies // Neurobiol. Aging. 2014. V. 35. № 4. P. 786–792. https://doi.org/10.1016/j.neurobiolaging.2013.10.084
5. *Chen W., Jin F., Cao G. et al.* ApoE4 may be a promising target for treatment of coronary heart disease and Alzheimer's disease // Current Drug Targets. 2018. V. 19. № 9. P. 1038–1044. https://doi.org/10.2174/1389450119666180406112050
6. *Wang F., Ji Y., Chen X. et al.* ABCA1 variants rs2230806 (R219K), rs4149313 (M8831I), and rs9282541 (R230C) are associated with susceptibility to coronary heart disease // J. Clin. Lab. Anal. 2019. V. 33. № 6. P. e22896. https://doi.org/10.1002/jcla.22896
7. *Fouladseresht H., Khazaee S., Javad Zibaeenezhad M. et al.* Association of ABCA1 haplotypes with coronary artery disease // Lab. Med. 2020. V. 51. № 2. P. 157–168. https://doi.org/10.1093/labmed/lmz031
8. *Fehér Á., Giricz Z., Juhász A. et al.* ABCA1 rs2230805 and rs2230806 common gene variants are associated with Alzheimer's disease // Neurosci. Lett. 2018. V. 664. P. 79–83. https://doi.org/10.1016/j.neulet.2017.11.027
9. *Pinero J., Bravo A., Rosinach N.Q. et al.* DisGeNET: A comprehensive platform, integrating information on human disease-associated genes and variants // Nucl. Acids. Res. 2017. V. 45. P. D833–D839. https://doi.org/10.1093/nar/gkw943
10. *Bindea G., Mlecnik B., Hackl H. et al.* ClueGO: A Cytoscape plug-into decipher functionally grouped gene ontology and pathway annotation networks // Bioinformatics. 2009. V. 25. № 8. P. 1091–1093. https://doi.org/10.1093/bioinformatics/btp101
11. *Ребров А.П., Воскобой И.В.* Роль воспалительных инфекционных факторов в развитии атеросклероза // Терапевтич. архив. 2004. Т. 76. № 1. С. 78–82.
12. *Ng A., Tam W.W., Zhang M.W. et al.* IL-1β, IL-6, TNF-α and CRP in elderly patients with depression or Alzheimer's disease: Systematic review and meta-analysis // Sci. Reports. 2018. V. 8. № 1. 12050. https://doi.org/10.1038/s41598-018-30487-6
13. *Kishimoto T.* IL-6: From its discovery to clinical applications // Int. Immunol. 2010. V. 22. № 5. P. 347–352. https://doi.org/10.1093/intimm/dxq030
14. *Buul J.D., Hordijk P.L.* Signaling in leukocyte transendothelial migration // Arterioscler. Thromb. Vasc. Biol. 2004. V. 24. № 5. P. 824–833. https://doi.org/10.1161/01.ATV.0000122854.76267.5c
15. *Becker B.F., Heindl B., Kupatt C., Zahler S.* Endothelial function and hemostasis // Z. Kardiol. 2000. V. 89. № 3. P. 160–167. https://doi.org/10.1007/PL00007320
16. *Toapanta F.R., Ross T.M.* Complement-mediated activation of the adaptive immune responses: Role of C3d in linking the innate and adaptive immunity // Immunol. Res. 2006. V. 36. № 1–3. P. 197–210. https://doi.org/10.1385/IR:36:1:197
17. *Tsoporis J.N., Marks A., Haddad A. et al.* S100b expression modulates left ventricular remodeling after myocardial infarction in mice // Circulation. 2005. V. 111. P. 598–560. https://doi.org/10.1161/01.CIR.0000154554.65287.F5
18. *Fang F., Lue L.-F., Yan S. et al.* RAGE-dependent signaling in microglia contributes to neuroinflammation, Aβ accumulation, and impaired learning/memory in a mouse model of Alzheimer's disease // FASEB J. 2010. V. 24. № 4. P. 1043–1055. https://doi.org/10.1096/fj.09-139634
19. *Jin Q., Chen H., Luo A. et al.* S100A14 stimulates cell proliferation and induces cell apoptosis at different concentrations via receptor for advanced glycation end products (RAGE) // PLoS One. 2011. V. 6. № 4. e19375. https://doi.org/10.1371/journal.pone.0019375
20. *Wan F., Lenardo M.J.* The nuclear signaling of NF-kappaB: Current knowledge, new insights, and future perspectives // Cell Res. 2010. V. 20. № 1. P. 24–33. https://doi.org/10.1038/cr.2009.137
21. *Allan S.M., Tyrrell P.J., Rothwell N.J.* Interleukin-1 and neuronal injury // Nat. Rev. Immunol. 2005. V. 5. P. 629–640. https://doi.org/10.1038/nri1664

22. *Ikushima H., Munakata Y., Ishii T. et al.* Internalization of CD26 by mannose 6-phosphate/insulin-like growth factor II receptor contributes to T cell activation // Proc. Natl Acad. Sci. USA. 2000. V. 97. P. 8439–8444. https://doi.org/10.1073/pnas.97.15.8439

23. *Hohsfield L.A., Humpel C.* Migration of blood cells to β-amyloid plaques in Alzheimer's disease // Exp. Gerontol. 2015. V. 65. P. 8–15. https://doi.org/10.1016/j.exger.2015.03.002

24. *Jarnagin K., Grunberger D., Mulkins M. et al.* Identification of surface residues of the monocyte chemotactic protein 1 that affect signaling through the receptor CCR2 // Biochemistry. 1999. V. 38. № 49. P. 16167–16177. https://doi.org/10.1021/bi9912239

25. *Rossignol P., Ho-Tin-Noe B., Vranckx R. et al.* Protease nexin-1 inhibits plasminogen activation-induced apoptosis of adherent cells // Biol. Chem. 2004. V. 279. № 11. P. 10346–10356. https://doi.org/10.1074/jbc.m310964200

26. *Dixelius J., Olsson A.-K., Thulin A. et al.* Minimal active domain and mechanism of action of the angiogenesis inhibitor histidine-rich glycoprotein // Cancer Res. 2006. V. 66. №4. P. 2089–2097. https://doi.org/10.1158/0008-5472.CAN-05-2217

27. *Lapicka-Bodzioch K., Bodzioch M., Kruell M. et al.* Homogeneous assay based on 52 primer sets to scan for mutations of the ABCA1 gene and its application in genetic analysis of a new patient with familial high-density lipoprotein deficiency syndrome // Biochim. Biophys. Acta. 2001. V. 1537. P. 42–48. https://doi.org/10.1016/s0925-4439(01)00053-9

28. *Marcil M., Brooks-Wilson A., Clee S.M. et al.* Mutations in the ABC1 gene in familial HDL deficiency with defective cholesterol efflux // Lancet. 1999. V. 354. P. 1341–1346. https://doi.org/10.1016/s0140-6736(99)07026-9

29. *Sanchez-Mejia R.O., Mucke L.* Phospholipase A2 and arachidonic acid in Alzheimer's disease // Biochim. Biophys. Acta. 2010. V. 1801. № 8. P. 784–790. https://doi.org/10.1016/j.bbalip.2010.05.013

30. *Alphey M.S., Yu W., Byres E. et al.* Structure and reactivity of human mitochondrial 2,4-dienoyl-CoA reductase: Enzyme-ligand interactions in a distinctive short-chain reductase active site // J. Biol. Chem. 2005. V. 280. P. 3068–3077. https://doi.org/10.1074/jbc.M411069200

31. *Kuo A., Stoica G., Riegel A., Wellstein A.* Recruitment of insulin receptor substrate-1 and activation of NF-κB essential for midkine growth signaling through anaplastic lymphoma kinase // Oncogene. 2007. V. 26. P. 859–869. https://doi.org/10.1038/sj.onc.1209840

32. *Orlando B.J., Malkowski M.G.* Substrate-selective inhibition of cyclooxygeanse-2 by fenamic acid derivatives is dependent on peroxide tone // Biol. Chem. 2016. V. 291. № 29. P. 15069–15081. https://doi.org/10.1074/jbc.M116.725713

33. *Musee J., Marnett L.J.* Prostaglandin H synthase-2-catalyzed oxygenation of 2-arachidonoylglycerol is more sensitive to peroxide tone than oxygenation of arachidonic acid // Biol. Chem. 2012. V. 287. № 44. P. 37383–37394. https://doi.org/10.1074/jbc.M112.381202

34. *Suzuki T., Watanabe K., Kanaoka Y. et al.* Induction of hematopoietic prostaglandin D synthase in human megakaryocytic cells by phorbol ester // Biochem. Biophys. Res. Commun. 1997. V. 241. P. 288–293. https://doi.org/10.1006/bbrc.1997.7803

35. *Gogvadze V., Orrenius S., Zhivotovsky B.* Multiple pathways of cytochrome c release from mitochondria in apoptosis // Biochim. Biophys. Acta (BBA) – Bioenergetics. 2006. V. 1757. № 5–6. P. 639–647. https://doi.org/10.1016/j.bbabio.2006.03.016

36. *Li M.O., Wan Y.Y., Sanjabi S. et al.* Transforming growth factor-β regulation of immune responses // Annu. Rev. Immunol. 2006. V. 24. № 1. P. 99–146. https://doi.org/10.1146/annurev.immunol.24.021605.090737

37. *Miyata Y., Akashi M., Nishida E.* Molecular cloning and characterization of a novel member of the MAP kinase superfamily // Genes Cells. 1999. V. 4. P. 299–309. https://doi.org/10.1046/j.1365-2443.1999.00261.x

38. *Deane R., Sagare A., Zlokovic B.V.* The role of the cell surface LRP and soluble LRP in blood-brain barrier Abeta clearance in Alzheimer's disease // Current Pharmaceutical Design. 2008. V. 14. № 16. P. 1601–1605. https://doi.org/10.2174/138161208784705487

39. *Kino T., Manoli I., Kelkar S. et al.* Glucocorticoid receptor (GR) β has intrinsic, GRα-independent transcriptional activity // Biochem. Biophys. Res. Communications. 2009. V. 381. № 4. P. 671–675. https://doi.org/10.1016/j.bbrc.2009.02.110

40. *Solanas-Barca M., Castro-Oros I., Mateo-Gallego R. et al.* Apolipoprotein E gene mutations in subjects with mixed hyperlipidemia and a clinical diagnosis of familial combined hyperlipidemia // Atherosclerosis. 2012. V. 222. P. 449–455. https://doi.org/10.1016/j.atherosclerosis.2012.03.011

41. *Corder E.H., Saunders A.M., Strittmatter W.J. et al.* Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families // Science. 1993. V. 261. P. 921–923. https://doi.org/10.1126/science.8346443

42. *Chen Z.-Y., Ieraci A., Teng H. et al.* Sortilin controls intracellular sorting of brain-derived neurotrophic factor to the regulated secretory pathway // J. Neurosci. 2005. V. 25. № 26. P. 6156–6166. https://doi.org/10.1523/JNEUROSCI.1017-05.2005

43. *Carey A.L., Febbraio M.A.* Interleukin-6 and insulin sensitivity: friend or foe? // Diabetologia. 2004. V. 47. P. 1135–1142. https://doi.org/10.1007/s00125-004-1447-y

Bioinformatic Annotation of Genes for Alzheimer's Disease and Ischemic Heart Disease

N. Yu. Chasovskikh[a], E. E. Chizhik[a, *], and A. A. Bobrysheva[a]

[a] *Siberian State Medical University , Tomsk, 634050 Russia*

**e-mail: evgenika06@gmail.com*

In this work, we performed a functional annotation for genes for Alzheimer's disease (AD) susceptibility and for coronary heart disease (CHD) susceptibility using Cytoscape v. 3.6.0. The identified genes are involved in the immune response, apoptosis, and regulate the processes of neurogenesis and angiogenesis. Based on the results of the functional annotation, genes for AD predisposition and genes for CHD susceptibility were assigned to terms in accordance with the gene ontology (GO) and combined into groups. The number of common groups of functions in which genes for susceptibility to Alzheimer's disease and coronary heart disease were involved was 107. Common genes of susceptibility *APOE*, *APOA1*, *ABCA1*, involved in fatty acid metabolism, can potentially participate in the mechanisms of association of the studied diseases. The results obtained can serve as a prerequisite for further studies of the contribution of hereditary factors to the joint manifestation of AD and CHD.

Keywords: coronary heart disease, Alzheimer's disease, GWAS, annotation of genes, predisposition genes.

ГЕНЕТИКА, 2021, том 57, № 11, с. 1286–1296

ГЕНЕТИКА ЧЕЛОВЕКА

УДК 575:599.9

РОЛЬ ГЕНОВ НЕЙРОТРАНСМИТТЕРНОЙ СИСТЕМЫ В РАЗВИТИИ ХРОНИЧЕСКОЙ ОБСТРУКТИВНОЙ БОЛЕЗНИ ЛЕГКИХ

© 2021 г. Г. Ф. Корытина[1, 2, *], Л. З. Ахмадишина[1], О. В. Кочетова[1], Т. Р. Насибуллин[1], Ю. Г. Азнабаева[2], Ш. Р. Зулкарнеев[2], С. М. Измайлова[2], Н. Ш. Загидуллин[2], Т. В. Викторова[2]

[1]Институт биохимии и генетики – обособленное структурное подразделение Уфимского федерального исследовательского центра Российской академии наук, Уфа, 450054 Россия

[2]Башкирский государственный медицинский университет, Уфа, 450000 Россия

**e-mail: guly_kory@mail.ru*

Поступила в редакцию 16.02.2021 г.
После доработки 21.03.2021 г.
Принята к публикации 11.05.2021 г.

Курение – основной фактор риска хронической обструктивной болезни легких (ХОБЛ), многофакторного воспалительного заболевания респираторной системы. Использовали образцы ДНК больных ($N = 601$) и здоровых ($N = 617$) индивидов. Полиморфные варианты генов *GRIK3* (rs534131), *GRIN2B* (rs7301328, rs1805476), *GRIA1* (rs2195450), *GRIN1* (rs6293), *GABBR2* (rs3750344), *BDNF* (rs6265, rs11030107), *ANKK1* (rs1800497) анализировали методом полимеразной цепной реакции в реальном времени. Показана ассоциация генов *GRIK3* (rs534131) ($P = 0.009$, OR = 1.42 в доминантной модели), *GRIA1* (rs2195450) ($P = 0.015$, OR = 1.35 в доминантной модели), *GRIN1* (rs6293) ($P = 0.036$, OR = 0.79 в лог-аддитивной модели), *GRIN2B* (rs7301328) ($P = 0.0009$, OR = 0.54 в рецессивной модели) с ХОБЛ. Среди курильщиков ассоциация подтверждена для генов *GRIK3* (rs534131) ($P = 0.0001$, OR = 1.68 в доминантной модели) и *GRIN2B* (rs7301328) ($P = 0.001$, OR = 0.52 в рецессивной модели). У некурящих ассоциации с ХОБЛ выявлены для генов *GRIN1* (rs6293) ($P = 0.0001$, OR = 0.36 в доминантной модели), *GRIA1* (rs2195450) ($P = 0.0018$, OR = 2.40 в лог-аддитивной модели) и *BDNF* (rs11030107) ($P = 0.005$, OR = 2.86 для генотипа *TT*). У индивидов с генотипами *GG* гена *GRIK3* (rs534131) и *GG* гена *GRIA1* (rs2195450) установлены более низкие показатели индекса курения ($P = 0.028$ и 0.0001 соответственно). Уровень никотиновой зависимости значимо выше у носителей редкого аллеля *A* гена *GRIK3* (rs534131) и генотипа *GC* гена *GRIN2B* (rs7301328) ($P = 0.011$ и 0.023 соответственно). С помощью алгоритма APSampler получены информативные сочетания генотипов, ассоциированные с ХОБЛ, включающие *GRIN2B* rs2268132**T*, *GRIN2B* rs7301328**G*, *GRIN2B* rs1805476**C*, *ANKK1* rs1800497**G*, *GABBR2* rs3750344**A*, *CHRNA5* rs16969968**T*, *CHRNA3* rs1051730**A*, *HTR2A* rs6313**CC*, *GRIA1* rs2195450**G*.

Ключевые слова: хроническая обструктивная болезнь легких, рецепторы глутамата, никотиновая зависимость, курение, ген-средовые взаимодействия.

DOI: 10.31857/S0016675821110060

Хроническая обструктивная болезнь легких (**ХОБЛ**) – многофакторное хроническое гетерогенное воспалительное заболевание респираторной системы, с преимущественным поражением дистальных отделов дыхательных путей и легочной паренхимы, которое является четвертой ведущей причиной смерти в мире, вызывая более трех миллионов смертей ежегодно [1]. Курение – основной фактор риска развития многих заболеваний, в том числе и ХОБЛ [2]. С табакокурением связывают появление патологических процессов в легких и развитие системных воспалительных реакций, окислительный стресс, дисфункцию эндотелия сосудов, возрастание активности прокоагулянтных факторов [1]. Табачная зависимость стала серьезной проблемой общественного здравоохранения; никотин является основным компонентом табака, вызывающим привыкание, и у большинства регулярно курящих людей развивается никотиновая зависимость [2]. Получены убедительные данные о вкладе генетических факторов в развитие никотиновой зависимости [3]. Полногеномные ассоциативные исследования и метаанализы позволили выявить ряд локусов, связанных с развитием никотиновой зависимости и количественными показателями, характеризующими поведение курильщика: на хромосоме 15q25 – кластер генов *CHRNA3/CHRNA5/CHRNB4*, кодирующих холинергические никотиновые рецепторы; LOC100188947 на хромосоме 10q25; в гене *EGLN2* на хромосоме 9q13;

на хромосоме 11p13 в гене *BDNF* [4]. К настоящему времени проведено много исследований по идентификации наследственных основ ХОБЛ; было показано, что ряд генов ассоциируют как с ХОБЛ, так и с никотиновой зависимостью [5]. Для некоторых SNP в этих генах, вовлеченных в нейротрансмиссию, выявлены значимые взаимодействия с курением, что указывает на важную роль генов нейротрансмиттеров в приверженности курению и формированию никотиновой зависимости [6]. Ранее нами было показано, что полиморфные варианты генов холинэргических никотиновых *CHRNA5* (rs16969968), *CHRNA3* (rs1051730), глутаматных *GRIK5* (rs8099939), *GRIN2B* (rs2268132) и серотонинового *HTR2A* (rs6313) рецепторов связаны с развитием ХОБЛ у курильщиков в популяции татар [7, 8]. В настоящей работе мы расширили список генов-кандидатов, включив в него гены, кодирующие ионотропный каинатный рецептор 3-го типа (*GRIK3*), ионотропный AMPA-рецептор (*GRIA1*), NMDA-рецептор глутамата тип 1 (*GRIN1*), рецептор гамма-аминомасляной кислоты B (*GABBR2*), нейротрофический фактор (*BDNF*) и протеинкиназу (*ANKK1*). Цель настоящего исследования – выявление ассоциации полиморфных вариантов генов *GRIK3*, *GRIN2B*, *GRIA1*, *GRIN1*, *GABBR2, BDNF* и *ANKK1* с ХОБЛ, анализ их вклада в вариабельность показателей, характеризующих уровень никотиновой зависимости и индекс курения, оценки взаимодействия генетических и средовых факторов при формировании ХОБЛ.

МАТЕРИАЛЫ И МЕТОДЫ

Дизайн исследования – кандидатное исследование по принципу случай–контроль. Использовали образцы ДНК неродственных индивидов, татар по этнической принадлежности, проживающих на территории Республики Башкортостан. Группа больных включала 601 индивида (из них 522 мужчин (86.85%) и 79 женщин (13.15%)), средний возраст составил 63.38 ± 11.81 лет. Среди больных ХОБЛ курильщиков и бывших курильщиков – 484 человека (80.53%), некурящих 117 (19.47%). Индекс курения у курильщиков и бывших курильщиков составил 44.58 ± 25.92 пачек/лет. Группа контроля включала 617 индивидов (из них 548 мужчин (88.88%) и 69 женщин (11.12%)), средний возраст составил 58.44 ± 14.79, курильщики и бывшие курильщики – 517 (83.79%) и некурящие – 100 (16.21%)); индекс курения у курильщиков составлял 38.54 ± 23.12 пачек/лет. Критерии включения и исключения в группу ХОБЛ и контроля, оценка уровня никотиновой зависимости с использованием теста Фагерстрема описаны нами ранее [8]. Исследование одобрено комитетом по этике ИБГ УНЦ РАН.

Генотипирование

ДНК выделяли из лейкоцитов периферической крови с использованием фенольно-хлороформной экстракции. Для анализа нами были выбраны полиморфные локусы генов, имеющие функциональную значимость и/или ранее связанные с формированием зависимости от психоактивных веществ и психическими заболеваниями. Учитывалась частота редкого аллеля в популяциях европеоидов по данным базы National Center for Biotechnology Information [9]. Для нашего исследования были выбраны следующие полиморфные локусы: *GRIK3* (rs534131, g.20608*A>G*), *GRIN2B* (rs7301328, c.366*C>G*, p.Pro122=), *GRIN2B* (rs1805476, c.*1354*C>A*), *GRIA1* (rs2195450, c.–750*G>A*), *GRIN1* (rs6293, c.789*A>G*, p.Pro263=), *GABBR2* (rs3750344, c.360*A>G*, p.Ala120=), *BDNF* (rs6265, c.283*G>A*, p.Val66Met), *BDNF* (rs11030107, c.–21-14703*T>C*), *ANKK1* (rs1800497, c.2137*G>A*, p.Glu713Lys). Функциональная значимость выбранных для исследования полиморфных локусов исследовалась по базам RegulomeDB Version 1.1 (https://regulomedb.org), SNPinfo Web Server (https://snpinfo.niehs.nih.gov) и HaploReg v3 [10]. Полиморфные варианты генов анализировали при помощи полимеразной цепной реакции (ПЦР) в реальном времени коммерческими наборами с флуоресцентной детекцией (https://www.oligos.ru, ООО “ДНК-Синтез”, Россия) на приборе BioRad CFX96™ (Bio-Rad Laboratories, Inc., США). Подробно методы анализа описаны нами ранее [7].

Статистическая обработка результатов

Статистическую обработку данных проводили, используя пакеты прикладных программ Statistica v 6.0 (StatSoft Inc., США) и PLINK v 1.07 [11]. Подробное описание методов статистического анализа приведено нами ранее [7, 8]. Вклад аллельных вариантов изучаемых генов-кандидатов в вариабельность количественных признаков (индекс курения и уровень никотиновой зависимости (FTND)) определяли с помощью линейной регрессии. Анализ ассоциаций сочетаний аллелей/генотипов с ХОБЛ осуществляли с помощью программы APSampler 3.6.1 (http://sourceforge.net/projects/apsampler/). Основной алгоритм описан в статье [12]. Поправку на множественное тестирование проводили с помощью метода оценки доли ложноположительных результатов FDR (False Discovery Rate) (Benjiamini Hochberg), используя программу (http//www.sdmproject.com/utilinies/?show=FDR) и получали новое значение $P_{FDR\text{-}cor}$.

РЕЗУЛЬТАТЫ

Прежде чем приступить к анализу ассоциации аллельных вариантов генов-кандидатов с развитием ХОБЛ была проведена проверка соответ-

ствия распределения частот генотипов равновесию Харди–Вайнберга. Для группы контроля были получены следующие результаты: *GRIK3* (rs534131) (P_{X-B} = 0.082), *GRIN2B* (rs7301328) (P_{X-B} = 0.66), *GRIN2B* (rs1805476) (P_{X-B} = 0.11), *GRIA1* (rs2195450) (P_{X-B} = 0.15), *GRIN1* (rs6293) (P_{X-B} = 0.14), *GABBR2* (rs3750344) (P_{X-B} = 0.08), *BDNF* (rs6265) (P_{X-B} = = 0.67), *BDNF* (rs11030107) (P_{X-B} = 0.16), *ANKK1* (rs1800497) (P_{X-B} = 0.88). В табл. 1 представлены данные по распределению частот генотипов и аллелей изученных локусов, значимость различий между группами, показатели отношения шансов, рассчитанные для редкого аллеля каждого локуса; в табл. 2 – статистически значимые результаты регрессионного анализа. Частота редкого аллеля *A* гена *GRIK3* (rs534131) была значимо выше в группе больных ХОБЛ (P = 0.017, OR = 1.21 95%CI 1.04–1.57). Ассоциация гена *GRIK3* (rs534131) с ХОБЛ была установлена в доминантной (P = 0.009, $P_{cor-FDR}$ = 0.044, OR = 1.42 95%CI 1.10–1.83) и лог-аддитивной (P = 0.018, $P_{cor-FDR}$ = 0.044, OR = 1.25 95%CI 1.07–1.51) моделях. Частота редкого аллеля *A* гена *GRIA1* (rs2195450) была значимо выше в группе больных ХОБЛ (22.13% против 18.15% в контроле, P = 0.016, OR = 1.28 95%CI 1.05–1.56). Ген *GRIA1* (rs2195450) ассоциировал с ХОБЛ в доминантной (P = 0.015, $P_{cor-FDR}$ = 0.044, OR = 1.35 95%CI 1.07–1.71) и лог-аддитивной (P = 0.02, $P_{cor-FDR}$ = 0.044, OR = 1.26 95%CI 1.01–1.57) моделях. Выявлены значимые различия по распределению частот аллелей полиморфного локуса *GRIN1* (rs6293) между исследованными группами (P = 0.04, OR = = 0.81 95%CI 0.66–0.98). Регрессионный анализ показал ассоциацию гена *GRIN1* (rs6293) с ХОБЛ в доминантной (P = 0.047, $P_{cor-FDR}$ = 0.0705, OR = 0.78 95%CI 0.62–0.98) и лог-аддитивной (P = 0.036, $P_{cor-FDR}$ = 0.061, OR = 0.79 95%CI 0.63–0.99) моделях, но после FDR-коррекции различия не достигали статистической значимости. В группе ХОБЛ частота редкого аллеля *C* гена *GRIN2B* (rs7301328) была значимо ниже, чем в контроле (36.36% против 41.17% в контроле, P = 0.021, OR = 0.81 95%CI 0.69–0.96). Ассоциация гена *GRIN2B* (rs7301328) с ХОБЛ была установлена в рецессивной (P = 0.0009, $P_{cor-FDR}$ = 0.0108, OR = 0.54 95%CI 0.37–0.78) и лог-аддитивной модели (P = 0.022, $P_{cor-FDR}$ = 0.044, OR = 0.81 95%CI 0.68–0.98); гомозиготный по редкому аллелю генотип *CC* гена *GRIN2B* (rs7301328) является маркером устойчивости к ХОБЛ. Сравнительный анализ распределения частот генотипов и аллелей между группами больных ХОБЛ и контроля по полиморфным вариантам генов *GRIN2B* (rs1805476), *GABBR2* (rs3750344), *BDNF* (rs6265), *BDNF* (rs11030107), *ANKK1* (rs1800497) статистически значимых различий не дал (табл. 1).

Анализ ассоциации генов-кандидатов с развитием ХОБЛ в группах, дифференцированных по статусу курения

В группе курильщиков ассоциация была подтверждена для генов *GRIK3* (rs534131) и *GRIN2B* (rs7301328) (см. табл. 3). Так же как и в общей группе, ген *GRIK3* (rs534131) значимо ассоциировал с ХОБЛ у курильщиков в доминантной (P = 0.0001, $P_{cor-FDR}$ = 0.00065, OR = 1.68 95%CI 1.27–2.24) и лог-аддитивной (P = 0.003, $P_{cor-FDR}$ = 0.0078, OR = = 1.31 95%CI 1.10–1.56) моделях. Ассоциация с развитием заболевания и геном *GRIN2B* (rs7301328) в группе курильщиков была показана в рецессивной (P = 0.001, $P_{cor-FDR}$ = 0.004, OR = 0.52 95%CI 0.36–0.77) и лог-аддитивной (P = 0.021, $P_{cor-FDR}$ = = 0.027, OR = 0.80 95%CI 0.67–0.96) моделях. Регрессионный анализ установил ассоциацию гена *GRIN1* (rs6293) с ХОБЛ у некурящих индивидов в доминантной (P = 0.0001, $P_{cor-FDR}$ = 0.00065, OR = = 0.36 95%CI 0.20–0.63) и лог-аддитивной (P = 0.02, $P_{cor-FDR}$ = 0.027, OR = 0.51 95%CI 0.31–0.84) моделях; риск развития заболевания связан с генотипом *AA* гена *GRIN1* (rs6293) (OR = 2.8 95%CI 1.57–4.97). Ген *GRIA1* (rs2195450) значимо ассоциировал с ХОБЛ у некурящих индивидов в доминантной (P = 0.0082, $P_{cor-FDR}$ = 0.015, OR = 2.57 95%CI 1.27–5.19), рецессивной (P = 0.0096, $P_{cor-FDR}$ = 0.015, OR = 6.40 95%CI 1.31–31.36) и лог-аддитивной (P = 0.0018, $P_{cor-FDR}$ = 0.005, OR = 2.40 95%CI 1.35–4.25) моделях. Значимые ассоциации с ХОБЛ только среди некурящих индивидов были получены для генотипа *TT* гена *BDNF* (rs11030107) (P = 0.005 $P_{cor-FDR}$ = 0.01, OR = 2.86 95%CI 1.40–5.81).

У гомозигот *GG* по гену *GRIA1* (rs2195450) установлены более низкие показатели индекса курения (29.54 пачек/лет, P = 0.028) (см. табл. 4). У гетерозигот и гомозигот по редкому аллелю *A* гена *GRIK3* (rs534131) показано значимое увеличение индекса курения (32.88 пачек/лет, P = 0.0001). Показатель, характеризующий уровень никотиновой зависимости, рассчитанный по шкале Фагерстрема, также был значимо выше (FTND = 5.51, P = 0.011) у носителей редкого аллеля *A* гена *GRIK3* (rs534131) и гетерозигот *GC* по гену *GRIN2B* (rs7301328) (FTND = 5.44, P = 0.023).

Анализ сочетаний аллелей и/или генотипов исследованных генов с развитием ХОБЛ

При помощи программы APSampler (Allelic Pattern Sampler) нами проведен поиск информативных предикторов развития ХОБЛ. В анализ помимо девяти исследованных полиморфных локусов были включены еще восемь, изученных нами ранее: *CHRNA5* (rs16969968), *CHRNA3* (rs1051730, rs6495309), *CHRNB4* (rs1948), *HTR4* (rs3995090), *HTR2A* (rs6313), *GRIK5* (rs8099939), *GRIN2B* (rs2268132) [7, 8]. Выявлены генетические паттерны,

Таблица 1. Распределение частот генотипов и аллелей полиморфных локусов генов нейротрансмиттеров в группах ХОБЛ и контроле

Ген, полиморфный локус	Редкий аллель	Генотипы, аллели	ХОБЛ *n* (%) (*N* = 601)	Контроль *n* (%) (*N* = 617)	*P*	OR (95%CI)
GRIK3 rs534131 *A>G*	*A*	*GG/GA/AA*	141/341/119 (23.46/56.74/19.80)	187/326/104 (30.31/52.84/16.86)	**0.022**	–
		G/A	623/579 (51.83/48.17)	700/534 (56.73/43.27)	**0.017**	1.21 (1.04–1.57)
GRIA1 rs2195450 *G>A*	*A*	*GG/GA/AA*	367/202/32 (61.06/33.61/5.32)	419/172/26 (67.91/27.88/4.21)	**0.044**	–
		G/A	936/266 (77.87/22.13)	1010/224 (81.85/18.15)	**0.016**	1.28 (1.05–1.56)
GRIN1 rs6293 *A>G*	*G*	*AA/AG/GG*	386/198/17 (64.23/32.95/2.83)	361/231/25 (58.51/37.44/4.05)	0.096	–
		A/G	970/232 (80.70/19.30)	953/281 (77.23/22.77)	**0.04**	0.81 (0.66–0.98)
GRIN2B rs7301328 *C>G*	*C*	*GG/GC/CC*	225/315/61 (37.44/52.41/10.15)	217/292/108 (35.17/47.33/17.50)	**0.001**	–
		G/C	765/437 (63.64/36.36)	726/508 (58.83/41.17)	**0.021**	0.81 (0.69–0.96)
GRIN2B rs1805476 *C>A*	*C*	*AA/AC/CC*	163/265/173 (27.12/44.09/28.79)	170/286/161 (27.55/46.35/26.09)	0.55	–
		A/C	591/611 (49.17/50.83)	626/608 (50.73/49.27)	0.465	1.06 (0.91–1.24)
GABBR2 rs3750344 *A>G*	*G*	*AA/AG/GG*	403/170/28 (67.05/28.29/4.66)	399/183/35 (64.67/29.66/5.67)	0.587	–
		A/G	976/226 (81.20/18.80)	981/253 (79.50/20.50)	0.315	0.89 (0.74–1.09)
BDNF rs6265 *G>A*	*A*	*GG/GA/AA*	446/141/14 (74.21/23.46/2.33)	457/147/13 (74.07/23.82/2.11)	0.958	–
		G/A	1033/169 (85.94/14.06)	1061/ 173 (85.98/14.02)	0.975	1.00 (0.79–1.26)
BDNF rs11030107 *T>C*	*C*	*TT/TC/CC*	461/136/4 (76.71/22.63/0.67)	475/138/4 (76.99/22.37/0.65)	0.993	–
		T/C	1058/144 (88.02/11.98)	1088/146 (88.17/11.83)	0.960	1.01 (0.79–1.29)
ANKK1 rs1800497 *G>A*	*A*	*GG/GA/AA*	336/237/28 (55.91/39.43/4.66)	363/220/34 (58.83/35.66/5.51)	0.359	–
		G/A	909/293 (75.62/24.38)	946/288 (76.66/23.34)	0.580	1.05 (0.83–1.27)

Примечание. *N* – число индивидов, включенных в анализ; *P* – значимость различий между группами по частотам аллелей и генотипов (тест χ^2 на гомогенность выборок); OR – показатель отношения шансов для редкого аллеля (базовый аллельный тест).

Таблица 2. Статистически значимые результаты анализа ассоциации полиморфных локусов генов-кандидатов с развитием ХОБЛ (лог-регрессионный анализ)

Ген, полиморфный локус	Редкий аллель	Модель	ХОБЛ *n* (%) (*N* = 601)	Контроль *n* (%) (*N* = 617)	OR_{adj} (95%CI)	P_{adj}	$P_{cor\text{-}FDR}$
GRIK3 rs534131	*A*	*GG* *GA* + *AA* Доминантная	141 (23.46) 460 (76.54)	187 (30.31) 430 (69.69)	1.00 1.42 (1.10–1.83)	0.009	0.044
		GG + *GA* *AA* Рецессивная	482 (80.20) 119 (19.80)	513 (83.16) 104 (16.86)	1.00 1.22 (0.88–1.70)	0.21	0.264
		Лог-аддитивная	–	–	1.25 (1.07–1.51)	0.018	0.044
GRIA1 rs2195450	*A*	*GG* *GA* + *AA* Доминантная	367 (61.06) 234 (38.94)	419 (67.91) 198 (32.09)	1.00 1.35 (1.07–1.71)	0.015	0.044
		GG + *GA* *AA* Рецессивная	569 (94.68) 32 (5.32)	591 (95.79) 26 (4.21)	1.00 1.27 (0.75–2.17)	0.438	0.445
		Лог-аддитивная	–	–	1.26 (1.01–1.57)	0.02	0.044
GRIN1 rs6293	*G*	*AA* *AG* + *GG* Доминантная	386 (64.23) 215 (35.77)	361 (58.51) 256 (41.49)	1.00 0.78 (0.62–0.98)	0.047	0.0705
		AA + *AG* *GG* Рецессивная	584 (97.17) 17 (2.83)	592 (95.95) 25 (4.05)	1.00 0.65 (0.32–1.30)	0.22	0.264
		Лог-аддитивная	–	–	0.79 (0.63–0.99)	0.036	0.061
GRIN2B rs7301328	*C*	*GG* *GC* + *CC* Доминантная	225(37.44) 376 (62.56)	217 (35.17) 400 (64.83)	1.00 0.91 (0.70–1.17)	0.445	0.445
		GG + *GC* *CC* Рецессивная	540 (89.85) 61(10.15)	509 (82.50) 108(17.50)	1.00 0.54 (0.37–0.78)	0.0009	0.0108
		Лог-аддитивная	–	–	0.81 (0.68–0.98)	0.022	0.044

Примечание. P_{adj} – значимость для теста отношения правдоподобия лог-регрессионной модели с учетом возраста, индекса массы тела, пола; OR_{adj} – отношение шансов с учетом всех факторов, 95%CI – 95%-ный доверительный интервал для OR; $P_{cor\text{-}FDR}$ – значимость теста после коррекции FDR; лог-аддитивная модель на дозу редкого аллеля – увеличение дозы редкого аллеля в ряду: гомозигота по частому аллелю (0)–гетерозигота (1)–гомозигота по редкому аллелю (2).

значимо ассоциированные с ХОБЛ; в табл. 5 отражены результаты наиболее значимых комбинаций с P_{FDR} менее 0.05 и OR более 2.5 для комбинаций риска или OR менее 0.33 для протективных комбинаций. Большинство наиболее значимых комбинаций, которые ассоциировали с риском развития ХОБЛ, включали генотип *GRIN2B* rs2268132**TT* или аллель *GRIN2B* rs2268132**T*, *GRIN2B* rs7301328**G*, а также *GRIN2B* rs1805476**C*. Аллель *GRIN2B* rs2268132**G* являлся обязательной частью комбинаций, имеющих протективный эффект. Локусы *ANKK1* rs1800497**G*, *GABBR2* rs3750344**A* встречались в нескольких комбинациях риска развития ХОБЛ. Комбинации риска включали гены *CHRNA5* rs16969968**T*, *CHRNA3* rs1051730**A* и *CHRNB4* rs1948**C* и гены *HTR2A* rs6313**CC* и *HTR4* rs3995090**A*, с которыми нами ранее были выявлены ассоциации с ХОБЛ в нашей выборке [7, 8]. Аллель *GRIK3* rs534131**G* встречался только в сочетаниях пониженного риска ХОБЛ. Анализ сочетаний аллелей/генотипов исследованных полиморфных локусов позволил вывить ассоциацию полиморфных локусов генов *ANKK1* (rs1800497), *GABBR2* (rs3750344), которые проявляли свой эффект только в комбинации с генами глутаматных и холинэргических никотиновых рецепторов.

Таблица 3. Статистически значимые результаты анализа ассоциации полиморфных локусов генов-кандидатов с развитием ХОБЛ в группах, дифференцированных по статусу курения

Ген, SNP	Редкий аллель	Модель	ХОБЛ абс. (%)	Контроль абс. (%)	OR (95%CI)	*P*	$P_{cor\text{-}FDR}$
Курильщики			(*N* = 484)	(*N* = 517)			
GRIK3 rs534131	*A*	*GG* *GA* + *AA* Доминантная	106 (21.90) 378 (78.10)	166 (32.11) 351 (67.89)	1.00 1.68 (1.27–2.24)	0.0001	0.00065
		GG + *GA* *AA* Рецессивная	388 (80.17) 96 (19.83)	430 (83.17) 87(16.83)	1.00 1.21 (0.84–1.72)	0.25	0.27
		Лог-аддитивная	–	–	1.31 (1.10–1.56)	0.003	0.0078
GRIN2B rs7301328	*C*	*GG* *GC* + *CC* Доминантная	180 (37.19) 304 (62.81)	176 (34.04) 341 (65.96)	1.00 0.87 (0.66–1.15)	0.34	0.34
		GG + *GC* *CC* Рецессивная	438 (90.50) 46 (9.50)	431 (83.37) 86 (16.63)	1.00 0.52 (0.36–0.77)	0.001	0.004
		Лог-аддитивная	–	–	0.80 (0.67–0.96)	0.021	0.027
Некурящие			(*N* = 117)	(*N* = 100)			
GRIN1 rs6293	*G*	*AA* *AG* + *GG* Доминантная	88 (75.21) 29 (24.79)	52 (52.00) 48 (48.00)	1.00 0.36 (0.20–0.63)	0.0001	0.00065
		AA + *AG* *GG* Рецессивная	113 (96.58) 4(3.42)	100 (100.00) 0	1.00 NA (0.00–NA)	0.037	0.043
		Лог-аддитивная	–	–	0.51 (0.31–0.84)	0.02	0.027
GRIA1 rs2195450	*A*	*GG* *GA* + *AA* Доминантная	59 (50.42) 58 (49.58)	72 (72.00) 28 (28.00)	1.00 2.57 (1.27–5.19)	0.0082	0.015
		GG + *GA* *AA* Рецессивная	101 (86.32) 116 (13.68)	98 (98.00) 2 (2.00)	1.00 6.40 (1.31–31.36)	0.0096	0.015
		Лог-аддитивная	–	–	2.40 (1.35–4.25)	0.0018	0.005
BDNF rs11030107	*C*	*TT* *TC*	103 (88.03) 14 (11.97)	72 (72.00) 28 (28.00)	2.86 (1.40–5.81) 0.34 (0.17–0.71)	0.005	0.010

ОБСУЖДЕНИЕ

В результате проведенного исследования нами установлена ассоциация полиморфных вариантов генов глутаматных рецепторов *GRIK3*, *GRIN2B*, *GRIN1* и *GRIA1* с развитием ХОБЛ в общей группе, в группах, дифференцированных по статусу курения, с индексом курения и уровнем никотиновой зависимости. Риск развития ХОБЛ в нашем исследовании был связан с редким аллелем *A* гена *GRIK3* (rs534131). Данная ассоциация была подтверждена только в группе курильщиков; более того, у носителей редкого аллеля *A* индекс курения и уровень никотиновой зависимости были значимо выше. Альтернативный алллель *GRIK3* rs534131**G* входил в состав сочетаний, связанных с пониженным риском развития ХОБЛ, с генами холинэргических никотиновых (*CHRNA5* и *CHRNA3*) и серотонинового (*HTR2A*) рецепторов. Ген *GRIK3* кодирует глутаматный ионотропный каинатный рецептор 3-го типа [9]. Согласно данным базы HaploReg v4.1, полиморфный локус *GRIK3* (rs534131) располагается в регионе ДНК, связывающегося с регуляторными белками. Полиморфные варианты этого гена ассоциируют с развитием

Таблица 4. Вклад генотипов полиморфных локусов генов-кандидатов в вариабельность показателя индекса курения и уровень никотиновой зависимости

Ген, полиморфный локус	Генотип	*n*	*M* ± S.E	*P* [a]	beta (95%CI)
Индекс курения (пачки/лет) в общей группе курильщиков (*N* = 1001)					
GRIA1 rs2195450	*GG*	653	29.54 (1.04)	0.028	0.00
	GA + *AA*	348	33.56 (1.58)		4.02 (0.44–7.61)
GRIK3 rs534131	*GG*	277	25.12 (1.38)	0.0001	0.00
	GA + *AA*	724	32.88 (1.07)		7.76 (3.97–11.55)
	GG + *AA*	458	27.69 (1.12)	0.0013	0.00
	GA	543	33.31 (1.29)		5.62 (2.21–9.04)
Уровень никотиновой зависимости по шкале Фагерстрема FTND (*N* = 1218)					
GRIK3 rs534131	*GG*	329	4.99 (0.16)	0.011	0.00
	GA + *AA*	889	5.51 (0.11)		0.51 (0.12–0.91)
GRIN2B rs7301328	*GG* + *GC*	1045	5.42 (0.09)	0.042	0.00
	CC	173	5.00 (0.24)		–0.16 (–0.31...–0.01)
	GG + *CC*	613	5.28 (0.12)	0.023	0.00
	GC	605	5.44 (0.12)		0.12 (0.02–0.23)

Примечание. *M* ± S.E – средние значения и стандартная ошибка среднего; *P*[a] – уровень значимости для уравнения регрессии; beta (95%CI) – коэффициент регрессии и 95%-ный доверительный интервал для коэффициента.

шизофрении, алкогольной зависимости, серьезными депрессивными расстройствами и суицидным поведением [13]. Вклад полиморфных вариантов гена *GRIK3* в развитие ХОБЛ не изучали. Маркером устойчивости к развитию ХОБЛ является генотип *CC* гена *GRIN2B* (rs7301328), кодирующего глутаматный ионотропный N-метиласпартатный рецептор. Значимость ассоциации была подтверждена в группе курильщиков, у индивидов с генотипом *CC* установлены более низкие показатели никотиновой зависимости. Среди полученных нами сочетаний полиморфных вариантов исследованных генов-кандидатов, ассоциированных с повышенным риском ХОБЛ, аллель *GRIN2B* rs7301328**G* встречался в четырех наиболее значимых комбинациях, и самыми информативными были сочетания с ранее изученным нами [8] локусом гена *GRIN2B* (rs2268132) и генами *ANKK1* (rs1800497) и *GABBR2* (rs3750344). В исследовании Grucza et al. (2010) показана ассоциация SNP гена *GRIN2B* с формированием никотиновой зависимости и возрастом начала курения [14]. В работе Vink et al. (2009) была установлена ассоциация нескольких генов, вовлеченных в глутаматный сигналинг с возрастом начала курения [15]. Ассоциации с развитием ХОБЛ установлены с полиморфными вариантами гена *GRIA1* (rs2195450), кодирующего глутаматный ионотропный AMPA-рецептор, который локализован на участке 5q33.2 [9]. GRIA1 в основном экспрессируется в переднем мозге и гиппокампе, областях, которые вовлечены в формирование памяти [9]. Риск развития ХОБЛ связан с редким аллелем *A* гена *GRIA1* (rs2195450). Ассоциация с ХОБЛ была подтверждена в группе некурящих индивидов. С другой стороны, у носителей редкого аллеля *A* регистрируются более высокие показатели индекса курения. Согласно данным RegulomeDB Version 1.1 и SNPinfo Web Server (https://snpinfo.niehs.nih.gov) локус *GRIA1* (rs2195450) локализован в 2KB области и содержит сайты связывания для нескольких транскрипционных факторов. Исследований, посвященных анализу ассоциации гена *GRIA1* с развитием ХОБЛ или никотиновой зависимости, ранее не проводилось. Ассоциация с развитием ХОБЛ в группе некурящих индивидов была установлена для полиморфного локуса гена *GRIN1* (rs6293*A*>*G*), кодирующего NMDA-рецептор глутамата типа 1, локализованного на хромосоме 9q34.3 [9]. Маркером риска является гомозиготный по частому аллелю генотип *AA* гена *GRIN1* (rs6293). Функциональный анализ показал, что по данным RegulomeDB Version 1.1 *GRIN1* (rs6293) имеет регуляторный ранг 1f и коэффициент 0.55436, что указывает на влияние данного полиморфизма на экспрессию гена. По данным HaploReg v3 и SNPinfo Web Server (https://snpinfo.niehs.nih.gov) полиморфный локус rs6293 расположен в регионе ДНК, связывающегося с гистонами, маркирующими энхансеры и промоторы (GM12878, K562), и участке ДНК, который связывается с регуляторным белком ZNF263 и гиперчувствителен к ДНКазе-1 в 30 различных тканях. В исследовании Orihara et al. (2018) пока-

Таблица 5. Сочетания аллелей и/или генотипов полиморфных локусов генов нейротрансмиттерной системы, значимо ассоциированные с ХОБЛ

Сочетания	ХОБЛ, %	Контроль, %	P_{value}	P_{FDR}	OR	95%CI
Рисковые						
GRIN2B rs2268132**TT* + *GRIN2B* rs7301328**G* + *ANKK1* rs1800497**G*	14.84	1.32	8.09e-11	1.35e-08	13.02	4.62–36.66
GRIN2B rs2268132**TT* + *GRIN2B* rs7301328**G*	14.24	1.58	3.04e-10	3.47e-08	10.46	4.10–26.67
CHRNA5 rs16969968**T* + *HTR2A* rs6313**CC*	15.84	2.89	3.78e-09	1.82e-07	6.74	3.16–4.40
GRIN2B rs7301328**G* + *HTR2A* rs6313**CC*	28.36	10.87	6.57e-09	2.95e-07	3.24	2.135–4.92
GRIN2B rs2268132**T* + *GRIN2B* rs7301328**G* + *ANKK1* rs1800497**G* + *GABBR2* rs3750344**A*	46.58	26.07	9.68e-08	2.79e-06	2.57	1.75–3.47
GRIN2B rs2268132**T* + *GRIN2B* rs7301328**G* + *GABBR2* rs3750344**A* + *BDNF* rs11030107**T*	48.97	29.49	9.36e-07	1.78e-05	2.59	1.63–3.22
GRIN2B rs2268132**T* + *GRIN2B* rs1805476**C* + *GABBR2* rs3750344**A* + *CHRNA5* rs16969968**T* + *CHRNA3* rs1051730**A*	21.51	7.31	1.12e-05	0.00013	3.47	1.90–6.33
GRIN2B rs2268132**T* + *ANKK1* rs1800497**G* + *CHRNA5* rs16969968**T* + *GRIK5* rs8099939**C*	19.62	8.22	2.71e-05	0.0002	2.96	1.72–5.11
GRIN2B rs2268132**T* + *GRIN2B* rs1805476**C* + *CHRNA5* rs16969968**T* + *GRIA1* rs2195450**G*	24.15	9.80	3.3e-05	0.0003	2.92	1.70–5.02
GRIN2B rs1805476**C* + *CHRNA3* rs1051730**A* + *HTR4* rs3995090**A* + *GRIA1* rs2195450**G*	19.44	6.63	3.54e-05	0.0003	3.39	1.81–6.34
GRIN2B rs2268132**T* + *CHRNA5* rs16969968**T* + *CHRNB4* rs1948**C* + *GRIK5* rs8099939**C* + *GRIA1* rs2195450**G*	18.24	6.45	5.69e-05	0.0004	3.23	1.74–6.01
Протективные						
GRIN2B rs2268132**G* + *CHRNA5* rs16969968**C* + *HTR2A* s6313**T*	48.20	78.73	$6.56e^{-15}$	1.64e-11	0.251	0.17–0.36
CHRNA5 rs16969968**C* + *CHRNA3* rs1051730**G* + *HTR2A* rs6313**T* + *GRIK3* rs534131**G*	43.29	75.21	3.31e-13	4.16e-10	0.252	0.17–0.37
GRIN2B rs2268132**G* + *CHRNA3* rs1051730**G* + *HTR2A* rs6313**T*	47.53	75.24	4.57e-13	3.83e-10	0.298	0.21–0.41
CHRNA5 rs16969968**C* + *HTR2A* rs6313**T* + *GRIK3* rs534131**G*	45.79	75.90	1.2e-12	3.01e-09	0.268	0.18–0.39

Примечание. P_{value} – уровень значимости по тесту Фишера; P_{FDR} – значимость теста после коррекции FDR; OR – отношение шансов, 95%CI – 95%-ный доверительный интервал для OR.

зано, что CD4+ Т-клетки активируют функциональные NMDA-рецепторы, это влияет на продукцию цитокинов, пролиферацию и жизнеспособность клеток [16]. Возможно, действие глутамата на иммунокомпетентные клетки может играть важную роль в патогенезе различных заболеваний, связанных с системным воспалением, в том числе и ХОБЛ.

Нами установлено, что генотип *TT* полиморфного локуса rs11030107 гена *BDNF* статистически значимо ассоциирует с развитием ХОБЛ у некурящих индивидов. Аллель *BDNF* rs11030107**T* входит в состав информативной комбинации риска развития ХОБЛ в сочетании с аллелями генов *GRIN2B* (rs2268132, rs7301328) и *GABBR2* (rs3750344). Ген *BDNF* кодирует нейротрофический фактор мозга, локализован на хромосоме 11p13 [9]. BDNF является членом семейства нейротрофинов, играющих ключевую роль в регуляции нейрогенеза и нейропластичности в целом [9]. Согласно данным функционального анализа rs11030107 имеет регуляторный ранг 2b, что указывает на влияние данного полиморфизма на экспрессию гена, по данным онлайн-ресурса GTEx (https://www.gtexportal.org) rs11030107 связан с изменением экспрессии гена в различных тканях, в том числе и в легочной, которая значимо выше у гомозигот *TT*. Ранее было показано, что *BDNF* (rs6265) связан с никотиновой зависимостью и ассоциирует с более высокими уровнями сывороточного BDNF [17]. Ohmoto et al. (2019) в своем исследовании показали ассоциацию *BDNF* (rs6265) с никотиновой зависимостью и возрастом начала курения в популяции Японии [18].

Ген *GABBR2*, локализованный на хромосоме 9q22.1-q22.3, кодирует рецептор гамма-аминомасляной кислоты, главного тормозящего нейромедиатора, участвующего в регуляции многих физиологических и психологических процессов в головном мозге [9]. В исследовании Beuten et al. (2005) было показано, что полиморфные варианты гена *GABAB2*, в том числе исследованный нами rs3750344, ассоциируют с никотиновой зависимостью в популяциях афро-американцев и европеоидов [19]. Имеются убедительные доказательства вовлеченности генов ГАМКергического сигнального пути в развитие никотиновой и алкогольной зависимости [20]. Нами не было выявлено ассоциации *GABBR2* (rs3750344) с развитием ХОБЛ, но полиморфный вариант *GABBR2* rs3750344**A* входил в состав трех значимых комбинаций риска развития ХОБЛ в сочетании с аллелями генов *GRIN2B* rs2268132**T*, *GRIN2B* rs7301328**G*, *ANKK1* rs1800497**G*, *BDNF* rs11030107**T*, *CHRNA5* rs16969968**T* и *CHRNA3* rs1051730**A*. Ранее исследований, посвященных анализу ассоциации гена *GABBR2* с развитием ХОБЛ, не проводилось.

Дофаминергическая система играет решающую роль в развитии различных видов зависимостей, в том числе никотиновой; наиболее изученным геном является ген *DRD2*, локализованный на хромосоме 11q23.2 [21]. Ранее интенсивно исследовался так называемый Taq1A полиморфизм гена *DRD2* (rs1800497); позже было показано, что он находится в экзоне 8 соседнего гена *ANKK1*, кодирующего протеинкиназу, и приводит к аминокислотной замене p.Glu713Lys [22]. По данным базы HaploReg v3 *ANKK1* (rs1800497) расположен в регионе ДНК, гиперчувствительном к ДНКазе-1 в шести различных тканях. Помимо тесного сцепления rs1800497 гена *ANKK1* с геном *DRD2*, в ряде исследований показано, что ANKK1 функционально связан с дофаминергической системой [23]. Liu et al. (2020) показали ассоциацию полиморфных вариантов, локализованных в кластере генов *ANKK1/DRD2*, с никотиновой зависимостью у мужчин из Китая [24]. Нами не выявлена ассоциация локуса rs1800497 гена *ANKK1* с развитием ХОБЛ. Значимые ассоциации с развитием ХОБЛ для полиморфного варианта rs1800497**G* гена *ANKK1* были выявлены только при анализе комбинаций всех исследованных генов в составе информативных сочетаний с полиморфными вариантами генов *GRIN2B* rs2268132**TT* и *GRIN2B* rs7301328**G*, *GABBR2* rs3750344**A*, *CHRNA5* rs16969968**T* и *GRIK5* rs8099939**C*.

В заключение следует отметить, что полученные нами данные подтверждают предположение о существенной роли генов нейротрансмиттеров в формирование предрасположенности к ХОБЛ, ключевым фактором развития которого является курение. Наиболее значимые ассоциации с ХОБЛ установлены для генов глутаматных рецепторов *GRIK3*, *GRIN2B*, *GRIA1*, *GRIN1*. Для генов *BDNF*, *GABBR2*, *ANKK1* ассоциации с развитием заболевания выявлены только в информативных сочетаниях с полиморфными вариантами генов глутаматных рецепторов (*GRIN2B*, *GRIK5*) и холинэргических никотиновых рецепторов (*CHRNA5*, *CHRNA3*), что может указывать на синергизм исследуемых генов. Определены патогенетически значимые взаимодействия полиморфных вариантов генов *GRIA1*, *GRIK3*, *GRIN2B* с уровнем никотиновой зависимости и индексом курения. Полученные результаты представляют интерес для понимания молекулярных механизмов развития заболеваний, связанных с курением.

Исследование проведено в рамках НИР № АААА-А21-121011990119-1; биологический материал (ДНК) для исследования взят из коллекции “Коллекция биологических материалов человека ИБГ УНЦ РАН” ИБГ УНЦ РАН, поддержанной программой биоресурсных коллекций ФАНО России; работа выполнена с использованием оборудования ЦКП “Биомика” и УНУ “КОДИНК” (ИБГ УФИЦ РАН).

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

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. From the Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2017. Available from: http://goldcopd.org.
2. *Broms U., Silventoinen K., Madden P.A. et al.* Genetic architecture of smoking behavior: A study of Finnish adult twins // Twin. Res. Hum. Genet. 2006. V. 9. № 1. P. 64–72. https://doi.org/10.1375/183242706776403046
3. *Bierut L.J.* Nicotine dependence and genetic variation in the nicotinic receptors // Drug Alcohol Depend. 2009. V. 104. Suppl. 1. P. S64–S69. https://doi.org/10.1016/j.drugalcdep.2009.06.003
4. Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior // Nat. Genet. 2010. V. 42. № 5. P. 441–447. https://doi.org/10.1038/ng.571
5. *Bierut L.J.* Convergence of genetic findings for nicotine dependence and smoking related diseases with chromosome 15q24-25 // Trends Pharmacol. Sci. 2010. V. 31. № 1. P. 46–51. https://doi.org/10.1016/j.tips.2009.10.004
6. *Wolock S.L., Yates A., Petrill S.A. et al.* Gene × smoking interactions on human brain gene expression: Finding common mechanisms in adolescents and adults // J. Child. Psychol. Psychiatry. 2013. V. 54. № 10. P. 1109–1119. https://doi.org/10.1111/jcpp.12119
7. *Korytina G.F., Akhmadishina L.Z., Viktorova E.V. et al. IREB2, CHRNA5, CHRNA3, FAM13A* & hedgehog interacting protein genes polymorphisms & risk of chronic obstructive pulmonary disease in Tatar population from Russia // Indian J. Med. Res. 2016. V. 144. № 6. P. 865–876. https://doi.org/10.4103/ijmr.IJMR_1233_14
8. *Корытина Г.Ф., Ахмадишина Л.З., Кочетова О.В. и др.* Полиморфные варианты генов глутаматных (*GRIK5, GRIN2B*) и серотонинового (*HTR2A*) рецепторов ассоциированы с хронической обструктивной болезнью легких // Мол. биология. 2017. Т. 51. № 4. С. 603–614. https://doi.org/10.7868/S0026898417040127
9. The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information (US). Available from: http://www.ncbi.nlm.nih.gov/projects/SNP/.
10. *Ward L.D., Kellis M.* HaploReg v4: Systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease // Nucl. Acids Res. 2016. V. 44. № D1. P. D877–D881. https://doi.org/10.1093/nar/gkv1340
11. *Purcell S., Neale B., Todd-Brown K. et al.* PLINK: A toolset for whole-genome association and population-based linkage analysis // Am. J. Hum. Genet. 2007. V. 81. № 3. P. 559–575. https://doi.org/10.1086/519795
12. *Favorov A.V., Andreewski T.V., Sudomoina M.A. et al.* A Markov chain Monte Carlo technique for identification of combinations of allelic variants underlying complex diseases in humans // Genetics. 2005. V. 171. № 4. P. 2113–2121. https://doi.org/10.1534/genetics.105.048090
13. *Smith D.J., Escott-Price V., Davies G. et al.* Genome-wide analysis of over 106000 individuals identifies 9 neuroticism-associated loci // Mol. Psychiatry. 2016. V. 21. № 6. P. 749–757. https://doi.org/10.1038/mp.2016.49
14. *Grucza R.A., Johnson E.O., Krueger R.F. et al.* Incorporating age at onset of smoking into genetic models for nicotine dependence: Evidence for interaction with multiple genes // Addict. Biol. 2010. V. 15. № 3. P. 346–357. https://doi.org/10.1111/j.1369-1600.2010.00220.x
15. *Vink J.M., Smit A.B., de Geus E.J. et al.* Genome-wide association study of smoking initiation and current smoking // Am. J. Hum. Genet. 2009. V. 84. № 3. P. 367–379. https://doi.org/10.1016/j.ajhg.2009.02.001
16. *Orihara K., Odemuyiwa S.O., Stefura W.P. et al.* Neurotransmitter signalling via NMDA receptors leads to decreased T helper type 1-like and enhanced T helper type 2-like immune balance in humans // Immunology. 2018. V. 153. № 3. P. 368–379. https://doi.org/10.1111/imm.12846
17. *Jamal M., Van der Does W., Elzinga B.M. et al.* Association between smoking, nicotine dependence, and BDNF Val66Met polymorphism with BDNF concentrations in serum // Nicotine Tob. Res. 2015. V. 17. № 3. P. 323–329. https://doi.org/10.1093/ntr/ntu151
18. *Ohmoto M., Takahashi T.* Effect of genetic polymorphism of brain-derived neurotrophic factor and serotonin transporter on smoking phenotypes: A pilot study of Japanese participants // Heliyon. 2019. V. 5. № 2. P. e01234. https://doi.org/10.1016/j.heliyon.2019.e01234
19. *Beuten J., Ma J.Z., Payne T.J. et al.* Single- and multilocus allelic variants within the GABA(B) receptor subunit 2 (GABAB2) gene are significantly associated with nicotine dependence // Am. J. Hum. Genet. 2005. V. 76. № 5. P. 859–864. https://doi.org/10.1086/429839
20. *Cui W.Y., Seneviratne C., Gu J. et al.* Genetics of GABAergic signaling in nicotine and alcohol dependence // Hum. Genet. 2012. V. 131. № 6. P. 843–855. https://doi.org/10.1007/s00439-011-1108-4

21. *Dani J.A.* Roles of dopamine signaling in nicotine addiction // Mol. Psychiatry. 2003. V. 8. № 3. P. 255–256. https://doi.org/10.1038/sj.mp.4001284
22. *Neville M.J., Johnstone E.C., Walton R.T.* Identification and characterization of ANKK1: A novel kinase gene closely linked to DRD2 on chromosome band 11q23.1 // Hum. Mutat. 2004. V. 23. № 6. P. 540–545. https://doi.org/10.1002/humu.20039
23. *Koeneke A., Ponce G., Troya-Balseca J. et al.* Ankyrin repeat and kinase domain containing 1 gene, and addiction vulnerability // Int. J. Mol. Sci. 2020. V. 21. № 7. P. 2516. https://doi.org/10.3390/ijms21072516
24. *Liu Q., Xu Y., Mao Y. et al.* Genetic and epigenetic analysis revealing variants in the NCAM1-TTC12-ANKK1-DRD2 cluster associated significantly with nicotine dependence in Chinese han smokers // Nicotine Tob. Res. 2020. V. 22. № 8. P. 1301–1309. https://doi.org/10.1093/ntr/ntz240

Role of the Neurotransmitter System Genes in Chronic Obstructive Pulmonary Disease Development

G. F. Korytina[a, b, *], L. Z. Akhmadishina[a], O. V. Kochetova[a], T. R. Nasibullin[a], Yu. G. Aznabaeva[b], Sh. R. Zulkarneev[b], S. M. Izmailova[b], N. Sh. Zagidullin[b], and T. V. Victorova[b]

[a] Institute of Biochemistry and Genetics – Subdivision of the Ufa Federal Research Centre of the Russian Academy of Sciences (IBG UFRC RAS), Ufa, 450054 Russia
[b] Bashkortostan State Medical University, Ufa, 450000 Russia
**e-mail: guly_kory@mail.ru*

Chronic obstructive pulmonary disease (COPD) is a complex chronic inflammatory disease of the respiratory system. Smoking is a major risk factor COPD. The goal of the present study was to investigate the association of COPD with neurotransmitter system genes. SNPs in *GRIK3* (rs534131), *GRIN2B* (rs7301328, rs1805476), *GRIA1* (rs2195450), *GRIN1* (rs6293), *GABBR2* (rs3750344), *BDNF* (rs6265, rs11030107), *ANKK1* (rs1800497) genes were genotyped by real-time polymerase chain reaction in a case-control study (601 COPD patients and 617 controls). We detected a significant associations of *GRIK3* (rs534131) (P = = 0.009, OR = 1.42 in dominant model), *GRIA1* (rs2195450) (P = 0.015, OR = 1.35 in dominant model), *GRIN1* (rs6293) (P = 0.036, OR = 0.79 in log-additive model), *GRIN2B* (rs7301328) (P = 0.0009, OR = 0.54 in recessive model) with COPD. *GRIK3* (rs534131) (P = 0.0001, OR = 1.68 in dominant model) and *GRIN2B* (rs7301328) (P = 0.001, OR = 0.52 in recessive model) were significantly associated with COPD only in smokers. The association of *GRIN1* (rs6293) (P = 0.0001, OR = 0.36 in dominant model), *GRIA1* (rs2195450) (P = 0.0018, OR = 2.40 in log-additive model) and *BDNF* (rs11030107) (P = 0.005, OR = 2.86 for TT) with COPD confirmed in non-smokers. The level of nicotine dependence was significantly higher in the carriers of the rare A allele *GRIK3* (rs534131A>G) and GC genotype for *GRIN2B* (rs7301328C>G) (P = 0.011 and P = 0.023). Using the APSampler algorithm, we obtained genotype/allele combinations that remained significantly associated with COPD; *GRIN2B* rs2268132*T, *GRIN2B* rs7301328*G, *GRIN2B* rs1805476*C, *ANKK1* rs1800497*G, *GABBR2* rs3750344*A, *CHRNA5* rs16969968*T, *CHRNA3* rs1051730*A, *HTR2A* rs6313*CC, *GRIA1* rs2195450*G loci were featured in combinations most significantly associated with COPD risk.

Keywords: chronic obstructive pulmonary disease (COPD), glutamate receptors, nicotine dependence, smoking, gene-environmental interactions.

ГЕНЕТИКА, 2021, том 57, № 11, с. 1297–1305

ГЕНЕТИКА ЧЕЛОВЕКА

УДК 616.127;577.21

ЭКСПРЕССИЯ ГЕНА β1-АДРЕНОРЕЦЕПТОРА (*ADRB1*) В МИОКАРДЕ БОЛЬНЫХ ХРОНИЧЕСКОЙ СЕРДЕЧНОЙ НЕДОСТАТОЧНОСТЬЮ

© 2021 г. Э. Ф. Муслимова[1], *, Т. Ю. Реброва[1], Д. С. Кондратьева[1], Э. Л. Сондуев[1], Б. Н. Козлов[1], С. А. Афанасьев[1]

[1]*Научно-исследовательский институт кардиологии, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, 634012 Россия*

**e-mail: muslimovef@yandex.ru*

Поступила в редакцию 16.02.2021 г.
После доработки 06.04.2021 г.
Принята к публикации 11.05.2021 г.

Проведен анализ уровня относительной экспрессии гена β1-адренорецептора *ADRB1* в миокарде 78 пациентов с хронической сердечной недостаточностью на фоне сочетанного развития ишемической болезни сердца и артериальной гипертензии. Выявлено значимое снижение уровня экспрессии гена *ADRB1* у пациентов с гипертрофией левого желудочка. В рассматриваемой когорте уровень экспрессии гена *ADRB1* не зависел от функционального класса сердечной недостаточности. Экспрессия *ADRB1* оказалась более высокой у пациентов, принимавших препараты ацетилсалициловой кислоты, но не показала значимой связи с приемом β-адреноблокаторов или ингибиторов АПФ.

Ключевые слова: β-адренорецепторы, *ADRB1*, экспрессия генов, сердечная недостаточность, гипертрофия левого желудочка.

DOI: 10.31857/S0016675821110084

Одним из важнейших звеньев нейрогуморальной регуляции хроно-инотропных возможностей миокарда являются β1-адренорецепторы (**β1-АР**), локализованные на мембранах кардиомиоцитов. β1-АР напрямую влияют на внутрисердечную гемодинамику и способность сердца переносить физическую нагрузку. Кратковременная активация β1-АР приводит к развитию мощного положительного инотропного эффекта за счет увеличения внутриклеточного содержания ионов Ca^{2+} [1, 2]. При стимуляции β1-АР повышается уровень цАМФ, что приводит к фосфорилированию фосфоламбана. Это убирает ингибирующее действие фосфоламбана на активность Ca^{2+}-АТФазы SERCA2a и способствует ускорению релаксации миокарда [3]. Таким образом, β1-АР играют важную роль в работе белков Ca^{2+}-транспортирующей системы саркоплазматического ретикулума кардиомиоцитов, нарушение которой может приводить к развитию сердечной недостаточности [4].

Однако под влиянием длительной или сильной стимуляции катехоламинами, например при хронической сердечной недостаточности (**ХСН**), изменяется количество и функциональное состояние β1-АР на мембране кардиомиоцитов [5]. Более того, стимуляция β1-АР может привести к кардиотоксическому эффекту – инициации апоптоза кардиомиоцитов, в основе которого лежит перегрузка клеток Ca^{2+} [6].

Преобладающие на кардиомиоцитах β1-АР кодируются геном *ADRB1* (MIM 109630), локализованным на 10-й хромосоме (10q25.3) [7, 8]. В настоящее время много внимания уделяется роли полиморфных вариантов гена *ADRB1* в развитии и прогрессировании сердечно-сосудистых заболеваний [9, 10]. Менее интенсивно исследуется связь между сократительной функцией миокарда и уровнем экспрессии гена *ADRB1.* Однако отмечено, что у больных ХСН экспрессия гена *ADRB1* повышалась совместно с увеличением сократительного резерва миокарда после сердечной ресинхронизирующей терапии [11]. Количественная оценка мРНК β1-АР, анализ изменения их уровня является важным инструментом в изучении патофизиологической регуляции этих рецепторов.

Целью нашего исследования была оценка уровня экспрессии гена β1-адренорецептора *ADRB1* в миокарде в зависимости от клинических параметров у больных хронической сердечной недостаточностью.

МАТЕРИАЛЫ И МЕТОДЫ

В исследование включено 78 пациентов с ХСН, развившейся на фоне сочетанного течения ишемической болезни сердца и артериальной гипертензии. Возраст пациентов составил 64 (59; 69) года. Выборка включала 63 (80.8%) мужчины

Таблица 1. Клиническая характеристика пациентов с хронической сердечной недостаточностью

Параметр	Значение
Инфаркт миокарда в анамнезе, *n* (%)	52 (66.7)
Возраст первого инфаркта миокарда, годы	62 (53; 67)
Сердечная недостаточность, ФК I/II/III, *n* (%)	4 (5.1)/46 (59.0)/28 (35.9)
Фракция выброса левого желудочка, %	62 (53; 65)
Конечный систолический объем, мл	45 (36; 59)
Конечный диастолический объем, мл	118 (101; 127)
Конечный систолический размер, мм	34 (31; 37)
Конечный диастолический размер, мм	51 (48; 54)
Индекс сферичности	0.56 (0.53; 0.59)
Пик Е, см/с	71 (61; 83)
Пик А, см/с	85 (64; 96)
Отношение Е/А	0.87 (0.69; 1.11)
Гипертрофия левого желудочка, *n* (%)	18 (23.1)
Дилатация левого желудочка, *n* (%)	7 (9.0)
Дилатация левого предсердия, *n* (%)	28 (35.9)
Диастолическая дисфункция, *n* (%)	53 (67.9)*
Митральная недостаточность I/II степени, *n* (%)	25 (32.1)/7 (9.0)
Трикуспидальная недостаточность I степени, *n* (%)	11 (14.1)
Нарушение толерантности к глюкозе/сахарный диабет 2-го типа, *n* (%)	8 (10.3)/27 (34.6)
Ожирение, *n* (%)	24 (30.8)
Ингибиторы АПФ, блокаторы рецепторов ангиотензина II, *n* (%)	61 (78.2)
β-адреноблокаторы, *n* (%)	63 (80.8)
Гиполипидемические препараты, *n* (%)	66 (84.6)
Ацетилсалициловая кислота 75–100 мг, *n* (%)	48 (61.5)**

Примечание. * У пяти пациентов не установлен диагноз; ** у семи пациентов не удалось выяснить информацию о приеме препаратов.

и 15 (19.2%) женщин. Клиническая характеристика пациентов представлена в табл. 1. Принимаемые лекарственные препараты указаны на момент поступления.

Функциональный класс (**ФК**) ХСН в соответствии с классификацией Нью-Йоркской кардиологической ассоциации (NYHA) был определен с помощью теста 6-минутной ходьбы. Пройденное расстояние в 426–550 и 301–425 м соответствует I и II ФК, а расстояние в 151–300 м соответствует III ФК [12].

Всем пациентам выполнено эхокардиографическое исследование сердца. Эхокардиографию проводили, используя аппарат Philips HD15 (Нидерланды) из стандартных позиций с оценкой размеров отделов сердца и фракцией выброса левого желудочка (**ЛЖ**) по методу Симпсона. Также диагностировалось наличие гипертрофии ЛЖ при увеличении индекса массы миокарда ЛЖ >115 у мужчин и >95 г/м^2 у женщин [12].

Для генетического исследования использовали интраоперационные биоптаты сердца – ушко правого предсердия, иссекаемые при подключении аппарата искусственного кровообращения во время плановой операции коронарного шунтирования. Сразу после иссечения биоптаты помещали в ледяной раствор Кребса–Хензеляйта следующего состава в мМ: NaCl – 120, KCl – 4.8; $CaCl_2$ – 2.0, $MgSO_4$ – 1.2, KH_2PO_4 – 1.2, $NaHCO_3$ – 20.0, глюкоза – 10.0. Данный раствор позволяет поддерживать жизнеспособность кардиомиоцитов с сохранением их сократительной функции [13]. В условиях стерильного бокса биоптаты освобождали от соединительной ткани и забирали образцы 20–25 мг. При невозможности немедленного выделения РНК образцы помещали в RNAlater RNA Stabilization Reagent (Кат. № 76104 QIAGEN, Германия) и хранили при –80°C. Образцы миокарда разрушали на гомогенизаторе TissueLyser LT (QIAGEN), и из гомогената выделяли тотальную РНК с помощью коммерческого набора RNeasy Fibrous Tissue Mini Kit (Кат. № 74704, QIAGEN). Качество образцов РНК оценивали спектрофотометрическим методом по отношению A260/A280

(NanoVue, Heaithcare Bio-Science, Швеция), которое варьировало в пределах от 1.93 до 2.03. Для синтеза кДНК использовали реактивы RevertAid First Strand cDNA Synthesis Kit (Кат. № K1622, Thermo Scientific, США).

Полимеразную цепную реакцию в режиме реального времени проводили на амплификаторе LightCycler 96 (Roche) с использованием праймеров, разработанных и синтезированных ООО "ДНК-Синтез" (Россия): hADRB1-F = CAGGTGAACTCGAAGCCC; hADRB1-R = CTCCCATCCCTTCCCAA; hADRB1-probe = FAM-AAAGCCACGGACCGTTGCAC-BHQ1. Подбор праймеров осуществлялся так, чтобы была исключена возможность амплификации геномной ДНК за счет расположения праймеров/зонда в разных экзонах. Для повышения эффективности реакции подбирались праймеры/зонд, не содержащие димеры и шпильки. Все праймеры/зонды проходят проверку в системе BLAST. В качестве референсного был использован ген глицеральдегид-3-фосфат-дегидрогеназы *GAPDH* (TaqMan® Gene Expression Assays, Hs03929097_g1 (VIC), Thermo Scientific), рекомендованный в качестве гена "домашнего хозяйства" при исследовании разных отделов и патологий сердца [14, 15].

Постановку реакции проводили в триплетах, в одном планшете параллельно проводили реакции для целевого гена и гена "домашнего хозяйства". Для мастер-микса использованы 10× DreamTaq Buffer (Кат. № B65, Thermo Scientific), DreamTaq DNA Polymerase (Кат. № EP0701, Thermo Scientific), Invitrogen dNTP Set 100 mM (Кат. № 10297117, Thermo Scientific). Расчет уровня экспрессии проводили с помощью программного обеспечения LightCycler 96 (Roche) с применением стандартной кривой, поправкой на эффективность реакции и калибратор постановок по методу Pfaffl [16]: Ratio = $= (E_{\text{target}})^{\Delta Ct,\ \text{target(calibrator} - \text{test)}}/(E_{\text{ref}})^{\Delta Ct,\ \text{ref(calibrator} - \text{test)}}$, где E – эффективность реакции; Ct – пороговый цикл генов мишеней (target) и референсного гена (ref); ΔCt, target(calibrator – test) – Ct гена мишени в калибраторе минус Ct гена мишени в опытном образце; ΔCt, ref(calibrator – test) – Ct референсного гена в калибраторе минус Ct референсного гена в опытном образце.

Статистический анализ проводили с помощью стандартного пакета программ SPSS версии 13 (IBM, США). Количественные данные были предварительно проверены на соответствие нормальному закону распределения с помощью критерия Шапиро–Уилка. Так как большая часть исследуемых параметров, включая уровень экспрессии гена *ADRB1*, не соответствовала нормальному закону распределения, то дальнейший анализ проводили с помощью критерия Манна–Уитни или теста Краскела–Уоллиса. Результаты представляли в виде медианы и интерквантильного размаха. Также оценивали силу линейной взаимосвязи между количественными показателями с помощью коэффициента ранговой корреляции Спирмена. Связь между качественными данными определяли с помощью χ^2 Пирсона или двустороннего точного теста Фишера. Уровень значимости различий принимали $p < 0.05$.

РЕЗУЛЬТАТЫ

При проведении ПЦР в режиме реального времени эффективность реакции для гена *ADRB1* в среднем составляла 1.94 ± 0.12 (94%), коэффициент детерминации R^2 – 0.98 ± 0.01; для гена *GAPDH* эффективность реакции составила 1.93 ± 0.1 (93%), коэффициент R^2 – 0.98 ± 0.02.

В исследуемой выборке относительная экспрессия гена *ADRB1* была сопоставима ($p = 0.526$) у мужчин и женщин – 1.2 (0.7; 1.7) и 1.3 (0.8; 1.7) – и линейно не связана с возрастом пациентов ($r = 0.154$, $p = 0.191$). Отсутствовали различия в уровне экспрессии гена в миокарде пациентов с ХСН ФК I, II, III, что показано на рис. 1.

Наличие инфаркта миокарда (ИМ) в анамнезе также не влияло на экспрессию *ADRB1* ($p = 0.556$). Экспрессия гена составила у пациентов без ИМ 1.3 (0.8; 1.7) и у больных с ИМ в анамнезе 1.2 (0.7; 1.7). В исследуемой выборке наблюдалась тенденция к более высокой экспрессии гена *ADRB1* при сахарном диабете 2-го типа – 1.5 (0.9; 2.1) против 1.2 (0.7; 1.6) у пациентов без гипергликемии, но различия не достигли статистической значимости ($p = 0.069$). У пациентов с ожирением и без него уровень экспрессии оказался одинаков ($p = 0.956$) – 1.2 (0.7; 1.7) и 1.2 (0.8; 1.7). Статистически значимых различий не выявлено у лиц с дилатацией левого предсердия (1.5 (0.9; 1.7) против 1.1 (0.7; 1.7), $p = 0.110$) или левого желудочка (1.2 (0.7; 2.2) против 1.2 (0.8; 1.7), $p = 0.984$) по сравнению с пациентами без изменений отделов сердца. Экспрессия гена *ADRB1* у пациентов с диастолической дисфункцией ЛЖ составила 1.2 (0.7; 1.6), без нее – 1.3 (0.8; 1.7) ($p = 0.369$). Коэффициент корреляции между уровнем экспрессии гена *ADRB1* и фракцией выброса левого желудочка составил $r = -0.202$ ($p = 0.085$). Имела место слабая прямая линейная взаимосвязь ($r = 0.248$, $p = 0.039$) между уровнем экспрессии и скоростью пика А, но корреляция с другими показателями ЭхоКГ, в том числе скоростью пика Е и отношением Е/А, отсутствовала.

Тем не менее выявлено значимое ($p = 0.022$) снижение экспрессии *ADRB1* на 21–35% в миокарде пациентов с гипертрофией ЛЖ. Результаты представлены на рис. 2. Хорошо известно, что к развитию гипертрофии ЛЖ приводит митральная недостаточность. Аналогичная зависимость получена и для нашей выборки (Fisher's Exact Test = 16.621,

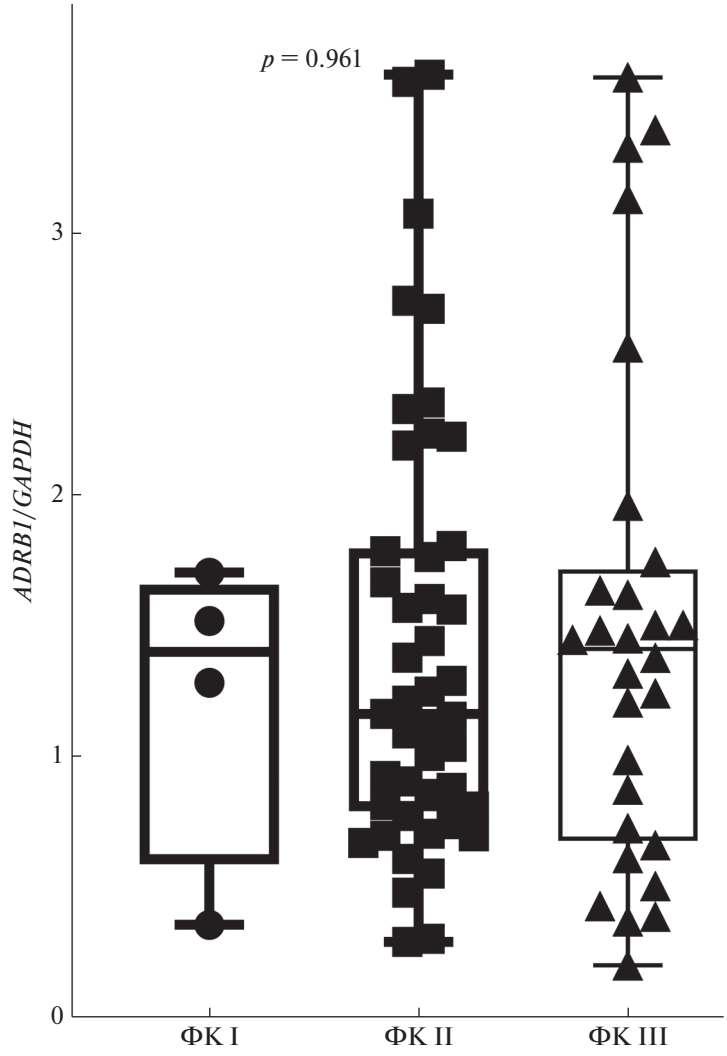

Рис. 1. Уровень экспрессии гена *ADRB1* относительно *GAPDH* у пациентов с ХСН ФК I (1.4 (0.8; 1.6)), ФК II (1.2 (0.8; 1.8)), ФК III (1.3 (0.7; 1.6)). При сравнении использован тест Краскела–Уоллиса.

$p < 0.001$, рис. 3,*а*). Но при этом уровень экспрессии *ADRB1* между лицами с митральной недостаточностью I, II степени и без нее оказался сопоставимым (рис. 3,*б*), как и при наличии и отсутствии трикуспидальной недостаточности (1.2 (0.9; 1.6) и 1.3 (0.7; 1.7) соответственно, $p = 0.818$).

Выполнен анализ уровня экспрессии гена *ADRB1* в зависимости от приема лекарственных препаратов на момент поступления пациента в стационар. В нашей выборке ни β-адреноблокаторы (1.2 (0.7; 1.7) у принимавших и 1.2 (0.9; 1.6) у не принимавших, $p = 0.353$), ни гиполипидемические препараты (1.2 (0.7; 1.6) против 1.3 (0.6; 1.8), $p = 0.868$), ни ингибиторы АПФ и блокаторы рецепторов ангиотензина II (1.2 (0.7; 1.6) против 1.4 (0.9; 2.0), $p = 0.115$) не показали значимого влияния на значения *ADRB1/GAPDH*. В то же время пациенты, принимавшие препараты ацетилсалициловой кислоты (АСК), отличались значимо более высокой экспрессией *ADRB1* ($p = 0.048$), что показано на рис. 4. При этом в группе лиц, принимавших препараты АСК, частота случаев гипертрофии ЛЖ (6 (12.5%) из 48 пациентов) была значимо ($p = 0.006$) ниже, чем среди пациентов, по каким-либо причинам не принимавших АСК (10 (43.5%) из 23 пациентов).

ОБСУЖДЕНИЕ

Классический путь внутриклеточной трансляции сигналов β1-АР хорошо изучен – это активация комплекса аденилатциклаза–Gs-белки, приводящая к повышению уровня цАМФ. Первичной мишенью для цАМФ является протеинкиназа А. Она фосфорилирует белки (Ca^{2+}-каналы L-типа, фосфоламбан, тропонин I, рианодиновые рецепторы и др.), которые влияют на сократимость кардиомиоцитов за счет увеличения притока Ca^{2+}, об-

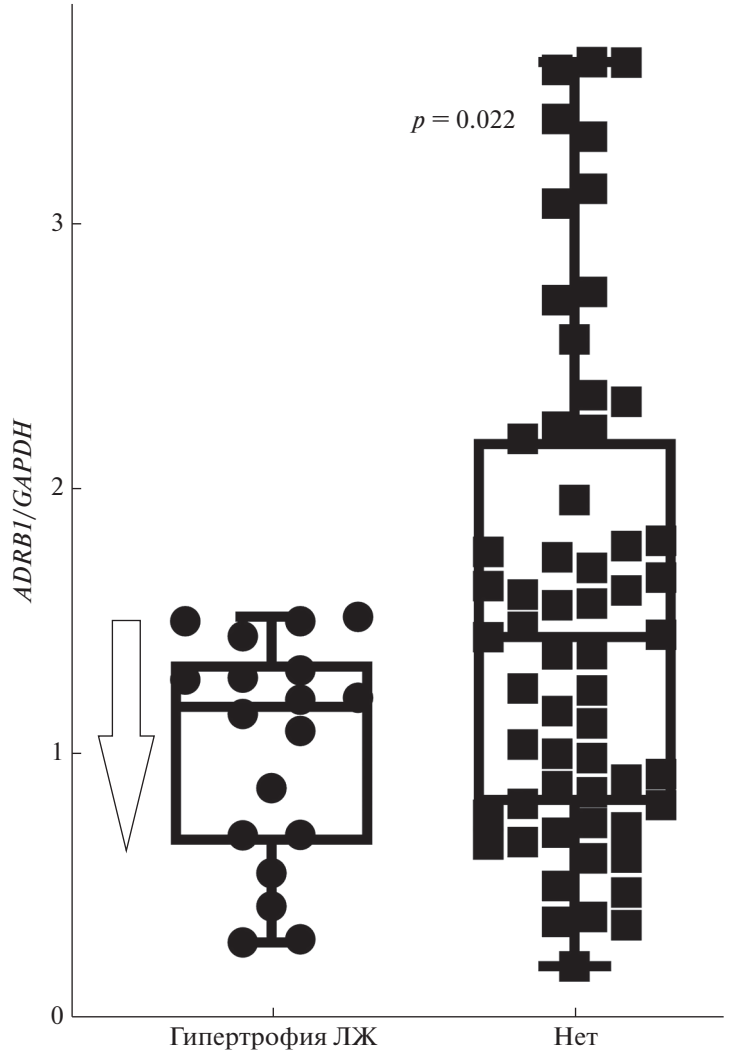

Рис. 2. Уровень экспрессии гена *ADRB1* относительно *GAPDH* у пациентов с гипертрофией ЛЖ (1.1 (0.6; 1.3)) и без нее (1.4 (0.8; 2.0)). При сравнении использован критерий Манна–Уитни.

ратного захвата Ca^{2+} в саркоплазматический ретикулум или за счет модулирования чувствительности миофиламентов к этим ионам [17].

Известно, что при ХСН из-за повышенной концентрации катехоламинов имеет место хроническая гиперсимпатикотония, характеризующаяся длительной активацией β1-АР. Это приводит к десенситизации, т.е. угнетению рецепторного ответа с течением времени в условиях наличия постоянного стимула высокой интенсивности, снижается экспрессия β1-АР на кардиомиоцитах. Как следствие, у пациентов с ХСН наблюдается уменьшение инотропного резерва миокарда и насосной функции сердца [2, 18].

Оценка экспрессии гена β1-АР *ADRB1* является одним из способов изучить влияние различных клинических параметров на состояние симпатоадреналовой системы при ХСН. Мы определяли относительный уровень экспрессии *ADRB1/GAPDH* в миокарде пациентов с ХСН ФК I, ФК II и ФК III, но не обнаружили статистически значимых различий. Отсутствовала линейная корреляция с фракцией выброса ЛЖ. Также не было различий по уровню экспрессии гена *ADRB1* у пациентов с диастолической дисфункцией ЛЖ и без нее. И хотя в исследуемой выборке была выявлена слабая прямая зависимость между уровнем экспрессии и скоростью кровотока в период предсердной систолы (пик А), мы не обнаружили связи между экспрессией и скоростью кровотока через митральный клапан в раннюю фазу диастолы (пик Е) и отношением Е/А, характеризующим трансмитральный кровоток.

Известно, что на скорость пика А оказывает влияние податливость ЛЖ и сократимость ЛП. Несмотря на отсутствие явной зависимости уровня экспрессии гена от тяжести ФК ХСН и корре-

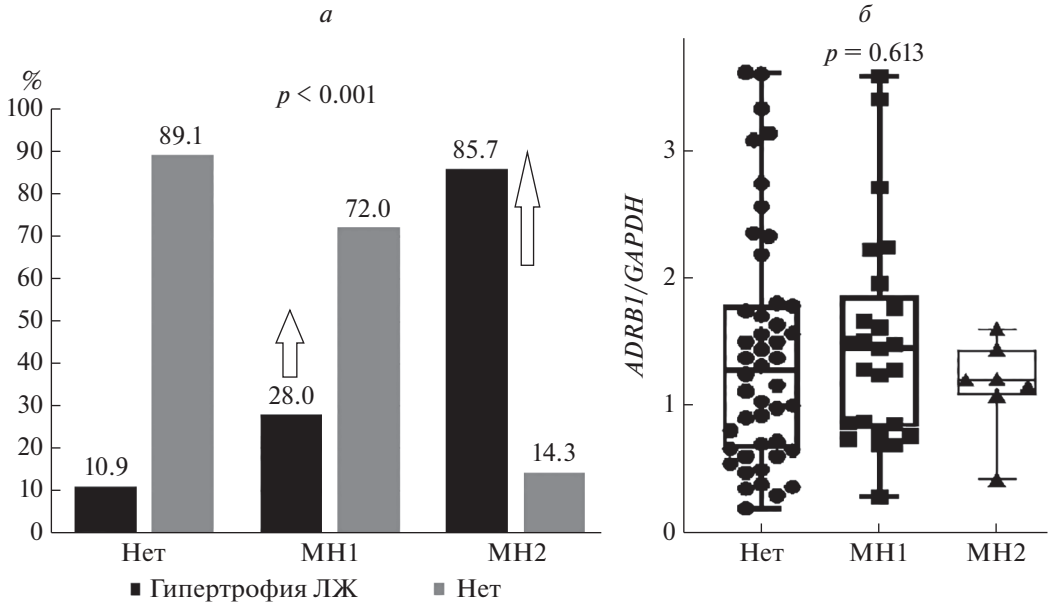

Рис. 3. Частота гипертрофии левого желудочка (*а*) и уровень экспрессии гена *ADRB1* относительно *GAPDH* (*б*) у пациентов с митральной недостаточностью (МН). Экспрессия гена при отсутствии МН – 1.3 (0.7; 1.7), при МН 1-й степени – 1.5 (0.8; 1.9), при МН 2-й степени – 1.2 (1.1; 1.3). При сравнении частот использован χ^2 Пирсона, при сравнении уровней экспрессии – тест Краскела–Уоллиса.

ляции с параметрами ЭхоКГ, экспрессия *ADRB1* оказалась значительно снижена в миокарде пациентов с гипертрофией ЛЖ. Более того, ранее в той же выборке пациентов мы оценили уровень экспрессии гена *ATP2A2,* кодирующего SERCA2a, и гена рианодиновых рецепторов *RYR2.* Относительная экспрессия генов *ATP2A2* и *RYR2* также была снижена у пациентов с гипертрофией ЛЖ, показатель *ATP2A2/GAPDH* коррелировал со скоростью пика А и *RYR2/GAPDH* [13]. При исследовании экспрессии гена *ADRB1* была обнаружена прямая линейная связь между изучаемым фактором и экспрессией гена *ATP2A2* ($r = 0.293, p = 0.030$). Однако отсутствовала статистически значимая линейная зависимость между уровнями экспрессии гена *ADRB1* и гена *RYR2* ($r = 0.165$, $p = 0.223$).

Стимуляция β1-АР через цАМФ приводит к фосфорилированию одного из ключевых белков Ca^{2+}-транспортирующей системы саркоплазматического ретикулума кардиомиоцитов – фосфоламбана. Дефосфорилированный фосфоламбан является ингибитором активности SERCA2a, осуществляющей обратный захват ионов Ca^{2+} из цитозоля в депо. Фосфорилирование снимает это торможение, соответственно стимуляция β1-АР ведет к увеличению активности SERCA2a [3]. Известно, что снижение активности или содержания SERCA2a приводит к ухудшению обратного захвата Ca^{2+} в саркоплазматический ретикулум, что может провоцировать развитие сердечной недостаточности [19]. Следовательно, меньший уровень экспрессии генов *ADRB1* и *ATP2A2* в миокарде пациентов с гипертрофией ЛЖ может свидетельствовать о неблагоприятных процессах, способствующих прогрессированию сократительной дисфункции миокарда.

С другой стороны, сниженная экспрессия гена *ADRB1* может быть компенсаторным ответом на повышенный уровень катехоламинов крови, характерный для пациентов с ХСН и коррелирующий со степенью тяжести дисфункции ЛЖ [2, 20]. Однако в исследуемой выборке почти 81% пациентов принимали β-адреноблокаторы. Такая терапия позволяет устранить избыточное влияние катехоламинов на β-адренергические рецепторы и уменьшить степень ремоделирования сердца, что улучшает клинический прогноз [12]. Оказалось, что по уровню экспрессии гена *ADRB1* группы пациентов, принимавших β-адреноблокаторы и не принимавших их, не различались.

При анализе влияния лекарственной терапии на уровень экспрессии гена *ADRB1* зависимость обнаружена только в отношении приема препаратов АСК. Пациенты, принимавшие АСК, имели значимо более высокую экспрессию гена *ADRB1.* Среди пациентов, принимавших препараты АСК, также реже встречалась гипертрофия ЛЖ. Однако препараты АСК применяются у больных ХСН при строгих показаниях, например после инфаркта миокарда или в случаях ангиопластики с применением стентов с лекарственным покрытием в рамках двойной антитромботической терапии [12].

Важным фактором развития гипертрофии ЛЖ является митральная недостаточность. Митральная

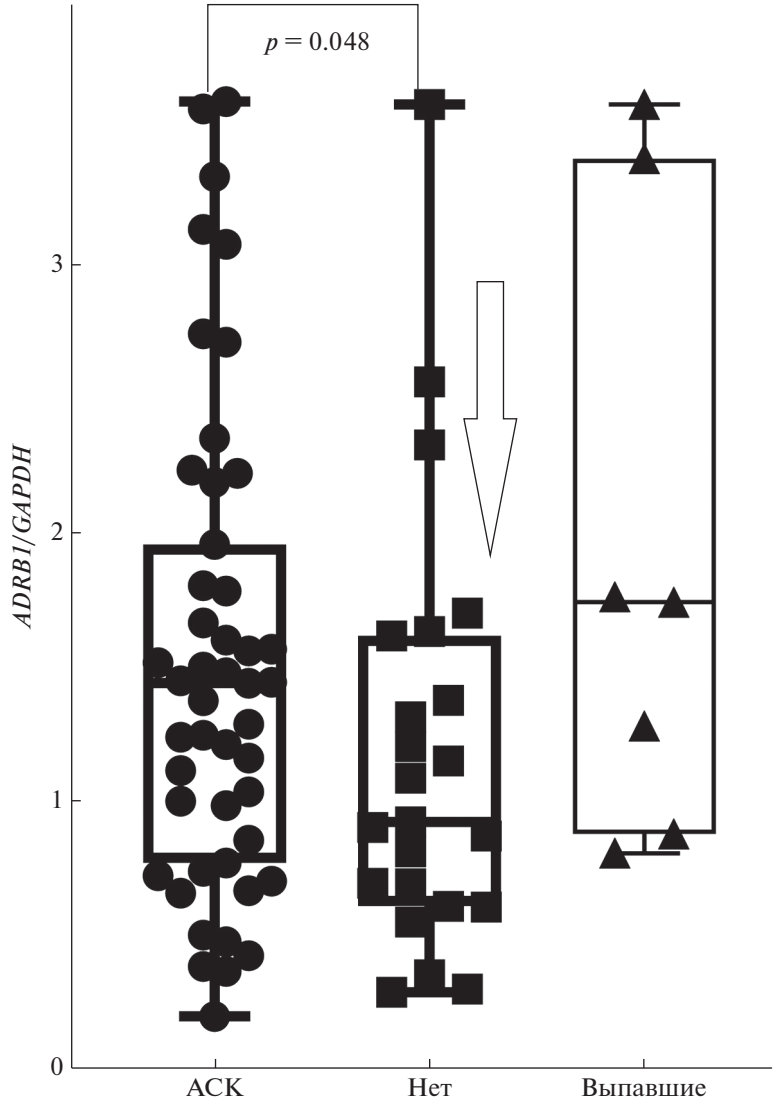

Рис. 4. Уровень экспрессии гена *ADRB1* относительно *GAPDH* у пациентов, принимавших препараты АСК (1.4 (0.8; 2.0)) и не принимавших АСК (0.9 (0.6; 1.4)). Группа "Выпавшие" – лица, у которых не удалось выяснить информацию о приеме препаратов АСК на момент поступления. При сравнении групп, принимавших и не принимавших препараты АСК, использован критерий Манна–Уитни.

недостаточность I степени характеризуется незначительной митральной регургитацией (МР), II степени – умеренной, III степени – выраженной, а IV степени – тяжелой митральной регургитацией. Пациенты с незначительной или умеренной МР могут длительное время оставаться бессимптомными на фоне небольших гемодинамических компенсаторных изменений. Известно, что МР приводит к компенсаторной гипертрофии ЛЖ и увеличению конечного диастолического объема, что позволяет восстановить эффективный сердечный выброс [21]. Исследуемая нами выборка не является исключением, пациенты с митральной недостаточностью чаще имели гипертрофию ЛЖ. Однако уровень экспрессии гена *ADRB1* оказался сопоставим между лицами с незначительной, умеренной МР и без МР.

Одним из ограничений нашего исследования является то, что мы не определили уровень катехоламинов у пациентов, адренореактивность организма и количество рецепторов на мембранах кардиомиоцитов, что затрудняет оценку связи между относительным уровнем экспрессии гена *ADRB1* и адренореактивностью организма больных ХСН.

Таким образом, в выборке пациентов с хронической сердечной недостаточностью, развившейся на фоне ишемической болезни сердца и артериальной гипертензии, выявлено значимое снижение уровня экспрессии гена β1-адренорецептора *ADRB1* в миокарде пациентов с гипертрофией левого же-

лудочка. В то же время пациенты с разным функциональным классом сердечной недостаточности имели сопоставимый уровень экспрессии гена *ADRB1*. Относительная экспрессия *ADRB1* оказалась более высокой у пациентов, принимавших препараты ацетилсалициловой кислоты, но не показала значимой связи с приемом β-адреноблокаторов или ингибиторов АПФ.

Исследование выполнено при поддержке гранта Российского научного фонда, проект № 20-75-00003: “Оценка ассоциации экспрессии генов Ca(2+)-транспортирующих белков саркоплазматического ретикулума кардиомиоцитов с тяжестью течения хронической сердечной недостаточности ишемического генеза”.

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

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. *Ranade K., Jorgenson E., Sheu W.H.-H. et al.* A polymorphism in the β1 adrenergic receptor is associated with resting heart rate // Am. J. Hum. Genet. 2002. V. 70. № 4. P. 935–942. https://doi.org/10.1086/339621
2. *Костюкевич М.В., Зыков К.А., Миронова Н.А. и др.* Роль аутоантител к ßJ-адренорецептору при сердечно-сосудистых заболеваниях // Кардиология. 2016. Т. 56. № 12. С. 82–91. https://doi.org/10.18565/cardio.2016.12.82-91
3. *Luo W., Grupp I.L., Harrer J. et al.* Targeted ablation of the phospholamban gene is associated with markedly enhanced myocardial contractility and loss of 3-agonist stimulation // Circ. Res. 1994. V. 75. № 3. P. 401–409. https://doi.org/10.1161/01.RES.75.3.401
4. *Zhihao L., Jingyu N., Lan L. et al.* SERCA2a: A key protein in the Ca^{2+} cycle of the heart failure // Heart Failure Rev. 2020. V. 25. P. 523–535. https://doi.org/10.1007/s10741-019-09873-3
5. *Najafi A., Sequeira V., Kuster D.W.D., Velden J. van der.* β-Adrenergic receptor signalling and its functional consequences in the diseased heart // Eur. J. Clin. Invest. 2016. V. 46. № 4. P. 362–374. https://doi.org/10.1111/eci.12598
6. *Zhu W.Z., Wang S.Q., Chakir K. et al.* Linkage of β1-adrenergic stimulation to apoptotic heart cell death through protein kinase A-independent activation of Ca^{2+}/calmodulin kinase II // J. Clin. Invest. 2003. V. 111. № 5. P. 617–625. https://doi.org/10.1172/JCI16326
7. *Frielle T., Collins S., Daniel K.W. et al.* Cloning of the cDNA for the human beta1-adrenergic receptor // Proc. Natl Acad. Sci. USA. 1987. V. 84. № 22. P. 7920–7924. https://doi.org/10.1073/pnas.84.22.7920
8. *Yang-Feng T.L., Xue F.Y., Zhong W.W. et al.* Chromosomal organization of adrenergic receptor genes // PNAS. 1990. V. 87. № 4. P. 1516–1520. https://doi.org/10.1073/pnas.87.4.1516
9. *Кузнецова О.О., Никулина С.Ю., Чернова А.А. и др.* Роль полиморфизма гена β-1-адренорецепторов в развитии дилатационной кардиомиопатии // РМЖ. Мед. обозрение. 2020. Т. 4. № 7. С. 394–398. https://doi.org/10.32364/2587-6821-2020-4-7-394-398
10. *Якушин С.С., Солодун М.В.* Полиморфизм гена β1-адренорецептора, постинфарктное ремоделирование миокарда и сердечно-сосудистый риск: есть ли взаимосвязь? // Мед. Совет. 2018. № 5. С. 42–46. https://doi.org/10.21518/2079-701X-2018-5-42-46
11. *Vanderheyden M., Mullens W., Delrue L. et al.* Endomyocardial upregulation of beta1 adrenoreceptor gene expression and myocardial contractile reserve following cardiac resynchronization therapy // J. Card. Fail. 2008. V. 14. № 2. P. 172–178. https://doi.org/10.1016/j.cardfail.2007.10.016
12. *Мареев В.Ю., Фомин И.В., Агеев Ф.Т. и др.* Клинические рекомендации ОССН – РКО – РНМОТ. Сердечная недостаточность: хроническая (ХСН) и острая декомпенсированная (ОДСН). Диагностика, профилактика и лечение // Кардиология. 2018. Т. 58. S6. С. 3–164. https://doi.org/10.18087/cardio.2475
13. *Муслимова Э.Ф., Реброва Т.Ю., Кондратьева Д.С. и др.* Экспрессия гена Ca^{2+}-АТФазы SERCA2A (*ATP2A2*) и гена рианодиновых рецепторов (*RYR2*) у больных хронической сердечной недостаточностью // Генетика. 2020. Т. 56. № 7. С. 819–825. https://doi.org/10.31857/S0016675820070103
14. *Li M., Rao M., Chen K. et al.* Selection of reference genes for gene expression studies in heart failure for left and right ventricles // Gene. 2017. V. 620. P. 30–35. https://doi.org/10.1016/j.gene.2017.04.006
15. *Molina C.E., Jacquet E., Ponien P. et al.* Identification of optimal reference genes for transcriptomic analyses in normal and diseased human heart // Cardiovascular Res. 2018. V. 114. № 2. P. 247–258. https://doi.org/10.1093/cvr/cvx182
16. *Pfaffl M.W.* Quantification strategies in real-time PCR // A-Z of Quantitative PCR. La Jolla, CA, USA: IUL, 2004. P. 87–112.
17. *Lohse M.J., Engelhardt S., Eschenhagen T.* What is the role of β-adrenergic signaling in heart failure? // Circ. Res. 2003. V. 93. № 10. P. 896–906. https://doi.org/10.1161/01.RES.0000102042.83024.CA
18. *Brodde O.-E.* β1- and β2-adrenoceptor polymorphisms and cardiovascular diseases // Fundam. & Clin. Pharmacology. 2008. V. 22. № 2. P. 107–125. https://doi.org/10.1111/j.1472-8206.2007.00557.x
19. *Greenberg B., Yaroshinsky A., Zsebo K.M. et al.* Design of a phase 2b trial of intracoronary administration of AAV1/SERCA2a in patients with advanced heart failure: The CUPID 2 Trial (calcium up-regulation by per-

cutaneous administration of gene therapy in cardiac disease phase 2b) // JACC: Heart Failure. 2014. V. 2. № 1. P. 84–92. https://doi.org/10.1016/j.jchf.2013.09.008

20. *Cohn J.N., Levine T.B., Olivari M.T. et al.* Plasma norepinephrine as a guide to prognosis in patients with chronic congestive heart failure // N. Engl. J. Med. 1984. V. 311. P. 819–823. https://doi.org/10.1056/NEJM198409273111303
21. Митральная регургитация. Клинические рекомендации. Министерство здравоохранения РФ, 2016. 25 с. https://racvs.ru/clinic/files/2016/mitral-regurg.pdf

Expression of the β1-Adrenoreceptor Gene (*ADRB1*) in the Myocardium of Patients with Chronic Heart Failure

E. F. Muslimova[a], *, T. Yu. Rebrova[a], D. S. Kondratieva[a], E. L. Sonduev[a], B. N. Kozlov[a], and S. A. Afanasiev[a]

[a] *Cardiology Research Institute, Tomsk National Research Medical Center, Russian Academy of Sciences, Tomsk, 634012 Russia*

**e-mail: muslimovef@yandex.ru*

The analysis of the level of relative expression of the β1-adrenergic receptor *ADRB1* gene in the myocardium of 78 patients with chronic heart failure on the background of the combined development of ischemic heart disease and arterial hypertension was carried out. A significant decrease in the level of *ADRB1* gene expression was revealed in patients with left ventricular hypertrophy. In the considered cohort, the expression level of the *ADRB1* gene did not depend on the functional class of heart failure. The *ADRB1* relative expression was higher in patients taking acetylsalicylic acid preparations, but did not show a significant association with the use of β-blockers or ACE inhibitors.

Keywords: β-adrenergic receptors, *ADRB1*, gene expression, heart failure, left ventricular hypertrophy.

ГЕНЕТИКА, 2021, том 57, № 11, с. 1306–1317

ГЕНЕТИКА ЧЕЛОВЕКА

УДК [576.316.353.74+575.224.23]:616-055.9

ТКАНЕВОЙ ГОНОСОМНЫЙ МОЗАИЦИЗМ У ПАЦИЕНТОВ С НАРУШЕНИЕМ ФОРМИРОВАНИЯ ПОЛА, СВЯЗАННЫМ С АНОМАЛИЯМИ ДИФФЕРЕНЦИРОВКИ ГОНАД

© 2021 г. Н. В. Опарина[1, *], Н. Ю. Райгородская[2], О. Ю. Латышев[3], Л. Н. Самсонова[3], Е. А. Володько[3], С. А. Суязова[4], В. Б. Черных[5]

[1]Московский областной научно-исследовательский клинический институт им. М.Ф. Владимирского, Москва, 129110 Россия

[2]Саратовский государственный медицинский университет им. В.И. Разумовского, Саратов, 410012 Россия

[3]Российская медицинская академия непрерывного профессионального образования, Москва, 125993 Россия

[4]Государственный университет управления, Москва, 109542 Россия

[5]Медико-генетический научный центр им. академика Н.П. Бочкова, Москва, 115522 Россия

**e-mail: nv_oparina@mail.ru*

Поступила в редакцию 25.03.2021 г.
После доработки 04.06.2021 г.
Принята к публикации 15.06.2021 г.

Представлены результаты исследования гоносомного мозаицизма, в том числе его межтканевой вариабельности в лимфоцитах периферической крови, буккальном эпителии и ткани гонад, у пациентов с нарушением формирования пола (НФП), связанным с аномалиями половых хромосом (гоносомные НФП, $n = 8$) и с аномалиями дифференцировки гонад (46,XY НФП, $n = 10$). При исследовании использовался комплекс методов анализа отдельных клеток: стандартное кариотипирование в комбинации с интерфазным FISH. По результатам исследования гоносомный мозаицизм выявлен у 12 (67%) из 18 обследованных пациентов. Скрытый, изолированный тканью гонад, соматический мозаицизм с Х-моносомной клеточной линией выявлен у 4/10 пациентов из группы 46,XY НФП. Истинный, межтканевой, мозаицизм с двумя клеточными линиями, одна из которых с Х-моносомией, выявлен у 8/8 пациентов группы гоносомных НФП. Впервые для пациентов с истинным мозаицизмом была произведена количественная оценка межтканевой вариабельности Х-моносомной клеточной линии. Для этого использовался коэффициент вариации (*CV*), диапазон выявленных значений которого составил от 2.3 до 49%. В соответствии с полученными значениями *CV* межтканевая вариабельность оценена как низкая, умеренная или высокая. При проведении корреляционного анализа выявлена статистически значимая ($p < 0.1$) обратная зависимость степени маскулинизации по шкале Прадера от доли Х-моносомной клеточной линии в лимфоцитах, буккальном эпителии и ткани гонад, а также прямая связь присутствия Х-моносомной клеточной линии в одной из тканей с ее наличием в других тканях. Поиск мозаицизма в тканях разного гистогенеза с использованием комплекса методов анализа отдельных клеток позволяет получить более полное представление о гоносомной конституции и выявить скрытый дисбаланс, что очень важно при обследовании пациентов с НФП. Предложенный и использованный нами в настоящей работе коэффициент вариации (*CV*) является удобным инструментом для количественной оценки вариабельности клеточных линий в разных тканях одного пациента и шагом к стандартизации результатов генетических исследований.

Ключевые слова: половые хромосомы, мозаицизм, нарушение формирования пола, половая дифференцировка, дисгенезия гонад.

DOI: 10.31857/S0016675821110102

Нарушение формирования пола (**НФП**) – обширная гетерогенная группа врожденных заболеваний и патологических состояний, обусловленных несоответствием между генетическим, гонадным и фенотипическим полом. Частота встречаемости НФП составляет 1 : 4500 новорожденных [1]. Согласно современной классификации выделено три группы НФП, каждая из которых отражает хромосомный (генетический) пол и включает отдельные клинические варианты, указывающие на генетическую этиологию или морфологическую форму возникших аномалий дифференцировки гонад и развития половых органов:

I группа: гоносомные НФП, такие как синдром Шерешевского–Тернера (**СШТ**), синдром Клайнфельтера, X/XY смешанная дисгенезия го-

над или овотестикулярная форма, XX/XY химеризм и другие формы.

II группа: 46,XY НФП, такие как полная, неполная дисгенезия гонад и овотестикулярная форма, тестикулярная регрессия и различные нарушения биосинтеза и действия андрогенов и другие формы.

III группа: 46,XX НФП, такие как овотестикулярная форма, тестикулярная дисгенезия гонад, а также состояния, связанные с избытком андрогенов, и другие формы [1].

Нарушение дифференцировки гонад может быть вызвано различными причинами: аномалиями половых хромосом, гоносомными CNV (copy number variation), изменением дозы генов, контролирующих дифференцировку пола, генными мутациями, эпигенетическими нарушениями. Причиной гоносомных НФП являются численные и/или структурные аномалии половых хромосом или гетерокариотипический (46,XX/46,XY) химеризм. При этом возникает дисбаланс X- и Y-сцепленных генов, который и приводит к нарушению формирования пола на уровне дифференцировки гонад. Это важное отличие гоносомных от других групп НФП, разные нозологические варианты которых затрагивают как дифференцировку гонад, так и последующие этапы развития половой системы [2, 3].

Аномалии гоносом могут быть представлены как немозаичными, так и мозаичными формами и могут являться причиной не только нарушений формирования пола, гипогонадизма, но и задержки физического и умственного развития, репродуктивных проблем, однако могут встречаться и у здоровых индивидуумов. Ограниченное количество доступных для исследования тканей является препятствием для обнаружения соматического, в том числе гонадного, мозаицизма. Поэтому многие вопросы, связанные с распространенностью, динамикой и закономерностями распределения гоносомного мозаицизма в различных тканях, его влияние на фенотип недостаточно изучены, а его диагностика и клиническая интерпретация является актуальной проблемой [4].

Цель работы – исследование гоносомного мозаицизма в нескольких тканях у пациентов с разными формами нарушения формирования пола, связанными с аномалиями половых хромосом (гоносомные НФП) и с аномалиями дифференцировки гонад (46,XY НФП).

МАТЕРИАЛЫ И МЕТОДЫ

Выборку составили 18 пациентов с НФП, связанным с нарушением дифференцировки гонад, в том числе восемь пациентов с хромосомным НФП, различные варианты гоносомного дисбаланса у которых выявлены с помощью стандартного цитогенетического исследования (**СЦИ**), и десять пациентов с 46,XY НФП и дисгенезией гонад. Все случаи нарушения формирования пола являлись спорадическими, т.е. впервые отмечены в родословной, сибсы пациентов здоровы, браки родителей неродственные.

Из исследования исключали пациентов, у которых были диагностированы моногенные формы НФП, в том числе синдромы Дениса–Дрэша, Фрезье, дефекты биосинтеза и действия андрогенов (синдром тестикулярной феминизации, дефицит 5a-редуктазы и др.), врожденная гиперплазия коры надпочечников (дефицит 21-гидроксилазы и другие формы ВГКН).

Возраст обследованных пациентов варьировал от 5 мес. до 17 лет (медиана возраста составила 7.5 лет, с межквартильным интервалом 12.8 года). Среди обследованных девять детей в возрасте от 0 до 4 лет, девять подростков в возрасте от 11 до 17 лет.

Наружные половые органы при рождении сформированы неправильно у 14 из 18 пациентов (степень маскулинизации от II до III по шкале Прадера). В мужском поле воспитания зарегистрировано 12 пациентов, в женском поле воспитания – пять пациентов. У одного пациента с 46,XY НФП в возрасте 1 года 8 мес. выполнена смена гражданского пола с мужского на женский с последующим проведением феминизирующей пластики гениталий. Гонадэктомия выполнена у 18 пациентов. Дериваты Мюллеровых протоков (**ДМ**) обнаружены у 16 пациентов. У двух пациентов с 46,XY НФП и одного пациента с гоносомным НФП диагностирована гонадобластома. Материалом для исследования являлись образцы тканей разного эмбрионального происхождения (лимфоциты периферической крови, клетки буккального эпителия и ткань биоптатов гонад).

Стандартное цитогенетическое исследование выполняли на ФГА-стимулированных лимфоцитах периферической крови в соответствии с общепринятым протоколом [5]. Для визуализации бэндинга хромосом использовали GTG-метод дифференциального окрашивания. Анализировали от 20 до 100 метафазных пластинок [6].

Флуоресцентную *in situ* гибридизацию (FISH) выполняли на препаратах интерфазных ядер культивированных и некультивированных лимфоцитов периферической крови, некультивированных клеток буккального эпителия, на депарафинизированных срезах гистологических образцов биопсийного материала. Для определения числа копий хромосом использовали ДНК-зонды, специфичные к центромерным районам X-/Y-хромосом, аутосом 7 и 18 (SE X (DXZ1), Y (DYZ3), 7 (D7Z1), 18 (D18Z1); Kreatech, Leica, Нидерланды). Сигналы от аутосомных зондов использовали в качестве внутреннего контроля. Предгибридизационную подготовку и гибридизацию выполняли

согласно протоколу производителя (Poseidon, Kreatech, Leica, Нидерланды). Количество проанализированных клеток для каждого образца составило от 100 до 1000 в каждой исследуемой ткани.

Для визуализации метафазных GTG-окрашенных хромосом и флуоресценции гибридизационных сигналов использовали микроскоп Nikon Eclipse Ci (Nikon Corporation, Япония) с программным обеспечением ВидеоТестКарио 3.1 (Видеотест, Россия) и LUCIA Cytogenetics – FISH (LUCIA Cytogenetics, Чехия). Результаты исследований записаны в соответствии с Международной цитогеномной номенклатурой [7].

Статистический анализ выполняли с помощью непараметрических методов с использованием Excel из пакета программ Microsoft Office 2016 (Microsoft Inc., США), программы Statistica, версия 12 (StatSoft, США), а также R версии 4.0.3 (The R Foundation for Statistical Computing, Австрия) [8] с пакетами vegan [9] и pwr [10]. Для количественной оценки межтканевой вариабельности клеточных линий нами предложено использовать коэффициент вариации (*CV*), который рассчитывали по формуле:

$$CV = \frac{S_x}{\overline{x}} \times 100\%,$$

где S_x и $\overline{x}$ – соответственно стандартное отклонение и среднее доли Х-моносомии по отдельному пациенту [11, 12]. Значения *CV* расценивались согласно рекомендуемым критериям [11, 12], где интервал 0–15% соответствует низкой вариабельности, 15–30% – умеренной, 30–100% – высокой вариабельности. Корреляционный анализ проводили с использованием *r*-критерия Спирмена и *W*-критерия Кендалла. Согласно проведенной с помощью пакета [10] предварительной оценке мощность критерия значимости коэффициента корреляции, равного 0.4 (умеренная взаимосвязь), при 18 наблюдениях и стандартном уровне значимости $\alpha = 0.05$ составляет $(1 - \beta) = 0.39$. Этот показатель ниже рекомендуемого значения $(1 - \beta) = 0.8$, поэтому уровень значимости повышен до $\alpha = 0.10$, что для коэффициента 0.4 даст мощность критерия $(1 - \beta) = 0.52$. Далее в анализе статистически значимыми считали корреляции при уровне вероятности $p < 0.10$ [9, 12].

РЕЗУЛЬТАТЫ

В группе гоносомных НФП ($n = 8$) пациенты имели различные варианты дисгенезии гонад и фенотипические черты СШТ (табл. 1). Четыре пациента данной группы имели феминное строение наружных половых органов при рождении и женский пол воспитания. Внутренние гениталии представляли собой гипоплазированную матку или маточный тяж и фаллопиевы трубы, вдоль которых располагались гонады в виде фиброзных тяжей. При гистологическом исследовании гонад помимо соединительной ткани определялись элементы овариальной стромы. Еще четыре пациента этой группы при рождении определены и воспитывались в мужском поле, имели двойственное строение наружных половых органов: расщепленную мошонку, промежностную гипоспадию, микропенис. У них выявлены ДМ в виде гипоплазированной матки и маточных труб. Гонады этих пациентов дифференцированы асимметрично (по типу смешанной дисгенезии гонад): яичко с одной стороны и стрек с другой стороны (контралатерально). При гистологическом исследовании в ткани удаленных гонад обнаружены элементы недифференцированной стромы, фиброз. Гонадобластома диагностирована у одного пациента.

В группе 46,XY НФП ($n = 10$) пациенты имели различные варианты дисгенезии гонад. Женский пол при рождении определен у одного пациента. При обследовании в возрасте 16 лет у этого пациента отмечено правильное феминное строение наружных половых органов, полная билатеральная дисгенезия тестикул, гипоплазированная матка и трубы с обеих сторон, гонадобластома. У остальных пациентов этой группы (9/10) выявлена неполная форма дисгенезии гонад. Наружные половые органы имели двойственное строение с различной степенью маскулинизации: проксимальную форму гипоспадии в сочетании с крипторхизмом, недоразвитием кавернозных тел. При визуализации внутренних половых органов обнаружено наличие ДМ – гипоплазированной матки и маточных труб. Гонады были представлены яичком с одной стороны и дисгенетичной гонадой контралатерально у 6/9 пациентов. По данным гистологического исследования, у 4/6 пациентов дисгенетичные гонады представляли собой стрек. У 2/6 пациентов – двойственно-дифференцированные гонады, состоящие из овариального и тестикулярного компонентов (овотестис). Наличие двустороннего овотестис установлено у 1/9 пациентов. Билатеральная паховая ретенция яичек с признаками их атрофии и проявлениями гипергонадотропного гипогонадизма выявлена у 2/9 пациентов этой группы.

При СЦИ у всех пациентов с гоносомным НФП ($n = 8$) выявлен мозаичный кариотип с двумя клеточными линиями, одна из которых с Х-моносомией (табл. 2). Во второй клеточной линии структурно измененная Х-хромосома выявлена у 2/8 пациентов, нормальная Y-хромосома – у 3/8 пациентов и структурно измененная Y-хромосома – у 3/8 пациентов. По результатам FISH исследования простой мозаицизм обнаружен во всех исследованных тканях у пациентов этой группы.

Изолированный тканью гонад соматический гоносомный мозаицизм с двумя клеточными ли-

Таблица 1. Клинические характеристики пациентов с НФП

Код пациента	Возраст	Паспортный пол	Наружные половые органы при рождении, маскулинизация по шкале Прадера		Локализация и макроморфология гонад по данным УЗИ		Гонадэктомия	ДМ	Гистологический диагноз морфологической структуры удаленных гонад
					правая	левая			
1	14 лет	Ж	Двойственные	II	Не определяется	Не определяется	Д	+	Фиброзные тяжи с обеих сторон
2	17 лет	Ж	Женские	0	Не определяется	Не определяется	Д	+	Фиброзные тяжи с обеих сторон
3	5 мес.	М	Двойственные	III	Не определяется	Яичко в мошонке	П	+	Фиброзный тяж справа
4	12 лет	Ж	Женские	0	Не определяется	Не определяется	Д	+	Фиброзные тяжи с обеих сторон
5	17 лет	М	Двойственные	III	Гипоплазированная гонада в малом тазу	Яичко в мошонке	П	+	Дисгенетичное яичко с придатком, с герминативными клетками, без признаков функциональной активности справа
6	7 мес.	М	Двойственные	II	Гипоплазированная гонада в малом тазу	Яичко в мошонке	П	+	Тяж с элементами ткани яичника, эктопией ткани надпочечника, герминативными клетками, гонадобластома справа
7	14 лет	Ж	Женские	0	Не определяется	Не определяется	Д	+	Фиброзные тяжи с обеих сторон
8	11 лет	М	Двойственные	III	Яичко в паховом канале	Не определяется	Л	+	Фиброзный тяж с элементами ткани яичка слева
9	16 лет	Ж	Женские	I	Не определяется	Не определяется	Д	+	Фиброзные тяжи с обеих сторон, гонадобластома
10	14 лет	М	Двойственные	III	Яичко в мошонке, микролитиаз	Гипоплазированное яичко в паховом канале	Д	–	Яички с атрофией эпителия семенных канальцев, гиперплазия кл. Лейдига с обеих сторон, сперматоцеле справа
11	2 года	М	Двойственные	III	Яичко в мошонке	Гипоплазированная гонада в паховом канале	Л	+	Тяж с элементами ткани яичника и кл. Лейдига слева
12	12 лет	М	Двойственные	III	Дисгенетичное яичко с признаками атрофии в паховом канале	Гипоплазированное яичко	П	–	Соединительная ткань с элементами атрофичной ткани сперматогенного эпителия яичка справа
13	1 год	М	Двойственные	III	Яичко в мошонке	Не определяется	Л	+	Тяж с элементами ткани яичника и герминативными клетками, гонадобластома слева
14	4 года	М	Двойственные	III	Яичко в мошонке	Не определяется	Л	+	Овотестис слева
15	2 года	М	Двойственные	III	Яичко в мошонке	Абдоминальная гонада	Л	+	Дисгенетичное яичко слева
16	5 мес.	М	Двойственные	III	Яичко в паховом канале	Не определяется	Л	+	Фиброзный тяж слева
17	2 года	М	Двойственные	II	Яичко в мошонке	Абдоминальная гонада	Л	+	Овотестис слева
18	1 год	М → Ж	Двойственные	II	Абдоминальная гонада	Абдоминальная гонада	Л	+	Овотестис с обеих сторон

Примечание. М → Ж – смена пола с мужского на женский. Гонадэктомия: Д – двусторонняя, П – правосторонняя, Л – левосторонняя.

Таблица 2. Результаты молекулярно-цитогенетического исследования разных тканей у 18 пациентов с НФП

Код пациента	Паспортный пол	Кариотип по данным СЦИ	Выявленные в исследованных тканях клеточные линии с различным набором гоносом по результатам FISH								
			лимфоциты, % клеток			буккальный эпителий, % клеток			гонады, % клеток		
			X	XX	XY	X	XX	XY	X	XX	XY
1	Ж	mos45,X[45]/46,X,psuidic(X)(p11.2)[7]	**90**	**10**	0	**60**	**40**	0	**75**	**25**	0
2	Ж	mos45,X[30]/46,X,i(X)(q10)[5]	**94**	**6**	0	**93**	**7**	0	**90**	**10**	0
3	М	mos45,X[38]/46,XY[8]	**81**	0	**19**	**66**	0	**34**	**65**	0	**35**
4	Ж	mos45,X[52]/46,XY[48]	**51**	0	**49**	**79**	0	**21**	**85**	0	**15**
5	М	mos45,X[17]/46,X,psuidic(Y)(q11.2)[13]	**65**	0	**35**	н.и.	н.и.	н.и.	**44**	0	**56**
6	М	mos45,X[24]/46,X,del(Y)(q11.2)[6]	**83**	0	**17**	**67**	0	**33**	**95**	0	**5**
7	Ж	mos45,X[12]/46,X,psuidic(Y)(q11.2)[8]	**65**	0	**35**	н.и.	н.и.	н.и.	**60**	0	**40**
8	М	mos45,X[6]/46,XY[24]	**43**	0	**57**	**60**	0	**40**	**20**	0	**80**
9	Ж	46,XY	0	0	100	0	0	100	**9**	0	**91**
10	М	46,XY	0	0	100	0	0	100	**7**	0	**93**
11	М	46,XY	0	0	100	0	0	100	0	0	100
12	М	46,XY	0	0	100	0	0	100	0	0	100
13	М	46,XY	0	0	100	0	0	100	**7**	0	**93**
14	М	46,XY	0	0	100	0	0	100	**29**	0	**71**
15	М	46,XY	0	0	100	0	0	100	0	0	100
16	М	46,XY	0	0	100	0	0	100	0	0	100
17	М	46,XY	0	0	100	0	0	100	0	0	100
18	М → Ж	46,XY	0	0	100	0	0	100	0	0	100

Примечание. М → Ж – смена пола с мужского на женский, н.и. – не исследовано. X – клеточная линия с одной X-хромосомой, вторая гоносома отсутствует; XX – клеточная линия с двумя X-хромосомами; XY – клеточная линия с одной X-хромосомой и одной Y-хромосомой в каждой клетке. При обозначении клеточных линий макроструктурные перестройки вовлеченных гоносом не учитывались. Истинный гоносомный генотип указан в столбце “кариотип”.

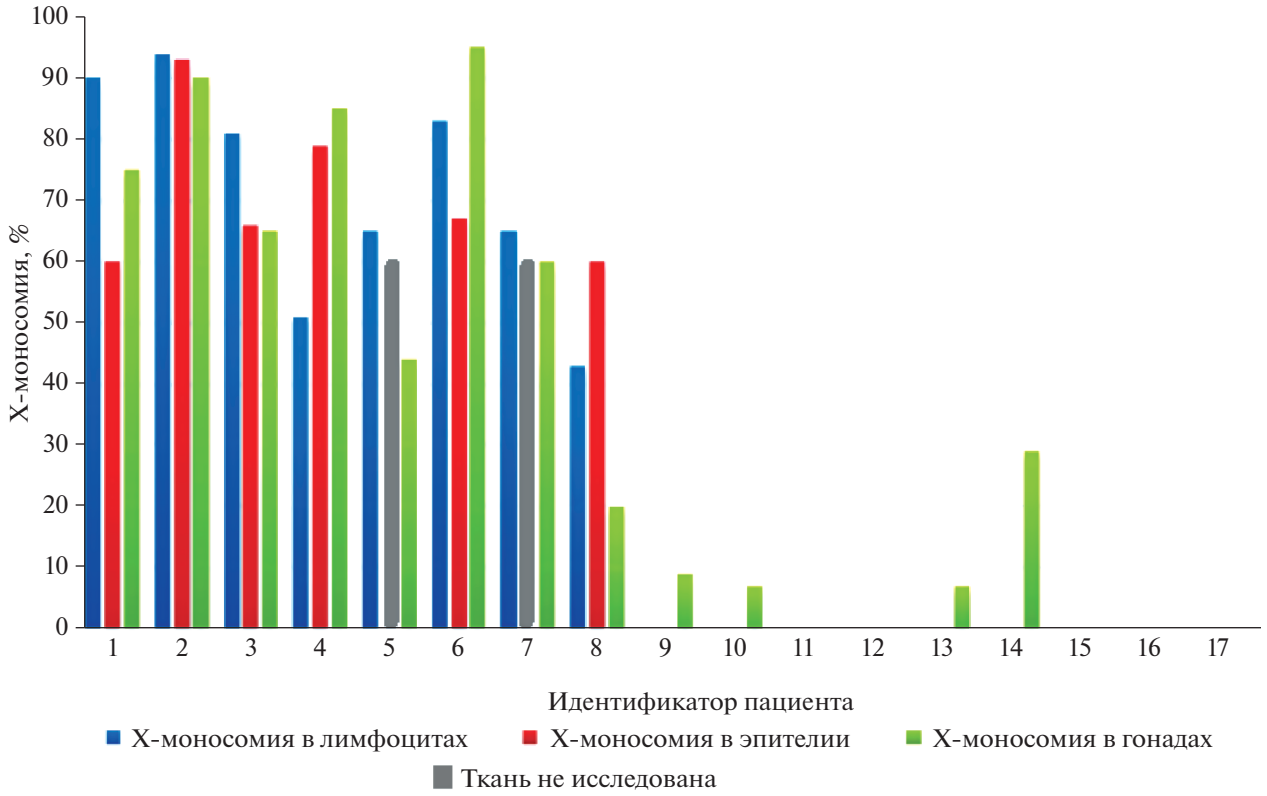

Рис. 1. Наличие и представленность X-моносомной клеточной линии в интерфазных ядрах клеток разных тканей в выборке из 18 пациентов с НФП. Каждому пациенту соответствуют три столбца разного цвета. Отсутствие столбца означает 0%. Столбец серого цвета у пациентов 5 и 7 имеет произвольную высоту и означает, что ткань не исследована.

ниями, одна из которых с X-моносомией, а вторая с XY, выявлен у 4/10 пациентов из группы 46,XY НФП.

У пациентов с истинным, межтканевым, гоносомным мозаицизмом из группы гоносомных НФП представленность X-моносомной клеточной линии варьировала от 43 до 94% в лимфоцитах, от 60 до 93% в буккальном эпителии, от 20 до 95% в ткани гонад (рис. 1).

Представленность X-моносомной клеточной линии у четырех пациентов с изолированным тканью гонад соматическим мозаицизмом из группы 46,XY НФП варьировала от 7 до 29%. Из них в трех наблюдениях (пациенты 9, 10, 13) представленность X-моносомной линии в ткани гонад составила менее 10%, в одном наблюдении (пациент 14) – 29%.

Для количественной оценки межтканевой вариабельности X-моносомной клеточной линии у пациентов с истинным, межтканевым, гоносомным мозаицизмом из группы гоносомных НФП ($n = 8$) нами предложен новый подход с использованием коэффициента вариации (*CV*) (рис. 2). В тех случаях, когда значение по какой-либо ткани отсутствовало (пациенты 5 и 7), расчет выполнен по исследованным тканям. Межтканевая вариабельность X-моносомной клеточной линии определена в диапазоне от 2.3 до 49%. У одного пациента межтканевая вариабельность расценена как высокая, у четырех – как умеренная и у трех – как низкая.

Самая низкая межтканевая вариабельность зарегистрирована у пациента женского пола (пациент 2), имеющего цитогенетически идентифицированные перестройки вовлеченной в мозаицизм X-хромосомы (табл. 3). При этом X-моносомная клеточная линия доминировала во всех исследованных тканях. Самая высокая межтканевая вариабельность зарегистрирована у пациента 8 с вовлеченной в мозаицизм Y-хромосомой без цитогенетически идентифицированных структурных изменений и с мужским полом воспитания. Доминирование X-моносомной клеточной линии отмечено только в одной из трех исследованных тканей: в буккальном эпителии.

С целью оценки клинической значимости цитогенетического исследования различных тканей был проведен корреляционный анализ. Для оценки возможной зависимости (как линейной, так и нелинейной) стадии маскулинизации по шкале Прадера от доли X-моносомной клеточной линии в различных тканях рассчитаны ранговые

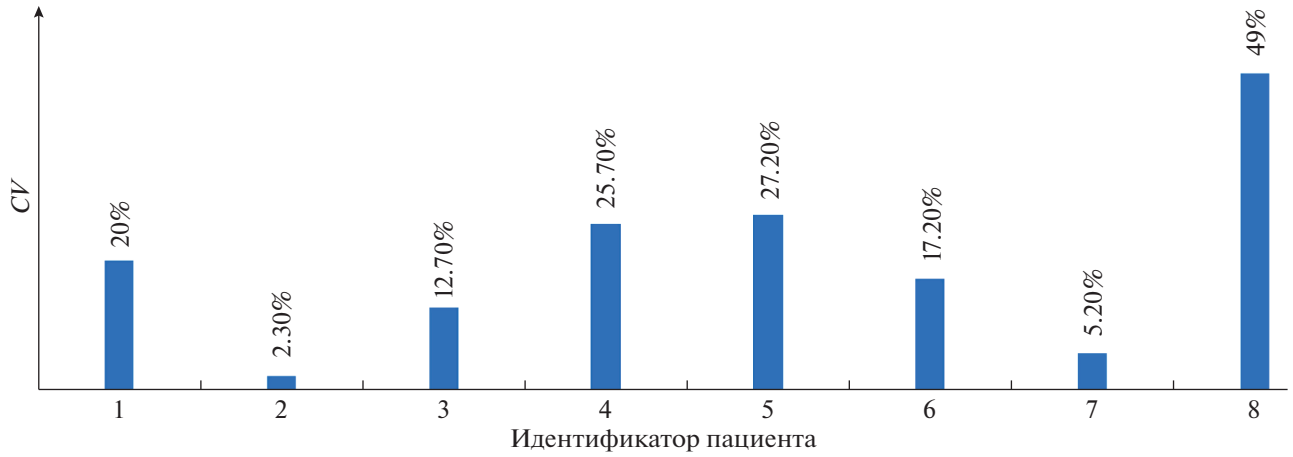

Рис. 2. Значения коэффициента вариации (*CV*), демонстрирующие межтканевую вариабельность доли Х-моносомной клеточной линии у пациентов из группы НФП гоносомные.

коэффициенты корреляции Спирмена (*r*) и конкордации Кендалла (*W*). В совершенной ранжировке доли Х-моносомной клеточной линии и коэффициента вариации наименьший ранг присваивался максимальному значению доли, а наибольший ранг присваивался наблюдениям с отсутствием Х-моносомной клеточной линии во всех трех исследованных тканях (11, 12, 15–18). В ранжировке шкалы Прадера наименьший ранг присвоен максимальному значению шкалы у пациентов описываемой нами группы. Тестировались двусторонние альтернативные гипотезы. Выявлена статистически значимая ($p < 0.1$) обратная зависимость степени маскулинизации по шкале Прадера от доли Х-моносомной клеточной линии в лимфоцитах ($r = -0.44$, $p = 0.0665$), буккальном эпителии ($r = -0.51$, $p = 0.0432$) и ткани гонад ($r = -0.49$, $p = 0.0397$), а также прямая связь присутствия Х-моносомной клеточной линии в одной из тканей с ее наличием в других тканях ($W = 0.78$, $p = 0.0010$).

Таким образом, с учетом исследования различных тканей гоносомный мозаицизм выявлен у 12 (67%) из 18 обследованных пациентов. Истинный, межтканевой, мозаицизм с двумя клеточными линиями, одна из которых с Х-моносомией, выявлен у 8/8 пациентов группы гоносомных НФП. Представленность Х-моносомной клеточной линии варьировала от 43 до 94% в лимфоцитах, от 60

Таблица 3. Данные о межтканевой вариабельности, доминировании Х-моносомной клеточной линии в различных тканях и о наличии макроструктурных перестроек в вовлеченной в мозаицизм гоносоме у пациентов с гоносомным НФП

Межтканевая вариабельность Х-моносомной клеточной линии	*CV*, %	Код пациента	Пол	Доминирование Х-моносомии в тканях			Изменение доли Х-моносомии в тканях	Вовлеченная в мозаицизм гоносома	Тип макроструктурной перестройки вовлеченной гоносомы	Точки разрыва
				Л	БЭ	Г				
Низкая	2.3	2	Ж	+	+	+	Нет	abnX	Изохромосома	q10
Низкая	5.7	7	Ж	+	н.и.	+	Нет	abnY	Псевдоизодицентрик	q11.2
Низкая	12.7	3	М	+	+	+	Нет	Y	Не выявлены	Отсутствуют
Умеренная	17.2	6	М	+	+	+	Г > Л > БЭ	abnY	Делеция	q11.2
Умеренная	20	1	Ж	+	+	+	Л > Г > БЭ	abnX	Псевдоизодицентрик	p11.2
Умеренная	25.7	4	Ж	–	+	+	Г > БЭ > Л	Y	Не выявлены	Отсутствуют
Умеренная	27.2	5	М	+	н.и.	–	Л > Г	abnY	Псевдоизодицентрик	q11.2
Высокая	49	8	М	–	+	–	БЭ > Л > Г	Y	Не выявлены	Отсутствуют

Примечание. н.и. – не исследовано, “+” – наличие параметра, “–” – отсутствие параметра, Л – лимфоциты, БЭ – буккальный эпителий, Г – гонады, abn – структурно измененная гоносома.

до 93% в буккальном эпителии, от 20 до 95% в ткани гонад. Межтканевая вариабельность Х-моносомной клеточной линии у 3/8 пациентов оценена как низкая, у 4/8 пациентов как умеренная и у 1/8 пациентов как высокая. Диапазон *CV* составил от 2.3 до 49%. Изолированный тканью гонад соматический мозаицизм с Х-моносомной клеточной линией выявлен у 4/10 пациентов из группы 46,XY НФП. Представленность Х-моносомной клеточной линии при этом варьировала от 7 до 29%.

Выявлена статистически значимая ($p < 0.1$) обратная зависимость степени маскулинизации по шкале Прадера от доли Х-моносомной клеточной линии в лимфоцитах ($r = -0.44$, $p = 0.0665$), буккальном эпителии ($r = -0.51$, $p = 0.0432$) и ткани гонад ($r = -0.49$, $p = 0.0397$), а также прямая связь присутствия Х-моносомной клеточной линии в одной из тканей с ее наличием в других тканях ($W = 0.78$, $p = 0.0010$).

ОБСУЖДЕНИЕ

Аномалии гоносом, возникающие на генном, хромосомном, геномном и эпигенетическом уровнях организации генома, могут изменять экспрессию Х- и Y- сцепленных про-тестикулярных генов и являться одной из причин нарушения дифференцировки гонад [4, 13, 14]. Гоносомный дисбаланс, возникающий на постзиготических стадиях из-за нарушения митотической сегрегации хромосом, может проявляться различными вариантами мозаицизма и широким спектром фенотипов [2, 3]. Сложность диагностики мозаицизма заключается как в ограниченном количестве доступных для исследования тканей, так и в отсутствии универсального метода.

В настоящей работе нами представлены результаты исследования гоносомного мозаицизма в лимфоцитах, буккальном эпителии и ткани гонад у 18 пациентов с 46,XY и дисгенезией гонад ($n = 10$) и гоносомной ($n = 8$) формами НФП. При исследовании использован комплекс методов анализа отдельных клеток: классическое кариотипирование в комбинации с интерфазным FISH. По результатам исследования у 4/10 пациентов с 46,XY НФП в ткани гонад выявлен скрытый мозаицизм, одна из линий которого с Х-моносомией. Представленность Х-моносомной клеточной линии варьировала от 7 до 29%.

46,XY НФП является высокогетерогенной группой, в которую включено множество клинико-генетических вариантов. В литературе описаны случаи выявления скрытого тканевого гоносомного мозаицизма с Х-моносомной клеточной линией у 46,XY пациентов с тестикулярной дисгенезией гонад и овотестикулярной формой, при тестикулярной регрессии, в том числе у пациентов с синдромом Свайера [15–19]. Наличие Х-моносомной клеточной линии в бипотенциальной гонаде может негативно отразиться на ее дальнейшей дифференцировке [20]. Однако постзиготическая потеря Y-хромосомы человека, являющаяся следствием мутаций в соматических клетках, связана с широким спектром заболеваний, таких как рак, болезнь Альцгеймера, патология сердца и сосудов [21]. Это необходимо учитывать для корректной интерпретации результатов настоящей работы, свидетельствующих о наличии изолированного соматического мозаицизма. Нами применялся таргетный подход поиска гоносомного дисбаланса с использованием ограниченного количества ДНК-зондов. Возможно, у пациентов с уровнем мозаицизма менее 10% нарушение генотипа клеток с выявленной Х-моносомией носит более сложный характер и отражает совсем другие процессы, например малигнизацию. Из гистологического описания известно, что гонадобластома диагностирована у пациентов 9 и 13. У пациента 10 обнаружено наличие тестикулярного микролитиаза, сперматоцеле, атрофии эпителия яичка. Такие процессы могут сопровождаться усилением апоптоза или онкозом клеток в поврежденном органе и стать одной из возможных причин появления анеуплоидной линии. В работе Faure-Conter [22] сообщается о семи случаях, когда выявление герминально-клеточного рака предшествовало диагнозу НФП, и о четырех случаях, когда диагноз НФП предшествовал выявлению онкологического процесса в гонадах. С другой стороны, представленность Х-моносомной клеточной линии в ткани гонад у пациента 14 выше, чем у пациента 8 с истинным, межтканевым, мозаицизмом и кариотипом 45,X/46,XY из группы гоносомных НФП (29 и 20% соответственно). Сочетание 46,XY кариотипа и овотестикулярной формы является одним из редких вариантов нарушения дифференцировки гонад и может возникать в результате изменения экспрессии про-тестикулярных (*SRY, SOX9* и др.) генов. Как вариант нарушения дифференцировки овотестикулярная форма встречается с частотой 1 : 100000 новорожденных, описана у пациентов с 46,XX кариотипом, с 46,XX/46,XY химеризмом или 45,X/46,XY мозаицизмом [23]. Очевидно, что тестикулярная регрессия, частичная дисгенезия гонад при кариотипе 46,XY и смешанная дисгенезия гонад при 45,X/46,XY гоносомном НФП представляют собой близкие, сходные, частично генетически и фенотипически перекрывающиеся клинико-морфологические варианты нарушения дифференцировки гонад и формирования пола, проявляющиеся

широким спектром фенотипов и дефицитом маскулинизации [18].

Нами выявлена статистически значимая ($p < 0.10$) обратная корреляция степени маскулинизации по шкале Прадера от доли Х-моносомной клеточной линии в лимфоцитах ($r = -0.44$, $p = 0.0665$), буккальном эпителии ($r = -0.51$, $p = 0.0432$) и ткани гонад ($r = -0.49$, $p = 0.0397$). Однако объем исследованной выборки не позволяет повысить уровень значимости выявленной зависимости.

По результатам исследования у пациентов с гоносомным НФП ($n = 8$) во всех исследованных тканях выявлен мозаицизм, состоящий из двух клеточных линий, одна из которых с Х-моносомией. Вторая клеточная линия у всех пациентов являлась эуплоидной и содержала Х- или Y-хромосому с макроструктурными изменениями (5/8) или без таковых (3/8). Других клеточных линий, отличающихся копийностью или дополнительными структурными изменениями гоносом, не выявлено. Это позволяет говорить об истинном, межтканевом, простом мозаицизме со стабильным количеством клеточных линий во всех исследованных тканях. Наиболее вероятный генез такого мозаицизма митотический, в результате анафазного отставания одной из гоносом в изначально эуплоидной зиготе. Структурные перестройки гоносом, дефекты центромеры, кинетохора и т.д. могут быть триггерами этого процесса [21].

Возникновение Х-моносомной клеточной линии может произойти на разных этапах формирования или развития эмбриона, а скорость накопления мутаций в тканях различна. Поэтому нами была исследована межтканевая вариабельность доли Х-моносомной клеточной линии у каждого пациента, что является важным аналитическим параметром, отражающим митотическую стабильность вовлеченной хромосомы и активность процесса накопления мутаций в тканях. Впервые произведена количественная оценка результатов с использованием коэффициента вариации (*CV*) как инструмента со стандартизованной интерпретацией полученных значений, исходя из которой межтканевая вариабельность может быть расценена как низкая, умеренная или высокая. Полученные значения *CV* находились в диапазоне от 2.3 до 49%. Только у 3/8 пациентов *CV* не превышал 15% и был расценен как низкий, что позволяет считать долю Х-моносомии во всех исследованных тканях идентичной. У 4/8 пациентов вариабельность расценена как умеренная при интервале *CV* от 15 до 30%. И низкая, и умеренная вариабельность наблюдалась у пациентов как с вовлеченной в мозаицизм структурно измененной гоносомой, так и с нормальной Y-хромосомой. Но у пациентов с низкой вариабельностью Х-моносомная клеточная линия доминировала во всех исследованных тканях. При умеренной вариабельности доминирование Х-моносомии во всех исследованных тканях наблюдали в двух случаях (пациенты 1 и 6). В одном случае Х-моносомия доминировала в буккальном эпителии и гонадах при равном соотношении клеточных линий в лимфоцитах (пациент 4). В одном случае Х-моносомия доминировала в лимфоцитах при равном соотношении клеточных линий в гонадах (пациент 5). Высокая вариабельность ($CV = 49\%$) выявлена у одного пациента (пациент 8) с вовлеченной в мозаицизм Y-хромосомой без цитогенетически идентифицированных структурных изменений. Х-моносомия доминировала только в буккальном эпителии этого пациента и составила 60%. В лимфоцитах соотношение Х-моносомной и Y-несущей клеточных линий составило 43 : 57% соответственно. А в ткани гонад Х-моносомия составила 20%. Это самый низкий показатель Х-моносомии в ткани гонад среди пациентов с гоносомным НФП. И этот показатель ниже, чем у пациента 14 с соматическим гоносомным мозаицизмом и овотестикулярной формой НФП.

Результаты ряда исследований свидетельствуют о межтканевой вариабельности доли клеточных линий у пациентов с гоносомным мозаицизмом. Так, Peek и соавт. [24], обследовав группу из 10 пациенток с СШТ, пришли к выводу, что мозаицизм в лимфоцитах, буккальном и уретральном эпителии не коррелирует с мозаицизмом в гонадах. Кроме того, сепарировав и отдельно исследовав ооциты, гранулезные и стромальные клетки, они выявили в них разную частоту встречаемости Х-моносомии. Reddy и Sulkova [25] описали пять пациентов с 45,X/46,XY кариотипом и выраженным различием доли Х-моносомии как между лимфоцитами и гонадами, так и между клетками Сертолли и Лейдига в дисгенетичных гонадах. Nomura и соавт. [26] сообщали о пациенте с 45,X/46,XY кариотипом и экстремальной вариабельностью Y-несущей клеточной линии: 0.2% в лимфоцитах и 50% в гонадах. Анализу межтканевой вариабельности посвящена работа Graff и соавт. [27], где сообщается, что при исследовании лимфоцитов и буккального эпителия от 142 взрослых с СШТ доля Х-моносомии между двумя тканями оказалась идентична в 32.4% наблюдений, а несоответствие выявлено в 41% наблюдений.

Наличие самого явления межтканевой вариабельности клеточных линий, особенно высоких и экстремальных вариантов, свидетельствует о том, что случаи межтканевого мозаицизма могут выявляться и интерпретироваться как изолированный мозаицизм. Нами выявлена статистически значимая связь присутствия Х-моносомной кле-

точной линии в одной из тканей с ее наличием в других тканях ($W = 0.78$, $p = 0.0010$). Ограниченное количество доступных для исследования тканей является одним из существенных препятствий при изучении мозаицизма.

Идентификация и анализ структурной целостности вовлеченной в мозаицизм гоносомы, так же как и идентификация сформировавшихся клеточных линий, имеют прогностическое значение. При выборе тактики лечебных мероприятий наличие нормальной или структурно аномальной Y-хромосомы у пациентов женского пола является показанием для проведения профилактической гонадэктомии из-за высокого риска возникновения и прогрессии неопластического процесса в дисгенетичной гонаде [28]. Учитывая возможность межтканевой вариабельности клеточных линий, поиск Y-хромосомы в тканях разного гистогенеза снижает риск ошибок диагностики. Пациентки с мозаицизмом X/XX имеют лучший репродуктивный прогноз и большую продолжительность жизни, чем пациентки со структурно измененной X-хромосомой, вовлеченной в мозаицизм [29]. В то же время при мозаицизме X/XX с низкой долей XX-эуплоидных клеток требуется уточнение происхождения каждой из X-хромосом для исключения однородительской дисомии. Поиск мозаицизма оправдан и при задержке роста, которая является одним из частых фенотипических эффектов X-моносомии. У пациентов с 45,X/46,XY мозаицизмом задержка роста чаще всего происходит после двух лет, при этом пациенты женского пола отстают в росте значительнее, чем пациенты мужского пола [30].

Сложность диагностики мозаицизма заключается в том, что на сегодняшний день нет универсального метода, а есть два принципиально разных подхода: анализ либо отдельных клеток, либо совокупности (массива) клеток. Преимуществом использованного нами комплексного подхода к анализу отдельных клеток является возможность оценить различные варианты хромосомной конституции (в пределах разрешающей способности метода), включая сбалансированные перестройки и низкоуровневый мозаицизм (менее 5%), корректно описать все присутствующие клеточные линии, в том числе динамический мозаицизм, а также проанализировать межтканевую вариабельность распределения клеточных линий. Следует отметить, что на сегодняшний день современные молекулярные методы анализа массива клеток, такие как сравнительная геномная гибридизация (arrayCGH) и секвенирование нового поколения (NGS), такой возможностью не обладают [31].

Таким образом, поиск мозаицизма в тканях разного гистогенеза с использованием комплекса методов анализа отдельных клеток позволяет получить более полное представление о гоносомной конституции и выявить скрытый дисбаланс, что очень важно при обследовании пациентов с НФП.

Впервые произведенная в настоящем исследовании количественная оценка межтканевой вариабельности клеточных линий при мозаицизме с использованием коэффициента вариации (*CV*) является шагом к стандартизации результатов генетических исследований, может способствовать выявлению новых закономерностей при изучении мозаицизма и облегчает клиническую интерпретацию полученных результатов.

Работа проведена без привлечения дополнительного финансирования со стороны третьих лиц.

Все процедуры, выполненные в исследовании с участием людей, соответствуют этическим стандартам институционального и/или национального комитета по исследовательской этике и Хельсинкской декларации 1964 г. и ее последующим изменениям или сопоставимым нормам этики. От каждого из включенных в исследование участников было получено информированное добровольное согласие.

Авторы заявляют, что у них нет конфликта интересов. Все авторы внесли существенный вклад в подготовку статьи, прочли и одобрили финальную версию.

СПИСОК ЛИТЕРАТУРЫ

1. *Hughes I.A., Houk C., Ahmed S.F. et al.* Consensus statement on management of intersex disorders // Arch. Dis. Child. 2006. V. 91. № 7. P. 554–563. https://doi.org/10.1136/adc.2006.098319
2. *Gardner R.J.M., David J.A.* Chromosome Abnormalities and Genetic Counseling. 5th ed. Oxford Univ. Press, 2018. 634 p.
3. *Курило Л.Ф.* Хромосомные заболевания органов половой системы // Клинич. и эксперим. морфология. 2015. № 1. С. 48–59.
4. *Курило Л.Ф., Андреева М.В., Коломиец О.Л. и др.* Генетические синдромы с нарушениями развития органов половой системы // Андрология и генитальная хирургия. 2013. № 4. С. 17–27.
5. *Рубцов Н.Б.* Методы работы с хромосомами млекопитающих: учебное пособие. Н.: Новосибирский гос. ун-т, 2006. 152 с.
6. *Кузнецова Т.В., Шилова Н.В., Творогова М.Г. и др.* Практические рекомендации по обеспечению качества и надежности цитогенетических исследований // Мед. генетика. 2019. Т. 18. № 5. С. 3–27. https://doi.org/10.25557/2073-7998.2019.05.3-27
7. *McGowan-Jordan J., Simons A., Schmid M.* ISCN 2016: An International System for Human Cytogenomic Nomenclature (2016) // Reprint of Cytogenetic and Genome Research. 2016. V. 149. № 1–2. 140 p.

8. R Core Team. [Электронный ресурс]. URL https://www.R-project.org/ R: A language and environment for statistical computing. R Foundation for Statistical Computing (02.12.2020).

9. Vegan: Community Ecology Package. R package version 2.5-7 [Электронный ресурс]. URL https://CRAN.R-project.org/package=vegan (02.12.2020).

10. pwr: Basic Functions for Power Analysis. R package version 1.3-0 [Электронный ресурс]. URL https://CRAN.R-project.org/package=pwr (02.12.2020).

11. *Ефимова М.Р., Петрова Е.В., Ганченко О.И., Михайлов М.Ф.* Общая теория статистики. Практикум. 4-е изд. М.: Юрайт, 2019. 355 с.

12. *Писарева О.М.* Методы прогнозирования развития социально-экономических систем. М.: Высш. шк., 2007. 591 с.

13. *Gomes N.L., Chetty T., Jorgensen A., Mitchell R.T.* Disorders of sex development–novel regulators, impacts on fertility, and options for fertility preservation // Int. J. Mol. Sci. 2020. V. 21. № 7. P. 2282. https://doi.org/10.3390/ijms21072282

14. *Veitia R.A., Salas-Corte's L., Ottolenghi C. et al.* Testis determination in mammals: More questions than answers // Mol. Cell. Endocrinology. 2001. V. 179. № 1–2. P. 3–16. https://doi.org/10.1016/s0303-7207(01)00460-9

15. *Röpke A., Kalinski T., Mohnike K. et al.* Distribution of sex chromosomes in dysgenetic gonads of mixed type // Cytogenet. Genome Res. 2007. V. 116. № 1–2. P. 146–151. https://doi.org/10.1159/000097435

16. *Röpke A., Pelz A.F., Volleth M. et al.* Sex chromosomal mosaicism in the gonads of patients with gonadal dysgenesis, but normal female or male karyotypes in lymphocytes // Am. J. Obstet. Gynecol. 2004. V. 190. № 4. P. 1059–1062. https://doi.org/10.1016/j.ajog.2003.09.053

17. *Berberoğlu M., Şıklar Z.* The evaluation of cases with y-chromosome gonadal dysgenesis: Clinical experience over 18 years // J. Clin. Res. Pediatr. Endocrinol. 2018. V. 10. № 1. P. 30–37. https://doi.org/10.4274/jcrpe.4826

18. *Andrade J.G., Guerra-Júnior G., Maciel-Guerra A.T.* 46,XY and 45,X/46,XY testicular dysgenesis: Similar gonadal and genital phenotype, different prognosis // Arq. Bras. Endocrinol. Metabol. 2010. V. 54. № 3. P. 331–334. https://doi.org/10.1590/s0004-27302010000300013

19. *Chand M.T., Turner S., Solomon L.A. et al.* A case of 45,X/46,XY mosaicism presenting as Swyer syndrome // J. Pediatr. Adolesc. Gynecol. 2020. V. 33. № 5. P. 575–580. https://doi.org/10.1016/j.jpag.2020.06.008

20. *Kamel A.K., Abd El-Ghany H.M., Mekkawy M.K. et al.* Sex chromosome mosaicism in the gonads of DSD patients: A karyotype/phenotype correlation // Sex. Dev. 2015. V. 9. № 5. P. 279–288. https://doi.org/10.1159/000442332

21. *Guo X., Dai X., Zhou T. et al.* Mosaic loss of human Y chromosome: what, how and why // Hum. Genet. 2020. V. 139. № 4. P. 421–446. https://doi.org/10.1007/s00439-020-02114-w

22. *Faure-Conter C., Orbach D., Fresneau B. et al.* Disorder of sex development with germ cell tumors: Which is uncovered first? // Pediatr. Blood Cancer. 2020. V. 67. № 4. P.e28169. https://doi.org/10.1002/pbc.28169

23. *Mekkawy M.K., Kamel A.K., Dessouky N. et al.* Cytogenetic spectrum of ovotesticular difference of sex development (OT DSD) among a large cohort of DSD patients and literature review // Sex. Dev. 2019. V. 13. № 5–6. P. 221–227. https://doi.org/10.1159/000508153

24. *Peek R., Schleedoorn M., Smeets D. et al.* Ovarian follicles of young patients with Turner's syndrome contain normal oocytes but monosomic 45,X granulosa cells // Hum. Reprod. 2019. V. 34. № 9. P. 1686–1696. https://doi.org/10.1093/humrep/dez135

25. *Reddy K.S., Sulcova V.* Pathogenetics of 45,X/46,XY gonadal mosaicism // Cytogenet. Cell Genet. 1998. V. 82. № 1–2. P. 52–57. https://doi.org/10.1159/000015064

26. *Nomura R., Miyai K., Okada M. et al.* A 45,X/46,XY DSD (Disorder of Sexual Development) case with an extremely uneven distribution of 46,XY cells between lymphocytes and gonads // Clin. Pediatr. Endocrinol. 2015. V. 24. № 1. P. 11–14. https://doi.org/10.1297/cpe.24.11

27. *Graff A., Donadille B., Morel H. et al.* Added value of buccal cell FISH analysis in the diagnosis and management of Turner syndrome // Hum. Reprod. 2020. V. 35. № 10. P. 2391–2398. https://doi.org/10.1093/humrep/deaa197

28. *Baer T.G., Freeman C.E., Cujar C. et al.* Prevalence and physical distribution of SRY in the gonads of a woman with Turner syndrome: Phenotypic presentation, tubal formation, and malignancy risk // Horm. Res. Paediatr. 2017. V. 88. № 3–4. P. 291–297. https://doi.org/10.1159/000477240

29. *Machiela M.J., Zhou W., Karlins E. et al.* Female chromosome X mosaicism is age-related and preferentially affects the inactivated X chromosome // Nat. Communications. 2016. № 7. P. 11843. https://doi.org/10.1038/ncomms11843

30. *Pan L., Su Z., Song J. et al.* Growth data and tumour risk of 32 Chinese children and adolescents with 45,X/46,XY mosaicism // BMC Pediatrics. 2019. V. 19. № 1. P. 143. https://doi.org/10.1186/s12887-019-1520-9

31. *Черных В.Б., Яманди Т.А., Сафина Н.Ю.* Новые молекулярные технологии в диагностике генетических причин мужского бесплодия // Андрология и генитальная хирургия. 2017. Т. 18. № 1. С. 10–22. https://doi.org/10.17650/2070-9781-2017-18-1-10-22

Inter-Tissue Gonosomal Mosaicism in Patients with Disorders Sex Development, Related to Gonadal Differentiation Abnormalities

N. V. Oparina[a, *], N. Yu. Raygorodskaya[b], O. Yu. Latyshev[c], L. N. Samsonova[c], E. A. Volodko[c], S. A. Suyazova[d], and V. B. Chernykh[e]

[a] *Vladimihsky Moscow Regional Scientific Research Clinical Institute, Moscow, 129110 Russia*
[b] *Razumovsky Saratov State Medical University, Saratov, 410012 Russia*
[c] *Russian Medical Academy of Continuous, Moscow, 125993 Russia*
[d] *The State University of Management, Moscow, 109542 Russia*
[e] *Research Centre for Medical Genetics, Moscow, 115522 Russia*
**e-mail: nv_oparina@mail.ru*

The aim of the study was to identify gonosomal mosaicism and study its interstitial variability in peripheral blood lymphocytes, buccal epithelium, and gonadal tissue in patients with disorders of sex development (DSD) associated with sex chromosome abnormalities (gonosomal DSD, $n = 8$) and gonad differentiation abnormalities (46,XY DSD, $n = 10$). The study used a set of methods for analyzing individual cells: standard karyotyping in combination with interphase FISH. According to the results of the study, gonosomal mosaicism was detected in 12 (67%) of the 18 examined patients. Hidden, isolated by gonadal tissue, somatic mosaicism with an X-monosomic cell line was detected in 4/10 patients from group 46,XY DSD. True interstitial mosaicism with two cell lines, one of them with X-monosomy, was detected in 8/8 of patients in the group of gonosomal DSD. For the first time in patients with true mosaicism, the inter-tissue variability of the X-monosomal cell line was quantified. For this purpose, the coefficient of variation (*CV*) was used, the range of detected values was from 2.3 to 49%. In accordance with the obtained *CV* values, the interstitial variability is estimated as low, moderate or high. The correlation analysis revealed a statistically significant ($p < 0.1$) inverse dependence of the degree of masculinization on the Prader's scale on the proportion of X-monosomal cell line in lymphocytes, buccal epithelium, and gonadal tissue, as well as a direct relationship between the presence of X-monosomic cell line in one of the tissues and its presence in other tissues. The search for mosaicism in tissues of different histogenesis using a set of methods for analyzing individual cells allows us to get a more complete picture of the gonosomal constitution and reveal a hidden imbalance, which is very important in examining patients with DSD. The coefficient of variation (*CV*) proposed and used in this study is a convenient tool for quantifying the variability of cell lines in different tissues of a single patient and a step towards standardizing the results of genetic studies.

Keywords: sex chromosomes, mosaicism, disorders of sex development, sex differentiation, gonadal dysgenesis

ГЕНЕТИКА, 2021, том 57, № 11, с. 1318–1329

МАТЕМАТИЧЕСКИЕ МОДЕЛИ И МЕТОДЫ

УДК 303.725.23:575.177:303.725.23

К АНАЛИЗУ СЛУЧАЙНЫХ ПРОЦЕССОВ ИЗОНИМИИ. II. ДИНАМИКА ДИВЕРГЕНЦИИ ПОПУЛЯЦИЙ

© 2021 г. В. П. Пасеков*

Вычислительный центр им. А.А. Дородницына Федерального исследовательского центра "Информатика и управление" Российской академии наук, Москва, 119991 Россия

**e-mail: pass40@mail.ru*

Поступила в редакцию 03.02.2021 г.
После доработки 21.04.2021 г.
Принята к публикации 27.04.2021 г.

Рассматривается случайная динамика фамильного состава популяции конечного размера в дискретном времени при неперекрывающихся поколениях. Наследование фамилий предполагается патрилинейным. Динамика анализируется на малом промежутке эффективного времени $t/NE(t)$, где $NE(t)$ – средний гармонический эффективный размер популяции за t поколений. Так как при этом систематическими давлениями можно пренебречь, то в результате фамильная микроэволюция приближенно соответствует процессу случайного генного дрейфа, синхронно протекающему в той же самой популяции с вчетверо меньшей, чем для фамилий, интенсивностью. Подобно модели дрейфа генов фамильный состав очередного поколения τ является случайной выборкой с возвращением размера $Ne(\tau)/2$ из фамилий мужской составляющей родительской популяции, т. е. размера в 4 раза меньшего, чем выборка $2Ne(\tau)$ гамет при дрейфе генов ($Ne(\tau)$ – эффективный размер популяции в поколении τ). Изучается динамика вероятности случайной встречи индивидуума с однофамильцем и поведение вероятности встречи индивидуумов с разными фамилиями. Данные вероятности аналогичны гомозиготам и гетерозиготам соответственно при анализе генетической структуры. Приведены точные зависимости от времени для указанных вероятностей, дисперсий концентраций фамилий и фамильного аналога коэффициента инбридинга. Дана аппроксимация точных зависимостей более простыми при малой величине эффективного времени $t/NE(t)$, где дивергенция фамилий идет вчетверо быстрей, чем генная. Результаты не предполагают монофилии фамилий и описывают умозрительную теоретическую совокупность популяций-реплик, как бы прошедших заново микроэволюционную историю рассматриваемой популяции в тех же условиях. Использование малого по сравнению с размером популяции времени оправдано недавним возникновением большинства фамилий в России и тем, что прошедшее время в поколениях много меньше типичных размеров популяций. В реальных подразделенных популяциях процедуры оценивания коэффициента инбридинга по концентрациям фамилий не позволяют различить ситуации механической смеси субпопуляций или их общего происхождения.

Ключевые слова: теоретическая популяционная генетика, популяции с неперекрывающимися поколениями, коэффициент инбридинга, дрейф фамилий, случайный генный дрейф, математические модели, динамика дивергенции фамильного состава, асимптотика на малом эффективном времени.

DOI: 10.31857/S0016675821110114

Фамильный состав популяции отражает ее этническую принадлежность, происхождение, генеалогию (всем хотелось бы разузнать побольше о своих корнях) и он любопытен сам по себе [1, 2]. Известно, что между передачей потомкам фамилии и родительских генов имеется сходство. Для нас наибольший интерес представляет связь фамильной и генетической структур [3–8].

Всплеск публикаций в этом направлении появился после статьи [3], посвященной оценке коэффициента инбридинга популяции по частоте изонимных браков (см. критические моменты использования данных по изонимным бракам в [9]). Библиография подобных работ приведена в [10]. В первой части серии [11] и в настоящей работе для анализа фамильной структуры привлекаются модификации соответствующих простейших популяционно-генетических моделей популяции ограниченного размера. Подробный анализ этой связи мы начали в первой части [11], в основном посвященной описанию структуры в статике (в фиксированный момент времени). В настоящей статье основное внимание уделено динамике фамильной структуры.

Хотя пришедший из древней Греции термин "изонимия" означает равенство всех перед законом,

в последующем расширении этого понятия под изонимией стали понимать совокупность одинаковых названий. Например, изонимными считаются браки, в которых фамилия супруга совпадает с девичьей фамилией жены (браки между однофамильцами). Применительно к области популяционной генетики широкое распространение получил метод оценивания коэффициента инбридинга популяции по частоте изонимных браков, но использование изонимии не ограничено только этим подходом.

Очевидно, изонимию также характеризуют состав и количество однофамильцев в популяции. Пусть k – количество разных фамилий, а m_i обозначает количество индивидуумов с i-й фамилией в популяции размера N. Доля i-й фамилии $p_i \equiv m_i/N$ показывает ее распространенность, а совокупность таких долей дает распределение фамилий в популяции (фамильное состояние), причем

$$\sum_{i=1}^{k} m_i = N, \quad p_i \equiv m_i/N, \quad \sum_{i=1}^{k} p_i = 1.$$

Распространенность (долю, концентрацию) однофамильцев с учетом всех фамилий можно было бы подсчитать как ее среднее значение вида

$$\left(\sum_{i=1}^{k} p_i\right)\Big/k = \sum_{i=1}^{k}(m_i/N)/k = (N/N)/k = 1/k.$$

Однако такая характеристика явно мало информативна, так как тождественно равна $1/k$ при любом распределении однофамильцев. Ее недостаток в том, что не учитывается "вес" отдельных фамилий, т.е. их разная распространенность. Например, если при $k = 2$ в популяции размера 1000 индивидуумов было 998 однофамильцев по первой фамилии и 2 по второй, то данный показатель распространенности равен 1/2 и будет таким же при одинаковом (по 500) числе однофамильцев по каждой из двух фамилий.

Подойдем к построению характеристики распространенности однофамильцев иначе. Выберем из всей популяции случайным образом некоторого индивидуума. У него окажется i-я фамилия с вероятностью p_i, очевидно равной ее концентрации в популяции. Независимо от результата первого наблюдения снова наугад выберем индивидуума. Вероятность, что у него та же фамилия, равна p_i. Значит вероятность $p_{iso,i}$ наблюдения пары однофамильцев с i-й фамилией (их случайной встречи) находится как p_i^2 независимо от монофилии фамилий (от условия, согласно которому у однофамильцев с конкретной фамилией один и тот же патрилинейный родоначальник). При суммировании по i находим вероятность Pr_{iso} для индивидуума случайно встретить в популяции своего однофамильца с учетом всех фамилий как

$$Pr_{iso} \equiv \sum_i p_{iso,i} = \sum_i p_i^2, \quad p_{iso,i} \equiv p_i^2.$$

Заметим, что если при наблюдении пар однофамильцев исключать пары типа индивидуум сам с собой, то вероятность выбрать второй раз индивидуума с i-й фамилией равна не p_i, а $(m_i - 1)/N = p_i - 1/N$. Величиной $1/N$ можно пренебречь, когда N не слишком мало. Учет поправки $1/N$ дает несмещенную оценку (см., например, [12]) вероятности встречи двух однофамильцев. Суть появления различий в том, используем ли мы выборку с возвращением (например, выборку фамилий из родительской популяции, когда отцом у каждого потомка при каждом выборе может быть любой из N мужчин, не имеющих в этом отношении преимуществ друг перед другом) или выборку без возвращения (например, при выборе следующего индивидуума из оставшихся в популяции).

Дополнительной к вероятности случайно наблюдать пару однофамильцев является вероятность случайной встречи индивидуумов с разными фамилиями, происходящих от заведомо разных родоначальников при патрилинейном наследовании фамилии и при отсутствии у них изменений ("мутаций"). Вероятность наблюдения индивидуума с i-й фамилией равна p_i, а вероятность его встречи с носителем другой фамилии $1 - p_i$, так что вероятность (Hs_i) наблюдения двух индивидуумов с i-й и иной фамилиями с учетом порядка равна $p_i(1 - p_i)$. Когда упорядоченность индивидуумов при встрече не имеет значения, такая вероятность равна $2p_i(1 - p_i)$ и складывается из $p_i(1 - p_i)$ и $(1 - p_i)p_i$, а при учете всех фамилий равна $\sum_i 2p_i(1 - p_i)$. Данную сумму можно интерпретировать как вероятность (Hs) встречи двух потомков разных родоначальников.

Согласно [11, 13] на малом по сравнению с размером популяции промежутке времени доминирующим фактором динамики является случайный дрейф. Именно на влиянии случайного дрейфа на указанном промежутке мы сосредоточимся далее. При этом будем использовать предпосылки стандартных широко использующихся популяционно-генетических моделей с неперекрывающимися поколениями. Так, наиболее известный вариант модели случайного генного дрейфа представлен моделью Райта–Фишера диплоидной популяции размера N со случайным скрещиванием и неперекрывающимися поколениями. Мы будем использовать модель с неперекрывающимися поколениями и дискретным временем для популяций человека, где присутствует перекрывание. Хотя очевидно, что такое применение противоречиво, но, например, при обработке данных по популяциям человека широко используется также выведенный для модели с неперекрывающимися поколениями закон Харди–Вайнберга. Его многократное употребление обычно подтверждает согласие закона с дан-

ными. Аналогично для анализа более сложных ситуаций в популяционной генетике используются как модели с дискретным, так и с непрерывным временем. Результаты их применения обычно не противоречат друг другу и могут рассматриваться как соответствующие аппроксимации.

При использовании далее модели со случайным скрещиванием Райта–Фишера генетическое состояние популяции в отношении аутосомного локуса с множественными аллелями описывается вектором концентраций аллелей, которые равноценны в том смысле, что у них нет преимуществ друг перед другом в передаче потомкам. В следующем поколении генетический состав популяции с численностью N представляет собой случайную выборку с возвращением размера $2N$ из пула гамет родительского поколения (по две гаметы на одного потомка).

При отсутствии давления систематических факторов, неслучайно изменяющего генетическую структуру, эта выборка извлекается из совокупности аллелей, в свою очередь полученных в предыдущем поколении как выборка аллелей из своего родительского поколения. Значит динамика состояния популяции представляет собой последовательность вложенных выборок, и в каждый момент времени нет систематического тренда в распространении одних и вымирании других аллелей.

Что касается модели динамики фамильной структуры, то фамилии равноценны в отношении передачи потомкам, предполагается их патрилинейное наследование и изменения (мутации) фамилий отсутствуют. После заключения брака жена принимает фамилию мужа. В результате фамильная структура женщин (матерей для следующего поколения) просто дублирует структуру мужчин (отцов) и не несет дополнительной информации. Более того, даже назначение женщинам любых фамилий не скажется на фамилиях следующего поколения, передающихся только от мужчин. На практике использование этой модели возможно и при отклонениях от данных предположений, но достаточно малых, чтобы ими было допустимо пренебречь. В итоге траектория состава фамилий в популяции в ряду поколений представляет собой последовательность составов вложенных выборок из фамилий мужчин соответствующих родительских поколений с объемами $N(\tau)/2$, и на нее не влияют фамилии женщин.

Однако отметим, что информация по *девичьим* фамилиям женщин может быть полезна, например, при использовании данных по изонимным бракам. Если дополнительно потребовать, чтобы фамилии комбинировались в брачных парах независимо, то вероятность изонимного брака совпадает с вероятностью случайной встречи двух однофамильцев.

ДИНАМИКА ФАМИЛЬНОГО СОСТОЯНИЯ ПОПУЛЯЦИИ

Состав фамилий в следующем поколении является случайной выборкой с возвращением из совокупности фамилий мужчин родительского поколения. Динамика фамильного состава популяции представляет собой последовательность вложенных выборок с возвращением из совокупности фамилий в соответствующих родительских поколениях. Каждая выборка определяет фамильный состав популяции на очередном шагу, и при равном соотношении полов ее размер совпадает с размером $N(\tau)/2$ мужской части популяции с общей численностью $N(\tau)$ в следующем поколении τ. Таким образом, динамика фамильной структуры формально совпадает с процессом генного дрейфа, синхронно идущего в той же самой популяции, но *при вчетверо меньших размерах выборок* ($N(\tau)/2$ для дрейфа фамилий и $2N(\tau)$ для дрейфа генов). При этом роль множественных аллелей локуса играют фамилии индивидуумов.

Уточним информацию о фамильном составе выборки. Он представляет собой результат $N(\tau)/2$ испытаний, а если рассматривать отдельную фамилию, то результат биномиальных испытаний. У нас “испытанию” соответствует рождение потомка (мужского пола, так как женщин не рассматриваем). У потомка будет i-я фамилия с вероятностью p_i, где p_i – ее концентрация в поколении родителей, или какая-либо другая (с вероятностью $1 - p_i$). По свойствам биномиальных испытаний с вероятностью успеха p_i у концентрации успехов x_i в выборке размера K будут следующие характеристики:

$$E\{x_i\} = E\{x_i|p_i\} = p_i,$$
$$V(x_i) \equiv E\left\{(x_i - E\{x_i\})^2\right\} = p_i(1 - p_i)/K. \quad (1)$$

Напомним, что $E\{.\}$ является символом операции получения математического ожидания случайной величины, заключенной в фигурные скобки (его можно понимать как среднее значение при неограниченном увеличении размера выборки), а $V(.)$ обозначает дисперсию случайной величины в круглых скобках.

Из (1) следует, что случайный дрейф как последовательность вложенных выборок не имеет тенденций в динамике концентраций “успехов”, оставляя их в среднем неизменными (т.е. у нас концентрации фамилий в среднем остаются равными начальным значениям). Однако при неизменности в среднем возможные концентрации фамилий как бы расплываются вокруг начальных значений. Поэтому обратимся к свойствам динамики фамильного дрейфа.

Результат 1. *Пусть динамика вектора* $\mathbf{x}(t)$ *концентраций фамилий* $\{x_i(t)\}$ *в популяции постоянного эффективного размера* N *рассматривается как по-*

следовательность вложенных случайных выборок с возвращением размера $N/2$ из фамилий родителей. На каждом шагу вероятности выбора фамилий равны их концентрациям среди родителей. Обозначим начальные концентрации фамилий через $\mathbf{p} = \{p_i\} = \{x_i(0)\}$ *и положим, что можно пренебречь их мутациями и миграциями.*

Тогда независимо от наличия или отсутствия монофилии фамилий динамика фамильного состава $\{x_i(t)\}$ *популяции характеризуется следующим образом*:

1. *В наугад выбранной популяции-реплике из умозрительной совокупности реплик (как бы повторяющих микроэволюцию рассматриваемой популяции в тех же условиях и с теоретически возможными фамильными состояниями) вероятность* $E\{Hs(\mathbf{x}(t))\}$ *случайной встречи двух потомков разных родоначальников и вероятность* $E\{Hs_i(\mathbf{x}(t))\}$, $i = 1, 2, \ldots$ *встречи потомков с i-й фамилией и иной убывают по поколениям* $t = 1, 2, \ldots$ *с одним и тем же темпом* $2/N$ *своей величины за шаг (поколение) независимо от предположения о монофилии фамилий. При этом*

$$E\{Hs_i(\mathbf{x}(t))\} \equiv E\{x_i(t)(1-x_i(t))\} = \\ = p_i(1-p_i)(1-2/N)^t \equiv Hs_i(\mathbf{p})(1-2/N)^t \underset{t\to\infty}{\to} 0,$$

$$E\{Hs(\mathbf{x}(t))\} \equiv E\left\{1-\sum_{i=1}^{k} x_i^2(t)\right\} = \\ = Hs(\mathbf{p})(1-2/N)^t \underset{t\to\infty}{\to} 0. \qquad (2)$$

2. *В пределе с течением времени популяция будет состоять только из однофамильцев и, более того, из потомков лишь одного родоначальника.*

Доказательство. **1.** Очевидно, потомки с разными фамилиями заведомо происходят от разных родоначальников независимо от монофилии фамилий. Сначала найдем значение $E\{x_i(t)(1-x_i(t))\}$ в первом поколении при отсутствии мутаций и миграций фамилий. Согласно (1) при $x_i(0) = p_i$ ожидаемое значение $E\{x_i(t)\}$ в первом поколении равно p_i и

$$E\{Hs_i(\mathbf{x}(1))\} \equiv E\{x_i(1)(1-x_i(1))\} = \\ = E\{x_i(1)\} - E\{x_i^2(1)\} = \\ = E\{x_i(1)\} - \left(Vs(x_i(1)) + (E\{x_i(1)\})^2\right) = \\ = p_i - Vs(x_i(1)) - p_i^2 = p_i(1-p_i) - Vs(x_i(1)),$$

так как для любой случайной величины x

$$V(x) \equiv E\{(x-E\{x\})^2\} = E\{x^2\} - (E\{x\})^2, \\ E\{x^2\} = (E\{x\})^2 + V(x). \qquad (3)$$

Подставим в полученное выражение для $E\{Hs_i(\mathbf{x}(1))\}$ значение дисперсии $Vs(x_i(1))$, соответствующее выборочному дрейфу фамилий, т.е. по (1) равное $p_i(1-p_i)\Big/\frac{N}{2}$:

$$E\{x_i(1)(1-x_i(1))\} = \\ = p_i(1-p_i)(1-V(x_i(1)))/(p_i(1-p_i)) = \\ = p_i(1-p_i)(1-2/N), \quad i = 1, 2, \ldots .$$

Здесь фамилия может быть произвольной, а ожидаемое значение $E\{Hs_i(x(1))\}$ вероятности $Hs_i(x(1))$ случайной встречи потомков с разными фамилиями уменьшилось за поколение (шаг) по сравнению с предыдущей величиной $Hs_i(p)$ в $1-2/N$ раз, какими бы ни были концентрация фамилий $x(0) \equiv p$ и сами фамилии. Поэтому еще за одно поколение при любом из полученных случайных значений концентраций фамилий $\{x_i(2)\}$ вероятность $E\{Hs_i(x(2))\}$ встречи двух потомков разных родоначальников на втором шагу также уменьшится в $1-2/N$ раз, откуда за два первых поколения уменьшение будет в $(1-2/N)^2$ раз.

Таким образом, за каждое поколение вероятность $Hs(\mathbf{x}(t))$ наблюдения пары потомков разных родоначальников с учетом всех фамилий уменьшается в $1-2/N$ раз, а через t шагов (поколений) в последовательности случайных выборок из соответствующих родительских популяций вплоть до родоначальной имеем

$$E\{Hs(\mathbf{x}(t))\} = Hs(\mathbf{p})(1-2/N)^t \to 0 \text{ при } t \to \infty.$$

Если эффективная численность изменяется и в поколении τ равна $Ne(\tau)$, то

$$E\{Hs(\mathbf{x}(t))\} = Hs(\mathbf{p})\prod_{\tau=1}^{t}(1-2/Ne(\tau)).$$

2. Когда среди конечного количества возможных вариантов реализации фамильного состава популяции среднее значение для неотрицательной вероятности $x_i(1-x_i)$ наблюдения двух потомков разных родоначальников равно нулю, то и на каждой реализации $x_i(1-x_i)$ будет равно нулю, что означает случайную утерю фамилий. В итоге вся популяция будет состоять из однофамильцев, в том числе и при росте размера популяции до некоторого предела, а также при отсутствии монофилии фамилий. В конечном итоге все индивидуумы будут потомками одного родоначальника. ◄

Следствие 2. *Динамика характеристик фамильной структуры популяции в рамках предположений предыдущего результата описывается следующими зависимостями от времени*:

$$Vs(x_i(t)) = p_i(1-p_i)(1-(1-2/N)^t) = \\ = Hs_i(\mathbf{p}) - E\{Hs_i(\mathbf{x}(t))\} \underset{t\to\infty}{\to} p_i(1-p_i) = Hs_i(\mathbf{p}),$$

$$E\left\{x_i^2(t)\right\} = p_i - p_i(1-p_i)(1-2/N)^t =$$
$$= p_i - E\{Hs_i(\mathbf{x}(t))\} \underset{t\to\infty}{\to} p_i, \quad i = 1, 2, \ldots k, \qquad (4)$$
$$Fs(\mathbf{x}(t)) = 1 - (1-2/N)^t \underset{t\to\infty}{\to} 1.$$

Здесь $Vs(x_i(t))$ – дисперсия концентрации x_i у i-й фамилии в теоретической совокупности популяций-реплик, $E\left\{x_i^2(t)\right\}$ – вероятность случайной встречи двух индивидуумов с i-й фамилией в популяции, случайно выбранной из теоретической совокупности, $Fs(\mathbf{x}(t))$ – фамильный аналог коэффициента инбридинга популяции в поколении t (в момент времени t).

Доказательство. 1. Для отыскания $Vs(x_i)$ воспользуемся выражением $Vs(x_i)$ через $E\{x_i(1-x_i)\}$ как при выводе (3) и (2):

$$Vs(x_i(t)) = p_i(1-p_i) - E\{x_i(t)(1-x_i(t))\} =$$
$$= p_i(1-p_i) - p_i(1-p_i)(1-2/N)^t =$$
$$= p_i(1-p_i)\left(1-(1-2/N)^t\right) \to$$
$$\to p_i(1-p_i) = Hs_i(\mathbf{p}) \text{ при } t \to \infty.$$

2. Вероятность случайной встречи однофамильцев с i-й фамилией при условии ее концентрации $x_i(t)$ в фиксированной популяции равна $x_i^2(t)$. При случайном выборе популяции-реплики вероятность такой встречи согласно (3) находится как $Vs(x_i) + (E\{x_i\})^2$, где подстановка $Vs(x_i) = p_i(1-p_i) - p_i(1-p_i)(1-2/N)^t$ из п. 1 дает

$$E\left\{x_i^2(t)\right\} = Vs(x_i) + (E\{x_i\})^2 =$$
$$= p_i - p_i(1-p_i)(1-2/N)^t = p_i - E\{Hs_i(\mathbf{x}(t))\} \underset{t\to\infty}{\to} p_i.$$

3. Из п. 1 следует, что

$$E\{x_i(1-x_i)\} = p_i(1-p_i) - Vs(x_i) =$$
$$= p_i(1-p_i)(1 - Vs(x_i)/(p_i(1-p_i))).$$

Правую часть данного выражения можно записать как $p_i(1-p_i)(1-Fs)$, где коэффициент $Fs \equiv Vs(x_i)/(p_i(1-p_i))$ является аналогом случайного коэффициента инбридинга популяции F (точнее F_{ST}) для данных по фамилиям. Поскольку $Fs(x_i(t)) \equiv Fs(t) = Vs(x_i(t))/(p_i(1-p_i))$, то подстановка сюда найденного выражения для $Vs(x_i)$ дает $Fs(t) = 1-(1-2/N)^t$ и $Fs(t) \to 1$ при $t \to \infty$. ◄

В более реальной ситуации с изменениями размера популяции вместо $(1-2/N)^t$ будет $\prod_{\tau=1}^{t}(1-2/Ne(\tau)) \to 0$ при $t \to \infty$, если у $Ne(\tau)$ существует конечный предел или $Ne(\tau)$ ограничено сверху константой. Отметим, что формулы (4) являются точными для модели случайного дрейфа, а далее займемся их аппроксимациями.

КОЭФФИЦИЕНТ ИНБРИДИНГА ПОПУЛЯЦИИ И СВОЙСТВА ДРЕЙФА ФАМИЛИЙ ПРИ МАЛОМ ЭФФЕКТИВНОМ ВРЕМЕНИ ДИВЕРГЕНЦИИ

Во многих исследованиях внимание фокусируется на микроэволюции, и желательно найти удобную и несложную аппроксимацию процессов динамики на небольших периодах эффективного времени $2t/N$. В отношении дрейфа фамилий это особенно актуально, так как подавляющее количество фамилий появилось в России сравнительно недавно после первой Всероссийской переписи населения в 1897 г., хотя они встречались еще в новгородских летописях. Отметим, что небольшой период должен быть таковым лишь по сравнению с размером популяции, достигающим сотен, тысяч и более. Так как мы время измеряем в поколениях, то для человека это уже тысячи и десятки тысяч лет.

Рассмотрим асимптотику дивергенции состояния одной популяции от начального значения в терминах Hs (или H) и Fs (F) на малом промежутке эффективного времени в поколениях. В стационарных условиях среды зависимость от времени t это зависимость от $\mathbf{x}(t)$, например $F(t) = F(\mathbf{x}(t))$. В популяционной генетике вероятности Hs соответствует концентрация H гетерозигот в случайно скрещивающейся популяции, и по формуле С. Райта

$$E\{H(\mathbf{x}(t))\} = H(\mathbf{p})(1-F(\mathbf{x}(t))),$$
$$\text{откуда } F(\mathbf{x}(t)) = \left(H(\mathbf{p}) - E\{H(\mathbf{x}(t))\}\right)/H(\mathbf{p}),$$

где $H(\mathbf{p})$ – начальное значение концентрации гетерозигот, $E\{H(\mathbf{x}(t))\}$ – ожидаемая концентрация гетерозигот в теоретической совокупности популяций-реплик с возможными генетическими состояниями $\mathbf{x}(t)$ при заданных условиях существования, $F(\mathbf{x}(t))$ – коэффициент инбридинга популяции (см., скажем, [12]).

Аналог коэффициента инбридинга F в случае фамилий мы обозначили ранее как Fs. Случайный генный дрейф отличается от фамильного дрейфа лишь тем, что для первого изменение состояния при смене поколений происходит за счет выборки размера $2N$ для аллелей, а для фамилий размер выборки $N/2$ вчетверо меньше. Значит, подстановка нужного значения $2N$ или $N/2$ в соответствующие формулы дает характеристики дрейфа генов или дрейфа фамилий. К сожалению, по данным обследования популяций, относящихся к одному моменту времени, входящие в эти формулы начальные значения $H(\mathbf{p}) = H(0)$ и $Hs(0)$ неизвестны.

Для дрейфа фамилий в случае постоянной по поколениям численности популяции N зависимость $Hs(\mathbf{x}(t))$ для фамилий согласно (2) довольно проста: $E\{Hs(\mathbf{x}(t))\} = Hs(\mathbf{p})(1-2/N)^t$, а для гетеро-

зигот $E\{H(\mathbf{x}(t))\} = H(\mathbf{p})(1 - 1/2N)^t$. На малом интервале эффективного времени эти зависимости еще более упрощаются:

$$(1 - 2/N)^t = e^{t\ln(1-2/N)} \sim e^{-2t/N},$$
$$(1 - 1/2N)^t \sim e^{-t/2N} \text{ при } 1/N \ll 1.$$

Отсюда следует, например, что в популяционной генетике приближенно

$$1 - F(t) = e^{-t/2N},$$
$$F(t) = 1 - e^{-t/2N} \sim t/2N, \quad t/2N \ll 1.$$

Уточним полученные выражения для ситуации неслучайного изменения численности N во времени и нарушения некоторых предпосылок модели случайного дрейфа. В таком случае размер популяции в каждом поколении τ следует заменить на эффективный размер $Ne(\tau)$) (см., например, [14, 15]), а при учете t поколений заменить N на средний гармонический эффективный размер популяции $\tilde{N}e(t)$ за весь период дивергенции t (см. нижеследующее замечание). При этом асимптотически на малом эффективном времени $t/2\tilde{N}e(t)$ зависимости рассматриваемых характеристик от $t/2\tilde{N}e(t)$ оказываются линейными.

Замечание 3. *Пусть в каждом поколении* $\tau = 1, 2, \ldots$ *фамильное состояние популяции с неперекрывающимися поколениями формируется как случайная выборка с возвращением размера* $\tilde{N}e(t)/2$ *из совокупности фамилий родителей, а вероятности выбора фамилий равны их концентрациям среди родителей.*

Тогда в t-м поколении при $\tilde{N}e(t)/t \to \infty$ $(t/\tilde{N}e(t) \to 0)$ *асимптотически*

$$E\{Hs(\mathbf{x}(t))\} \sim Hs(\mathbf{x}(0))(1 - 2t/\tilde{N}e(t)) = Hs(\mathbf{x}(0))(1 - Fs(\mathbf{x}(t))),$$

$$Fs(\mathbf{x}(t)) \sim 2t/\tilde{N}e(t), \quad \tilde{N}e(t) \equiv t\Big/\sum_{\tau=1}^{t}(1/Ne(\tau)),$$

$$F(\mathbf{x}(t)) \sim 1/4 \cdot Fs(\mathbf{x}(t)) \sim t/2\tilde{N}e(t). \qquad (5)$$

Здесь $\tilde{N}e(t)$ *– средний гармонический эффективный размер популяции за t поколений дивергенции, F – случайный коэффициент инбридинга.*

Доказательство. В случае фамилий величина $E\{Hs_i(\mathbf{x}(t))\} \equiv E\{x_i(t)(1 - x_i(t))\}$ уменьшается за поколение τ в $1 - 2/Ne(\tau)$ раз согласно ранее доказанному. Значит через 2 поколения $E\{Hs_i(\mathbf{x}(2))\} = Hs(\mathbf{x}(0))(1 - 2/Ne(1))(1 - 2/Ne(2))$, а через t поколений

$$E\{x_i(t)(1 - x_i(t))\} = p_i(1 - p_i)\prod_{\tau=1}^{t}(1 - 2/Ne(\tau)) =$$
$$= p_i(1 - p_i)e^{\sum_{\tau=1}^{t}\ln(1-2/Ne(\tau))} \sim p_i(1 - p_i)e^{-2\sum_{\tau=1}^{t}1/Ne(\tau))} \approx$$
$$\approx p_i(1 - p_i)e^{-2t/\tilde{N}e(t)} \sim p_i(1 - p_i)(1 - 2t/\tilde{N}e(t)),$$
$$t/\tilde{N}e(t) \ll 1,$$

поскольку $\ln(1 - 2/Ne(\tau)) \sim -2/Ne(\tau)$, $\sum_{\tau=1}^{t}1/Ne(\tau)) \approx t/\tilde{N}e(t)$ при малом значении $t/\tilde{N}e(t)$. В результате суммирования $E\{Hs_i(\mathbf{x}(t))\}$ по i при $t/\tilde{N}e(t) \ll 1$ получаем

$$E\{Hs(\mathbf{x}(t))\} \sim Hs(\mathbf{p})(1 - 2t/\tilde{N}e(t)) = Hs(\mathbf{p})(1 - Fs(t)),$$
$$Fs(t) \sim 2t/\tilde{N}e(t), \quad t/\tilde{N}e(t) \ll 1 \ (F,\ Fs \ll 1).$$

Очевидно, что при увеличении $\tilde{N}e(t)$ в 4 раза согласно размеру выборки гамет мы получим соответствующие результаты для генного дрейфа, и $F = 1/4 \cdot Fs = t/2\tilde{N}e(t)$. ◄

Таким образом, *коэффициент инбридинга F приблизительно в 4 раза меньше его фамильного аналога Fs*. Однако это простое соотношение становится неверным при большом количестве поколений t, отделяющих рассматриваемую популяцию от родоначальной (когда коэффициент инбридинга нельзя считать малым). Например, и тот и другой показатели в пределе с течением времени теоретически станут равны единице и не будут отличаться в 4 раза.

Выражение $2t/\tilde{N}e(t)$ показывает, что параметр $\tilde{N}e$ играет роль масштабирования времени. Например, возрастание его в 10 раз равносильно уменьшению времени в 10 раз и такому же уменьшению скорости его течения. Подобный результат верен для любой фамилии. Поэтому можно говорить, что *при относительно малом количестве поколений процесс дрейфа фамилий отличается от генного дрейфа тем, что для первого из них время течет приближенно в 4 раза быстрее.*

Ремарка 4. Выражение $2t/\tilde{N}e(t)$ для процесса дрейфа фамилий мы назвали *эффективным временем.* Мы видим, что оно входит в асимптотические формулы в качестве одного неразделимого параметра. Так как асимптотически $Fs(t) \sim 2t/\tilde{N}e(t)$ согласно (5), то *фамильный коэффициент инбридинга совпадает с эффективным временем*. Он монотонно увеличивается вместе с привычным временем в силу роста инбредности в последовательности поколений. Поэтому можно сказать, что приведенные *асимптотические формулы верны при малом коэффициенте инбридинга* (что типично для популяций человека).

Подчеркнем еще раз, что сказанное справедливо на относительно малом промежутке времени. Когда численность популяции изменяется по поколениям и допускаются некоторые отклонения от идеализированных предпосылок модели случайного дрейфа, то течение процесса дрейфа лучше характеризуется эффективным временем, в котором в качестве численности популяции N фигурирует ее средний гармонический эффективный размер $\tilde{N}e(t)$, учитывающий особенности репродукции и структуры популяции, а также динамику ее численности.

Следствие 5. *Пусть фамильное состояние популяции с неперекрывающимися поколениями формируется как случайная выборка с возвращением размера* $Ne(\tau)/2$ *из фамилий родителей в каждом поколении* $\tau = 1, 2, \ldots$, *и вероятности выбора фамилий равны их концентрациям среди родителей.*

Тогда асимптотически при $\tilde{N}e(t)/t \to \infty$ ($t/\tilde{N}e(t) \to 0$) *дисперсия концентраций фамилий* $Vs(x_i(t))$, $i = 1, 2, \ldots$ *линейно зависит от эффективного времени как*

$$Vs(x_i(t)) \sim p_i(1-p_i) \times 2t/\tilde{N}e(t) = Hs_i(\mathbf{x}(\mathbf{p})) \times 2t/\tilde{N}e(t) = Hs_i(\mathbf{x}(\mathbf{p})) \times Fs(t), \quad (6)$$

$$\tilde{N}e(t) \equiv t \Big/ \sum_{\tau=1}^{t} 1/Ne(\tau).$$

Здесь $Ne(\tau)$ *– дисперсионный* (*см., например,* [14]) *эффективный размер популяции в поколении* τ, $\tilde{N}e(t)$ *– средняя гармоническая эффективная численность популяции за t поколений дивергенции.*

Доказательство. Дисперсия $Vs(x_i(t))$ концентрации i-й фамилии характеризует дивергенцию популяций-реплик в теоретической совокупности мыслимых вариантов фамильного состава популяции. Согласно (4) и (5) она представима в виде

$$Vs(x_i(t)) = Hs_i(\mathbf{p}) - E\{Hs_i(\mathbf{x}(t))\} \sim Hs_i(\mathbf{p}) - Hs_i(\mathbf{x}(\mathbf{p}))(1 - 2t/\tilde{N}e(t)) = Hs_i(\mathbf{x}(\mathbf{p})) \times 2t/\tilde{N}e(t). \blacktriangleleft$$

Следствие 6. *Пусть в рамках предыдущего следствия размеры выборок фамилий при каждой смене поколений* τ *достаточно велики для аппроксимации распределений выборочных отклонений* $\{\delta_i(\tau)\}$ *концентрации i-й фамилии нормальными.*

Тогда в t-м поколении распределение этой концентрации $x_i(t)$ *приближенно будет нормальным со средним значением* p_i *и дисперсией* $p_i(1-p_i) \times 2t/\tilde{N}e(t)$.

Доказательство. Значение концентрации $x_i(t)$ у i-й фамилии в t-м поколении представляет собой при $\tilde{N}e(t)/t \to \infty$ сумму приближенно нормально распределенных некоррелирующих выборочных отклонений $\delta_i(\tau)$ с нулевыми средними плюс фиксированная начальная концентрация p_i. Поэтому аппроксимацией распределения $x_i(t)$ также будет нормальное распределение со средним значением $p_i = x_i(0)$ и найденной дисперсией $p_i(1 - p_i) \times 2t/\tilde{N}e(t)$. ◄

К сожалению, здесь по-прежнему фигурирует начальная концентрация p_i, которую нельзя найти по данным о текущем состоянии популяции (как, впрочем, и $\tilde{N}e(t)$).

ОПИСАНИЕ НЕЗАВИСИМОЙ ДИВЕРГЕНЦИИ ДВУХ ПОПУЛЯЦИЙ

До сих пор мы рассматривали дивергенцию фамильного состояния одной популяции от начального состояния. В практическом отношении, возможно, более интересно исследование дивергенции друг от друга нескольких популяций, имеющих общее происхождение. Примером могут служить популяции, состояния которых соответствуют одному и тому же моменту времени (часто моменту обследования). При общем происхождении дивергенция состояний таких популяций напоминает дивергенцию в теоретической совокупности популяций-реплик. Ее основные черты отражают начальный этап микроэволюции популяций, когда доминирующим фактором является случайный дрейф.

Рассмотрим какую-либо пару популяций, возникшую при разделении t поколений тому назад участка родословного древа на две ветви, существующие далее изолированно. Некоторые свойства фамильной дивергенции одной популяции от начального состояния мы уже рассмотрели. Фамильное состояние другой популяции аналогично претерпевает независимые изменения в результате процесса случайного дрейфа. В итоге наблюдается фамильная дивергенция популяций друг от друга. Независимо от степени дивергенции (ее длительности и размеров популяций) имеют место следующие соотношения для концентраций $x^{(1)}(t)$ и $x^{(2)}(t)$ какой-либо фамилии в первой и второй популяциях соответственно в поколении t при одинаковой у них начальной концентрации p:

$$E\{x^{(1)}(t)\} = E\{x^{(2)}(t)\} = p,$$
$$E\{x^{(1)}(t)\,x^{(2)}(t)\} = E\{x^{(1)}(t)\} \times E\{x^{(2)}(t)\} = p^2,$$
$$E\{x^{(1)}(t) - x^{(2)}(t)\} = E\{x^{(1)}(t)\} - E\{x^{(2)}(t)\} = 0.$$

Данные соотношения (кроме свойства произведения) верны для любого множества популяций с общим происхождением *без предположения о независимости дивергенции*. Таким образом, приведенные показатели не дают информации о течении процесса дивергенции (не зависят от времени). Эту зависимость характеризует, например, средний квадрат расстояния (дисперсия) между концентрациями фамилии $x^{(1)}(t)$ и $x^{(2)}(t)$,

который *при независимых изменениях* $x^{(1)}(t)$ и $x^{(2)}(t)$ равен сумме их дисперсий и согласно (5)–(6) при малом эффективном времени дивергенции находится как:

$$E\left\{\left(x^{(1)}(t)-x^{(2)}(t)\right)^2\right\}=Vs\left(x^{(1)}(t)\right)+Vs\left(x^{(2)}(t)\right)\sim$$
$$\sim p(1-p)\left(2t/\tilde{N}e_1(t)+2t/\tilde{N}e_2(t)\right)=$$
$$=p(1-p)\times 4t/\tilde{N}e(t)=p(1-p)\times$$
$$\times\left(Fs^{(1)}(t)+Fs^{(2)}(t)\right)=p(1-p)\times 2\bar{F}s(t).$$

Здесь $\tilde{N}e(t)=2/(1/\tilde{N}e_1(t)+1/\tilde{N}e_2(t))$ – среднее гармоническое значение для пары $\tilde{N}e_1(t)$ и $\tilde{N}e_2(t)$, $\bar{F}s(t)$ – среднее арифметическое значение коэффициентов фамильного инбридинга $Fs^{(1)}(t)$ и $Fs^{(2)}(t)$.

ОБСУЖДЕНИЕ

Отметим, что в приведенном анализе не принималась во внимание возможная инбредность общего предка в точке разделения ветвей, ведущих к рассматриваемой паре популяций. Поэтому, вообще говоря, в полученной выше формуле фигурируют не коэффициенты фамильного инбридинга популяций, а скорее их приращения за период t поколений от момента разветвления.

Дивергенция по отдельной фамилии между двумя популяциями на родословном древе (средний квадрат расстояния между концентрациями фамилий, т.е. дисперсия возможных различий концентраций) пропорциональна времени дивергенции или среднему арифметическому значению коэффициентов фамильного инбридинга популяций. С другой стороны, дивергенция обратно пропорциональна средней гармонической численности $\tilde{N}e(t)=2/(1/\tilde{N}e_1(t)+1/\tilde{N}e_2(t))$ популяций. Полученные зависимости подтверждают интуитивно ожидаемый качественный характер связей (их направления), но даны количественно.

К сожалению, здесь фигурирует обычно неизвестное начальное состояние p у концентраций фамилий, влияния которого желательно избежать. В точках ветвления, играющих роль ближайшего общего предка соответствующих пар, концентрации фамилии также отличаются от значений в корне родословного древа. Поэтому коэффициенты $p(1-p)$ в формуле для ожидаемого квадрата расстояния принимают случайные значения, свои для каждого узла ветвления, и у конкретного родословного древа их трудно учесть.

Коснемся проблем, возникающих при анализе данных, получаемых в результате обследования генетической структуры реальных популяций. В реальной подразделенной популяции, рассматриваемой как теоретическая совокупность, вместо начального состояния (математического ожидания) обычно используют среднее значение для распределения концентраций аллелей по субпопуляциям. При этом получают оценку коэффициента инбридинга в качестве *статистической корреляции* гомологичных генов объединяющихся гамет как в случае субпопуляций с общим происхождением, так и в случае произвольной группы субпопуляций, у которых нет идентичных по происхождению генов. Найденная *оценка не позволяет выявить, какой из этих случаев имеет место*, но статистическая корреляция сама по себе имеет важное значение. Она влияет на концентраций генотипов и тем самым результаты естественного и искусственного отбора, уровень наследственной отягощенности и др.

Обратим внимание на специфические черты процесса случайного выборочного дрейфа. Ожидаемым значением концентрации фамилии (аллеля) в новом поколении будет прежнее значение, т.е. *у случайного дрейфа нет преимущественного направления*. Поэтому величину дивергенции за поколение (“скорость” ненаправленной эволюции) можно измерять, скажем, дисперсией, а не просто средним отклонением, которое при ненаправленной эволюции равно нулю. При одинаковых прочих условиях выборочная дисперсия $x(t)$ при смене поколений, как характеристика скорости ненаправленной дивергенции, обратно пропорциональна размеру популяции, а также определяется текущим значением $x(t)$. Тем самым темп дивергенции для концентрации аллеля зависит от значения $x(t)$, что затрудняет интерпретацию величины наблюдаемых различий между популяциями с общим происхождением. Поэтому желательно использовать подходы, стабилизирующие темп дивергенции.

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

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

СПИСОК ЛИТЕРАТУРЫ

1. *Бужилова А.П.* География русских фамилий // Восточные славяне. Антропология и этническая история. М.: Научный мир, 1999. С. 135–152.
2. *Балановская Е.В., Романов А.Г., Балановский О.П.* Однофамильцы или родственники. Подходы к изучению связи между гаплогруппами Y-хромосомы и фамилиями // Мол. биология. 2011. Т. 45. № 3. С. 473–485.
3. *Crow J.F., Mange A.P.* Measurement of inbreeding from the frequency of marriages between persons of the same surname // Soc. Biol. 1982. V. 29. № 1/2. P. 101–105.
4. *Lasker W.G.* Surnames and Genetic Structure. Cambridge: Cambr. Univ. Press, 2005. 148 p.
5. *Ревазов А.А., Парадеева Г.М., Русакова Г.И.* Пригодность русских фамилий в качестве квазигене-

тического маркера // Генетика. 1986. Т. 22. № 4. С. 699–703.
6. *Tarskaia L., El'chinova G., Scapoli C. et al.* Surnames in Siberia: A study of the population of Yakutia through isonymy // Am. J. Phys. Anthropol. 2009. V. 138. P. 190–198. https://doi.org/10.1002/ajpa.20918
7. *Сорокина И.Н., Чурносов М.И., Балтуцкая И.В., и др.* Антропогенетическое изучение населения центральной России. М.: Издательство РАМН, 2014. 336 с.
8. *Lasker G.W.* A coefficient of relationship by isonymy: A method fore stimating the genetic relationship between populations // Hum. Biol. 1977. V. 49. № 3. P. 489–493.
9. *Rogers A.R.* Doubts about isonymy // Hum. Biol. 1991. V. 63. № 5. P. 663–668.
10. *Сорокина И.Н., Рудых Н.А., Крикун Е.Н., Сокорев С.Н.* Применение фамилий в популяционно-генетических исследованиях (на примере зарубежных популяций) // Науч. ведомости БелГУ. Сер. Медицина. Фармация. 2016. № 19(240). Вып. 35. С. 5–10.
11. *Пасеков В.П.* К анализу случайных процессов изонимии. I. Структура изонимии // Генетика. 2021. Т. 57. № 10. С. 1194–1204.
12. *Вейр Б.* Анализ генетических данных: дискретные генетические признаки. М.: Мир, 1995. 400 с.
13. *Малютов М.Б., Пасеков В.П.* Об одной статистической задаче популяционной генетики // Теория вероятностей и ее применения. 1971. Т. 16. Вып. 3. С. 579–581.
14. *Хедрик Ф.* Генетика популяций. М.: Техносфера, 2003. 592 с.
15. *Ли Ч.* Введение в популяционную генетику. М.: Мир, 1978. 555 с.

ПРИЛОЖЕНИЕ

СВОДКА РЕЗУЛЬТАТОВ ДЛЯ ПОПУЛЯЦИОННО-ГЕНЕТИЧЕСКОГО АНАЛИЗА

Чтобы было удобно видеть сходство и различия основных черт фамильной и генетической картин состояния популяции, просто перечислим в одном месте разбросанные по настоящей статье (иногда неявные) результаты для популяционно-генетической модели элементарной диплоидной популяции с неперекрывающимися поколениями, ограниченной численностью и случайным скрещиванием. Пусть анализируется генетический состав диплоидной популяции по одному аутосомному локусу с k аллелями. Тогда генетическое состояние (состав) популяции можно описывать вектором $\mathbf{x} = \{x_i\}$ концентраций аллелей. Динамика состояния в общем случае обусловлена систематическими давлениями и случайным дрейфом генов. *При малом количестве поколений микроэволюции t по сравнению со средней гармонической численностью $\tilde{N}e(t)$ популяции для этого периода времени мы можем пренебречь давлением систематических факторов на генетические состояния популяции и решающим будет влияние случайного генного дрейфа.* В результате дрейфа происходит дивергенция как состояния популяции от первичного, так и популяций с общим происхождением друг от друга. Данный этап является начальным для микроэволюции исходно одинаковых родственных популяций.

Описание дивергенции в статике

В фиксированный момент времени картину дивергенции можно охарактеризовать следующими показателями. Пусть p_i начальная концентрация i-го аллеля в популяции со свободным скрещиванием, подверженной случайному генному дрейфу. Тогда через некоторый промежуток времени ожидаемая концентрация гетерозигот $E\{H_i(\mathbf{x}(t))\}$ с i-м аллелем в популяции, *случайно выбранной из умозрительной совокупности популяций-реплик* (как бы повторивших микроэволюцию рассматриваемой) с теоретически возможными состояниями, находится как

$$E\{H_i(\mathbf{x}(t))\} \equiv E\{(x_i(t)(1-x_i(t))\} = \\ = p_i(1-p_i) - V(x_i(t)) = p_i(1-p_i)(1-F(t)).$$

Здесь $E\{.\}$ символ математического ожидания (операции получения среднего значения) величины в фигурных скобках, $V(.)$ дисперсия случайной величины, стоящей в скобках (у нас дисперсия концентрации в репликах), F – коэффициент инбридинга популяции, $p_i = x_i(0)$. В то же время при условии, что концентрация у i-го аллеля в конкретной популяции этой совокупности равна x_i, доля рассматриваемых гетерозигот в ней с учетом порядка аллелей по закону Харди–Вайнберга будет теоретически равна $x_i(1-x_i)$.

Ожидаемые концентрации гетерозигот всех типов $H(\mathbf{x})$ и гомозигот в наугад выбранной популяции из умозрительной совокупности находятся как

$$E\{H(\mathbf{x})\} \equiv E\left\{1 - \sum_{i=1}^{k} x_i^2\right\} = H(\mathbf{p})(1-F),$$

$$F = \sum_{i=1}^{k} V(x_i)/H(\mathbf{p}).$$

$$E\{x_i^2\} = p_i^2 + Fp_i(1-p_i) = \\ = p_i^2 + V(x_i) = p_i^2 + FH_i(\mathbf{p}), \quad i = 1, 2, \ldots, k,$$

$$E\left\{\sum_{i=1}^{k} x_i^2\right\} = \sum_{i=1}^{k} p_i^2 + V(\mathbf{x}) = \sum_{i=1}^{k} p_i^2 + FH(\mathbf{p}) =$$
$$= 1 - H(\mathbf{p}) + FH(\mathbf{p}), \quad V(\mathbf{x}) \equiv \sum_{i=1}^{k} V(x_i).$$

Здесь $V(x_i)$ – дисперсия распределения значений $\{x_i\}$, $i = 1, 2, \ldots k$ концентраций i-го аллеля в популяциях-репликах. Коэффициент инбридинга популяции F выражается через дисперсию $V(x_i)$ концентрации отдельного аллеля x_i, а также с учетом концентраций всех аллелей в теоретической совокупности популяций-реплик как

$$F = V(x_i)/(p_i(1-p_i)), \quad i = 1, 2, \ldots, k,$$
$$F = (H_i(\mathbf{p}) - E\{H_i(\mathbf{x})\})/H(\mathbf{p}),$$
$$F = V(\mathbf{x})\Big/\left(1 - \sum_{i=1}^{k} p_i^2\right) =$$
$$= V(\mathbf{x})/H(\mathbf{p}), \quad V(\mathbf{x}) \equiv \sum_{i=1}^{k} V(x_i).$$

Таким же образом дивергенцию реальных субпопуляций в подразделенной популяции можно описать как в терминах дисперсии распределения концентрации отдельных аллелей, так и в терминах случайного коэффициента инбридинга F.

Приведенные формулы характеризуют инбредную популяцию. *Формально она является как бы подразделенной* – состоит либо из умозрительных популяций-реплик, либо из реальных субпопуляций. В последнем случае приведенные формулы в равной степени приложимы как к совокупности популяций с общим происхождением, так и к произвольному механическому набору популяций со случайным скрещиванием. При этом подразумевается, что математические ожидания и дисперсии концентраций относятся к рассматриваемой подразделенной популяции, играющей роль теоретической совокупности. Для реальной группы субпопуляций математические ожидания и дисперсии вычисляются стандартным образом как средние значения и средние квадраты отклонений от этих средних. В результате в подразделенной популяции как едином целом, несмотря на выполнение в отдельных субпопуляциях соотношений Харди–Вайнберга, эти соотношения нарушаются (эффект Валунда). В ней будет наблюдаться дефицит гетерозигот по сравнению с ожидаемым при всеобщей панмиксии. Дефициту, вызванному различиями (дисперсией) концентраций аллелей между субпопуляциями, соответствует некоторое значение случайного коэффициента инбридинга F (точнее, F_{ST}).

Как для механической смеси популяций, так и для группы популяций с общим происхождением этот коэффициент инбридинга численно равен статистической корреляции между гомологичными генами объединяющихся в генотипах гамет. Получение ее оценки *не позволяет определить с какой из ситуаций сталкивается исследователь*. Лишь в случае общего происхождения популяций при их независимой и одинаковой микроэволюционной истории статистическая корреляция прямо связана с вероятностью идентичности по происхождению пары аллелей аутосомного локуса диплоидного генотипа. В общем случае родственные популяции не являются независимыми, и процедура оценивания коэффициента инбридинга как идентичности по происхождению гомологичных аллелей должна учитывать характер происхождения популяций и условия их микроэволюции.

Отметим, что статистическая корреляция между гомологичными генами объединяющихся гамет важна, так как с ее помощью определяются концентрации генотипов и тем самым, например, изменение генетической структуры в результате отбора.

Динамика основных характеристик дивергенции

Напомним, что t – время в поколениях, $\tilde{N}e(t) \equiv t\Big/\sum_{\tau=1}^{t}(1/Ne(\tau))$ – средний гармонический эффективный размер популяции за t поколений, $Ne(\tau)$ – эффективный размер популяции в поколении τ. При достаточно малой величине *эффективного времени* $t/\tilde{N}e(t)$ модель микроэволюции аппроксимируется процессом случайного генного дрейфа. Данный процесс приводит к случайному отклонению состояния популяции от начального, росту ее коэффициента инбридинга и увеличению дивергенции возможных состояний (дивергенции популяций с общим происхождением). При этом выполняются следующие зависимости от времени основных характеристик дивергенции (точные в рамках модели)

$$E\{H_i(\mathbf{x}(t))\} \equiv E\{x_i(t)(1-x_i(t))\} =$$
$$= p_i(1-p_i)\prod_{\tau=1}^{t}(1 - 1/2Ne(\tau)) \underset{t\to\infty}{\to} 0, \quad i = 1, 2, \ldots k,$$

$$E\{H(\mathbf{x}(t))\} \equiv E\left\{1 - \sum_{i=1}^{k} x_i^2(t)\right\} =$$
$$= H(\mathbf{p})\prod_{\tau=1}^{t}(1 - 1/2\,Ne(\tau)) \underset{t\to\infty}{\to} 0.$$

$$E\left\{x_i^2(t)\right\} = p_i - p_i(1-p_i)\prod_{\tau=1}^{t}(1-1/2\,Ne(\tau)) \underset{t\to\infty}{\to} p_i,$$

$$V(x_i(t)) = p_i(1-p_i)\times$$

$$\times\left(1-\prod_{\tau=1}^{t}(1-1/2\,Ne(\tau))\right)\underset{t\to\infty}{\to} p_i(1-p_i),$$

$$i = 1, 2, \ldots k,$$

$$V(\mathbf{x}(t)) \equiv \sum_{i=1}^{k} V(x_i(t)) =$$

$$= H(\mathbf{p})\left(1-\prod_{\tau=1}^{t}(1-1/2\,Ne(\tau))\right)\underset{t\to\infty}{\to} H(\mathbf{p}),$$

$$Fs(\mathbf{x}(t)) = 1-\prod_{\tau=1}^{t}\left(1-1/2\,Ne(\tau)\right)\underset{t\to\infty}{\to} 1.$$

Здесь пределы существуют при постоянной (N) и растущей $Ne(\tau)$ численности популяции (когда, например, она ограничена сверху константой). Формулы будут проще (и более привычны) при постоянной численности N. Тогда $\prod_{\tau=1}^{t}(1-1/2Ne(\tau)) = (1-1/2N)^t$.

Аппроксимация результатов при малом эффективном времени дивергенции

На относительно малом этапе дивергенции (при малом $t/\tilde{N}e(t)$) данные зависимости еще упрощаются до линейных по эффективному времени $t/2\tilde{N}e(t)$:

$$E\{H(\mathbf{x}(t))\} \sim H(\mathbf{p})\left(1-t/2\,\tilde{N}e(t)\right) =$$

$$= H(\mathbf{p})(1-F(\mathbf{x}(t))),\quad F(\mathbf{x}(t)) \sim t/2\,\tilde{N}e(t),$$

$$V(x_i(t)) \sim p_i(1-p_i)\times t/2\,\tilde{N}e(t),$$

$$V(\mathbf{x}(t)) \sim \left(1-\sum_{i=1}^{k} p_i^2\right)\times t/2\,\tilde{N}e(t) =$$

$$= H(\mathbf{p})\times t/2\,\tilde{N}e(t) = H(\mathbf{p})F(\mathbf{x}(t)).$$

Когда численность популяции постоянна, скажем, равна N, то $\tilde{N}e(t)$ заменяется на N.

Так как асимптотически $F(t) \sim t/2\tilde{N}e(t)$, то коэффициент инбридинга совпадает с эффективным временем и монотонно увеличивается вместе с временем t в поколениях. Поэтому можно сказать, что *эффективное время измеряется величиной коэффициента инбридинга и наоборот, т.е. данная асимптотика верна при малых коэффициентах инбридинга*, присущих популяциям человека.

Случайный процесс выборочного дрейфа остается таковым в отношении произвольных подгрупп аллелей, в частности, когда одна группа состоит из единственного аллеля, скажем, с концентрацией x, а в другой все остальные. Концентрация последней группы равна $1-x$, ее можно отбросить как зависимую переменную и сфокусироваться на изучении частного случая динамики концентраций аллелей по отдельности. При этом возможно выразить коэффициент инбридинга $F(t)$ через ожидаемую концентрацию $E\{x(t)(1-x(t))\}$ гетерозигот (с учетом порядка аллелей) в случайно выбранной популяции из теоретической совокупности популяций-реплик с одной и той же демографической историей:

$$F(t) = \left(p(1-p) - E\{x(t)(1-x(t))\}\right)/p(1-p) =$$

$$= \left(H(p) - E\{H(x(t))\}\right)/H(p),\quad p = x(0).$$

Здесь $E\{x(t)(1-x(t))\}$ означает усреднение $x(t)(1-x(t))$ по возможным значениям $x(t)$ в популяциях-репликах теоретической совокупности. Отметим, что последнее выражение для $F(t)$ верно и при учете всех аллелей.

Подчеркнем, что приведенная формула *не позволяет оценить инбридинг по типичным данным только о текущих концентрациях гетерозигот* $x(t)(1-x(t))$ в изучаемой популяции. Она относится к *ожидаемой* концентрации гетерозигот в наугад выбранной популяции из *теоретического множества популяций-реплик*. Возможные генетические составы реплик случайно различаются – при повторении в одних и тех же условиях микроэволюционной истории изучаемой популяции произойдет дивергенция реплик друг от друга в силу выборочной природы случайного дрейфа с присущей ей выборочными ошибками.

Кроме того, найденное выражение для $F(t)$ *зависит от неизвестного начального состояния p обследуемой популяции* (и ее умозрительных реплик), которое нельзя найти по текущему состоянию.

Случайный коэффициент инбридинга можно также представить как

$$F(t) = V(x(t))/p(1-p) = V(x(t))/H(x(p)).$$

Здесь V межпопуляционная дисперсия концентраций аллелей *в теоретическом множестве реплик* (дисперсия возможных результатов реализаций процесса дрейфа в одних и тех же условиях с одинаковыми начальными состояниями). Данный подход также требует сведений о неизвестном начальном значении $x(0) = p$.

To Analysis of Random Processes of Isonymy. II. Dynamics of Population Divergence

V. P. Passekov*

Dorodnitsyn Computing Center of the Federal Research Center "Informatics and Management" of the Russian Academy of Sciences, Moscow, 119991 Russia

**e-mail: pass40@mail.ru*

Random dynamics of the surname composition of a population of finite size with non-overlapping generations is considered in discrete time. Inheritance of surnames is assumed to be patrilineal. Dynamics is analyzed on a small $t/NE(t)$ effective time interval, where the $NE(t)$ is an average harmonic effective population size over t generations. Since here systematic pressures can be neglected, as a result, surname microevolution is approximated by the process of random gene drift, synchronously proceeding in the same population with four times less intensity than for surnames. Like the genetic drift model, the surname composition of the next generation is a random sample with replacement of the size $Ne(\tau)/2$ from the surnames of the male component of the parent population, that is, four times smaller than the sample of $2Ne(\tau)$ gametes under the genetic drift ($Ne(\tau)$ effective population size in generation τ). Dynamics of probability of random encounter of individual with namesake and the behavior of encounter probability for individuals with different surnames are studied. These probabilities are similar to homozygotes and heterozygotes, respectively, in the analysis of genetic structure. Exact time dependencies for those probabilities, for variances of concentrations of surnames and the surname analogue of inbreeding coefficient are given. An approximation of the exact dependencies by simpler ones is given over a small effective time $t/NE(t)$, where the surname divergence is four times faster than the genetic divergence. The results do not suggest surname monophilia and describe a speculative theoretical set of replica populations, as if having repeated the microevolutionary history of the population in question under the same conditions. The use of a small time compared to the size of the population is justified by the recent emergence of most surnames in Russia and the fact that the past time in generations is much smaller than the typical size of populations. In real subdivided populations, procedures for estimating the inbreeding coefficient using surname concentrations do not allow us to distinguish between situations of a mechanical mixture of subpopulations or their common origin.

Keywords: theoretical population genetics, populations with non-overlapping generations, inbreeding coefficient, random surname drift, random genetic drift, mathematical models, dynamics of surname composition divergence, asymptotics over small effective time.

ГЕНЕТИКА, 2021, том 57, № 11, с. 1330–1332

КРАТКИЕ СООБЩЕНИЯ

УДК 575.1:591:599.323.4

КОНКОРДАНТНОСТЬ КОЛИЧЕСТВЕННЫХ ПРИЗНАКОВ НЕСАХАРНОГО ДИАБЕТА И ВЫЖИВАЕМОСТИ ЕДИНОУТРОБНЫХ ПОТОМКОВ У ГИБРИДОВ КРЫС ЛИНИЙ *Brattleboro* И *WAG*

© 2021 г. И. И. Хегай*

Федеральный исследовательский центр, Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, 630090 Россия

**e-mail: khegay@bionet.nsc.ru*

Поступила в редакцию 29.12.2020 г.
После доработки 09.02.2021 г.
Принята к публикации 11.02.2021 г.

У крыс линии *Brattleboro* зафиксирована мутация *di* (*diabetes insipidus*), детерминирующая развитие несахарного диабета с высоким уровнем потребления воды. Нами исследован эффект переноса мутантного гена *di* в линию *WAG* под действие гена *c* (*альбино*). Гены *di* и *альбино* локализованы в различных хромосомах и наследуются как рецессивные аллели. Исходную родительскую пару составили самки линии *WAG* генотипа *c* $di^+/c\ di^+$ и самцы линии *Brattleboro* генотипа *C di*/*C di*. В F_2 дигибридного скрещивания выщепились четыре фенотипических класса. Исследование у самок F_2 индивидуального потребления воды и размеров пометов крысят к моменту отъятия от грудного вскармливания выявило высокий уровень конкордантности с коэффициентом корреляции 0.996.

Ключевые слова: вазопрессин, несахарный диабет, потребление воды, *не-агути*, *альбино*, размер помета.

DOI: 10.31857/S0016675821110047

Крысы линии *Brattleboro* являются носителями точечной делеции в гене вазопрессина. Мутация блокирует трансляцию белкового препрогормона вследствие исчезновения в нуклеотидной последовательности стоп-кодона [1]. Физиологически мутация проявляется в форме гипоталамического несахарного диабета с высоким уровнем потребления воды. Данный признак хорошо воспроизводится и используется для идентификации мутантного генотипа [2]. Мутантный ген вазопрессина наследуется как рецессивный аллель и обозначается символом *di* (*diabetes insipidus*). Крысы *Brattleboro* представляют единственную экспериментальную линию крыс, моделирующую наследственный несахарный диабет. В литературе подробно описаны особенности водно-электролитного обмена у крыс *Brattleboro* [3, 4]. В то же время, вопрос об экспрессии мутантного гена вазопрессина в ином генетическом окружении в значительной степени остается открытым. Нами было предпринято и осуществлено селекционно-генетическое исследование по переносу и закреплению мутации на генетический фон крыс линии *WAG* (*Wistar Albino Glaxo*).

Исходные крысы линии *Brattleboro* характеризуются черным окрасом шерсти, обусловленным действием гена *не-агути*. У крыс линии *WAG* действие гена *не-агути* маскируется мутантным геном *альбино*. Крысы линии *WAG* имеют белый окрас и одновременно являются носителями нормальных аллелей гена вазопрессина. Ген *альбино*, как и мутантный ген *di*, наследуется в скрещиваниях как рецессивный аллель. Нами было выполнено классическое дигибридное скрещивание относительно пары локусов *di* и *альбино*. Локус *альбино* зафиксирован в 1-ой хромосоме крысы [5]. Действие мутации связано с нарушением работы гена тирозиназы [6]. Нормальная форма гена тирозиназы доминирует и обозначается в геноме крыс большой литерой *C*, соответственно, мутантный аллель обозначают как *c* малое. Ген *di* нами локализован в хромосоме 3 [7]. Нормальный аллель гена вазопрессина принято обозначать символом di^+ [8].

Родительские пары были составлены из самок линии *WAG* генотипа *c* $di^+/c\ di^+$ и самцов линии *Brattleboro* генотипа *C di*/*C di*. На протяжении всего эксперимента у крыс измеряли индивидуальный уровень потребления воды. Все особи F_1 имели единообразный гетерозиготный генотип *c* $di^+/C\ di$ и фенотип *не-агути* с уровнем потребления воды, характерным для носителей нормального гена вазопрессина. В поколении F_2 произошло ожидаемое расщепление по двум парам признаков – *не-агути*/*альбино* и *диабет*/*норма*. В случае диабета

Таблица 1. Расщепление потомков F_2 по фенотипическим классам

Фенотип	*не-агути/норма*	*не-агути/диабет*	*альбино/норма*	*альбино/диабет*
Теоретическое расщепление	74.25	24.75	24.75	8.25
Опытное расщепление	76	23	26	7
χ^2	0.41	$p > 0.05$		

суточное потребление воды превышало 30% от веса тела, норма характеризовалась на порядок более низким уровнем потребления воды. Всего было получено 132 гибридных особи, из них 30 крыс с диабетом и 33 крысы с фенотипом *альбино*. Распределение гибридов по фенотипическим классам и сравнение с теоретически ожидаемым представлено в табл. 1, $\chi^2 = 0.416$ и соответствует условиям недостоверности различий между теоретическим и опытным распределением. В дальнейших скрещиваниях участвовали самки всех выщепившихся фенотипов и самцы исключительно гомозиготного генотипа *di/di* с целью закрепления мутантного генотипа. В результате, в следующем поколении было получено и исследовано 35 пометов у самок фенотипа *не-агути/норма*, 13 пометов у самок *не-агути/диабет*, 14 пометов у самок *альбино/норма* и пять пометов у самок *альбино/диабет*, выборка общей численностью 438 крысят. В данном исследовании у крыс различного фенотипа сравнивались размеры пометов на момент отъятия крысят от вскармливания молоком в возрасте одного месяца. Количественные данные по среднесуточному потреблению воды и средним размерам пометов представлены на рис. 1. В дополнение к ожидаемым достоверным различиям между диабетом и нормой по потреблению воды у самок обнаружены такие же достоверные различия по размерам пометов. Если самки с нормальным уровнем потребления воды вскармливали в среднем восемь крысят, то у самок с несахарным диабетом этот показатель равен трем. Коэффициент корреляции между количеством потребляемой воды и количеством крысят в потомстве составляет 0.996.

Таким образом, существует высокий уровень конкордантности между признаками несахарного диабета и наличием ограниченного количества потомков у крыс. При сравнении крыс *альбино* и нормально окрашенных крыс такой конкордантности не обнаружено. Результат можно интерпретировать как пониженную жизнеспособность

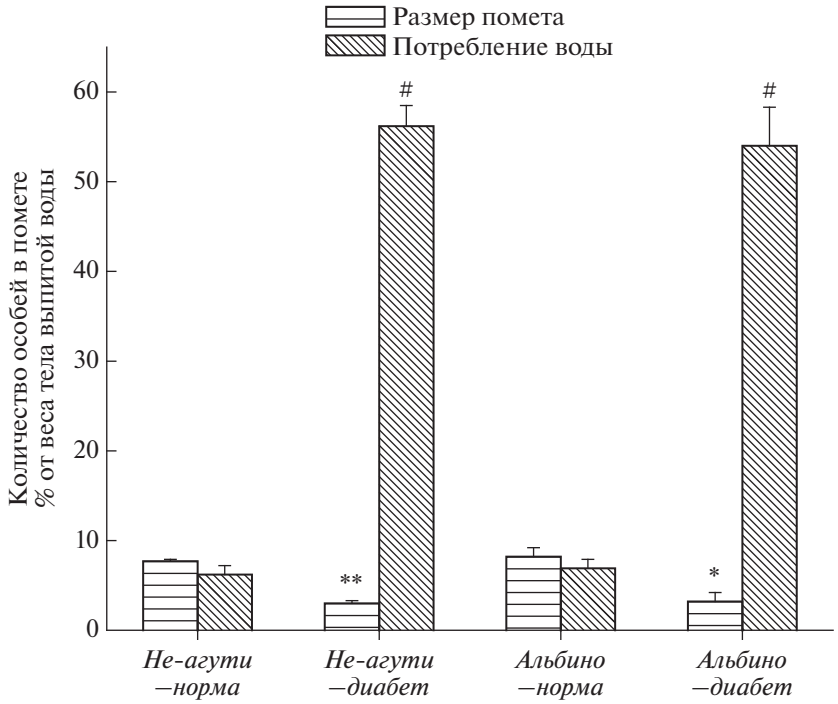

Рис. 1. Размер пометов на момент прекращения вскармливания молоком и среднесуточное потребление воды у различных фенотипических классов самок. Достоверность отличия от фенотипического класса *не-агути—норма*: по размеру помета: * – $p < 0.01$, ** – $p < 0.001$; по уровню потребления воды: # – $p < 0.001$.

крыс генотипа *di*/*di*, с высокой вероятностью, сохраняющейся при переносе мутантного гена под действие гена *альбино*. Физиологический механизм снижения фертильности не совсем очевиден. Возможно это является следствием гомологии вазопрессина и окситоцина, способных оказывать взаимные регуляторные эффекты на пролактин и уровень лактации [9, 10].

Работа выполнена при поддержке бюджетного проекта ФИЦ ИЦиГ СО РАН № 0259-2021-0014.

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

СПИСОК ЛИТЕРАТУРЫ

1. *Schmale H., Richter D.* Single base deletion in the vasopressin gene is the cause of *diabetes insipidus* in *Brattleboro* rats // Nature. 1984. V. 308. № 5961. P. 705–709. https://doi.org/10.1038/308705a0
2. *Хегай И.И.* Фенотипическое проявление мутантного гена *diabetes insipidus* у крыс и критерии генотипирования по фенотипу // Генетика. 2003. Т. 39. № 1. С. 70–74.
3. *Хегай И.И.* Взаимодействие аллелей локуса *diabetes insipidus* и гена почечного белка 120 кДа в онтогенезе крысы // Генетика. 2003. Т. 39. № 3. С. 397–401.
4. *Katkova L.E., Baturina G.S., Solenov E.I., Ivanova L.N.* Regulation of V2 receptor by vasopressin is different in *Brattleboro* rats compared with *Wistar* // Dokl. Biochem. Biophys. 2015. V. 463. P. 247–250. https://doi.org/10.1134/S1607672915040158
5. *Yamada J., Kuramoto T., Serikawa T.* A rat genetic linkage map and comparative maps for mouse or human homologous rat genes // Mammalian Genome. 1994. V. 5. № 2. P. 63–83.
6. *Kuramoto T., Nakanishi S., Ochiai M. et al.* Origins of albino and hooded rats: Implications from molecular genetic analysis across modern laboratory rat strains // PLoS One. 2012. V. 7. № 8. P. e43059. https://doi.org/10.1371/journal.pone.0043059
7. *Khegay I.I.,* Localization of the vasopressin gene to rat chromosome 3 // Mammalian Genome. 1996. V. 7. № 11. P. 867. https://doi.org/10.1007/s003359900300
8. *Valtin H., North W.G., Edwards B.R., Gellai M.* Animal models of diabetes insipidus // Front. Horm. Res. 1985. V. 13. P. 105–126.
9. *Uvnäs Moberg K., Prime D.K.* Oxytocin effects in mothers and infants during breastfeeding // Infant. 2013. V. 9. № 6. P. 201–206.
10. *Erickson E.N., Sue Carter C., Emeis C.L.* Oxytocin, vasopressin and prolactin in new breastfeeding mothers: Relationship to clinical characteristics and infant weight loss // J. Hum. Lact. 2020. V. 36. № 1. P. 136–145. https://doi.org/10.1177/0890334419838225

Concordance of Quantitative Traits of Diabetes Insipidus and Litters Survival in Hybrids of *Brattleboro* and *WAG* Rat Lines

I. I. Khegai*

Federal Research Center, Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 630090 Russia

**e-mail: khegay@bionet.nsc.ru*

Mutation *di* determining diabetes insipidus with high level of water consumption is fixed in *Brattleboro* line of rats. We studied effect of mutant *di* gene replacement into *WAG* line of rats under gene *c* (*albino*) action. Gene *di* and *albino* inherited as recessive alleles. Alternative dominant alleles di^+ and *C* represent normal copies of genes. Original parent pairs were composed by females of *WAG* line with *c* di^+/*c* di^+ genotype and males of *Brattleboro* line with *C di*/*C di* genotype. The F_2 generation splitted into four phenotypic classes. Investigation of individual water consumption and litter sizes of F_2 females at the moment of litter weaning revealed high level of concordance with coefficient correlation – 0.996.

Keywords: vasopressin, *diabetes insipidus*, water consumption, *non-agouti*, *albino*, litter size.

ГЕНЕТИКА, 2021, том 57, № 11, с. 1333–1338

КРАТКИЕ СООБЩЕНИЯ

УДК 575;575.1;575.22

ИДЕНТИФИКАЦИЯ ПАТОГЕННЫХ CNV С ПОМОЩЬЮ ЭКЗОМНОГО СЕКВЕНИРОВАНИЯ ПРИ НЕОБЪЯСНЕННЫХ ЗАДЕРЖКАХ РАЗВИТИЯ: ИССЛЕДОВАНИЕ СЕМЕЙНОГО ТРИО

© 2021 г. О. Ю. Наумова[1, 2, *], П. В. Добрынин[1, 3], Е. А. Гибитова[3], М. А. Жукова[2, 4], С. Ю. Рычков[1], О. В. Жукова[1], Е. Л. Григоренко[2, 4]

[1]*Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, 119991 Россия*

[2]*Хьюстонский Университет, Хьюстон, Техас, 77204 США*

[3]*Университет информационных технологий, механики и оптики, Санкт-Петербург, 197101 Россия*

[4]*Санкт-Петербургский государственный университет, Санкт-Петербург, 119034 Россия*

**e-mail: oksana.yu.naumova@gmail.com*

Поступила в редакцию 20.01.2021 г.
После доработки 25.03.2021 г.
Принята к публикации 26.03.2021 г.

На примере исследования отдельного семейного случая, показано успешное применение экзомного секвенирования семейных трио (ребенок-пробанд и родители) в качестве первого метода геномной диагностики множественных необъясненных нарушений и задержек развития, с целью скриннинга на наличие патогенных однуклеотидных замен и структурных геномных вариантов (Copy Number Variation – CNV). В конкретном исследованном случае установлены клинически значимые структурные вариации *de novo* в геноме ребенка, связанные с синдромом дупликации 17p11.2, или синдромом Потоки–Люпски. Установленные геномные перестройки 17p11.2 были подтверждены методами FISH и MLPA в сертифицированных клинических лабораториях.

Ключевые слова: множественные нарушения развития, секвенирование экзома, семейный геномный анализ, структурные геномные вариации.

DOI: 10.31857/S0016675821110096

Публикации последних лет в области медицинской генетики предоставляют многочисленные свидетельства успешности и продуктивности применения полногеномного и экзомного секвенирования в клинической практике для персонализированной медицины [1]. Убедительно показана особенная ценность геномного секвенирования в трио пробанд–родители, как эффективного диагностического инструмента первого выбора в случаях сложных и труднодиагностируемых заболеваний, таких как детские неврологические заболевания [2] и задержки развития [3]. Так, недавний метаанализ релевантной литературы показал бесспорные преимущества секвенирования над методами геномного скрининга с помощью ДНК-чипов [4]. С активным развитием технологий геномного секвенирования стоимость такого анализа ежедневно сокращается, что делает метод все более доступным для клинической практики. Это в свою очередь побуждает ряд стран к разработке основ и правил внедрения этого типа анализа в государственные медицинские программы. Одним из таких примеров является прошлогодняя публикация результатов исследования одной из канадских организаций (Health Quality Ontario), разрабатывающих стандарты качества и рекомендации для здравоохранения, в которой представлена подробная оценка технологий здравоохранения относительно использования полногеномного секвенирования для пациентов с необъяснимыми нарушениями развития или множественными врожденными аномалиями [5].

В настоящем исследовании мы приводим одно из доказательств успешности применения геномного (в данном конкретном случае экзомного) секвенирования семейного трио – родителей и пробанда, в качестве первого метода геномного скриннинга случаев множественных необъясненных нарушений и задержек развития ребенка. Кроме того, мы показываем, что экзомное секвенирование позволяет успешно идентифицировать как потенциально патогенные однонуклеотидные варианты, так и клинически значимые структурные перестройки в геноме.

Участники исследования – семья славянского происхождения: родители и ребенок – девочка в возрасте 12 лет с нарушениями физического и

психического развития неизвестной этиологии. На момент рождения ребенка возраст матери и отца был 34 и 30 лет соответственно. Беременность проходила без осложнений. Однако роды, наступившие на 42-ой неделе беременности, проходили с некоторыми осложнениями: применялась искусственная стимуляция родовой деятельности, у плода было обнаружено замедленное сердцебиение и жидкость в легких. Оценка новорожденного по шкале Апгар составляла 7–8 балов, вес при рождении – 2920 г, врожденных патологий отмечено не было. Мать с новорожденной были выписаны из родильного отделения на пятый день после родов.

С первых лет жизни у ребенка отмечались множественные нарушения и задержки физического и психического развития: замедленный рост, недостаток веса, задержка развития моторных функций, нарушение сна и кормления. В анамнезе ребенка присутствуют нарушения работы желудочно-кишечного тракта и дефекации, фебрильные судороги и афективно-респираторные приступы с потерей сознания, проблемы со зрением и слухом. Со стороны психоречевого развития у ребенка наблюдаются: сенсорно-моторная алалия, дислексия, синдром дефицита внимания и нарушения обучения. По результатам психологического тестирования уровень адаптивного поведения ребенка отстает от такового у сверстников, с наиболее выраженными проблемами в навыках самообслуживания. Значительное отставание было также отмечено в речевом и познавательном развитии, оцененными по шкалам PLS-5 [7] и MSEL [8] соответственно. В этих доменах общие показатели развития соответствовали пятилетнему уровню, т.е. на семь лет меньше биологического возраста ребенка.

Общая клиническая картина расстройств и задержек развития предполагала высокую долю генетической компоненты в этиологии нарушений. В семейной истории случаев патологий и нарушений развития не отмечено. Проведенный ранее анализ кариотипа ребенка патологий не выявил, был установлен нормальный женский кариотип 46,XX. Других типов молекулярной диагностики ребенку не назначалось; диагноз поставлен не был. Все это позволяло предположить наличие *de novo* микроизменений в геноме ребенка, не улавливаемых классическими цитогенетическими методами, что, в свою очередь, обусловливало необходимость полногеномного скриннинга ребенка и родителей. Будучи заинтересованы в таком анализе, родители проявили инициативу стать участниками нашего исследовательского проекта, нацеленного на геномные исследования нарушений развития с помощью методов секвенирования. Семья предоставила свое информированное согласие на использование геномных данных и последующую публикацию результатов.

Источником ДНК являлись образцы слюны от родителей и ребенка. Геномный скриннинг семейного трио проводился методом полноэкзомного секвенирования. Для приготовления библиотек использовалась панель и набор реагентов для экзомного обогащения Illumina TruSeq Exome. Секвенирование проводилось на базе РЦ СПбГУ на платформе Illumina HiSeq4000. Для анализа данных секвенирования использовались хорошо апробированные и широко применяемые методы и инструменты. Контроль качества данных секвенирования выполнялся с помощью FastQC [9] и оценки распределения к-меров [10]. Выравнивание фрагментов и аннотирование к GRCh37/hg19 проводилось с использованием BWA [11] (версия 0.7.17-r1188) с последующей сортировкой, дедупликацией и повторной калибровкой в соответствии со стандартами и рекомендациями GATK Best Practices [12]. Обнаружение однонуклеотидных вариантов (Single Nucleotide Variant, SNV) проводилось с помощью HaplotypeCaller с последующей рекалибровкой вариантов с помощью VariantRecalibrator, а определение *de novo* SNV – с помощью VariantAnnotator (GATK 4.1.6). Кроме того, на основе данных экзомного секвенирования и с помощью алгоритмов сегментации cn.MOPS [13] были идентифицированы структурные геномные вариации или вариации числа копий (Copy Number Variation, CNV).

По результатам анализа данных секвенирования и аннотирования к референсному геному в семейном трио было обнаружено свыше 30 тыс. SNV. С помощью инструментов ANNOVAR [14] была проведена фильтрация SNV на основе основных моделей наследования (аутосомно-доминантного, рецессивного и Х-сцепленного), с последующей функциональной аннотацией и приоритезацией вариантов в контексте ассоциации с фенотипом задержки развития. Потенциальных вариантов-кандидатов обнаружено не было. Сходный результат был получен для анализа *de novo* SNV в геноме ребенка. Таковых было обнаружено 23 SNV в состоянии гетерозиготы, 10 из которых являлись миссенс-мутациями, но только две из них: rs77696629 и rs1255143876 – замены в 10-ом экзоне гена карбоксилэфирной липазы *CEL* и 26-ом экзоне муцина *MUC5B* соответственно, могут быть отнесены ко вредным мутациям (табл. 1). Оба варианта являются достаточно редкими (Minor allele frequency – MAF < 0.05) и затрагивают гены, мутации в которых известно ассоциированы с развитием таких заболеваний как MODY-диабет (*CEL*), заболеваний дыхательной системы – астма и идиопатический легочный фиброз (*MUC5B*). Однако следует отметить, что ни один из выявленных *de novo* SNV не являлся клинически значимым согласно результатам аннотации к релевантной базе данных ClinVar [15].

Таблица 1. Миссенс-замены *de novo*, выявленные у ребенка на основе данных экзомного секвенирования семейной триады

Позиция (GRCh37)	Ген	Аллель	Код dbSNP	Замена	Значимость*	Частота**	
						ExAC	gnomAD
chr4:145041720	*GYPA*	A > G	rs7682260	ex2:c.T59C:p.L20S	N	0.0502	0.1846
chr6:170627739	*FAM120B*	G > C	rs17860761	ex2:c.G1297C:p.A433P	N	0.025	0.0105
chr9:135946015	*CEL*	T > C	rs77696629	ex10:c.T1463C:p.I488T	D	0.0163	0.0263
chr11:1093412	*MUC2*	C > T	rs113330607	ex31:c.C5231T:p.T1744M	N	0.2426	0.0746
chr11:1260248	*MUC5B*	G > A	rs1255143876	ex26:c.G3445A:p.D1149N	D		0.0037
chr17:78063996	*CCDC40*	A > G	rs7210679	ex18:c.A2891G:p.H964R	N	0.2679	0.2652
chr19:9025654	*MUC16*	T > G	rs201931630	ex15:c.A36800C:p.K12267T	N	0.0019	0.0011
chr19:56274503	*RFPL4A*	T > G	rs140087406	ex3:c.T826G:p.S276A	N	0.0001	0.0005
chr19:56274506	*RFPL4A*	G > A	rs147984855	ex3:c.G829A:p.A277T	N	0.0011	0.0167
chr19:56274507	*RFPL4A*	C > G	rs149792632	ex3:c.C830G:p.A277G	N	0.0001	0.0514

Примечание. * – функциональная значимость SNV: нейтральный вариант – N, вредная мутация – D; определялась по результатам оценок шестью алгоритмами: SIFT [16], PolyPhen-2, PROVEAN [17], MutationTaster [18], M-CAP [19] и CADD [20]; статус D присуждался при условии, что три из шести алгоритмов предсказывают достоверно высокий уровень вредоносности мутации. ** – приведена частота варианта в европейской популяции по данным агрегаторов консорциумов gnomAD [25] и ExAC [26].

При фактическом отсутствии патогенных SNV анализ структурных вариаций позволил установить несколько потенциально клинически значимых перестроек в геноме ребенка. Всего в семейном трио было выявлено 54 CNV, из них 22 – в геноме ребенка, включая 11 микроделеций и дупликаций *de novo*, отмеченных в восьми хромосомных сегментах (табл. 2). Большинство структурных вариаций может быть отнесено к распространенным полиморфным вариантам: например CNV в регионе сегментальной дупликации генов на хромосоме 1 (1q21.2) и в высокополиморфном локусе на хромосоме 6 (6p21.3), несущем гены главного комплекса гистосовместимости HLA. Однако микроделеции 15q11.2 и дупликации 17p11.2 могут являться маркерами клинически значимых структурных вариаций генома. Также следует отметить, что в данных хромосомных сегментах было обнаружено по два блока делеций и дупликаций соответственно (табл. 2), что при условии дискретности экзомных данных в контексте покрытия генома может являться маркером гораздо более крупных перестроек в данных сегментах.

Будучи выявлены как микроперестройки, не затрагивающие основные синдромальные гены-кандидаты, делеции 15q11.2 и дупликации 17p11.2 тем не менее пересекаются с регионами и типами структурных изменений, связаными с одноименными синдромами (табл. 2), проявляющимися прежде всего как множественные нарушения и задержки развития. Так, делеция NC_000015.9:g.23444002_23600066del была отмечена в районе дупликации гена *HERC2*, фланкирующего консервативный участок BP1–BP2 (Break Point) критического региона Прадера–Вилли/ Ангельмана 15q11-13. Делеция 300–500 тпн в этом участке, несущем неимпринтные гены *CYFIP1*, *NIPA1*, *NIPA2* и *TUBGCP5*, связана с синдромом микроделеции 15q11.2. Выявленные инсерции 17p11.2 длиной 64 и 95 тпн пересекаются с CNV NC_000017.10:g.16757111-20219651)dup, обусловливающем классический синдром дупликации 17p11.2, или синдром Потоки–Люпски (табл. 2). Следует отметить, что комплекс нарушений, как при синдроме микроделеции 15q11.2 – неврологические дисфункции, задержка речевого и моторного развития и аутистическое расстройство, так и при синдроме Потоки–Люпски – задержки психо-физического развития, гипотонус мышц, проблемы с кормлением и расстройства поведения, в высокой степени соответствуют клиническому фенотипу исследованного ребенка.

Как следствие, CNV в 15q11.2 и 17p11.2 были рассмотрены в качестве ключевых для определения этиологии/диагностики данного случая нарушений развития. Полученные в результате нашего исследования данные позволили семье инициировать направленные клинические тесты. Было проведено молекулярно-цитогенетическое исследование хромосом 15 и 17 методом *in situ* FISH-гибридизации в одной из сертифицированных клинических лабораторий г. Москвы. Выявленные нами микроделеции в локусе 15q11.2 FISH-методом подтверждены не были, тогда как наличие у ребенка дупликации района 17p11.2 было подтверждено как присутствие двух копий локуса *LIS* и трех копий локуса *RAI1* в каждом ядре [nuc ish(LISx2,RAI1x3); ДНК-зонд MD Miller-Dieker LIS (17p13)/MD Smith-Magenis RAI (17p11), KREATECH]. Помимо этого дупликация 17p11.2 была дополнительно валидиро-

Таблица 2. Структурные геномные вариации *de novo*, выявленные у ребенка на основе данных экзомного секвенирования семейной триады

Хромомсома, сегмент	Структурный вариант	Длина (п.о.)	Структурная перестройка	Гены
1q21.2	NC_000001.10:g.144617148_144677049del	59902	Делеция	*NBPF8, NBPF9, NBPF20, PDE4DIP*
6p21.33	NC_000006.11:g.31083750_31106558del	22809	То же	*PSORS1C1, PSORS1C2, CDSN*
	NC_000006.11:g.31129221_31165659del	36439	»	*TCF19, POU5F1, PSORS1C3, HCG27*
8p23.1	NC_000008.10:g 12274164_12286409del	12246	»	*FAM66A, FAM86B2*
8p11.22	NC_000008.10:g.39311490_39349532del	38043	»	*ADAM3A*
14q32.33	NC_000014.8:g.106452668_106494604del	41937	»	Вариабельная область Ig
15q11.1	NC_000015.9:g 20636293_20663214del	26922	»	*HERC2P3*
15q11.2*	NC_000015.9:g.20776030_20778064del	2035	»	*GOLGA8CP*
	NC_000015.9:g.23444002_23600066del	156065	»	*GOLGA8EP, HERC2P2*
17p11.2**	NC_000017.10:g.18325571_18389514dup	63944	Дупликация	*KRT16P1, LGALS9C*
	NC_000017.10:g.18395987_18491218dup	95232	То же	*LGALS9C, USP32P2*

Примечание. Звездочками отмечены хромосомные локусы, структурые перестройки в которых ассоциированы с клиническими фенотипами; приведено по данным интегрированной БД заболеваний и их аннотаций MalaCards [21]. * – синдром микроделеции 15q11.2: делеция 300–500 тпн в консервативном неимпринтном участке BP1–BP2 критического региона Прадера–Вилли/Ангельмана (NC_000015.9:g.22765628-23317514del), гены-кандидаты – *CYFIP1*, *NIPA1*, *NIPA2* и *TUBGCP5* [15]. ** – синдром микродупликации 17p11.2, или синдром Потоки–Люпски: размер дупликации варьирует от 1.3 до 15.2 Мб, наиболее распространенный вариант дупликации ~3.5 Мб NC_000017.10:g.16757111_20219651 dup включает 49 генов, основные гены-кандидаты – *FLCN* и *RAI1* [15].

вана другой диагностической лабораторией методом мультиплексной амплификации лигированных зондов (MRC-Holland SALSA MLPA probemix P369-A2). Таким образом, она установлена как дупликация g.(?_16852223)_(20130903_?)dup (GRCh37/hg19), включающая гены *TNFRSF13B*, *COPS3*, *MIR33B*, *TOM1L2*, *DRC3*, *LLGL1*, *PRPSAP2*, *MFAP4*, *ALDH3A1*, *AKAP10*, *SPECC1* и основные гены-кандидаты клинического фенотипа – *FLCN* и *RAI1*.

В результате, обследованному ребенку, на 13 году его жизни, был поставлен диагноз – вновь приобретенный синдром Потоки–Люпски, что для семьи означает определенность в прогнозах для здоровья и развития дочери и помощь в принятии решения о рождении второго ребенка. Потоки–Люпски (OMIM:610883) – это редкое (1:25 тыс. новорожденных) генетическое заболевание, впервые открытое в 1996 г. [22] и описанное в 2007 г. [23]. Прежде всего, оно связано с нарушением развития ЦНС и пороками сердца. Как многие другие генетические заболевания данная патология характеризуется вариативными клиническими фенотипами [24], что значительно усложняет диагностику в отсутствие генетического анализа.

Таким образом, нами еще раз показана успешность применения экзомного секвенирования для первичной диагностики множественных нарушений развития. Мы уверены, что в условиях постоянно снижающейся стоимости молекулярного и биоинформатического анализов, связанных с геномным секвенированием, подобное тестирование семейных трио в скором будущем найдет свое применение в широкой клинической практике. Также полагаем, что изложенные в данном сообщении результаты анализа CNV и значимости их валидации могут быть полезны для других коллективов, использующих в своих исследованиях методы таргетного экзомного секвенирования для выявления геномных вариаций.

Работа выполнена при поддержке гранта РФФИ-17-29-02384, руководитель проекта Е.Л. Григоренко.

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

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. *Tărlungeanu D.C., Novarino G.* Genomics in neurodevelopmental disorders: an avenue to personalized medicine // Ex. Mol. Med. 2018. V. 50. № 8. P. 1–7. https://doi.org/10.1038/s12276-018-0129-7
2. *Bowling K.M., Thompson M.L., Amaral M.D. et al.* Genomic diagnosis for children with intellectual disability and/or developmental delay // Genome Med. 2017. V. 9. № 1. P. 43. https://doi.org/10.1186/s13073-017-0433-1
3. *Han J.Y., Lee I.G.* Genetic tests by next-generation sequencing in children with developmental delay and/or intellectual disability // Clin. Exp. Pediatr. 2020. V. 63. № 6. P. 195–202. https://doi.org/10.3345/kjp.2019.00808
4. *Srivastava S., Love-Nichols J.A., Dies K.A. et al.* Meta-analysis and multidisciplinary consensus statement: Exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders // Genet. Med. 2019. V. 21. № 11. P. 2413–2421. https://doi.org/10.1038/s41436-019-0554-6
5. *Ontario Health Q.* Genome-wide sequencing for unexplained developmental disabilities or multiple congenital anomalies: A health technology assessment // Ont. Health Technol. Assess. Ser. 2020. V. 20. № 11. P. 1–178.
6. *Sparrow S.S., Cicchetti D., Balla D.V.* Vineland Adaptive Behavior Scales. San Antonio, TX: Pearson Assessment, 2005.
7. *Zimmerman I.L., Steiner V.G., Pond R.A.* The Preschool Language Scale-5. San Antonio, TX, Pearson, 2011.
8. *Mullen E.M.* Mullen Scales of Early Learning. Circle Pines, MN: American Guidance Service Inc., 1995.
9. *Andrews S.* FastQC: A quality control tool for High Throughput Sequence Data [Online]. 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
10. *Marçais G., Kingsford C.* A fast, lock-free approach for efficient parallel counting of occurrences of k-mers // Bioinformatics. 2011. V. 27. № 6. P. 764–770. https://doi.org/10.1093/bioinformatics/btr011
11. *Li H., Durbin R.* Fast and accurate short read alignment with Burrows–Wheeler transform // Bioinformatics. 2009. V. 25. № 14. P. 1754–1760. https://doi.org/10.1093/bioinformatics/btp324
12. *Van der Auwera G.A., Carneiro M.O., Hartl C. et al.* From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline // Curr. Protoc. Bioinformatics. 2013. V. 43. № 1110. P. 11.10.1–11.10.33. https://doi.org/10.1002/0471250953.bi1110s43
13. *Klambauer G., Schwarzbauer K., Mayr A. et al.* cn.MOPS: Mixture of Poissons for discovering copy number variations in next-generation sequencing data with a low false discovery rate // Nucl. Ac. Res. 2012. V. 40. № 9. P. e69. https://doi.org/10.1093/nar/gks003
14. *Wang K., Li M., Hakonarson H.* ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data // Nucl. Ac. Res. 2010. V. 38. № 16. P. e164. https://doi.org/10.1093/nar/gkq603
15. *Landrum M.J., Lee J.M., Benson M. et al.* ClinVar: Improving access to variant interpretations and supporting evidence // Nucl. Ac. Res. 2018. V. 46. № D1. P. D1062–D1067. https://doi.org/10.1093/nar/gkx1153
16. *Kumar P., Henikoff S., Ng P.C.* Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm // Nat. Protoc. 2009. V. 4. № 7. P. 1073–1081. https://doi.org/10.1038/nprot.2009.86
17. *Choi Y., Sims G.E., Murphy S. et al.* Predicting the functional effect of amino acid substitutions and indels // PLoS One. 2012. V. 7. № 10. P. e46688. https://doi.org/10.1371/journal.pone.0046688
18. *Schwarz J.M., Cooper D.N., Schuelke M., Seelow D.* MutationTaster2: Mutation prediction for the deep-sequencing age // Nat. Methods. 2014. V. 11. № 4. P. 361–362. https://doi.org/10.1038/nmeth.2890
19. *Jagadeesh K.A., Wenger A.M., Berger M.J. et al.* M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity // Nat. Genet. 2016. V. 48. № 12. P. 1581–1586. https://doi.org/10.1038/ng.3703
20. *Rentzsch P., Witten D., Cooper G.M. et al.* CADD: Predicting the deleteriousness of variants throughout the human genome // Nucl. Ac. Res. 2018. V. 47. № D1. P. D886–D894. https://doi.org/10.1093/nar/gky1016
21. *Rappaport N., Twik M., Plaschkes I. et al.* MalaCards: An amalgamated human disease compendium with diverse clinical and genetic annotation and structured search // Nucl. Ac. Res. 2017. V. 45. № D1. P. D877–D887. https://doi.org/10.1093/nar/gkw1012
22. *Brown A., Phelan M.C., Patil S. et al.* Two patients with duplication of 17p11.2: The reciprocal of the Smith-Magenis syndrome deletion // Am. J. Med. Genet. 1996. V. 63. № 2. P. 373–377. https://doi.org/10.1002/(SICI)1096-8628(19960517)63:2<373::AID-AJMG9>3.0.CO;2-U
23. *Potocki L., Bi W., Treadwell-Deering D. et al.* Characterization of Potocki–Lupski syndrome (dup(17)(p11.2p11.2)) and delineation of a dosage-sensitive critical interval that can convey an autism phenotype // Am. J. Hum. Genet. 2007. V. 80. № 4. P. 633–649. https://doi.org/10.1086/512864
24. *Potocki L., Neira-Fresneda J., Yuan B.* Potocki-Lupski Syndrome // Gene Reviews® [Internet]. Seattle (WA): University of Washington, Seattle, 2017. https://www.ncbi.nlm.nih.gov/books/NBK447920/
25. *Karczewski K.J., Francioli L.C., Tiao G. et al.* The mutational constraint spectrum quantified from variation in 141,456 humans // Nature. 2020. V. 581. P. 434–443. https://doi.org/10.1038/s41586-020-2308-7
26. *Lek M., Karczewski K.J., Minikel E.V. et al.* Analysis of protein-coding genetic variation in 60,706 humans // Nature. 2016. V. 536. P. 285–291. https://doi.org/10.1093/nar/gkw971

Identification of Pathogenic CNVs in Unexplained Developmental Disabilities Using Exome Sequencing: A Family Trio Study

O. Yu. Naumova[a, b, *], P. V. Dobrynin[a, c], E. A. Gibitova[c], M. A. Zhukova[b, d], S. Yu. Rychkov[a], O. V. Zhukova[a], and E. L. Grigorenko[b, d]

[a] *Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, 119991 Russia*
[b] *University of Houston, Houston, TX, 77204 USA*
[c] *University of Information Technologies, Mechanics and Optics, Saint Petersburg, 197101 Russia*
[d] *Saint Petersburg State University, Saint Petersburg, 119034 Russia*
**e-mail: oksana.yu.naumova@gmail.com*

This short report on a family case-study provides evidence of the effectiveness of exome sequencing of family trios (proband-parents) as the first-tier test in genomic diagnostics of unexplained developmental disorders and delays, as a genomic screening for the pathogenic single-nucleotide variants and copy number variations (CNVs). Several clinically significant structural genomic variations *de novo* were identified in the proband's genome, specifically those associated with the 17p11.2 duplication (or Potocki–Lupski) syndrome. These genome rearrangements have been confirmed in the certified clinical laboratories using both FISH and MLPA techniques.

Keywords: unexplained developmental disorders and delays, family study, exome sequencing, structural genomic variation.

ГЕНЕТИКА, 2021, том 57, № 11, с. 1339–1342

КРАТКИЕ СООБЩЕНИЯ

УДК 61:575:17

РАЗНООБРАЗИЕ МОНОГЕННЫХ НАСЛЕДСТВЕННЫХ БОЛЕЗНЕЙ СРЕДИ РУССКОГО НАСЕЛЕНИЯ ВОЛГО-УРАЛЬСКОГО РЕГИОНА

© 2021 г. Р. А. Зинченко[1, 2, *], Е. К. Гинтер[1, 3], В. В. Кадышев[1], С. С. Амелина[4], Т. П. Васильева[2], О. Ю. Александрова[2], С. И. Куцев[1]

[1]Медико-генетический научный центр им. академика Н.П. Бочкова, Москва, 115522 Россия

[2]Национальный научно-исследовательский институт общественного здоровья им. Н.А. Семашко, Москва, 105064 Россия

[3]Российская медицинская академия непрерывного профессионального образования, Москва, 125993 Россия

[4]Ростовский государственный медицинский университет, Ростов-на-Дону, 344022 Россия

**e-mail: renazinchenko@mail.ru*

Поступила в редакцию 25.01.2021 г.
После доработки 01.03.2021 г.
Принята к публикации 16.03.2021 г.

Представлены результаты разнообразия моногенных наследственных болезней (МНБ) среди русского населения десяти регионов Российской Федерации с различной этнической экстракцией: пять с коренным обитанием, пять в этнических популяциях Поволжья. Сравнительный анализ проведен с разнообразием МНБ у десяти этнических групп РФ. Суммарная численность обследованного населения составила 2.837 млн чел. (включая 1.624 млн русских и 1.213 млн представителей десяти различных этнических групп). Разнообразие МНБ составило 480 заболеваний, выявленных у 8355 больных. Определен спектр и распространенность частых заболеваний. Проведен сравнительный анализ геногеографических взаимоотношений между двадцатью этносами/популяциями по распространенности 201 аутосомно-рецессивного заболевания. Показано, что несмотря на интенсивную метисацию русских с коренным населением Поволжья, при положительной этнической брачной ассортативности генофонд русских по генам МНБ сохраняется.

Ключевые слова: генетическая эпидемиология, разнообразие менделирующей наследственной патологии, распространенность, русские, Поволжье.

DOI: 10.31857/S0016675821110163

Изучение разнообразия моногенных наследственных болезней (**МНБ**), их территориального распространения, этнической приуроченности, аллельной и локусной гетерогенности является актуальной задачей медицинской генетики.

Целью настоящей публикации является геногеографический анализ разнообразия МНБ у русского населения различных регионов, включая полиэтнические (пять республик Волго-Уральского региона), а также оценка влияния метисации на возможные изменения патогенома русских.

Русские являются титульным, самым многочисленной народом России (более 80%) и Европы. Несмотря на то что русские популяции являются довольно однородными в антропологическом отношении, однако, на основе различий региона проживания, традиций, особенностей культуры, выделяют русских Кавказа, Дона, центральной, северной и азиатской частей России. Проведен сравнительный анализ разнообразия наследственных заболеваний у русского населения, как в популяциях коренного обитания русских (пять регионов) в западной, центральной и южной частях РФ, так и в популяциях с различной этнической экстракцией – в Волго-Уральском регионе (Марий Эл, Чувашия, Удмуртия, Башкортостан, Татарстан). В анализ для сравнения также включено разнообразие МНБ у десяти этнических групп Российской Федерации – пяти народов Поволжья (марийцев Марий Эл, чувашей Чувашии, удмуртов Удмуртии, башкир Башкортостана и татар Татарстана), и пяти народов Северного Кавказа (карачаевцев, черкесов, абазин, ногайцев Карачаево-Черкессии и адыгейцев Адыгеи). Исследование проведено по оригинальному протоколу, разработанному в лаборатории генетической эпидемиологии ФГБНУ “МГНЦ”, предусматривающему комплексный анализ генетического портрета популяций одновременно через различные генетические и биологические системы [1, 2]. Методы сбора и статистической обработки данных остаются неизменны на всем протяжении исследова-

Таблица 1. Нозологический спектр и распространенность частой (чаще, чем 1 : 50000) наследственной патологии среди русского населения и десяти этнических групп европейской части РФ

№	OMIM [6]	Диагноз	Распространенность на 100000, человек						
			РРУ	РРЧ	РМЭ	РРБ	РРТ	РПоп	ЭПоп
Аутосомно-доминантное наследование									
1	PS118220	Наследственная моторно-сенсорная нейропатия	3.57	5.65	3.38			5.48	7.50
2	#162200	Нейрофиброматоз, тип I	2.38	4.24	6.76		3.70	**3.39**	**7.34**
3	178300	Врожденный птоз	1.19	2.83	1.13			2.71	5.03
4	PS268000	Пигментный ретинит				2.83		2.59	3.46
5	PS116200	Врожденная катаракта			6.76	5.65	7.40	4.56	7.01
6	#146000	Гипохондроплазия		2.83	1.13	5.65	3.70	4.31	5.36
7	PS166200	Несовершенный остеогенез	14.27		1.13			2.28	1.98
8	151900	Множественный липоматоз	3.57				11.10	2.89	3.46
9	#146700	Вульгарный ихтиоз	2.38	2.83	16.91	5.65	14.80	**13.12**	**26.79**
10	#154700	Синдром Марфана	10.70	4.24	5.64		3.70	2.03	4.78
11	PS130000	Синдром Элерса-Данло	2.38			8.48	3.70	2.77	26.70
12	PS124900	Несиндромальная тугоухость	4.76		2.25	16.96	14.80	3.88	4.37
Аутосомно-рецессивное наследование									
1	PS253600	Поясно-конечностная мышечная дистрофия	3.57	2.99	1.13			2.61	1.81
2	#253300	Спинальная мышечная атрофия	1.19		2.25			1.97	2.14
3	PS251200	Микроцефалия, олигофрения	3.57	1.50	6.76	16.96	3.70	3.17	6.68
4	PS249500	Нарушение интеллектуального развития	10.70		1.13	31.09	22.20	**4.11**	**27.61**
5	PS268000	Пигментный ретинит		16.46	4.51			4.58	3.54
6	PS116200	Врожденная катаракта		1.50	2.25			2.69	3.46
7	#242100	Ихтиозиформная эритродермия		2.99	2.25			2.52	2.06
8	#261600	Фенилкетонурия	2.38	1.50	4.51	5.65	3.70	2.57	7.01
9	PS220290	Несиндромальная тугоухость	20.21	26.93	22.40	39.57	62.89	23.43	35.61
Х-сцепленное наследование*									
1	#310200	Миопатия Дюшенна		2.83	9.02			2.96	2.97
2	PS309530	Нарушение интеллектуального развития	2.38	5.65	6.76	45.23	73.99	**2.89**	**23.74**
3	#308100	Ихтиоз	19.02	2.83	4.51	5.65		**3.88**	**6.26**
4	PS310700	Врожденный нистагм			0.25			0.86	5.77

Примечание. В графе OMIM – номера заболеваний по международному каталогу В. Мак Кьюсика; PS – фенотипическая серия заболеваний гетерогенной группы по OMIM; РРУ – русские Удмуртии, РРЧ – русские Чувашии, РМЭ – русские Марий Эл, РРБ – русские Республики Башкортостан, РРТ – русские Республики Татарстан, РПоп – средние значения распространенности заболевания по результатам генетико-эпидемиологических исследований европейской части РФ для русских десяти регионов, ЭПоп – средние значения распространенности заболевания для десяти этнических групп Поволжья и Северного Кавказа; * – распространенность на мужское население.

ний и подробно описаны ранее [1–3]. Суммарная численность обследованного населения составила 2.837 млн чел. (включая 1.624 млн русских и 1.213 млн представителей десяти этнических групп).

Разнообразие МНБ составило 480 заболеваний, выявленное у 8355 больных: 238 (4771 пациентов) с аутосомно-доминантным (**АД**) типом наследования, 201 (2918 больных) с аутосомно-рецессивным (**АР**) и 41 (666 больных) с Х-сцепленным типом наследования (**Х-сц**.). Нозологический спектр МНБ среди русского населения составил 305 заболеваний (166 с АД, 117 с АР и 22 с Х-сц. типом наследования). Для каждого региона определены различия в разнообразии и точечной распространенности МНБ, оцененные при обследовании населения. Для русских не было выявлено статистически значимых эндемичных МНБ по регионам. Наибольшие различия определены между распространенностью МНБ среди русских и дру-

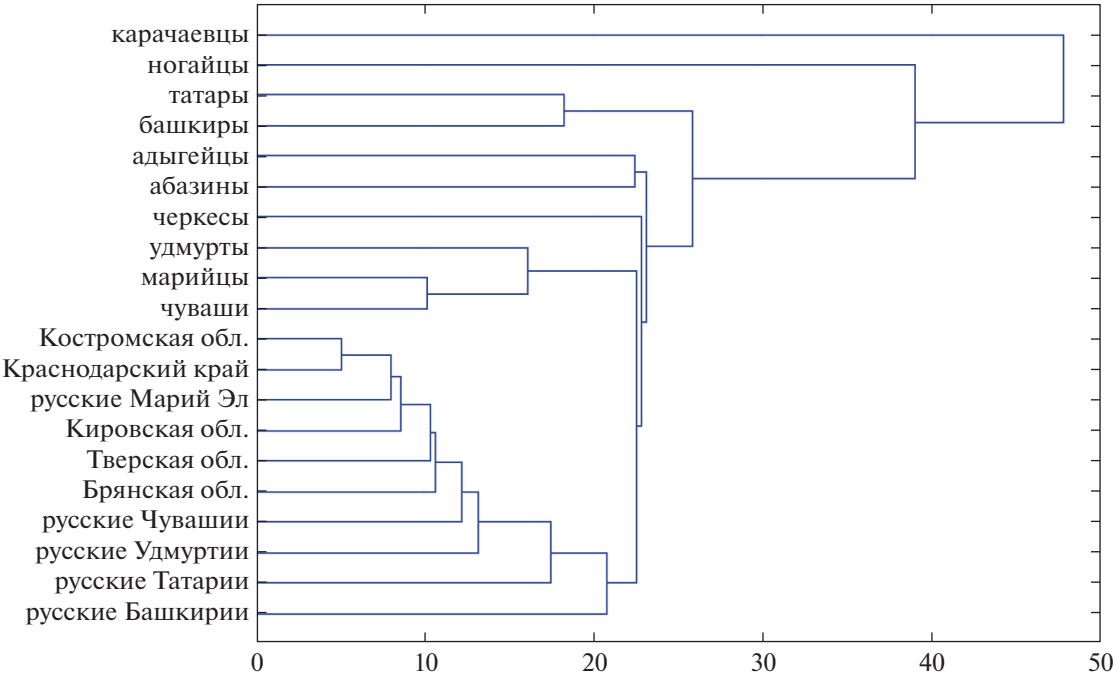

Рис. 1. Дендрограмма генетических расстояний, построенная по распространенности 201 АР-заболевания.

гих этнических групп (эти заболевания выделены в табл. 1 жирным шрифтом). Такая картина характерна для многих популяций мира – особенности распространения МНБ напрямую зависят от генетической структуры населения, от действия эффективного дрейфа в популяциях, миграционных процессов и естественного отбора, что подтверждалось для каждой обследованной популяции [2–7].

Рассчитана кумулятивная (АД, АР и Х-сц.) распространенность МНБ для русских десяти регионов, которая составила 1 : 532 человека, для представителей этнических групп – 1 : 230, что более, чем в два раза выше. Выявленные различия характерны не только для регионов РФ, но и для других популяций мира [4, 5, 8, 9].

Определено разнообразие частых заболеваний (распространенность чаще, чем 1 : 50000). Группа частых заболеваний аккумулировала наибольшее число больных: для заболеваний АД – 59% пациентов, для АР – 66.48%, для Х-сц. – 68.32%. В табл. 1 представлен спектр и распространенность частых заболеваний у русского населения в пяти популяциях Волго-Уральского региона, суммарная характеристика по русским десяти регионов (включая пять популяций коренного проживания русских), а также у этнических групп РФ [1, 9].

Проведен сравнительный анализ геногеографических взаимоотношений между двадцатью рассматриваемыми этносами/популяциями по распространенности 201 АР-заболевания, наиболее контрастно отражающих изменения патогенома [5, 7, 8]. Выполнен кластерный анализ среднесвязывающим методом с использованием программного пакета “Statistica 16”.

На дендрограмме (рис. 1) представлены полученные результаты кластерного анализа. Как следует из рис. 1, на самом первом этапе объединяются русские различных регионов, включая русских Поволжья. Дальнейшее формирование кластеров отражает этногенез народов. Отдельные кластеры формируются из финно-угорских народов (марийцы-удмурты-чуваши) и адыго-абхазских народов (адыгейцы-абазины-черкесы). Замыкают анализ кластеры тюркоязычных народов – “башкиры-татары” и “ногайцы-карачаевцы”. Таким образом, несмотря на интенсивную метисацию русских Поволжья (52.5%), при положительной этнической брачной ассортативности [10] генофонд русских в его патологической части сохраняется. Теоретически, рассматривая популяции с различной этнической экстракцией, от процесса метисации можно ожидать эффект выравнивания генных частот с одновременным расширением генетического разнообразия для обеих популяций. Однако необходимо учитывать и другие факторы динамики генофонда – миграции и естественный отбор, постоянно действующие в популяциях, которые могут изменить этот процесс, приводя к формированию нового равновесного состояния генных частот.

Работа выполнена при финансовой поддержке РФФИ, проект № 20-015-00061 и государственного задания министерства образования и науки РФ.

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

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. *Зинченко Р.А., Куцев С.И., Александрова О.Ю., Гинтер Е.К.* Основные методологические подходы к выявлению и диагностике моногенных наследственных заболеваний и проблемы в организации медицинской помощи и единых профилактических программ // Пробл. соц. гигиены, здравоохранения и истории медицины. 2019. Т. 27. № 5. С. 865–877. https://doi.org/10.32687/0869-866X-2019-27-5-865-877
2. *Zinchenko R.A., Makaov A.Kh., Marakhonov A.V. et al.* Epidemiology of hereditary diseases in Karachay-Cherkess Republic // Int. J. Mol. Sci. 2020. V. 21(1). P. 325–329. https://doi.org/10.3390/ijms21010325
3. Методология генетико-эпидемиологического изучения наследственных болезней и врожденных пороков развития. Учебно-методическое пособие / Под. ред. Зинченко Р.А., Амелиной С.С., Куцева С.И., Гинтера Е.К. Белгород: “Политерра”, 2020. 569 с. https://doi.org/10.34851/PHP.2020.31.777
4. *Ferreira C.R.* The burden of rare diseases // Am. J. Med. Genet. 2019. V. 179A. P. 885–892. https://doi.org/10.1002/ajmg.a.61124
5. *Wu D.D., Zhang Ya.P.* Different level of population differentiation among human genes // BMC Evol. Biol. 2011. V. 11(16). https://doi.org/10.1186/1471-2148-11-16
6. *Зинченко Р.А., Ельчинова Г.И., Гинтер Е.К.* Факторы, определяющие распространение наследственных болезней в российских популяциях // Мед. генетика. 2009. Т. 8. № 12. С. 7–23.
7. *Гинтер Е.К., Мамедова Р.А., Ельчинова Г.И., Брусинцева О.В.* Генетическая структура популяций Кировской области и особенности территориального распределения в ней аутосомно-рецессивных заболеваний // Генетика. 1994. Т. 30. № 1. С. 107–111.
8. Procedural document on Epidemiology of rare disease in Orphanet, Orphanet, February 2019. Version 01. https://www.orpha.net/orphacom/cahiers/docs/GB/Epidemiology_in_Orphanet_R1_Ann_Epi_EP_05.pdf
9. Online Mendelian Inheritance in Man. URL: http://www.ncbi.nlm.nih.gov/OMIM (Updated: January, 2020).
10. *Ельчинова Г.И., Зинченко Р.А.* Русские Поволжья. Интенсивность метисации в конце XX века // Генетика. 2021. Т. 57. № 1. С. 112–115. https://doi.org/10.31857/S0016675821010057

Diversity of Monogenic Hereditary Diseases among the Russian of the Volga-Ural Region

R. A. Zinchenko[a, b, *], E. K. Ginter[a, c], V. V. Kadyshev[a], S. S. Amelina[d], T. P. Vasilieva[b], O. Yu. Alexandrova[b], and S. I. Kutsev[a]

[a] *Research Centre for Medical Genetics, Moscow, 115522 Russia*

[b] *Semashko National Research Institute of Public Health, Moscow, 105064 Russia*

[c] *Russian Medical Academy of Continuous Professional Education, Moscow, 125993 Russia*

[d] *Rostov State Medical University, Rostov-on-Don, 344022 Russia*

**e-mail: renazinchenko@mail.ru*

We present the results of the diversity of monogenic hereditary diseases (MHDs) among the Russian population of 10 regions of the Russian Federation with different ethnic extraction – five with native populations and five in ethnic populations of the Volga region. A comparative analysis was performed with the diversity of MHDs in 10 ethnic groups of the RF. The total population surveyed was 2.837 million people (including 1.624 million Russians and 1.213 million representatives of 10 ethnic groups). The diversity of the MHDs was 480 diseases detected in 8355 patients. The spectrum and prevalence of frequent diseases were determined. A comparative analysis of the genogeographic relationships between twenty ethnicities/populations was performed on the prevalence of 201 AR diseases. It was shown that despite intensive mestization of Russians with the native population of the Volga region with positive ethnic mating assortativity, the gene pool of Russians according to the MHDs genes is preserved.

Keywords: genetic epidemiology, diversity of mendelian hereditary pathology, prevalence, Russian, Volga-Ural region.