

ISSN 0016-6758

Том 59, Номер 12

Декабрь 2023



ГЕНЕТИКА



www.sciencejournals.ru



СОДЕРЖАНИЕ

Том 59, номер 12, 2023

Обзорные и теоретические статьи

Генетика и эпигенетика несуицидального самоповреждающего поведения:
нарративный обзор

А. А. Кибитов, Г. Э. Мазо 1347

Генетика и эпигенетика преждевременного полового созревания

Е. А. Саженова, С. А. Васильев, Л. В. Рычкова,
Е. Е. Храмова, И. Н. Лебедев 1360

Общая генетика

Регуляция одноцепочечных и двуцепочечных кластеров рiРНК в герминальных
и соматических тканях *Drosophila melanogaster* зависит от гена *rhino*

П. А. Миляева, А. Р. Лавренов, И. В. Кузьмин, А. И. Ким, Л. Н. Нефедова 1372

Генетика растений

Биологическое разнообразие генов гомологов дефензинов пшеницы

М. П. Слезина, Е. А. Истомина, Т. И. Одинцова 1382

Генетика животных

Популяционно-генетическая структура белого медведя (*Ursus maritimus*)
в морях российской Арктики

П. А. Сорокин, Е. Ю. Звычайная, Е. А. Иванов, И. А. Мизин,
И. Н. Мордвинцев, Н. Г. Платонов, А. И. Исаченко,
Р. Е. Лазарева, В. В. Рожнов 1393

Невыраженность популяционно-генетической структуры серого журавля *Grus grus* L.

Е. А. Мудрик, Ю. М. Маркин, К. А. Постельных, К. Д. Кондракова,
Т. А. Кащенко, А. В. Шатохина, П. А. Казимиров,
В. Ю. Ильяшенко, Е. И. Ильяшенко, Д. В. Политов 1407

Генетика человека

Метилирование подсемейств ретротранспозона LINE-1 в ворсинах хориона
при невынашивании беременности

С. А. Васильев, В. В. Деменева, Е. Н. Толмачева, С. А. Филатова, А. С. Зуев,
А. С. Ушакова, О. Ю. Васильева, И. В. Лушников,
Е. А. Саженова, Т. В. Никитина, И. Н. Лебедев 1419

Транскриптомное исследование клеток плаценты при преэклампсии
как эффективный инструмент персонализированной медицины

Е. А. Трифонова, А. А. Бабовская, А. А. Зарубин, В. Н. Сереброва,
М. М. Гавриленко, М. Г. Сваровская, Е. В. Ижойкина,
И. Г. Куценко, В. А. Степанов 1427

Математические модели и методы

Описание дивергенции субпопуляций в иерархической системе при анализе изонимии.
III. Коэффициент фамильного инбридинга

В. П. Пасеков 1440

Краткие сообщения

Полиморфизм коровых и симбиотически специализированных генов
у политипических видов клубеньковых бактерий

*Н. А. Проворов, А. К. Кимеклис, Е. С. Карасев, С. Хосид,
О. П. Онищук, О. Н. Курчак, Е. Е. Андронов*

1457

Вариации синапсиса хромосом в профазе I мейоза у слепушонок
Ellobius tancrei гетерозиготных по робертсоновским транслокациям

*С. Н. Матвеевский, Ю. Ф. Богданов, Е. А. Ляпунова,
И. Ю. Баклушинская, О. Л. Коломиец*

1462

Contents

Vol. 59, No. 12, 2023

Reviews and Theoretical Articles

- Genetics and Epigenetics of Nonsuicidal Self-Injury: a Narrative Review
A. A. Kibitov and G. E. Mazo 1347

- Genetics and Epigenetics of Precocious Puberty
E. A. Sazhenova, S. A. Vasilyev, L. V. Rychkova, E. E. Khramova, and I. N. Lebedev 1360

General Genetics

- Regulation of Uni-Strand and Dual-Strand piRNA Clusters in Germ and Somatic Tissues in *Drosophila melanogaster* under Control of *rhino*
P. A. Milyaeva, A. R. Lavrenov, I. V. Kuzmin, A. I. Kim, and L. N. Nefedova 1372

Plant Genetics

- Biological Diversity of Genes Encoding Wheat Defensin Homologues
M. P. Slezina, E. A. Istomina, and T. I. Odintsova 1382

Animal Genetics

- Population Genetic Structure in Polar Bears (*Ursus maritimus*) from the Russian Arctic Seas
P. A. Sorokin, E. Yu. Zvychnaynaya, E. A. Ivanov, I. A. Mizin, I. N. Mordvintsev, N. G. Platonov, A. I. Isachenko, R. E. Lazareva, and V. V. Rozhnov 1393

- Weak Population Genetic Structure of the Eurasian Crane *Grus grus* L.
E. A. Mudrik, Yu. M. Markin, K. A. Postelnykh, K. D. Kondrakova, T. A. Kashentseva, A. V. Shatokhina, P. A. Kazimirov, V. Yu. Ilyashenko, E. I. Ilyashenko, and D. V. Politov 1407

Human Genetics

- Methylation of the Retrotransposon LINE-1 Subfamilies in Chorionic Villi of Miscarriages
S. A. Vasilyev, V. V. Demeneva, E. N. Tolmacheva, S. A. Filatova, A. S. Zuev, A. S. Ushakova, O. Yu. Vasilyeva, I. V. Lushnikov, E. A. Sazhenova, T. V. Nikitina, and I. N. Lebedev 1419

- Transcriptomic Profiling of Placental Cells in Preeclampsia as an Effective Tool for Personalized Medicine
E. A. Trifonova, A. A. Babovskaya, A. A. Zarubin, V. N. Serebrova, M. M. Gavrilenko, M. G. Svarovskaya, E. V. Izhyokina, I. G. Kutsenko, and V. A. Stepanov 1427

Mathematical Models and Methods

- Description of the Divergence of Subpopulations in the Hierarchical System under the Analysis of Isonymy. III. Surname Inbreeding Coefficient
V. P. Passekov 1440

Short Communications

Polymorphism of Core and Symbiotically Specialized Genes
in the Polytypic Species of Nodule Bacteria

*N. A. Provorov, A. K. Kimeklis, E. S. Karasev, S. Khosid,
O. P. Onishchuk, O. N. Kurchak, and E. E. Andronov*

1457

Variations in Chromosome Synapsis at Meiotic Prophase I of Mole Voles
Ellobius tancrei Heterozygous for Robertsonian Translocations

*S. N. Matveevsky, Yu. F. Bogdanov, E. A. Lyapunova,
I. Yu. Bakloushinskaya, and O. L. Kolomiets*

1462

**ОБЗОРНЫЕ
И ТЕОРЕТИЧЕСКИЕ СТАТЬИ**

УДК 575.1, 616.89

ГЕНЕТИКА И ЭПИГЕНЕТИКА НЕСУИЦИДАЛЬНОГО САМОПОВРЕЖДАЮЩЕГО ПОВЕДЕНИЯ: НАРРАТИВНЫЙ ОБЗОР

© 2023 г. А. А. Кибитов^{1, 2, *}, Г. Э. Мазо²

¹Научный центр психического здоровья, Москва, 115522 Россия

²Национальный медицинский исследовательский центр психиатрии и неврологии им. В.М. Бехтерева,
Санкт-Петербург, 192019 Россия

*e-mail: andreykibitov18@gmail.com

Поступила в редакцию 30.05.2023 г.

После доработки 18.07.2023 г.

Принята к публикации 20.07.2023 г.

Несуицидальное самоповреждающее поведение (НССП), или селфхарм, широко распространено, особенно среди людей молодого возраста. Тем не менее, несмотря на актуальность данной проблемы, НССП исследовалось в основном с психологических и социальных позиций, в то время как биологические основы НССП изучались в значительно меньшем объеме. Цель данного нарративного обзора — анализ всей доступной на данный момент литературы по генетике и эпигенетике НССП, являющихся одними из наиболее перспективных областей биологических исследований. Обсуждены и проанализированы все этапы развития генетических исследований НССП: от близнецовых исследований и изучения самоповреждений в рамках наследственных заболеваний к поиску генов-кандидатов, полногеномным исследованиям и изучению эпигенетических механизмов и взаимодействия “ген—среда”. Продемонстрирован значимый вклад генетических и эпигенетических факторов в развитие НССП, обсуждены преимущества и ограничения проведенных исследований, намечены перспективы развития дальнейших исследований в данной области.

Ключевые слова: несуицидальное самоповреждающее поведение, селфхарм, генетика, эпигенетика, наследуемость.

DOI: 10.31857/S0016675823120032, **EDN:** QDPGGW

Несуицидальное самоповреждающее поведение (НССП), или селфхарм, — это умышленное самоповреждение тканей тела без суицидального замысла и с целями, не согласующимися с социальными и культурными нормами [1]. По данным метаанализа эпидемиологических исследований НССП 2014 г., хотя бы единожды наносили самоповреждения 17.2% подростков и 5.5% молодых взрослых [2]. Более поздние данные, полученные в ходе пандемии COVID-19, свидетельствуют о сохранении и даже некотором увеличении распространенности НССП. Так, по данным метаанализа Н. Cheng с соавт. 2023 г. распространенность самоповреждающего поведения (у подростков и у взрослых) составила 15.8% [3]. Таким образом, из-за широкой распространенности данного феномена изучение НССП становится чрезвычайно актуальной задачей.

Кроме того, несмотря на наличие в термине слова “несуицидальное”, НССП тесно связано с суицидальным поведением, хотя и остается несколько обособленным от него феноменом. В ли-

тературе неоднократно было продемонстрировано увеличение риска совершения суицидальной попытки при наличии в анамнезе НССП [4].

Несуицидальное самоповреждающее поведение является результатом комплексного взаимодействия психологических, социальных и биологических факторов. Хотя вклад психологических и социальных факторов в развитие НССП изучен сравнительно хорошо, биологическим механизмам, лежащим в основе данного феномена, было уделено значительно меньшее внимание [5].

Особый интерес среди исследований биологических механизмов НССП представляют генетические исследования, так как поиск валидных генетических и эпигенетических маркеров НССП даст возможность лучше понять нейробиологические механизмы, обуславливающие возникновение данного феномена, а также возможные различия в его характеристиках. Несмотря на это, на сегодняшний день опубликовано относительно небольшое количество генетических исследований НССП. Тем не менее среди имеющихся гене-

тических исследований самоповреждений можно выделить ряд парадигм, отражающих общие тенденции развития генетических исследований в медицине: от близнецовых исследований и изучения самоповреждений в рамках наследственных заболеваний к поиску генов-кандидатов, полногеномным исследованиям и изучению эпигенетических механизмов и взаимодействия “ген—среда”.

Нами не было обнаружено обзоров литературы на русском языке, посвященных генетическим и эпигенетическим исследованиям НССП. В связи с этим мы проанализировали доступную в международных базах литературу и включили ее в настоящий нарративный обзор, описывающий развитие и современное состояние исследований, посвященных вкладу генетических факторов в развитие НССП.

МАТЕРИАЛЫ И МЕТОДЫ

Поиск доступной литературы проводился в базах научной литературы Pubmed (<https://pubmed.ncbi.nlm.nih.gov>) и Российском Индексе Научного Цитирования (РИНЦ; https://elibrary.ru/project_risc.asp) в декабре 2022 г. Поиск в базе Pubmed проводили при помощи использования MeSH-термина “Self-Injurious Behavior” с подзаголовком “Genetics” (поисковое предложение: “Self-Injurious Behavior/genetics” [Mesh]). В базе данных РИНЦ поиск проводился при помощи поисковых предложений “самоповреждение ген”, “селфхарм ген”, “несуицидальное самоповреждение генетика”. Затем в полученных поисковых выдачах были отобраны оригинальные исследования, посвященные изучению наследственности и роли генетических и эпигенетических факторов в формировании самоповреждающего поведения.

В настоящий обзор не были включены исследования, посвященные изучению сугубо суицидальных попыток. Тем не менее во многих работах изучались самоповреждения в целом, без уточнения преследовали ли они суицидальную цель или нет. Для обозначения этого “широкого” феномена, изучавшегося в данных работах, далее по тексту используется термин “самоповреждение”. Если же в работе исследовалось именно несуйцидальное самоповреждающее поведение, то для его обозначения далее по тексту используется соответствующий термин. Кроме того, в рамках данного нарративного обзора не исследовалась в полном объеме литература о самоповреждающем поведении при различных наследственных заболеваниях, сопряженных с умственной отсталостью. Несмотря на существование ряда гипотез, согласно которым самоповреждения при умственной отсталости и НССП *per se* у людей без особенностей интеллектуального развития могут иметь некоторое количество общих “патогенетических звеньев”, все же разнородность данных феноменов не позволяет объединить их в

одно целое [6]. На основании анализа литературы отобранные статьи были сгруппированы в тематические группы, отражающие различные аспекты генетических исследований НССП, и приведены в хронологическом порядке, отражающем историю развития данной области исследований.

РЕЗУЛЬТАТЫ

Результаты поиска литературы

Всего в поисковой выдаче в базе Pubmed имелось 189 публикаций, в поисковой выдаче РИНЦ — 330 публикаций. После отбора согласно указанным выше критериям из базы Pubmed была отобрана 31 публикация. В поисковой выдаче РИНЦ нами не было обнаружено публикаций, подходящих под критерии. Были проанализированы полнотекстовые версии отобранной 31 публикации, после чего все они вошли в настоящий нарративный обзор (рис. 1). Информация о статьях, вошедших в данный обзор, приведена в табл. 1.

Самоповреждающее поведение при наследственных заболеваниях

Впервые поиск генетических причин развития самоповреждающего поведения был начат с изучения наследственных заболеваний. Известно, что при различных наследственных заболеваниях, ассоциированных с задержкой интеллектуального развития, может возникать самоповреждающее поведение. В число таких заболеваний входят синдром Ангельмана (делеция участка копии 15-й хромосомы), синдром Прадера—Вилли (делеция участка 15q11-13 отцовской копии 15-й хромосомы), синдром Корнелии де Ланге (мутации генов *NIPBL*, *SMC1A*), синдром Дауна (трисомия по 21-й хромосоме), синдром ломкой X-хромосомы (мультипликация повторяющегося участка CGG гена *FMRI*, расположенного в X-хромосоме), синдром Лоу (мутация гена *OCRL*), синдром Ретта (мутация гена *MECP2*), синдром Смит—Магенис (делеция участка 17-й хромосомы), туберозный склероз (мутация генов *TSC1*, *TSC2*), синдром Уильямса (делеция участка 7-й хромосомы) [7]. Однако самоповреждающее поведение при данных генетических заболеваниях, хотя и различается по частоте возникновения, является относительно неспецифическим и может рассматриваться как следствие задержки интеллектуального развития в целом [8].

Отдельно стоит выделить синдром Леша—Нихана — наследственное заболевание, связанное с мутацией гена *HPRT1*, приводящее к нарушению обмена пуриновых оснований и увеличению синтеза мочевой кислоты [9]. Синдром Леша—Нихана проявляется задержкой интеллектуального развития, хореоатетозом, спастическим церебральным параличом, а также крайне выраженным самоповреждающим поведением, начинающим

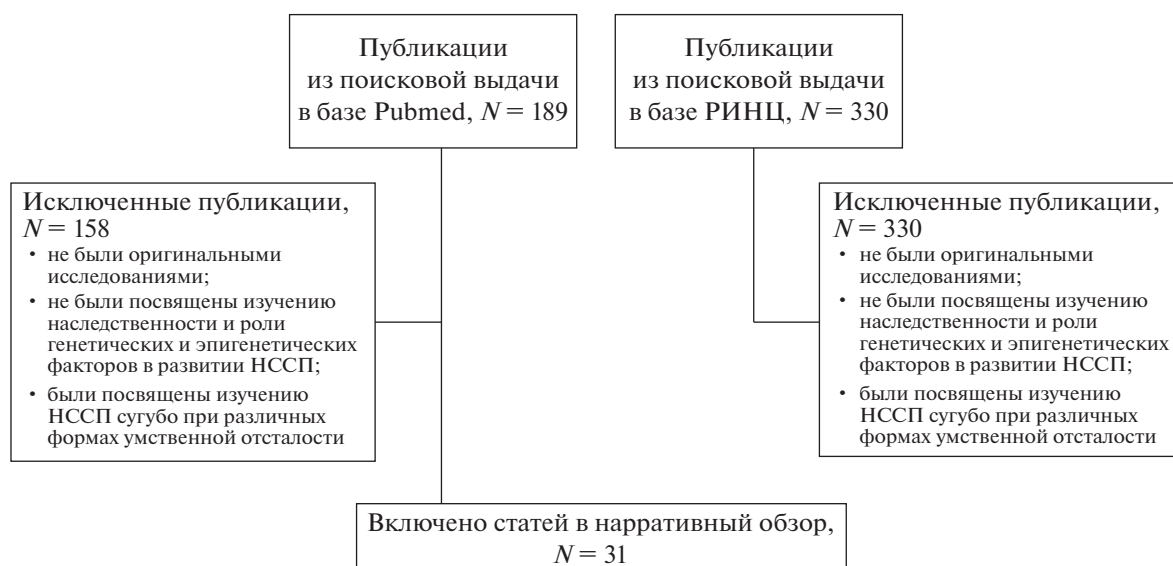


Рис. 1. Блок-схема процесса отбора статей для данного нарративного обзора.

проявляться в раннем детском возрасте и отличающимся многообразием и брутальностью его форм [10]. Выраженность самоповреждающего поведения при синдроме Леша–Нихана, отличающая его от других наследственных заболеваний, связанных с задержкой интеллектуального развития, дала возможность предположить, что на развитие самоповреждающего поведения самого по себе, как транснозологического клинического феномена, могут влиять определенные генетические факторы.

Одной из первых данную точку зрения высказала R.M. Palmour в опубликованном в 1983 г. обзоре исследований генетических факторов, влияющих на агрессивное и аутоагрессивное поведение. Автор предположила, что синдром Леша–Нихана может расцениваться как некая “модель”, наглядно демонстрирующая роль генетических аномалий в возникновении самоповреждающего поведения. По мнению автора, механизмом, обуславливающим развитие самоповреждающего поведения даже вне рамок наследственных заболеваний и задержки интеллектуального развития, может являться пуринергическая система. В частности, автор приводит результаты собственного исследования на крысах, продемонстрировавшие усиление самоповреждающего поведения при использовании антагонистов аденозиновых рецепторов [11].

Тем не менее роль пуринергической системы и ее генов в развитии НССП вне рамок синдрома Леша–Нихана и задержки интеллектуального развития не была изучена. Нами не было обнаружено опубликованных литературных данных по этому вопросу, хотя имеются сообщения о ее роли в развитии ассоциированных с НССП клинических феноменов: расстройств настроения [12], суицидальности [13], агрессивности [14].

Семейные и близнецовые исследования самоповреждающего поведения

Исследования совокупного генетического вклада в развитие какого-либо феномена начинается с установления факта его *наследуемости*, для чего проводятся семейные и близнецовые исследования. По нашим данным, на сегодняшний день не опубликовано полноценных эпидемиологических исследований НССП с семейным дизайном. Однако имеются некоторые исследования, посвященные изучению ассоциации НССП и семейной отягощенности по данному феномену.

В исследовании T.L. Deliberto и M.K. Nock, опубликованном в 2008 г., было показано, что риск развития НССП возрастает при наличии семейной отягощенности зависимостями от ПАВ, суицидальными идеями и высоким уровнем насилия. Примечательно, что значимых ассоциаций с семейной отягощенностью самоповреждающим поведением, а также с семейной отягощенностью конкретными психическими расстройствами обнаружено не было [15].

В более позднем исследовании S. Kandsperger с соавт., опубликованном в 2021 г., также не было обнаружено значимой ассоциации между риском НССП и семейной отягощенностью НССП и другими психическими расстройствами [16].

Тем не менее в отличие от семейных исследований НССП работ с близнецовым дизайном было проведено достаточное количество. Впервые попытка оценить наследуемость самоповреждающего поведения (иными словами, какая доля вариативности по данному феномену обуславливается генетическими, а не средовыми факторами) при помощи близнецового метода была предпринята

Таблица 1. Данные и основные результаты включенных в обзор исследований, представленные в хронологическом порядке

Группа исследований	Размер выборки	Основной вывод	Ссылка
Пуринергическая гипотеза НССП	—	Первая попытка выявить связь между пуринергической системой и НССП	[11]
Ген-кандидат	384 госпитальных пациентов с самоповреждениями и 346 здоровых контролей	Не было обнаружено ассоциации между <i>SLC6A4</i> STin2 VNTR и НССП	[24]
Близнецовое исследование	247 пар близнецов (183 — монозиготы, 64 — однополые дизиготы)	Значимой наследуемости для НССП обнаружено не было	[18]
Ген-кандидат	467 пациентов с НССП	Было обнаружено, что мужчины с сериновым вариантом полиморфизма 5-HT2c (Cys23Ser) были более импульсивны (по шкале I-V-E), чем мужчины-носители цистеинового варианта	[25]
Ген-кандидат	195 пациентов с депрессией (из них 46 (24%) с НССП)	T-аллель гена <i>GAB3</i> оказался ассоциирован с более высоким риском НССП	[26]
Близнецовое исследование	5533 близнеца (1879 полных пар) и 1202 сиблинга	Наследуемость НССП была оценена в 26%	[19]
Ген-кандидат	399 женщин с РПП (из них 156 (39%) — с НССП)	Носители высокофункциональных аллелей MAOAuVNTR чаще сообщали о НССП. Прямой ассоциации с 5HTTLPR обнаружено не было	[27]
Близнецовое исследование	6265 пар близнецов	Наследуемость мыслей о самоповреждениях и суициде у женщин была оценена в 74%, у мужчин — 45%	[20]
Ген-кандидат	269 женщин с булимией	Носители одновременно низкофункционального аллеля (7 повторов) по полиморфизму <i>DRD4</i> 7R и обоих аллелей <i>Val</i> по полиморфизму <i>COMT</i> rs4680 имеют большую тенденцию к самоповреждениям	[28]
Ген—среда	392 человека с криминальной историей (26% — с НССП)	Ассоциация между эмоциональной детской травмой и НССП наблюдалась только у гомозигот по аллелю <i>Val</i> полиморфизма Val66Met <i>BDNF</i>	[31]
Ген-кандидат	147 пациентов (из них 22 (15%) — с НССП)	Генотип <i>SS</i> по полиморфизму 5HTTLPR был ассоциирован с комплексом личностных черт (высокие уровни враждебности, поиска новизны и избегания опасности при оценке по TPQ), связанных с самоповреждениями	[29]
Близнецовое исследование	10678 близнецов	Генетические факторы определяют значительную часть вариативности по НССП (37% у мужчин и 59% у женщин)	[21]

Таблица 1. Продолжение

Группа исследований	Размер выборки	Основной вывод	Ссылка
Эпигенетика	281 пациент с ПРЛ	Метилирование промоторного участка гена <i>NR3C1</i> у пациентов с ПРЛ ассоциировано как с детской травмой, так и с тяжестью самоповреждающего поведения	[36]
Ген-кандидат	64 ребенка и подростка с РАС	Выраженность самоповреждающего поведения увеличивается с увеличением количества аллелей <i>La</i> полиморфизма 5HTTLPR	[30]
Ген—среда	271 ребенок и подросток (8.6% — с НССП)	Было обнаружено, что у детей, перенесших выраженный межличностный стресс, носительство короткого аллеля 5HTTLPR увеличивает риск НССП	[32]
Близнецовое исследование	6082 близнеца	Наследуемость НССП составила 54%	[22]
Ген-кандидат/Ген—среда	702 заключенных мужского пола (из них 145 (20.7%) — с НССП)	Носители высоко- и среднеактивных генотипов 5HTTLPR (<i>LaLa</i> , <i>LaLg</i> , <i>SLa</i>) по сравнению с носителями низкоактивных (<i>LgLg</i> , <i>SLg</i> , <i>SS</i>) чаще проявляли аутоагрессию. Эффектов взаимодействия генотипа по 5HTTLPR и детской травмы обнаружено не было	[33]
Анализ данных GWAS	6237 взрослых участников (из них НССП — у 3.2%)	PRS для депрессивных симптомов не предсказывал вариабельности по НССП	[40]
Анализ данных GWAS	307 родственников пациентов с БАП и 166 участников из группы контроля в возрасте от 12 до 21 года	PRS для БАП не различались между участниками с НССП и без	[42]
Ген-кандидат/Ген—среда	258 пациентов с аффективными расстройствами	Была обнаружена ассоциация с НССП (на уровне тенденции) гаплотипа GAT по полиморфизмам rs737865, rs6269, rs4633. Также не было обнаружено значимого взаимодействия этих полиморфизмов с детской травмой	[34]
Близнецовое исследование	9526 близнецов (5.6% — с НССП)	У мужчин наследуемость составила 33–36%, у женщин — 46%	[23]
Анализ данных GWAS	105222 человека из когорты UK Biobank	PRS для аутизма значимо ассоциирован с НССП (max $R^2 = 0.13\%$)	[44]
Анализ данных GWAS	126291 человек из когорты UK Biobank	PRS для употребления каннабиса значимо предсказывал вариабельность по НССП ($R^2 = 0.21\%$), как и PRS для депрессии ($R^2 = 1.1\%$)	[45]

Таблица 1. Окончание

Группа исследований	Размер выборки	Основной вывод	Ссылка
GWAS	3308 женщин	Был проведен GWAS для фенотипа “идеи о суициде и самоповреждении”. Отдельных значимых локусов обнаружено не было. PRS для депрессии значимо предсказывал часть вариативности по данному фенотипу ($R^2 = 0.3\%$)	[46]
GWAS	153 560 человек из когорты UK Biobank (4.37% хотя бы раз наносили самоповреждения (как с суицидальной, так и с несуйцидальной целью))	Была обнаружена обратная ассоциация самоповреждений с С-РБ. Генетические факторы, ассоциированные с повышением С-РБ, снижали риск самоповреждений	[47]
Анализ данных GWAS	125 925 человек из когорты UK Biobank (из них 5520 — самоповреждения)	Часть вариативности по самоповреждениям значимо обуславливают PRS по депрессии, СДВГ, шизофрении, алкогольной зависимости и употреблению каннабиса	[48]
GWAS	Когорта UK Biobank (156 700 человек) и собственная когорта (8703 человека)	Единственным значимым геном, ассоциированным как с идеями, так и с действиями, является ген <i>DCC</i> . Кроме того, PRS для мыслей о самоповреждениях, но не PRS для самих самоповреждений, значимо предсказывал часть вариативности НССП	[49]
Ген-кандидат/Ген—среда	269 китайских подростков мужского пола (51.2% — с НССП)	У носителей аллеля <i>T</i> по полиморфизму T941G и аллеля <i>Met</i> по полиморфизму Val158Met наличие детской травмы не приводило к повышению риска НССП, в отличие от носителей остальных генотипов	[35]
GWAS	4959 человек из когорты ALSPAC (из них 17% — НССП)	SNP-наследуемость НССП по результатам GWAS была оценена в 0.132	[51]
Эпигенетика	15 подростков с депрессией и НССП и 15 здоровых контролей	У пациентов из группы депрессии + НССП был повышен уровень метилирования промоторного участка гена <i>SIRT1</i> . Плазменная концентрация белка Sirt1 была также значимо ниже	[37]
Ген—среда/Анализ данных GWAS	74482 пациента с самоповреждениями	В когорте пациентов с самоповреждениями ни один SNP не был значимо ассоциирован со средовым показателем	[52]

в исследовании K.L. Jang с соавт., опубликованном в 1996 г. Авторы изучали наследуемость множества феноменов, ассоциированных с расстройствами личности, на выборке из 483 пар близнецов. Согласно результатам исследования, значимой наследуемости самих самоповреждающих действий обнаружено не было, однако мысли о самоповреждениях, напротив, имели значимую наследуемость, оцененную авторами в 36% [17].

В исследовании P.A. Vernon с соавт., опубликованном в 1999 г., изучалась наследуемость различных видов агрессии (в том числе и аутоагрессия — самоповреждающее поведение). В исследовании приняли участие 247 пар близнецов. В ходе анализа авторами не было обнаружено значимого влияния генетических факторов на самоповреждающее поведение [18].

В более позднем близнецовом исследовании НССП M.A. Distel с соавт., опубликованном в

2010 г. и посвященном наследуемости пограничного расстройства личности (ПРЛ), авторы показали, что наследуемость НССП, рассматриваемого авторами в контексте ПРЛ, составляет 26% [19]. Стоит отметить, что данное исследование включало значительно большее, чем в работе P.A. Verhulst с соавт., число участников: 5533 близнеца и 1202 сиблинга.

В 2012 г. было опубликовано исследование R.R. Althoff с соавт., в котором в качестве исследуемого фенотипа были выбраны “мысли о самоповреждениях и самоубийстве”. Авторы включили 6265 пар близнецов и оценили наследуемость данного фенотипа в 74% у женщин и в 45% у мужчин [20].

В близнецовом исследовании НССП и суицидальных мыслей, опубликованном D.F. Maciejewski с соавт. в 2014 г., проведенном с участием 10678 близнецов из Австралийского регистра близнецов, было обнаружено, что генетические факторы обуславливают значительную часть вариативности по НССП (37% у мужчин и 59% у женщин), при этом была обнаружена значительная генетическая корреляция между НССП и суицидальными идеями [21].

В 2015 г. было опубликовано исследование L.R. Few с соавт., также проведенное с участием близнецов из Австралийского регистра (6082 однополых пар). Работа была посвящена изучению различий в наследуемости НССП и суицидальных попыток у близнецов, кон- и дискондартных в отношении раннего начала употребления алкоголя (до 15 лет). Была подтверждена наследуемость НССП (составившая 54%), а также обнаружена генетическая корреляция между НССП и ранним началом употребления алкоголя [22].

Последнее на сегодняшний день близнецовое исследование НССП было опубликовано L.S. Richmond-Rakerd с соавт. в 2019 г. Данное исследование также было проведено с участием близнецов из Австралийского регистра (9526 близнецов) и было посвящено изучению не только НССП и суицидальности, но и влияния детской травмы. Наследуемость НССП была оценена в 33–36% у мужчин и в 46% у женщин. Кроме того, была обнаружена генетическая корреляция между НССП (вместе с суицидальностью) и подверженностью детской травме. Авторы отмечают, что, по-видимому, общие генетические факторы влияют как на “восприимчивость” к травматичному детскому опыту, так и на развитие самоповреждений [23].

Таким образом, несмотря на некоторые данные об отсутствии ассоциации между НССП и семейной отягощенностью по данному феномену, результаты масштабных близнецовых исследований демонстрируют, что генетические факторы все же играют важную роль в предрасположенности к развитию НССП.

Исследования генов-кандидатов

Следующим этапом развития генетических исследований являлись молекулярно-генетические исследования, позволившие генотипировать выборку по заранее определенным генетическим полиморфизмам. Такой подход получил название “исследование генов-кандидатов”, так как предполагает предварительный выбор на основе уже имеющихся гипотез отдельных генов и поиск различий по исследуемым характеристикам между носителями различных генотипов по полиморфизмам данных генов. На сегодняшний день опубликовано относительно небольшое количество исследований ассоциации полиморфизмов различных генов-кандидатов с НССП. Данные исследования были сконцентрированы, главным образом, на полиморфизмах генов нейромедиаторных систем, при этом в части работ оценивались самоповреждения в целом (без разделения на суицидальные попытки и НССП).

Первым, согласно нашим данным, исследованием ассоциации конкретного полиморфизма с самоповреждениями является опубликованная в 1997 г. работа J. Evans с соавт., посвященная изучению влияния полиморфизма STin2 VNTR гена серотонинового транспортера *SLC6A4*. Данное исследование было проведено при участии 384 госпитальных пациентов с самоповреждениями (без разделения на суицидальные и несуйцидальные) и 346 здоровых контрольных испытуемых. Авторами не было обнаружено значимой ассоциации данного полиморфизма с самоповреждениями, однако среди пациентов выраженность тревоги была выше у носителей более коротких вариантов (STin2.9, STin2.10) [24].

В 2000 г. J. Evans с соавт. опубликовали исследование, посвященное влиянию полиморфизма гена триптофангидроксилазы *TPH* intron7 и полиморфизма Cys23Ser гена серотонинового рецептора 5-HT_{2c} *HTR2C* на импульсивность у пациентов с самоповреждениями (467 пациентов). Было обнаружено, что мужчины с сериновым вариантом полиморфизма Cys23Ser *HTR2C* были более импульсивны, чем мужчины-носители цистеинового варианта. Значимого влияния полиморфизма intron7 *TPH* обнаружено не было [25].

P.R. Joyce с соавт. в исследовании, опубликованном в 2006 г., на выборке из 195 пациентов с депрессией, показали наличие значимой ассоциации аллеля T по полиморфизму rs5443 гена *GNB3* (кодирует G-белок β_3 , участвующий во внутриклеточной передаче сигнала) с более высоким риском НССП. Авторы выбрали данный полиморфизм для изучения в связи с данными о его ассоциации с ответом на антидепрессивную терапию у пациентов с депрессией [26].

В 2011 г. H. Steiger с соавт. опубликовали исследование, посвященное влиянию полиморфиз-

ма 5HTTLPR гена серотонинового транспортера *SLC6A4* и полиморфизма MAOAuVNTR гена моноаминоксидазы *MAOA* на НССП у пациенток с расстройствами пищевого поведения (РПП). Всего было включено 399 женщин с различными РПП. Авторы обнаружили, что носители высокофункциональных аллелей MAOAuVNTR чаще сообщали о НССП. Прямой ассоциации с 5HTTLPR обнаружено не было [27].

Ассоциации генетических полиморфизмов с НССП у пациенток с РПП оценивались и в исследовании L. Thaler с соавт. 2012 г. На выборке из 269 женщин с нервной булимией изучался ряд полиморфизмов генов дофаминергической системы: *DRD2* Taq1A (rs1800497), *DRD4* 7R, *COMT* Val158Met (rs4680), *DAT1* VNTR. Значимых влияний каждого из данных полиморфизмов по отдельности обнаружено не было, однако было показано, что носители одновременно низкофункционального аллеля (7 повторов) по полиморфизму *DRD4* 7R и обоих аллелей Val по полиморфизму *COMT* rs4680 имеют большую тенденцию к самоповреждениям [28].

В исследовании L. Dell'Osso с соавт., опубликованном в 2013 г., влияние генетических полиморфизмов оценивалось не только на само НССП, но и на ассоциированные с ним характерологические и личностные черты. В исследование были включены 147 пациентов и были изучены полиморфизмы 5HTTLPR гена *SLC6A4*, 218C>A гена *TPH* и Val158Met гена *COMT*. Значимого влияния данных полиморфизмов на НССП обнаружено не было. Однако генотип *SS* по полиморфизму 5HTTLPR был ассоциирован с комплексом личностных черт, связанных, по мнению авторов, с самоповреждениями: высокие уровни враждебности, поиска новизны и избегания опасности [29].

A. Kolevzon с соавт. в опубликованном в 2014 г. исследовании изучали влияние полиморфизма 5HTTLPR гена *SLC6A4* на НССП у детей и подростков с расстройствами аутистического спектра (РАС). В исследование было включено 64 человека в возрасте до 18 лет с диагнозом РАС. Было обнаружено, что выраженность самоповреждающего поведения увеличивается с увеличением количества аллелей *La* по полиморфизму 5HTTLPR [30].

Исследования взаимодействия “ген–среда”

В некоторых исследованиях влияние генетических полиморфизмов на НССП оценивалось не только изолированно, но и с учетом их взаимодействия с различными средовыми факторами, так как согласно парадигме взаимодействия “ген–среда” (gene-environment interaction, GxE) генетические факторы могут не только сами по себе влиять на вероятность наличия определен-

ного фенотипа, но и модифицировать влияние ассоциированных с ним средовых факторов, иными словами изменять “чувствительность” к факторам риска. В контексте изучения НССП наиболее исследованным средовым фактором являлась пережитая детская травма.

Так, K. Bresin с соавт. в опубликованном в 2013 г. исследовании влияния детской травмы и полиморфизма Val66Met гена мозгового нейротрофического фактора *BDNF* на риск развития НССП, проведенном на выборке из 392 человек, показали, что ассоциация между эмоциональной детской травмой и НССП наблюдалась только у гомозигот по аллелю Val полиморфизма Val66Met *BDNF*, при этом значимого влияния самого полиморфизма обнаружено не было [31].

Аналогичным образом, согласно данным исследования B.L. Hankin с соавт. 2015 г., проведенном на выборке из 271 ребенка и подростка, ассоциация между выраженным межличностным стрессом и повышением риска НССП в большей степени проявлялась у носителей короткого аллеля по полиморфизму 5HTTLPR гена *SLC6A4* [32].

Влияние данного полиморфизма и детской травмы на аутоагрессивное поведение было также исследовано E. Gorodetsky с соавт. на выборке из 702 заключенных мужского пола. Было обнаружено, что носители высоко- и среднеактивных генотипов 5HTTLPR (*LaLa*, *LaLg*, *SLa*) по сравнению с носителями низкоактивных (*LgLg*, *SLg*, *SS*) чаще проявляли аутоагрессию. При этом эффектов взаимодействия генотипа по 5HTTLPR и детской травмы обнаружено не было [33].

A. Bernegger с соавт. в работе 2018 г. на выборке из 258 пациентов с аффективными расстройствами изучали влияние на НССП детской травмы и пяти полиморфизмов гена *COMT*: rs4633, rs4680 (Val158Met), rs737865, rs6269, rs165599. Влияния отдельных полиморфизмов на НССП, как и эффектов взаимодействия этих полиморфизмов с детской травмой, обнаружено не было. Однако была обнаружена ассоциация с НССП (на уровне тенденции) гаплотипа GAT по полиморфизмам rs737865, rs6269, rs4633 [34].

Влияние на НССП полиморфизма rs4680 (Val158Met) *COMT*, а также полиморфизма T941G *MAOA* и детской травмы изучалось на выборке из 269 подростков мужского пола в исследовании Y. Gao с соавт., опубликованном в 2021 г. Прямых ассоциаций полиморфизмов T941G *MAOA* и Val158Met *COMT* с НССП обнаружено не было. Однако у носителей аллеля *T* по полиморфизму T941G и аллеля *Met* по полиморфизму Val158Met наличие детской травмы не приводило к повышению риска НССП, в отличие от носителей остальных генотипов [35].

Таким образом, результаты исследований генотипов-кандидатов представляются противоречивыми

и сделать однозначный вывод о значимом влиянии исследуемых полиморфизмов затруднительно. Однако некоторые полиморфизмы, например 5HTTLPR гена *SLC6A4*, Val66Met гена *BDNF*, могут обуславливать свое влияние на риск развития НССП путем взаимодействия со средовыми факторами риска, главным образом с детской травмой. Эти данные позволяют рассматривать НССП как продукт комплексного взаимодействия “ген–среда”.

Эпигенетические исследования самоповреждений

Парадигма взаимодействия “ген–среда” предполагает не только уже рассмотренное выше влияние “гена на среду” – модулирующее воздействие генетических факторов на восприимчивость к средовым факторам риска, но и влияние “среды на ген” – воздействие средовых факторов на экспрессию генов посредством эпигенетических механизмов. По нашим данным, на сегодняшний день опубликовано всего два эпигенетических исследования самоповреждений.

В исследовании А. Martin-Blanco с соавт., опубликованном в 2014 г., на выборке из 281 пациента с ПРЛ было показано, что в образцах периферической крови метилирование промоторного участка гена *NR3C1* (кодирует рецептор к глюкокортикоидам) у пациентов с ПРЛ ассоциировано как с детской травмой, так и с тяжестью самоповреждающего поведения (как НССП, так и суицидных попыток). Таким образом, изменение экспрессии данного гена может являться одним из возможных механизмов, обуславливающих связь детской травмы и самоповреждений [36].

В 2021 г. были опубликованы результаты исследования L. Wang с соавт., посвященного изучению ассоциации метилирования промоторного участка гена *SIRT1*, ассоциированного с серотонинергической передачей, и депрессии с НССП. Исследование было проведено на небольшой выборке из 15 подростков с депрессией и имевшимся в ее рамках НССП и 15 здоровых контролей. Авторы обнаружили, что в группе пациентов с депрессией и НССП уровень метилирования промоторного участка *SIRT1* выше (на основании анализа образцов периферической крови). Однако в связи с особенностями дизайна исследования сделать вывод об ассоциации данной находки с НССП как трансоэологическим феноменом, вне связи с депрессией, затруднительно [37].

Результаты данных исследований позволяют предположить определенный вклад эпигенетических механизмов в опосредование взаимодействий “ген–среда”, приводящих к развитию НССП. Тем не менее изучение конкретных эпигенетических механизмов и локусов риска является задачей для будущих исследований.

Полногеномные исследования самоповреждений

Следующий этап развития генетических исследований – полногеномные ассоциативные исследования (Genome-Wide Association Study, GWAS), позволяющие комплексно оценивать ассоциированные с интересующим фенотипом генетические полиморфизмы сразу во всем геноме, а не только изучать вклад отдельных генов-кандидатов, определенных на основании предварительных допущений и гипотез. Одним из важнейших практических результатов GWAS является вычисление полигенных шкал риска (Polygenic Risk Score, PRS). PRS рассчитываются путем умножения аллелей риска (показавших по результатам GWAS значимый вклад в увеличение риска наличия интересующего фенотипа), которые несет человек, на величину эффекта каждого варианта, а затем их суммирования. Иными словами, PRS позволяет оценить какой риск наличия интересующего фенотипа имеет человек, исходя из его индивидуальных генетических особенностей [38]. Помимо этого PRS могут применяться для оценки наличия или отсутствия общих генетических факторов для двух фенотипов. Если PRS по одному фенотипу значимо предсказывает вариабельность по другому фенотипу, можно предположить наличие общей “генетической основы”. Первые исследования НССП с использованием данных GWAS были выполнены именно в таком ключе и были посвящены возможности применения PRS для психических расстройств (по результатам соответствующих GWAS) для оценки риска НССП.

Так, в исследовании 2017 г. D.F. Maciejewski с соавт. попробовали применить PRS для депрессивных симптомов [39] для оценки риска НССП на выборке из 6237 человек. Было обнаружено, что данные PRS для депрессивных симптомов значимо не предсказывали вариабельности по НССП [40].

В опубликованном в том же году исследовании Н.С. Wilcox с соавт. на выборке из 307 родственников пациентов с БАП и 166 контролей было показано, что PRS для БАП [41] не различались между участниками с НССП и без [42].

Тем не менее V. Warriar с соавт. в 2019 г. на выборке из 105222 человек из когорты UK Biobank показали, что PRS для аутизма [43] значимо ассоциированы с самоповреждениями ($\max R^2 = 0.13\%$) и что между двумя данными фенотипами существует соответствующая генетическая корреляция ($r_g = 0.48$). Необходимо отметить, что для фенотипирования по самоповреждениям в когорте UK Biobank был использован вопрос: “Вы когда-либо нанесли себе умышленно самоповреждения, вне зависимости от того, хотели ли вы покончить с собой или нет? (Have you deliberately harmed yourself, whether or not you meant to end your life?)”. Подобный подход к фенотипированию не позволяет

разграничить самоповреждения с несуицидальными и суицидальными целями, что является ограничением этого и других исследований на базе этой когорты [44].

К. Hodgson с соавт. в своем исследовании, опубликованном в 2020 г., также показали на выборке из когорты UK Biobank (126291 человек), что вариабельность по самоповреждениям (как с несуицидальными, так и с суицидальными целями) значимо предсказывают PRS для депрессии ($R^2 = 1.1\%$) и PRS для употребления каннабиса ($R^2 = 0.21\%$) [45].

В 2020 г. были опубликованы результаты сразу нескольких полногеномных исследований фенотипов, имеющих отношения к НССП. Н. Shen с соавт. провели GWAS для фенотипа “мысли о суициде и самоповреждениях” на выборке, состоящей из 3308 женщин латиноамериканского происхождения. Отдельных значимых локусов обнаружено не было, но PRS для депрессии значимо предсказывал часть вариабельности по данному фенотипу ($R^2 = 0.3\%$) [46].

А.Е. Russell с соавт. провели GWAS по самоповреждениям (как несуицидальным, так и суицидальным) на выборке из 153560 человек из когорты UK Biobank и изучали генетическую корреляцию с факторами воспаления. Ни одного локуса с полногеномной значимостью для самоповреждений обнаружено не было. Ассоциации между генетическими факторами, связанными с интерлейкином-6 (когорта SardiNIA) и самоповреждениями, обнаружено не было. При этом была обнаружена обратная ассоциация с С-реактивным белком: генетические факторы, ассоциированные с повышением С-РБ (когорта 1000G), снижали риск самоповреждений [47].

К.Х. Lim с соавт. на выборке из 125925 человек из когорты UK Biobank использовали данные 24 GWAS, чтобы выявить потенциальные факторы риска самоповреждающего поведения. Было обнаружено, что с риском самоповреждений ассоциированы PRS для депрессии, синдрома дефицита внимания и гиперактивности, биполярного аффективного расстройства, шизофрении, алкогольной зависимости и употребления каннабиса [48].

В исследовании А.И. Campos с соавт. были исследованы когорта UK Biobank (156700 человек) и собственная когорта взрослых жителей Австралии (8703 человека) [49]. Был проведен GWAS как по самим самоповреждениям (как несуицидальным, так и суицидальным), так и по мыслям о них. Было обнаружено, что с мыслями о самоповреждениях ассоциированы полиморфизмы rs4865733 и rs7721698, а с самими самоповреждениями — rs567805973. В ходе данной работы при использовании тестов ассоциаций с генами (gene-based association tests) был обнаружен ряд генов, ассоциированных с идеями о самоповреждениях (*SYT14*,

RPP14, *FAM172A*, *SEMA3D*, *DCC*, *DDX27*, *ZNFXT1*) и с самими самоповреждениями (*LINGO2*, *DCC*, *FBXO27*, *WRB*). Единственным значимым геном, ассоциированным как с идеями, так и с действиями, являлся ген *DCC*, кодирующий рецептор к нетрину-1 — белок, играющий важную роль в формировании префронтальной коры [50]. Кроме того, при анализе собственной когорты было показано, что PRS для мыслей о самоповреждениях, но не PRS для самих самоповреждений, значимо предсказывал часть вариабельности именно по несуицидальному самоповреждающему поведению [49].

Единственный на сегодняшний день GWAS именно НССП был проведен А.Е. Russell с соавт. на выборке из 4959 человек из когорты Avon Longitudinal Study of Parents and Children (ALSPAC). Авторами не были приведены данные об ассоциациях конкретных полиморфизмов с НССП, однако общая SNP-наследуемость (single nucleotide polymorphism, однонуклеотидный полиморфизм) НССП была оценена в 0.132 (13.2% вариабельности обуславливаются генетическими факторами). Также была показана значимая ассоциация PRS для коэффициента интеллекта, употребления алкоголя, образования (только для НССП, но не для суицидальных попыток); PRS для попытки суицида, СДВГ, РАС, нейротизма (и для НССП, и для суицидальных попыток). PRS для депрессии был сильнее ассоциирован с суицидальными самоповреждениями, но не с НССП [51].

Наконец, в недавнем исследовании С. Li с соавт. на полногеномном уровне изучалось взаимодействие “ген—среда”. Была проведена оценка влияния совокупности средовых факторов (курение, потребление алкоголя, физическая активность, средний доход), объединенных в общий “средовой показатель” (environmental score), на различные генетические локусы риска самоповреждений (как суицидальных, так и не суицидальных) на когорте UK Biobank (74482 участника). Авторы не обнаружили ни одного локуса, который бы значимо взаимодействовал с данным средовым показателем [52]. Можно предположить, что в развитии НССП взаимодействие каких-либо генетических факторов с исследуемыми средовыми факторами не играет существенной роли.

Таким образом применение PRS позволило выявить генетическую корреляцию между самоповреждениями и рядом психических расстройств. Тем не менее практически во всех GWAS самоповреждений использовался слишком широкий фенотип, включавший, помимо НССП, также суицидальные попытки. Соответственно возникает проблема оценки релевантности полученных результатов в отношении именно несуицидальных самоповреждений, которые представляют собой отдельный, не связанный прямо с суицидальными

ми попытками, феномен. В связи с этим все еще существует необходимость проведения полногеномных исследований НССП на достаточно больших выборках, что позволит выявить конкретные локусы риска развития данного феномена.

ОБСУЖДЕНИЕ

Опубликованные на сегодняшний день данные литературы подтверждают предположение о значимом вкладе генетических факторов в развитие несуицидального самоповреждающего поведения. Данные близнецовых исследований демонстрируют высокие уровни наследуемости (30–70%); кроме того, в GWAS НССП A.E. Russel с соавт. показан достаточный уровень SNP-наследуемости (13.2%). Результаты полногеномных исследований и исследований генов-кандидатов позволяют предположить наличие определенных локусов риска НССП (например, гена рецептора к нетрину-1 DCC), а также “генетическое перекрытие” НССП с различными психиатрическими расстройствами.

При этом, помимо прямого влияния, генетические факторы взаимодействуют со средовыми, формируя комплексное взаимодействие “ген–среда”, обуславливающее развитие НССП. При этом, с одной стороны, генетические факторы могут влиять на чувствительность к средовым факторам, с другой стороны, экспрессия самих генов может изменяться под воздействием средовых факторов посредством эпигенетических механизмов.

Тем не менее сравнительно небольшое количество генетических исследований НССП, а также неоднородность использовавшихся фенотипов не дают возможности выделить конкретные и однозначные локусы риска НССП. Необходимы дальнейшие молекулярно-генетические, в том числе полногеномные, и эпигенетические исследования с использованием узкого фенотипа НССП (без объединения с суицидными попытками) для изучения конкретных генетических и эпигенетических факторов, ассоциированных с НССП.

ЗАКЛЮЧЕНИЕ

Нами был проведен первый, по нашим данным, нарративный обзор генетических и эпигенетических исследований НССП на русском языке. Несмотря на небольшой объем литературы по данному вопросу, опубликованные исследования демонстрируют значительную роль генетических и эпигенетических механизмов, а также взаимодействия “ген–среда” в развитии НССП. Были очерчены направления для будущих исследований, которые позволят с большей степенью достоверности выявить конкретные локусы риска НССП, что в перспективе может быть полезным для понимания патогенетических механизмов развития

данного феномена и выявления потенциальных специфических фармакотерапевтических мишеней.

Работа поддержана грантом РНФ № 23-15-00347.

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

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. International Society for the Study of Self-Injury. What is Non-Suicidal Self-Injury? 2022. <https://www.itriples.org/what-is-nssi>
2. Swannell S.V., Martin G.E., Page A. et al. Prevalence of nonsuicidal self-injury in nonclinical samples: Systematic review, meta-analysis and meta-regression // *Suicide Life-Threatening Behav.* 2014. V. 44. № 3. P. 273–303. <https://doi.org/10.1111/sltb.12070>
3. Cheng H., Wang D., Wang L. et al. Global prevalence of self-harm during the COVID-19 pandemic: A systematic review and meta-analysis // *BMC Psychol.* 2023. V. 11. № 1. P. 149. <https://doi.org/10.1186/s40359-023-01181-8>
4. Grandclerc S., De Labrouhe D., Spodenkiewicz M. et al. Relations between nonsuicidal self-injury and suicidal behavior in adolescence: A systematic review // *PLoS One.* 2016. V. 11. № 4. <https://doi.org/10.1371/journal.pone.0153760>
5. Kaess M., Hooley J.M., Klimes-Dougan B. et al. Advancing a temporal framework for understanding the biology of nonsuicidal self-injury: An expert review // *Neurosci. Biobehav. Rev.* 2021. V. 130. P. 228–239. <https://doi.org/10.1016/j.neubiorev.2021.08.022>
6. Oliver C., Richards C. Self-injurious behaviour in people with intellectual disability // *Curr. Opin. Psychiatry.* 2010. V. 23. № 5. P. 412–416. <https://doi.org/10.1097/YCO.0b013e32833cfb80>
7. Huismans S., Mulder P., Kuijk J. et al. Self-injurious behavior // *Neurosci. Biobehav. Rev.* 2018. V. 84. P. 483–491. <https://doi.org/10.1016/j.neubiorev.2017.02.027>
8. Arron K., Oliver C., Moss J. et al. The prevalence and phenomenology of self-injurious and aggressive behaviour in genetic syndromes // *J. Intellect. Disabil. Res.* 2011. V. 55. № 2. P. 109–120. <https://doi.org/10.1111/j.1365-2788.2010.01337.x>
9. Lesch M., Nyhan W.L. A familial disorder of uric acid metabolism and central nervous system function // *Am. J. Med.* 1964. V. 36. № 4. P. 561–570. [https://doi.org/10.1016/0002-9343\(64\)90104-4](https://doi.org/10.1016/0002-9343(64)90104-4)
10. Harris J.C. Lesch–Nyhan syndrome and its variants: Examining the behavioral and neurocognitive phenotype // *Curr. Opin. Psychiatry.* 2018. V. 31. № 2. P. 96–102. <https://doi.org/10.1097/YCO.0000000000000388>
11. Palmour R.M. Genetic models for the study of aggressive behavior // *Prog. Neuro-Psychopharmacology Biol. Psychiatry.* 1983. V. 7. № 4–6. P. 513–517. [https://doi.org/10.1016/0278-5846\(83\)90018-0](https://doi.org/10.1016/0278-5846(83)90018-0)

12. Ortiz R., Ulrich H., Zarate C.A., Machado-Vieira R. Purinergic system dysfunction in mood disorders: A key target for developing improved therapeutics // *Prog. Neuro-Psychopharmacology Biol. Psychiatry*. 2015. V. 57. P. 117–131. <https://doi.org/10.1016/j.pnpbp.2014.10.016>
13. Gonçalves M.C.B., Corrêa-Velloso J., Naaldijk Y. et al. Purinergic modulation of pathways associated to suicidal behavior // *Mol. Psychiatry*. 2020. V. 25. № 3. P. 514–516. <https://doi.org/10.1038/s41380-018-0088-3>
14. Mrug S., Mrug M. Uric acid excretion predicts increased aggression in urban adolescents // *Physiol. Behav.* 2016. V. 163. P. 144–148. <https://doi.org/10.1016/j.physbeh.2016.05.014>
15. Deliberto T.L., Nock M.K. An exploratory study of correlates, onset, and offset of non-suicidal self-injury // *Arch. Suicide Res.* 2008. V. 12. № 3. P. 219–231. <https://doi.org/10.1080/13811110802101096>
16. Kandsperger S., Jarvers I., Ecker A. et al. Emotional reactivity and family-related factors associated with self-injurious behavior in adolescents presenting to a child and adolescent psychiatric emergency service // *Front. Psychiatry*. 2021. V. 12. <https://doi.org/10.3389/fpsy.2021.634346>
17. Jang K.L., Livesley W.J., Vernon P.A., Jackson D.N. Heritability of personality disorder traits: A twin study // *Acta Psychiatr. Scand.* 1996. V. 94. № 6. P. 438–444. <https://doi.org/10.1111/j.1600-0447.1996.tb09887.x>
18. Vernon P.A., McCarthy J.M., Johnson A.M. et al. Individual differences in multiple dimensions of aggression: A univariate and multivariate genetic analysis // *Twin Res.* 1999. V. 2. № 1. P. 16–21. <https://doi.org/10.1375/136905299320566068>
19. Distel M.A., Willemsen G., Ligthart L. et al. Genetic covariance structure of the four main features of borderline personality disorder // *J. Pers. Disord.* 2010. V. 24. № 4. P. 427–444. <https://doi.org/10.1521/pedi.2010.24.4.427>
20. Althoff R.R., Hudziak J.J., Willemsen G. et al. Genetic and environmental contributions to self-reported thoughts of self-harm and suicide // *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* 2012. V. 159 B. № 1. P. 120–127. <https://doi.org/10.1002/ajmg.b.32010>
21. Maciejewski D.F., Creemers H.E., Lynskey M.T. et al. Overlapping genetic and environmental influences on nonsuicidal self-injury and suicidal ideation // *JAMA Psychiatry*. 2014. V. 71. № 6. P. 699. <https://doi.org/10.1001/jamapsychiatry.2014.89>
22. Few L.R., Werner K.B., Sartor C.E. et al. Early onset alcohol use and self-harm: A discordant twin analysis // *Alcohol. Clin. Exp. Res.* 2015. V. 39. № 11. P. 2134–2142. <https://doi.org/10.1111/acer.12889>
23. Richmond-Rakerd L.S., Trull T.J., Gizer I.R. et al. Common genetic contributions to high-risk trauma exposure and self-injurious thoughts and behaviors // *Psychol. Med.* 2019. V. 49. № 3. P. 421–430. <https://doi.org/10.1017/S0033291718001034>
24. Evans J., Battersby S., Ogilvie A.D. et al. Association of short alleles of a VNTR of the serotonin transporter gene with anxiety symptoms in patients presenting after deliberate self-harm // *Neuropharmacology*. 1997. V. 36. № 4–5. P. 439–443. [https://doi.org/10.1016/S0028-3908\(97\)00027-0](https://doi.org/10.1016/S0028-3908(97)00027-0)
25. Evans J., Reeves B., Platt H. et al. Impulsiveness, serotonin genes and repetition of deliberate self-harm (DSH) // *Psychol. Med.* 2000. V. 30. № 6. P. 1327–1334. <https://doi.org/10.1017/S0033291799002822>
26. Joyce P.R., McKenzie J.M., Mulder R.T. et al. Genetic, developmental and personality correlates of self-mutilation in depressed patients // *Aust. N. Z. J. Psychiatry*. 2006. V. 40. № 3. P. 225–229. <https://doi.org/10.1111/j.1440-1614.2006.01778.x>
27. Steiger H., Fichter M., Bruce K.R. et al. Molecular-genetic correlates of self-harming behaviors in eating-disordered women: Findings from a combined Canadian-German sample // *Prog. Neuro-Psychopharmacology Biol. Psychiatry*. 2011. V. 35. № 1. P. 102–106. <https://doi.org/10.1016/j.pnpbp.2010.09.012>
28. Thaler L., Groleau P., Badawi G. et al. Epistatic interactions implicating dopaminergic genes in bulimia nervosa (BN): Relationships to eating- and personality-related psychopathology // *Prog. Neuro-Psychopharmacology Biol. Psychiatry*. 2012. V. 39. № 1. P. 120–128. <https://doi.org/10.1016/j.pnpbp.2012.05.019>
29. Dell'Osso L., Mandelli L., Carlini M. et al. Temperamental and genetic predictors of suicide attempt and self-mutilation // *Neuropsychobiology*. 2013. V. 68. № 4. P. 250–257. <https://doi.org/10.1159/000356228>
30. Kolevzon A., Lim T., Schmeidler J. et al. Self-injury in autism spectrum disorder: An effect of serotonin transporter gene promoter variants // *Psychiatry Res.* 2014. V. 220. № 3. P. 987–990. <https://doi.org/10.1016/j.psychres.2014.09.018>
31. Bresin K., Sima Finy M., Verona E. Childhood emotional environment and self-injurious behaviors: The moderating role of the BDNF Val66Met polymorphism // *J. Affect. Disord.* 2013. V. 150. № 2. P. 594–600. <https://doi.org/10.1016/j.jad.2013.01.050>
32. Hankin B.L., Barrocas A.L., Young J.F. et al. 5-HTTLPR × interpersonal stress interaction and nonsuicidal self-injury in general community sample of youth // *Psychiatry Res.* 2015. V. 225. № 3. P. 609–612. <https://doi.org/10.1016/j.psychres.2014.11.037>
33. Gorodetsky E., Carli V., Sarchiapone M. et al. Predictors for self-directed aggression in Italian prisoners include externalizing behaviors, childhood trauma and the serotonin transporter gene polymorphism 5-HTTLPR // *Genes, Brain Behav.* 2016. V. 15. № 5. P. 465–473. <https://doi.org/10.1111/gbb.12293>
34. Bernegger A., Kienesberger K., Carlberg L. et al. The impact of COMT and childhood maltreatment on suicidal behaviour in affective disorders // *Sci. Rep.* 2018. V. 8. № 1. P. 1–9. <https://doi.org/10.1038/s41598-017-19040-z>
35. Gao Y., Xiong Y., Liu X., Wang H. The effects of childhood maltreatment on non-suicidal self-injury in male adolescents: The moderating roles of the monoamine oxidase a (maoa) gene and the catechol-o-methyltransferase (comt) gene // *Int. J. Environ. Res. Public Health*. 2021. V. 18. № 5. P. 1–13. <https://doi.org/10.3390/ijerph18052598>
36. Martín-Blanco A., Ferrer M., Soler J. et al. Association between methylation of the glucocorticoid receptor gene, childhood maltreatment, and clinical severity in borderline personality disorder // *J. Psychiatr. Res.* 2014. V. 57. № 1. P. 34–40. <https://doi.org/10.1016/j.jpsychires.2014.06.011>

37. Wang L., Zheng D., Liu L. et al. Relationship between SIRT1 gene and adolescent depressive disorder with nonsuicidal self-injury behavior: Based on gene methylation and mRNA expression // *Medicine (Baltimore)*. 2021. V. 100. № 31. <https://doi.org/10.1097/MD.00000000000026747>
38. Kullo I.J., Lewis C.M., Inouye M. et al. Polygenic scores in biomedical research // *Nat. Rev. Genet.* 2022. V. 23. № 9. P. 524–532. <https://doi.org/10.1038/s41576-022-00470-z>
39. Hek K., Demirkan A., Lahti J. et al. A genome-wide association study of depressive symptoms // *Biol. Psychiatry*. 2013. V. 73. № 7. P. 667–678. <https://doi.org/10.1016/j.biopsych.2012.09.033>
40. Maciejewski D.F., Renteria M.E., Abdellaoui A. et al. The association of genetic predisposition to depressive symptoms with non-suicidal and suicidal self-injuries // *Behav. Genet.* 2017. V. 47. № 1. P. 3–10. <https://doi.org/10.1007/s10519-016-9809-z>
41. Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4 // *Nat. Genet.* 2011. V. 43. № 10. P. 977–983. <http://www.nature.com/articles/ng.943doi:10.1038/ng.943>
42. Wilcox H.C., Fullerton J.M., Glowinski A.L. et al. Traumatic stress interacts with bipolar disorder genetic risk to increase risk for suicide attempts // *J. Am. Acad. Child Adolesc. Psychiatry*. 2017. V. 56. № 12. P. 1073–1080. <https://doi.org/10.1016/j.jaac.2017.09.428>
43. Grove J., Ripke S., Als T.D. et al. Identification of common genetic risk variants for autism spectrum disorder // *Nat. Genet.* 2019. V. 51. № 3. P. 431–444. <https://doi.org/10.1038/s41588-019-0344-8>
44. Warriar V., Baron-Cohen S. Childhood trauma, lifetime self-harm, and suicidal behaviour and ideation are associated with polygenic scores for autism // *Mol. Psychiatry*. 2019. V. 26. № 5. P. 1670–1684. <https://doi.org/10.1038/s41380-019-0550-x>
45. Hodgson K., Coleman J.R.I., Hagenaars S.P. et al. Cannabis use, depression and self-harm: phenotypic and genetic relationships // *Addiction*. 2020. V. 115. № 3. P. 482–492. <https://doi.org/10.1111/add.14845>
46. Shen H., Gelaye B., Huang H. et al. Polygenic prediction and GWAS of depression, PTSD, and suicidal ideation/self-harm in a Peruvian cohort // *Neuropsychopharmacology*. 2020. V. 45. № 10. P. 1595–1602. <https://doi.org/10.1038/s41386-020-0603-5>
47. Russell A.E., Ford T., Gunnell D. et al. Investigating evidence for a causal association between inflammation and self-harm: A multivariable Mendelian randomisation study // *Brain. Behav. Immun.* 2020. V. 89. P. 43–50. <https://doi.org/10.1016/j.bbi.2020.05.065>
48. Lim K.X., Rijdsdijk F., Hagenaars S.P. et al. Studying individual risk factors for self-harm in the UK Biobank: A polygenic scoring and Mendelian randomisation study // *PLoS Med.* 2020. V. 17. № 1. <https://doi.org/10.1371/journal.pmed.1003137>
49. Campos A.I., Verweij K.J.H., Statham D.J. et al. Genetic aetiology of self-harm ideation and behaviour // *Sci. Rep.* 2020. V. 10. № 1. P. 1–11. <https://doi.org/10.1038/s41598-020-66737-9>
50. Manitt C., Eng C., Pokinko M. et al. DCC orchestrates the development of the prefrontal cortex during adolescence and is altered in psychiatric patients // *Transl. Psychiatry*. 2013. V. 3. № 12. P. e338–e338. <https://doi.org/10.1038/tp.2013.105>
51. Russell A.E., Hemani G., Jones H.J. et al. An exploration of the genetic epidemiology of non-suicidal self-harm and suicide attempt // *BMC Psychiatry*. 2021. V. 21. № 1. P. 207. <https://doi.org/10.1186/s12888-021-03216-z>
52. Li C., Liang X., Cheng S. et al. A multi-environments-gene interaction study of anxiety, depression and self-harm in the UK Biobank cohort // *J. Psychiatr. Res.* 2022. V. 147. P. 59–66. <https://doi.org/10.1016/j.jpsychires.2022.01.009>

Genetics and Epigenetics of Nonsuicidal Self-Injury: a Narrative Review

A. A. Kibitov^{a, b, *} and G. E. Mazo^b

^a*Mental Health Research Center, Moscow, 115522 Russia*

^b*Bekhterev National Medical Research Center for Psychiatry and Neurology, Saint-Petersburg, 192019 Russia*

^{*}*e-mail: andreykibitov18@gmail.com*

Nonsuicidal self-injury (NSSI), or self-harm is widely spread, especially among the young people. However, despite the problem's relevance, NSSI was investigated mainly from the psychological and social perspective, while a fairly small number of biological studies of NSSI have been published to date. The aim of this narrative review was to analyze all the currently available publications on the genetics and epigenetics of NSSI – one of the most promising areas in biological research. We discussed and analyzed all stages of genetic research of NSSI: from twin studies and studies of self-harm in the framework of hereditary diseases to candidate genes, genome-wide, epigenetic and gene-environment interactions studies. We demonstrated data indicating significant contribution of genetic and epigenetic factors to the development of NSSI, discussed advantages and limitations of analyzed studies, outlined prospects for further research in this area.

Keywords: nonsuicidal self-injury, self-harm, genetics, epigenetics, heritability.

ОБЗОРНЫЕ
И ТЕОРЕТИЧЕСКИЕ СТАТЬИ

УДК 575.113.15

ГЕНЕТИКА И ЭПИГЕНЕТИКА ПРЕЖДЕВРЕМЕННОГО
ПОЛОВОГО СОЗРЕВАНИЯ

© 2023 г. Е. А. Саженова¹*, С. А. Васильев¹, Л. В. Рычкова², Е. Е. Храмова², И. Н. Лебедев¹

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

²Научный центр проблем здоровья семьи и репродукции человека, Иркутск, 664003 Россия

*e-mail: elena.sazhenova@medgenetics.ru

Поступила в редакцию 02.06.2023 г.

После доработки 05.07.2023 г.

Принята к публикации 12.07.2023 г.

Центральное преждевременное половое созревание (ППС) вызвано преждевременной реактивацией гипоталамо-гипофизарно-гонадной оси. В определении сроков полового созревания решающую роль играют генетические, эпигенетические и экологические факторы. В последние годы варианты в генах *KISS1*, *KISS1R*, *MKRN3* и *DLK1* были идентифицированы как наследственные причины ППС. Гены *MKRN3* и *DLK1* являются импринтированными, в связи с чем эпигенетические модификации, изменяющие экспрессию данных генов, также рассматриваются в качестве причины преждевременного полового созревания. При прогрессировании ППС эпигенетические факторы, такие как метилирование ДНК, посттрансляционные модификации гистонов и некодирующие РНК, могут опосредовать взаимосвязь между влиянием генетических вариантов и окружающей среды. ППС связано и с другими краткосрочными и долгосрочными неблагоприятными последствиями для здоровья. Это является основанием для исследований, направленных на понимание генетических и эпигенетических причин ППС. Цель настоящего обзора — обобщение данных литературы о молекулярно-генетических и эпигенетических механизмах формирования ППС.

Ключевые слова: центральное преждевременное половое созревание, гонадотропин-рилизинг-гормон (ГнРГ), гипоталамо-гипофизарно-гонадная ось (ГПГ), геномный импринтинг, *DLK1*, *KISS1*, *KISS1R*, *MKRN3*.

DOI: 10.31857/S001667582312010X, EDN: QCSJXMF

Подростковый возраст — один из наиболее важных периодов развития человека в связи со сложными нейроэндокринными и другими биологическими изменениями, на фоне которых физическое и половое развитие максимально приближается к состоянию взрослого организма [1]. На основании данных крупного эпидемиологического исследования National Health and Nutrition Examination Survey (NHANES) разработаны нормативы физического и полового развития представителей различных этнических групп североамериканской популяции, которые в настоящее время используются как референтные в большинстве стран мира, включая Россию [2]. С учетом международных регламентирующих документов определяющими критериями начала полового созревания являются 2-я стадия развития молочных желез у девочек и 2-я стадия развития наружных половых органов у мальчиков. Экспертами NHANES был определен средний возраст начала пубертатного периода: для девочек европеоидной расы — 10.4 (10.1–10.7) года, негроидной расы —

9.5 (9.1–9.8) года; для юношей — 10.0 (9.6–10.4) лет и 9.2 (8.62–9.64) года [2, 3].

Преждевременное половое созревание (ППС) клинически определяется появлением вторичных половых признаков в возрасте до 8 лет у девочек и 9 лет у мальчиков [4]. Частота заболевания оценивается в 3.7 случаев на 10000 человек. В последние годы данная патология наблюдается все чаще, с преобладанием у девочек, и довольно изменчива в разных географических регионах, варьируя от 0.22 до 26.28 на 10000 девочек и от 0.02 до 0.9 на 10000 мальчиков [5, 6]. Так, частота ППС значительно увеличилась в период с 1998 по 2017 г. в Дании, достигнув 9.2 на 10000 девочек и 0.9 на 10000 мальчиков [7, 8]. Эпидемиологическое исследование в Испании (1997–2009 гг.) показало, что распространенность ППС составляет 1.9 на 10000 человек (девочки 3.7; мальчики 0.046). Ежегодный рост заболеваемости колебался от 0.002 до 0.107 новых случаев на 10000 человек, с заметным увеличением с 2000 г. среди девочек [9]. Такая разница между частотой ППС среди мальчиков и

девочек объясняется тем, что у девочек причиной ППС является нарушение гипоталамо-гипофизарной системы, в то время как у мальчиков ППС обусловлено органическими поражениями головного мозга.

Неопровержимые доказательства влияния генетических факторов на время наступления полового созревания предоставлены популяционными исследованиями [10]. Роль генетических факторов также иллюстрируется одинаковым возрастом менархе у матерей и дочерей среди членов одной этнической группы, а также большей согласованностью периода полового созревания у монозиготных, чем у dizиготных близнецов. Анализ сегрегации аллельных вариантов генов показал, что 27.5% случаев ППС являются семейными и предполагают аутомно-доминантное наследование с неполной пенетрантностью [11, 12].

За последнее десятилетие идентифицированы гены в сложной сети тормозных, стимулирующих и разрешающих нейроэндокринных факторов, участвующих в контроле наступления полового созревания, — *KISS1*, *GPR54* (*KISS1R*), а также два импринтированных гена, экспрессирующихся в норме только с отцовского гомолога, — *DLK1* и *MKRN3* [13]. Геномный импринтинг — эпигенетический феномен, дифференциально маркирующий материнские и отцовские копии генов в геноме и обеспечивающий их моноаллельную экспрессию в зависимости от родительского происхождения. В геноме человека из более 20 тыс. генов около 150 — импринтированные [14, 15]. Известно, что импринтированные гены играют основополагающую роль как в развитии головного мозга [16], так и в развитии плаценты как питающего эмбрион органа [17]. В выявлении причин развития ППС эти два органа имеют большое значение. Так, центральное ППС связано с нарушением гипоталамо-гипофизарной системы мозга. Эта система в эмбриональном периоде отвечает за развитие эндокринной системы плода, которая функционирует совместно с эндокринной системой материнского организма через плаценту. Гипофиз плода отвечает за синтез соматотропного фолликулостимулирующего, лютеинизирующего и тиреотропного гормонов, которые играют важную роль в процессах фетального роста и последующего полового развития, а также в формировании нервной системы. В результате нарушения гипоталамо-гипофизарной системы в эмбриональном развитии могут приводить к ограничению внутриутробного роста, а в постэмбриональном периоде — стать причиной раннего пубертатного развития [18].

Установлено, что патологии, связанные с геномным импринтингом, в некоторых случаях сопровождаются мультилокусными дефектами метилирования импринтированных генов (MLID —

multi-locus imprinting disturbances), которые были продемонстрированы как при остановке эмбрионального развития, так и при болезнях геномного импринтинга, сопровождающихся в том числе ППС [19–21].

Генетические нарушения объясняют примерно 50–80% вариативности начала и прогрессирования полового созревания. Детское ожирение также влияет на раннее начало пубертатного развития, хотя не существует статистических данных о разнице в возрасте наступления менархе между девочками с ожирением и нормальным весом [22]. Патологии, связанные с поражением центральной нервной системы, также могут влиять на развитие ППС [23]. К ним относятся: гамартома гипоталамуса [24], нейрофиброматоз I типа, сочетающийся с глиомой зрительных путей [25], супраселлярные арахноидальные кисты [26], менингиомиелоцеле [27].

К настоящему моменту накоплен большой массив данных о генетической компоненте возникновения и патогенеза преждевременного полового созревания. Цель настоящего обзора — обобщение данных литературы о молекулярно-генетических и эпигенетических механизмах формирования ППС.

КЛИНИЧЕСКАЯ КАРТИНА ПРЕЖДЕВРЕМЕННОГО ПОЛОВОГО СОЗРЕВАНИЯ

Преждевременное половое созревание (E30.1 по МКБ 10, MIM 604161, 603856) — это начало полового созревания раньше возрастной нормы. ППС может быть центральным (истинным, или гонадотропин-зависимым), периферическим (ложным) с механизмом развития, независимым от гонадотропин-рилизинг-гормона (**ГнРГ**), а также изолированным (появление железистой ткани молочных желез или лобкового оволосения у девочек до 8 лет без других клинико-лабораторных признаков ППС) [4]. Центральное ППС встречается примерно в 80% случаев по сравнению с периферическим и чаще у девочек, чем у мальчиков. Примерно треть всех случаев центрального ППС может быть связана с органическими поражениями головного мозга — гамартомой гипоталамуса, супраселлярными арахноидальными кистами и гидроцефалией [28]. Так, примерно у 50% мальчиков с центральным ППС обнаруживают органическое поражение головного мозга. В некоторых случаях причину центрального ППС у пациентов невозможно установить. В таком случае говорят об идиопатическом центральном ППС. Этот диагноз устанавливают у девочек в 70%, а у мальчиков — в 10–20% случаев ППС.

При центральной форме заболевания ранняя активация секреции ГнРГ приводит к увеличению

выработки гипофизарных гонадотропинов, которые, в свою очередь, активируют синтез половых гормонов в яичниках и яичках, что обуславливает развитие признаков пубертата. Центральное ППС характеризуется изменениями в гипоталамо-гипофизарной системе, которые происходят в процессе полового созревания у здоровых детей, поэтому клинически имитирует нормальный препубертатный и пубертатный периоды, являясь полным (со всеми признаками половой зрелости) и изосексуальным. Половое развитие у девочек включает увеличение молочных желез (телархе), лобковое (пубархе) и аксиллярное оволосение, а также менархе. По сравнению с нормальным пубертатом половое оволосение характеризуется меньшей интенсивностью. Другие андрогензависимые признаки (акне, жирная себорея, активность потовых желез) также в большинстве случаев появляются после 6–7 лет, в возрасте физиологической активации андрогеновой функции надпочечников. Развитие менархе зависит от активности патологического процесса. При идиопатическом ППС менструальная реакция в большинстве случаев наблюдается через 5–6 лет после появления вторичных половых признаков, при гипоталамической гамартомой — через несколько месяцев, при ложной форме ППС — одновременно с увеличением молочных желез. У девочек при полной форме заболевания формируется женская фигура [29].

Симптомами ППС центрального генеза у мальчиков являются увеличение наружных половых органов, половое оволосение (рост усов и бороды), акне, специфическое потоотделение, увеличение мышечной массы и сперматогенез. У них появляется грубый мужской голос. Периферическая форма ППС отличается прогрессированием вторичных половых признаков в отсутствие увеличения объема яичек [30]. Ускорение костного возраста считается наиболее точным маркером интенсивности полового созревания. При истинном ППС конечный рост составляет в среднем 146–150 см у девочек и 150–156 см у мальчиков. Максимальное ускорение костного возраста наблюдается у девочек с гипоталамической гамартомой, при которой конечный рост колеблется в пределах 135–148 см [30].

Психоневрологический статус при ППС характеризуется двигательной и эмоциональной расторможенностью, эйфоричностью, у мальчиков возможны агрессивность, аффективные реакции. Интеллектуальное развитие, как правило, не страдает.

Основными диагностическими мероприятиями являются: оценка физического и полового развития, рентгенография кистей рук с определением костного возраста, УЗИ органов малого таза, МРТ визуализация головного мозга, сельлярной области, определение уровня гормонов и проведение

стимуляционной пробы с ГнРГ. При проведении антропометрии и подсчете скорости роста за предшествующие 6–12 мес. увеличение скорости роста более чем на 2SD также может свидетельствовать о ППС. При истинном ППС базальный уровень гонадотропинов достигает пубертатных значений, однако в 30–50% случаев не превышает препубертатного диапазона. Повышение лютеинизирующего гормона (ЛГ) в ходе пробы считается главным гормональным маркером истинного ППС. Содержание тестостерона у мальчиков возрастает, а концентрация эстрогенов у девочек может колебаться от допубертатного до пубертатного уровня. Таким образом, характерными признаками центрального ППС являются: пубертатный базальный уровень гонадотропинов и пубертатное повышение их уровня, прежде всего ЛГ, в ответ на стимуляцию ГнРГ, увеличение яичников и матки у девочек и пубертатный объем яичек и уровень тестостерона у мальчиков, а также объемное образование головного мозга при опухолевом генезе заболевания. В дифференциальной диагностике различных форм ППС имеют значение также анамнез, особенности клинического статуса, результаты анализа стероидной секреции, цитогенетический анализ, рентгенографическое определение костного возраста, данные офтальмологического исследования.

При отсутствии лечения ППС сопровождается развитием таких осложнений как низкорослость (результат преждевременного закрытия зон роста трубчатых костей), диспластическое телосложение (короткие конечности, длинное туловище, широкий таз), ранняя сексуальность, навязчивость, нарушения функций репродуктивной системы. У девочек отмечаются нарушения менструального цикла по типу аномальных маточных кровотечений, формирование синдрома поликистозных яичников и преждевременное истощение овариального резерва [4]. Более ранний возраст менархе у девочек также связан с повышенным риском рака молочных желез, рака эндометрия, ожирения, диабета II типа и сердечно-сосудистых заболеваний. У мальчиков выявляются усиление сосудистого рисунка, гиперпневматизация пазухи скелета. Может проявляться неврологическая симптоматика: головная боль, быстрая утомляемость, булимия, полидипсия, симптомы поражения глазодвигательных нервов.

ГнРГ-независимое (периферическое) ППС намного менее распространено. Вторичные половые признаки развиваются в результате высокого уровня циркулирующих эстрогенов или андрогенов без активации гипоталамо-гипофизарной системы. Причиной такой дисфункции могут быть опухолевые заболевания этих желез, а также тестотоксикоз и синдром МакКьюна–Олбрайта [31].

ГЕНЕТИКА ПРЕЖДЕВРЕМЕННОГО ПОЛОВОГО СОЗРЕВАНИЯ

Известно, что начало полового созревания является результатом реактивации гипоталамо-гипофизарно-гонадной (ГПГ) оси. Эта ось включает в себя гипоталамус, гипофиз и гонадные железы и играет важную роль в развитии и регуляции репродуктивной и иммунной систем организма. Нейронами гипоталамуса секретируется ГнРГ, который проходит вниз по передней части гипофиза через гипофизарную портальную систему и связывается с рецепторами на секреторных клетках аденогипофиза. В ответ на стимуляцию ГнРГ эти клетки вырабатывают лютеинизирующий и фолликулостимулирующий гормоны, а гонады — эстроген и тестостерон. Эти гормоны играют важную роль во взаимодействии с гонадами. У женщин они активируют яичники для выработки эстрогена и ингибина для регуляции менструального и овариального циклов.

Ось ГПГ активна и в эмбриональном, и в neonatalном периоде развития. Она отвечает за формирование и развитие эндокринной системы плода в тесном взаимодействии с эндокринной системой материнского организма и плаценты. Типичным примером в этом отношении является синтез эмбрионом соматотропного гормона гипофиза начиная с 7–8-й нед. внутриутробного развития. Возрастание концентрации этого гормона в крови плода наблюдается до 20–24 нед. беременности, что коррелирует с показателями роста плода. Продукция гипофизом гонадотропных гормонов происходит в следующей последовательности: фолликулостимулирующий гормон синтезируется с 13-й нед. онтогенеза, лютеинизирующий — с 18-й, пролактин — с 19-й. Окситоцин и вазопрессин начинают синтезироваться задней долей гипофиза плода с 23-й нед. внутриутробного развития. Наивысшие концентрации окситоцина обнаруживают в конце плодного периода развития. Продукция тиреотропного гормона гипофиза начинается приблизительно на 12-й нед. беременности, что сопровождается началом образования тиреоидных гормонов Т3 и Т4, которые играют важную роль в процессах фетального роста и развития, особенно в ossификации костей скелета и зубов, а также в формировании нервной системы. В результате нарушения гипоталамо-гипофизарной системы в эмбриональном развитии могут приводить к ограничению внутриутробного роста [18]. После neonatalного периода ГПГ подавлена до наступления полового созревания, когда она реактивируется, что приводит к пульсирующей секреции ГнРГ, который регулируется активаторами нейронов, такими как киспептин (KISS1), нейрокинин В (NKB) и динорфин (KND) [32].

Ген *KISS1* (MIM 603286, 1q32.1, 3 экзона, размер мРНК 414 пн) кодирует белок киспептин, который стимулирует секрецию ГнРГ, тем самым участвуя в регуляции эндокринной функции и наступлении половой зрелости. Кроме того, *KISS1* является супрессором метастазирования, так как подавляет метастазы меланом и карцином молочной железы, не влияя на канцерогенность. Кодированный белок может ингибировать хемотаксис и инвазию и тем самым ослаблять метастазирование при злокачественных меланом. В настоящее время выявлен ряд генетических вариантов при ППС в этом гене, тем не менее мажорных вариантов не обнаружено [8, 33].

GPR54 (ген *KISS1R*, MIM 604161, 19p13.3, 5 экзонов, размер мРНК 1194 пн) — киспептин/рецептор-54, связанный с G-белком. Продукт гена *KISS1R* является регулятором выработки в гипоталамусе ГнРГ и ключевым фактором в инициации полового созревания, действуя как мощный стимулятор секреции ГнРГ-зависимого лютеинизирующего гормона. Данный рецептор широко экспрессируется в головном мозге, гипофизе и плаценте. В 2003 г. две исследовательские группы, независимо друг от друга, впервые сообщили об инактивирующих генетических вариантах гена *KISS1R* у пациентов с идиопатическим гипогонадотропным гипогонадизмом [34, 35]. Были идентифицированы аутомно-доминантные миссенс-варианты в пятом экзоне гена *KISS1R* у девочек с идиопатическим ППС: с.1157G>C (p.Arg386Pro) и с.1091T>A [36, 37]. Эти исследования подтвердили критическую роль, которую играет система *KISS1/KISS1R* в активации оси гипоталамус–гипофиз–гонады. Кроме того, недавние исследования продемонстрировали участие *KISS1* и *KISS1R* в процессах имплантации и плацентации [38]. Самый высокий уровень экспрессии *KISS1* в плаценте у людей был обнаружен в течение первого триместра эмбрионального развития и совпадал со временем инвазии трофобласта — миграции ворсинчатого цитотрофобласта из оснований якорных ворсин вглубь базального эндометрия, т.е. в зону контакта плаценты и стенки матки. В это время инвазивные клетки целенаправленно проникают в стенки спиральных артерий эндометрия, затем радиальных артерий миометрия. У человека *KISS1* в основном экспрессируется в ворсинчатом трофобласте и играет ингибирующую роль в миграции трофобласта через паракринный/аутокринный механизм в ткани плаценты [38]. *KISS1R* эволюционно консервативен у млекопитающих, при этом сходство этого гена у коз, свиней, крупного рогатого скота и овец выше, чем у других видов млекопитающих, таких как человек, крыса и мышь [39].

Импринтированный и не экспрессирующийся на материнском гомологе хромосомы 14 (14q32) ген *DLK1* (MIM 176290, 5 экзонов, размер мРНК

1149 пн) кодирует EGF-подобный фактор роста. Это мембраносвязывающий белок, который участвует в сигнальном пути Notch и способствует передаче сигналов для клеточной пролиферации во время нейрогенеза. В то же время у мыши его активность ингибируется геном *Numb*, способствуя дифференцировке нейроэндокринных клеток. Продукты этого гена также принимают участие в остеогенезе, адипогенезе, кроветворении и пролиферации гепатоцитов [40]. Показано, что у мышей *Dlk1* экспрессируется пренатально в нейроэндокринных тканях, включая гипофиз, и постнатально в гипоталамусе, включая срединно-базальный гипоталамус — центр контроля секреции ГнРГ. Продукт этого гена также важен для гомеостаза жировой ткани [13]. Полногеномные ассоциативные исследования (GWAS) показали, что однонуклеотидные полиморфизмы гена *DLK1*, унаследованные от отца, связаны с более ранним началом менархе [41]. Был описан ребенок с центральным ППС и сложной перестройкой, сопровождающейся делецией ~14 тпн и дупликацией 269 пн *DLK1*. Гетерозиготная делеция затрагивала 5'-нетранслируемый регион и первый экзон *DLK1*, включая стартовый сайт трансляции, а дупликация была представлена 269-нуклеотидным сегментом интрона 3 данного гена, который оказался вставлен между точками разрыва геномной делеции. Ребенок унаследовал данный вариант от отца [42]. В группе из 60 детей с ППС L.G. Gomes с соавт. выявили три варианта сдвига рамки считывания в гене *DLK1* у пяти пациентов из трех семей: p.Gly199Alafs*11, p.Val271Cysfs*14 и p.Pro160Leufs*50 [40]. Интересно, что пациенты с патогенетически значимыми вариантами в гене *DLK1* по сравнению с пациентами с патогенетически значимыми вариантами в генах *KISS1*, *GPR54* и *MKRN3* имеют центральное ожирение, раннее начало непереносимости глюкозы/диабет II типа и гиперлипидемию. Эти метаболические изменения были сходны с фенотипом дефицита *Dlk1*, ранее описанным на нулевой мышинной модели, при этом наличия ППС у мыши отмечено не было [43].

Импринтированный также на материнском гомологе ген *MKRN3* (MIM 603856, 15q11.2, 3 экзона, безынтронный ген, размер мРНК 1521 пн) кодирует макорин-белок RING-finger 3, который относится к семейству макоринов и участвует в управлении началом полового созревания. В дуговидном ядре гипоталамуса мыши активны участвующие в пубертатном развитии гены *Kiss1* и *Tac3*, экспрессия которых увеличивается по мере снижения уровня *Mkrm3* [44, 45]. Белок MKRN3 блокирует высвобождение ГнРГ из гипоталамуса, тем самым задерживая наступление половой зрелости [46].

Семейство белков макоринов отличается характерной комбинацией мотивов “цинковые пальцы” C3H, которые участвуют в связывании

РНК, и “цинковые пальцы” RING, обнаруживающиеся в большинстве убиквитинлигаз E3. MKRN3 ответствен за убиквитинизацию белков, при которой фрагмент убиквитина присоединяется к внутриклеточному белку для перемещения его в протеасому. Убиквитинирование также может быть индикатором передачи сигнала регуляции клеточного цикла, дифференцировки и морфогенеза. Широко распространенная консервативность семейства белков макоринов предполагает, что они играют важную роль в клетках с высоким уровнем экспрессии этих белков в развивающейся нервной системе. С другой стороны, *MKRN3* сохраняется только у терианских млекопитающих, и его точная функция до конца не определена [39].

Делеция региона хромосомы 15q11-q13 включает импринтированный ген *MKRN3* и вносит определенный вклад в синдром Прадера—Вилли. Было описано два пациента с основными признаками синдрома Прадера—Вилли, у которых *MKRN3* не был вовлечен в делецию [47]. В то же время в этом же исследовании описан пациент с унаследованной от отца делецией, включающей гены *MKRN3*, *MAGEL2* и *NDN*, у которого были лишь некоторые признаки синдрома Прадера—Вилли: ожирение, задержка в развитии и высокий болевой порог, а также признаки ППС в возрасте 7 лет 6 мес. Пациенту был поставлен диагноз центрального ППС, подтвержденный реакцией на лечение трипторелином. В данном случае делеция *MKRN3*, вероятно, являлась причиной раннего полового созревания у этого пациента.

В нескольких исследованиях было показано, что высокий уровень белка MKRN3 в сыворотке крови наблюдается у девочек и мальчиков до наступления половой зрелости. Уровень MKRN3 в сыворотке крови был снижен у девочек с центральным ППС по сравнению с контрольной группой, соответствующей по возрасту [48]. У детей с ППС в гене *MKRN3* описаны также варианты со сдвигом рамки считывания, нонсенс- и миссенс-варианты, которые могли привести к снижению его функции [11, 49, 50]. Кроме того, метаанализ показал, что варианты в *MKRN3* при центральном ППС выявляются гораздо чаще (~46%), чем при спорадическом (~4%), и являются наиболее распространенной причиной семейного ППС [51]. И действительно, в трех семейных случаях центрального ППС были описаны варианты гена *MKRN3*: нонсенс-вариант p.E40X, миссенс-вариант c.343T>A p.C115S и c.1091G>C p.C364S, расцененные как патогенные [52]. Также в этом же гене идентифицированы в двух семейных случаях патогенетические варианты: со сдвигом рамки считывания p.Ala162Glyfs*14 и миссенс-вариант p.Arg365Ser [42].

Патогенетические варианты гена *LHCGR* рецептора лютеинизирующего гормона были опи-

саны при гонадотропин-независимой форме ППС (табл. 1). Это ауtosомно-доминантные или редко *de novo* варианты, приводящие к повышению экспрессии данного гена. Также были описаны соматические варианты этого гена при изолированных аденомах клеток Лейдига, приводящие к двустороннему диффузному образованию опухоли этих клеток [56]. Данное заболевание обычно проявляется в возрасте от 2 до 4 лет исключительно у мальчиков. У пациентов наблюдаются раннее развитие вторичных и третичных половых признаков, а также ускоренный рост скелета с последующей ранней его остановкой [61].

Ген *LHCGR* (MIM 152790, 2p16.3, 11 экзонов) кодирует рецептор лютеинизирующего гормон-хорионического гонадотропина, который связан с G-белком. Патогенетические варианты этого гена приводят к повышению уровня половых стероидов при низком содержании лютеинизирующего гормона [62].

ЭПИГЕНЕТИКА ПРЕЖДЕВРЕМЕННОГО ПОЛОВОГО СОЗРЕВАНИЯ

Эпигенетические изменения определяются как митотически и/или мейотически стабильные молекулярные факторы и процессы, которые регулируют активность генома, не влияя на последовательность ДНК. Эпигенетическая информация у млекопитающих может передаваться с помощью таких основных механизмов как митотически стабильное метилирование ДНК, посттрансляционные модификации гистоновых белков, ремоделирование хроматина и некодирующих РНК (например, микроРНК и длинных некодирующих РНК) [63, 64].

В последние годы появляется все больше доказательств участия эпигенетических механизмов в контроле полового созревания, что указывает на их потенциальную модулирующую роль в сроках полового созревания человека [23]. Так, переключение на экспрессию киспептина в дугообразных нейронах является основным механизмом, лежащим в основе начала полового созревания [63]. Выявлен эпигенетический механизм подавления транскрипции, который регулирует начало половой зрелости у крыс. Основными участниками этого механизма подавления были белки-регуляторы группы Polycomb, влияющие на генную сеть, которая участвует в стимулирующем контроле полового созревания, включая ген *Kiss1* и гены глутаматергических рецепторов [65]. Фермент *Kdm6b* был идентифицирован как центральный узел этой сети путем удаления Polycomb-зависимой модификации гистонов H3K27me3. Было высказано предположение, что подавление *Kdm6b* может быть механизмом, используемым белками группы Polycomb для модуляции активирующих половое созревание генных сетей

[65]. Структуры цинковых пальцев также вовлечены в качестве эпигенетического модулятора полового созревания посредством подавления транскрипции [56]. Так, у обезьян и грызунов обнаружен ген цинковых пальцев *GATAD1*, экспрессия которого снижается в перипубертатном периоде. Этот ген был идентифицирован как основной в подавлении *Kiss1* и *Tac3* в дугообразных нейронах и сдерживающий половое созревание [66]. Группа белков Trithorax представляет собой комплекс активаторов транскрипции. Примечательно, что, как и белки группы Polycomb, Trithorax являются белками хроматина, которые противодействуют белкам группы Polycomb путем осуществления метилирования гистоновых меток в H3K4 [63]. Исследования на мышах, оценивающие взаимодействие с промотором *Kiss1*, показали, что перед началом полового созревания, по мере снижения экспрессии данного гена, опосредованного белками Polycomb, активация Trithorax увеличивается, точно настраивая процесс эпигенетической регуляции сроков полового созревания [67].

Исследование на мышах показало, что микроРНК могут регулировать выработку гипоталамического ГнРГ до наступления половой зрелости, причем miR-200 и miR-155 идентифицированы как два важных компонента этого процесса, регулирующих транскрипционные факторы ГнРГ и вмешивающихся в правильное начало полового созревания [68]. Также эксперименты на крысах продемонстрировали увеличение экспрессии miR30-b в гипоталамусе во время постнатального созревания, что сопровождалось снижением экспрессии *Mkrn3*. В работах *in vitro* было показано сильное репрессивное действие miR-30b на 3'UTR *Mkrn3* и, следовательно, подавление экспрессии этого гена [69].

Также были описаны случаи ППС при болезнях геномного импринтинга, таких как синдромы Прадера–Вилли, Темпла, Рассела–Сильвера и других (табл. 1). Так, у трех из 36 пациентов с идиопатическим центральным ППС было выявлено гипометилирование центра импринтинга *DLK1*, характерное для синдрома Темпла [70]. Кроме ППС пациенты имели задержку эмбрионального и постнатального развития, низкий рост и вес при рождении и лицевые дисморфии, которые могут быть связаны с синдромом Темпла. Отсутствие продукта гена *DLK1* является одной из причин данного синдрома, который в 86% случаев сопровождается центральным ППС [54]. ППС также было описано у детей с синдромом Рассела–Сильвера, характеризующимся пренатальной и постнатальной задержкой роста. Молекулярными причинами данного синдрома были: однопородительское наследование хромосомы 7 от матери и гипометилирование центра импринтинга *IGF2/H19*. В то же время ген, обуславливающий

Таблица 1. Наследственные синдромы с центральным преждевременным половым созреванием

Синдром ОМIM, регион	Молекулярные нарушения	Распространенность ППС	Другие основные клинические признаки	Ген, вовлеченный в ППС	Ссылка
Синдром Рассела–Сильвера #180860, 11p15.5 #618905, 7p13-q32	Гипометилирование ЦИ <i>IGF2/H19</i> ОРД7 ^{mat}	5–15%	Пrenатальная и постнатальная задержка роста, относительная макроцефалия, выпуклый лоб, асимметрия тела, трудности кормления	Не установлен	[53]
Синдром Темпла #616222, 14q32.2	ОРД14 ^{mat} ; гипометилирование ЦИ <i>DLK1/MEG3</i> ; делеция 14q32.2 на отцовском гомологе	86%	Пrenатальная и постнатальная задержка роста, гипотония, маленькие руки и ноги, ожирение, задержка двигательной активности	<i>DLK1</i>	[54, 55]
Синдром Прадера–Вилли #176270, 15q11-q13	Делеция 15q11-q13 отцовского гомолога; ОРД15 ^{mat}	4%	Гипотония, ожирение, задержка роста, когнитивные нарушения, гипогонадизм	<i>MKRN3</i>	[47, 52]
Соматическая аденома клеток Лейдига (тестотоксикоз) #176410, 2p16.3	Мутации в гене <i>LHCGR</i>	АД тип наследования	Раннее развитие вторичных и третичных половых признаков у мальчиков, а также ускоренный рост скелета с последую- щей ранней его остановкой	<i>LHCGR</i>	[56]
Синдром Вильямса–Бойрена #194050, 7q11.23	Делеция 7q11.23	3–18%	Сердечно-сосудистые заболевания, низкий рост, умственная отсталость, гиперсоциальность	Не установлен	[57, 58]
Синдром дупликации Xp11.23-p11.22 #300881, Xp11.23-p11.22	Дупликация Xp11.23-p11.22 с доминантным наследованием	Женщины: 70%; мужчины: 11%	Умственная отсталость, задержка речи, аномалии электроэнцефалограммы, избыточный вес, аномалии скелета	Не установлен	[59]
Гипотония, эпилепсия и преждевременное половое развитие #612777, 2p16.3	Не установлены	Описаны три ребенка в одной семье со схожей клинической картиной	Гипотония, судороги и дисморфические особенности (брахицефалия, плагиоце- фалия и экзофтальм)	Не установлен	[60]

Примечание. ЦИ — центр импринтинга; ОРД — однородительское наследование; mat — материнский гомолог; АД — аутосомно-доминантный тип наследования.

ППС у данного пациента, не идентифицирован [71, 72].

В некоторых случаях болезни геномного импринтинга сопровождаются MLID. Так, описан двухлетний ребенок, имеющий гипометилирование одновременно в локусе *KCNQ1OT1* (11p15), характерное для синдрома Видемана—Беквита, и *DLK1* (14q32), определяющее развитие синдрома Тэмпла [73]. Клинические признаки у ребенка были нетипичны для этих синдромов: относительно небольшой рост, выступающий лоб, относительная макроцефалия, трудности речи и асимметрия скелета, характерная для синдрома Рассела—Сильвера. Поиск патогенетических вариантов в генах с предположительным контролем импринтинга, таких как *ZFP57*, не проводился. ППС у ребенка также не отмечено, возможно в силу возраста. Эпимутации в гене *DLK1*, сопровождающиеся MLID, также были продемонстрированы при остановке эмбрионального развития [20, 21, 74]. Так, ранее было показано гипометилирование *DLK1*, сопровождающееся MLID у спонтанных эмбрионов I триместра беременности от женщин как со спорадическим абортom, так и при привычном невынашивании беременности [21, 75]. Поэтому в случае выявления эпимутаций в *DLK1* и *MKRN3* существует большая вероятность обнаружения у этих больных MLID.

MLID в свою очередь поднимает вопрос о причинах возникновения данного феномена, указывая на возможность существования особых регуляторных механизмов, централизованно контролирующих эпигенетический статус нескольких импринтированных локусов генома. В настоящее время определены гены, предположительно контролируемые импринтинг [76, 77]. Это гены *NLRP7* (*NALP7*), *NLRP5*, *NLRP2*, *KHDC3L*, *ZFP57* и некоторые другие (список периодически пополняется), мутации в которых сопровождаются MLID как при остановке эмбрионального развития, так и при наследственных синдромах, обусловленных нарушениями геномного импринтинга, в том числе проявляющимися интеллектуальными отклонениями [78, 79]. Поэтому при ППС, сопровождающемся эпимутациями в генах *DLK1* и (или) *MKRN3*, актуально определение статуса метилирования в других импринтированных генах для выявления мультилокусных дефектов метилирования, а в случае их обнаружения — поиск патогенетических вариантов в генах контроля импринтинга.

ZFP57 (6p22.1) относится к генам, которые предположительно контролируют импринтинг [76]. Этот ген кодирует белок, который имеет KRAB-домен, ассоциированный с цинковыми пальцами типа Cys2His2. KRAB — домен транскрипционной репрессии, действующий через белок-белковые взаимодействия с KAP1/TIF1b/Trim28, ко-

торые могут начать *de novo* метилирование ДНК мыши во время эмбриогенеза. *Zfp57* необходим для поддержания статуса метилирования импринтированных генов, особенно в период эпигенетических волн перепрограммирования во время преимплантационного развития [79].

Анализ профилей метилирования ДНК лимфоцитов у 10 девочек с центральным ППС и 33 здоровых девочек препубертатного и пубертатного возраста показал, что в препубертатном и пубертатном возрасте в лимфоцитах происходит изменение метилирования генов цинковых пальцев, включая гипометилирование промотора гена *ZFP57*, что указывает на потенциальную репрессию геном-мишенной *ZFP57* [80]. Чтобы определить, сопровождается ли гипометилирование *ZFP57* в лимфоцитах периферической крови повышенной экспрессией этого гена в гипоталамусе во время полового созревания, авторы измерили уровни мРНК *ZFP57* в медиальном базальном гипоталамусе у самок макак-резусов до- и перипубертатного периода. И действительно, в гипоталамусе самок обезьян в период полового созревания показано увеличение экспрессии *ZFP57*. По-видимому, важной функцией *ZFP57* в нейроэндокринной системе мозга является подавление активности транскрипционных репрессоров полового созревания, таких как комплекс Polycomb или другие ZNF [66]. Кроме того, отрицательная корреляция, обнаруженная между увеличением уровней мРНК *ZFP57* и снижением экспрессии *ZNF597*, которая происходит в гипоталамусе во время полового созревания обезьян, предполагает, что *ZNF597* может быть одним из транскрипционных ингибиторов полового созревания, подавляемых *ZFP57*.

Импринтированные гены необходимы для роста и развития как в пренатальном, так и в постнатальном периоде. В пренатальном периоде через функционирование в плаценте они участвуют в регуляции имплантации, дифференцировке трофобласта, процессе получения эмбрионом стабильного питания через ремоделирование спиральных артерий матери. В гипоталамо-гипофизарной системе плода они отвечают за процессы фетального роста и развития через формирование эндокринной системы плода. Важно отметить, что гипоталамус плода и плацента развиваются в период, критически значимый для формирования плацентарного кровоснабжения, обеспечивая тем самым питание плода и выработку гормонов плаценты, которые регулируют гипоталамус матери. Таким образом, плацента участвует в развитии гипоталамуса плода, а гипоталамус матери предоставляет ресурсы для этого [81]. В результате эволюции геномного импринтинга проходила одновременно по пути участия в развитии мозга и регуляции плаценты, в формировании головного мозга под влиянием плаценты и давления отбора. В результате нару-

шения импринтинга могли стать, в том числе, и причиной аномалий полового созревания.

ЗАКЛЮЧЕНИЕ

Начало полового созревания контролируется взаимодействием между генетическими, эпигенетическими и ненаследственными факторами, а ППС является результатом преждевременной их активации. Обнаружение патогенетически значимых вариантов в генах *KISS1*, *KISS1R*, *MKRN3* и *DLK1* вносит заметный вклад в установление генетических причин ППС. В то же время генетические варианты в этих генах выявляются только у 50–80% детей с данной патологией, что поднимает вопрос о возможном влиянии минорных генетических вариантов в других генах. Гены *MKRN3* и *DLK1*, являясь импринтированными, с исключительно отцовской экспрессией в норме, могут вносить вклад в формирование ППС также через эпимутации промоторных регионов. Эпимутации в импринтированных генах часто сопровождаются мультилокусными дефектами метилирования в других импринтированных локусах. При обнаружении MLID возникает необходимость проводить поиск патогенетических вариантов в генах с предположительным контролем импринтинга, таких как *NLRP7* (*NALP7*), *NLRP5*, *NLRP2*, *KHDC3L*, *ZFP57*. Сегодня генетическая диагностика ППС ограничивается только секвенированием генов *KISS1*, *KISS1R*, *MKRN3* и *DLK1*, в то же время для определения нарушений генома у таких пациентов, в случае отсутствия патогенетических вариантов в этих генах, необходимо определять статус метилирования *MKRN3* и *DLK1*, а при выявлении эпимутаций в них проводить поиск MLID в других импринтированных локусах с последующим поиском мутаций в генах контроля импринтинга. Кроме того, остается невыясненной ситуация о возможном участии других генов в патогенезе ППС, а для этого появляется необходимость полного экзомного и геномного секвенирования. Необходимость поиска других генов в установлении причин ППС объясняется описанием синдромов с участием ППС без выявления конкретного гена, вовлеченного в эту патологию (табл. 1). Такой всесторонний подход к выявлению генетических и эпигенетических причин ППС позволит в полной мере подойти к пониманию роли генетических и эпигенетических факторов в формировании данной патологии.

Исследование выполнено при финансовой поддержке гранта РНФ в рамках научного проекта № 23-25-00359.

Настоящая статья не содержит каких-либо исследований, выполненных с использованием биологических материалов.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Дедов И.И., Семичева Т.В., Петеркова В.А. Половое развитие детей: норма и патология. М.: Колор Ит. Студио, 2002. 232 с.
2. US Department of Health and Human Services. Third National Health and Nutrition Examination Survey, 1988–1994 // Hyattsville. MD: National Center for Health Statistics, Centers for Disease Control and Prevention. 1999. https://www.academia.edu/6706974/Third_National_Health_and_Nutrition_Examination_Survey
3. Bleil M.E., Booth-LaForce C., Benner A.D. Race disparities in pubertal timing: Implications for cardiovascular disease risk among African American women // *Popul. Res. Policy Rev.* 2017. V. 36. P. 717–738. <https://doi.org/10.1007/s11113-017-9441-5>
4. Петеркова В.А., Алимова И.Л., Башнина Е.Б. и др. Клинические рекомендации “преждевременное половое развитие” // *Проблемы эндокринологии.* 2021. Т. 67. № 5. С. 84–103. <https://doi.org/10.14341/probl12821>
5. Kim Y.J., Kwon A., Jung M.K. et al. Incidence and prevalence of central precocious puberty in Korea: An epidemiologic study based on a national database // *J. Pediatr.* 2019. V. 208. P. 221–228.
6. Eckert-Lind C., Busch A.S., Petersen J.H. et al. Worldwide secular trends in age at pubertal onset assessed by breast development among girls: A systematic review and meta-analysis // *JAMA Pediatr.* 2020. V. 174. № 4. P. e195881. <https://doi.org/10.1001/jamapediatrics.2019.5881>
7. Brauner E.V., Busch A.S., Eckert-Lind C. et al. Trends in the incidence of central precocious puberty and normal variant puberty among children in Denmark, 1998 to 2017 // *JAMA Netw. Open.* 2020. V. 3. № 10. P. e2015665. <https://doi.org/10.1001/jamanetworkopen.2020.15665>
8. Maione L., Bouvattier C., Kaiser U.B. Central precocious puberty: Recent advances in understanding the etiology and in the clinical approach // *Clin. Endocrinol.* 2021. V. 95. № 4. P. 542–555. <https://doi.org/10.1111/cen.14475>
9. Soriano-Guillen L., Corripio R., Labarta K. et al. Central precocious puberty in children living in Spain: incidence, prevalence, and influence of adoption and immigration // *J. Clin. Endocrinol. Metab.* 2010. V. 95. № 9. P. 4305–4313.
10. Zhu J., Kusa T.O., Chan Y.M. Genetics of pubertal timing // *Curr. Opin. Pediatr.* 2018. V. 30. P. 532–540. <https://doi.org/10.1097/MOP.0000000000000642>
11. Valadares L.P., Meireles C.G., De Toledo I.P. et al. *MKRN3* mutations in central precocious puberty: A systematic review and meta-analysis // *J. Endocr. Soc.* 2019. V. 3. P. 979–995. <https://doi.org/10.1210/js.2019-00041>
12. Varimo T., Wang Y., Miettinen P.J. et al. Circulating miR-30b levels increase during male puberty // *Eur. J. Endocrinol.* 2021. V. 184. № 5. P. K11–K14.

13. Shim Y.S., Lee H.S., Hwang J.S. Genetic factors in precocious puberty // Clin. Exp. Pediatr. 2022. V. 65. № 4. P. 172–181.
<https://doi.org/10.3345/cep.2021.00521>
14. <http://igc.otago.ac.nz> — каталог импринтированных генов и родительских эффектов у человека и животных.
15. Renfree M.B., Hore T.A., Shaw G. Evolution of genomic imprinting: Insights from marsupials and monotremes // Annu. Rev. Genomics Hum. Genet. 2009. V. 10. P. 241–262.
16. Tucci V., Isles A.R., Kelsey G., Ferguson-Smith A.C. Genomic imprinting and physiological processes in mammals // Cell. 2019. V. 176. № 5. P. 952–965.
<https://doi.org/10.1016/j.cell.2019.01.043>
17. Eggermann T., Davies J.H., Tauber M. et al. Growth restriction and genomic imprinting—overlapping phenotypes support the concept of an imprinting network // Genes (Basel). 2021. V. 12. № 4. P. e585.
<https://doi.org/10.3390/genes12040585>
18. Canton A.P., Krepisch A.C., Montenegro L.R. et al. Insights from the genetic characterization of central precocious puberty associated with multiple anomalies // Hum. Reprod. 2021. V. 36. № 2. P. 506–518.
<https://doi.org/10.1093/humrep/deaa306>
19. Fontana L., Bedeschi M.F., Maitz S. et al. Characterization of multi-locus imprinting disturbances and underlying genetic defects in patients with chromosome 11p15.5 related imprinting disorders // Epigenetics. 2018. V. 13. № 9. P. 897–909.
<https://doi.org/10.1080/15592294.2018.1514230>
20. Sparago A., Verma A., Patricelli M.G. et al. The phenotypic variations of multi-locus imprinting disturbances associated with maternal-effect variants of *NLRP5* range from overt imprinting disorder to apparently healthy phenotype // Clin. Epigenetics. 2019. V. 11. P. e190.
<https://doi.org/10.1186/s13148-019-0760-8>
21. Sazhenova E.A., Nikitina T.V., Vasilyev S.A. et al. *NLRP7* variants in spontaneous abortions with multilocus imprinting disturbances from women with recurrent pregnancy loss // J. Assist. Reprod. Genet. 2021. V. 38. № 11. P. 2893–2908.
<https://doi.org/10.1007/s10815-021-02312-z>
22. Faienza M.F., Urbano F., Mosconi L.A. et al. Genetic, epigenetic and environmental influencing factors on the regulation of precocious and delayed puberty // Front. Endocrinol. (Lausanne). 2022. V. 13. P. e1019468.
<https://doi.org/10.3389/fendo.2022.1019468>
23. Brito V.N., Canton A.P., Seraphim C.E. et al. The congenital and acquired mechanisms implicated in the etiology of central precocious puberty // Endocr. Rev. 2023. V. 44. № 2. P. 193–221.
<https://doi.org/10.1210/endrev/bnac020>
24. Cantas-Orsdemir S., Garb J.L., Allen H.F. Prevalence of cranial MRI findings in girls with central precocious puberty: A systematic review and meta-analysis // J. Pediatr. Endocrinol. Metab. 2018. V. 31. № 7. P. 701–710.
<https://doi.org/10.1515/jpem-2018-0052>
25. Imperial R., Toor O.M., Hussain A. et al. Comprehensive pancancer genomic analysis reveals (RTK)-RAS-RAF-MEK as a key dysregulated pathway in cancer: its clinical implications // Semin. Cancer Biol. 2019. V. 54. P. 14–28.
<https://doi.org/10.1016/j.semcancer.2017.11.016>
26. Savas Erdeve S., Ocal G., Berberoglu M. et al. The endocrine spectrum of intracranial cysts in childhood and review of the literature // J. Pediatr. Endocrinol. Metab. 2011. V. 24. № 11–12. P. 867–875.
<https://doi.org/10.1515/jpem.2011.263>
27. Almutlaq N., O'Neil J., Fuqua J.S. Central precocious puberty in spina bifida children: guidelines for the care of people with spina bifida // J. Pediatr. Rehabil. Med. 2020. V. 13. № 4. P. 557–563.
28. Vurali D., Ozon A., Gonc E.N. et al. Gender-related differences in etiology of organic central precocious puberty // Turk. J. Pediatr. 2020. V. 62. № 5. P. 763–769.
<https://doi.org/10.24953/turkjped.2020.05.007>
29. Фархутдинова Л.М. Преждевременное половое созревание центрального происхождения // Архив внутренней медицины. 2017. № 4. С. 245–251.
30. Витебская А.В., Амишинская Д.Р., Шуминов О.В. Гонадотропинзависимое преждевременное половое созревание у девочек. Описание клинических случаев // Сеченовский вестник. 2017. Т. 1. № 27. С. 36–40.
31. Gangat M., Radovick S. Precocious puberty // Minerva Pediatr. 2020. V. 72. № 6. P. 491–500.
<https://doi.org/10.23736/S0026-4946.20.05970-8>
32. Campbell R.E., Coolen L.M., Hoffman G.E., Hrabovszky E. Highlights of neuroanatomical discoveries of the mammalian gonadotropin-releasing hormone system // J. Neuroendocrinol. 2022. V. 34. № 5. P. e13115.
<https://doi.org/10.1111/jne.13115>
33. Vazquez M.J., Toro C.A., Castellano J.M. et al. *SIRT1* mediates obesity- and nutrient-dependent perturbation of pubertal timing by epigenetically controlling *Kiss1* expression // Nat. Commun. 2018. V. 9. № 1. P. e4194.
<https://doi.org/10.1038/s41467-018-06459-9>
34. De Roux N., Genin E., Carel J.C. et al. Hypogonadotropic hypogonadism due to loss of function of the *Kiss1*-derived peptide receptor *GPR54* // Proc. Natl Acad. Sci. USA. 2003. V. 100. № 19. P. 10972–10976.
35. Seminara S.B., Messager S., Chatzidaki E.E. et al. The *GPR54* gene as a regulator of puberty // N. Engl. J. Med. 2003. V. 349. № 17. P. 1614–1627.
36. Teles M.G., Bianco S.D., Brito V.N. et al. *AGPR54*-activating mutation in a patient with central precocious puberty // N. Engl. J. Med. 2008. V. 358. № 7. P. 709–715.
37. Pagani S., Calcaterra V., Acquafredda G. et al. *MKRN3* and *KISS1R* mutations in precocious and early puberty // Ital. J. Pediatr. 2020. V. 46. № 1. P. e39.
<https://doi.org/10.1186/s13052-020-0808-6>
38. Hu K.L., Chang H.M., Zhao H.C. et al. Potential roles for the kisspeptin/kisspeptin receptor system in implantation and placentation // Hum. Reprod. Update. 2019. V. 25. № 3. P. 326–343.
<https://doi.org/10.1093/humupd/dmy046>
39. Abreu A.P., Toro C.A., Song Y.B. et al. *MKRN3* inhibits the reproductive axis through actions in kisspeptin-expressing neurons // J. Clin. Invest. 2020. V. 130. № 8. P. 4486–4500.
<https://doi.org/10.1172/JCI136564>
40. Gomes L.G., Cunha-Silva M., Crespo R.P. et al. *DLK1* is a novel link between reproduction and metabolism // J.

- Clin. Endocrinol. Metab. 2019. V. 104. № 6. P. 2112–2120.
41. Perry J.R., Day F., Elks C.E. et al. Parent-of-origin specific allelic associations among 106 genomic loci for age at menarche // Nature. 2014. V. 514. P. 92–97.
 42. Dauber A., Cunha-Silva M., Macedo D.B. et al. Paternally inherited *DLK1* deletion associated with familial central precocious puberty // J. Clin. Endocrinol. Metab. 2017. V. 102. № 5. P. 1557–1567.
<https://doi.org/10.1210/jc.2016-3677>
 43. Moon Y.S., Smas C.M., Lee K. et al. Mice lacking paternally expressed *Pref-1/Dlk1* display growth retardation and accelerated adiposity // Mol. Cell. Biol. 2002. V. 22. № 15. P. 5585–5592.
<https://doi.org/10.1128/MCB.22.15.5585-5592.2002>
 44. Li C., Han T., Li Q. et al. MKRN3-mediated ubiquitination of Poly(A)-binding proteins modulates the stability and translation of *GNRH1* mRNA in mammalian puberty // Nucl. Acids Res. 2021. V. 49. № 7. P. 3796–3813.
 45. Yellapragada V., Liu X., Lund C. et al. MKRN3 interacts with several proteins implicated in puberty timing but does not influence *GNRH1* expression // Front. Endocrinol. (Lausanne). 2019. V. 10. P. e48.
<https://doi.org/10.3389/fendo.2019.00048>
 46. Garcia J.P., Guerriero K.A., Keen K.L. et al. Kisspeptin and neurokinin B signaling network underlies the pubertal increase in GnRH release in female rhesus monkeys // Endocrinology. 2017. V. 158. № 10. P. 3269–3280.
<https://doi.org/10.1210/en.2017-00500>
 47. Kanber D., Giltay J., Wiczorek D. et al. A paternal deletion of *MKRN3*, *MAGEL2* and *NDN* does not result in Prader–Willi syndrome // Eur. J. Hum. Genet. 2009. V. 17. № 5. P. 582–590.
 48. Jeong H.R., Lee H.J., Shim Y.S. et al. Serum Makorin ring finger protein 3 values for predicting Central precocious puberty in girls // Gynecol. Endocrinol. 2019. V. 35. P. 732–736.
<https://doi.org/10.1080/09513590.2019.1576615>
 49. Fanis P., Skordis N., Toumba M. et al. Central precocious puberty caused by novel mutations in the promoter and 5'-UTR region of the imprinted *MKRN3* // Front. Endocrinol. (Lausanne). 2019. V. 10. P. e677.
<https://doi.org/10.3389/fendo.2019.00677>
 50. Maione L., Naule L., Kaiser U.B. Makorin RING finger protein 3 and central precocious puberty // Curr. Opin. Endocr. Metab. Res. 2020. V. 14. P. 152–159.
<https://doi.org/10.1016/j.coemr.2020.08.003>
 51. Ayca N., Savas-Erdeve S., Cetinkaya S. et al. Investigation of *MKRN3* mutation in patients with familial central precocious puberty // J. Clin. Res. Pediatr. Endocrinol. 2018. V. 10. P. e223–e229.
<https://doi.org/10.4274/jcrpe.5506>
 52. Зубкова Н.А., Колодкина А.А., Макрецькая Н.А. и др. Клиническая и молекулярно-генетическая характеристика 3 семейных случаев гонадотропинзависимого преждевременного полового развития, обусловленного мутациями в гене *MKRN3* // Проблемы эндокринологии. 2021. Т. 7. № 3. С. 55–61.
 53. Patti G., Malerba F., Calevo M.G. et al. Pubertal timing in children with Silver–Russell syndrome compared to those born small for gestational age // Front. Endocrinol. (Lausanne). 2022. V. 24. № 13. P. e975511.
<https://doi.org/10.3389/fendo.2022.975511>
 54. Ioannides Y., Lokulo-Sodipe K., Mackay D. et al. Temple syndrome: Improving the recognition of an underdiagnosed chromosome 14 imprinting disorder: an analysis of 51 published cases // J. Med. Genet. 2014. V. 51. P. 495–501.
 55. Juriaans A.F., Kerkhof G.F., Mahabier E.F. et al. Syndrome: Clinical findings, body composition and cognition in 15 patients // J. Clin. Med. 2022. V. 11. № 21. P. e6289.
<https://doi.org/10.3390/jcm11216289>
 56. Flippo C., Kolli V., Andrew M. et al. Precocious puberty in a boy with bilateral leydig cell tumors due to a somatic gain-of-function *LHCGR* variant // J. Endocr. Soc. 2022. V. 6. P. e10.
<https://doi.org/10.1210/jendso/bvac127>
 57. Partsch C.J., Japing I., Siebert R. et al. Central precocious puberty in girls with Williams syndrome // J. Pediatr. 2002. V. 141. № 3. P. 441–444.
 58. Kozel B.A., Barak B., Kim C.A. et al. Williams syndrome // Nat. Rev. Dis. Primers. 2021. V. 7. № 1. P. e42.
<https://doi.org/10.1038/s41572-021-00276-z>
 59. Nizon M., Andrieux J., Rooryck C. et al. Phenotype-genotype correlations in 17 new patients with an Xp11.23p11.22 microduplication and review of the literature // Am. J. Med. Genet. A. 2015. V. 167A. № 1. P. 111–122.
<https://doi.org/10.1002/ajmg.a.36807>
 60. Smith A., Leask K., Tomlin P., Donnai D. A familial dysmorphic condition with hypotonia, seizures and precocious puberty // Clin. Dysmorph. 2008. V. 17. P. 161–164.
 61. Luckie T.M., Danzig M., Zhou S. et al. A multicenter retrospective review of pediatric Leydig cell tumor of the testis // J. Pediatr. Hematol. Oncol. 2019. V. 41. № 1. P. 74–76.
<https://doi.org/10.1097/mp.0000000000001124>
 62. Menon K.M., Menon B. Structure, function and regulation of gonadotropin receptors – a perspective // Mol. Cell. Endocrinol. 2012. V. 356. № 1–2. P. 88–97.
<https://doi.org/10.1016/j.mce.2012.01.021>
 63. Lomniczi A., Ojeda S.R. The emerging role of epigenetics in the regulation of female puberty // Endocr. Dev. 2016. V. 29. P. 1–16.
<https://doi.org/10.1159/000438840>
 64. Feinberg A.P. The key role of epigenetics in human disease prevention and mitigation // N. Engl. J. Med. 2018. V. 378. № 14. P. 1323–1334.
<https://doi.org/10.1056/NEJMr1402513>
 65. Wright H., Aylwin C.F., Toro C.A. et al. Polycomb represses a gene network controlling puberty via modulation of histone demethylase *Kdm6b* expression // Sci. Rep. 2021. V. 11. № 1. P. e1996.
<https://doi.org/10.1038/s41598-021-81689-4>
 66. Lomniczi A., Wright H., Castellano J.M. et al. Epigenetic regulation of puberty via zinc finger protein-mediated transcriptional repression // Nat. Commun. 2015. V. 6. P. e10195.
<https://doi.org/10.1038/ncomms10195>

67. *Toro C.A., Wright H., Aylwin C.F. et al.* Trithorax dependent changes in chromatin landscape at enhancer and promoter regions drive female puberty // *Nat. Commun.* 2018. V. 9. № 1. P. e57.
<https://doi.org/10.1038/s41467-017-02512-1>
68. *Messina A., Langlet F., Chachlaki K. et al.* A microRNA switch regulates the rise in hypothalamic GnRH production before puberty // *Nat. Neurosci.* 2016. V. 19. № 6. P. 835–844.
<https://doi.org/10.1038/nn.4298>
69. *Heras V., Sangiao-Alvarellos S., Manfredi-Lozano M. et al.* Hypothalamic miR-30 regulates puberty onset via repression of the puberty-suppressing factor *Mkrn3* // *PLoS Biol.* 2019. V. 17. № 11. P. e3000532.
<https://doi.org/10.1371/journal.pbio.3000532>
70. *Canton A.P., Krepischi A.C., Montenegro L.R. et al.* Insights from the genetic characterization of central precocious puberty associated with multiple anomalies // *Hum. Reprod.* 2021. V. 25. № 36. P. 506–518.
<https://doi.org/10.1093/humrep/deaa306>
71. *Le Goff L.J., Cachin O., Rappaport R.* Precocious puberty associated with Silver's syndrome // *Arch. Fr. Pédiatr.* 1977. V. 34. № 9. P. 899–905.
72. *Wakeling E.L., Brioude F., Lokulo-Sodipe O. et al.* Diagnosis and management of Silver–Russell syndrome: first international consensus statement // *Nat. Rev. Endocrinol.* 2017. V. 13. № 2. P. 105–124.
<https://doi.org/10.1038/nrendo.2016.138>
73. *Grosvenor S.E., Davies J.H., Lever M. et al.* A patient with multilocus imprinting disturbance involving hypomethylation at 11p15 and 14q32, and phenotypic features of Beckwith–Wiedemann and Temple syndromes // *Am. J. Med. Genet. A.* 2022. V. 188. № 6. P. 1896–1903.
<https://doi.org/10.1002/ajmg.a.62717>
74. *Docherty L.E., Rezwan F.I., Poole R.L.* Mutations in *NLRP5* are associated with reproductive wastage and multilocus imprinting disorders in humans // *Nat. Commun.* 2015. V. 6. P. e8086.
<https://doi.org/10.1038/ncomms9086>
75. *Саженова Е.А., Никитина Т.В., Скрябин Н.А. и др.* Эпигенетический статус импринтированных генов в плаценте при привычном невынашивании беременности // *Генетика.* 2017. Т. 53. № 3. С. 364–377.
<https://doi.org/10.7868/S0016675817020096>
76. *Anvar Z., Chakchouk I., Demond H. et al.* DNA methylation dynamics in the female germline and maternal-effect mutations that disrupt genomic imprinting // *Genes (Basel).* 2021. V. 12. № 8. P. e1214.
<https://doi.org/10.3390/genes12081214>
77. *Pignata L., Cecere F., Verma A. et al.* Novel genetic variants of *KHDC3L* and other members of the subcortical maternal complex associated with Beckwith–Wiedemann syndrome or Pseudohypoparathyroidism 1B and multi-locus imprinting // *Clin. Epigenetics.* 2022. V. 14. P. e71.
<https://doi.org/10.1186/s13148-022-01292-w>
78. *Begemann M., Rezwan F.I., Beygo J. et al.* Maternal variants in *NLRP* and other maternal effect proteins are associated with multilocus imprinting disturbance in offspring // *J. Med. Genet.* 2018. V. 55. № 7. P. 497–504.
<https://doi.org/10.1136/jmedgenet-2017-105190>
79. *Monteagudo-Sanchez A., Hernandez Mora J.R., Simon C. et al.* The role of *ZFP57* and additional KRAB–zinc finger proteins in the maintenance of human imprinted methylation and multi-locus imprinting disturbances // *Nucl. Acids Res.* 2020. V. 48. № 20. P. 11394–11407.
<https://doi.org/10.1093/nar/gkaa837>
80. *Bessa D.S., Maschietto M., Aylwin C.F. et al.* Methylome profiling of healthy and central precocious puberty girls // *Clin. Epigenetics.* 2018. V. 10. № 1. P. e146.
<https://doi.org/10.1186/s13148-018-0581-1>
81. *Holland A., Manning K., Whittington J.* The paradox of Prader–Willi syndrome revisited: Making sense of the phenotype // *EBioMedicine.* 2022. V. 78. P. e103952.
<https://doi.org/10.1016/j.ebiom.2022.103952>

Genetics and Epigenetics of Precocious Puberty

E. A. Sazhenova^{a,*}, S. A. Vasilyev^a, L. V. Rychkova^b, E. E. Khramova^b, and I. N. Lebedev^a

^aScientific Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, 634050 Russia

^bScientific Center for Family Health and Human Reproduction, Irkutsk, 664003 Russia

*e-mail: elena.sazhenova@medgenetics.ru

Central precocious puberty (CPP) is caused by premature reactivation of the hypothalamic–pituitary–gonadal axis. Genetic, epigenetic and environmental factors play a crucial role in determining the timing of puberty. In recent years, the *KISS1*, *KISS1R*, *MKRN3* and *DLK1* variants have been identified as hereditary causes of CPP. The *MKRN3* and *DLK1* genes are imprinted, and therefore epigenetic modifications that alter the expression of these genes are also considered as a cause of precocious puberty. With the progression of CPP, epigenetic factors such as DNA methylation, posttranslational modifications of histones and non-coding RNAs can mediate the relationship between the influence of genetic variants and the environment. CPP is also associated with other short- and long-term adverse health effects. This is the basis for research aimed at understanding the genetic and epigenetic causes of CPP. The purpose of this review is to summarize the literature data on the molecular genetic and epigenetic mechanisms of the formation of CPP.

Keywords: central precocious puberty (CPP), gonadotropin-releasing hormone (GnRH), hypothalamic–pituitary–gonadal axis (HPG), genomic imprinting, *DLK1*, *KISS1*, *KISS1R*, *MKRN3*.

РЕГУЛЯЦИЯ ОДНОЦЕПОЧЕЧНЫХ И ДВУЦЕПОЧЕЧНЫХ КЛАСТЕРОВ
piРНК В ГЕРМИНАЛЬНЫХ И СОМАТИЧЕСКИХ ТКАНЯХ*Drosophila melanogaster* ЗАВИСИТ ОТ ГЕНА *rhino*© 2023 г. П. А. Миляева^{1,2}, А. Р. Лавренов^{1,3}, И. В. Кузьмин¹, А. И. Ким^{1,2}, Л. Н. Нефедова^{1, *}¹Московский государственный университет им. М.В. Ломоносова, Москва, 119234 Россия²Биологический факультет, университет МГУ-ППИ в Шэньчжэне, Шэньчжэнь, 518172 Китай³Институт проблем экологии и эволюции им. А.Н. Северцова Российской академии наук, Москва, 119071 Россия

*e-mail: nefedova@mail.bio.msu.ru

Поступила в редакцию 05.05.2023 г.

После доработки 17.06.2023 г.

Принята к публикации 29.06.2023 г.

Drosophila melanogaster — широко используемый объект генетики для изучения механизмов РНК-интерференции и регуляции мобильных генетических элементов (МГЭ), в частности системы piРНК-интерференции. Считается, что у *D. melanogaster* система piРНК-интерференции работает только в гонадах и существуют два пути биогенеза piРНК — с помощью одноцепочечных и двуцепочечных кластеров, содержащих дефектные копии МГЭ, транскрибируемых с одной или двух цепей соответственно. Предполагается, что функционирование двуцепочечных кластеров зависит от функции гена *rhino*, в то время как одноцепочечные кластеры транскрибируются и затем сплайсируются независимо от *rhino*. В настоящей работе мы показали, что ген *rhino* участвует в регуляции сплайсинга одноцепочечного кластера *flamenco*. Более того, система piРНК-интерференции необходима для регуляции некоторых ретротранспозонов не только в гонадах, но и в других органах.

Ключевые слова: *Drosophila melanogaster*, piРНК, *rhino*, ретротранспозоны, РНК-интерференция.

DOI: 10.31857/S0016675823120056, EDN: QDLMAL

У эукариот известны три системы РНК-интерференции, необходимые для подавления транспозиционной активности мобильных генетических элементов (МГЭ), — это системы mi-, si- и piРНК-интерференции [1–3]. Источником miРНК и siРНК могут быть одиночные вставки последовательностей МГЭ и инфицирующих клетку вирусов, в то время как источником piРНК служат как одиночные копии МГЭ, так и кластеры — скопления, как правило, дефектных копий МГЭ [4, 5].

Система piРНК-интерференции, участвующая в подавлении ретротранспозонов, хорошо изучена в тканях яичников *Drosophila melanogaster*. В клетках зародышевого пути большая часть работающих piРНК-кластеров является двуцепочечной (*42AB*, *38C*, *80EF*) [6]. Такие кластеры представлены МГЭ, не имеющими единого направления транскрипции. Их транскрипция запускается с обеих цепей с неканонических сайтов. За счет этого получают длинные транскрипты, отдельные участки которых являются антисмысловыми по отношению к мРНК МГЭ. Транскрипты двуцепочечных кластеров без сплайсинга отправляются в цитоплазму, где в околоядерной области “Нюаж” вовлекаются в цикл “пинг-понг” с последующим процессингом и пре-

образованием в piРНК [7]. Блокирование сплайсинга транскриптов таких кластеров необходимо, так как именно в интронах содержатся последовательности, необходимые для получения специфичных piРНК.

Важную роль в транскрипции и блокировании сплайсинга транскриптов двуцепочечных кластеров играет комплекс белков Rhino (Rhi), Deadlock (Del) и Cutoff (Cuff). Rhi, представитель семейства гетерохроматиновых белков Hrp1, связывается с кластером, распознавая гистоновые метки H3K9me3. Rhi взаимодействует с Del, который, в свою очередь, связывается с Cuff, и вместе они образуют комплекс RDC. Данный комплекс способствует посадке РНК-полимеразы на неканонические старты транскрипции, а также препятствует кэпированию и полиаденилированию транскриптов таких кластеров. Активный сплайсинг транскриптов двуцепочечных кластеров без работы комплекса RDC, наоборот, приводит к неправильной работе системы piРНК-интерференции и повышению экспрессии некоторых МГЭ [8].

Ген *rhino* *D. melanogaster* является настолько быстро эволюционирующим геном, что Rhi самого эволюционно близкого вида, *D. simulans*, не спо-

способен образовывать комплекс с Del *D. melanogaster* [9]. Стремительная эволюция *rhino* объясняется высокой скоростью изменчивости МГЭ. В отличие от других белков семейства Hrp1 Rhi экспрессируется в основном в гонадах как самцов, так и самок дрозофилы, и отсутствие этого белка ведет к активному сплайсингу транскриптов двуцепочечных кластеров и повышению уровня транспозиции МГЭ [6].

В соматических клетках яичников дрозофилы piRNA образуются из длинных транскриптов одноцепочечных кластеров, таких как *flamenco*. Длинные транскрипты процессируются путем альтернативного сплайсинга и последующего нарезания на piRNA. Считается, что *flamenco* транскрибируется независимо от комплекса RDC и его транскрипты перемещаются из ядра в белковый комплекс “Уб-тельце”, локализованный на внешней мембране митохондрий. После процессинга пре-piRNA модифицируются метилтрансферазой Hen1 и транспортируются в ядро в составе комплекса с белком PIWI, где происходит подавление активности ретротранспозонов на уровне транскрипции с привлечением белков для гетерохроматинизации Asterix (Arx), Hrp1a и Maelstrom (Mael) [10]. Стоит отметить, что метилтрансфераза Hen1 является неотъемлемым компонентом процессинга любых si- и piRNA как в соматических клетках зародышевой линии, так и в цикле “пинг-понг” в клетках зародышевого пути. Более того, этот фермент выполняет ту же функцию в нервной системе взрослых особей нематоды *Caenorhabditis elegans* [11]. Таким образом, кластер *flamenco* участвует исключительно в создании геновых piRNA и не принимает участия в интерференции. Вероятно, поэтому считается, что кластер в норме подвергается альтернативному сплайсингу и процессируется независимо от *rhino*. На нервной системе личинок дрозофилы было продемонстрировано, что снижение сплайсинга транскриптов этого кластера приводит к повышению количества пре-piRNA и снижению зрелых piRNA [12], однако эффект от повышения количества сплайсированных форм этого кластера для МГЭ ранее не рассматривался.

Считается, что пути siRNA и piRNA работают в гонадах, в то время как путь siRNA работает еще и в соматических тканях [4]. Однако исследования системы piRNA-интерференции на разных модельных объектах в последнее десятилетие позволили выявить функционирование этой системы не только в клетках зародышевого пути, но и в нервной системе. Так, нормальный процессинг транскриптов piRNA-кластеров необходим для обеспечения двигательной активности как личинок, так и взрослых мух *D. melanogaster*. Основными участниками контроля piRNA в нервной системе, согласно исследованию [12], являются гены *aubergine* (*aub*), *cabeza* (*caz*) и *rhino*, а также одноцепочечный кластер *flamenco*. У позвоночных и

беспозвоночных piRNA принимают участие в ответе на повреждения периферической нервной системы: у нематоды piRNA участвуют в подавлении роста аксонов в ответ на повреждение, а у мыши экспрессия piRNA повышается при повреждении нервов, и одна из piRNA отвечает за миграцию шванновских клеток [13].

Ранее было показано, что МГЭ *gypsy* повышает экспрессию в тканях головы *D. melanogaster* при мутациях в генах *aub* и *argonaute 3* (*ago3*) [14]. Ген *rhino* комплементарно взаимодействует с этими генами благодаря участию в транскрипции и сплайсинге кластеров piRNA. Таким образом, для отдельных компонентов системы piRNA было показано, что они вносят вклад в подавление экспрессии отдельных МГЭ в соматических тканях с помощью цикла “пинг-понг”.

Ранее нами была показана корреляция между снижением экспрессии *rhino* и повышением экспрессии сплайсированных форм *flamenco* у линии с фенотипом *flamenco*, SS7K, характеризующейся нарушением контроля транспозиции МГЭ *gypsy* [15]. В настоящей работе мы провели нокдаун гена *rhino* у *D. melanogaster* и измерили экспрессию наиболее изученных кластеров piRNA и МГЭ в тканях яичников, головы, корпуса (тканей, оставшихся после отделения головы и яичников) у семидневных самок, чтобы оценить роль комплекса RDC в регуляции системы piRNA-интерференции за пределами гонад, а также установить роль Rhi в регуляции синтеза piRNA не только двуцепочечных кластеров, которые работают в клетках зародышевого пути, но и одноцепочечных кластеров, которые, как считается, работают исключительно в соматических тканях гонад.

МАТЕРИАЛЫ И МЕТОДЫ

Линии Drosophila melanogaster и условия культивирования

Для нокдауна гена *rhino* использовали систему UAS-GAL4, основанную на скрещивании драйверной и трейсерной линий *D. melanogaster*, несущих соответственно конструкцию с геном, который кодирует белок GAL4 и находится под тканеспецифичным промотором, и конструкцию со шпилькой для siRNA-интерференции под промотором UAS. Для создания гибридов с нокдауном исследуемого гена использовали самок линии *rhi* (Vienna Drosophila Resource Center (VDRC) ID 101620) и самцов линии *tubP-Gal4* (*y1w**; P{*tubP-GAL4*}LL7/TM3, Sb1Ser1; Bloomington Drosophila Stock Center BL5138; запускает нокдаун во всех тканях). В качестве контроля использовали гибрид самок линии *w*[118] (на ее основе была получена линия *rhi*) и самцов *tubP-Gal4*. Гибриды от скрещивания драйверной и трейсерной линий культивировали при температуре 27°C. Данная

температура была оптимальной для работы драйверной конструкции. Эксперимент проводили в трех повторах.

Выделение РНК, обратная транскрипция и ПЦР в реальном времени

Выделение РНК семидневных самок с нокдауном гена *rhino* и самок контрольных гибридов проводили из тканей яичников, головы и корпуса с помощью реактива ExtractRNA (Евроген) после вскрытия мух в буфере PBS. Перед постановкой реакции обратной транскрипции образцы РНК обрабатывали ДНКазой I (Thermo Fisher Scientific). Для обратной транскрипции использовали набор MMLV-RT Kit (Евроген). Обратную транскрипцию ставили со случайным праймером (так как экспрессия большинства кластеров в нашем опыте не превышает экспрессию МГЭ). ПЦР проводили в присутствии SYBR Green I (Евроген) в амплификаторе MiniOpticon Real-Time PCR System (Bio-Rad). В опыте анализировали относительную экспрессию ретротранспозонов *gypsy*, *Tirant*, *roo*, *scopia*, *blood*, *HeT-A*, *TART-A*, *TART-B*, *TART-C* и кластеров *38C*, *20A*, *42AB*, *flamenco*, нормированную на экспрессию генов α *Tub84D*, *Rpl40* и *EloB*. Для контроля нокдауна гена *rhino* использовали праймеры к соответствующему гену (табл. 1). Статистическую обработку результатов ПЦР проводили с помощью критерия Манна–Уитни в среде программирования R. В каждой выборке использовали от пяти до 14 образцов. Праймеры для кластера *42AB* были взяты из работы Zh. Zhang и соавт. [6], праймеры для кластеров *38C* и *20A* были взяты из работы J. Nig и соавт. [16], праймеры для кластера *flamenco* были подобраны нами к сплайсированной форме транскрипта согласно сиквенсу из базы данных NCBI (Gene ID: 26067356) [17]. Все праймеры подобраны к специфичным частям кластеров и не отжигаются на копиях мобильных элементов, находящихся в других точках генома.

Для оценки экспрессии piРНК картировали piРНК из базы данных (<https://www.pirnadb.org/>) на аннотированную последовательность кластера *flamenco* и выбрали piРНК *11620*, так как она оказалась картирована в интрон кластера. В качестве референсной РНК использовали miРНК *U6*. Для выделения микроРНК использовали набор для выделения суммарной РНК и микроРНК из реагента “Лира” (Biolabmix). Выборки содержали до шести образцов. Обратную транскрипцию проводили с помощью специфичного шпилечного праймера. Для количественной ПЦР использовали праймеры к шпильке и к целевой РНК (табл. 1). ПЦР проводили в присутствии SYBR Green I (Евроген) в амплификаторе MiniOpticon Real-Time PCR System (Bio-Rad). В качестве референсной РНК использовали miРНК *U6*.

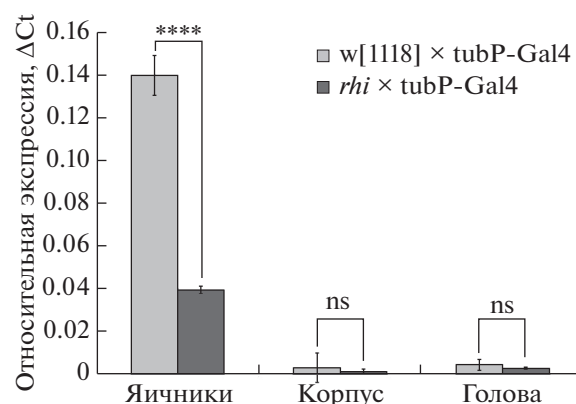


Рис. 1. Относительная экспрессия гена *rhino* в тканях яичников, корпуса и головы у семидневных самок гибридов с нокдауном гена *rhino* и контрольных гибридов (**** — $p < 0.0001$ по критерию Манна–Уитни; ns, not significant — статистически незначимое изменение).

РЕЗУЛЬТАТЫ

Зависимость сплайсинга кластеров piРНК от уровня экспрессии гена *rhino*

Для подтверждения роли гена *rhino* в контроле активности ретротранспозонов разных семейств мы получили гибриды с его нокдауном от скрещивания линий *D. melanogaster* *rhi* и *tubP-Gal4* и проанализировали экспрессию этого гена в тканях яичников, головы и корпуса в сравнении с теми же тканями у контрольных гибридов (*w[1118] × tubP-Gal4*) (рис. 1). Как и ожидалось, в контрольных гибридах экспрессия гена *rhino* была наиболее высокой в тканях яичников, где система piРНК-интерференции работает наиболее активно. Примечательно, что мы также обнаружили экспрессию гена *rhino* в небольшом количестве в тканях корпуса и головы. У гибридов с нокдауном гена *rhino* мы увидели подавление экспрессии этого гена во всех тканях, однако статистически значимым оно было только в яичниках. Мы предполагаем, что это могло быть вызвано низкой экспрессией *rhino* в тканях корпуса и головы, а также различиями в контроле. Так как в контроле видим стабильно низкую экспрессию *rhino* считаем, что в этих тканях нокдаун этого гена был осуществлен успешно.

Далее мы провели сравнительный анализ экспрессии сплайсированных и несплайсированных форм предшественников piРНК, транскрибирующихся с одно- и двуцепочечных кластеров и участвующих в сайленсинге МГЭ. Для анализа мы использовали наиболее изученные кластеры piРНК (*flamenco* (одноцепочечный) и *42AB* (двуцепочечный)) [18, 19] (рис. 2, а, б).

Для плохо изученных кластеров *20A* (одноцепочечный) и *38C* (двуцепочечный) мы провели анализ только несплайсированных форм, по-

Таблица 1. Праймеры для обратной транскрипции (ОТ) и ПЦР

Название	Прямой праймер	Обратный праймер
Гены		
<i>αTub84D</i>	5'-GTGCATGTTGTCCAAACASSAC-3'	5'-AGAACTCTCCCTCCCTCCATA-3'
<i>EloB</i>	5'-GCACAAACATACACATCAGC-3'	5'-TTTCTACTTTCGCTTGCAACC-3'
<i>RpL40</i>	5'-CTGCGTGTGGTATCATTG-3'	5'-CAGGTTGTTGGTGTGTCC-3'
<i>rhino</i>	5'-CTTCGAGTCGGAGGTCCTTCA-3'	5'-CTGCACCGACTTGGAAATGTT-3'
Кластеры		
<i>flamenco-spliced</i>	5'-CATCAGCTCAGCAGCAGTGTA-3'	5'-AGTCCAAATGCTCGCTTGAAA-3'
<i>flamenco-unspliced</i>		5'-GACTTAACACTTACCGCTTGAAA-3'
<i>42AB-spliced</i>	5'-GCAGTTGCCGTCCTCTCCCTT-3'	5'-TGGGTCAAAAGTGCAGCAGTT-3'
<i>42AB-unspliced</i>		5'-TACGGGAATATAATCGCAGCAGTT-3'
<i>38C 1</i>	5'-AATGGCTAGTTGCTACCAGACAG-3'	5'-CGGGTCTTCTCTCAAAAGCGCAATC-3'
<i>38C 2</i>	5'-SACAAAATGCCGCTGGAAA-3'	5'-AGCCAAAACCCCTGTGTTGTA-3'
<i>20A</i>	5'-GCCTACGCAGAGGCCCTAAGT-3'	5'-CAGATGTGGTCCAGTTGTGTC-3'
Ретротранспозоны		
<i>gyrus</i>	5'-CTGCTGAAAAGACGGCATTTATC-3'	5'-AGAACTTTGCCTTGCCSAGAT-3'
<i>Tirant</i>	5'-AACGCTATTCAATTCTGCAAC-3'	5'-AGGGTTCTCCTAACTACGTC-3'
<i>copia</i>	5'-CTTCAGTGATGGACAAC TG-3'	5'-CAGTGTAACTCTCATGGTCAAT-3'
<i>roo</i>	5'-ACAGATTCCCACCTTCTTTG-3'	5'-AGCCGAGGAGTAATGAGTA-3'
<i>blood</i>	5'-ATGCGACAACGATACAGG-3'	5'-TTTGTCAAGCATTTATCGGCG-3'
<i>HeT-A</i>	5'-TCAGGATACAGGCAAGGAC-3'	5'-CGACTATCCACGAGCATGG-3'
<i>TART-A</i>	5'-GCAGCAGTTTCCATTTGAC-3'	5'-AAGAACACACCATCGCTTG-3'
<i>TART-B</i>	5'-ATAGGGTATGGCACGAAAGG-3'	5'-GGGCAGACCAAGTTTATCC-3'
<i>TART-C</i>	5'-ATAGGGTATGGCACGAAAGG-3'	5'-AAGTAGTGGTCCGAGATTGTC-3'
Шпилечные праймеры для ОТ		
<i>U6</i>	5'-CTCAACTGGTGTCTGAGTCGGCAATTCAGTTGAGAAAAATGTGG-3'	
<i>11620</i>	5'-CTCAACTGGTGTCTGAGTCGGCAATTCAGTTGAGGATGGCA-3'	
Праймеры для микроРНК		
Универсальный праймер для шпильки		5'-ACTGGTGTCTGTTGGAGTCS-3'
<i>U6</i>	5'-GCTGGGATACAGAGAAAGATTAG-3'	
<i>11620</i>	5'-GCTGGGTCAACTAGTATTCTTG-3'	

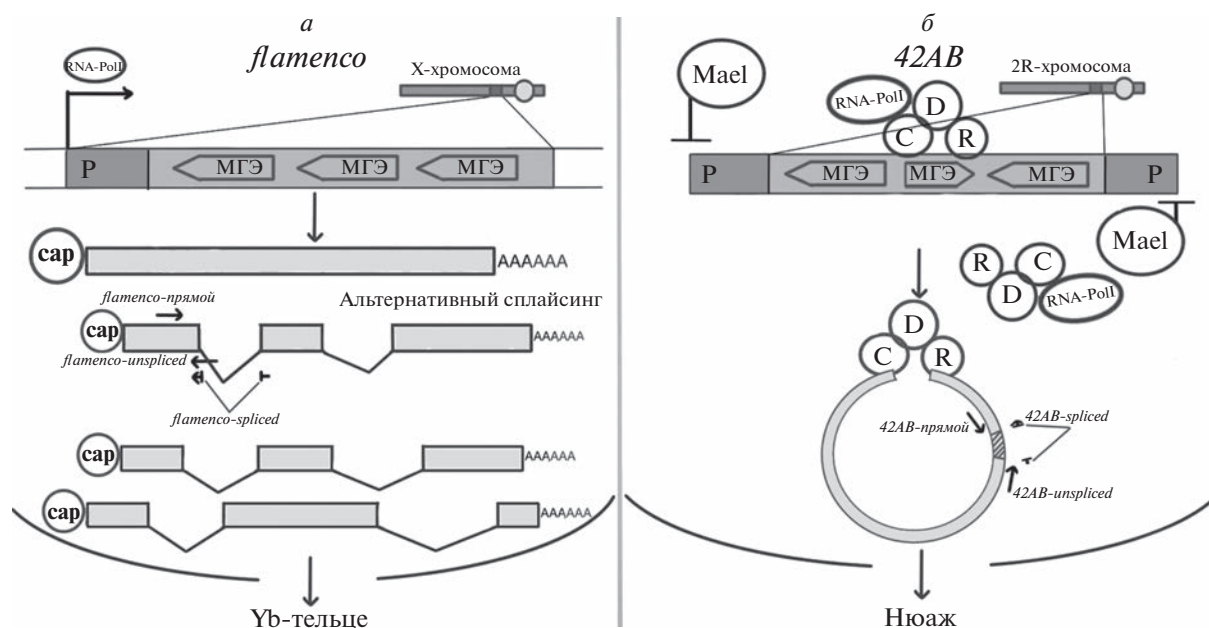


Рис. 2. Регуляции одноцепочечных и двуцепочечных кластеров *flamenco* и *42AB* и схемы локализации праймеров для выявления сплайсированных и несплайсированных форм транскриптов. *a* – особенности транскрипции одноцепочечных кластеров: транскрипция идет с канонического промотора в одном направлении, транскрипты сплайсируются (праймеры для отличия сплайсированных и несплайсированных форм: *flamenco-прямой* (общий праймер), *flamenco-spliced* (для сплайсированной формы), *flamenco-unspliced* (для несплайсированной формы)); *б* – особенности транскрипции двуцепочечных кластеров: транскрипция идет с неканонических промоторов в обоих направлениях за счет комплекса RDC и белка Mael, транскрипты не сплайсируются, так как RDC блокирует экзипирование и полиаденилирование транскриптов (праймеры для отличия сплайсированных и несплайсированных форм: *42AB-прямой* (общий праймер), *42AB-spliced* (для сплайсированной формы), *42AB-unspliced* (для несплайсированной формы)). Р – промотор.

сколько альтернативный сплайсинг для них пока не показан. Анализ для кластера *38C* проводили в двух точках, чтобы определить, равнозначно ли отразится на экспрессии разных частей кластера снижение экспрессии *rhino*.

Измерив относительную экспрессию несплайсированных форм кластеров *20A*, *38C*, *42AB* и *flamenco* между тканями контрольных гибридов и гибридов с нокдауном *rhino*, мы в большинстве случаев не обнаружили значимого изменения экспрессии этих кластеров при снижении экспрессии *rhino* (рис. 3,а). Исключение составило изменение экспрессии кластеров *42AB*, *38C* и *20A* в тканях корпуса. Возможно, это связано с особенностями регуляции кластеров в этих тканях. Примечательно, что в яичниках, основном месте работы кластеров рiРНК, при нокдауне *rhino* повышается экспрессия только сплайсированных форм кластеров *flamenco* и *42AB*. Это наводит на мысль, что *rhino* может регулировать и сплайсинг *flamenco*.

При сравнении относительной экспрессии сплайсированных и несплайсированных форм кластера *42AB*, которые являются мишенью Rhi, мы увидели значительное повышение количества сплайсированных форм во всех типах тканей гибридов с нокдауном *rhino* относительно контроля

(рис. 3,б). Более того, значительное повышение экспрессии сплайсированных форм по отношению к экспрессии несплайсированных мы также обнаружили для кластера *flamenco*, который, в отличие от остальных кластеров, у взрослых самок экспрессируется только в яичниках. Однако при подсчете отношения количества сплайсированных форм к несплайсированным мы обнаружили, что при нокдауне *rhino* это соотношение повышается не только для кластера *42AB* в яичниках и голове, но и для кластера *flamenco* в яичниках, хотя уровень значимости в случае последнего ниже.

Мы также оценили экспрессию рiРНК *11620*, которая картирована в интроне кластера *flamenco*, однако не нашли различий между контролем и мутантом с нокдауном *rhino* (рис. 3,в).

Таким образом, роль гена *rhino* не ограничивается только генеративными и соматическими тканями яичников, но и распространяется на другие соматические ткани, где *rhino* так же, как и в тканях яичников, регулирует сплайсинг кластеров рiРНК. Помимо этого мы не исключаем, что процессинг транскриптов кластера *flamenco* может зависеть от *rhino* так же, как процессинг транскриптов кластера *42AB*. Однако на уровень экспрессии рiРНК из этого кластера повышенная частота сплайсинга, вероятно, не влияет.

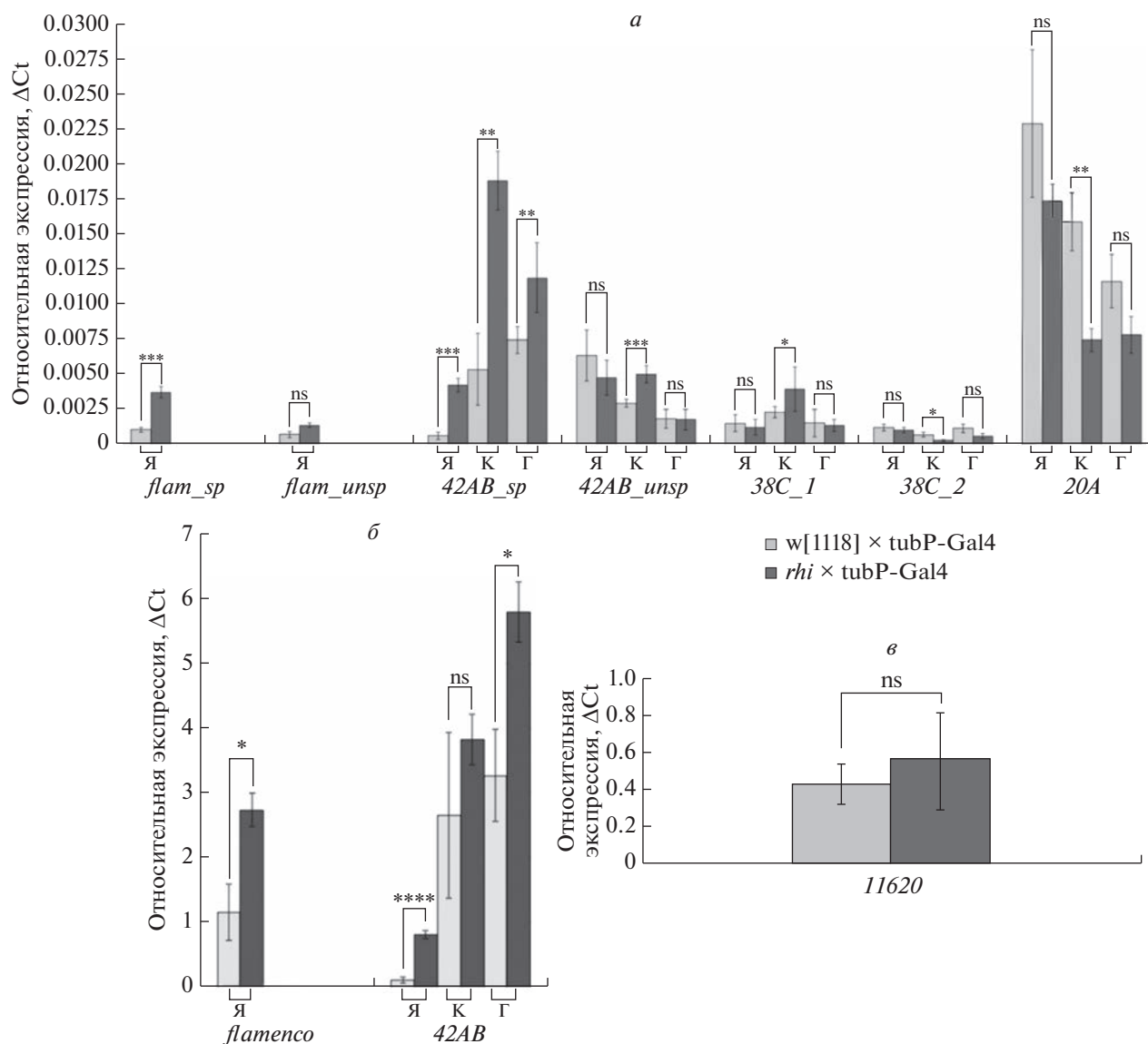


Рис. 3. Относительный уровень транскрипции различных кластеров piРНК в тканях яичников, головы и корпуса. *а* – относительный уровень транскрипции сплайсированных и несплайсированных форм одноцепочечных и двуцепочечных кластеров; *б* – соотношение уровней транскрипции сплайсированных и несплайсированных форм транскриптов кластеров *flamenco* (одноцепочечный) и *42AB* (двуцепочечный); *в* – относительный уровень транскрипции piРНК 11620 (кластер *flamenco*) в яичниках. Для рис. 3, 4: Я – яичники, К – корпус, Г – голова (* – $p < 0.05$, ** – $p < 0.01$, *** – $p < 0.001$, **** – $p < 0.0001$ по критерию Манна–Уитни; ns, not significant – статистически незначимое изменение).

Зависимость уровня экспрессии ретротранспозонов от уровня экспрессии гена *rhino*

Далее мы провели исследование транскрипции МГЭ в тканях нокдаун-мутанта по гену *rhino* и проанализировали взаимосвязь транскрипции МГЭ с уровнем сплайсированных/несплайсированных форм исследуемых кластеров. Для анализа были выбраны ДКП-ретротранспозоны, которые ранее демонстрировали повышение экспрессии в ответ на стресс либо повышали экспрессию в тканях головы при мутациях центральных участников цикла “пинг-понг”, *aub* либо *ago3*, по данным

P. Perrat и соавт. [14]: *gypsy*, *roo*, *copia*, *Tirant* и *blood*. Также мы взяли несколько теломерных LINE-элементов, которые, как показано ранее, по-разному реагируют на снижение экспрессии *rhino*: *HeT-A*, *TART-A*, *TART-B* и *TART-C* [20].

С помощью базы данных FlyBase (<http://flybase.org>) [21] мы проверили наличие копий изучаемых нами ретротранспозонов в исследуемых кластерах (*flamenco*, *42AB*, *20A* и *38C*) и обнаружили, что в этих кластерах присутствуют только копии ДКП-ретротранспозонов, в то время как теломерные LINE-элементы в них не представлены. Кла-

стер *flamenco* содержит ДКП-ретротранспозон *gypsy*, в кластере 20A есть копии *roo*, *copia* и *blood*, кластер 42AB содержит *gypsy*, *roo* и *Tirant*, а в кластере 38C находятся *gypsy*, *copia* и *roo*.

ДКП-ретротранспозоны и теломерные LINE-элементы осуществляют транспозицию путем репликативного синтеза, однако в отличие от ДКП-ретротранспозонов элементы *HeT-A*, *TART-A*, *TART-B* и *TART-C* встраиваются после районов TAS в теломеры, замещая таким образом функцию теломеразы. LINE-элемент *HeT-A* в ходе эволюции утерял рамку считывания, кодирующую эндонуклеазу и обратную транскриптазу, поэтому способен осуществлять транспозицию только за счет белков, кодируемых родственными LINE-элементами. В настоящее время считается, что *rhino* запускает транскрипцию не только с кластеров, но и с одиночных копий как ДКП-ретротранспозонов, так и LINE-элементов, однако его роль в транскрипции одноцепочечных кластеров и в соматических тканях за пределами гонад остается неизученной [22].

Мы обнаружили, что в контрольных гибридах ретротранспозоны по-разному экспрессируются в тканях яичников, корпуса и головы. Большинство исследованных МГЭ закономерно экспрессируются на более низком уровне в яичниках по сравнению с другими тканями. Однако исключение составил МГЭ *Tirant*, который практически не работает в тканях корпуса и головы, в отличие от яичников. При этом он, скорее всего, содержится в двуцепочечном кластере 42AB. Также следует отметить, что большинство ретротранспозонов имеют различный уровень транскрипции в голове и тканях корпуса, что может указывать на различный характер регуляции транскрипции МГЭ в разных соматических тканях.

При нокдауне *rhino* ДКП-ретротранспозоны *gypsy*, *copia* и теломерные LINE-элементы *HeT-A* и *TART-C* значительно повышают уровень экспрессии в тканях яичников и корпуса. Более того, теломерный LINE-элемент *TART-B* значительно повышает свою экспрессию не только в тканях яичников, но и во всех соматических тканях, в то время как другой теломерный ретротранспозон, *TART-A*, только в тканях головы (рис. 4, а, б). Для ДКП-ретротранспозона *blood* и *Tirant* мы обнаружили падение экспрессии только в одной из исследуемых тканей. Для ДКП-ретротранспозона *roo* мы обнаружили падение экспрессии в тканях корпуса, в то время как в тканях головы изменений не было.

Таким образом, изменения в считывании кластеров могут приводить к повышению уровня транскрипции одних ретротранспозонов и к его снижению у других. Возможно, что при снижении экспрессии *rhino* транскрипция кластеров изменяется и либо сплайсированные транскрип-

ты способны быть источником рiРНК против некоторых ретротранспозонов, либо эти МГЭ активно подавляются системой siРНК-интерференции, куда могут направляться транскрипты кластеров после сплайсинга. Более того, мы увидели повышение транскрипции одних участков кластеров и снижение других, что может быть причиной усиления образования пре-si и рiРНК против *roo*, но падения количества антисмысловых транскриптов против *gypsy*, *copia*, *TART-B*, *TART-C* и *HeT-A*.

ОБСУЖДЕНИЕ

Белок Rhi входит в состав комплекса RDC, взаимодействуя гетерохроматиновыми метками H3K9Me3, маркирующими кластеры рiРНК [10]. С помощью хромотеневого (chromoshadow) домена Rhi взаимодействует с белком Del, который, в свою очередь, связывается с белком Cuff. Роль Cuff заключается в конкурировании за кэп на 5'-конце и препятствовании полиаденилированию транскрипта рiРНК. Кроме того, Cuff взаимодействует с комплексом TREX, который отвечает за транспорт пре-рiРНК из ядра. Таким образом, Rhi является фундаментом для построения комплекса RDC, который необходим для запуска транскрипции двуцепочечных кластеров с неканонических сайтов, участвуя в транскрипции рiРНК как с двуцепочечных кластеров, так и с единичных вставок ретротранспозонов. Новые ретротранспозоны становятся кластерами после присоединения меток H3K9me3 благодаря привлечению гетерохроматиновых белков с помощью PIWI и препятствию началу транскрипции с канонических сайтов белком Mael [9, 19, 23–25].

В нашей работе мы продемонстрировали, что функция гена *rhino* необходима для функционирования как двуцепочечных, так и одноцепочечных кластеров. Мы обнаружили, что снижение экспрессии *rhino* приводит к активному сплайсингу транскриптов одноцепочечного кластера *flamenco* так же, как и двуцепочечного кластера 42AB (при этом экспрессия несплайсированных форм кластеров в яичниках остается на прежнем уровне), что указывает на возможную зависимость процессинга транскриптов *flamenco* от комплекса RDC, фундаментом которого является *rhino*. Однако изменений в экспрессии рiРНК этого кластера мы не наблюдаем. Подобные эффекты наблюдали и другие исследователи, изучая вопрос взаимодействия *rhino* с кластером *flamenco* с помощью ChIP-seq [6, 26]. В данных работах взаимодействие *flamenco* и *rhino* не было достоверно установлено, однако в нашем случае мы видим слабое влияние нокдауна *rhino*, оказываемое на сплайсинг. Так как транскрипция кластеров — это сложный процесс, который не до конца изучен, мы не исключаем, что кластер *flamenco* может обладать минорными

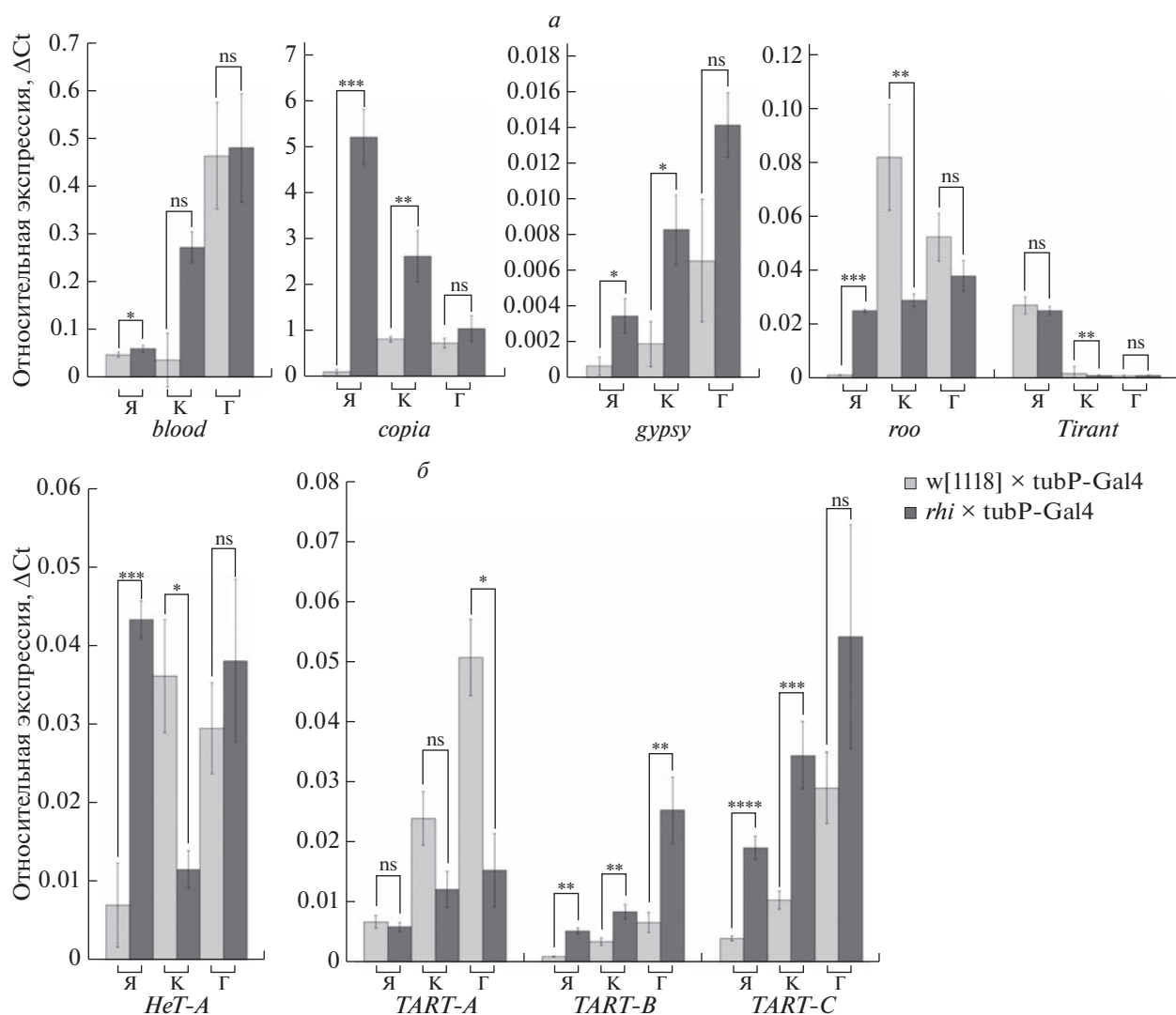


Рис. 4. Относительная экспрессия различных ретротранспозонов в тканях яичников, головы и корпуса. *а* — относительная экспрессия ДКП-ретротранспозонов; *б* — относительная экспрессия теломерных LINE-элементов.

механизмами регуляции. Например, кластер *42AB* транскрибируется с неканонических сайтов только в одном направлении и требует присутствия комплекса RDC только для транскрипции одной из цепей [27]. Таким образом, в нормальных условиях присутствует некоторое количество сплайсированных и несплайсированных форм кластеров. При нокдауне *rhino* баланс сдвигается в сторону увеличения количества сплайсированных форм, и вероятно происходит вырезание участков, которые в норме используются для процессинга piPHK. Повышение количества сплайсированных форм кластера *42AB* в тканях яичников и головы в ответ на понижение экспрессии *rhino* подтверждает наше предположение, что кластеры могут выполнять роль источника piPHK и в соматических тканях. В работах других исследователей было показано, что *rhino* экспрессируется в

соматических клетках и в клетках зародышевого пути яичников дрозофилы [28]. Считается, что наличие сплайсинга приводит к отправке транскриптов по пути siPHK, что было показано для кластеров *42AB* и *flamenco* в голове у дрозофилы [29]. Возможно, при отсутствии *rhino* система siPHK-интерференции частично замещает piPHK-интерференцию.

Показано, что одиночные копии I-элемента, относящегося к классу LINE-элементов, могут играть роль источников piPHK, причем эти кластеры делятся на сильные и слабые в зависимости от обогащения данных участков генома белком Rhi. Более того, старые копии I-элемента в линии R (чувствительна к I-элемент-индуцируемому гибридному дисгенезу) могут являться источниками первых piPHK для установки меток H3K9me3 на новых инсерциях [22]. Также источниками

piРНК может служить сателлитная ДНК с метками H3K9me3; транскрипты с таких участков присутствуют не только в гонадах, но и различных тканях тела [18]. Также известно, что у дрозофилы некоторые кластеры транскрибируются полоспецифично [30].

В настоящей работе мы увидели, что снижение экспрессии *rhino* сказывается неоднозначно на экспрессии различных ретротранспозонов в соматических тканях. Большинство ретротранспозонов отреагировали на снижение экспрессии данного гена только в яичниках. Ретротранспозоны *coria*, *gypsy*, *TART-C* и *HeT-A* повысили уровень экспрессии только в гонадах и корпусе, а ретротранспозон *TART-B* начал активно экспрессироваться во всех тканях.

Известно, что теломерные сателлитные районы (TAS), а также теломерные транспозоны *HeT-A*, *TAHRE* и *TART* являются *rhino*-зависимыми источниками piРНК и транскрибируются с двух цепей [22]. Более того, гетерохроматиновые метки H3K9me3 обнаруживаются у дрозофилы и на теломерах в соматических тканях [20]. Повышение активности теломерных транспозонов не обязательно приводит к удлинению теломер за счет активной транспозиции. Так, мутация гена *hiphop* (привлекает белок Hrp1a для метилирования участков, содержащих МГЭ) приводит к значительному увеличению экспрессии теломерных ретротранспозонов без повышения транспозиционной активности [22]. Показано, что мутации генов *spnE* и *enok* системы piРНК-интерференции приводят к повышению экспрессии различных семейств ретротранспозонов, включая теломерные [8].

Регуляция различных ретротранспозонов может быть тканеспецифичной, что мы наблюдаем в случае ретротранспозонов *coria*, *TART-C*, *HeT-A*, *gypsy* и *TART-B*, первые четыре из которых регулируются компонентами системы piРНК только в яичниках и корпусе, а последний — во всем теле. Мы предполагаем, что различия в экспрессии ретротранспозонов в разных тканях единого организма могут зависеть как от работы кластеров в этих тканях, так и от дополнительных факторов транскрипции.

Авторы выражают благодарность старшему научному сотруднику Института молекулярной генетики РАН М.С. Кленову за предоставление линии *ghi* и старшему научному сотруднику Института биологии развития им. Н.К. Кольцова РАН Ю.Е. Воронцовой за предоставление линии *tubP-Gal4*.

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

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Park E.G., Ha H., Lee D.H. et al. // Genomic analyses of non-coding RNAs overlapping transposable elements and its implication to human diseases // Intern. J. Mol. Sciences. 2022. V. 23. P. 8950. <https://doi.org/10.3390/ijms23168950>
2. Schwarz D.S., Hátvágner G., Haley B., Zamore P.D. Evidence that siRNAs function as guides, not primers, in the drosophila and human RNAi pathways // Mol. Cell. 2002. V. 10. P. 537–548. [https://doi.org/10.1016/S1097-2765\(02\)00651-2](https://doi.org/10.1016/S1097-2765(02)00651-2)
3. Guerreiro G.M.P. What makes transposable elements move in the *Drosophila* genome // Heredity. 2012. V. 108. P. 461–468. <https://doi.org/10.1038/hdy.2011.89>
4. Carthew R.W., Sontheimer E.J. Origins and mechanisms of miRNAs and siRNAs // Cell. 2009. V. 136. P. 642–655. <https://doi.org/10.1016/j.cell.2009.01.035>
5. Yamanaka S., Siomi M.C., Siomi H. PiRNA clusters and open chromatin structure // Mobile DNA. 2014. V. 5. P. 22. <https://doi.org/10.1186/1759-8753-5-22>
6. Zhang Zh., Wang J., Schultz N. et al. The HP1 homolog Rhino anchors a nuclear complex that suppresses piRNA precursor splicing // Cell. 2014. V. 157. № 6. P. 1353–1363. <https://doi.org/10.1016/j.cell.2014.04.030>
7. Claycomb J.M. Emerging from the Clouds: Vasa helix sheds light on piRNA amplification // Dev. Cell. 2014. V. 29. P. 632–634. <https://doi.org/10.1016/j.devcel.2014.06.009>
8. Tsai S.-Y., Huang F. Acetyltransferase Enok regulates transposon silencing and piRNA cluster transcription // PLoS Genetics. 2021. V. 17. № 2. <https://doi.org/10.1371/journal.pgen.1009349>
9. Yu B., Lin Y.A., Parhad S.S. et al. Structural insights into Rhino–Deadlock complex for germline piRNA cluster specification // EMBO Reports. 2018. V. 19. <https://doi.org/10.15252/embr.201745418>
10. Théron E., Dennis C., Brassat E., Vauray C. Distinct features of the piRNA pathway in somatic and germ cells: From piRNA cluster transcription to piRNA processing and amplification // Mobile DNA. 2014. V. 5. № 28. <https://doi.org/10.1186/s13100-014-0028-y>
11. Kim K.W., Tang N.H., Andrusiak M.G. et al. A neuronal piRNA pathway inhibits axon regeneration in *C. elegans* // Neuron. 2018. V. 97(3). P. 511–519. <https://doi.org/10.1016/j.neuron.2018.01.014>
12. Wakisaka K.T., Tanaka R., Hirashima T. et al. Novel roles of drosophila FUS and Aub responsible for piRNA biogenesis in neuronal disorders // Brain Research. 2019. V. 1708. P. 207–219. <https://doi.org/10.1016/j.brainres.2018.12.028>
13. Kim K.W. PIWI proteins and piRNAs in the nervous system // Mol. Cells. 2019. V. 42(12). P. 828–835. <https://doi.org/10.1016/j.neuron.2018.01.014>
14. Perrat P.N., DasGupta Sh., Wang J. et al. Transposition driven genomic heterogeneity in the drosophila brain // Science. 2013. V. 340(6128). P. 9195. <https://doi.org/10.1126/science.1231965>

15. Лавренов А.Р., Неведова Л.Н., Романова Н.И., Ким А.И. Экспрессия генов семейства HP1 и их возможная роль в формировании фенотипа flamenco у *D. melanogaster* // Биохимия. 2014. Т. 79. № 11. С. 1554–1560.
16. Hur J.K., Luo Y., Moon S. et al. Splicing-independent loading of TREX on nascent RNA is required for efficient expression of dual-strand piRNA clusters in *Drosophila* // Genes & Development. 2016. V. 30. P. 840–855. <https://doi.org/10.1101/gad.276030.115>
17. Sayers E.W., Bolton E.E., Brister J.R. et al. Database resources of the national center for biotechnology information // Nucl. Ac. Res. 2022. V. 50. (D1):D20–D26. <https://doi.org/10.1093/nar/gkab1112>
18. Wei X., Eickbush D.G., Speece I., Larracuenta A.M. Heterochromatin-dependent transcription of satellite DNAs in the *Drosophila melanogaster* female germline // eLife. 2021. V. 10. <https://doi.org/10.7554/eLife.62375>
19. Chang T.H., Mattei E., Gainetdinov I. et al. Maelstrom represses canonical polymerase II transcription within bi-directional piRNA clusters in *Drosophila melanogaster* // Mol. Cell. 2019. V. 73. P. 291–303. <https://doi.org/10.1016/j.molcel.2018.10.038>
20. Radion E., Morgunova V., Ryazansky S. et al. Key role of piRNAs in telomeric chromatin maintenance and telomere nuclear positioning in *Drosophila* germline // Epigenetics & Chromatin. 2018. V. 11:40. <https://doi.org/10.1186/s13072-018-0210-4>
21. Gramates L.S., Agapite J., Attrill H. et al. FlyBase: A guided tour of highlighted features // Genetics. 2022. V. 220. № 4. <https://doi.org/10.1093/genetics/iyac035>
22. Cui M., Bai Y., Li K., Rong Y.S. Taming active transposons at *Drosophila* telomeres: The interconnection between HipHop's roles in capping and transcriptional silencing // PLoS Genetics. 2021. V. 17(11). <https://doi.org/10.1371/journal.pgen.1009925>
23. Sato K., Siomi M.C. Two distinct transcriptional controls triggered by nuclear Piwi-piRISCs in the *Drosophila* piRNA pathway // Curr. Op. in Structural Biology. 2018. V. 53. P. 69–76. <https://doi.org/10.1016/j.sbi.2018.06.005>
24. Zhang G., Tu S., Yu T. et al. Co-dependent assembly of *Drosophila* piRNA precursor complexes and piRNA cluster heterochromatin // Cell Reports. 2018. V. 24. P. 3413–3422. <https://doi.org/10.1016/j.celrep.2018.08.081>
25. Akulenko N., Ryazansky S., Morgunova V. et al. Transcriptional and chromatin changes accompanying de novo formation of transgenic piRNA clusters // RNA. 2018. V. 24. № 4. P. 574–584. <https://doi.org/10.1261/rna.062851.117>
26. Klattenhoff C., Xi H., Li C. et al. The *Drosophila* HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters // Cell. 2009. V. 138. № 6. P. 1137–1149. <https://doi.org/10.1016/j.cell.2009.07.014>
27. Parhad S.S., Yu T., Zhang G. et al. Adaptive evolution targets a piRNA precursor transcription network // Cell Reports. 2020. V. 30. № 8. P. 2672–2685. <https://doi.org/10.1016/j.celrep.2020.01.109>
28. Volpe A.M., Horowitz H., Grafer C.M. et al. *Drosophila* rhino encodes a female-specific chromo-domain protein that affects chromosome structure and egg polarity // Genetics. 2001. V. 159. P. 1117–1134. <https://doi.org/10.1093/genetics/159.3.1117>
29. Chung W.-J., Okamura K., Martin R., Lai E.C. Endogenous RNA interference provides a somatic defense against *Drosophila* transposons // Curr. Biology. 2008. V. 18. P. 795–802. <https://doi.org/10.1016/j.cub.2008.05.006>
30. Chen P., Luo Y., Aravin A.A. RDC complex executes a dynamic piRNA program during *Drosophila* spermatogenesis to safeguard male fertility // PLoS Genetics. 2021. V. 17(9). <https://doi.org/10.1371/journal.pgen.1009591>

Regulation of Uni-Strand and Dual-Strand piRNA Clusters in Germ and Somatic Tissues in *Drosophila melanogaster* under Control of *rhino*

P. A. Milyaeva^{a, b}, A. R. Lavrenov^{a, c}, I. V. Kuzmin^a, A. I. Kim^{a, b}, and L. N. Nefedova^{a, *}

^aLomonosov Moscow State University, Moscow, 119234 Russia

^bFaculty of Biology, Shenzhen MSU-BIT University, Longgang District, Shenzhen, 518172 China

^cSevertsov Institute of Ecology and Evolution of Russian Academy of Sciences, Moscow, 119071 Russia

*e-mail: nefedova@mail.bio.msu.ru

Drosophila melanogaster is a common genetic object for research of RNA-interference pathways and mobile elements regulation. Nowadays taking a part in control of retrotransposon expression the system of piRNA-interference well studied in ovary tissues. It is strongly believed that *D. melanogaster* piRNA-interference is used for retrotransposon suppression only in gonads, and two distinct pathways of piRNA biogenesis exist. Both mechanisms use transcripts of piRNA-clusters (accumulations of truncated and defect mobile elements copies): from unstrand clusters in the first case and from dualstrand clusters in the second, transcribed with one or both DNA chains correspondingly. It is well-known that proper dualstrand clusters function depends on the gene *rhino*, while unstrand clusters are transcribed *rhino*-independent and transcripts are spliced. In this paper we show that *rhino* participates in unstrand flamenco transcripts splicing and the piRNA-interference significance for regulation of several retrotransposons not only in gonads, but in other organs.

Keywords: *Drosophila melanogaster*, piRNA, *rhino*, retrotransposons, RNA-interference.

БИОЛОГИЧЕСКОЕ РАЗНООБРАЗИЕ ГЕНОВ
ГОМОЛОГОВ ДЕФЕНЗИНОВ ПШЕНИЦЫ© 2023 г. М. П. Слезина¹, Е. А. Истомина¹, Т. И. Одинцова¹, *¹Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, 119991 Россия

*e-mail: odintsova2005@rambler.ru

Поступила в редакцию 27.06.2023 г.

После доработки 06.07.2023 г.

Принята к публикации 11.07.2023 г.

Антимикробные пептиды (АМП) являются основными компонентами врожденного иммунитета растений и животных. Гены АМП обладают значительным внутри- и межвидовым полиморфизмом, роль которого практически не изучена. Ранее нами методом глубокого секвенирования транскриптомов растений пшеницы были выявлены гены дефензинов, экспрессия которых усиливалась при заражении патогенным грибом *Fusarium oxysporum* и/или обработке индукторами устойчивости. В настоящей работе был проведен биоинформатический поиск в базах данных NCBI пептидов-гомологов этих дефензинов пшеницы по последовательностям их γ-коров — участков молекул, ответственных за антимикробную активность. Гомологи дефензина DEFL1-16 были выявлены у 95 видов покрытосеменных растений, относящихся к 48 семействам и принадлежащих к 30 порядкам одно- и двудольных растений. Повсеместное распространение этого дефензина у покрытосеменных растений позволяет высказать предположение о его участии не только в защите, но и в других процессах в цветковых растениях. Гомологи других дефензинов, индуцированных заражением, обнаружены только у растений семейства Poaceae, что предполагает наличие у мятликовых специфичного для этого семейства механизма защиты от патогенов, связанного с экспрессией этих дефензинов. Среди выявленных в ходе исследования вариантов γ-коров дефензинов дикорастущих растений могут быть пептиды с лучшей по сравнению с пшеницей антимикробной активностью, которые представляют несомненный интерес для разработки новых антибиотиков для медицины и сельского хозяйства.

Ключевые слова: иммунитет растений, антимикробные пептиды, дефензины, биоразнообразие генов.

DOI: 10.31857/S0016675823120111, **EDN:** QCCUZR

Каждый живой организм находится в постоянном контакте с различными патогенными микроорганизмами. Одним из эволюционно древних путей борьбы с ними является синтез антимикробных пептидов (АМП) — коротких полипептидных молекул, обладающих широким спектром антимикробного действия и разнообразным набором мишеней — от мембраны и клеточной стенки до трансляционного аппарата клетки. АМП обнаружены во всех царствах живого мира [1]. Считается, что в системе врожденного иммунитета они формируют “первую линию защиты” от патогенных микроорганизмов и вредителей. АМП обладают значительным внутри- и межвидовым полиморфизмом, роль которого практически не изучена. АМП растений подразделяют на несколько семейств: дефензины, тионины, снакины, липид-переносящие белки, циклотиды, гевеиноподобные и кноттиноподобные пептиды, α-гарпинины [2–5]. Каждое семейство АМП включает, как правило, несколько пептидов, различающихся по аминокислотной последовательности и активности.

Большинство геномов растений и животных содержат 5–10 семейств генов АМП [6]. Таким образом, каждый вид содержит свой уникальный набор АМП, от которого зависит спектр патогенов для этого вида микроорганизмов, а также состав микробиома.

Наблюдаемое в настоящее время развитие устойчивости у патогенов к современным антибиотикам, связанное с их бесконтрольным использованием в медицине и сельском хозяйстве, представляет глобальную угрозу человечеству и требует поиска новых эффективных соединений для борьбы с инфекционными заболеваниями человека, сельскохозяйственных животных и растений. АМП в этом плане привлекают особое внимание исследователей, поскольку они сохранили свой антимикробный потенциал в непрекращающейся в ходе эволюции борьбе с различными видами патогенов [7]. Основным преимуществом АМП по сравнению с современными антибиотика-

ми является значительно более низкая скорость возникновения устойчивости к ним у патогенов и меньшее влияние на экологию, поскольку они менее стабильны в окружающей среде и распадаются до аминокислот.

Для того чтобы понять роль разнообразия АМП в защите растения-хозяина от различных патогенов, а также оценить их терапевтический потенциал для практического использования необходимо определить молекулярные детерминанты активности и понять их механизм действия в природном и эволюционном контекстах. Исследование детерминант антимикробной активности АМП показало, что антимикробная активность преимущественно связана с так называемым γ -кором — участком молекулы, обладающим мотивом $GXCX_nC$ и характерной пространственной структурой, состоящей из двух β -тяжей, образующих шпильку [8].

Для выяснения роли АМП в иммунной системе растений ранее нами методом высокопроизводительного секвенирования транскриптомов был исследован молекулярный ответ растений пшеницы высокоустойчивого вида *Triticum kiharae* Dorof. et Migush. на заражение фитопатогенным грибом *Fusarium oxysporum*, а также на обработку индуктором устойчивости — метаболитами непатогенного штамма гриба *F. sambucinum* [9]. Нами было установлено, что как инфицирование патогеном, так и обработка индуктором устойчивости (элиситором) приводят к активации экспрессии не отдельных генов, а групп генов дефензинов различных подсемейств. Определив антимикробную активность γ -коров этих дефензинов, мы показали, что подавляющее большинство из них *in vitro* подавляет рост патогенного гриба *F. oxysporum* и целого ряда других патогенов растений и человека [10].

В связи с тем что дефензины относятся к эволюционно консервативным АМП, являясь неизменной составной частью врожденного иммунитета как растений, так и животных, представляло интерес выяснить насколько широко распространены в царстве растений гомологи дефензинов пшеницы, экспрессия генов которых индуцировалась заражением и/или индукторами устойчивости. В связи с этим целью настоящей работы был биоинформатический поиск гомологов этих дефензинов пшеницы у растений. Полученные в ходе подобного исследования сведения важны для понимания функционирования и эволюции иммунной системы растений, а также для практического использования пептидов на основе γ -коров дефензинов в качестве антибиотиков в медицине и сельском хозяйстве.

МАТЕРИАЛЫ И МЕТОДЫ

Поиск гомологов дефензинов пшеницы осуществляли по базе данных NCBI (National Center for Biotechnology Information) с помощью программы BLAST [11]. Определение сигнальной последовательности предшественников пептидов проводили с помощью программы SignalP 5.0 [12]. Выравнивание аминокислотных последовательностей и построение филогенетического дерева проводили с помощью программы MEGA7 (Molecular Evolutionary Genetics Analysis) [13].

В работе использовали классификацию растений по системе APG IV (всего в системе APG IV признано 416 семейств, отнесенных к 64 порядкам) [14].

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Для поиска гомологов по последовательности γ -коров были выбраны следующие дефензины пшеницы *T. kiharae*: DEFL1-16, DEFL1-32, DEFL1-11 и DEFL1-12. Экспрессия генов DEFL1-32, DEFL1-11 и DEFL1-12 индуцировалась заражением *F. oxysporum* [9]. Экспрессия гена дефензина DEFL1-16 активировалась при обработке растений пшеницы биогенными индукторами устойчивости, полученными из непатогенного штамма *F. sambucinum* [9]. Последовательности γ -коров выбранных пептидов были синтезированы методом твердофазного синтеза [10]. Они различались по физико-химическим свойствам, влияющим на антимикробную активность: длине, заряду, гидрофильности и амфифильности [10]. По данным молекулярного моделирования все пептиды содержали α -спиральный участок.

Определение антимикробной активности против девяти патогенов (фитопатогенных бактерий и грибов, а также дрожжеподобных патогенов человека и животных) показало, что самой высокой ингибирующей активностью и самым широким спектром антимикробного действия обладал γ -кор DEFL1-16 [10]. Процент ингибирования всех тестированных патогенов варьировал от 50 до 100%. Ингибирование *F. oxysporum* достигало 70%. γ -Коры DEFL1-32 и DEFL1-12 подавляли рост гриба *F. oxysporum* несколько слабее — примерно на 60%. γ -Кор DEFL1-32 не действовал на *F. solani* и *F. verticillioide*s, а DEFL1-12 был неактивным в отношении *F. verticillioide*s. Самой слабой ингибирующей активностью обладал γ -кор DEFL1-11 [10]. Из девяти патогенов только три ингибировались этим пептидом, причем степень ингибирования не превышала 18%. Интересно, что γ -кор DEFL1-11 не подавлял рост четырех видов фузариевых грибов, включая *F. oxysporum* в тестах *in vitro*, хотя заражение этим патогеном сопровождалось усилением экспрессии его гена [9].

Поскольку, как отмечалось выше, γ -кор считается функционально важным элементом, определяющим антимикробную активность дефензинов и других цистеин-богатых пептидов, мы провели поиск гомологов дефензинов DEFL1-16, DEFL1-32 и DEFL1-11 в базах данных NCBI, используя в качестве запроса последовательности их γ -коров.

DEFL1-16

Длина γ -кора DEFL1-16 составляет 18 аминокислотных остатков, он обладает максимальным положительным зарядом (+8) среди выбранных γ -коров (рис. 1) [10]. В результате проведенного поиска гомологи DEFL1-16 были обнаружены нами у 95 видов покрытосеменных растений, относящихся к 48 семействам и принадлежащих к 30 порядкам одно- и двудольных растений (рис. 1). В некоторых семействах было исследовано несколько видов, относящихся к одному или нескольким родам. В ряде случаев у одного вида растений было выявлено несколько гомологов DEFL1-16. У голосеменных растений и низших растений гомологи DEFL1-16 обнаружены не были. Из класса однодольных гомологи DEFL1-16 были найдены у растений восьми семейств: наиболее многочисленного семейства Poaceae (порядок Poales), Bromeliaceae (порядок Poales), Cyperaceae (порядок Poales), а также таких семейств как Arecaceae (порядок Arecales), Dioscoreaceae (порядок Dioscoreales), Asparagaceae (порядок Asparagales), Zosteraceae (порядок Alismatales) и Araceae (порядок Alismatales). У двудольных растений гомологи DEFL1-16 были обнаружены у растений 40 семейств. Жизненные формы растений-носителей гомологов DEFL1-16 самые разнообразные и включают деревья, кустарники и травы. Интересно, что гомологи DEFL1-16 выявлены как у наземных, так и у водных растений. К последним относятся так называемая «морская трава» *Zostera marina* (взморник морской, семейство Zosteraceae) — многолетнее травянистое растение, приспособленное к жизни в соленой воде морей и океанов [15]; *Spirodela intermedia* (многокоренник обыкновенный, или ряска многокоренниковая, семейство Araceae) — одно из наиболее быстро растущих цветковых растений, распространенное в умеренном поясе обоих полушарий в водоемах со стоячей или медленно текущей водой, а также *Nelumbo nucifera* (лотос орехоносный, семейство Nelumbonaceae) — тропический реликтовый вид, который растет в заболоченных местах, озерах и медленно текущих речках [16]. Гомологи DEFL1-16 обнаружены у видов как с широким ареалом распространения, так и приспособленных к жизни в экстремальных условиях высокогорья и повышенной засоленности почвы. Так, виды семейства Poaceae, одного из самых многочисленных и широко распространенных семейств однодольных растений, вклю-

чающего около 11 000 видов, произрастают во всех климатических зонах и на всех континентах. Представители семейства Rosaceae также заселяют самые разнообразные природные зоны — тундру, высокогорье, тропические леса, болота и полупустыни [16]. Лапчатка гусиная *Potentilla anserina* встречается в умеренной зоне по всему земному шару [17]. В то же время представитель этого же семейства — груша уссурийская *Pyrus ussuriensis*, в которой также был обнаружен гомолог DEFL1-16, распространена в Приморье и Приамурье и является самой морозостойкой грушей в мире, поскольку способна выдерживать температуру до -50°C [18]. Представители семейства Malvaceae хотя тоже широко распространены, но преобладают в тропиках и практически отсутствуют в арктической зоне. В качестве примера можно привести представителей рода *Gossypium*, которые растут в тропической и субтропической зонах [16]. В то же время ряд растений-носителей гена DEFL1-16 имеют очень ограниченный ареал распространения. Так, *Cinnamomum micranthum* f. *kanehirae* (камфорное дерево, семейство Lauraceae) является эндемиком Тайваня, произрастающим на высоте до 2000 м [19]. Реликтовым деревом с ограниченным ареалом, растущим в Китае, Вьетнаме и восточных Гималаях на высоте 1100–3500 м над уровнем моря, является и тетрацентрон китайский *Tetracentron sinense* (семейство Trochodendraceae) [20]. Хлебная зерновая культура *Chenopodium quinoa* (киноа, семейство Amaranthaceae) также адаптирована к жизни на высоте 3000–4000 м над уровнем моря в условиях бедных почв и сурового климата и хорошо переносит резкие перепады температур, засуху и заморозки [16]. Приспособлена к жизни в условиях высокогорья и *Kingdonia uniflora* — многолетнее травянистое растение, растущее в горных районах Западного и Северного Китая [21]. К растениям-носителям гена DEFL1-16 относятся и *Arabis nemorensis* (резуха Жерара, семейство Brassicaceae) — растение с уникальной экологией, семена которого способны выдерживать длительные периоды затопления и засухи [16], и *Eutrema salsugineum* (эвтрема солонцовая, семейство Brassicaceae) — галофильный вид, адаптированный к жизни на засоленной почве [22]. В экстремальных условиях произрастает и *Parasponia andersonii*, заселяющая вулканические склоны Малайского архипелага [23]. Чрезвычайно широкое распространение растений-носителей гена DEFL1-16 подтверждается и тем, что он был обнаружен в экологической группе полупаразитических растений семейства Orobanchaceae — у вшивосемянника японского *Phtheirospermum japonicum*, распространенного в Восточной и Северо-Восточной Азии [24]. Обнаружен он и у насекомоядного растения — цефалота мешочкового (*Cephalotus follicularis*, семейство Cephalotaceae), произрастающего в Западной Австралии [25].

Среди растений-носителей генов гомологов DEFL1-16 много хозяйственно ценных и медицинских растений. К первым относятся, прежде всего, хлебные злаки (семейство Poaceae), такие как пшеница, ячмень, кукуруза, овес, просо, сорго и рис, и масличные культуры, такие как подсолнечник, маслина, арахис, кунжут, масличная пальма [16, 26]. Многие виды, в которых были обнаружены гомологи DEFL1-16, относятся к медицинским растениям, среди них такое уникальное растение как женьшень настоящий *Panax ginseng* семейства Araliaceae, встречающийся на Дальнем Востоке [16]. Экстракты многих растений-носителей генов гомологов DEFL1-16 обладают терапевтическим действием. Так, показано, что этанольный экстракт листьев *Cinnamom micranthum* f. *kanehirae* обладает антимикробными свойствами и цитотоксическим действием на раковые клетки [19]. Экстракт *Momordica charantia* обладает цитотоксическим действием на клетки рака поджелудочной железы, а также используется при лечении малярии и ВИЧ [27]. *Handroanthus impetiginosus* используют для лечения целого ряда заболеваний, в том числе воспалительных процессов, грибных инфекций и опухолей [28]. Лапчатка гусиная *P. anserina* эффективна при лечении воспалительных заболеваний кожи, горла и полости рта [17]. Роль гомологов DEFL1-16 в лечебных свойствах растений-носителей этого пептида не изучена.

Подводя итог краткой характеристике видов – носителей гена DEFL1-16 можно утверждать, что среди покрытосеменных растений ген дефензина DEFL1-16 распространен очень широко, если не повсеместно.

Аминокислотные последовательности гомологов предшественника DEFL1-16 приведены на рис. 1. Известно, что в зависимости от структуры предшественника дефензины растений подразделяются на два класса: дефензины класса 1 синте-

зируются в виде предшественников, состоящих из сигнального пептида и зрелого пептида. Дефензины класса 2 содержат еще и С-концевой продомен. Обнаруженные нами дефензины-гомологи DEFL1-16 относятся к дефензинам класса 1. Сравнение аминокислотных последовательностей всех предшественников показывает, что сигнальный пептид, который ответствен за транспорт дефензина в апопласт, является наименее консервативной областью предшественников (рис. 1). Последовательности зрелого пептида достаточно консервативны. Причем в пределах одного семейства (у представителей разных родов) сходство последовательностей гомологов DEFL1-16 выше, чем между гомологами, принадлежащими разным семействам. Однако если сравнить последовательности гомологов всех изученных семейств растений, то помимо консервативных остатков цистеина в N-концевой области молекулы (остатки 1–36) лишь четыре остатка являются неизменными у всех гомологов: это R6, S12, N24 и E32 (рис. 2). Напротив, находящийся в С-концевой области пептида γ -кор оказывается наиболее консервативным участком всей молекулы дефензина. В 10-членном мотиве γ -кора GXCX_nC помимо G в положении 37, G в положении 41 и RRR в положениях 43–45 встречаются у всех гомологов DEFL1-16 (рис. 2). Более того, R в положении 40 встречается почти во всех последовательностях. Лишь в двух в этом положении находится K и в одном – H. Таким образом, в положении 40 всегда расположена основная аминокислота. В положении 42 в подавляющем большинстве случаев находится F, в шести последовательностях – L и в одной – I. Таким образом, в положении 42 всегда расположена гидрофобная аминокислота. Чрезвычайно высокий консерватизм последовательности γ -кора у гомологов DEFL1-16 у огромного числа филогенетически удаленных видов свиде-

Рис. 1. Множественное выравнивание последовательностей гомологов предшественника DEFL1-16 пшеницы, представленных в базе данных NCBI. Остатки цистеина в зрелом пептиде выделены белым цветом на черном фоне. Консервативные аминокислотные остатки выделены серым цветом. Сигнальная последовательность и область зрелого пептида разделены пробелом. Последовательность γ -кора обведена рамкой. Использованы следующие сокращения: C.e., *Coffea eugenoides*; C.a., *Coffea arabica*; C.c., *Coffea canephora*; P.f., *Paulownia fortunei*; S.i., *Sesamum indicum*; Ca.r., *Catharanthus roseus*; N.n., *Nelumbo nucifera*; Te.s., *Telopea speciosissima*; Ma.c., *Macleaya cordata*; N.s., *Nyssa sinensis*; T.s., *Tetracentron sinense*; Q.r., *Quercus robur*; M.t., *Medicago truncatula*; P.m., *Panicum miliaceum*; P.h., *Panicum hallii*; P.v., *Panicum virgatum*; M.l., *Miscanthus lutarioriparius*; D.e., *Digitaria exilis*; S.b., *Sorghum bicolor*; Z.m., *Zea mays*; L.r., *Lolium rigidum*; Ae.t., *Aegilops tauschii* subsp. *strangulata*; T.k., *Triticum kiharae*; T.e., *Thinopyrum elongatum*; T.a., *Triticum aestivum*; H.v., *Hordeum vulgare*; T.d., *Triticum dicoccoides*; Sp.i., *Spirodela intermedia*; A.o., *Asparagus officinalis*; D.l., *Diospyros lotus*; K.u., *Kingdonia uniflora*; C.q., *Chenopodium quinoa*; A.t., *Amaranthus tricolor*; Be.v., *Beta vulgaris*; Q.s., *Quillaja saponaria*; P.a., *Parasponia andersonii*; S.o., *Spinacia oleracea*; O.m., *Oryza meyeriana*; M.r., *Morella rubra*; C.m., *Cinnamomum micranthum* f. *kanehirae*; Ac.n., *Acer negundo*; Zo.m., *Zostera marina*; A.f., *Aristolochia fimbriata*; M.c., *Momordica charantia*; H.a., *Helianthus annuus*; J.c., *Jatropha curcas*; V.f., *Vernicia fordii*; Ri.c., *Ricinus communis*; P.g., *Panax ginseng*; E.g., *Elaeis guineensis*; D.z., *Dioscorea zingerensis*; R.b., *Rhynchospora breviuscula*; R.t., *Rhynchospora tenuis*; S.s., *Solanum stenotomum*; S.l., *Solanum lycopersicum*; S.c., *Solanum commersonii*; P.j., *Phtheirospermum japonicum*; V.d., *Vaccinium darrowii*; B.c., *Brugmansia* $\times$ *candida*; S.ch., *Solanum chacoense*; A.c., *Ananas comosus*; N.a., *Nicotiana glauca*; A.n., *Acer negundo*; A.a., *Artemisia annua*; C.r., *Capsella rubella*; T.h., *Tarenaya hassleriana*; Ph.v., *Phaseolus vulgaris*; Eu.s., *Eutrema salsugineum*; S.p., *Solanum pennellii*; G.h., *Gerbera hybrida*; O.e., *Olea europaea* var. *sylvestris*; Ci.c., *Citrus clementina*; V.v., *Vitis vinifera*; Eu.g., *Eucalyptus grandis*; Co.c., *Corymbia citriodora*; G.b., *Gossypium barbadense*; G.r., *Gossypium raimondii*; G.a., *Gossypium arboreum*; C.f., *Cephalotus follicularis*; Po.a., *Potentilla anserina*; R.c., *Rosa chinensis*; P.u., *Pyrus ussuriensis*; Pr.a., *Prunus avium*; P.p., *Prunus persica*; A.p., *Acacia pycnantha*; R.s., *Raphanus sativus*; H.i., *Handroanthus impetiginosus*; Z.j., *Ziziphus jujuba* var. *spinosa*; R.m., *Rhododendron molle*; S.d., *Salix dunnii*; H.p., *Heliosperma pusillum*; A.i., *Arachis ipaensis*; A.d., *Arachis duranensis*; B.v., *Bauhinia variegata*; Ar.a., *Artemisia annua*. * Последовательность выведена на основе геномных данных [10].

[illegible]

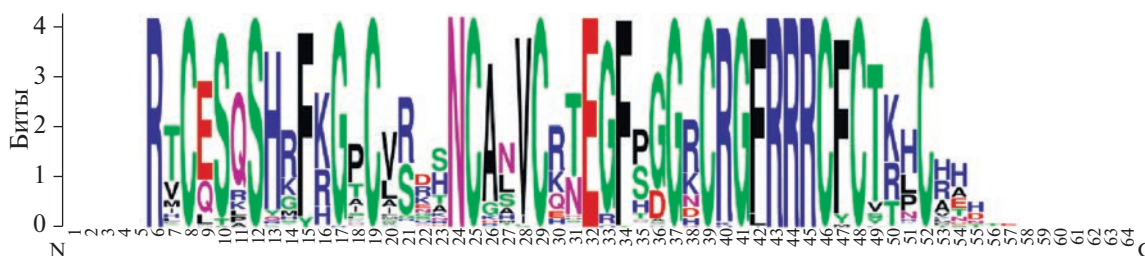


Рис. 2. Логотип последовательностей дефензинов-гомологов DEFL1-16. Высота каждого аминокислотного остатка отражает степень консервативности. Числа по шкале абсцисс определяют положение в молекуле дефензина. По оси ординат представлено информационное содержание.

тельствует о его решающей роли для функций этого дефензина.

На рис. 3 приведено филогенетическое дерево, построенное по аминокислотным последовательностям зрелых гомологов DEFL1-16. Видно, что гомологи DEFL1-16 образуют два основных кластера, один из которых состоит из двух крупных подкластеров. Основные группы гомологов DEFL1-16 обозначены нами цифрами I, II, III (рис. 3). В группу I вошли гомологи DEFL1-16 двудольных растений и один представитель однодольных семейства Bromeliaceae *Ananas comosus*. Эта группа гомологов включает все виды семейств Rosaceae, Malvaceae, Myrtaceae и Solanaceae (образующие несколько подгрупп). Виды семейства Brassicaceae также образуют одну основную группу в этом кластере. Кстати, в нее же попал и представитель семейства Cleomaceae, который так же как и Brassicaceae относится к порядку Brassicales. Сюда же входят и большинство видов семейств Asteraceae и Fabaceae, оба вида семейства Ericaceae и два из трех видов семейства Pedaliaceae. В группу II вошли гомологи DEFL1-16 однодольных растений семейств Poaceae, Arecaceae, Asparagaceae, Zosteraceae, Cyperaceae, Dioscoreaceae, Arecaceae. В этот же кластер вошли гомологи DEFL1-16 двудольных растений семейств Amaranthaceae и Euphorbiaceae, а также представителей семейств Lauraceae и Aristolochiaceae, относящихся к порядкам Laurales и Piperales подкласса магнолииды (Magnoliidae), которые считаются наиболее архаичными из ныне живущих растений. Из изученных семейств семейство Poaceae представлено наибольшим числом видов (14). Наблюдаемое сходство последовательностей гомологов DEFL1-16 отражает филогенетическое родство видов согласно современным молекулярным представлениям о филогении видов семейства Poaceae [29]. Так, последовательности гомологов DEFL1-16 видов, относящихся к подсемействам Pooideae и Panicoideae, заметно различаются между собой. В пределах подсемейства Pooideae наиболее сходные последовательности у филогенетически наиболее близких представителей родов *Triticum*, *Aegilops*, *Thinopyrum* и *Hordeum*, относящихся к супертрибе Triticodae [29]. В пределах подсемейства Panicoideae очевидно сходство последовательностей у фи-

логенетически родственных видов родов *Panicum* и *Digitaria*, а также *Sorghum* и *Miscanthus*. Наиболее отличающаяся последовательность гомолога DEFL1-16 у риса *Oryza meyeriana*, относящегося к подсемейству Oryzoideae. В пределах одного рода различия аминокислотных последовательностей минимальные, что следует, например, из сравнения последовательностей гомологов DEFL1-16 у трех видов рода *Panicum*.

В группу III попали гомологи видов семейства Rubiaceae (рис. 3). Сюда же вошли гомологи филогенетически родственных по классификации APG IV семейств Papaveraceae, Proteaceae, Nelumbonaceae и Trochodendraceae.

Таким образом, большинство семейств, в которых было изучено несколько видов, образуют на филогенетическом дереве отдельные подкластеры. К таким семействам относятся Poaceae, Rosaceae, Myrtaceae, Malvaceae, Rubiaceae и Euphorbiaceae. В то же время в некоторых семействах, таких как Solanaceae, Fabaceae и Brassicaceae, гомологи DEFL1-16 образуют несколько групп родственных пептидов, что может свидетельствовать о дивергенции функций между пептидами разных групп.

Механизм действия дефензинов-гомологов DEFL1-16 слабо изучен. Имеющиеся данные указывают на то, что защитное действие γ -кора этого дефензина связано с увеличением проницаемости мембраны микроорганизмов [10, 30]. Само ли нарушение проницаемости мембран приводит к гибели патогена, или, разрушив мембраны, пептид проникает внутрь клеток, где он взаимодействует с внутриклеточными мишенями и блокирует их действие, еще предстоит установить.

Полученные нами ранее транскриптомные данные свидетельствуют о том, что роль гомологов DEFL1-16 в защите от патогенов у разных видов растений и у разных паралога в пределах одного вида различается. Так, нами транскриптомным анализом у пшеницы *T. kiharae* выявлено два паралога DEFL1-15 и DEFL1-16 [9]. Ни тот, ни другой ген не активировался при заражении *F. oxysporum*. Однако экспрессия DEFL1-16 усиливалась в ответ на обработку индукторами устойчивости. Таким образом, в природе DEFL1-16 не

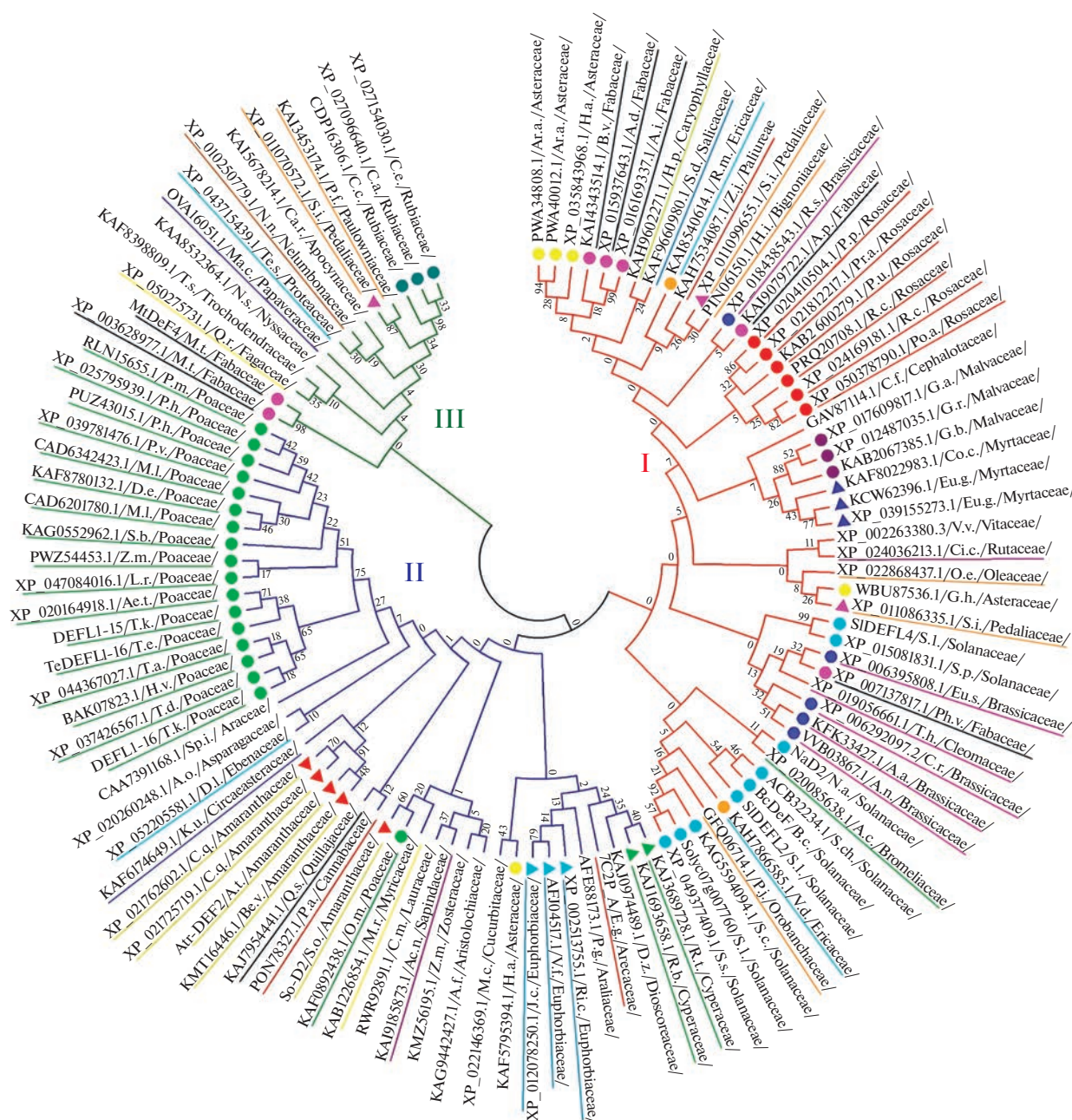


Рис. 3. Филогенетическое дерево дефензинов-гомологов DEFL1-16, построенное методом присоединения соседей (Neighbor-joining) с помощью программы MEGA7 [13]. Статистический анализ дерева был проведен методом бутстрэп (Bootstrap) с 10000 повторами. Дефензины представителей семейства Rosaceae отмечены салатовыми кружками, семейства Fabaceae – розовыми кружками, семейства Rubiaceae – темнорозовыми кружками, семейства Ericaceae – оранжевыми кружками, семейства Brassicaceae – синими кружками, семейства Asteraceae – желтыми кружками, семейства Malvaceae – фиолетовыми кружками, семейства Pedaliaceae – розовыми треугольниками, семейства Amaranthaceae – красными треугольниками, семейства Euphorbiaceae – бирюзовыми треугольниками, семейства Myrtaceae – синими треугольниками. Нижним подчеркиванием одинакового цвета отмечены семейства одного порядка: зеленым – Poales, фиолетовым – Sapindales, синим – Ranunculales, оранжевым – Lamiales, черным – Fabales, розовым – Brassicales, красным – Rosales, бирюзовым – Ericales, голубым – Malpighiales, коричневым – Proteales, желтым – Fagales, салатовым – Caryophyllales. Группы гомологов обозначены разным цветом ветвей: группа I – красным, группа II – синим, группа III – зеленым. Сокращения см. подпись к рис. 1.

	1		87	
DEFL1-11/T.k./	-----MLVLLLLSSYDGGG-----	VGVAEA	RICTGKSQHHSFFCISDKSCKTKCLGEHGAKWTA	GYCKISRCCTCQREC
XP_020155345.1/Ae.t./	MEFKPKATVCAVMLVLLLLSSYDGGG-----	VGVAEA	RICTGKSQHHSFACISDKSCKTKCLGEHGAKWTA	GYCKISRCCTCQREC
*TeDEFL1-32/T.e. /	MEFKPKVTVCVAVMLVLLLLSSDDGGG-----	VGVAEA	RICTGKSQHHSFPCISDKSCKTKCLGEHGAKWTA	GYCKIRRCCTCQREC
DEFL1-32/T.k./	-----MLVLLLLSSYDGGG-----	VGVAEA	RICTGKSQHHSFFCISDKSCKTKCLGEHGAKWTA	GYCKIRRCCTCQREC
*TeDEFL1-11/T.e./	MEFKPKATVCAVMLVLLLLSSYSDSGG-----	IGVAEA	RICTGKSQHHSFFCISDKSCKTKCLGEHGAKWTA	GYCKIRRCCTCQREC
AU282778.1/A.s./	MEFKPKATVCAVMLVLLLLSSYSDSGG-----	IGVAEA	RICTGKSQHHSFPCISDKSCKTKCLGEHGAKWTA	GYCKIRRCCTCQREC
VAI35182.1/T.t./	MEFKPKATVCAVMLVLLLLSSYSGGGG-----	IGVAEA	RICTGKSQHHSFFCISDKSCKTKCLGEHGAKWTA	GYCKIRRCCTCQREC
XP_037431333.1/T.d./	MEFKPKATVCAVMLVLLLLSSYSGGGG-----	IGVAEA	RICTGKSQHHSFFCISDKSCKTKCLGEHGAKWTA	GYCKIRRCCTCQREC
XP_044389907.1/T.a./	MEFKPKATVCAVMLVLLLLSSYSGGGG-----	IGVAEA	RICTGKSQHHSFFCISDKSCKTKCLGEHGAKWTA	GYCKIRRCCTCQREC
AST11412.1/A.s./	MEFKKATVCAVALLMLLLLLSSYDS-G-----	SMVEA	RVCTGKSQHHSFFCISDRCLCSHTCTKEGGGWTA	GYCHLRFCTCQKQC
XP_047073612.1/L.r./	MEFKPKATVCAVALLMLLLLLSSYDGGG-----	TAEA	RVCTGKSQHHSFFCISDRCKICAGCTKEEGSHWTA	GYCHRRVCTCQKRC
KAG8039429.1/Z.p./	MEATSKATACAVLLMLLLISHNGGGG-----	GVAEA	RVCTGKSQHHSFFCISDRCSQCAKQRG-GWTA	GYCHFRVCTCQKSC
TVU43273.1/E.c./	MELKATVCAVLLMLVLSHDDNGGG-----	VSVAEA	RVCTGKSQHHSFFCISDRCLCSNQCVKQRG-GWTA	GYCHLRFCTCQKAC
XP_025819879.1/P.h./	MELKSRAAACALLLLLLVLTHGGSSDVGG-----	MVAEA	RVCTGKSQHHSFFCISDRCLCTRNQVNEGGG-WTA	GYCHRRVCTCQKAC
XP_034601423.1/S.v./	MELKSRAATVALLLLLLVLTHDGGYGGGGGGGRLAVEA	RVCTGKSEHHSFPCISDRHCSDDCIKQKRG-WTA	GYCHRRVCTCQKAC	
KA8804473.1/H.v./	MEFKTKATVCAALLVLLLLSSYDGGDGG-----	VGVAEA	RICTGKSQHHSFFCISDKSCARNQVSEHGAWTA	GYCHLRFCTCQREC
XP_044950506.1/H.v./	MEFKTKATVCAALLVLLLLSSYDGGDGG-----	VGVAEA	RICTGKSQHHSFFCISDKSCARNQVSEHGAWTA	GYCHLRFCTCQREC
XP_037438255.1/T.d./	MEFKPKATVCAVMLVLLLLSSYDGGG-----	VGVAEA	RICTGKSQHHSFPCISDKSCKTKCLGEHGAKWTA	GYCKIRRCCTCQREC

Рис. 4. Множественное выравнивание последовательностей гомологов предшественников DEFL1-11 и 1-32 пшеницы, представленных в базе данных NCBI. Остатки цистеина в зрелом пептиде выделены белым цветом на черном фоне. Консервативные аминокислотные остатки выделены серым цветом. Сигнальная последовательность и область зрелого пептида разделены пробелом. Последовательность γ -кора обведена рамкой. Использованы следующие сокращения: T.k., *T. kiharae*; Ae.t., *Ae. tauschii* subsp. *strangulata*; T.e., *Th. elongatum*; A.s., *Avena sativa*; T.t., *Triticum turgidum* subsp. *durum*; T.d., *T. dicoccoides*; T.a., *T. aestivum*; L.r., *L. rigidum*; Z.p., *Zizania palustris*; E.c., *Eragrostis curvula*; P.h., *P. hallii*; S.v., *Setaria viridis*; H.v., *H. vulgare*. * Последовательность выведена на основе геномных данных [10].

участвует в иммунном ответе растений пшеницы на заражение *F. oxysporum*. Для этого используются другие дефензины или АМП, или иные защитные соединения. Нами установлено, что заражение пшеницы приводило к усилению экспрессии четырех генов дефензинов DEFL1-11, DEFL1-12, DEFL1-32 и DEFL1-43 [9]. Однако в устойчивости, индуцированной непатогенными штаммами микроорганизмов или другими элиситорами, дефензин DEFL1-16, несомненно, участвует. Роль паралога дефензина DEFL1-15 остается неясной. Возможно, он участвует в защите от других патогенов или выполняет какие-то иные (не защитные) функции. Таким образом, уже в пределах одного вида происходит разделение функций у паралогов DEFL1-16. Сходная ситуация наблюдается у томатов. В транскриптомах томатов нами выявлено три гена гомологов DEFL1-16 пшеницы: SIDEFL1, -2 и -4 [31]. Экспрессия генов дефензинов SIDEFL1 и SIDEFL2, как и DEFL1-16 пшеницы, усиливалась лишь при обработке индукторами устойчивости, и только экспрессия гена SIDEFL4 многократно усиливалась при заражении грибом *F. oxysporum*. Эти данные свидетельствуют о том, что лишь SIDEFL4 участвует в защите от гриба *F. oxysporum*. Как и в случае DEFL1-15, возможно, что другие паралоги связаны с защитой от других патогенов (или стрессовых факторов).

Чрезвычайно широкое распространение гомологов DEFL1-16 среди покрытосеменных растений наводит на мысль о том, что этот дефензин может выполнять и другие, не связанные с защитой, жизненно важные функции. Наличие у дефензинов более одной функции (полифункциональность) была показана для некоторых из них [32–34]. Полученные нами данные по дифференциальной экспрессии генов дефензинов у пшеницы и томата также свидетельствуют в пользу этого

предположения. Поскольку гомологи DEFL1-16 были обнаружены нами только у покрытосеменных растений, возможно, что этот пептид участвует в процессах развития цветка или двойного оплодотворения.

DEFL1-11 и DEFL1-32

Длина γ -коров DEFL1-11 и DEFL1-32 составляет 14 остатков, они различаются лишь двумя аминокислотными остатками: IS у γ -кора DEFL1-11 заменен на FR у DEFL1-32 (рис. 4). Петля между первым и вторым цистеинами у γ -коров DEFL1-11 и DEFL1-32 короче, чем у DEFL1-16, на два остатка. Заряд γ -коров DEFL1-32 и DEFL1-11 также отличался +3 и +2 соответственно [10]. Поиск гомологов близкородственных дефензинов DEFL1-11 и DEFL1-32 показал, что в отличие от DEFL1-16 они выявляются только у растений семейства Poaceae (рис. 4). Более того, гомологи с точно таким же γ -кором, как у DEFL1-11 и DEFL1-32, обнаружены только у представителей рода *Triticum* и у *Ae. tauschii*, который считается донором генома D полиплоидным пшеницам. Как отмечалось выше, неактивный DEFL1-11 отличается от высокоактивного DEFL1-32 двумя аминокислотными остатками в γ -коре: IS у DEFL1-11 заменен на FR у DEFL1-32. Таким образом, эта замена приводит к потере активности. У представителей рода *Triticum* (*T. turgidum*, *T. dicoccoides*, *T. aestivum*), а также у пырея и овса S заменен на R. Такая замена увеличивает заряд γ -кора, поэтому можно предположить, что она приводит к усилению антимикробных свойств, т.е. пептиды с IR, вероятно, обладают антимикробными свойствами, в отличие от пептидов с IS, таких как DEFL1-11. Однако это предположение нуждается в экспериментальной проверке. Вместо K4 встречаются другие основ-

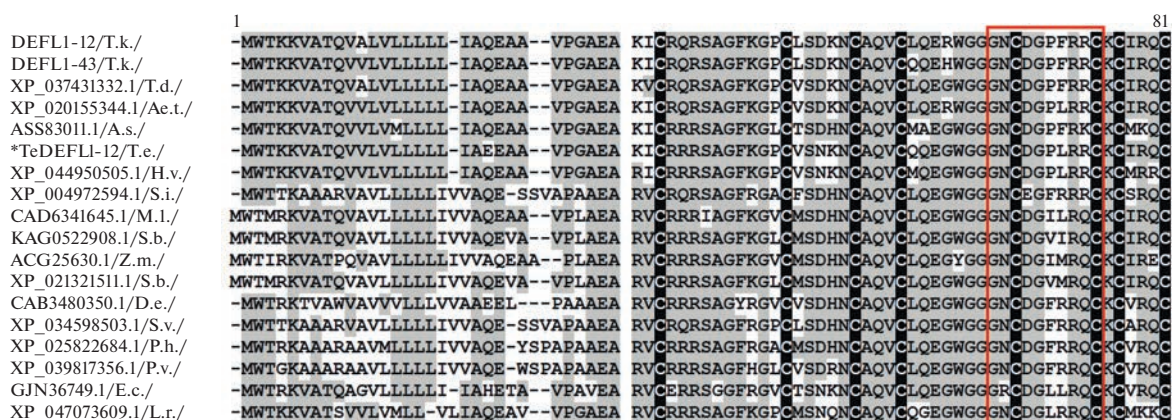


Рис. 5. Множественное выравнивание последовательностей гомологов предшественников DEFL1-12 и 1-43 пшеницы, представленных в базе данных NCBI. Остатки цистеина в зрелом пептиде выделены белым цветом на черном фоне. Консервативные аминокислотные остатки выделены серым цветом. Сигнальная последовательность и область зрелого пептида разделены пробелом. Последовательность γ -кора обведена рамкой. Использованы следующие сокращения: T.k., *T. kiharae*; T.d., *T. dicoccoides*; Ae.t., *Ae. tauschii* subsp. *strangulata*; A.s., *A. sativa*; T.e., *Th. elongatum*; H.v., *H. vulgare*; S.i., *S. italica*; M.l., *Miscanthus lutarioriparius*; S.b., *S. bicolor*; Z.m., *Z. mays*; D.e., *Digitaria exilis*; S.v., *S. viridis*; P.h., *P. hallii*; P.v., *P. virgatum*; E.c., *Eleusine coracana* subsp. *coracana*; L.r., *L. rigidum*. * Последовательность выведена на основе геномных данных [10].

ные аминокислоты — Н или R. Непонятно, влияет ли это на активность пептида. Вместо F5 встречается I, L или W. Поскольку все эти аминокислоты относятся к гидрофобным, возможно, что пептиды с I, L или W обладают такой же активностью, как DEFL1-32. Следует, однако, отметить, что триптофан является мембраноактивной аминокислотой, поэтому не исключена вероятность того, что γ -кор с триптофаном, как у *P. hallii*, обладает более высокой активностью, чем DEFL1-32. Интересно, что в отличие от DEFL1-16, у которого γ -кор является более консервативной частью молекулы, чем другие участки, у гомологов DEFL1-11 и DEFL1-32 одновременно с заменами в γ -коре наблюдаются и замены в других частях молекул дефензинов.

Экспрессия обоих генов DEFL1-11 и DEFL1-32 усиливалась при инфицировании *F. oxysporum* и в зараженных растениях после обработки элиситорами. Поскольку DEFL1-32 подавлял рост *F. oxysporum* в тестах *in vitro*, усиление экспрессии его гена свидетельствует в пользу его участия в подавлении роста этого патогена в растениях пшеницы (*in planta*). Роль DEFL1-11 остается неясной, поскольку он не подавлял рост *F. oxysporum in vitro*, хотя экспрессия его гена и усиливалась при заражении грибом. Можно предположить, что DEFL1-11 действует в синергизме с другими дефензинами (АМП или другими защитными белками), усиливая их антифунгальное действие. Нельзя исключить существование и еще каких-то (не антимикробных) функций у DEFL1-11. И наконец, экспрессия “коктейлей” АМП в ответ на заражение патогеном, в которых есть пептиды, неактивные в отношении именно этого патогена, может иметь смысл в природном контексте. Инфицирование растений в природных условиях часто про-

исходит не одним, а сразу несколькими патогенами, в связи с чем синтез сразу целого набора АМП с различными спектрами действия приобретает смысл. Более того возможно и другое объяснение. Растение может не совсем точно идентифицировать “врага”, а синтез антимикробных “коктейлей” позволит все же защититься.

DEFL1-12/DEFL1-43

Мы провели также поиск гомологов других дефензинов, индуцированных заражением *F. oxysporum*. Это дефензины DEFL1-12 и DEFL1-43. Их γ -кору длиной 16 аминокислотных остатков и зарядом +3 одинаковы, а в зрелом пептиде есть различие лишь по одному аминокислотному остатку (рис. 5). Как и в случае с DEFL1-11 и DEFL1-32, гомологи этих дефензинов выявляются только у злаков (рис. 5). Последовательность γ -коров этих дефензинов менее консервативна, чем у DEFL1-16. В последовательности пепти между двумя остатками цистеина 10-членного мотива γ -кора GXCX_nC у всех выявленных гомологов одинаковы лишь две аминокислоты: G в положении 5 и R в положении 8. В положении 4 в большинстве последовательностей находится D. Наблюдаемые замены в последовательности γ -кора у гомологов приводят либо к увеличению положительного заряда пептида, либо к усилению гидрофобных свойств. Каким образом эти изменения влияют на антимикробную активность пептида, еще предстоит установить.

Подводя итог полученным результатам, можно сказать, что из изученных дефензинов пшеницы только гомологи DEFL1-16 были обнаружены у очень большого числа видов покрытосеменных

растений, что позволяет предположить наличие у этого дефензина уникальных, ключевых для растений функций, возможно связанных не только с защитой от патогенов. Высокий консерватизм последовательности γ -кора этого дефензина свидетельствует о том, что он подвержен строгим функциональным ограничениям, важным для функционирования содержащих эту последовательность дефензинов, и находится под действием сильного стабилизирующего отбора. Повсеместное распространение гомологов DEFL1-16 у покрытосеменных растений делает этот дефензин особо привлекательным для дальнейших исследований его биологических функций и роли выявленных мутаций в γ -коре в биологической активности, что послужит основой для разработки на его основе пептидных антибиотиков нового поколения. Гомологи других дефензинов пшеницы, индуцированных заражением патогенным грибом, обнаружены только у растений семейства Poaceae. Это дает основание предполагать наличие у них общего и характерного только для этого семейства механизма защиты от патогенов, связанного с экспрессией этих генов дефензинов. Однако это предположение нуждается в экспериментальной проверке и предполагает транскриптомный анализ различных представителей этого семейства в ответ на заражение. Последовательности γ -коров дефензинов, индуцированных заражением, менее консервативны в эволюции, чем γ -кор DEFL1-16. Можно предположить, что эти последовательности находятся под действием движущего отбора, который усиливает антимикробные свойства пептида или изменяет его функцию. В связи с этим особый интерес представляют исследования активности γ -коров дефензинов не культурных, а устойчивых к патогенам дикорастущих видов, таких как пырей. Работы в этом направлении нами уже ведутся. Как показали наши исследования, мутации в γ -коре могут приводить и к потере антимикробных свойств, как в случае DEFL1-11, что может свидетельствовать о происходящей неофункционализации, которая может быть оправдана в определенных условиях в природе при снижении инфекционной нагрузки.

Работа выполнена при финансовой поддержке гранта РНФ № 22-16-00010.

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

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Zaslöff M. Antimicrobial peptides of multicellular organisms // *Nature*. 2002. № 415. P. 389–395. <https://doi.org/10.1038/415389a>
2. Tam J.P., Wang S., Wong K.H., Tan W.L. Antimicrobial peptides from plants // *Pharmaceuticals*. 2015. V. 8.

- № 4. P. 711–757. <https://doi.org/10.3390/ph8040711>
3. Li J., Hu S., Jian W. et al. Plant antimicrobial peptides: structures, functions, and applications // *Bot. Stud.* 2021. V. 62. № 1. <https://doi.org/10.1186/s40529-021-00312-x>
4. Lima A.M., Azevedo M.I.G., Sousa L.M. et al. Plant antimicrobial peptides: An overview about classification, toxicity and clinical applications // *Int. J. Biol. Macromol.* 2022. V. 214. P. 10–21. <https://doi.org/10.1016/j.ijbiomac.2022.06.043>
5. Zou F., Tan C., Shinali T.S. et al. Plant antimicrobial peptides: A comprehensive review of their classification, production, mode of action, functions, applications, and challenges // *Food Funct.* 2023. V. 14. № 12. P. 5492–5515. <https://doi.org/10.1039/d3fo01119d>
6. Lazzaro B.P., Zasloff M., Rolff J. Antimicrobial peptides: Application informed by evolution // *Science*. 2020. V. 368. № 6490. <https://doi.org/10.1126/science.aau5480>
7. Zhu Y., Hao W., Wang X. et al. Antimicrobial peptides, conventional antibiotics, and their synergistic utility for the treatment of drug-resistant infections // *Med. Res. Rev.* 2022. V. 42. № 4. P. 1377–1422. <https://doi.org/10.1002/med.21879>
8. Yount N.Y., Yeaman M.R. Multidimensional signatures in antimicrobial peptides // *Proc. Natl Acad. Sci. USA*. 2004. V. 101. № 19. P. 7363–7368. <https://doi.org/10.1073/pnas.0401567101>
9. Odintsova T.I., Slezina M.P., Istomina E.A. et al. Defensin-like peptides in wheat analyzed by whole-transcriptome sequencing: A focus on structural diversity and role in induced resistance // *PeerJ*. 2019. V. 7. <https://doi.org/10.7717/peerj.6125>
10. Slezina M.P., Istomina E.A., Kulakovskaya E.V. et al. The γ -core motif peptides of AMPs from grasses display inhibitory activity against human and plant pathogens // *Int. J. Mol. Sci.* 2022. V. 23. № 15. <https://doi.org/10.3390/ijms23158383>
11. <https://blast.ncbi.nlm.nih.gov/Blast.cgi>
12. Bendtsen J.D., Nielsen H., von Heijne G., Brunak S. Improved prediction of signal peptides: SignalP 3.0 // *J. Mol. Biol.* 2004. V. 340. № 4. P. 783–795. <https://doi.org/10.1016/j.jmb.2004.05.028>
13. Kumar S., Stecher G., Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets // *Mol. Biol. Evol.* 2016. V. 33. № 7. P. 1870–1874. <https://doi.org/10.1093/molbev/msw054>
14. The Angiosperm Phylogeny Group, Chase M.W., Christenhusz M.J.M. et al. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV // *Bot. J. Linn. Soc.* 2016. V. 181. № 1. P. 1–20. <https://doi.org/10.1111/boj.12385>
15. Вехов В.Н. Зостера морская Белого моря. М.: МГУ, 1992. 144 с.
16. Жизнь растений. Т. 5. Ч. 2. Цветковые растения / Под ред. Тахтаджяна А.Л. М.: Просвещение, 1981. 511 с.
17. Попов А.П. Лекарственные растения в народной медицине. Киев: Здоров'я, 1967. 316 с.
18. Усенко Н.В. Деревья, кустарники и лианы Дальнего Востока. Хабаровск: Хабаровское кн. изд-во, 1984. С. 110–111.
19. Li P.-H., Shih Y.-J., Lu W.-C. et al. Antioxidant, antibacterial, anti-inflammatory, and anticancer properties

- of *Cinnamomum kanehirae* Hayata leaves extracts // Arab. J. Chem. 2023. V. 16. № 7. 104873. <https://doi.org/10.1016/j.arabjc.2023.104873>
20. Endress P.K. Trochodendraceae // Flowering Plants Dicotyledons. The Families and Genera of Vascular Plants. V. 2. Berlin; Heidelberg: Springer, 1993. P. 599–602. https://doi.org/10.1007/978-3-662-02899-5_74
 21. Sun Y., Deng T., Zhang A. et al. Genome sequencing of the endangered *Kingdonia uniflora* (Circaceasteraceae, Ranunculales) reveals potential mechanisms of evolutionary specialization // iScience. 2020. V. 23. № 5. <https://doi.org/10.1016/j.isci.2020.101124>
 22. Li C., Duan C., Zhang H. et al. Adaptive mechanisms of halophytic *Eutrema salsugineum* encountering saline environment // Front. Plant Sci. 2022. V. 13. <https://doi.org/10.3389/fpls.2022.909527>
 23. Dupin S.E., Geurts R., Kiers E.T. The non-legume *Parasponia andersonii* mediates the fitness of nitrogen-fixing rhizobial symbionts under high nitrogen conditions // Front. Plant Sci. 2020. V. 10. <https://doi.org/10.3389/fpls.2019.01779>
 24. Clarke C.R., Timko M.P., Yoder J.I. et al. Molecular dialog between parasitic plants and their hosts // Annu. Rev. Phytopathol. 2019. V. 57. P. 279–299. <https://doi.org/10.1146/annurev-phyto-082718-100043>
 25. Conran J.G. Cephalotaceae // Flowering Plants Dicotyledons. The Families and Genera of Vascular Plants. Berlin; Heidelberg: Springer, 2004. V. 6. P. 65–68. https://doi.org/10.1007/978-3-662-07257-8_7
 26. Жизнь растений. Т. 6. Цветковые растения / Под ред. Тахтаджяна А.Л. М.: Просвещение, 1982. 543 с.
 27. Li C.J., Tsang S.F., Tsai C.H. et al. *Momordica charantia* extract induces apoptosis in human cancer cells through caspase- and mitochondria-dependent pathways // Evid. Based Complement. Alternat. Med. 2012. V. 2012. <https://doi.org/10.1155/2012/261971>
 28. Zhang J., Hunto S.T., Yang Y. et al. *Tabebuia impetiginosa*: A comprehensive review on traditional uses, phytochemistry, and immunopharmacological properties // Molecules. 2020. V. 25. № 18. <https://doi.org/10.3390/molecules25184294>
 29. Huang W., Zhang L., Columbus J.T. et al. A well-supported nuclear phylogeny of Poaceae and implications for the evolution of C4 photosynthesis // Mol. Plant. 2022. V. 15. № 4. P. 755–777. <https://doi.org/10.1016/j.molp.2022.01.015>
 30. Slezina M.P., Istomina E.A., Kulakovskaya E.V. et al. Synthetic oligopeptides mimicking γ -core regions of cysteine-rich peptides of *Solanum lycopersicum* possess antimicrobial activity against human and plant pathogens // Curr. Issues Mol. Biol. 2021. V. 43. № 3. P. 1226–1242. <https://doi.org/10.3390/cimb43030087>
 31. Slezina M.P., Istomina E.A., Korostyleva T.V. et al. Molecular insights into the role of cysteine-rich peptides in induced resistance to *Fusarium oxysporum* infection in tomato based on transcriptome profiling // Int. J. Mol. Sci. 2021. V. 22. № 11. <https://doi.org/10.3390/ijms22115741>
 32. Stotz H.U., Spence B., Wang Y. A defensin from tomato with dual function in defense and development // Plant Mol. Biol. 2009. V. 71. № 1–2. P. 131–143. <https://doi.org/10.1007/s11103-009-9512-z>
 33. Allen A., Snyder A.K., Preuss M. et al. Plant defensins and virally encoded fungal toxin KP4 inhibit plant root growth // Planta. 2008. V. 227. № 2. P. 331–339. <https://doi.org/10.1007/s00425-007-0620-1>
 34. Mith O., Benhamdi A., Castillo T. et al. The antifungal plant defensin AhPDF1.1b is a beneficial factor involved in adaptive response to zinc overload when it is expressed in yeast cells // Microbiologyopen. 2015. V. 4. № 3. P. 409–422. <https://doi.org/10.1002/mbo3.248>

Biological Diversity of Genes Encoding Wheat Defensin Homologues

M. P. Slezina^a, E. A. Istomina^a, and T. I. Odintsova^a, *

^aVavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, 119991 Russia

*e-mail: odintsova2005@rambler.ru

Antimicrobial peptides (AMPs) are major components of innate immunity in plants and animals. AMP genes have significant intra- and interspecific polymorphism, the role of which is poorly understood. Previously, by high throughput transcriptome sequencing of wheat plants, we identified defensin genes up-regulated upon infection with the pathogenic fungus *Fusarium oxysporum* and/or treatment with resistance inducers. In the present work, a bioinformatic search in NCBI databases for peptide homologues of these defensins was carried out using the sequences of their γ -cores, the sites of the molecules responsible for antimicrobial activity. DEFL1-16 homologues were identified in 95 species of angiosperms belonging to 48 families and 30 orders of monocotyledonous and dicotyledonous plants. The ubiquitous distribution of this defensin in angiosperms suggests its involvement not only in defense, but also in other processes in flowering plants. Homologues of other defensins induced by infection were found only in plants of the Poaceae family, which suggests the existence of a Poaceae-specific defense mechanism associated with the expression of these defensins. Among the γ -core variants of wild plant defensins identified in the study, the peptides with better antimicrobial activity compared to wheat, might be present, which are of considerable interest for the development of new antibiotics for medicine and agriculture.

Keywords: plant immunity, antimicrobial peptides, defensins, gene biodiversity.

ПОПУЛЯЦИОННО-ГЕНЕТИЧЕСКАЯ СТРУКТУРА БЕЛОГО МЕДВЕДЯ (*Ursus maritimus*) В МОРЯХ РОССИЙСКОЙ АРКТИКИ

© 2023 г. П. А. Сорокин¹*, Е. Ю. Звычайная¹, Е. А. Иванов¹, И. А. Мизин², И. Н. Мордвинцев¹,
Н. Г. Платонов¹, А. И. Исаченко³, Р. Е. Лазарева³, В. В. Рожнов¹

¹Институт проблем экологии и эволюции им. А.Н. Северцова Российской академии наук, Москва, 119071 Россия

²Национальный парк “Русская Арктика”, Архангельск, 163051 Россия

³ООО “Арктический Научный Центр”, Москва, 119333 Россия

*e-mail: sorokin-p@yandex.ru

Поступила в редакцию 30.05.2023 г.

После доработки 16.07.2023 г.

Принята к публикации 18.07.2023 г.

Рассмотрена популяционно-генетическая структура белого медведя (*Ursus maritimus*) на модельных участках в морях российской Арктики по материалам, собранным в период 2010–2021 гг. Получены данные по полиморфизму 17 микросателлитных локусов ядерной ДНК и фрагмента D-петли мтДНК длиной 610 пн для 93 животных. Для исследованной выборки взрослых белых медведей обнаружено высокое генетическое разнообразие ядерной ДНК и низкое значение нуклеотидной изменчивости π по митохондриальной ДНК. По всем генетическим маркерам обнаружена дифференциация медведей из южной части Баренцева моря от животных севера Баренцева и Карского морей. Данные группировки различаются по распределению митохондриального маркера ($\theta_{st} = 0.270$) и слабо дифференцируются по ядерным локусам ($R_{st} = 0.018$).

Ключевые слова: белый медведь, *Ursus maritimus*, контрольный регион, микросателлиты, полиморфизм, Баренцево море, Карское море.

DOI: 10.31857/S0016675823120123, **EDN:** VDKHCW

Белый медведь (*Ursus maritimus* Phipps, 1774) — эволюционно относительно молодой вид, имеющий циркумполярный ареал и адаптировавшийся к жизни и охоте на льдах [1, 2]. Международные природоохранные организации, исходя из соображений удобства природоохранного менеджмента, рассматривают всех белых медведей как единую популяцию, разделенную на 19 субпопуляций [3].

Многочисленные исследования внутривидовой структуры белого медведя с использованием различных генетических методов (секвенирования фрагментов мтДНК, фрагментного анализа микросателлитных локусов, исследования однонуклеотидного полиморфизма SNP) и перемещений животных с помощью радиотелеметрии и спутниковых передатчиков показали различную степень дифференциации популяций — от слабо-выраженных до статистически значимых различий, подтверждающих их пространственную обособленность. Так, анализ перемещений белых медведей и их генетической изменчивости выявил резкий контраст между их группировками с плохо выраженной генетической структурой, окружающими Полярный бассейн, и четырьмя генетически изолятами в канадской Арктике, соответствующими отдельным популяциям медведей.

Предполагается, что такие различия обусловлены сезонным распределением морского льда и тюленей [4]. Группировки медведей Гудзонова залива и Канадского Арктического архипелага генетически хорошо различаются, при этом особый интерес представляет небольшая генетически изолированная группировка Норвежского залива, ограниченного толстым льдом, сушей и полыньями [4, 5]. А наименее генетически различающимися оказались белые медведи Чукотского моря и моря Бофорта. Данные радиотелеметрии показали, что хотя они выделены в отдельные субпопуляции, участки обитания животных в них перекрываются, а это предполагает существование здесь единой популяции белых медведей [6].

В последнее время генетическая структура глобальной популяции белых медведей пересматривается. Одни авторы [7] выделяют три–четыре основных генетических кластера: Канадский Арктический архипелаг, Южная Канада и Полярный бассейн (последний подразделяется на восточный и западный субкластеры). При этом в связи с происходящими изменениями климата предполагается существование направленного потока

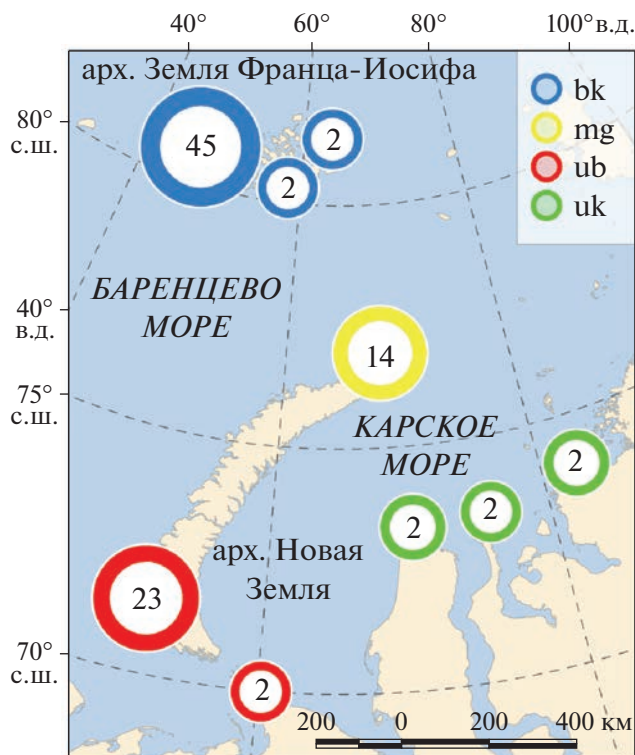


Рис. 1. Места сбора проб для генетического анализа: bk — север Баренцева и Карского морей, uk — юг Карского моря, ub — юг Баренцева моря, mg — мыс Желания. Цифры в кружках — объем выборки.

генов, идущего преимущественно через самцов, из генетического кластера южной Канады и восточного субкластера Полярного бассейна в генетический кластер Канадского Арктического архипелага, который рассматривается в качестве межледникового рефугиума. Другие авторы [8] выделяют шесть генетических кластеров белых медведей: Гудзонова залива, западного и восточного участков Канадского Арктического архипелага, западного и восточного Полярного бассейнов и Норвежского залива. Хотя на численность, распространение и структуру популяции белых медведей негативно влияет быстрая потеря арктического морского льда, эти генетические данные, в отличие от предыдущей работы, не подтверждают существования сильного направленного потока генов в ответ на происходящее изменение климата.

В акваториях Карского и Баренцева морей белые медведи встречаются практически повсеместно, за исключением юго-западной части Баренцева моря, которое также рассматривается в качестве межледникового рефугиума. На основе данных телеметрии в этом регионе выделяют несколько обособленных группировок этого вида: две на севере Баренцева моря и две (северную и южную) в Карском море. Имеющиеся данные о перемещениях белых медведей в этом регионе

свидетельствуют об относительной обособленности белых медведей, обитающих в районе острова Южный архипелага Новая Земля [9]. Подавляющее большинство данных о перемещениях получено для взрослых самок, которые привержены достаточно четко ограниченному местообитанию и местам обустройства родовых берлог [10, 11]. При такой выраженной и стабильной пространственной структуре группировок белого медведя остаются неизученными механизмы, которые обеспечивают перемешивание и генетическую однородность мировой популяции. Возможно, она поддерживается за счет расселения молодых особей или миграции самцов, о чем известно значительно меньше, чем о перемещениях самок. Отдельные работы, выполненные на небольшом количестве животных, показывают, что самцы могут перемещаться более активно, оставаясь при этом в пределах территории, характерной для локальной группировки самок [10], или даже используя меньшую территорию [12]. Однако данные о перемещении самцов обычно ограничены несколькими месяцами с момента мечения животного [10, 13]. Имеющиеся данные о генетической структуре и о перемещениях белых медведей карско-баренцевоморского региона были получены в основном на медведях из северной части Баренцева моря. Группировка медведей, обитающая в южной части Баренцева моря, ранее не исследовалась.

Цель данной работы состояла в оценке степени генетической изоляции медведей из разных участков карско-баренцевоморского региона с использованием данных анализа микросателлитных локусов и фрагмента D-петли митохондриальной ДНК.

МАТЕРИАЛЫ И МЕТОДЫ

На исследованной территории (архипелаги Земля Франца-Иосифа и Новая Земля, побережье Карского моря от о-ва Вайгач до п-ова Таймыр) были собраны образцы от отловленных и идентифицированных белых медведей (рис. 1). Все манипуляции с животными проводились после химической иммобилизации. Инъекция смеси вводилась с помощью дистанционного пневматического инъектора DanInject JM-25. Забор крови проводился однократно из бедренной вены с помощью одноразового шприца либо вакуумной системы. Во время проведения процедур у всех животных контролировался пульс, оксигенация и температура тела. Далее животному вводили препарат для вывода из наркоза. Данная методика одобрена комиссией по этике ИПЭЭ РАН, протокол № 37 от 25 мая 2020 г.

Были проанализированы генетическая структура и генетическое разнообразие группировок животных на разных участках карско-баренцево-

морского региона, а также генетическое разнообразие самок и самцов в разных группировках для выявления возможных различий в активности расселения особей за пределы обитания локальной группировки.

Проведен молекулярно-генетический анализ 93 образцов белого медведя, собранных на разных участках карско-баренцевоморского региона. Получены данные по микросателлитным локусам для 89 животных, по митохондриальной ДНК для 93 животных. Информация об использованных в анализе пробах представлена в (табл. 1).

Образцы крови замораживали в пробирках и транспортировали в замороженном виде или наносили на специальные ФТА карты или бинт и высушивали. Остальные образцы консервировали в поваренной соли или 96%-ном этаноле.

ДНК выделяли при помощи набора для выделения DNeasy Blood and Tissue Kit (Qiagen, Германия). Полимеразную цепную реакцию (ПЦР) фрагмента контрольного региона мтДНК проводили с использованием набора Master Mix 5x (Диалат, Россия) и праймеров Bed1 5'-AGCAACAGCTCCACTACCAG-3' и Bed3 5'-CGATTTAGTGGCGTTTATGTAC-3' в следующем режиме: 94°C 3 мин (1 цикл – предварительная денатурация), 94°C 1 мин, 60°C 30 с, 72°C 2 мин (40 циклов), 72°C 5 мин [14]. Очистку продуктов ПЦР проводили методом ферментативной очистки с помощью набора Exo/SAP Go (Grisp, Португалия). Нуклеотидные последовательности определяли с помощью автоматических генетических анализаторов ABI PRISM 3130 и 3500 (Applied Biosystems, США) и применения набора BigDye Terminator v 3.1 (Applied Biosystems). Секвенирование проводили с праймерами, использованными для амплификации. Выравнивание и редактирование последовательностей проводили вручную с помощью программы BioEdit 7.05 [15]. Построение сетей гаплотипов проводилось с помощью программы Network 10.2 [16]. В качестве аутосомных молекулярно-генетических маркеров использовали специально разработанные для медведей 17 микросателлитных локусов из коммерческого набора: COrDIS Bear, включающего 17 пар аутосомных и одну пару половых праймеров (Гордиз, Россия). ПЦР проводили с помощью амплификатора Bio-Rad T 100 (Bio-Rad, США) в объеме 25 мкл, при условиях, рекомендуемых производителем набора. Длины микросателлитных фрагментов определяли на автоматическом генетическом анализаторе Нанофор (Синтол, Россия) и ABI 3500 (Applied Biosystems) с добавлением стандарта длины SD 550 (Гордиз) и программы GeneMapper v 4.1 (Applied Biosystems). Молекулярную изменчивость для микросателлитных локусов AMOVA, соответствие распределения генотипов микросателлитных локусов равновесию Харди–Вайнберга, парные генетические ди-

станции (R_{st}) и значения достоверности, ожидаемую гетерозиготность (H_e) и наблюдаемую гетерозиготность (H_o), гаплотипическое разнообразие (H), нуклеотидную изменчивость (π), парные генетические дистанции (θ_{st}) и среднее число аллелей на локус (N_a) вычисляли в программе Arlequin 3.5.1.2 [17] и программе Ms tool [18]. Для оценки степени генетической дифференциации между популяциями (θ_{st}) использовали молекулярный дисперсионный анализ (AMOVA) с эволюционной моделью Тамура–Неи ($TrN + I + G$) и значением γ равным 0.8. Эта модель была выявлена как наиболее подходящая для наших данных по информационному критерию Акаике после анализа в программе Modeltest 3.7 [19]. Для тестирования генетической обособленности животных использовали также кластерный анализ Байеса, реализованный в программе Structure 2.3.4 [20], и модель Admixture с коррелированными частотами аллелей для K равного от одного до семи с пятью повторами. Применяли 500 000 повторов и период разогрева 50 000. Для выбора оптимального количества кластеров K , суммирования полученных кластеров и построения графического отображения использовали программу Clumpak [21]. Вероятность наличия нуль-аллелей оценивали в программе Cervus 3.0.3 [22].

РЕЗУЛЬТАТЫ

Анализ митохондриальной ДНК

ДНК успешно выделена из 93 образцов. Для всех образцов получены нуклеотидные последовательности фрагмента контрольного региона мтДНК. Длина выравнивания составила 610 пн. Всего обнаружено 15 вариативных сайтов и описан 21 гаплотип. Анализировались четыре предполагаемые группы: север Баренцева и Карского морей (bk), юг Карского моря (uk), юг Баренцева моря (ub), мыс Желания (mg) (табл. 1).

Данные выборки анализировали также отдельно для взрослых животных: самцов и самок. Результаты анализа значения генетической дифференциации θ_{st} представлены в табл. 2–4.

Из-за слабых генетических различий мы объединили выборки bk, uk и mg в одну. При анализе двух группировок достоверное значение θ_{st} составило 0.270. При анализе только самцов $\theta_{st} = 0.343$, только самок $\theta_{st} = 0.251$. Нуклеотидное разнообразие для исследованных выборок составило π (ub) = 0.0031 ± 0.0020 , π (bk + uk + mg) = 0.0047 ± 0.0028 . Гаплотипическое разнообразие H (ub) = 0.480 ± 0.117 , H (bk + uk + mg) = 0.916 ± 0.016 . По величине θ_{st} группировка юга Баренцева моря (ub) значительно отличается от всех остальных. При этом вклад самцов в эту генетическую изоляцию выше чем самок. Это свидетельствует о том,

Таблица 1. Информация об использованных в анализе пробах

Номер	Пол	Предполагаемая группировка	мтДНК	№ NCBI	Микросателлиты
um003	m	bk	+	OP923918	+
um5	m	bk	+	OP923919	+
um6	f	bk	+	OP923920	+
um7	f	bk	+	OP923921	+
um8	m	bk	+	OP923922	+
um9	f	bk	+	OP923923	+
um10	f	bk	+	OP923924	+
um12	m	bk	+	OP923925	+
zfi4	f	bk	+	OP923926	+
um001	m	bk	+	OP923933	+
um002	m	bk	+	OP923934	+
zfi10	f	bk	+	OP923940	+
zfi11	f	bk	+	OP923941	+
zfi12	m	bk	+	OP923942	+
zfi24	f	bk	+	OP923943	+
zfi25	m	bk	+	OP923944	+
zfi26	m	bk	+	OP923945	+
zfi27	f	bk	+	OP923946	+
zfi28	m	bk	+	OP923947	+
zfi29	m	bk	+	OP923948	+
zfi7	f	bk	+	OP923949	+
zfi8	m	bk	+	OP923950	+
zfi9	f	bk	+	OP923951	+
zfi15	f	bk	+	OP923952	+
zfi16	f	bk	+	OP923953	+
zfi17	m	bk	+	OP923954	+
zfi18	f	bk	+	OP923955	+
zfi19	m	bk	+	OP923956	+
zfi20	f	bk	+	OP923957	+
zfi21	f	bk	+	OP923958	+
zfi22	f	bk	+	OP923959	+
zfi30	m	bk	+	OP923984	+
zfi31	m	bk	+	OP923985	+
zfi32	m	bk	+	OP923986	+
zfi33	m	bk	+	OP923987	+
zfi34	f	bk	+	OP923988	+
zfi35	f	bk	+	OP923989	+
zfi36	f	bk	+	OP923990	+
zfi37	m	bk	+	OP923991	+
zfi38	m	bk	+	OP923992	+
zfi39	m	bk	+	OP923998	+
zfi40	f	bk	+	OP923999	+
zfi41	f	bk	+	OP924000	+
zfi42	m	bk	+	OP924001	+

Таблица 1. Продолжение

Номер	Пол	Предполагаемая группировка	мтДНК	№ NCBI	Микросателлиты
zfi43	m	bk	+	OP924002	+
um68	f	bk	+	OP924003	+
zfi3	f	bk	+	OP923968	+
zfi5	m	bk	+	OP923992	+
zfi6	m	bk	+	OP923926	+
nz5	m	mg	+	OP923964	+
nz6	m	mg	+	OP923965	+
nz7	f	mg	+	OP923966	+
nz8	f	mg	+	OP923967	+
um43	f	mg	+	OP923977	+
um44	m	mg	+	OP923978	+
um46	—	mg	+	OP923979	—
um47	m	mg	+	OP923982	—
um53	—	mg	+	OP923983	—
nz9	m	mg	+	OP923993	+
nz10	f	mg	+	OP923994	+
nz11	m	mg	+	OP923995	+
nz12	m	mg	+	OP923996	+
nz13	f	mg	+	OP923997	+
bg5	f	ub	+	OP923927	+
bg2	f	ub	+	OP923928	+
bg3	m	ub	+	OP923929	+
bg1	m	ub	+	OP923930	+
um34	f	ub	+	OP923935	+
um36	f	ub	+	OP923936	+
um38	f	ub	+	OP923937	+
bg4	f	ub	+	OP923960	+
bg6	f	ub	+	OP923961	+
bg7	f	ub	+	OP923962	+
bg8	m	ub	+	OP923963	+
um55	m	ub	+	OP923970	+
um58	m	ub	+	OP923971	+
um59	f	ub	+	OP923972	+
bg9	f	ub	+	OP923973	+
bg10	f	ub	+	OP923974	+
bg11	m	ub	+	OP923975	+
um56	f	ub	+	OP923980	+
um57	m	ub	+	OP923981	+
um39	m	ub	+	OP923931	+
um40	f	ub	+	OP923926	+
b_1	m	ub	+	OP923939	+
b_2	f	ub	+	OP923926	+
t3	m	uk	+	OP923931	+
t4	f	uk	+	OP923932	+

Таблица 1. Окончание

Номер	Пол	Предполагаемая группировка	мтДНК	№ NCBI	Микросателлиты
ya3	f	uk	+	OP923938	+
ya4	f	uk	+	OP923939	+
ya5	m	uk	+	OP923968	+
ya6	f	uk	+	OP923969	+
um30	m	ub	+	OP923926	—

Примечание для табл. 1—4 и 6—8. bk — север Баренцева и Карского морей, mg — мыс Желания, ub — юг Баренцева моря, uk — юг Карского моря.

Таблица 2. Значения генетической дифференциации θ_{st} (под диагональю) и значения достоверности (над диагональю) по данным анализа фрагмента D-петли мтДНК длиной 610 пн для взрослых животных

Группировки	bk	ub	uk	mg
bk		0	0.432	0.180
ub	0.297*		0.027	0
uk	0	0.199*		0.811
mg	0.0326	0.286*	0	

Примечание для табл. 2—4 и 6—8. * Достоверные значения θ_{st} ($p < 0.05$) помечены звездочкой.

Таблица 3. Значения генетической дифференциации θ_{st} (под диагональю) и значения достоверности (над диагональю) по данным анализа фрагмента D-петли мтДНК длиной 610 пн для самок

Группировки	bk	ub	uk	mg
bk		0	0.450	0.396
ub	0.301*		0.126	0.009
uk	0	0.101		0.676
mg	0	0.316*	0	

Таблица 4. Значения генетической дифференциации θ_{st} (под диагональю) и значения достоверности (над диагональю) по данным анализа фрагмента D-петли мтДНК длиной 610 пн для самцов

Группировки	bk	ub	uk	mg
bk		0	0.901	0.477
ub	0.245*		0.216	0.018
uk	0	0.267		0.892
mg	0	0.256*	0	

что взрослые самцы не покидают территории отлова или проводят значительную часть времени на этой территории. Также возможным объяснением может быть эффект основателя и родство животных из этой группировки между собой, так как подавляющее число животных из этой выборки относится к одному гаплотипу. Изменчивость группы и сходство по выбранному митохондриальному маркеру можно оценить по медианной сети гаплотипов (рис. 2) и распределению одинаковых гаплотипов по исследованной территории (рис. 3).

Микросателлитный анализ

Для проведения микросателлитного анализа использована выборка из 89 образцов взрослых белых медведей. По всем 17 локусам для выборки описано 178 аллелей (в среднем 10.47 аллелей на локус). Частоты аллелей исследованных локусов приведены в табл. 5.

Значения генетической дифференциации R_{st} по данным анализа микросателлитных локусов для взрослых животных ($n = 89$) представлены в табл. 6—8. Слабую, но достоверную дифференци-

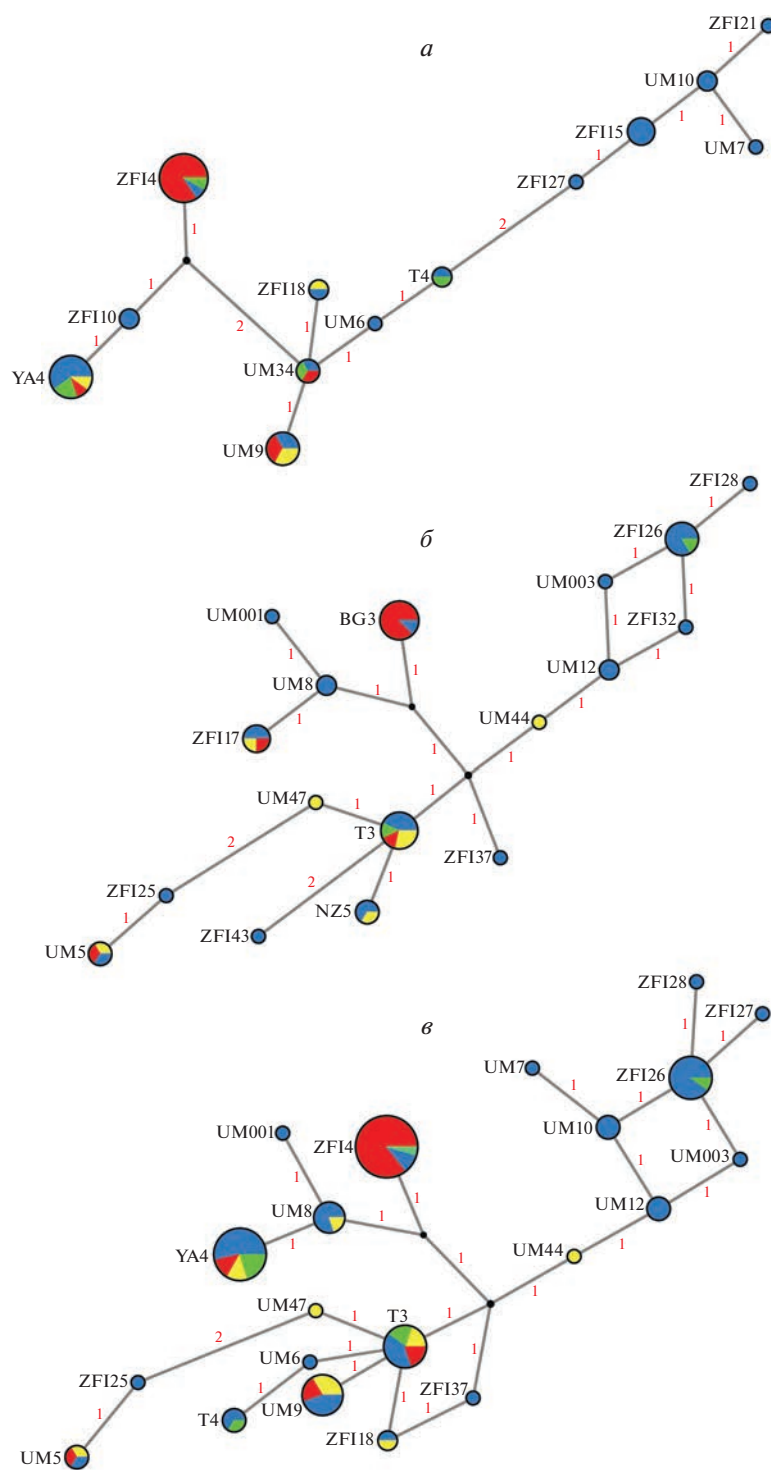


Рис. 2. Медианная сеть гаплотипов исследованных образцов белого медведя от самок (*а*), самцов (*б*) и смешанной выборки (*в*), построенная при помощи программы Network 10.2 на основании выравнивания нуклеотидных последовательностей контрольного региона мтДНК длиной 610 пн. Длина ветвей пропорциональна количеству мутаций, обозначенных красными цифрами у ветвей. Размер окружностей соответствует числу образцов определенного гаплотипа. Цветом обозначены гаплотипы из различных группировок. Цвета выборки соответствуют рис. 1.

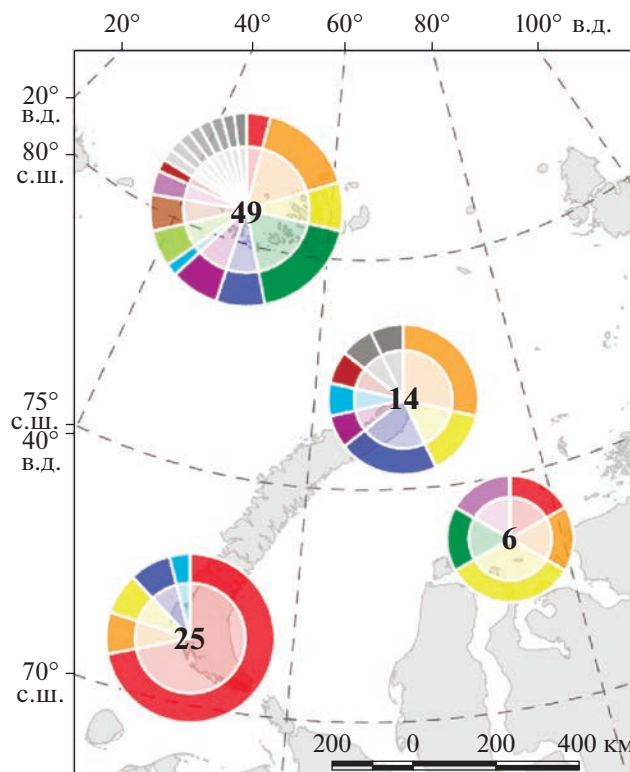


Рис. 3. Распределение белых медведей с выявленными гаплотипами по исследованной территории. Цветами обозначены различные гаплотипы, единичные гаплотипы показаны оттенками серого.

ацию показывают группировки юг Баренцева моря (ub) и север Баренцева и Карского морей (bk). Для выборки самцов и смешанной группы эти различия достоверны, в отличие от выборок самок, что свидетельствует о большем вкладе в изоляцию распределения самцов по исследованной территории.

При различных вариантах объединения исследуемых выборок достоверный результат получился только при объединении выборок bk, uk и mg в одну и сравнении ее с выборкой ub. Обе полученные выборки по всем локусам находились в равновесном состоянии генотипов согласно закону Харди–Вайнберга. Небольшое превышение вероятности наличия нуль-аллелей $F > 0.05$ наблюдается для локуса 7126_411 в выборке bk + uk + mg и для локуса 7045_229 в выборке ub. При анализе двух таких группировок достоверное значение R_{st} составило 0.018 для смешанной выборки, $R_{st} = 0.022$ для самцов и $R_{st} = 0.018$ для самок.

Обработка микросателлитных данных по взрослым животным с помощью кластерного анализа Байеса в программе Structure 2.3.4 с учетом распределения особей по выборкам для наиболее вероятных значений числа генетических кластеров $K = 2$ не выявила различий в популяционной структуре. Однако минимальному значению

$\log \Pr(X|K)$ соответствует $K = 1$, что свидетельствует об отсутствии выраженной генетической структуры популяции (рис. 4). Средние значения наблюдаемой и ожидаемой гетерозиготности по всем локусам для выборки ub составляют: $H_o = 0.810 \pm 0.072$, $H_e = 0.823 \pm 0.050$, для объединенной выборки bk + uk + mg: $H_o = 0.834 \pm 0.049$, $H_e = 0.834 \pm 0.035$.

Мы провели сравнение наших результатов с опубликованными ранее данными по баренцево-морской популяции [11, 23], популяциям моря Бофорта Аляски [24] и Гудзонова залива Канады [25]. Несмотря на то что в цитируемых исследованиях использовались, разные микросателлитные локусы и частично разные фрагменты митохондриальной ДНК, сравнение средних значений ожидаемой (H_e) и наблюдаемой (H_o) гетерозиготности по всем локусам и гаплотипического разнообразия (H) дало сравнимые значения. Исключение составляет значительно меньшее значение H для выборки юга Баренцева моря (ub), вероятно из-за родства животных в выборке (табл. 9).

ОБСУЖДЕНИЕ

Анализ генетических данных с использованием микросателлитов и контрольного региона митохондриальной ДНК для количественной оценки генетической дифференциации позволяет вы-

Таблица 5. Частоты аллелей использованных микросателлитных локусов в общей выборке

Локус	Алель	Частота	Локус	Алель	Частота	Локус	Алель	Частота	Локус	Алель	Частота	Локус	Алель	Частота
7162_072	209	0.023	7204_064	265	0.011	7097_019	301	0.039	7056_179	227	0.011	7031_181	339	0.035
	213	0.148		269	0.062		305	0.118		231	0.017		343	0.070
	217	0.131		273	0.247		309	0.270		235	0.039		347	0.145
	221	0.097		277	0.230		313	0.292		239	0.084		351	0.180
	225	0.244		281	0.225		317	0.197		243	0.225		355	0.256
	229	0.102		285	0.129		321	0.051		247	0.124		359	0.145
	233	0.131		289	0.045		325	0.017		251	0.225		363	0.105
	237	0.102		293	0.039		329	0.011		255	0.140		367	0.029
	241	0.023		297	0.011		337	0.006		259	0.079		371	0.012
7027_138	251	0.017	7206_095	236	0.006	7128_106	147	0.102	7128_139	263	0.051	7200_159	375	0.017
	255	0.045		240	0.006		149	0.011		271	0.006		379	0.006
	259	0.079		244	0.011		151	0.142		166	0.011		183	0.039
	263	0.393		248	0.067		153	0.006		170	0.011		187	0.022
	267	0.146		252	0.135		155	0.170		174	0.219		191	0.051
	271	0.197		256	0.096		159	0.244		178	0.258		195	0.264
	275	0.101		260	0.163		163	0.216		182	0.264		199	0.348
	279	0.017		264	0.213		167	0.063		186	0.163		203	0.124
	283	0.006		268	0.140		171	0.028		190	0.056		207	0.067
7126_411	285	0.035	7079_531	272	0.056	7060_084	175	0.017	7045_229	194	0.011	7148_319	211	0.051
	289	0.052		276	0.067		209	0.017		198	0.006		215	0.022
	293	0.140		280	0.011		213	0.051		235	0.017		219	0.011
	297	0.093		284	0.022		217	0.129		239	0.023		294	0.017
	301	0.209		288	0.006		221	0.197		243	0.131		298	0.006
	305	0.174		229	0.039		225	0.118		247	0.256		302	0.012
	309	0.151		233	0.073		229	0.163		251	0.170		306	0.116
	313	0.047		237	0.146		233	0.101		255	0.142		310	0.099
	317	0.058		241	0.230		237	0.073		259	0.085		314	0.151
7032_021	321	0.017		245	0.247		241	0.039	7211_071	263	0.119		318	0.203
	325	0.023		249	0.174		245	0.039		267	0.057		322	0.174
	298	0.051		253	0.056		249	0.051		202	0.006		326	0.070
	302	0.040		257	0.011		255	0.006		206	0.017		330	0.081
	306	0.136		261	0.017		257	0.011		210	0.185		334	0.041
	310	0.205		341	0.006		261	0.006		214	0.124		338	0.017
	314	0.227								218	0.298		342	0.006
	318	0.136								222	0.202		350	0.006
	322	0.051								226	0.112			
	326	0.063								230	0.056			
	330	0.045												
	334	0.040												
	338	0.006												

Таблица 6. Значения генетической дифференциации R_{st} (под диагональю) и значения достоверности (над диагональю) по данным анализа 17 микросателлитных локусов для взрослых животных

Группировки	mg	uk	ub	bk
mg		0.270	0.099	0.351
uk	0.005		0.252	0.108
ub	0.019	0.014		0
bk	0.003	0.025	0.021*	

Таблица 7. Значения генетической дифференциации R_{st} (под диагональю) и значения достоверности (над диагональю) по данным анализа 17 микросателлитных локусов для самок

Группировки	mg	uk	ub	bk
mg		0.432	0.504	0.919
uk	0		0.090	0.171
ub	0	0.039		0.081
bk	0	0.028	0.015	

Таблица 8. Значения генетической дифференциации R_{st} (под диагональю) и значения достоверности (над диагональю) по данным анализа 17 микросателлитных локусов для самцов

Группировки	mg	uk	ub	bk
mg		0.198	0.387	0.171
uk	0.020		0.586	0.594
ub	0.003	0		0.036
bk	0.017	0	0.026*	

делить в исследованном нами регионе только две популяции (табл. 9). Данные группировки различаются по распределению митохондриального маркера ($\theta_{st} = 0.270$) и очень слабо дифференцируются по ядерным локусам ($R_{st} = 0.018$), что не позволяет отличать с помощью байесовского кластерного анализа животных из этих группировок. Возможное объяснение генетической дифференциации белых медведей по митохондриальной ДНК на юге архипелага Новая Земля — это влияние антропогенных факторов. При отсутствии возможности охоты на тюленей медведи имеют легкий доступ к большому количеству пищевых отходов человека на свалке ТБО. Результаты двухлетнего мониторинга и мечения белых медведей на архипелаге Новая Земля в районе Белушней Губы и анализ литературных источников также позволяют предположить, что здесь обитает и размножается отдельная группировка медведей. Изучение перемещений меченых самок и визуальные наблюдения показали, что медведи ежегодно скапливаются в районе свалки ТБО вблизи Белушней Губы в конце осени в количестве десят-

ков особей и уходят на лед, как только он образуется. Помеченные медведицы зимой и весной не покидали ледовых местообитаний в морских акваториях о-ва Южный и о-ва Вайгач, а после таяния льда много времени проводили в глубине о-ва Южный. В конце осени они вновь перемещались в район Белушней Губы [26].

Различия мтДНК, вероятно, отражают как генетический дрейф, так и историческую динамику колонизации территории южной части Баренцева моря. В отличие от этого дифференциация, получаемая с помощью микросателлитов, находится только в масштабе сотен лет и отражает отсутствие современных препятствий для потока генов. Похожие результаты были получены для животных из залива Бутия и пролива Мак-Клинтон Канадского Арктического архипелага [27]. Возможно, слабая генетическая дифференциация будет усиливаться при использовании других генетических маркеров — однонуклеотидных замен (SNP) с более низкой изменчивостью и более чувствительных к эффекту основателя. Подтверждение этому можно увидеть в исследовании обширной выборки из

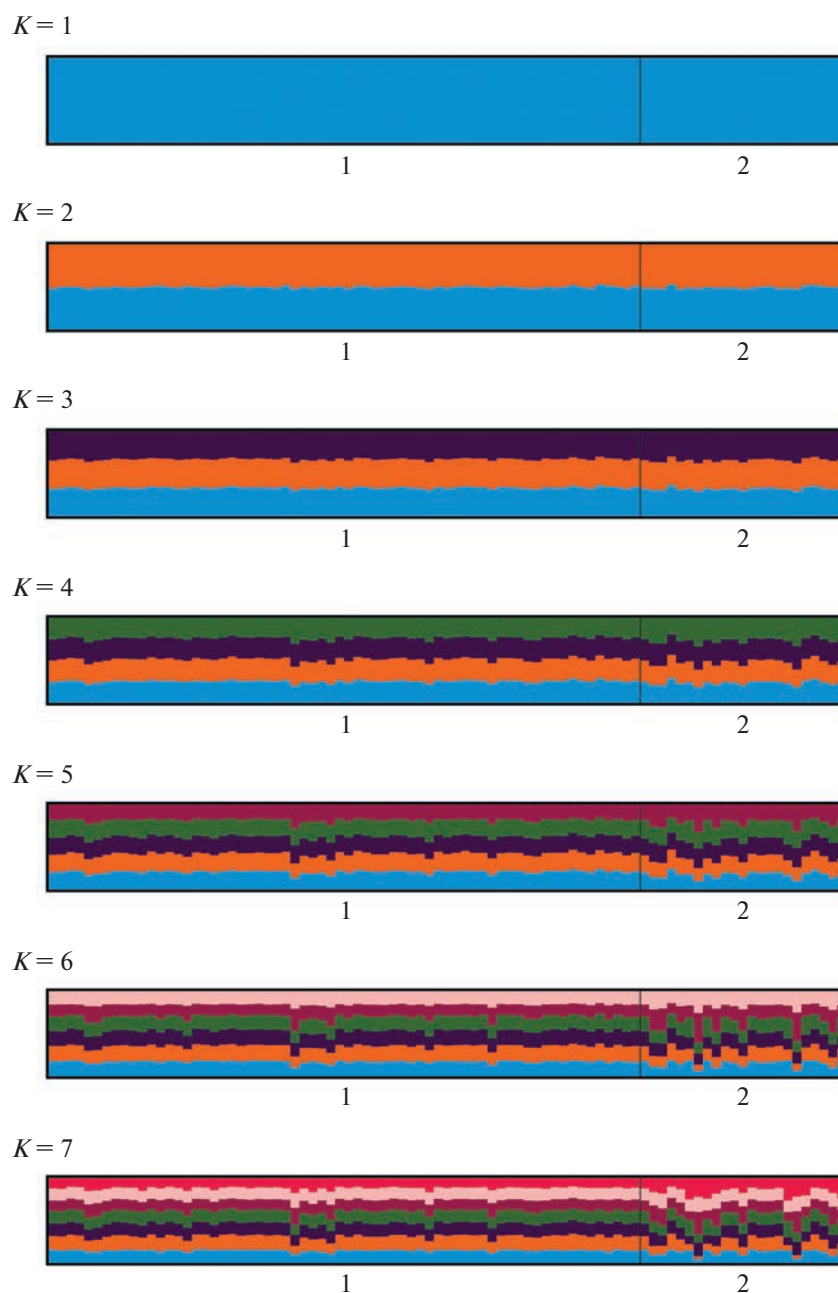


Рис. 4. Генетическая структура популяции белого медведя, исходя из анализа микросателлитных данных в программе Structure 2.3.4, где K равно от 1 до 7 с пятью повторами. Для построения использованы 500 000 повторов и период разогрева 50 000.

Чукотского и Баренцева морей, южной части моря Бофорта и южной части Гудзонова залива. Эти данные указывают на наличие генетической структуризации среди нескольких существующих популяций, расположенных с востока на запад в Полярном бассейне. F_{st} находился в диапазоне от 0.004 до 0.105 ($P < 0.005$) среди всех пар популяций, кроме популяций из южной части моря Бофорта и Чукотского моря ($F_{st} = 0.004$, $P = 0.2$). При этом выборки из Баренцева моря (Шпицбер-

ген) и южной части Гудзонова залива значительно отличаются от выборок из Аляски [28]. Более сильное воздействие на генетическую изоляцию определяют рельеф и распределение пищевых ресурсов белого медведя. Например, в работе о дифференциации группировки Гудзонова залива с использованием 26 микросателлитных локусов и генетических профилей от 377 белых медведей были выявлены три кластера и значительная их дифференциация по индексу F_{st} , что позволяет

Таблица 9. Значения индексов генетического разнообразия по литературным и собственным данным. Нуклеотидное разнообразие (π), гаплотипическое разнообразие (H), наблюдаемая гетерозиготность (H_o), ожидаемая гетерозиготность (H_e), среднее число аллелей на локус N_a

Выборка	H	π	H_o	H_e	N_a
Популяция юга Баренцева моря (ub)	0.480 ± 0.117	0.0031 ± 0.0020	0.810 ± 0.072	0.823 ± 0.050	7.82 ± 1.70
bk + uk + mg	0.916 ± 0.016	0.0047 ± 0.0028	0.834 ± 0.049	0.834 ± 0.035	10.23 ± 1.92
Популяция севера Баренцева и Карского морей [11, 23]	0.902 ± 0.014	0.0032 ± 0.0018	0.61 ± 0.24	0.62 ± 0.24	8.04 ± 3.07
Популяция моря Бофорта Аляски [24]	—	—	0.708	0.706	7.9
Популяция Гудзонова залива Канады [25]	—	—	0.665	0.691	7.65 ± 2.17

предположить дифференциацию белых медведей, взятых с островов в заливе Джеймс [24]. Эти данные были дополнены и уточнены исследованиями с помощью анализа 2603 однонуклеотидных полиморфизмов. Структура популяций ожидаемо отличалась от микросателлитного исследования. Был выделен дополнительный генетический кластер внутри группировки — бассейн Фокс (FB) [29]. Дальнейшие работы по этому региону уже на 13488 однонуклеотидных заменах показали соответствие с более ранними работами [30]. Исследование исторических выборок с 1995 по 2016 гг. на микросателлитных данных на архипелаге Шпицберген в Баренцевом море показывает усиление генетической дифференциации на 200% и потерю генетического разнообразия на 3–10% из-за инбридинга отдельных группировок, что, вероятно, связано с изменением ледяного покрова и фрагментацией мест обитания [31]. В нашем исследовании мы также видим большее значение коэффициента инбридинга ($F_{is} = 0.015$) для группировки юга Баренцева моря по сравнению с группировкой севера Баренцева и Карского морей ($F_{is} = 0$), что также может быть в какой-то степени обусловлено изменением климата и деятельностью человека.

Публикация подготовлена при выполнении работ по анализу и обобщению результатов исследований белого медведя в российской Арктике, проведенных в 2010–2021 гг. Постоянно действующей экспедицией РАН по изучению животных Красной книги Российской Федерации и других особо важных животных фауны России, включенной в состав Института проблем экологии и эволюции им. А.Н. Северцова РАН, в рамках гранта Русского географического общества “Изучение редких видов животных (амурский тигр, дальневосточный леопард, ирбис (снежный барс), белуха, белый медведь)” и в рамках работ по изучению и мониторингу белого медведя и моржа как индикаторов устойчивого состояния

морских арктических экосистем по заказу ПАО “НК “Роснефть” (сбор и анализ данных 2020–2021 гг.), а также при поддержке проекта Международного экологического фонда “Чистые моря” в рамках проекта “Хозяин Арктики-2021”.

Авторы благодарят за разработку и возможность использования набора для идентификации медведей российскую компанию “Гордиз”.

Работа выполнена в ЦКП “Инструментальные методы в экологии” ИПЭЭ РАН.

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. Liu S., Lorenzen E.D., Fumagalli M. et al. Population genomics reveal recent speciation and rapid evolutionary adaptation in polar bears // *Cell*. 2014. V. 157. P. 785–794. <https://doi.org/10.1016/j.cell.2014.03.054>
2. Hassanin A. The role of Pleistocene glaciations in shaping the evolution of polar and brown bears. Evidence from a critical review of mitochondrial and nuclear genome analyses // *C. R. Biol.* 2015. V. 338. P. 494–501. <https://doi.org/10.1016/j.crv.2015.04.008>
3. Taylor M.K., Akeagok S., Andriashek D. et al. Delineating Canadian and Greenland polar bear (*Ursus maritimus*) populations by cluster analysis of movements // *Can. J. Zool.* 2001. V. 79. P. 690–709. <https://doi.org/10.1139/z01-028>
4. Paetkau D., Amstrup S.C., Born E.W. et al. Genetic structure of the world's polar bear populations // *Mol. Ecol.* 1999. V. 8. P. 1571–1584. <https://doi.org/10.1046/j.1365-294x.1999.00733.x>
5. Malenfant R.M., Coltman D.W., Davis C.S. Design of a 9K Illumina BeadChip for polar bears (*Ursus maritimus*) from RAD and transcriptome sequencing // *Mol. Ecol. Resour.* 2015. V. 15(3). P. 587–600. <https://doi.org/10.1111/1755-0998.12327>

6. Cronin M.A., Amstrup S.C., Scribner K.T. Microsatellite DNA and mitochondrial DNA variation in polar bears in the Beaufort and Chukchi seas, Alaska. // Can. J. Zool. 2006. № 84. P. 655–660. <https://doi.org/10.1139/Z06-039>
7. Peacock E., Sonsthagen S.A., Obbard M.E. et al. Implications of the circumpolar genetic structure of polar bears for their conservation in a rapidly warming arctic // PLoS One. 2015 V. 10(1). <https://doi.org/10.1371/journal.pone.0112021>
8. Malenfant R.M., Davis C.S., Cullingham C.I., Coltman D.W. Circumpolar genetic structure and recent gene flow of polar bears: a reanalysis // PLoS One. 2016. V. 11(3). <https://doi.org/10.1371/journal.pone.0148967>
9. Mauritzen M., Derocher A.E., Wiig Ø. et al. Using satellite telemetry to define spatial population structure in polar bears in the Norwegian and western Russian Arctic // J. Appl. Ecol. 2002. V. 39. P. 79–90. <https://doi.org/10.1046/j.1365-2664.2002.00690.x>
10. Amstrup S.C., Durner G.M., McDonald T.L. et al. Comparing movement patterns of satellite-tagged male and female polar bears // Can. J. Zool. 2001. V. 79. P. 2147–2158. <https://doi.org/10.1139/z01-174>
11. Zeyl E., Ehrich D., Aars J. et al. Denning-area fidelity and mitochondrial DNA diversity of female polar bears (*Ursus maritimus*) in the Barents Sea // Can. J. Zool. 2010. V. 88. P. 1139–1148. <https://doi.org/10.1139/Z10-078>
12. Laidre K., Born E., Gurarie E. et al. Females roam while males patrol: Divergence in breeding season movements of pack-ice polar bears (*Ursus maritimus*) // Proc. Biol. Sciences. The Royal Society. 2013. V. 280. <https://doi.org/10.1098/rspb.2012.2371>
13. Wiig Ø., Born E., Laidre K. et al. Performance and retention of lightweight satellite radio tags applied to the ears of polar bears (*Ursus maritimus*) // Animal Biotelemetry. 2017. V. 5. P. 1–11. <https://doi.org/10.1186/s40317-017-0124-0>
14. Matsushashi T., Masuda R., Mano T. et al. Microevolution of the mitochondrial DNA control region in the Japanese brown bear (*Ursus arctos*) population // Mol. Biol. Evol. 1999. V. 16(5). P. 676–684. <https://doi.org/10.1093/oxfordjournals.molbev.a026150>
15. Hall T.A. Bioedit: A user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT // Nucl. Acids Symp. Series. 1999. V. 41. P. 95–98.
16. Bandelt H.J., Forster P., Rohl A. Median-Joining networks for inferring intraspecific phylogenies // Mol. Biol. Evol. 1999. V. 16. № 1. P. 37–48. <https://doi.org/10.1093/oxfordjournals.molbev.a026036>
17. Excoffier L.G., Lischer H.E. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows // Mol. Ecol. Resour. 2010. V. 10. P. 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
18. Park S.D.E. Trypanotolerance in West African Cattle and the Population Genetic Effects of Selection. Ph.D. thesis Univ. Dublin, 2001. 53 p.
19. Posada D., Crandall K.A. Modeltest: Testing the model of DNA substitution // Bioinformatics. 1998. V. 14. № 9. P. 817–818. <https://doi.org/10.1093/bioinformatics/14.9.817>
20. Pritchard J.K., Stephens M., Donnelly P. Inference of population structure using multilocus genotype data // Genetics. 2000. V. 155. P. 945–959. <https://doi.org/10.1093/genetics/155.2.945>
21. Kopelman N.M., Mayzel J., Jakobsson M. et al. Clumpak: A program for identifying clustering modes and packaging population structure inferences across K // Mol. Ecol. Resour. 2015. V. 15(5). P. 1179–1191. <https://doi.org/10.1111/1755-0998.12387>
22. Kalinowski S.T., Taper M.L., Marshall T.C. et al. Revisiting how the computer program CERVUS accommodates genotyping error increases success in paternity assignment // Mol. Ecol. 2007. V. 16. P. 1099–1106. <https://doi.org/10.1111/j.1365-294X.2007.03089.x>
23. Zeyl E., Aars J., Ehrich D., Wiig Ø. Families in space: Relatedness in the Barents sea population of polar bears (*Ursus maritimus*) // Mol. Ecol. 2009. V. 18(4). P. 735–749. <https://doi.org/10.1111/j.1365-294X.2008.04049.x>
24. Cronin M.A., Amstrup S.C., Talbot S.L. et al. Genetic variation, relatedness, and effective population size of polar bears (*Ursus maritimus*) in the southern Beaufort sea, Alaska // J. Hered. 2009. № 100. P. 681–690. <https://doi.org/10.1093/jhered/esp061>
25. Crompton A.E., Obbard M.E., Petersen S.D., Wilson P.J. Population genetic structure in polar bears (*Ursus maritimus*) from Hudson Bay, Canada: implications of future climate change // Biol. Conserv. 2008. V. 141(10). P. 2528–2539. <https://doi.org/10.1016/j.biocon.2008.07.018>
26. Платонов Н.Г., Мизин И.А., Иванов Е.А. и др. Использование белым медведем (*Ursus maritimus*) местообитаний вдоль береговой линии в течение года по данным спутникового мониторинга // Исследование Земли из космоса. 2019. № 3. С. 80–91. <https://doi.org/10.31857/S0205-96142019380-91>
27. Campagna L., Van Coeverden de Groot P.J., Saunders B.L. Extensive sampling of polar bears (*Ursus maritimus*) in the Northwest Passage (Canadian Arctic Archipelago) reveals population differentiation across multiple spatial and temporal scales // Ecol. Evol. 2013. V. 3(9). P. 3152–3165. <https://doi.org/10.1002/ece3.662>
28. Miller W., Schuster S.C., Welch A.J. et al. Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change // Proc. Natl. Acad. Sci. USA. 2012. V. 109(36). P. E2382–E2390. <https://doi.org/10.1073/pnas.1210506109>
29. Viengkone M., Derocher A.E., Richardson E.S. et al. Assessing polar bear (*Ursus maritimus*) population structure in the Hudson Bay region using SNPs // Ecol. Evol. 2016. V. 6. P. 8474–8484. <https://doi.org/10.1002/ece3.2563>
30. Jensen E.L., Tschirner C., de Groot P.V.C. et al. Canadian polar bear population structure using genome-wide markers // Ecol. Evol. 2020. P. 1–9. <https://doi.org/10.1002/ece3.615>
31. Maduna S., Aars J., Fløystad I. et al. Sea ice reduction drives genetic differentiation among Barents sea polar bears // Proc. R. Soc. B. 2021. <https://doi.org/10.1098/rspb.2021.1741>

Population Genetic Structure in Polar Bears (*Ursus maritimus*) from the Russian Arctic Seas

P. A. Sorokin^{a, *}, E. Yu. Zvychnaya^a, E. A. Ivanov^a, I. A. Mizin^b, I. N. Mordvintsev^a,
N. G. Platonov^a, A. I. Isachenko^c, R. E. Lazareva^c, and V. V. Rozhnov^a

^aSevertsov Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow, 119071 Russia

^b“Russian Arctic” National Park, Arkhangelsk, 163051 Russia

^cLTD “Arctic Research Center”, Moscow, 119333 Russia

*e-mail: sorokin-p@yandex.ru

Population genetic structure in polar bear (*Ursus maritimus*) from model areas in the Russian Arctic is considered based on materials collected in the period 2010–2021. Data on polymorphism of 17 microsatellite loci of nuclear DNA and a 610 nucleotide long mtDNA D-loop fragment were obtained for 93 animals. For the studied sample of adult polar bears, a high genetic diversity of nuclear DNA and a low value of nucleotide variability π for mitochondrial DNA were found. For all genetic markers, differentiation of bears from the southern part of the Barents Sea from animals from the north of the Barents and Kara seas was found. These groups differ in the distribution of the mitochondrial marker ($\theta_{st} = 0.270$) and are weakly differentiated by nuclear loci ($R_{st} = 0.018$).

Keywords: polar bear, *Ursus maritimus*, control region, microsatellite, polymorphism, Barents Sea, Kara Sea.

НЕВЫРАЖЕННОСТЬ ПОПУЛЯЦИОННО-ГЕНЕТИЧЕСКОЙ СТРУКТУРЫ СЕРОГО ЖУРАВЛЯ *Grus grus* L.

© 2023 г. Е. А. Мудрик^{1, *}, Ю. М. Маркин², К. А. Постельных²,
К. Д. Кондракова^{1, 3}, Т. А. Кашенцева², А. В. Шатохина¹, П. А. Казимиров¹,
В. Ю. Ильяшенко³, Е. И. Ильяшенко^{1, 3}, Д. В. Политов¹

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

²Окский государственный природный биосферный заповедник, Рязанская обл., Брыкин Бор, 391072 Россия

³Институт проблем экологии и эволюции им. А.Н. Северцова Российской академии наук, Москва, 119011 Россия

*e-mail: mudrik@vigg.ru

Поступила в редакцию 10.07.2023 г.

После доработки 15.07.2023 г.

Принята к публикации 18.07.2023 г.

В статье представлен расширенный по сравнению с предыдущими исследованиями анализ популяционно-генетической структуры мигрирующего широкоареального вида с высокой численностью — серого журавля *Grus grus* L. С использованием семи высокополиморфных микросателлитных локусов получены высокие значения показателей генетического разнообразия, одинаковые в выборках западного (*G. g. grus*) и восточного (*G. g. lilfordi*) подвидов, и низкие коэффициенты генетической дифференциации между этими подвидами ($F_{ST} = 0.008$, $G_{ST} = 0.002$). По данным AMOVA, 99% генетической изменчивости сосредоточено внутри особей *G. grus*. С использованием алгоритма Байесовской кластеризации не выявлено четкой популяционно-генетической структуризации вида согласно STRUCTURE, однако визуализация пространственных паттернов генетической изменчивости в Geneland показала наличие кластера “чистых” подвидов *G. g. grus* и *G. g. lilfordi*, окружающего кластер особей из зоны интерградации этих подвидов. Такой результат на фоне очень низких значений F -статистики генетической дифференциации может указывать на едва уловимые генетические отличия журавлей из этой зоны, возможно, имеющие мутационный характер. Более низкое аллельное разнообразие и отсутствие уникальных аллелей у восточного подвида *G. g. lilfordi* позволяет выдвинуть предположение о его относительной эволюционной молодости и недавнем происхождении от восточных маргинальных популяций номинативного западного подвида.

Ключевые слова: Gruidae, генетическая изменчивость, генетическая дифференциация, микросателлитные локусы, Россия.

DOI: 10.31857/S0016675823120068, **EDN:** QDIEVG

Серый журавль (*Grus grus* Linnaeus, 1758) — один из самых широкораспространенных и многочисленных видов журавлей. Его мировая численность превышает 700 тыс. особей, при этом в европейской части ареала она в последние десятилетия увеличивается в связи с потеплением климата и обилием кормовых ресурсов [1], что в целом определяет состояние вида как вызывающее наименьшие опасения [2]. Серый журавль гнездится почти на всей территории Северной Евразии от тундры на севере до степей и гор на юге. Во внутривидовой структуре *G. grus* выделяют четыре подвида [3–5], два из которых — западный (*G. g. grus* Linnaeus, 1758) и восточный (*G. g. lilfordi* Sharpe, 1894) — населяют, соответственно, Европу и азиатскую часть России, равнинную часть Казахстана и север Монголии и Китая, и разделены естественной географической грани-

цей, проходящей по Уралу [5]. Морфологически эти два подвида различаются в основном размерами тела и яиц, окраской глаз, “короны” и третьестепенных маховых перьев [3]. Птиц, гнездящихся в России, можно условно отнести к трем группировкам: европейской, западносибирско-центрально-азиатской и восточносибирской. У них разные пролетные пути, а места зимовок преимущественно приурочены к Северо-Восточной Африке, странам Ближнего Востока, Индии и Китаю [1]. Места обитания серого журавля в России занимают самую большую площадь гнездовой части ареала вида [6] (рис. 1,а), при этом численность птиц и плотность населения серых журавлей здесь гораздо более низкая, чем в странах Западной и Центральной Европы. Так, на 500 тыс. западноевропейских серых журавлей приходится около 80 тыс. особей в европейской части России

[7, 8] и 125 тыс. особей восточного подвида в азиатской части ареала (Россия, Казахстан, Монголия, Северный Китай) [9]. Протяженная территория, разрыв ареала в зоне Уральских гор, подразделенность на подвиды, использование разных пролетных путей, выраженная гнездовая филопатрия могут ограничивать поток генов и способствовать популяционно-генетической структуризации серого журавля в гнездовой части ареала.

В 2015 г. нами были получены пилотные данные о слабой генетической дифференциации серого журавля с использованием микросателлитных локусов из нескольких панелей, разработанных для других видов журавлей [10–12], и небольшого количества птиц природного происхождения, содержащихся в зоопарках и центрах разведения [13]. В период 2016–2022 гг. мы собрали в природе более репрезентативный биологический материал от большего количества птиц с более обширной территории, а также освоили более информативные маркеры и более чувствительные методы статистического анализа.

Цель работы – анализ генетической изменчивости и дифференциации серого журавля на значительной территории российской части ареала с акцентом на выявление потенциальных преград для потока генов между западным и восточным подвидами.

МАТЕРИАЛЫ И МЕТОДЫ

Биологический материал и выделение ДНК

В качестве биологического материала мы использовали индивидуальные образцы перьев и крови от 127 особей серого журавля из мест обитания западного (100 особей) и восточного (27) подвидов, собранные в пределах территории Российской Федерации (табл. 1, рис. 1,б). Большая часть образцов (98), преимущественно в европейской части России, собрана при мечении птиц на местах гнездования и осенних предмиграционных скоплений. Последующее дистанционное слежение по сигналам GPS-GSM-передатчиков свидетельствовало, что осенние скопления в основном были представлены местными птицами, за исключением скопления в Актанышском р-не Республики Татарстан, где были взяты образцы от 13 особей. С помощью передатчиков выявлены места гнездования четырех из них – двух в Пермском крае и двух в Республике Коми [14]. От остальных девяти журавлей данных о местах их гнездования не получено. Один журавль, помеченный осенью 2019 г. в долине р. Маныч в Ставропольском крае, в течение двух последующих после мечения лет гнезвился в Западном Казахстане [14]. Для большей географической репрезентативности, особенно в отношении восточного подвида, мы дополнили анализируемые выборки

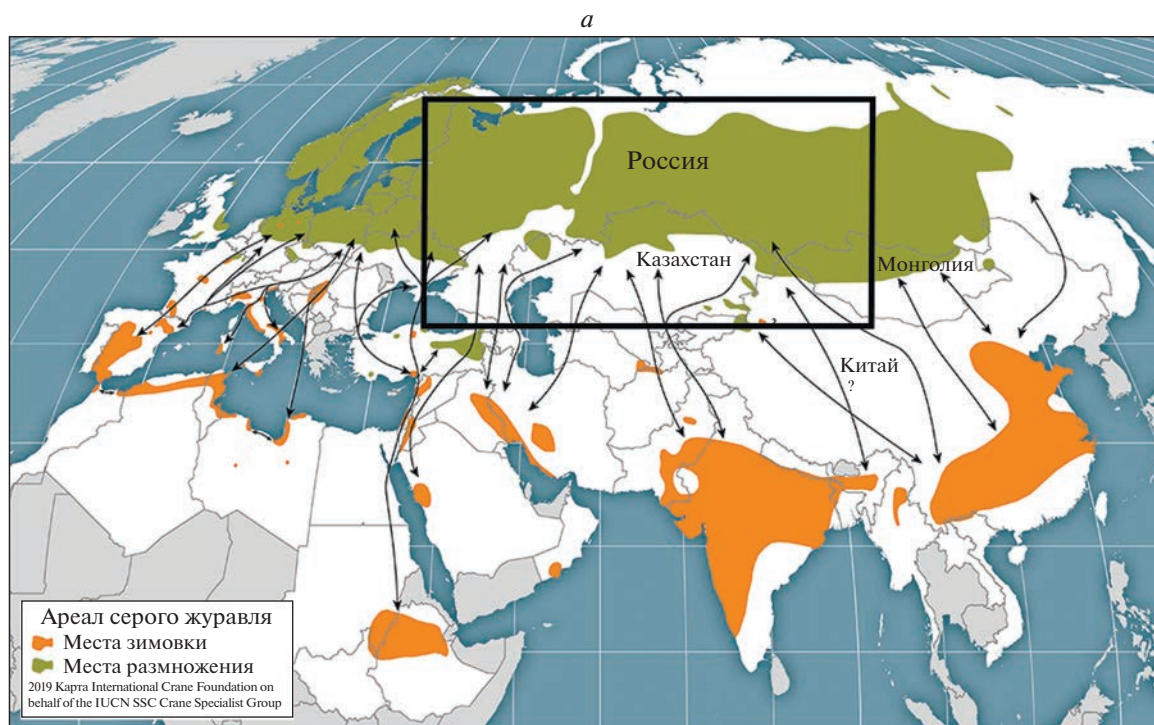
образцами от вольерных птиц с известным происхождением (табл. 1). Как правило это травмированные птицы, попавшие в региональные зоопарки в молодом возрасте, что с большой вероятностью указывает на их местное происхождение. Экстракцию ДНК осуществляли с применением наборов D1Atom™ DNAPrep100 Kit (ООО “Лаборатория Изоген”, Россия) и K-сорб (НПК “Синтол”, Россия) по протоколам производителей.

Микросателлитный анализ

В связи с отсутствием специфических микросателлитных локусов для серого журавля в данной работе мы впервые протестировали и применили гетерологичные локусы с 4–6-нуклеотидными повторами из панели *Lia* (табл. 2) [15], изолированные из генома одного из ближайших родственных серому журавлю [16] видов – японского журавля (*G. japonensis* Muller, 1776). Из 12 протестированных локусов нами было отобрано восемь полиморфных, однако локус *Lia31298* впоследствии был исключен из анализа из-за сильного отклонения от равновесия Харди–Вайнберга, обусловленного высоким процентом содержания нуль-аллелей (>10%). Остальные семь высокоизменчивых локусов (*Lia3745*, *Lia60455*, *Lia62171*, *Lia47103*, *Lia20751*, *Lia4342*, *Lia11091*) характеризовались равновесным состоянием генотипов и отсутствием нуль-аллелей. При проведении ПЦР мы использовали набор реагентов GenPak PCR Core (ООО “Лаборатория Изоген”, Россия) и амплификатор Dyad Peltier Thermal Cycler (“Bio-Rad”, США). Условия ПЦР взяты из первоисточника [15]. Для определения аллельного состава анализируемых локусов полученные фрагменты ДНК были подвергнуты электрофоретическому разделению в 6%-ном полиакриламидном геле в трис-ЭДТА-боратной буферной системе с использованием вертикальных камер модели VE-20 (ООО “Компания Хеликон”, Россия). Визуализацию ПЦР-продуктов осуществляли окрашиванием гелей бромистым этидием и обработкой УФ-светом в системе гель-регистрации Kodak Edas 290 (США). Размеры полученных фрагментов ДНК (аллелей микросателлитных локусов) определяли с помощью программы GelAnalyzer 19.1 [17].

Статистический анализ

Наличие и частоту нуль-аллелей для каждого микросателлитного локуса определяли в программе MICRO-CHECKER 2.2.3 [18]. Параметры генетического разнообразия, равновесие генотипов по Харди–Вайнбергу, *F*-статистики межпопуляционной дифференциации, поток генов (*N_m*) и молекулярную дисперсию (AMOVA) рассчитывали с помощью надстройки для электронной таблицы MSEXcel – GenAlEx 6.503 [19]. Для



Mirande C.M., Harris J.T., editors. 2019. Crane Conservation Strategy. Baraboo, Wisconsin, USA: International Crane Foundation. С изменениями

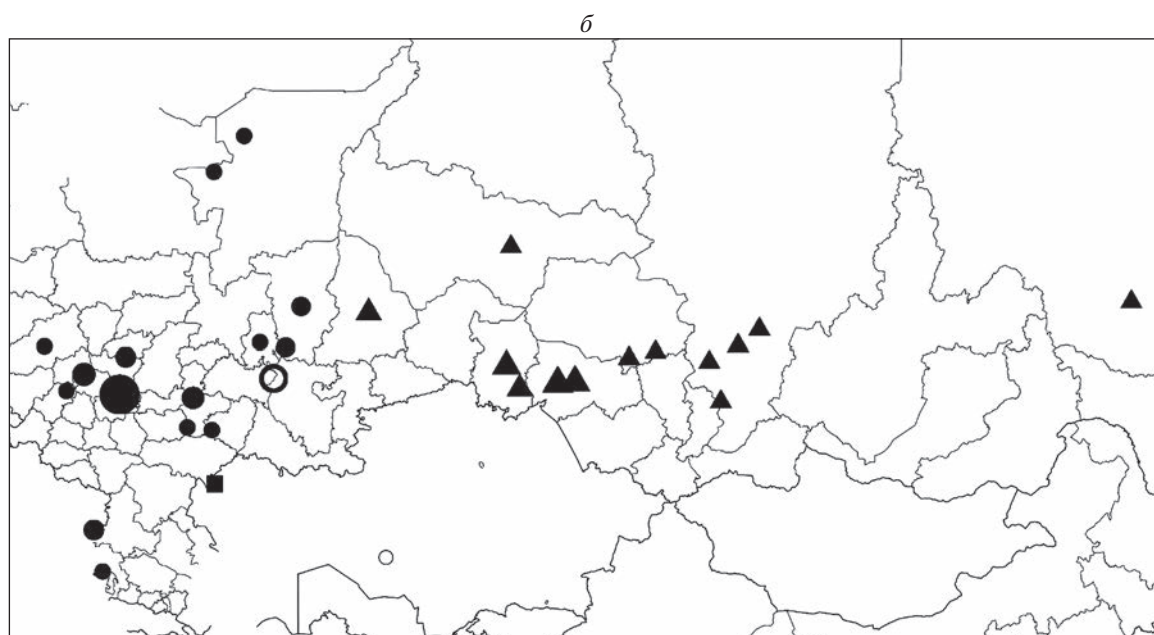


Рис. 1. Ареал серого журавля по [6] с выделенной областью сбора биологических образцов (а) и карта происхождения образцов серого журавля на выделенной области (б). Круглые метки обозначают образцы западного подвида *G. g. grus*, треугольные — восточного *G. g. lilfordi*; квадратная метка — место гнездования птицы в Западной Казахстане, помеченной на осеннем скоплении в Ставропольском крае. Незакрашенный кружок обозначает место сбора образцов от серых журавлей на осеннем предмиграционном скоплении в Татарстане (места их гнездования не известны). Размеры меток коррелируют с количеством особей.

выявления популяционной структуры использовали алгоритм Байесовской кластеризации в программах STRUCTURE 2.3.4 [20] и Geneland 3.3.0 [21–28]. В обеих программах для определения ве-

роятного числа генетических кластеров K по методу Эванно [29] применяли одинаковые параметры: период разогрева (burn-in) — 100 000 итераций; количество повторов MCMC (Markov chain

Таблица 1. Происхождение биологических образцов западного (*Grus grus grus*) и восточного (*G. g. lilfordi*) подви-
дов серого журавля

Регион	Количество особей		Место содержания
	из природы	вольерные	
G. g. grus			
Московская обл.	7	—	—
Владимирская обл.	3	—	—
Калужская обл.	—	1	Парк птиц “Воробьи”
Тверская обл.	—	1	
Рязанская обл.	60	—	Питомник редких видов журавлей
Ростовская обл.	3	—	
Краснодарский край	—	1	Ростовский-на-Дону зоопарк
Пензенская обл.	—	1	Пензенский зоопарк
Самарская обл.	—	1	Самарский зоопарк
Ульяновская обл.	5	—	—
Пермский край*	2	—	—
	—	2	Пермский зоопарк
Республика Коми*	2	—	—
Удмуртская Республика	—	1	Ижевский зоопарк
Республика Татарстан**	9	—	—
Западный Казахстан"	1	—	—
Всего	92	8	
G. g. lilfordi			
Ханты-Мансийский АО	—	1	Мини-зоопарк г. Сургут
Свердловская обл.	—	2	Екатеринбургский зоопарк
Омская обл.	—	3	Детский экоцентр г. Омск
	—	3	Большереченский зоопарк
Новосибирская обл.	6	—	—
	—	4	Новосибирский зоопарк
	—	1	Якутский зоопарк “Орто Дойду”
Томская обл.	—	2	Северский зоопарк
Красноярский край	—	1	Парк флоры и фауны “Роев Ручей”
	—	3	Абаканский зоопарк
Республика Якутия	—	1	Якутский зоопарк “Орто Дойду”
Всего	6	21	

Примечание. * — места гнездования выявлены с помощью дистанционного слежения серых журавлей, помеченных в Актанышском р-не Республики Татарстан. ** — место мечения и сбора образцов западного подвида на осеннем предмиграционном скоплении в Актанышском р-не Республики Татарстан (места гнездования не выявлены). " — место гнездования серого журавля, помеченного в долине р. Маныч в Ставропольском крае.

Monte Carlo) после разогрева — 1000000; модель “admixture”; число кластеров *K* от 1 до 7, число повторов для каждого значения *K* — три для STRUCTURE и пять для Geneland. Визуализацию результатов STRUCTURE осуществляли с использованием веб-сервиса CLUMPAK [30], для интерпретации выбирали наиболее вероятное и несколько соседних значений *K*. Визуализацию пространственных паттернов популяционно-генетической структуры серого журавля по результатам расчетов в Geneland проводили с помощью встроенных в пакет функций, а также программы GeoDa 1.20.0.36 [31]. Карта происхождения образцов подготовлена с использованием пакета PBSMapping 2.73.2 [32] для среды R [33].

РЕЗУЛЬТАТЫ

Характеристика микросателлитных локусов

В общей выборке (127 особей) серых журавлей по семи полиморфным локусам идентифицировано от 11 (*Lia47103*) до 22 (*Lia62171*) аллелей (табл. 2). В выборке западного подвида по каждому локусу и в целом выявлено больше аллелей, чем у восточного, что может объясняться разными размерами изучаемых выборок. Так, общее число аллелей в выборках *G. g. grus* и *G. g. lilfordi* составило, соответственно, 114 и 83. У восточного подвида не выявлено уникальных аллелей, то есть 83 аллеля, идентифицированные у *G. g. lilfordi*, являются общими для обоих подвигов, а аллельное разнообразие западного подвида выше на 31 аллель.

Таблица 2. Характеристика микросателлитных локусов, количество аллелей и частота наиболее распространенного аллеля в выборках западного (*Grus grus grus*) и восточного (*G. g. lilfordi*) подвидов и в общей выборке изученных особей серого журавля *G. grus*

Локус	Мотив	Размер, пн	<i>G. g. grus</i> (<i>n</i> = 100)		<i>G. g. lilfordi</i> (<i>n</i> = 27)		<i>G. grus</i> (<i>n</i> = 127)	
			N_A	наиболее частый аллель	N_A	наиболее частый аллель	N_A	наиболее частый аллель
<i>Lia3745</i>	GGAA ₍₁₁₎	128–184	13	148 (0.230)*	10	148 (0.204)	13	148 (0.224)
<i>Lia60455</i>	CTTT ₍₁₃₎	104–160	15	108 (0.180)	10	132 (0.167)	15	108 (0.169)
<i>Lia62171</i>	CCTT ₍₁₂₎	88–188	22	136 (0.165)	13	132 (0.222)	22	136 (0.150)
<i>Lia47103</i>	AAAAC ₍₉₎	115–165	11	125 (0.310)	10	125 (0.259)	11	125 (0.299)
<i>Lia20751</i>	TATTC ₍₂₄₎	140–235	18	140 (0.165)	14	140 (0.204)	18	140 (0.173)
<i>Lia4342</i>	AAAAC ₍₁₁₎	135–240	20	190 (0.120)	14	205 (0.185)	20	190 (0.130)
<i>Lia11091</i>	CTTTTC ₍₈₎	106–196	15	210 (0.120) 148 (0.140)	12	148 (0.222)	15	148 (0.157)

Примечание. N_A — число аллелей на локус, * — в скобках указана частота аллеля.

Таблица 3. Показатели популяционно-генетической изменчивости серого журавля *G. grus*

Локус	<i>G. g. grus</i> (<i>n</i> = 100)			<i>G. g. lilfordi</i> (<i>n</i> = 27)			<i>G. grus</i> (<i>n</i> = 127)					
	H_O	H_E	F	H_O	H_E	F	H_O	H_E	F_{IS}	F_{ST}	G_{ST}	N_m
<i>Lia3745</i>	0.880	0.872	–0.009	0.852	0.863	0.013	0.874	0.873	0.002	0.006	0.000	44.946
<i>Lia60455</i>	0.860	0.890	0.034	0.889	0.876	–0.015	0.866	0.890	0.010	0.005	–0.001	54.232
<i>Lia62171</i>	0.860	0.908	0.052	0.852	0.876	0.027	0.858	0.911	0.040	0.017	0.011	14.305
<i>Lia47103</i>	0.870	0.833	–0.044	0.852	0.854	0.002	0.866	0.840	–0.021	0.004	–0.002	59.886
<i>Lia20751</i>	0.910	0.905	–0.005	0.963	0.889	–0.083	0.921	0.905	–0.044	0.006	0.000	40.640
<i>Lia4342</i>	0.920	0.924	0.004	0.926	0.890	–0.040	0.921	0.922	–0.018	0.009	0.003	27.802
<i>Lia11091</i>	0.900	0.904	0.004	0.963	0.874	–0.102	0.913	0.902	–0.048	0.008	0.003	29.385
В среднем	0.886 ± 0.009	0.891 ± 0.011	0.005 ± 0.012	0.899 ± 0.019	0.874 ± 0.005	-0.028 ± 0.019	0.893 ± 0.010	0.883 ± 0.006	-0.011 ± 0.012	0.008 ± 0.002	0.002 ± 0.002	38.742 ± 6.042

Примечание. H_O — наблюдаемая гетерозиготность, H_E — ожидаемая гетерозиготность, F — индекс фиксации Райта, F_{IS} — внутривидовой коэффициент инбридинга, F_{ST} — межвидовой коэффициент генетической дифференциации, G_{ST} — несмещенная оценка межвидового коэффициента генетической дифференциации.

Наиболее частые аллели по локусам *Lia3745*, *Lia47103*, *Lia20751* и *Lia11091* были общими у обоих подвидов, а по локусам *Lia60455*, *Lia62171* и *Lia4342* совпадали с наиболее частыми аллелями у западного подвида. И хотя среднее число на локус в обеих выборках *G. g. grus* и *G. g. lilfordi* различалось (16.286 ± 1.475 и 11.857 ± 0.705 соответственно), эффективное число аллелей на локус в них было близким, хотя различия оставались достоверными (9.673 ± 0.867 и 8.040 ± 0.312). При этом уникальные для сравниваемых группировок аллели (“приватные” аллели) отсутствовали у восточного серого журавля, в то время как у типового подвида их было найдено 31, причем отмечены они были во всех семи анализируемых локусах.

Популяционно-генетическая изменчивость и внутривидовая дифференциация

Полокусные и средние значения наблюдаемой (H_O) и ожидаемой (H_E) гетерозиготности в выборках западного и восточного серого журавля и в целом для вида были высокими — более 0.850. Так, для *G. g. grus* значения этих показателей составили в среднем: $H_O = 0.886 \pm 0.009$, $H_E = 0.891 \pm 0.011$, для *G. g. lilfordi*: $H_O = 0.899 \pm 0.019$, $H_E = 0.874 \pm 0.005$ (табл. 3). Несмещенные оценки ожидаемой гетерозиготности (H_E) у обоих подвидов были близкими и в среднем составили 0.893 ± 0.006 . Средние значения наблюдаемой и ожидаемой гетерозиготности для вида в целом равны $H_O = 0.893 \pm 0.010$, $H_E = 0.883 \pm 0.006$. Значения индекса

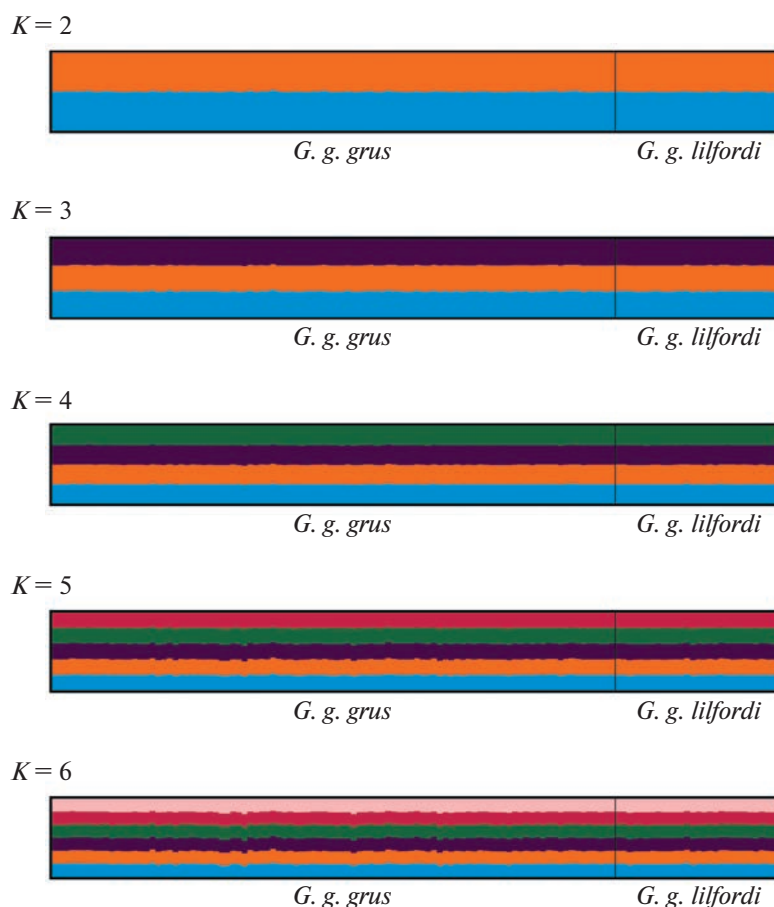


Рис. 2. Результаты кластеризации индивидуальных генотипов выборок западного *G. g. grus* и восточного *G. g. lilfordi* подвидов серого журавля по микросателлитным локусам в программе STRUCTURE для наиболее вероятных значений числа генетических кластеров $K = 4$.

фиксации варьировали по локусам незначительно и в среднем показали небольшой дефицит гетерозигот в выборке западного подвида, нивелирующийся ошибкой ($F = 0.005 \pm 0.012$), и небольшой эксцесс гетерозигот в выборке восточного подвида ($F = -0.028 \pm 0.019$) (табл. 3).

Внутрипопуляционный коэффициент инбридинга ($F_{IS} = -0.011 \pm 0.012$) свидетельствует о незначительном отклонении от нуля и в целом равновесном состоянии генотипов во всей изученной выборке серых журавлей. Значение показателей межпопуляционных коэффициентов генетической дифференциации были низкими: $F_{ST} = 0.008 \pm 0.002$, $G_{ST} = 0.002 \pm 0.002$. По данным AMOVA 99% генетической изменчивости *G. grus* сосредоточено внутри особей и только 1% — между ними. Эти значения ниже полученных нами для серого журавля ранее ($F_{ST} = 0.011$, межпопуляционная дисперсия по AMOVA — 3%) [13]. Поток генов (Nm) между изучаемыми подвидами измеряется 38.742 $\pm$ 6.042 мигрантами на поколение (табл. 3).

Пространственная структура

Оптимальное число генетических кластеров по методу Эванно [29] для программы STRUCTURE было установлено как $K = 4$, однако различий в популяционной структуре серого журавля в западной и восточной частях ареала не выявлено ни при этом значении K , ни при соседних (рис. 2). Для визуализации географических паттернов генетических различий в программе Geneland мы исключили из анализа особи из Татарстана, происхождение которых неизвестно. Наиболее вероятное число генетических кластеров по Эванно для Geneland равнялось $K = 1$ (рис. 3,а, б), что свидетельствует о гомогенности генофонда серого журавля на всей изучаемой территории. Однако при допущении, что изучаемые нами особи с 5%-ной вероятностью относятся к двум популяциям, Geneland выделяет два кластера, обсуждение которых не лишено смысла (рис. 3,в). Так, на карте вероятностей отнесения анализируемых особей к каждому из них видно, что западные и восточные серые журавли относятся к одному генетическому кластеру, а особи из “переходной”

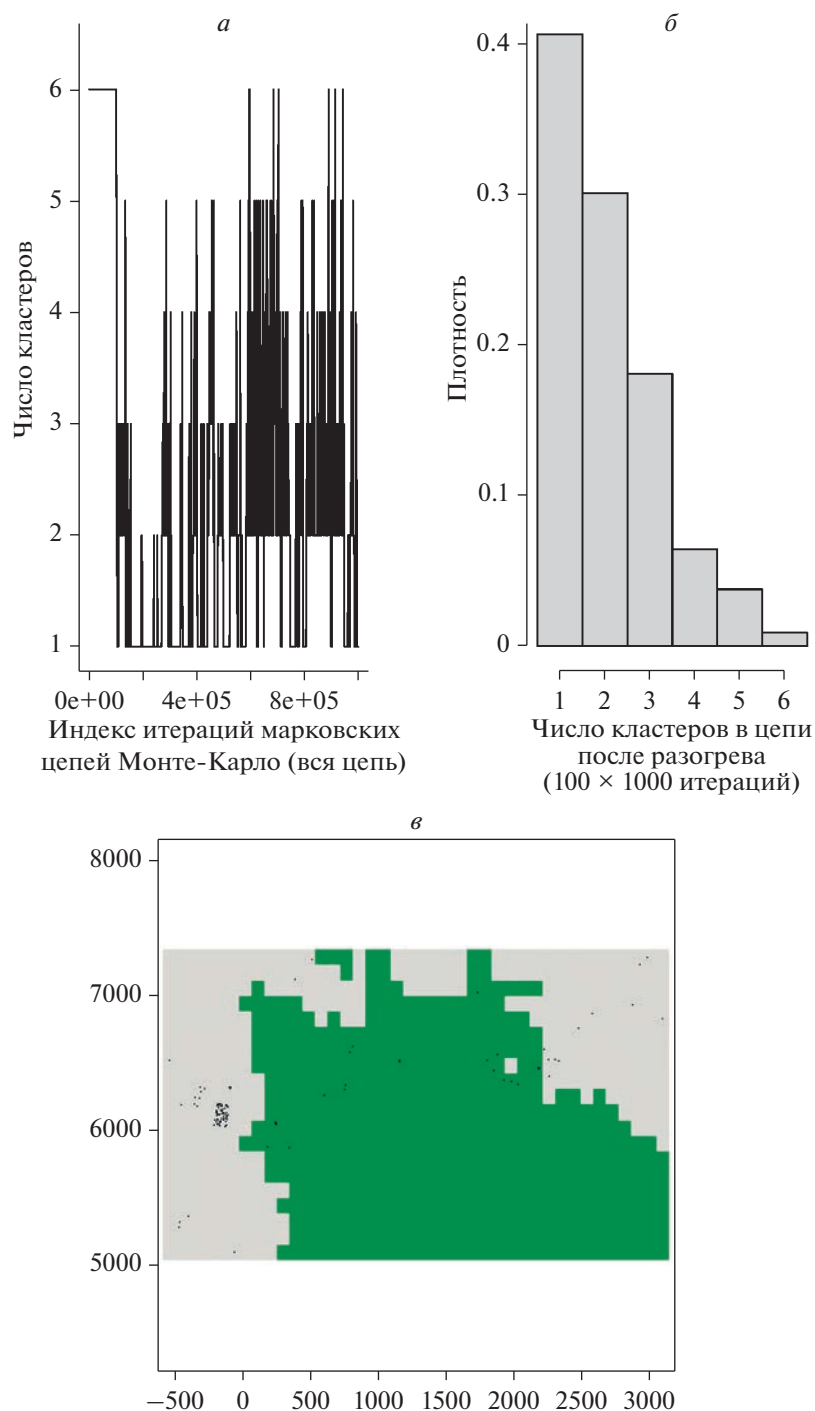


Рис. 3. Определение оптимального значения K для расчетов в Geneland (а, б) и карта наиболее вероятных географических кластеров при допущении $\pm 5\%$ того, что анализируемые птицы относятся к двум популяциям (в).

зоны, находящейся в Предуралье и Зауралье — к другому (рис. 4, 5).

ОБСУЖДЕНИЕ

Серые журавли — моногамные мигрирующие птицы, населяющие протяженную территорию и использующие несколько мест зимовки и разные

миграционные маршруты. Размножающиеся пары характеризуются сильно выраженной гнездовой филопатрией — как правило, из года в год они возвращаются на свой гнездовой участок для размножения и защищают его от конкурентов [34]. Молодые птицы до достижения половой зрелости (возраста трех лет и больше) и образования пар более свободны в своих перемещениях. На при-

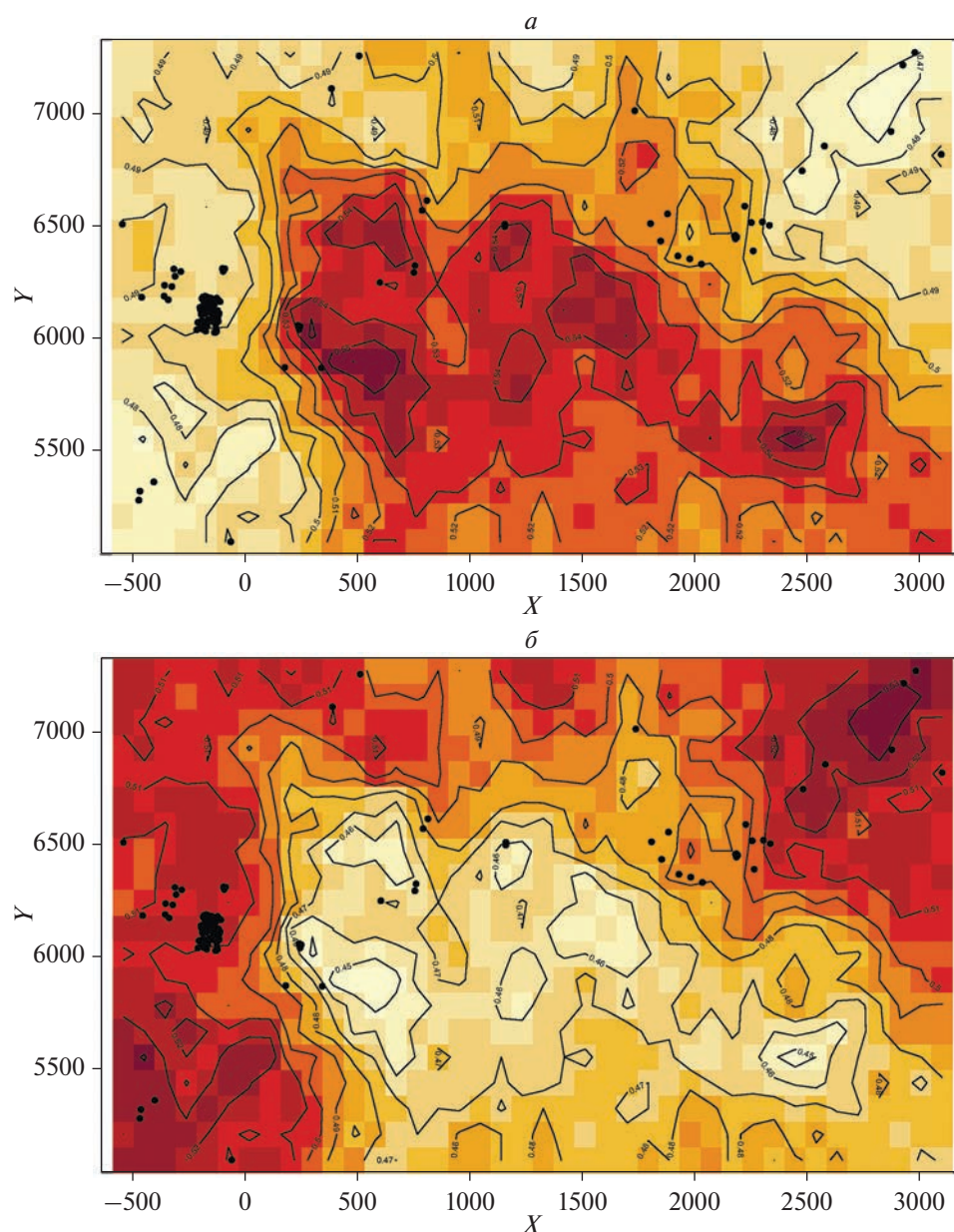


Рис. 4. Карта вероятностей отнесения изученных особей серого журавля к двум географическим кластерам (а, б).

мере меченых GPS-GSM-передатчиками птенцов серого журавля в центральной части России показано, что три самца в годовалом возрасте разлетались в радиусе от 1 до 956 км от гнездовой территории родителей, а возрасте двух, трех и четырех лет они находились в 1–20 км от нее [35]. Один самец каждый год прилетал на территорию, где был помечен птенцом на осеннем скоплении, и загнездились там же в возрасте трех лет [36]. Данные о natalной филопатрии самок серого журавля отсутствуют. Возможно они более дисперсные, чем самцы, как это показано у канадского журавля (*G. canadensis* Linnaeus, 1758) [37, 38].

С использованием других маркеров на расширенных выборках мы получили очередные свидетельства того, что несмотря на принадлежность к разным пролетным путям серый журавль в западной и восточной частях ареала характеризуется гомогенным генофондом. Генетическая дифференциация серого журавля на территории протяженностью около 9 тыс. км еще более низкая, чем было показано нами ранее [13]. Подобный результат также был получен в результате микросателлитного анализа *G. grus* в странах Западной Европы ($F_{ST} = 0.012$) для птиц, использующих два

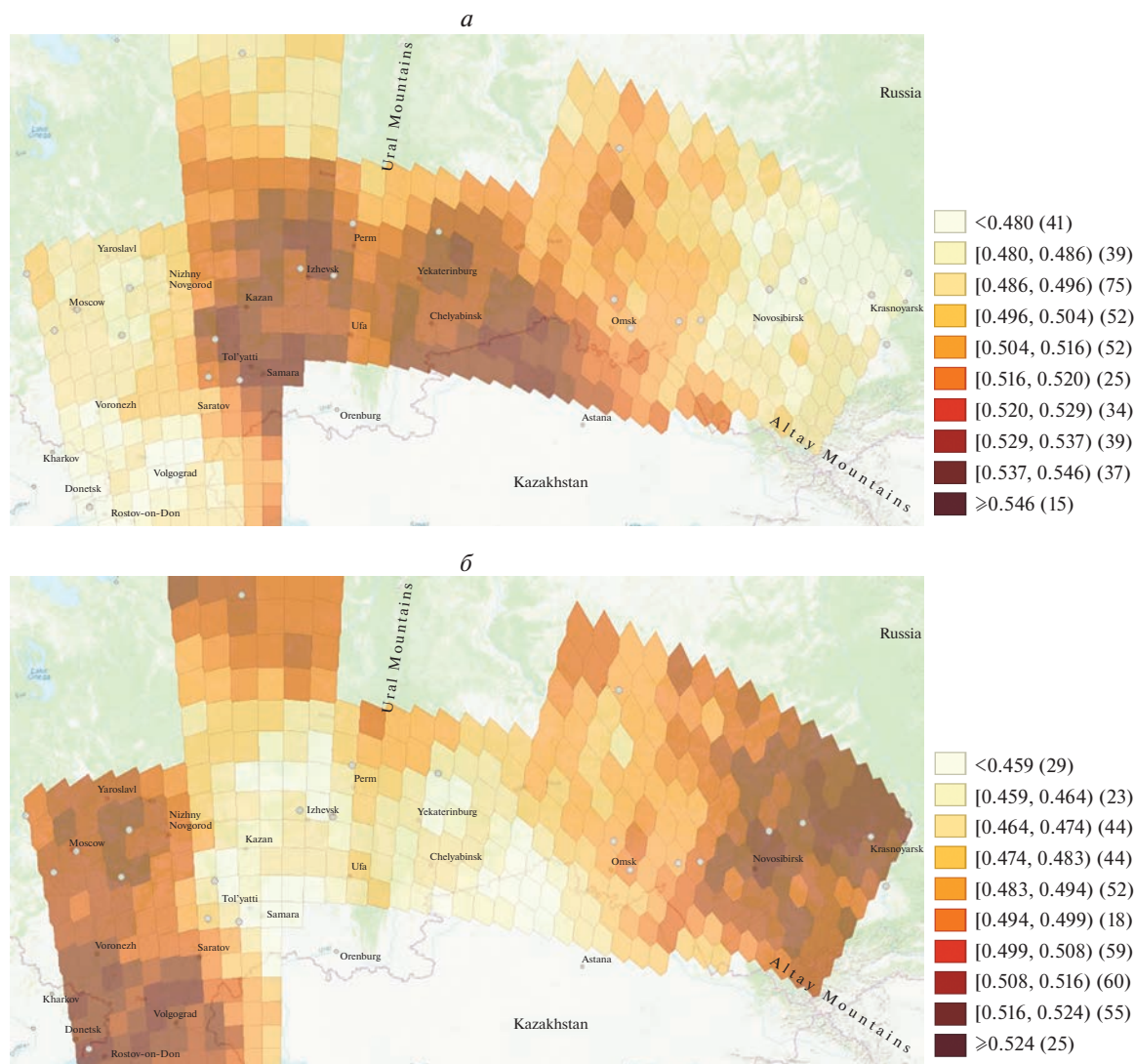


Рис. 5. Результаты пространственной кластеризации изученных особей, полученные при расчетах в Geneland, перенесенные на топографическую карту (*a, б*).

пролетных пути [39], но гораздо более плотно населяющих территорию, чем в данном исследовании.

Вместе с тем, несмотря на однородность генофонда и очень низкую внутривидовую дифференциацию, статистическое допущение величиной $\pm 5\%$ о принадлежности анализируемых птиц к двум генетическим популяциям дало интересный результат, согласно которому выборки не разделились по Уралу, а сформировали кластер “чистых” подвидов, куда вошли и *G. g. grus* и *G. g. lilfordi*, и кластер птиц из “переходной” зоны с обеих сторон от Уральских гор. Возможно едва уловимые генетические отличия птиц в этой зоне могут формироваться за счет мутационных событий в зоне интрградации двух подвидов серого журавля.

ЗАКЛЮЧЕНИЕ

Наше исследование пространственной популяционно-генетической структуры серого журавля, на настоящий момент самое широкое по охвату ареала, показало невыраженную пространственную дифференциацию у этого вида, взрослые особи которого характеризуются гнездовым консерватизмом (филопатрией). Возможно на российской части ареала перемешивающий эффект потока генов обеспечивается нательной миграцией молодых особей и отсутствием естественных преград как между выделяемыми на основании морфологических признаков подвидами, так и в целом между гнездовыми группировками. Таким образом, на фоне высокой внутрипопуляционной изменчи-

ности, что характерно не только для серого журавля в России [13], но и для популяций данного вида в Европе [39] и других мигрирующих видов журавлей [40–42], пространственная структура оказалась не выраженной. Подобная картина характерна для многих видов птиц [43], в том числе с практически непрерывным ареалом. Она наблюдается в случаях активного потока генов и в отсутствии трофической (эврифагия в пределах характерного спектра кормов на всем ареале) и других видов экологической специализации. Таксономическая значимость наблюдаемых различий в морфологических признаках западных и восточных журавлей очевидно, не имеет глубокой генетической основы и не выявляется с помощью анализа микросателлитных локусов, так же, как и анализа контрольного региона мтДНК [44]. Более низкое аллельное разнообразие и отсутствие уникальных аллелей у *G. g. lilfordi* позволяет выдвинуть предположение об относительной эволюционной молодости и недавнем происхождении восточного подвида серого журавля от типового западного подвида — вероятно, от одной из его восточных маргинальных популяций. Обеднение аллельного разнообразия в этом случае обусловлено генетическим дрейфом по механизму “эффекта основателя”. Потеря части аллелей вследствие прохождения стадии резкого падения численности также может рассматриваться в данном контексте, но в этом случае отсутствие приватных аллелей объяснить сложнее. Анализ генетической дифференциации на ареале серого журавля необходимо продолжать на расширенных по численности и охвату территории выборках (что в настоящее время осложняется проблемой неинвазивного получения образцов от большого числа особей) с помощью различных классов генетических маркеров и современных геномных подходов.

Авторы благодарят: И.П. Арюлину, О.С. Гринченко, М.В. Корепова, Т.В. Свиридову, Т.В. Селезневу, Г.В. Носаченко, сотрудников Федерального природного заказника “Кирзинский” (Новосибирская обл.) за помощь в сборе биологического материала. Также выражают большую благодарность: Институту биологии им. Ал. Зильбермана Еврейского университета (Иерусалим, Израиль), Институту поведения животных М. Планка (Германия), Исследовательскому центру по экологии и окружающей среде Китайской Академии наук, Университету Китайской Академии наук (Китай) за обеспечение передатчиками GPS-GSM, благодаря которым выявлены места гнездования меченых серых журавлей.

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

Отлов птиц осуществлялся по разрешениям региональных органов Росприроднадзора.

Исследование одобрено Локальным комитетом по биоэтике при Институте общей генетики им. Н.И. Вавилова РАН (протокол № 1 от 18.05.2023 г.).

Авторы заявляют, что у них нет конфликта интересов.

Исследование выполнено за счет гранта Российского научного фонда № 23-24-00613, <https://rscf.ru/project/23-24-00613/>.

СПИСОК ЛИТЕРАТУРЫ

1. Prange H., Ilyashenko E.I. Eurasian crane // Crane Conservation Strategy. Baraboo, Wisconsin, USA: Int. Crane Foundation, 2019. P. 397–423.
2. Bird Life International. 2016. *Grus grus*. The IUCN Red List of Threatened Species 2016. <https://doi.org/10.2305/IUCN.UK.2016-3.RLTS.T2-2692146A86219168.en>
3. Ильяшенко В.Ю. О систематике серого журавля // Журавли Евразии (биология, распространение, миграция, управление). 2011. Вып. 4. С. 93–103.
4. Ильяшенко В.Ю., Белялов О.В. Новый подвид серого журавля *Grus grus korelovi* ssp. n. (Aves: Gruidae) из Центрального и Восточного Тянь-Шаня // Русс. орнитол. журн. 2011. Т. 20. № 687. С. 1803–1811.
5. Ильяшенко В.Ю., Касабян М.Г., Маркин Ю.М. Морфологическая изменчивость серого журавля — *Grus grus* (Linnaeus, 1758) (Aves: Gruidae) // Журавли Евразии (биология, распространение, миграции). 2008. Вып. 3. С. 50–82.
6. Crane conservation strategy / Eds Mirande C.M., Harris J.T. Baraboo, Wisconsin, USA: Int. Crane Foundation, 2019. 454 p.
7. Ilyashenko E., Markin Y. Changing of the Eurasian crane staging areas distribution in the European part of Russia from 1982 to 2007 // Proc. of the Cranes, Agriculture, and Climate Change Workshop (Muraviovka Park, Russia, 28 May–3 June, 2010). 2012. P. 88–99.
8. Пранге Х. Распространение и миграции серого журавля на западноевропейском пролетном пути // Журавли Евразии (биология, распространение, разведение). М.—Н. Пасучей: Изд-во “Белый ветер”. 2015. Вып. 5. С. 287–312.
9. Ильяшенко Е.И. Оценка численности журавлей (Gruiformes, Gruidae) Северной Евразии в начале 21 века // Зоол. журнал. 2016. Т. 95. № 8. Р. 976–980. <https://doi.org/10.7868/S0044513416080043>
10. Hasegawa O., Ishibashi Y., Abe S. Isolation and characterization of microsatellite loci in the red-crowned crane *Grus japonensis* // Mol. Ecol. 2000. V. 9. № 10. P. 1677–1678.
11. Jones K.L., Henkel J.R., Howard J.J. et al. Isolation and characterization of 14 polymorphic microsatellite DNA loci for the endangered whooping crane (*Grus americana*) and their applicability to other crane species // Conserv. Gen. Res. 2010. V. 2. № 1. P. 251–254. <https://doi.org/10.1007/s12686-010-9196-3>
12. Meares K., Dawson D., Horsburgh G. et al. Characterisation of 14 blue crane *Grus paradisea* (Gruidae, AVES) microsatellite loci for use in detecting illegal trade // Conserv. Genet. 2008. V. 9. P. 1363–1367. <https://doi.org/10.1007/s10592-007-9490-0>

13. Мудрик Е.А., Каиенцева Т.А., Редчук П.С., Политов Д.В. Данные по микросателлитной изменчивости подтверждают низкую генетическую дифференциацию западного и восточного подвидов серого журавля (*Grus grus* L.) // Мол. биология. 2015. Т. 49. № 2. С. 297–304.
<https://doi.org/10.7868/S002689841502010X>
14. Ильяшенко Е.И., Ильяшенко В.Ю., Викельски М., Цао Л. Предварительные результаты слежения за серыми журавлями, помеченными в европейской части России и Западной Сибири в 2019–2021 гг. // Инф. бюлл. РГЖЕ. 2022. № 16. С. 157–169.
15. Zhang L., Zhang Z., Shen F. et al. Identification and characterization of polymorphic microsatellite loci in the red-crowned crane // Genet. Mol. Res. 2015. V. 14. № 4. P. 15169–15176.
<https://doi.org/10.4238/2015.November.25.5>
16. Krajewski C., Sipiorski J.T., Anderson F.E. Complete mitochondrial genome sequences and the phylogeny of cranes (Gruiformes: Gruidae) // Auk. 2010. V. 127. № 2. P. 440–452.
<https://doi.org/10.1525/auk.2009.09045>
17. Lazar I., Jr., Lazar I., Sr. GelAnalyzer 19.1. www.gelanalyzer.com
18. Van Oosterhout C., Hutchinson W.F., Wills D.P., Shipley P. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data // Mol. Ecol. Notes. 2004. V. 4. № 3. P. 535–538.
<https://doi.org/10.1111/J.1471-8286.2004.00684.X>
19. Peakall R., Smouse P.E. GenA1Ex 6.5: Genetic analysis in Excel. Population genetic software for teaching and research – an update // Bioinformatics. 2012. № 28. P. 2537–2539.
<https://doi.org/10.1093/bioinformatics/bts460>
20. Pritchard J.K., Stephens M., Donnelly P. Inference of population structure using multilocus genotype data // Genetics. 2000. V. 155. № 2. P. 945–959.
<https://doi.org/10.3410/f.1015548.197423>
21. Guillot G., Estoup A., Mortier F., Cosson J.F. A spatial statistical model for landscape genetics // Genetics. 2005. V. 170. № 3. P. 1261–1280.
<http://doi.org/doi:10.1534/genetics.104.033803>
22. Guillot G., Mortier F., Estoup A. Geneland: A program for landscape genetics // Mol. Ecol. Notes. 2005. V. 5. № 3. P. 712–715.
<http://doi.org/doi:10.1111/j.1471-8286.2005.01031.x>
23. Guillot G. Inference of structure in subdivided populations at low levels of genetic differentiation. The correlated allele frequencies model revisited // Bioinformatics. 2008. V. 24. P. 2222–2228.
<https://doi.org/10.1093/bioinformatics/btn419>
24. Guillot G., Renaud S., Ledevin R. et al. Unifying model for the analysis of phenotypic, genetic and geographic data // System. Biol. 2012. V. 61. № 6. P. 897–911.
<https://doi.org/10.1093/sysbio/sys038>
25. Guillot G., Santos F. A computer program to simulate multilocus genotype data with spatially auto-correlated allele frequencies // Mol. Ecol. Res. 2009. V. 9. № 4. P. 1112–1120.
26. Guillot G., Santos F. Using AFLP markers and the Geneland program for the inference of population genetic structure // Mol. Ecol. Res. 2010. V. 10. № 6. P. 1082–1084.
<https://doi.org/10.1111/j.1755-0998.2010.02864.x>
27. Guillot G., Santos F., Estoup A. Analysing georeferenced population genetics data with Geneland: A new algorithm to deal with null alleles and a friendly graphical user interface // Bioinformatics. 2008. V. 24. № 11. P. 1406–1414.
<https://doi.org/10.1093/bioinformatics/btn136>
28. Guillot G., Santos F., Estoup A. Population genetics analysis using R and the Geneland program. Lyngby, Denmark: Technical University of Denmark, 2011.
29. Evanno G., Regnaut S., Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study // Mol. Ecol. 2005. V. 14. № 8. P. 2611–2620.
<https://doi.org/10.1111/j.1365-294X.2005.02553.x>
30. Kopelman N.M., Mayzel J., Jakobsson M. et al. Clumpak: A program for identifying clustering modes and packaging population structure inferences across K // Mol. Ecol. Res. 2015. V. 15. № 5. P. 1179–1191.
<https://doi.org/10.1111/1755-0998.12387>
31. Anselin L., Syabri I., Kho Y. GeoDa: An introduction to spatial data analysis // Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications. Berlin, Heidelberg: Springer, 2010. P. 73–89.
https://doi.org/10.1007/978-3-642-03647-7_5
32. Schnute J.T., Boers N., Haigh R. et al. PBS mapping: Mapping Fisheries Data and Spatial Analysis Tools // Fisheries and Oceans Canada. 2022.
<https://github.com/pbs-software/pbs-mapping>
33. R CoreTeam. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria: 2021. <https://www.R-project.org>
34. Мевес В. Постоянство использования гнездовых участков журавлями в округе Мекленбург-Западная Померания (Северо-Восточная Германия) // Журавли Евразии: биология, охрана, управление. М., Цасучей: Изд-во “Белый ветер”, 2015. Вып. 5. С. 68–76.
35. Кондракова К.Д., Маркин Ю.М., Постельных К.А. и др. Перемещения неполовозрелых серых журавлей в центре европейской части России // Орнитология. 2021. Т. 45. С. 75–80.
36. Sviridova T.V., Grinchenko O.S., Wikelski M., Ilyashenko E.I. Geographical connectivity, migration routes, and wintering grounds of the common crane in the Northern Moscow Region // Arid Ecosystems. 2023. V. 13. № 2. P. 196–207.
<https://doi.org/10.1134/S2079096123020142>
37. Hayes M.A. Dispersal and Population Genetic Structure in Two Flyways of Sandhill Cranes (*Grus canadensis*). Madison, Wisconsin (US): The University of Wisconsin-Madison. 2015. 277 p.
38. Nesbitt S.A., Schwikert S.T., Folk M.J. Natal dispersal in Florida sandhill cranes // J. Wildl. Manage. 2002. P. 349–352.
39. Haase M., Hölte H., Blahy B. et al. Shallow genetic population structure in an expanding migratory bird with high breeding site fidelity, the western eurasian crane *Grus grus grus* // J. Ornithol. 2019. V. 160. P. 965–972.
<https://doi.org/10.1007/s10336-019-01688-1>

40. Mudrik E.A., Ilyashenko E.I., Goroshko O.A. et al. The demoiselle crane (*Anthropoides virgo*) population genetic structure in Russia // Vavilov J. Genetics and Breeding. 2018. V. 22. № 5. P. 586–592. <https://doi.org/10.18699/VJ18.398>
41. Мудрик Е.А., Горошко О.А., Сурмач С.Г. и др. Однородность генофонда западной и восточной популяций даурского журавля *Antigone vipio* на разных пролетных путях // Генетика. 2022. Т. 58. № 5. С. 570–580. <https://doi.org/10.31857/S001667582205006X>
42. Мудрик Е.А., Политов Д.В. Молекулярно-генетические подходы в изучении и сохранении популяционных генофондов журавлей (Gruidea, Aves) // Усп. соврем. биологии. 2022. Т. 142. № 5. P. 477–486. <https://doi.org/10.31857/S004213242205009X>
43. Parau L.G., Wink M. Common patterns in the molecular phylogeography of western palearctic birds: A comprehensive review // J. Ornithol. 2021. V. 162. P. 937–959. <https://doi.org/10.1007/s10336-021-01893-x>
44. Haase M., Ilyashenko V. A glimpse on mitochondrial differentiation among four currently recognized subspecies of the common crane *Grus grus* // Ardeola. 2012. V. 59. № 1. P. 131–136. <https://doi.org/10.13157/arla.59.1.2012.131>

Weak Population Genetic Structure of the Eurasian Crane *Grus grus* L.

E. A. Mudrik^{a, *}, Yu. M. Markin^b, K. A. Postelnykh^b, K. D. Kondrakova^{a, c}, T. A. Kashentseva^b,
A. V. Shatokhina^a, P. A. Kazimirov^a, V. Yu. Ilyashenko^c, E. I. Ilyashenko^{a, c}, and D. V. Politov^a

^aVavilov Institute of General Genetics Russian Academy of Sciences, Moscow, 119991 Russia

^bOka State Nature Biosphere Reserve, Brykin Bor, 391072 Russia

^cSevertsov Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow, 119011 Russia

*e-mail: mudrik@vigg.ru

The article presents an extended (as compared with our previous studies) analysis of the population genetic structure of a migratory wide-range bird species with a high abundance, the Eurasian Crane *Grus grus* L. Using seven highly polymorphic microsatellite loci, high and very similar values of genetic diversity parameters were obtained in the samples of both the western (*G. g. grus*) and eastern (*G. g. lilfordi*) subspecies. Coefficients of genetic differentiation between these subspecies ($F_{ST} = 0.008$, $G_{ST} = 0.002$) were also found to be low. According to AMOVA, 99% of genetic variation of *G. grus* is attributed to individual levels. Using the Bayesian clustering algorithm implemented in STRUCTURE software, no clear population-genetic structuring of the species was revealed. However, visualization of spatial patterns of genetic variability in Geneland showed the presence of a cluster of “pure” subspecies *G. g. grus* and *G. g. lilfordi* surrounding a cluster of individuals from the zone of intergradation of these subspecies. This result, along with the observed very low F_{ST} values, may indicate subtle genetic differences between cranes from the studied area, possibly having a mutational nature. The lower allelic richness and lack of private alleles in the eastern subspecies suggest that *G. g. lilfordi* is relatively evolutionary young and suggest its recent origin from some eastern marginal populations of the nominative western subspecies.

Keywords: Gruidae, genetic variability, genetic differentiation, microsatellite loci, Russia.

МЕТИЛИРОВАНИЕ ПОДСЕМЕЙСТВ РЕТРОТРАНСПОЗОНА LINE-1 В ВОРСИНАХ ХОРИОНА ПРИ НЕВЫНАШИВАНИИ БЕРЕМЕННОСТИ

© 2023 г. С. А. Васильев^{1, *}, В. В. Деменева¹, Е. Н. Толмачева¹, С. А. Филатова^{1, 2},
А. С. Зуев¹, А. С. Ушакова^{1, 2}, О. Ю. Васильева¹, И. В. Лушников²,
Е. А. Саженова¹, Т. В. Никитина¹, И. Н. Лебедев¹

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

²Национальный исследовательский Томский государственный университет, Томск, 634050 Россия

*e-mail: stanislav.vasilyev@medgenetics.ru

Поступила в редакцию 25.06.2023 г.

После доработки 21.07.2023 г.

Принята к публикации 31.07.2023 г.

Невынашивание беременности потенциально связано с нарушениями эпигенетической регуляции генов, ответственных за развитие эмбриона и плаценты. Цель настоящей работы — анализ уровня метилирования различных подсемейств ретротранспозона LINE-1, составляющего около 17% всего генома, в ворсинах хориона спонтанных абортусов первого триместра беременности с различным кариотипом, включая наиболее распространенные анеуплоидии. Был проведен анализ профиля метилирования в промоторе ретротранспозона LINE-1 с помощью таргетного бисульфитного массового параллельного секвенирования в ворсинах хориона медицинских абортусов ($n = 39$), спонтанных абортусов с нормальным кариотипом ($n = 173$), трисомией 16 ($n = 62$) и моносомией X ($n = 46$), а также в лимфоцитах периферической крови здоровых добровольцев ($n = 17$). Уровень метилирования подсемейств ретротранспозона LINE-1 в контрольных группах лимфоцитов взрослых индивидов и в ворсинах хориона медицинских абортусов был наибольшим для эволюционно молодых подсемейств L1HS, меньшим — для более древних подсемейств L1PA2, L1PA3 и минимальным для еще более древнего подсемейства L1PA4. В группах спонтанных абортусов наблюдался повышенный уровень метилирования LINE-1, причем этот эффект был более выражен для более древних подсемейств LINE-1. Выявленные закономерности указывают на меньший контроль за более древними подсемействами ретротранспозона LINE-1 в геноме человека, которые потенциально могут использоваться в качестве регуляторных элементов для расположенных рядом генов, участвующих в эмбриональном развитии. Повышение уровня метилирования таких последовательностей может нарушать развитие плаценты и эмбриона и вносить определенный вклад в невынашивание беременности.

Ключевые слова: ретротранспозон LINE-1, метилирование ДНК, ворсины хориона, невынашивание беременности, анеуплоидия, бисульфитное секвенирование, спонтанные абортусы.

DOI: 10.31857/S0016675823120147, **EDN:** PQAIPV

Невынашивание беременности является распространенным явлением, затрагивающим около 15% всех беременностей [1]. Существует много известных причин выкидыша, главной из которых является анеуплоидия. Тем не менее, почти 50% случаев потери беременности на ранних сроках остаются необъясненными [2].

Плацента выполняет ведущую роль в раннем эмбриональном развитии, обеспечивая питание зародыша и взаимодействие между ним и матерью. К окончанию первого триместра беременности критичным для дальнейшего развития эмбриона становится нормальное протекание процесса ремоделирования спиральных артерий, так как до девятой недели беременности пробки из клеток вневорсинчатого трофобласта ограничивают поступление материнской крови в пространство

между ворсинками хориона. Закупорка помогает поддерживать состояние физиологической гипоксии на ранних стадиях процесса плацентации, способствуя пролиферации цитотрофобласта, а не дифференцировке и инвазии [3]. После распада этих пробок примерно на девятой неделе беременности маточно-плацентарные спиральные артерии начинают снабжать материнской кровью межворсинчатое пространство [4]. Эмбрионы, которые не могут обеспечить нормальное протекание этого процесса, по-видимому погибают из-за гипоксии и недостатка питательных веществ.

Фундаментальную роль в контроле развития плаценты играют эпигенетические механизмы регуляции экспрессии генов [5]. Плацента обладает специфическим метиломом, значительно отличающим ее от эмбриональных тканей и тканей

взрослого организма [6]. Ключевым его свойством является низкий уровень метилирования различных повторяющихся последовательностей, включая известные ретровирусы и ретротранспозоны [7]. Известно, что гены, имеющие ретровирусное происхождение, играют важную роль в формировании и функционировании плаценты [8]. Ошибки регуляции таких последовательностей, в основном осуществляемой за счет метилирования ДНК, может потенциально объяснять нарушения функционирования плаценты на ранних сроках беременности и приводить к невынашиванию беременности.

Кроме того, важную роль на ранних стадиях развития эмбриона и плаценты играют различные мобильные генетические элементы, включая эндогенные ретровирусы человека и ретротранспозоны [9, 10]. Подсемейства LINE-1 различаются между собой отдельными заменами нуклеотидов, которые являются специфическими для всех членов отдельного подсемейства. В геномах гоминиоидных приматов распространились пять подсемейств ретротранспозона LINE-1 (L1P1, L1PA2, L1PA3, L1PA4, L1PA5) [11, 12]. Еще одно подсемейство (L1HS) является специфичным только для человека [13]. Ранее нами были выявлены нарушения уровня метилирования ретротранспозона LINE-1 в ворсинах хориона у спонтанных абортусов первого триместра беременности с анеуплоидией по различным хромосомам [14].

Цель настоящего исследования — анализ уровня метилирования различных подсемейств ретротранспозона LINE-1 в ворсинах хориона спонтанных абортусов первого триместра беременности с нормальным кариотипом и с наиболее частыми анеуплоидиями — трисомией 16 и моносомией X.

МАТЕРИАЛЫ И МЕТОДЫ

В качестве материала для исследования были отобраны образцы ворсин хориона медицинских абортусов (МА) ($n = 39$, гестационный возраст 8.3 ± 1.8 нед.), спонтанных абортусов (СА) с нормальным кариотипом ($n = 173$, гестационный возраст 7.6 ± 1.9 нед.), трисомией по хромосоме 16 ($n = 62$, гестационный возраст 6.9 ± 1.9 нед.) и моносомией X ($n = 46$, гестационный возраст 8.8 ± 1.2 нед.). Образцы, взятые из биобанка "Биобанк населения Северной Евразии" НИИ медицинской генетики Томского НИМЦ, были получены в период с 1993 по 2022 гг. До начала исследования образцы хранили при температуре -80°C без размораживания. В качестве группы сравнения использовались лимфоциты периферической крови здоровых добровольцев ($n = 17$, возраст 30.8 ± 2.7 лет). Исследование проводилось с соблюдением этических норм в соответствии с Хельсинской декларацией Всемирной медицинской ассоциации. Для всех образцов из биобанка было получено информированное согласие родителей на использование биоматериала для био-

банкирования и проведения исследований. Также были получены информированные согласия от здоровых добровольцев. Исследование было одобрено комитетом по биомедицинской этике НИИ медицинской генетики Томского НИМЦ (09.11.2020/№ 7).

Для определения кариотипа был проведен стандартный цитогенетический анализ на прямых препаратах ворсин хориона и культурах фибробластов экстраэмбриональной мезодермы [15]. Результаты кариотипирования всех спонтанных абортусов с анеуплоидией были подтверждены методом флуоресцентной гибридизации *in situ* (FISH). Уровень мозаицизма по анеуплоидии оценивали с нижним порогом 10% и верхним порогом 90%. Для анализа моносомии по X-хромосоме были использованы центромеро-специфичные ДНК-зонды на хромосому X, а для анализа трисомии по 16-ой хромосоме были использованы субтеломерные ДНК-зонды (16q и 16p). Анализ проводился по описанной ранее методике [16]. Выделение геномной ДНК из ворсин хориона и лимфоцитов периферической крови проводили с помощью фенол-хлороформного метода.

Праймеры для таргетного обогащения части промотора ретротранспозона LINE-1 с помощью ПЦР были разработаны на основе референсной последовательности ретротранспозона LINE-1 подсемейства L1HS, взятой из базы данных GenBank (X58075.1). Анализ проводился в соответствии с разработанным ранее протоколом [17].

Таргетное бисульфитное массовое параллельное секвенирование было проведено на приборе MiSeq (Illumina, США) с использованием набора Nano Kit (2×250). Оценка качества прочтений была выполнена с помощью FastQC v0.11.8, после чего был проведен тримминг оставшихся последовательностей адаптеров и низкого качества прочтений с помощью Trim-Galore. Затем прочтения были картированы на бисульфит-конвертированные таргетные последовательности с помощью инструмента bwa-meth v0.2.2 (<https://github.com/brentp/bwa-meth>) с параметрами по умолчанию. Данные по метилированию в контексте CpG были извлечены из полученных файлов BAM с помощью инструмента MethylDackel (<https://github.com/dpryan79/MethylDackel>). Уровень метилирования анализировался только в CpG-сайтах с покрытием более 10.

Сайты посадки использованных праймеров присутствуют в последовательностях нескольких подсемейств LINE-1. Для определения степени совпадения был проведен анализ BLAST (<https://blast.ncbi.nlm.nih.gov/>) с использованием бисульфит-конвертированной последовательности участка промотора LINE-1, получаемого с помощью разработанных в работе праймеров, и бисульфит-конвертированных консенсусных последовательностей повторов в геноме человека, полученных из UCSC Repeat Browser [18] (рис. 1). Результаты анализа показали, что в совпадающих с L1HS после-

довательностях L1P2, L1PA6, LTR65, HERVL18 и MER110-int отсутствует последовательность прямого праймера, а в последовательностях L1PA7 и L1PA10 — последовательность обратного праймера. Поэтому для этих подсемейств с использованием разработанных праймеров не могло быть получено продукта ПЦР. В результате полное перекрытие последовательностей обнаружено только для следующих подсемейств LINE-1: L1HS, L1P1, L1PA2, L1PA3, L1PA4, L1PA5. Были выделены три участка с повышенной вариабельностью без CpG-сайтов, SNP в которых могут быть использованы для последовательной классификации прочтений, относящихся к разным подсемействам LINE-1 (рис. 1). Первый участок позволяет провести фильтрацию подсемейств L1HS и L1PA2, второй участок — подсемейства L1PA5, третий участок — подсемейств L1PA3, L1PA4, L1P1. Для каждого нераспределенного прочтения был проведен анализ BLAST по сравнению с консенсусными последовательностями подсемейств LINE-1, по результатам которого прочтение было отнесено к подсемейству с максимальной степенью совпадения. При проведении сравнений позиции цитозинов были маскированы.

Результаты представлялись в виде уровня метилирования, равного отношению числа цитозинов к суммарному числу цитозинов и тимининов в отдельном CpG-сайте. Кроме того был вычислен средний уровень метилирования вдоль всей области интереса. Статистический анализ был выполнен с помощью программного пакета Statistica 10.0 (StatSoft, США). Ранговый критерий Манна–Уитни был использован для сравнения уровня метилирования между группами образцов. Для определения границ выбросов в группе медицинских абортусов использовались формулы $Q_1 - 1.5IQR$ и $Q_3 + 1.5IQR$, где Q_1 и Q_3 — первый и третий квартиль соответственно, IQR — межквартильное расстояние. Отличия считались значимыми при $p < 0.05$.

РЕЗУЛЬТАТЫ

После фильтрации всех прочтений по подсемействам LINE-1 было обнаружено, что наиболее представленным подсемейством оказалось подсемейство L1PA3 (21%), а большая часть прочтений относилась к подсемействам L1HS, L1PA2 и L1PA3 (суммарно 55 и 54% в группах СА и МА соответственно) (рис. 2, а, б). На оставшиеся подсемейства приходилось в сумме менее 1% всех прочтений. В подсемейство L1PA5 попало менее 0.1% прочтений, в связи с чем в дальнейшем анализе оно не использовалось. Около 44% прочтений не удалось классифицировать как относящиеся к одному из подсемейств LINE-1 (рис. 2, а, б). Возможно, это связано с генетическими вариантами, возникшими в отдельных сайтах в регионах, по которым проводилась классификация подсемейств. Поэтому для нераспределенных прочтений была опреде-

лена степень близости их последовательности к консенсусным последовательностям различных подсемейств LINE-1. Было обнаружено, что большая часть всех нераспределенных последовательностей (77.6%) была наиболее близка к подсемейству L1HS (рис. 2, в). Таким образом, около 48% всех прочтений относилось к подсемейству L1HS.

Уровень метилирования в лимфоцитах периферической крови взрослых индивидов и ворсинках хориона медицинских и спонтанных абортусов отличался между разными подсемействами LINE-1, будучи самым высоким для эволюционно наиболее молодого подсемейства L1HS, специфичного для человека и сохранившего способности к ретротранспозиции копии LINE-1 (рис. 3, 4). Уровень метилирования был ниже для более эволюционно древних подсемейств L1PA2 и L1PA3, и минимальным — для самого древнего подсемейства L1PA4. Исключением стало эволюционно молодое подсемейство L1P1, для которого уровень метилирования был на уровне более древних подсемейств (рис. 3, 4). Выявленный тренд к снижению уровня метилирования различных подсемейств LINE-1 от эволюционно молодых к более древним (за исключением L1P1) в лимфоцитах периферической крови взрослых индивидов и ворсинках хориона соответствовал таковому в образцах опухолей различной локализации [19]. Это указывает на общую тенденцию в тканях разной локализации, включая опухолевые клетки.

Ранее нами был обнаружен повышенный уровень метилирования LINE-1 в спонтанных абортусах с анеуплоидным кариотипом [14]. При анализе отдельно в каждом подсемействе индекс метилирования LINE-1 значимо не отличался между медицинскими абортусами и объединенной выборкой всех спонтанных абортусов для более молодых подсемейств L1HS (МА — $51.1 \pm 3.1\%$; СА — $52.6 \pm 5.0\%$, $p = 0.12$) и L1PA2 (МА — $36.6 \pm 2.8\%$; СА — $37.9 \pm 5.4\%$, $p = 0.3$). Индекс метилирования более древних подсемейств LINE-1 был выше в объединенной выборке спонтанных абортусов по сравнению с медицинскими абортусами: L1PA3 (МА — $37.4 \pm 2.3\%$; СА — $39.4 \pm 5.0\%$, $p = 0.02$), L1PA4 (МА — $23.5 \pm 7.1\%$; СА — $30.8 \pm 9.9\%$, $p < 0.001$). Индекс метилирования в недифференцированных подсемействах был также выше у спонтанных абортусов (МА — $40.7 \pm 2.0\%$; СА — $43 \pm 4.5\%$, $p = 0.001$). Исключением стало эволюционно молодое подсемейство L1P1 (МА — $20.8 \pm 8.4\%$; СА — $23.5 \pm 8.0\%$, $p = 0.004$), для которого уровень метилирования был также повышен в группе спонтанных абортусов.

При сравнении уровня метилирования LINE-1 каждой отдельной подгруппы спонтанных абортусов с медицинскими абортусами результаты различались. Спонтанные абортусы с трисомией 16 имели значимо более высокий средний уровень метилирования LINE-1 и уровни метилирования всех подсемейств LINE-1 по сравнению с меди-

X58075	1	TATTAGGGAGTG-TTAGATAGTGGGCGTAGGTTATTGTGTGC-GCGTATCGTGCGGAGTCG	60
L1HS	113	-.....G.....T-.....	172
L1PA2	112	-.....G..G...-.....	171
L1PA3	112	-.....A..G..G...TA.....	172
L1PA4	113	...G.....-.....T...A..G..G...TA.....T.....	173
L1P1	110	...G.....-.....T...A..G..G...TA.....A...T.....	170
L1PA5	108	...G...TT...-CG.....T...A..G..G...TA.TT...A.T.T.....	168
L1P2	147		152
L1PA6	149	T.....	154
L1PA10	105	-...T..G.....T...T...-----..G.----.G.N...A.	151
L1PA7	110	-...T..G.....T...T...-----..G.----.G.....	156
X58075	61	AAGTAGGGCGAGGTATTGTTTTATTGGGAAGCGTAAGGGGTTA-GGGAGnnnnnnnnnnCGA	121
L1HS	173	-.....	234
L1PA2	172	T.....-.....	233
L1PA3	173	C.....-.....TT.	234
L1PA4	174	C.....-.....A.....TT.	235
L1P1	171	C.....C.....TA.....-.....A.....TT.	232
L1PA5	169	C.....C.....N.....-.....A.....TT.	230
L1P2	153	G...C.....C.....T.....CG-..GA.....TT.	214
L1PA6	155	G...C.....C.....G...T.....TT.	216
L1PA10	152	T.G.....C.....N.....-.....A.....TT.	213
L1PA7	157	T.G..CG.C.....C.....CG-...A.....TT.	218
X58075	122	GTTAAAGAAAGGGGTGACGGA-C-GTATTTGGAATCGGGTTATTTTATTTCG--AATATT	179
L1HS	235	-.....-.....	292
L1PA2	234	-G.....-.....	291
L1PA3	235	TA..-G.....TT--.....	292
L1PA4	236	TA..TG-.....TT--.....	293
L1P1	233	TA..TG-.....TT--.....	290
L1PA5	231	G.G...C.....TA..-G.....TT--.....	288
L1P2	215	G.G...TC...TA..-TT.....A.....TA--.....	272
L1PA6	217	G.G...TC...TA..TT.....AT...A.....CG..TA--.....	274
L1PA10	214	G.G...	223
L1PA7	219	G.G...TTN...G...TT-..G..AT..GG.A...TG....CGG..TA--G....	276
HERVL18	1915	TT.....-.....	1938
X58075	180	GCGTTTTTTTAGATCGGTTTAAGAAACGGCGTATTACGAGATTATATTTTATATTGGTTTAGAGG	244
L1HS	293	C..CG.....A.....CG.....CG...	356
L1PA2	292	C..CG.....A.....G.....CG.....CG...	356
L1PA3	293	C..CG.....A.....G.....CG.....CG...	357
L1PA4	294	TA..G.....T.....TA.....G.....CG.....CG...	358
L1P1	291	TA..G..T...GT...TA.....G.....CG.....CG...	355
L1PA5	289	TA..G..T...GT...TA.....G.....CGCG.....CG...	353
L1P2	273	TA..G..T...GT...T..TA.....G.....CG.G.....	331
L1PA6	275	TT..G..T...GT...TT..TAG....G.....T.T...CG.G.....T..	340
L1PA7	277	AT.....TT.....	294
LTR65	317	T...T..A..G..T..TG..T..A.....AT....	362
HERVL18	1939	-.....	1943
MER110	4329		4346

Рис. 1. Результаты анализа BLAST бисульфит-конвертированной последовательности участка промотора LINE-1, получаемого с помощью разработанных в работе праймеров, и бисульфит-конвертированных консенсусных последовательностей повторов в геноме человека, полученных из UCSC Repeat Browser. В верхней строке текста приведена референсная последовательность, взятая из базы данных GenBank (X58075.1), ниже — последовательности повторов в геноме человека, полученные из UCSC Repeat Browser. Точками отмечены сайты, совпадающие с референсной последовательностью, буквами — несовпадающие позиции. Желтым цветом выделены сайты гибридизации праймеров; зеленым цветом выделены вариабельные регионы, по присутствию которых проводилась фильтрация прочтений, специфичных для различных подсемейств LINE-1.

цинскими абортусами (рис. 4). Для спонтанных абортусов с моносомией X значимо более высокий уровень метилирования LINE-1 по сравнению с медицинскими абортусами отмечался для подсемейств L1PA4, L1P1, нераспределенных прочтений и среднего уровня метилирования LINE-1. Спонтанные абортусы с нормальным кариотипом име-

ли повышенный уровень метилирования лишь для подсемейств L1PA4, L1P1 и нераспределенных прочтений (рис. 4).

Во всех группах спонтанных абортусов был в два раза выше коэффициент вариации (CV) среднего уровня метилирования LINE-1 (CA НК — 11.6%, Моно X — 11.4%, Три 16 — 10.55%) по срав-

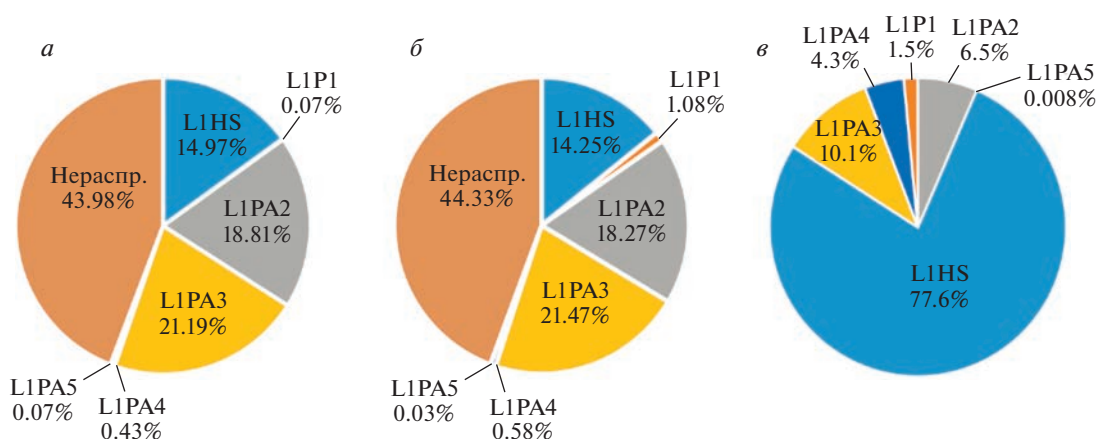


Рис. 2. Распределение прочтений между подсемействами LINE-1. *а* — в группе спонтанных абортусов (СА), *б* — в группе медицинских абортусов (МА), *в* — представленность подсемейств LINE-1 в группе нераспределенных прочтений.

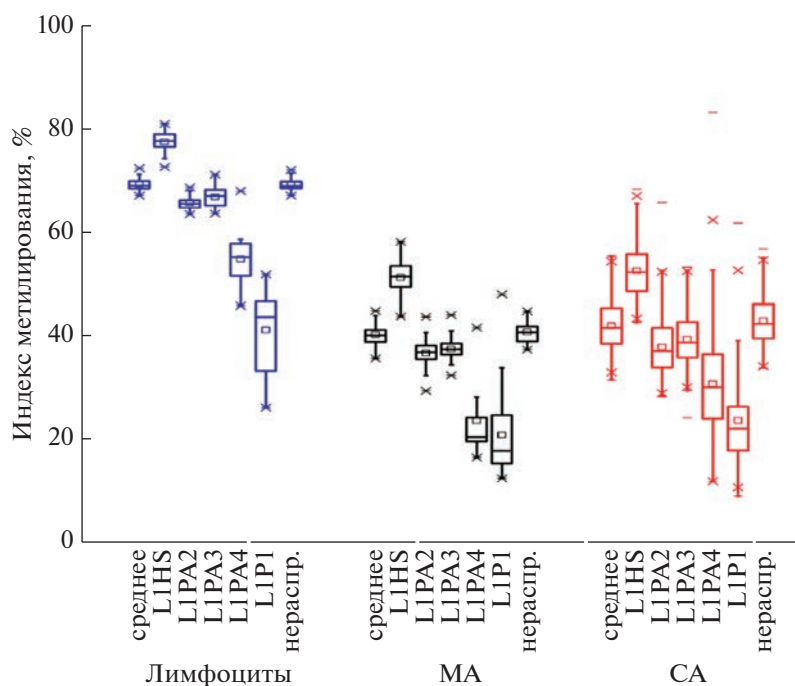


Рис. 3. Средний индекс метилирования ретротранспозона LINE-1 и индекс метилирования отдельных подсемейств LINE-1 в лимфоцитах периферической крови взрослых индивидов и в ворсинах хориона медицинских абортусов (МА), и объединенной группы спонтанных абортусов с различным кариотипом (СА). Бокс — 25–75% проценти, усы — границы выбросов; полосой в боксе отмечена медиана, квадратом — среднее значение, крестами отмечены значения 1 и 99%, горизонтальными чертами — минимальное и максимальное значения.

нению с медицинскими абортусами (5.5%) (рис. 4). В результате значительная часть спонтанных абортусов имела уровень метилирования LINE-1 выше границы выбросов в группе медицинских абортусов (табл. 1). Доля спонтанных абортусов с уровнем метилирования LINE-1, выходящим за границы нормальной вариации в группе медицинских абортусов, возрастала с увеличением эволюционного возраста подсемейства LINE-1 в ряду L1HS, L1PA2, L1PA3.

ОБСУЖДЕНИЕ

Анализ уровня метилирования различных подсемейств ретротранспозона LINE-1 подтвердил наши более ранние результаты [14] и показал, что в ворсинах хориона спонтанных абортусов повышен как средний уровень метилирования ретротранспозона LINE-1, так и уровень метилирования отдельных подсемейств ретротранспозона LINE-1. При этом для эволюционно более древних подсемейств повышение уровня метили-

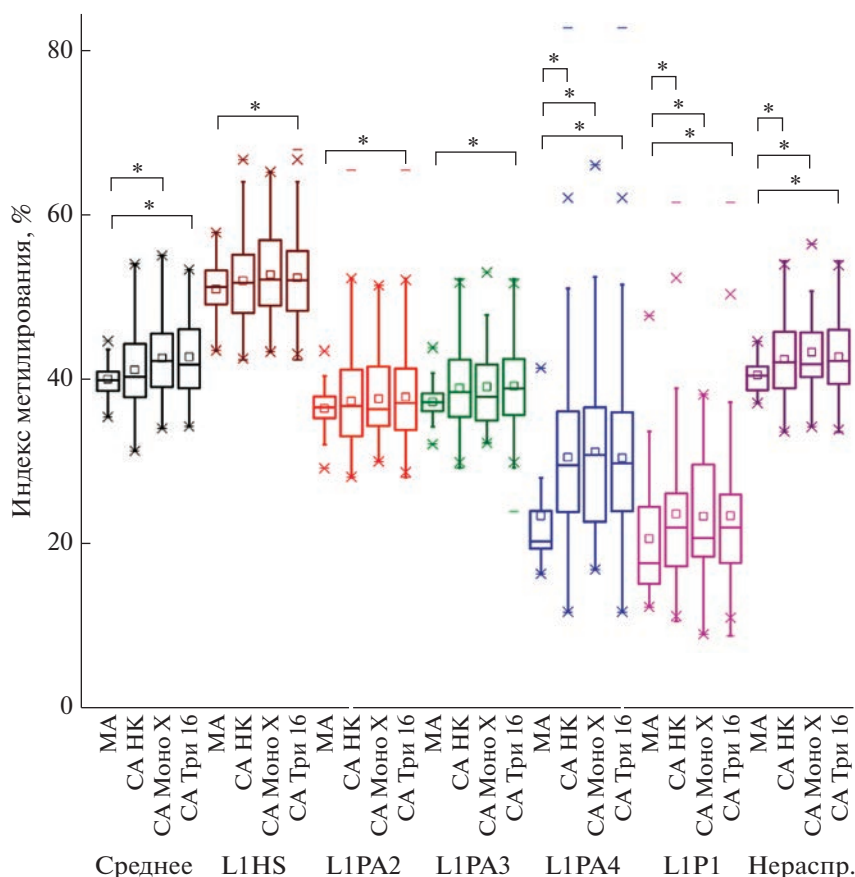


Рис. 4. Средний индекс метилирования ретротранспозона LINE-1 и индекс метилирования отдельных подсемейств LINE-1 в ворсинах хориона медицинских абортусов (МА), и спонтанных абортусов с нормальным кариотипом (СА НК), с моносомией X (СА Моно X), с трисомией 16 (СА Три 16). Бокс – 25–75% процентиля, усы – границы выбросов, полосой в боксе отмечена медиана, квадратом – среднее значение, крестами отмечены значения 1 и 99%, горизонтальными чертами – минимальное и максимальное значения, нераспр. – прочтения, нераспределенные на группы по результатам анализа; * – $p < 0.05$.

рования LINE-1 относительно уровня, характерного для контрольной группы медицинских абортусов, было более выраженным по сравнению с эволюционно молодыми семействами. Это может быть связано с тем, что наиболее эволюционно молодое подсемейство LINE-1 (L1HS) еще способно к ретротранспозиции и находится под более жестким контролем за счет метилирования ДНК в промоторе [19], что подтверждается и результатами настоящего исследования.

Роль ретротранспозона LINE-1 в плаценте не ясна. Несмотря на обнаруженный в настоящем исследовании повышенный уровень метилирования у сравнительно молодых подсемейств LINE-1 (L1HS, L1PA2 и L1PA3), они являются наиболее транскрипционно активными в плаценте [20]. Показано, что среди генов, способных к использованию промоторов LINE-1, обогащены гены, экспрессирующиеся в мозге и плаценте [21]. Кроме того, ретротранспозоны из подсемейства L1PA2 могут предоставлять свои бессмысловые промоторы в качестве альтернативных промоторов длинных некодирующих РНК в плаценте [22]. Возможно,

что неспособные к ретротранспозиции элементы LINE-1 постепенно начинают использоваться геномом в качестве отдельных функциональных элементов (в первую очередь, промоторов) для рядом расположенных генов [23]. В этом случае повышение уровня метилирования LINE-1 может приводить к нарушению экспрессии рядом расположенных генов, как было показано нами недавно [24], и нарушать развитие плаценты.

Повышенный уровень метилирования был сильнее выражен для спонтанных абортусов с анеупloidией (трисомия 16, моносомия X), однако также отмечался и у пятой части спонтанных абортусов с нормальным кариотипом. Это указывает на возможную роль нарушений метилирования генома как самостоятельного и независимого от наличия анеупloidии фактора, ассоциированного с гибелью эмбриона на ранних сроках беременности.

Исследование выполнено при поддержке гранта Российского научного фонда № 23-15-00341.

Таблица 1. Доля спонтанных абортусов с уровнем метилирования LINE-1, выходящим за границы нормальной вариации в группе медицинских абортусов

		Среднее	L1HS	L1PA2	L1PA3	Нераспр.	Всего
Границы индекса метилирования у МА (non-outlier range)		33.37–45.73%	39.82–61.32%	29.96–42.52%	32.88–42.73%	35.62–46.76%	
МА	гипер	0	0	1 (2.6%)	1 (2.6%)	0	39
	гипо	0	0	1 (2.6%)	1 (2.6%)	0	
СА НК	гипер	34 (19.7)	6 (3.5)	27 (15.6)	41 (23.7)	28 (16.2)	173
	гипо	4 (2.3)	0	7 (4.0)	19 (11.0)	9 (5.2)	
СА Моно X	гипер	11 (23.9)	5 (10.9)	10 (21.7)	10 (21.7)	11 (23.9)	46
	гипо	0	0	0	4 (8.7)	1 (2.2)	
СА Три 16	гипер	18 (29.0)	2 (3.2)	17 (27.4)	16 (25.8)	15 (24.2)	62
	гипо	0	0	1 (1.6)	2 (3.2)	0	
Итого		67 (20.9)	13 (4.1)	64 (20.0)	94 (29.4)	64 (20.0)	320

Примечание. Представлены количество и доля (%) абортусов с уровнем метилирования выше, чем верхняя граница (гипер) и ниже, чем нижняя граница (гипо) диапазона вариации в группе медицинских абортусов. МА – медицинские абортусы, СА НК – спонтанные абортусы с нормальным кариотипом, СА Моно X – спонтанные абортусы с моносомией X (СА Моно X), СА Три 16 – спонтанные абортусы с трисомией 16.

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

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. *Quenby S., Gallos I.D., Dhillon-Smith R.K. et al.* Miscarriage matters: the epidemiological, physical, psychological, and economic costs of early pregnancy loss // *Lancet*. 2021. V. 397. № 10285. P. 1658–1667. [https://doi.org/10.1016/S0140-6736\(21\)00682-6](https://doi.org/10.1016/S0140-6736(21)00682-6)
2. *Li T.C., Makris M., Tomsu M. et al.* Recurrent miscarriage: aetiology, management and prognosis // *Hum. Reprod. Update*. 2002. V. 8. № 5. P. 463–481. <https://doi.org/10.1093/humupd/8.5.463>
3. *Red-Horse K., Zhou Y., Genbacev O. et al.* Trophoblast differentiation during embryo implantation and formation of the maternal-fetal interface // *J. Clin. Invest.* 2004. V. 114. № 6. P. 744–754. <https://doi.org/10.1172/JCI22991>
4. *Jauniaux E., Poston L., Burton G.J.* Placental-related diseases of pregnancy: Involvement of oxidative stress and implications in human evolution // *Hum. Reprod.* Update. 2006. V. 12. № 6. P. 747–755. <https://doi.org/dml016>
5. *Shridhar V., Chu T., Simhan H. et al.* High-resolution analysis of the human placental DNA methylome in early gestation // *Prenat. Diagn.* 2020. V. 40. № 4. P. 481–491. <https://doi.org/10.1002/pd.5618>
6. *Robinson W.P., Price E.M.* The human placental methylome // *Cold Spring Harbor Perspect. Med.* 2015. V. 5. № 5. <https://doi.org/10.1101/cshperspect.a023044>
7. *Vlahos A., Mansell T., Saffery R. et al.* Human placental methylome in the interplay of adverse placental health, environmental exposure, and pregnancy outcome // *PLoS Genet.* 2019. V. 15. № 8. <https://doi.org/10.1371/journal.pgen.1008236>
8. *Etchegaray E., Naville M., Volff J.N. et al.* Transposable element-derived sequences in vertebrate development // *Mob. DNA*. 2021. V. 12. № 1. P. 1. <https://doi.org/10.1186/s13100-020-00229-5>
9. *Grow E.J., Flynn R.A., Chavez S.L. et al.* Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells // *Nature*. 2015. V. 522. № 7555. P. 221–225. <https://doi.org/10.1038/nature14308>
10. *Reiss D., Zhang Y., Mager D.L.* Widely variable endogenous retroviral methylation levels in human placenta // *Nucl. Ac. Res.* 2007. V. 35. № 14. P. 4743–4754. <https://doi.org/10.1093/nar/gkm455>
11. *Lee J., Cordaux R., Han K. et al.* Different evolutionary fates of recently integrated human and chimpanzee LINE-1 retrotransposons // *Gene*. 2007. V. 390. № 1–

2. P. 18–27.
<https://doi.org/10.1016/j.gene.2006.08.029>
12. Khan H., Smit A., Boissinot S. Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates // *Genome Res.* 2006. V. 16. № 1. P. 78–87.
<https://doi.org/10.1101/gr.4001406>
13. Boissinot S., Sookdeo A. The Evolution of LINE-1 in Vertebrates // *Genome Biol. Evol.* 2016. V. 8. № 12. P. 3485–3507.
<https://doi.org/10.1093/gbe/evw247>
14. Vasilyev S.A., Tolmacheva E.N., Vasilyeva O.Y. et al. LINE-1 retrotransposon methylation in chorionic villi of first trimester miscarriages with aneuploidy // *J. Assist. Reprod. Genet.* 2021. V. 38. № 1. P. 139–149.
<https://doi.org/10.1007/s10815-020-02003-1>
15. Lebedev I.N., Ostroverkhova N.V., Nikitina T.V. et al. Features of chromosomal abnormalities in spontaneous abortion cell culture failures detected by interphase FISH analysis // *Eur. J. Hum. Genet.* 2004. V. 12. № 7. P. 513–520.
<https://doi.org/10.1038/sj.ejhg.5201178>
16. Vasilyev S.A., Timoshevsky V.A., Lebedev I.N. Cytogenetic mechanisms of aneuploidy in somatic cells of chemonuclear industry professionals with incorporated plutonium-239 // *Russ. J. Genetics.* 2010. V. 46. № 11. P. 1381–1385.
<https://doi.org/10.1134/s1022795410110141>
17. Vasilyev S.A., Markov A.V., Vasilyeva O.Y. et al. Method of targeted bisulfite massive parallel sequencing of the human LINE-1 retrotransposon promoter // *MethodsX.* 2021. V. 8.
<https://doi.org/10.1016/j.mex.2021.101445>
18. Fernandes J.D., Zamudio-Hurtado A., Clawson H. et al. The UCSC repeat browser allows discovery and visualization of evolutionary conflict across repeat families // *Mob. DNA.* 2020. V. 11. P. 13.
<https://doi.org/10.1186/s13100-020-00208-w>
19. Zheng Y., Joyce B.T., Liu L. et al. Prediction of genome-wide DNA methylation in repetitive elements // *Nucl. Ac. Res.* 2017. V. 45. № 15. P. 8697–8711.
<https://doi.org/10.1093/nar/gkx587>
20. Zadora J., Singh M., Herse F. et al. Disturbed placental imprinting in preeclampsia leads to altered expression of DLX5, a human-specific early trophoblast marker // *Circulation.* 2017. V. 136. № 19. P. 1824–1839.
<https://doi.org/10.1161/CIRCULATIONAHA.117.028110>
21. Criscione S.W., Theodosakis N., Micevic G. et al. Genome-wide characterization of human L1 antisense promoter-driven transcripts // *BMC Genomics.* 2016. V. 17. P. 463.
<https://doi.org/10.1186/s12864-016-2800-5>
22. Chishima T., Iwakiri J., Hamada M. Identification of transposable elements contributing to tissue-specific expression of long non-coding RNAs // *Genes (Basel).* 2018. V. 9. № 1.
<https://doi.org/10.3390/genes9010023>
23. Pourrajab F., Hekmatimoghaddam S. Transposable elements, contributors in the evolution of organisms (from an arms race to a source of raw materials) // *Heliyon.* 2021. V. 7. № 1.
<https://doi.org/10.1016/j.heliyon.2021.e06029>
24. Demeneva V.V., Tolmacheva E.N., Nikitina T.V. et al. Expression of the NUP153 and YWHAB genes from their canonical promoters and alternative promoters of the LINE-1 retrotransposon in the placenta of the first trimester of pregnancy // *Vavil. Zh. Genet. Selektcii.* 2023. V. 27. № 1. P. 63–71.
<https://doi.org/10.18699/VJGB-23-09>

Methylation of the Retrotransposon LINE-1 Subfamilies in Chorionic Villi of Miscarriages

S. A. Vasilyev^{a,*}, V. V. Demeneva^a, E. N. Tolmacheva^a, S. A. Filatova^{a, b}, A. S. Zuev^a, A. S. Ushakova^{a, b},
O. Yu. Vasilyeva^a, I. V. Lushnikov^b, E. A. Sazhenova^a, T. V. Nikitina^a, and I. N. Lebedev^a

^aResearch Institute of Medical Genetics, Tomsk National Research Medical Center
of Russian Academy of Sciences, Tomsk, 634050 Russia

^bNational Research Tomsk State University, Tomsk, 634050 Russia

*e-mail: stanislav.vasilyev@medgenetics.ru

Miscarriage is potentially associated with abnormal epigenetic regulation of genes responsible for the development of the embryo and placenta. The aim of this work was to analyze the methylation level of various subfamilies of the LINE-1 retrotransposon, which makes up about 17% of the entire genome, in chorionic villi of spontaneous abortions of the first trimester of pregnancy with different karyotypes, including the most common aneuploidies. The methylation profile in the LINE-1 retrotransposon promoter was analyzed using targeted bisulfite massive parallel sequencing in chorionic villi of induced abortions ($n = 39$), spontaneous abortions with normal karyotype ($n = 173$), trisomy 16 ($n = 62$) and monosomy X ($n = 46$), and peripheral blood lymphocytes of healthy volunteers ($n = 17$). The level of methylation of the LINE-1 retrotransposon subfamilies in the control groups of adult lymphocytes and chorionic villi of induced abortions was the highest for evolutionarily young L1HS subfamilies, lower for the more ancient L1PA2 and L1PA3 subfamilies, and the lowest for the even more ancient L1PA4 subfamily. In the groups of spontaneous abortions, an increased level of LINE-1 methylation was observed, and this effect was more pronounced for the older LINE-1 subfamilies. The revealed patterns indicate less control over the older subfamilies of the LINE-1 retrotransposon in the human genome, which can potentially be used as regulatory elements for nearby genes involved in embryonic development. An increase in the level of methylation of such sequences can disrupt the development of the placenta and embryo and make a certain contribution to miscarriage.

Keywords: LINE-1 retrotransposon, DNA methylation, chorionic villi, miscarriage, aneuploidy, bisulfite sequencing, spontaneous abortions.

ТРАНСКРИПТОМНОЕ ИССЛЕДОВАНИЕ КЛЕТОК ПЛАЦЕНТЫ ПРИ ПРЕЭКЛАМПСИИ КАК ЭФФЕКТИВНЫЙ ИНСТРУМЕНТ ПЕРСОНАЛИЗИРОВАННОЙ МЕДИЦИНЫ

© 2023 г. Е. А. Трифонова¹, *, А. А. Бабовская¹, А. А. Зарубин¹, В. Н. Сереброва¹,
М. М. Гавриленко¹, М. Г. Сваровская¹, Е. В. Ижойкина², И. Г. Куценко², В. А. Степанов¹

¹Научно-исследовательский институт медицинской генетики Томского национального исследовательского
медицинского центра, Томск, 634050 Россия

²Сибирский государственный медицинский университет, Томск, 634050 Россия

*e-mail: ekaterina-trifonova@medgenetics.ru

Поступила в редакцию 29.06.2023 г.

После доработки 19.07.2023 г.

Принята к публикации 08.08.2023 г.

На сегодняшний день транскриптомика — одна из самых быстро развивающихся областей молекулярной биологии, позволяющая получить детальную информацию о функциональной активности генома как в норме, так и при патологии. В представленной работе мы использовали современные транскриптомные технологии для всесторонней характеристики полногеномного профиля экспрессии генов клеток синцитиотрофобласта (СТБ) плаценты человека при физиологической и осложненной преэклампсией (ПЭ) беременности. В результате проведенного нами анализа выявлены 26 генов, экспрессия которых статистически значимо различается в клетках СТБ женщин с ПЭ и физиологическим течением беременности. Кластер дифференциально-экспрессирующихся генов (ДЭГ) содержит не только известные гены-кандидаты, выявленные ранее во многих зарубежных полногеномных исследованиях плаценты (к примеру, *LEP*, *INHBA* и *FLT1*), но и новые гены (*AC098613.1*, *AC087857.1*, *FCRLB*, *TENM4*, *PTP4A1P7*, *LINC01225* и др.), которые могут рассматриваться в качестве новых биологических маркеров ПЭ и представляют интерес для дальнейшего изучения. Результаты функциональной аннотации ДЭГ показывают, что с развитием ПЭ на уровне СТБ могут быть связаны сигнальные пути регуляции гормональной секреции, MAPK-каскада, ERK1 и ERK2 каскада, положительной регуляции клеточной адгезии и пролиферации эндотелиальных клеток. Полученные результаты анализа альтернативного сплайсинга гена *FLT1* свидетельствуют о важной роли в патогенетике ПЭ данного механизма процессинга РНК за счет существенного увеличения транскрипционного разнообразия генов в клетках СТБ. Уровень экспрессии транскрипта, кодирующего белковую изоформу FLT-1 e15a, был значительно повышен у пациенток с ПЭ по сравнению с контрольной группой. Это исследование расширяет представление о задействованных в ПЭ молекулярных механизмах и может служить основой для разработки профилактических, прогностических и терапевтических стратегий в области персонифицированного акушерства.

Ключевые слова: полнотранскриптомное секвенирование, плацента, синцитиотрофобласт, ген, sFlt-1, альтернативный сплайсинг, MAPK-каскад.

DOI: 10.31857/S0016675823120135, **EDN:** QERSXD

Персонализированная направленность современной медицины требует инновационных подходов к изучению молекулярных механизмов заболеваний, а также разработки эффективных методов таргетной терапии, основанных на экспериментальных результатах, полученных с использованием омиксных технологий. В данном контексте значительный интерес для исследователей представляет анализ транскриптома различных тканей и клеток, поскольку нарушение экспрессионной активности генов вовлечено в патогенез многочисленных заболеваний. На сегодняшний день

транскриптомика — одна из самых быстро развивающихся областей молекулярной биологии, позволяющая получить детальную информацию о функциональной активности генома человека. Транскриптом отражает профиль экспрессии генов в данный момент времени и в отличие от генома обладает динамичностью, активно меняется в зависимости от внешних факторов, потребностей окружающих тканей и клеток, что делает его удобным инструментом для изучения механизмов реализации генетической информации в процессе адаптации клеток к изменяющимся условиям

среды, а также путей их взаимодействия в норме и при патологии. Ввиду этого транскриптомные технологии нашли широкое применение во всевозможных областях предиктивной медицины, к примеру в разработке методов диагностики, стратегий прогнозирования и профилактики различных заболеваний, включая и акушерскую патологию, где в качестве основного объекта изучения выступает плацента. Необходимо отметить, что изучение молекулярных процессов, происходящих в плацентарной ткани и связанных с транскрипционной регуляцией генной экспрессии, рассматривается на сегодняшний день как наиболее перспективное направление исследований для раскрытия патофизиологии осложненного течения беременности, в частности такого тяжелого заболевания как преэклампсия [1].

Молекулярные механизмы развития преэклампсии (ПЭ) являются предметом фундаментальных и прикладных исследований в области медицины и биологии более 40 лет. Однако, несмотря на некоторый прогресс в этой области, попытки сформировать единую теорию этиопатогенеза ПЭ до сих пор так и не увенчались успехом. За последнее десятилетие с использованием методов биохимии и молекулярной генетики удалось выявить отдельные предиктивные маркеры развития ПЭ (к примеру, PIGF и sFLT1), данные о прогностической значимости которых варьируют в зависимости от исследований [2, 3]. На сегодняшний день не дискутируется только один этиологический момент в развитии этого гестационного осложнения — наличие беременности и нарушения формирования спиральных артерий плаценты. Предположительно в возникновении ПЭ могут играть существенную роль разнообразные пусковые механизмы, которые обуславливают включение множества патогенетических звеньев в развитие этой тяжелой патологии, проявляющейся в конечном итоге полиорганной недостаточностью и вовлечением в патологический процесс различных тканей и типов клеток [4]. Совершенствование высокопроизводительных методов транскриптомики сделало возможным анализ полногеномного профиля мРНК в отдельных клетках человека, что позволяет выявить новые биомаркеры, функционально значимые пути сигнальной трансдукции, а также сети межмолекулярных и межклеточных взаимодействий, и тем самым расширить представления о молекулярных механизмах данного тяжелого акушерского заболевания, а также разработать новые персонализированные стратегии терапевтических и профилактических мероприятий.

Общепризнано, что основным структурным элементом и регулятором функций плаценты является синцитиотрофобласт (СТБ), выстилающий всю площадь поверхности ворсин плаценты. Благодаря непосредственному контакту с материнской кровью СТБ осуществляет многие функции пла-

центы, включая газообмен, транспорт питательных веществ, защиту плода и секрецию ряда гормонов и медиаторов, которые впоследствии попадают в материнский кровоток [5]. При этом любые нарушения функционирования СТБ, начиная с процессов инвазии, могут явиться причиной развития последующих осложнений беременности, включая и ПЭ. Показано, что медиаторы, выделяющиеся из плацентарной ткани в кровь матери, связаны с такими клиническими симптомами ПЭ как повышенное артериальное давление и содержание белка в моче. Более того, обнаружено, что при окислительном стрессе СТБ высвобождает в материнский кровоток ряд сложных веществ, включая провоспалительные цитокины, экзосомы и антиангиогенные факторы, которые могут нарушать функцию эндотелия матери, приводить к системному воспалительному ответу и формированию патологического течения беременности [5, 6].

Стоит отметить, что к настоящему времени проведены многочисленные зарубежные и единичные отечественные исследования, посвященные как характеристике профиля генной экспрессии при физиологической беременности, так и поиску паттернов транскриптома лейкоцитов крови и плаценты, связанных с различными гестационными осложнениями [1, 6–11]. Однако до сих пор неясно, какие изменения происходят в профилях экспрессии генов различных типов клеток плацентарной ткани в условиях нормального и патологического течения беременности. В представленной работе мы использовали современные транскриптомные технологии для всесторонней характеристики полногеномного профиля экспрессии генов клеток СТБ плаценты человека при физиологической и осложненной ПЭ беременности.

МАТЕРИАЛЫ И МЕТОДЫ

Проведение настоящего исследования одобрено Комитетом по биомедицинской этике НИИ медицинской генетики Томского НИМЦ. Материал для исследования представлен биоптатами материнской части плаценты женщин с физиологическим течением беременности ($N = 8$) и преэклампсией тяжелой степени ($N = 8$). Сбор биоптатов осуществлялся по стандартной методике [12] акушерами-гинекологами на базе ОГАУЗ “ОПЦ им. И.Д. Евтушенко” г. Томска. Полученный материал немедленно промывался физиологическим раствором и помещался в криопробирки, которые далее хранились в жидком азоте, чтобы свести к минимуму деградацию РНК. У всех пациентов срезы плацентарной ткани были охарактеризованы гистологически с окрашиванием гематоксилин-эозином. Лазерная микродиссекция с целью получения клеток СТБ осуществлена на оборудо-

вании PALM (Carl Zeiss, Германия) с технологией автоматизированного захвата фрагментов (Laser Capture Microdissection).

Для выделения тотальной РНК использован набор Single Cell RNA Purification Kit (Norgen, США). Концентрация и качество РНК были оценены с помощью Agilent 2100 Bioanalyzer. Приготовление библиотек проведено по протоколу SMARTer Stranded Total RNA-Seq Kit v2 (Takara, США). Массовое параллельное секвенирование выполнено на приборе Next-seq 2000 (Illumina). Выравнивание на референсный геном (hg38) было выполнено с помощью программы STAR. Аннотацию транскрипционных активных областей, полученных в настоящей работе в результате секвенирования клеток СТБ, проводили в программной среде R (<http://www.r-project.org>) с использованием базы данных gencode.v39. Фильтрация и нормализация данных осуществлялись в пакете edgeR среды для статистических вычислений R согласно стандартным параметрам. Статистический анализ дифференциальной экспрессии также проводили с помощью пакета edgeR. Порогом для фильтрации дифференциально экспрессированных РНК был выбран $|\log_2FC| > 1$ и $FDR < 0.05$ как оптимальный для минимизации ошибки второго рода, так как методы последующих статистического и интегративного анализов устойчивы к ошибкам первого рода [13].

Функциональный анализ кластера дифференциально-экспрессирующихся генов (ДЭГ) проводили с помощью веб-инструмента в ресурсе Molecular Signatures Database (MSigDB) программного обеспечения GSEA и базы данных GeneMANIA [14, 15]. Анализ событий альтернативного сплайсинга выполнен с помощью пакета “MAJIQ” [16], позволяющего прогнозировать и количественно оценивать события сплайсинга на основе данных полногеномного РНК-секвенирования, используя входные файлы BAM и GFF3. Восстановление последовательностей альтернативных изоформ проводили с помощью Illumina DRAGEN (Dynamic Read Analysis for GENomics) Bio-IT Platform. Экспериментальные исследования были выполнены на базе Центра коллективного пользования научно-исследовательским оборудованием “Медицинская геномика” НИИ медицинской генетики Томского НИМЦ.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Для установления молекулярных механизмов, которые могут лежать в основе формирования преэклампсии на уровне нарушения функционирования ворсин плацентарной ткани, был проведен транскриптомный анализ клеток СТБ пациенток с ПЭ и женщин с физиологической беременностью с помощью технологии RNA-seq.

В рамках данной работы было получено 16 транскриптомов клеток СТБ из образцов плацентарной ткани восьми пациенток с ПЭ и восьми женщин из контрольной группы. По меньшей мере 86% прочтений последовательности были однозначно выровнены на геном человека, равномерность отображения считывания по транскриптам всех образцов была хорошей. После выравнивания на геном человека среднее покрытие на образец составило 43 млн прочтений. Всего в анализ было включено 22625 транскриптов, для которых количество прочтений (CPM) составляло не менее пяти для 10% и более образцов. Аннотацию транскрипционных активных областей, полученных в настоящей работе в результате секвенирования клеток СТБ, проводили в программной среде R с использованием базы данных gencode.v39 (см. табл. 1, приводятся усредненные данные по всем анализируемым образцам). Большинство из анализируемых транскриптов (66.5%) соответствуют протеинкодирующим регионам генома человека (15046). В оставшемся кластере преобладают длинные некодирующие РНК и процессированные псевдогены.

Анализ дифференциальной экспрессии генов синцитиотрофобласта

Дифференциальную экспрессию генов СТБ определяли с помощью пакета edgeR, который вычисляет дисперсию показателя экспрессии для каждого гена. Дифференциально экспрессируемыми считали гены с различиями в абсолютном значении экспрессии в логарифмической шкале $|\log_2FC| > 2$ и уровнем значимости с поправкой на множественные сравнения менее 0.05 (рис. 1).

В результате проведенного нами анализа выявлены 26 генов, экспрессия которых статистически значимо различается ($FDR < 0.05$; $|\log_2FC| > 1$) в клетках СТБ женщин с ПЭ и физиологическим течением беременности (рис. 2). Причем только для двух генов — *IGHG1* и *TUBG2*, кодирующих константную область тяжелой цепи гамма-глобулина 1 и цепь гамма-тубулина соответственно, было обнаружено снижение транскрипционной активности в группе больных по сравнению с контрольной выборкой, в то время как гиперэкспрессия у пациенток с ПЭ была показана для 24 генов, что составляет более 92% от всех идентифицированных ДЭГ.

Кластер ДЭГ, экспрессия которых повышена при ПЭ в клетках СТБ, включает не только известные гены-кандидаты, выявленные ранее во многих полногеномных исследованиях профилей экспрессии генов плаценты при данной патологии (к примеру, *LEP*, *FLT1*, *INHBA*), но также новые потенциальные генетические маркеры (*PTP4A1P7*, *AC087857.1*, *SLC46A3*, *RPS18P9*, *STAMBPL1*,

Таблица 1. Аннотация транскриптов клеток СТБ, полученных в результате RNA-seq

Идентифицированные транскрипты (категории GENCODE)	Количество транскриптов
Белок-кодирующие (protein_coding)	15046
Длинные некодирующие РНК (lincRNA)	4034
Процессированные псевдогены (processed_pseudogene)	1618
Транскрибируемые непротессированные псевдогены (transcribed_unprocessed_pseudogene)	346
Транскрипты с неустановленной функцией (ТЕС)	356
Непротессированные псевдогены (unprocessed_pseudogene)	248
Неидентифицированная РНК (misc_RNA)	203
Транскрибируемые процессированные псевдогены (transcribed_processed_pseudogene)	172
Малые ядерные РНК (snRNA)	156
МикроРНК (miRNA)	109
Малые ядрышковые РНК (snoRNA)	153
Транскрибируемые унитарные псевдогены (transcribed_unitary_pseudogene)	71
Инактивированные гены рРНК (rRNA_pseudogene)	52
Унитарные псевдогены (unitary_pseudogene)	9
Митохондриальная тРНК (Mt_tRNA)	22
Гены варибельной цепи Т-клеточного рецептора (TR_V_gene)	12
Гены варибельной цепи иммуноглобулина (IG_C_gene)	7
рРНК (rRNA)	7
Инактивированный ген иммуноглобулина (IG_C/V_pseudogene)	1
Митохондриальная рРНК (Mt_rRNA)	1
Транслированные процессированные псевдогены (translated_processed_pseudogene)	1
Транслированные непротессированные псевдогены (translated_unprocessed_pseudogene)	1

SULT1C2, *AC073410.2*, *ASRGL1*, *TP53TG1* и др.), связь которых с развитием ПЭ установлена впервые в настоящей работе. Стоит отметить также большое количество некодирующих РНК в данном кластере ДЭГ, функциональную значимость которых в механизмах патологического течения беременности еще предстоит охарактеризовать. В табл. 2 представлены данные о наиболее значимых ДЭГ ($FDR < 0.01$).

С целью оценить потенциальную биологическую значимость выявленных ДЭГ в молекулярных механизмах ПЭ мы провели функциональную аннотацию и анализ сетевых взаимодействий идентифицированных генов с помощью вычислительного метода GSEA (Gene Set Enrichment Analysis) и базы данных GeneMANIA. В представленной работе при проведении функциональной аннотации в ресурсе Molecular Signatures Database (MSigDB) программного обеспечения GSEA учитывали категории, имеющие FDR (false discovery rate) < 0.05 (в данном случае FDR представляет собой вероятность неслучайного попадания группы генов в категорию с поправкой на множественные сравнения). Нами было выявлено восемь категорий, включающих девять генов из 26 изучен-

ных. Необходимо отметить, что большая часть из данных генов одновременно вовлечена в реализацию нескольких генных онтологий и молекулярных процессов (например, локусы *FGB*, *FGG*, *LEP*, *INHBA*, *TPBG* и *FLT1* включены минимум в три и более из обнаруженных категорий). Наиболее значимые биологические пути, процессы и молекулярные функции, обогащенные данными ДЭГ ($FDR < 0.005$), связаны с регуляцией гормональной секреции, регуляцией MAPK-каскада и положительной регуляцией процессов в многоклеточном организме (табл. 3).

Важно отметить, что среди ДЭГ сверхпредставлены гены, связанные с регуляцией MAPK-каскада. Предыдущие исследования показали, что окислительный стресс активирует передачу сигналов митоген-активируемой протеинкиназы (MAPK) и увеличивает экспрессию рецепторов апоптоза [17]. Активность p38 MAPK была значительно повышена в плаценте пациенток с ПЭ по сравнению с женщинами с нормальной беременностью. Предположительно активные формы кислорода (АФК) могут активировать сигнальный путь p38 MAPK в плацентарной ткани с последующей гиперэкспрессией sFlt-1 и sEng в материнской

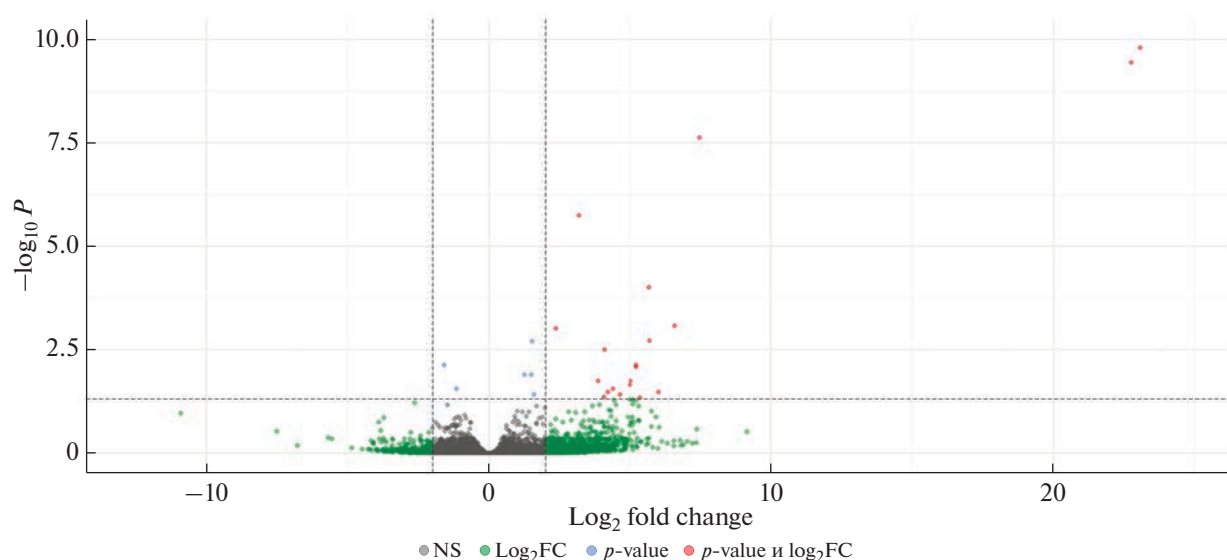


Рис. 1. Volcano-график, демонстрирующий зависимость статистической значимости (уровень FDR) от кратности различий в экспрессии генов между группами. Топовые дифференциально экспрессируемые при ПЭ и физиологической беременности гены ($|\log_2 FC| > 2$ и $FDR < 0.05$) отмечены красными точками.

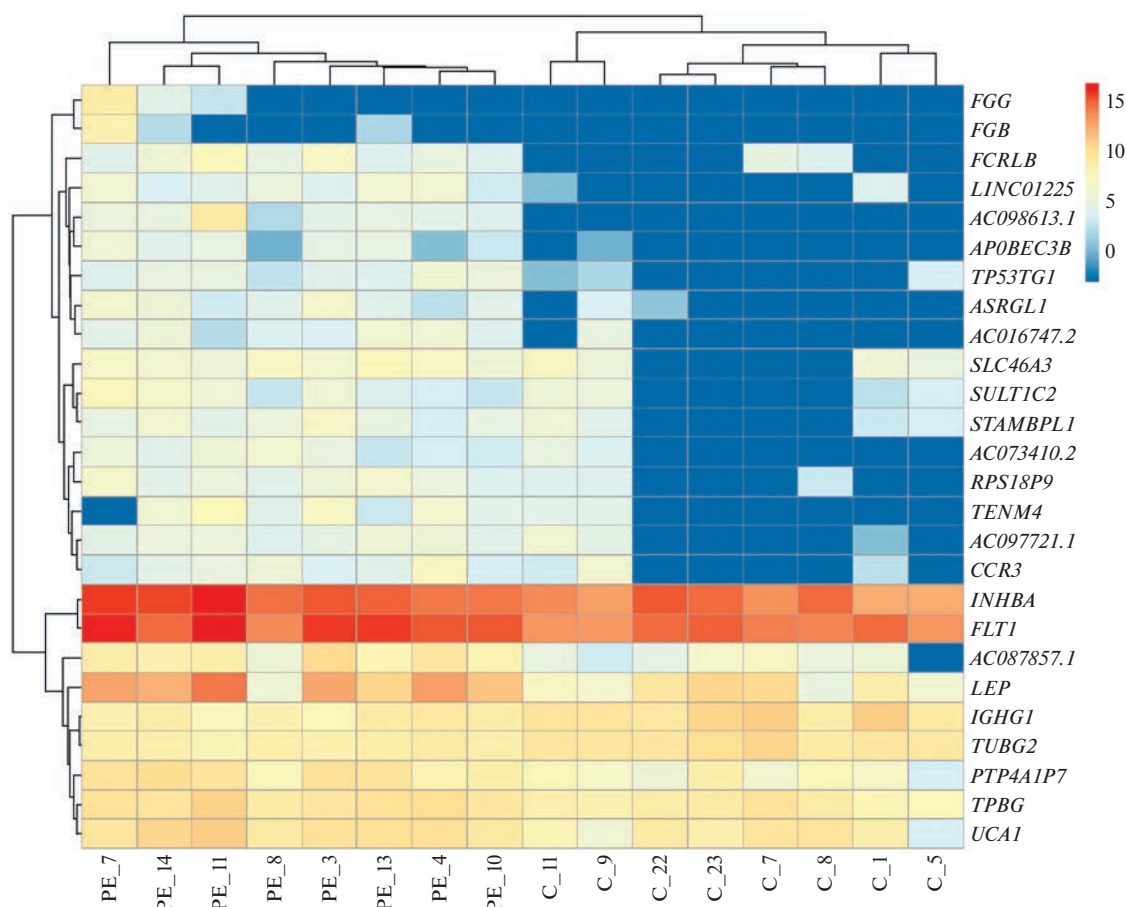


Рис. 2. Теплокорта ДЭГ, полученных при сравнительном анализе полнотранскриптомных профилей клеток СТБ в группах с физиологической беременностью (принадлежность к данной группе обозначена на рисунке С) и ПЭ (образцы в данном случае маркированы как PE).

Таблица 2. Гены, экспрессия которых наиболее существенно различается в клетках СТБ у женщин с ПЭ и физиологической беременностью

№	Ген	Локализация на хромосоме	Продукт гена	log ₂ FC	FDR
1	<i>FGG</i>	4q32.1	Гамма-цепь фибриногена	23.08	1.6×10^{-10}
2	<i>FGB</i>	4q31.3	Бета-цепь фибриногена	22.75	3.6×10^{-10}
3	<i>AC098613.1</i>	3p21.31	Некодирующая РНК	7.48	2.4×10^{-8}
4	<i>AC087857.1</i>	3p26.1	Некодирующая РНК	3.19	1.8×10^{-6}
5	<i>FCRLB</i>	1q23.3	Рецептор Fc, цепь B	5.68	9.8×10^{-5}
6	<i>TENM4</i>	11q14.1	Трансмембранный белок 4 тенейрин	6.58	0.0008
7	<i>PTP4A1P7</i>	1q25.2	РТР4A1 псевдоген 7	2.36	0.0009
8	<i>LINC01225</i>	1p35.2	Длинная некодирующая РНК	5.72	0.0019
9	<i>INHBA</i>	7p14.1	Ингибин Бета А субъединица	1.52	0.0019
10	<i>LEP</i>	7q32.1	Лептин	4.08	0.0031
11	<i>AC097721.1</i>	2q22.1	Некодирующая РНК	5.24	0.0073
12	<i>IGHG1</i>	14q32.33	Тяжелая цепь гамма-1 иммуноглобулина	1.58	0.0073
13	<i>TP53TG1</i>	7q21.12	Белок TP53 фрагмент 1	5.23	0.0081
14	<i>TPBG</i>	6q14.1	Гликопротеин трофобласта	1.26	0.0101
15	<i>FLT1</i>	13q12.3	Рецептор сосудистого эндотелиального фактора роста 1	1.50	0.0112

сыворотке и системной воспалительной реакцией, вызывая эндотелиальную дисфункцию и клинические симптомы ПЭ [18]. Показано, что инактивация сигнального пути Ras-МАРК в клетках трофобласта является одним из основных патогенетических факторов идиопатического привычного невынашивания беременности за счет ингибирования способности клеток трофобласта к пролиферации и инвазии [19].

Полученная с помощью web-ресурса GeneMANIA сеть взаимодействий продуктов ДЭГ подтверждает взаимосвязи между генами, выявленные при исследовании биологических путей и процессов. Анализ данной сети, включающей 15 протеинов, кодируемых изучаемыми генами, указывает на важное место в молекулярных механизмах ПЭ на уровне клеток СТБ белок-белковых взаимодействий продуктов ДЭГ и ко-экспрессии ряда генов (рис. 3). Функциональная аннотация кластера ДЭГ, связанных сетевыми взаимодействиями, указывает на высокостоверную сверхпредставленность данных генов ($FDR < 0.001$) в таких категориях как пролиферация эндотелиальных и эпителиальных клеток, регуляция пролиферации эндотелиальных клеток, положительная регуляция клеточной адгезии, регуляция связывания цитокиновых рецепторов, каскада ERK1/ERK2 и межклеточной адгезии, а также негативная регуляция процесса апоптоза.

Центральное место в данной сети с наибольшим числом взаимодействий занимают гены *FGB*, *FGG*, *LEP* и *FLT1*, кодирующие гамма и бета

цепи фибриногена, лептин и рецептор сосудистого эндотелиального фактора роста 1-го типа соответственно. Продукты некоторых из этих генов, согласно известным на сегодняшний день данным об их функциональных особенностях, могут играть ведущую роль в молекулярных механизмах исследуемой патологии.

Так, современные исследования демонстрируют, что нарушение баланса между про- и антиангиогенными факторами роста имеет большое значение в процессах инвазии цитотрофобласта при ПЭ [3, 20]. Показано, что растворимый антиангиогенный фактор (sFlt-1) – продукт гена *FLT1*, при физиологическом течении беременности уменьшает скорость инвазии цитотрофобласта, при этом концентрация его в начале беременности относительно низка и начинает повышаться в последнем триместре беременности. S.A. Kagimanchi с коллегами продемонстрирована значительная гиперэкспрессия в плацентарной ткани изоформы гена *FLT1*, кодирующей sFlt-1 [20]. Наряду с этим методами иммуногистохимии выявлены значительные изменения экспрессии VEGF, плацентарного фактора роста (PlGF) и антиангиогенного фактора (Flt-1) в плацентах женщин с ПЭ. У мышей с нокдауном в геноме данных локусов обнаруживается нарушение плацентарного васкулогенеза и увеличение внутриутробной смертности плодов [21]. Результаты, полученные в настоящей работе, также свидетельствуют о ключевой роли гиперэкспрессии гена *FLT1* в молекулярных механизмах ПЭ на уровне клеток СТБ, в связи с чем

Таблица 3. Наиболее значимые категории, обогащенные ДЭГ, полученные с помощью ресурса MSigDB

№	Категория в базе данных	Описание основных процессов категории	Гены, входящие в категорию	FDR
1	Позитивная регуляция MAPK-каскада (GO_POSITIVE_REGULATION_OF_MAPK_CASCADE)	Процессы, активирующие повышение частоты, скорости или степени передачи сигнала, опосредованной MAPK-каскадом	<i>FGB, FGG, LEP, INHBA, TPBG, FLT1</i>	0.008
2	Положительная регуляция секреции гормонов (GOBP_POSITIVE_REGULATION_OF_HORMONE_SECRETION)	Любые процессы, которые активируют или увеличивают частоту, скорость или степень секреции гормона из клетки	<i>FGB, FGG, LEP, INHBA</i>	0.009
3	Положительная регуляция процессов в многоклеточном организме (GOBP_POSITIVE_REGULATION_OF_MULTICELLULAR_ORGANISMAL_PROCESS)	Любые процессы, которые активируют или увеличивают частоту, скорость или длительность процесса в организме, любой из процессов, относящихся к функционированию организма выше клеточного уровня, включая интегрированные процессы тканей и органов	<i>FGB, FGG, LEP, INHBA, TPBG, FLT1, CCR3, TENM4</i>	0.009
4	Менометроррагия (HP_MENOMETRORRHAGIA)	Сочетание метроррагии и меноррагии	<i>FGB, FGG</i>	0.01
5	Эпидуральное кровоотечение (HP_EPIDURAL_HEMORRHAGE)	Эпидуральное кровоотечение	<i>FGB, FGG</i>	0.01
6	Регуляция вазоконстрикции (GOBP_REGULATION_OF_VASOCONSTRICTION)	Любой процесс, который модулирует частоту, темп или степень уменьшения диаметра кровеносных сосудов	<i>FGB, FGG, LEP</i>	0.01
7	Положительная регуляция секреции гонадотропина (GOBP_POSITIVE_REGULATION_OF_GONADOTROPIN_SECRETION)	Процессы, которые активируют или увеличивают степень регулируемого высвобождения гонадотропина	<i>LEP, INHBA</i>	0.01
8	Положительная регуляция каскада ERK1/ERK2 (GOBP_POSITIVE_REGULATION_OF_ERK1_AND_ERK2_CASCADE)	Любой процесс, который активирует или увеличивает скорость передачи сигнала, опосредованного каскадом ERK1/ERK2	<i>FGB, FGG, INHBA, TPBG</i>	0.01

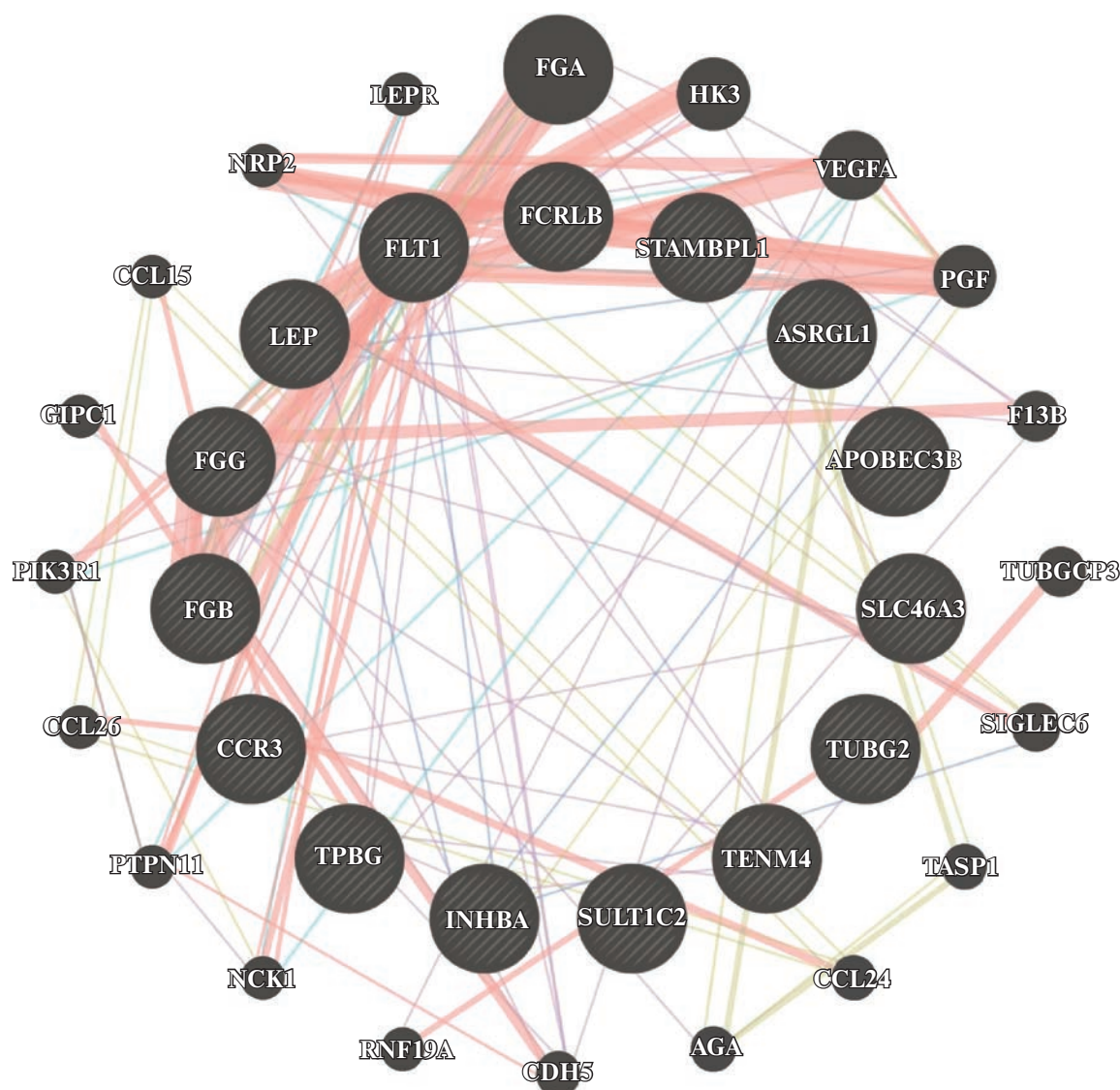


Рис. 3. Сеть белок-белковых взаимодействий продуктов ДЭГ. Сеть потенциальных взаимодействий создана с помощью базы данных GeneMANIA (<http://www.genemania.org>). Соответственно каждая цветная линия представляет определенный тип функциональной ассоциации, доступный в литературе, полученный из различных биологических систем и принятый алгоритмом GeneMANIA. Физическое взаимодействие – розовый цвет, ко-экспрессия – фиолетовый, общий белковый домен – желтый, ко-локализация – синий, общий патологический путь – светло-голубой. Чем толще линия, тем ближе вес ассоциации двух данных генов друг к другу.

нами был проведен детальный анализ событий альтернативного сплайсинга данного локуса.

Анализ альтернативного сплайсинга гена FLT1

Анализ событий альтернативного сплайсинга (АС) в клетках СТБ обследованных женщин проводили с помощью программы MAJIQ [16]. Данная программа идентифицирует комплексные (LSV (Local Splice Variation) события) и бинарные (семь “классических” типов событий) события АС; аннотированные и новые, используя данные RNA-seq. LSVs определяются и легко визуализи-

руются как расщепления (множественные ребра) в графе сплайсинга, где несколько ребер либо входят в один экзон, либо исходят из него. В частности, ранее определенные “классические” события АС отображаются как особые случаи расщепления двоичного графа (например, включение или пропуск экзона), в то время как LSVs фиксируют неклассические двоичные расщепления и расщепления, включающие более двух соединений. Количественная оценка событий АС в конкретных условиях основана на предельном процентном индексе сплайсинга (PSI, Ψ) от 0.05

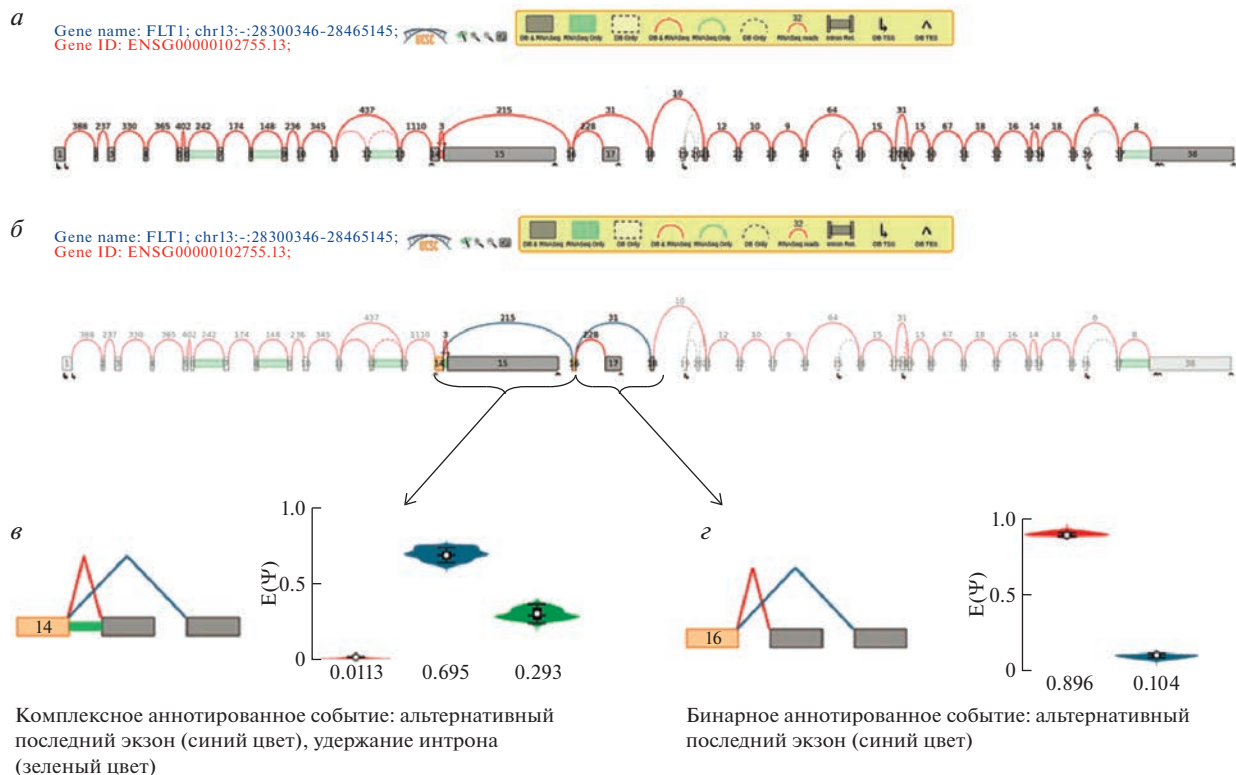


Рис. 4. Визуализация событий АС для гена *FLT1* в контрольной группе с помощью программы MAJIQ VOILA. *а* – интрон-экзонная структура гена *FLT1*; *б* – статистически значимые события в интрон-экзонной структуре; *в* – аннотированное комплексное событие; *г* – аннотированное бинарное событие.

до 0.95 для каждого элемента, участвующего в событии.

В контрольной группе для гена *FLT1* идентифицированы два аннотированных события. Одно из них является комплексным LSV и включает в себя альтернативный последний экзон с уровнем $PSI = 0.695$ и удержание интрона с уровнем $PSI = 0.293$ (рис. 4, *в*), красным цветом указано конститутивное событие соединения экзон–экзон, которое имеет низкий процент включения в транскрипт и $PSI = 0.0113$. Второе событие является классическим бинарным и соответствует альтернативному последнему экзону с $PSI = 0.104$, тогда как конститутивное событие экзон–экзон имеет $PSI = 0.896$ и более вероятно будет включено в транскрипт (рис. 4, *г*).

В группе с преэклампсией для гена *FLT1* идентифицированы два аннотированных и два *de novo* события. Если рассмотреть группу аннотированных событий, то они имеют ту же структуру, что и в контрольной группе, но незначительно отличающийся процентный индекс сплайсинга. Одно из них – комплексное LSV и включает в себя альтернативный последний экзон с уровнем $PSI = 0.702$ и удержание интрона с уровнем $PSI = 0.285$ (рис. 5, *з*), красным цветом указано конститутивное событие соединения экзон–экзон, которое

имеет низкий процент включения в транскрипт и $PSI = 0.0131$. Второе событие является классическим бинарным и соответствует альтернативному последнему экзону с $PSI = 0.338$, тогда как конститутивное событие экзон–экзон имеет $PSI = 0.662$ и более вероятно будет включено в транскрипт (рис. 5, *д*). Два события *de novo* являются бинарными и соответствуют удержанию интрона между экзонами 6 и 7 (рис. 5, *е*), между экзонами 37 и 38 (рис. 5, *е*). Оба события имеют невысокий процентный индекс сплайсинга, равный 0.0662 и 0.117 соответственно.

При оценке уровня экспрессии отдельных изоформ гена *FLT1* во всех 16 образцах были представлены следующие транскрипты (рис. 6): ENST00000282397.9 (кодирует белок FLT-1), ENST00000541932.5 (sFLT-1 e15b), ENST00000615840.4 (sFLT-i13), ENST00000639477.1 (sFLT-e15a) и ENST00000540678.2, который не имеет белкового продукта. Согласно U-критерию Манна–Уитни статистически значимо между контрольной группой и группой с преэклампсией отличается экспрессия транскриптов ENST00000282397.9 ($Z = -2.941$, $p = 0.0019$) и ENST00000639477.1 ($Z = -2.836$, $p = 0.0029$).

Примечательно, что уровень экспрессии транскрипта, кодирующего белковую изоформу FLT-1

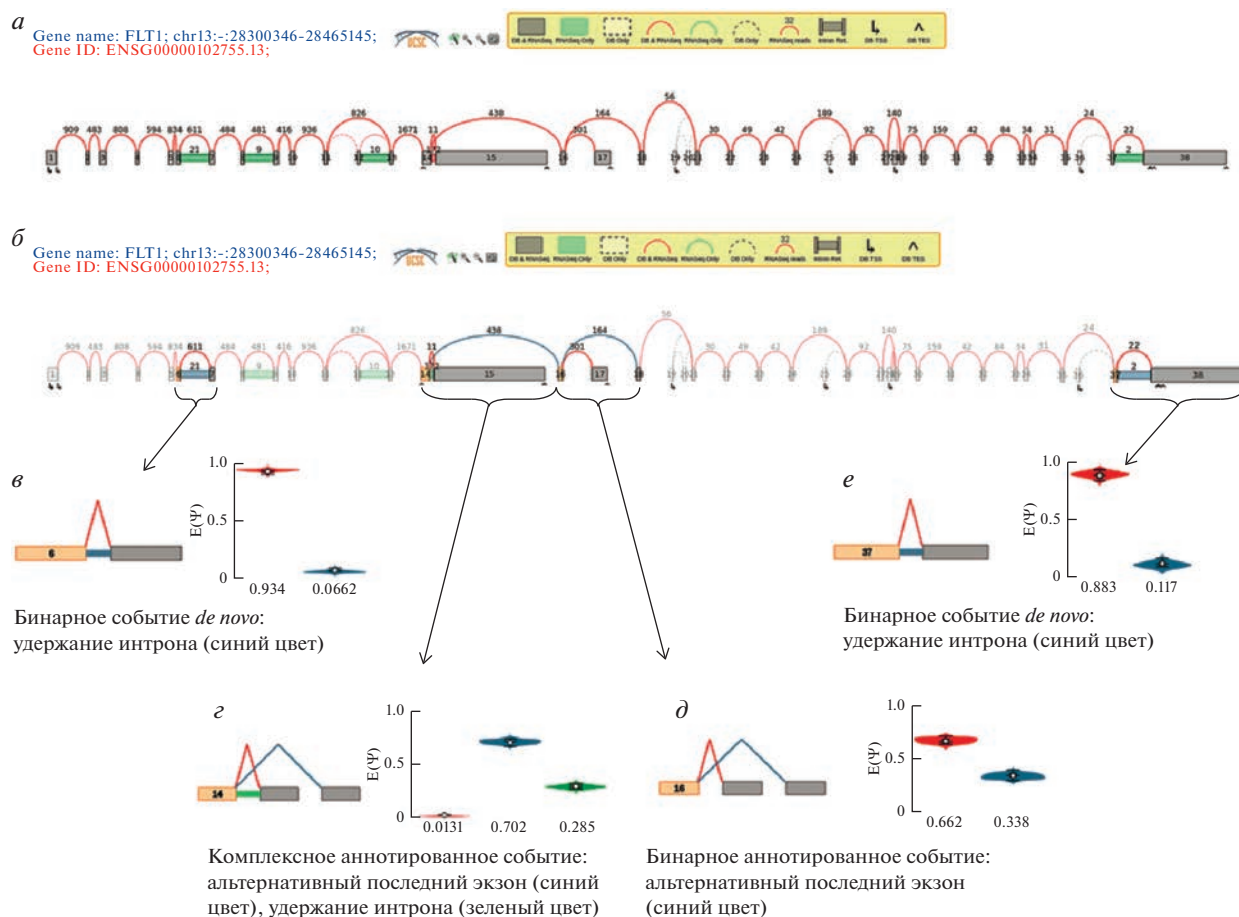


Рис. 5. Визуализация событий АС для гена *FLT1* в группе с преэклампсией с помощью программы MAJIQ VOILA. *a* – интрон-экзонная структура гена *FLT1*; *б* – статистически значимые события в интрон-экзонной структуре; *в* – *de novo* бинарное событие 1; *г* – аннотированное комплексное событие; *д* – аннотированное бинарное событие; *е* – *de novo* бинарное событие 2.

е15а, которая согласно литературным данным гиперэкспрессируется в крови и синцитиотрофобласте при ПЭ [22], был значительно повышен у пациенток с ПЭ по сравнению с контрольной группой (так, среднее количество CPM составило 88.21 в контрольной группе и 390.15 в группе с преэклампсией). В экспериментах *in vitro* на клеточных культурах протеин sFlt-1e15a снижал миграцию эндотелиальных клеток, инвазию и ингибировал образование трубочек в эндотелиальных клетках [22]. Предполагается, что именно эта изоформа FLT-1 может вызывать эндотелиальную дисфункцию и дисфункцию органов-мишеней, поражающихся при ПЭ [23], что позволяет рассматривать данный вариант как перспективный в отношении разработки современных методов диагностики и таргетной терапии этого тяжелого гестационного заболевания.

Транскриптомные исследования многофакторных заболеваний являются одним из приоритетных направлений современной персонализированной медицины, в том числе и в области аку-

шерства. Несмотря на значительные успехи в исследовании этиопатогенеза осложненного течения беременности, многие вопросы, связанные с молекулярными механизмами данного состояния, остаются открытыми и требуют детального изучения. Современные методы транскриптомики предоставляют подходящую методологическую базу для реализации подобных исследований. Основным звеном патогенеза преэклампсии считается нарушение плацентации с недостаточной инвазией трофобласта в структуры матки и неполноценным ремоделированием спиральных артерий, приводящее к снижению кровоснабжения и ишемии плаценты. Показано, что основным структурным элементом и регулятором функций сформированной плаценты является синцитиотрофобласт, выступающий ворсины плаценты и непосредственно контактирующий с материнской кровью. Представленная работа – первое в России полнотранскриптомное исследование клеток синцитиотрофобласта плаценты человека. Выявлены 26 генов, экспрессия которых статистиче-

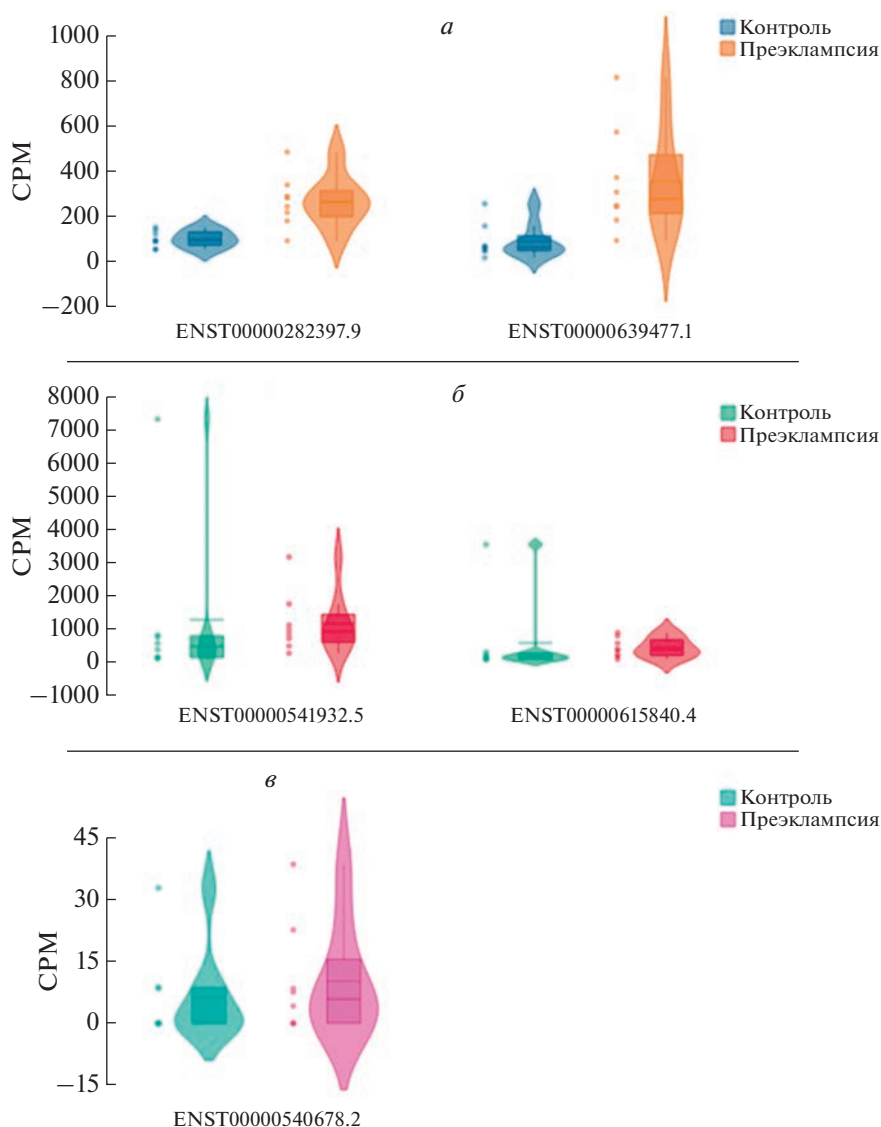


Рис. 6. Вариабельность изоформ гена *FLT1* в клетках СТБ.

ски значительно различается в клетках СТБ женщин с ПЭ и с физиологическим течением беременности. Кластер ДЭГ содержит не только известные гены-кандидаты, выявленные ранее во многих зарубежных полногеномных исследованиях плаценты (к примеру, *LEP*, *INHBA* и *FLT1*), но и новые гены (*AC098613.1*, *AC087857.1*, *FCRLB*, *TENM4*, *PTP4A1P7*, *LINC01225* и др.), которые могут рассматриваться в качестве новых биологических маркеров ПЭ и представляют интерес для дальнейшего изучения. Результаты функциональной аннотации ДЭГ показывают, что с развитием ПЭ на уровне СТБ могут быть связаны сигнальные пути регуляции гормональной секреции, MAPK-каскада, ERK1 и ERK2 каскада, положительной регуляции клеточной адгезии, пролиферации эндотелиальных клеток и др. Полученные результа-

ты анализа АС гена *FLT1* свидетельствуют о важной роли в патогенетике ПЭ данного механизма процессинга РНК, который существенно увеличивает транскрипционное разнообразие генов-кандидатов ПЭ в клетках СТБ.

Профилирование плацентарного транскриптома на уровне единичных клеток и отдельных субпопуляций клеток плацентарной ткани, а также интегративный анализ с геномными, протеомными и метиломными данными представляет собой огромный потенциал для расширения знаний о функционировании плаценты как при физиологической, так и патологической беременности, выявления основных причинно-следственных механизмов развития ПЭ, а также для прогнозирования и улучшения материнских и перинаталь-

ных исходов при осложненном течении беременности.

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Gong S., Gaccioli F., Dopierala J. et al. The RNA landscape of the human placenta in health and disease // Nat. Communications. 2021. V. 12. № 1. P. 2639. <https://doi.org/10.1038/s41467-021-22695-y>
2. Dimitriadis E., Rolnik D.L., Zhou W. et al. Pre-eclampsia // Nat. Reviews Disease Primers. 2023. V. 9. № 1. P. 8. <https://doi.org/10.1038/s41572-023-00417-6>
3. Verlohren S., Brennecke S.P., Galindo A. et al. Clinical interpretation and implementation of the sFlt-1/PIGF ratio in the prediction, diagnosis and management of preeclampsia // Pregnancy Hypertension. 2022. V. 27. P. 42–50. <https://doi.org/10.1016/j.preghy.2021.12.003>
4. Jena M.K., Sharma N.R., Petitt M. et al. Pathogenesis of preeclampsia and therapeutic approaches targeting the placenta // Biomolecules. 2020. V. 10. № 6. P. 953. <https://doi.org/10.3390/biom10060953>
5. Redman C.W.G., Staff A.C., Roberts J.M. Syncytiotrophoblast stress in preeclampsia: the convergence point for multiple pathways // Am. J. Obstetrics and Gynecology. 2022. V. 226. № 2. P. S907–S927. <https://doi.org/10.1016/j.ajog.2020.09.047>
6. Tsang J.C.H., Vong J.S., Ji L. et al. Integrative single-cell and cell-free plasma RNA transcriptomics elucidates placental cellular dynamics // Proc. Natl Acad. Sci. USA. 2017. V. 114. № 37. P. E7786–E7795. <https://doi.org/10.1073/pnas.1710470114>
7. Li H., Huang Q., Liu Y., Garmire L.X. Single cell transcriptome research in human placenta // Reproduction (Cambridge, England). 2020. V. 160. № 6. P. R155. <https://doi.org/10.1530/REP-20-0231>
8. Benton S.J., Leavey K., Grynspan D. et al. The clinical heterogeneity of preeclampsia is related to both placental gene expression and placental histopathology // Am. J. Obstetrics and Gynecology. 2018. V. 219. № 6. P. 604.e1–604.e25. <https://doi.org/10.1016/j.ajog.2018.09.036>
9. Trifonova E.A., Gabidulina T.V., Ershov N.I. et al. Analysis of the placental tissue transcriptome of normal and preeclampsia complicated pregnancies // Acta Naturae. 2014. V. 6. № 2(21). P. 71–83. <https://doi.org/10.32607/20758251-2014-6-2-71-83>
10. Vashukova E.S., Glotov A.S., Fedotov P.V. et al. Placental microRNA expression in pregnancies complicated by superimposed pre-eclampsia on chronic hypertension // Mol. Med. Reports. 2016. V. 14. № 1. P. 22–32. <https://doi.org/10.3892/mmr.2016.5268>
11. Liu S., Xie X., Lei H. et al. Identification of key circRNAs/lncRNAs/miRNAs/mRNAs and pathways in preeclampsia using bioinformatics analysis // Medical Science Monitor: Intern. Med. J. Experim. Clin. Research. 2019. V. 25. P. 1679. <https://doi.org/10.12659/MSM.912801>
12. Robson S.C., Simpson H., Ball E. et al. Punch biopsy of the human placental bed // Am. J. Obstetrics and Gynecology. 2002. V. 187. № 5. P. 1349–1355. <https://doi.org/10.1067/mob.2002.126866>
13. Robinson M.D., McCarthy D.J., Smyth G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data // Bioinformatics. 2010. V. 26. № 1. P. 139–140. <https://doi.org/10.1093/bioinformatics/btp616>
14. Liberzon A., Subramanian A., Pinchback R. et al. Molecular signatures database (MSigDB) 3.0 // Bioinformatics. 2011. V. 27. № 12. P. 1739–1740. <https://doi.org/10.1093/bioinformatics/btr260>
15. Franz M., Rodriguez H., Lopes C. et al. GeneMANIA update 2018 // Nucl. Acids Res. 2018. V. 46. № W1. P. W60–W64. <https://doi.org/10.1093/nar/gky311>
16. Vaquero-Garcia J., Aicher J.K., Jewell S. et al. RNA splicing analysis using heterogeneous and large RNA-seq datasets // Nat. Communications. 2023. V. 14. № 1. P. 1230. <https://doi.org/10.1038/s41467-023-36585-y>
17. Wang Z., Zhao G., Zibrila A.I. et al. Acetylcholine ameliorated hypoxia-induced oxidative stress and apoptosis in trophoblast cells via p38 MAPK/NF-κB pathway // Mol. Hum. Reproduction. 2021. V. 27. № 8. <https://doi.org/10.1093/molehr/gaab045>
18. Guo L., Liu M., Duan T. Hydrogen suppresses oxidative stress by inhibiting the p38 MAPK signaling pathway in preeclampsia // Adv. Clin. & Experim. Medicine. 2023. V. 32. № 3. P. 357–367. <https://doi.org/10.17219/acem/154623>
19. Zhang J., Liu X., Gao Y. Abnormal H3K27 histone methylation of RASA1 gene leads to unexplained recurrent spontaneous abortion by regulating Ras-MAPK pathway in trophoblast cells // Mol. Biol. Reports. 2021. V. 48. № 6. P. 5109–5119. <https://doi.org/10.1007/s11033-021-06507-6>
20. Karumanchi S.A. Angiogenic factors in preeclampsia: from diagnosis to therapy // Hypertension. 2016. V. 67. № 6. P. 1072–1079. <https://doi.org/10.1161/HYPERTENSIONAHA.116.06421>
21. Hiratsuka S., Minowa O., Kuno J. et al. Flt-1 lacking the tyrosine kinase domain is sufficient for normal development and angiogenesis in mice // Proc. Natl Acad. Sci. USA. 1998. V. 95. № 16. P. 9349–9354. <https://doi.org/10.1073/pnas.95.16.9349>
22. Palmer K.R., Kaitu'u-Lino T.U.J., Hastie R. et al. Placental-specific sFLT-1 e15a protein is increased in preeclampsia, antagonizes vascular endothelial growth factor signaling, and has antiangiogenic activity // Hypertension. 2015. V. 66. № 6. P. 1251–1259. <https://doi.org/10.1161/HYPERTENSIONAHA.115.05883>
23. Palmer K. Assessing the circulating placental-specific anti-angiogenic protein sFLT-1 e15a in preeclampsia // Preeclampsia: Methods and Protocols. 2018. P. 27–37. https://doi.org/10.1007/978-1-4939-7498-6_3

Transcriptomic Profiling of Placental Cells in Preeclampsia as an Effective Tool for Personalized Medicine

E. A. Trifonova^{a, *}, A. A. Babovskaya^a, A. A. Zarubin^a, V. N. Serebrova^a,
M. M. Gavrilenko^a, M. G. Svarovskaya^a, E. V. Izhoikina^b, I. G. Kutsenko^b, and V. A. Stepanov^a

^aResearch Institute of Medical Genetics, Tomsk National Research Medical Center, Tomsk, 634050 Russia

^bSiberian State Medical University, Tomsk, 634050 Russia

*e-mail: ekaterina-trifonova@medgenetics.ru

At present transcriptomics is one of the fastest developing fields of molecular biology, which allows to obtain detailed information about the functional activity of the genome both in normal and pathological conditions. We used modern transcriptomic technologies to comprehensively characterize the whole genome gene expression profile of human placental syncytiotrophoblast cells (STB) in physiological pregnancy and preeclampsia (PE). As a result of our analysis, we identified 26 differentially expressed genes (DEGs) in the STB cells between healthy and diseased states. The cluster of DEGs contains not only well-known candidate genes identified earlier in many foreign whole genome studies of the placenta (for example, *LEP*, *INHBA* and *FLT1*), but also new genes (*AC098613.1*, *AC087857.1*, *FCRLB*, *TENM4*, *PTP4A1P7*, *LINC01225*, etc.) that can be considered as new biological markers of PE and are of interest for further study. Functional enrichment annotation indicated that most of the DEGs were implicated in the signaling pathways of regulation of hormonal secretion, MAPK cascade, ERK1 and ERK2 cascade, positive regulation of cell adhesion and proliferation of endothelial cells. These processes may be associated with the development of PE at the STB cells level. Additionally, we revealed that alternative splicing of the *FLT1* gene indicate the important role of this RNA processing mechanism in the pathogenetics of PE due to a significant increase in the transcriptional diversity of genes in STB cells. The expression level of the transcript encoding the protein isoform FLT-1 e15a was significantly increased in patients with PE compared to the control group. This study expands understanding of the molecular mechanisms involved in PE and can serve as a basis for developing of preventive, prognostic and therapeutic strategies in the field of personalized obstetrics.

Keywords: whole transcriptome sequencing, placenta, syncytiotrophoblast, gene, sFlt-1, alternative splicing, MAPK cascade.

ОПИСАНИЕ ДИВЕРГЕНЦИИ СУБПОПУЛЯЦИЙ В ИЕРАРХИЧЕСКОЙ
СИСТЕМЕ ПРИ АНАЛИЗЕ ИЗОНИМИИ. III. КОЭФФИЦИЕНТ
ФАМИЛЬНОГО ИНБРИДИНГА

© 2023 г. В. П. Пасеков*

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

Поступила в редакцию 03.05.2023 г.

После доработки 30.06.2023 г.

Принята к публикации 04.07.2023 г.

Рассматривается характерная для человека иерархическая структура метапопуляций с их подразделенностью на части (субпопуляции), обычно классифицируемые на основе административно-территориального деления, принципах биологической систематики и др. Целью настоящей работы является продолжение анализа общих свойств распределения концентрации фамилии по субпопуляциям при их иерархической подразделенности. Внимание концентрируется на описании фамильной дивергенции субпопуляций. Через показатели дивергенции выражается коэффициент фамильного инбридинга F_s , аналогичный традиционному коэффициенту инбридинга в популяционной генетике. Получено его разложение по уровням иерархии на сумму приращений коэффициента фамильного инбридинга, соответствующих отдельным уровням. Продемонстрирована связь найденного разложения с разложением дисперсии распределения концентрации фамилии по субпопуляциям. При дополнительном предположении о независимости концентрации фамилии и F_s найдена факторизация F_s по уровням иерархии в русле идей С. Райта. Она упрощается при малой фамильной дивергенции субпопуляций, когда коэффициент F_s всей метапопуляции равен сумме средних значений F_s на отдельных уровнях иерархии. Получена величина занижения фамильного инбридинга, когда в качестве единиц наблюдения вместо неподразделенных субпопуляций первого уровня иерархии служат субпопуляции более высокого ранга. Данные результаты являются статистическими характеристиками иерархической структуры, а не особенностью конкретной популяционной системы и не вытекают из той или иной модели микроэволюции. Они вычислительно не зависят от изучаемой иерархической системы, но позволяют характеризовать ее неоднородность количественно. Полученные результаты относятся к сельским и городским иерархическим метапопуляциям как отдельным компонентам всего населения. В Приложении, которое можно читать независимо, приведены без доказательств аналогии найденных свойств применительно к генетической структуре метапопуляции.

Ключевые слова: иерархическая структура популяций, метапопуляции, концентрации фамилии в субпопуляциях человека, характеристики неоднородности метапопуляций, разложение вероятности неизонимных встреч и коэффициента фамильного инбридинга по уровням иерархии.

DOI: 10.31857/S001667582312007X, **EDN:** PSWQUJ

Важность исследований фамильной структуры связана со сходством процессов генного и фамильного дрейфа, что позволяет рассматривать фамилии как квазигенетический маркер [1], а также косвенно судить о генетической структуре населения, важной с медико-биологической точки зрения [2]. Данные по фамилиям, в частности по частоте изонимных браков [3], дают возможность оценивать коэффициент инбридинга (см. критический анализ этого подхода в [4]). Связи фамилий с генетической структурой отражены в монографии [5], исследованиям фамильной структуры населения России посвящена, например,

монография [6]. В работе [7] приведена обширная библиография зарубежных исследований.

Параллельно и одновременно со случайным генным дрейфом в популяциях с общим происхождением протекает их фамильная дивергенция в результате дрейфа фамилий [8, 9], что приводит, в частности, к увеличению дисперсии концентрации фамилии и дисперсии концентрации гена в родственных субпопуляциях. В предыдущей части [10] данной работы внимание фокусировалось на описании свойств дисперсии распределения концентрации фамилии по субпопуляциям в иерархически подразделенной метапопуляции как

характеристике фамильной дивергенции. В следующей публикации [11] подробно рассмотрена общая (полная) вероятность неизоимной встречи (столкновения) индивидуума с фиксированной фамилией с носителем какой-либо другой в субпопуляции, случайно выбранной из рассматриваемой метапопуляции с иерархической структурой.

Целью настоящей работы является продолжение анализа в [11] общих свойств распределения концентрации фамилии по субпопуляциям при их иерархической подразделенности с использованием и обобщениями традиционных подходов популяционной генетики. Рассматривается задача выделения показателей дивергенции, свойственных собственно иерархической системе и независимых от истории ее формирования и взаимодействия между подсистемами, например от внутренних миграционных связей. В данной статье акцент лежит на анализе свойств коэффициента фамильного инбридинга, аналогичного коэффициенту инбридинга в популяционной генетике [12].

Особенности конкретного процесса микроэволюции приводят к своим значениям показателей дивергенции субпопуляций, но на них сказывается и популяционная структура. Поэтому важная задача состоит в том, чтобы выделить свойства показателей дивергенции, присущие самой иерархической структуре и независимые от характера микроэволюционного процесса, изолированности субпопуляций или существования миграций между ними.

Напомним, что под *иерархической системой* понимается система с многоуровневой структурой, элементы которой связаны отношениями подчинения, причем отдельный элемент, в свою очередь, является иерархической системой (более низкого ранга). У нас иерархической системой является метапопуляция, подразделенная на субпопуляции различного уровня иерархии, а подчинение означает вхождение субпопуляции в качестве составной части в субпопуляцию более высокого ранга.

Приведем пояснения относительно обозначений и терминологии. Под концентрациями фамилий в популяции подразумеваются концентрации однофамильцев (носителей данной фамилии). Векторы набраны полужирным шрифтом, к обозначениям фамильных аналогов популяционно-генетических характеристик (дисперсии, коэффициента фамильного инбридинга) добавлено окончание *s* (*Us* и *Fs* соответственно). Символ тождества ($\equiv$) у нас используется в смысле равенства по определению, а символ $\blacktriangleleft$ обозначает конец доказательства.

Дадим обозначения используемых далее характеристик метапопуляции с иерархической структурой и укажем на некоторые из их свойств:

s_i — субпопуляция *i*-го уровня иерархии (с единицами наблюдения $\{s_i\}$);

$s_{i,n}$ — субпопуляция s_i с единицами наблюдения $\{s_n\}$;

$x(s_i)$ — концентрация рассматриваемой фамилии в субпопуляции s_i ;

$x(s_i|s_j)$ — случайная величина, концентрация рассматриваемой фамилии в субпопуляции s_i , выбираемой наугад из s_j ;

$E_{s_j}\{x(s_i|s_j)\}$ — среднее значение случайной величины $x(s_i|s_j)$ (ее математическое ожидание, *E* — его символ, нижний индекс указывает на используемую при усреднении субпопуляцию); по концентрациям $\{x(s_i)\}$ находится среднее значение для $x(s_i|s_j)$, равное $x(s_j)$, или по рекуррентной формуле

$$x(s_j) \equiv E_{s_1}\{x(s_1)|s_j\}, \quad x(s_j) \equiv E_{s_{j-1}}\{x(s_{j-1})|s_j\}; \quad (1)$$

$\{x(s_1)|s_m\}$ — начальные данные для концентрации носителей интересующей фамилии в субпопуляциях первого уровня иерархии внутри метапопуляции s_m ;

$V(x(s_i|s_j))$ — дисперсия случайной величины $x(s_i|s_j)$;

$Hs(x(s_j))$, или $Hs(s_j)$ — вероятность (концентрация, доля) неизоимных встреч в s_j индивидуумов с рассматриваемой фамилией и с какой-нибудь иной с учетом их порядка, находящаяся по определению как среднее значение для субпопуляций $\{s_i\}$ в s_j , или по рекуррентной формуле

$$Hs(s_j) \equiv E_{s_1}\{Hs(s_1)|s_j\}, \\ Hs(s_j) \equiv E_{s_{j-1}}\{Hs(s_{j-1})|s_j\}; \quad (2)$$

$\{Hs(s_1)\}$ — начальные значения вероятностей рассматриваемых неизоимных встреч;

$Hs_0(s_i) \equiv x(s_i)(1 - x(s_i))$ — вероятность такой чисто случайной встречи в s_i с учетом порядка ее участников.

Концентрация $x(s_j)$ рассматриваемого признака в группе s_j (фамилии, аллеля или другого признака) согласно формуле полного математического ожидания находится при произвольном разбиении s_j на подгруппы $\{s_i\}$ как

$$x(s_j) \equiv E_{s_1}\{x(s_1)|s_j\} = \\ = E_{s_1}\{E_{s_i}\{x(s_1)|s_i\}|s_j\} = E_{s_1}\{x(s_i)|s_j\}, \quad (3)$$

т.е. как среднее значение для концентраций $\{x(s_i) = E_{s_i}\{x(s_1)|s_i\}\}$ в подгруппах $\{s_i\}$.

Для субпопуляции $s_j \equiv \{s_i\}$ эту формулу можно представить как

$$E_{s_i}\{E_{s_1}\{T(s_1)|s_i\}|s_j\} = E_{s_1}\{T(s_1)|s_j\}, \\ s_j \equiv \{s_i\}, \quad s_i \equiv \{s_1\}, \quad (4)$$

где *T* — обозначение рассматриваемого признака.

В популяции *s*, подразделенной на субпопуляции (группы) $\{s_i\}$ с межгрупповой дисперсией

$V_{s_{\text{betw}}}(x(s_1|s))$ и случайными встречами в s_1 с вероятностями $Hs_0(s_1) = x(s_1)(1 - x(s_1))$, согласно эффекту Валунда общая (полная) вероятность рассматриваемой неизонимной встречи $Hs(s)$ в выбранной наугад из s субпопуляции s_1 не совпадает с $Hs_0(s) = x(s)(1 - x(s))$, а находится как

$$Hs(s) = Hs_0(x(s)) - V_{s_{\text{betw}}}(x(s_1|s)). \quad (5)$$

Иначе эту формулу можно записать в виде

$$\begin{aligned} Hs(s) &= Hs_0(x(s)) \left(1 - \frac{V_{s_{\text{betw}}}(x(s_1|s))}{Hs_0(s)} \right) = \\ &= Hs_0(x(s))(1 - Fs(s)), \\ Fs(s) &\equiv V_{s_{\text{betw}}}(x(s_1|s))/Hs_0(s), \end{aligned} \quad (6)$$

где $Fs(s)$ — коэффициент фамильного инбридинга субпопуляции s .

Если субпопуляции $\{s_1\}$ дополнительно сгруппированы в $\{s_2\}$, т.е. s разбита на $\{s_2\}$, а каждая из s_2 на $\{s_1\}$, то по правилу сложения дисперсий полная дисперсия $V_{s_{\text{tot}}}(x(s_1|s))$ концентраций $\{x(s_1)\}$ равна сумме межгрупповой $V_{s_{\text{betw}}}(x(s_2|s))$ и средней внутригрупповой $E_{s_2}\{V_{s_{\text{in}}}(x(s_1|s_2))|s\}$ дисперсий:

$$\begin{aligned} V_{s_{\text{tot}}}(x(s_1|s)) &= \\ &= V_{s_{\text{betw}}}(x(s_2|s)) + E_{s_2}\{V_{s_{\text{in}}}(x(s_1|s_2))|s\}. \end{aligned} \quad (7)$$

Эта формула верна, если в многоуровневой субпопуляции s_m вместо второго уровня рассматривать k -й уровень иерархии:

$$\begin{aligned} V_{s_{\text{tot}}}(x(s_1|s_m)) &= V_{s_{\text{betw}}}(x(s_k|s_m)) + \\ &+ E_{s_k}\{V_{s_{\text{in}}}(x(s_1|s_k))|s_m\}, \quad 2 < k < m. \end{aligned} \quad (8)$$

Данное правило играет большую роль в дальнейшем.

Очевидно, при ходьбе пройденное расстояние складывается из приращений на каждом шагу. Аналогично любую числовую величину можно разбить на последовательность частей, границы которых играют роль своего рода делений измерительной линейки. Расстояния между соседними границами дают приращения измеряемой величины при движении вдоль такой “линейки”, а сумма приращений дает измеряемую величину. Если рассматривать полную дисперсию $V_{s_{\text{tot}}}(x(s_1|s_m))$ и в качестве формальных границ частей взять значения дисперсий $Vs(x(s_k|s_m))$ на уровнях иерархии, то сумма приращений дает разложение полной дисперсии. Согласно этому в предыдущей части [11] такое разложение полной дисперсии концентраций $\{x(s_1)\}$ в метапопуляции s_m с m уровнями иерархии представлено в виде

$$\begin{aligned} V_{s_{\text{tot}}}(x(s_1|s_m)) &= \sum_{i=1}^{m-1} \Delta_i Vs(s_m) \equiv \\ &\equiv \sum_{i=1}^{m-1} (Vs(x(s_i|s_m)) - Vs(x(s_{i+1}|s_m))) = \\ &= \sum_{i=1}^{m-1} E_{s_{i+1}}\{Vs(x(s_i|s_{i+1}))|s_m\}. \end{aligned} \quad (9)$$

Дальнейшее изложение придерживается следующего плана. На основании разложения полной дисперсии концентраций $\{x(s_1)\}$ в метапопуляции s_m с m уровнями иерархии получено разложение коэффициента фамильного инбридинга метапопуляции по уровням иерархии. Его изменения от одного уровня к соседнему определяются различиями между дисперсиями концентраций фамилии на смежных уровнях. Для субпопуляции каждого ранга введено понятие коэффициента фамильного инбридинга при условии заданного уровня у единиц наблюдения. Рассмотрены разложения коэффициента фамильного инбридинга метапопуляции в предположении случайных встреч в субпопуляциях, служащих единицами наблюдения, и при нарушении этого предположения. Исследовано понижение оценки дивергенции субпопуляций, когда не учитывается неоднородность (подрозделенность) субпопуляций, служащих единицами наблюдения.

Далее обобщен подход С. Райта к анализу иерархически подразделенной метапопуляции при дополнительном условии независимости концентрации аллеля и коэффициента фамильного инбридинга Fs в контексте анализа фамильной структуры. Получена факторизация Fs по уровням иерархии. Для нее найдена простая аппроксимация в случае малых коэффициентов. Полученные результаты относятся к сельским и городским иерархическим метапопуляциям как отдельным компонентам всего населения. В Приложении приведены без доказательств аналоги найденных в настоящей серии работ результатов по анализу фамильной структуры применительно к данным популяционной генетики. Оно предназначено для интересующихся популяционно-генетическими аспектами иерархически структурированных подразделенных популяций. Приложение можно читать независимо от предыдущего анализа фамильной структуры.

РАЗЛОЖЕНИЕ КОЭФФИЦИЕНТА ФАМИЛЬНОГО ИНБРИДИНГА ПО УРОВНЯМ ИЕРАРХИИ

При чисто случайных столкновениях индивидов в популяции s вероятность $Hs(s)$ встречи неизонимной пары лиц (одного с заданной фамилией с концентрацией $x(s)$ и другого с любой иной фамилией) с учетом порядка в паре имеет вид $Hs_0(s)$

$= x(s)(1 - x(s))$. Предположим, что столкновения неслучайны и вероятность $Hs(s)$ встреч описывается с помощью коэффициента фамильного инбридинга Fs как

$$Hs(s|Fs(s)) = x(s)(1 - x(s))(1 - Fs(s)) \equiv Hs_0(s)(1 - Fs(s)).$$

Отсюда коэффициент фамильного инбридинга выражается как

$$Fs(s) \equiv (Hs_0(s) - Hs(s))/Hs_0(s). \quad (10)$$

На разложении дисперсии и вероятности встреч основывается и разложение коэффициента фамильного инбридинга в соответствии с уровнями иерархии. Выбор для вероятности встреч в единице наблюдения s_1 (с концентрацией $x(s_1)$ интересующей фамилии) значения $Hs_0(s_1) = x(s_1)(1 - x(s_1))$, соответствующего чисто случайным столкновениям, приводит к следующему обобщению эффекта Валунда.

Результат 1. Пусть в иерархически подразделенной метапопуляции s_m с t уровнями иерархии в качестве единиц наблюдения служат субпопуляции $\{s_1\}$ первого уровня. Положим, что концентрации интересующей фамилии в $\{s_1\}$ равны $\{x(s_1)\}$, а вероятности встречи в них пары индивидуумов с этой фамилией и с какой-нибудь иной с учетом их порядка находятся как

$$\{Hs(s_1) = Hs_0(s_1) \equiv x(s_1)(1 - x(s_1))\}.$$

Тогда разложения коэффициента фамильного инбридинга $Fs(s_m)$ по компонентам отдельных уровней иерархии можно записать как

$$\begin{aligned} Fs(s_m) &= \sum_{i=1}^{m-1} \frac{Vs(x(s_i|s_m)) - Vs(x(s_{i+1}|s_m))}{Hs_0(s_m)} = \\ &= \sum_{i=1}^{m-1} \frac{E_{s_{i+1}} \{Vs_{in}(x(s_i|s_{i+1}))|s_m\}}{Hs_0(s_m)} = \\ &= \sum_{i=1}^{m-1} \Delta_i Fs(s_m), \quad (11) \\ \Delta_i Fs(s_m) &\equiv \frac{Vs(x(s_i|s_m)) - Vs(x(s_{i+1}|s_m))}{Hs_0(s_m)} = \\ &= \frac{E_{s_{i+1}} \{Vs_{in}(x(s_i|s_{i+1}))|s_m\}}{Hs_0(s_m)} \equiv \\ &\equiv Fs(s_{m,i}) - Fs(s_{m,i+1}) \geq 0. \end{aligned}$$

Вероятности $Hs(s_m)$ рассматриваемой встречи в наугад выбранной из s_m субпопуляции s_1 представляемы в виде

$$\begin{aligned} Hs(s_m) &= Hs_0(s_m) \times \\ &\times \left(1 - \sum_{i=1}^{m-1} \frac{Vs(x(s_i|s_m)) - Vs(x(s_{i+1}|s_m))}{Hs_0(s_m)} \right) = \\ &= Hs_0(s_m) \left(1 - \sum_{i=1}^{m-1} \frac{E_{s_{i+1}} \{Vs_{in}(x(s_i|s_{i+1}))|s_m\}}{Hs_0(s_m)} \right) = \\ &= Hs_0(s_m) \left(1 - \sum_{i=1}^{m-1} \Delta_i Fs(s_m) \right). \end{aligned} \quad (12)$$

Здесь

$$\begin{aligned} \Delta_i Fs(s_m) &\equiv Fs(s_{m,i}) - Fs(s_{m,i+1}) = \\ &= E_{s_{i+1}} \{Vs_{in}(x(s_i|s_{i+1}))|s_m\} / Hs_0(s_m), \\ Fs(s_m) &\equiv Vs(x(s_1|s_m)) / Hs_0(s_m), \\ Fs(s_{m,i}) &\equiv Vs_{in}(x(s_i|s_m)) / Hs_0(s_m), \\ Hs(s_m) &\equiv E_{s_1} \{Hs(s_1|s_m)\}, \\ Hs_0(s_m) &= x(s_m)(1 - x(s_m)), \end{aligned} \quad (13)$$

$Fs(s_m)$ и $Fs(s_{m,i})$ обозначают коэффициенты фамильного инбридинга в метапопуляции s_m и в $s_{m,i}$, когда в качестве единиц наблюдения служат субпопуляции $\{s_1\}$ и $\{s_i\}$ с уровнями иерархии 1 и i при условии, что вероятности рассматриваемых встреч в них равны $\{Hs_0(s_1)\}$ и $\{Hs_0(s_i)\}$ соответственно; $Hs(s_m)$ — общая вероятность рассматриваемых встреч в s_m .

Доказательство. В иерархически подразделенной метапопуляции s_m при $\{Hs(s_1) \equiv Hs_0(s_1) = x(s_1)(1 - x(s_1))\}$ согласно эффекту Валунда (5) и разложению (9) дисперсии $Vs(x(s_1|s_m))$

$$\begin{aligned} Hs(s_m) &= Hs_0(s_m) - Vs(x(s_1|s_m)) = \\ &= Hs_0(s_m) - \sum_{i=1}^{m-1} (Vs(x(s_i|s_m)) - Vs(x(s_{i+1}|s_m))) = \\ &= Hs_0(s_m) \left(1 - \sum_{i=1}^{m-1} \frac{Vs(x(s_i|s_m)) - Vs(x(s_{i+1}|s_m))}{Hs_0(s_m)} \right) = \\ &= Hs_0(s_m) \left(1 - \sum_{i=1}^{m-1} \frac{E_{s_{i+1}} \{Vs_{in}(x(s_i|s_{i+1}))|s_m\}}{Hs_0(s_m)} \right). \end{aligned}$$

В последнем равенстве использовано правило сложения дисперсий (8), когда в метапопуляции $s = s_m$ рассматриваются только три уровня иерархии (m , $i+1$, i). Тогда дисперсии $Vs_{tot}(x(s_i|s_m))$, $Vs_{betw}(x(s_{i+1}|s_m))$ и $E_{s_{i+1}} \{Vs_{in}(x(s_i|s_{i+1}))|s_m\}$ рассматриваются как полная, межгрупповая и средняя внутригрупповая дисперсии соответственно. В силу этого правила разность первых двух равна третьей, т.е. уменьшение межгрупповой дисперсии при переходе от i к смежному более высокому уровню равно средней внутригрупповой дисперсии для $\{s_{i+1}\}$:

$$Vs(x(s_i|s_m)) - Vs(x(s_{i+1}|s_m)) = E_{s_{i+1}} \{Vs_{in}(x(s_i|s_{i+1}))|s_m\}.$$

Согласно данному выше определению (13) коэффициента фамильного инбридинга $Fs(s_m)$ и (9) находим

$$\begin{aligned} Fs(s_m) &\equiv \frac{Vs(x(s_1|s_m))}{Hs_0(s_m)} = \\ &= \sum_{i=1}^{m-1} \frac{E_{s_{i+1}}\{Vs(x(s_i|s_{i+1})|s_m)\}}{Hs_0(s_m)} = \\ &= \sum_{i=1}^{m-1} \left(\frac{Vs(x(s_i|s_m))}{Hs_0(s_m)} - \frac{Vs(x(s_{i+1}|s_m))}{Hs_0(s_m)} \right). \end{aligned}$$

С учетом определения $\Delta_i Fs(s_m)$ подстановка данного разложения для Fs в $Hs_0(1 - Fs)$ дает (12). ◀

Рассмотрим компоненты i -го слагаемого в правой части последней формулы:

$$\begin{aligned} \frac{Vs(x(s_i|s_m))}{Hs_0(s_m)} &\equiv Fs(s_{m,i}), \\ \frac{Vs(x(s_{i+1}|s_m))}{Hs_0(s_m)} &\equiv Fs(s_{m,i+1}). \end{aligned}$$

Здесь отношение $Vs(x(s_i|s_m))/Hs_0(s_m)$ согласно определению (13) равно коэффициенту фамильного инбридинга $Fs(s_{m,i})$ метапопуляции s_m , где в качестве единиц наблюдения используются субпопуляции $\{s_i\}$ на i -м уровне иерархии и встречи (столкновения) внутри них чисто случайны. Данный коэффициент характеризует подразделенность метапопуляции на i -м уровне (интерпретируется как коэффициент фамильного инбридинга $Fs(s_m)$ у метапопуляции s_m при отсутствии в ней подразделенности на уровне ниже i). Аналогично $Fs(s_{m,i+1})$ характеризует подразделенность метапопуляции на уровне $i+1$, а разность $Fs(s_{m,i}) - Fs(s_{m,i+1})$ дает вклад $\Delta_i Fs(s_m)$ в коэффициент фамильного инбридинга $Fs(s_m)$.

Случайность встреч внутри любой из субпопуляций $\{s_i\}$ означает однородность каждой в смысле отсутствия препятствий для встреч. При этом формально s_i едина, а субпопуляции $\{s_{i-1}\}$ как бы условны (границы, при существовании которых столкновения происходят только между индивидами одной и той же субпопуляции, у них отсутствуют). Базовыми единицами являются такие субпопуляции, внутри которых столкновения случайны (значения концентраций в них служат начальными данными). Их группировка не оказывает влияния на общую (полную) вероятность встреч во всей метапопуляции подобно тому, как количество гетерозигот в метапопуляции складывается из количества в ее подразделениях независимо от их группировок при подсчетах (на сумму не влияет группировка слагаемых).

Рассмотрим погрешности выражений для вероятности рассматриваемых встреч и для коэффициента фамильного инбридинга, вносимые использованием вместо начальных данных (концентрации

интересующих встреч в единицах наблюдения $\{s_1\}$ со случайными столкновениями) агрегированных сведений для субпопуляций более высокого ранга $\{s_n\}$. В общем случае выбор значений $x(s_n)(1 - x(s_n))$ для вероятности встреч в субпопуляциях $\{s_n\}$, взятых в качестве единиц наблюдения уровня n , не обоснован. Использование для вероятности встреч в единице наблюдения s_n (с концентрацией $x(s_n)$ интересующей фамилии) значения $Hs_0(s_n) = x(s_n) \times (1 - x(s_n))$, соответствующего чисто случайным столкновениям, вместо неизвестного реального значения приводит к следующему утверждению.

Следствие 2. Пусть в иерархически подразделенной метапопуляции s_m с t уровнями иерархии в качестве единиц наблюдения вместо субпопуляций первого уровня $\{s_1\}$ с концентрациями фамилии $\{x(s_1)\}$ и случайными столкновениями рассматриваются субпопуляции уровня n ($1 < n < t - 1$) с концентрациями $\{x(s_n)\}$. Предположим, что при отсутствии или игнорировании сведений о наличии в них препятствий для встреч индивидов принимается решение о случайных столкновениях в каждой из субпопуляций $\{s_n\}$, т.е. вероятность $Hs(s_{m,n})$ встреч в s_m пары индивидов с интересующей фамилией и с какой-нибудь иной с учетом их порядка (в наугад выбранной из s_m субпопуляции s_n) находится как

$$\begin{aligned} Hs(s_{m,n}) &\equiv E_{s_n}\{Hs(s_n)|s_m\} = E_{s_n}\{Hs_0(s_n)|s_m\}, \\ Hs_0(s_n)(s_m) &\equiv x(s_n)(1 - x(s_n)), \end{aligned}$$

где $s_{m,n}$ обозначает метапопуляцию s_m с единицами наблюдения $\{s_n\}$.

Тогда коэффициент фамильного инбридинга $Fs(s_{m,n})$, определяемый из условия $Hs(s_{m,n}) = Hs_0(s_m) \times (1 - Fs(s_{m,n}))$, можно разложить по компонентам на уровнях иерархии как

$$\begin{aligned} Fs(s_{m,n}) &= \sum_{i=n}^{m-1} \frac{Vs(x(s_i|s_m)) - Vs(x(s_{i+1}|s_m))}{x(s_m)(1 - x(s_m))} = \\ &= \sum_{i=n}^{m-1} \frac{E_{s_{i+1}}\{Vs(x(s_i|s_{i+1})|s_m)\}}{x(s_m)(1 - x(s_m))} \equiv \sum_{i=n}^{m-1} \Delta_i Fs(s_m), \\ \Delta_i Fs(s_m) &\equiv \frac{Vs(x(s_i|s_m)) - Vs(x(s_{i+1}|s_m))}{x(s_m)(1 - x(s_m))} = \\ &= \frac{E_{s_{i+1}}\{Vs(x(s_i|s_{i+1})|s_m)\}}{x(s_m)(1 - x(s_m))} \geq 0. \end{aligned}$$

Коэффициент $Fs(s_{m,n})$ меньше значения $Fs(s_m)$, соответствующего единице наблюдения $\{s_1\}$, на сумму вкладов неучтенных уровней иерархии от 1 до $n - 1$:

$$\begin{aligned}
Fs(s_m) - Fs(s_{m,n}) &= \\
&= \sum_{i=1}^{m-1} \frac{Vs(x(s_i|s_m)) - Vs(x(s_{i+1}|s_m))}{x(s_m)(1-x(s_m))} - \\
&- \sum_{i=n}^{m-1} \frac{Vs(x(s_i|s_m)) - Vs(x(s_{i+1}|s_m))}{x(s_m)(1-x(s_m))} = \\
&= \sum_{i=1}^{n-1} \frac{E_{s_{i+1}}\{Vs(x(s_i|s_{i+1}))|s_m\}}{x(s_m)(1-x(s_m))} = \sum_{i=1}^{n-1} \Delta_i Fs(s_m) \geq 0.
\end{aligned}$$

Доказательство. Воспользуемся эффектом Валунда (5), согласно которому полная вероятность рассматриваемых встреч в метапопуляции s_m , где в качестве единиц наблюдения служат субпопуляции $\{s_1\}$, находится как

$$\begin{aligned}
Hs(s_m) &\equiv E_{s_1}\{Hs_0(s_1)|s_m\} = \\
&= Hs_0(s_m) \left(1 - \frac{Vs_{\text{betw}}(x(s_1|s_m))}{Hs_0(s_m)}\right) = \\
&= Hs_0(s_m)(1 - Fs(s_m)), \\
Fs(s_m) &\equiv \frac{Vs_{\text{betw}}(x(s_1|s_m))}{Hs_0(s_m)} \geq 0.
\end{aligned}$$

Пусть при анализе метапопуляции s в качестве единиц наблюдения выбраны субпопуляции $\{s_n\}$ на n -м уровне иерархии, например из-за отсутствия информации о существующей в них подразделенности или о наличии других препятствий для встреч индивидуумов, или же из-за игнорирования таких сведений. При упрощенном предположении о случайных столкновениях в каждой из субпопуляций $\{s_n\}$ получаем вероятность встреч $Hs(s_{m,n})$ в s_m заменой в приведенной формуле s_1 на s_n :

$$\begin{aligned}
Hs(s_{m,n}) &\equiv E_{s_n}\{Hs_0(s_n)|s_m\} = \\
&= Hs_0(s_m) \left(1 - \frac{Vs_{\text{betw}}(x(s_n|s_m))}{Hs_0(s_m)}\right) = \\
&= Hs_0(s_m)(1 - Fs(s_{m,n})), \\
Fs(s_{m,n}) &\equiv \frac{Vs_{\text{betw}}(x(s_n|s_m))}{Hs_0(s_m)} \geq 0.
\end{aligned}$$

В результате выбора единицы наблюдения $\{s_n\}$ вместо $\{s_1\}$ вероятность встреч увеличится согласно (5)–(6) на

$$\begin{aligned}
Hs(s_{m,n}) - Hs(s_m) &= \\
&= Vs_{\text{tot}}(x(s_1|s_m)) - Vs_{\text{betw}}(x(s_n|s_m)) = \\
&= Hs_0(s_m)(Fs(s_m) - Fs(s_{m,n})) = \\
&= E_{s_n}\{Vs_{\text{in}}(x(s_1|s_n))|s_m\} \geq 0,
\end{aligned}$$

где мы учли выражение для разности между общей и межгрупповой дисперсиями согласно (8).

С другой стороны, в соответствии с представлением $Hs = Hs_0(1 - Fs)$ коэффициент фамилного инбридинга уменьшится на

$$\begin{aligned}
Fs(s_m) - Fs(s_{m,n}) &= \\
&= E_{s_n}\{Vs_{\text{in}}(x(s_1|s_n))|s_m\} / Hs_0(s_m) \geq 0. \blacktriangleleft
\end{aligned}$$

Следствие 3. В условиях предыдущего следствия разложение вероятности встреч $Hs(s_{m,n}) \equiv E_{s_n}\{Hs(s_n)|s_m\}$ имеет вид

$$\begin{aligned}
Hs(s_{m,n}) &= x(s_m)(1-x(s_m))(1-Fs(s_{m,n})) = \\
&= Hs_0(s_m) \left(1 - \sum_{i=n}^{m-1} \frac{E_{s_{i+1}}\{Vs(x(s_i|s_{i+1}))|s_m\}}{x(s_m)(1-x(s_m))}\right) \equiv \\
&\equiv Hs_0(s_m) \left(1 - \sum_{i=n}^{m-1} \Delta_i Fs(s_m)\right).
\end{aligned}$$

Вероятность встреч $Hs(s_{m,n})$ больше значения $Hs(s_m)$ при единицах наблюдения $\{s_1\}$ первого уровня иерархии, где $Hs(s_1) = Hs_0(s_1) \equiv x(s_1)(1-x(s_1))$, на величину

$$\begin{aligned}
&\sum_{i=1}^{n-1} (Vs(x(s_i|s_m)) - Vs(x(s_{i+1}|s_m))) = \\
&= \sum_{i=1}^{n-1} E_{s_{i+1}}\{Vs(x(s_i|s_{i+1}))|s_m\} = \\
&= Vs(x(s_1|s_n)) \geq 0.
\end{aligned}$$

Отклонение $Hs(s_{m,n}) - Hs_0(s_m)$ возрастает на $E_{s_{n+1}}\{Vs(x(s_n|s_{n+1}))|s_m\} \geq 0$ при повышении ранга единицы наблюдения субпопуляции от n до следующего уровня.

Коэффициент фамилного инбридинга с повышением ранга единицы наблюдения от n до следующего уровня, напротив, уменьшается на $E_{s_{n+1}}\{Vs_{\text{in}}(x(s_n|s_{n+1}))|s_m\} / Hs_0(s_m)$.

Доказательство. Суммирование неотрицательных слагаемых $\Delta_i Fs(s_m)$ в разложении (11) коэффициента фамилного инбридинга $Fs(s_{m,n})$ начинается с номера иерархии у единицы наблюдения, т.е. не с единицы, как для $Fs(s_m)$, а с n . Следовательно, $Fs(s_{m,n})$ оказывается меньше значения $Fs(s_m)$ на величину суммы вкладов неучтенных уровней иерархии от 1 до $n-1$:

$$\begin{aligned}
Fs(s_{m,n}) &= \sum_{i=n}^{m-1} \Delta_i Fs(s_m), \\
Fs(s_m) - Fs(s_{m,n}) &= \sum_{i=1}^{n-1} \Delta_i Fs(s_m) = \\
&= \sum_{i=1}^{n-1} E_{s_{i+1}}\{Vs(x(s_i|s_{i+1}))|s_m\} / Hs_0(s_m) \geq 0.
\end{aligned}$$

Вероятность встреч $Hs(s_{m,n})$ оказывается больше значения $Hs(s_m)$ при единице наблюдения $\{s_1\}$:

$$\begin{aligned}
Hs(s_{m,n}) &= x(s_m)(1 - x(s_m))(1 - Fs(s_{m,n})) = \\
&= Hs_0(s_m) \left(1 - \sum_{i=n}^{m-1} \Delta_i Fs(s_m) \right), \\
Hs(s_{m,n}) - Hs_0(s_m) &= Hs_0(s_m)(1 - Fs(s_{m,n})) - \\
&\quad - Hs_0(s_m)(1 - Fs(s_m)) \quad (14) \\
&= Hs_0(s_m)(Fs(s_m) - Fs(s_{m,n})) = \\
&= E_{s_n} \{ Vs_{in}(x(s_i|s_n)) | s_m \} / Hs_0(s_m) = \\
&= \sum_{i=1}^{n-1} E_{s_{i+1}} \{ Vs(x(s_i|s_{i+1})) | s_m \} / Hs_0(s_m) \geq 0.
\end{aligned}$$

Очевидно расхождение $Hs(s_{m,n}) - Hs(s_m)$ при увеличении единицы наблюдения субпопуляции от n до $n + 1$ возрастает на величину $\Delta_n Hs(s_m) \equiv Hs_0(s_m) \Delta_n Fs(s_m) = E_{s_{n+1}} \{ Vs(x(s_n|s_{n+1})) | s_m \}$. Отсюда соответствующим изменением коэффициента фамильного инбридинга будет

$$\begin{aligned}
\Delta_n Fs(s_m) &= \Delta_n Hs(s_m) / Hs_0(s_m) = \\
&= E_{s_{n+1}} \{ Vs(x(s_n|s_{n+1})) | s_m \} / Hs_0(s_m). \blacktriangleleft
\end{aligned}$$

ФАКТОРИЗАЦИЯ КОЭФФИЦИЕНТА ФАМИЛЬНОГО ИНБРИДИНГА ПО УРОВНЯМ ИЕРАРХИИ ПРИ ЕГО НЕЗАВИСИМОСТИ ОТ КОНЦЕНТРАЦИИ ФАМИЛИИ

Другой вариант анализа иерархически подразделенной метапопуляции следует подходу С. Райта [12] и *требует дополнительного предположения о независимости концентрации x фамилии и коэффициента фамильного инбридинга Fs* . Отметим, что они разным образом описывают субпопуляции. Если концентрации характеризуют распространенность фамилий по отдельности, то коэффициент фамильного инбридинга характеризует распространенность пар фамилий при встречах носителей (как бы отталкивание или притяжение). Она может принимать разные значения даже при фиксированных концентрациях отдельных фамилий. Поэтому указанное предположение о независимости кажется естественным.

Согласно фамильному аналогу (5) эффекта Валуна в метапопуляции s_2 , подразделенной на субпопуляции первого уровня иерархии $\{s_1\}$ с концентрациями заданной фамилии $\{x(s_1)\}$ и случайными встречами индивидов в них, вероятность $Hs(s_2)$ встречи в s_2 лиц с заданной фамилией и с любой иной находится как

$$\begin{aligned}
Hs(s_2) &= x(s_2)(1 - x(s_2)) - Vs(x(s_1|s_2)) = \\
&= x(s_2)(1 - x(s_2))(1 - Fs(s_2)), \\
Fs(s_2) &\equiv \frac{Vs(x(s_1|s_2))}{x(s_2)(1 - x(s_2))} \geq 0.
\end{aligned}$$

Далее для лучшей ориентации в соответствующих формулах используется обозначение вероятности рассматриваемых встреч как $Hs(s_{m,i})$, где $s_{m,i}$ относится к метапопуляции s_m , в которой единицами наблюдения являются субпопуляции $\{s_i\}$ на i -м уровне иерархии со случайными встречами индивидов в них. Сохраним обозначение $Hs(s_m)$ за вероятностью встреч при единицах наблюдения первого уровня. Когда s_m разбита на $\{s_i\}$ и встречи в $\{s_i\}$ случайны, то $Hs(s_i) = Hs_0(s_i) = x(s_i)(1 - x(s_i))$, а

$$\begin{aligned}
Hs(s_{m,i}) &\equiv E_{s_i} \{ x(s_i)(1 - x(s_i)) | s_m \} = \\
&= x(s_m)(1 - x(s_m)) - Vs(x(s_i|s_m)) = \\
&= x(s_m)(1 - x(s_m))(1 - Fs(s_{m,i})), \quad (15) \\
Fs(s_{m,i}) &\equiv \frac{Vs(x(s_i|s_m))}{x(s_m)(1 - x(s_m))} = \frac{Vs(x(s_i|s_m))}{Hs_0(s_m)}.
\end{aligned}$$

Повторим, что коэффициент фамильного инбридинга $Fs(s_{m,i})$ субпопуляции s_m характеризует дивергенцию субпопуляций $\{s_i\}$ внутри s_m . Его нижний индекс i указывает на служащие единицами наблюдения субпопуляции $\{s_i\}$, встречи в которых предполагаются случайными. В обозначении $Fs(s_i)$ подразумевается, что единицами наблюдения являются субпопуляции $\{s_i\}$.

Рассмотренная выше подразделенная популяция s_2 может быть одной из частей популяции более высокого ранга s_3 . В случае трех уровней иерархии (т.е. при $m = 3$) по определению $Hs(s_3) \equiv E_{s_1} \{ Hs(s_1) | s_3 \}$ и в соответствии с (2) $Hs(s_3) = E_{s_2} \{ Hs(s_2) | s_3 \}$, т.е. концентрация $Hs(s_3)$ является средним значением как для концентраций Hs в $\{s_1\}$, так и (рекуррентно) для концентраций Hs в $\{s_2\}$. В терминах коэффициентов фамильного инбридинга Fs имеем

$$\begin{aligned}
Hs(s_3) &= E_{s_2} \{ Hs(s_2) | s_3 \} = \\
&= E_{s_2} \{ Hs_0(s_2)(1 - Fs(s_{2,1})) | s_3 \}.
\end{aligned}$$

При независимости Fs и концентрации фамилии x эта формула для $Hs(s_3)$ упрощается, поскольку тогда математическое ожидание произведения равно произведению математических ожиданий его сомножителей:

$$\begin{aligned}
Hs(s_3) &= E_{s_2} \{ Hs_0(s_2) | s_3 \} E_{s_2} \{ (1 - Fs(s_{2,1})) | s_3 \} = \\
&= Hs_0(s_3)(1 - Fs(s_{3,2}))(1 - E_{s_2} \{ Fs(s_{2,1}) | s_3 \}).
\end{aligned}$$

Рассмотрим последовательность значений Hs при малых уровнях иерархии:

$$\begin{aligned}
Hs(s_1) &= Hs_0(s_1), \quad Hs(s_2) = Hs_0(s_2)(1 - Fs(s_{2,1})), \\
Hs(s_3) &= Hs_0(s_3)(1 - Fs(s_{3,2}))(1 - E_{s_2} \{ Fs(s_{2,1}) | s_3 \}).
\end{aligned}$$

Из нее просматривается некоторая закономерность зависимости вида факторизации $Hs(s_m)$ от количества уровней иерархии. Первым сомножителем в факторизации является $Hs_0(s_m)$, который при рекуррентном переходе к $Hs(s_{m+1})$ дает согласно фа-

мильному аналогу (5) эффекта Валунда вместо себя два сомножителя: $Hs_0(s_{m+1})$ и $1 - Fs(s_{m+1, m})$. Вторым сомножителем является $1 - Fs(s_{m, m-1})$, заменяемый в результате усреднения по m при рекуррентном переходе к $Hs(s_{m+1})$ на $1 - E_{s_m}\{Fs(s_{m, m-1})|s_{m+1}\} = 1 - E_{s_{i+1}}\{Fs(s_{i+1, i})|s_{m+1}\}_{i=m-1}$. Прочие сомножители представимы как $1 - E_{s_{i+1}}\{Fs(s_{i+1, i})|s_m\}$ и заменяются на $1 - E_{s_{i+1}}\{Fs(s_{i+1, i})|s_{m+1}\}$, где $E_{s_{i+1}}\{Fs(s_{i+1, i})|s_{m+1}\}$ интерпретируется как средний коэффициент фамильного инбридинга субпопуляций $\{s_{i+1, i}\}$ на уровне $i+1$ с соответствующими единицами наблюдения уровня i в популяции s_{m+1} .

Таким образом, при очередном повышении уровня иерархии у s_m (включении s_m в качестве составной части в метапопуляцию s_{m+1}) повышается на единицу уровень иерархии аргумента у Hs_0 и количество сомножителей в факторизации $Hs(s_{m+1})$. При этом добавленный сомножитель равен разности между единицей и коэффициентом фамильного инбридинга $Fs(s_{m+1, m})$ у метапопуляции s_{m+1} (у остальных сомножителей того же типа, также включаемых в s_{m+1} , фигурирует средний коэффициент фамильного инбридинга для субпопуляций соответствующего уровня). Для метапопуляции s_3 разложение на множители можно представить в виде

$$\begin{aligned} Hs(s_3) &= Hs_0(s_3)(1 - Fs(s_{3,2})) \times \\ &\times \prod_{j=1}^{3-2} (1 - E_{s_{j+1}}\{Fs(s_{j+1,j})|s_3\}), \\ Fs(s_{3,2}) &\equiv Vs(x(s_2|s_3))/Hs_0(s_3). \end{aligned}$$

Здесь выражение с символом произведения $\prod$ используется как подсказка для обобщения формулы на произвольный уровень иерархической подразделенности.

Результат 4. Пусть рассматривается метапопуляция s_m с иерархической структурой подразделенности на субпопуляции. Положим, что концентрации заданной фамилии в субпопуляциях $\{s_i\}$ самого низкого уровня иерархии (единицы наблюдения) внутри s_m равны $\{x(s_i|s_m)\}$, вероятность $Hs(s_i|Fs_i)$ встречи в s_i пары индивидуумов с этой фамилией и с какой-нибудь другой (с учетом их порядка) равна

$$x(s_i)(1 - x(s_i))(1 - Fs(s_i)) = Hs_0(s_i)(1 - Fs(s_i)),$$

где $Fs(s_i)$ характеризует нарушение случайности встреч в $\{s_i\}$.

Тогда общая (полная) вероятность $Hs(s_m|Fs_1)$, т.е. вероятность такой встречи в наугад выбранной из s_m субпопуляции s_i при независимости в совокупности концентрации фамилии x и коэффициентов фамильного инбридинга Fs , факторизуется как

$$\begin{aligned} Hs(s_m|Fs_1) &= Hs_0(s_m)(1 - E_{s_1}\{Fs(s_1)|s_m\}) \times \\ &\times (1 - Fs(s_{m,m-1})) \prod_{i=1}^{m-2} (1 - E_{s_{i+1}}\{Fs(s_{i+1,i})|s_m\}), \\ Fs(s_{i+1,i}) &\equiv Vs(x(s_{i+1}|s_i))/Hs_0(s_{i+1}). \end{aligned} \quad (16)$$

Доказательство. Выясненные выше свойства рекуррентного преобразования для $Hs(s_m)$, согласно которым при увеличении m на единицу происходит увеличение количества специфического вида сомножителей в факторизации $Hs(s_m)$, подсказывают (16). Докажем по индукции предполагаемую формулу зависимости Hs от уровня иерархии m метапопуляции s_m . Согласно предыдущим результатам очевидно существует m , при котором $Hs(s_m)$ представимо в требуемом виде. Покажем, что данный вид сохраняется и для $Hs(s_{m+1})$. По рекуррентной формуле (2)

$$\begin{aligned} Hs(s_{m+1}|Fs_1) &\equiv E_{s_m}\{Hs(s_m)|s_{m+1}\} = \\ &= E_{s_m}\left\{Hs_0(s_m)(1 - E_{s_1}\{Fs(s_1)|s_m\}) \times \right. \\ &\times (1 - Fs(s_{m,m-1})) \prod_{i=1}^{m-2} (1 - E_{s_{i+1}}\{Fs(s_{i+1,i})|s_m\})|s_{m+1}\Big\}, \\ Hs_0(s_m) &\equiv x(s_m)(1 - x(s_m)). \end{aligned}$$

При независимости в совокупности x и Fs правая часть данного выражения равна произведению математических ожиданий соответствующих сомножителей по отдельности. Рассмотрим по очереди математические ожидания для сомножителей $Hs_0(s_m) = x(s_m)(1 - x(s_m))$, $1 - E_{s_1}\{Fs(s_1)|s_m\}$, $1 - Fs(s_{m,m-1})$ и для отдельного сомножителя $1 - E_{s_{i+1}}\{Fs(s_{i+1,i})|s_m\}$ произведения $\prod$. Согласно предыдущим результатам

$$\begin{aligned} 1. E_{s_m}\{x(s_m)(1 - x(s_m))|s_{m+1}\} &= \\ &= x(s_{m+1})(1 - x(s_{m+1}))(1 - Fs(s_{m+1,m})). \\ 2. E_{s_m}\{1 - E_{s_1}\{Fs(s_1)|s_m\}|s_{m+1}\} &= \\ &= 1 - E_{s_m}\{E_{s_1}\{Fs(s_1)|s_m\}|s_{m+1}\} = \\ &= 1 - E_{s_1}\{Fs(s_1)|s_{m+1}\}. \\ 3. E_{s_m}\{1 - Fs(s_{m,m-1})|s_{m+1}\} &= \\ &= 1 - E_{s_m}\{Fs(s_{m,m-1})|s_{m+1}\} = \\ &= 1 - E_{s_{i+1}}\{Fs(s_{i+1,i})|s_{m+1}\}_{i=m-1}. \\ 4. E_{s_m}\{1 - E_{s_{i+1}}\{Fs(s_{i+1,i})|s_m\}|s_{m+1}\} &= \\ &= 1 - E_{s_m}\{E_{s_{i+1}}\{Fs(s_{i+1,i})|s_m\}|s_{m+1}\} = \\ &= 1 - E_{s_{i+1}}\{Fs(s_{i+1,i})|s_{m+1}\}. \end{aligned}$$

Поясним эти равенства.

1. Правая часть п. 1 следует из формулы (6) эффекта Валунда с коэффициентом фамильного

инбридинга при соответствующих уровнях иерархии. В доказываемом выражении для $Hs(s_{m+1})$ первый сомножитель $x(s_{m+1})(1 - x(s_{m+1}))$ в п. 1 входит в $Hs(s_{m+1})$ как $Hs_0(s_{m+1}|Fs_1)$ до символа произведения $\prod$. Полученный в п. 1 второй сомножитель $1 - Fs(s_{m+1}, m)$ также входит в $Hs(s_{m+1})$ до символа произведения $\prod$.

2. В п. 2 сомножитель $1 - E_{s_m}\{E_{s_1}\{Fs(s_1)|s_m\}|s_{m+1}\}$, очевидно, равен $1 - E_{s_1}\{Fs(s_1)|s_{m+1}\}$ согласно (4), что соответствует его значению для разложения Hs в метапопуляции s_{m+1} , а не s_m . Он сохраняет свое место до символа произведения $\prod$.

3. Сомножитель $E_{s_m}\{1 - Fs(s_{m,m-1})|s_{m+1}\}$ равен $1 - E_{s_{i+1}}\{Fs(s_{i+1}, i)|s_{m+1}\}|_{i=m-1}$, т.е. его можно переместить под знак произведения $\prod$, где он соответствует значению i , равному $m-1$. В результате верхний предел произведения увеличится на единицу и станет равным $m-1$ вместо $m-2$.

4. Каждый сомножитель произведения $\prod$ в $Hs(s_m)$ представим в виде $1 - E_{s_{i+1}}\{Fs(s_{i+1}, i)|s_m\}$ и при переходе к s_{m+1} преобразуется как

$$\begin{aligned} E_{s_m}\{1 - E_{s_{i+1}}\{Fs(s_{i+1}, i)|s_m\}|s_{m+1}\} &= \\ = 1 - E_{s_m}\{E_{s_{i+1}}\{Fs(s_{i+1}, i)|s_m\}|s_{m+1}\} &= \\ = E_{s_{i+1}}\{Fs(s_{i+1}, i)|s_{m+1}\}. \end{aligned}$$

Подстановка полученных выражений в $E_{s_m}\{Hs(s_m)|s_{m+1}\}$ при сделанном предположении относительно факторизации $Hs(s_m)$ дает

$$\begin{aligned} Hs(s_{m+1}|Fs_1) &\equiv E_{s_m}\{Hs(s_m)\} = \\ &= Hs_0(s_{m+1})(1 - E_{s_1}\{Fs(s_1)|s_{m+1}\}) \times \\ &\times (1 - Fs(s_{m,m-1})) \prod_{i=1}^{m-1} (1 - E_{s_{i+1}}\{Fs(s_{i+1}, i)|s_{m+1}\}), \end{aligned}$$

что совпадает с (15) для $Hs(s_{m+1})$ и доказывает предположение индукции. $\blacktriangleleft$

Доказанный результат в сущности является более детальной формой выводов С. Райта [12] при анализе иерархически подразделенных популяций.

Ремарка 5. В популяционной генетике известна [12] следующая формула С. Райта для концентрации гетерозигот $H(s)$ в подразделенной метапопуляции s :

$$H(s) = H_0(s)(1 - F_{IS}(s))(1 - F_{ST}(s)),$$

где $F_{IS}(s)$ и $F_{ST}(s)$ — неслучайный и случайный коэффициент инбридинга соответственно. Аналогично у нас отдельные составляющие выражения (15) для $Hs(s_m)$ имеют сходную интерпретацию. Среднюю фамилную инбредность субпопуляций первого уровня иерархии в s_m характеризует $E_{s_1}\{Fs(s_1)|s_m\}$, а

подразделенности метапопуляции отвечает произведение $(1 - Fs(s_{m,m-1})) \prod (\dots)$, отражающее фамилную дивергенцию субпопуляций в s_m . Компонент $E_{s_1}\{Fs(s_1)|s_m\}$ соответствует неслучайному коэффициенту инбридинга $F_{IS}(s_m)$ в популяционной генетике, а произведение $(1 - Fs(s_{m,m-1})) \prod (\dots)$ дает $1 - F_{ST}(s_m)$, где F_{ST} — аналог случайного коэффициента инбридинга. В терминах общего $F_{IT}(s_m)$, случайного и неслучайного коэффициентов фамилного инбридинга для выражения общей (полной) вероятности встреч $Hs(s_m)$ в метапопуляции s_m имеем

$$\begin{aligned} Hs(s_m) &= Hs_0(s_m)(1 - F_{IS}(s_m)) \times \\ &\times (1 - F_{ST}(s_m)) = Hs_0(s_m)(1 - F_{IT}(s_m)), \\ F_{IS}(s_m) &\equiv E_{s_1}\{Fs(s_1)|s_m\}, \\ 1 - F_{ST}(s_m) &\equiv (1 - Fs(s_{m,m-1})) \times \\ &\times \prod_{i=1}^{m-2} (1 - E_{s_{i+1}}\{Fs(s_{i+1}, i)|s_m\}). \end{aligned}$$

В анализируемых данных может отсутствовать информация о наличии преград для встреч в субпопуляциях, служащих единицами наблюдения. В связи с этим желательно проанализировать последствия реального существования препятствий встречам в единицах наблюдения, когда принято решение об их отсутствии. Рассмотрим данную ситуацию в предположении независимости коэффициентов фамилного инбридинга и концентраций фамилии.

Следствие 6. Пусть в условиях предыдущего результата в метапопуляции s_m с t уровнями иерархии и нулевым коэффициентом неслучайного фамилного инбридинга в качестве единицы наблюдения служат субпопуляции $\{s_n\}$ уровня n ($1 < n < t$), в которых концентрации интересующей фамилии равны $\{x(s_n)\}$, а вероятности $\{Hs(s_n)\}$ рассматриваемой неизонимной встречи неизвестны. Положим, что в качестве ее значения используется вероятность случайных столкновений индивидумов $Hs_0(s_n)$.

Тогда $Hs(s_m)$ будет меньше в $\prod_{i=1}^{n-1} (1 - E_{s_{i+1}}\{Fs(s_{i+1}, i)|s_m\})$ раз значения $Hs(s_{m,n})$, где $Fs(s_{i+1}, i)$ — неизвестные коэффициенты фамилного инбридинга в субпопуляциях.

Доказательство. Напомним, что согласно только что установленному результату полная вероятность встреч во всей метапопуляции s_m находится как

$$\begin{aligned} Hs(s_m) &= Hs_0(s_m)(1 - Fs(s_{m,m-1})) \times \\ &\times \prod_{i=1}^{m-2} (1 - E_{s_{i+1}}\{Fs(s_{i+1}, i)|s_m\}). \end{aligned}$$

Здесь подразумевается, что единицей наблюдения является субпопуляция s_1 первого уровня иерархии с

нулевым коэффициентом неслучайного фамильного инбридинга, и индекс i произведения начинается с единицы. Когда уровень единицы наблюдения равен n , то начальное значение i равно n . Соответственно при факторизации $Hs(s_{m,n})$ в отличие от $Hs(s_m)$ вместо $\prod_{i=1}^{m-1}$ фигурирует $\prod_{i=n}^{m-1}$ с другим начальным значением индекса i . Тем самым здесь опущено произведение $\prod_{i=1}^{n-1}$. По условию коэффициенты фамильного инбридинга у субпопуляций уровня иерархии n (у единиц наблюдения) приняты равными нулю, что возможно только когда такие коэффициенты равны нулю в каждой из субпопуляций с уровнями иерархии до n . В этом случае соответствующие им множители равны единице, и их можно вычеркнуть в формуле факторизации $Hs(s_m)$. Вычеркнутые множители при ненулевых в действительности коэффициентах фамильного инбридинга дают произведение $\prod_{i=1}^{n-1} (1 - E_{s_{i+1}}\{Fs(s_{i+1,i})|s_m\})$, показывающее во сколько раз изменилось значение $Hs(s_m)$ из-за пренебрежения возможным отклонением $Hs(s_n)$ от $Hs_0(s_n)$.

Таким образом получаем, что игнорирование инбредности субпопуляций первого уровня иерархии и подразделенности субпопуляций более высоких уровней до уровня единицы наблюдения приводит к завышению Hs согласно доказываемому выражению. ◀

Как известно, коэффициент инбридинга в популяциях человека обычно порядка сотых долей или тысячных. Если коэффициенты фамильного инбридинга также малы, то можно пренебречь их произведениями. В результате получим простую аппроксимацию для коэффициентов фамильного инбридинга в таких метапопуляциях.

Результат 7. В иерархически подразделенной метапопуляции s при малых коэффициентах фамильного инбридинга в составляющих ее субпопуляциях общая вероятность (15) случайной встречи $Hs(s_m|Fs)$ индивидуумов с данной фамилией и какой-нибудь другой с учетом их порядка в наугад выбранной из s_m единице наблюдения s_1 аппроксимируется следующим выражением:

$$\begin{aligned} Hs(s_m|Fs_1) &\sim x(s_m)(1 - x(s_m)) \times \\ &\times \left(1 - E_{s_1}\{Fs(s_1)|s_m\} - Fs(s_{m,m-1}) - \right. \\ &\quad \left. - \sum_{i=1}^{m-2} E_{s_{i+1}}\{Fs(s_{i+1,i})|s_m\} \right) \equiv \\ &\equiv x(s_m)(1 - x(s_m))(1 - Fs(s_m)), \\ Fs(s_m) &\equiv E_{s_1}\{Fs(s_1)|s_m\} + Fs(s_{m,m-1}) + \\ &+ \sum_{i=1}^{m-2} E_{s_{i+1}}\{Fs(s_{i+1,i})|s_m\} \geq 0, \end{aligned}$$

где $Fs(s_m)$ — коэффициент фамильного инбридинга субпопуляции уровня m , неотрицательные коэффициенты $\{Fs(s_{i+1,i})\}$ характеризуют подразделенность субпопуляций, а $E_{s_1}\{Fs(s_1)|s_m\}$ — среднюю инбредность субпопуляций, единиц наблюдения.

Таким образом, приближенно коэффициент инбридинга иерархически подразделенной метапопуляции равен сумме средних коэффициентов инбридинга субпопуляций на отдельных уровнях иерархии.

Доказательство. По определению коэффициента фамильного инбридинга $Hs(s_m|Fs_1) = Hs_0(s_m)(1 - Fs(s_m))$. С другой стороны, согласно (15)

$$\begin{aligned} Hs(s_m|Fs_1) &= Hs_0(s_m)(1 - E_{s_1}\{Fs(s_1)|s_m\}) \times \\ &\times (1 - Fs(s_{m,m-1})) \prod_{i=1}^{m-2} (1 - E_{s_{i+1}}\{Fs(s_{i+1,i})|s_m\}). \end{aligned}$$

Эти два выражения для $Hs(s_m)$ должны быть равны друг другу, т.е.

$$\begin{aligned} 1 - Fs(s_m) &= (1 - E_{s_1}\{Fs(s_1)|s_m\})(1 - Fs(s_{m,m-1})) \times \\ &\times \prod_{i=1}^{m-2} (1 - E_{s_{i+1}}\{Fs(s_{i+1,i})|s_m\}). \end{aligned}$$

При раскрытии скобок в правой части данного равенства в случае малых по условию коэффициентов фамильного инбридинга субпопуляций можно пренебречь их произведениями. Очевидно, в результате этого остающиеся члены дают требуемую аппроксимацию для $Fs(s_m)$, которая лаконично выражается с помощью единого суммарного коэффициента инбридинга. ◀

Из доказанного также ясно, что при $Fs(s_m) > 0$ вероятность Hs меньше значения, соответствующего случайным встречам. Отклонение увеличивается с каждым повышением уровня иерархии на величину разности единицы и Fs на соответствующем уровне иерархии.

ЗАКЛЮЧЕНИЕ

Проведенный анализ фамильной структуры мотивирован имеющимися параллелями в передаче фамилии и генов от родителя детям. На популяционном уровне передача фамилий и генов описывается одним и тем же процессом выборочного дрейфа, но отличающимся интенсивностью [8, 9]. На этом и на других аргументах основаны сотни статей в базе данных <https://pubmed.ncbi.nlm.nih.gov> по изучению фамильной структуры и ее связи с генетической. Указанные источники и приведенные ниже в списке литературы для первоначального ознакомления позволяют получить достаточно подробное представление о важности фамильных данных в популяционной генетике.

Существенной чертой обитания современного и первобытного человека является подразделен-

ность популяций, поддерживающаяся в течение длительного времени. При этом иерархическая подразделенность в силу территориального деления, происхождения, лингвистической классификации и т.д. является типичной. Более всего иерархическая структура характерна для семейных данных, часто представленных в соответствии с иерархическим административно-территориальным делением.

Важным показателем неоднородности иерархической системы является дивергенция составляющих ее субпопуляций. Настоящая серия статей автора [10, 11] посвящена анализу показателей дивергенции, присущих самой иерархической структуре и вычислительно независимых от характера микроэволюционного процесса, изолированности субпопуляций или существования миграций между ними.

Процесс выборочного дрейфа приводит к генетической и семейной дивергенции первоначально сходных популяций. Поэтому основное внимание мы уделяем показателям дивергенции субпопуляций в подразделенной метапопуляции. Таким показателем является нарушение случайного характера встреч (столкновений) индивидумов в метапопуляции и ее подразделениях. Формально вероятности встреч аналогичны концентрациям гетерозигот, поскольку в обоих случаях речь идет о вероятностях комбинаций пар (семей при встрече и генов в диплоидных генотипах).

В данной статье рассматривается семейный аналог коэффициента инбридинга, важной характеристики в популяционной генетике. В изложении результатов сделан акцент на одной из сторон коэффициента семейного инбридинга, а именно на нем как характеристике нарушения случайности встреч индивидумов в метапопуляции. В контексте популяционной генетики это свойство аналогично отклонению концентрации гетерозигот в метапопуляции от значения, соответствующего закону Харди–Вайнберга. В указанном отношении получено специфическое для иерархической подразделенности обобщение эффекта Валунда в виде разложения коэффициента семейного инбридинга по компонентам, соответствующим отдельным уровням иерархии.

Конкретной целью всей серии работ автора по теоретическому изучению семейной структуры является адаптация и обобщение традиционных популяционно-генетических подходов к анализу и описанию свойств семейной структуры в иерархически подразделенной метапопуляции. Полученные результаты применимы и в соответствующей области популяционной генетики при минимальных коррекциях, и в Приложении, которое можно читать независимо, сформулирова-

ны популяционно-генетические аналоги найденных обобщений.

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.

ПРИЛОЖЕНИЕ

Как упоминалось ранее, сформулированные выше результаты для распределения концентрации фамилии дают представление о статистических свойствах числовых характеристик иерархически подразделенной на группы совокупности объектов любой природы и не зависят от закономерностей формирования групп. Например, они приложимы к биологическим популяциям, характеризующимся количественными признаками и иерархически подразделенным согласно территориальной принадлежности, происхождению, родственным связям и т.д. Близкой к изучаемой картине должна быть ситуация с концентрациями x аллеля аутосомного локуса в иерархически подразделенной метапопуляции, которые подобны данным по концентрации фамилии при анализе структуры изонимии у человека.

Настоящее приложение содержит приведенные без доказательств аналоги полученных в данной серии работ результатов по анализу семейной структуры применительно к материалам популяционной генетики. Объектом изучения в популяционной генетике является генетическая структура популяции (концентрации аллелей и генотипов). Типичные популяционно-генетические материалы охватывают ряд субпопуляций, образующих метапопуляцию. Ее генетическая структура определяется структурами субпопуляций, и важно исследовать связь между ними в типичных ситуациях.

Такой характерной для материалов по человеку ситуацией является иерархическая подразделенность метапопуляции. Формулируемые ниже свойства иерархически подразделенной метапопуляции в зависимости от генетической структуры субпопуляций относятся к ее текущему генетическому состоянию и вычислительно не зависят от существования миграций между субпопуляциями или их полной изолированности, они не определяются происхождением субпопуляций от общей родоначальной группы или их составом (вплоть до произвольной механической смеси) и моделями микроэволюции, которые влияют на результаты количественно. Наше внимание фокусируется на характеристиках дивергенции субпопуляций, которые определяют такую важную черту метапопуляции, как ее коэффициент инбридинга.

Читателям, чей интерес лежит только на генетических аспектах иерархической структуры ме-

тапопуляций, можно посоветовать сразу обратиться к данному приложению и просматривать соответствующие (близкие) доказательства в предшествующем тексте при желании глубже вникнуть в рассматриваемые проблемы. Чтобы приложение можно было читать независимо от предыдущего изложения, напомним используемые обозначения и приведем их модификации применительно к популяционно-генетическому контексту приложения.

Уровень иерархии m всей метапопуляции s_m считаем самым высоким, а субпопуляции первого уровня $\{s_i\}$ рассматриваем как неподразделенные с выполнением закона Харди–Вайнберга или с некоторым коэффициентом инбридинга, вызванным какими-либо причинами. Обозначения популяционно-генетических характеристик (дисперсии V , концентрации гетерозигот H , коэффициента инбридинга F) отличаются от использованных выше V_s , H_s , F_s для их фамильных аналогов отсутствием окончания s . Векторы набраны полужирным шрифтом, Символ s_i относится к субпопуляции на уровне (ранге) иерархии i , $\{x(s_i)\}$ дает концентрации носителей рассматриваемой характеристики, например, концентрации конкретного аллеля, гетерозиготы и т.д. в субпопуляции s_i (далее для краткости просто концентрации), $x(s_i|s_m)$, обозначает случайное значение рассматриваемой концентрации x в субпопуляции s_i , получаемое при выборе наугад s_i из содержащей ее (наряду с другими) субпопуляции s_m более высокого уровня m ($i < m$). Иначе говоря, $\{x(s_i|s_m)\}$ случайная величина, принимающая значения $\{x(s_i)\}$, равные концентрациям рассматриваемого далее какого-нибудь аллеля в субпопуляциях уровня i , входящих в s_m (при условии выбора наугад s_i из s_m). Ее средняя величина и дисперсия обозначаются как $x(s_m)$ и $V(x(s_i|s_m))$ соответственно.

Рассмотрим свойства распределения концентраций в подразделенной метапопуляции.

Замечание А1. Пусть анализируется подразделенная популяция s_m с размером $N(s_m)$ и с численностью $T_N(s_m)$ носителей рассматриваемой характеристики T (скажем, интересующего аллеля, гетерозиготы и т.п.) в s_m . Тогда концентрация $T(s_m) \equiv T_N(s_m)/N(s_m)$ данных носителей в популяции s_m , произвольно разбитой на группы $\{s_i\}$ с численностями $\{N(s_i)\}$, находится как

$$T(s_m) = \sum_{s_i \in s_j} T(s_i) \frac{N(s_i)}{N(s_m)} = \sum_{s_i \in s_j} T(s_i) Pr(s_i|s_m) = \\ = E_{s_i} \{T(s_i)|s_m\}, \quad Pr(s_i|s_m) \equiv \frac{N(s_i)}{N(s_m)},$$

т.е. концентрация $T(s_m)$ во всей популяции s_m равна среднему значению (математическому ожиданию) случайной величины $T(s_i|s_m)$, концентрации T в s_i

при выборе наугад s_i из s_m с вероятностью $Pr(s_i|s_m) \equiv N(s_i)/N(s_m)$. Здесь E является символом операции получения математического ожидания, нижний индекс у E указывает на переменную, служащую для усреднения.

Значение концентрации $T(s_m)$ не зависит от разбиения s_m на $\{s_i\}$, и $T(s_m)$ не меняется при изменении разбиения.

Типичным признаком в популяционной генетике служит концентрация x какого-либо аллеля. Как и для характеристики T , среднее значение $x(s_m)$ для концентрации аллеля во всей популяции s_m равно средней величине для его концентраций в составляющих s_m субпопуляциях. Скажем, если s_m разбита на $\{s_i\}$, то $x(s_m) = E_{s_i} \{x(s_i)|s_m\}$. То же самое можно сказать о разбиении s_m на $\{s_j\}$. Математическое ожидание случайной величины $x(s_i|s_m)$, обозначаемое как $x(s_m)$, равняется концентрации x в s_m : $x(s_m) = E_{s_i} \{x(s_i)|s_m\}$.

Данное свойство является частным случаем правила полного математического ожидания, согласно которому у нас концентрация $x(s)$ соответствующих носителей (аллеля, гетерозиготы и т.д.) в разбитой на субпопуляции $\{s_i\}$ метапопуляции s является средней величиной для концентраций в отдельных субпопуляциях (концентраций при условиях рассмотрения субпопуляций по одной), каково бы ни было данное разбиение. Для случайной величины x формула полного математического ожидания означает, что

$$x(s) \equiv E_{s_i} \{x(s_i)|s\} = E_{s_i} \{x(s_i)|s_i\}, \\ i = 1, 2, 3, \dots, \quad x(s_i) \equiv E_{s_i} \{x(s_i)|s_i\},$$

т.е. среднее значение концентрации x в некоторой иерархически подразделенной метапопуляции s является средней величиной для концентраций в разбивающих s подгруппах $\{s_i\}$. Приведенная формула показывает в явном виде равенство концентрации $x(s)$ рассматриваемого аллеля во всей метапопуляции s средней величине для его концентраций в отдельных частях $\{x(s_i)\}$ разбиения s одного и того же уровня иерархии i , какими бы ни были эти части и уровень.

Сформулированные результаты справедливы не только для концентрации аллеля, но и для концентрации H гетерозигот. Концентрация $H(s_i)$ в субпопуляции s_i является средним взвешенным значением $E_{s_i} \{H(s_i)|s_i\}$ концентраций гетерозигот в отдельных субпопуляциях-единицах наблюдения $\{s_1\}$, содержащихся в s_i , причем веса можно понимать как вероятности случайного выбора соответствующих субпопуляций $\{s_1\}$ из s_i . Другими словами, $H(s_i)$ находится как ожидаемое значение $H(s_1)$ при выборе наугад s_1 из s_i (т.е. средней вели-

чине для результатов, получаемых при выборе значений $H(s_1)$ с учетом вероятностей выбора s_1 .

Результат А2 (эффект Валунда). Пусть дана популяция s , разбитая на субпопуляции $\{s_1\}$, где концентрации рассматриваемого аллеля равны $\{x(s_1)\}$, а концентрации гетерозигот с этим аллелем с учетом порядка аллелей находятся как $\{x(s_1)(1 - x(s_1))\}$ по закону Харди–Вайнберга.

Тогда общая (полная) концентрация гетерозигот $H(s)$ с данным аллелем в наугад выбранной из s субпопуляции s_1 имеет вид

$$\begin{aligned} H(s) &\equiv E_{s_1} \{H(s_1)|s\} = \\ &= \sum_{s_1} x(s_1)(1 - x(s_1))Pr(s_1|s) = \\ &= x(s)(1 - x(s)) - V_{\text{betw}}(x(s_1|s)), \\ V_{\text{betw}}(x(s_1|s)) &= V(x(s_1|s)) = \\ &= \sum_{s_1} (x(s_1) - x(s))^2 Pr(s_1|s). \end{aligned}$$

Здесь $V_{\text{betw}}(x(s_1|s))$ обозначает межгрупповую дисперсию значений концентрации $\{x(s_1|s)\}$ интересующего аллеля в субпопуляциях (группах) $\{s_1\}$, составляющих популяцию s ; $Pr(s_1|s)$ вероятность выбора наугад субпопуляции s_1 из популяции s .

Иным образом общую (полную) концентрацию $H(s)$ можно представить как

$$\begin{aligned} H(s) &= x(s)(1 - x(s)) \times \\ &\times \left(1 - \frac{V_{\text{betw}}(x(s_1|s))}{x(s)(1 - x(s))}\right) \equiv H_0(s)(1 - F(s)), \\ F(s) &\equiv \frac{V_{\text{betw}}(x(s_1|s))}{x(s)(1 - x(s))} = \frac{V_{\text{betw}}(x(s_1|s))}{H_0(s)} \geq 0, \\ H_0(s) &\equiv x(s)(1 - x(s)), \end{aligned}$$

где $F(s)$ является коэффициентом инбридинга С. Райта, $H_0(s)$ обозначает концентрацию рассматриваемых гетерозигот в метапопуляции s при выполнении в ней закона Харди–Вайнберга.

Дисперсия значений $\{x(s_1|s)\}$ характеризует дивергенцию концентраций аллеля между субпопуляциями $\{s_1\}$. Чем она больше, тем больше дивергенция концентрации аллеля между субпопуляциями и тем меньше общая концентрация гетерозигот в подразделенной популяции. Коэффициент инбридинга также характеризует генетическую дивергенцию субпопуляций, и далее мы интерпретируем его как показатель отклонения концентрации гетерозигот от значения, соответствующего закону Харди–Вайнберга.

Обратим внимание, что результат Валунда обобщается на случай, когда в подразделенной популяции нарушаются соотношения Харди–Вайнберга в составляющих ее субпопуляциях $\{s_1\}$, где $H(s_1) \equiv x(s_1)(1 - x(s_1))(1 - F(s_1))$ с учетом порядка аллелей в гетерозиготах.

Следствие А3. Пусть дана популяция s_2 , разбитая на субпопуляции $\{s_1\}$ с концентрациями $\{x(s_1)\}$ рассматриваемого аллеля. Предположим, что в s_1 концентрацию $H(s_1)$ гетерозигот с данным аллелем с учетом порядка аллелей возможно описать с помощью коэффициента инбридинга $F(s_1)$ как

$$\begin{aligned} H(s_1) &= x(s_1)(1 - x(s_1))(1 - F(s_1)) = \\ &= H_0(s_1)(1 - F(s_1)), \quad s_2 = \{s_1\}, \end{aligned}$$

где $H_0(s_1) = x(s_1)(1 - x(s_1))$ обозначает концентрацию гетерозигот при выполнении закона Харди–Вайнберга в субпопуляции s_1 .

Тогда общая (полная) концентрация рассматриваемых гетерозигот $H(s_2|F_{s_1})$ в наугад выбранной из s_2 субпопуляции s_1 при независимости концентрации x рассматриваемого аллеля и коэффициента инбридинга $F(s_1)$ в субпопуляциях $\{s_1\}$ имеет вид

$$\begin{aligned} Hs(s_2|F_{s_1}) &= \\ &= (x(s_2)(1 - x(s_2)) - V_{s_{\text{betw}}}(x(s_1|s_2))) \times \\ &\times (1 - E_{s_1} \{F(s_1)|s_2\}). \end{aligned}$$

Здесь $V_{s_{\text{betw}}}(x(s_1|s_2))$ обозначает межгрупповую дисперсию концентраций $\{x(s_1|s_2)\}$ рассматриваемого аллеля в субпопуляциях (группах) $\{s_1\}$, составляющих популяцию s_2 .

Иным образом общая (полная) концентрация $H(s_2|F_{s_1})$ представима как

$$\begin{aligned} H(s_2|F_{s_1}) &= H_0(s_2) \left(1 - \frac{V_{s_{\text{betw}}}(x(s_1|s_2))}{x(s_2)(1 - x(s_2))}\right) \times \\ &\times (1 - E_{s_1} \{F(s_1)|s_2\}) = Hs_0(s_2) \times \\ &\times (1 - F(x(s_2)))(1 - E_{s_1} \{F(s_1)|s_2\}), \end{aligned}$$

где $F(x(s_2)) \equiv V_{s_{\text{betw}}}(x(s_1|s_2))/Hs_0(s_2) \geq 0$ является случайным коэффициентом инбридинга F_{ST} , а средний коэффициент инбридинга $E_{s_1} \{F(s_1)|s_2\}$ для субпопуляций $\{s_1\}$ в s_2 является коэффициентом неслучайного инбридинга F_{IS} согласно результатам С. Райта.

Мы видим, что дисперсия концентраций аллеля в субпопуляциях метапопуляции играет важную роль для величины коэффициента инбридинга. Поэтому представляет интерес выяснение структуры дисперсии концентрации аллеля в иерархически подразделенной метапопуляции. При этом для наших целей анализа структуры дисперсии существенную роль играет правило сложения дисперсии, согласно которому в совокупности из нескольких групп произвольных объектов с каким-либо числовым признаком x общая (полная) дисперсия x во всей совокупности равна сумме межгрупповой дисперсии (дисперсии распределения средних значений x в группах) и средней внутригрупповой дисперсии (средней дисперсии x внутри групп).

Понятно, что в соответствии с приведенной формулировкой данное статистическое правило

не зависит от природы объектов и принципов их объединения в группы, которые могут быть произвольными. Правило выполняется не только для дисперсий, но и при совместном изучении нескольких признаков для их матриц ковариаций. В случае популяционно-генетического анализа подразделенной популяции данное правило принимает следующий вид.

Замечание А4 (правило сложения дисперсий). Пусть метапопуляция s с тремя уровнями иерархии разбита на субпопуляции $\{s_2\}$, каждая из которых включает неподразделенные группы $\{s_1\}$, являющиеся единицами наблюдения с концентрациями $\{x(s_1)\}$ рассматриваемого аллеля в них.

Тогда полная дисперсия $V_{\text{tot}}(x(s_1|s))$ распределения концентрации аллеля по несгруппированным единицам наблюдения $\{s_1\}$ во всей метапопуляции s равна сумме межгрупповой дисперсии $V_{\text{betw}}(x(s_2|s))$, характеризующей дивергенцию концентраций аллеля $\{x(s_2)\}$ между субпопуляциями $\{s_2\}$, и средней внутригрупповой дисперсии $W(x(s_1|s_2)|s) \equiv E_{s_2}\{V_{\text{in}}(x(s_1|s_2))|s\}$, характеризующей среднюю дивергенцию концентраций $x(s_1)$ субпопуляций $\{s_1\}$ внутри $\{s_2\}$:

$$V_{\text{tot}}(x(s_1|s)) = V_{\text{betw}}(x(s_2|s)) + E_{s_2}\{V_{\text{in}}(x(s_1|s_2))|s\}.$$

Результат анализа свойств дисперсии для метапопуляции с тремя уровнями иерархии обобщается на ситуацию произвольного количества уровней.

Следствие А5. В иерархически подразделенной метапопуляции s_m с t уровнями иерархии полная дисперсия $V(x(s_1|s_m))$ распределения концентрации аллеля по субпопуляциям первого уровня допускает разложение по компонентам на уровнях иерархии в виде

$$\begin{aligned} V(x(s_1|s_m)) &= \sum_{i=1}^{m-1} E_{s_{i+1}}\{V(x(s_i|s_{i+1}))|s_m\} = \\ &= \sum_{i=1}^{m-1} (V(x(s_i|s_m)) - V(x(s_{i+1}|s_m))). \end{aligned}$$

Использование данного разложения дисперсии позволяет найти разложение ожидаемой концентрации гетерозигот по компонентам, соответствующим отдельным уровням иерархии.

Результат А6. Рассмотрим иерархически подразделенную метапопуляцию s_m с уровнями иерархии $i = 1, 2, \dots, t$. Положим, что в субпопуляциях $\{s_1\}$ концентрации рассматриваемого аллеля равны $\{x(s_1)\}$, а концентрации несущих его гетерозигот удовлетворяют закону Харди–Вайнберга и с учетом порядка аллелей находятся как $H(s_1) = H_0(s_1) \equiv x(s_1)(1 - x(s_1))$.

Если из s_m наугад выбрана субпопуляция s_1 первого уровня иерархии, то ожидаемая концентрация данных гетерозигот $H(s_m)$ в s_1 (общая, или полная концентрация) допускает следующее разложение:

$$\begin{aligned} H(s_m) &\equiv E_{s_1}\{H(s_1)|s_m\} = x(s_m)(1 - x(s_m)) - \\ &- V_{\text{betw}}(x(s_1|s_m)) = H_0(s_m) - \\ &- \sum_{i=1}^{m-1} E_{s_{i+1}}\{V_{\text{in}}(x(s_i|s_{i+1}))|s_m\}. \end{aligned}$$

Таким образом, ожидаемая концентрация гетерозигот $H(s_m)$ равна разности между значением $H_0(s_m) \equiv x(s_m)(1 - x(s_m))$, соответствующим выполнению закона Харди–Вайнберга во всей метапопуляции s_m , и разложением полной дисперсии $V(x(s_1|s_m))$ на средние внутригрупповые дисперсии $E_{s_{i+1}}\{V_{\text{in}}(x(s_i|s_{i+1}))|s_m\} = V(x(s_i|s_m)) - V(x(s_{i+1}|s_m))$, соответствующие отдельным уровням иерархии i .

На разложении дисперсии и концентрации гетерозигот основывается и разложение коэффициента инбридинга в соответствии с уровнями иерархии.

Результат А7. Пусть в иерархически подразделенной метапопуляции s_m с t уровнями иерархии в качестве единиц наблюдения служат субпопуляции $\{s_1\}$ первого уровня. Положим, что концентрации интересующего аллеля в $\{s_1\}$ равны $\{x(s_1)\}$, а концентрации его несущих гетерозигот с учетом порядка аллелей находятся в $\{s_1\}$ как $\{H(s_1) = H_0(s_1) \equiv x(s_1)(1 - x(s_1))\}$.

Тогда разложение коэффициента инбридинга $F(s_m)$ в наугад выбранной из s_m субпопуляции s_1 можно записать как

$$\begin{aligned} F(s_m) &= \sum_{i=1}^{m-1} \frac{V(x(s_i|s_m)) - V(x(s_{i+1}|s_m))}{H_0(s_m)} = \\ &= \sum_{i=1}^{m-1} \frac{E_{s_{i+1}}\{V_{\text{in}}(x(s_i|s_{i+1}))|s_m\}}{H_0(s_m)} = \sum_{i=1}^{m-1} \Delta_i F(s_m), \\ \Delta_i F(s_m) &\equiv \frac{V(x(s_i|s_m)) - V(x(s_{i+1}|s_m))}{H_0(s_m)} = \\ &= \frac{E_{s_{i+1}}\{V_{\text{in}}(x(s_i|s_{i+1}))|s_m\}}{H_0(s_m)} \equiv F(s_{m,i}) - F(s_{m,i+1}) \geq 0. \end{aligned}$$

Общая концентрация $H(s_m)$ рассматриваемых гетерозигот в наугад выбранной из s_m субпопуляции s_1 представима в виде

$$\begin{aligned} H(s_m) &= H_0(s_m) \times \\ &\times \left(1 - \sum_{i=1}^{m-1} \frac{V(x(s_i|s_m)) - V(x(s_{i+1}|s_m))}{H_0(s_m)} \right) = \\ &= H_0(s_m) \left(1 - \sum_{i=1}^{m-1} \frac{E_{s_{i+1}}\{V_{\text{in}}(x(s_i|s_{i+1}))|s_m\}}{H_0(s_m)} \right) = \\ &= H_0(s_m) \left(1 - \sum_{i=1}^{m-1} \Delta_i F(s_m) \right). \end{aligned}$$

Здесь

$$\begin{aligned}\Delta_i F(s_m) &\equiv F(s_{m,i}) - F(s_{m,i+1}) = \\ &= E_{s_{i+1}} \{V_{in}(x(s_i|s_{i+1})) | s_m\} / H_0(s_m), \\ F(s_m) &\equiv V(x(s_1|s_m)) / H_0(s_m), \\ F(s_{m,i}) &\equiv V_{in}(x(s_i|s_m)) / H_0(s_m), \\ H(s_m) &\equiv E_{s_1} \{H(s_1) | s_m\}, \\ H_0(s_m) &= x(s_m)(1 - x(s_m)),\end{aligned}$$

$F(s_m)$ и $F(s_{m,i})$ обозначают коэффициенты инбридинга в метапопуляции s_m и в $s_{m,i}$, когда в качестве единиц наблюдения в s_m служат субпопуляции $\{s_1\}$ и $\{s_i\}$ с уровнями иерархии 1 и i при условии, что концентрации рассматриваемых гетерозигот в них равны $\{H_0(s_1)\}$ и $\{H_0(s_i)\}$ соответственно; $H(s_m)$ общая концентрация рассматриваемых гетерозигот в s_m .

Обратим внимание на то, что когда в качестве единиц наблюдения используются субпопуляции $\{s_n\}$ с уровнем иерархии n , то в общем случае выбор значений $x(s_n)(1 - x(s_n))$ для концентрации гетерозигот в $\{s_n\}$ по закону Харди–Вайнберга не обоснован. Рассмотрим погрешности выражений для концентрации гетерозигот и коэффициента инбридинга, вносимые использованием вместо начальных данных (концентрации гетерозигот с интересующим аллелем в единицах наблюдения с нулевым коэффициентом инбридинга) агрегированных сведений для субпопуляций более высокого ранга.

Следствие А8. Пусть в иерархически подразделенной метапопуляции s_m с t уровнями иерархии в качестве единиц наблюдения вместо субпопуляций первого уровня $\{s_1\}$ с концентрациями интересующего аллеля $\{x(s_1)\}$ и случайным скрещиванием рассматриваются субпопуляции уровня n ($1 < n < t - 1$) с концентрациями $\{x(s_n)\}$. Предположим, что при отсутствии или игнорировании сведений о нарушении в них закона Харди–Вайнберга принимается решение об отсутствии нарушений в каждой из субпопуляций $\{s_n\}$, т.е. концентрация $H(s_{m,n})$ в s_m несущих интересующий аллель гетерозигот с учетом порядка аллелей (в наугад выбранной из s_m субпопуляции s_n) находится как

$$\begin{aligned}H(s_{m,n}) &\equiv E_{s_n} \{H(s_n) | s_m\} = E_{s_n} \{H_0(s_n) | s_m\}, \\ H_0(s_n)(s_m) &\equiv x(s_n)(1 - x(s_n)),\end{aligned}$$

где $s_{m,n}$ обозначает метапопуляцию s_m с единицами наблюдения $\{s_n\}$.

Тогда коэффициент инбридинга $F(s_{m,n})$, определяемый из условия $H(s_{m,n}) = H_0(s_m)(1 - F(s_{m,n}))$ можно разложить по компонентам на уровнях иерархии как

$$\begin{aligned}F(s_{m,n}) &= \sum_{i=n}^{m-1} \frac{V(x(s_i|s_m)) - V(x(s_{i+1}|s_m))}{x(s_m)(1 - x(s_m))} = \\ &= \sum_{i=n}^{m-1} \frac{E_{s_{i+1}} \{V(x(s_i|s_{i+1})) | s_m\}}{x(s_m)(1 - x(s_m))} \equiv \sum_{i=n}^{m-1} \Delta_i F(s_m), \\ \Delta_i F(s_m) &\equiv \frac{V(x(s_i|s_m)) - V(x(s_{i+1}|s_m))}{x(s_m)(1 - x(s_m))} = \\ &= \frac{E_{s_{i+1}} \{V(x(s_i|s_{i+1})) | s_m\}}{x(s_m)(1 - x(s_m))} \geq 0.\end{aligned}$$

Коэффициент $F(s_{m,n})$ меньше значения $F(s_m)$, соответствующего единице наблюдения $\{s_1\}$, на сумму вкладов неучтенных уровней иерархии от 1 до $n - 1$:

$$\begin{aligned}F(s_m) - F(s_{m,n}) &= \sum_{i=1}^{m-1} \frac{V(x(s_i|s_m)) - V(x(s_{i+1}|s_m))}{x(s_m)(1 - x(s_m))} - \\ &- \sum_{i=n}^{m-1} \frac{V(x(s_i|s_m)) - V(x(s_{i+1}|s_m))}{x(s_m)(1 - x(s_m))} = \\ &= \sum_{i=1}^{n-1} \frac{E_{s_{i+1}} \{V(x(s_i|s_{i+1})) | s_m\}}{x(s_m)(1 - x(s_m))} = \sum_{i=1}^{n-1} \Delta_i F(s_m) \geq 0.\end{aligned}$$

Описанные результаты базируются на разложении дисперсии концентрации интересующего аллеля в иерархически подразделенной метапопуляции.

Структуру показателей инбридинга такой метапопуляции можно также исследовать и с помощью подхода С. Райта, основанного на *дополнительном предположении о независимости коэффициента инбридинга и концентрации аллеля* в субпопуляциях. Рассмотрим данное предположение в условиях, допускающих нарушение закона Харди–Вайнберга в субпопуляциях уровня единицы наблюдения.

Результат А9. Пусть рассматривается метапопуляция s_m с иерархической структурой подразделения на субпопуляции. Положим, что концентрации интересующего аллеля в субпопуляциях $\{s_i\}$ самого низкого уровня иерархии (единицы наблюдения) внутри s_m равны $\{x(s_i|s_m)\}$, концентрация $H(s_1)$ гетерозигот с этим аллелем в s_1 (с учетом порядка) находится как $x(s_1)(1 - x(s_1))(1 - F(s_1)) = H_0(s_1)(1 - F(s_1))$.

Тогда общая (полная) концентрация $H(s_m)$ гетерозигот в наугад выбранной из s_m субпопуляции s_1 при независимости в совокупности концентрации аллелей и коэффициентов инбридинга F факторизуется как

$$\begin{aligned}H(s_m) &= H_0(s_m)(1 - E_{s_1} \{F(s_1) | s_m\})(1 - F(s_{m,m-1})) \times \\ &\times \prod_{i=1}^{m-2} (1 - E_{s_{i+1}} \{F(s_{i+1,i}) | s_m\}), \\ F(s_{i+1,i}) &\equiv V(x(s_{i+1}|s_i)) / H_0(s_{i+1,i}), \\ H_0(s_i) &\equiv x(s_i)(1 - x(s_i)).\end{aligned}$$

Отдельные составляющие данного выражения имеют следующую интерпретацию. Инбредность

субпопуляций первого уровня иерархии в s_m характеризует $E_{s_1}\{F(s_1)|s_m\}$, а подразделенность метапопуляции отражает произведение $(1 - F(s_{m,m-1})) \prod (...)$, характеризующее дивергенцию субпопуляций в s_m . Компонент $E_{s_1}\{F(s_1)|s_m\}$ соответствует неслучайному коэффициенту инбридинга $F_{IS}(s_m)$, а $F_{ST}(s_m) \equiv (1 - F(s_{m,m-1}))(1 - \prod_{i=1}^{m-2} (1 - E_{s_{i+1}}\{F(s_{i+1,i})|s_m\}))$ случайному коэффициенту инбридинга F_{ST} .

В реальных данных может быть не представлена информация о нарушениях закона Харди–Вайнберга в субпопуляциях, служащих единицами наблюдения. В связи с этим желательно проанализировать последствия существования нарушений в единицах наблюдения (например, из-за скрытой подразделенности, либо недоступности наблюдению, либо простого игнорирования имеющихся сведений), когда принято решение об их отсутствии. Эта проблема рассматривалась выше при разложении концентрации гетерозигот и коэффициента инбридинга по уровням иерархии. Теперь обратимся к ней при факторизации данных показателей в предположении независимости концентраций аллеля и коэффициентов инбридинга.

Следствие A10. Пусть в условиях предыдущего результата в метапопуляции s_m с t уровнями иерархии в качестве единицы наблюдения служат субпопуляции $\{s_n\}$ уровня n ($1 < n < t$), в которых концентрации интересующего аллеля равны $\{x(s_n)\}$, а концентрации несущих его гетерозигот неизвестны. Пусть в качестве значения неизвестных концентраций гетерозигот используется $H_0(s_n)$.

Тогда $H(s_m)$ будет меньше значения $H(s_{m,n})$ в $\prod_{i=1}^{n-1} (1 - E_{s_{i+1}}\{F(s_{i+1,i})|s_m\})$ раз, где $F(s_{i+1,i})$ неизвестные коэффициенты инбридинга в субпопуляциях.

Поскольку обычно коэффициенты инбридинга в популяциях человека порядка тысячных, максимум сотых долей, то разумно аппроксимировать полученные формулы более простыми выражениями, пренебрегая произведениями коэффициентов.

Результат A11. В иерархически подразделенной метапопуляции s при малых коэффициентах инбридинга в ее субпопуляциях общая концентрация $H(s_m)$ гетерозигот с заданным аллелем в наугад выбранной из s_m единицы наблюдения s_1 с коэффициентом инбридинга $F(s_1)$ аппроксимируется следующим выражением:

$$\begin{aligned} H(s_1|s_m) &\sim x(s_m)(1 - x(s_m))(1 - E_{s_1}\{F(s_1)|s_m\} - \\ &- F(s_{m,m-1}) - \sum_{i=1}^{m-2} E_{s_{i+1}}\{F(s_{i+1,i})|s_m\}) \equiv \\ &\equiv x(s_m)(1 - x(s_m))(1 - F(s_m)) \equiv \\ &\equiv H_0(s_1|s_m)(1 - F(s_m)), \\ F(x(s_m)) &\equiv E_{s_1}\{F(s_1)|s_m\} + \\ &+ F(s_{m,m-1}) + \sum_{i=1}^{m-2} E_{s_{i+1}}\{F(s_{i+1,i})|s_m\}, \end{aligned}$$

где $F(s_m)$ — коэффициент инбридинга метапопуляции s_m , неотрицательные коэффициенты $\{F(s_{i+1,i})\}$ характеризуют подразделенность субпопуляций. Таким образом, приближенно коэффициент инбридинга иерархически подразделенной метапопуляции равен сумме средних коэффициентов инбридинга субпопуляций на отдельных уровнях иерархии.

СПИСОК ЛИТЕРАТУРЫ

1. Ревазов А.А., Парадеева Г.М., Русакова Г.И. Пригодность русских фамилий в качестве квазигенетического маркера // Генетика. 1986. Т. 22. № 4. С. 699–703.
2. Гинтер Е.К., Зинченко Р.А., Ельчинова Г.И. и др. Роль факторов популяционной динамики в распространении наследственной патологии в российских популяциях // Мед. генетика. 2004. Т. 3. № 12. С. 548–555.
3. Crow M.F., Mange A.P. Measurement of inbreeding from the frequency of marriages between persons of the same surname // Soc. Biology. 1982. V. 29. № 1/2. P. 101–105.
4. Rogers A.R. Doubts about isonymy // Hum. Biology. 1991. V. 63. № 5. P. 663–668.
5. Lasker W.G. Surnames and Genetic Structure. Cambridge: Cambr. Univ. Press, 1985. 148 p.
6. Сорокина И.Н., Чурносоев М.И., Балтуцкая И.В. и др. Антропогенетическое изучение населения центральной России. М.: Изд-во РАМН, 2014. 336 с.
7. 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.
8. Пасеков В.П. К анализу случайных процессов изонимии. I. Структура изонимии // Генетика. 2021. Т. 57. № 10. С. 1194–1204. <https://doi.org/10.31857/S001667582110009X>
9. Пасеков В.П. К анализу случайных процессов изонимии. II. Динамика дивергенции популяций // Генетика. 2021. Т. 57. № 11. С. 1318–1329. <https://doi.org/10.31857/S0016675821101114>

10. Пасеков В.П. Описание дивергенции субпопуляций в иерархической системе при анализе изонимии. I. Дисперсия как показатель дивергенции // Генетика. 2022. Т. 58. № 6. С. 713–727.
11. Пасеков В.П. Описание дивергенции субпопуляций в иерархической системе при анализе изонимии. II. Вероятности неизонимных встреч // Генетика. 2023. Т. 59. № 11. С. 1326–1340.
12. Wright S. The interpretation of population structure by F statistics with special regard to systems of mating // Evolution. 1965. V. 19. P. 395–420.

Description of the Divergence of Subpopulations in the Hierarchical System under the Analysis of Isonymy. III. Surname Inbreeding Coefficient

V. P. Passekov*

Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Moscow, 119991 Russia

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

The hierarchical structure of metapopulations characteristic of humans with their subdivision into parts (subpopulations), usually classified on the basis of administrative-territorial division, principles of biological systematics, etc., is considered. Attention is focused on the description of surname divergence of subpopulations. The surname inbreeding coefficient F_s is expressed through the divergence indices, which is similar to the traditional inbreeding coefficient in population genetics. Its expansion by hierarchy levels into the sum of increments of the surname inbreeding coefficient corresponding to separate levels is obtained. The relationship between the found decomposition and the decomposition of the dispersion of the distribution of the surname concentration by subpopulations is demonstrated. With an additional assumption about the independence of the concentration of the surname and F_s , the factorization of F_s by hierarchy levels is found in line with the ideas of S. Wright. It is simplified with a small surname divergence of subpopulations when the F_s coefficient of the entire metapopulation is equal to the sum of the average F_s values at individual levels of the hierarchy. The underestimation of surname inbreeding was obtained, when subpopulations of a higher rank serve as units of observation instead of undivided subpopulations of the first level of the hierarchy. These results are statistical characteristics of the hierarchical structure, and not a feature of a particular population system, and do not follow from one or another model of microevolution. They are computationally independent of the hierarchical system under study, but allow characterizing their heterogeneity quantitatively. The results obtained refer to rural and urban hierarchical metapopulations as separate components of the entire population. In an appendix that can be read independently, genetic analogues of the properties found are given without proof in respect to the genetic structure of a metapopulation.

Keywords: hierarchical structure of populations, metapopulations, surname concentrations in human subpopulations, characteristics of heterogeneity of metapopulations, decomposition of the probability of non-isonimic encounters and the coefficient of surname inbreeding by hierarchy levels.

КРАТКИЕ
СООБЩЕНИЯ

УДК 576.6:575.8+631.461.52

ПОЛИМОРФИЗМ КОРОВЫХ И СИМБИОТИЧЕСКИ
СПЕЦИАЛИЗИРОВАННЫХ ГЕНОВ У ПОЛИТИПИЧЕСКИХ ВИДОВ
КЛУБЕНЬКОВЫХ БАКТЕРИЙ

© 2023 г. Н. А. Проворов^{1, *}, А. К. Кимеклис^{1, 2}, Е. С. Карасев¹, С. Хосид¹,
О. П. Онищук¹, О. Н. Курчак¹, Е. Е. Андронов^{1, 3}

¹Всероссийский научно-исследовательский институт сельскохозяйственной микробиологии,
Санкт-Петербург, 196608 Россия

²Санкт-Петербургский государственный университет, Санкт-Петербург, 199034 Россия

³Почвенный институт им. В.В. Докучаева, Москва, 109017 Россия

*e-mail: provorovnik@yandex.ru

Поступила в редакцию 25.05.2023 г.

После доработки 04.07.2023 г.

Принята к публикации 12.07.2023 г.

Политипические виды клубеньковых бактерий *Rhizobium leguminosarum* (включает биовары *viciae* и *trifolii*) и *Neorhizobium galegae* (биовары *orientalis* и *officinalis*) различаются по нуклеотидному полиморфизму генов “домашнего хозяйства” (*hkg*) и симбиотически специализированных генов (*sym*), контролирующих образование N₂-фиксирующих клубеньков у бобовых растений. У *R. leguminosarum* величины *p*-distance по *sym*-генам выше, чем по *hkg*-генам у штаммов, относящихся как к одному, так и к разным биоварам. У *N. galegae* различия между биоварами по *sym*-генам выше, чем по *hkg*-генам, однако у штаммов одного биовара полиморфизм *sym*-генов ниже, чем *hkg*-генов. Коэффициенты дифференциации биоваров по обеим группам генов выше у *N. galegae*, чем у *R. leguminosarum*, что может быть связано с пространственной изоляцией биоваров *N. galegae*. У данного вида ризобий конгруэнтность (сходство) филогений *sym*- и *hkg*-генов более выражена, чем у *R. leguminosarum*. Это указывает на активный перенос *sym*-генов в популяциях *R. leguminosarum*, который может быть важным фактором глубокой диверсификации этого вида ризобий по признакам симбиоза.

Ключевые слова: клубеньковые бактерии (ризобии), нуклеотидный полиморфизм, эволюция симбиоза, *Rhizobium leguminosarum*, *Neorhizobium galegae*, конгруэнтность филогений и перенос генов.

DOI: 10.31857/S0016675823120081, **EDN:** QCSMBNH

Клубеньковые бактерии (ризобии), N₂-фиксирующие симбионты бобовых растений обладают сложно устроенными геномами, которые состоят из консервативных коровых и варибельных акцессорных частей [1]. Гены коровой части генома отвечают за функции “домашнего хозяйства” — основной метаболизм, развитие и размножение, матричные процессы (*hkg*, от house-keeping genes), и их дивергенция определяет видообразование. Акцессорные гены кодируют различные адаптивные функции, включая симбиоз с растениями (*sym*-гены), и их дивергенция характеризует микроэволюционные процессы [2]. Многие виды ризобий являются политипическими: они состоят из биоваров, контрастно различающихся по хозяйской специфичности [3]. Среди этих видов наиболее изучены *Rhizobium leguminosarum*, включающий bv. *viciae* (симбионты растений трибы Fabeae, роды *Lathyrus*, *Lens*, *Pisum*, *Vavilovia*, *Vicia*), bv. *trifolii* (симбионты рода *Trifolium* из трибы Trifolieae), bv.

phaseoli (симбионты *Phaseolus vulgaris* из трибы Phaseoleae) и *Neorhizobium galegae*, разделяемый на bv. *orientalis* и bv. *officinalis* — симбионты козлятника восточного (*Galega orientalis*) и лекарственного (*G. officinalis*).

Ранее мы показали [4–6], что штаммы *R. leguminosarum* и *N. galegae* существенно варьируют по нуклеотидным последовательностям *sym*- и *hkg*-генов. Поскольку эти виды ризобий различаются по геномной организации [2] и по ареалам распространения, мы решили сопоставить полиморфизм *sym*- и *hkg*-генов у *R. leguminosarum* и *N. galegae*.

Для этого мы изучили нуклеотидный полиморфизм (*p*-distance) у выборки штаммов ризобий из базы данных NCBI (<https://www.ncbi.nlm.nih.gov>), включающей 16 штаммов *R. leguminosarum* (10 — bv. *viciae*, 6 — bv. *trifolii*) и 14 штаммов *N. galegae* (8 — bv. *officinalis*, 6 — bv. *orientalis*) по конкатенированным (объединенным) последовательностям 4 *sym*-генов (*nodA*, *nodC*, *nodD*, *nifH*) и 4 *hkg*-генов

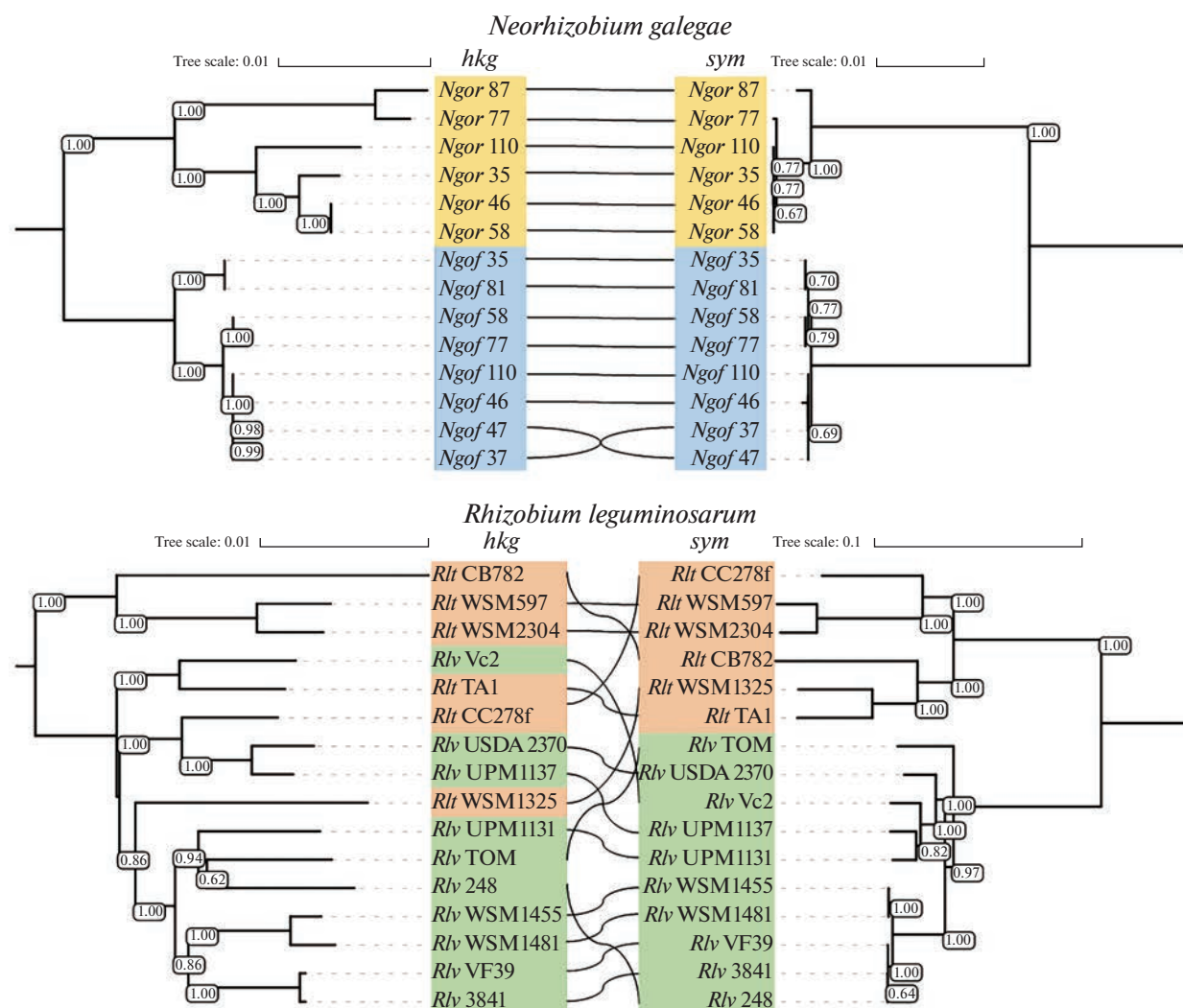


Рис. 1. Сопоставление филогений конкатенированных (объединенных) последовательностей генов “домашнего хозяйства” (*hkg*) и симбиотически специализированных генов (*sym*) у биоваров, входящих в виды *Neorhizobium galegae* (*Ngor* – *bv. orientalis* выделен желтым, *Ngof* – *bv. officinalis* выделен синим) и *Rhizobium leguminosarum* (*Rlt* – *bv. trifolii* выделен оранжевым, *Rlv* – *bv. viciae* выделен зеленым). Филогении генов построены с применением метода Neighbor joining и изображены с сохранением эволюционных расстояний, на узлах представлены значения bootstrap в долях от 1. Линии соединяют названия одного и того же штамма на *hkg*- и *sym*-дендрограммах.

(16S рДНК, *dnaK*, *glnA*, *gltA*); обозначения штаммов даны на рис. 1. Для построения филогений использовали метод Neighbor Joining [7], эволюционные расстояния рассчитывали с помощью метода *p*-distance [8], bootstrap-тест [9] проведен с 1000 итераций в программе Mega X [10], конечная обработка данных проведена с помощью онлайн-утилиты iTOL [11].

Сопоставление двух политипических видов ризобий по степени полиморфизма нуклеотидных последовательностей показало (табл. 1), что у *R. leguminosarum* внутри биоваров наиболее изменчивы *sym*-гены, а у *N. galegae* – *hkg*-гены. Расхождение биоваров по *hkg*-генам одинаково для обоих видов, тогда как *sym*-гены наиболее сильно различаются у биоваров *R. leguminosarum*. У обоих

видов ризобий различия биоваров по уровню полиморфизма сильнее выражены для *sym*-генов, чем для *hkg*-генов, причем у *R. leguminosarum* эти различия более существенны, чем у *N. galegae*. Коэффициенты дифференциации (КД) биоваров по нуклеотидным последовательностям *sym*- и *hkg*-генов (отношения различий между биоварами к общей изменчивости генов [12]) у *N. galegae* выше, чем у *R. leguminosarum*, по обеим группам генов, причем наиболее сильно различие КД выражено для *hkg*-генов.

Анализ конгруэнтности (совпадения) филогений *sym*- и *hkg*-генов показал (рис. 1), что среди 14 изученных штаммов *N. galegae* одинаковое положение на дендрограммах этих генов занимают 12 (85.7%) штаммов, тогда как среди 16 штаммов

Таблица 1. Нуклеотидный полиморфизм (p -distance) генов симбиоза (sym) и домашнего хозяйства (hkg) у политипических видов ризобий

Гены	<i>Rhizobium leguminosarum</i>		<i>Neorhizobium galegae</i>	
	bv. <i>viciae</i>	bv. <i>trifolii</i>	bv. <i>orientalis</i>	bv. <i>officinalis</i>
Сравнение штаммов одного биовара				
<i>sym</i>	0.0393 <	0.1351	0.0018 >	0.0007
<i>hkg</i>	0.0192 <	0.0307	0.0177 >	0.0035
<i>sym/hkg</i>	2.05	4.40	0.10	0.20
<i>t</i> _{St} (<i>P</i> ₀)	9.52 (<0.001)	27.13 (<0.001)	10.52 (<0.001)	4.87 (<0.01)
Сравнение штаммов из разных биоваров				
<i>sym</i>	0.2485 >		0.0444	
<i>hkg</i>	0.0290 ≈		0.0313	
<i>sym/hkg</i>	8.48		1.42	
<i>t</i> _{St} (<i>P</i> ₀)	36.92 (<0.001)		3.34 (<0.01)	
Коэффициенты дифференциации биоваров				
<i>sym</i>	0.4404 <		0.9492	
<i>hkg</i>	0.0228 <		0.4842	
<i>sym/hkg</i>	19.31		1.96	
<i>t</i> _{St} (<i>P</i> ₀)	29.12 (<0.001)		17.61 (<0.001)	

Примечание. Средние значения параметров полиморфизма sym - и hkg -генов сравнивали с помощью критерия Стьюдента (t_{St}). Знаки "<" и ">" указывают на достоверность различий ($P_0 < 0.01$) между находящимися в соседних колонках значениями параметров, вычисленными для разных биоваров или видов ризобий ("≈" – различие недостоверно, $P_0 > 0.05$).

R. leguminosarum такое положение занимают лишь 2 (12.5%) штамма ($t_{St} = 5.86$, $P_0 < 0.001$). Это позволяет предположить, что выявленные нами межвидовые различия по степени полиморфизма sym - и hkg -генов связаны с более высокой интенсивностью переноса sym -генов в популяциях *R. leguminosarum*, что согласуется с различием видов по величинам КД. Известно, что геном этого вида ризобий состоит из хромосомы и плазмид, одна из которых ($pSym$), имеющая размер 200–500 тпн, содержит sym -гены и обычно способна к конъюгативному переносу между разными штаммами [1]. У *N. galegae* sym -гены расположены на хромиде, имеющих размер более 1600 тпн и не способных к переносу. Эти хромиды имеют системы репликации плазмидного типа *repABC*, однако содержат гены $pRHK$ и $tRHK$, которые у большинства видов ризобий расположены в хромосомах [13].

Ограниченный перенос генов в популяциях *N. galegae* может быть связан с тем, что *Galega orientalis* и *G. officinalis* обычно произрастают раздельно [14], что определяется конкуренцией родственных видов козлятника за общие биотопы, а также разной экологической специализацией этих видов. В то же время хозяева биоваров *R. leguminosarum*, относящиеся к разным трибам бобовых (например, *Trifolium*, *Lathyrus*, *Vicia*), часто произрастают совместно, что создает возможность для стабильного сосуществования bv. *viciae* и bv. *trifolii* и для их обмена генами.

Ранее было показано, что эколого-генетические механизмы эволюции sym - и hkg -генов у ризобий существенно различаются. Эволюция sym -генов происходит под действием селективных факторов (различные формы индивидуального и группового отбора), индуцируемых растениями-хозяевами, тогда как эволюция hkg -генов определяется не изученными пока факторами, которые могут быть связаны с почвенно-климатическими условиями [2]. Представленные нами данные позволяют предположить, что у *N. galegae* благодаря ограниченному переносу генов в популяциях, индуцируемый растениями дизруптивный отбор, определяющий формирование политипической структуры вида, не ограничен sym -генами и распространяется на hkg -гены.

Причиной межвидовых различий ризобий по уровню полиморфизма может быть также разная степень их дивергенции хозяев, которые у *N. galegae* относятся к одному роду бобовых, а у *R. leguminosarum* – к разным родам и трибам. Анализ величин КД (табл. 1) позволяет предположить, что у *R. leguminosarum* популяции обладают большей, чем у *N. galegae*, панмиктичностью, что согласуется с результатами филогенетического анализа (рис. 1). Интересно отметить, что у обоих видов ризобий дифференциация биоваров менее выражена для hkg -генов, чем для sym -генов. Низкая дифференциация биоваров *R. leguminosarum* по hkg -генам (КД = 0.0228) указывает на то, что интенсивность

внутривидового переноса для этих хромосомных генов может быть достаточно высокой, например в связи с характерными для ризобий процессами трансформации и трансдукции [15].

Таким образом, мы показали, что у политипических видов *N. galegae* и *R. leguminosarum* нуклеотидный полиморфизм *hkg*- и *sym*-генов существенно различается, поскольку: а) изменчивость генов обеих групп у *R. leguminosarum* выше, чем у *N. galegae*, причем это различие наиболее выражено для *sym*-генов (отношения *sym/hkg* для всех показателей полиморфизма у *R. leguminosarum* выше, чем у *N. galegae*); б) величины КД по обоим группам генов у *N. galegae* выше, чем у *R. leguminosarum*; в) у *N. galegae* конгруэнтность филогений *sym*- и *hkg*-генов более выражена, чем у *R. leguminosarum*.

Таким образом, выявленная нами видоспецифичность нуклеотидного полиморфизма генов ризобий, по-видимому, определяется особенностями организации геномов *N. galegae* и *R. leguminosarum*, влияющими на мобильность *sym*-генов. Наиболее велика она у *R. leguminosarum*, что соответствует глубокой диверсификации этого вида ризобий по хозяйской специфичности, определяющей более широкое географическое распространение *R. leguminosarum* по сравнению с *N. galegae*. Логично предположить, что в эволюции *sym*-генов, наряду с индуцируемым хозяевами отбором существенную роль играет перенос этих генов в бактериальных популяциях, интенсивность которого определяется как их пространственной структурой, так и геномной локализацией *sym*-генов. Поэтому влияние географической изоляции растений-хозяев на дивергенцию их симбионтов может быть не менее значимым, чем влияние индуцируемых растениями селективных факторов.

Механизмы связи между интенсивностью переноса генов в популяциях ризобий и их полиморфизмом представляют большой интерес для дальнейших исследований. Особое внимание привлекают механизмы эволюции коровой части генома, которая: а) происходит независимо от симбиоза — наиболее изученного, и, как долго считалось, основного фактора эволюции ризобий; б) включает перенос хромосомных *hkg*-генов в популяциях обоих видов ризобий, который, как показал анализ КД, может происходить не менее интенсивно, чем перенос вне хромосомных *sym*-генов в связи с характерными для этих бактерий процессами трансформации и трансдукции [15].

Работа поддержана грантом РНФ 19-16-00081П.

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

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Young J.P., Crossman L.C., Johnston A.W.B. et al. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components // *Genome Biol.* 2006. V. 7. <https://doi.org/10.1186/gb-2006-7-4-r34>
2. Provorov N.A., Andronov E.E., Kimeklis A.K. et al. Microevolution, speciation and macroevolution in rhizobia: Genomic mechanisms and selective patterns // *Front. Plant Sci.* 2022. V. 13. <https://doi.org/10.3389/fpls.2022.1026943>
3. Rogel M.A., Ormeño-Orrillo E., Martínez-Romero E. Symbiovars in rhizobia reflect bacterial adaptation to legumes // *Syst. Appl. Microbiol.* 2011. V. 34. № 2. P. 96–104. <https://doi.org/10.1016/j.syapm.2010.11.015>
4. Карасев Е.С., Андронов Е.Е., Аксенова Т.С. и др. Эволюция ризобий козлятника (*Neorhizobium galegae*): анализ полиморфизма генов фиксации азота и развития клубеньков // *Генетика.* 2019. Т. 55. № 2. С. 234–238. <https://doi.org/10.1134/S0016675819020085>
5. Кимеклис А.К., Кузнецова И.Г., Сазанова А.Л. и др. Дивергентная эволюция симбиотических бактерий: ризобии реликтового бобового *Vavilovia formosa* формируют обособленную группу в пределах вида *Rhizobium leguminosarum* bv. *viciae* // *Генетика.* 2018. Т. 54. № 7. С. 851–855. <https://doi.org/10.1134/S0016675818070068>
6. Kimeklis A., Chirak E., Kuznetsova I. et al. Rhizobia isolated from the relict legume *Vavilovia formosa* represent a genetically specific group within *Rhizobium leguminosarum* biovar *viciae* // *Genes.* 2019. V. 10. <https://doi.org/10.3390/genes10120991>
7. Saitou N., Nei M. The Neighbor-joining method: A new method for reconstructing phylogenetic trees // *Mol. Biol. Evol.* 1987. V. 4. № 4. P. 406–425. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>
8. Nei M., Kumar S. *Molecular Evolution and Phylogenetics.* N. Y.: Oxford Univ. Press, 2000. 248 p.
9. Felsenstein J. Confidence limits on phylogenies: An approach using the bootstrap // *Evolution.* 1985. V. 39. № 4. P. 783–791. <https://doi.org/10.1111/j.1558-5646.1985.tb00420>
10. Kumar S., Stecher G., Li M. et al. MEGA X: Molecular evolutionary genetics analysis across computing platforms // *Mol. Biol. Evol.* 2018. V. 35. № 6. P. 1547–1549. <https://doi.org/10.1093/molbev/msy096>
11. Letunic I., Bork P. Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation // *Nucl. Ac. Res.* 2021. V. 47(W1). <https://doi.org/10.1093/nar/gkz239>
12. Verity R., Nichols R. What is genetic differentiation, and how should we measure it — GST, D, neither or both? // *Mol. Ecol.* 2014. V. 23. № 17. P. 4216–4225. <https://doi.org/10.1111/mec.12856>

13. Österman J., Marsh J., Laine P.K. et al. Genome sequencing of two *Neorhizobium galegae* strains reveals a *noe* T gene responsible for the unusual acetylation of the nodulation factors // BMC Genomics. 2014. V. 15. № 1.
https://doi.org/10.1186/1471-2164-15-500
14. Andronov E., Terefework Z., Roumiantseva M. et al. Symbiotic and genetic diversity of *Rhizobium galegae* isolates collected from the *Galega orientalis* gene center in the Caucasus // Appl. Environ. Microbiol. 2003. V. 69. № 2. P. 1067–1074.
https://doi.org/10.1128/AEM.69.2.1067–1074.2003
15. Beringer J.E., Brewin N.J., Johnston A.W.B. The genetic analysis of *Rhizobium* in relation to symbiotic nitrogen fixation // Heredity. 1980. V. 45. № 2. P. 161–186.

Polymorphism of Core and Symbiotically Specialized Genes in the Polytypic Species of Nodule Bacteria

N. A. Provorov^{a, *}, A. K. Kimeklis^{a, b}, E. S. Karasev^a, S. Khosid^a,
O. P. Onishchuk^a, O. N. Kurchak^a, and E. E. Andronov^{a, c}

^aAll-Russia Research Institute of Agricultural Microbiology, Saint Petersburg, 196608 Russia

^bSaint Petersburg State University, Saint Petersburg, 199034 Russia

^cDokuchaev Soil Institute, Moscow, 109017 Russia

*e-mail: provorovnik@yandex.ru

The polytypic species of nodule bacteria *Rhizobium leguminosarum* (includes biovars *viciae* and *trifolii*) and *Neorhizobium galegae* (biovars *orientalis* and *officinalis*) differ in the nucleotide polymorphism of housekeeping genes (*hkg*) and of symbiotically specialized genes (*sym*) that control the formation of N₂-fixing nodules in leguminous plants. In *R. leguminosarum*, *p*-distance values for *sym* genes are higher than for *hkg* genes in strains from the same and from different biovars. In *N. galegae*, differences between biovars in *sym* genes are higher than in *hkg* genes while within biovars, polymorphism in *sym* genes is lower than in *hkg* genes. Coefficients of biovar differentiation for both groups of genes are higher in *N. galegae* than in *R. leguminosarum*, possibly reflecting the spatial isolation of *N. galegae* biovars. In these species of rhizobia, the phylogenetic congruence of *sym* and *hkg* genes is more pronounced in *N. galegae* than in *R. leguminosarum*. This difference indicates an active transfer of *sym* genes in *R. leguminosarum* populations possibly representing an important factor of the deep diversification for symbiotic traits in this rhizobia species.

Keywords: nodule bacteria (rhizobia), nucleotide polymorphism, evolution of symbiosis, *Rhizobium leguminosarum*, *Neorhizobium galegae*, phylogenetic congruence and transfer of genes.

ВАРИАЦИИ СИНАПСИСА ХРОМОСОМ В ПРОФАЗЕ I МЕЙОЗА У СЛЕПУШОНОК *Ellobius tancrei* ГЕТЕРОЗИГОТНЫХ ПО РОБЕРТСОНОВСКИМ ТРАНСЛОКАЦИЯМ

© 2023 г. С. Н. Матвеевский¹, *, Ю. Ф. Богданов¹, Е. А. Ляпунова²,
И. Ю. Баклушинская², О. Л. Коломиец¹, **

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

²Институт биологии развития им. Н.К. Кольцова Российской академии наук, Москва, 119334 Россия

*e-mail: sergey8585@mail.ru

**e-mail: olkolomiets@mail.ru

Поступила в редакцию 16.07.2023 г.

После доработки 20.07.2023 г.

Принята к публикации 25.07.2023 г.

Установлено, что у гетерозигот по робертсоновским перестройкам в профазе мейоза I могут формироваться различные варианты сочетаний мейотических конфигураций (различное количество тривалентов и разные по структуре цепочки хромосом) у одной и той же особи. В работе были исследованы два типа экспериментальных гибридов восточной слепушонки *Ellobius tancrei*, гетерозиготных по четырем робертсоновским транслокациям. Вместо ожидаемых четырех тривалентов были выявлены разные типы мейотических конфигураций на стадии пахитены, вплоть до цепочек из десяти элементов. Мы предполагаем, что вероятность прохождения мейоза в разных клетках зависит от структуры сформировавшихся цепочек и возможности их коррекции. Такие вариации могут приводить к снижению продукции гамет, но не к полной остановке гаметогенеза, что является основой для сохранения робертсоновских транслокаций в популяции.

Ключевые слова: синаптонемный комплекс, мейоз, тривалент, хромосомные перестройки.

DOI: 10.31857/S0016675823120044, **EDN:** QDPBWW

Крупные хромосомные мутации, такие как парацентрические и перичцентрические инверсии, реципрокные и робертсоновские (Rb) транслокации, ведут к трансформации геномов, участвуя в формо- и видообразовании. В этом контексте хромосомные перестройки имеют важные эволюционные последствия, поскольку могут реорганизовывать архитектуру ядер, изменять рекомбинационные паттерны и вести к снижению плодовитости.

Возникновение *de novo* Rb-транслокаций [1] у человека может служить причиной генетических заболеваний и нарушения фертильности [2–4], тогда как у представителей различных отрядов млекопитающих наблюдается полиморфизм и, в ряде случаев, формирование кариоморф, например у обыкновенной бурозубки, домового мыши, некоторых видов летучих мышей, слепышей, слепушонок и других видов [5]. Многообразие таких природных феноменов вступает в противоречие со сложившимися представлениями о том, что нарушения синапсиса хромосом и формирование сложных фигур в профазе мейоза I ведут к стерильности и препятствуют наследованию Rb-

хромосом. Однако накапливаются данные о том, что абсолютная стерильность наблюдается достаточно редко даже в случае гибридов между формами с частичной (монобрахимальной) гомологией Rb-метацентриков, например, у обыкновенной бурозубки [6, 7], домового мыши [8, 9], лемуров [10], слепушонок [11] и межвидовых гибридов валлаби [12]. Одним из объяснений этого феномена может быть формирование не одного, четко детерминированного паттерна синапсиса в профазе мейоза I, а некоего спектра вариантов, часть из которых обеспечивает прохождение профазы I и следующих стадий мейоза, вплоть до формирования эуплоидных по числу плеч гамет, которые могут иметь разный набор хромосом и, соответственно, разное гаплоидное число. Ранее нами была высказана гипотеза о том, что формирование цепочек, состоящих из открытых тривалентов, в результате коррекции синапсиса — распада на отдельные триваленты и их последующее закрытие в результате синаптической пригонки [13] — возможно только при условии позднего кроссинговера [14]. При этом не рассматривалась возможность формирования различных типов

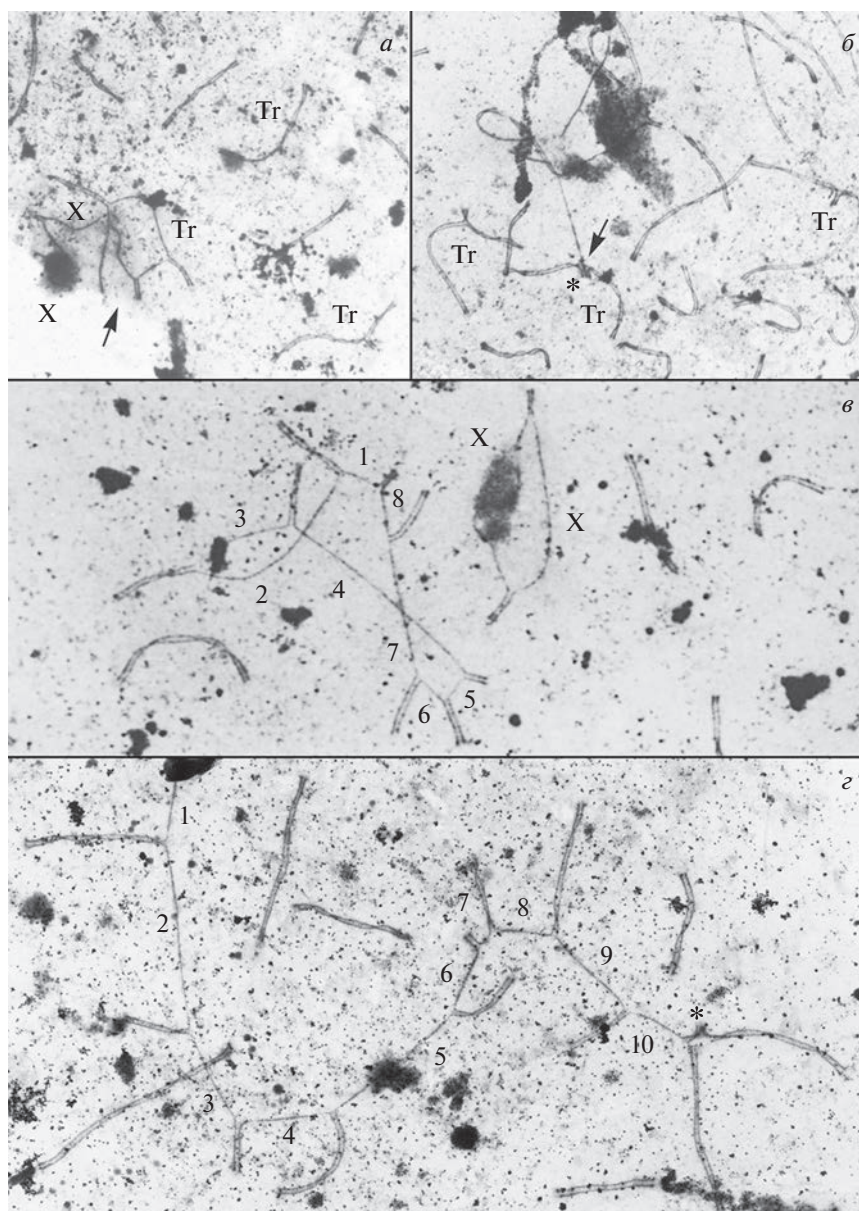


Рис. 1. Хромосомный синapsис в сперматocyтах первого порядка на стадии ранней–средней пахитены у экспериментальных гибридов восточной слепушонки *E. tancrei*. а–в – гибрид с $2n = 50$, г – гибрид с $2n = 48$; Tr – СК-тривалент, X – X-хромосома, * – тройной синapsис, стрелка – ассоциации тривалента с XX; цифры обозначают количество осей в СК-цепочке.

мультивалентов у одной и той же особи в разных клетках.

Для проверки этой гипотезы мы проанализировали особенности организации и поведения синаптонемных комплексов (СК) – нуклеопротеидных скелетов, формирующихся между гомологичными хромосомами в профазе I мейоза [15, 16], – у двух разнохромосомных полужерильных гибридов восточной слепушонки *Ellobius tancrei* (Blasius, 1884) (Rodentia, Mammalia). Следует подчеркнуть, что для *E. tancrei* характерна вариабельность диплоидного числа от 54 до 30 при постоянном числе

хромосомных плеч $NF = 56$, обусловленная различиями в количестве Rb-метацентриков [17, 18]. Другой удивительной особенностью этого вида является изоморфизм половых (XX) хромосом не только у самок, но и у самцов [19].

Гибриды были получены нами в серии прямых и возвратных скрещиваний низкохромосомных ($2n = 36$) и стандартной формы $2n = 54$ особей восточной слепушонки *E. tancrei*. В кариотипе самца № 20703 ($2n = 50$) были выявлены четыре Rb-метацентрика – 1Rb(5.9), 1Rb(6.12), 1Rb(10.19), 1Rb(15.22). В кариотипе самца № 20643 ($2n = 48$)

были идентифицированы шесть Rb-метацентриков — 1Rb(1.8), 1Rb(5.9), 1Rb(4.14), 1Rb(6.12), 2(10.19) (номенклатура по [18]). Оба животных являются гетерозиготами по четырем Rb-транслокациям и, соответственно, ожидалось, что на стадии пахитены будут выявлены четыре СК-тривалента. Однако электронно-микроскопический анализ распластанных ядер сперматоцитов, полученных по методу J. Navarro с соавт. [20], показал широкую вариабельность синаптических конфигураций у этих животных (рис. 1). У гибрида № 20703 наблюдалась вариация от двух или трех сформированных СК-тривалентов, один из которых находился в ассоциации с XX-бивалентом (рис. 1, а, б) до СК-цепочки из восьми осевых элементов (рис. 1, в). У гибрида № 20643 также выявлены различные хромосомные конфигурации: СК-цепочки варьировали по количеству осевых элементов, вплоть до десяти (рис. 1, г). Отклонения от ожидаемых конфигураций вероятно обусловлены наличием прицентромерного гетерохроматина, за счет которого осуществляются негомологичные контакты коротких плеч акроцентриков [21]. Такая интерпретация была высказана нами при изучении 44-х хромосомных гибридов *E. tancrei* гетерозиготных по десяти Rbs, у которых было показано формирование СК-цепочек [22, 23]. Объединение тривалентов рассматривалось как один из путей синаптической пригонки с последующим формированием закрытых СК-тривалентов. Действительно, в районе коротких плеч хромосом у разных гетерозиготных особей был выявлен электронно-плотный гетерохроматин или H3K9me3-позитивный гетерохроматин [24, 25], благодаря которому осуществлялось формирование закрытых СК-тривалентов. Этот факт в некоторой степени пояснял, почему такие гибриды не полностью стерильны.

Мы также показали, что в другой экспериментальной линии гибридов слепушонок часть СК-тривалентов могут быть не полностью закрытыми, при этом область коротких плеч акроцентриков и асинаптированной части метацентрика оказывается транскрипционно инактивирована [26], так же как ранее было выявлено у гетерозиготных мышей [27]. В последнем случае гибриды мышей не были полностью стерильными, в связи с чем было высказано предположение, что в области инактивации нет генов, критически необходимых для прогрессии мейоза. Вероятно, особенности синапсиса, обуславливающие формирование цепочек хромосом в сперматоцитах гибридов, не всегда приводят к пахитенному аресту — феномену блокировки прогрессии мейоцитов на стадии пахитены из-за наличия дефектных паттернов в структуре и поведении хромосом [28, 29] — и, в случае прохождения мейоза, не препятствуют формированию сбалансированных гамет.

Именно разнообразие вариантов синапсиса хромосом при формировании мультивалентов приводит к снижению продукции гамет, но не к полной остановке гаметогенеза. Стоит отметить, что важную роль в упорядочивании синапсиса хромосом в цепочках играют три важных фактора: прикрепление хромосом к ядерной оболочке, формирование хиазм в только что просинаптированных участках, даже при наличии широких зон асинапсиса, и негомологичные контакты коротких плеч акроцентриков. Сохранение плодовитости у гибридов позволяет поддерживать поток генов между формами и видами, но это не означает, что виды, вступающие в гибридизацию, не могут рассматриваться как “хорошие” таксономические виды. Более того, даже небольшое снижение плодовитости гибридов может вести к видообразованию [30], а теоретические представления о “видообразовании с потоком генов” [31] находят все новые подтверждения на разных объектах [12, 32, 33], в том числе и на слепушонках.

Работа выполнена при поддержке Российского научного фонда, грант № 22-24-01163 (О.Л. Коломиец).

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

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Robertson W.M.R.B. Chromosome studies. I. Taxonomic relationships shown in the chromosomes of Tettigidae and Acrididae: V-shaped chromosomes and their significance in Acrididae, Locustidae and Grillidae: Chromosomes and variations // J. of Morphology. 1916. V. 27. P. 179–331.
2. Baccetti B., Capitani S., Collodel G. et al. Infertile spermatozoa in a human carrier of Robertsonian translocation 14;22 // Fertility and Sterility. 2002. V. 78. № 5. P. 1127–1130. [https://doi.org/10.1016/S0015-0282\(02\)03379-4](https://doi.org/10.1016/S0015-0282(02)03379-4)
3. Mack H., Swisshelm K. Robertsonian translocations // Brenner's Encyclopedia of Genetics. 2013. V. 6. P. 301–305. <https://doi.org/10.1016/B978-0-12-374984-0.01357-7>
4. Wilch E.S., Morton C.C. Historical and clinical perspectives on chromosomal translocations // Chromosome Translocation. Advances in Experimental Medicine and Biology. Singapore: Springer, 2018. P. 1–14. https://doi.org/10.1007/978-981-13-0593-1_1
5. King M. Species Evolution: The Role of Chromosome Change. Cambridge: Cambridge Univ. Press, 1993. 336 p.
6. Narain Y., Fredga K. Spermatogenesis in common shrews *Sorex araneus* from a hybrid one with extensive Robertsonian polymorphism // Cytogenetic and Cell Genetics. 1998. V. 80. P. 158–164. <https://doi.org/10.1159/000014973>

7. *Jadwiszczak K.A., Banaszek A.* Fertility in the male common shrews *Sorex araneus* from the extremely narrow hybrid zone between chromosome races // *Mammalian Biol.* 2006. V. 71. № 5. P. 257–267. <https://doi.org/10.1016/j.mambio.2006.02.004>
8. *Johannisson R., Winking H.* Synaptonemal complexes of chains and rings in mice heterozygous for multiple Robertsonian translocations // *Chromosome Res.* 1994. V. 2. P. 137–145. <https://doi.org/10.1007/BF01553492>
9. *Ribagorda M., Berriós S., Solano E. et al.* Meiotic behavior of a complex hexavalent in heterozygous mice for Robertsonian translocations: insights for synapsis dynamics // *Chromosoma*. 2019. V. 128. P. 149–163. <https://doi.org/10.1007/s00412-019-00695-8>
10. *Dutrillaux B., Rumpler Y.* Chromosomal evolution in Malagasy lemurs. // *Cytogenetic and Cell Genetics*. 1977. V. 18. № 4. P. 197–211. <https://doi.org/10.1159/000130763>
11. *Matveevsky S., Bakloushinskaya I., Tambovtseva V. et al.* Analysis of meiotic chromosome structure and behavior in Robertsonian heterozygotes of *Ellobius tancrei* (Rodentia, Cricetidae): A case of monobrachial homology // *Comparative Cytogenetics*. 2015. V. 9. № 4. P. 691–697. <https://doi.org/10.3897/CompCytogen.v9i4.5674>
12. *Potter S., Bragg J.G., Turakulov R. et al.* Limited introgression between rock-wallabies with extensive chromosomal rearrangements // *Mol. Biol. and Evol.* 2022. V. 39. № 1. <https://doi.org/10.1093/molbev/msab333>
13. *Moses M.J., Poorman P.A.* Synaptonemal complex analysis of mouse chromosomal rearrangements: II. Synaptic adjustment in a tandem duplication // *Chromosoma*. 1981. V. 81. № 4. P. 519–535. <https://doi.org/10.1007/BF00285847>
14. *Коломиец О.Л., Ляпунова Е.А., Мазурова Т.Ф. и др.* Различные пути формирования тривалентов СК у гибридов гетерозиготных по Rb-транслокациям // Молекулярные механизмы генетических процессов: молекулярная генетика, эволюция и молекулярно-генетические основы селекции. М.: Наука, 1985. С. 72–84.
15. *Fawcett D.W.* The fine structure of chromosomes in the meiotic prophase of vertebrate spermatocytes // *J. of Cell Biol.* 1956. V. 2. № 4. P. 403–406. <https://doi.org/10.1083/jcb.2.4.403>
16. *Moses M.J.* Chromosomal structures in crayfish spermatocytes // *J. of Biophys. and Biochem. Cytology*. 1956. V. 2. № 2. <https://doi.org/10.1083/jcb.2.2.215>
17. *Lyapunova E.A., Vorontsov N.N., Korobitsina K.V. et al.* A Robertsonian fan in *Ellobius talpinus* // *Genetica*. 1980. V. 52/53. P. 239–247. <https://doi.org/10.1007/BF00121833>
18. *Romanenko S.A., Lyapunova E.A., Saidov A.S. et al.* Chromosome translocations as a driver of diversification in mole voles *Ellobius* (Rodentia, Mammalia) // *Int. J. Mol. Sci.* 2019. V. 20. № 18. <https://doi.org/10.3390/ijms20184466>
19. *Vorontsov N.N.* The evolution of the sex chromosomes // *Cytotaxonomy and Vertebrate Evolution* N. Y.: Acad. Press, 1973. P. 619–657.
20. *Navarro J., Vidal F., Guitart M., Egozcue J.* A method for the sequential study of synaptonemal complexes by light and electron microscopy // *Hum. Genet.* 1981. V. 59. P. 419–421. <https://doi.org/10.1007/BF00295483>
21. *Sharp P.J.* Synaptic adjustment at a C-band heterozygosity // *Cytogenetic and Genome Res.* 1986. V. 41. № 1. P. 56–57.
22. *Коломиец О.Л., Ляпунова Е.А., Мазурова Т.Ф. и др.* Участие гетерохроматина в формировании цепочек синаптомемных комплексов у животных, гетерозиготных по множественным робертсоновским транслокациям // *Генетика*. 1986. Т. 22. № 2. С. 273–280.
23. *Bogdanov Y.F., Kolomiets O.L., Lyapunova E.A. et al.* Synaptonemal complexes and chromosome chains in the rodent *Ellobius talpinus* heterozygous for ten Robertsonian translocations // *Chromosoma*. 1986. V. 94. P. 94–102. <https://doi.org/10.1007/BF00286986>
24. *Matveevsky S., Tretiakov A., Kashintsova A. et al.* Meiotic nuclear architecture in distinct mole vole hybrids with Robertsonian translocations: Chromosome chains, stretched centromeres, and distorted recombination // *Int. J. Mol. Sci.* 2020. V. 21. № 20. <https://doi.org/10.3390/ijms21207630>
25. *Matveevsky S., Bakloushinskaya I., Tambovtseva V. et al.* Nonhomologous chromosome interactions in prophase I: Dynamics of bizarre meiotic contacts in the Alay mole vole *Ellobius alaicus* (Mammalia, Rodentia) // *Genes*. 2022. V. 13. № 12. <https://doi.org/10.3390/genes13122196>
26. *Kolomiets O., Bakloushinskaya I., Pankin M. et al.* Irregularities in meiotic prophase I as prerequisites for reproductive isolation in experimental hybrids carrying Robertsonian translocations // *Diversity*. 2023. V. 15. № 3. <https://doi.org/10.3390/d15030364>
27. *Manterola M., Page J., Vasco C. et al.* A high incidence of meiotic silencing of unsynapsed chromatin is not associated with substantial pachytene loss in heterozygous male mice carrying multiple simple Robertsonian translocations // *PLoS Genetics*. 2009. V. 5. № 8. <https://doi.org/10.1371/journal.pgen.1000625>
28. *Roeder G.S.* Meiotic chromosomes: It takes two to tango // *Genes & Development*. 1997. V. 11. № 20. P. 2600–2621. <https://doi.org/10.1101/gad.11.20.2600>
29. *Roeder G.S., Bailis J.M.* The pachytene checkpoint // *Trends in Genetics*. 2000. V. 16. № 9. P. 395–403. [https://doi.org/10.1016/S0168-9525\(00\)02080-1](https://doi.org/10.1016/S0168-9525(00)02080-1)
30. *Bazykin A.D.* Hypothetical mechanism of speciation // *Evolution*. 1969. V. 23. № 4. P. 684–687. <https://doi.org/10.2307/2406862>
31. *Feder J.L., Egan S.P., Nosil P.* The genomics of speciation-with-gene-flow // *Trends in Genetics*. 2012. V. 28. № 7. P. 342–350. <https://doi.org/10.1016/j.tig.2012.03.009>
32. *Gimenez M.D., White T.A., Hauffe H.C. et al.* Understanding the basis of diminished gene flow between hybridizing chromosome races of the house mouse // *Evolution*. 2013. V. 67. № 5. P. 1446–1462. <https://doi.org/10.1111/evo.12054>
33. *Tigano A., Khan R., Omer A.D. et al.* Chromosome size affects sequence divergence between species through the interplay of recombination and selection // *Evolution*. 2022. V. 76. № 4. P. 782–798. <https://doi.org/10.1111/evo.14467>

Variations in Chromosome Synapsis at Meiotic Prophase I of Mole Voles *Ellobius tancrei* Heterozygous for Robertsonian Translocations

S. N. Matveevsky^{a, *}, Yu. F. Bogdanov^a, E. A. Lyapunova^b,
I. Yu. Bakloushinskaya^b, and O. L. Kolomiets^{a, **}

^aVavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, 119991 Russia

^bKoltzov Institute of Developmental Biology, Russian Academy of Sciences, Moscow, 119334 Russia

*e-mail: sergey8585@mail.ru

**e-mail: olkolomiets@mail.ru

We have shown that different combinations of meiotic configurations (different number of trivalents and different chromosome chain structure) in the same individual can be formed in heterozygotes with the Robertsonian translocations in the meiotic prophase I. Two types of experimental hybrids of the eastern mole vole *Ellobius tancrei* which are heterozygous for four Robertsonian translocations were studied here. Instead of the expected four trivalents, different types of meiotic configurations at the pachytena stage up to 10-element chains were identified. We suggest that the probability of passing meiosis in different cells depends on the structure of the formed chains and possibility of their correction. Such variations in chromosome synapsis during multivalent formation may lead to a decrease in gametes production but not to a complete stop of gametogenesis, which provides the background for the maintenance of Robertsonian translocations in the population.

Keywords: synaptonemal complex, meiosis, trivalent, chromosomal rearrangements.