

ISSN 0016-6758

Том 59, Номер 11

Ноябрь 2023



ГЕНЕТИКА



www.sciencejournals.ru



СОДЕРЖАНИЕ

Том 59, номер 11, 2023

Обзорные и теоретические статьи

Гены рецепторов гормонов, влияющих на яичную продуктивность и воспроизводительные качества кур

Е. И. Куликов, Л. Г. Коршунова, Р. В. Карапетян, А. С. Комарчев, А. К. Кравченко, Д. М. Дмитренко, В. А. Попов, В. Н. Мартынова, Л. И. Малахеева, Д. Н. Ефимов 1203

Изменение продолжительности жизни как интегральный ответ на иммунный статус организма и активность мобильных элементов

М. В. Тростников, Д. Р. Малышев, Е. Г. Пасюкова 1212

Генетика растений

Локализация хромосомных регионов, определяющих содержание магния и кальция у образцов риса

Ю. К. Гончарова, В. В. Симонова, С. В. Гончаров, Н. А. Очкас 1219

Распространение и сила аллелей генов гибридного некроза в генотипах канадских сортов яровой мягкой пшеницы (*Triticum aestivum* L.)

В. А. Пухальский, В. С. Рубец, Е. Н. Билинская, А. М. Кудрявцев 1230

Генетика животных

Причины морфологической и генетической неоднородности сигов *Coregonus lavaretus* sensu lato арктической части бассейна р. Лены

Н. А. Бочкарев, Д. С. Сендек, Е. И. Зуйкова, Л. А. Пестрякова, Л. А. Ушницкая, Е. С. Захаров, А. А. Эверстова, Н. Н. Захарова, Д. В. Политов 1235

Генетическая оценка прогнозируемого остаточного потребления корма и экспрессия значимых генов-кандидатов свиней породы дюрок и товарных помесей второго поколения

А. А. Белоус, А. А. Сермягин, Н. А. Зиновьева 1253

Изучение генетического разнообразия и популяционной структуры осетинской породы в сравнении с другими породами грубошерстных жирнохвостых овец на основе данных SNP-генотипирования

Т. Е. Денискова, А. В. Доцев, М. И. Селионова, Н. А. Зиновьева 1270

Генетика человека

Оценка ассоциации степени метилирования ДНК и частоты хромосомных aberrаций лимфоцитов человека при однократном облучении крови *in vitro*

О. С. Цымбал, Д. С. Исубакова, Е. В. Брониковская, А. Ф. Николаева, В. О. Сигин, А. И. Калинин, Ж. А. Старцева, Н. В. Литвяков, И. В. Мильто, Р. М. Тахауов 1282

Молекулярный ландшафт рака полости рта у молодых пациентов

Е. А. Простакишина, Т. Д. Дампилова, Л. А. Кононова, П. С. Ямщиков, М. Р. Патышева, Е. С. Колегова, Е. Л. Чойнзонов, Е. В. Денисов 1290

Определение полиморфизма *G/C* в позиции *xp20:37352001* в препаратах ДНК из крови человека методами *GlaI*- и *Bst2UI*-ПЦР анализа

А. Г. Акишев, Н. А. Нетесова, М. А. Абдурашитов, С. Х. Дегтярев 1303

Характеристика спектра нуклеотидных замен митохондриальной ДНК в популяциях человека в условиях высокогорья

Б. А. Малярчук 1313

Особенности распространения гаплогрупп Y-хромосомы в населении Санкт-Петербурга в связи с проблемой создания референтных баз данных <i>И. Г. Удина, А. С. Грачева, С. А. Боринская, О. Л. Курбатова</i>	1319
--	------

Математические модели и методы

Описание дивергенции субпопуляций в иерархической системе при анализе изонимии. II. Вероятности неизонимных встреч <i>В. П. Пасеков</i>	1326
--	------

Краткие сообщения

Поиск генов-кандидатов живой массы у овец южной мясной породы <i>В. С. Шевцова, А. Я. Куликова, Л. В. Гетманцева, А. В. Усатов</i>	1341
---	------

Contents

Vol. 59, No. 11, 2023

Reviews and Theoretical Articles

Genes of Hormone Receptors Affecting Egg Productivity and Reproductive Qualities of Chickens

E. I. Kulikov, L. G. Korshunova, R. V. Karapetyan, A. S. Komarchev, A. K. Kravchenko, D. M. Dmitrenko, V. A. Popov, V. N. Martynova, L. I. Malakheeva, and D. N. Efimov 1203

Changes in Lifespan as an Integral Response to the Organism's Immune Status and Mobile Elements Activity

M. V. Trostnikov, D. R. Malyshev, and E. G. Pasyukova 1212

Plant Genetics

Localization of Chromosomal Regions Determining Magnesium and Calcium Content in Rice Varieties

J. K. Goncharova, V. V. Simonova, S. V. Goncharov, and N. A. Ochkas 1219

Distribution and Allele Strength of Hybrid Necrosis Genes in the Genotypes of Canadian Cultivars of Spring Bread Wheat (*Triticum aestivum* L.)

V. A. Pukhalskiy, V. S. Rubets, E. N. Bilinskaya, and A. M. Kudryavtsev 1230

Animal Genetics

Causes of Morphological and Genetic Heterogeneity of White Fish *Coregonus lavaretus* sensu lato in the Arctic Part of the Lena River Basin

N. A. Bochkarev, D. S. Sendek, E. I. Zuikova, L. A. Pestryakova, L. A. Ushnitskaya, E. S. Zakharov, A. A. Everstova, N. N. Zakharova, and D. V. Politov 1235

Genetic Assessment of Projected Residual Feed Consumption and Expression of Significant Candidate Genes in Duroc Pigs and Second-Generation Commercial Blends

A. A. Belous, A. A. Sermyagin, and N. A. Zinovieva 1253

Study of Genetic Diversity and Population Structure of the Osetin Breed in Comparison with Other Coarse Wool Fat-Tailed Sheep Breeds Based on SNP Genotyping Data

T. E. Deniskova, A. V. Dotsev, M. I. Selionova, and N. A. Zinovieva 1270

Human Genetics

Assessing the Association of the Degree of DNA Methylation and the Frequency of Chromosomal Aberrations in Human Lymphocytes in a Single Irradiation of Blood *in vitro*

O. S. Tsymbal, D. S. Isubakova, E. V. Bronikovskaya, A. F. Nikolaeva, V. O. Sigin, A. I. Kalinkin, Zh. A. Startseva, N. V. Litviakov, I. V. Milto, and R. M. Takhaouov 1282

Molecular Landscape of Oral Cancer in Young Adults

E. A. Prostackishina, T. D. Dampilova, L. A. Kononova, P. S. Iamshchikov, M. R. Patysheva, E. S. Kolegova, E. L. Choinzonov, and E. V. Denisov 1290

Determination of <i>G/C</i> Polymorphism at chr20:37352001 Position in Human Blood DNA Preparations by <i>GlaI</i> - and <i>Bst2UI</i> -PCR Analyses Methods	
<i>A. G. Akishev, N. A. Netesova, M. A. Abdurashitov, and S. Kh. Degtyarev</i>	1303

Characterization of the Spectrum of Mitochondrial DNA Nucleotide Substitutions in Human Populations in High Altitude Environments	
<i>B. A. Malyarchuk</i>	1313

Distribution Peculiarities of Y-Chromosome Haplogroups in the Population of St. Petersburg in Connection with the Problem of Creation Reference Data Bases	
<i>I. G. Udina, A. S. Gracheva, S. A. Borinskaya, and O. L. Kurbatova</i>	1319

Mathematical Models and Methods

Description of Divergence of Subpopulations in the Hierarchical System under the Analysis of Isonymy. II. Probabilities of Non-Isonymic Encounters	
<i>V. P. Passekov</i>	1326

Short Communications

Search for Candidate Genes for Live Weight in Southern Meat Breed Sheep	
<i>V. S. Shevtsova, A. Ya. Kulikova, L. V. Getmantseva, and A. V. Usatov</i>	1341

ОБЗОРНЫЕ
И ТЕОРЕТИЧЕСКИЕ СТАТЬИ

УДК 636.082.12

ГЕНЫ РЕЦЕПТОРОВ ГОРМОНОВ, ВЛИЯЮЩИХ НА ЯИЧНУЮ
ПРОДУКТИВНОСТЬ И ВОСПРОИЗВОДИТЕЛЬНЫЕ КАЧЕСТВА КУР

© 2023 г. Е. И. Куликов¹, *, Л. Г. Коршунова¹, Р. В. Карапетян¹, А. С. Комарчев¹,
А. К. Кравченко¹, Д. М. Дмитренко¹, В. А. Попов¹, В. Н. Мартынова¹,
Л. И. Малахеева¹, Д. Н. Ефимов¹

¹Федеральный научный центр “Всероссийский научно-исследовательский и технологический институт
птицеводства” Российской академии наук, Московская область, Сергиев Посад, 141311 Россия

*e-mail: kulikovegor33@yandex.ru

Поступила в редакцию 29.03.2023 г.

После доработки 18.05.2023 г.

Принята к публикации 02.06.2023 г.

Яичная продуктивность кур носит сложный полигенный тип наследования, контролируется многими генами, и является результатом сложного процесса, регулируемого гипоталамо-гипофизарно-гонадной системой. В обзоре обобщена информация о влиянии полиморфизмов генов рецепторов гормонов: фолликулостимулирующего (FSHR), лютеинизирующего (LHCGR), прогестерона (PR) и пролактина (PRLR) на яичную продуктивность и воспроизводительные качества кур. Представленные данные показывают, что полиморфизмы этих генов являются перспективными для использования в селекционных программах с целью улучшения яичной продуктивности и воспроизводительных качеств кур.

Ключевые слова: ген, гормон, рецептор, полиморфизм, куры, яйценоскость.

DOI: 10.31857/S0016675823110073, **EDN:** NJHKEA

Дальнейшее развитие отечественного птицеводства невозможно без внедрения новых, базирующихся непосредственно на анализе наследственной информации, методов оценки продуктивности птицы. Благодаря распространению новых методов молекулярной биологии в современной генетике стало возможным значительно ускорить процесс селекции в птицеводстве за счет обнаружения и изучения полиморфизмов генов, связанных с продуктивными и хозяйственно полезными признаками птицы. Это позволяет проводить отбор птицы по генотипу, а не только на основе анализа фенотипа как в традиционной селекции, и тем самым полнее учесть генетический потенциал птицы [1, 2], невзирая на возраст, физиологическое состояние и условия ее содержания. Яичная продуктивность кур носит сложный полигенный тип наследования и контролируется многими генами. Методами молекулярной генетики на сегодняшний день выявлено 66 локусов, связанных с особенностями яйцекладки и 223 локуса, связанных с качеством яиц [2].

Яйценоскость птиц является результатом сложного процесса, регулируемого гипоталамо-гипофизарно-гонадной системой. К гормонам гипофиза, непосредственно участвующим в регуляции функционирования репродуктивной системы,

относят фолликулостимулирующий (FSH) и лютеинизирующий (LH) гормоны. Оба эти гормона принадлежат семейству гликопротеинов и имеют между собой ярко выраженное структурное сходство. Также к гормонам, влияющим на яичную продуктивность, относят прогестерон (PG) и пролактин (PRL).

Действие гормона пролактина опосредовано его рецептором (PRLR), ген которого является важным геном-регулятором роста и дифференцировки клеток. Выявленные полиморфизмы этого гена в основном рассматриваются с точки зрения признаков яйценоскости [3]. Предполагается, что ген PRLR может быть основным геном, ответственным за достижение птицей половой зрелости [4]. PRLR тесно связан с рецептором гормона роста и является членом семейства цитокиновых рецепторов [5]. Это — единственный трансмембранный белок, принадлежащий к классу I рецептора цитокинов [6]. В соответствии с его разнообразными эффектами PRLR экспрессирует во многих тканях у кур и индеек [7, 8], а уровни рецептора в тканях-мишенях повышаются и понижаются в ответ на изменения концентрации циркулирующего PRL. Примерно во время вывода, секреция PRL значительно увеличивается у цыплят и индюшат в соответствии с увеличением

PRLR. Предположительно, увеличение циркулирующего **PRL** связано с адаптацией эмбриона к жизни *ex ovo*, хотя в каком качестве — пока неясно [9]. Таким образом, пролактин и его рецептор участвуют в росте и развитии, контроле водного и электролитного баланса, репродукции, эндокринной сигнализации и метаболизме. В связи с различной биологической активностью, приписываемой **PRL** и **PRLR**, они могут быть использованы в качестве основных генов-кандидатов в программах молекулярной селекции животных [10].

С целью идентификации генов, лежащих в основе признаков яйцекладки как у кур, так и других видов домашней птицы, важно изучение молекулярного механизма иерархии фолликулов, которая регулируется **FSH** посредством связывания с его рецептором (**FSHR**) [11]. Поскольку **FSH** действует исключительно через **FSHR**, механизмы, контролирующие экспрессию рецепторов, определяют популяцию клеток, реагирующих на **FSH**, и влияют на их чувствительность к гормонам. Таким образом, экспрессия **FSHR** определяет не только мишени, но и степень действия **FSH**, в конечном итоге направляя гормональный ответ на клетки гранулезы в яичнике и клетки Сертоли. В гранулезных клетках яичников временные изменения в передаче сигналов **FSH** регулируют ряд транскрипционных, метаболических и гормональных активностей, которые важны для событий пролиферации и дифференцировки, необходимых для роста фолликулов и созревания ооцитов. В тестикулярных клетках Сертоли действие **FSH** меняется по мере развития семенника [3, 12].

Не менее важным гормоном, влияющим на размножение кур, является прогестерон (**PG**) — гормон желтого тела. Было показано, что прогестерон положительно связан с яйценоскостью у индеек и фазанов [2]. Синтез прогестерона инициируется переносом холестерина из наружной митохондриальной мембраны во внутреннюю с помощью стероидогенного острого регуляторного белка (**StAR**). Расщепление боковой цепи холестерина цитохрома **P450 (P450_{scc})** превращает холестерин в прегненолон, который выходит из митохондрий и попадает в гладкий эндоплазматический ретикулум и, таким образом, превращается в прогестерон 3β -гидроксистероиддегидрогеназой (**3 β -HSD**) [12, 13]. Данный гормон действует через свой специфический рецептор (**PR**), который действует как ядерный фактор транскрипции. Прогестерон участвует в синтезе белков клеток яйцевода и эпителия. Уровень рецептора прогестерона (**PRs**) повышается после обработки кур эстрадиолом. Помимо прямого воздействия на процесс синтеза прогестерона, существует множество других причин, которые могут косвенно влиять на синтез прогестерона. Пролиферация гранулезных клеток тесно связана с секрецией **PG** [13]. Достоверно не известно, где расположены линии

клеток, экспрессирующие рецептор прогестерона у кур.

Известно, что **LH** является одним из основных эндокринных факторов, играющих важную роль в контроле фолликулогенеза и функций яичников. Овуляторный цикл домашней курицы имеет характерную продолжительность 24–25 ч и является результатом сложных взаимодействий между двумя регуляторными системами в рамках гипоталамо-гипофизарно-гонадной оси. Гипоталамус и гипофиз работают согласованно и реагируют на стероиды яичников. Ближе к завершению созревания фолликулов, фолликул **F1**, предназначенный для овуляции, увеличивает выработку прогестерона, чтобы в результате произвести овуляторный сигнал, закодированный как всплеск **LH**. В свою очередь, выход яйцеклетки из фолликула зависит от всплеска **LH** [14].

В исследовании D. Li с соавт. [15] в качестве гена-кандидата для репродуктивных признаков был рассмотрен ген рецептора лютеинизирующего гормона — хориогонадотропина (**LHCGR**). Он пересекает плазматическую мембрану с семью высоко консервативными α -спиралями, ориентированными с внеклеточным аминоконцом и внутриклеточным карбоксиконцом. После связывания с лигандами (лютеинизирующим гормоном) он активирует аденилилциклазу и стимулирует дифференцировку клеток и стероидогенез в фолликулярных клетках [15].

Благодаря активному развитию современных методов генотипирования и полногеномного ассоциативного анализа стало возможным выявлять множество однонуклеотидных замен (**SNP**) связанных с яйценоскостью кур. Это определяет целесообразность и необходимость изучения полиморфизмов генов, как самих гормонов, так и их рецепторов, от которых зависит физиологический ответ. В табл. 1 представлена сводная информация по изучаемым рецепторам гормонов, влияющих на яичную продуктивность и воспроизводительные качества кур. На рис. 1 представлено расположение изучаемых рецепторов в геноме кур. Рис. 1 выполнен с использованием архива данных **NCBI**.

Ген **PRLR** расположен в **Z**-хромосоме, содержит 15 экзонов, 14 интронов, и его общая длина составляет 20890 пн. Обнаружены и изучены **SNP** в различных участках гена и их связь с продуктивными признаками кур. Так, например аллельный вариант **BamHI+** в пятом экзоне гена положительно коррелирует с яичной продуктивностью [16]. Ген **PRLR** экспрессирует во многих тканях, например, в семенниках, яичниках, семявыносящих протоках, яйцеводе, почках, в тонком и толстом кишечнике, а также в гипоталамусе. Также предполагается, что ген **PRLR** может быть основным геном, ответственным за достижение поло-

Таблица 1. Рецепторы гормонов, связанные с воспроизводительными признаками и яичной продуктивностью кур

Рецептор	Характеристика	Функции	Авторы
PRLR	Расположен в Z-хромосоме; содержит 15 экзонов, 14 интронов, его общая длина составляет 20890 пн; экспрессируется в семенниках, яичниках, семявыносящих протоках, яйцеводе, почках, тонком и толстом кишечнике, гипоталамусе	Является важным геном-регулятором роста и дифференцировки клеток, влияет на возраст наступления половой зрелости у кур	[3, 6, 16–19]
PR	Расположен в хромосоме 1; имеет 2 изоформы: рецептор прогестерона-A (A-субъединица, PR-A) и рецептор прогестерона-B (B-субъединица, PR-B), которые возникают из одного гена путем поочередной инициации трансляции с одного гена и функционируют как регуляторы транскрипции генов, чувствительных к прогестерону	Обеспечивает регуляцию репродуктивной функции; две изоформы рецептора функционируют как регуляторы транскрипции генов, чувствительных к прогестерону	[7, 9, 20]
LHCGR	Расположен в хромосоме 3 в позиции 45.35; связан с белками G, которые опосредуют действие лютеинизирующего гормона (LH) и хорионического гонадотропина (HG)	Активирует аденилилциклазу, стимулирует дифференцировку клеток и стероидогенез в фолликулярных клетках после связывания с LH; индуцирует овуляцию из зрелого фолликула в яичнике, стимулирует выработку прогестерона в желтом теле	[14, 15, 35]
FSHR	Располагается в хромосоме 3, состоит из 15 экзонов, существует 5 изоформ; входит в состав гликопротеинового семейства рецепторов, связанных с G-белком, находится на клетках мишенях; экспрессируется в клетках Сертоли и клетках гранулозы яичников	Влияет на яичную продуктивность; уровень его экспрессии тесно связан с дифференцировкой и созреванием половых клеток, определяет не только мишени, но и степень действия FSH, в конечном итоге направляя гормональный ответ на клетки гранулозы в яичнике и клетки Сертоли	[11–13, 21–27, 29]

вой зрелости. Исследования, проведенные на сегодняшний день, показали, что SNP, наблюдаемые в разных местах гена *PRLR*, связаны в основном с признаками яйценоскости. Например полиморфизм G1836C в шестом экзоне был достоверно связан с возрастом снесения первого яйца. Таким образом, ген *PRLR* может быть либо основным геном, влияющим на возраст кур при снесении первого яйца, либо молекулярным маркером в тесном сцеплении с таким геном [6].

В преовуляторных фолликулах *PRLR* экспрессировался на более высоких уровнях в гранулезах, чем в слоях теки. Гранулеза и тека-клетки — это группы эндокринных клеток в яичнике, состоящей из соединительной ткани, окружающей фолликул. *FSH* оказывал наибольший стимулирую-

щий эффект на экспрессию *PRLR* и *Star* в культивированных гранулезных клетках фолликулов размером 6–8 мм, но этот эффект снижался по мере созревания фолликулов до размера в 10–12 мм (F1). В отличие от этого, LH не изменял экспрессию *PRLR* в гранулезных клетках всех классов фолликулов, но повышал уровень *Star* в гранулезных клетках с размерами от 8 до 12 мм (F2 и F1). Эти результаты показывают, что *PRLR* дифференцированно распределяется и регулируется *FSH* или *PRL* независимо или в комбинации в фолликулярной иерархии [17].

Ранее была обнаружена инсерция длиной 24 пн, происходящая в промоторе, –377 ~ –354 *PRL* (GenBank accession no. AB011434). Особи без инсерционной последовательности на данном

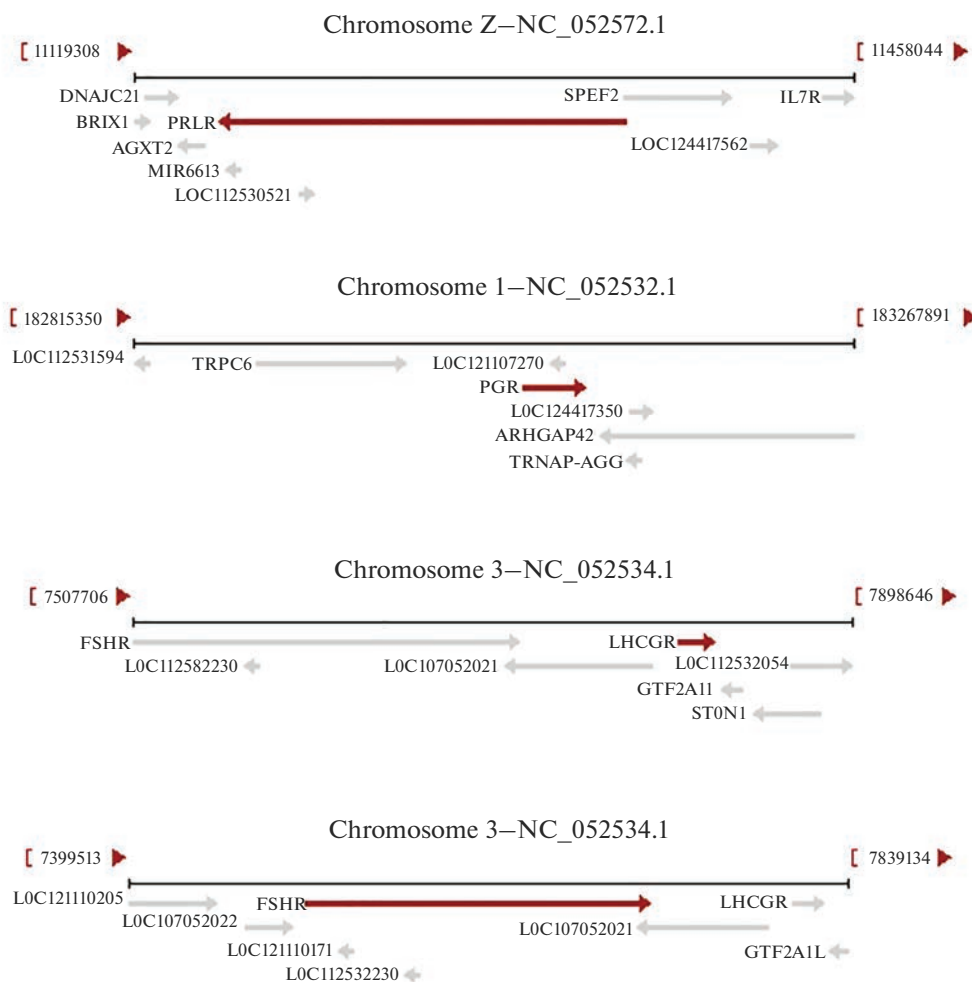


Рис. 1. Расположение рецепторов гормонов в геноме *Gallus gallus*.

участке были ассоциированы с проявлением эффекта насиживания ($p < 0.05$). Кроме того, все белые леггорны были $+/+$ для данной инсерции, в то время как местные породы с очень сильным эффектом насиживания почти все были $-/-$. Гомозиготная вставка последовательности 24 пн в *PRL*-промоторе может снизить экспрессию *PRL*. Результаты показали, что данная инсерция может быть генетическим маркером в селекции против эффекта насиживания кур [18].

Был определен локус (G-10T), расположенный в нуклеотиде десятого участка 5'-нетранслируемой области гена *PRL* у гусей. Породы гусей с более высокими частотами аллелей *T* имели более высокую яичную продуктивность, чем породы с более высокими частотами аллеля *G* [19]. Также было проведено исследование экзона 10 гена *PRLR* и определено три генотипа. Последовательность генотипа *AB* совпадает с последовательностью оригинальной последовательности (NCBI

Accession: DQ660982). В экзоне 10 генотипа *AA* имеется пять мутаций: C → A на 840 пн, C → T на 862 пн, T → C на 875 пн, T → A при 963 пн, A → T на 989 пн. Эти мутации приводят к следующим аминокислотным мутациям: His → Asn, Thr → Ile, Asn → Lys, Thr → Ser и синонимичной мутации в 875 пн. В экзоне 10 генотипа *AC* имеется пять мутаций: генотип: G → T на 816 пн, A → T на 861 пн, C → T в 862 пн, T → C при 875 пн, A → G на уровне 948 пн. Они также вызывают аминокислотные мутации: Val → Phe, Thr → Phe, синонимичные мутации в 875 и 948 пн. Годовая яйценоскость гусей генотипа *AB* в среднем значительно выше, чем у гусей генотипов *AA* и *AC* ($p < 0.05$) [19].

Фолликулостимулирующий гормон и его рецептор играют важную физиологическую роль в репродуктивной функции животных. FSH представляет собой гликопротеин, синтезируемый и секретируемый гонадотропными клетками передней доли гипофиза [20]. В 1996 г. последова-

тельность кДНК *FSHR* была впервые успешно клонирована из ткани куриного яичника [21]. Исследования показали, что *FSHR* избирательно экспрессируется в клетках Сертоли и клетках гранулозы яичников, и уровень его экспрессии тесно связан с дифференцировкой и созреванием половых клеток. Полиморфизмы в промоторе гена *FSHR* кур могут также влиять на транскрипцию *FSHR* и влиять на яичную продуктивность у кур [22].

Анализ транскриптомов куриных фолликулов, различающихся по экспрессии мРНК *FSHR*, показал, что Wnt — сигнальный путь, являющийся важным регулятором эмбриогенеза и клеточной дифференцировки у позвоночных и насекомых, был значительно обогащен в фолликулах с наибольшим изменением экспрессии *FSHR* [23]. WNT- β -катениновая сигнализация участвует во множестве процессов развития и поддержания гомеостаза тканей, регулируя пролиферацию, дифференцировку, миграцию, генетическую стабильность и апоптоз клеток, а также поддерживая стволовые клетки в плюрипотентном состоянии.

Путем комплексного анализа транскриптомов куриных фолликулов, член семейства доменов анкириновых повторов А (гена *SOWAHA*) был определен как сетевой ген, дифференциально экспрессируемый между фолликулами с выраженной экспрессией *FSHR*. Дальнейший анализ *in vitro* показал, что на *SOWAHA* влияет *FSH* в клетках гранулозы фолликулов яичников. *SOWAHA* влияет на экспрессию генов, участвующих в иерархии куриных фолликулов, и ингибирует пролиферацию гранулезных клеток, что позволяет предположить ингибирующую роль в отборе куриных фолликулов [24].

В промоторе гена *FSHR* было выявлено шесть SNP на позициях 10, 51, 59, 121, 233, 331 нуклеотидов от его начала. Самая высокая яйценоскость у цыплят помеси BC1 наблюдалась у гаплотипа TTGCYA, а самая низкая яйценоскость у гаплотипа TTGYYG. По результатам корреляционного анализа была выявлена положительная корреляция между гаплотипом TTGCYA и яйценоскостью [25].

При анализе промотора гена *FSHR* другой исследовательской группой было выявлено 49 вариаций, из которых 39 являются однонуклеотидными заменами, одна нуклеотидная замена (TTG) на (CAC) и девять инделов. Полиморфизм на участке —874 (индел длиной 200 пн) был идентифицирован и проанализирован с влиянием на признаки яичной продуктивности, а также на экспрессию генов. В этом сайте аллель I+ был доминантным у Lohmann Brown и Синьян Браун, но очень редким в трех китайских местных породах кур, а именно Jining Bairei, Wenchang, Zang. В популяции Синьян Браун полиморфизм был связан с возрастом первого яйца (AFE) ($p < 0.05$), но его влияние на количество яиц в возрасте 37 нед.

(E37) и количество яиц в возрасте 57 нед. (E57) не было значительно ($p > 0.05$) [26].

Другие исследователи обнаружили пять полиморфизмов, включая индел —869, C—1684T, C—1608T, G—368A и T—238A у цыплят. Возраст снесения первого яйца значительно различался ($p < 0.01$) для разных генотипов индела —869 у сукенских цыплят. Для SNP C—1684T у кур Донгсианг, генотип CC имел значительно большее количество яиц в возрасте 43 нед. (E43), чем генотип TC ($p < 0.05$). Для SNP C—1684T у цыплят Сукен, генотип TC имел более высокие показатели возраста наступления половой зрелости (AFE), чем генотип CC ($p < 0.05$). Для SNP C—1608T у сукенских цыплят, генотип CC имел более высокую AFE, чем генотип CC ($p < 0.05$). Для SNP G—368A у сукенских кур, генотип AG имел более высокий AFE, чем генотип GG ($p < 0.05$) [22].

Нами была проанализирована однонуклеотидная замена rs317093289 в гене *FSHR* у исходной линии CM9 породы плимутрок бройлерного кросса “Смена 9”. Наиболее часто встречающимся (42% среди исследованных кур) был генотип TA, генотип TT встречался с частотой 24%, а генотип AA — 34%. По массе яиц в 210-дневном возрасте птица с генотипом TA превосходила птицу с генотипом AA на 2.4%, группа с генотипом TT была приближена к группе TA, разность между группами TT и TA не достоверна. Изучаемый SNP оказал значимый эффект на показатели яйценоскости. Генотип TA превосходил генотип TT по количеству снесенных яиц за 210 и 308 дней на 15.0 и 2.8% соответственно [27].

Прогестерон (PR), являясь одним из наиболее важных гормонов, влияющих на размножение кур, регулирует репродуктивную функцию через рецептор, у которого описаны две изоформы: рецептор прогестерона-А (А-субъединица, PR-A) и рецептор прогестерона-В (В-субъединица, PR-B), которые возникают из одного гена путем поочередной инициации трансляции и функционируют как регуляторы транскрипции генов, чувствительных к прогестерону [7, 9]. PR-B (110 кДа) является полной изоформой, а PR-A (79 кДа) — N-концевой усеченной изоформой, в которой отсутствуют аминоконцевые 128 аминокислот. PR-изоформы происходят из одного гена и образуются с альтернативных сайтов начала трансляции. Соотношение экспрессии изоформ PR варьирует в тканях-мишенях при различном гормональном фоне и в разнообразных условиях окружающей среды, например, в период половой зрелости птиц и в разные сезоны года [22, 28, 29]. Следует отметить и влияние прогестерона в сочетании с эстрадиолом на половое поведение у птиц путем активации соответствующих участков мозга [22].

Обширный структурно-функциональный анализ рецептора прогестерона у кур путем исследо-

ваний делеции (утрата сегмента ДНК, размер которого может быть от одного нуклеотида до субхромосомного фрагмента из нескольких генов) и направленного на сайт мутагенеза помог выделить и определить три основных функциональных домена в молекуле рецептора. Он состоит из переменного аминотерминального домена (который, как полагают, оказывает модулирующее действие на транскрипцию генов-мишеней), а также короткого, богатого цистеином центрального ДНК-связывающего домена, который является высококонсервативным среди всех стероидных рецепторов и карбокси-концевого лиганд-связывающего домена [29].

Важные результаты о работе гена прогестерона были получены на других видах животных. Так, исследования на козах показали, что SNP G109519T в гене *PGR* может быть использован в качестве молекулярного маркера для признака плодовитости и производства молока у Ляонинских кашемировых коз [30].

В исследовании на кроликах было показано, что ассоциация SNP 2464G > A, обнаруженного в промоторной области гена рецептора прогестерона, была связана с экспрессией рецептора прогестерона. Генотип *GG* показал меньшую экспрессию PR-B и PR-A, чем генотип *AA* в яйцевом (GG/AAPR-B = 0.81 и GG/AAPR-A = 0.73) и матке (около 0.70 для обеих изоформ). Генотип *GA* показал аналогичную экспрессию PR-A в обеих тканях, а также аналогичную экспрессию PR-B в яйцевом по сравнению с генотипом *GG*. И наоборот, генотип *GG* показал меньшую экспрессию PR-B, чем генотип *GA* в матке (GG/GAPR-B = 0.79). Сходная экспрессия обеих изоформ PR была обнаружена в матке на втором и третьем днях беременности; в то же время в яйцевом наблюдалось увеличение обеих изоформ. Была получена тенденция на снижение выживаемости, наблюдаемой у крольчих с генотипом *AA* по сравнению с генотипом *GG* [31].

При исследовании на курах основной упор делается на соотношение изоформ рецептора прогестерона, но комплексно не рассматриваются однонуклеотидные полиморфизмы в данном гене, в то время как при исследованиях на людях многие ученые [32] говорят о взаимосвязи ряда полиморфизмов данного гена с риском развития рака молочной железы. Частоты минорных аллелей полиморфизмов rs1042838, rs590688 и rs10895068 гена *PGR* были значительно выше у больных раком молочной железы по сравнению с контрольной группой. Пациенты, несущие генотипы rs1042838 G/T, rs590688 C/C и rs10895068 G/A, имели более высокий риск развития рака молочной железы, в то время как носительство генотипа rs3740753 G/G было связано с незначительным снижением риска развития рака молочной желе-

зы [32]. Таким образом, данный ген является возможным перспективным для исследования его влияния на репродуктивные показатели кур.

Лютеинизирующий гормон играет важную роль в процессе полового созревания и яйцекладки у кур. Его рецептор относится к рецепторам, связанным с белками G, которые опосредуют действие LH и хорионического гонадотропина (HG). *LHCGR* экспрессируется преимущественно в гонадах обоих полов, а также в других тканях, включая мозг. Рецептор регулирует развитие гонад и производство гамет. В яичнике он индуцирует овуляцию из зрелого фолликула и стимулирует выработку прогестерона в лютеиновом теле, которое отвечает за предотвращение наступления цикла овуляции и подготовку к оплодотворению. Дефектные рецепторы LH приводят к нарушениям овуляции и оплодотворения. Фактически, многочисленные мутации вызывают бесплодие, делая рецептор рецессивно неактивным или доминантно активным. Пока не ясно, являются ли мутации в рецепторе LH курицы гена ответственными за влияние на яичную продуктивность [15].

Проведены исследования, в результате которых были обнаружены мутации в гене *LHCGR* и оценена их связь с суперовуляцией. Изучены полиморфизмы в *LHCGR* и генотипы, связанные с признаками суперовуляции у 127 китайских голштинских телок. Полиморфизм G/T (ss52050737) в экзоне 11 был значительно связан с общим количеством яйцеклеток и количеством переносимых эмбрионов [33]. Четыре SNPs G51656T, A51703G, A51726G и G51737A были идентифицированы в интроне 9 гена *LHCGR* у 171 китайской голштинской коровы, подвергнутой суперовуляции, и оценили их ассоциации с суперовуляторным ответом. Ассоциативный анализ показал, что эти четыре SNP оказывают значительное влияние на общее количество яйцеклеток ($p < 0.05$). Более того, полиморфизмы A51703G и A51726G значительно ассоциировались с количеством переносимых эмбрионов ($p < 0.05$) [34].

Сегрегационный анализ показал, что ген рецептора лютеинизирующего гормона/хориогонадотропина курицы расположен на хромосоме 3 [35]. В одном из исследований [15] в гене *LHCGR* обнаружены SNP, связывающие такие признаки, как живая масса при снесении первого яйца и возраст наступления половой зрелости (AFE). Гетерозигота может быть доминантной, так как особи с генотипами *AG* или *TG* имели более высокие значения признаков, чем гомозиготные птицы, демонстрируя, что особи с генотипами *AG* и *TG* сносят первое яйцо в более раннем возрасте и имеют более высокие показатели по количеству снесенных яиц за 300 дней (EN), чем особи с генотипами *AA*, *GG* и *TT*. Таким образом, эти гетерозиготные особи имеют более высокую приспособ-

собленность, чем гомозиготные. Это очевидное репродуктивное преимущество генотипов *TG*, имеющих более низкий AFE и более высокий EN, согласуется с избытком гетерозиготных генотипов в этом локусе, что указывает на сильный балансирующий отбор (сверхдоминирование) [15]. Было показано, что чем меньше живая масса кур в начале кладки и чем раньше наступает AFE, тем больше влияние на общее количество яиц в возрасте 300 дней. Сделан вывод, что ген *LHCGR* значительно связан как с AFE, так и с EN у кур [15].

Таким образом, исследованные в настоящем обзоре гены рецепторов гормонов представляются перспективными для использования в практической селекции кур. В настоящее время активно ведутся исследования по генотипированию пород кур мясного, яичного и комбинированного направления продуктивности. Проводится поиск полиморфных вариантов ключевых генов, которые улучшают показатели яйценоскости. Накоплен значительный объем информации по значимым ассоциациям различного типа ДНК-маркеров с продуктивными и хозяйственно ценными показателями у кур локальных и некоторых коммерческих пород в течение продуктивного цикла, что может и должно быть использовано в маркерной селекции кур, а также при разработке селекционно-генетических программ с целью обеспечения конкурентоспособности отечественного птицеводства.

Статья подготовлена в рамках выполнения государственного задания Министерства науки и высшего образования РФ по теме “Разработать селекционно-генетические методы повышения выхода племенной и товарной продукции от сельскохозяйственной птицы” (№ гос. рег. 121030100022-8).

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. Алтухов Ю.П., Салменкова Е.А. Полиморфизм ДНК в популяционной генетике // Генетика. 2002. Т. 38. № 9. С. 1173–1195.
2. Новгородова И.П. Взаимосвязь яичной продуктивности кур и перепелов с локусами количественных признаков // Эффективное животноводство. 2018. № 7(146). С. 24–26.
3. Kulibaba R.A. Polymorphism of growth hormone, growth hormone receptor, prolactin and prolactin receptor genes in connection with egg production in Poltava clay chicken // Agricultural Biology. 2015. V. 50. № 2. P. 198–207.
<https://doi.org/10.15389/agrobiology.2015.2.198eng>
4. Ling-B.L., Di-Yan L., Xiao-Ling et al. Polymorphism of prolactin receptor gene and its association with egg production traits in erlang mountainous chicken // Asian J. of Animal and Veterinary Adv. 2012. № 7. P. 1183–1190.
5. Bole-Feysot C., Goffin V., Edery M. et al. Prolactin (PRL) and its receptor: Actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice // Endocr Rev. 1998. V. 19. № 3. P. 225–268.
<https://doi.org/10.1210/edrv.19.3.0334>
6. Wilkanowska A., Mazurowski A., Mroczkowski S., Kokozyński D. Prolactin (PRL) and prolactin receptor (PRLR) genes and their role in poultry production traits // Folia Biologica. 2014. V. 62. № 1. P. 1–8.
7. Leclerc B., Zadworny D., Bédécarrats G., Kuhnlein U. Development of a real-time (Q) PCR assay to measure variation in expression of prolactin receptor mRNA in the hypothalamus and pituitary gland during late embryogenesis in turkeys and chickens // Gen. Comp. Endocrinol. 2007. V. 150. № 2. P. 319–325.
<https://doi.org/10.1016/j.ygcen.2006.08.007>
8. Leclerc B., Zadworny D., Bédécarrats G., Kuhnlein U. Ontogenesis of the expression of prolactin receptor messenger ribonucleic acid during late embryogenesis in turkeys and chickens // Poult Sci. 2007. V. 86. № 6. P. 1174–1179.
<https://doi.org/10.1093/ps/86.6.1174>
9. Hiyama G., Kansaku N., Kinoshita M. et al. Changes in post-translational modifications of prolactin during development and reproductive cycles in the chicken // Gen. Comp. Endocrinol. 2009. V. 161. P. 238–245.
10. Lü A., Hu X., Chen H. et al. Single nucleotide polymorphisms of the prolactin receptor (PRLR) gene and its association with growth traits in chinese cattle // Mol. Biol. 2011. Rep. 38. P. 261–266.
<https://doi.org/10.1007/s11033-010-0103-5>
11. Wang Y., Chen Q., Liu Z. et al. Transcriptome analysis on single small yellow follicles reveals that Wnt4 is involved in chicken follicle selection // Front. Endocrinol. 2017. V. 8.
<https://doi.org/10.3389/fendo.2017.00317>
12. Johnson P.A. Follicle selection in the avian ovary // Reprod. Domestic Animals. 2012. V. 47. S. 4. P. 283–287.
<https://doi.org/10.1111/j.1439-0531.2012.02087.x>
13. Johnson A.L., Woods D.C. Dynamics of avian ovarian follicle development: cellular mechanisms of granulosa cell differentiation // Gen. Comp. Endocrinol. 2009. V. 163. P. 12–17.
<https://doi.org/10.1016/j.ygcen.2008.11.012>
14. Tischkau S.A., Howell R.E., Hickok J.R. et al. The luteinizing hormone surge regulates circadian clock gene expression in the chicken ovary // Chronobiology International. 2010. V. 28. № 1. P. 10–20.
15. Li D.Y., Zhang L., Yang M.Y. et al. Effect of luteinizing hormone/choriogonadotropin receptor (LHCGR)

- gene on chicken reproductive traits // *Mol. Biol. Reports*. 2013. V. 40. № 12. P. 7111–7116.
16. *Rashidi H., Rahimi-Mianji G., Farhadi A., Gholizadeh M.* Association of prolactin and prolactin receptor gene polymorphisms with economic traits in breeder hens of indigenous chickens of Mazandaran province // *Iranian J. Biotechnol.* 2012. V. 10. № 2. P. 129–135.
 17. *Hu S., Duggavathi R., Zadworny D.* Regulatory mechanisms underlying the expression of prolactin receptor in chicken granulosa cells // *PLoS One*. 2017. V. 12. № 1. <https://doi.org/10.1371/journal.pone.0170409>
 18. *Jiang R.S., Xu G.Y., Zhang X.Q., Yang N.* Association of polymorphisms for prolactin and prolactin receptor genes with broody traits in chickens // *Poultry Sci.* 2005. V. 84. № 6. P. 839–845. <https://doi.org/10.1093/ps/84.6.839>
 19. *Hong-Quan C., Han-Qin W., Jie Q., Hua C.* The novel genetic change in 5'-untranslated region of goose prolactin gene and their distribution pattern in different goose breeds // *Asian J. of Animal and Veterinary Advances*. 2011. V. 6. P. 1069–1075. <https://doi.org/10.3923/ajava.2011.1069.1075>
 20. *Pierce J.G., Parsons T.F.* Glycoprotein hormones: Structure and function // *Annual Rev. of Biochemistry*. 1981. V. 50. P. 465–495.
 21. *You S., Bridgham J.T., Foster D.N., Johnson A.L.* Characterization of the chicken follicle-stimulating hormone receptor (cFSH-R) complementary deoxyribonucleic acid, and expression of cFSH-R messenger ribonucleic acid in the ovary // *Biology of Reproduction*. 1996. V. 55. № 5. P. 1055–1062. <https://doi.org/10.1095/biolreprod55.5.1055>
 22. *Li X., Lu Y., Liu X. et al.* Identification of chicken FSHR gene promoter and the correlations between polymorphisms and egg production in Chinese native hens // *Reprod. Domest. Anim.* 2019. V. 54. № 4. P. 702–711. <https://doi.org/10.1111/rda.13412>
 23. *Kang L., Cui X., Zhang Y. et al.* Identification of miRNAs associated with sexual maturity in chicken ovary by illumina small RNA deep sequencing // *BMC Genomics*. 2013. V. 14. <https://doi.org/10.1186/1471-2164-14-352>
 24. *Zhong C., Liu Z., Qiao X. et al.* Integrated transcriptomic analysis on small yellow follicles reveals that soson-dowah ankyrin repeat domain family member A inhibits chicken follicle selection // *Anim. Biosci.* 2021. V. 34. № 8. P. 1290–1302.
 25. *Rikha R.K., Indra L., Aditya R.E. et al.* The association of follicle stimulating hormone receptor (FSHR) gene polymorphism of on egg productivity in hybrid chicken (*Gallus gallus gallus*, Linnaeus 1758) // *Biodiv. J. of Biol. Diversity*. 2021. V. 22. <https://doi.org/10.13057/biodiv/d220318>
 26. *Kang L., Zhang N., Zhang Y. et al.* Molecular characterization and identification of a novel polymorphism of 200 bp indel associated with age at first egg of the promoter region in chicken follicle-stimulating hormone receptor (FSHR) gene // *Mol. Biol. Rep.* 2012. V. 39. № 3. P. 2967–2973. <https://doi.org/10.1007/s11033-011-1058-x>
 27. *Куликов Е.И., Каранетян Р.В., Коршунова Л.Г. и др.* Влияние однонуклеотидной замены rs317093289 в гене рецептора фолликулостимулирующего гормона на продуктивность исходной линии породы плимутрок бройлерного кросса “Смена 9” // *Птицеводство*. 2022. № 11. С. 4–8. <https://doi.org/10.33845/0033-3239-2022-71-11-4-8>
 28. *Johnson A.L., Solovieva E.V., Bridgham J.T.* Relationship between steroidogenic acute regulatory protein expression and progesterone production in hen granulosa cells during follicle development // *Biology of Reproduction*. 2002. V. 67. № 4. P. 1313. <https://doi.org/10.1095/biolreprod67.4.1313>
 29. *Woods D., Johnson A.L.* Regulation of follicle-stimulating hormone-receptor messenger RNA in hen granulosa cells relative to follicle selection // *Biology of Reproduction*. 2005. V. 72. P. 643–650. <https://doi.org/10.1095/biolreprod.104.033902>
 30. *Rui C., Jichang L., Yurou Z. et al.* Association analysis between reproduction genes INHA, PGR, RARG with lamb and other traits of Liaoning cashmere goats // *Animal Biotechnology*. 2022. V. 27. P. 1–12. <https://doi.org/10.1080/10495398.2022.2077212>
 31. *Peiró R., Herrler A., Santacreu M.A. et al.* Expression of progesterone receptor related to the polymorphism in the PGR gene in the rabbit reproductive tract // *J. of Animal Sci.* 2013. V. 88. № 2. P. 421–427. <https://doi.org/10.2527/jas.2009-1955>
 32. *Ghali R.M., Al-Mutawa M.A., Ebrahim B.H. et al.* Progesterone receptor (PGR) gene variants associated with breast cancer and associated features: a case-control study // *Pathol. Oncol. Res.* 2020. V. 26. P. 141–147. <https://doi.org/10.1007/s12253-017-0379-z>
 33. *Yu Y., Pang Y., Zhao H. et al.* Association of a missense mutation in the luteinizing hormone/choriogonadotropin receptor gene (LHCGR) with superovulation traits in Chinese Holstein heifers // *J. Anim. Sci. Biotechnol.* 2012. V. 3. № 1. P. 35. <https://doi.org/10.1186/2049-1891-3-35>
 34. *Yang W.C., Tang K.Q., Li S.J. et al.* Polymorphisms of the bovine luteinizing hormone/choriogonadotropin receptor (LHCGR) gene and its association with superovulation traits // *Mol. Biol. Rep.* 2012. V. 39. № 3. P. 2481–2487. <https://doi.org/10.1007/s11033-011-0999-4>
 35. *Ge S.F., Romanov M.N., Sharp P.J. et al.* Mapping of the luteinizing hormone/choriogonadotropin receptor gene (LHCGR) to chicken chromosome 3 // *Animal Genetics*. 2008. V. 32. № 1. P. 50.

Genes of Hormone Receptors Affecting Egg Productivity and Reproductive Qualities of Chickens

**E. I. Kulikov^{a,*}, L. G. Korshunova^a, R. V. Karapetyan^a, A. S. Komarchev^a, A. K. Kravchenko^a,
D. M. Dmitrenko^a, V. A. Popov^a, V. N. Martynova^a, L. I. Malakheeva^a, and D. N. Efimov^a**

*^aFederal Scientific Center "All-Russian Research and Technological Poultry Institute" of Russian Academy of Sciences,
Moscow oblast, Sergiev Posad, 141311 Russia*

**e-mail: kulikovegor33@yandex.ru*

The egg production of chickens is a complex polygenic type of inheritance and is controlled by many genes. The laying performance of chickens is the result of a complex process regulated by the hypothalamic-pituitary-gonadal system. This review summarizes the information on the effect of polymorphisms of follicle stimulating hormone (FSHR), luteinizing hormone (LHCGR), progesterone (PR) and prolactin (PRLR) receptor genes on the egg production and reproductive performance of chickens. The data presented show that the polymorphisms of these genes are promising for use in breeding programs to improve egg production and reproductive performance of chickens.

Keywords: gene, hormone, receptor, polymorphism, chickens, egg production.

**ОБЗОРНЫЕ
И ТЕОРЕТИЧЕСКИЕ СТАТЬИ**

УДК 575.162+575.176+577.24

ИЗМЕНЕНИЕ ПРОДОЛЖИТЕЛЬНОСТИ ЖИЗНИ КАК ИНТЕГРАЛЬНЫЙ ОТВЕТ НА ИММУННЫЙ СТАТУС ОРГАНИЗМА И АКТИВНОСТЬ МОБИЛЬНЫХ ЭЛЕМЕНТОВ

© 2023 г. М. В. Тростников^{1, 2, *}, Д. Р. Малышев¹, Е. Г. Пасюкова¹

¹Национальный исследовательский центр “Курчатовский институт”, Москва, 123182 Россия

²Сколковский институт науки и технологий, Московская область, Сколково, 121205 Россия

*e-mail: mikhail.trostnikov@gmail.com

Поступила в редакцию 12.05.2023 г.

После доработки 29.05.2023 г.

Принята к публикации 01.06.2023 г.

Одной из ключевых задач при изучении молекулярно-генетических основ многих патологий является поиск триггеров, влияние на которые могло бы положительно сказаться на частоте возникновения зависимых от возраста заболеваний и в целом темпах старения. Возможной причиной зависимой от возраста деградации функций организма, индуцирующей старение, является иммуносенесценция. Известно, что наблюдаемое с возрастом повышение активности мобильных элементов может не только влиять на уровень стабильности генома, но и играть ключевую роль в формировании иммунного ответа. В то же время давно доказана ключевая роль нервной системы в контроле продолжительности жизни, а недавно показано, что компоненты аппарата, регулирующего активность мобильных элементов, функционируют в нервной системе, и их работа влияет на развитие нейродегенеративных заболеваний. В мини-обзоре представлены факты, указывающие на комплексную регуляцию старения со стороны нервной и иммунной систем при участии систем контроля активности мобильных элементов и предложена гипотетическая схема их совместного влияния на продолжительность жизни.

Ключевые слова: продолжительность жизни, иммунная система, мобильные элементы, нервная система.

DOI: 10.31857/S0016675823110140, **EDN:** NDJCID

Считается, что биологической основой старения — увеличения вероятности смерти с возрастом [1, 2] — является постепенная деградация и потеря функций отдельных органов и организма в целом, приводящая к смерти [3]. Выяснение фундаментальных молекулярно-генетических и эволюционных механизмов старения является одной из наиболее актуальных задач биологии и медицины. Современные данные свидетельствуют о том, что такой сложный универсальный признак как старение, не может находиться под контролем единственного регулятора и, напротив, формируется в результате сложных взаимодействий различных высоко консервативных молекулярных регуляторных каскадов и генных сетей, играющих решающую роль в поддержании гомеостаза. Среди причин, вносящих наибольший вклад в процесс старения, можно выделить несколько наиболее важных: нарушение обмена основных метаболитов и энергетического гомеостаза, нарушение иммунитета, нарушение структуры основных макромолекул и ослабление систем их восстановления [4]. К наиболее значимым молеку-

лярно-генетическим регуляторам старения относят каскады IGF-1, TOR, JNK, системы иммунного ответа, детоксикации, репарации, аутофагии и другие [4–8]. Несмотря на большое количество данных, свидетельствующих об активной роли каждого из этих регуляторных путей в поддержании гомеостаза на уровне организма, остается открытым вопрос об оптимальных уровнях их активности с точки зрения замедления старения и продления жизни и факторах, оказывающих наибольшее влияние на их дисбаланс.

Однако в настоящее время наиболее интересными представляются вопросы о том, как взаимодействуют между собой различные системы регуляции старения, какова иерархия этих взаимодействий, как осуществляется обратная связь между разнонаправленными молекулярно-генетическими сигналами, существуют ли пути регуляции, которые играют наиболее значимую, возможно лимитирующую роль в определении скорости зависимой от возраста деградации организма. Принципиальную важность имеет также вопрос о том, какие типы клеток и тканей, а также возрастные перио-

ды в наибольшей степени влияют на длительность жизни, темпы старения и здоровое долголетие организма.

ИММУННЫЙ СТАТУС ОРГАНИЗМА ОПРЕДЕЛЯЕТ СКОРОСТЬ СТАРЕНИЯ И РАЗВИТИЯ ЗАВИСИМЫХ ОТ ВОЗРАСТА ПАТОЛОГИЙ

В качестве одного из глобальных свойств организма, связывающих многие регуляторные сети, в современных работах часто рассматривают иммунный статус. Интересно, что такие метаболические пути, как IGF-1 и TOR, являются консервативными плейотропными регуляторами не только старения и продолжительности жизни, но и иммунитета [9]. Значимую роль генов/белков иммунного ответа в контроле продолжительности жизни доказывают эксперименты, демонстрирующие их достоверную ассоциацию с долгожительством при направленном отборе на увеличение продолжительности жизни [10].

Среди белков иммунного ответа, связанных с контролем продолжительности жизни, наиболее хорошо исследованы **Toll-рецепторы**, играющие важную роль в обеспечении антигрибковой, антимикробной и противовирусной защиты [11–14]. Семейство Toll-рецепторов было впервые открыто в работах с плодовой мушкой, а впоследствии эти рецепторы были найдены и у млекопитающих [15, 16]. Считается, что активность Toll-рецепторов может быть связана со старением, причем эта связь может быть обусловлена как напрямую, с влиянием их функционального состояния на успешное противостояние патогенам, так и косвенно, с их взаимодействием с сигнальными путями, ассоциированными со старением, такими как IGF-1 [17].

Как и многие глобальные регуляторы гомеостаза, иммунная система в целом и отдельные ее компоненты, например, высоко консервативные Toll-рецепторы, по-разному работают на разных стадиях развития и в разном возрасте [18, 19]. Кроме того, в некоторых случаях гены/белки эволюционно достаточно давно сформировавшегося канонического пути врожденного иммунного ответа могли приобрести дополнительные функции: так, у плодовой мушки Toll участвует в контроле эмбрионального развития [20]. В целом изучение функций генов/белков иммунной системы, в частности Toll-сигналинга, в эмбриогенезе имеет большое значение, поскольку эти функции могут определить не только особенности иммунитета, но также специфику работы органов и тканей, формирующихся в процессе развития во взрослом организме. Так, в ряде работ было продемонстрировано, что у плодовой мушки гемоциты (фагоциты, иммунные клетки беспозвоночных) переходят во взрослую особь непосредственно из

эмбриона и личинки [21, 22]. Эти результаты свидетельствуют о том, что свойства иммунной системы взрослого организма, а следовательно и темпы старения, могут быть, по крайней мере, частично заданы особенностями развития. Понимание генетических и молекулярных механизмов, связывающих становление иммунитета на ранних стадиях развития с темпами старения и частотой возникновения возрастных патологий, может играть важную роль в разработке геропротекторных технологий.

Иммунный статус организма существенно меняется с возрастом, и в старости нарушение иммунитета ассоциировано с повышенной уязвимостью к патогенам и повышенным риском развития онкологических заболеваний [23, 24]. Меры по усилению иммунного ответа в старости представляются в связи с этим очевидно полезными и ведущими к увеличению продолжительности жизни. Однако известно, что, например, гиперактивация продукции антимикробных пептидов приводит к сокращению продолжительности жизни плодовой мушки [25]. Показано, что возрастная гиперактивация и дисбаланс иммунного ответа, по-видимому, приводят к ослаблению иммунитета и, в том числе, к стимулированию нейродегенерации [26, 27]. Таким образом, усиление иммунной функции может иметь вредные последствия для организма. В этом случае подавление иммунной функции, нарушенной при старении, может оказаться полезным для терапии возрастных воспалительных процессов и, следовательно, для увеличения продолжительности жизни [28, 29].

ВОЗРАСТНЫЕ ОСОБЕННОСТИ ИММУННОГО ОТВЕТА ЗАВИСЯТ ОТ АКТИВНОСТИ МОБИЛЬНЫХ ЭЛЕМЕНТОВ

Функция и эволюция систем врожденного и адаптивного иммунного ответа тесно связаны с контролем активности мобильных элементов [30–32]. Зависят ли возрастные нарушения регуляции иммунного ответа, в частности упомянутая выше негативная гиперактивация, от сопровождающих старение изменений контроля подвижности мобильных элементов? Хорошо известно, что перемещения мобильных элементов могут активироваться в старости, что приводит к повреждению ДНК, накоплению мутаций и общей геномной нестабильности, которая провоцирует деградацию клеточных функций с возрастом [33–35]. Более того, повышение активности мобильных элементов может не только влиять на уровень стабильности генома, но и играть ключевую роль в формировании иммунного ответа [30, 36–38]. Учитывая сходство между отдельными мобильными элементами и некоторыми вирусами, это можно объяснить тем, что усиление активности ряда мобильных

элементов, приводящее к усиленному образованию вирусоподобных частиц, по-видимому может стимулировать иммунную систему и тем самым либо улучшать ответ организма на вирусные атаки, либо вызывать негативную гиперактивацию. Действительно, недавние публикации указывают на достоверную связь активации мобильных элементов, усиления вирусной нагрузки и активации Toll-пути: так, у человека активность эндогенного ретровируса K активирует Toll-сигналинг [39]. Кроме того, активность мобильных элементов влияет на репродуктивные функции [33–35], тесно связанные с продолжительностью жизни. В целом, несмотря на все выявленные факты, роль мобильных элементов в контроле продолжительности жизни и механизмы их влияния на темпы старения остаются предметом дискуссий [3, 40, 41].

Известно, что активность мобильных элементов, из которых на 40% состоит геном человека и на 16% – плодовой мушки, в норме репрессируется системой коротких РНК, кодируемых специфическим кластером в геноме [42]. Ключевую роль в контроле образования таких коротких РНК и подавления активности мобильных элементов играет эволюционно консервативные белки семейства *Piwi*. С возрастом система подавления работает с меньшей эффективностью, что и приводит к активации мобильных элементов и росту уровня нестабильности генома [33, 43, 44]. В то же время, показано, что нокаут *Piwi* в герминальных тканях плодовой мушки приводит к активации сборки вирусоподобных частиц, также влияя на Toll [39, 45, 46]. В контексте поиска причин старения обращает на себя внимание тот факт, что зависимый от возраста рост активности LINE1-элементов провоцирует воспаление, которое может быть подавлено веществами, воздействующими на обратную транскрипцию [47].

Как было отмечено выше, появляются данные, свидетельствующие о том, что процессы, происходящие в период раннего развития иммунной системы, могут влиять на состояние иммунитета взрослого организма. В связи с этим особый интерес приобретает тот факт, что роль мобильных элементов в контроле иммунитета, по всей видимости, также может формироваться еще в ходе развития. Известно, что активность семейства мобильных элементов *Gypsy* в процессе метаморфоза плодовой мушки может детерминировать противовирусную активность взрослой особи [48].

Еще один интересный аспект зависимости иммунного ответа от транспозиционной активности мобильных элементов связан с участием нервной системы в этом взаимодействии. Нервная система, отвечая на изменение внешних и внутренних факторов, играет связующую и контролируемую роль в отношении других тканей и органов. Молекулярные аспекты взаимодействия нервной и

иммунной систем важны для понимания закономерностей поддержания гомеостаза организма в целом [49, 50]. Известно, что сила иммунного ответа и баланс между его полезными и вредными последствиями в значительной мере регулируются нервной системой посредством специфических нейротрансмиттеров и нейропептидов, причем роль отдельных типов нейронов как в регуляции иммунной системы, так и в чувствительности к активации иммунного ответа, по-видимому не одинакова [51–54].

Активность белков, являющихся компонентами клеточной системы подавления мобильных элементов, чаще ассоциируют с герминальными тканями. Однако относительно недавние данные продемонстрировали присутствие полипептидов, соответствующих белкам *Piwi*-пути, в гранулах глиальных клеток мозга плодовой мушки [55]. В связи с этим интересно отметить, что спектр нарушений нервной системы, характерный для пожилого возраста, в ряде случаев аналогичен изменениям, которые сопутствуют дерегуляции белков *Piwi*-пути. Так, при анализе пациентов с болезнью Альцгеймера было выявлено 103 коротких РНК (синтез которых контролирует *Piwi*), являющихся потенциальными факторами риска возникновения данной патологии [56, 57]. У пациентов с болезнью Паркинсона также выявлена значительная дерегуляция образования коротких РНК [56, 57]. Стоит отметить, что у плодовой мушки эффект повышения экспрессии *Aub* – одного из белков *Piwi*-пути, усиливает эффекты нокаута гена *caz*, нарушение в работе которого у человека ассоциировано с боковым амиотрофическим склерозом [58]. Факты, свидетельствующие о связи между активностью мобильных элементов в нервной системе и развитием нейродегенеративных, в том числе старческих болезней, пока достаточно разрознены, однако они все больше указывают на комплексную регуляцию старения со стороны нервной и иммунной систем при участии систем контроля активности мобильных элементов.

ЗАКЛЮЧЕНИЕ

Рассмотренные в настоящем обзоре факты дают возможность судить о том, насколько сложным и неочевидным образом могут быть переплетены молекулярные механизмы контроля продолжительности жизни, темпов старения и возникновения возрастных патологий. Пока остается неясным вопрос о том, существует ли событие или фактор, который можно рассматривать как некий триггер, запускающий комплекс зависимых от возраста изменений в различных молекулярно-генетических путях, регулирующих гомеостаз организма. Возможной причиной зависимой от возраста деградации функций организма, индуцирующей старение,

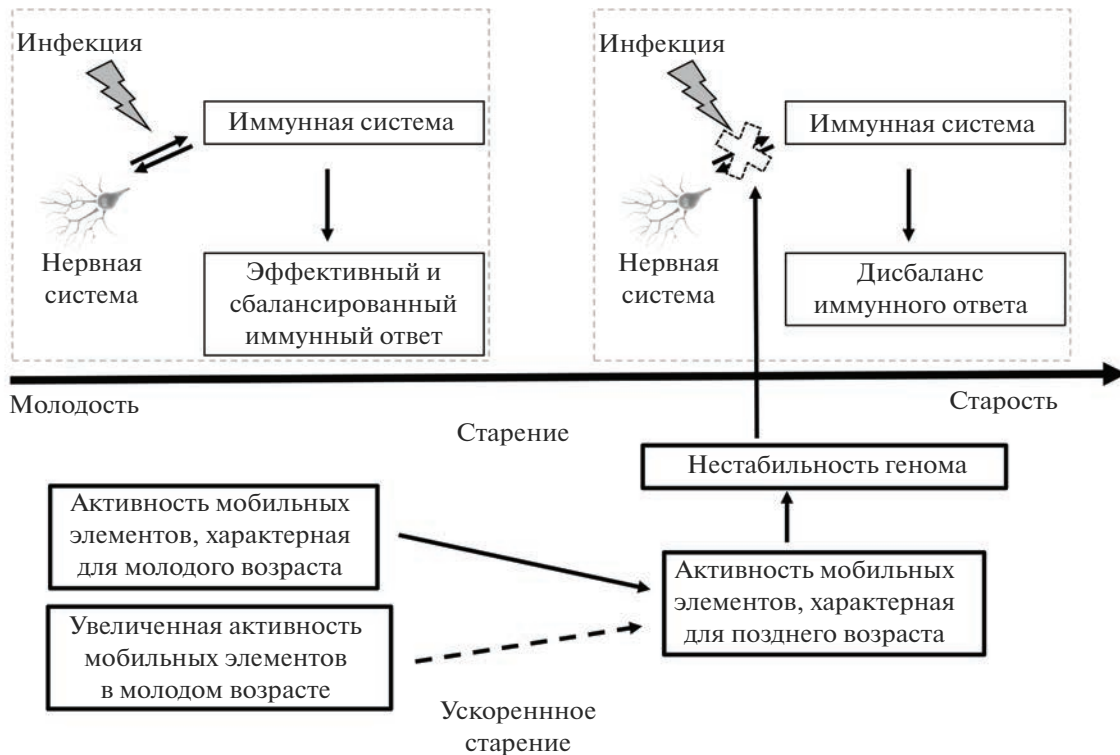


Рис. 1. Гипотетический механизм влияния повышенной активности мобильных элементов при старении и в раннем возрасте на продолжительность жизни.

является иммуносенесцентность (зависимое от возраста снижение иммунного ответа), вызванная действием окружающей среды и образа жизни (питания, физических упражнений и лекарств, принимаемые в течение жизни) [59]. В то же время давно доказана ключевая роль нервной системы в контроле продолжительности жизни [60]. Тем более интересными выглядят факты, свидетельствующие о взаимосвязи иммунной и нервной систем в контроле старения [38, 39].

Каковы молекулярные механизмы такой взаимосвязи? Полученные в последнее время данные позволяют выдвинуть новые и несколько неожиданные предположения на этот счет. Оказалось, что активность мобильных элементов генома в существенной степени определяет свойства как иммунной, так и нервной систем [32, 38]. Более того, отдельные представленные выше факты позволяют предположить, что на продолжительность жизни и темпы старения взрослых особей могут влиять особенности взаимодействия всех трех компонентов в период развития организма. Гипотетический механизм совместного влияния иммунной и нервной систем, при участии систем контроля активности мобильных элементов, на продолжительность жизни и старение представлен на рис. 1.

В норме у молодых особей взаимодействие между нервной и иммунной системами под действием инфекционных агентов стимулирует возникновение иммунного ответа, например очага воспаления, к которому затем привлекаются иммунные клетки, способствующие заживлению поврежденной ткани. При этом активность мобильных элементов находится на низком уровне. Нормальное старение сопровождается активацией мобильных элементов, что ведет к геномной нестабильности, нарушающей нормальную связь между нервной системой, инфекционным стимулом и иммунным ответом. В результате может возникнуть дисбаланс иммунного ответа, приводящий к ослаблению иммунитета, стимулированию нейродегенерации и ускорению старения. В определенных условиях активация мобильных элементов может произойти у молодых особей, в том числе в процессе развития, и далее закрепиться благодаря эпигенетическому наследованию дестабилизированного статуса. Нормальная связь между нервной системой, инфекционным стимулом и иммунным ответом нарушится и приведет к ускоренному старению.

Таким образом, взаимодействие нервной и иммунной систем в контроле продолжительности жизни может быть опосредовано одним из ста-

рейших механизмов, определяющих эволюцию генома живых организмов. Этот механизм, являясь “эгоистичным”, направленным на распространение повторяющихся последовательностей в геноме и приводящим к его нестабильности, может тем не менее участвовать в контроле важнейших признаков — гомеостаза, старения, продолжительности жизни. Изучение тонких молекулярных механизмов, обеспечивающих баланс между функциональным статусом нервной и иммунной систем и системой контроля мобильных элементов генома может стать основой для понимания эволюционных основ контроля продолжительности жизни и старения, а также для выработки эффективных подходов к терапии зависимых от возраста патологий.

Мы выражаем глубокую признательность центру культур *Drosophila* (Блумингтон, США, <https://bdsc.indiana.edu/index.html>) за многолетнее содействие нашим исследованиям.

Работа выполнена с использованием оборудования Центра коллективного пользования научным оборудованием НИЦ “Курчатовский институт”.

Работа была поддержана грантом РНФ № 22-74-00065.

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. Kirkwood T.B.L. Deciphering death: a commentary on Gompertz (1825) ‘On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies’ // *Philos. Trans. Royal Soc. B Biol. Sci.* 2015. V. 370. № 1666. P. 20140379. <https://doi.org/10.1098/rstb.2014.0379>
2. Liochev S.I. Which is the most significant cause of aging? // *Antioxidants (Basel)*. 2015. V. 4. № 4. P. 793–810. <https://doi.org/10.3390/antiox4040793>
3. López-Otín C., Blasco M.A., Partridge L. et al. Hallmarks of aging: an expanding universe // *Cell*. 2023. V. 186. № 2. P. 243–278. <https://doi.org/10.1016/j.cell.2022.11.001>
4. Guo J., Huang X., Dou L. et al. Aging and aging-related diseases: from molecular mechanisms to interventions and treatments: 1 // *Sig. Transduct. Target Ther.* 2022. V. 7. № 1. P. 1–40. <https://doi.org/10.1038/s41392-022-01251-0>
5. Yu M., Zhang H., Wang B. et al. Key signaling pathways in aging and potential interventions for healthy aging: 3 // *Cells Multidisciplinary Digital Publ. Inst.* 2021. V. 10. № 3. P. 660. <https://doi.org/10.3390/cells10030660>
6. Mogilenko D.A., Shchukina I., Artyomov M.N. Immune ageing at single-cell resolution: 8 // *Nat. Rev. Immunol.* 2022. V. 22. № 8. P. 484–498. <https://doi.org/10.1038/s41577-021-00646-4>
7. Gan T., Fan L., Zhao L. et al. JNK signaling in *Drosophila* aging and longevity: 17 // *Int. J. Mol. Sci.* 2021. V. 22. № 17. P. 9649. <https://doi.org/10.3390/ijms22179649>
8. Hayat R., Manzoor M., Hussain A. Wnt signaling pathway: A comprehensive review // *Cell Biol. International*. 2022. V. 46. № 6. P. 863–877. <https://doi.org/10.1002/cbin.11797>
9. Fabian D.K., Fuentealba M., Dönertaş H.M. et al. Functional conservation in genes and pathways linking ageing and immunity // *Immunity & Ageing*. 2021. V. 18. № 1. P. 23. <https://doi.org/10.1186/s12979-021-00232-1>
10. Fabian D.K., Garschall K., Klepsatel P. et al. Evolution of longevity improves immunity in *Drosophila* // *Evol. Letters*. 2018. V. 2. № 6. P. 567–579. <https://doi.org/10.1002/evl3.89>
11. Mafi S., Mansoori B., Taeb S. et al. mTOR-mediated regulation of immune responses in cancer and tumor microenvironment // *Front. in Immunology*. 2022. V. 12. <https://doi.org/10.3389/fimmu.2021.774103>
12. Kirchheis R., Planz O. The role of Toll-like receptors (TLRs) and their related signaling pathways in viral infection and inflammation // *Int. J. Mol. Sci.* 2023. V. 24. № 7. <https://doi.org/10.3390/ijms24076701>
13. Haseeb M., Pirzada R.H., Ain Q.U. et al. Wnt signaling in the regulation of immune cell and cancer therapeutics // *Cells*. 2019. V. 8. № 11. <https://doi.org/10.3390/cells8111380>
14. Duan T., Du Y., Xing C. et al. Toll-like receptor signaling and its role in cell-mediated immunity // *Front. Immunology*. 2022. V. 13. <https://doi.org/10.3389/fimmu.2022.812774>
15. Lemaitre B., Nicolas E., Michaut L. et al. The dorsoventral regulatory gene cassette *spätzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults // *Cell*. 1996. V. 86. № 6. P. 973–983. [https://doi.org/10.1016/s0092-8674\(00\)80172-5](https://doi.org/10.1016/s0092-8674(00)80172-5)
16. Medzhitov R., Preston-Hurlburt P., Janeway C.A. A human homologue of the *Drosophila* Toll protein signals activation of adaptive immunity // *Nature*. 1997. V. 388. № 6640. P. 394–397. <https://doi.org/10.1038/41131>
17. Salminen A., Kaarniranta K., Kauppinen A. Insulin/IGF-1 signaling promotes immunosuppression via the STAT3 pathway: Impact on the aging process and age-related diseases // *Inflamm. Res.* 2021. V. 70. № 10–12. P. 1043–1061. <https://doi.org/10.1007/s00011-021-01498-3>
18. Kubiak M., Tinsley M.C. Sex-specific routes to immune senescence in *Drosophila melanogaster*: 1 // *Sci. Rep.* 2017. V. 7. № 1. P. 10417. <https://doi.org/10.1038/s41598-017-11021-6>

19. Yu S., Luo F., Xu Y. et al. *Drosophila* innate immunity involves multiple signaling pathways and coordinated communication between different tissues // Front. Immunology. 2022. V. 13
20. Nüsslein-Volhard C. The Toll gene in *Drosophila* pattern formation // Trends in Genet. Elsevier. 2022. V. 38. № 3. P. 231–245.
<https://doi.org/10.1016/j.tig.2021.09.006>
21. Boulet M., Renaud Y., Lapraz F. et al. Characterization of the *Drosophila* adult hematopoietic system reveals a rare cell population with differentiation and proliferation potential // Frontiers in Cell and Developmental Biology. 2021. V. 9.
22. Sanchez Bosch P., Makhijani K., Herbozo L. et al. Adult *Drosophila* lack hematopoiesis but rely on a blood cell reservoir at the respiratory epithelia to relay infection signals to surrounding tissues // Dev. Cell. 2019. V. 51. № 6. P. 787–803.e5.
<https://doi.org/10.1016/j.devcel.2019.10.017>
23. Ventura M.T., Casciaro M., Gangemi S. et al. Immunosenescence in aging: between immune cells depletion and cytokines up-regulation // Clin. Mol. Allergy. 2017. V. 15.
<https://doi.org/10.1186/s12948-017-0077-0>
24. Sadighi Akha A.A. Aging and the immune system: An overview // J. Immunol. Meth. 2018. V. 463. P. 21–26.
<https://doi.org/10.1016/j.jim.2018.08.00>
25. Badinloo M., Nguyen E., Suh W. et al. Over-expression of antimicrobial peptides contributes to aging through cytotoxic effects in *Drosophila* tissues // Arch. Insect Biochem. Physiol. 2018. V. 98. № 4.
<https://doi.org/10.1002/arch.21464>
26. Lucin K.M., Wyss-Coray T. Immune activation in brain aging and neurodegeneration: too much or too little? // Neuron. 2009. V. 64. № 1. P. 110–122.
<https://doi.org/10.1016/j.neuron.2009.08.039>
27. Sanuki R., Tanaka T., Suzuki F. et al. Normal aging hyperactivates innate immunity and reduces the medical efficacy of minocycline in brain injury // Brain, Behavior, and Immunity. 2019. V. 80. P. 427–438.
<https://doi.org/10.1016/j.bbi.2019.04.023>
28. Weyand C.M., Goronzy J.J. Aging of the immune system. Mechanisms and therapeutic targets // Ann. Am. Thorac. Soc. 2016. V. 13. Suppl. 5. P. S422–S428.
<https://doi.org/10.1513/AnnalsATS.201602-095AW>
29. Lee K.-A., Flores R.R., Jang I.H. et al. Immune senescence, immunosenescence and aging // Frontiers in Aging. 2022. V. 3.
30. Koonin E.V., Krupovic M. Evolution of adaptive immunity from transposable elements combined with innate immune systems: 3 // Nat. Rev. Genet. 2015. V. 16. № 3. P. 184–192.
<https://doi.org/10.1038/nrg3859>
31. Broecker F., Moelling K. Evolution of immune systems from viruses and transposable elements // Frontiers in Microbiology. 2019. V. 10.
32. Gázquez-Gutiérrez A., Witteveldt J., Heras S.R. et al. Sensing of transposable elements by the antiviral innate immune system // RNA. 2021. V. 27. № 7. P. 735–752.
<https://doi.org/10.1261/rna.078721.121>
33. Li W., Prazak L., Chatterjee N. et al. Activation of transposable elements during aging and neuronal decline in *Drosophila* // Nat. Neurosci. 2013. V. 16. № 5. P. 529–531.
<https://doi.org/10.1038/nn.3368>
34. Van Meter M., Kashyap M., Rezaeadeh S. et al. SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age // Nat. Commun. 2014. V. 5.
<https://doi.org/10.1038/ncomms6011>
35. Brunet T.D.P., Doolittle W.F. Multilevel selection theory and the evolutionary functions of transposable elements // Genome Biol. Evol. 2015. V. 7. № 8. P. 2445–2457.
<https://doi.org/10.1093/gbe/evv152>
36. Santos D., Verdonckt T.-W., Mingels L. et al. PIWI proteins play an antiviral role in lepidopteran cell lines // Viruses. 2022. V. 14. № 7.
<https://doi.org/10.3390/v14071442>
37. Kolliopoulou A., Santos D., Taning C.N.T. et al. PIWI pathway against viruses in insects // WIREs RNA. 2019. V. 10. № 6.
<https://doi.org/10.1002/wrna.1555>
38. Takahashi T., Heaton S.M., Parrish N.F. Mammalian antiviral systems directed by small RNA // PLoS Pathogens Publ. Library of Sci. 2021. V. 17. № 12.
<https://doi.org/10.1371/journal.ppat.1010091>
39. Rolland A., Jouvin-Marche E., Viret C. et al. The envelope protein of a human endogenous retrovirus-W family activates innate immunity through CD14/TLR4 and promotes Th1-like responses // J. Immunol. 2006. V. 176. № 12. P. 7636–7644.
<https://doi.org/10.4049/jimmunol.176.12.7636>
40. López-Otín C., Blasco M.A., Partridge L. et al. The hallmarks of aging // Cell. 2013. V. 153. № 6. P. 1194–1217.
<https://doi.org/10.1016/j.cell.2013.05.039>
41. Moskalev A.A., Shaposhnikov M.V., Plyusnina E.N. et al. The role of DNA damage and repair in aging through the prism of Koch-like criteria // Ageing Res. Rev. 2013. V. 12. № 2. P. 661–684.
<https://doi.org/10.1016/j.arr.2012.02.001>
42. Iwasaki Y.W., Siomi M.C., Siomi H. PIWI-Interacting RNA: Its biogenesis and functions // Annu. Rev. Biochem. 2015. V. 84. P. 405–433.
<https://doi.org/10.1146/annurev-biochem-060614-034258>
43. Chen H., Zheng X., Xiao D. et al. Age-associated de-repression of retrotransposons in the *Drosophila* fat body, its potential cause and consequence // Aging Cell. 2016. V. 15. № 3. P. 542–552.
<https://doi.org/10.1111/accel.12465>
44. Cecco M.D., Criscione S.W., Peterson A.L. et al. Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues // Aging. 2013. V. 5. № 12. P. 867–883.
<https://doi.org/10.18632/aging.100621>
45. Lin K.-Y., Wang W.-D., Lin C.-H. et al. Piwi reduction in the aged niche eliminates germline stem cells via Toll-GSK3 signaling: 1 // Nat. Commun. 2020. V. 11. № 1. P. 3147.
<https://doi.org/10.1038/s41467-020-16858-6>
46. Rolland A., Jouvin-Marche E., Saresella M. et al. Correlation between disease severity and in vitro cytokine production mediated by MSRV (multiple sclerosis associated retroviral element) envelope protein in patients

- with multiple sclerosis // *J. Neuroimmunol.* 2005. V. 160. № 1–2. P. 195–203.
<https://doi.org/10.1016/j.jneuroim.2004.10.019>
47. *De Cecco M., Ito T., Petrashen A.P. et al.* L1 drives IFN in senescent cells and promotes age-associated inflammation // *Nature.* 2019. V. 566. № 7742. P. 73–78.
<https://doi.org/10.1038/s41586-018-0784-9>
 48. *Wang L., Tracy L., Su W. et al.* Retrotransposon activation during *Drosophila* metamorphosis conditions adult antiviral responses: 12 // *Nat. Genet.* 2022. V. 54. № 12. P. 1933–1945.
<https://doi.org/10.1038/s41588-022-01214-9>
 49. *Dantzer R.* Neuroimmune interactions: from the brain to the immune system and vice versa // *Physiol. Rev.* 2018. V. 98. № 1. P. 477–504.
<https://doi.org/10.1152/physrev.00039.2016>
 50. *Daëron M.* The immune system as a system of relations // *Front. Immunol.* 2022. V. 13.
<https://doi.org/10.3389/fimmu.2022.984678>
 51. *Kawli T., He F., Tan M.-W.* It takes nerves to fight infections: insights on neuro-immune interactions from *C. elegans* // *Disease Models & Mechanisms.* 2010. V. 3. № 11–12. P. 721–731.
<https://doi.org/10.1242/dmm.003871>
 52. *Shukla A.K., Spurrier J., Kuzina I. et al.* Hyperactive innate immunity causes degeneration of dopamine neurons upon altering activity of Cdk5 // *Cell Rep.* 2019. V. 26. № 1. P. 131–144.e4.
<https://doi.org/10.1016/j.celrep.2018.12.025>
 53. *Kounatidis I., Chtarbanova S.* Role of glial immunity in lifespan determination: A *Drosophila* perspective // *Front. Immunol.* 2018. V. 9.
<https://doi.org/10.3389/fimmu.2018.01362>
 54. *Otarigho B., Aballay A.* Immunity-longevity tradeoff neurally controlled by GABAergic transcription factor PITX1/UNC-30 // *bioRxiv.* Cold Spring Harbor Lab. 2021.
<https://doi.org/10.1101/2021.02.25.432801>
 55. *Tindell S.J., Rouchka E.C., Arkov A.L.* Glial granules contain germline proteins in the *Drosophila* brain, which regulate brain transcriptome: 1 // *Commun. Biol.* 2020. V. 3. № 1. P. 1–12.
<https://doi.org/10.1038/s42003-020-01432-z>
 56. *Qiu W., Guo X., Lin X. et al.* Transcriptome-wide piRNA profiling in human brains of Alzheimer's disease // *Neurobiol. Aging.* 2017. V. 57. P. 170–177.
<https://doi.org/10.1016/j.neurobiolaging.2017.05.020>
 57. *Schulze M., Sommer A., Plötz S. et al.* Sporadic Parkinson's disease derived neuronal cells show disease-specific mRNA and small RNA signatures with abundant deregulation of piRNAs // *Acta Neuropathol. Commun.* 2018. V. 6. № 1. P. 58.
<https://doi.org/10.1186/s40478-018-0561-x>
 58. *Wakisaka K.T., Tanaka R., Hirashima T. et al.* Novel roles of *Drosophila* FUS and Aub responsible for piRNA biogenesis in neuronal disorders // *Brain Res.* 2019. V. 1708. P. 207–219.
<https://doi.org/10.1016/j.brainres.2018.12.028>
 59. *Lathe R., St Clair D.* Programmed ageing: Decline of stem cell renewal, immunosenescence, and Alzheimer's disease // *Biol. Rev. Camb. Philos. Soc.* 2023.
<https://doi.org/10.1111/brv.12959>
 60. *Alcedo J., Flatt T., Pasyukova E.G.* The role of the nervous system in aging and longevity // *Front. Genet.* 2013. V. 4.
<https://doi.org/10.3389/fgene.2013.00124>

Changes in Lifespan as an Integral Response to the Organism's Immune Status and Mobile Elements Activity

M. V. Trostnikov^{a, b, *}, D. R. Malyshev^a, and E. G. Pasyukova^a

^aNational Research Center “Kurchatov Institute”, Moscow, 123182 Russia

^bSkolkovo Institute of Science and Technology, Moscow oblast, Skolkovo, 121205 Russia

*e-mail: mikhail.trostnikov@gmail.com

One of the key goals in studying the molecular and genetic basis of many pathologies is to find triggers, which could positively influence the incidence of age-dependent diseases and the rate of aging in general. A possible cause of age-dependent degradation of the organism functions inducing aging is immunosenescence. It is known that the increased activity of mobile elements observed with age may not only affect the level of genome stability, but also play a crucial role in the development of immune response. At the same time, the pivotal role of the nervous system in controlling lifespan has long been proven, and recently it has been shown that components of the machinery that regulate mobile element activity do function in the nervous system, and their functioning affects the development of neurodegenerative diseases. This mini-review presents evidence for complex aging regulation by the nervous and immune systems, involving systems that control mobile element activity and proposes a hypothetical scheme for their joint impact on lifespan.

Keywords: lifespan, immune system, mobile elements, nervous system.

ЛОКАЛИЗАЦИЯ ХРОМОСОМНЫХ РЕГИОНОВ, ОПРЕДЕЛЯЮЩИХ
СОДЕРЖАНИЕ МАГНИЯ И КАЛЬЦИЯ У ОБРАЗЦОВ РИСА© 2023 г. Ю. К. Гончарова^{1, 3, *}, В. В. Симонова¹, С. В. Гончаров², Н. А. Очкас³¹Федеральный научный центр риса, Краснодарский край, пос. Белозерный, 350921 Россия²Кубанский государственный аграрный университет им. И.Т. Трубилина, Краснодар, 350044 Россия³ООО “Аратай”, участник инновационного центра Сколково, Москва, 143026 Россия

*e-mail: yuliya_goncharova_20@mail.ru

Поступила в редакцию 13.05.2023 г.

После доработки 14.06.2023 г.

Принята к публикации 15.06.2023 г.

Отмечено широкое варьирование по признаку “содержание микро- и макроэлементов” как между подвидами, так и внутри видов и сортов риса. Магний как кофактор участвует более чем в 300 ферментативных реакциях и нужен для роста и развития как растений, так и человека. Он влияет на обмен углеводов и белковых веществ. От его концентрации зависят производство аденозинтрифосфата (АТФ), синтез нуклеотидов и глюкозы, регуляция окисления липидов. Кальций также необходим для формирования структуры клеточных стенок и их деления. Высокие концентрации кальция изменяют состав и состояние мембран эритроцитов, морфологические характеристики клетки. В то же время варибельность по признакам отечественных сортов риса до сих пор не изучалась. Также не было исследований о локализации хромосомных регионов, ответственных за формирование признаков “содержание магния и кальция” у российских образцов риса. Установлено широкое варьирование содержания кальция (0.07–2.33%) и магния (2–14%) у отечественных образцов риса. Данные по фенотипированию отечественных сортов риса по признакам качества использовали для выделения контрастных по этим признакам групп сортов и источников. Проведен поиск хромосомных регионов, определяющих качество отечественных образцов, с использованием 58 распределенных по геному риса молекулярных маркеров (SSR) как сцепленных с признаком, так и нейтральных. Выделены пять локусов, определяющих содержание магния и кальция у отечественных сортов риса. Локусы, определяющие содержание магния, расположены на хромосомах 5, 6, 7, 8. Связанных с содержанием кальция было выделено четыре локуса: один на второй хромосоме (при уровне значимости 0.05) и по одному на второй, восьмой и пятой (при уровне значимости 0.09). На пятой хромосоме он расположен (RM 13; 28.6 cM) в непосредственной близости к региону расположения маркера RM 405 (28 cM), который связан с содержанием магния.

Ключевые слова: генетика, рис, варибельность, локусы количественных признаков, качество, молекулярное маркирование, кальций, магний.

DOI: 10.31857/S001667582311005X, **EDN:** NIVOKY

Для функционирования как животных, так и растительных организмов нужны минералы. Для человека наиболее важные из них железо, цинк, магний [1, 2]. На содержание минералов в растении влияют как генетические характеристики образца, так и внешние факторы: в первую очередь почвы, на которых оно произрастает [3–5]. Рис может быть источником кальция, магния, фосфора, селена, цинка, марганца, железа, калия. Отмечено широкое варьирование по признаку “содержание микро- и макроэлементов” как между подвидами, так и внутри видов [6].

Магний как кофактор участвует более чем в 300 ферментативных реакциях. От его концентрации зависят как производство аденозинтрифос-

фата (АТФ), так и синтез нуклеотидов и глюкозы, регуляция окисления липидов [7–10]. Катионы Mg^{2+} являются наиболее важными элементами, обеспечивающими сокращение мышц, проведение нервных импульсов и процесс формирования костей у позвоночных. Важна роль магния также в регуляции сердечно-сосудистых процессов, поддержании гомеостаза организма. Также ионы магния Mg^{2+} способны корректировать воспалительные и окислительные процессы, приводящие к сердечно-сосудистым заболеваниям (ССЗ) [11]. Повышенный риск развития сахарного диабета 2-го типа и нарушения сердечного ритма также связаны с гипомagneмией. Катионы двухвалентного магния регулируют эффективность работы

ионных каналов, определяющих поступление натрия, калия и кальция через клеточные мембраны [12, 13]. Наличие ионов магния важно в обмене ионов K^+ , Na^+ . Его недостаток приводит к избыточному выделению кальция, увеличению концентрации ионов K^+ , Na^+ внутри клеток [14–16]. Высокие концентрации кальция изменяют состав и состояние мембран эритроцитов, морфологические характеристики клетки и приводят к разрушению ядра, а также количественно изменяют состав фосфолипидов мембран эритроцитов; препятствие этим процессам могут оказать флавоноиды [17]. Ишемическая болезнь сердца сопровождается перегрузкой клеток кальцием. Магний как антагонист кальция снижает кальциевую нагрузку при ишемии миокарда, чем ограничивает размер инфаркта, спазмы артерий, уменьшает постинфарктные повреждения [18]. Защитные функции магния включают также стимуляцию процессов метаболизма в миокарде, регуляцию сопротивления сосудов и агрегации тромбоцитов [19]. Дефицит магния связан с изменением активности Na - K -АТФазы, что приводит к снижению поглощения калия и его выделению из организма [20]. Коррекция дефицита магния во многих случаях способствует нормализации содержания калия в клетке, устранению риска гиперкалиемии. Содержание магния в сыворотке крови менее 0.65 ммоль/л — один из факторов риска внезапной сердечной смерти [18–20]. Так, показано при изучении 86000 женщин, что риск летального исхода повышается на 37% [20, 21]. Ограниченное на 50% (100–120 мг) поступление ионов магния с пищей индуцирует нарушения ритма и проводимости сердца, сосудистого тонуса и углеводного обмена и повышает риск инсульта [10, 22, 23]. Для снижения риска осложнений ИБС показано ежедневное потребление 250 мг магния.

Содержание металлов влияет не только на питательную ценность растений, но и на их устойчивость к стрессам. Так, выявлено, что катионный состав изменяет оводненность цитоплазмы. Гидратация цитоплазмы положительно коррелировала с содержанием щелочных металлов, представителем которых является калий, и отрицательно с количеством щелочноземельных (кальций, магний) [11–13]. Коллоидные частицы белка, несущие отрицательный заряд, как и катионы, окружены гидратной оболочкой. Заряженные частицы с большей массой Mg^{2+} , Ca^{2+} сильнее притягиваются к отрицательно заряженным белковым молекулам, нейтрализуя их заряд, что вызывает уменьшение количества молекул воды, окружающих белок. Одновалентно заряженные частицы слабее влияют на гидратацию, повышая оводненность тканей [14–16]. Соотношение Ca^{2+}/Mg^{2+} имеет большое значение для жизнедеятельности

растений и регулирует многие процессы обмена веществ [17–19]. При анализе селекционного материала надо учитывать, что концентрация микроэлементов до 10 раз выше в зародыше, отрубях, шелушенном рисе, чем в полированном белом рисе, который обычно употребляют в пищу. Так, магния в отрубях было 5.12%, а в полированном рисе 0.51% [20–22].

Данных по локализации хромосомных регионов определяющих содержание микро- и макроэлементов в сортах риса, пока немного и они противоречивы, так как используются различные маркеры и популяции для молекулярного маркирования [23–25]. Так, три локуса, определяющих содержание железа, группой K. Anuradha с соавт. локализованы на хромосомах 1, 5, 12, C. James с соавт. выявлены на хромосомах 2 и 8, L. Kaiyang с соавт. их установили на хромосомах 1 и 11 [25–27]. Содержание кальция в работе [23] определялось локусами первой, третьей и пятой хромосом, магния — третьей. В работе [27] авторы выявили локусы, определяющие содержание кальция, на третьей хромосоме. M. Zhang с соавт. показали, что содержание магния контролировали хромосомные регионы на хромосомах 6, 10, 11 [28]. Показано, что дикие виды *Oryza rufipogon*, *O. meridionalis*, *O. nivara* и *O. longistaminata* являются источниками QTLs, повышающих содержание таких элементов как железо, цинк, марганец, магний, медь, кальций, фосфор [28–30]. Наиболее полная карта расположения локусов, определяющих качество риса, с этой точки зрения дана в работах [23, 24]. В нескольких работах также показана роль транспортеров OsZIP1 и OsHMA3 в поддержании гомеостаза микроэлементов у образцов риса [31–33]. Рассмотрены методы биофортификации селекционного материала [34–36].

МАТЕРИАЛЫ И МЕТОДЫ

В настоящей работе использованы как коллекционные образцы из УНУ “Коллекция ФНЦ риса”, сорта селекции ООО “Аратай”, так и перспективные и допущенные к использованию в производстве сорта риса отечественной и зарубежной селекции (48 образцов). Среди них сорта различные по качеству: короткозерные, среднезерные, крупнозерные, длиннозерные, с различными кулинарными и технологическими свойствами. Растения выращивались в сосудах на оптимальном фоне минерального питания ($N_{120}P_{60}K_{60}$), в трех повторностях, густота стояния 10 растений на один сосуд. Контролем служил сорт Флагман с той же густотой стояния. При недостатке растений или их гибели подсаживался маркер с фиолетовыми листьями.

Определение содержания кальция и магния проводилось на приборе Капель 105. Для анализа отбирали материал от 20 растений каждого образца. Определяли процент вещества в пробе. Подго-

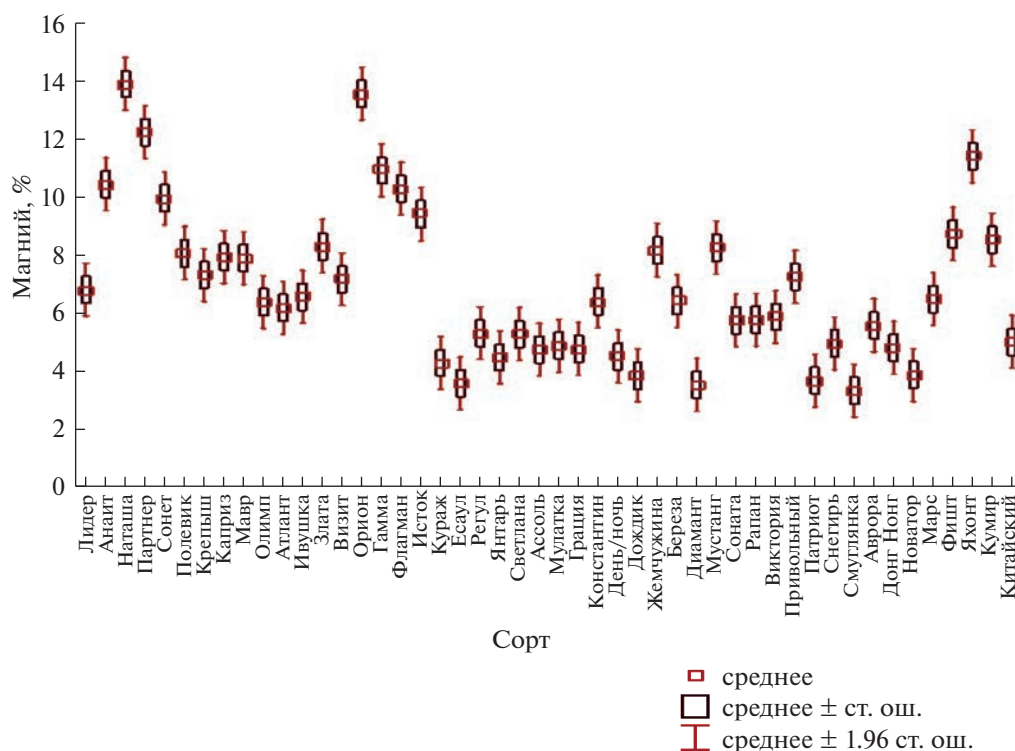


Рис. 1. Вариабельность сортов риса по содержанию магния.

товка проб для получения данных по содержанию катионов включала отбор растительного материала 1 г образца (от 20 растений), его гомогенизацию, фиксацию 10%-ным спиртом, выдержку в холодильнике одни сутки, центрифугирование при 10–15 тыс. оборотов в минуту, отбор супернатанта в чистую 2-мл пробирку до проведения анализа.

Молекулярное маркирование сортов для локализации хромосомных регионов, определяющих питательную ценность, проводили с использованием микросателлитных SSR-маркеров. Изучен полиморфизм этих сортов с использованием 58 маркеров, распределенных по семи хромосомам риса: 1, 2, 3, 5, 6, 8, 10. В работе использовали маркеры как связанные с локусами, контролирующими содержание микроэлементов, так и нейтральные, распределенные по геному риса.

ДНК риса выделяли из этиолированных проростков и листьев с помощью STAB-метода в различных модификациях. Постановку полимеразной цепной реакции (ПЦР) и визуализацию продуктов амплификации проводили по методике Международного института риса [37–39].

Параметры ПЦР, использованные в данном эксперименте: 5 мин при 94°C — начальная денатурация; следующие 35 циклов: 1 мин — денатурация при 94°C, 1 мин — отжиг праймеров при 55°C, 2 мин — синтез при 72°C; последний его цикл —

7 мин при 72°C. ПЦР-смесь включала 40 нг ДНК (2 мкл), 1 мкл (1 mM) дезоксинуклеотидтрифосфатов (dNTPs), 3.7 мкл H₂O, 1 мкл буферного раствора по 0.5 мкл (5 мкM) каждого праймера, 1 мкл (1.5 ед.) Taq-полимеразы, в общем объеме 10 мкл. Продукты амплификации разделяли электрофорезом в полиакриламидном геле при напряжении 100 В. Для статистической обработки полученных данных использовали программу Statistica [40].

РЕЗУЛЬТАТЫ

Вариабельность образцов по содержанию катионов кальция и магния

Содержание магния в образцах риса варьировало от двух до 14% (рис. 1). В 40 образцах его содержание не превышало 10%, в десяти образцах было выше 10%, но меньше 14%, что позволяет их рекомендовать как источники по признаку, это сорта: Наташа, Орион, Партнер, Гамма. Кальций также очень важен для организма, так как он стабилизирует клеточные мембраны, при его недостатке особенно страдает корневая система. Содержание кальция варьировало в образцах риса от 0.07 до 2.33% (рис. 2). В 36 образцах его содержание не превышало 1%, в семи образцах было выше 1.5% (Полевик, Крепыш, Ивушка, Флагман, Кураж, Наташа, Смуглянка). На основании про-

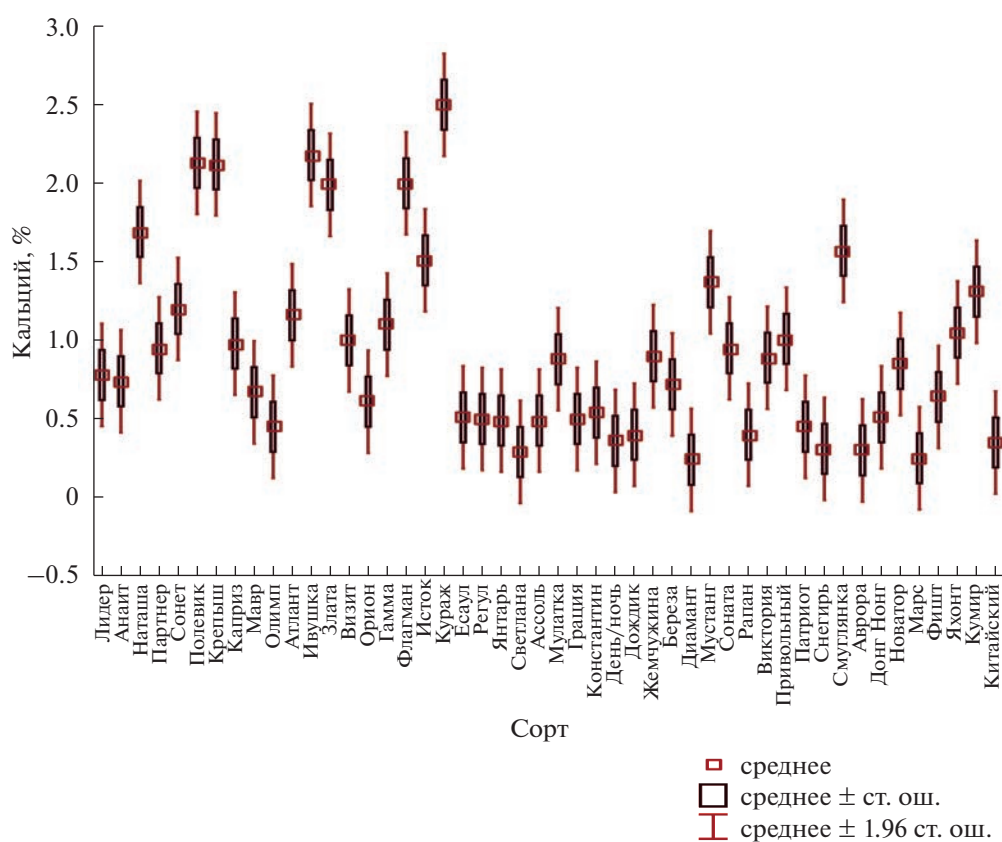


Рис. 2. Вариабельность сортов риса по содержанию кальция.

веденных исследований образцы разделили на контрастные группы по признакам: содержание магния и содержание кальция (табл. 1).

Поиск хромосомных регионов, определяющих содержание магния у отечественных сортов риса

Поиск хромосомных регионов, определяющих качество отечественных образцов, проводили с использованием 58 молекулярных маркеров, распределенных по геному риса. Данные по фенотипированию отечественных сортов риса по признаку “содержание магния” использовали для разделения их на группы с максимальным значением признака (1) и минимальным его значением (2). С помощью дисперсионного анализа установили связь между фенотипическим проявлением признака и генотипом образца.

Достоверно на уровне значимости, принятом для биологических исследований, контрастные по содержанию магния группы сортов разделяли четыре маркера: RM 542, RM 284, RM 405, RM 136 (табл. 2). Рис. 3, 4 иллюстрируют разделение

групп с использованием выделенных в исследовании маркеров. Размер продуктов амплификации при использовании различных маркеров во многих группах сортов варьирует. В первой группе мог быть меньше, чем во второй, также наблюдалась и обратная зависимость.

Локализация хромосомных регионов, определяющих содержание кальция

Данные по фенотипированию отечественных сортов риса по признаку “содержание кальция” также использовали для разделения их на группы с максимальным значением признака (1) и минимальным его значением (2) (табл. 1). Связь между фенотипическим проявлением признака и генотипом образца установили с помощью дисперсионного анализа. В табл. 2 представлены результаты дисперсионного анализа по разделению групп сортов с контрастным значением признака по содержанию кальция.

Достоверно на уровне значимости, принятом для биологических исследований, контрастные

Таблица 1. Содержание магния и кальция у разных групп сортов

Сорт	Кальций, %	Группа сортов	Сорт	Магний, %	Группа сортов
Мустанг	1.37	1	Грация	4.95	1
Исток	1.51	1	Исток	9.57	1
Смутлянка	1.57	1	Сонет	10.10	1
Наташа	1.69	1	Флагман	10.43	1
Злата	1.99	1	Анаит	10.58	1
Флагман	2.00	1	Гамма	11.06	1
Крепыш	2.12	1	Яхонт	11.54	1
Полевик	2.13	1	Партнер	12.37	1
Ивушка	2.18	1	Орион	13.69	1
Кураж	2.50	1	Наташа	14.03	1
Диамант	0.24	2	Смутлянка	3.52	2
Марс	0.25	2	Диамант	3.72	2
Светлана	0.29	2	Есаул	3.77	2
Аврора	0.30	2	Патриот	3.86	2
Снегирь	0.31	2	Дождик	4.04	2
Китайский	0.35	2	Новатор	4.06	2
День/ночь	0.36	2	Кураж	4.47	2
Дождик	0.40	2	Янтарь	4.66	2
Рапан	0.40	2	День/ночь	4.69	2
Олимп	0.45	2	Ассоль	4.92	2
НСР	0.17		НСР	0.21	

Таблица 2. Результаты дисперсионного анализа по разделению групп сортов с контрастным значением признака

Маркер	Сумма квадратов внутригрупповая	Средний квадрат	Сумма квадратов межгрупповая	Средний квадрат	<i>F</i>	<i>p</i>
По содержанию магния						
RM 284	1.125	1.125	0.750	0.125	9.000	0.024
RM 405	2.000	2.000	1.500	0.250	8.000	0.030
RM 136	1.125	1.125	0.750	0.125	9.000	0.024
RM 542	1.125	1.125	0.750	0.125	9.000	0.024
По содержанию кальция						
RM 154	2.408	2.408	3.467	0.578	4.168	0.087
RM 13	1.875	1.875	2.000	0.333	5.625	0.055
RM 2770	10.800	10.800	7.200	1.200	9.000	0.024
RM 3155	2.408	2.408	3.467	0.578	4.168	0.087

Примечание. *F* – критерий Фишера.

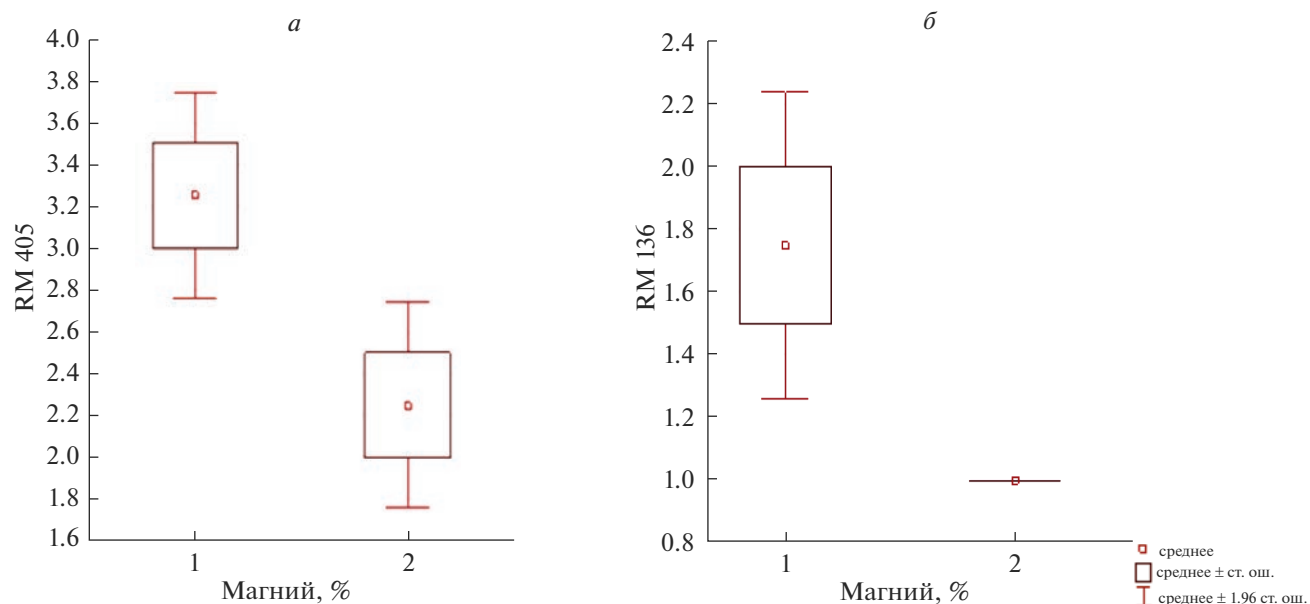


Рис. 3. Разделение контрастных по содержанию магния групп сортов с использованием маркеров RM 405, RM 136.

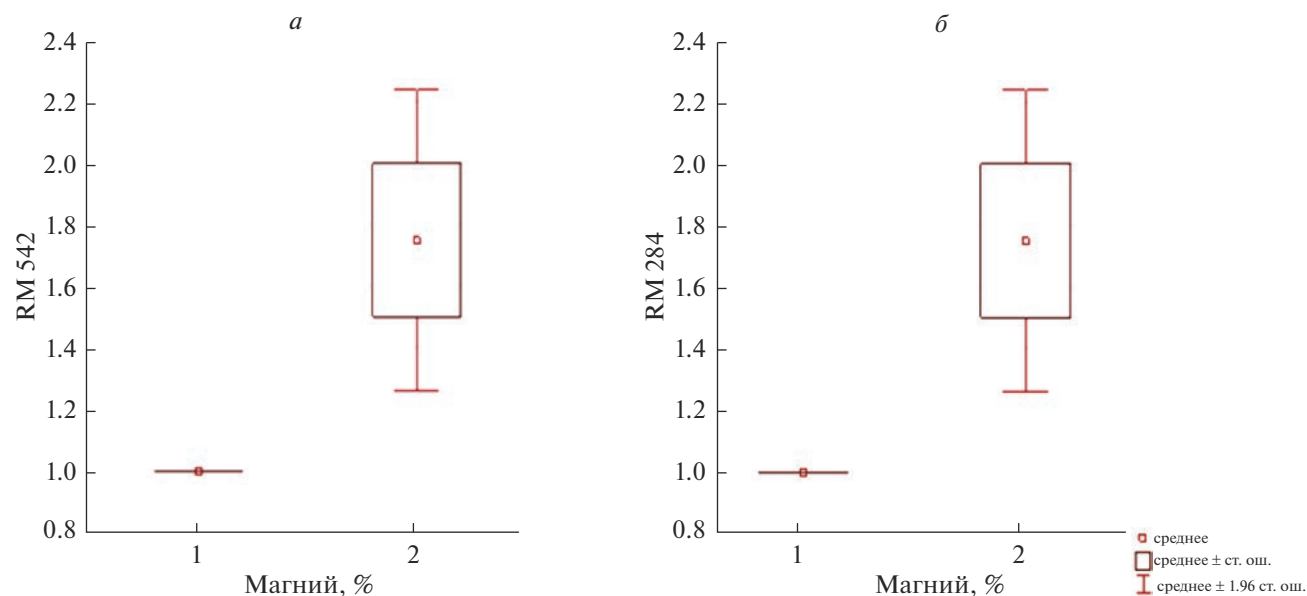


Рис. 4. Разделение контрастных по содержанию магния групп сортов с использованием маркеров RM 542, RM 284.

группы сортов разделял только маркер RM 2770. Рис. 5 иллюстрирует разделение групп с его использованием. Размер продуктов амплификации в первой группе сортов варьирует, во второй одинаков, но меньше чем в первой. Данный маркер также был сцеплен с локусом, определяющим содержание калия.

Поскольку выборка сортов в группах была небольшая, только локусы, определяющие значительные эффекты на формирование признака, могут быть выделены. Также снижает эффективность метода наличие нескольких локусов, определяющих признак, которые также влияют на фенотип образца. При снижении уровня значимости до 0.9 выде-

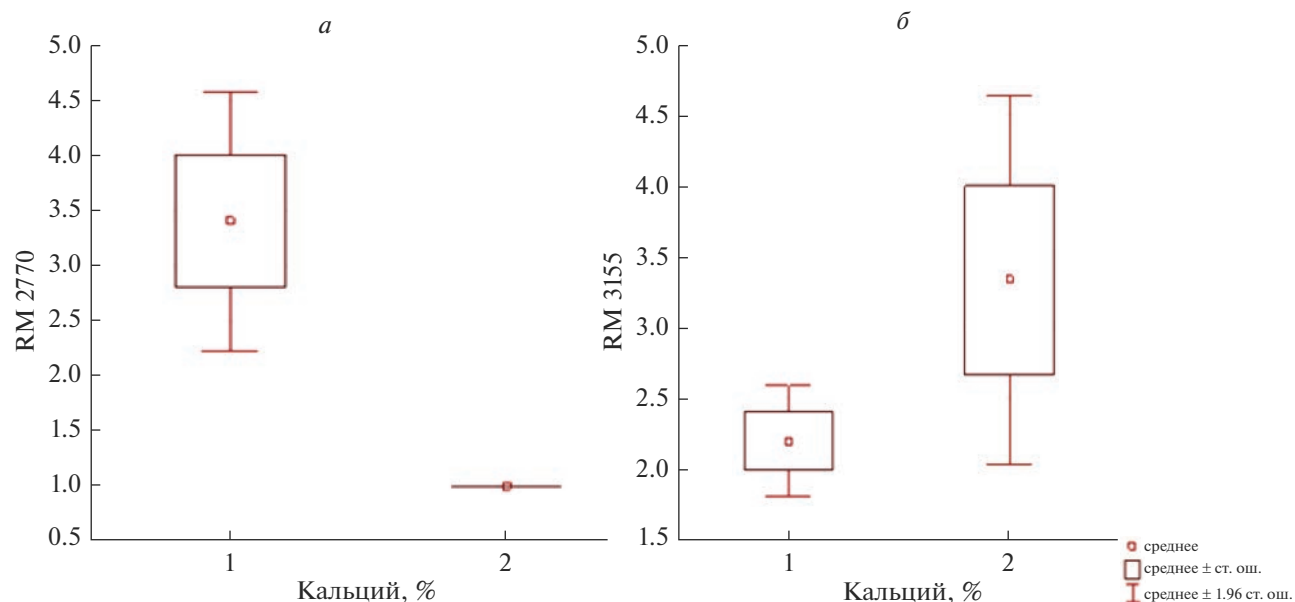


Рис. 5. Разделение контрастных по содержанию кальция групп сортов с использованием маркеров RM 2770, RM 3155.

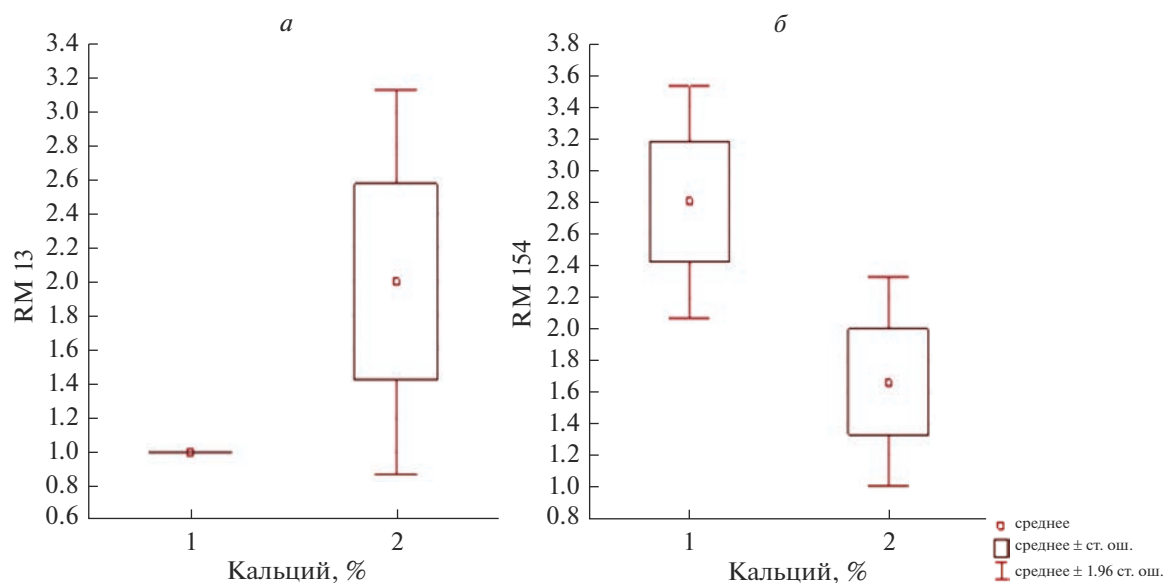


Рис. 6. Разделение контрастных по содержанию кальция групп сортов с использованием маркеров RM 13, RM 154.

ляются еще три локуса, возможно определяющих признак: RM 154, RM 13, RM 3155 (рис. 6).

В дальнейшем необходимо будет провести дополнительные исследования с большей выборкой сортов, чтобы подтвердить достоверность влияния этих локусов.

ОБСУЖДЕНИЕ

Нами выделены восемь локусов, определяющих содержание микроэлементов у отечественных сортов риса. Используемый метод позволяет выделить только самые значимые локусы, так как для анализа использовались контрастные

Таблица 3. Локусы, определяющие содержание магния и кальция

SSR-маркер	Хромосома	Размер продукта амплификации (пн); положение на карте, расстояние сМ	Ассоциация с признаком
Локусы, определяющие содержание магния			
RM 136	6	101; 51	Mg ²⁺ , пролин, тирозин
RM 542	7	113; 12.7	Mg ²⁺ , пролин
RM 284	8	141; 104	Mg ²⁺
RM 405	5	110; 28	Mg ²⁺ , лейцин
Локусы, определяющие содержание кальция			
RM 3155	8	174; 119	Ca ²⁺ , лейцин
RM 154	2	183; 4.8–13.8	Ca ²⁺
RM 2770	2	184; 4.7	Ca ²⁺ , калий
RM 13	5	141; 28.6	Ca ²⁺ , пролин

группы сортов по признаку [41–43]. Три из выделенных локусов связаны с содержанием пролина. Пролин — аминокислота, во многих случаях определяющая засухоустойчивость образца. Содержание магния и кальция также определяет гидратацию тканей, что косвенно подтверждает результаты исследования. Локусы, определяющие содержание магния, расположены на хромосомах 5, 6, 7, 8, размер продуктов амплификации выделенных локусов — от 101 до 141 пары нуклеотидов. Лocus на шестой хромосоме, в районе расположения маркера RM 136 (51 сМ) был связан с содержанием сразу двух аминокислот — пролина и тирозина. На пятой хромосоме locus в районе расположения маркера RM 405 (28 сМ) был связан с содержанием лейцина (табл. 3).

Связанных с содержанием кальция было выделено четыре локуса: один на второй хромосоме (при уровне значимости 0.05) и по одному на второй, восьмой и пятой (при уровне значимости 0.09). На пятой хромосоме locus расположен (RM 13, 28.6 сМ) в непосредственной близости к региону расположения маркера RM 405 (28 сМ), который связан с содержанием магния. Возможно, эти маркеры фланкируют регион, определяющий содержание микроэлементов. Подтверждает

достоверность полученных данных верификация результата в опубликованных источниках, ранее в этом регионе на пятой хромосоме также были локализованы QTL, определяющие эти два признака [23]. Для маркирования в этих работах использовали популяцию интрогрессивных линий, полученную при гибридизации сорта Teging и дикого вида риса *Oryza rufipogon*. На восьмой хромосоме, в районе близком к расположению маркеров, выявленных нами, также выявили locus, определяющий содержание микроэлементов. На второй хромосоме разными авторами также были выявлены локусы, определяющие содержание различных микроэлементов. В работе A. Garcia-Oliveira и соавт. [23] — это железо и марганец; C. James и соавт. [25] выявили locus, определяющий содержание железа. В работах [24, 28] этот регион также определял содержание меди и железа. То есть полученные нами данные подтверждаются данными других исследователей, полученными на других популяциях. Возможно, в этих регионах расположены кластеры генов, определяющие качество и содержание микроэлементов. Полученные данные позволят более эффективно переносить гены, определяющие качество образ-

цов, как из диких видов риса, так и перспективных селекционных образцов.

Работа поддержана грантом Фонда содействия инновациям № 4654ГС2/48601.

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. Гончарова Ю.К., Харитонов Е.М., Малюченко Е.А., Бушман Н.Ю. Молекулярное маркирование признаков, определяющих качество зерна у российских сортов риса // Вавил. журн. генетики и селекции. 2018. Т. 26. № 1. С. 79–87. <https://doi.org/10.18699/VJ18.334>
2. Фотев Ю.В., Пивоваров В.Ф., Артемьева А.М. и др. Концепция создания Российской национальной системы функциональных продуктов питания // Вавил. журн. генетики и селекции. 2018. Т. 22. № 7. С. 776–783. <https://doi.org/10.18699/VJ18.421>
3. Харитонов Е.М., Гончарова Ю.К. Взаимосвязь между устойчивостью к высоким температурам и стабильностью урожаев у риса // Аграрная Россия. 2008. Т. 3. № 1. С. 2–24. <https://doi.org/10.30906/1999-5636-2008-3-22-24>
4. Харитонов Е.М., Гончарова Ю.К., Иванов А.Н. Применение кластерного анализа для разделения сортов риса по реакции на изменение условий среды // Вестник Росс. акад. с.-х. наук. 2014. № 6. С. 32–35.
5. Liu X.S., Feng S.J., Zhang B.Q. et al. OsZIP1 functions as a metal efflux transporter limiting excess zinc, copper and cadmium accumulation in rice // BMC Plant Biol. 2019. № 19. P. 283. <https://doi.org/10.1186/s12870-019-1899-3>
6. Fan F., Long W., Liu M. et al. Quantitative trait locus mapping of the combining ability for yield-related traits in wild rice *Oryza longistaminata* // J. Agric. Food Chem. 2019. № 67. P. 8766–8772. <https://doi.org/10.1021/acs.jafc.9b02224>
7. Iseri L.T., Allen B.J., Ginkel M.L., Brodsky M.A. Ionic biology and ionic medicine in cardiac arrhythmias with particular reference to magnesium // Am. Heart J. 1992. V. 123. № 5. P. 1404–1409. [https://doi.org/10.1016/0002-8703\(92\)91059-a](https://doi.org/10.1016/0002-8703(92)91059-a)
8. Mubagwa K., Gwanyanya A., Zakharov S., Macianskiene R. Regulation of cation channels in cardiac and smooth muscle cells by intracellular magnesium // Arch. Biochem. Biophys. 2007. V. 458. № 1. P. 73–89. <https://doi.org/10.1016/j.abb.2006.10.014>
9. Qu X., Jin F., Hao Y. et al. Magnesium and the risk of cardiovascular events: A meta-analysis of prospective cohort studies // PLoS One. 2013. V. 8. № 3. <https://doi.org/10.1371/journal.pone.0057720>
10. Сыркин А.Л., Салагаев Г.И., Сыркина Е.А., Лысенко А.В. Преимущества оротата магния для коррекции магни-дефицитных состояний у больных с различными формами нарушений ритма сердца // Кардиология и сердечно-сосудистая хирургия. 2019. Т. 12. № 4. С. 308–313. <https://doi.org/10.17116/kardio201912041308>
11. Sheehan J.P., Seelig M.S. Interactions of magnesium and potassium in the pathogenesis of cardiovascular disease // Magnesium. 1984. V. 3. № 4–6. P. 301–314. Corpus ID: 2095870.
12. Leor J., Kloner R.A. An experimental model examining the role of magnesium in the therapy of acute myocardial infarction // Am. J. Cardiology. 1995. V. 75. № 17. P. 1292–1293. [https://doi.org/10.1016/s0002-9149\(99\)80787-5](https://doi.org/10.1016/s0002-9149(99)80787-5)
13. Paravicini T.M., Chubakov V., Gudermann T. TRPM7: A unique channel involved in magnesium homeostasis // Intern. J. Biochem. Cell Biol. 2012. V. 44. № 8. P. 1381–1384. <https://doi.org/10.1016/j.biocel.2012.05.010>
14. Whang R., Welt L.A. Observations in experimental magnesium depletion // J. Clin. Investigation. 1963. V. 42. № 3. P. 305–313. <https://doi.org/10.1172/JCI104717>
15. Solomon R. The relationship between disorders of K⁺ and Mg²⁺ homeostasis // Semin Nephrol. 1987. V. 7. № 1. P. 253–262.
16. Weiner I.D., Wingo C.S. Hypokalemia: Consequences, causes, and correction // J. Am. Soc. Nephrology. 1997. V. 8. № 7. P. 1179–1188. <https://doi.org/10.1681/ASN.V871179>
17. Сюсин И.В. Влияние ионов Ca²⁺ на фосфолипидный состав, состояние и морфологические характеристики эритроцитов: Дисс. ... канд. биол. наук. Саранск, 2021. 123 с.
18. Wong N.L., Sutton R.A., Navichak V. et al. Enhanced distal absorption of potassium by magnesium-deficient rats // Clin. Science. 1985. V. 69. № 5. P. 625–630. <https://doi.org/10.1042/cs0690625>
19. Shechter M. Magnesium and cardiovascular system // Magnesium Res. 2010. V. 23. № 2. P. 60–72. <https://doi.org/10.1684/mrh.2010.0202>
20. Громова О.А., Торшин И.Ю., Калачева А.Г. и др. О некоторых ролях калия и магния в терапевтической практике // Лечебное дело. 2019. Т. 2. С. 21–31. <https://doi.org/10.24411/2071-5315-2019-12108>
21. Song Y., He K., Levitan E.B. et al. Effects of oral magnesium supplementation on glycaemic control in Type 2 diabetes: A meta-analysis of randomized double-blind controlled trials // Diabet Medicine. 2006. V. 23. № 10. P. 1050–1056. <https://doi.org/10.1111/j.1464-5491.2006.01852.x>

22. Ohira T., Peacock J.M., Iso H. et al. Serum and dietary magnesium and risk of ischemic stroke: the atherosclerosis risk in communities study // *Am. J. Epidemiology*. 2009. V. 169. № 12. P. 1437–1444. <https://doi.org/10.1093/aje/kwp071>
23. Garcia-Oliveira A.L., Tan L., Fu Y., Sun C. Genetic identification of quantitative trait loci for contents of mineral nutrients in rice grain // *J. Integrative Plant Biol.* 2009. V. 51. № 1. P. 84–92. <https://doi.org/10.1111/j.1744-7909.2008.00730.x>
24. Mahender A., Anandan A., Kumar S. et al. Rice grain nutritional traits and their enhancement using relevant genes and QTLs through advanced approaches // *Springer Plus*. 2016. V. 5. № 1. P. 2086. <https://doi.org/10.1186/s40064-016-3744-6>
25. James C.R., Huynh B.L., Welch R.M. et al. Quantitative trait loci for phytate in rice grain and their relationship with grain micronutrient content // *Euphytica*. 2007. V. 154. № 3. P. 289–294. <https://doi.org/10.1007/s10681-006-9211-7>
26. Anuradha K., Agarwal S., Rao Y.V. et al. Mapping QTLs and candidate genes for iron and zinc concentrations in unpolished rice of Madhukar * Swarna RILs // *Gene*. 2012. V. 508. № 2. P. 233–240. <https://doi.org/10.1016/j.gene.2012.07.054>
27. Kaiyang L., Lanzhi L., Xingfei Z. et al. Quantitative trait loci controlling Cu, Ca, Zn, Mn and Fe content in rice grains // *J. Genet.* 2008. V. 87. № 3. P. 305–310. <https://doi.org/10.1007/s12041-008-0049-8>
28. Zhang M.W., Guo B.J., Peng Z.M. Genetic effects on Fe, Zn, Mn and P contents in Indica black pericarp rice and their genetic correlations with grain characteristics // *Euphytica*. 2004. V. 135. № 3. P. 315–323.
29. Гончарова Ю.К., Гончаров С.В. Изучение и использование дикорастущего африканского *Oryza longistaminata* в качестве донора признаков аллогамии в селекции культурного риса на гетерозис // В сб. тезисов межд. конф. “Генетические ресурсы культурных растений. Проблемы мобилизации, инвентаризации, сохранения и изучения генофонда важнейших сельскохозяйственных культур для решения приоритетных задач селекции”, 13–16 ноября 2001 г. СПб., 2001. С. 253–255.
30. Liu X., Fan F., Liu M. Quantitative trait loci mapping of mineral element contents in brown rice using backcross inbred lines derived from *Oryza longistaminata* // *Front. Plant Sci.* 2020. V. 11. P. 1229. <https://doi.org/10.3389/fpls.2020.01229>
31. Liu X.S., Feng S.J., Zhang B.Q. et al. OsZIP1 functions as a metal efflux transporter limiting excess zinc, copper and cadmium accumulation in rice // *BMC Plant Biol.* 2019. V. 19. P. 283. <https://doi.org/10.1186/s12870-019-1899-3>
32. Cai H., Huang S., Che J. et al. The tonoplast-localized transporter OsHMA3 plays an important role in maintaining Zn homeostasis in rice // *J. Exp. Bot.* 2019. V. 70. P. 2717–2725. <https://doi.org/10.1093/jxb/erz091>
33. Wu Q., Liu C., Wang Z. et al. Zinc regulation of iron uptake and translocation in rice (*Oryza sativa* L.): Implication from stable iron isotopes and transporter genes // *Environ. Pollut.* 2022. V. 297. <https://doi.org/10.1016/j.envpol.2022.118818>
34. D'Amato R., Fontanella M. C., Falcinelli B. et al. Selenium biofortification in rice (*Oryza sativa* L.) sprouting: Effects on Se yield and nutritional traits with focus on phenolic acid profile // *J. Agric. Food Chem.* 2018. V. 66. P. 4082–4090. <https://doi.org/10.1021/acs.jafc.8b00127>
35. He J., Rossner N., Hoang M. et al. Transport, functions, and interaction of calcium and manganese in plant organellar compartments // *Plant Physiol.* 2021. V. 187. P. 1940–1972. <https://doi.org/10.1093/plphys/kiab122>
36. Adeva C., Yun Y.-T., Shim K.-C. et al. QTL mapping of mineral element contents in rice using introgression lines derived from an interspecific cross // *Agronomy*. 2023. V. 13. P. 76. <https://doi.org/10.3390/agronomy13010076>
37. Гончарова Ю.К., Харитонов Е.М. Генетический контроль признаков, связанных с усвоением фосфора у сортов риса (*Oryza sativa* L.) // Вавил. журн. генетики и селекции. 2015. Т. 19. № 2. С. 197–204. <https://doi.org/10.18699/VJ15.025>
38. Fan F., Long W., Liu M. et al. Quantitative trait locus mapping of the combining ability for yield-related traits in wild rice *Oryza longistaminata* // *J. Agric. Food Chem.* 2019. V. 67. P. 8766–8772. <https://doi.org/10.1021/acs.jafc.9b02224>
39. Jin J., Long W., Wang L. et al. QTL mapping of seed vigor of backcross inbred lines derived from *Oryza longistaminata* under artificial aging // *Front. Plant Sci.* 2018. № 9. P. 1909. <https://doi.org/10.3389/fpls.2018.01909>
40. Goncharova Yu.K., Goncharov S.V., Chicharova E.E. Localization of chromosome regions controlling high photosynthetic potential in Russian rice cultivars // *Russ. J. Genet.* 2018. V. 54. № 7. P. 796–804. <https://doi.org/10.1134/S0016675818070032>
41. Butardo V.M., Jr., Sreenivasulu N., Juliano B.O. Improving rice grain quality: State-of-the-art and future prospects // *Methods Mol. Biol.* 2019. V. 19. P. 55. https://doi.org/10.1007/978-1-4939-8914-0_2
42. Birla D.S., Kapil Malik, Manish Sainger et al. Progress and challenges in improving the nutritional quality of rice (*Oryza sativa* L.) // *Crit. Rev. Food Sci. Nutr.* 2017. V. 57. P. 2455–2481. <https://doi.org/10.1080/10408398.2015.1084992>
43. Zhang L., Ma B., Bian Z. et al. Grain size selection using novel functional markers targeting 14 genes in rice // *Rice*. 2020. V. 13. P. 63. <https://doi.org/10.1186/s12284-020-00427-y>

Localization of Chromosomal Regions Determining Magnesium and Calcium Content in Rice Varieties

J. K. Goncharova^{a, c, *}, V. V. Simonova^a, S. V. Goncharov^b, and N. A. Ochkas^c

^a*Federal Scientific Center of Rice, Krasnodarskiy krai, p. Belozerny, 350921 Russia*

^b*Trubilin Kuban State Agrarian University, Krasnodar, 350044 Russia*

^c*“Aratai” LLC, Participant of Skolkovo Innovation Center, Moscow, 143026 Russia*

^{*}*e-mail: yuliya_goncharova_20@mail.ru*

There has been a wide variation in the content of micro and macro elements, both between subspecies and within rice species and varieties. Magnesium as a cofactor is involved in more than 300 enzymatic reactions and is necessary for growth and development of both plants and humans. It affects carbohydrate and protein metabolism. Both adenosine triphosphate (ATP) production, nucleotide and glucose synthesis, and lipid oxidation regulation depend on its concentration. Calcium is also essential for the formation of cell wall structure and cell division. High concentrations of calcium change the composition and condition of red blood cell membranes and cell morphological characteristics. At the same time the variability in signs of domestic rice varieties has not yet been studied. There have been no studies on the localization of chromosomal regions responsible for the formation of traits of magnesium and calcium content in rice samples. Wide variation in calcium content (0.07–2.33%) and magnesium content in rice samples (2–14%) was established. Contrasting groups of varieties and sources by signs were identified. A search for chromosomal regions determining the quality of domestic samples was carried out using 58 molecular markers distributed across the rice genome (SSR). The data on phenotyping of native rice varieties by quality signs were used to divide them into groups with maximum trait value (1) and minimum trait value (2). The relationship between phenotypic manifestation of the trait and the genotype of the sample was established: by means of analysis of variance. Eight loci determining magnesium and calcium content in native rice varieties – were identified. The loci determining magnesium content are located on chromosomes 5, 6, 7, 8. Four loci associated with calcium content were identified two on the second chromosome and one each on the eighth and fifth chromosomes. On the fifth chromosome, it is located (RM 13, 28.6 cM) in close proximity to the RM 405 marker region (28 cM), which is connected with magnesium content.

Keywords: genetics, rice, variability, quantitative trait loci, molecular marking, quality, calcium, magnesium.

РАСПРОСТРАНЕНИЕ И СИЛА АЛЛЕЛЕЙ ГЕНОВ ГИБРИДНОГО НЕКРОЗА В ГЕНОТИПАХ КАНАДСКИХ СОРТОВ ЯРОВОЙ МЯГКОЙ ПШЕНИЦЫ (*Triticum aestivum* L.)

© 2023 г. В. А. Пухальский^{1, *}, В. С. Рубец², Е. Н. Билинская¹, А. М. Кудрявцев¹

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

²Российский государственный аграрный университет — МСХА им. К.А. Тимирязева, Москва, 127434 Россия

*e-mail: pukhalsk@vigg.ru

Поступила в редакцию 08.06.2023 г.

После доработки 22.06.2023 г.

Принята к публикации 28.06.2023 г.

Исследовано генетическое разнообразие 30 канадских сортов яровой мягкой пшеницы (*Triticum aestivum* L.). Показано наличие двух типов сортов яровой мягкой пшеницы Канады, различающихся по некрозным генотипам. Одни сорта имели генотип *ne1ne1Ne2Ne2*, а вторые — генотип *ne1ne1ne2ne2*. Генотип *Ne1Ne1ne2ne2* отсутствует. Гены *Ne2* представлены аллелями средней силы — *ms* и сильными — *s*. По родословным определены сорта-доноры гена *Ne2*.

Ключевые слова: гены гибридного некроза, комплементарное взаимодействие генов, сила аллелей, яровая мягкая пшеница, *Triticum aestivum* L.

DOI: 10.31857/S0016675823110115, EDN: NNKKRK

Оптимизация селекционного процесса при создании новых сортов пшеницы напрямую зависит от результатов изучения геномов исходного селекционного материала с использованием различных генетических маркеров. К таким маркерам можно отнести и гены гибридного некроза, позволяющие выявлять как видовую специфичность этих генов [1], так и их распределение по геномам в различных регионах и странах [2–5]. В результате появляется возможность понимания микроэволюционных процессов, происходящих в популяциях сортов пшеницы в результате селекции и влияния на нее антропогенного фактора [6].

Фенотипически гибридный некроз проявляется в летальности или сублетальности растений первого гибридного поколения в результате отмирания и недоразвития листьев в разные фазы онтогенеза пшеничного растения (табл. 1). Гибридный некроз определяется комплементарным взаимодействием двух доминантных генов *Ne1* и *Ne2* [7–9]. Оба гена находятся в геноме В: ген *Ne1* — на хромосоме 5BL на расстоянии от центromеры 10.5 ± 2.0 единиц генетической карты, а ген *Ne2* — на хромосоме 2BS на расстоянии 9.4 ± 1.5 единиц генетической карты от центromеры [10, 11]. Эти данные подтверждены созданием молекулярных карт гибридного некроза [12–14]. Для генов *Ne1* и *Ne2* установлена серия из трех аллелей: слабый *w* (*weak*), средний *m* (*moderate*) и сильный *s* (*strong*). Для гена *Ne2* описаны еще два аллеля промежу-

точной силы — *wm* (*moderately*) и *ms* (*moderately strong*). Степень проявления гибридного некроза зависит от сочетания в генотипах первого поколения различных по силе аллелей генов гибридного некроза (табл. 2).

Силу аллелей генов гибридного некроза определяют только в первом гибридном поколении, так как в F_2 существуют различия между растениями в проявлении признаков некроза, зависящие от дозы генов некроза в генотипах [6].

В целом, проведение исследований по установлению частоты встречаемости сортов пшеницы, отягощенных генами гибридного некроза, обусловлено возможностью проводить на этой основе дальнейшую оптимизацию генетической теории селекции.

В настоящем сообщении приводятся данные по сравнительному исследованию отягощенности генотипов генами гибридного некроза современных канадских сортов яровой мягкой пшеницы, допущенных для возделывания в разных регионах Канады.

МАТЕРИАЛЫ И МЕТОДЫ

Материалом для наших исследований послужили образцы 30 сортов яровой мягкой пшеницы (*Triticum aestivum* L.) Канады, полученные из коллекции Всероссийского института генетических

Таблица 1. Степень проявления гибридного некроза ([8] с сокращением)

Степень проявления	Фаза проявления первого симптома	Степень угнетения	Масса 1000 зерен
0	Нет симптомов	Нормальное созревание	Норма
1	Максимальная высота растения	Нормальное созревание	Норма
2	Колошение	Преобладают нормальные растения	В большинстве случаев норма
3	Выход в трубку	Несколько ослабленный рост	Менее 30 г
4	Кушение	Преждевременное прекращение роста	15–30 г
5	Начало кушения	Уменьшенное число колосьев, колосков	Менее 15 г
6	2–3 листа	Мало колосьев, все без зерен	–
7	2 листа	Несколько побегов, не образующих колосья	–
8	1–2 листа	3–6 листьев, побег	–

ресурсов растений им. Н.И. Вавилова. В качестве тестеров использовали сорта яровой мягкой пшеницы Marquillo (генотип $Ne1^fNe1^sne2ne2$), Triso ($ne1ne1Ne2^sNe2^s$), Granni ($ne1ne1Ne2^sNe2^s$), Мисс ($ne1ne1Ne2^sNe2^s$). Скрещивание сортов с тестерами проводили в полевых условиях твел-методом с изоляцией колосьев. Наличие в генотипах генов гибридного некроза и силу их аллелей определяли в F_1 на разных стадиях онтогенеза растений, используя критерии, представленные в табл. 1, 2.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Полученные данные приведены в табл. 3. Как видно из приведенных данных, 18 из изученных сортов имеют генотип $ne1ne1Ne2Ne2$ (носители гена $Ne2$), а 12 сортов являются неносителями генов некроза (генотип $ne1ne1ne2ne2$).

Анализ родословных изучавшихся сортов позволил установить сорта, передавшие в результате скрещиваний ген некроза вновь созданным сортам [15]. Так, в геном сорта Katerwa ген некроза попал в результате скрещивания от сорта Неерава, дважды встречавшегося в родословной. Сорт Нараяо получил ген некроза от сорта Maniton, который в свою очередь унаследовал этот ген от сорта Thatcher. Сорт Leader получил ген некроза от сорта

Chris (генотип $ne1ne1Ne2Ne2$). В генотип сорта Roblin ген $Ne2$ попал или от одной из линий – RL-4359 или RL-4353, в родословную которых входит сорт Неерава (генотип $ne1ne1Ne2Ne2$). Возможно сорт Roblin унаследовал по одной хромосоме 2BS с геном $Ne2$ от каждой из этих линий. Сорт Colymbus, входящий в родословную сортов AC Minto, AC Kenzie, Pasqua, передал в их генотипы ген $Ne2$, который в свою очередь был им получен от сорта Неерава. От сорта Неерава ген $Ne2$ вошел и в родословную сорта Kenyon. Ген $Ne2$ в генотип сорта Cutler мог попасть из генотипов двух сортов. Это сорта Ciano-67 (генотип $ne1ne1Ne2^wNe2^w$) и Sonora-64 (генотип $ne1ne1Ne2^sNe2^s$), входящие в его родословную [16, 17]. Эти же два сорта входят в родословную сорта Oslo, но как и в предыдущем случае, трудно определить донора гена $Ne2$. Скорее всего – это сорт Sonora-64, так как сорт Ciano-67 несет слабый аллель w гена $Ne2$, а сорт Oslo – сильный аллель s .

В генотип сорта Biggar ген $Ne2$ попал от сорта Tobari-66. Сорт AC Taber унаследовал ген $Ne2$ от линии HY-320 (генотип $ne1ne1Ne2^mNe2^m$). Линия RL-4386 (генотип $ne1ne1Ne2^mNe2^m$), получившая ген $Ne2$ от сорта Неерава, передала этот ген сорту CDC Merlin. При создании сорта AC Majestic в скрещиваниях дважды использовались сорта Colymbus и Неерава, несущие ген $Ne2$. Поэтому трудно определить, какой из этих сортов является в данном случае донором этого гена.

У гибридов первого поколения от скрещивания тестеров с сортами Katerwa, Chester, Wildecatt, Roblin, AC Minto, CDC Teal, Pasqua, Kenyon, Cutler, Oslo, MC Kenzie, AC Taber, AC Majestic и AC Domain симптомы гибридного некроза в F_1 проявились в начале появления третьего листа, а затем растения погибали, что свидетельствует о присутствии в их геноме сильного аллеля гена

Таблица 2. Степень гибридного некроза в баллах в зависимости от сочетания аллелей генов $Ne1$ и $Ne2$ [9]

$Ne1$	$Ne2$				
	w	wm	m	ms	s
w	0	0	0–1	1–2	2–3
m	1–2	3–4	4–5	5–6	6
s	3–4	5–6	6–7	7–8	8

Таблица 3. Генотипы по генам гибридного некроза канадских сортов яровой мягкой пшеницы

№ п.п.	Каталог ВИР	Сорт	Родословная	Генотип*
1	к-49261	Glenlea	UM-530/(Vex)CB-100	<i>ne 1ne2</i>
2	к-59026	Katepwa	Neepawa*6/RL-2438/3/Neepawa*6//CI-81542/2*-Frocor	<i>ne 1Ne2^s</i>
3	к-60583	Napayo	Manitou*2/RL-4124	<i>ne 1Ne2^{ms}</i>
4	к-61224	Chester	Renown/s-615//Rescue/3/kendee/4/Mida/Cadet	<i>ne 1Ne2^s</i>
5	к-61227	Leader	Fortuna/Chris	<i>ne 1Ne2^{ms}</i>
6	к-62153	Altar	Tobart-66(SIB)//Desconocido/ Frocor	<i>ne 1ne2</i>
7	к-62599	Bluesky	Potam-70/Gleulea	<i>ne 1ne2</i>
8	к-62600	Wildcat	NB-113/Glenlea	<i>ne 1Ne2^s</i>
9	к-62853	Laura	BW-15/BW-517	<i>ne 1ne2</i>
10	к-62855	Roblin	RL-4302/RL-4356//RL-4359/RL-4353	<i>ne 1Ne2^s</i>
11	к-62878	Ac Minto	Columbus/BW-63//Katepwa/BW-552	<i>ne 1Ne2^s</i>
12	к-62898	CDC Teal	BW-514/Benito//BW38(CAN)	<i>ne 1Ne2^s</i>
13	к-63179	Pasqua	BW-63*2/Columbus	<i>ne 1Ne2^s</i>
14	к-63482	Kenyon	Neepawa*5/Bick-Manantial	<i>ne 1Ne2^s</i>
15	к-64150	BW90	9505-LP-03-A/Jorney//Lillian	<i>ne 1ne2</i>
16	к-64447	Cutler	Ciano-67/4/Sonora-64/Yaqui-50-E-5//Gaboto/3/Inia-66	<i>ne 1Ne2^s</i>
17	к-64448	Oslo	Sonora-64/Yaqui-50-E//Guajolote/3/Inia-66/4/Ciano-67//Elgin/Sonora-64	<i>ne 1Ne2^s</i>
18	к-64561	Biggar	Tobari-66/Romany-66	<i>ne 1Ne2^{ms}</i>
19	к-64562	AC Nanda	FB-42/B-109	<i>ne 1ne2</i>
20	к-64563	MC Kenzie	Columbus/Amidon	<i>ne 1Ne2^s</i>
21	к-64565	AC Cadillac	BW-90*3/BW553	<i>ne 1ne2</i>
22	к-64596	AC Barrie	Neepawa/Cjlumbus(Cid-188385)//BW-90	<i>ne 1ne2</i>
23	к-64699	AC Taber	HY-320*3/BW-553	<i>ne 1Ne2^s</i>
24	к-64976	CDC Merlin	RL-4386//BW-525/BW-37	<i>ne 1Ne2^{ms}</i>
25	к-64978	AC Phil	PT-303/Dirkwin//Keniya-321/Fieldwin	<i>ne 1ne2</i>
26	к-64979	AC Majestic	Columbus*2//Saric-70/Neepawa/3/Columbus*5//Saric-70/Neepawa	<i>ne 1Ne2^s</i>
27	к-64980	AC Corinne	Glenlea*6/RL-4137	<i>ne 1ne2</i>
28	к-65005	AC Cabriel	Frontana/2*Belvedere	<i>ne 1ne2</i>
29	к-62877	AC Reed	PT-303/Dirkwin//Kenya-321/Feldwin	<i>ne 1ne2</i>
30	62919	AC Domain	BW-83/ND-585	<i>ne 1Ne2^s</i>

* — в таблице представлен гаплоидный генотип.

Ne2 — *s*. Аллель средней силы *ms* характерен для генов гибридного некроза сортов Napayo, Ltader, Biggar, CDC Merlin. Растения этих сортов или образывали побеги, не доходящие до образования колоса, или же гибнущие в фазу 4–5 листьев.

Полученные данные показали, что за длительный период селекции (около, и даже более, 100 лет) селекционерам Канады не удалось полностью освободить селекционные популяции пшеницы от генов гибридного некроза. Объясняется этот феномен сцеплением некротических генов (осо-

бенно гена *Ne2*) с генами, детерминирующими устойчивость к бурой листовой ржавчине пшеницы [18]. Так, в хромосоме 2В, несущей ген *Ne2*, содержится четыре гена — *Lr13*, *Lr16*, *Lr23* и *Lr35*. Этим генам свойственно аддитивное взаимодействие, обуславливающее высокую устойчивость к бурой листовой ржавчине [19]. Среди этих генов особое значение в условиях Канады имеет ген устойчивости взрослых растений *Lr13*, не теряющий свою эффективность более пятидесяти лет [20]. Как считают А.Р. Roelts [21] и J.A. Koloner

[22], вообще гены *Lr13* и *Lr34* (хромосома 7D) наиболее эффективно детерминируют устойчивость сортов пшеницы к бурой листовой ржавчине в большинстве регионов мира. По имеющимся у нас данным [15] сорта с генотипом *ne1ne1Ne2Ne2* (табл. 3) несут в своем геноме ген *Lr13* в сочетании, в зависимости от генома, и с другими генами устойчивости к бурой листовой ржавчине. В геномах сортов с генотипом *ne1ne1Ne2ne2* ген *Lr13* установлен только у пяти сортов: Glenela, CDC Teal, AC Barrie, AC Reed, AC Cabriel, получившими этот ген в результате кроссинговера.

Интересно, что (по литературным данным [6]) некрозные генотипы канадских сортов яровой мягкой пшеницы, созданных в период с 1907 по 1974 гг., выглядели следующим образом. Группа созданных сортов, но не допущенных к коммерческому использованию, состояла из 11 образцов и включала 3 сорта с генотипом *ne1ne1Ne2Ne2* — два сорта с генотипом *Ne1Ne1Ne2ne2* и шесть сортов с генотипом *ne1ne1Ne2ne2*. Группа сортов, допущенных к коммерческому использованию, состояла из 14 сортов, из которых генотип *ne1ne1Ne2Ne2* имели 4 сорта, генотип *Ne1Ne1Ne2ne2* — 2 сорта и генотип *ne1ne1Ne2ne2* — 8 сортов. Одновременно было отобрано и зафиксировано 27 селекционных линий, из которых генотип *Ne1Ne1Ne2ne2* имели 10 линий (37%), *ne1ne1Ne2Ne2* — 8 линий (30%) и генотип *ne1ne1Ne2ne2* — 9 линий (33%). В дальнейшем линии с генотипами *ne1ne1Ne2Ne2* и *ne1ne1Ne2ne2* широко использовались в селекционных программах. Сравнение этих данных с полученными нами (табл. 3) свидетельствует, что селекционные программы Канады привели к изъятию из геномов сортов яровой мягкой пшеницы гена *Ne1* и существенному увеличению генотипов, несущих ген *Ne2*, а вместе с ним — к увеличению эффективных в условиях ряда регионов Канады генов устойчивости к бурой листовой ржавчине *Lr13*, *Lr16*, *Lr23* и *Lr35*.

При этом селекционеры Канады, как и селекционеры Австралии, Китая и России [13, 23], часто не зная, несет ли генотип сорта ген *Ne2*, при выборе исходного материала для скрещивания ориентируются прежде всего на устойчивость сортов к листовой бурой ржавчине, наиболее эффективно определяемой геном *Lr13* в сочетании с генами *Lr16* и *Lr23*. Одновременно может иметь значение и факт селективного преимущества генотипа *ne1ne1Ne2Ne2* над генотипом *Ne1Ne1Ne2ne2* — наличием генов, определяющих склонность растений к влажным условиям произрастания [24], а также обладающими другими хозяйственно ценными признаками [25, 26]. Все это в целом показывает, что ген *Ne2* оказывает определенное влияние на изменение подходов при составлении программ селекции пшеницы в определенных регионах мира.

Однако все вышеизложенное не снимает вопрос о возможном возникновении проблем, вызванных необходимостью использования в селекционных программах форм с генотипом *Ne1Ne1Ne2ne2*.

Авторы выражают искреннюю благодарность доктору биологических наук Т.И. Одинцовой и М.П. Слезиной за помощь при оформлении данной рукописи в печать.

Работа выполнена при бюджетном финансировании.

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

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

Авторы заявляют, что у них нет конфликтов интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Pukhalskiy V.A. Data on hybrid necrosis genes in the genus *Triticum* L. // Russian J. Genetics. 1996. V. 32(4). P. 469–471.
2. Tsunewaki K. Geographical distribution of necrosis genes in common wheat. I. Genetic basis of necrosis // Annu. Rep. Nat. Inst. Genet. Jap. 1964. V. 15. P. 74–79.
3. Tsunewaki K., Nishikava K.N. Frequencies of *Ne1* and *Ne2* genes in emmer and common wheats // Annu. Rep. Nat. Inst. Genet. Jap. 1963. V. 14. P. 64–69.
4. Zeven A.C. Geographical distribution of genes of hybrid necrosis in wheat // Euphytica. 1966. V. 15(3). P. 281–284.
5. Zeven A.C. Eighth supplementary list of wheat varieties classified hybrid necrosis // Euphytica. 1981. V. 30(1). P. 521–539.
6. Пухальский В.А., Мартынов С.П., Добротворская Т.В. Гены гибридного некроза пшениц (Теория вопроса и каталог носителей летальных генов). Москва: Изд-во ТСХА, 2002. 315 с.
7. Hermesen J.G.T. Quantitative investigations on progressive necrosis in wheat hybrids // Euphytica. 1960. V. 9(1). P. 141–172.
8. Hermesen J.G.T. The genetic basis of hybrid necrosis in wheat // Genetica (Netherlands). 1963. V. 12. P. 245–287.
9. Hermesen J.G.T. Hybrid necrosis as a problem for the wheat breeder // Euphytica. 1963. V. 12. P. 1–16.
10. Tsunewaki K. Monosomic and conventional gene analyses in common wheat. III. Lethality // Jap. J. Genet. 1960. V. 35. № 4. P. 71–75.
11. Nishikava K. Mapping of necrosis genes *Ne1* and *Ne2* // EWAC Newslett. 1974. № 4. P. 73–74.
12. Chu C.-G., Faris J.D., Friesen T.L., Xu S.S. Molecular mapping of hybrid necrosis genes *Ne1* and *Ne2* in hexaploid wheat using microsatellite markers // Theor. Appl. Genet. 2006. V. 112. P. 1374–1381.

13. Si Y., Zheng S., Niu J. et al. Fine mapping of hybrid necrosis gene *Ne1* in common wheat (*Triticum aestivum* L.) // Theor. Appl. Genet. 2021. V. 134. № 8. P. 2603–2611.
14. Zhang M., Luv S., Wang S. et al. Fine mapping and distribution analysis of hybrid necrosis genes *Ne1* and *Ne2* in wheat in China // Theor. Appl. Genet. 2022. 135(4). P. 1177–1189.
15. Мартынов С.П., Доброворская Т.В., Крупнов В.А. Анализ распространения генетического материала *Triticum timopheevi* Zhuk. в сортах мягкой пшеницы (*Triticum aestivum* L.) // Генетика. 2018. Т. 54. № 2. С. 177–186.
<https://doi.org/10.7868/S0016675818020121>
16. Gill K.S., Ghai B., Gupta M. Genetic and development analysis of hybrid necrosis in wheat // Ind. J. Genet. Plant Breeding. 1972. V. 32(1). P. 12–17.
17. Петросян Дж.А., Бекназарян Л.Г. Распределение генов некроза у низкостебельных сортов *Triticum aestivum* L. Пятнадцатый список летальных генов // Тр. ин-та земледелия Арм. ССР. Серия Пшеница. 1974. № 2. С. 21–24.
18. McIntosh R.A., Hart G.E., Gale M.D. Catalogue of gene symbols for wheat. Suppl. // Cer. Res. Comm. 1990. V. 18(1–2). P. 141–157.
19. Samborski D.Y., Dyck P.L. Enhancement of resistance to *Puccinia recondite* by interactions of resistance genes in wheat // Can. J. Plant Pathol. 1982. V. 4. P. 152–156.
20. Михайлова Л.А. Устойчивость пшеницы к бурой ржавчине // Идентифицированный генофонд растений и селекция. Санкт-Петербург: Наука, 2005. С. 513–527.
21. Roelfs A.P. Resistance to leaf and stem rusts in wheat // Breeding Strategies for Resistance to Rust of Wheat. Mexico: Int. Maize and Wheat Improvement Center, 1988. P. 10–19.
22. Kolmer J.A. Genetics of resistance to wheat leaf rust // Annu. Rev. Phytopath. 1996. V. 34. P. 435–455.
23. Pukhalskiy V.A., Martynov S.P., Bilinskaya E.N. Distribution of hybrid necrosis genes in common wheat cultivars of Australia // Ann. Wheat Newsletter. 2003. V. 14. P. 125–127.
24. Pukhalskiy V.A., Martynov S.P., Dobrovorskaya T.V. Analysis of geographical and breeding-related distribution of hybrid necrosis in bread wheat (*Triticum aestivum* L.) // Euphytica. 2000. V. 114. P. 233–240.
25. Рункина Л.А. О возможной связи генов гибридного некроза с генами, контролирующими другие селекционно ценные признаки // Бюлл. ВИР. 1991. Вып. 210. С. 24–26.
26. Мережко А.Ф., Рункина Л.А. Изогенные линии яровой мягкой пшеницы по генам гибридного некроза // Сб. тр. ИГиГ СО РАН “Генетические основы признаков продуктивности растений”. Новосибирск: Наука, 1992. С. 119–121.

Distribution and Allele Strength of Hybrid Necrosis Genes in the Genotypes of Canadian Cultivars of Spring Bread Wheat (*Triticum aestivum* L.)

V. A. Pukhalskiy^{a,*}, V. S. Rubets^b, E. N. Bilinskaya^a, and A. M. Kudryavtsev^a

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

^bTimiryazev Russian State Agrarian University, Moscow, 127434 Russia

*e-mail: pukhalsk@vigg.ru

The occurrence of two types of cultivars of spring bread wheat of Canada that differ in their necrotic genotypes was shown. The cultivars of one type had the *ne1ne1Ne2Ne2* genotype, and the cultivars of the second type had the *ne1ne1ne2ne2* genotype. The *Ne1Ne1ne2ne2* genotype was not found. The *Ne2* genes were represented by medium strength, *ms*, and strong alleles, *s*. By pedigree analysis, the donors of the *Ne2* gene were identified.

Keywords: hybrid necrosis genes, complementary gene interaction, allele strength, spring bread wheat, *Triticum aestivum* L.

ПРИЧИНЫ МОРФОЛОГИЧЕСКОЙ И ГЕНЕТИЧЕСКОЙ
НЕОДНОРОДНОСТИ СИГОВ *Coregonus lavaretus sensu lato*
АРКТИЧЕСКОЙ ЧАСТИ БАССЕЙНА р. ЛЕНЫ

© 2023 г. Н. А. Бочкарев¹, *, Д. С. Сендек², Е. И. Зуйкова¹, Л. А. Пестрякова³, Л. А. Ушницкая³,
Е. С. Захаров³, А. А. Эверстова³, Н. Н. Захарова³, Д. В. Политов⁴

¹Институт систематики и экологии животных Сибирского отделения
Российской академии наук, Новосибирск, 630091 Россия

²Государственный научно-исследовательский институт озерного и речного рыбного хозяйства,
Санкт-Петербург, 199053 Россия

³Северо-Восточный федеральный университет, Якутск, 677980 Россия

⁴Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, 119991 Россия

*e-mail: nikson_1960@mail.ru

Поступила в редакцию 27.04.2023 г.

После доработки 05.06.2023 г.

Принята к публикации 06.06.2023 г.

Представлены результаты морфологического и генетического анализа популяций/форм/видов пыжьяновидных сигов из среднего и нижнего течения р. Лены и сопредельных рек. Есть основания полагать, что несмотря на значительную неоднородность сига в Сибири в водоемах сибирского Заполярья обитает небольшое число *устойчивых* форм/видов. Это восточносибирский (*C. lavaretus pidschian* n. *brachymystax*), ледниково-равнинный (*C. lavaretus pidschian* n. *glacialis*) и юкагирский сиг (*C. lavaretus pidschian* n. *jucagiricus*), а также их гибридные популяции. Ледниково-равнинный сиг и его производные отличаются от восточносибирского сига по пластическим признакам, но по меристическим они идентичны. Молекулярно-генетические исследования показали, что в большинстве ледниково-равнинный сиг формирует обширные слабосвязанные сети гаплотипов, что соответствует длительному существованию популяций в благоприятных условиях. Гаплотипы восточносибирского сига тесно связаны с гаплотипами форм/видов сигов из водоемов Южной Сибири и имеющих радиальную структуру медианной сети. Подобная структура сетей свидетельствует о недавнем сокращении численности восточносибирского сига с последующей экспансией, что напрямую связано с событиями четвертичных оледенений.

Ключевые слова: *Coregonus lavaretus*, экологическая форма, гаплотипы, формообразование, Палеарктика, Плейстоцен, р. Лена, р. Енисей, р. Попига́й, р. Омолóй.

DOI: 10.31857/S0016675823110036, **EDN:** TNYIZN

Среди сиговых рыб рода *Coregonus* в Палеарктике наиболее распространены сиги комплекса *Coregonus lavaretus sensu lato* с множеством форм и подвидов, которым авторы часто придают статус вида. В основе их дифференциации обычно рассматривают различия по числу жаберных тычинок, хотя низкая таксономическая стоимость этого признака обсуждается уже длительное время [1–4]. Соответственно таксономическая структура пыжьяновидных сигов являлась предметом дискуссий, а устоявшихся взглядов на их разнообразие до сих пор не выработано [5–11]. По результатам морфологических исследований в водоемах Европы было описано до 100 морфологически дистантных форм/видов сигов [5–7, 9, 11]. Еще 20–25 форм/видов описано из водоемов Сибири, и об-

щее число форм/видов сигов достигло значительных величин, приближаясь к числу видов эндемичных цихлид, описанных из древних южноафриканских озер [12]. К концу XX в. стала доминировать точка зрения, согласно которой на территории Евразии обитает только один вид “настоящих” сигов, в его состав было включено шесть подвидов, два из которых обитают в сибирских водоемах. Это широко распространенный сиг-пыжьян *C. lavaretus pidschian* (Gmelin, 1789) и узкоареальный баунтовский сиг *C. lavaretus baunti* или *C. baunti* Mukhomedyarov, 1948 [8, 13]. Однако такой “объединительский” подход не стал общепризнанным, и придание аллопатрическим и симпатрическим популяциям, формам сигов внутривидового или видового статуса продолжа-

ло находить поддержку. В современной научной литературе встречаются восточносибирский сиг (*C. lavaretus pidschian* n. *brachymystax* Smitt, 1886) и ледниково-равнинный сиг (*C. lavaretus pidschian* n. *glacialis* Kirillov, 1972), описанный Ф.Н. Кирилловым в водоемах Якутии. Упоминаются симпатрические сиги (*C. lavaretus pidschian* n. *taimyrensis* и *C. lavaretus pidschian* n. *logaschevi*), описанные из оз. Таймыр [14, 15]. Следует упомянуть особую форму сига из р. Оленёк (*C. lavaretus pidschian* n. *oleneki*), юкагирского сига (*C. lavaretus pidschian* n. *jucagiricus*) из р. Колыма, симпатрических сигов (*C. lavaretus pidschian* и *C. anaulorum*) из р. Анадырь. Продолжают использоваться названия сигов, описанных из водоемов Западной и Центральной Сибири: *C. lavaretus pidschian* n. *gydanus*, *C. lavaretus pidschian* n. *norilensis*, *C. lavaretus pidschian* n. *mokschevgor* [16–21].

Современный прогресс в изучении эволюционных и таксономических взаимоотношений сигов связан с использованием молекулярно-генетических методов. За относительно короткий период времени изучения изменчивость некоторых генов мтДНК в популяциях пыжьяновидных сигов из водоемов Фенноскандии и Южной Сибири [22–28]. В результате было показано, что ряд форм/видов пыжьяновидных сигов, обитающих на окраинах ареала, генетически отличаются от типичного сига-пыжьяна из р. Сось (бассейн р. Обь), другие формы/виды отличаются незначительно или вообще не отличаются от типичных сигов по структуре митохондриальных генов [26, 28]. Изучение генетической изменчивости многочисленных форм сигов из верхнего течения р. Лены (Баунтовская система озер) показало, что их современное разнообразие в большей степени обусловлено гибридизацией между несколькими филогенетическими линиями, произошедшей в связи с изменением гидрологической сети водоемов на фоне неравномерного поднятия и опускания суши в периоды активного горнообразования [24, 28]. На основании этих выводов было сделано предположение о том, что в основе всех баунтовских форм сигов лежит форма, близкая к реликтовому “ряпушковидному” сигу (в некоторых источниках описан как *C. sardinella baunti* или *C. karasjovi*), гибридизация которого с дистантными филогенетическими линиями пыжьяновидных сигов привела к образованию “букета” форм/видов. Аналогичные процессы формообразования были отмечены в р. Оленёк, где тоже была обнаружена гибридизация между двумя арктическими формами/видами сигов [27, 29].

Цель нашей работы — анализ филогенетических и филогеографических связей между современными формами/видами сигов из водоемов сибирской Арктики и в первую очередь из среднего и нижнего течения р. Лены. Привлечение дополнительных морфологических и генетиче-

ских данных по другим формам/видам пыжьяновидных сигов из водоемов Сибири позволило выявить некоторые аспекты их эволюционной истории в постледниковый период.

МАТЕРИАЛЫ И МЕТОДЫ

Характеристика района исследований. В работе были использованы выборки сигов комплекса *C. lavaretus* sensu lato из бассейна р. Лены, а именно: 1) из р. Марха (включая приток р. Моркока, бассейн р. Вилюй); 2) из устья р. Буотама (правый приток р. Лены); 3) из р. Нюя (левый приток р. Лены); 4) из р. Олекма (правый приток р. Лены). Дополнительно в анализ включены морфологические и генетические данные по сигам из р. Амга и оз. Б. Токо (бассейн р. Алдан) (рис. 1), озер Кутарамакан и Хантайское (бассейн р. Курейка), оз. Каракуль и р. Б. Абакан (бассейн р. Абакан), рек Хатанга, Попигай, Анабар, Оленёк, Омолой, Хрома и Яна (ледовитоморский бассейн в пределах сибирской части Субарктики). Полный список использованного материала приведен в табл. 1. В качестве внешней группы в работу были включены гаплотипы юкагирского сига *C. l. pidschian* n. *jucagiricus* из оз. Илирней (бассейн р. Малый Анюй, бассейн р. Колымы) и оз. Ичилях (бассейн р. Индигирки) [30].

Сбор материала и морфологический анализ. Лов большинства форм/видов сигов в реках и озерах проводили в летний период. В реках Марха, Моркока, устье р. Буотама, Вилюй, Оленёк, в озерах Б. Токо, Курейка, Хантайское, Кутарамакан, Ичилях, Илирней для лова применяли жаберные сети с ячейей 30–45 мм, при этом для проведения дальнейшего морфологического анализа рыб фотографировали сразу после вылова. Сигов из рек Яна, Хрома, Омолой, Попигай, Оленёкская и Быковская протоки дельты р. Лены получали в замороженном виде и обрабатывали в лабораторных условиях. Рыбу сразу после лова (разморозки) с расправленными плавниками выкладывали на пенополиуретановый коврик и фотографировали с помощью цифрового фотоаппарата Nikon D500. Все измерения рыб проводил один оператор по фотографиям, согласно стандартной схеме в программе AxioVision v 3.1 (Carl Zeiss Vision GmbH) [31, 32]. В работе использовали 33 пластических признака [26]. Подсчет жаберных тычинок проводили одновременно с фотографированием. Число прободенных чешуй в боковой линии подсчитывали по фотографиям. Для выявления различий по числу прободенных чешуй в боковой линии и жаберных тычинок на первой жаберной дуге использовали однофакторный дисперсионный анализ для неравных по размеру выборок (критерий Тьюки). Поскольку в случае аллометрического роста форма тела зависит от размера, перед проведением анализа по пластическим

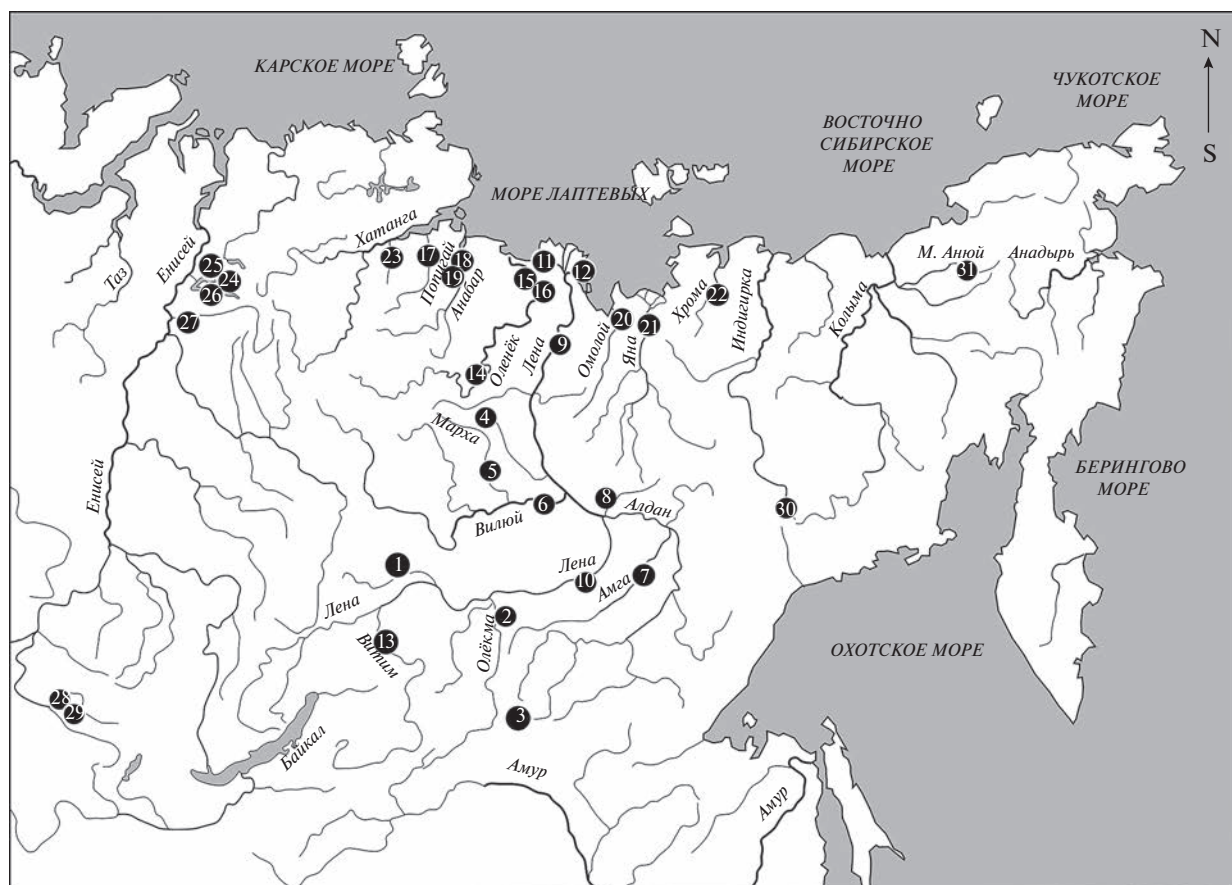


Рис. 1. Карта-схема сбора ихтиологического материала 1–31 (точки сбора соответствуют номерам в табл. 1).

признакам все данные были скорректированы с учетом размеров особей с последующим их логарифмированием [33, 34]. Кластеризацию средних популяционных величин по имеющимся 19 пригодным для анализа выборкам проводили методом UPGMA. Оценку достоверности различий между выборками проводили с помощью многомерного дисперсионного анализа MANOVA в программе PAST v 4.3 [35].

Анализ мтДНК. Геномную ДНК выделяли фенольно-хлороформным методом из фиксированной 96%-ным этанолом печени сига. Амплификацию митохондриального гена *ND1* проводили с использованием ранее разработанных праймеров [27]. Полученные продукты очищали с помощью набора реактивов “БИОСИЛИКА” (Новосибирск, Россия) и секвенировали как в прямом, так и в обратном направлении в компании “Синтол” (Москва, Россия, www.syntol.ru). В результате получены последовательности длиной в 1091 нуклеотидных оснований. Последовательности выравнивали с помощью алгоритма ClustalW, редактировали вручную и депонировали в базу данных GenBank (табл. 1).

Для анализа генетического полиморфизма вычисляли: число полиморфных (сегрегирующих) сайтов (S), число гаплотипов (h), гаплотипическое разнообразие (H_d), нуклеотидное разнообразие (π), среднее число нуклеотидных различий (k), выполняли в программе DnaSP v 5.10 [35]. Оценку степени межпопуляционной дифференциации F_{ST} и числа мигрантов N_m между популяциями проводили в программе ARLEQUIN v 3.5 [36, 37]. Для оценки нейтральности эволюции были использованы тесты Таджимы ($Tajima's, D$) и Фу ($Fu's, F_s$) [38, 39]. Генеалогические связи между гаплотипами предполагаемых форм/видов сига установлены на основе медианной сети (алгоритм Median-Joining, MJ) в программе Network v 4.5 [38]. Географическое распределение гаплотипов реконструировано по результатам анализа медианной сети [40]. Для анализа гипотезы популяционной экспансии оценивали среднестатистическое отклонение (SSD) и индекс Харпендинга (r) [37, 41]. Достоверность модели оценивалась при сравнении сумм квадратов отклонений (SSD) между наблюдаемым распределением и распределением, полученным для каждого итерационного набора данных.

Таблица 1. Места и координаты лова рыб, вид/форма, объем выборок, номера нуклеотидных последовательностей мтДНК в международной базе данных GenBank (NCBI)

Место лова, вид/форма	Объем выборки	Широта	Долгота	Номера доступа в GenBank
Бассейн р. Лены				
1. р. Нюя (приток р. Лены), в/с	3	60°31′	116°18′	OQ606836–OQ606838
2. р. Олёкма (приток р. Лены), в/с	7	53°44′	117°20′	OQ606839–OQ606844
3. оз. Б. Токо (бассейн р. Алдан), в/с	11	56°07′	130°82′	KM013410–KM013417, MT862534
4. р. Моркока (приток р. Марха), в/с	7	65°10′	115°51′	OQ606845–OQ606848
5. р. Марха (приток р. Вилюй), в/с	9	65°06′	116°42′	KM013405–KM013409, MT862525
6. р. Вилюй, в/с	5	64°22′	126°24′	OQ606849–OQ606852
7. р. Амга (приток р. Алдан), в/с	15	62°37′	134°55′	OQ606853–OQ606857
8. р. Тумара (приток р. Алдан), в/с	3	64°53′	130°27′	OQ606858, OQ606859
9. р. Лена (с. Кюсюр), в/с	5	70°41′	127°21′	OQ606860–OQ606862
10. р. Лена (устье р. Буотама), в/с	10	61°14′	128°34′	MT862527–MT862533
11. Оленёкская протока (р. Лена), л/р	17	73°11′	122°14′	MT885382–MT885390
12. Быковская протока (р. Лена), л/р	—	72°34′	127°74′	—
13. р. Витим, в/с	11	57° 07′	116°42′	OK018185–OK018189
Реки арктического бассейна Сибири				
14. р. Оленёк (среднее течение), в/с	18	67°54′	105°16′	MT885362–MT885376
15. р. Оленёк (с. Таймылыр), в/с	11	72°37′	121°53′	MT885346, MT885348–MT885353
16. р. Оленёк (с. Таймылыр), л/р	12	72°37′	121°53′	MT885347–MT885361
17. р. Попигай, л/р	8	72°57′	106°09′	MT885377–MT885381
18. р. Анабар, в/с	17	70°46′	113°10′	KU948937–KU948951
19. р. Анабар, л/р	20	70°46′	113°10′	KU948952–KU948962
20. р. Омолой, л/р	6	71°13′	131°59′	MT862514–MT862516, MT862536
21. р. Яна, л/р	3	70°79′	136°21′	MT862505–MT862507
22. р. Хрома, л/р	9	70°84′	143°59′	MT862508–MT862513, MT862487, MT862498
Бассейн р. Хатанги				
23. р. Рассоха, л/р	5	72°07′	101°08′	KU948963–KU948966
Бассейн р. Енисей				
24. оз. Кутарамакан, в/с	20	68°42′	91°43′	MK777973–MK777976, MK777978
25. оз. Кутарамакан, л/р	15	68°42′	91°43′	MK777971–MK777973, MK777977
26. оз. Хантайское, ?	4	68°24′	91°28′	KU948967–KU948970
27. р. Курейка, ?	7	66°57′	88°15′	KU948985–KU948988
28. р. Б. Абакан, ?	24	52°00′	87°65′	KJ742910–KJ742918, HM538404, HM538405, JN628999, JN629000
29. оз. Каракуль, ? (бассейн р. Абакан)	11	52°10′	87°65′	HM538401–HM538403, KJ742909
Бассейны рек Колымы и Индигирки				
30. оз. Илirianей, Ju (Колыма)	9	67°36′	168°34′	MT862482–MT862484, MT862479, MT862493, MT862494, MT862496
31. оз. Ичилях, Ju (Индигирка)	10	62°36′3	142°46′	MT862483, MT862484, MT862493, MT862494

Примечание. в/с — восточносибирский сиг, л/р — ледниково-равнинный сиг, определенные по внешним морфологическим признакам; Ju — юкагирский сиг, ? — форма не установлена.

РЕЗУЛЬТАТЫ

Морфология. Все изученные формы/виды из водоемов сибирской Арктики, как и группировки из среднего и нижнего течения р. Лены, относятся к малочешуйчатым ($lim = 77-85$) и малотычинковым ($lim = 17-25$) сигам (табл. 2). Попарное сравнение числа *прободенных чешуй* в боковой линии по критерию Тьюки в большинстве случаев не выявило достоверных различий (4.5%). Только восточносибирский сиг из устья р. Курейка ($ll = 84.16 \pm 0.43$) и восточносибирский сиг из оз. Кутарамакан ($ll = 85.86 \pm 0.65$) демонстрируют достоверные отличия от других популяций. Различия по числу жаберных тычинок встречаются значительно чаще (17.5%). Сравнения между объединенными выборками **восточносибирских** и **ледниково-равнинных** сигов (диагностированные визуально) выявили достоверные различия **по обоим признакам** ($p = 0.002$ и $p = 0.023$ соответственно). Многомерный дисперсионный анализ (MANOVA) по **пластическим** признакам показал, что большинство исследованных выборок сигов хорошо дифференцированы (Wilk's $\Lambda = 0.000092$, $F = 7.844$; $d.f.1 = 576$, $d.f.2 = 5599$) и достоверно отличаются друг от друга ($p < 0.001$). Сравнения, проведенные между выборками визуально определенных восточносибирских и ледниково-равнинных сигов, тоже выявили достоверные различия (Wilk's $\Lambda = 0.3707$, $F = 14.64$; $d.f.1 = 32$, $d.f.2 = 276$, $p = 7.697E-43$). Голова восточносибирского речного сига имеет простую конусообразную форму, плавно переходящую в туловище, тогда как голова ледниково-равнинного сига имеет более сложную форму и в затылочной части переходит в горб той или иной степени выраженности (рис. 2). Дендрограмма сходства, построенная на основе пластических признаков, демонстрирует структурированность массива (с высокой степенью поддержки) на два кластера (I и II), которые соответствуют дифференциации сигов по форме тела, т.е. на восточносибирских “востряков” и ледниково-равнинных “горбунов” (рис. 3). Кластер I объединяет выборки восточносибирских сигов из рек Оленёк, Лена (с. Кюсюр, устье р. Бутама), притоков Лены — реки Амга, Тумара, Марха, Витим, из оз. Б. Токо, а также р. Анабар, нижнего течения р. Оленёк (с. Таймылыр) и оз. Кутарамакан 1. Второй кластер сформирован выборками озерных сигов (ледниково-равнинных “горбунов”) из бассейнов рек Омолой, Анабар, Яна, Рассоха, из Оленёкской и Быковской протоки дельты р. Лены и ледниково-равнинных сигов из оз. Кутарамакан 2. В основе кластера II расположен близкий по форме тела к ледниково-равнинным сигам *юкагирский сиг* из оз. Илirianей. Исключением стал сиг из нижнего течения р. Оленёк (с. Таймылыр), первоначально определенный как ледниково-равнинный, но по резуль-

татам данного анализа объединившийся с выборками восточносибирских сигов (кластер I).

Полиморфизм мтДНК. Большинство популяций сигов из бассейна р. Лена характеризуются высоким уровнем генетического полиморфизма митохондриального гена *ND1* (табл. 3). В бассейне р. Лены максимальные значения показателей зарегистрированы для популяции сигов из Оленёкской протоки дельты р. Лены ($H_d = 0.985$; $\pi = 0.004$; $k = 4.221$), а минимальные — для популяции сигов из р. Амга ($H_d = 0.629$; $\pi = 0.0013$; $k = 1.238$). За пределами Ленского бассейна наиболее низкие показатели генетической изменчивости выявлены в популяции сигов из оз. Каракуль ($H_d = 0.182$; $\pi = 0.0002$; $k = 0.182$) бассейна р. Енисей. Соседняя популяция из р. Б. Абакан, напротив, имеет очень высокие генетические показатели ($H_d = 0.906$; $\pi = 0.0032$; $k = 3.152$). Выборки юкагирского сига из бассейна рек Колымы и Индигирки (ограничивающие Ленский бассейн с востока) тоже имеют низкие показатели ($H_d = 0.222$; $\pi = 0.0007$; $k = 0.667$ и $H_d = 0.378$; $\pi = 0.0006$; $k = 0.600$ соответственно), сравнимые с показателями сига из оз. Каракуль (табл. 3).

Реконструкция генеалогических взаимоотношений между гаплотипами сигов из некоторых арктических водоемов Сибири, среднего и нижнего течения р. Лены и бассейна р. Енисей выявила структурирование медианной сети на четыре большие гаплогруппы (I–IV) (рис. 4). В основе гаплогруппы I лежит звездообразная структура (H26), сформированная гаплотипами сигов из р. Енисей, из р. Моркока, оз. Лангтибейто и р. Оленёк. В гаплогруппу I входит также множество уникальных, удаленных друг от друга гаплотипов, связанных с гаплогруппой II циклическими связями.

Гаплогруппа II связана с первой через три мутационных шага, один гипотетический и один уникальный гаплотип, принадлежащий ледниково-равнинному сигу. Основу второй звездообразной структуры (H16) с множеством минорных гаплотипов первого, второго и третьего порядка составляют гаплотипы сигов притоков р. Лены, Оленёкской протоки, оз. Кутарамакан и р. Курейка. Несколько небольших звездообразных структур, состоящих из гаплотипов сигов Оленёкской протоки и р. Анабар (H10, H84), тесно связаны с главной звездообразной структурой, с центральным гаплотипом H16. Кроме того, в гаплогруппу входят несколько слабо связанных с центральным гаплотипом структур (H14, 17, 48, 59, 96, 129) и множество уникальных гаплотипов. Третья гаплогруппа с центральным гаплотипом H108 связана со второй гаплогруппой через четыре мутационных шага, два гипотетических гаплотипа и один гаплотип ледниково-равнинного сига из оз. Кутарамакан. Гаплогруппа состоит из одной звездообразной структуры (H108), в состав которой входят

Таблица 2. Места лова, вид/форма, объем выборок, число прободенных чешуй в боковой линии (*ll*) и число жаберных тычинок (*Sp.br.*)

Место лова, вид, форма	Объем выборки		<i>ll</i>	<i>Sp.br.</i>
	по пластическим признакам	по меристическим признакам		
1. р. Нюя (приток р. Лены), в/с	—	—	—	—
2. р. Олёкма (приток р. Лена), в/с	—	—	—	—
3. оз. Б. Токо (бассейн р. Алдан), в/с	15	15	81.07 ± 0.64	21.20 ± 0.60
4. р. Моркока (приток р. Марха), в/с	—	—	—	—
5. р. Марха (приток р. Виллюй), в/с	10	18	80.89 ± 0.86	19.38 ± 0.37
6. р. Виллюй, в/с	—	—	—	—
7. р. Амга (приток р. Алдан), в/с	19	19	80.42 ± 0.58	20.05 ± 0.41
8. р. Тумара (приток р. Алдан), в/с	6	6	77.33 ± 0.76	19.66 ± 0.21
9. р. Лена (с. Кюсюр), в/с	12	12	79.17 ± 0.60	19.83 ± 0.41
10. р. Лена (устье р. Буотамы), в/с	26	26	80.96 ± 0.78	19.28 ± 0.22
11. р. Лена (Оленёкская протока), л/р	12	12	80.75 ± 0.73	22.08 ± 0.50
12. р. Лена (Быковская протока), л/р	10	8	82.00 ± 0.27	21.00 ± 0.74
13. р. Витим, в/с	25	25	81.12 ± 0.60	21.80 ± 0.45
14. р. Оленёк среднее течение, в/с	69	66	82.91 ± 0.42	22.98 ± 0.16
15. р. Оленёк (с. Таймылыр), в/с	—	5	81.60 ± 0.45	19.46 ± 0.40
16. р. Оленёк (с. Таймылыр), л/р	9	15	82.20 ± 0.37	21.60 ± 0.87
17. р. Попигай, л/р	—	8	81.25 ± 1.11	20.87 ± 0.85
18. р. Анабар, в/с	56	56	80.65 ± 0.71	20.08 ± 0.31
19. р. Анабар, л/р	25	23	82.14 ± 0.44	19.17 ± 0.19
20. р. Омолой, л/р	5	12	79.83 ± 0.51	19.67 ± 0.50
21. р. Яна, л/р	11	11	79.00 ± 0.74	19.54 ± 0.55
22. р. Хрома, л/р	—	9	81.11 ± 0.73	20.55 ± 0.47
23. р. Рассоха, л/р	4	8	79.75 ± 1.01	19.00 ± 0.60
24. оз. Кутарамакан, л/р	20	29	82.10 ± 0.63	22.32 ± 0.23
25. оз. Кутарамакан в/с	40	40	85.86 ± 0.65	21.93 ± 0.28
26. оз. Хантайское, ?	—	4	82.16 ± 0.68	21.37 ± 0.32
27. р. Курейка, ?	—	32	84.16 ± 0.43	—
28. р. Абакан, ?	—	26	82.83 ± 0.57	21.67 ± 0.21
29. р. Каракуль, ?	—	137	82.84 ± 0.30	25.34 ± 0.11
30. оз. Илirianей, Ju	31	22	78.82 ± 1.74	20.18 ± 0.48
31. оз. Ичилях, Ju	—	20	81.25 ± 0.53	20.80 ± 0.47

несколько уникальных гаплотипов, удаленных от центрального на разное число замен. Все гаплотипы (за исключением МТ86248 из р. Хрома) принадлежат сигам из оз. Илirianей бассейна р. Колымы и оз. Ичилях бассейна р. Индигирки. Четвертая гаплогруппа (IV) связана со второй через множество мутаций и гипотетических гаплотипов и представляет собой сеть слабо связанных друг с другом уникальных гаплотипов (вероятно, ледниково-равнинных сигов) из р. Анабар и дельты р. Лены (KU948941, KU948948, KU948949,

KU948952, МТ885363, МТ885366, МТ885375, МТ885376).

Анализ молекулярного разнообразия (AMOVA) показал, что при условном разделении популяций сигов из р. Лены на русловые группировки и сохранении остальных выборок без изменения значимых различий по генетическому разнообразию **между группами** не обнаруживается (19.30% от разнообразия), а внутривидовые различия составляют 73.26%. При любых других вариантах анализа степень межгруппового разнообразия только уменьшается до 8%, при этом растет

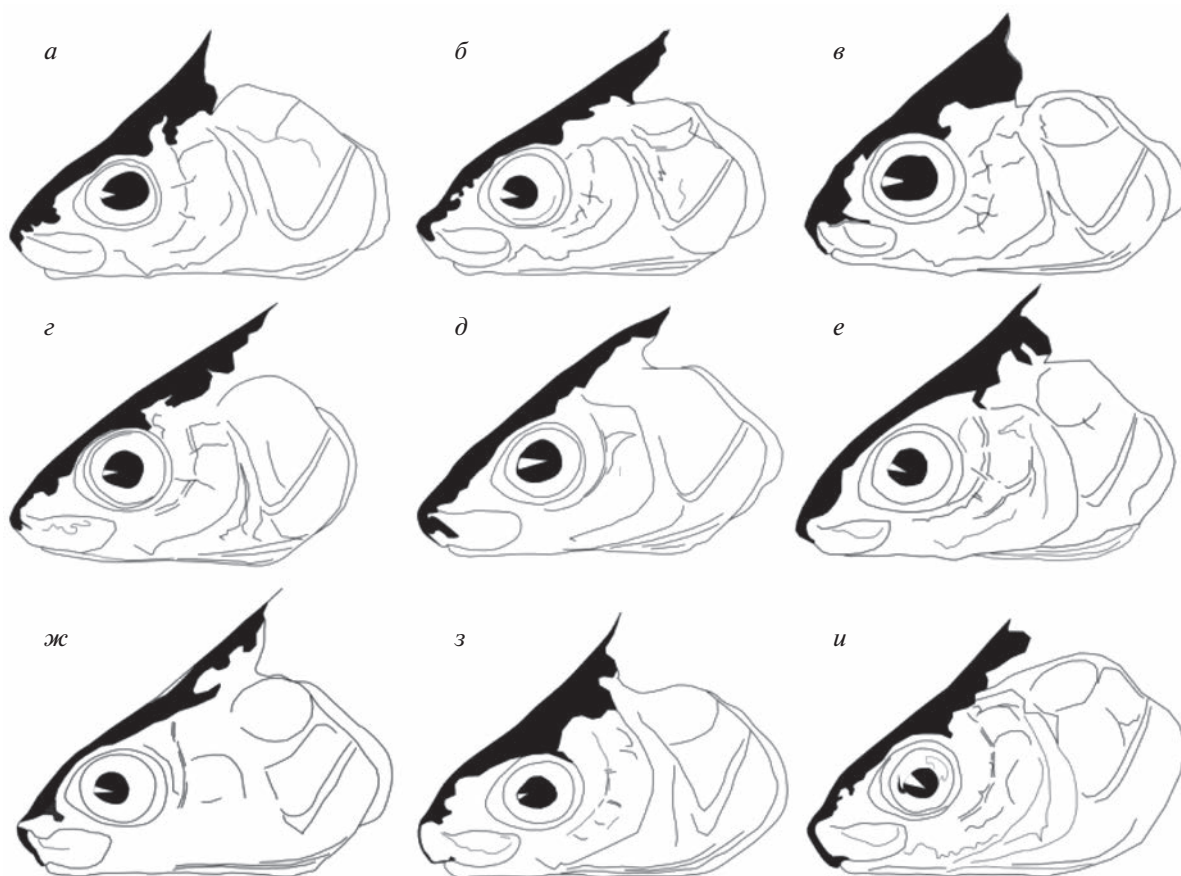


Рис. 2. Абрис голов некоторых восточносибирских (*в, з, д, е*) и ледниково-равнинных сигов (*а, б, ж, з*). *а* — Оленёкская протока р. Лены; *б* — Быковская протока р. Лены; *в* — с. Кюсюр, нижнее течение р. Лены; *з* — р. Амга, приток р. Алдан; *д* — р. Оленёк, среднее течение; *е* — с. Таймылыр, нижнее течение р. Оленёк; *ж* — р. Яна; *з* — р. Хрома; *и* — юкагирский сиг из оз. Илirianей (бассейн р. Колымы).

процент разнообразия между популяциями внутри групп (табл. 4).

Тест на нейтральность эволюции. Тесты Таджи-мы D и F_s для большинства исследованных популяций принимают отрицательные, и для четверти выборок достоверные значения (28.57%). Положительные недостоверные значения выявили для выборок сигов из нижнего течения р. Лены с. Кюсюр, р. Рассоха бассейна р. Хатанга и рек Омолдой, Хрома. По всей видимости, это обусловлено в выборках небольшим числом гаплотипов. После их объединения в одну группу получили отрицательные и достоверные величины ($D = -0.923$, $p = 0.178$; $F_s = -6.359$, $p = 0.005$). Объединенная выборка сигов из бассейна р. Лены и выборка ледниково-равнинных сигов (определенных по морфологическим признакам) характеризуются отрицательными и достоверными значениями обоих тестов. Тесты D и F_s для совокупной выборки сигов из бассейна р. Енисей тоже имели отрицательные достоверные значения (табл. 5).

Генетическая дифференциация. Оценка подразделенности полиморфизма мтДНК с помощью критерия F_{ST} подтвердила некоторую степень дифференциации между выборками (рис. 5). Анализ выявил низкие (редко средние и высокие) значения индекса F_{ST} (от 0.005–0.589) в пределах Ленского бассейна, до 0.939 при сравнении с выборками гаплотипов сигов бассейна р. Енисей. Максимально высокие показатели F_{ST} получены при сравнении выборки сигов из оз. Каракуль ($F_{ST} = 0.803–0.939$) с выборками гаплотипов сигов из притоков р. Лены (Вилуей, Марха, Амга, оз. Токо и др.). В своем большинстве оценки F_{ST} достоверны. Недостоверные оценки связаны с небольшими по размеру выборками. Оценка генетической дифференциации между популяциями сигов, выделенных по бассейновому признаку (кроме бассейна р. Енисей), выявила низкие статистически значимые показатели (0.06–0.20). Минимально существенные оценки F_{ST} наблюдаются между объединенными выборками гаплотипов сигов из бассейна р. Енисей и выборками из

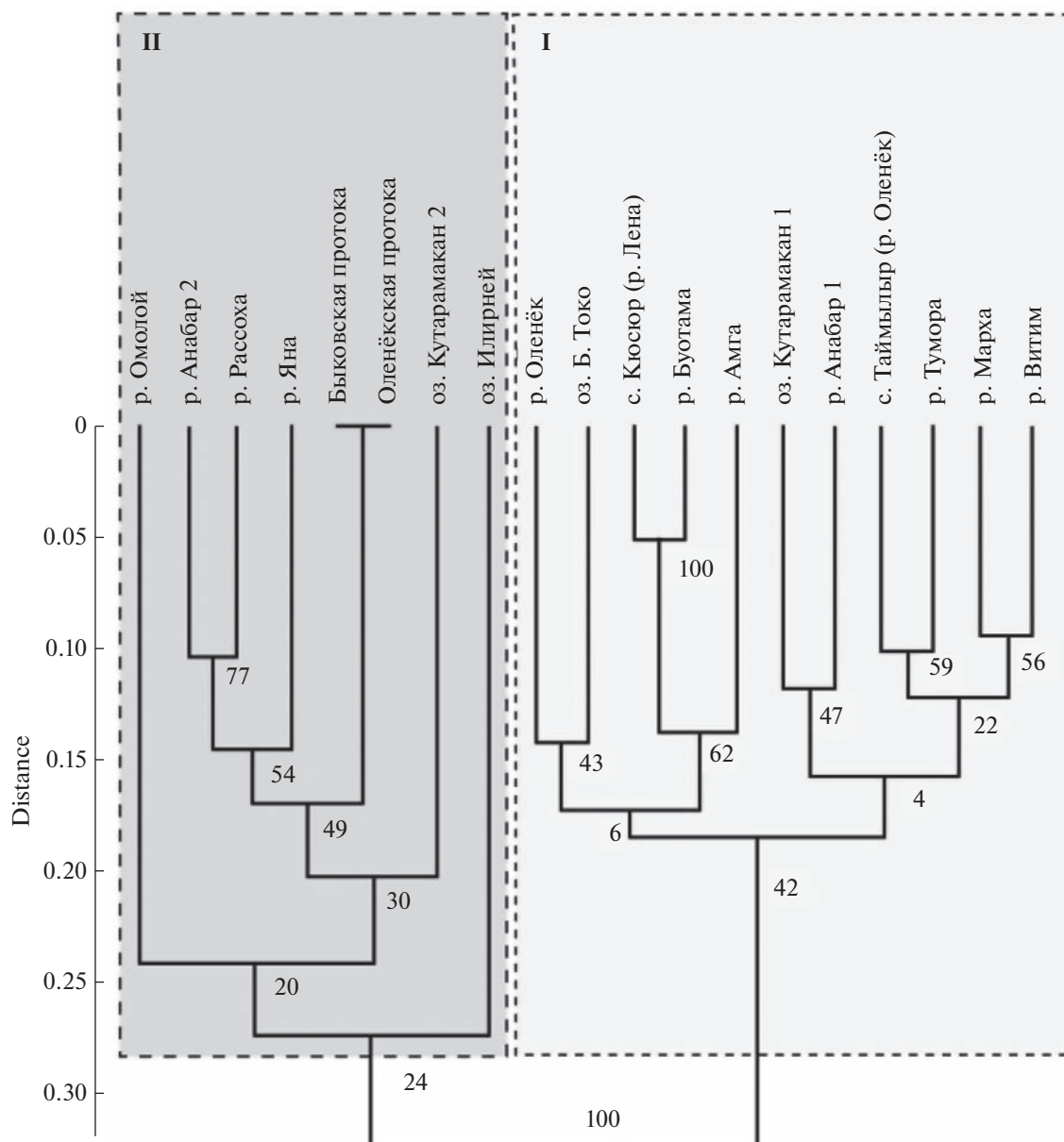


Рис. 3. UPGMA-дендрограмма выборок сигов из среднего и нижнего течения р. Лены и некоторых рек арктического бассейна Сибири по средневыворочным данным пластических признаков. I – ледниково-равнинный сиг; II – восточносибирский сиг. В узлах ветвления указано значение бутстреп-поддержки при числе репликаций 100000.

рек Яна, Хрома, Омолой ($0.2048, p < 0.001$). Среднее число мигрантов (N_m) в бассейне р. Лены за одно поколение сильно изменяется – от 0.350 до 71.360. Самые низкие числа мигрантов за одно поколение наблюдаются между популяциями из основных бассейнов рек (0.098–0.950). Внутри бассейна р. Оленёк (среднее и нижнее течение) $N_m = 71.36$.

Демографический анализ. Согласно значению показателей вероятности суммы квадратичного отклонения p (SSD_{obs}), вероятности индекса Хар-

пендинга p (r) модель пространственной экспансии сигов из бассейна р. Лены и рек Яна и Хрома оказалась более предпочтительной (рис. 6). Анализ распределения частот попарных нуклеотидных различий выявил унимодальный характер кривой для объединенной выборки гаплотипов сигов из среднего и нижнего течения р. Лена (табл. 6). Тогда как наличие двух пиков указывает на некоторую генетическую гетерогенность объединенных популяций сигов из рек Яна и Хрома, что позволяет предполагать вторичную интерградацию.

Таблица 3. Показатели полиморфизма по гену *ND1* митохондриальной ДНК в популяциях различных форм/видов сигов из водоемов Сибири

Место лова	<i>n</i>	<i>S</i>	<i>h</i>	<i>H_d</i>	<i>π</i>	<i>k</i>
1. р. Нюя (приток р. Лены)	3	7	3	1.000	0.005	4.667
2. р. Олёкма (приток р. Лены)	7	7	6	0.952	0.003	2.571
3. оз. Б. Токо (бассейн р. Алдан)*	11	9	7	0.881	0.002	2.036
4. р. Моркока (приток р. Марха)	7	8	5	0.857	0.003	2.952
5. р. Марха (приток р. Виллой)*	9	6	5	0.722	0.002	1.889
6. р. Виллой	5	7	4	0.900	0.003	3.200
7. р. Амга (приток р. Алдан)	15	5	4	0.629	0.001	1.238
8. р. Тумара (приток р. Алдан)	3	1	2	0.667	0.001	0.667
9. р. Лена (с. Кюсюр), в/с	5	7	3	0.700	0.004	3.800
10. р. Лена (устье р. Буотамы), в/с	10	9	7	0.933	0.003	3.089
11. Оленёкская протока (р. Лена), л/р	17	17	15	0.985	0.004	4.221
12. Быковская протока (р. Лена), л/р						
13. р. Витим*	12	7	6	0.879	0.002	2.121
14. р. Оленёк (среднее течение)*	21	32	20	0.995	0.007	7.343
15. р. Оленёк (с. Таймылыр), в/с	20	20	13	0.953	0.004	3.821
16. р. Оленёк (с. Таймылыр), л/р						
17. р. Попигай	8	11	7	0.964	0.004	4.143
18. р. Анабар, в/с	35	31	24	0.976	0.006	5.644
19. р. Анабар, л/р						
20. р. Омолой	6	7	4	0.800	0.003	2.800
21. р. Яна	3	12	3	1.000	0.008	8.000
22. р. Хрома	9	17	8	0.972	0.007	6.639
23. р. Рассоха*	5	9	4	0.900	0.005	4.600
24. оз. Кутарамакан, в/с*	31	12	8	0.733	0.003	3.187
25. оз. Кутарамакан, л/р*						
26. оз. Хантайское (бассейн р. Енисей)*	4	5	4	1.000	0.003	2.500
27. р. Курейка (бассейн р. Енисей)*	7	6	4	0.714	0.003	2.667
28. р. Б. Абакан (бассейн р. Енисей)*	24	13	12	0.906	0.0032	3.152
29. оз. Каракуль (бассейн р. Енисей)*	11	1	2	0.182	0.000	0.182
30. оз. Илирней (бассейн р. Колымы)*	9	3	2	0.222	0.001	0.667
31. оз. Ичилях (бассейн р. Индигирки)*	10	3	3	0.378	0.001	0.600
32. Бассейн р. Лены	94	47	53	0.968	0.004	3.583
33. Бассейн р. Енисей	77	27	25	0.883	0.004	3.995
34. Реки Арктики, кроме бассейнов рек Енисей, Лена	125	76	85	0.990	0.006	5.889

Примечание. *n* — число образцов, *S* — число полиморфных (сегрегирующих сайтов), *h* — число гаплотипов, *H_d* — гаплотипическое разнообразие, *π* — нуклеотидное разнообразие, *k* — среднее число нуклеотидных различий. * Данные, опубликованные ранее. № 11 — 11 + 12, № 15 — 15 + 16, № 18 — 18 + 19, № 24 — 24 + 25 — объединенные выборки (для табл. 3 и 5).

Таблица 4. Иерархический анализ молекулярной дисперсии (AMOVA) для изученных форм/видов сигов, объединенных согласно популяциям (1–23)

Уровень иерархии	<i>d.f.</i>	Процент изменчивости	Индекс фиксации (<i>p</i>)
Между группами	20	19.30	0.09219 (0.000)
Между популяциями в пределах групп	4	7.44	0.26737 (0.004)
В пределах популяций	267	73.26	0.19297 (0.002)

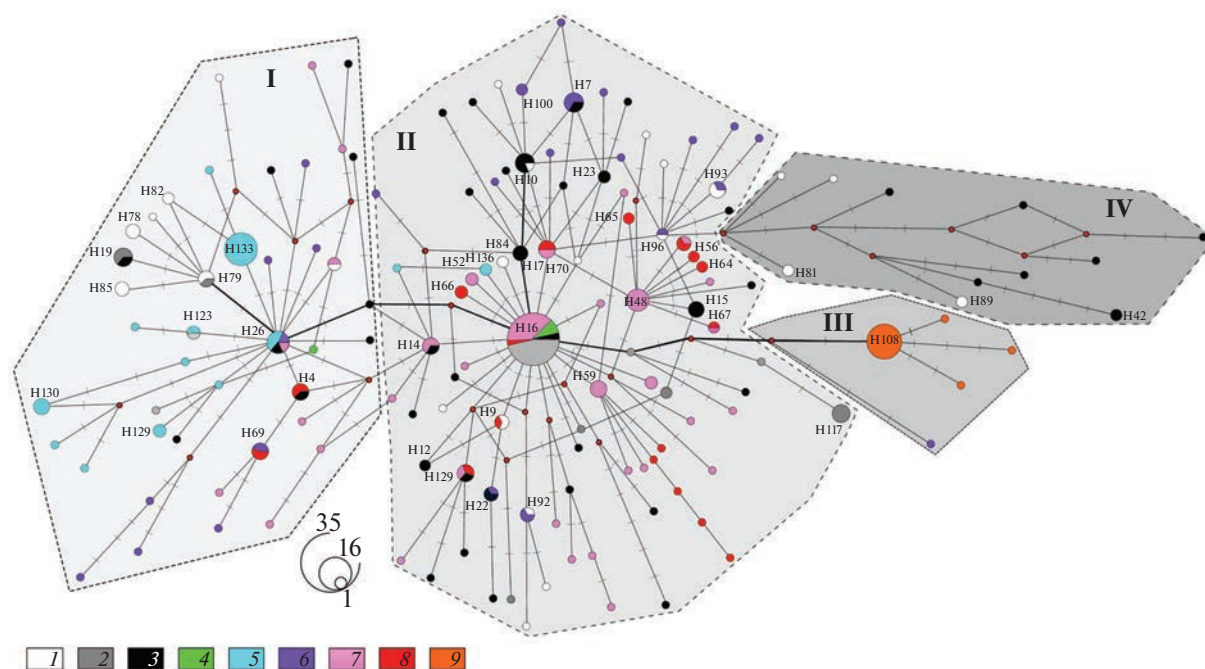


Рис. 4. Медианные сети гаплотипов (ген *ND1* мтДНК) сигов из нижнего и среднего течения р. Лены. Размер узла пропорционален частоте встречаемости гаплотипа. I–IV – гаплогруппы. 1 – р. Анабар; 2 – озера Кутарамакан, Хантайское; 3 – Оленёкская протока р. Лены, р. Оленёк; 4 – р. Курейка; 5 – р. Абакан, оз. Каракуль; 6 – реки Попигай, Омоллой, Яна, Хрома; 7 – реки Алдан, Виллой, Олекма; 8 – русло р. Лены, р. Витим; 9 – юкагирский сиг из озера Ичилях, Илирней.

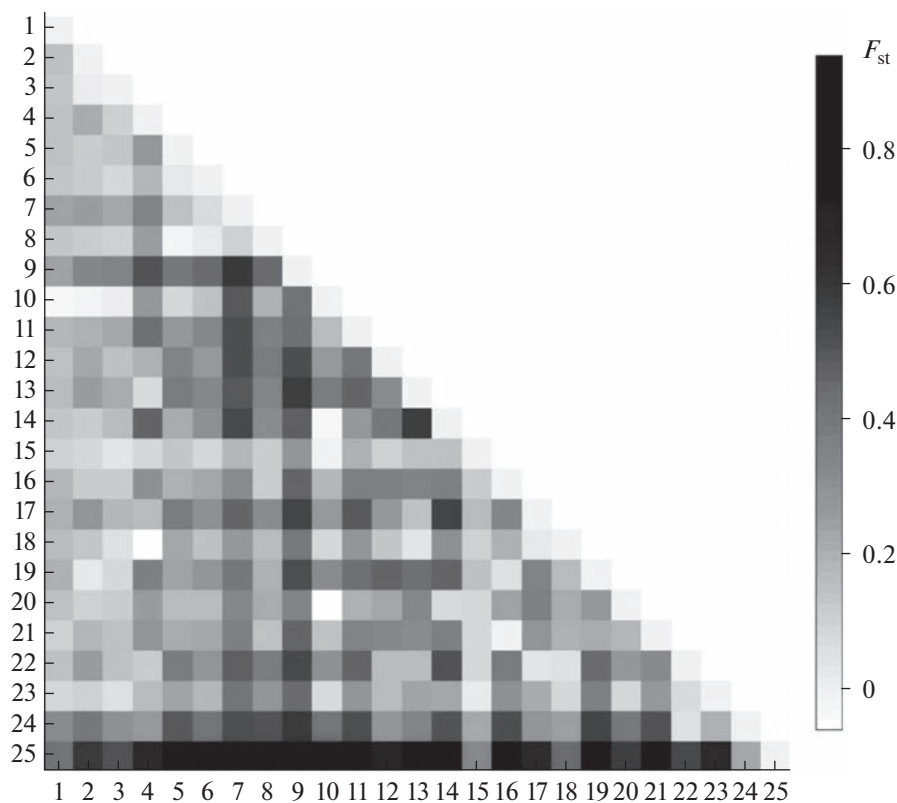


Рис. 5. Матрица F_{ST} для различных выборок гаплотипов сигов из бассейна р. Лены и сопредельных водоемов. 1 – р. Оленёк, 2 – с. Таймылыр (р. Оленёк), 3 – Оленёкская протока, 4 – с. Кюсюр (р. Лена), 5 – р. Буотама, 6 – р. Маркока, 7 – р. Марха, 8 – р. Виллой, 9 – р. Витим, 10 – р. Тумара, 11 – оз. Б. Токо, 12 – р. Олекма, 13 – р. Нюя, 14 – р. Амга, 15 – р. Анабар, 16 – р. Попигай, 17 – р. Яна, 18 – р. Хрома, 19 – р. Омоллой, 20 – оз. Кутарамакан, 21 – р. Рассоха, 22 – оз. Лангтибейто, 23 – р. Курейка, 24 – р. Абакан, 25 – оз. Каракуль.

Таблица 5. Тесты на селективную нейтральность по *NDI* мтДНК для исследованных популяций, групп популяций сигов

Место лова	Объем выборки	Тест Таджимы, (<i>D</i>)	<i>p</i>	Тест Φ_u , (F_s)	<i>p</i>
1. р. Нюя (приток р. Лены)	3	—	—	—	—
2. р. Олёкма (приток р. Лены)	7	−0.518	0.350	−2.613	0.026
3. оз. Б. Токо (бассейн р. Алдан)	11	−1.417	0.093	−2.818	0.017
4. р. Моркока (приток р. Марха)	7	−0.503	0.366	−0.737	0.234
5. р. Марха (приток р. Виллой)	9	−1.629	0.286	−0.976	0.186
6. р. Виллой	5	1.573	—	−3.304	—
7. р. Амга (приток р. Алдан)	20	−0.728	0.275	−2.198	0.087
8. р. Тумара (приток р. Алдан)	3	—	—	—	—
9. с. Кюсюр (нижнее течение р. Лены)	5	0.913	0.781	1.775	0.791
10. р. Буотама (приток р. Лены)	10	−0.127	0.468	−1.969	0.077
11. р. Лена (Оленёкская протока)	17	−0.627	0.305	−10.213	0.000
12. р. Лена (Быковская протока)					
13. р. Витим	12	−0.332	0.413	−1.177	0.182
14. р. Оленёк (среднее течение)	21	−0.679	0.273	−13.210	0.000
15. р. Оленёк (с. Таймылыр)	20	−1.224	0.104	−5.234	0.005
16. р. Оленёк (с. Таймылыр)					
17. р. Попигай	8	−0.117	0.479	−2.369	0.036
18. р. Анабар	36	−0.860	0.198	−12.247	0.000
19. р. Анабар					
20. р. Омолуй	6	−0.504	0.358	0.110	0.470
21. р. Яна	3	—	—	—	—
22. р. Хрома	9	0.299	0.679	−1.974	0.108
23. р. Рассоха	5	0.461	0.669	0.357	0.506
24. оз. Кутарамакан	31	0.198	0.643	0.350	0.591
25. оз. Кутарамакан					
26. оз. Хантайское (бассейн р. Енисей)	4	−0.798	0.168	−1.514	0.056
27. р. Курейка (бассейн р. Енисей)	7	0.452	0.643	0.426	0.573
28. р. Абакан (бассейн р. Енисей)	24	−0.328	0.399	−3.920	0.041
29. оз. Каракуль (бассейн р. Енисей)	11	−1.128	0.181	−0.410	0.122
30. оз. Илирней (бассейн р. Колымы)	9	−1.513	0.057	1.318	0.711
31. оз. Ичилях (бассейн р. Индигирки)	19	−2.046	0.002	−1.093	0.122
32. Бассейн р. Лены	94	−1.943	0.008	−26.106	0.000
33. Бассейн р. Енисей	77	−0.847	0.196	−9.592	0.004
34. Реки Арктики, кроме бассейнов рек Енисей и Лена	125	−1.993	0.004	−25.126	0.000
Для всех популяций	312	−2.253	0.010	4.898	0.020

ОБСУЖДЕНИЕ

Морфология. Все обитающие в бассейне р. Лены и в менее крупных бассейнах рек сибирской Арктики пыжьяновидные сиги относятся к малочешуйчатым и малотычинковым формам/видам. В целом наши собственные данные не отличаются от данных литературы [9, 14, 18, 24–28, 42]. Не-

сколько большее число чешуй в боковой линии в популяциях сигов из оз. Кутарамакан и устья р. Курейка позволяет сделать предположение о гибридной природе этих популяций [42]. Близкими значениями этого показателя характеризуются некоторые популяции сигов из озер Тоджинской котловины и сига-мокчегора из Пясинских

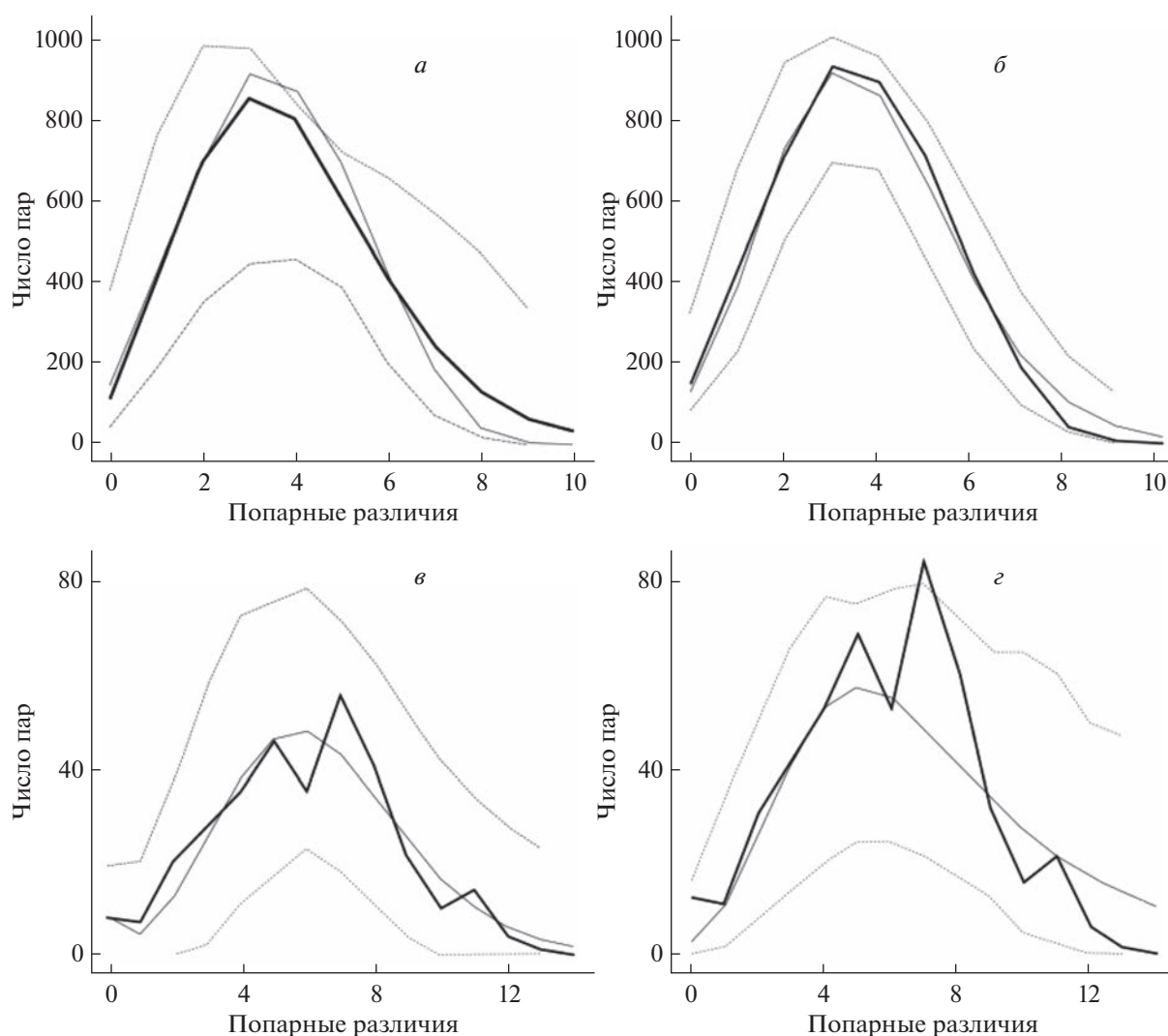


Рис. 6. График распределения числа попарных различий между нуклеотидными последовательностями в изучаемых популяциях сигов. *а, в* – *demographic expansion model*; *б, г* – *spatial expansion model*; черная линия – наблюдаемое распределение; серая линия – ожидаемое распределение; *а, б* – р. Лена с притоками; *в, г* – реки Яна и Хрома; пунктирные линии – 95%-ный доверительный интервал.

озер ($lim = 81.00–98.00$) [17, 42–45]. Несколько большее число прободенных чешуй в боковой линии зарегистрировано у озерного байкальского сига *C. lavaretus baicalensis* и сига Исаченко *C. fluviatilis*. В связи с тем, что мы уже обнаруживали обмен мтДНК между удаленными видами сигов, можно предполагать гибридизацию и вторичную интерградацию между сигом Исаченко и фоновыми популяциями сигов [46, 47]. Анализ популяций сигов по пластическим признакам подтвердил существование двух хорошо различных **морфологических** форм, идентифицируемых соответственно как восточносибирский и ледниково-равнинный сиги. Необходимо отметить, что в средних по размеру реках морфооблик сигов обеих форм/видов выражен более четко, чем в нижнем течении больших рек (р. Лена). Высокая кон-

трастность условий среды на относительно небольшом протяжении малых рек (в диапазоне от реки полугорного типа в среднем течении до лимана в нижнем течении) позволяет этим формам/видам существовать достаточно обособленно. Несмотря на то что выраженность видоспецифичных признаков (форма головы, горбатость) в некоторых случаях была незначительной, классификация UPGMA по пластическим признакам подтвердила морфологическую уникальность этих форм.

Сети гаплотипов и филогеография. Следует признать, что в отличие от югайского сига ледниково-равнинный сиг в целом не имеет видоспецифичной мтДНК. Гаплотипы сигов из разных локальностей исследованного региона сильно перемешаны. Гаплогруппа I, состоящая из гаплотипов сигов из

Таблица 6. Параметры *mismatch*-распределения для объединенных выборок сигов из р. Лены с притоками и рек Яна и Хрома

Популяция	Тестирование модели демографической экспансии			
	<i>SSD</i>	<i>p</i> (<i>SSD</i>)	<i>r</i>	<i>p</i> (<i>r</i>)
р. Лена с притоками	0.00162	0.09900	0.02038	0.19600
Реки Яна, Хрома	0.00277	0.24400	0.00780	0.28900
Популяция	Тестирование модели пространственной экспансии			
	<i>SSD</i>	<i>p</i> (<i>SSD</i>)	<i>r</i>	<i>p</i> (<i>r</i>)
р. Лена с притоками	0.00066	0.32900	0.02038	0.24600
Реки Яна, Хрома	0.00052	0.55300	0.00780	0.74100

бассейна р. Енисей и тесно связанная с гаплотипами восточносибирского сига из р. Анабар, отличается большей однородностью. В эту гаплогруппу входит и значительная часть *уникальных гаплотипов* сигов из арктических водоемов (реки Попигай, Омолой, Яна, Хрома). Гаплогруппа II сформирована гаплотипами сигов из бассейна р. Лена, оленекских, анабарских сигов, а также гаплотипами сигов из бассейна р. Енисей. Гаплогруппа III сформирована гаплотипами юкагирских сигов, относительно недавнее происхождение которых от одной из филогенетических линий ледниково-равнинных сигов обсуждалось ранее [30, 48, 49]. Кроме того, в медианной сети выделяется гаплогруппа IV, сформированная слабо связанными между собой гаплотипами сигов из рек Оленёк и Анабар. Подобные структуры характерны для арктических популяций ледниково-равнинных сигов, но редко встречаются в реконструкциях гаплотипов сигов для водоемов Южной Сибири. Данный факт косвенно подтверждает гипотезу о присутствии в арктических водоемах форм/видов сигов более древнего происхождения, чем **современные** дериваты филогенетических линий восточносибирских сигов из рек Обь и Анабар. Следует отметить, что восточносибирские сиги из р. Анабар генетически не тождественны сигам из русловой части р. Лены, хотя по внешним морфологическим признакам они слабо различимы и относятся к одной экологической группе [49, 50]. В отличие от медианных сетей, построенных для сигов Южной Сибири, в настоящей сети (рис. 4) доминируют слабо связанные цепи уникальных гаплотипов. Похожие структуры наблюдаются при анализе гаплотипов уссурийского *C. ussuriensis* сига из бассейна р. Амур, где следов четвертичного оледенения обнаружено не было [51–53]. Мы полагаем, что и популяции ледниково-равнинного сига (в отличие от восточносибирского) в последнее время не подвергались значительному сокращению численности, что подтверждается анализом на селективную нейтральность. Иерархический анализ молекулярной дисперсии AMOVA

показал, что большая часть генетической изменчивости приходится на внутривидовую составляющую (72%), что подтверждает высокую степень смешения гаплотипов сигов, очевидно вследствие их высокой миграционной активности в четвертичный период. Низкие значения индекса F_{ST} между популяциями свидетельствуют в пользу интенсивного потока генов между ними. В то же время поток генов между географически удаленными популяциями из бассейнов рек Енисей и Лена сильно ограничен.

Эволюционная история. Общие генетические показатели популяций сигов из бассейна р. Лены в целом уменьшаются от верхнего течения (по всем гаплотипам сигов из Баунтовской системы озер) к среднему течению реки [28]. В популяциях из нижнего течения показатели генетического полиморфизма значительно выше. Самые высокие значения π отмечены в популяциях арктических рек – Оленёк, Яна, Хрома и Оленёкской протоке дельты р. Лены. Самые низкие генетические показатели обнаружены в популяции сигов из относительно изолированного оз. Б. Токо. Близкие показатели были обнаружены в популяциях сигов из горных озер, Доронг, Б. Капылюши [28]. Эти озера лежат в пределах верхнего течения р. Ципа – одного из главных притоков р. Витим бассейна р. Лены. Наиболее контрастно (по показателям генетического полиморфизма) относительно ленских популяций сигов выглядят популяции сигов из оз. Каракуль бассейна р. Абакан (приток р. Енисей). В популяции из оз. Каракуль выявлены крайне низкие значения общих генетических показателей, указывающие на их недавнее происхождение (что подтверждается и геологической историей региона). Тесты на нейтральность (особенно F_u F_s), проведенные для большого числа популяций сигов из водоемов Сибири, показали, что они почти всегда принимают отрицательные и частично достоверные значения. Учитывая значительную долю достоверных отрицательных значений F_u F_s и то, что этот тест считается более чувствительным для выявления недавней экс-

пансии популяции, можно предполагать, что в недавнем прошлом эти популяции существенно отклонились от равновесного состояния. Подобные результаты тестов предполагают расширение, очищающий отбор, или генетический хитчайкинг в недавнем прошлом. Положительные значения в наших исследованиях почти всегда связаны с незначительными по объему выборками. Для некоторых популяций бассейна р. Лены зафиксированы положительные, но недостоверные, значения F_s , предполагающие недавнее прохождение популяции “бутылочного горлышка” либо ее фрагментацию.

Анализ распределения частот гаплотипов в популяциях показал, что модель пространственной экспансии сигов из бассейна р. Лены и рек Яны и Хрома оказалась более предпочтительной. Однако недостоверные значения среднеквадратичного отклонения и индексов не позволяют отдать предпочтение той или иной демографической модели, что отмечалось ранее и для других арктических популяций сигов [27, 28, 42]. *Mismatch*-распределение близкое к унимодальному соответствуют модели исторической “взрывной” демографической экспансии или пространственной экспансии с высоким уровнем миграции между соседними популяциями. Пространственной экспансии, как правило, предшествует демографическая экспансия, что и накладывает отпечаток на результаты MMD-анализа [51–53]. Аналогичная тенденция ранее была отмечена и для популяций сигов из бассейна р. Енисей и некоторых других арктических водоемов. Паттерн распределения частот попарных нуклеотидных различий, полученный для исследованных популяций из меньших по размеру рек, характеризуется наличием двух ярко выраженных максимумов значений, т.е. является бимодальным, что, по-видимому, определяется наличием пары удаленных гаплотипов. Такое распределение свидетельствует в пользу предположения, что исследованные структуры сформировались в результате слияния генетически дистантных, конспецифичных популяций.

ЗАКЛЮЧЕНИЕ

В отличие от верхнего течения р. Лены (Баунтовская система озер) морфологическое разнообразие сигов в среднем и нижнем течении р. Лены значительно ниже. В русле р. Лены доминирует низкотелый восточносибирский сиг. В дельте реки преобладают высокотелые ледниково-равнинные сиги. Сходная картина наблюдается во всех рассмотренных нами арктических реках, где восточносибирский сиг доминирует в верхнем и среднем течении, а ледниково-равнинный сиг тяготеет к приустьевым участкам рек и лиману [27,

49]. Наличие популяций ледниково-равнинного сига во всех изученных нами бассейнах рек и озер дает основания для предположения о *доминировании данной формы/вида в лиманах* всей сибирской Арктики. Поскольку ареал пыжьяновидных сигов не ограничен Сибирью, уместно предположить, что в основе множества ранее описанных внутривидовых форм сигов из водоемов Северо-Западной Европы лежат сибирские филогенетические линии, тождественные ледниково-равнинному сигу.

Отрицательные значения большинства тестов на нейтральность эволюции свидетельствуют о недавней пространственной экспансии популяций сигов, что согласуется с другими рассмотренными характеристиками. Однако такая гипотеза не исключает и вероятности прохождения через бутылочное горлышко или “эффекта основателя” для сигов более изолированных от основной (оз. Кутарамакан, Илирней) популяции.

Анализ распределения частот гаплотипов свидетельствует о том, что в различных по гидрологическим характеристикам водоемах популяции сигов могут претерпевать разнонаправленные эволюционные процессы. MMD-анализ для выборок сигов из *средних по размерам рек* арктического бассейна выявил мультимодальное распределение частот гаплотипов, что может свидетельствовать о смешанной выборке. Тогда как унимодальное распределение частот гаплотипов для сигов из нижнего течения р. Лены, наряду с тестами на селективную нейтральность и радиальную структуру медианной сети, вернее служит свидетельством экспансии ледниково-равнинных сигов глубоко в русло р. Лены, где экологические условия для этой формы продолжают быть оптимальными. Подобное расселение и гибридизацию ледниково-равнинных сигов мы уже наблюдали в р. Оленёк [27].

Учитывая в целом относительно невысокие различия в структуре исследованного фрагмента мтДНК между сигами из четырех основных гаплогрупп, можно констатировать, что ледниково-равнинный сиг, как и муксун, не имеет видоспецифичной мтДНК. Аналогичные закономерности наблюдаются и для пеляди, чья мтДНК тождественна ряпушковой. При этом хорошо диагностируемые по морфологическим и экологическим свойствам виды могут достаточно надежно различаться по другим генетическим признакам. Например, аллель локуса идитолдегидрогеназы *sIDDH**^f, встречающийся в популяциях ледниково-равнинного сига и юкагирского сига, совсем не отмечается в популяциях восточносибирского сига и у *разных экологических форм муксуна* [51]. На основании этого наблюдения можно сделать предположение о регулярных заменах мтДНК малочисленных аллопатрических линий и форм на мтДНК видов, домини-

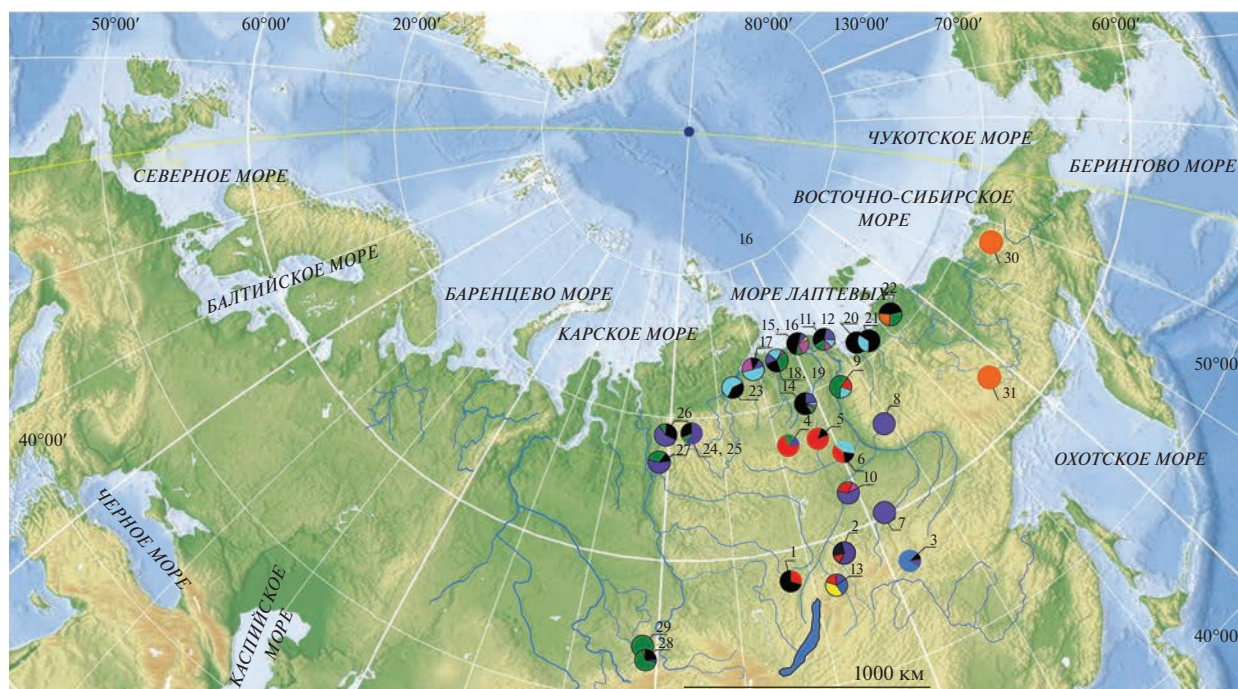


Рис. 7. Географическое распределение гаплотипов (*ND1* мтДНК) сигов в некоторых водоемах Сибири. Точки 1–31 соответствуют номерам в табл. 1.

рующих на ареалах в определенные временные отрезки. В связи с проявлением подобного эффекта “омоложения” по мтДНК в филогенетическом анализе может возникать путаница, маскирующая часть хороших видов под “экологические формы” и приводящая к неверным выводам. Комплексный анализ экологических, морфологических и разнокачественных генетических признаков у исследуемых популяций/форм, относящихся к комплексным видам, позволит решить эту проблему.

Очевидно, что становление современных популяций сигов в Сибири в четвертичное время связано с событиями кардинальных изменений водных связей в периоды похолоданий и потеплений климата, морских трансгрессий и регрессий. Высокая перемешанность гаплотипов в наиболее крупных гаплогруппах из водоемов сибирского Заполярья указывает на то, что именно в районах, приуроченных к низовьям рек и их приустьевым морским участкам, генетический обмен между представителями разных филогенетических линий на разных временных отрезках происходил наиболее интенсивно. Например, в стадию Каргинского оледенения около 50–60 тыс. лет назад значительный обмен элементами пресноводной фауны между сибирскими реками мог происходить на осушенных шельфовых участках Карского, Лаптевых, Восточно-Сибирского и Чукотского морей [54]. Кроме того, в приустьевых речных участках распресненных арктических морей с до-

статочно разнообразной и богатой кормовой базой сига, принадлежащие к экологической группе ледниково-равнинных сигов, обитали в условиях экологического оптимума, и их популяции не подвергались значительным колебаниям численности в связи с изменением внешних факторов среды. В отличие от них генетический обмен между изолированными популяциями сигов из верхних и средних течений смежных речных систем в четвертичное время мог происходить в связи с гидрологическими перестройками из-за локальных покровных оледенений, масштаб которых в нижнем плейстоцене в Центральной Сибири был ограничен по сравнению с севером Западной Сибири и Европы. Именно поэтому популяции сигов, приуроченные к средним и верхним течениям рек, сохраняют большую генетическую обособленность, крайний случай которой проявляется в виде формирования аллопатрических юагирских сигов в бассейнах рек Колымы и Индигирки [30]. Все отмеченные выше закономерности в выявлении родственных отношений сигов разных форм/видов и их распределении на исследуемом ареале хорошо иллюстрируются картой распределения гаплотипов в современных популяциях (рис. 7).

Работа выполнена при поддержке гранта РФФИ FSRG–2020–0019 и частичной поддержке гранта РФФИ (Монг_а) 20-44-20-54-44017.

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. Gasovska M. Genus *Coregonus* L. discussed in connection with a new systematic feature that of shape and proportion of os maxillare and os supramaxillare // *Annal. Zool. Warszawa*. 1960. V. XVIII. № 26. P. 471–513.
2. Kahilainen K. Ecology of Sympatric Whitefish (*Coregonus lavaretus* (L.) Forms in a Subarctic. Finland, 2004. 44 p.
3. Bitner D., Excoffier L., Largiadèr C.R. Patterns of morphological changes and hybridization between sympatric whitefish morphs (*Coregonus* spp.) in a Swiss lake: A role for eutrophication? // *Mol. Ecol.* 2010. V. 19. P. 2152–2167
<https://doi.org/10.1111/j.1365-294X.2010.04623.x>
4. Зуйкова Е.И., Бочкарев Н.А. Особенности строения и функционирования жаберно-челюстного аппарата сига Правдина *Coregonus lavaretus pravdinellus* Dulkeit // *Вопр. ихтиологии*. 2008. Т. 48. № 6. С. 767–776.
5. Берг Л.С. Рыбы пресных вод СССР и сопредельных стран. Ч. I. М.: Изд-во АН СССР, 1948. 468 с.
6. Бозуцкая Н.Г., Насека А.М. Каталог бесчелюстных и рыб пресных и солоноватых вод России с номенклатурными и таксономическими комментариями. М.: Тов. науч. изд. КМК, 2004. 389 с.
7. Гундризер А.Н., Иоганзен Б.Г., Кафанова В.В. и др. Рыбы Телецкого озера. Новосибирск: Наука, 1981. 159 с.
8. Решетников Ю.С. Экология и систематика сиговых рыб. М.: Наука, 1980. 300 с.
9. Скрябин А.Г. Сиговые рыбы юга Сибири. Новосибирск: Наука, 1979. 230 с.
10. Kottelat M., Freyhof J. Handbook of European Freshwater Fishes. Cornol. Switzerland; Berlin, Germany, 2007. 640 p.
11. Подлесный А.В., Лобовикова А.А. Рыбы Таймырского озера // *Вопросы географии Сибири*. Томск: Изд-во Томск. ун-та, 1951. № 2. С. 269–292.
12. Seehausen O. Explosive speciation rates and unusual species richness in haplochromine cichlid fishes: effects of sexual selection // *Adv. Ecol.* 2000. V. 31. P. 237–274.
13. Мухомедияров Ф.Б. Ряпушка – *Coregonus sardinella baunti* subsp. nova из системы Ципа–Ципиканских озер бассейна реки Витим // *Докл. на 1-й науч. сессии Якутской базы АН СССР*. Якутск, 1948. С. 270–280.
14. Кириллов Ф.Н. Рыбы Якутии. М.: Наука, 1972.
15. Романов В.И., Зуйкова Е.И., Бочкарев Н.А. Морфологическая и экологическая дифференциация симпатрических видов рода *Coregonus* из оз. Таймыр // *Сиб. экол. журн.* 2016. № 3. С. 343–351.
16. Михин В.С. Сиг-пыжьян реки Оленёк // *Вопр. ихтиологии* 1959. Вып. 13. С. 71–74.
17. Остроумов Н.А. Рыбы и рыбный промысел р. Пясины // *Тр. Поляр. Комиссии*. 1937. Вып. 30. 115 с.
18. Шапошникова Г.Х. Сиг-пыжьян (*Coregonus lavaretus pidschian natio gydanus* nova) Гыданского залива // *Труды Ин-та полярного земледелия. Серия Промысловое хоз-во*. 1941. Вып. 15. С. 63–82.
19. Новиков А.С. Рыбы реки Колымы. М.: Наука, 1966. 134 с.
20. Кагановский А.Г. Промысловые рыбы р. Анадыря и Анадырского лимана // *Вестник ДВ филиала АН СССР*. 1933. № 1/2. С. 137–139.
21. Карантонис Ф.Э., Кириллов Ф.Н., Мухомедияров Ф.Б. Рыбы среднего течения р. Лены // *Тр. Ин-та Биологии Якутского филиала АН СССР*. 1956. Вып. 2. С. 3–144.
22. Østbye K. Evolutionary history of the European whitefish *Coregonus lavaretus* (L.) species complex as inferred from mtDNA phylogeography and gill-raker numbers // *Mol. Ecol.* 2005. V. 14. P. 4371–4388.
<https://doi.org/10.1111/j.1365-294x.2005.02737.x>
23. Østbye K., Næsje T.F., Bernatchez L. et al. Morphological divergence and origin of sympatric populations of European whitefish (*Coregonus lavaretus* L.) in Lake Femund, Norway // *J. Evol. Biol.* 2004. V. 18. P. 683–702.
<https://doi.org/10.1111/j.1420-9101.2004.00844.x>
24. Смирнов В.В., Смирнова-Залуми Н.С., Суханова Л.С. Микроэволюция байкальского омуля *Coregonus autumnalis migratorius* (Georgi). Новосибирск: Изд-во СО РАН, 2009. 246 с.
25. Bochkarev N.A., Zuykova E.I., Katokhin A.V. Morphology and mitochondrial DNA variation of the Siberian whitefish *Coregonus lavaretus pidschian* (Gmelin) in the upstream water bodies of the Ob and Yenisei rivers // *Evol. Ecol.* 2011. V. 25. P. 557–572.
<https://doi.org/10.1007/s10682-010-9437-7>
26. Bochkarev N.A., Zuykova E.I., Abramov S.A. et al. Morphological, biological and mtDNA sequences variation of coregonid species from the Baunt lake system (the Vitim river basin) // *Adv. Limnology*. 2013. V. 64. P. 257–277.
<https://doi.org/10.1127/1612-166x/2013/0064-0025>
27. Бочкарев Н.А., Сендек Д.С., Зуйкова Е.И. и др. Популяционная структура и происхождение некоторых экологических форм *Coregonus lavaretus pidschian* из р. Оленёк // *Генетика*. 2021. Т. 57. № 7. С. 797–809.
<https://doi.org/10.1134/s1022795421070048>
28. Бочкарев Н.А., Сендек Д.С., Катохин А.В. и др. Морфологическая, экологическая и генетическая изменчивость сигов *Coregonus lavaretus sensu lato* из верхнего и среднего течения р. Лены // *Генетика*.

2022. Т. 58. № 11. С. 1292–1310.
<https://doi.org/10.31857/S0016675822110030>
29. Bogdanov B.E., Knizhin I.B. The Salmonid fishes of Lake Baikal and its adjacent water systems: Annotated checklist with new taxa description // *Limnology and Freshwater Biology*. 2022. V. 6. P. 1688–1704.
<https://doi.org/10.31951/2658-3518-2022-A-6-1688>
30. Bochkarev N.A., Zuykova E.I., Pestryakova L.A. et al. Intraspecific structure of the *Coregonus lavaretus* complex in water bodies of Siberia: A case of postglacial allopatric origin of Yukagirian whitefish // *Can. J. Zool.* 2021. V. 99. P. 1040 – 1053.
<https://doi.org/10.1139/cjz-2021-0045>
31. Бочкарев Н.А., Зуйкова Е.И., Политов Д.В. Дополнительные возможности сбора и регистрации морфологических данных у рыб // *Биология, биотехника разведения и состояние запасов сиговых рыб* / Под ред. Литвиненко А.И., Решетникова Ю.С. Тюмень: ФГУП Госрыбцентр, 2013. С. 32–36.
32. Правдин И.Ф. Руководство по изучению рыб. М.: Пищепромиздат, 1966. 376 с.
33. Darroch J.N., Mosimann J.E. Canonical and principal components of shape // *Biometrika*. 1985. V. 72. № 2. P. 241–252.
34. Mosimann J.E. Size allometry: size and shape variables with characteristics of the log normal and generalized gamma distributions // *J. Am. Stat. Association*. 1970. V. 65. P. 930–945.
35. Hammer Ø., Harper D.A.T., Ryan P.D. PAST: Paleontological statistics software package for education and data analysis // *Palaeontologia Electronica*. 2001. V. 4(1) P. 1–9.
36. Librado P., Rozas J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data // *Bioinformatics*. 2009. V. 25. P. 1451–1452.
<https://doi.org/10.1093/bioinformatics/btp187>
37. Excoffier L., Lischer H.E.L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows // *Mol. Ecol. Res.* 2010. V. 10. P. 564–567.
<https://doi.org/10.1111/j.1755-0998.2010.02847.x>
38. Fu Y.-X., Li W.-H. Statistical tests of neutrality of mutations // *Genetics*. 1993. V. 133. P. 693–709.
39. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism // *Genetics*. 1989. V. 123. № 3. P. 585–595.
40. Bandelt H.-J., Forster P., Röhl A. Median-joining networks for inferring intraspecific phylogenies // *Mol. Biol. Evol.* 1999. V. 16. P. 37–48.
41. Rogers A.R., Harpending H. Population growth makes waves in the distribution of pair-wise genetic differences // *Mol. Biol. Evol.* 1992. V. 9. № 3. P. 552–569.
42. Бочкарёв Н.А., Зуйкова Е.И., Романов В.И. и др. Морфологическая и генетическая изменчивость симпатрических сигов комплекса *Coregonus lavaretus pidschian* из оз. Кутармакан Хантайской гидро- системы (п-ов Таймыр) // *Генетика*. 2020. Т. 56. № 5. С. 571–583.
<https://doi.org/10.31857/S0016675820050033>
43. Шапошникова Г.Х. Сиг-пыжьян *Coregonus lavaretus* (Gmelin) водоемов Советского Союза // *Вопр. ихтиологии*. 1974. Т. 14. Вып. 5(88). С. 749–768.
44. Бочкарев Н.А., Зуйкова Е.И. Популяционная структура сига-пыжьяна *Coregonus lavaretus pidschian* (Gmelin, 1789) (Coregonidae) в озерах Тоджинской котловины и верхнем течении реки Большой Енисей (Тува) // *Зоол. журн.* 2009. Т. 88. № 1. С. 47–60.
45. Бочкарев Н.А., Зуйкова Е.И. Популяционная структура и возможные пути дифференциации сига-пыжьяна (*Coregonus lavaretus pidschian*, Coregonidae) в Додотских озерах Тоджинской котловины // *Сиб. экологический журн.* 2010. Т. 1. С. 41–55.
46. Крозиус Ф.В. Материалы по систематике и биологии сига оз. Байкал // *Тр. Байкал. лимнол. станции АН СССР*. 1933. Т. 5. С. 5–154.
47. Исаченко В.Л. Новый вид сига из бассейна р. Енисей // *Тр. Сибирской ихтиол. лаб.* 1925. Т. II. Вып. 2. 18 с.
48. Бочкарев Н.А., Зуйкова Е.И., Соловьев М.М. Вторичная интерградация различных форм сигов (*Coregonus lavaretus sensu lato*, Coregonidae) в водоемах Алтае-Саянской горной страны // *Экол. генетика*. 2017. Т. 15. № 2. С. 31–43.
<https://doi.org/10.17816/ecogen15231-43>
49. Бочкарев Н.А., Пестрякова Л.А., Захаров Е.С. и др. Сиг-пыжьян (*Coregonus lavaretus pidschian*, Coregonidae) р. Анабар: морфогенетическая структура популяций // *Генетика*. 2018. Т. 54. № 9. С. 1057–1067.
<https://doi.org/10.1134/S0016675818090047>
50. Сендек Д.С., Иванов Е.В. О причинах генетической неоднородности сига-пыжьяна *Coregonus lavaretus pidschian* реки Анабар // *Экол. генетика*. 2017. Т. 15. № 3. С. 20–26.
<https://doi.org/10.17816/ecogen15320-26>
51. Sendek D.S., Ivanov Ye.V., Khodulov V.V. et al. Genetic differentiation of coregonids populations in Subarctic areas // *Adv. Limnology*. 2013. V. 64. P. 223–246.
52. Готванский В.И., Махинов А.Н. Новые данные о распространении четвертичного оледенения на Сихоте-Алине // *Изв. Всесоюз. геогр. об-ва*. 1983. Т. 115. Вып. 2. С. 154–159.
53. Bochkarev N.A., Zuykova E.I., Abramov S.A. et al. The sympatric whitefishes *Coregonus ussuriensis* and *C. chadary* from the Amur river basin: morphology, biology and genetic diversity // *Fundam. Appl. Limnol.* 2017. V. 189. № 3. P. 193–207.
<https://doi.org/10.1127/fal/2016/0801>
54. Excoffier L. Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model // *Mol. Ecol.* 2004. V. 13. P. 853–864.
<https://doi.org/10.1046/j.1365-294X.2003.02004.x>

Causes of Morphological and Genetic Heterogeneity of White Fish *Coregonus lavaretus* sensu lato in the Arctic Part of the Lena River Basin

N. A. Bochkarev^{a, *}, D. S. Sendek^b, E. I. Zuikova^a, L. A. Pestryakova^c, L. A. Ushnitskaya^c,
E. S. Zakharov^c, A. A. Everstova^c, N. N. Zakharova^c, and D. V. Politov^d

^aFederal Research Center, Institute of Systematics and Ecology of Animals, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 630091 Russia

^bState Research Institute of Lake and River Fishery, Saint Petersburg, 199053 Russia

^cAmnosov North-Eastern Federal University, Yakutsk, 677980 Russia

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

*e-mail: nikson_1960@mail.ru

The results of the morphological and genetic analysis of populations/forms/species of tanned whitefish from the middle and lower reaches of the river are presented. Lena, and adjacent rivers. There are reasons to believe that, despite the significant heterogeneity of whitefish in Siberia, a small number of stable forms/species inhabit the water bodies of the Siberian Arctic. These are East Siberian (*C. lavaretus pidschian* n. *brachymystax*), glacial plain (*C. lavaretus pidschian* n. *glacialis*), and Yukagir whitefish (*C. lavaretus pidschian* n. *jucagiricus*), as well as their hybrid populations. The glacier-plain whitefish and its derivatives differ from the East Siberian whitefish in plastic features, but they are identical in meristic ones. Molecular genetic studies have shown that in most glacier-plain whitefish form extensive loosely connected networks of haplotypes, which corresponds to the long-term existence of populations in favorable conditions. The East Siberian whitefish haplotypes are closely related to the whitefish haplotypes from the water bodies of Southern Siberia and have a radial structure. Such a structure of median networks indicates a recent decline in the number of East Siberian whitefish followed by expansion, which is directly related to the events of the Quaternary glaciations.

Keywords: whitefish, mtDNA, haplotypes, evolution, morphogenesis, Siberia, Palearctic, Pleistocene, Lena river, B. Yenisei river.

ГЕНЕТИЧЕСКАЯ ОЦЕНКА ПРОГНОЗИРУЕМОГО ОСТАТОЧНОГО ПОТРЕБЛЕНИЯ КОРМА И ЭКСПРЕССИЯ ЗНАЧИМЫХ ГЕНОВ-КАНДИДАТОВ СВИНЕЙ ПОРОДЫ ДЮРОК И ТОВАРНЫХ ПОМЕСЕЙ ВТОРОГО ПОКОЛЕНИЯ

© 2023 г. А. А. Белоус¹, *, А. А. Сермягин¹, Н. А. Зиновьева¹

¹Федеральный исследовательский центр животноводства — ВИЖ им. академика Л.К. Эрнста,
пос. Дубровицы, Московская обл., 142132 Россия

*e-mail: belousa663@gmail.com

Поступила в редакцию 27.04.2023 г.

После доработки 29.05.2023 г.

Принята к публикации 02.06.2023 г.

Остаточное потребление корма (RFI) — одна из основных и сложных кормовых характеристик, которая экономически важна для животноводства. Однако генетические и биологические механизмы, регулирующие данный признак, у свиней в значительной степени неизвестны. Таким образом, настоящее исследование было направлено на выявление полногеномных однонуклеотидных полиморфизмов (SNP), генов-кандидатов, участвующих в регуляции RFI, их биологических путей и кластеризации, с использованием полногеномного анализа ассоциации (GWAS). Исследование проводилось на свиньях породы дюрок ($n = 783$) и их товарных гибридах второго поколения ($n = 250$), проходящих тестовый откорм на автоматических кормовых станциях индивидуального учета. В результате были получены значимые по онтологии биологических функций и по экспрессии в тканях и органах гены, имеющие связь с RFI. К таким генам-кандидатам отнесены гены, кодирующие адгезию—рецептор, связанный с белком G6 (*ADGRG6*), центромерный белок S (*APITD1*), карбокси-пептидазу E (*CPE*), трансмембранный кальций-связывающий белок (*SYTL2*), молекулу клеточной адгезии 1 (*CADMT1*), протоонкоген Fli-1, фактор транскрипции ETS (*FLI1*), трансмембранный белок 3 теневина (*TENM3*), простагландин E4 (*PTGER4*) и член 2 подсемейства D калиевых потенциал-зависимых каналов (*KCND2*). Также анализ полученных данных по кластеризации показал разделение на биологическую, функциональную и молекулярную библиотеки и данные, опубликованные в PubMed. Объединяя полученную информацию, можно сказать, что генетическая составляющая показателя прогнозируемого остаточного потребления корма важна, о чем было указано в предыдущих исследованиях. В связи с чем возникает необходимость создания молекулярных диагностик и разработки расчетов геномной оценки, в совокупности с конверсией корма, что позволит улучшить показатели продуктивности в племенных стадах свиней и улучшить качество производимой продукции.

Ключевые слова: эффективность использования корма, прогнозируемое остаточное потребление корма, геномная оценка, полногеномное ассоциативное исследование, экспрессия генов, хрячки породы дюрок, помеси второго поколения, свиноводство.

DOI: 10.31857/S0016675823110024, **EDN:** NNVFQI

Корма являются основным экономическим фактором в свиноводстве и составляют 60–70% от общей стоимости производства. Стратегии, повышающие эффективность кормления, являются приоритетными для отрасли свиноводства и направлены, в первую очередь, на снижение кормовых затрат и выделение питательных веществ. В связи с этим показатель остаточного потребления корма (RFI) считается наиболее подходящей мерой для изучения биологических факторов, влияющих на эффективность корма, поскольку в рамках этой меры учитываются различия в производственных

характеристиках, таких как живая масса, скорость роста, состав тела и условия содержания [1]. В последние годы были предприняты значительные усилия для понимания физиологической и молекулярно-генетической характеристики RFI у свиней разных пород. Результаты исследований выявили ряд процессов, лежащих в основе формирования и всасывание питательных веществ, обмен белка, энергетический метаболизм и иммунную функцию [2–5]. Также было установлено, что изменения в функциях митохондрий влияют на эф-

фективность использования корма, что соответствует их центральной роли в энергетическом метаболизме [6, 7]. С другой стороны, и изменение иммунного ответа также было определено как фактор, влияющий на данный признак [4, 8, 9].

Независимо от биологических процессов, влияющих на RFI, получение ранних предикторов показателей эффективности использования корма имеет весомую значимость, поскольку измерение индивидуального потребления корма невозможно в коммерческих производственных системах, без использования автоматических кормовых станций индивидуального откорма. В этом отношении геномная селекция является актуальной и необходимой, поскольку позволяет проводить раннюю селекцию без необходимости прямого измерения потребления корма [10]. Для оптимизации как научной, так и производственной работы возрастает потребность в разработке и расширении панелей для функционально-биологически важных SNP, которые повысят точность геномных прогнозов по таким сложным признакам, как RFI. Для работы по данному этапу возникает существенная потребность в поиске достоверных генов-кандидатов методом полногеномного анализа, для последующего выявления их экспрессии и включения в тест-системы. Показатель RFI имеет умеренную наследуемость (0.21–0.33), поэтому геномная оценка поспособствует получению достоверной информации генетического прогресса, по сравнению с традиционными методами селекции, основанными на оценке фенотипа и родословной [11–14]. Необходимо выявлять гены, которые изменяют мотивы сайтов связывания факторов транскрипции в промоторных областях и дифференциально экспрессируются у различных животных. При таком подходе важно, чтобы транскриптомный анализ выполнялся в тканях с наибольшим влиянием на признак прогнозируемого остаточного потребления корма, таких как печень и мышцы. Гены, идентифицированные в метаболически значимых тканях, изученные в разных популяциях свиней, являются более достоверными для изучаемого признака. Для показателя прогнозируемого остаточного потребления корма наиболее ключевыми являются печень и мышцы [6, 15].

В связи с вышеизложенной актуальностью цель исследований заключается в выявлении метаболически важных генов-кандидатов методом полногеномного ассоциативного исследования, имеющих значимость для показателя прогнозируемого остаточного потребления корма у чистопородных свиней и их помесей.

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Научно-исследовательская работа проводилась на генотипированных хрячках породы дюрок

($n = 783$) и их помесей (F2, $n = 250$), проходящих тестовый откорм на автоматических кормовых станциях индивидуального учета. Расчет признака прогнозируемого остаточного потребления корма для каждой популяции был по следующей модели множественной регрессии [16]:

для свиней породы дюрок:

$$RFI = ADFI -$$

$$- (412.42 + 56.18MWT^{0.75} + 0.59ADG);$$

для товарных помесей:

$$RFI = ADFI -$$

$$- (1870.04 + 75.07MWT^{0.75} - 0.81ADG),$$

где RFI – прогнозируемое остаточное потребление корма, г; ADFI – среднесуточное потребление корма, г; $MWT^{0.75}$ – средняя метаболическая масса животного, кг; ADG – среднесуточный прирост, г.

В качестве материала для молекулярно-генетических исследований использовали пробы ткани (ушные выщипы). Выделение ДНК проводили с помощью наборов для геномной ДНК серии “ДНК-Экстрен” (ЗАО “Синтол”, Россия), в соответствии с протоколом фирмы-производителя. Концентрацию двухцепочечной ДНК определяли с помощью флуориметра Qubit 2.0 (Invitrogen/Life Technologies, США). Для определения качества ДНК измеряли соотношение OD260/OD280 (спектрофотометр NanoDrop8000; ThermoFisher Scientific, США). Для анализа использовали ДНК с OD260/OD280 = 1.6–1.8. Полногеномное генотипирование проводили с использованием высокоплотных ДНК-чипов Porcine GGP HD (платформа GeneSeek Genomic Profiler; Neogene, США), содержащих ≈ 70 тыс. SNP. Контроль качества и фильтрацию данных генотипирования для каждого SNP и каждого образца выполняли с использованием программного пакета PLINK 1.9 (<http://zzz.bwh.harvard.edu/plink/>), применяя следующие фильтры:

- call-rate по всем исследуемым SNP для индивидуального образца не ниже 90%;

- call-rate для каждого из исследованных SNP по всем генотипированным образцам не ниже 90%;

- частота встречаемости минорных аллелей (MAF) ≥ 0.05 ;

- отклонение генотипов по SNP от распределения по Харди–Вайнбергу в совокупности протестированных образцов.

Для поиска генов-кандидатов, локализованных в области идентифицированных SNP, и их экспрессии, использовали геномный ресурс Sscrofa11.1 (https://www.ncbi.nlm.nih.gov/assembly/GCF_000003025.6) и Ensembl (www.ensembl.org). Функциональные аннотации и выявление обогащения генов вы-

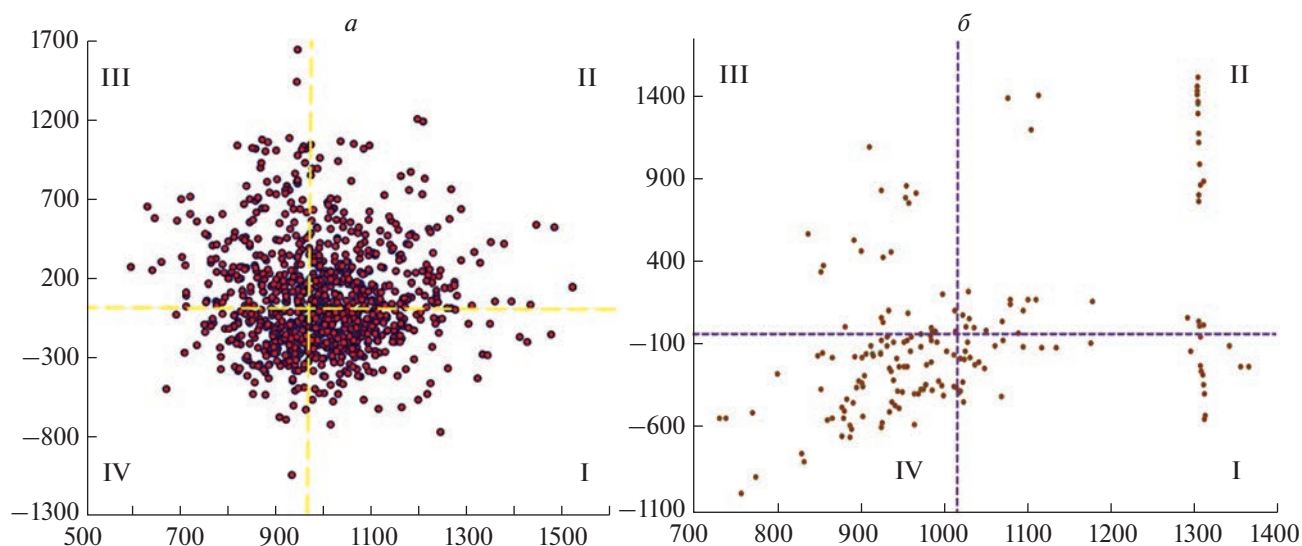


Рис. 1. Распределение прогнозируемого остаточного потребления корма по группам I–IV в зависимости от среднесуточного прироста. *а* – популяция хрячков породы дюрок; *б* – популяция помесей второго поколения.

полняли с привлечением базы данных DAVID (<https://david.ncifcrf.gov/summary.jsp>) и STRING (string-db.org).

Кластеризацию полученных генов-кандидатов проводили в программе TopCluster (<https://topcluster.cchmc.org/>) с применением биологических, молекулярных библиотек и публикаций PubMed.

РЕЗУЛЬТАТЫ

Статистическая обработка фенотипических данных

Порода дюрок используется как отцовская при выведении товарных помесей второго поколения. В связи с чем сравнительный аспект на генетическом уровне показателя прогнозируемого остаточного потребления корма необходим для понимания и улучшения селекционного отбора. Изначально провели исследование по соотношению фактических данных среднесуточного прироста и RFI (рис. 1). На рис. 1, *а* хрячки породы дюрок в большей части распределены в группу I ($ADG > 960$, $RFI < -0$), что указывает на хорошую усвояемость корма, в связи с чем можно их рекомендовать в племенное ядро стада. Большая часть товарных помесей, представленных на рис. 1, *б*, принадлежит IV (мясной) группе ($ADG < 1020$, $RFI < -0$).

Свиньи породы дюрок и их помеси считаются скороспелыми животными, по данным исследований среднесуточный прирост составил 960 и 1020 г соответственно. Регуляция генетического материала по хозяйственно полезным признакам необходима для выхода качественной мясной

продукции и получения быстрорастущих товарных свиней.

Полногеномное ассоциативное исследование прогнозируемого остаточного потребления корма

После фильтрации данных осталось 43 212 SNP по породе дюрок и 43 687 SNP по товарным помесям второго поколения (рис. 2).

Прогнозируемое остаточное потребление корма характеризуется выявленными 178 SNP на 17 из 18 SSA (порода дюрок) и 27 SNP на 10 из 18 SSA (F2). SNP, выходящие за полногеномный порог, обнаружены на SSA6 (WU_10.2_6_29045127), SSA10 (ASGA0048559) и SSA12 (MARC0093419) у чистопородных хрячков и на SSA1 (ALGA0006952) – у боровов. WU_10.2_6 был ранее выявлен в исследованиях An M. [17] как маркер, влияющий на репродуктивные аномалии у свиноматок, связанные со смертностью поросят. На позиции 15567222 BP данного SNP выявлен предген *ZFH3* (Zinc finger homeobox 3). SNP ALGA0006952 был обнаружен как маркер, влияющий на выработку андростенона (стероидный феромон, содержащийся в слюне кабана) и скатола (образуется в кишечнике животных в результате разложения триптофана – одной из α -аминокислот, входящих в состав белков) в организме кабана и свиней [18].

Структурная и функциональная аннотация генов-кандидатов

Структурная аннотация выявила наличие 358 генов-кандидатов, из них 345 – у свиней породы дюрок и 13 – у товарных помесей второго поколения. После фильтрации генов по их экспрессии и

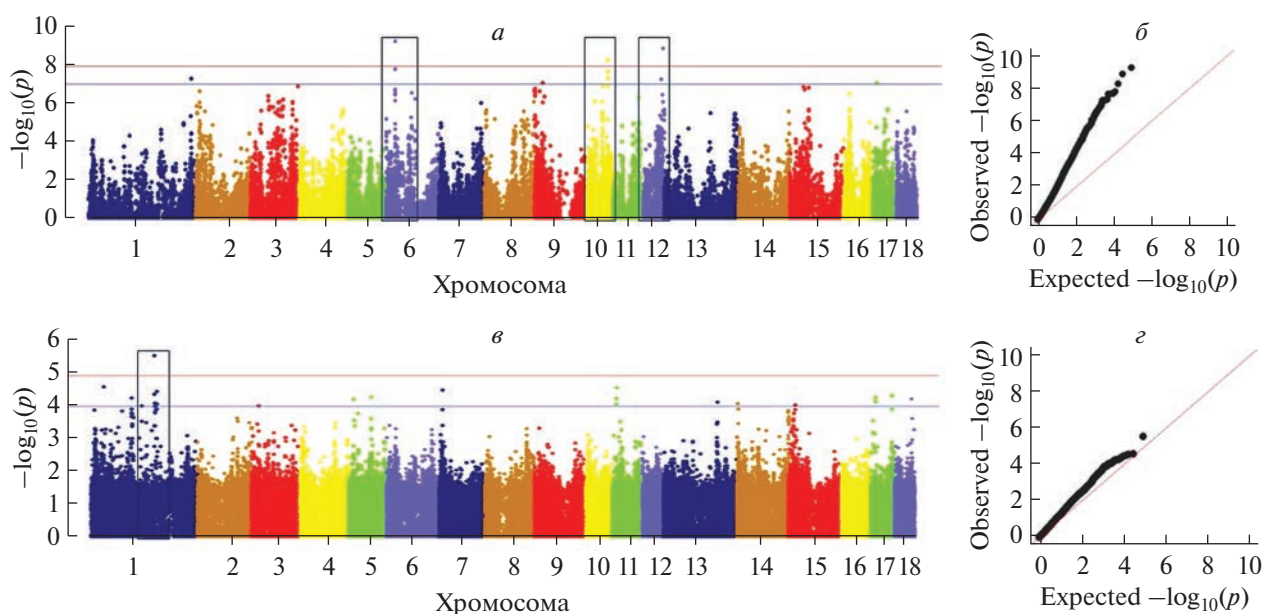


Рис. 2. Манхэттен- и QQ-графики показателя прогнозируемого остаточного корма: *а* – по породе дюрок, *б* – QQ plot по породе дюрок, *в* – по товарным гибридам второго поколения (F2), *з* – QQ plot F2. На графике выделены хромосомные регионы, выходящие за порог полногеномных исследований.

биологическим библиотекам в программе DAVID (табл. 1) было получено 60 генов-кандидатов, имеющих функциональное значение в данном исследовании. Мы описали наиболее значимые гены по биологическим функциям и по которым есть опубликованные исследования.

Ген *EPM2A* (глюканфосфатаза, лафорин), расположен на SSA1. Мутация данного гена, более чем у 95%, вызывает болезнь Лафора – аутосомно-рецессивную прогрессирующую миоклонус-эпилепсию, которая впервые была описана в 1911 г. Прогноз болезни всегда прогрессирующий и фатальный, приводящий к летальному исходу через 5–10 лет после начала клинической картины [19].

Ген *ADGRG6* выявлен с основным влиянием на качественные показатели состава туши у бройлеров, полученных от быстрорастущей белоперой бройлерной линии Б [20]. Ген *SMPDL3A* экспрессируется в первичных макрофагах и секреция из них стимулируется нагрузкой холестерином, лигандами X-рецепторов печени, циклическим АМФ и N-гликозилированным белком. Данный ген обнаруживается в свежеснятом образце крови [21]. Следующий ген *MAN1A1*, расположенный на SSA1, выявленный у помесных свиней, связан с фертильностью самок крупного рогатого скота [27], а по программе DAVID – отвечает за процесс углеводного обмена.

Ген *KCNQ5*, расположенный на SSA1, является важным регулятором транспорта ионов калия и предположительно связан с близорукостью. Известно, что калий играет жизненно важную роль в

поддержании нормальных функций клеток, поэтому любой дисбаланс ионов калия отрицательно сказывается на здоровье людей и животных [28]. Следующий ген поло-подобной киназы 1 (*PLK1*) является ключевым регулятором митотического деления, включая: созревание центросомы, образование биполярного веретена, сегрегацию хроматид, активацию комплекса/циclosомы, способствующего анафазе (APC/C), и выход из митоза. Также известно, что ген *PLK1* контролирует многие немитотические события (такие как репликация, ответ на повреждение ДНК и восстановление контрольной точки G2), динамику хромосом и микротрубочек [29]. Проведенные клинические исследования и полученные по ним данные свидетельствуют о том, что данный ген играет ключевую роль в развитии рака у человека и может стать мишенью для разработки противораковых препаратов [30, 31]. Ген *PLK1* высоко экспрессируется в широком спектре злокачественных опухолей человека, что связано со значительной степенью злокачественности опухоли, и коррелирует с неблагоприятным прогнозом для пациента [30–32].

В наших исследованиях выявлено значительное количество генов, чья экспрессия связана с раковым образованием клеток и иммунитетом в организме. Qiong J. et al. [33] установили, что ген *CHP2* (изоформа 2 белка, гомологичного кальциневрину В) экспрессируется в различных злокачественных клеточных линиях, включая рак яичников. В исследованиях Machuka E.M. et al. [34], проведенных на кенийской породе свиней, отме-

Таблица 1. Структурная и функциональная аннотация генов-кандидатов

Порода	Хромосома	SNP-позиция	Ген-протяженность	Функциональная аннотация
F2	1	ALGA0006973 ^{22491434...22891434}	<i>ERM2A</i> ^{19806363..19893350}	Отвечает за нейродегенерацию, атаксию, миоклонус-эпилепсию и нарушение ответа на поведенческую функцию
			<i>ADGRG6</i> ^{22662009..22802166}	Отвечает за развитие нервной системы
			<i>SMPDL3A</i> ^{39848641...39869184}	Связан с метаболизмом жира в качестве гена-кандидата для специфической регуляции внутримышечного отложения жира
			<i>PKIB</i> ^{39913753..40022971}	Играет решающую роль в уничтожении внутримитохондрий, ухудшает внутриклеточный рост и выживаемость <i>Salmonella</i>
		ALGA0006943 ^{42603383...43003383}	<i>MAN1A</i> ^{42729722..42896290}	Отвечает за процесс углеводного метаболизма
		MARC0033388 ^{52514383...52914383}	<i>KCNQ5</i> ^{52435332..52939705}	Предположительно связан с близорукостью
			<i>ORC3</i> ^{55982132..56044244}	Стимулирует развитие нервной системы
			<i>PLK1</i> ^{22537814..22550901}	Индуктирует хромосомную нестабильность и подавляет контрольные точки клеточного цикла, чтобы стимулировать онкогенез
			<i>CHP2</i> ^{22473241..22478697}	Играет важную роль в патогенезе вирусов
		Дюрок	3	DRGA0004061 ^{30481388...30881388}
<i>CPPED1</i> ^{30362725...30484250}	Регулирует путь фосфатидилинозитол-3-киназы (PI3K), протеинкиназы B (AKT)			
<i>SLX4</i> ^{38646904...38670173}	Выявлен как супрессор опухолей, который активирует специфичность нуклеазы XPF-ERCC1 в репарации перекрестных связей ДНК			
<i>NLRC3</i> ^{38673590...38719188}	Связан с иммуноглобулином и колоректальным раком			
ASGA0014571 ^{42592381...42992381}	<i>WNK2</i> ^{42711388..42837650}			Отвечает за рост и метастазирование опухолей
ALGA0019052 ^{44445487...44845487}	<i>NINJ1</i> ^{42654536..42672830}			Способствует патогенезу нейровоспаления
	<i>MERTK</i> ^{44349377..44483400}			Отвечает за фагоцитарный клиренс апоптотических клеток и антагонизм врожденных иммунных ответов
F2				ALGA0109270 ^{46575456...46975456}
		<i>MGAT4A</i> ^{55560076...55667221}	Избыточная экспрессия необходима в молочной железе свиноматки для улучшения качества клинических исследований	

Таблица 1. Продолжение

Порода	Хромосома	SNP-позиция	Ген-протяженность	Функциональная аннотация
Дюрок	3	ASGA0093631 ^{69529486...69929486}	<i>SFXN5</i> ^{69701021..69833971}	Связан с метаболизмом железа в организме
			<i>RAB11FIP5</i> ^{69656660..69699542}	Связан с дерматологическими и иммунологическими заболеваниями, повреждениями и аномалиями организма
			<i>FGD4</i> ^{41683265..41879197}	Отвечает за клеточное развитие, рост и пролиферацию, морфологию органов
			<i>TNFRSF9</i> ^{68590562..68609784}	Отвечает за иммунный ответ организма
			<i>ERRFI1</i> ^{68667651..68682629}	Способствует развитию альвеол легких, развитию сосудистой сети и эпителия легких
			<i>APTD1</i> ^{70700802..70711719}	Отвечает за рост и метастазирование опухолей
	6	ASGA0100674 ^{70511450...70911450}	<i>KIF1B</i> ^{70495143..70663464}	Отвечает за транспорт везикул, развитие нервной системы
			<i>FES</i> ^{33472947..53485084}	Отвечает за врожденный иммунный ответ
			<i>SLC41A3</i> ^{53672131..53741033}	Переносчик магния
			<i>DCAF16</i> ^{12751199..12763777}	Связан со среднесуточным приростом
			<i>SLIT2</i> ^{14758426..15149515}	Связан с морфогенезом клеток
			<i>PCDH7</i> ^{23459475..23872204}	Дифференциально экспрессировался в первичных эпителиальных клетках толстой кишки
	8	ALGA0049113 ^{23522374...23922374}	<i>ATP10D</i> ^{37407677..37531707}	Регулирует ожирение [22]
			<i>RASL11B</i> ^{39703399..39708174}	Вовлечен в клеточный метаболизм, такой как сборка коллагеновых волокон или расположение цитоскелета
			<i>CPE</i> ^{43568744..43695475}	Связан с диареей у поросят
			<i>MSMO1</i> ^{43726834..43745784}	Связь с репродуктивной способностью свиней и размером помета
			<i>BTC</i> ^{70645721..70697538}	Выявлен в исследованиях по RFI
			<i>PARM1</i> ^{70788134..70892642}	Играет важную роль в кардиомиогенной дифференцировке [23], играет важную роль в развитии мышц
9	ALGA0109134 ^{19517492...19917492}	MARC0068822 ^{42498489...42898489}	<i>SYTL2</i> ^{19617287..19763439}	Отвечает за развитие эпителиальных клеток
			<i>CADM1</i> ^{42644487..42984934}	Связан с апоптозом и воспалительными процессами
			<i>MPZL3</i> ^{45536024..45564210}	Отвечает за аномалии кожи и волосяного покрова
			<i>JAML</i> ^{45506496..45537304}	Кодирует трансмембранные белки Т-клеток [24–26]

Таблица 1. Окончание

Порода	Хромосома	SNP-позиция	Ген протяженности	Функциональная аннотация
Дюрок	9	ALGA0109741 ^{46564464...46964464}	<i>USP2</i> ^{46563469...46590392}	Ген антипатогена
		ASGA0094126 ^{48588375...48988375}	<i>THY1</i> ^{46607022...46612691}	Отвечает за сперматогенез, обогащение используется для лечения бесплодия
		H3GA0025902 ^{55442496...55842496}	<i>SORL1</i> ^{48433350...48653670}	Мутации гена приводят к появлению болезни Альцгеймера в раннем возрасте свиньи
			<i>FLI1</i> ^{55612905...55749784}	Связан с гемопоэзом
			<i>ETS1</i> ^{55377264...55512180}	Отвечает за положительную регуляцию миграции эндотелиальных клеток кровеносных сосудов и положительную регуляцию дифференцировки эритроцитов
		MARC0061961 ^{66448413...66848413}	<i>PLEKHA6</i> ^{64876803...65027455}	Отвечает за формирование переднего окорока
			<i>RAB7B</i> ^{6663536...66666034}	Связан с врожденным иммунным ответом
			<i>SLC41A1</i> ^{66473369...66500367}	Связан с внутриклеточным гомеостазом Mg (IMH), является основной клеточной системой оттока Mg ²⁺
		DRGA0010601 ^{40468370...40868370}	<i>STEAP4</i> ^{68549142...68570350}	Связан с ожирением и резистентностью к инсулину, выявлен у F2
			<i>MAP3K8</i> ^{640714004...40746390}	Отвечает за рост и метастазирование опухолей
10	12	ALGA0067326 ^{54590433...54990433}	<i>RUBCNL</i> ^{20995439...21036429}	Отвечает за мышечную массу
			<i>MILR1</i> ^{14612177...14637946}	Связан с ожирением
			<i>TEX14</i> ^{34815472...34975901}	Отвечает за сперматогенез
		ASGA0069521 ^{43568475...43968475}	<i>GAS7</i> ^{54752436...54974535}	Регулирует аутофагию и гибель клеток
			<i>MLH1</i> ^{21791276...21877789}	Связан с опухолью
			<i>TTC21A</i> ^{225797518...23830334}	Отвечает за сперматогенез
		ALGA0089404 ^{25476449...25876449}	<i>TENM3</i> ^{41818190...44401747}	Отвечает за порок глаза
			<i>PTGER4</i> ^{5670832...25686556}	Формирует нейроактивное лиганд-рецепторное взаимодействие
			<i>SNAI1</i> ^{51546446...51552498}	Формирует клетки ткани
		ASGA0079424 ^{26402504...26802504}	<i>KCND2</i> ^{26139457...26624210}	Связан с серотонинергическими синапсами, транспортом ионов металлов

Примечание. Жирным шрифтом выделены гены, локализованные внутри SNP.

чается, что данный ген играет важную роль в патогенезе вирусов и отвечает за иммунный статус организма. Ген *PCDH7*, расположенный на *SSA7* свиней породы дюрок, представляет собой трансмембранный рецептор и является членом суперсемейства кадгеринов, гиперэкспрессируется при аденокарциноме (железистый рак) легкого. В то же время данный ген замедляет метастазы в головной мозг при раке легкого и молочной железы. Снижение экспрессии гена *PCDH7* было зарегистрировано при колоректальном раке, раке желудка и инвазивном раке мочевого пузыря. Проведенные исследования предполагают контекстно-зависимые функции гена в различных типах опухолей. Учитывая, что ген *PCDH7* является потенциально целевой молекулой на поверхности раковых клеток, дальнейшее изучение его роли в онкогенезе *in vivo* необходимо для оценки терапевтического потенциала ингибирования (подавления активности изучаемых веществ) [35]. Ген *TNFRSF9*, находящийся на *SSA3*, играет роль во многих аспектах врожденного и адаптивного иммунитета, включая иммунологию рака и аутоиммунные заболевания. Примечательно, что экспрессия данного гена может также происходить в неиммунных типах клеток, таких как эндотелиальные, нейронные, астроциты и клетки микроглии [36]. Ген *NLRC3* был идентифицирован при скрининге генов, кодирующих белки, несущие домены, связывающие нуклеотиды. Ген *NLRC3* состоит из N-концевого домена активации и рекрутирования каспаз, центрального домена, связывающего нуклеотиды, и C-концевого домена. Он высоко экспрессируется в иммунных клетках человека и мыши [37].

Достижения в области метаболизма железа, в начале 2000-х гг., пролили свет на многие основные процессы всасывания и распределения железа у млекопитающих [38]. Ген *SFXN1* считается геном-кандидатом наследственной анемии. Накопление митохондриального железа является отличительной чертой сидеробластных анемий, которые обычно возникают в результате дефектов биосинтеза или других путей, приводящих к аномальному использованию эритроидного митохондриального железа [39]. Ген *ARID5A* входит в семейство генов, отвечающих за функции в развитии организма. Экспрессия происходит в тканях [40]. Модификация гликанов важна в фармацевтической промышленности. Особенно сиаловая кислота влияет на биологическую активность и стабильность лекарства. Молоко от свиноматок используется в качестве биореактора для производства различных фармацевтических белков. Ген-кандидат *MGAT4A* сильно экспрессируется в печени и молочной железе свиней, слабо — в тонком кишечнике, желудке и мочевом пузыре. Анализируя полученные данные, выяснилось, что избыточная экспрессия данного гена необходима в

молочной железе свиньи для улучшения качества проводимых клинических исследований [41]. Исследования Chen Y. [42] на африканских и шаанбейских белых кашемировых козах показали, что ген *SNX29* (сортирующий нексин 29) связан с качеством мяса и признаками роста.

Исследования de Baaij J. [43] охарактеризовали функциональную роль гена *SLC41A3* в гомеостазе двухвалентного магния у мышей. Регуляция баланса Mg^{2+} в организме происходит в дистальных извитых канальцах, где трансцеллюлярная реабсорбция определяет окончательную экскрецию Mg^{2+} с выделением жидкости из мочевого пузыря. Способность кишечника к абсорбции Mg^{2+} измеряли с использованием стабильного изотопа у мышей, получавших диету с низким содержанием магния. Ген *SLC41A3* был установлен как новый фактор для формирования Mg^{2+} . Следующий ген *DCAF16* ранее был получен в исследованиях Zhang W. [44] методом мультистратегического GWAS, который показал, что область гена является сопряженным локусом среднесуточного прироста у популяции крупного рогатого скота. Ген *SLIT2* интересен тем, что его мутации приводят к врожденным аномалиям почек и мочеточников, необходимо изучение данного гена для нормализации развития почек [45]. Ген *ERRF1* является механизмом раннего ответа организма на агрессивную внешнюю среду, и его продукт легко индуцируется широким спектром внеклеточных и внутриклеточных стимулов [46]. Ген *BTC* был ранее выявлен в исследованиях по прогнозируемому остаточному потреблению корма у свиней и имеет биологический процесс (GO:0043066), а именно отрицательную регуляцию апоптоза — регулируемый процесс клеточной гибели в организме [47]. Заслуживает внимания и ген *SORL1*. В исследованиях, проведенных в Орхусском университете, мутация данного гена была выявлена у клонированных модельных свиней. У людей данная мутация обнаруживается в 2–3% всех случаев раннего начала болезни Альцгеймера [48]. Еще один ген *JAML* связан с клинической картиной рассеянного склероза, который представляет собой иммуноопосредованное заболевание центральной нервной системы (ЦНС) и характеризуется множественными очагами лейкоцитарной инфильтрации, демиелинизации и повреждения аксонов. Данный ген также экспрессируется эндотелием, где он гомодимеризуется в цис-положении [25]. В исследованиях Sironen A. et al. [49], проведенных на популяции финских йоркширских свиней методом полногеномного анализа на дефект сперматогенеза, обнаружен ген *TEX14*. Секвенирование полученного гена-кандидата выявило вставку 51 пн в экзоне 27, которая вызвала дифференциальный сплайсинг экзона и создала кодон остановки преждевременной трансляции.

Одним из немаловажных генов является *PARM1*, отвечающий за развитие длиннейшей мышцы спины у свиней [50]. Ген *SYTL2* отвечает за развитие эпителиальных клеток [51], а ген *CADM1* отвечает за макрофаги, находящиеся в костном мозге [52]. Ген *ATP10D*, по исследованиям Ramdas M. et al. [22], регулирует ожирение и участвует в процессе превращения эндоплазматического ретикулаума в церамид Гольджи [53]. Ген *MSMO1*, полученный в исследованиях Ran X. [54], коррелирует с созревaniem яйцеклеток у сяньских свиней и со значительным размером помета. Высказано предположение, что данный ген, в совокупности с остальными, может играть важную роль в увеличении размера помета за счет повышения уровня поступления стероидных и пептидных гормонов через яичники, облегчения овуляции ооцитов и оплодотворения *in vivo*. Ген-кандидат *PLEKHA6* отвечает за вес переднего окорока [55], а ген *ETSI* связан с гликолитическим потенциалом [56].

Сохранение энергетического гомеостаза в присутствии стрессоров, таких как провоспалительные цитокины и перегрузка питательными веществами, имеет решающее значение для поддержания нормальной клеточной функции. Считается, что шеститрансмембранный эпителиальный антиген простаты 4 (*STEAP4*), участвующий в гомеостазе железа и меди, играет потенциально важную роль в клеточном ответе на воспалительный стресс. Полногеномные ассоциативные исследования показали различные мутации в гене *STEAP4*, которые влияют на развитие метаболических нарушений, таких как ожирение, метаболический синдром и диабет II типа. Экспрессия данного гена модулируется воспалительными цитокинами, гормонами и другими индикаторами клеточного стресса, он может защищать клетки от повреждений, помогая поддерживать нормальную метаболическую функцию. Ген *STEAP4* особенно актуален для метаболически ориентированных клеток, таких как адипоциты, гепатоциты и клетки поджелудочной железы [57].

Экспрессия выявленных генов

Для изучения данного раздела исследований были выбраны 13 генов, имеющих биологические функции, значимую экспрессию и локализованные внутри изучаемых SNP (табл. 2).

Рецептор гена *ADGRG6*, связанный с G-белком адгезии, играет важную роль в развитии, поддержании и восстановлении миелина периферической нервной системы [58]. Мутация в данном гене вызывает дефектную миелинизацию (процесс образования билипидного миелинового слоя вокруг аксона) периферической нервной системы человека [59]. Экспрессия гена *ADGRG6* происходит в легких и селезенке. Исследования Hong C. [60] сообщают об открытии генетического и эпи-

генетического направления гена *WINK2* в глиальных опухолях взрослых людей, который играет потенциальную роль в качестве супрессора опухолей. Krona C. в 2004 г. [61] охарактеризовал ген *APITD1* как один из белков, участвующих в регуляции клеточного цикла. Снижение уровня транскрипции гена *APITD1* препятствует клетке организма осуществлять нормальную регуляцию во время онкогенеза.

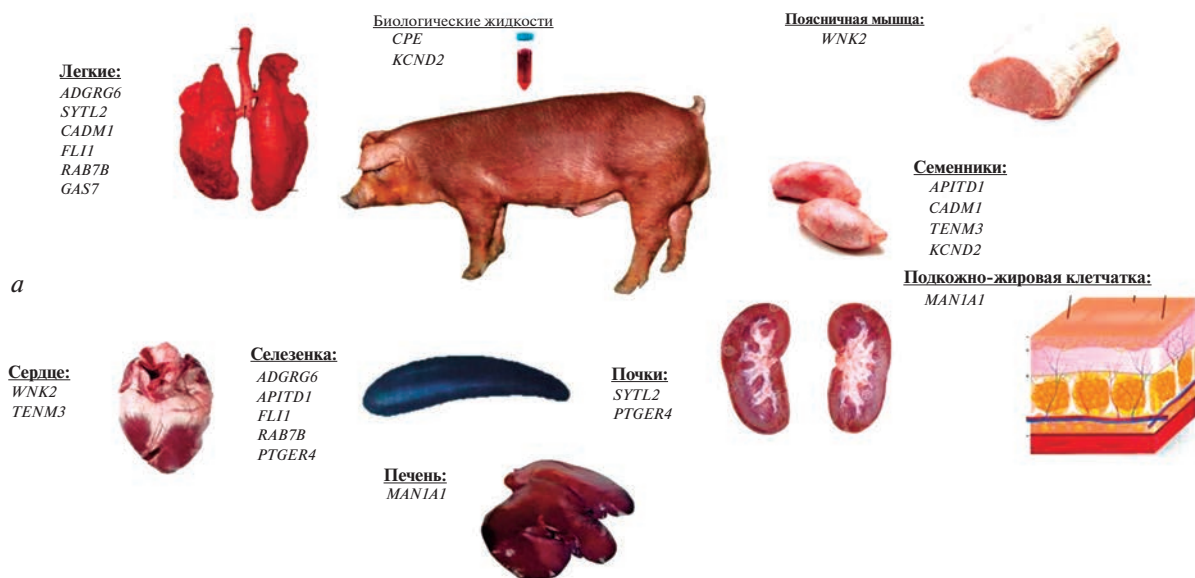
Ген *CPE* (carboxypeptidase E) кодирует многофункциональный периферический белок, который играет несколько неферментативных ролей в эндокринных тканях и центральной нервной системе, с наибольшей концентрацией в головном мозге и гипофизе [62]. Более того, данный фермент обрабатывает прогормоны, которые участвуют в биосинтезе многочисленных пептидных гормонов и нейротрансмиттеров [63], включая процесс превращения проинсулина в инсулин [64]. Изучаемый ген консервативен у разных видов [65], а полиморфизмы в определенных регионах связаны с диабетом II типа и ожирением у людей, ожирением и бесплодием у мышей и свиней [66]. Также он связан с качеством туши и мяса у разных пород крупного рогатого скота [62]. На экспрессию гена *CPE* в головном мозге напрямую влияют различные физиологические или экологические стрессовые условия, которым подвергаются животные, что приводит к разным уровням дегенерации нейронов гиппокампа, коры головного мозга, миндалин и выживанию нейронов [64]. Кроме того, ген был связан с недостаточной усвояемостью обучения и концентрации памяти [65], моделью настроения и эмоциональными реакциями [66] у мышей.

Ген *SYTL2* кодирует синаптотагмин-подобный белок (SLP), содержащий домен C2. Наблюдались множественные альтернативно приправленные варианты транскриптов, кодирующие различные изоформы. Экспрессируется, по данным NCBI, в почках и легких. Ген *CADM1*, также известный как ген-супрессор опухолевого роста 1, экспрессируется во всех миелоидных клетках. На функциональном уровне делеция *CADM1* в клетках мышинной линии Lineage⁻ Sca1⁺ Kit⁺ модифицирует соотношение лимфоидных и миелоидных клеток в костном мозге (хотя и не изменяет их потенциала восстановления кроветворения) после сингенной трансплантации (пересадка от донора — идентичного близнеца). Вместе с этим ген *CADM1* может быть важен в физиопатологии, расширяя свою роль в качестве гена-супрессора от солидных до гемопоэтических злокачественных новообразований [67]. Выявленный ген-кандидат *FLI1* у людей при слиянии с геном *EWSR1* характеризуется прогрессирующей саркомой Юинга [68], а сайты связывания гена способствуют дифференцировке гемопоэтических стволовых клеток в тромбоциты и показывают сильное обогащение

Таблица 2. Гены-кандидаты, локализованные внутри SNP, и их экспрессия

Порода	Хромосома	SNP	Ген-кандидат	Биологическая функция	Экспрессия*
F2	1	ALGA0006973	<i>ADGRG6</i> ^{22662009..22802166}	Развитие нервной системы, сперматогенез	В легких (RPKM 14.1), селезенке (RPKM 11.8)
F2	1	ALGA0006943	<i>MANIA1</i> ^{42729722..42896290}	Процесс углеводного метаболизма	В печени (RPKM 17.4), подкожно-жировой клетчатке (RPKM 17.3)
Дюрок	3	ASGA0014571	<i>WNK2</i> ^{42711388..42837650}	Отвечает за рост и метастазы опухолей	В поясничной мышце (RPKM 8.2), сердце (RPKM 8.2)
Дюрок	6	ASGA0100674	<i>APITD1</i> ^{70700802..70711719}	Отвечает за рост и метастазы опухолей	В семенниках (RPKM 10.0), селезенке (RPKM 2.3)
Дюрок	8	ALGA0049113	<i>CPE</i> ^{43568744..43695475}	Связан с диареей у поросят	В биологических жидкостях (RPKM 51.4), легких (RPKM 27.7)
Дюрок	9	ALGA0109134	<i>SYTL2</i> ^{19617287..19763439}	Развитие эпителиальных клеток	В почках (RPKM 10.9), легких (RPKM 3.3)
Дюрок	9	MARC0068822	<i>CADM1</i> ^{42644487..42984934}	Связан с апоптозом (клеточная гибель) и воспалительными процессами в организме	В легких (RPKM 51.9), семенниках (RPKM 16.1)
Дюрок	9	H3GA0025902	<i>FLII</i> ^{55612905..55749784}	Связан с гемопоэзом (кроветворение)	В селезенке (RPKM 27.3), легких (RPKM 27.1)
Дюрок	9	MARC0061961	<i>RAB7B</i> ^{66635636..66666034}	Врожденный иммунный ответ	В селезенке (RPKM 7.3), легких (RPKM 3.6)
Дюрок	12	ALGA0067326	<i>GAS7</i> ^{4752436..54974535}	Регулирует аутофагию и гибель клеток	В семенниках (RPKM 15.1), легких (RPKM 8.5)
Дюрок	15	ASGA0069521	<i>TENM3</i> ^{41818190..44401747}	Отвечает за недостаток экстерьера: порок глаза	В сердце (RPKM 2.6), семенниках (RPKM 2.5)
Дюрок	16	ALGA0089404	<i>PTGER4</i> ^{25670832..25686556}	Лиганд-рецепторное взаимодействие (иммунитет)	В селезенке (RPKM 10.6), почках (RPKM 4.5)
Дюрок	18	ASGA0079424	<i>KCND2</i> ^{26139457..26624210}	Связан с серотонинергическими синапсами, транспортом ионов металлов	В биологических жидкостях (RPKM 1.4), семенниках (RPKM 0.7)

Таблица 2. Окончание



Примечание. * Протоколы секвенирования по экспрессии генов. *a* — экспрессия выявленных генов и их локализация в организме свиньи (*Sus scrofa*).

($P = 1.5 \times 10^{-6}$), что в итоге приводит к мозаичной потере хромосомы Y [69]. Ген *RAB7B* связан с признаками жирных кислот у крупного рогатого скота породы ангус [70].

Ген *TENM3* (трансмембранный белок 3 теней-рин) кодирует большой трансмембранный белок, участвующий в развитии нервной системы, регулируя установление надлежащих связей в нервной системе [71–73]. Было обнаружено, что он играет роль в развитии человеческого глаза, регулируя формирование ипсилатерального отображения сетчатки как на дорсальное латеральное коленчатое ядро, так и на верхнее двухолмие [74–76]. Гомозиготный нулевой вариант впервые был зарегистрирован в кровнородственной семье из Саудовской Аравии с несиндромальным двусторонним колобома-тотным микрофтальмом [77, 78]. Ген *PTGER4* (рецептор простагландина E4) участвует в иммунной регуляции и костном метаболизме [79].

Ген *KCND2* кодирует потенциал-зависимый калиевый канал, экспрессия происходит в тканях мозга мышей. Данный ген относится к онтологиям GO “регуляция мембранного потенциала”, “позитивная регуляция ионного транспорта”, “транспортный комплекс” и “ГАМК-эргический синапс”, которые сильно обогащены, связаны с нервной системой и поведением. Было показано, что ген *KCND2* важен для регуляции синаптической пластичности [80, 81], а нокаутные мыши проявляют повышенную чувствительность к механическим стимулам [81]. Мутации гена у людей

связаны с аутизмом [82, 83]. Все эти данные свидетельствуют о том, что он необходим для функционирования нервной системы и участвует в изменениях поведения в результате одомашнивания мышей [84].

Кластеризация генов

Приложение веб-сервера ToppCluster использует мощный анализ обогащения и базовую среду данных для сравнительной оценки нескольких списков генов. Он генерирует тепловые карты или сети подключения, выявляющие функциональные особенности, общие или специфичные для нескольких списков генов. ToppCluster использует гипергеометрические тесты для получения *P*-значений обогащения. В наших исследованиях были взяты библиотеки биологической, функциональной и молекулярной дискуссии и данные, опубликованные в PubMed (рис. 3).

Выявлены две биологические функции, влияющие на общую нормализацию и на положительную регуляцию секреции слизи в организме, гены-кандидаты — *PTGER4* (prostaglandin E receptor 4) и *SYTL2* (synaptotagmin like 2). К молекулярной функции относится связывание белка семейства нейрексинов, из-за которого происходят сложности при дрессировке. Семейство нейрексинов представляет собой группу трансмембранных белков, обеспечивающих взаимодействие нервных клеток, их дифференцировку в коре больших полу-

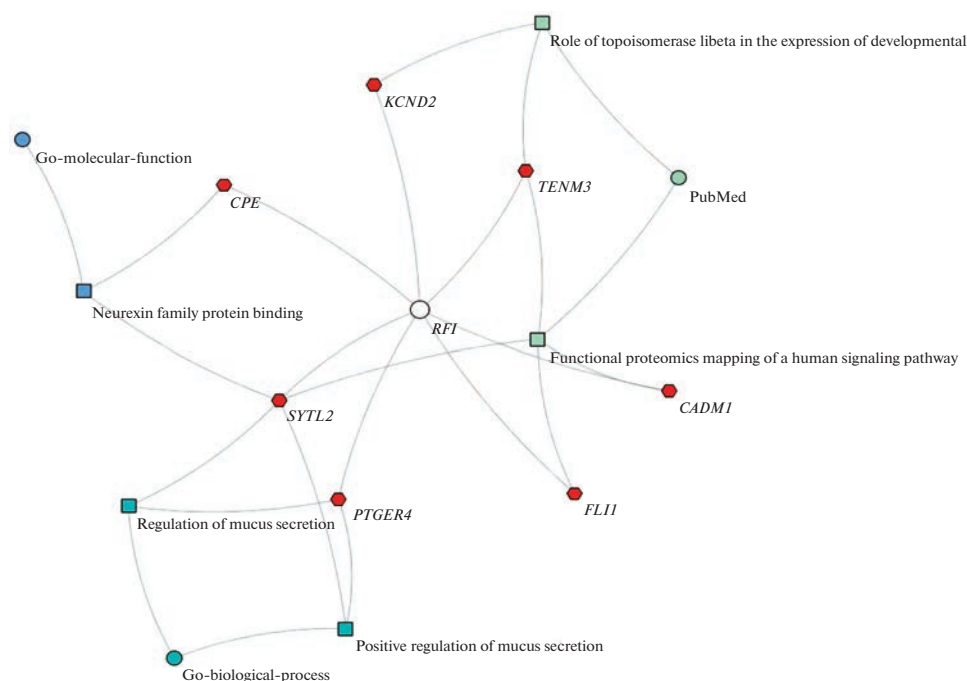


Рис. 3. Сравнительная кластеризация обогащения и сетевого анализа биологических систем изучаемого списка генов. RFI – прогнозируемое остаточное потребление корма.

шарий головного мозга и экспрессию в ионных каналах, тем самым участвуя в стабилизации аксональной проводимости и поддержании потенциалов действия [85]. Нейрексины представляют собой высокополиморфное семейство синаптических рецепторов, диверсифицированных за счет обширного альтернативного сплайсинга [86]. Гены-кандидаты, входящие в данную функцию, по нашим исследованиям – *CPE* (carboxypeptidase E) и *SYTL2* (synaptotagmin like 2).

По исследованиям других авторов (библиотека PubMed) выделены следующие гены:

– *KCND2* и *TENM3*, у которых отмечена роль экспрессии топоизомеразы *libeta* в развитии организма [87].

– *TENM3*, *CADM1*, *FLII* и *SYTL2*, имеющие связь функционального протеомного картирования сигнального пути у человека [88].

Анализируя полученные данные, мы видим, что ген *SYTL2* имеет взаимосвязь между кластеризацией обогащения и сетевым анализом биологических систем. У свиней выявленный ген не имел апробации, но у людей мутация гена *SYTL2* связана с клинической ассоциацией и генетической основой колоректального рака [89], способствует метастатическому потенциалу рака яичников [90], имеет влияние на остеосаркому позвоночника [91]. В связи с высоким влиянием мутации гена *SYTL2* на злокачественные новообразования различного

спектра у человека выявление его у свиней позволит планировать создание модели для клинических исследований препаратов и геномного редактирования.

Для достоверного понимания практической значимости в дальнейшем будет проведен экспериментальный анализ по выявленным генам, проанализирован механизм их экспрессии, будут произведены апробации и скрининг полученных тест-систем.

Исследования проведены при поддержке Министерства науки и высшего образования Российской Федерации, регистрационный номер темы Государственного задания № FGGN-2022-0007 и при финансовой поддержке Министерства науки и высшего образования Российской Федерации, грант № 075-15-2021-1037 (внутренний № 15. БРК.21.0001).

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. Koch R.M., Swiger L.A., Chambers D., Gregory K.E. Efficiency of feed use in beef cattle // J. Anim. Sci. 1963.

- V. 22. № 2. P. 486–494.
<https://doi.org/10.2527/jas1963.222486x>
2. *Vigors S., Sweeney T., O'Shea C.J. et al.* Pigs that are divergent in feed efficiency, differ in intestinal enzyme and nutrient transporter gene expression, nutrient digestibility and microbial activity // *Animal*. 2016. V. 10. № 11. P. 1848–1855.
<https://doi.org/10.1017/S1751731116000847>
 3. *Patience J.F., Rossoni-Serao M.C., Gutierrez N.A.* A review of feed efficiency in swine: biology and application // *J. Anim. Sci. Biotechnol.* 2015. V. 6. № 1. P. 33.
<https://doi.org/10.1186/s40104-015-0031-2>
 4. *Rakhshandeh A., Dekkers J.C., Kerr B.J. et al.* Effect of immune system stimulation and divergent selection for residual feed intake on digestive capacity of the small intestine in growing pigs // *J. Anim. Sci.* 2012. V. 90. № 4. P. 233–255.
<https://doi.org/10.2527/jas.53976>
 5. *Grubbs J.K., Huff-Lonergan E., Gabler N.K. et al.* Liver and skeletal muscle mitochondria proteomes are altered in pigs divergently selected for residual feed intake // *J. Anim. Sci.* 2014. V. 92. № 5. P. 1995–2007.
<https://doi.org/10.2527/jas.2013-7391>
 6. *Fu L., Xu Y., Hou Y. et al.* Proteomic analysis indicates that mitochondrial energy metabolism in skeletal muscle tissue is negatively correlated with feed efficiency in pigs // *Sci. Rep.* 2017. № 7.
<https://doi.org/10.1038/srep45291>
 7. *Jing L., Hou Y., Wu H. et al.* Transcriptome analysis of mRNA and miRNA in skeletal muscle indicates an important network for differential residual feed intake in pigs // *Sci. Rep.* 2015. № 5.
<https://doi.org/10.1038/srep11953>
 8. *Vigors S., O'Doherty J.V., Kelly A.K. et al.* The effect of divergence in feed efficiency on the intestinal microbiota and the intestinal immune response in both unchallenged and lipopolysaccharide challenged Ileal and colonic explants // *PLoS One*. 2016. V. 11. № 2.
<https://doi.org/10.1371/journal.pone.0148145>
 9. *Mani V., Harris A.J., Keating A.F. et al.* Intestinal integrity, endotoxin transport and detoxification in pigs divergently selected for residual feed intake // *J. Anim. Sci.* 2013. V. 91. № 5. P. 2141–2150.
<https://doi.org/10.2527/jas.2012-6053>
 10. *Hayes B.J., Lewin H.A., Goddard M.E.* The future of livestock breeding: Genomic selection for efficiency, reduced emissions intensity, and adaptation // *Trends Genet.* 2013. V. 29. № 4. P. 206–214.
<https://doi.org/10.1016/j.tig.2012.11.009>
 11. *Saintilan R., Mérour I., Brossard L. et al.* Genetics of residual feed intake in growing pigs: Relationships with production traits, and nitrogen and phosphorus excretion traits // *J. Anim. Sci.* 2013. V. 91. № 6. P. 2542–2554.
 12. *Zhang C., Kemp R.A., Stothard P. et al.* Genomic evaluation of feed efficiency component traits in Duroc pigs using 80K, 650K and whole-genome sequence variants // *Genet. Sel. Evol.* 2018. V. 50. № 1.
<https://doi.org/10.1186/s12711-018-0387-9>
 13. *Onteru S.K., Gorbach D.M., Young J.M. et al.* Whole genome association studies of residual feed intake and related traits in the pig // *PLoS One*. 2013. V. 8. № 6.
<https://doi.org/10.1371/journal.pone.0061756>
 14. *Do D.N., Ostersen T., Strathe A.B. et al.* Genome-wide association and systems genetic analyses of residual feed intake, daily feed consumption, backfat and weight gain in pigs // *BMC Genet.* 2014. V. 15. № 27.
<https://doi.org/10.1186/1471-2156-15-27>
 15. *Reyer H., Oster M., Magowan E. et al.* Strategies towards improved feed efficiency in pigs comprise molecular shifts in hepatic lipid and carbohydrate metabolism // *Int. J. Mol. Sci.* 2017. V. 18. № 8.
<https://doi.org/10.3390/ijms18081674>
 16. Белоус А.А., Требунских Е.А., Сермягин А.А., Зиновьева Н.А. Методические рекомендации по расчету и использованию в селекции свиней показателя прогнозируемого остаточного потребления корма (RFI). Дубровицы: ФГБНУ ФИЦ ВИЖ им. Л.К. Эрнста, 2022. 32 с.
 17. *An M., Zhou G., Li Y. et al.* Multi-breed genetic parameters and genome-wide association studies for mortality rate at birth in pigs // *Res. Square*. 2021.
<https://doi.org/10.21203/rs.3.rs-146253/v1>
 18. *Mignon G.L., Iannuccelli N., Robic A. et al.* Fine mapping of quantitative trait loci for androstenedione and skatole levels in pig // *Res. Gate*. 2011. <https://hal.inrae.fr/hal-02816807>
 19. *Turnbull J., Tiberia E., Striano P. et al.* Lafora disease // *Epileptic Disord.* 2016. V. 18. № S2. P. 38–62.
<https://doi.org/10.1684/epd.2016.0842>
 20. *Yang X., Sun J., Zhao G. et al.* Identification of major loci and candidate genes for meat production-related traits in broilers // *Front. Genet.* 2021. V. 12.
<https://doi.org/10.3389/fgene.2021.645107>
 21. *Traini M., Quinn C.M., Sandoval C. et al.* Sphingomyelin phosphodiesterase acid-like 3A (*SMPDL3A*) is a novel nucleotide phosphodiesterase regulated by cholesterol in human macrophages // *J. Biol. Chem.* 2014. V. 21. № 289. P. 32895–32913.
<https://doi.org/10.1074/jbc.M114.612341>
 22. *Ramdas M., Harel C., Armoni M., Karnieli E.* AHNK KO mice are protected from diet-induced obesity but are glucose intolerant // *Horm. Metab. Res.* 2015. V. 47. P. 265–272.
<https://doi.org/10.1055/s-0034-1387736>
 23. *Nakanishi N., Takahashi T., Ogata T. et al.* PARM-1 promotes cardiomyogenic differentiation through regulating the BMP/Smad signaling pathway // *Biochem. Biophys. Res. Commun.* 2012. V. 428. P. 500–505.
<https://doi.org/10.1016/j.bbrc.2012.10.078>
 24. *Zhu Y., Wang D., Wang F. et al.* A comprehensive analysis of GATA-1-regulated miRNAs reveals miR-23a to be a positive modulator of erythropoiesis // *Nucl. Acids*

- Res. 2013. V. 41. № 7. P. 4129–4143.
<https://doi.org/10.1093/nar/gkt093>
25. *Alvarez J.I., Kébir H., Cheslow L. et al.* JAML mediates monocyte and CD8 T-cell migration across the brain endothelium // *Ann. Clin. Transl. Neurol.* 2015. V. 2. № 11. P. 1032–1037.
<https://doi.org/10.1002/acn3.255>
 26. *Forbes M.K., Wright A.G.C., Markon K.E., Krueger R.F.* Evidence that psychopathology symptom networks have limited replicability // *J. Abnormal Psychol.* 2017. V. 126. № 7. P. 969–988.
<https://doi.org/10.1037/abn0000276>
 27. *Tarekegn A.A., Mengistu M.Y., Mirach T.H.* Health professionals' willingness to pay and associated factors for cervical cancer screening program at College of Medicine and Health Sciences, University of Gondar, Northwest Ethiopia // *PLoS One.* 2019. V. 14. № 4.
<https://doi.org/10.1371/journal.pone.0215904>
 28. *Zhang Y., Wildsoet C.F.* RPE and choroid mechanisms underlying ocular growth and myopia // *Prog. Mol. Biol. Transl. Sci.* 2015. V. 134. P. 221–240.
<https://doi.org/10.1016/bs.pmbts.2015.06.014>
 29. *Zitouni S., Nabais C., Jana S.C. et al.* Polo-like kinases: Structural variations lead to multiple functions // *Nat. Rev. Mol. Cell. Biol.* 2014. V. 15. P. 433–452.
<https://doi.org/10.1038/nrm3819>
 30. *Liu Z., Sun Q., Wang X.* PLK1, A potential target for cancer therapy // *Transl. Oncol.* 2017. V. 10. P. 22–32.
<https://doi.org/10.1016/j.tranon.2016.10.003>
 31. *Strebhardt K., Ullrich A.* Targeting polo-like kinase 1 for cancer therapy // *Nat. Rev. Cancer.* 2006. V. 6. P. 321–330.
<https://doi.org/10.1038/nrc1841>
 32. *Gheghiani L., Wang L., Zhang Y. et al.* PLK1 induces chromosomal instability and overrides cell-cycle checkpoints to drive tumorigenesis // *Cancer Res.* 2021. V. 81. № 5. P. 1293–1307.
<https://doi.org/10.1158/0008-5472.CAN-20-1377>
 33. *Qiong J., Beihua K., Xingsheng Y. et al.* Overexpression of CHP2 enhances tumor cell growth, invasion and metastasis in ovarian cancer // *In Vivo.* 2007. V. 21. № 4. P. 593–598.
 34. *Machuka E.M., Juma J., Muigai A.W.T. et al.* Transcriptome profile of spleen tissues from locally-adapted Kenyan pigs (*Sus scrofa*) experimentally infected with three varying doses of a highly virulent African swine fever virus genotype IX isolate: Ken12/busia.1 (ken-1033) // *BMC Genomics.* 2022. V. 23. № 522.
<https://doi.org/10.1186/s12864-022-08754-8>
 35. *Zhou X., Padanad M.S., Evers B.M. et al.* Modulation of mutant krasg12d -driven lung tumorigenesis in vivo by gain or loss of PCDH7 function // *Mol. Cancer Res.* 2019. V. 17. № 2. P. 594–603.
<https://doi.org/10.1158/1541-7786.MCR-18-0739>
 36. *Eckstrum K., Bany B.M.* Tumor necrosis factor receptor subfamily 9 (Tnfrsf9) gene is expressed in distinct cell populations in mouse uterus and conceptus during im-plantation period of pregnancy // *Cell Tissue Res.* 2011. № 344. P. 567–576.
<https://doi.org/10.1007/s00441-011-1171-0>
 37. *Karki R., Malireddi R.K.S., Zhu Q., Kanneganti T.D.* NLRC3 regulates cellular proliferation and apoptosis to attenuate the development of colorectal cancer // *Cell Cycle.* 2017. V. 16. № 13. P. 1243–1251.
<https://doi.org/10.1080/15384101.2017.1317414>
 38. *Andrews N.C.* Iron homeostasis: Insights from genetics and animal models // *Nat. Rev. Genet.* 2000. № 3. P. 208–217.
<https://doi.org/10.1038/35042073>
 39. *Fleming M.D., Campagna D.R., Haslett J.N. et al.* A mutation in a mitochondrial transmembrane protein is responsible for the pleiotropic hematological and skeletal phenotype of flexed-tail (f/f) mice // *Genes Dev.* 2001. V. 15. № 6. P. 652–657.
<https://doi.org/10.1101/gad.873001>
 40. *Núñez Y., Radović Č., Savić R. et al.* Muscle transcriptome analysis reveals molecular pathways related to oxidative phosphorylation, antioxidant defense, fatness and growth in mangalitsa and moravka pigs // *Animals.* 2021. V. 11.
<https://doi.org/10.3390/ani11030844>
 41. *Ji-Youn K., Hwang H., Hak-Jae C. et al.* Identification and functional analysis of pig β -1,4-N-Acetylglucosaminyltransferase A (MGAT4A) // *J. Life Sci.* 2016. № 26(3). P. 275–281.
<https://doi.org/10.5352/JLS.2016.26.3.275>
 42. *Chen Y., Yang L., Lin X. et al.* Effects of genetic variation of the sorting nexin 29 (SNX29) gene on growth traits of xiangdong black goat // *Animals.* 2022. V. 12.
<https://doi.org/10.3390/ani12243461>
 43. *de Baaij J., Arjona F., van den Brand M. et al.* Identification of SLC41A3 as a novel player in magnesium homeostasis // *Sci. Rep.* 2016. № 6.
<https://doi.org/10.1038/srep28565>
 44. *Zhang W., Li J., Guo Y. et al.* Multi-strategy genome-wide association studies identify the DCAF16-NCAPG region as a susceptibility locus for average daily gain in cattle // *Sci. Rep.* 2016. № 6.
<https://doi.org/10.1038/srep38073>
 45. *Li J., Geraldo L.H., Dubrac A. et al.* Slit2-robo signaling promotes glomerular vascularization and nephron development // *J. Am. Soc. Nephrol.* 2021. V. 32. № 9. P. 2255–2272.
<https://doi.org/10.1681/ASN.2020111640>
 46. *Xu D., Li C.* Gene 33/Mig6/ERRFI1, an adapter protein with complex functions in cell biology and human diseases // *Cells.* 2021. V. 10. № 1574.
<https://doi.org/10.3390/cells10071574>
 47. *Messad F., Louveau I., Koffi B. et al.* Investigation of muscle transcriptomes using gradient boosting machine learning identifies molecular predictors of feed efficiency in growing pigs // *BMC Genomics.* 2019. V. 20. № 659.
<https://doi.org/10.1186/s12864-019-6010-9>

48. Andersen O.M., Bogh N., Landau A.M. et al. A genetically modified minipig model for Alzheimer's disease with SORL1 haploinsufficiency // *Cell Rep. Med.* 2022. V. 3. № 9.
<https://doi.org/10.1016/j.xcrm.2022.100740>
49. Sironen A., Uimari P., Venhoranta H. et al. An exonic insertion within *Tex14* gene causes spermatogenic arrest in pigs // *BMC Genomics*. 2011. V. 12.
<https://doi.org/10.1186/1471-2164-12-591>
50. Xue Y., Li C., Duan D. et al. Genome-wide association studies for growth-related traits in a crossbreed pig population // *Anim. Genet.* 2021. V. 52. P. 217–222.
<https://doi.org/10.1111/age.13032>
51. Jaing C., Rowland R.R., Allen J.E. et al. Gene expression analysis of whole blood RNA from pigs infected with low and high pathogenic African swine fever viruses // *Sci. Rep.* 2017. V. 7. № 1.
<https://doi.org/10.1038/s41598-017-10186-4>
52. Kapetanovic R., Fairbairn L., Downing A. et al. The impact of breed and tissue compartment on the response of pig macrophages to lipopolysaccharide // *BMC Genomics*. 2013. V. 14.
<https://doi.org/10.1186/1471-2164-14-581>
53. Zhang L., Huang Y., Wang M. et al. Development and genome sequencing of a laboratory-inbred miniature pig facilitates study of human diabetic disease // *Science*. 2019. V. 19. P. 162–176.
<https://doi.org/10.1016/j.isci.2019.07.025>
54. Ran X., Hu F., Mao N. et al. Differences in gene expression and variable splicing events of ovaries between large and small litter size in Chinese Xiang pigs // *Porc. Health Manag.* 2021. V. 7. № 52.
<https://doi.org/10.1186/s40813-021-00226-x>
55. Diao S., Huang S., Chen Z. et al. Genome-wide signatures of selection detection in three south china indigenous pigs // *Genes*. 2019. V. 10.
<https://doi.org/10.3390/genes10050346>
56. Liu X., Zhang J., Xiong X. et al. An integrative analysis of transcriptome and GWAS data to identify potential candidate genes influencing meat quality traits in pigs // *Front. Genet.* 2021. V. 12.
<https://doi.org/10.3389/fgene.2021.748070>
57. Scarl R.T., Lawrence C.M., Gordon H.M., Nunemaker C.S. STEAP4: Its emerging role in metabolism and homeostasis of cellular iron and copper // *J. Endocrinol.* 2017. V. 234. № 3. P. R123–R134.
<https://doi.org/10.1530/JOE-16-0594>
58. Henzi A., Senatore A., Lakkaraju A.K. et al. Soluble dimeric prion protein ligand activates Adgrg6 receptor but does not rescue early signs of demyelination in PrP-deficient mice // *PLoS One*. 2020. V. 15. № 11.
<https://doi.org/10.1371/journal.pone.0242137>
59. Torregrosa-Carrión R., Piñeiro-Sabaris R., Siguero-Álvarez M. et al. Adhesion G protein-coupled receptor Gpr126/Adgrg6 is essential for placental development // *Sci. Adv.* 2021. V. 7. № 46.
<https://doi.org/10.1126/sciadv.abj5445>
60. Hong C., Moorefield K.S., Jun P. et al. Epigenome scans and cancer genome sequencing converge on WNK2, a kinase-independent suppressor of cell growth // *Proc. Natl Acad. Sci. USA*. 2007. V. 104. № 26.
<https://doi.org/10.1073/pnas.0700683104>
61. Krona C., Ejeskär K., Carén H. et al. A novel 1p36.2 located gene, *APITD1*, with tumour-suppressive properties and a putative p53-binding domain, shows low expression in neuroblastoma tumours // *Br. J. Cancer*. 2004. V. 91. P. 1119–1130.
<https://doi.org/10.1038/sj.bjc.6602083>
62. Shin S.C., Chung E.R. Association of SNP marker in the leptin gene with carcass and meat quality traits in Korean cattle // *Asian-Aust. J. Anim. Sci.* 2007. V. 20. P. 1–6.
<https://doi.org/10.5713/ajas.2007.1>
63. Cawley N.X., Wetsel W.C., Murthy S.R. et al. New roles of carboxypeptidase E in endocrine and neuronal function and cancer // *Endocr. Rev.* 2012. V. 33. P. 216–253.
<https://doi.org/10.1210/er.2011-1039>
64. Wang J., Zhang Y., Yang Z. et al. Association of human carboxypeptidase E exon5 gene polymorphisms with angiographical characteristics of coronary atherosclerosis in a Chinese population // *Acta Pharmacol. Sin.* 2008. № 29. P. 736–744.
<https://doi.org/10.1111/j.1745-7254.2008.00798.x>
65. Koshimizu H., Senatorov V., Loh Y.P., Gozes I. Neuroprotective protein and carboxypeptidase E // *J. Mol. Neurosci.* 2009. № 39(1–2). P. 1–8.
<https://doi.org/10.1007/s12031-008-9164-5>
66. Valente T.S., Baldi F., Sant'Anna A.C. et al. Genome-wide association study between single nucleotide polymorphisms and flight speed in nellore cattle // *PLoS One*. 2016. V. 11. № 6.
<https://doi.org/10.1371/journal.pone.0156956>
67. Lafage-Pochitaloff M., Gerby B., Baccini V. et al. The CADM1 tumor suppressor gene is a major candidate gene in MDS with deletion of the long arm of chromosome 11 // *Blood Adv.* 2022. V. 6. № 2. P. 386–398.
<https://doi.org/10.1182/bloodadvances.2021005311>
68. Machiela M.J., Grünwald T.G.P., Surdez D. et al. Genome-wide association study identifies multiple new loci associated with Ewing sarcoma susceptibility // *Nat. Commun.* 2018. V. 9. № 1.
<https://doi.org/10.1038/s41467-018-05537-2>
69. Terao C., Momozawa Y., Ishigaki K. et al. GWAS of mosaic loss of chromosome Y highlights genetic effects on blood cell differentiation // *Nat. Commun.* 2019. V. 10. № 1.
<https://doi.org/10.1038/s41467-019-12705-5>
70. Dawood M., Kramer L.M., Shabbir M.I., Reecy J.M. Genome-wide association study for fatty acid composition in american angus cattle // *Animals (Basel)*. 2021. V. 11. № 8.
<https://doi.org/10.3390/ani11082424>

71. Ben-Zur T., Feige E., Motro B., Wides R. The mammalian *odg* gene family: Homologs of a drosophila pair-rule gene with expression implying distinct yet overlapping developmental roles // *Dev. Biol.* 2000. V. 217. № 1. P. 107–120.
<https://doi.org/10.1006/dbio.1999.9532>
72. Berns D.S., DeNardo L.A., Pederick D.T., Luo L. Tenascin-3 controls topographic circuit assembly in the hippocampus // *Nature*. 2018. V. 554. № 7692. P. 328–333.
<https://doi.org/10.1038/nature25463>
73. Takano I., Takeshita N., Yoshida M. et al. Ten-m/Odz3 regulates migration and differentiation of chondrogenic ATDC5 cells via RhoA-mediated actin reorganization // *J. Cell Physiol.* 2021. V. 236. № 4. P. 2906–2919.
<https://doi.org/10.1002/jcp.30058>
74. Carr O.P., Glendinning K.A., Leamey C.A., Marotte L.R. Retinal overexpression of ten-m3 alters ipsilateral retinogeniculate projections in the wallaby (*Macropus eugenii*) // *Neurosci. Lett.* 2014. № 566. P. 167–171.
<https://doi.org/10.1016/j.neulet.2014.02.048>
75. Glendinning K.A., Liu S.C., Nguyen M. et al. Downstream mediators of ten-m3 signalling in the developing visual pathway // *BMC Neurosci.* 2017. V. 18. № 1.
<https://doi.org/10.1186/s12868-017-0397-5>
76. Young T.R., Bourke M., Zhou X. et al. Ten-m2 is required for the generation of binocular visual circuits // *J. Neurosci.* 2013. V. 33. № 30. P. 12490–12509.
<https://doi.org/10.1523/JNEUROSCI.4708-12.2013>
77. Aldahmesh M.A., Mohammed J.Y., Al-Hazzaa S., Alkuraya F.S. Homozygous null mutation in *ODZ3* causes microphthalmia in humans // *Genet. Med.* 2012. V. 14. № 11. P. 900–904.
<https://doi.org/10.1038/gim.2012.718>
78. Lu F., Xu X., Zheng B. et al. Case report: Expansion of phenotypic and genotypic data in *TENM3*-related syndrome: Report of two cases // *Front. Pediatr.* 2023. № 11.
<https://doi.org/10.3389/fped.2023.1111771>
79. Rodriguez-Rodriguez L., Ivorra-Cortes J., Carmona F.D. et al. PTGER4 gene variant rs76523431 is a candidate risk factor for radiological joint damage in rheumatoid arthritis patients: a genetic study of six cohorts // *Arthritis Res. Ther.* 2015. V. 17. № 306.
<https://doi.org/10.1186/s13075-015-0830-z>
80. Losonczy A., Makara J.K., Magee J.C. Compartmentalized dendritic plasticity and input feature storage in neurons // *Nature*. 2008. № 452. P. 436–441.
<https://doi.org/10.1038/nature06725>
81. Aceto G., Colussi C., Leone L. et al. Chronic mild stress alters synaptic plasticity in the nucleus accumbens through GSK3 β -dependent modulation of Kv4.2 channels // *Proc. Natl Acad. Sci. USA*. 2020. № 117. P. 8143–8153.
<https://doi.org/10.1073/pnas.1917423117>
82. Lin M.A., Cannon S.C., Papazian D.M. Kv4.2 autism and epilepsy mutation enhances inactivation of closed channels but impairs access to inactivated state after opening // *Proc. Natl Acad. Sci. USA*. 2018. № 115. P. E3559–E3568.
<https://doi.org/10.1073/pnas.1717082115>
83. Lee H., Lin M.C., Kornblum H.I. et al. Exome sequencing identifies de novo gain of function missense mutation in *KCND2* in identical twins with autism and seizures that slows potassium channel inactivation // *Hum. Mol. Genet.* 2014. № 23. P. 3481–3489.
<https://doi.org/10.1093/hmg/ddu056>
84. Liu M., Yu C., Zhang Z. et al. Whole-genome sequencing reveals the genetic mechanisms of domestication in classical inbred mice // *Genome Biol.* 2022. V. 23. № 1.
<https://doi.org/10.1186/s13059-022-02772-1>
85. Казанцева А.В., Еникеева Р.Ф., Романова А.Р. и др. Гены семейства нейрексина (*CNTNAP2* и *NRXN1*): их роль в развитии математической тревожности // *Мед. генетика*. 2016. Т. 15. № 11(173). С. 17–23.
86. Iijima T., Wu K., Witte H. et al. SAM68 regulates neuronal activity-dependent alternative splicing of neurexin-1 // *Cell*. 2011. V. 147. № 7. P. 1601–1614.
<https://doi.org/10.1016/j.cell.2011.11.028>
87. Lyu Y.L., Lin C.P., Azarova A.M. et al. Role of topoisomerase II β in the expression of developmentally regulated genes // *Mol. Cell Biol.* 2006. V. 26. № 21. P. 7929–7941.
<https://doi.org/10.1128/MCB.00617-06>
88. Colland F., Jacq X., Trouplin V. et al. Functional proteomics mapping of a human signaling pathway // *Genome Res.* 2004. V. 14. № 7. P. 1324–1332.
<https://doi.org/10.1101/gr.2334104>
89. Zhao Q., Wang F., Chen Y.X. et al. Comprehensive profiling of 1015 patients' exomes reveals genomic-clinical associations in colorectal cancer // *Nat. Commun.* 2022. V. 13. № 1.
<https://doi.org/10.1038/s41467-022-30062-8>
90. Sung H.Y., Han J., Ju W., Ahn J.H. Synaptotagmin-like protein 2 gene promotes the metastatic potential in ovarian cancer // *Oncol. Rep.* 2016. V. 36. № 1. P. 535–541.
<https://doi.org/10.3892/or.2016.4835>
91. Zhou X., Li J., Teng J. et al. Long noncoding RNA BSN-AS2 induced by E2F1 promotes spinal osteosarcoma progression by targeting miR-654-3p/SYTL2 axis // *Cancer Cell Int.* 2020. V. 20. № 133.
<https://doi.org/10.1186/s12935-020-01205-y>

Genetic Assessment of Projected Residual Feed Consumption and Expression of Significant Candidate Genes in Duroc Pigs and Second-Generation Commercial Blends

A. A. Belous^{a,*}, A. A. Sermyagin^a, and N. A. Zinovieva^a

^a*Ernst Federal Research Center for Animal Husbandry, Moscow oblast, Dubrovitsy, 142132 Russia*

^{*}*e-mail: belousa663@gmail.com*

Residual feed intake (RFI) is one of the basic and complex feed characteristics that is economically important for livestock production. However, the genetic and biological mechanisms governing this trait in pigs are largely unknown. Therefore, the study aimed to identify genome-wide single nucleotide polymorphisms (SNPs), candidate genes involved in RFI regulation, their biological pathways and clustering, using genome-wide association analysis (GWAS). The study was carried out on Duroc pigs ($n = 783$) and their commercial hybrids of the second generation ($n = 250$), undergoing test fattening at automatic feed stations for individual accounting. As a result, genes that are significant in terms of the orthology of biological functions and in terms of expression in tissues and organs and are associated with RFI were obtained. These candidate genes include: adhesion receptor G6 (*ADGRG6*), centromeric protein S (*APITD1*), carboxypeptidase E (*CPE*), transmembrane calcium-binding protein (*SYTL2*), cell adhesion molecule 1 (*CADMI*), Fli proto-oncogene-1, transcription factor ETS (*FLII*), teneurin transmembrane protein 3 (*TENM3*), prostaglandin E4 (*PTGER4*), and Potassium voltage-gated channel D subfamily member 2 (*KCND2*). In addition, the analysis of the obtained data on clustering showed the division into biological, functional and molecular libraries and data published in PubMed. Combining the information obtained, it can be said that the genetic component of the predicted residual feed intake is important, as indicated in previous and current studies. In this connection, there is a need to create molecular diagnostics and develop calculations for genomic assessment, in conjunction with feed conversion, which will improve productivity in pig breeding herds and improve the quality of products.

Keywords: feed utilization efficiency, predicted residual feed intake, genomic assessment, genome-wide association study, gene expression, Duroc boars, second-generation crossbreds, pig breeding.

ИЗУЧЕНИЕ ГЕНЕТИЧЕСКОГО РАЗНООБРАЗИЯ И ПОПУЛЯЦИОННОЙ СТРУКТУРЫ ОСЕТИНСКОЙ ПОРОДЫ В СРАВНЕНИИ С ДРУГИМИ ПОРОДАМИ ГРУБОШЕРСТНЫХ ЖИРНОХВОСТЫХ ОВЕЦ НА ОСНОВЕ ДАННЫХ SNP-ГЕНОТИПИРОВАНИЯ

© 2023 г. Т. Е. Денискова^{1, *}, А. В. Доцев¹, М. И. Селионова², Н. А. Зиновьева^{1, **}

¹Федеральный исследовательский центр животноводства – ВИЖ им. академика Л.К. Эрнста, городской округ Подольск, пос. Дубровицы, 142132 Россия

²Российский государственный аграрный университет – МСХА им. К.А. Тимирязева, Москва, 127434 Россия

*e-mail: horarka@yandex.ru

**e-mail: n_zinovieva@mail.ru

Поступила в редакцию 05.05.2023 г.

После доработки 15.06.2023 г.

Принята к публикации 20.06.2023 г.

Впервые проведен анализ генетического разнообразия осетинской породы овец с использованием полногеномных SNP-профилей в сравнительном аспекте с другими российскими грубошерстными породами. Установлено, что осетинская порода отличается высоким генетическим и аллельным разнообразием. Впервые рассчитаны современные и исторические эффективные размеры популяции осетинской породы. Показаны близкие генетические взаимосвязи между осетинской и карачаевской породами овец, что подтверждается низким значением дифференциации ($F_{st} = 0.009$) и формированием единого кластера в структуре генетической сети. Выявлена схожесть популяционной структуры осетинской и карачаевской пород.

Ключевые слова: SNP, локальные породы, домашние овцы, генетическое разнообразие, эффективный размер популяции.

DOI: 10.31857/S0016675823110048, **EDN:** NNRNBV

Овцеводство — это популярное направление животноводческой деятельности на территориях с преобладающим горным рельефом и с эпохи неолита главный источник существования населения древнего Кавказа [1]. В Республике Осетии овцеводство — это основное традиционное занятие, обеспечивающее осетин бараниной, молоком, сыром, шерстью и кожей [2]. Живые овцы и продукты овцеводства активно использовались для натурального обмена на различные товары, а также в качестве выплаты феодальных податей [2]. Известно, что особенно ценились животные черно-бурой масти, так как их шерсть была основой для изготовления высококачественного сукна — “осетинского сукна” [1].

Неприхотливая и выносливая осетинская порода овец приспособлена к суровым условиям обитания и характеризуется хорошей мясной продуктивностью и высокими вкусовыми качествами баранины без специфического запаха. Кроме того, до сих пор шерсть используется для изготовления бурок, валяной обуви и иных изделий [3]. Численность осетинской породы претер-

пела довольно большие изменения за последние 20 лет: 2200 голов в 2000 г., 3400 голов в 2010 г., 4300 голов в 2011 г., 3400 голов в 2012 г. и 1300 голов в 2019 г. [4, 5].

Считается, что осетинская порода произошла от средневековых пород или популяций овец, которых разводили аланы — ираноязычные кочевые племена скифо-сарматского происхождения, в послемонгольский период [1, 6]. Согласно историческим сведениям, до оттеснения в горы аланы-осетины активно практиковали систему отгонного овцеводства, которая была возрождена при переселении на равнины [6].

Для разработки научно обоснованных программ по сохранению генофонда уникальной осетинской породы необходимо проводить генетический мониторинг. Тем не менее практически ничего неизвестно о генетическом разнообразии, популяционной структуре и демографической истории осетинской породы овец.

Внедрение высокопроизводительных технологий для проведения массового параллельного генотипирования однонуклеотидных полиморфиз-

мов (SNP) открыло эру геномных исследований сельскохозяйственных животных. Среди различных направлений применения ДНК-чипов особую популярность нашло их использование для геномной характеристики локальных пород овец [7, 8]. Так, E. Ciani et al. [7] провели масштабное геномное изучение 19 итальянских пород овец с использованием ДНК-чипов, в результате чего было показано наличие активного потока генов между этими породами в прошлом и незадокументированной интрогрессии с диким муфлоном. На основе применения чипа Illumina OvineSNP50 для анализа популяционной структуры 18 аборигенных валлийских пород овец было продемонстрировано, что изучаемые овцы не образовывали генетически однородной группы (попарные значения показателя F_{st} между породами варьировали от 0.107 до 0.201) [8]. С помощью ДНК-чипов были рассчитаны значения эффективного размера популяции киргизских локальных пород овец, которые варьировали от 176 в алайской породе до 563 в айкольской породе [9]. Геномные исследования десяти алжирских и марокканских пород овец выявили, что некоторые породы разного происхождения генетически неразличимы. Полученные результаты поставили под сомнение целостность геномных ресурсов этих местных пород, сохранение которых имеет решающее значение для будущего сектора животноводства региона Магриба [10]. SNP-анализ пяти шведских пород показал, что эти породы представляют собой пять различных пород, в то время как гуте и готланд более тесно связаны друг с другом [11]. Кроме того, использование ДНК-чипа позволило оценить популяционную структуру и генетическое разнообразие бразильских пород овец санта-инес, морада нова, бразильская чернобрюхая и бразильская жирнохвостая, разводимых экстенсивным способом [12].

В связи с этим цель настоящей работы — характеристика генетического разнообразия и анализ популяционной структуры осетинской породы овец на основе полногеномных SNP-профилей.

МАТЕРИАЛЫ И МЕТОДЫ

Выборка осетинской породы была представлена 30 образцами (OSET, $n = 30$). В качестве групп сравнения были использованы образцы овец грубошерстных жирнохвостых овец, в том числе: карачаевской, лезгинской, тушинской, андийской, каракульской, эдильбаевской и тувинской короткожирнохвостой пород. Краткая характеристика изучаемых пород [13] представлена в табл. 1.

Исследуемые образцы овец были генотипированы с использованием ДНК-чипов OvineSNP50 BeadChip (Illumina, San Diego, CA, США) или Ovine Infinium HD BeadChip (Illumina, San Diego, CA, США) [14]. После объединения генотипиче-

ских данных в анализе осталось всего 42 230 ауто-сомных SNP, которые перекрывались между двумя ДНК-чипами.

Проведены контроль качества генотипирования и выбор SNP для проведения дальнейшего анализа. SNP, расположенные на половых хромосомах или с неизвестной локацией, а также отсутствующие у менее чем 90% образцов, были удалены из анализа. Дополнительно были отброшены SNP-маркеры, для которых частота минорного аллеля (MAF) составила менее 5% и которые были идентифицированы как неполиморфные.

Параметры генетического разнообразия, включая наблюдаемую гетерозиготность (H_o), несмещенную ожидаемую гетерозиготность (uH_e) и несмещенный коэффициент инбридинга (uF_{is}) с 95%-ным доверительным интервалом (CI95%), и рарифицированное аллельное разнообразие (A_r) были рассчитаны в R “diveRsity” [15].

Эффективный размер популяции (N_e) был рассчитан на основе анализа неравновесия по сцеплению (LD) в программе SNeP [16]. При расчетах были установлены параметры по умолчанию, за исключением поправки на размер выборки, частоты мутаций ($\alpha = 2.2$ [17]) и скорости рекомбинации между парой генетических маркеров по [18].

Генетическая дифференциация между изучаемыми популяциями овец проводилась путем расчета попарных значений показателя F_{st} [19] с помощью пакета R StAMPP [20].

Анализ главных компонент (Principal Component Analysis, PCA) был выполнен в PLINK v1.9 [21] и визуализирован с помощью пакета R “ggplot2” [22].

Генетическая сеть Neighbor-Net, основанная на матрице попарных значений F_{st} , была построена в программе SplitsTree 4.14.5 [23].

Геномная кластеризация была изучена с использованием программного обеспечения Admixture v1.3 [24] и визуализирована с помощью пакета R “pophelper” [25].

РЕЗУЛЬТАТЫ

Анализ параметров генетического разнообразия показал, что наблюдаемая гетерозиготность не сильно варьировала в изучаемых популяциях грубошерстных локальных пород: от 0.367 в андийской породе до 0.388 в лезгинской породе соответственно (табл. 2). Уровень наблюдаемой гетерозиготности, рассчитанный для выборки осетинской породы, приближался к высоким значениям среди изучаемых популяций ($H_o = 0.386$). В осетинской породе отмечена незначительная тенденция к недостатку гетерозигот, так как значение коэффициента инбридинга было положительным, но не достоверным.

Таблица 1. Краткая характеристика овец осетинской и других грубошерстных пород (*Ovis aries*), используемых в исследовании

№	Порода	Аббревиатура	Объем выборки	Краткая характеристика	Регион отбора проб
1	Осетинская	OSET	30	Мясо-шерстно-молочная. Живая масса 45–50 кг (♀), 65–90 кг (♂)	Республика Северная Осетия-Алания
2	Карачаевская	KRCH	41	Мясо-шерстно-молочная. Живая масса 45–50 кг (♀), 65–90 кг (♂). Молоко – сырье для брынзы. Высокая валкость шерсти	Карачаево-Черкесская Республика
3	Лезгинская	LEZG	42	Мясо-шерстно-молочная. Живая масса 40–45 кг (♀), 60–65 кг (♂). Шерсть – сырье для производства ковров. Круглогодичное пастбищное содержание	Республика Дагестан
4	Тушинская	TUSH	17	Мясо-шерстно-молочная. Живая масса 40–45 кг (♀), 60–65 кг (♂). Приспособлены к длительным переходам	Республика Дагестан
5	Андийская	ANDB	17	Мясо-шерстно-молочная. Живая масса 40–42 кг (♀), 46–55 кг (♂). Высокая насыщенность цвета шерсти (белый и черный)	Республика Дагестан
6	Каракульская	KARA	41	Мясо, овчины, смушки. Живая масса 45–50 кг (♀), 65–90 кг (♂). Красивый узор меха. Выносливы к засушливому климату	Республика Калмыкия
7	Эдильбаевская	EDLB	41	Мясосальная. Живая масса 65–70 кг (♀), 110–120 кг (♂). Крепкая конституция, высокая выносливость	Волгоградская область
8	Тувинская короткожирно-хвостая	TUVA	41	Мясная. Живая масса 45–56 кг (♀), 67–78 кг (♂). Приспособлены к суровым условиям	Республика Тыва

Примечание. Краткая характеристика пород взята из книги “Справочник пород и типов сельскохозяйственных животных, разводимых в Российской Федерации”, подготовленной под руководством И.М. Дунина и А.Г. Данкверта [13].

Аллельное разнообразие в осетинской породе (наряду с лезгинской породой) было наиболее высоким среди исследуемых пород: $A_r = 1.992$ по сравнению с вариациями от 1.972 в андийской породе до 1.990 в карачаевской и тувинской короткожирнохвостой породах.

Значения эффективного размера популяции, рассчитанного для трех поколений назад, варьировали от 98 в андийской породе до 516 в тувинской короткожирнохвостой породе (табл. 2). Современный эффективный размер популяции осетинской породы был довольно высок и составлял 406 рассчитанных особей.

Анализ изменений эффективного размера популяций, рассчитанного для 50 поколений назад, показал, что снижение этого показателя с течени-

ем поколений наблюдалось у всех исследуемых групп овец (рис. 1).

Дополнительно были проанализированы тенденции изменений исторического эффективного размера популяций, рассчитанного для 520 поколений назад (рис. 2). У осетинской породы наблюдалось несколько пиков и спадов в разрезе исторических поколений. Так, например, 506 поколений назад спад наблюдался как у осетинской ($N_{e506} = 5264$ по сравнению с $N_{e520} = 6069$), так и у других изучаемых пород овец, за которым последовал рост ($N_{e430} = 6269$). Исторический эффективный размер популяции осетинской породы был максимальным 235 поколений назад ($N_{e235} = 6360$). Снижение эффективного размера численности

Таблица 2. Генетическое разнообразие в осетинской и других изучаемых грубошерстных породах овец (*Ovis aries*)

Порода	<i>n</i>	H_o	uH_e	uF_{is}	A_r	N_{e3}	N_{e5}	N_{e15}
OSET	30	0.386	0.387	0.002 [0; 0.004]	1.992	406	560	1035
KRCH	41	0.386	0.385	−0.003 [−0.005; −0.001]	1.990	351	469	893
LEZG	42	0.388	0.389	0.002 [0; 0.004]	1.992	371	556	1233
TUSH	17	0.383	0.386	0.007 [0.005; 0.009]	1.991	211	324	764
ANDB	17	0.367	0.364	−0.007 [−0.009; −0.005]	1.972	98	129	250
KARA	41	0.380	0.378	−0.007 [−0.009; −0.005]	1.985	290	412	854
EDLB	41	0.383	0.379	−0.008 [−0.01; −0.006]	1.985	255	357	710
TUVA	41	0.385	0.385	0 [−0.002; 0.002]	1.990	516	706	1442

Примечание. *n* – число образцов; H_o – наблюдаемая гетерозиготность; uH_e – несмещенная ожидаемая гетерозиготность; uF_{is} – несмещенный коэффициент инбридинга; A_r – аллельное разнообразие; N_{e3} , N_{e5} и N_{e15} – эффективный размер популяции, рассчитанный для трех, пяти и 15 поколений назад. В квадратных скобках обозначен размах изменчивости uF_{is} при доверительном интервале 95%. Стандартные ошибки (SE) составляли ± 0.001 для показателей H_o , uH_e , A_r .

Таблица 3. Попарные значения F_{st} , рассчитанные для осетинской и других изучаемых грубошерстных пород овец (*Ovis aries*)

	OSET	KRCH	LEZG	TUSH	ANDB	KARA	EDLB	TUVA
OSET								
KRCH	0.009							
LEZG	0.012	0.017						
TUSH	0.013	0.019	0.008					
ANDB	0.033	0.037	0.032	0.034				
KARA	0.030	0.034	0.025	0.026	0.051			
EDLB	0.025	0.030	0.021	0.023	0.047	0.026		
TUVA	0.020	0.023	0.017	0.019	0.042	0.023	0.017	

снижается со 180 поколений назад вплоть до современных значений.

Анализ главных компонент показал, что первая главная компонента (PC1), ответственная за 10.92% генетической изменчивости, разделяет исследуемые породы на несколько групп (рис. 3). Первая группа была сформирована осетинской, каракаевской, лезгинской, тушинской и андийской породами. Вторая группа включала в себя эдильбаевскую, тувинскую короткожирнохвостую (левый верхний квадрант) и каракульскую (левый нижний квадрант) породы.

Генетическая сеть, построенная для осетинской и других изучаемых грубошерстных пород овец на основе расчета попарных значений F_{st} , показала наличие двух кластеров (рис. 4). В первый кластер входили каракульская, эдильбаевская

и тувинская короткожирнохвостая породы. Вторым кластер был представлен группой осетинской и каракаевской пород, независимой ветвью андийской породы и группой лезгинской и тушинской пород.

Наименьшее значение F_{st} было идентифицировано между осетинской и каракаевской породами ($F_{st} = 0.009$), а наибольшее – между осетинской и андийской и между осетинской и каракульской породами ($F_{st} = 0.033$ и $F_{st} = 0.03$ соответственно). Значения F_{st} составили 0.012 и 0.013 для пар осетинская и лезгинская, осетинская и тушинская соответственно (табл. 3).

Кластерный анализ продемонстрировал, что при $K = 2$ породы распределились по двум кластерам, соответствующим детектированным ранее с помощью PCA-анализа и построения генетической

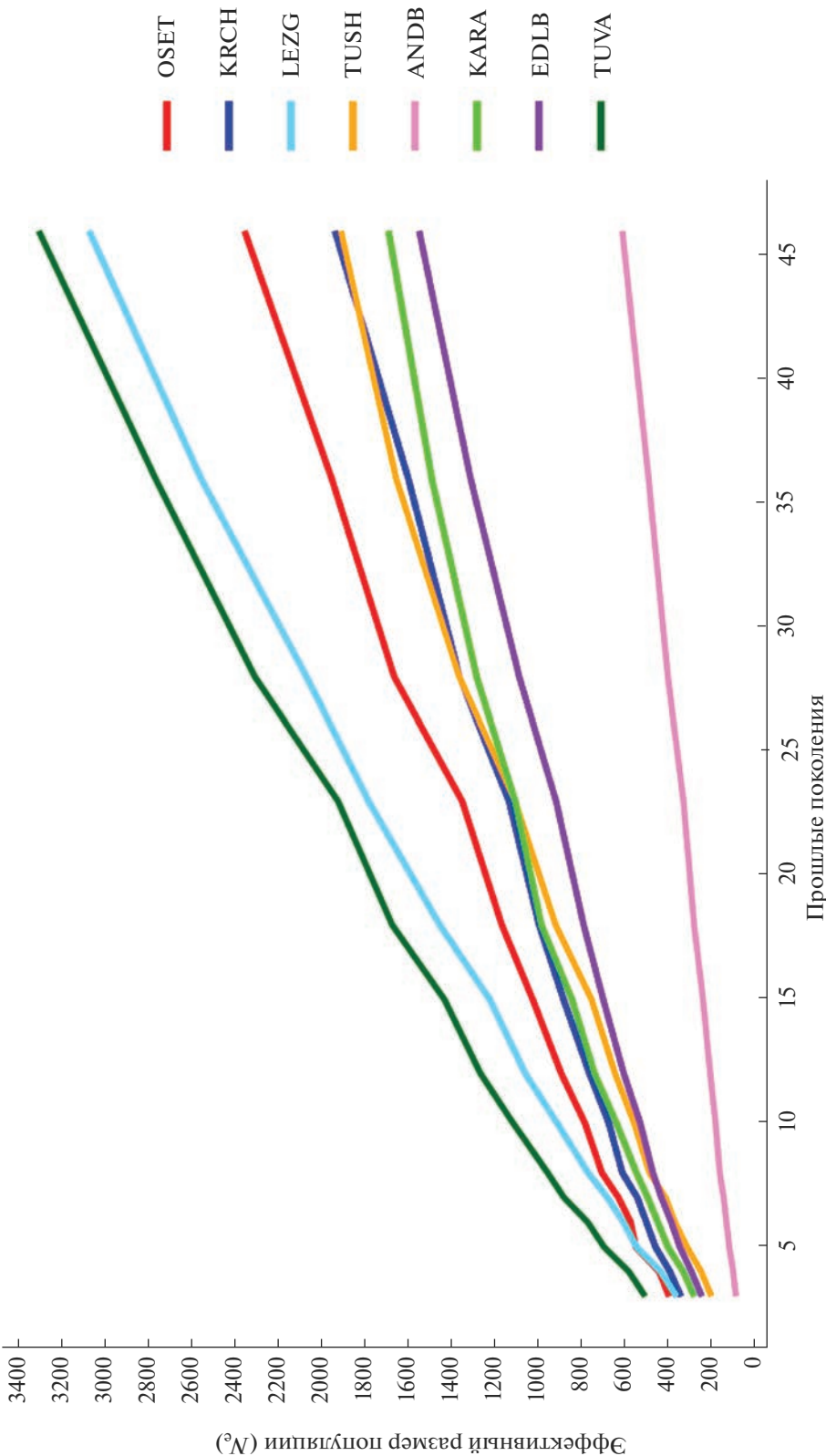


Рис. 1. Эффективный размер популяции, рассчитанный для 50 поколений назад (N_{e50}).

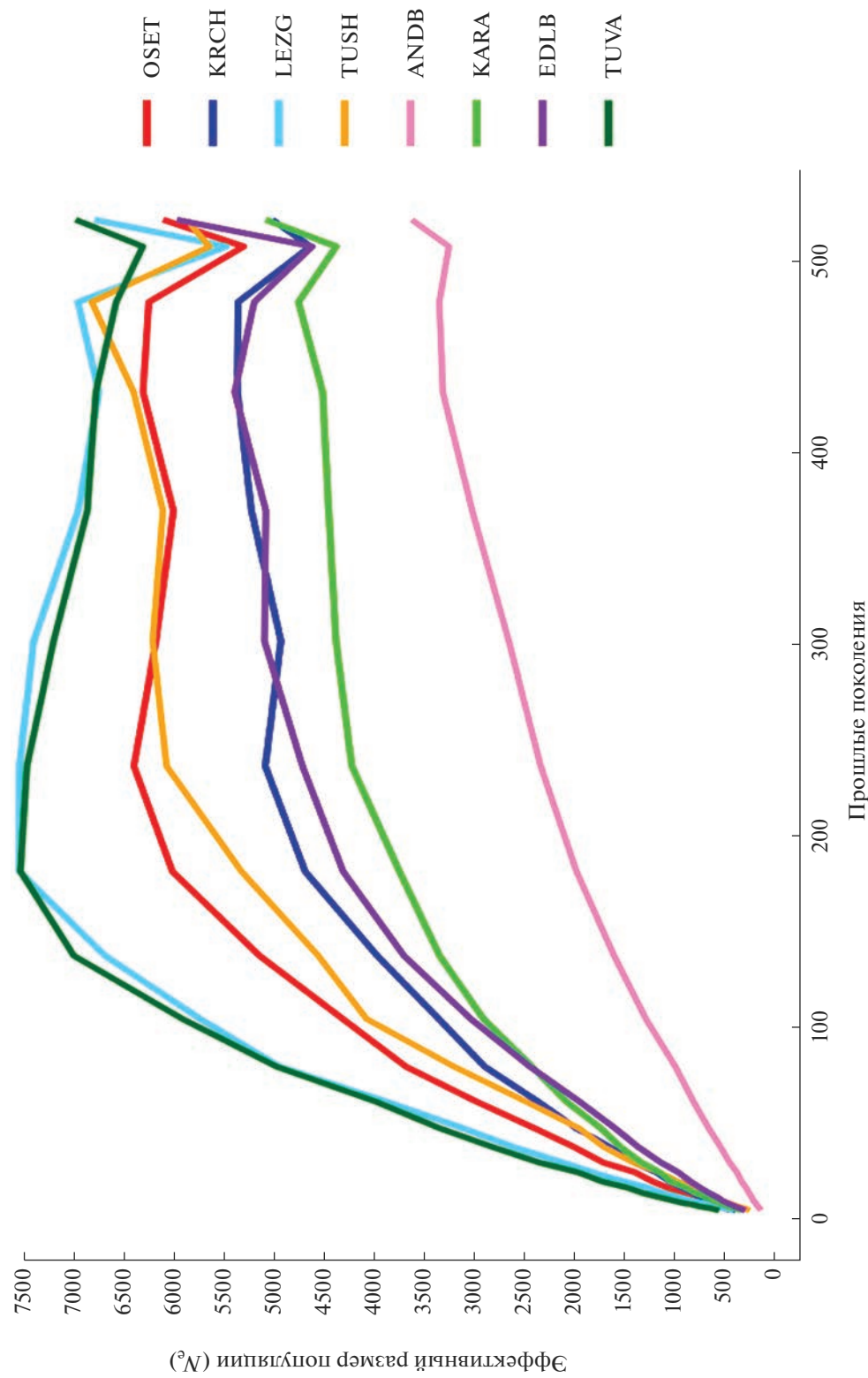


Рис. 2. Эффективный размер популяции, рассчитанный для 520 поколений назад (N_{e520}).

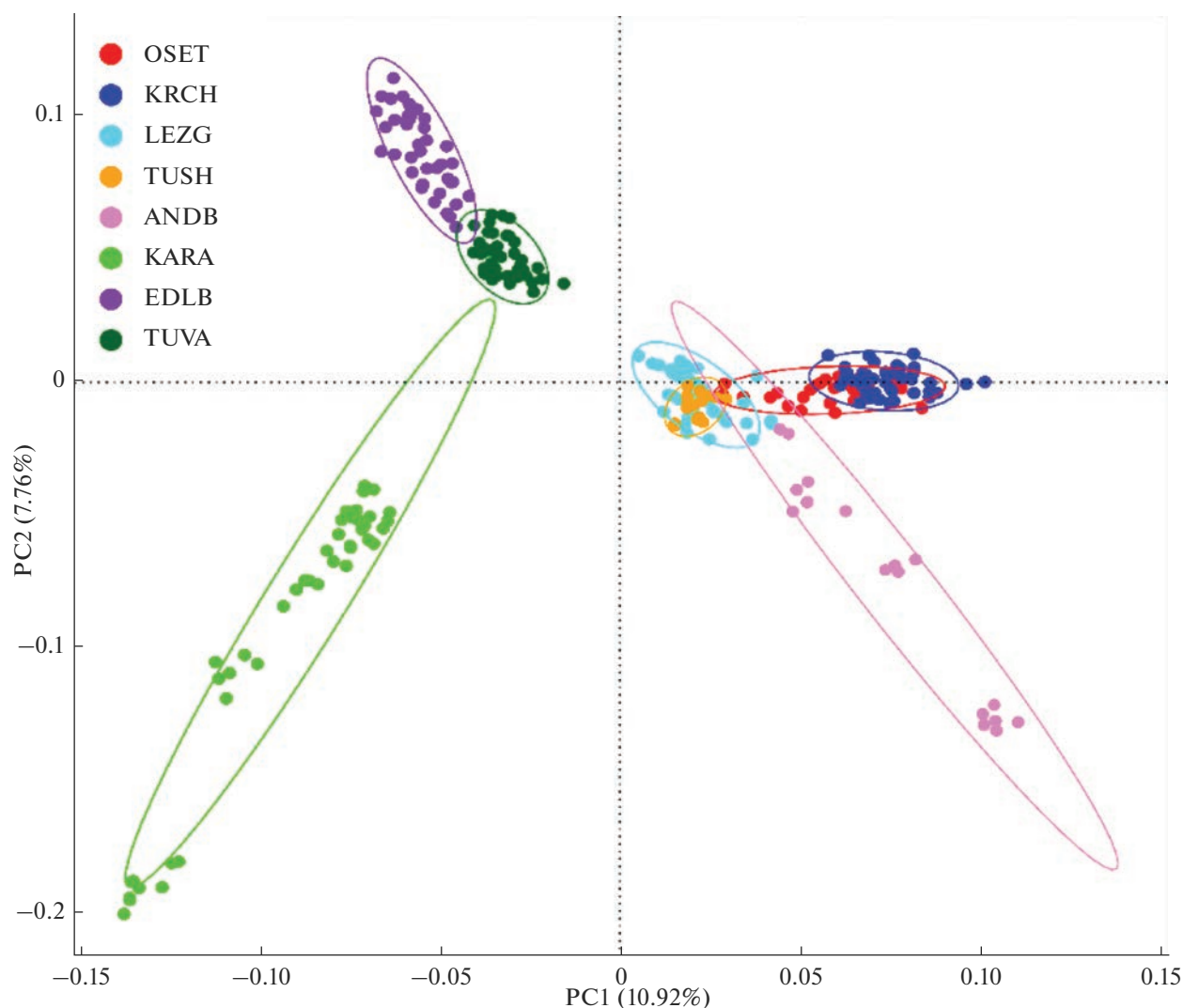


Рис. 3. Анализ главных компонент, проведенный для осетинской и других изучаемых грубошерстных пород овец (*Ovis aries*).

сети (рис. 5). Начиная с этого значения K , осетинская и карачаевская породы характеризовались большой схожестью популяционной структуры. С увеличением числа K лезгинская, тушинская и андийская породы заметнее обособляются от осетинской и карачаевской пород.

ОБСУЖДЕНИЕ

Аборигенные породы овец уникальны тем, что адаптированы к разведению в специфических, зачастую суровых условиях (например, разреженный воздух, скудные пастбища, сложный рельеф). Следует отметить, что условия горной местности не всегда равнозначны, поэтому в регионах разводят грубошерстные породы овец, приспособленные к конкретной местности. Так, андийская и лезгинская породы популярны в Республике Дагестан;

осетинская, тушинская и карачаевская породы наиболее распространены в Республике Северная Осетия-Алания; карачаевская порода — одна из часто встречающихся групп овец в республиках Карачаево-Черкесия и Кабардино-Балкария [4]. В связи с этим генетические ресурсы каждой аборигенной породы по-своему уникальны.

Генетические взаимосвязи осетинской породы с другими породами овец Северного Кавказа недостаточно изучены. Например, была предпринята попытка исследования межпородной дифференциации северокавказских грубошерстных овец — карачаевской, лезгинской, осетинской, тушинской и андийской пород [26]. На основании маркеров групп крови породы распределились по двум кластерам: кластер А, включающий тушинскую и осетинскую породы, и Б, сформированный из карачаевской, лезгинской, андийской и осетинской пород.

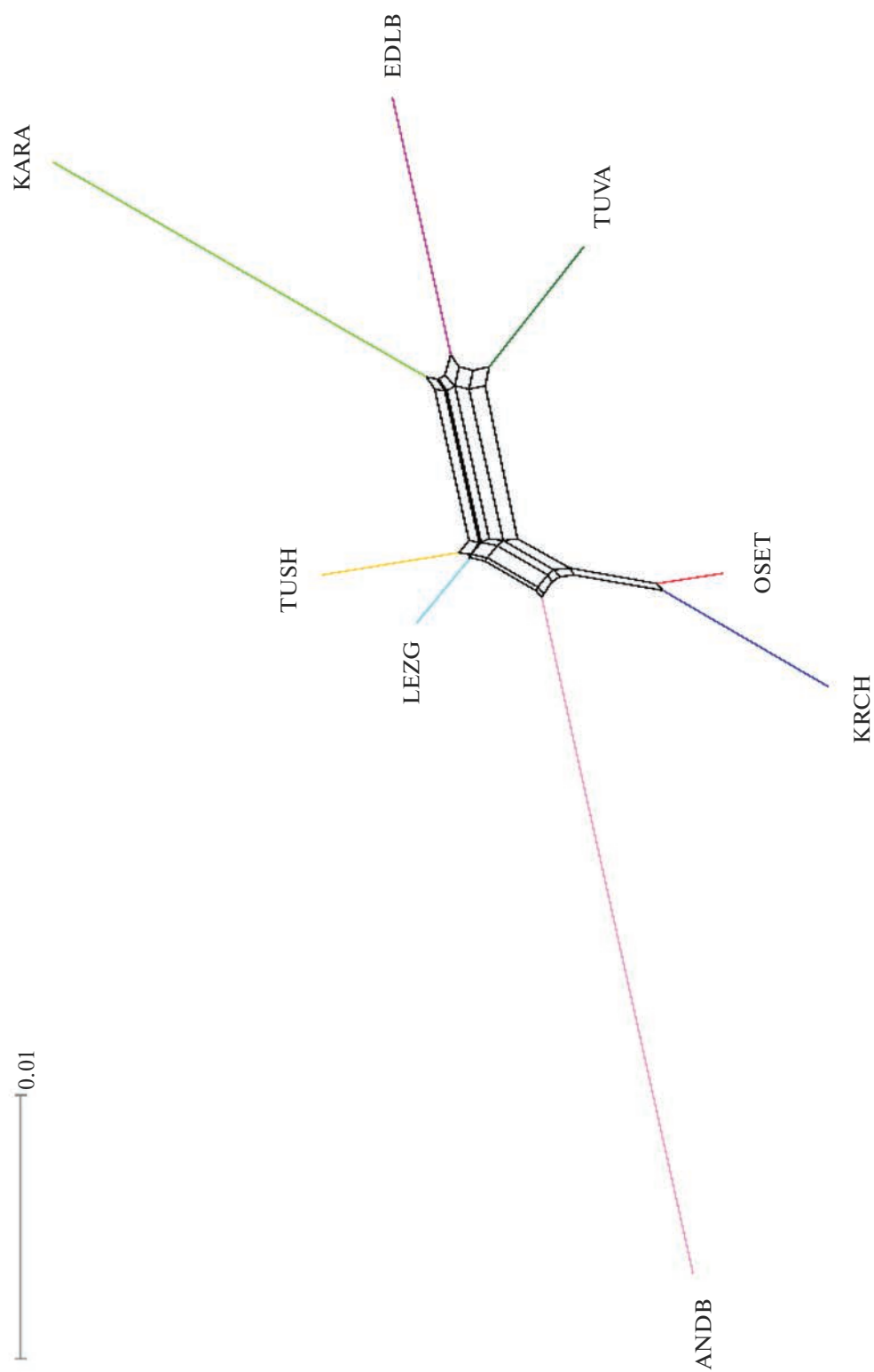


Рис. 4. Генетическая сеть, построенная для осетинской и других изучаемых грубошерстных пород овец (*Ovis aries*) на основе расчета попарных значений F_{st} .

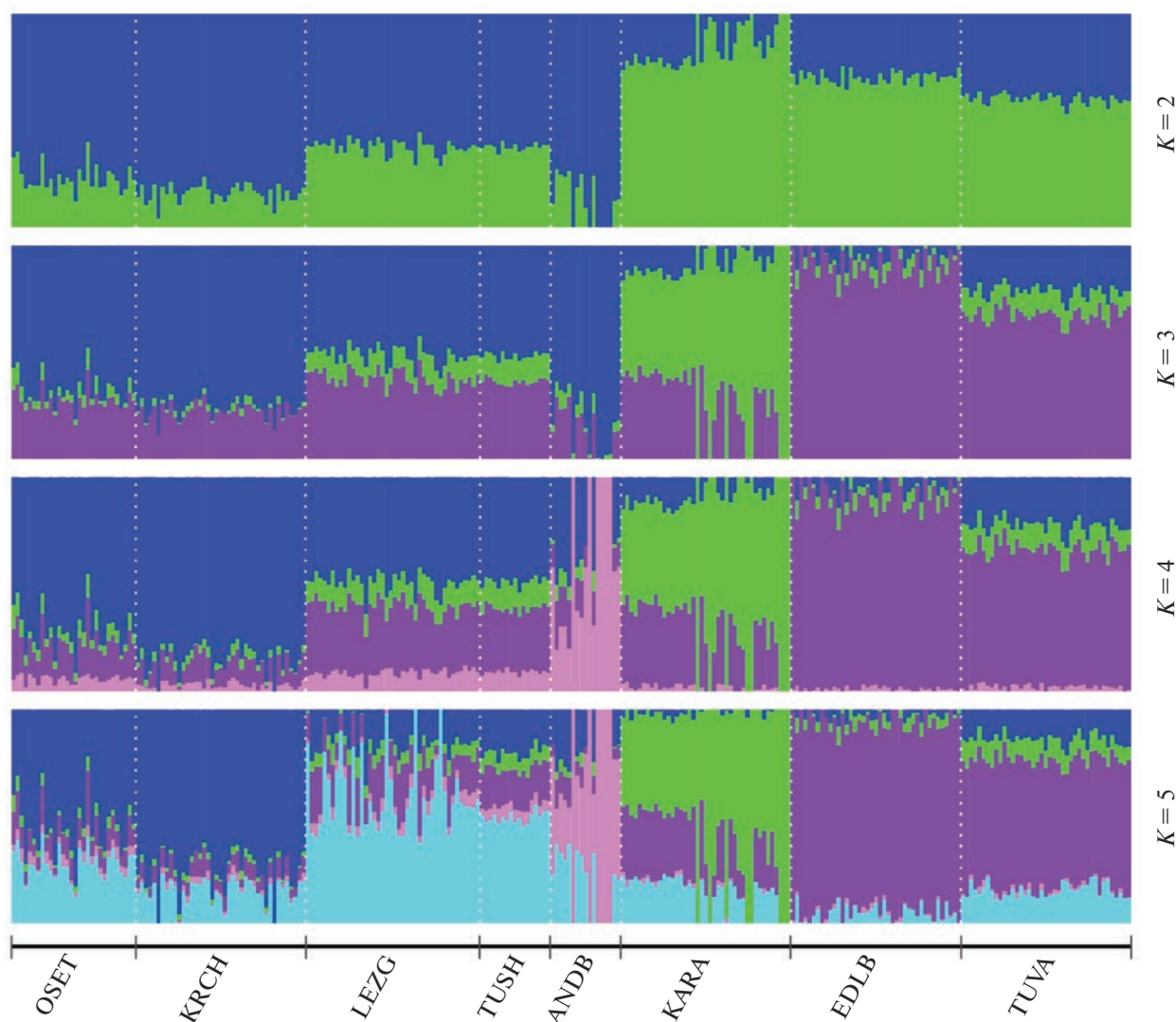


Рис. 5. Кластерный анализ, проведенный для осетинской и других изучаемых грубошерстных пород овец (*Ovis aries*), для вероятного числа предковых популяций (K) от 2 до 5.

рованный карачаевской, андийской и лезгинской породами. Наши результаты не поддерживают эти данные, так как белковые маркеры для проведения популяционных исследований характеризуются недостаточной разрешающей способностью.

Следует отметить, что SNP-профили осетинской породы были использованы в наших предыдущих исследованиях для проведения ряда специфических анализов. Так, ранее был изучен паттерн распространения сегментов гомозиготности в геномах российских пород овец, включая осетинскую [27]. В результате проведенного анализа вариации числа копий (CNV) в выборке овец осетинской породы были выявлены 337 регионов CNV [28]. В геномных регионах, перекрываю-

щихся с CNV, у осетинской породы были выявлены интересные гены, ассоциированные с адаптацией к гипоксии и мясными качествами, не идентифицированные в других российских породах, в том числе и в карачаевской. Эти экспериментальные данные еще раз подтверждают то, что уникальный генофонд осетинской породы необходимо поддерживать и совершенствовать.

В настоящей работе проведено изучение генетического разнообразия осетинской породы, а также изучена демографическая история породы на основе расчета современного и исторического эффективного размера популяции.

Несмотря на то что осетинская порода развивается в узком ареале, эта группа овец характеризуется довольно высокими значениями параметров,

характеризующих генетическое разнообразие. Так, значения ожидаемой гетерозиготности, рассчитанные для итальянских местных пород овец (H_e от 0.33 у альматураны до 0.37 у латичуда и соправиссаны) [7] и для валлийских пород овец (H_e от 0.306 у керри хил до 0.380 у баджер фейс) [8], были ниже рассчитанных в настоящей работе для осетинской породы. Аллельное разнообразие осетинской породы было выше по сравнению со значениями этого показателя у киргизских пород овец ($A_r = 1.992$ против A_r от 1.964 у алайской до 1.987 у айкольской пород) [9].

Современный эффективный размер популяции, рассчитанный для пяти поколений назад, также довольно высок у осетинской породы и был соизмерим со значениями, рассчитанными для других анализируемых пород в настоящей работе ($N_{e5} = 129-706$), а также для гиссарской и айкольской пород ($N_{e5} = 660$ и 563) [9] и для валлийских пород ($N_{e5} = 88-825$) [8].

Тенденция к снижению эффективного размера популяции с уменьшением числа поколений отмечена у многих пород овец. Так, максимальные значения эффективного размера популяции итальянских пород наблюдались между 50–60 или более 60 поколений назад [7]. Тем не менее значения N_{e50} для этих пород существенно ниже (224–657) по сравнению со значением аналогичного показателя у осетинской породы. Следует отметить, что пик эффективного размера популяции осетинской породы приходился более чем на 200 поколений назад (а именно на 235-е поколение назад). Подобная тенденция была отмечена в гиссарской породе [9]. Гиссарская порода, как и осетинская, характеризуется древним происхождением и разведением преимущественно по системе классического пасторализма, что, возможно, нашло отражение в паттернах изменения эффективного размера популяции.

Расчет попарных значений F_{st} , построение генетической сети и кластерный анализ показали близкие генетические взаимосвязи осетинской и карачаевской пород овец. Известно, что карачаевская порода была также издавна распространена в Осетии [6]. Высказываются предположения, что карачаевская порода – это одна из разновидностей осетинской породы [29]. Исследуемые выборки пород в настоящей работе недостаточны для подтверждения или опровержения этого предположения. Тем не менее можно дискутировать о том, что осетинская и карачаевская породы произошли от общей предковой группы овец или что происходит, с меньшей долей вероятности, неконтролируемое скрещивание этих пород.

Исследуя африканские магрибские породы овец, I. Belabdi et al. [10], основываясь на рассчитанном значении $F_{st} = 0.0024$, сделали вывод о генетической гомогенизации местных пород. В настоящем исследовании между породами, обитающими на территории Северного Кавказа (осетинская, карачаевская, лезгинская, тушинская), также детектированы невысокие значения дифференциации ($F_{st} =$ от 0.009 до 0.013), что дает основания предполагать о родственном происхождении этих пород.

Таким образом, осетинская порода овец характеризуется высоким уровнем генетического разнообразия и высоким значением современного эффективного размера популяции. Следует отметить, что несмотря на сравнительно небольшую численность осетинская порода не уступает другим грубошерстным, в частности северокавказским, породам по генетическому разнообразию. Вероятно, порода еще не находится в критическом состоянии и может быть успешно сохранена и улучшена для будущих поколений горцев с помощью программ консервации.

Финансирование проводилось за счет средств Министерства науки и высшего образования Российской Федерации, грант № 075-15-2021-1037 (внутренний № 15. БРК.21.0001).

Полногеномные профили овец каракульской и эдильбаевской пород были получены при финансовой поддержке проекта РНФ № 21-46-00001.

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. Калоев Б.А. Скотоводство народов Северного Кавказа: с древнейших времен до начала XX в. М.: Наука, 1993. 227 с.
2. Агузарова О.М. Из истории овцеводства Осетии // Современная наука: теоретический и практический взгляд. Материалы I Международной научно-практ. конф. ООО “НОУ “Вектор науки”. М.: Перо, 2014. С. 83–86.
3. Елканова Р.Э. Сравнительная характеристика мясной продуктивности овец грубошерстных пород // Вестник научных трудов молодых ученых, аспирантов, магистрантов и студентов ФГБОУ ВО “Горский государственный аграрный университет”. Владикавказ: Горский гос. аграрный ун-т, 2018. С. 127–128.

4. Гаджиев З.К., Велибеков Р.А., Мусалаев Х.Х. Состояние и перспективы развития грубошерстного овцеводства на юге России // Овцы, козы, шерстяное дело. 2013. № 2. С. 40–42.
5. Дунин И.М., Амерханов Х.А., Сафина Г.Ф. и др. Овцеводство России и его племенные ресурсы // Ежегодник по племенной работе в овцеводстве и козоводстве в хозяйствах Российской Федерации (2018 г.). Лесные Поляны, 2019. С. 3–14.
6. Калоев Б.А. Осетины: историко-этнографическое исследование. М.: Наука, 2004. 472 с.
7. Ciani E., Crepaldi P., Nicoloso L. et al. Genome-wide analysis of Italian sheep diversity reveals a strong geographic pattern and cryptic relationships between breeds // Anim. Genet. 2014. V. 45. № 2. P. 256–266. <https://doi.org/10.1111/age.12106>
8. Beynon S.E., Slavov G.T., Farré M. et al. Population structure and history of the Welsh sheep breeds determined by whole genome genotyping // BMC Genet. 2015. № 16. P. 65. <https://doi.org/10.1186/s12863-015-0216-x>
9. Deniskova T., Dotsev A., Lushihina E. et al. Population structure and genetic diversity of sheep breeds in the Kyrgyzstan // Front. Genet. 2019. № 10. P. 1311. <https://doi.org/10.3389/fgene.2019.01311>
10. Belabdi I., Ouhrouch A., Lafri M. et al. Genetic homogenization of indigenous sheep breeds in Northwest Africa // Sci. Rep. 2019. V. 9. № 1. P. 7920. <https://doi.org/10.1038/s41598-019-44137-y>
11. Rochus C.M., Jonas E., Johansson A.M. Population structure of five native sheep breeds of Sweden estimated with high density SNP genotypes // BMC Genet. 2020. V. 21. № 1. P. 27. <https://doi.org/10.1186/s12863-020-0827-8>
12. Paim T.P., Paiva S.R., de Toledo N.M. et al. Origin and population structure of Brazilian hair sheep breeds // Anim. Genet. 2021. V. 52. № 4. P. 492–504. <https://doi.org/10.1111/age.13093>
13. Дунин И.М., Данкверт А.Г. Справочник пород и типов сельскохозяйственных животных, разводимых в Российской Федерации. М.: ВНИИплем, 2013. 500 с.
14. Kijas J.W., Porto-Neto L., Dominik S. et al. Linkage disequilibrium over short physical distances measured in sheep using a high-density SNP chip // Anim. Genet. 2014. V. 45. № 5. P. 754–757. <https://doi.org/10.1111/age.12197>
15. Keenan K., McGinnity P., Cross T.F. et al. DiveRsity: An R package for the estimation of population genetics parameters and their associated errors // Methods. Ecol. Evol. 2013. № 4. P. 782–788. <https://doi.org/10.1111/2041-210X.12067>
16. Barbato M., Orozco-terWengel P., Tapio M., Bruford M.W. SNeP: A tool to estimate trends in recent effective population size trajectories using genome-wide SNP data // Front. Genet. 2015. № 6. P. 109. <https://doi.org/10.3389/fgene.2015.00109>
17. Corbin L.J., Liu A.Y., Bishop S.C., Woolliams J.A. Estimation of historical effective population size using linkage disequilibria with marker data // J. Anim. Breed. Genet. 2012. V. 129. № 4. P. 257–270. <https://doi.org/10.1111/j.1439-0388.2012.01003.x>
18. Sved J.A., Feldman M.W. Correlation and probability methods for one and two loci // Theor. Popul. Biol. 1973. V. 4. № 1. P. 129–132. [https://doi.org/10.1016/0040-5809\(73\)90008-7](https://doi.org/10.1016/0040-5809(73)90008-7)
19. Weir B.S., Cockerham C.C. Estimating F-statistics for the analysis of population structure // Evolution. 1984. V. 38. № 6. P. 1358–1370. <https://doi.org/10.1111/j.1558-5646.1984.tb05657.x>
20. Pembleton L.W., Cogan N.O., Forster J.W. StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations // Mol. Ecol. Resour. 2013. V. 13. № 5. P. 946–952. <https://doi.org/10.1111/1755-0998.12129>
21. Chang C.C., Chow C.C., Tellier L.C. et al. Second-generation PLINK: Rising to the challenge of larger and richer datasets // Gigascience. 2015. V. 4. P. 7. <https://doi.org/10.1186/s13742-015-0047-8>
22. Wickham H. ggplot2: Elegant Graphics for Data Analysis. N. Y.: Springer-Verlag, 2009. 212 p.
23. Huson D.H., Bryant D. Application of phylogenetic networks in evolutionary studies // Mol. Biol. Evol. 2006. V. 23. № 2. P. 254–267. <https://doi.org/10.1093/molbev/msj030>
24. Alexander D.H., Novembre J., Lange K. Fast model-based estimation of ancestry in unrelated individuals // Genome. Res. 2009. V. 19. № 9. P. 1655–1664. <https://doi.org/10.1101/gr.094052.109>
25. Francis R.M. pophelper: An R package and web app to analyse and visualize population structure // Mol. Ecol. Resour. 2017. V. 17. № 1. P. 27–32. <https://doi.org/10.1111/1755-0998.12509>
26. Ольховская Л.В., Криворучко С.В. Межпородная дифференциация грубошерстных овец Северного Кавказа // Вестник Росс. акад. сельскохозяйственных наук. 2011. № 4. С. 63–65.
27. Deniskova T., Dotsev A., Selionova M. et al. Biodiversity of Russian local sheep breeds based on pattern of runs of homozygosity // Diversity. 2021. V. 13. № 8. P. 360. <https://doi.org/10.3390/d13080360>
28. Igoshin A.V., Deniskova T.E., Yurchenko A.A. et al. Copy number variants in genomes of local sheep breeds from Russia // Anim. Genet. 2022. V. 53. № 1. P. 119–132. <https://doi.org/10.1111/age.13163>
29. Рчеулишвили М.Д. К истории овцеводства Грузии. Тбилиси: Госиздат. ГССР, 1953. 220 с.

Study of Genetic Diversity and Population Structure of the Osetin Breed in Comparison with Other Coarse Wool Fat-Tailed Sheep Breeds Based on SNP Genotyping Data

T. E. Deniskova^{a, *}, A. V. Dotsev^a, M. I. Selionova^b, and N. A. Zinovieva^{a, **}

^a*Ernst Federal Research Center for Animal Husbandry, Moscow oblast, pos. Dubrovitsy, 142132 Russia*

^b*Russian State Agrarian University – Moscow Timiryazev Agricultural Academy, Moscow, 127434 Russia*

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

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

For the first time, an analysis of the genetic diversity of the Osetin sheep breed was performed using genome-wide SNP profiles in a comparison with other Russian coarse wool breeds. We established that the Osetin breed has high genetic and allelic diversity indicators. For the first time, the recent and historical population effective sizes of the Osetin breed were calculated. Close genetic relationships between the Osetin and Karachay sheep breeds were shown, that is confirmed by the low value of differentiation ($F_{st} = 0.009$) and the formation of a joint cluster at the Neighbor Net graph. The similarity of the population structure of the Osetin and Karachay breeds was revealed as well.

Keywords: SNP, local breeds, domestic sheep, genetic diversity, effective population size.

ОЦЕНКА АССОЦИИ СТЕПЕНИ МЕТИЛИРОВАНИЯ ДНК И ЧАСТОТЫ ХРОМОСОМНЫХ АБЕРРАЦИЙ ЛИМФОЦИТОВ ЧЕЛОВЕКА ПРИ ОДНОКРАТНОМ ОБЛУЧЕНИИ КРОВИ *in vitro*

© 2023 г. О. С. Цымбал^{1, *}, Д. С. Исубакова¹, Е. В. Брониковская¹, А. Ф. Николаева², В. О. Сигин², А. И. Калинин², Ж. А. Старцева³, Н. В. Литвяков^{1, 3}, И. В. Мильто^{1, 4}, Р. М. Тахауов^{1, 4}

¹Северский биофизический научный центр Федерального медико-биологического агентства, Томская обл., Северск, 636013 Россия

²Медико-генетический научный центр им. академика Н.П. Бочкова, Москва, 115478 Россия

³Томский национальный исследовательский медицинский центр Российской академии наук, Томск, 634009 Россия

⁴Сибирский государственный медицинский университет, Томск, 634050 Россия

*e-mail: olga-tsymbal@mail.ru

Поступила в редакцию 21.04.2023 г.

После доработки 30.05.2023 г.

Принята к публикации 05.07.2023 г.

Наиболее чувствительной к радиационному воздействию биомолекулой является ДНК, повреждение которой проявляются, в том числе, в виде хромосомных aberrаций (ХА). Метилирование ДНК участвует в регуляции экспрессии генов, репликации, репарации ДНК и др. Изменение процессов метилирования ДНК подвержено воздействию γ -излучения. Цель исследования – оценить связь степени метилирования ДНК и частоты ХА после острого облучения *in vitro* лимфоцитов крови человека γ -излучением. В исследовании приняли участие десять условно здоровых работников Сибирского химического комбината, в лимфоцитах крови которых оценивали степень метилирования CpG-динуклеотидов (широкогеномное бисульфитное секвенирование, XmaI-Reduced representation bisulfite sequencing – XmaI-RRBS) и частоту ХА (цитогенетическое исследование) после острого облучения крови *in vitro* в дозе 1.5 Гр. Выявили, что острое облучение крови γ -излучением привело к увеличению в лимфоцитах частоты aberrантных клеток, дицентрических хромосом, хроматидных и хромосомных фрагментов. Корреляционный анализ статуса метилирования CpG-динуклеотидов и частоты ХА выявил изменения степени метилирования 97 генов, которые с высокой силой положительно (56 генов) или отрицательно (41 ген) связаны с повышенной частотой ХА. В ходе широкогеномного скрининга генов, метилирование которых коррелирует с повышенной частотой ХА, выявлены гены, перспективные в качестве потенциальных маркеров радиационного воздействия и изучения механизмов формирования радиочувствительности организма, а также радиорезистентности опухолей при лучевой терапии.

Ключевые слова: метилирование ДНК, хромосомные aberrации, острое облучение, γ -излучение, лимфоциты крови, широкогеномное бисульфитное секвенирование (XmaI-RRBS).

DOI: 10.31857/S0016675823110152, EDN: NGTOJL

Ионизирующее излучение (ИИ) охватывает разные сферы деятельности человека: медицина, производство электро- и тепловой энергии, военная промышленность и др. Последствия радиационного воздействия проявляются на организменном и клеточно-молекулярном уровнях. Наиболее чувствительной к воздействию ИИ является молекула ДНК, изменение структуры которой приводит к различным генным, хромосомным и геномным нарушениям [1]. Признанной моделью для выявления повреждений ДНК является анализ хромосомных aberrаций (ХА) лимфоцитов крови [2]. Широко распространено изучение спектра и

частоты ХА после воздействия разного по продолжительности γ -излучения. Исследования острого [3–6] и хронического [7, 8] воздействия ИИ на организм и клетки человека показали, что частота ХА возрастает с увеличением дозы облучения, но в диапазоне “малых” доз эта зависимость имеет нелинейный характер [9]. Изучение влияния “малых” доз (менее 0.1 Гр) на организм человека после компьютерной томографии [10] показало, что количество дицентрических и кольцевых хромосом в соматических клетках увеличивается, при этом следует отметить значи-

тельную индивидуальную изменчивость этого показателя.

ИИ также влияет на эпигенетические процессы, из которых особый интерес представляет метилирование ДНК. Степень метилирования ДНК — довольно чувствительный показатель, на который могут оказывать влияние экзогенные и эндогенные факторы. Например, окислительный стресс вызывает подавление транскрипции генов-супрессоров опухолей посредством гиперметилирования ДНК и, наоборот, активация онкогенов путем деметилирования ДНК приводит к инициации канцерогенеза [11]. Хроническое воздействие ИИ приводит к aberrантному метилированию CpG-островков промоторов генов [12], которое имеет дозозависимый характер и сохраняется в течение нескольких лет после длительного внешнего (γ -излучение) и комбинированного (α - и γ -излучения) облучения [13]. В другом исследовании, наоборот, отмечается в большей степени гипометилирование CpG-островков, которое также сохраняется после острого облучения ИИ в дозе 1 Гр [14].

Исследование радиационного воздействия на организм человека показало связь степени метилирования ДНК и высокой частоты ХА. Y. Lee и соавт. [15] у работников объектов использования атомной энергии выявили глобальное снижение степени метилирования ДНК, по сравнению с контрольной группой. При этом низкая степень глобального метилирования ДНК коррелировала с высокой частотой ХА [15]. Нами было определено, что степень метилирования *Bak1* в лимфоцитах крови работников Сибирского химического комбината (СХК) связана с дозой облучения и повышенной частотой хроматидных (непарных) фрагментов [16]. Также выявлена отрицательная корреляция степени метилирования промоторов генов апоптоза с дозой внешнего облучения (*HRK*, *BAD*, *BID*), повышенной частотой aberrантных клеток (*BAX*) и дицентрических хромосом (*APAF1*) [17].

Таким образом, целью настоящего исследования является оценка связи степени метилирования ДНК и повышенной частоты ХА лимфоцитов при однократном облучении ИИ крови *in vitro*.

МАТЕРИАЛЫ И МЕТОДЫ

В исследование было включено десять условно здоровых мужчин — работников СХК, в возрасте от 18 до 34 лет, которые не подвергались радиационному воздействию и не страдали хроническими заболеваниями. От каждого донора было получено добровольное информированное согласие на сбор биоматериала (цельной крови) и обработку персональных данных.

Материалом для исследования являлась венозная кровь. После получения кровь от каждого

донора была разделена на две части — контрольный и экспериментальный образцы (контрольная и экспериментальная группы соответственно). Экспериментальные образцы подвергались однократному воздействию γ -излучения *in vitro* в дозе 1.5 Гр. Облучение крови проводили с помощью γ -терапевтического аппарата Theratron Equinox (Канада), предназначенного для дистанционного облучения (радиоактивный источник ^{60}Co , мощность дозы 1.93 Гр/мин, $E\gamma \approx 1.25$ МэВ). Облучение и культивирование образцов обеих групп осуществляли в день получения крови. После культивирования в лимфоцитах крови оценивали спектр и частоту ХА, а также степень метилирования ДНК.

Культивирование. Культивирование образцов цельной крови контрольной и экспериментальной групп проводили в стерильных условиях с использованием ламинарного шкафа (Kojair, Финляндия). Кровь в объеме 4 мл смешивали с 16 мл питательной среды RPMI 1640 (ПанЭко, Россия), с предварительно добавленным L-глутамином (ПанЭко, Россия), 4 мл эмбриональной телячьей сыворотки (ПанЭко, Россия) и 0.4 мл фитогемаглютинаина (ПанЭко, Россия). После чего кровь инкубировали в культуральных флаконах при 37°C в орбитальном шейкере-инкубаторе (Biosan Sia, Латвия) в течение 48 ч.

Из культурального флакона каждого образца за 1.5 ч до окончания инкубации половину культуральной суспензии отбирали в две отдельные пробирки, которые затем использовали для выделения ДНК и бисульфитного секвенирования. К оставшейся в культуральных флаконах суспензии добавляли 180 мкл колхицина (ПанЭко, Россия) и продолжали культивирование 1.5 ч при 37°C.

Выделение ДНК и метод HpaI-RRBS. Обе пробирки с культуральной суспензией (без колхицина) центрифугировали в течение 6 мин при 1500 об./мин (Eppendorf Centrifuge 5702R, Германия) и трижды отмывали физиологическим раствором для удаления гепарина. Осадок пипетировали и аликвотировали в пробирки типа эппендорф по 700 мкл, затем замораживали (–80°C).

Выделение ДНК из лимфоцитов крови проводили с использованием набора QIAamp DNA mini Kit (Qiagen, Германия) в соответствии с инструкцией производителя, после предварительной промывки замороженных образцов 1-кратным SSC (saline sodium citrate — смесь хлорида натрия и цитрата натрия).

Концентрацию и чистоту ДНК оценивали с помощью спектрофотометра NanoDrop-1000 (Thermo Scientific, США).

Степень метилирования ДНК в образцах контрольной и экспериментальной групп определяли бисульфитным секвенированием ограниченного набора геномных локусов — RRBS (reduced

representation bisulfite sequencing) с использованием фермента рестрикции XmaI (метод XmaI-RRBS) [18, 19]. Секвенирование геномных библиотек XmaI-RRBS осуществляли с использованием Ion Personal Genome Machine (PGM), набора для секвенирования Ion PGM 200 Sequencing Kit и Ion 318 Chip (Thermo Fisher Scientific, США) в соответствии с инструкциями производителя. Концентрацию геномных библиотек измеряли с использованием флуориметра Qubit 4 (Thermo Fisher Scientific, США).

Цитогенетический метод. После 48-часовой инкубации в культуральную суспензию (с колхицином) добавляли 0.56%-ный раствор KCl. Процесс гипотонизации длился 45 мин при 37°C. Полученную суспензию фиксировали 3%-ной уксусной кислотой, затем смесью этанола и ледяной уксусной кислоты в соотношении 3 : 1. После отмывки фиксатором и осветления раствора суспензию наносили на предметные стекла и окрашивали красителем Гимза (ПанЭко, Россия), растворенным в фосфатном буфере. Цитогенетический анализ был проведен на микроскопе Leica DM2500 (Leica, Германия; увеличение $\times 1000$) у 20 образцов (10 образцов контрольной и 10 образцов экспериментальной групп). У каждого образца оценивали минимум 300 метафазных пластинок на наличие ХА. Учитывали общее количество aberrантных клеток, а также хроматидные и хромосомные фрагменты, хроматидные обмены, кольцевые и дицентрические хромосомы, полиплоидные и мультиабберрантные (имеющие более пяти ХА) клетки.

Обработка полученных результатов. Данные секвенирования обрабатывали с использованием стандартного программного обеспечения Ion Torrent Suite™. С помощью программного пакета Bismark 0.20 выполняли выравнивание обработанных бисульфитом натрия прочтений и получение данных о метилированных и неметилированных цитозинах. Дифференциально метилированные цитозины отбирали из данных XmaI-RRBS с использованием пакета ROTS (reproducibility-optimized test statistic) с возможностью парного тестирования [20]. Расчеты проводили с использованием статистического языка программирования R 3.6.3 [21].

Результаты цитогенетического исследования выражали в виде частоты aberrантных клеток и различных типов ХА на 100 клеток (метафазных пластинок). Статистическую обработку результатов проводили с использованием программного обеспечения Statistica 10 (StatSoft, США). Для оценки статистически значимых различий между группами использовался U-критерий Манна–Уитни.

Для оценки связи частоты ХА и степени метилирования генов (*b*-value) каждого образца экспе-

риментальной группы использовали коэффициент корреляции Спирмена. Положительной или отрицательной корреляцией высокой силы считали значения от ± 0.7 до ± 1 , в соответствии со шкалой Чеддока. Анализ обогащения по функциональной принадлежности (gene ontology, далее GO) был выполнен с использованием пакета clusterProfiler функцией enrichGO. В качестве предельного значения для дифференциально метилированных цитозinov было установлено изменение среднего значения метилирования *b*-value на ± 0.20 (20%) между контрольной и экспериментальной группами. Различия считали статистически значимыми при $p < 0.05$.

РЕЗУЛЬТАТЫ

Цитогенетический анализ

Цитогенетический анализ, проведенный для обеих групп, показал, что облучение *in vitro* в дозе 1.5 Гр приводит к увеличению частоты aberrантных клеток, хроматидных и хромосомных фрагментов, а также дицентрических хромосом (табл. 1). Повышение частоты кольцевых хромосом в экспериментальной группе находится на уровне тенденции ($p = 0.064$). Мультиабберрантные клетки были исключены из анализа, поскольку не обнаружены в образцах обеих групп.

Корреляция среднего значения степени метилирования ДНК (*b*-value) с частотой ХА

Корреляционный анализ среднего значения степени метилирования генов и частоты ХА в образцах экспериментальной группы выявил 56 генов (*PPM1J*, *BCAN*, *GNPAT*, *DMBX1*, *C1orf87*, *LRRC37A6P*, *SLC16A9*, *TSPAN15*, *NPFFR1*, *ODF3*, *SMTNL1*, *STX3*, *C11orf80*, *DTX1*, *PGAM5*, *ATP11A*, *ATP12A*, *CCDC169-SOHLH2*, *HTR2A*, *PSMC1*, *FAN1*, *C2CD4A*, *CPLX3*, *ANPEP*, *CIB1*, *RHBDF1*, *TK2*, *FAM173A*, *COTL1*, *SPG7*, *MIR22HG*, *SLC25A52*, *CCDC159*, *WIZ*, *THEG5*, *MICOS13*, *CLEC4G*, *MZT2B*, *METTL21A*, *ANKMY1*, *LGALS1*, *BCL2L1*, *TXNRD2*, *SFMBT1*, *SMAD1*, *PIGG*, *CTNND2*, *FGF18*, *PAIP1*, *HACE1*, *HCG11*, *PKD1L1*, *ADCY8*, *DMRT1*, *LOC100129316* и *NUDT11*), у которых значение *b*-value положительно, с высокой силой, коррелировало с повышенной частотой различных типов ХА. Среднее значение степени метилирования 41 гена (*RBP7*, *CLCNKB*, *ZNF695*, *TTC39A*, *NET1*, *SIK2*, *C11orf44*, *TNKS1BP1*, *AP5B1*, *EFEMP2*, *ZNF839*, *SLC25A21*, *MPI*, *RFWD3*, *DOC2B*, *IMPA2*, *SMAD2*, *ZNF236-DT*, *VSTM2B*, *APBA3*, *PSG3*, *TMEM160*, *GAD1*, *ACSL3*, *HJURP*, *ZNF2*, *ADRA2B*, *FOXA2*, *FRG1BP*, *CHAF1B*, *AATBC*, *BID*, *PRR14L*, *SLC6A11*, *CRIPAK*, *HELT*, *SLC25A48*, *STC2*, *TCPI0*, *ODF2* и *RNF224*) отрицательно, с высокой силой, коррелировало с повышенной частотой ХА. Перечень топ-10 генов, среднее зна-

Таблица 1. Спектр и частота ХА лимфоцитов крови при остром облучении γ -излучением *in vitro* крови работников СХК

Тип ХА	Частота ХА на 100 клеток, Ме ($Q_{25}; Q_{75}$)		Уровень значимости, p
	контрольная группа, 0 Гр ($n = 10$)	экспериментальная группа, 1.5 Гр ($n = 10$)	
Аберрантные клетки	1.954 (1.329; 2.365)	6.181 (5.611; 6.803)	0.0002
Хроматидные фрагменты	0.660 (0.327; 0.974)	1.570 (0.952; 1.987)	0.0257
Хромосомные фрагменты	0.802 (0.331; 1.307)	1.794 (1.650; 4.923)	0.0073
Кольцевые хромосомы	0.000 (0.000; 0.000)	0.165 (0.000; 0.345)	0.0640
Дицентрические хромосомы	0.312 (0.000; 0.327)	2.397 (1.980; 2.721)	0.0002
Хроматидные обмены	0.000 (0.000; 0.000)	0.000 (0.000; 0.000)	0.7337
Полиплоидные клетки	0.000 (0.000; 0.000)	0.000 (0.000; 0.000)	0.6776

Примечание. Ме — медиана; ($Q_{25}; Q_{75}$) — межквартильный интервал; p — уровень статистической значимости; полужирным шрифтом выделены статистически значимые отличия.

Таблица 2. Гены, среднее значение степени метилирования (b -value) которых положительно коррелирует с повышенной частотой ХА у образцов экспериментальной группы (1.5 Гр)

Тип ХА	Гены (топ-10)	Суммарное количество генов
Аберрантные клетки	<i>CLEC4G, DMBX1, PGAM5, CCDC169-SOHLH2, RHBDF1, SFMBT1, PKD1L1, LOC100129316, NUDT11, PCBP1-AS1</i>	269
Хроматидные фрагменты	<i>NPFFR1, SLC25A52, WIZ, SMAD1, HCG11, DMRT1, C1orf87, ATP12A, CPLX3, TK2</i>	331
Хромосомные фрагменты	<i>HTR2A, ANPEP, THEG5, SFMBT1, CTNND2, MOV10, TMEM38A, GPAAI, ODF3, LGALS</i>	342
Кольцевые хромосомы	<i>FAM173A, GNPAT, LRRC37A6P, SLC16A9, STX3, C11orf80, CCDC159, MICOS13, MZT2B, METTL21A</i>	436
Дицентрические хромосомы	<i>SMTNL1, CIB1, ADCY8, PPM1J, BCAN, DMBX1, TSPAN15, DTX1, HTR2A, PSMC1</i>	370

Примечание. Перечень топ-10 генов включает гены, корреляция которых составляет от +0.7 до +1.0.

чение степени метилирования которых показало положительную и отрицательную корреляцию высокой силы с повышенной частотой ХА, представлен в табл. 2 и 3 соответственно.

Анализ GO биологических процессов показал, что гены, метилирование ДНК которых положительно коррелирует с повышенной частотой ХА, в основном участвуют в пролиферации половых клеток, организации внутренней мембраны митохондрий, проникновении вируса в клетки, регуляции локализации белка на плазматической мембране и организации плазмолеммы, регуляции сужения сосудов и ангиогенеза, а также в различных клеточных реакциях и процессах развития. В свою очередь, гены, метилирование которых отрицательно коррелирует с повышенной частотой ХА, участвуют в транспорте кофакторов ферментов, регуляции концентрации и секреции гормонов, реакции на воздействие ИИ, регуляции перехода G1/S клеточного цикла, регуляции

транскрипции и др. На рис. 1 показаны топ-20 биологических процессов для генов, значение метилирования ДНК которых положительно и отрицательно коррелирует с повышенной частотой ХА.

ОБСУЖДЕНИЕ

Анализ спектра и частоты ХА широко используется для оценки радиационного воздействия, маркерами которого являются дицентрические и кольцевые хромосомы (нестабильный тип ХА), а также транслокации (стабильный тип ХА), сохраняющиеся длительное время после облучения [22].

В настоящем исследовании, после облучения крови *in vitro* частота аберрантных клеток и нестабильных типов ХА (хроматидные и хромосомные фрагменты, дицентрические хромосомы) увеличивается, что также наблюдалось в других исследованиях. О.Г. Чердиченко и Е.Г. Губицкая [23]

Таблица 3. Гены, среднее значение степени метилирования (*b*-value) которых отрицательно коррелирует с повышенной частотой ХА у образцов экспериментальной группы (1.5 Гр)

Тип ХА	Гены (топ-10)	Суммарное количество генов
Аберрантные клетки	<i>PHRF1, R3HDM1, CLGN, VEGFA, SCRNI, SIK2, ACADS, CAMKK2, LARP1B, LENEPI</i>	305
Хроматидные фрагменты	<i>PUS1, ZNF385A, ZFYVE1, HERC2, NEO1, ACD, EPB41L3, DAPK3, GFPT1, SNAP25-AS1</i>	401
Хромосомные фрагменты	<i>UBE2J2, AK5, COTL1, CLPTM1L, ZNF282, ANGPT1, ZNF33B, HAUS1, CPT1C, LOC389765</i>	275
Кольцевые хромосомы	<i>ATF3, ABCB10, PER3, IFIT5, SRGAP1, DNM1P35, PIP4K2B, NT5C3B, ZBTB7A, ZDHHC8</i>	342
Дицентрические хромосомы	<i>PHRF1, DAZAP1, ZNF142, RALGAPB, TTC39A, CCDC6, ACADS, SCN4A, TEX2, FAM20A</i>	313

Примечание. Перечень топ-10 генов включает гены, корреляция которых составляет от -0.7 до -1.0 .

отмечали, что наибольший вклад в ХА вносили транслокации, дицентрические и кольцевые хромосомы, частота которых повышалась с увеличением дозы облучения. И.И. Пелевина и соавт. [4] показали, что увеличение количества аберрант-

ных клеток и ХА линейно зависит от дозы облучения. Увеличение частоты хроматидных фрагментов, наблюдавшееся также при остром [4] и хроническом [9] облучении ИИ, способствовало увеличению количества аберрантных клеток после

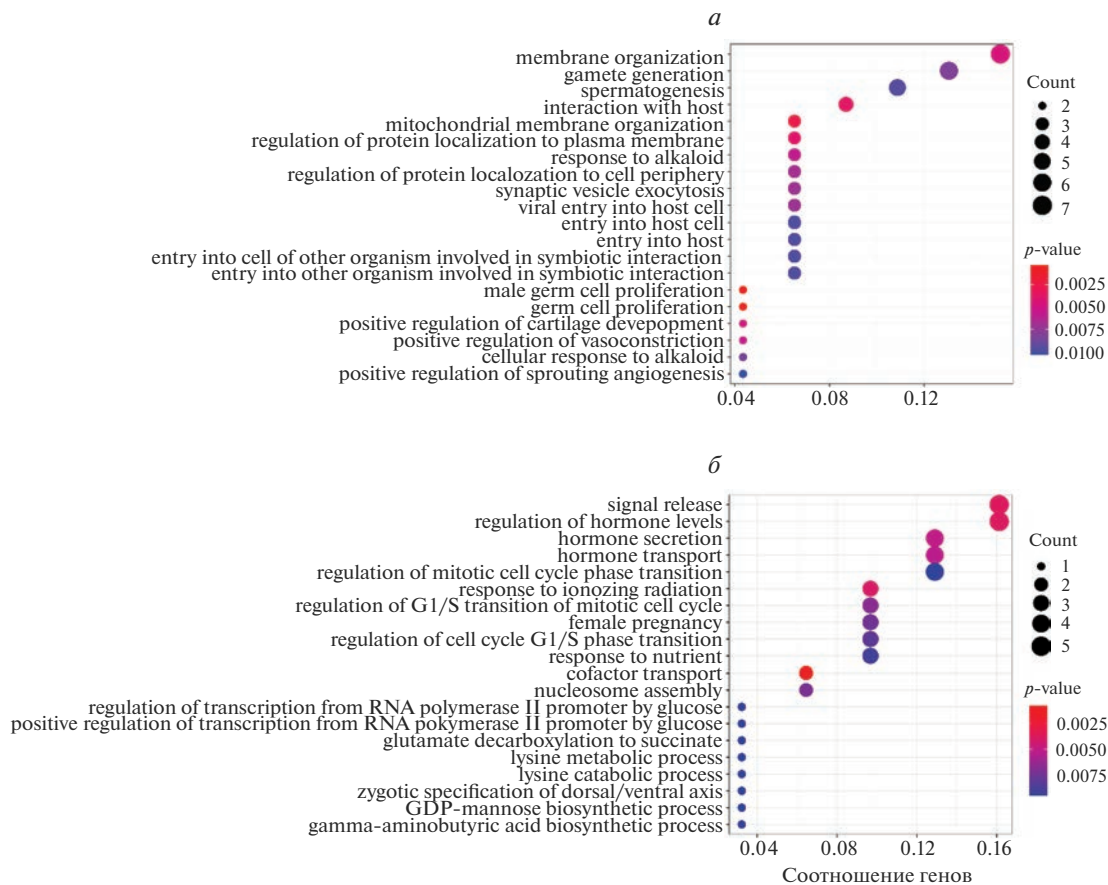


Рис. 1. Топ-20 биологических процессов онтологии генов (GO) с положительной (*а*) и отрицательной (*б*) корреляцией среднего значения метилирования (*b*-value) с частотой ХА в образцах экспериментальной группы. Count — число выявленных генов, которые участвуют в биологическом процессе.

воздействия γ -излучения в дозе 1.5 Гр. Однако радиогенное происхождение хроматидных фрагментов дискуссионно [24].

Корреляция степени метилирования ДНК с частотой ХА. Для исследования участия генов, которые с высокой силой, положительно и отрицательно коррелируют с повышенной частотой ХА в образцах экспериментальной группы, был проведен анализ обогащения по функциональной принадлежности GO. Положительно коррелирующие гены перепредставлены в следующих процессах: пролиферация мужских половых клеток, пролиферация клеток зародыша, организация мембран, проникновение вируса, регуляция сужения сосудов и др., что может свидетельствовать об участии этих генов в таких заболеваниях как мужское бесплодие, сердечно-сосудистые и вирусные инфекционные.

Было отмечено, что *FAN1*, степень метилирования которого положительно коррелирует с повышенной частотой ХА, участвует в реакции на повреждение ДНК и репарации как одноцепочечных, так и двухцепочечных повреждений ДНК [25]. Гипометилирование промотора этого гена и снижение экспрессии *FAN1* отмечалось при патологиях плаценты [26].

Из генов, которые отрицательно коррелируют с повышенной частотой ХА после облучения, представляют интерес *NET1*, *TNKS1BP1* и *RFWD3*, поскольку эти гены принимают участие в восстановлении повреждений двухцепочечной ДНК и межцепочечных перекрестных связей ДНК после воздействия ИИ. Например, белок *NET1* активирует ядерный белок RhoA, который контролирует многие клеточные процессы, включая выживание клеток, экспрессию генов и миграцию. ИИ способствует активации ядерного пула RhoA *NET1*-зависимым способом в ответ на повреждение ДНК [27]. Благодаря нашему исследованию становится понятным механизм активации ядерного пула RhoA в ответ на воздействие ИИ, показанный А.Д. Dubash и соавт. [27], он происходит за счет гипометилирования *NET1*. Поскольку система RhoA является драйвером развития рака и участвует в резистентности и прогрессии опухоли [28], ее активация под действием лучевой терапии за счет гипометилирования *NET1* может обуславливать радиорезистентность многих опухолей. Белок *TNKS1BP1* активируется в ответ на воздействие ИИ, взаимодействует с ДНК-зависимой протеинкиназой и поли(АДФ-рибоза) полимеразой 1 (PARP-1), что приводит к эффективной репарации двухцепочечных разрывов ДНК [29]. *RFWD3* участвует в репликации, обеспечивает успешную гомологичную рекомбинацию двухцепочечных разрывов ДНК после воздействия ИИ, что повышает радиочувствительность клеток [30].

Интерес представляет также *BID*, средняя степень метилирования которого сильно и отрицательно коррелирует с повышенной частотой ХА (хроматидные фрагменты). Проведенное нами ранее исследование выявило только отрицательную корреляцию степени метилирования этого гена с дозой внешнего облучения [17]. Других исследований ассоциации степени метилирования *BID* и частоты ХА не обнаружено.

Таким образом, проведен широкогеномный скрининг генов, метилирование которых коррелирует с повышенной частотой ХА после острого облучения γ -излучением *in vitro*. Корреляцию степени метилирования ДНК с частотой ХА проявили 97 генов, из которых положительная корреляция высокой силы отмечалась у 56 генов, а отрицательная — у 41 гена. Эти гены участвуют в важных клеточных процессах, таких как репарация и репликация ДНК, регуляция экспрессии генов, пролиферация, рост, дифференцировка клеток и др. На гены *NET1*, *TNKS1BP1*, *RFWD3* и *BID*, участвующие в восстановлении повреждений двухцепочечной ДНК после воздействия ИИ, следует обратить внимание, поскольку они могут стать потенциальными маркерами радиационного воздействия.

Изменение степени метилирования генов-индикаторов, отмеченных в настоящем исследовании, с использованием наиболее простого скринингового метода анализа метилирования ДНК — метилчувствительной ПЦР совместно с цитогенетическим анализом может быть перспективным для оценки индивидуальной радиочувствительности и предрасположенности к развитию радиационно-индуцированных заболеваний у персонала, контактирующего в процессе профессиональной деятельности с ИИ.

Проведение дополнительного исследования на выборке с разными спектрами доз (от 0 до 1.5 Гр) позволит выявить дозозависимые изменения степени метилирования ДНК, что может быть положено в основу биологической дозиметрии нового поколения.

Авторы благодарны И.А. Милойчиковой и Н.Д. Тургуновой (медицинским физикам отделения лучевой терапии научно-исследовательского института онкологии Томского национального исследовательского медицинского центра Российской академии наук) за помощь в облучении крови.

Исследование проводилось на базе Северского биофизического научного центра, Северск (забор крови у доноров, культивирование клеток, выделение ДНК), Научно-исследовательского института онкологии, Томск (облучение крови γ -терапевтическим аппаратом Theratron Equinox) и Медико-генетического научного центра им. академика Н.П. Бочкова, Москва (широкогеномное секвенирование методом XmaI-RRBS).

Исследование выполнено в рамках Федеральной целевой программы “Обеспечение ядерной и радиационной безопасности на 2016–2020 годы и на период до 2035 года”, Государственный контракт № 56.003.21.2 от 15.06.2021.

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

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Кужахметова Д.А. Механизмы образования радиационно-индуцированных повреждений хромосом // Вестник совета молодых ученых и специалистов Челябинской области. 2020. Т. 1. № 2(29). С. 18–24.
2. Бочков Н.П. Хромосомы человека и облучение. М.: Атомиздат, 1971. 168 с.
3. Пелевина И.И., Алещенко А.В., Антошина М.М. и др. Молекулярные и клеточные последствия аварии на ЧАЭС // Радиационная биология. Радиоэкология. 2011. Т. 51. № 1. С. 154–161.
4. Пелевина И.И., Алещенко А.В., Антошина М.М. и др. Хромосомные aberrации как показатель реакции лимфоцитов периферической крови человека на облучение *in vitro* при наличии в организме злокачественной опухоли // Радиационная биология. Радиоэкология. 2016. Т. 56. № 6. С. 565–569.
5. Pajic J., Rakic B., Rovcanin B. et al. Inter-individual variability in the response of human peripheral blood lymphocytes to ionizing radiation: Comparison of the dicentric and micronucleus assays // Radiation and Environmental Biophysics. 2015. V. 54. P. 317–325.
6. Pajic J., Rovcanin B., Kekic D. et al. The influence of redox status on inter-individual variability in the response of human peripheral blood lymphocytes to ionizing radiation // Intern. J. Rad. Biology. 2018. V. 94. № 6. P. 569–575.
<https://doi.org/10.1007/s00411-015-0596-3>
7. Druzhinin V.G., Sinitsky M.Yu., Larionov A.V. et al. Assessing the level of chromosome aberrations in peripheral blood lymphocytes in long-term resident children under conditions of high exposure to radon and its decay products // Mutagenesis. 2015. V. 30. № 5. P. 677–683.
<https://doi.org/10.1093/mutage/gev029>
8. Qian Q.-Z., Cao X.-K., Shen F.-H., Wang Q. Effects of ionising radiation on micronucleus formation and chromosomal aberrations in Chinese radiation workers // Rad. Protection Dosimetry. 2016. V. 168. № 2. P. 197–203.
<https://doi.org/10.1093/rpd/ncv290>
9. Литвяков Н.В., Фрейдин М.Б., Халюзова М.В. и др. Частота и спектр цитогенетических нарушений у работников Сибирского химического комбината // Радиационная биология. Радиоэкология. 2014. Т. 54. № 3. С. 283–296.
10. Shi L., Fujioka K., Sakurai-Ozato N. et al. Chromosomal abnormalities in human lymphocytes after computed tomography scan procedure // Radiat. Res. 2018. V. 190. № 4. P. 424–432.
<https://doi.org/10.1667/RR14976.1>
11. Franco R., Schoneveld O., Georgakilas A.G., Panayiotidis M.I. Oxidative stress, DNA methylation and carcinogenesis // Cancer Lett. 2008. V. 266. № 1. P. 6–11.
<https://doi.org/10.1016/j.canlet.2008.02.026>
12. Kuzmina N.S., Lapteva N.Sh., Rubanovich A.V. Hypermethylation of gene promoters in peripheral blood leukocytes in humans long term after radiation exposure // Environmental Res. 2016. V. 146. P. 10–17.
<https://doi.org/10.1016/j.envres.2015.12.008>
13. Кузьмина Н.С., Лантева Н.Ш., Русинова Г.Г. и др. Дозовая зависимость гиперметилирования промоторов генов в лейкоцитах крови лиц, подвергшихся облучению в результате профессиональной деятельности // Соврем. проблемы радиац. генетики: Материалы Росс. конф. с международным участием, 27–28 июня. Дубна, 2019. С. 76–78.
14. Kennedy E.M., Powell D.R., Li Z. et al. Galactic cosmic radiation induces persistent epigenome alterations relevant to human lung cancer // Sci. Rep. 2018. V. 8. № 1. P. 6709.
<https://doi.org/10.1038/s41598-018-24755-8>
15. Lee Y., Kim Y.J., Choi Y.J. et al. Radiation-induced changes in DNA methylation and their relationship to chromosome aberrations in nuclear power plant workers // Intern. J. Radiat. Biol. 2015. V. 91. № 2. P. 142–149.
<https://doi.org/10.3109/09553002.2015.969847>
16. Цымбал О.С., Исубакова Д.С., Брониовская Е.В. и др. Роль метилирования *Bak1* в индукции хромосомных aberrаций при хроническом низкоинтенсивном внешнем облучении // Мед. радиология и радиац. безопасность. 2020. Т. 65. № 5. С. 29–34.
17. Исубакова Д.С., Цымбал О.С., Брониовская Е.В. и др. Метилирование промоторов генов апоптоза в лимфоцитах крови работников, подвергавшихся профессиональному внешнему облучению // Бюлл. эксперим. биологии и медицины. 2021. Т. 171. № 3. С. 339–343.
18. Танас А.С., Кузнецова Е.Б., Борисова М.Э. и др. Дизайн метода бисульфитного секвенирования ограниченных наборов геномных локусов (RRBS) для анализа метилирования CpG-островков человека в больших выборках // Мол. биология. 2015. Т. 49. № 4. С. 689–699.
19. Tanas A.S., Borisova M.E., Kuznetsova E.B. et al. Rapid and affordable genome-wide bisulfite DNA sequencing by XmaI-reduced representation bisulfite sequencing // Epigenomics. 2017. V. 9. № 6. P. 833–847.
<https://doi.org/10.2217/epi-2017-0031>
20. Suomi T., Seyednasrollah F., Jaakkola M.K. et al. ROTS: An R package for reproducibility-optimized statistical testing // PLoS Comput. Biol. 2017. V. 13. № 5. P. e1005562.
<https://doi.org/10.1371/journal.pcbi.1005562>

21. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. 2020. URL: <https://www.R-project.org/>
22. Аклеев А.В., Шведов В.Л., Костюченко В.А. Медико-биологические и экологические последствия радиоактивного загрязнения реки Теча. М.: Медбио-экстрем. 2001. 530 с.
23. Чередниченко О.Г., Губицкая Е.Г. Цитогенетический анализ медицинских работников контактирующих с источниками ионизирующей радиации // Научно-техн. журн. нац. ядерного центра Республики Казахстан. 2016. Т. 1. № 65. С. 112–117.
24. Севаньяев А.В., Деденков А.Н. Актуальные проблемы современной радиобиологии в свете оценки и прогнозирования последствий аварии на Чернобыльской АЭС // Радиобиология. 1990. Т. 30. № 5. С. 579–582.
25. Jin H., Cho Y. Structural and functional relationships of FAN1 // DNA Repair. 2017. V. 56. P. 135–143. <https://doi.org/10.1016/j.dnarep.2017.06.016>
26. Norton C., Clarke D., Holmstrom J. et al. Altered epigenetic profiles in the placenta of preeclamptic and intra-uterine growth restriction patients // Cells. 2023. V. 12. № 8. P. 1130. <https://doi.org/10.3390/cells12081130>
27. Dubash A.D., Guilluy Ch., Srougi M.C. et al. The small GTPase RhoA localizes to the nucleus and is activated by Net1 and DNA damage signals // PLoS One. 2011. V. 6. № 2. P. e17380. <https://doi.org/10.1371/journal.pone.0017380>
28. Schaefer A., Der Ch.J. RHOA takes the RHOad less traveled to cancer // Trends in Cancer. 2022. V. 8. № 8. P. 655–669. <https://doi.org/10.1016/j.trecan.2022.04.005>
29. Zou L.-H., Shang Z.-F., Tan W. et al. TNKS1BP1 functions in DNA double-strand break repair through facilitating DNA-PKcs autophosphorylation dependent on PARP-1 // Oncotarget. 2015. V. 6. № 9. P. 7011–7022. <https://doi.org/10.18632/oncotarget.3137>
30. Katsuki Y., Jeggo P.A., Uchihara Y. et al. DNA double-strand break end resection: A critical relay point for determining the pathway of repair and signaling // Genome Instability & Disease. 2020. V. 1. P. 155–171. <https://doi.org/10.1007/s42764-020-00017-8>

Assessing the Association of the Degree of DNA Methylation and the Frequency of Chromosomal Aberrations in Human Lymphocytes in a Single Irradiation of Blood *in vitro*

O. S. Tsymbal^{a,*}, D. S. Isubakova^a, E. V. Bronikovskaya^a, A. F. Nikolaeva^b, V. O. Sigin^b, A. I. Kalinkin^b, Zh. A. Startseva^c, N. V. Litviakov^{a, c}, I. V. Milto^{a, d}, and R. M. Takhauov^{a, d}

^aSeversk Biophysical Research Center of the Federal Medical-Biological Agency, Seversk, 636013 Russia

^bResearch Centre for Medical Genetics, Moscow, 115478 Russia

^cTomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, 634009 Russia

^dSiberian State Medical University, Tomsk, 634050 Russia

*e-mail: olga-tsymbal@mail.ru

The most sensitive biomolecule under radiation exposure is DNA, whose damage manifests itself in the form of chromosomal aberrations (CA). The processes of DNA methylation, which are involved in the regulation of gene expression, replication, DNA repair, etc., are also affected by gamma radiation. The aim of the study was to evaluate the relationship between the degree of DNA methylation and the frequency of CA after acute *in vitro* irradiation of human blood lymphocytes with gamma radiation. The study involved 10 conditionally healthy workers of the Siberian Chemical Combine, in whose blood lymphocytes the degree of methylation of CpG-dinucleotides (wide-genome bisulfite sequencing, XmaI-Reduced representation bisulfite sequencing – XmaI-RRBS) and the frequency of CA (cytogenetic study) after acute *in vitro* blood irradiation with doses of 0 and 1.5 Gy were evaluated. After acute exposure to gamma radiation in lymphocytes, the frequency of aberrant cells, dicentric chromosomes, chromatid and chromosomal fragments increased. Correlation analysis of the status of CpG-dinucleotide methylation and the frequency of CA revealed changes in the degree of methylation of 97 genes, which strongly correlated positively (56 genes) or negatively (41 genes) with an increased frequency of CA. A primary genome-wide screening of genes whose methylation is correlates with a high frequency of CA was carried out. Many of the identified genes are promising as potential markers of radiation exposure and to study the mechanisms of formation of radiosensitivity of the body and radiore-sistance of tumors during radiation therapy.

Keywords: DNA methylation, chromosomal aberrations, acute irradiation, gamma radiation, blood lymphocytes, wide-genome bisulfite sequencing (XmaI-RRBS).

МОЛЕКУЛЯРНЫЙ ЛАНДШАФТ РАКА ПОЛОСТИ РТА У МОЛОДЫХ ПАЦИЕНТОВ¹

© 2023 г. Е. А. Простакишина¹, Т. Д. Дампилова^{1, 2}, Л. А. Кононова², П. С. Ямщиков¹,
М. Р. Патышева¹, Е. С. Колегова^{1, *}, Е. Л. Чойнзонов¹, Е. В. Денисов¹

¹Научно-исследовательский институт онкологии, Томский национальный исследовательский медицинский центр
Российской академии наук, Томск, 634009 Россия

²Сибирский государственный медицинский университет, Томск, 634050 Россия

*e-mail: elenakolegova@oncology.tomsk.ru

Поступила в редакцию 12.05.2023 г.

После доработки 19.06.2023 г.

Принята к публикации 22.06.2023 г.

Рак полости рта (РПР) — наиболее распространенный тип злокачественных новообразований головы и шеи. При этом рак языка (РЯ) представляет собой наиболее часто встречающуюся форму РПР и характеризуется более агрессивным течением и высокой вероятностью прогрессирования. РПР и РЯ считаются заболеваниями пожилых людей, однако их встречаемость среди лиц моложе 45 лет растет с каждым годом. Этиологические факторы и патогенетические механизмы молодого рака неизвестны. В настоящей работе на основании данных базы Атласа генома рака изучены особенности мутационного профиля, метилома, транскриптома, протеома и микробиома РПР и РЯ у молодых пациентов ($n = 127$) по сравнению с возрастными больными. В ткани РПР и РЯ обнаружены снижение мутационной нагрузки, активация Rar1, PI3K-Akt, MAPK, cGMP-PKG сигнальных путей и сигналинга Fc-гамма R-опосредованного фагоцитоза и специфический микробиомный профиль. В отличие от РПР для РЯ характерны активация сигнальных путей JAK-STAT, иммунного ответа на инфекционные и паразитарные заболевания и PD-L1/PD-1-опосредованной иммуносупрессии и ингибирование сигналинга фагоцитоза. Полученные результаты свидетельствуют о том, что РПР у молодых пациентов обладает молекулярными особенностями, отличными от таковых у больных старше 45 лет, при этом РЯ отличается от других локализаций РПР по молекулярному профилю и, вероятно, должен рассматриваться как отдельная клиническая форма.

Ключевые слова: рак полости рта, молодой возраст, геном, эпигеном, транскриптом, протеом, микробиом.

DOI: 10.31857/S0016675823110103, **EDN:** NMECVY

Рак головы и шеи занимает седьмое место среди онкологических нозологий во всем мире, почти в половине случаев поражая полость рта [1]. Ежегодно во всем мире диагностируется около 350000 новых случаев рака полости рта (РПР) и умирает более 180000 больных, большинство из которых находится в трудоспособном возрасте [1, 2]. К классическим факторам риска развития рака ротовой полости относят употребление табака и алкоголя, а также вирус папилломы человека (ВПЧ) [3].

Из всех видов РПР рак языка (РЯ) считается более агрессивным и чаще прогрессирующим [4]. В настоящее время РЯ предлагают рассматривать как отдельную клиническую форму [4, 5]. Несмотря

на постоянное усовершенствование методов диагностики и лечения, пятилетняя выживаемость при РЯ не превышает 50% [4]. Возникновение РЯ также связывают с употреблением табака (курительный и жевательный), алкоголя и жевательных орехов бетеля [4].

Первоначально рак головы и шеи считался болезнью пожилых людей. Однако у лиц моложе 45 лет заболеваемость РПР в общем и РЯ в частности с каждым годом растет [6–8]. К группе риска относят людей европеоидной и монголоидной рас женского пола [6, 7]. При этом клиническое течение заболевания у молодых людей довольно неоднородно и отличается быстрым прогрессированием по сравнению с людьми старшего возраста [9]. Классические факторы рака головы и шеи, такие как табак, алкоголь и ВПЧ-положительный статус, не связаны с развитием РПР в молодом возрасте [3]. Различные исследования показали специфические

¹ Дополнительная информация для этой статьи доступна по doi 10.31857/S0016675823110103 для авторизованных пользователей.

Таблица 1. Клинико-морфологические характеристики пациентов с РПР и РЯ

Характеристики		РПР			РЯ		
		возраст ≤45 лет, число случаев (%) N = 18	возраст >45 лет, число случаев (%) N = 109	p-value	возраст >45 лет, число случаев (%) N = 15	возраст ≤45 лет, число случаев (%) N = 70	p-value
Стадия	I	3 (16.6)	12 (11.0)	0.212	3 (20.0)	10 (14.3)	0.299
	II	1 (5.6)	24 (22.0)		1 (6.7)	16 (22.8)	
	III	1 (5.6)	14 (12.8)		1 (6.7)	10 (14.3)	
	IV	12 (66.6)	50 (45.9)		9 (59.9)	27 (38.6)	
	NA	1 (5.6)	9 (8.3)		1 (6.7)	7 (10.0)	
Пол	мужской	13 (72.2)	66 (60.6)	0.494	10 (66.7)	42 (60.0)	0.850
	женский	5 (27.8)	43 (39.4)		5 (33.3)	28 (40.0)	
Размер опухоли (Т)	T1	4 (22.2)	14 (12.8)	0.876	4 (26.7)	12 (17.1)	0.646
	T2	5 (27.8)	33 (30.3)		5 (33.3)	22 (31.4)	
	T3	4 (22.2)	31 (28.4)		2 (13.3)	22 (31.4)	
	T4	4 (22.2)	25 (22.9)		3 (20.0)	9 (12.9)	
	Tx	1 (5.6)	6 (5.6)		1 (6.7)	5 (7.2)	
Метастатическое поражение лимфоузлов (N)	N0	6 (33.3)	64 (58.7)	0.197	5 (33.3)	38 (54.3)	0.483
	N2	10 (55.6)	34 (31.2)		8 (53.4)	24 (34.3)	
	Nx	2 (11.1)	11 (10.1)		2 (13.3)	8 (11.4)	
Степень дифференцировки (G)	G1	6 (33.3)	14 (12.8)	0.054	5 (33.3)	7 (10.0)	0.543
	G2	11 (61.1)	73 (67.0)		9 (60.0)	50 (71.4)	
	G3	1 (5.6)	21 (19.3)		1 (6.7)	12 (17.1)	
	NA	—	1 (0.9)		—	10 (14.3)	

Примечание. *p*-value указывает на значимость различий между молодыми и возрастными больными согласно критерию χ^2 . NA (not available) — данные не доступны; Nx — регионарные лимфатические узлы не могут быть оценены; Tx — первичная опухоль не может быть оценена.

молекулярные и генетические особенности РПР и РЯ у молодых пациентов [8]. Однако полученные данные противоречивы и получены на небольших группах пациентов.

Цель настоящего исследования — анализ особенностей мутационного профиля, профиля метилирования, транскриптома, протеома и микробиома РПР и РЯ у молодых пациентов.

МАТЕРИАЛЫ И МЕТОДЫ

Группы исследования. На основании базы данных Атласа генома рака (The Cancer Genome Atlas, <https://www.cancer.gov/tcga>) с помощью инструмента cBioPortal [10] в исследование отобрано 127 пациентов с РПР, принадлежащих к европеоидной расе. В работу были включены только пациенты, имеющие отрицательный ВПЧ-статус, представляющие основную долю случаев РПР

среди молодых пациентов. Пациенты были разделены на четыре группы по возрасту на момент постановки диагноза (до и после 45 лет) и типу заболевания (РПР в целом и РЯ в частности). Первая группа включала пациентов до 45 лет включительно с диагнозом РПР, C01-06 по МКБ-10. У пациентов второй группы РПР (C01-06) был диагностирован после 45 лет. В третью группу входили пациенты с РЯ (C01-02) в возрасте до 45 лет включительно. Четвертую группу составили пациенты с РЯ (C01-02) старше 45 лет. Клинико-морфологические параметры пациентов в рамках групп исследования представлены в табл. 1. Не было обнаружено различий между молодыми и возрастными пациентами с РПР и РЯ по полу, стадии заболевания, размеру опухолевого узла (Т), степени метастатического поражения лимфатических узлов (N) и степени дифференцировки опухоли (G).

Генетический анализ. Анализ однонуклеотидных вариантов и аббераций числа копий ДНК проводили с помощью онлайн-инструмента cBioPortal на основании данных полногеномного секвенирования РПР и РЯ, полученных из базы данных Атласа генома рака.

Транскриптомный анализ. Данные РНК-секвенирования ткани РПР и РЯ были получены из базы данных Атласа генома рака. Идентификацию дифференциально-экспрессирующихся генов (ДЭГ) проводили с использованием пакета DESeq2 в среде R. ДЭГ со значениями $\log_2FC \geq 0.75$ (гиперэкспрессия) и $\log_2FC \leq -0.75$ (гипоэкспрессия) и $p_{adj} < 0.05$ использовались в дальнейшем анализе выраженности сигнальных путей. Визуализацию полученных результатов проводили с помощью инструментов ggplot2 (ggplot2, RRID: SCR_014601), pheatmap (pheatmap, RRID: SCR_016418) и онлайн-ресурса Phantasus [11].

Анализ метилирования. Анализ профиля метилирования ДНК тканей РПР и РЯ проводили с помощью онлайн-инструмента cBioPortal на основании микроматричного профилирования, полученных из базы данных Атласа генома рака. Дифференциально-метилированные гены (ДМГ) со значениями $\log_2FC > 0$ (гипометилирование) и $\log_2FC < 0$ (гиперметилирование) и $p_{adj} \leq 0.05$ использовались в дальнейшем анализе выраженности сигнальных путей.

Протеомный анализ. Анализ экспрессии белков в тканях РПР и РЯ проводили с помощью онлайн-инструмента cBioPortal на основании количественного микрочипового профилирования, полученных из базы данных Атласа генома рака. Дифференциально-экспрессирующиеся белки (ДЭБ) со значениями $\log_2FC > 0$ (гиперэкспрессия) и $\log_2FC < 0$ (гипоэкспрессия) и $p \leq 0.05$ использовались в дальнейшем анализе выраженности сигнальных путей.

Анализ микробиома. Анализ микробиомного профиля РПР и РЯ проводили с помощью онлайн-инструмента cBioPortal на основании РНК-секвенирования, полученных из базы данных Атласа генома рака. В исследовании использовался показатель log-cpm, который обозначает логарифм количества прочтений микробных сигнатур, приходящихся на 1 млн картированных прочтений [12].

Функциональное аннотирование генов. Онлайн-инструмент Enrichr использовался для функционального аннотирования ДЭГ, ДМГ, ДЭБ и генов, несущих мутации, с целью идентификации сигнальных путей, характерных для РПР и РЯ у молодых пациентов [13]. Аннотацию сигнальных путей проводили на основании электронной базы данных Киотской энциклопедии генов и геномов (KEGG).

Статистический анализ. Сравнение клинико-патологических характеристик больных РПР и РЯ молодого и старшего возраста проводили с помощью критерия χ^2 . Сравнение экспрессии белков, профиля метилирования и микробиомного профиля РПР и РЯ в зависимости от возраста проводили с помощью t -критерия Стьюдента. Сравнение мутационного профиля РПР и РЯ в зависимости от возраста проводили с помощью точного одностороннего критерия Фишера. Различия принимались статистически значимыми при $p \leq 0.05$ и на уровне тенденции при $p \leq 0.1$.

РЕЗУЛЬТАТЫ

1. Мутационный профиль РПР у молодых пациентов

Рак полости рта. У молодых больных по сравнению с возрастными пациентами обнаружено значимое снижение количества мутаций (mutation count) в ткани опухоли: 45 (33; 58) против 92 (64; 137) соответственно ($p < 0.001$). Также у молодых пациентов была значимо ниже мутационная нагрузка опухоли (tumor mutation burden, TMB): 1.5 (1.1; 1.9) мутаций на мегабазу (Мб) против 3.4 (2.2; 4.8) у возрастных пациентов ($p < 0.001$).

Только у молодых пациентов обнаружены одонуклеотидные варианты (single nucleotide variations, SNVs) в 342 генах; из них нарушения 11 генов встречались более чем у одного пациента: *GZF1*, *HAO1*, *HTT*, *KATNIP*, *KIF4B*, *MAGI3*, *NLRP2*, *PLA2R1*, *SCEL*, *SLF2* и *ZNF473*. Среди аббераций числа копий ДНК (copy number alterations, CNAs) только у молодых пациентов и более чем в одном случае встречались амплификации, затрагивающие гены *ASPHD2*, *CRYBB3*, *CRYBB2*, *CRYBB2P1*, *CRYBA4*, *CRYBB1*, *KIAA1671*, *MYO18B*, *SGSM1* и *TOP1P2*. Кроме того, 12 генов с SNVs (*FAT1*, *GZF1*, *HAO1*, *HTT*, *KATNIP*, *KIF4B*, *MAGI3*, *NLRP2*, *PLA2R1*, *SCEL*, *SLF2*, *ZNF473*) (табл. П1 приведена в Online Resource 1) и 10 генов с CNAs (*BORA*, *DERL3*, *DIS3*, *MZT1*, *SLC2A11*, *AMELX*, *ARHGAP6*, *ASPHD2*, *CRYBA4*, *CRYBB1*) (табл. П3 приведена в Online Resource 1) значимо различались по частоте среди молодых и возрастных пациентов.

Рак языка. У молодых пациентов с РЯ также обнаружено снижение количества мутаций (34 (30; 49) и 81 (58; 124)) и мутационной нагрузки (1.2 (1.0; 1.6) и 3.0 (1.9; 4.3)) по сравнению с возрастными больными ($p < 0.001$).

SNVs в 396 генах обнаружены только у молодых пациентов с РЯ, среди них нарушения 15 генов встречались более чем у одного больного: *ADARB2*, *DCAF4L2*, *GHR*, *HTT*, *ITGA8*, *KATNIP*, *KIF4B*, *MAGI3*, *NLRP2*, *ORC2*, *PLSCR4*, *POLG*, *SCEL*, *SZT2* и *ZNF473*. Среди CNAs только у молодых пациентов и более чем в одном случае обнаружены амплификации и делеции, затрагиваю-

Таблица 2. Топ 10 гипометилированных генов в ткани РПР у пациентов моложе 45 лет

Ген	Регион	log ₂ FC	p.adj
<i>PRDM12</i>	cg09191327	0.14	1.55E-05
<i>SLC7A8</i>	cg27035169	0.13	2.66E-03
<i>PCDHB14</i>	cg06327515	0.13	1.87E-04
<i>MYADM</i>	cg05832051	0.12	1.55E-05
<i>ZNF585B</i>	cg03751813	0.12	0.0341
<i>AQP5</i>	cg23855989	0.11	0.0426
<i>FLJ40125-RTN2</i>	cg00116838	0.11	4.02E-06
<i>TMPRSS2</i>	cg02613803	0.11	4.75E-06
<i>RAP1GAP</i>	cg07138512	0.10	0.0182
<i>BARD1</i>	cg12091944	0.10	6.24E-04

Таблица 3. Топ 10 гипометилированных генов у пациентов с РЯ моложе 45 лет

Ген	Участок метилирования	log ₂ FC	p.adj
<i>ZFP42 (REX1)</i>	cg06274159	0.23	0.0235
<i>NEUROG1</i>	cg04330449	0.21	0.0484
<i>TRPC4</i>	cg21432954	0.19	0.0247
<i>ADD3</i>	cg25341032	0.19	0.0182
<i>TOX2</i>	cg11319389	0.18	5.19E-03
	cg04369341	0.17	4.21E-03
<i>ATP6VIC2</i>	cg22243662	0.17	5.98E-04
<i>GATA4</i>	cg21073927	0.17	5.19E-03
<i>TNFRSF8</i>	cg19896198	0.17	7.50E-03
<i>ST6GALNAC5</i>	cg06201642	0.16	0.0471
<i>FAM5B</i>	cg13589108	0.16	0.0188

щие гены *ASPHD2*, *CRYBB3*, *CRYBB2*, *CRYBB2P1*, *CRYBA4*, *CRYBB1*, *KIAA1671*, *MYO18B*, *SGSM1* и *TOP1P2*. Помимо этого, частота SNVs 17 генов (*MAP1A*, *FAT1*, *ADARB2*, *DCAF4L*, *GHR*, *HTT*, *ITGA8*, *KATNIP*, *KIF4B*, *MAGI3*, *NLRP2*, *ORC2*, *PLSCR4*, *POLG*, *SCEL*, *SZT2* и *ZNF473*) (табл. П2 приведена в Online Resource 1) и CNAs 10 генов (таких же, как у пациентов с РПР) (табл. П3 приведена в Online Resource 1) значимо различалась между молодыми и возрастными больными РЯ.

2. Транскриптомный профиль РПР у молодых пациентов

Рак полости рта. Обнаружено 382 ДЭГ между молодыми и возрастными больными РПР, включая 318 гипохеэкспрессирующихся и 64 гиперхекспрессирующихся генов (таблица приведена в Online Resource 2). Топ 10 генов, гипохеэкспрессирующихся в опухоли молодых пациентов, был представлен

BPIFB1, *RPL10P6*, *PIGR*, *BPIFB2*, *FDCSP*, *ODAM*, *DMBT1*, *TMPRSS3*, *IDO1*, *MUC5B* (p.adj < 0.0003 и log₂FC ≤ -2.1), в то время как топ 10 гиперхекспрессирующихся генов включал *HCG22*, *DCDC2*, *KRT4*, *CFAP58*, *PMEL*, *LINC02901*, *PPP1R3C*, *LGR6*, *FMN2*, *PLP1* (p.adj < 0.004 и log₂FC ≥ 1.1; рис. 1,а, б).

По данным функционального аннотирования для гиперхекспрессирующихся у молодых пациентов генов наблюдалась тенденция активации сигнальных путей функционирования синаптического везикулярного цикла, ГАМК-ергических синапсов и секреции продуктов поджелудочной железы (рис. 1,в). В то же время гипохеэкспрессирующиеся гены вовлечены в сигнальные пути взаимодействия цитокиновых рецепторов, каскада системы комплемента и свертывания крови, нарушения регуляции транскрипции при раке, JAK-STAT, метаболизма триптофана и ряда иммунных реакций (рис. 1,г).

Рак языка. У молодых больных РЯ был дифференциально экспрессирован 91 ген, включая 66 гипохеэкспрессирующихся и 25 гиперхекспрессирующихся (таблица приведена в Online Resource 2). Топ 10 генов, гипохеэкспрессирующихся в опухоли молодых пациентов, был представлен *BPIFB1*, *RPL10P6*, *FDCSP*, *PIGR*, *DMBT1*, *MUC5B*, *DMRTA2*, *ODAM*, *ALGIL1P*, *KRT7* (p.adj < 0.007 и log₂FC ≤ -1.2), а топ 10 гиперхекспрессирующихся генов включал *HCG22*, *ALB*, *PMEL*, *KRT4*, *PPP1R3C*, *EMPI1*, *CLGN*, *THSD4*, *PLP1*, *FMN2* (p.adj < 0.03 и log₂FC ≥ 1.2; рис. 2,а, б).

Гиперхекспрессирующиеся в опухоли молодых пациентов гены на уровне тенденции задействованы в процессах регуляции синтеза гормонов щитовидной железы (рис. 2,в). В то же время гипохеэкспрессирующиеся гены вовлечены в сигнальные пути взаимодействия цитокиновых рецепторов и в процессы, характерные для коклюша и туберкулеза (рис. 2,г).

3. Профиль метилирования РПР у молодых пациентов

Рак полости рта. Между тканями РПР молодых и возрастных больных обнаружен 161 ДМГ (p.adj ≤ 0.05; рис. 3,а) (таблица приведена в Online Resource 3). Топ 10 гипометилированных генов представлен *PRDM12*, *SLC7A8*, *PCDHB14*, *MYADM*, *ZNF585B*, *AQP5*, *FLJ40125-RTN2*, *TMPRSS2*, *RAP1GAP* и *BARD1* (табл. 2; рис. 3,б-л), тогда как гиперметилирование затрагивало только генный комплекс *DTX3L-PRP9* (рис. 3,м).

Рак языка. Между тканями РЯ молодых и возрастных больных обнаружено 226 ДМГ (p.adj ≤ 0.05; рис. 4,а) (таблица приведена в Online Resource 3). Топ 10 гипометилированных генов представлен *ZFP42 (REX1)*, *NEUROG1*, *TRPC4*, *ADD3*, *TOX2*, *ATP6VIC2*, *GATA4*, *TNFRSF8*,

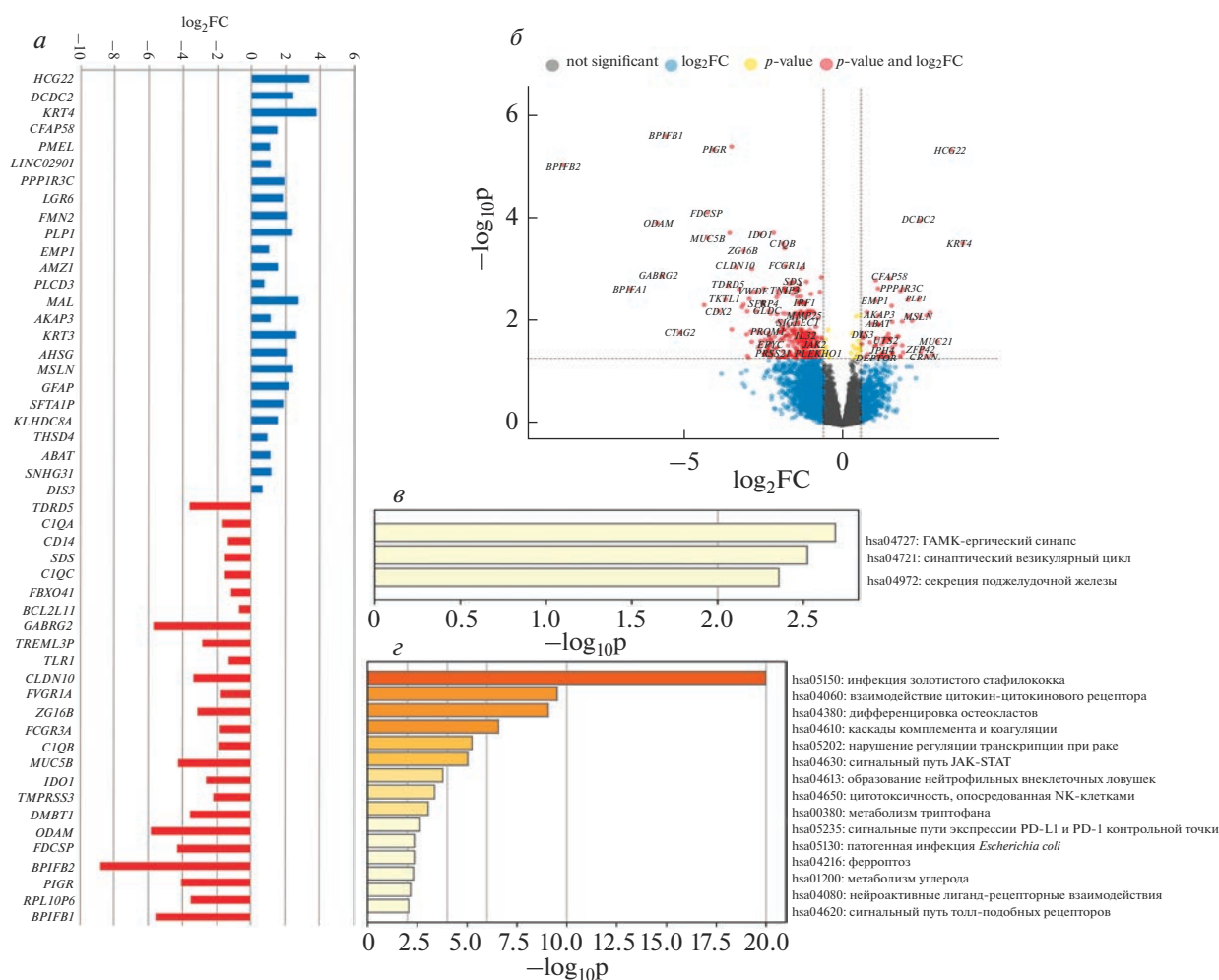


Рис. 1. Различия транскриптомного профиля генов РПП между молодыми и возрастными пациентами. *а, б* — тепловая карта и диаграмма рассеяния разности экспрессии (volcano plot) генов; *в* — сигнальные пути KEGG, выраженные на основе гиперэкспрессирующихся генов, у молодых пациентов ($p_{adj} \leq 0.1$); *г* — сигнальные пути KEGG, выраженные на основе гипоекспрессирующихся генов, у молодых пациентов ($p_{adj} \leq 0.05$).

ST6GALNAC5, *FAM5B* (табл. 3; рис. 4, *б–м*). Статистически значимых гиперметилированных генов не обнаружено у молодых пациентов.

4. Протеомный профиль РПП у молодых пациентов

Рак полости рта. По сравнению с возрастными больными у пациентов молодого возраста с РПП обнаружена гиперэкспрессия белков BAK1, PRKCB_PS660, MYC, RICTOR_PT1135, STAT3_PY705, RICTOR, BCL2L11, PEA15_PS116, CHEK1_PS345, PRKCD_PS664, NOTCH1, RPS6KA1_PT359_S363, YWHAЕ, TSC2_PT1462, PRKCA_PS657, AKT3_PS473, AKT2_PS473, AKT1_PS473 и гипоекспрессия TP53 (рис. 5, *а*).

Рак языка. По сравнению с возрастными больными у пациентов молодого возраста с РЯ обнару-

жена гипоекспрессия белков RAB11A и RAB11B (рис. 5, *б*) и гиперэкспрессия BAK1, PRKCB_PS660, MYC, RICTOR_PT1135, STAT3_PY705, RICTOR, PEA15_PS116, RPS6KA1_PT359_S363, NOTCH1, TSC2_PT1462, AKT3_PS473, AKT2_PS473, AKT1_PS473 (рис. 5, *б*).

5. Микробиомный профиль РПП у молодых пациентов

Рак полости рта. Статистически значимые различия микробиомного профиля РПП между молодыми и возрастными пациентами выявлены для 25 родов бактерий. Среди них шесть родов преобладали у возрастных пациентов: *Ilumato-bacter*, *Lachnobacterium*, *Alicyclobacillus*, *Hylemonella*, *Desulfotomaculum*, *Reyranella*. В то время как остальные 19 родов чаще встречались у молодых больных и включали представителей *Allosalinalacti-*

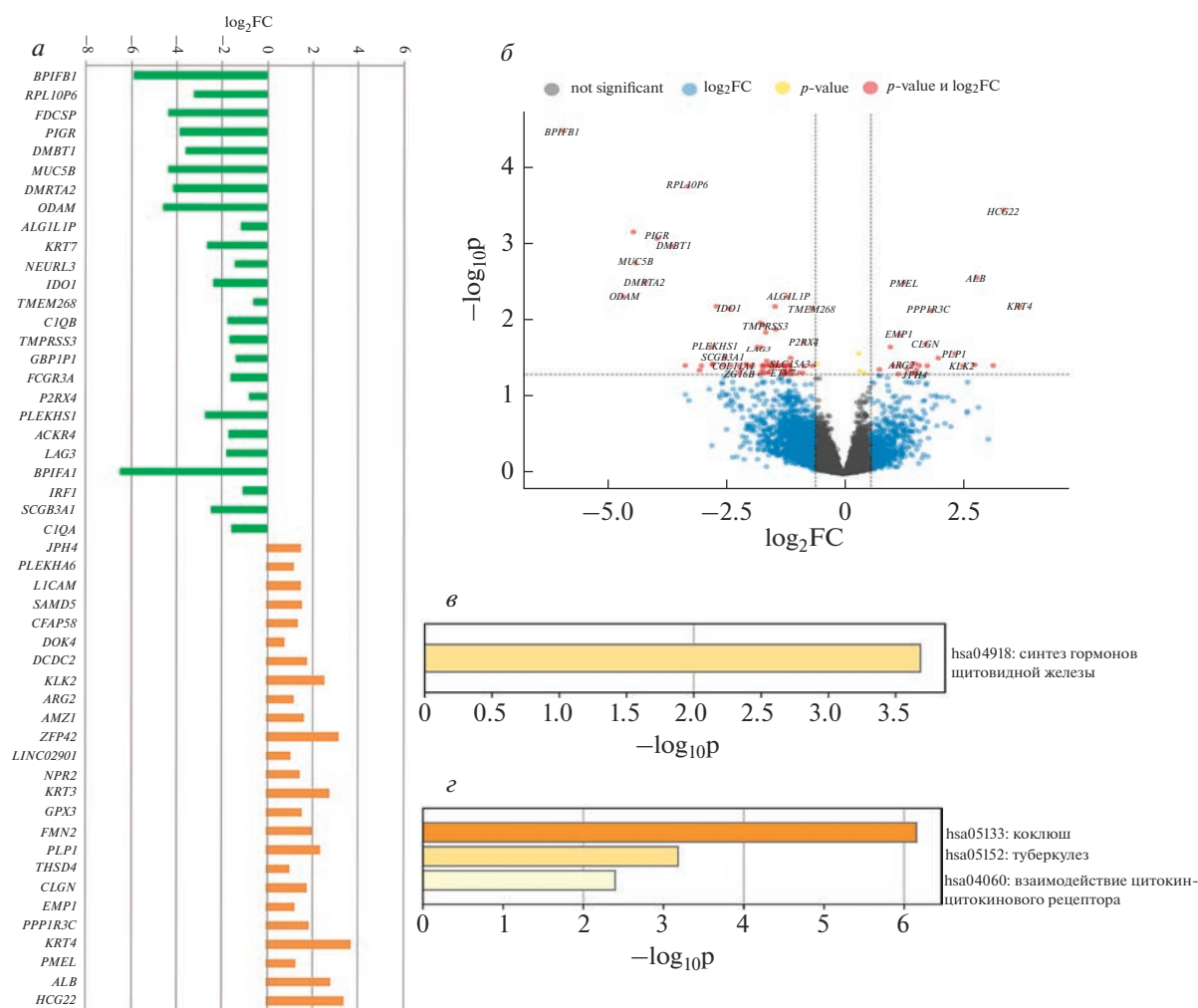


Рис. 2. Различия транскриптомного профиля генов РЯ между группами молодых и возрастных пациентов. а, б – тепловая карта и диаграмма рассеяния разности экспрессии (volcano plot) генов; в – сигнальные пути KEGG, выраженные на основе гиперэкспрессирующихся генов, у молодых пациентов ($p_{adj} \leq 0.1$); г – сигнальные пути KEGG, выраженные на основе гипоекспрессирующихся генов, у молодых пациентов ($p_{adj} \leq 0.05$).

nospora, Riemeirella, Ulvibacter, Zooshikella, Demequina, Desulfovermiculus, Ruegeria, Cylindrospermopsis, Rhodoferrax, Kyrpidia, Mastigocladopsis, Ethanologenens, Thiorhodospira, Ferrofum, Arcobacter, Bibersteinia, Cesiribacter, Haemophilus и Actibacterium.

Рак языка. Для молодых и возрастных пациентов с РЯ обнаружены отличия в 26 родах бактерий. Из них пять родов преобладали у возрастных пациентов, в частности *Ferrimonas, Lachnobacterium, Hylemonella, Thermacetogenium* и *Haematobacter*, тогда как 21 род чаще обнаруживался у молодых пациентов: *Riemeirella, Actibacterium, Allosalinactinospora, Methylobacterium, Ruegeria, Pararhodospirillum, Demequina, Knoellia, Collimonas, Zooshikella, Ferrofum, Actinokineospora, Kyrpidia, Mitsuraria, Arcobacter, Hydrocarboniphaga, Bibersteinia, Chromohalobacter, Haemophilus, Ethanologenens* и *Angustibacter*.

ОБСУЖДЕНИЕ

Актуальной проблемой современной онкологии является увеличение числа онкологических больных молодого возраста, что характерно и для РПР в частности [14]. Несмотря на обновление спектра имеющихся методов терапии, современные схемы лечения молодых больных РПР малоэффективны, что связывают с недостаточным уровнем знаний о молекулярных механизмах развития этой нозологии. При этом из всех видов РПР РЯ считается более агрессивным и чаще прогрессирующим, и предлагается рассматривать его как отдельный тип рака [4]. В связи с этим исследование молекулярных особенностей РПР в общем и РЯ в частности у пациентов молодого возраста является актуальной задачей. В представленной работе изучены молекулярные осо-

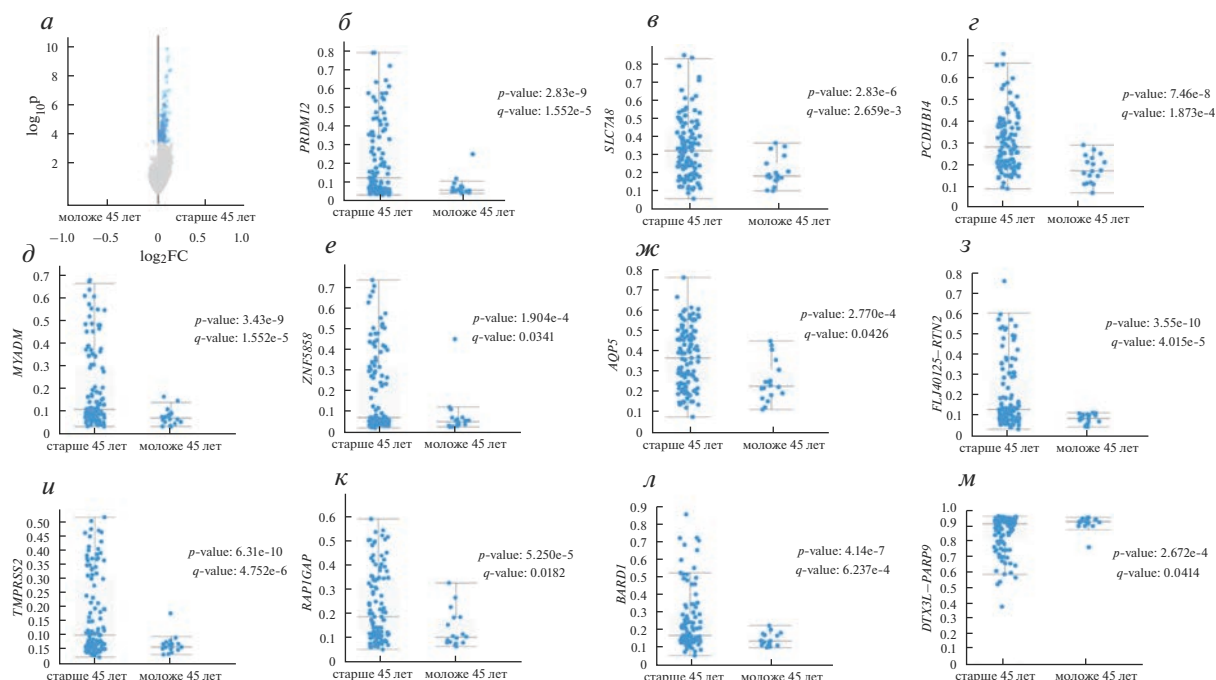


Рис. 3. Различия в профиле метилирования генов РПР (ось ординат) между молодыми и возрастными пациентами. *а* — дифференциально-метилированные гены (синим цветом указаны статистически значимые гены); *б–л* — гипометилированные гены у молодых пациентов; *м* — гиперметилированные гены комплекса *DTX3L–PARP9* у молодых пациентов.

бенности РПР и РЯ на уровне генома, эпигенома, транскриптома, протеома и микробиома.

Для РПР характерен сложный мутационный ландшафт, который включает хромосомные аномалии, aberrации числа копий ДНК и генные мутации [15]. В настоящем исследовании у молодых пациентов как с РПР, так и с РЯ обнаружено снижение мутационной нагрузки в опухоли по сравнению с пациентами старше 45 лет. В настоящее время хорошо известно, что мутационная нагрузка отражает антигенный профиль опухоли, является предиктором для назначения иммунотерапии, в т.ч. у пациентов с раком головы и шеи, и связана с прогнозом РПР при проведении хирургического лечения [16, 17]. При этом степень мутационной нагрузки ниже 11.5 мутаций/Мб является неблагоприятным прогностическим фактором для общей выживаемости пациентов с РПР, которым не проводилась иммунотерапия [16]. Таким образом, результаты настоящего исследования могут указывать на худший прогноз РПР и РЯ у пациентов молодого возраста, что должно учитываться при проведении лечения данной категории больных.

На уровне транскриптома гипохеэкспрессирующиеся гены в ткани РПР и РЯ у молодых пациентов главным образом вовлечены в регуляцию иммунных реакций и связаны с инфекционными заболеваниями, тогда как гиперэкспрессирующиеся гены в большинстве своем ассоциированы с гомеоста-

зом организма. При этом обнаружена гиперэкспрессия генов *HCG22*, *KRT4* и *EMP1*, которые могут ингибировать пролиферацию, рост и инвазию злокачественных клеток и индуцировать апоптоз в клетках РПР [18–20]. *HCG22* также способен ингибировать сигнальные пути Akt, mTOR и Wnt/ β -катенин, ассоциированные с онкогенезом и прогрессированием РПР, рака желудка и колоректального рака [18]. На основании этих результатов можно заключить, что РПР у молодых больных имеет транскрипционные особенности, отличные от таковых при РПР в старшем возрасте. Наиболее важным, на наш взгляд, является сниженная активность иммунно-воспалительных реакций в РПР у молодых больных.

Профиль метилирования как РПР, так и РЯ у молодых пациентов представлен в основном гипометилированными локусами. Обнаружено гипометилирование генов *AQP5* и *ZNF583B* при РПР, которые по данным литературы гиперметилированы при раке головы и шеи и раке молочной железы у пациентов старше 45 лет [21, 22]. Другие гены, в частности *ZFP42*, *GATA4* и *ST6GALNAC5*, гипометилированные при РЯ у молодых пациентов, ранее были описаны как гиперметилированные при плоскоклеточном раке легкого, глиобластоме и колоректальном раке у лиц старше 55 лет [23–25].

Обнаруженные ДЭБ при РПР у молодых пациентов участвуют главным образом в регуляции клеточного цикла и апоптоза посредством изме-

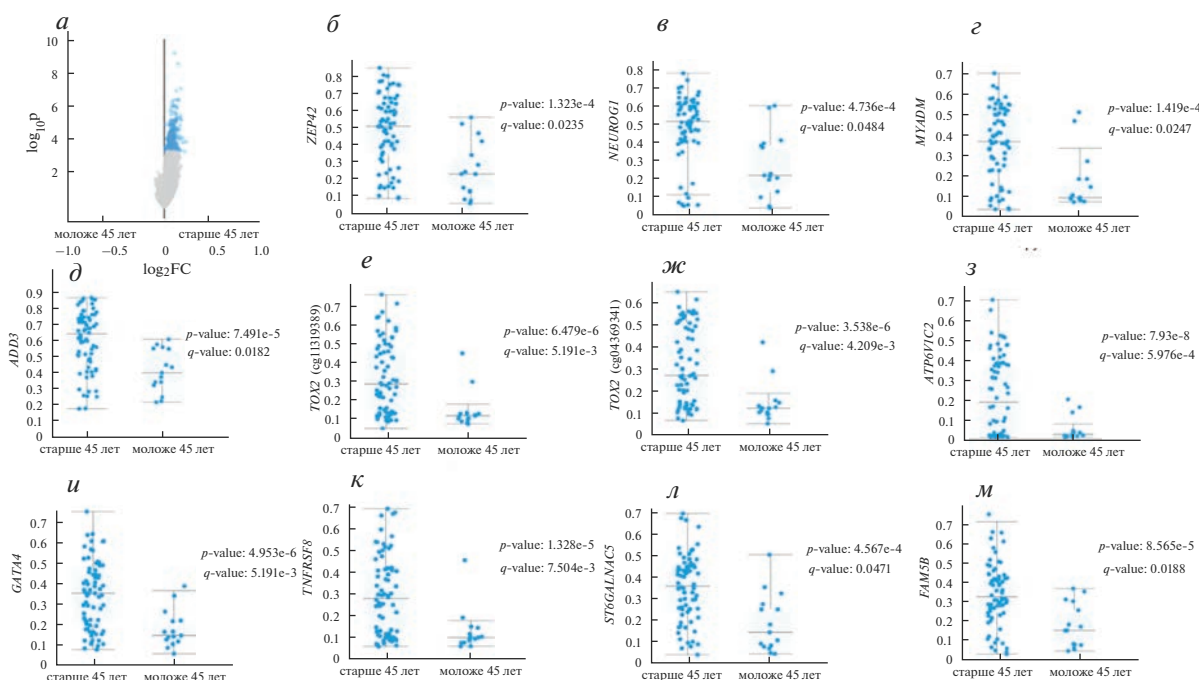


Рис. 4. Различия в профиле метилирования генов опухоли языка (ось ординат) между молодыми и возрастными пациентами. *а* — дифференциально-метилированные гены (синим цветом указаны статистически значимые гены); *б–м* — гипометилированные гены у молодых пациентов.

нения активности генов *TP53* и *BCL2* [26]. Помимо этого *NOTCH1*, гиперэкспрессирующийся в ткани РПР молодых больных, входит в состав комплекса *MAML1–RBP–J*карра–*ICN1*, регулирующего клеточную дифференцировку [27]. Другие белки, гиперэкспрессирующиеся в ткани РПР молодых пациентов, участвуют в эпигенетической регуляции ядерных процессов: *PRKCA* и *PRKCB* регулируют активность гистона *H3T6* [28], а *CHEK1* является треонин-киназой для гистона *H3T11* [29]. Особенностью РЯ у молодых пациентов является гипокспрессия белков *RAB11A* и *RAB11B* по сравнению с возрастными пациентами. Известно, что данные белки активируют пролиферацию и подвижность опухолевых клеток пищевода через сигнальный путь *Wnt* [30].

Интегральный анализ генов с SNVs и CNAs, ДЭГ, ДМГ и ДЭБ позволил обнаружить общие сигнальные пути, характерные для РПР и РЯ у молодых пациентов (рис. 6). Молекулярный профиль РПР и РЯ у молодых больных в основном представлен гиперэкспрессией генов и белков, а также гипометилированием генов, участвующих в активации *Rap1*, *PI3K–Akt*, *MAPK* и *cGMP–PKG* сигнальных путей и *Fc-гамма R*-опосредованного фагоцитоза. Известно, что *PI3K–Akt* индуцирует экспрессию транскрипционных факторов, ответственных за прогрессию клеточного цикла и выживание клеток [31, 32]. *cGMP–PKG*-сигнальный путь подавляет рост и пролиферацию опухолевых

клеток [33]. *MAPK*-сигнальный путь является ключевым игроком в опухолях разных локализаций и вовлечен в регуляцию пролиферации, дифференцировки и миграции клеток [34]. Также активация *MAPK* может приводить к злокачественной трансформации, а экспрессия компонентов данного пути имеет тенденцию к увеличению при РПР [35]. Для РПР у молодых пациентов также характерны *Rap1*-сигнальный путь, способствующий по данным литературы инвазии и миграции клеток рака предстательной железы и плоскоклеточного рака головы и шеи [36], и сигналинг *Fc-гамма R*-опосредованного фагоцитоза, играющего центральную роль в противоопухолевом иммунитете за счет фосфорилирования остатков тирозина иммунных рецепторов [37].

Особенностью РЯ в молодом возрасте является активация сигнальных путей, ответственных за гомеостаз организма и его развитие, например посредством *JAK–STAT*-сигналинга и регуляции метаболизма гормонов и холина [38–41], за функционирование контрольных точек *PD–L1* и *PD-1* и ответ иммунной системы на инфекционные и паразитарные заболевания (болезнь Шагаса) [42–44]. Как следствие активация данных путей влияет на онкогенез и способствует прогрессии рака [39–41]. Интересно, что для РЯ также характерен сигнальный путь, ассоциированный с инфицированием ВПЧ, хотя в исследуемой группе все пациенты имели негативный ВПЧ-статус.



Рис. 5. Дифференциально-экспрессирующиеся белки между молодыми и возрастными пациентами с РПП (а) и РЯ (б). Слева — тепловая карта и справа — линейчатая гистограмма, отражающие средний уровень экспрессии белков у молодых и возрастных пациентов в ткани РПП и РЯ. На тепловой карте красным выделена гипоэкспрессия, зеленым — гиперэкспрессия белков.

Кроме того, в ткани РЯ наблюдается снижение активности рецепторов, ответственных за сборку фагосом. По данным литературы ингибирование фагоцитоза приводит к хроническому воспалению и повреждению тканей, а также обуславливает “ускользание” опухоли от иммунного ответа [45].

За последние десятилетия стала очевидной роль микробиома в поддержании здоровья человека и развитии различных заболеваний. В ряде

работ описан вклад микробиома в развитие заболеваний полости рта, в том числе РПП и РЯ [46, 47]. Тем не менее молекулярные механизмы роли микробиома в формировании РПП непонятны [46]. В настоящем исследовании обнаружено, что в ткани РПП молодых больных чаще представлены бактерии родов *Allosalinactinospora*, *Riemerella*, *Zooshikella*, *Demequina*, *Kyrpidia*, *Ethanoligenens*, *Ferrovum*, *Arcobacter*, *Bibersteinia*, *Haemophilus* и *Ac-*

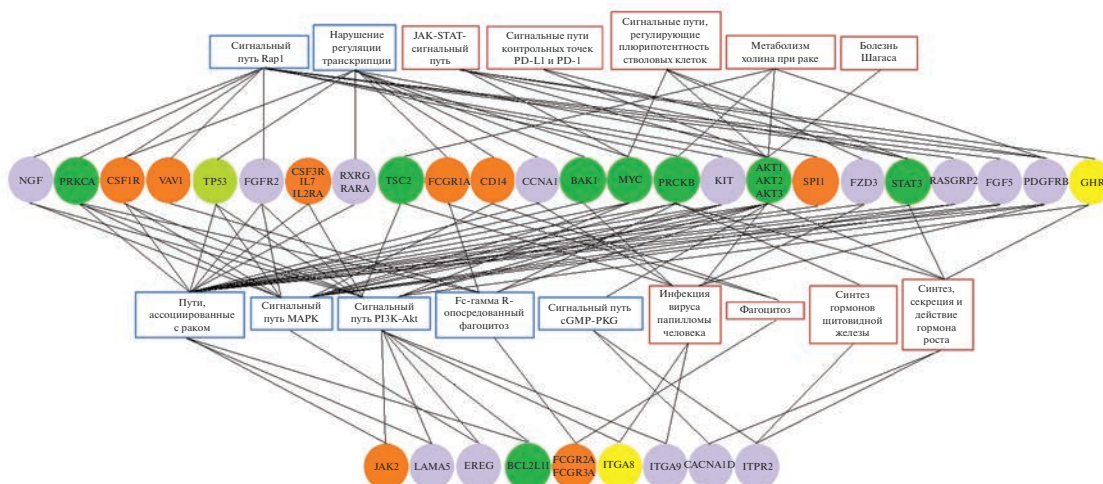


Рис. 6. Сигнальные пути, выраженные при РПР и РЯ у молодых пациентов. В прямоугольниках приведены сигнальные пути, при этом синий цвет указывает на пути, характерные для РПР в целом, красный — для РЯ. Гены и белки, составляющие сигнальные пути, приведены в кружках: желтым — гены с SNVs, фиолетовым — гипометилированные гены, оранжевым — гипоэкспрессирующиеся гены, темно-зеленым — гиперэкспрессирующиеся белки и светло-зеленым — гипоэкспрессирующиеся белки.

tibacterium и реже *Lachnobacterium* и *Hylemonella*. Большинство из перечисленных родов бактерий слабо изучены, особенно в отношении связи с раком. Есть данные о том, что бактерии рода *Haemophilus* в норме населяют ротовую полость [48]. При этом численность бактерий *Haemophilus* снижается у больных РПР по сравнению со здоровыми индивидами, у которых медиана возраста больных с РПР составила 66 лет [49]. Интересно, что микробиота РЯ у молодых пациентов отличается от таковой при других локализациях РПР. Так, в ткани РЯ отмечено увеличение бактерий родов *Methylobacterium*, *Pararhodospirillum*, *Knoellia*, *Collimonas*, *Actinokineospora*, *Mitsuaria*, *Hydrocarboniphaga*, *Chromohalobacter*, *Angustibacter* и снижение *Ferriomonas*, *Thermacetogenium*, *Haematobacter*. Тем не менее информация о роли данных микроорганизмов в развитии и прогрессировании злокачественных новообразований практически отсутствует в доступной литературе. Известно, что для бактерий рода *Chromohalobacter* (*Chromohalobacter salexigens* и *Chromohalobacter israelensis*) и рода *Actinokineospora* характерна противоопухолевая активность [50].

Таким образом, РПР в общем и РЯ в частности у молодых пациентов характеризуются снижением мутационной нагрузки, активацией Rap1, PI3K-Akt, MAPK и cGMP-PKG сигнальных путей и сигналинга Fc-гамма R-опосредованного фагоцитоза, а также специфическими особенностями микробиома. В отличие от РПР для РЯ характерна активация сигнальных путей JAK-STAT, иммунного ответа на инфекционные и паразитарные заболевания и PD-L1/PD-1-опосредованной

иммуносупрессии и ингибирование сигналинга фагоцитоза. Полученные данные свидетельствуют о том, что РПР у молодых пациентов обладает молекулярными особенностями, отличными от таковых у больных старше 45 лет, при этом РЯ отличается от других локализаций РПР по молекулярному профилю и, вероятно, должен рассматриваться как отдельная клиническая форма.

Все данные, использованные в настоящем исследовании, были загружены из общедоступного ресурса Атласа генома рака, а получение и анализ данных осуществлялись в соответствии с рекомендациями Атласа генома рака по публикации и политикой доступа к ним. Таким образом, дополнительного одобрения со стороны местного комитета по этике не требовалось.

Работа выполнена при финансовой поддержке гранта РНФ № 22-15-00308.

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

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

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Slootweg P.J., El-Naggar A.K. World Health Organization 4th edition of head and neck tumor classification: insight into the consequential modifications // Virchows Archiv. 2018. V. 3. P. 311–313. <https://doi.org/10.1007/s00428-018-2320-6>

2. Bray F., Ferlay J., Soerjomataram I. et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries // CA: a cancer journal for clinicians. 2018. V. 6. P. 394–424. <https://doi.org/10.3322/caac.21492>
3. Maroun C.A., Zhu G., Fakhry C. et al. An immunogenomic investigation of oral cavity squamous cell carcinoma in patients aged 45 years and younger // The Laryngoscope. 2021. V. 2. P. 304–311. <https://doi.org/10.1002/lary.28674>
4. Li Z., Jiang C., Yuan Y. TCGA based integrated genomic analyses of ceRNA network and novel subtypes revealing potential biomarkers for the prognosis and target therapy of tongue squamous cell carcinoma // PLoS One. 2019. V. 5. P. e0216834. <https://doi.org/10.1371/journal.pone.0216834>
5. Karatas O.F., Oner M., Abay A. et al. MicroRNAs in human tongue squamous cell carcinoma: From pathogenesis to therapeutic implications // Oral Oncology. 2017. P. 124–130. <https://doi.org/10.1016/j.oraloncology.2017.02.015>
6. Dos Santos Costa S.F., Brennan P.A., Gomez R.S. et al. Molecular basis of oral squamous cell carcinoma in young patients: Is it any different from older patients? // J. Oral Pathology & Med.: off. publ. Intern. association of oral pathologists and Am. Acad. oral pathology. 2018. V. 6. P. 541–546. <https://doi.org/10.1111/jop.12642>
7. Gonzalez M., Riera March A. Tongue Cancer, StatPearls Publ., 2023. Treasure Island (FL).
8. Kolegova E.S., Patysheva M.R., Larionova I.V. et al. Early-onset oral cancer as a clinical entity: Aetiology and pathogenesis // Intern. J. Oral and Maxillofacial Surgery. 2022. V. 12. P. 1497–1509. <https://doi.org/10.1016/j.ijom.2022.04.005>
9. Федорова И.К., Кульбакин Д.Е., Колегова Е.С. и др. Лечение молодых пациентов с раком полости рта: текущее состояние и перспективы (консорциум “Этиология и патогенез рака полости рта у пациентов молодого возраста”). Часть 1 // Голова и шея. 2022. Т. 10. С. 83–91. <https://doi.org/10.25792/HN.2022.10.4.83-91>
10. Cerami E., Gao J., Dogrusoz U. et al. The cBio cancer genomics portal: An open platform for exploring multi-dimensional cancer genomics data // Cancer Discovery. 2012. V. 5. P. 401–404. <https://doi.org/10.1158/2159-8290.cd-12-0095>
11. Kleverov M., Zenkova D., Kamenev V. et al. Phantasus: Web-application for visual and interactive gene expression analysis // bioRxiv. 2022. <https://doi.org/10.1101/2022.12.10.519861>
12. Poore G.D., Kopylova E., Zhu Q. et al. Microbiome analyses of blood and tissues suggest cancer diagnostic approach // Nature. 2020. V. 7800. P. 567–574. <https://doi.org/10.1038/s41586-020-2095-1>
13. Kuleshov M.V., Jones M.R., Rouillard A.D. et al. Enrichr: A comprehensive gene set enrichment analysis web server 2016 update // Nucl. Acids Res. 2016. V. W1. P. W90–W97. <https://doi.org/10.1093/nar/gkw377>
14. Ugai T., Sasamoto N., Lee H.-Y. et al. Is early-onset cancer an emerging global epidemic? Current evidence and future implications // Nat. Rev. Clin. Oncology. 2022. V. 10. P. 656–673. <https://doi.org/10.1038/s41571-022-00672-8>
15. Campbell B.R., Chen Z., Faden D.L. et al. The mutational landscape of early- and typical-onset oral tongue squamous cell carcinoma // Cancer. 2021. V. 4. P. 544–553. <https://doi.org/10.1002/cncr.33309>
16. Moreira A., Poulet A., Masliah-Planchon J. et al. Prognostic value of tumor mutational burden in patients with oral cavity squamous cell carcinoma treated with upfront surgery // ESMO Open. 2021. V. 4. P. 100178. <https://doi.org/10.1016/j.esmoop.2021.100178>
17. Marabelle A., Fakih M., Lopez J. et al. Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: Prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study // The Lancet. Oncology. 2020. V. 10. P. 1353–1365. [https://doi.org/10.1016/s1470-2045\(20\)30445-9](https://doi.org/10.1016/s1470-2045(20)30445-9)
18. Wang M., Feng Z., Li X. et al. Assessment of multiple pathways involved in the inhibitory effect of HCG22 on oral squamous cell carcinoma progression // Mol. Cell. Biochemistry. 2021. V. 6. P. 2561–2571. <https://doi.org/10.1007/s11010-021-04091-8>
19. Wang L., Ma X., Yu J. et al. Negative regulation of miR-1288-3p/KRT4 axis through a circular RNA in oral cancer // J. Biochem. Mol. Toxicology. 2022. V. 8. P. e23118. <https://doi.org/10.1002/jbt.23118>
20. Zhang J., Cao W., Xu Q. et al. The expression of EMP1 is downregulated in oral squamous cell carcinoma and possibly associated with tumour metastasis // J. Clin. Pathology. 2011. V. 1. P. 25–29. <https://doi.org/10.1136/jcp.2010.082404>
21. Kiely M., Tse L.A., Koka H. et al. Age-related DNA methylation in paired normal and tumour breast tissue in Chinese breast cancer patients // Epigenetics. 2021. V. 6. P. 677–691. <https://doi.org/10.1080/15592294.2020.1819661>
22. Gaykalova D.A., Vatapalli R., Wei Y. et al. Outlier analysis defines zinc finger gene family DNA methylation in tumors and saliva of head and neck cancer patients // PLoS One. 2015. V. 11. P. e0142148. <https://doi.org/10.1371/journal.pone.0142148>
23. Chmelarova M., Kos Š., Dvorakova E. et al. Importance of promoter methylation of GATA4 and TP53 genes in endometrioid carcinoma of endometrium // Clinical chemistry and laboratory medicine: CCLM / FESCC. 2014. 52. <https://doi.org/10.1515/cclm-2013-0162>
24. Hulbert A., Jusue-Torres I., Stark A. et al. Early detection of lung cancer using dna promoter hypermethylation in plasma and sputum // Clin. cancer Res.: an off. j. Am. association for Cancer Res. 2017. V. 8. P. 1998–2005. <https://doi.org/10.1158/1078-0432.ccr-16-1371>
25. Oster B., Thorsen K., Lamy P. et al. Identification and validation of highly frequent CpG island hypermethylation in colorectal adenomas and carcinomas // Intern. J. Cancer. 2011. V. 12. P. 2855–2866. <https://doi.org/10.1002/ijc.25951>

26. *Taherkhani A., Dehto S.S., Jamshidi S. et al.* Pathogenesis and prognosis of primary oral squamous cell carcinoma based on microRNAs target genes: A systems biology approach // *Genomics & Informatics*. 2022. V. 3. P. e27.
<https://doi.org/10.5808/gi.22038>
27. *Shen H., McElhinny A.S., Cao Y. et al.* The Notch co-activator, MAML1, functions as a novel coactivator for MEF2C-mediated transcription and is required for normal myogenesis // *Genes & Development*. 2006. V. 6. P. 675–688.
<https://doi.org/10.1101/gad.1383706>
28. *Metzger E., Imhof A., Patel D. et al.* Phosphorylation of histone H3T6 by PKC β (I) controls demethylation at histone H3K4 // *Nature*. 2010. V. 7289. P. 792–796.
<https://doi.org/10.1038/nature08839>
29. *Shimada M., Niida H., Zineldeen D.H. et al.* Chk1 is a histone H3 threonine 11 kinase that regulates DNA damage-induced transcriptional repression // *Cell*. 2008. V. 2. P. 221–232.
<https://doi.org/10.1016/j.cell.2007.12.013>
30. *Zhao D., Wang B., Chen H.* RAB11A mediates the proliferation and motility of esophageal cancer cells via WNT signaling pathway // *Acta Biochimica Polonica*. 2020. V. 4. P. 531–538.
https://doi.org/10.18388/abp.2020_5392
31. *Fan Q.C., Tian H., Wang Y. et al.* Integrin- α 5 promoted the progression of oral squamous cell carcinoma and modulated PI3K/AKT signaling pathway // *Arch. Oral Biology*. 2019. P. 85–91.
<https://doi.org/10.1016/j.archoralbio.2019.03.007>
32. *Noorolyai S., Shajari N., Baghban E. et al.* The relation between PI3K/AKT signalling pathway and cancer // *Gene*. 2019. P. 120–128.
<https://doi.org/10.1016/j.gene.2019.02.076>
33. *Peng Y., Zeng X., Lian M. et al.* FXYD3 promotes the proliferation, migration, and invasion of pancreatic cancer cells by regulating the cGMP-PKG signaling pathway // *Mol. & Cell. Toxicology*. 2022. V. 3. P. 371–381.
<https://doi.org/10.1007/s13273-021-00207-9>
34. *Kim E.K., Choi E.-J.* Pathological roles of MAPK signaling pathways in human diseases // *Biochimica et Biophysica Acta (BBA) – Molecular Basis of Disease*. 2010. V. 4. P. 396–405.
<https://doi.org/10.1016/j.bbadis.2009.12.009>
35. *Peng Q., Deng Z., Pan H. et al.* Mitogen-activated protein kinase signaling pathway in oral cancer (Review) // *Oncol Lett*. 2018. V. 2. P. 1379–1388.
<https://doi.org/10.3892/ol.2017.7491>
36. *Looi C.K., Hii L.W., Ngai S.C. et al.* The Role of Ras-associated protein 1 (Rap1) in cancer: Bad actor or good player? // *Biomedicine*. 2020. V. 9.
<https://doi.org/10.3390/biomedicine8090334>
37. *Qian D., Liu H., Zhao L. et al.* Novel genetic variants in genes of the Fc gamma receptor-mediated phagocytosis pathway predict non-small cell lung cancer survival // *Translational Lung Cancer Res*. 2020. V. 3. P. 575–586.
<https://doi.org/10.21037/tlcr-19-318>
38. *Harvey S.* Extrapituitary growth hormone // *Endocrine*. 2010. V. 3. P. 335–359.
<https://doi.org/10.1007/s12020-010-9403-8>
39. *Hu X., Li J., Fu M. et al.* The JAK/STAT signaling pathway: From bench to clinic // *Signal Transduction and Targeted Therapy*. 2021. V. 1. P. 402.
<https://doi.org/10.1038/s41392-021-00791-1>
40. *Rosen M.D., Privalsky M.L.* Thyroid hormone receptor mutations in cancer and resistance to thyroid hormone: Perspective and prognosis // *J. Thyroid Res*. 2011. P. 361304.
<https://doi.org/10.4061/2011/361304>
41. *Glunde K., Bhujwalla Z.M., Ronen S.M.* Choline metabolism in malignant transformation // *Nat. Rev. Cancer*. 2011. V. 12. P. 835–848.
<https://doi.org/10.1038/nrc3162>
42. *Kraus J.M., Verlinde C.L.M.J., Karimi M. et al.* Rational modification of a candidate cancer drug for use against chagas disease // *J. Med. Chemistry*. 2009. V. 6. P. 1639–1647.
<https://doi.org/10.1021/jm801313t>
43. *Lehoux M., D'Abramo C.M., Archambault J.* Molecular mechanisms of human papillomavirus-induced carcinogenesis // *Publ. Health Genomics*. 2009. V. 5–6. P. 268–280.
<https://doi.org/10.1159/000214918>
44. *Alsaab H.O., Sau S., Alzhrani R. et al.* PD-1 and PD-L1 checkpoint signaling inhibition for cancer immunotherapy: mechanism, combinations, and clinical outcome // *Frontiers in Pharmacology*. 2017. V. 8.
<https://doi.org/10.3389/fphar.2017.00561>
45. *Feng M., Jiang W., Kim B.Y.S. et al.* Phagocytosis checkpoints as new targets for cancer immunotherapy // *Nat. Rev. Cancer*. 2019. V. 10. P. 568–586.
<https://doi.org/10.1038/s41568-019-0183-z>
46. *Healy C.M., Moran G.P.* The microbiome and oral cancer: More questions than answers // *Oral Oncology*. 2019. V. 89. P. 30–33.
<https://doi.org/10.1016/j.oraloncology.2018.12.003>
47. *Yang C.-Y., Yeh Y.-M., Yu H.-Y. et al.* Oral microbiota community dynamics associated with oral squamous cell carcinoma staging // *Frontiers in Microbiology*. 2018. V. 9.
<https://doi.org/10.3389/fmicb.2018.00862>
48. *Avila M., Ojcius D.M., Yilmaz O.* The oral microbiota: Living with a permanent guest // *DNA and Cell Biology*. 2009. V. 8. P. 405–411.
<https://doi.org/10.1089/dna.2009.0874>
49. *Guerrero-Preston R., Godoy-Vitorino F., Jedlicka A. et al.* 16S rRNA amplicon sequencing identifies microbiota associated with oral cancer, human papilloma virus infection and surgical treatment // *Oncotarget*. 2016. V. 32. P. 51320–51334.
<https://doi.org/10.18632/oncotarget.9710>
50. *Sagar S., Esau L., Holtermann K. et al.* Induction of apoptosis in cancer cell lines by the Red Sea brine pool bacterial extracts // *BMC Complementary and Alternative Medicine*. 2013. P. 344.
<https://doi.org/10.1186/1472-6882-13-344>

Molecular Landscape of Oral Cancer in Young Adults

E. A. Prostakishina^a, T. D. Dampilova^{a, b}, L. A. Kononova^b, P. S. Iamshchikov^a,
M. R. Patysheva^a, E. S. Kolegova^{a, *}, E. L. Choinzonov^a, and E. V. Denisov^a

^a*Cancer Research Institute, Tomsk National Research Medical Center, Russian Academy of Sciences, Tomsk, 634009 Russia*

^b*Federal Siberian State Medical University, Tomsk, 634050 Russia*

**e-mail: elenakolegova@oncology.tomsk.ru*

Oral cancer (OC) is the most common cancer of the head and neck. Tongue cancer (TC) is the most frequently diagnosed form of OC and is characterized by a high aggressiveness and progression. OC and TC are considered diseases of the elderly, but the incidence among young patients (under 45 years) is increasing every year. The etiological factors and pathogenetic mechanisms of early-onset OC and TC remain unclear. In the present study, based on The Cancer Genome Atlas database, we analyzed the mutational profile, methylome, transcriptome, proteome, and microbiome of OC and TC in young adults ($n = 127$) compared with older patients. Early-onset OC and TC demonstrated a decrease in the mutation burden, activation of Rap1, PI3K-Akt, MAPK, cGMP-PKG signaling pathways and signaling of Fc-gamma R-mediated phagocytosis, and a specific microbiome profile. In contrast to OC, TC was characterized by activation of such signaling pathways as JAK-STAT, immune response to infectious and parasitic diseases, and PD-L1/PD-1-mediated immunosuppression and inhibition of phagocytosis signaling. The obtained results indicate that early-onset OC demonstrates molecular features different from those in elderly patients, while TC differs from OC in molecular profile and should probably be considered a separate clinical form.

Keywords: oral cancer, young adults, genome, epigenome, transcriptome, proteome, microbiome.

ОПРЕДЕЛЕНИЕ ПОЛИМОРФИЗМА G/C В ПОЗИЦИИ **xp20:37352001** В ПРЕПАРАТАХ ДНК ИЗ КРОВИ ЧЕЛОВЕКА МЕТОДАМИ **GlaI-** И **Bst2UI-**ПЦР АНАЛИЗА

© 2023 г. А. Г. Акишев¹, Н. А. Нетесова², М. А. Абдурашитов¹, С. Х. Дегтярев^{1, 3, *}

¹Научно-производственное объединение “СибЭнзим”, Новосибирск, 630117 Россия

²ООО “Эпигенлаб”, Новосибирск, 630559 Россия

³Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, 119991 Россия

*e-mail: degt@sibenzyme.com

Поступила в редакцию 17.05.2023 г.

После доработки 09.06.2023 г.

Принята к публикации 20.06.2023 г.

Однонуклеотидный полиморфизм (SNP) заключается в замене одного нуклеотида на другой, что часто приводит к возникновению (или исчезновению) сайта узнавания определенной рестриктазы. В результате при амплификации фрагмента ДНК с праймеров, окружающих точку SNP (содержащую либо *N1*, либо *N2*-нуклеотид), и последующего гидролиза ампликона данной рестриктазой картины расщепления ДНК будут разные для трех возможных вариантов в диплоидной ДНК (генотипы *N1/N1*, *N1/N2* и *N2/N2*). Данный метод определения полиморфизма длин фрагментов рестрикции (RFLP) широко используется в практике генетических исследований. Ранее мы разработали методы *GlaI*- и *FatI*-ПЦР анализа, в которых проводится ПЦР в реальном времени вместо электрофореза, и показали его применимость для определения SNP *T/C*. В настоящей работе предложен новый способ определения однонуклеотидного полиморфизма *G/C* методом *Bst2UI*-ПЦР анализа. *GlaI*- и *Bst2UI*-ПЦР анализ использовали для определения частоты вариантов полиморфизма *G/C* в положении **xp20:37352001** (по геномной сборке GRCh38.p14) в препаратах ДНК, выделенной из клеток крови 161 человека. Исследование включало: 1) выделение лейкоцитарной ДНК из клеток крови; 2) проведение *GlaI*- и *Bst2UI*-ПЦР анализа фрагмента ДНК **xp20:37351957–37352083**; 3) определение цитозина и гуанина в позиции **xp20:37352001** в анализируемых препаратах ДНК; 4) сравнительный анализ полученных результатов. Показано, что 68 доноров (42.2%) имеют гетерозиготный набор *G/C* в положении **xp20:37352001**, 89 доноров (55.3%) гомозиготны по *G*, а четыре донора (2.5%) — по *C*. Таким образом, принимая во внимание, что клетки крови имеют диплоидный набор хромосом, замена *G* на *C* встречается в 76 из 322 проанализированных вариантов (23.6%). При этом из полученных результатов следует, что цитозин, комплементарный *G* в положении **xp20:37352001**, в большей части молекул ДНК находится в метилированной форме (5-метилцитозин) как у гомо-, так и у гетерозигот. Предложенный метод *Bst2UI*-ПЦР анализа расширяет возможности определения SNP с помощью ПЦР в реальном времени.

Ключевые слова: однонуклеотидный полиморфизм *G/C*, частоты аллелей, ДНК из крови человека, *GlaI*-ПЦР анализ, *Bst2UI*-ПЦР анализ.

DOI: 10.31857/S0016675823110012, **EDN:** NGJDEA

Ранее с помощью *GlaI*- и *FatI*-ПЦР анализа нами было проведено изучение частот аллелей для полиморфизмов *C/T* в положении **xp16:75033884** [1] и **xp1:245618129** [2] в препаратах ДНК 92 доноров. В данных позициях находится последовательность *CACGC*, которая превращается в последовательность *CATGC* с возникновением сайта узнавания рестриктазы *FatI* (*CATG*). В настоящей работе мы применили методы *GlaI*- и *Bst2UI*-ПЦР анализа для изучения однонуклеотидного полиморфизма *G/C* в третьем положении последовательности *AC(G/C)TGG* (**rs11697215** по базе данных “dbSNP”

[3]) в положении **xp20:37352001** в препаратах ДНК 161 донора. Данный SNP расположен в интроне протоонкогена *SRC*, кодирующего тирозин-специфичную протеинкиназу и его анализ не может проводиться методом *FatI*-ПЦР анализа ввиду отсутствия последовательности *CATG*. На рис. 1 представлена нуклеотидная структура участка геномной ДНК, включающая в себя анализируемый полиморфизм. Суть методов *GlaI*- и *Bst2UI*-ПЦР анализа заключается в альтернативном расщеплении ДНК ферментами *GlaI* или *Bst2UI* соответственно, и последующей ПЦР с праймеров, лежа-

CCCAGACTTCTGGCTGCTGAGTAATTTGTGGGGATCTGATTCTTATTAAGTGTGTACCCAACGTGGCTAAAA
 CTTTTCCCAGGTATAGGAGGGGTGCGGTGCTGGGGGCAGGTGACCCCTGACTCTCTGGGTTCCAGCTCTGGCT
 GATGGATTCCCCAGTCGGACGTGGAGTACTGGGATCTTGGGTGCTCACGGATGCTTATGCCATAAAGGGTAA
 TGGAAGTAGCAAAGCCTTTGAAGCCAGTTGGTTCTGGGTTCAAATCGTGGCTTGGCCACTTGTTTGCTGTGT
 ATGATTGTACAGCTTAGAACCTCAGTTTCCCCATCTGT

Рис. 1. Нуклеотидная последовательность анализируемого участка геномной ДНК хр20:37351837–37352162. Волнистой линией выделены сайты узнавания рестриктазы BstDEI (CTNAG), окаймляющие амплифицируемый участок. Зоны связывания прямого и обратного праймеров, а также флуоресцентно меченного TaqMan-зонда выделены серым фоном. Изучаемый сайт ACGTGG, содержащий метилируемый сайт ACGT [5], обведен рамкой. Подчеркнут гуаниновый остаток в сайте ACGTGG, который может замещаться цитозином.

щих до и после искомой последовательности. При расщеплении ДНК ферментами происходит следующее: рестриктаза Bst2UI (сайт узнавания CC(A/T)GG) расщепляет ДНК по последовательности ACCTGG и не расщепляет другие сайты, а ДНК-эндонуклеаза GlaI (сайт узнавания R(5mC)GY) [4] расщепляет только сайт A(5mC)GTGG. При гидролизе ДНК ферментом Bst2UI и последующей ПЦР происходит амплификация фрагмента при наличии последовательности A(5mC)GTGG или ACGTGG и не наблюдается продукта ПЦР в случае сайта ACCTGG. При гидролизе ДНК ферментом GlaI последующая амплификация фрагмента идет при наличии последовательностей ACCTGG и ACGTGG и не наблюдается продукта ПЦР в случае сайта A(5mC)GTGG. Для проведения контрольной ПЦР [5] в этом случае был подобран фермент рестрикции BstDEI, сайты узнавания которого (CTNAG) расположены вне участка амплификации (рис. 1).

При анализе диплоидной ДНК в случае гомозиготы по *G* фермент Bst2UI не расщепляет искомый фрагмент, и контрольная ПЦР после расщепления BstDEI и Bst2UI-ПЦР дают практически одинаковые результаты. В случае гомозиготы по *C* Bst2UI расщепляет ДНК близко к 100% и разница в результатах Bst2UI-ПЦР анализа и контрольной ПЦР составляет более восьми циклов (гидролиз сайта CCTGG более 99.5%). При этом в случае гетерозиготы расщепление ДНК ферментом Bst2UI происходит на 50% и разница результатов Bst2UI-ПЦР анализа с результатами контрольной ПЦР должна быть близка к одному циклу. Аналогичный результат будет наблюдаться и в случае GlaI-ПЦР-анализа, если значительная часть цитозин в положении, комплементарном гуанину в третьем положении последовательности ACGCTGG, метилирована. Тогда в случае гомозиготы по *C* GlaI не будет расщеплять ДНК, в случае гетерозиготы будет наблюдаться около

50% гидролиза и в случае гомозиготы по *G* глубина гидролиза будет зависеть от полноты метилирования цитозина. Результаты обоих методов должны соответствовать друг другу: в случае гетерозиготы оба метода должны давать разницу с контрольным ПЦР в один цикл, а в случае гомозигот один метод должен давать разницу с контролем около 0, при этом второй метод должен давать разницу в значительное число циклов.

Таким образом, мы определяем частоты вариантов полиморфизма *G/C* в положении хр20:37352001 (по геномной сборке GRCh38.p14) в препаратах ДНК, выделенной из клеток крови 161 человека, и контрольных образцах ДНК, используя два метода. Поскольку методы GlaI- и Bst2UI-ПЦР анализа определяют наличие разных нуклеотидов и являются независимыми, их совместное применение подтверждает достоверность результатов, полученных каждым из них.

МАТЕРИАЛЫ И МЕТОДЫ

Для проведения Bst2UI- и GlaI-ПЦР анализа использовались образцы ДНК 161 условно здорового донора, выделенной из лейкоцитов периферической крови. Методика выделения и использованные реактивы описаны ранее в работе [1]. В качестве контроля метилирования мы использовали TaqI-гидролизат ДНК HeLa, метилированный SssI метилазой (НПО «СибЭнзим»). Возраст и пол участников исследования, а также концентрации полученных препаратов ДНК и данные их спектрофотометрической чистоты приведены в табл. 1.

Ниже указаны структуры праймеров и флуоресцентно меченного зонда, использованных для проведения ПЦР-анализа участка хр20:37351957–37352083:

SRCd 5'-ACTCTCTGGGTTCCAGCTCTGGCT-3' (24)

SRCr 5'-CCAACTGGCTTCAAAGGCTTTGC-3' (23)

SRCz 5'-FAM-AGCATCCGTGAGCACCCAAGATCCC-BHQ1-3' (25).

Таблица 1. Значения Cq в случае гидролиза ДНК по сайтам CTNAG (BstDEI), CCWGG (Bst2UI-ПЦР анализ) и A(5mC)GT (GlaI-ПЦР анализ) и значения разницы величины Cq для двух последних реакций относительно первой ($\Delta Cq_C = Cq_{CCWGG} - Cq_{CTNAG}$ и $\Delta Cq_G = Cq_{A(5mC)GT} - Cq_{CTNAG}$) соответственно

1	2	3	4	5	6	7	8	9	10
Донор	Пол	Возраст, лет	Концентрация ДНК, нг/мкл	A260/280	Cq_BstDEI	Cq_Bst2UI-ПЦР анализ	Cq_GlaI-ПЦР анализ	ΔCq_C	ΔCq_G
1172	ж	14	376	1.6	20.66 ± 0.05	20.76 ± 0.13	25.57 ± 0.09	0.10	4.91
1164	ж	17	252	1.7	20.19 ± 0.09	20.21 ± 0.04	24.05 ± 0.21	0.02	3.86
1147	м	18	180	2.0	21.31 ± 0.09	22.50 ± 0.04	22.25 ± 0.12	1.19	0.94
1133	м	18	55	1.7	20.10 ± 0.03	20.11 ± 0.07	23.83 ± 0.32	0.01	3.73
1181	ж	18	90	1.8	20.16 ± 0.03	29.68 ± 0.65	20.32 ± 0.04	9.52	0.16
1303	м	18	129	1.8	20.05 ± 0.00	20.04 ± 0.06	24.07 ± 0.09	−0.01	4.02
1187	м	20	304	1.8	21.29 ± 0.02	22.12 ± 0.04	22.11 ± 0.11	0.83	0.82
1502	ж	20	461	1.7	20.49 ± 0.07	20.38 ± 0.09	25.08 ± 0.03	−0.11	4.59
1487	ж	22	406	1.7	20.62 ± 0.06	21.78 ± 0.06	21.70 ± 0.04	1.16	1.08
1304	ж	23	129	1.7	21.08 ± 0.14	29.98 ± 0.38	21.11 ± 0.10	8.90	0.03
1218	м	24	317	1.7	21.88 ± 0.09	21.85 ± 0.13	26.13 ± 0.03	−0.03	4.25
1379	ж	24	241	1.8	20.00 ± 0.10	20.00 ± 0.03	23.36 ± 0.09	0.00	3.36
1243	м	24	280	1.7	21.50 ± 0.04	22.45 ± 0.07	22.62 ± 0.04	0.95	1.12
1332	м	25	507	1.9	21.06 ± 0.05	22.01 ± 0.12	21.98 ± 0.02	0.95	0.92
1263	м	26	252	1.9	20.56 ± 0.07	21.61 ± 0.15	21.41 ± 0.09	1.05	0.85
1173	ж	27	232	1.7	20.37 ± 0.06	21.51 ± 0.12	21.55 ± 0.12	1.14	1.18
1200	м	27	271	1.6	20.67 ± 0.06	21.50 ± 0.06	21.54 ± 0.03	0.83	0.87
1203	м	27	309	1.6	20.95 ± 0.13	22.02 ± 0.09	22.11 ± 0.07	1.07	1.16
1209	м	27	204	1.7	20.27 ± 0.06	21.15 ± 0.05	21.24 ± 0.03	0.88	0.97
1214	ж	27	265	1.6	20.49 ± 0.09	21.49 ± 0.09	21.40 ± 0.04	1.00	0.91
1329	м	27	488	1.8	20.25 ± 0.09	20.27 ± 0.12	24.10 ± 0.09	0.02	3.85
1224	м	28	213	1.9	21.22 ± 0.06	22.11 ± 0.05	22.06 ± 0.08	0.89	0.84
1225	м	28	114	1.9	21.69 ± 0.06	21.73 ± 0.07	24.38 ± 0.14	0.04	2.69
1226	ж	30	297	1.9	21.19 ± 0.04	21.30 ± 0.09	24.41 ± 0.03	0.11	3.22
1240	ж	30	252	1.8	21.38 ± 0.02	21.26 ± 0.04	24.98 ± 0.14	−0.12	3.60
1221	ж	31	263	1.7	20.22 ± 0.02	21.37 ± 0.04	21.33 ± 0.09	1.15	1.11
1134	м	32	53	1.7	19.66 ± 0.04	19.79 ± 0.07	24.21 ± 0.09	0.13	4.55
18л	ж	33	140	1.8	21.10 ± 0.01	21.06 ± 0.02	25.71 ± 0.06	−0.04	4.61
1161	м	33	197	1.7	20.34 ± 0.08	20.52 ± 0.03	24.50 ± 0.12	0.18	4.16
1192	ж	33	110	1.7	20.16 ± 0.12	20.23 ± 0.10	23.68 ± 0.14	0.07	3.52
1184	ж	34	305	1.7	21.50 ± 0.10	21.55 ± 0.08	25.00 ± 0.14	0.05	3.50
1300	м	34	127	1.9	20.79 ± 0.08	20.93 ± 0.07	24.52 ± 0.13	0.14	3.73
17л	ж	35	72	1.8	20.07 ± 0.03	21.18 ± 0.06	21.09 ± 0.04	1.11	1.02
1208	м	35	207	1.6	21.27 ± 0.06	21.27 ± 0.14	25.76 ± 0.19	0.00	4.49
1227	м	35	101	1.8	21.26 ± 0.04	22.42 ± 0.04	22.17 ± 0.09	1.16	0.91
272	ж	36	49	1.9	22.43 ± 0.12	23.48 ± 0.09	23.26 ± 0.09	1.05	0.83
285	м	36	20	2.0	22.41 ± 0.03	23.54 ± 0.05	23.42 ± 0.11	1.13	1.01
1086	м	36	193	1.7	21.37 ± 0.14	22.36 ± 0.15	22.26 ± 0.07	0.99	0.89

Таблица 1. Продолжение

1	2	3	4	5	6	7	8	9	10
Донор	Пол	Возраст, лет	Концентрация ДНК, нг/мкл	A260/280	Cq_ BstDEI	Cq_Bst2UI-ПЦР анализ	Cq_GlaI-ПЦР анализ	ΔCq_C	ΔCq_G
1162	м	36	181	1.8	21.12 ± 0.06	22.17 ± 0.05	22.20 ± 0.11	1.05	1.08
1217	м	36	407	1.6	21.20 ± 0.09	22.11 ± 0.09	22.00 ± 0.09	0.91	0.80
1247	ж	36	286	1.7	21.10 ± 0.06	21.04 ± 0.13	25.19 ± 0.08	−0.06	4.09
273	ж	37	25	1.9	22.27 ± 0.09	22.22 ± 0.05	25.97 ± 0.15	−0.05	3.70
1186	ж	37	175	1.8	20.60 ± 0.00	20.57 ± 0.04	24.21 ± 0.07	−0.03	3.61
1235	ж	37	168	1.7	21.10 ± 0.14	22.11 ± 0.05	21.93 ± 0.05	1.01	0.83
1345	м	37	420	1.8	20.50 ± 0.10	21.32 ± 0.03	21.35 ± 0.09	0.82	0.85
1094	ж	38	172	1.7	20.89 ± 0.06	22.08 ± 0.06	22.04 ± 0.09	1.19	1.15
1236	м	39	196	1.8	20.02 ± 0.08	21.15 ± 0.01	20.98 ± 0.11	1.13	0.96
1376	ж	40	225	1.9	20.26 ± 0.07	20.22 ± 0.04	24.64 ± 0.06	−0.04	4.38
243	ж	41	67	1.9	21.88 ± 0.13	22.80 ± 0.08	22.82 ± 0.09	0.92	0.94
1219	ж	41	315	1.8	21.18 ± 0.09	21.25 ± 0.03	25.99 ± 0.07	0.07	4.81
1335	ж	41	782	1.9	20.34 ± 0.10	20.39 ± 0.11	24.88 ± 0.15	0.05	4.54
1381	ж	41	500	1.7	20.83 ± 0.04	20.71 ± 0.05	24.64 ± 0.09	−0.12	3.81
657	м	42	145	1.8	20.42 ± 0.04	21.21 ± 0.05	21.18 ± 0.11	0.79	0.76
1207	м	42	220	1.6	20.81 ± 0.04	20.67 ± 0.09	24.61 ± 0.14	−0.14	3.80
1331	ж	42	230	1.8	20.86 ± 0.14	20.66 ± 0.03	23.77 ± 0.08	−0.20	2.91
1323	м	42	300	1.8	20.60 ± 0.04	20.54 ± 0.11	24.23 ± 0.23	−0.06	3.63
1411	м	42	569	1.8	20.74 ± 0.07	20.73 ± 0.02	24.49 ± 0.18	−0.01	3.75
23л	м	44	25	1.7	21.25 ± 0.11	22.28 ± 0.09	22.18 ± 0.12	1.03	0.93
1374	ж	44	101	1.8	20.99 ± 0.11	21.06 ± 0.04	25.22 ± 0.08	0.07	4.23
661	ж	45	67	1.8	19.54 ± 0.05	20.74 ± 0.14	20.56 ± 0.05	1.20	1.02
1145	ж	45	323	1.6	20.66 ± 0.04	20.61 ± 0.05	25.30 ± 0.12	−0.05	4.64
1171	ж	45	67	1.6	21.23 ± 0.09	21.10 ± 0.04	24.53 ± 0.09	−0.13	3.30
1371	ж	46	233	1.8	20.27 ± 0.07	20.15 ± 0.10	23.78 ± 0.09	−0.12	3.51
1491	ж	46	498	1.7	20.39 ± 0.02	20.29 ± 0.11	23.99 ± 0.33	−0.10	3.60
123л	м	48	34	1.8	22.31 ± 0.10	22.46 ± 0.10	25.61 ± 0.12	0.15	3.30
1155	м	48	110	1.6	19.86 ± 0.12	19.92 ± 0.03	23.66 ± 0.13	0.06	3.80
1175	ж	49	202	1.9	20.35 ± 0.03	21.47 ± 0.05	21.40 ± 0.15	1.12	1.05
1313	м	49	157	1.8	19.52 ± 0.07	19.53 ± 0.06	22.77 ± 0.06	0.01	3.25
1317	ж	49	316	1.9	20.59 ± 0.08	21.51 ± 0.05	21.40 ± 0.04	0.92	0.81
1338	м	49	696	1.8	20.28 ± 0.04	21.21 ± 0.05	21.15 ± 0.05	0.93	0.87
1339	ж	49	506	1.9	20.20 ± 0.09	21.18 ± 0.11	21.14 ± 0.01	0.98	0.94
1099	м	50	144	1.7	21.91 ± 0.09	21.73 ± 0.10	25.69 ± 0.10	−0.18	3.78
1111	ж	50	141	1.7	20.63 ± 0.04	21.54 ± 0.07	21.64 ± 0.12	0.91	1.01
1151	ж	50	225	1.7	19.95 ± 0.09	21.06 ± 0.13	21.07 ± 0.04	1.11	1.12
1248	ж	50	154	1.7	19.97 ± 0.07	20.95 ± 0.10	20.88 ± 0.01	0.98	0.91
1321	м	50	112	1.8	21.22 ± 0.07	31.97 ± 0.42	21.12 ± 0.06	10.75	−0.10
1093	м	51	201	1.7	19.98 ± 0.08	21.12 ± 0.07	21.03 ± 0.10	1.14	1.05
1129	ж	51	72	1.7	19.14 ± 0.05	19.04 ± 0.02	22.95 ± 0.11	−0.10	3.81

Таблица 1. Продолжение

1	2	3	4	5	6	7	8	9	10
Донор	Пол	Возраст, лет	Концентрация ДНК, нг/мкл	A260/280	Cq_ BstDEI	Cq_Bst2UI-ПЦР анализ	Cq_GlaI-ПЦР анализ	ΔCq_C	ΔCq_G
1165	м	51	356	1.6	20.39 ± 0.01	21.62 ± 0.05	21.46 ± 0.06	1.23	1.07
1337	ж	51	563	1.9	19.48 ± 0.14	19.48 ± 0.08	24.21 ± 0.18	0.00	4.73
1367	ж	51	188	1.9	20.29 ± 0.10	21.29 ± 0.03	21.09 ± 0.04	1.00	0.80
1230	ж	52	723	1.9	22.19 ± 0.14	22.30 ± 0.05	25.61 ± 0.18	0.11	3.42
1350	м	52	231	1.8	20.70 ± 0.10	21.96 ± 0.02	21.73 ± 0.04	1.26	1.03
6л	ж	53	288	1.8	21.09 ± 0.10	21.20 ± 0.08	26.13 ± 0.04	0.11	5.04
202	м	53	248	1.8	20.14 ± 0.02	20.10 ± 0.06	24.09 ± 0.21	−0.04	3.95
1159	м	53	358	1.7	20.50 ± 0.03	21.50 ± 0.08	21.40 ± 0.07	1.00	0.90
1239	ж	53	230	1.7	20.43 ± 0.14	20.41 ± 0.02	24.50 ± 0.16	−0.02	4.07
1358	м	53	784	1.8	21.73 ± 0.11	22.63 ± 0.12	22.54 ± 0.06	0.90	0.81
1361	ж	53	762	1.9	20.03 ± 0.10	20.24 ± 0.07	24.41 ± 0.08	0.21	4.38
53л	ж	54	88	1.9	21.17 ± 0.07	21.15 ± 0.08	26.02 ± 0.21	−0.02	4.85
1113	ж	55	109	1.9	20.49 ± 0.08	20.54 ± 0.02	24.78 ± 0.06	0.05	4.29
1246	ж	55	191	1.7	20.86 ± 0.06	21.01 ± 0.09	24.00 ± 0.14	0.15	3.14
1342	м	55	119	1.9	20.74 ± 0.08	21.82 ± 0.13	21.81 ± 0.10	1.08	1.07
54л	м	56	113	1.9	20.70 ± 0.01	20.79 ± 0.09	24.47 ± 0.03	0.09	3.77
1280	м	56	436	1.8	21.59 ± 0.03	21.43 ± 0.05	25.29 ± 0.05	−0.16	3.70
1127	ж	57	123	1.7	20.38 ± 0.13	20.21 ± 0.02	24.38 ± 0.09	−0.17	4.00
1135	ж	57	127	1.6	19.22 ± 0.02	20.15 ± 0.06	20.10 ± 0.06	0.93	0.88
1163	ж	57	157	1.7	20.48 ± 0.14	20.51 ± 0.12	24.42 ± 0.12	0.03	3.94
1183	ж	57	662	1.7	20.51 ± 0.10	20.56 ± 0.07	24.60 ± 0.05	0.05	4.09
1360	ж	57	570	1.9	20.34 ± 0.05	20.48 ± 0.14	24.48 ± 0.24	0.14	4.14
19л	ж	58	54	1.8	21.44 ± 0.13	21.35 ± 0.05	25.32 ± 0.10	−0.09	3.88
1130	ж	58	59	1.7	20.28 ± 0.10	20.24 ± 0.04	23.78 ± 0.08	−0.04	3.50
1351	ж	58	260	1.8	20.10 ± 0.02	21.09 ± 0.15	21.10 ± 0.06	0.99	1.00
297	ж	59	48	1.9	23.10 ± 0.12	24.24 ± 0.04	23.91 ± 0.10	1.14	0.81
1137	ж	59	299	1.6	21.42 ± 0.11	21.48 ± 0.05	25.00 ± 0.15	0.06	3.58
1182	ж	59	136	1.9	21.36 ± 0.01	21.47 ± 0.06	25.39 ± 0.15	0.11	4.03
1259	ж	59	111	1.8	20.72 ± 0.08	21.48 ± 0.05	21.66 ± 0.01	0.76	0.94
1299	ж	59	252	1.8	20.35 ± 0.14	21.30 ± 0.06	21.12 ± 0.03	0.95	0.77
1326	ж	59	380	1.8	19.71 ± 0.14	19.65 ± 0.14	23.15 ± 0.04	−0.06	3.44
1344	ж	59	381	1.7	20.26 ± 0.02	21.43 ± 0.05	21.21 ± 0.04	1.17	0.95
1340	ж	60	562	1.7	20.34 ± 0.09	20.27 ± 0.09	25.20 ± 0.18	−0.07	4.86
1348	м	60	761	1.7	20.35 ± 0.07	20.43 ± 0.08	24.51 ± 0.23	0.08	4.16
1353	ж	60	250	1.9	20.15 ± 0.02	20.17 ± 0.13	24.93 ± 0.20	0.02	4.78
298	ж	61	29	1.9	22.47 ± 0.02	23.39 ± 0.00	23.28 ± 0.01	0.92	0.81
1144	ж	61	147	1.8	20.27 ± 0.11	20.51 ± 0.08	24.48 ± 0.16	0.24	4.21
1234	ж	61	178	1.7	20.18 ± 0.05	21.29 ± 0.08	21.12 ± 0.09	1.11	0.94
1364	м	61	178	1.9	20.11 ± 0.10	21.22 ± 0.08	21.05 ± 0.16	1.11	0.94
1202	ж	62	111	1.8	20.48 ± 0.09	20.37 ± 0.14	24.22 ± 0.06	−0.11	3.74

Таблица 1. Продолжение

1	2	3	4	5	6	7	8	9	10
Донор	Пол	Возраст, лет	Концентрация ДНК, нг/мкл	A260/280	Cq_ BstDEI	Cq_Bst2UI-ПЦР анализ	Cq_GlaI-ПЦР анализ	ΔCq_C	ΔCq_G
1347	м	62	590	1.9	20.34 ± 0.05	21.21 ± 0.12	21.14 ± 0.06	0.87	0.80
1349	м	62	355	1.8	21.57 ± 0.15	22.48 ± 0.03	22.44 ± 0.04	0.91	0.87
1250	м	63	271	1.8	21.40 ± 0.04	21.25 ± 0.05	24.86 ± 0.05	−0.15	3.46
1294	м	63	173	1.6	20.63 ± 0.11	20.57 ± 0.05	24.03 ± 0.05	−0.06	3.40
1369	ж	63	301	1.9	20.35 ± 0.12	21.21 ± 0.09	21.25 ± 0.09	0.86	0.90
206	ж	64	90	1.9	21.36 ± 0.06	22.30 ± 0.05	22.11 ± 0.04	0.94	0.75
264	м	64	35	2.0	24.27 ± 0.27	25.38 ± 0.14	25.26 ± 0.10	1.11	0.99
1211	ж	64	323	1.4	19.50 ± 0.04	19.52 ± 0.03	22.67 ± 0.09	0.02	3.17
1255	ж	64	357	1.9	21.00 ± 0.05	21.04 ± 0.08	23.95 ± 0.13	0.04	2.95
1277	ж	64	330	1.7	20.37 ± 0.03	21.32 ± 0.04	21.39 ± 0.05	0.95	1.02
16л	ж	65	95	1.8	20.97 ± 0.09	21.94 ± 0.10	21.89 ± 0.08	0.97	0.92
268	м	65	88	1.9	21.25 ± 0.04	22.28 ± 0.06	22.27 ± 0.05	1.03	1.02
299	м	65	56	1.9	21.41 ± 0.05	21.38 ± 0.14	25.04 ± 0.40	−0.03	3.63
1152	ж	65	112	1.7	19.78 ± 0.06	19.58 ± 0.06	23.05 ± 0.12	−0.20	3.27
1205	м	65	420	1.7	21.11 ± 0.09	21.19 ± 0.03	24.55 ± 0.08	0.08	3.44
1238	ж	66	346	1.7	20.28 ± 0.08	21.07 ± 0.09	21.07 ± 0.04	0.79	0.79
266	м	67	18	1.8	22.13 ± 0.14	22.08 ± 0.04	26.07 ± 0.14	−0.05	3.94
296ф	ж	67	175	1.8	21.43 ± 0.04	22.54 ± 0.03	22.33 ± 0.05	1.11	0.90
1143	ж	67	84	1.6	20.51 ± 0.09	21.58 ± 0.10	21.40 ± 0.08	1.07	0.89
1372	ж	67	221	1.8	20.23 ± 0.11	20.03 ± 0.07	23.81 ± 0.27	−0.20	3.58
1428	м	67	433	1.8	20.65 ± 0.04	20.61 ± 0.11	23.86 ± 0.12	−0.04	3.21
295	м	68	219	1.8	20.66 ± 0.01	20.66 ± 0.07	23.93 ± 0.14	0.00	3.27
1365	ж	68	204	1.8	20.30 ± 0.04	20.29 ± 0.05	24.03 ± 0.07	−0.01	3.73
173	м	68	352	1.7	19.98 ± 0.07	19.89 ± 0.02	23.39 ± 0.09	−0.09	3.41
1437	м	68	213	1.8	20.17 ± 0.05	20.19 ± 0.04	24.13 ± 0.18	0.02	3.96
237	м	69	31	1.8	21.60 ± 0.11	21.75 ± 0.06	25.14 ± 0.25	0.15	3.54
240	м	69	18	1.9	21.27 ± 0.09	21.41 ± 0.11	25.08 ± 0.13	0.14	3.81
293	м	69	359	1.8	20.02 ± 0.01	20.08 ± 0.05	23.06 ± 0.02	0.06	3.04
1146	ж	70	149	1.8	19.56 ± 0.06	20.76 ± 0.07	20.41 ± 0.08	1.20	0.85
1413	м	70	522	1.9	19.77 ± 0.02	19.82 ± 0.06	23.65 ± 0.17	0.05	3.88
40	ж	72	179	1.8	20.22 ± 0.03	20.20 ± 0.04	24.10 ± 0.12	−0.02	3.88
205а	ж	72	461	1.8	21.00 ± 0.03	20.90 ± 0.07	24.56 ± 0.18	−0.10	3.56
643	м	72	51	2.0	23.09 ± 0.02	23.08 ± 0.09	26.71 ± 0.10	−0.01	3.62
1115	м	72	78	1.7	21.24 ± 0.05	22.28 ± 0.10	22.21 ± 0.08	1.04	0.97
1576	м	72	307	1.8	21.06 ± 0.04	22.28 ± 0.10	21.98 ± 0.11	1.22	0.92
178	м	73	182	1.7	20.92 ± 0.05	20.92 ± 0.05	24.64 ± 0.13	0.00	3.72
1669	м	74	351	1.7	21.09 ± 0.02	20.90 ± 0.08	25.14 ± 0.07	−0.19	4.05
121	ж	75	163	1.7	20.36 ± 0.07	21.37 ± 0.07	21.19 ± 0.08	1.01	0.83
1168	ж	79	239	1.7	20.88 ± 0.12	20.96 ± 0.08	24.47 ± 0.11	0.08	3.59
1174	ж	79	271	1.6	21.05 ± 0.11	29.12 ± 0.41	20.98 ± 0.05	8.07	−0.07

Таблица 1. Окончание

1	2	3	4	5	6	7	8	9	10
Донор	Пол	Возраст, лет	Концентрация ДНК, нг/мкл	A260/280	Cq_BstDEI	Cq_Bst2UI-ПЦР анализ	Cq_GlaI-ПЦР анализ	ΔCq_C	ΔCq_G
1160	ж	80	195	1.7	21.20 ± 0.04	22.31 ± 0.02	22.12 ± 0.06	1.11	0.92
55	ж	81	221	1.7	19.17 ± 0.05	19.11 ± 0.03	22.35 ± 0.01	−0.06	3.18
56	м	81	414	1.7	20.09 ± 0.03	21.31 ± 0.06	21.18 ± 0.12	1.22	1.09
HeLa					20.77 ± 0.05	20.98 ± 0.11	23.81 ± 0.17	0.21	3.04
HeLa/Taq/Sss					20.63 ± 0.12	20.82 ± 0.16	27.52 ± 0.37	0.19	6.89

ПЦР в реальном времени проводили согласно протоколу производителя [6] в объеме 20 мкл в триплетах на детектирующем амплификаторе CFX-96 (Bio-Rad, США) по следующей программе: при 95°C – 3 мин; 5 “слепых” циклов: 95°C – 10 с, 63°C – 20 с, 72°C – 10 с; 40 циклов: 95°C – 10 с, 63°C – 20 с (с детекцией флуоресцентного сигнала в канале FAM), 72°C – 10 с.

По завершении ПЦР, при помощи программного обеспечения амплификатора Bio-Rad CFX Manager v.2.1 устанавливали среднее значение Cq и значение среднеквадратичного отклонения для анализируемых образцов.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В табл. 1 показаны значения Cq , полученные при ПЦР-анализе гидролизатов ДНК 161 донора крови и препаратов ДНК HeLa и HeLa/TaqI, метилированной ферментом SssI (столбцы 6–8). В столбцах 9 и 10 указаны значения величин Cq , полученные в Bst2UI- и GlaI-ПЦР анализе, за вычетом значений Cq в столбце 6 (контрольная ПЦР на BstDEI-гидролизате), обозначенные как ΔCq_G и ΔCq_5mC соответственно.

Анализ данных, представленных в табл. 1, показывает, что как и в предыдущих работах [1, 2] полученные значения ΔCq_G и ΔCq_C соответствуют друг другу для всех образцов ДНК в выборке. При этом выявляются три группы образцов ДНК. В первой из них значения ΔCq_G и ΔCq_C варьируют от 0.75 до 1.26, что приблизительно соответствует двукратному уменьшению количества нерасщепленных копий амплифицируемого участка. Таким образом, образцы первой группы представляют собой гетерозиготы по аллелям *G* и *C* в позиции xp20:37352001.

Во второй группе значения ΔCq_C варьируют от 8.07 до 10.75, в то время как значения ΔCq_G минимальны (от −0.1 до 0.16). Очевидно, что образцы ДНК второй группы полностью расщепляются ферментом Bst2UI и не содержат сайтов узнавания GlaI. Это свидетельствует о том, что в исследуемой позиции SNP находится лишь цитозинный остаток (гомозиготы по *C*).

Для третьей группы образцов ДНК характерно повышенное значение ΔCq_G (2.69–5.04) при небольших значениях ΔCq_C (от −0.2 до 0.24). Соответственно данные образцы можно рассценивать как содержащие лишь гуаниновый остаток в позиции xp20:37352001 (гомозиготы по *G*), причем комплементарный ему цитозинный остаток находится преимущественно в метилированной форме.

Значения ΔCq_G (2.69–5.04) существенно меньше величин ΔCq_C (от 8.07 до 10.75), что, вероятно, связано с наличием неметилированного комплементарного цитозина, доля которого может достигать почти четверти. Для ДНК HeLa значение ΔCq_G составляет 3.04 и соответствует 12% неметилированной ДНК, тогда как при полностью метилированном сайте в этой ДНК (HeLa/TaqI, метилированной SssI) величина ΔCq_G составляет 6.89, что соответствует 99%-ному метилированию сайта.

На рис. 2 результаты, представленные в табл. 1, отсортированы в порядке уменьшения ΔCq_G .

Таким образом, гетерозиготный набор *G/C* имеется в 68 образцах ДНК, четыре образца представляют собой гомозиготу *C/C*, а 89 образцов являются гомозиготными по аллелю *G*. Частота аллеля *C* в исследованной выборке составляет 23.6%. Полученные данные показывают, что во всех образцах ДНК цитозинные остатки на обеих цепях

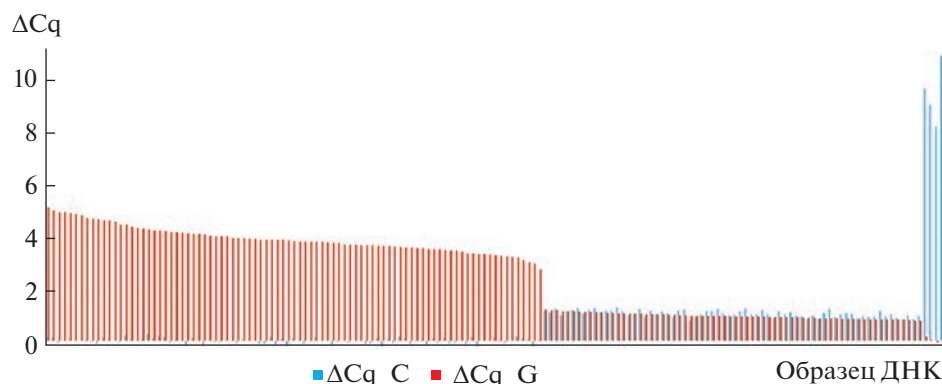


Рис. 2. Значения ΔCq_G и ΔCq_C для 161 образца донорской ДНК лейкоцитов (красный и синий цвет соответственно).

сайта ACGT в последовательности ACGTGG преимущественно метилированы. Как было показано ранее, метилированная форма CG-динуклеотидов преобладает во всех компартментах генома за исключением CpG-островков и первых экзонов, выполняющих функцию регуляции экспрессии генов [7]. Варибельный нуклеотид в положении chr20:37352001 расположен в первом интроне гена *SRC* и в случае аллеля *G* образующийся сайт ACGT метилирован по цитозиновым остаткам в обеих цепях ДНК.

В табл. 2 показано число генотипов и их доля от общего количества проанализированных препаратов ДНК доноров крови.

В табл. 3 дано сравнение частот аллелей *G* и *C* в различных популяциях по данным проектов, представленных в базе “dbSNP”.

Как видно из табл. 3, полученная нами частота аллеля *C* в 24% близка к значениям 21–22% для всей мировой когорты обследованных без учета географического проживания или этнического происхождения (“Global”) в проектах “1000Genomes” и совпадает с частотой в проекте “TopMed”. Что же касается исследованных ранее жителей Сибири, то завышенная (0.57) частота аллеля *C*, скорее всего, связана с небольшим размером выборки ($n = 14$). Теоретически определенная по формуле Харди–Вайнберга [8] величина числа гомозигот *C/C* для нашей выборки равна 9, что существенно

больше экспериментально полученного значения 4, однако это различие может быть вызвано недостаточным числом обследованных в настоящей работе.

Интересно отметить, что аллель *C* в исследованной позиции считается “предковым” по отношению к аллелю *G* (по данным сравнения геномов гоминид, дополнительная информация к [9]). Гомозиготность по *C* выявлена также для денисовского генома и неандертальского генома из Виндии (<https://www.eva.mpg.de/genetics/genome-projects/>). Превалирование варианта *G* у современных людей вне зависимости от расы и этноса свидетельствует о том, что нуклеотидная замена *C* на *G* произошла на ранних стадиях становления *Homo sapiens*, причем в дальнейшем более эволюционно успешными оказались носители аллеля *G*.

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

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

Авторы заявляют, что у них нет конфликта интересов.

Таблица 2. Распределение диплоидных наборов аллелей в позиции chr20:37352001 в общей выборке

Всего доноров	<i>C/C</i>		<i>G/C</i>		<i>G/G</i>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
161	4	2.48%	68	42.24%	89	55.28%

Таблица 3. Частоты аллелей G и C для SNP в положении chr20:37352001 по данным различных проектов

Проект	Когорта	Число обследованных	Частота G	Частота C
1000Genomes	Глобальная	5008	0.79	0.21
	Африка	1322	0.73	0.27
	Восточная Азия	1008	0.83	0.17
	Европа	1006	0.80	0.20
	Южная Азия	978	0.78	0.22
	Америка	694	0.83	0.17
1000Genomes_30x	Глобальная	6404	0.78	0.22
	Африка	1786	0.73	0.27
	Европа	1266	0.79	0.21
	Южная Азия	1202	0.78	0.22
	Восточная Азия	1170	0.83	0.17
	Америка	980	0.83	0.17
TopMed	Глобальная	264690	0.76	0.24
The Avon Longitudinal Study	Родители и дети	3854	0.73	0.27
Genetic variation in the Estonian population	Эстония	4480	0.84	0.16
Northern Sweden	Лен Вестерботтен	600	0.85	0.16
Genome of the Netherlands Release 5	Нидерланды	998	0.73	0.27
The Danish reference pan genome	Дания	40	0.72	0.28
UK 10K study – Twins	Близнецы	3708	0.73	0.27
Qatari	Глобальная	216	0.71	0.29
A Vietnamese Genetic Variation Database	Глобальная	214	0.85	0.15
Korean population from KRGDB	Корея	2930	0.83	0.17
Siberian	Глобальная	14	0.43	0.57
Настоящая работа	Новосибирская область	161	0.76	0.24

СПИСОК ЛИТЕРАТУРЫ

1. Акишев А.Г., Нетесова Н.А., Абдурашитов М.А., Дегтярев С.Х. Определение полиморфизма 5mC/T в повторе AluSx (Chr16: 75033884) в препаратах ДНК из крови человека методами GluI- и FatI-ПЦР анализа // Эпигенетическая ДНК диагностика. 2021. Т. 1. С. 1–12. <https://doi.org/10.26213/SE.2021.11.76.001>
2. Акишев А.Г., Нетесова Н.А., Абдурашитов М.А., Дегтярев С.Х. Определение полиморфизма 5mC/T в позиции Chr1: 245618129 в препаратах ДНК из крови человека методами GluI- и FatI-ПЦР анализа // Эпигенетическая ДНК диагностика. 2021. Т. 1. С. 13–23. <https://doi.org/10.26213/SE.2019.69.42836>
3. Sherry S.T., Ward M.H., Kholodov M. et al. dbSNP: the NCBI database of genetic variation // Nucl. Acids Res. 2001. V. 29. P. 308–311. <https://doi.org/10.1093/nar/29.1.308>
4. Tarasova G.V., Nayakshina T.N., Degtyarev S.K. Substrate specificity of new methyl-directed DNA endonuclease GluI // BMC Molecular Biol. 2008. V. 9: 7. <https://doi.org/10.1186/1471-2199-9-7>
5. Акишев А.Г., Нетесова Н.А., Абдурашитов М.А., Дегтярев С.Х. Выщепление рестриктазой амплифицируемого фрагмента ДНК как способ исключения ошибочных результатов ПЦР в реальном времени с TaqMan зондом // ДНК-узнающие ферменты. 2021. Т. 1. С. 1–13. <https://doi.org/10.26213/SE.2021.45.29.001>
6. CFX96 and CFX384 Real-Time PCR Detection Systems. Instruction Manual // <https://www.biorad.com/webroot/web/pdf/lsr/literature/10010424.pdf>
7. Edwards J.R., Yarychivska O., Boulard M., Bestor T.H. DNA methylation and DNA methyltransferases // Epigenet. Chromatin. 2017. V. 10: 23. <https://doi.org/10.1186/s13072-017-0130-8>
8. Stark A.E., Seneta E. A reality check on Hardy–Weinberg // Twin Res. Hum. Genet. 2013. V. 16. P. 782–789. <https://doi.org/10.1017/thg.2013.40>
9. The 1000 Genomes Project Consortium. A global reference for human genetic variation // Nature 2015. V. 526. P. 68–74. <https://doi.org/10.1038/nature15393>

Determination of *G/C* Polymorphism at chr20:37352001 Position in Human Blood DNA Preparations by *GlaI*- and *Bst2UI*-PCR Analyses Methods

A. G. Akishev^a, N. A. Netesova^b, M. A. Abdurashitov^a, and S. Kh. Degtyarev^{a, c, *}

^a*Scientific and Production Association "SibEnzyme", Novosibirsk, 630117 Russia*

^b*"EpigenLab" Ltd, Novosibirsk, 630559 Russia*

^c*Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, 119991 Russia*

**e-mail: degt@sibenzyme.com*

Single nucleotide polymorphism (SNP) is a change of one nucleotide by another. This change often leads to an emergence (or disappearance) of a site recognized by a certain restriction endonuclease. As a result amplification of DNA fragment using primers surrounding SNP point (containing either *N1* or *N2* nucleotide) followed by hydrolysis of the amplicon with this restriction enzyme will be different for three possible variants in a diploid genome (genotypes *N1/N1*, *N1/N2* and *N2/N2*). This method of restriction fragments length polymorphism (RFLP) is widely used in the genetic studies. Earlier we have developed *GlaI*- and *FatI*-PCR analyses methods which allowed to carry out real-time PCR and showed its applicability for SNP determination. In the current work a new way to determine the single nucleotide polymorphism *G/C* by the *Bst2UI*-PCR analysis is proposed. *GlaI*- and *Bst2UI*-PCR analyses have been used to determine the frequency of *G/C* polymorphism variants at the chr20:37352001 position (according to GRCH38.p14 genomic assembly) in the blood DNA samples of 161 donors. The study included: 1) the isolation of leukocyte DNA from blood cells; 2) *GlaI*- and *Bst2UI*-PCR analyses of the DNA fragment chr20:37351957-37352083, 3) determination of cytosine and guanine at the chr20:37352001 position in the analyzed DNA preparations, and 4) comparative analysis of the obtained results. It has been shown that 68 donors (42.2%) have a heterozygous set of *G/C* at the chr20:37352001 position, 89 donors (55.3%) are homozygous by *G*, and 4 donors (2.5%) are homozygous by *C*. Thus, taking into account that blood cells have a diploid set of chromosomes, *G* to *C* replacement occurs in 76 out of 322 analyzed cases (23.6%). At the same time, from the results obtained it follows that the cytosine residue complementary to *G* at the chr20:37352001 position exists in methylated form (5-methylcytosine) in most of the DNA molecules, both in homo- and heterozygotes. The proposed method of *Bst2UI*-PCR analyses extends the possibilities of SNP determination using real-time PCR.

Keywords: single nucleotide polymorphism *G/C*, allele frequency, blood DNA, *GlaI*-PCR analysis, *Bst2UI*-PCR analysis.

ХАРАКТЕРИСТИКА СПЕКТРА НУКЛЕОТИДНЫХ ЗАМЕН
МИТОХОНДРИАЛЬНОЙ ДНК В ПОПУЛЯЦИЯХ ЧЕЛОВЕКА
В УСЛОВИЯХ ВЫСОКОГОРЬЯ© 2023 г. Б. А. Малярчук¹, *¹Институт биологических проблем Севера Дальневосточного отделения
Российской академии наук, Магадан, 685000 Россия

*e-mail: malyarchuk@ibpn.ru

Поступила в редакцию 29.05.2023 г.

После доработки 14.06.2023 г.

Принята к публикации 29.06.2023 г.

С помощью филогенетического анализа нуклеотидных последовательностей целых митохондриальных геномов (мтДНК) реконструированы спектры генеративных нуклеотидных замен (по L-цепи мтДНК) в высокогорных популяциях Памира и Тибета в сравнении с региональными группами коренного населения Западной Азии, Северо-Восточной и Южной Сибири. Различий в распределении частот нуклеотидных замен в спектрах мтДНК в зависимости от высоты проживания популяций не обнаружено. Во всех спектрах мтДНК преобладают пиримидиновые транзиции, а из них – замены $T \rightarrow C$. Вторыми по частоте в большинстве региональных групп следуют замены $A \rightarrow G$, в памирской и северо-восточноазиатской группах – замены $G \rightarrow A$. Из трансверсий во всех исследованных группах населения преобладают замены $C \rightarrow A$, кроме тибетской, где чаще замены $A \rightarrow C$. Отсутствие различий в распределении мутаций мтДНК у населения высокогорных и не-горных районов свидетельствуют о том, что структура спектров нуклеотидных замен мтДНК в популяциях человека не зависит от интенсивности окислительного стресса в митохондриях.

Ключевые слова: митохондриальный геном, спектр нуклеотидных замен, генеративные мутации, популяции человека, гипоксия.

DOI: 10.31857/S0016675823110085, EDN: NLIAAK

Митохондриальный геном (мтДНК) человека кодирует основные белковые субъединицы системы окислительного фосфорилирования митохондрий и имеет, таким образом, важнейшее значение в обеспечении энергетических потребностей клеток. В митохондриях проходят интенсивные окислительные процессы, связанные с перекисным окислением липидов и белков, реакции карбонилирования и окисления ДНК активными формами кислорода (АФК). Последние включают в себя чрезвычайно реакционно-способные ионы кислорода, свободные радикалы и перекиси. Установлено, что продукция АФК существенно возрастает в условиях гипоксии [1]. Несмотря на меньшую доступность кислорода при гипоксии, митохондрии демонстрируют более высокую скорость генерирования супероксида своей электронно-транспортной системой [2–4].

Продукция АФК при гипоксии настолько высока, что для предотвращения развития связанных с гипоксией патологий требуется компенсаторная активация системы антиоксидантных ферментов [5]. Гипоксия вызывает также изменения в мор-

фологии митохондрий, которые способствуют повышению их устойчивости к метаболическому стрессу [6]. Между тем клеточный ответ на гипоксию направлен на снижение окислительного метаболизма в пользу анаэробного производства АТФ, и одновременно с этим клетки при гипоксии активизируют защитные механизмы для предотвращения окислительного стресса, повреждения митохондрий и клеток [3]. Таким образом, предполагается, что митохондрии играют важную роль в сигнализации (посредством продукции АФК) о состоянии гипоксии, что способствует формированию цепей обратной связи, необходимых для поддержания клеточного окислительно-восстановительного гомеостаза [4].

Широко распространено мнение о том, что молекулы ДНК в митохондриях подвергаются воздействию активных форм кислорода, и поэтому в зависимости от интенсивности аэробного метаболизма можно ожидать изменений мутационных спектров мтДНК [7–9]. Если предположить, что основным источником повреждений ДНК в митохондриях является окислительный

стресс, то это должно приводить к одно- и двухцепочечным разрывам ДНК, к окислению или утрате азотистых оснований в составе нуклеотидов. Среди окисленных оснований наиболее частым в мтДНК предположительно является 8-оксигуанин [10]. Поэтому в условиях окислительного стресса ожидается избыток трансверсий $C \rightarrow A$ и $G \rightarrow T$ в тех случаях, когда гуанин в $C:G$ парах окисляется до 8-оксигуанина, и трансверсий $A \rightarrow C$ и $T \rightarrow G$ в тех случаях, когда 8-окси-ГТФ используется в качестве нуклеотидного субстрата во время репликации [11, 12]. Однако исследования изменчивости мтДНК в популяциях человека показали, что в спектрах нуклеотидных замен наблюдаются, главным образом, транзиции, а доля трансверсий составляет всего несколько процентов от всех мутаций [13–17]. Из трансверсий, действительно, чаще всего наблюдаются замены $C \rightarrow A$ (по L-цепи мтДНК), но их доля составляет всего около одного процента от всех мутаций мтДНК [17]. Поэтому, по всей видимости, мутации, индуцированные 8-оксигуанином, не являются главным источником мутагенеза мтДНК.

Для более детального исследования вопроса о влиянии окислительного стресса на геном митохондрий в настоящей работе сопоставлены спектры нуклеотидных замен мтДНК в группах населения, проживающих на различной высоте – примерно до 1 км и в высокогорье, где максимально проявляется гипоксия тканей из-за недостатка кислорода. Генетические исследования популяций человека показали, что существует целый ряд ядерных генов, вовлеченных в высокогорную адаптацию и отвечающих за самые разные физиологические функции (регуляцию транспорта кислорода, гликолиз, кровообращение и др.) [5]. Однако результаты исследований генетических вариантов мтДНК в связи с высокогорной гипоксией довольно противоречивы [18–21], а спектры нуклеотидных замен митохондриальных геномов человека в условиях высокогорья ранее не изучались.

МАТЕРИАЛЫ И МЕТОДЫ

Проанализированы представленные в генетической базе данных GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) нуклеотидные последовательности целых митохондриальных геномов у населения Тибета и Памира, проживающего на высоте более 4 км над уровнем моря. Популяция Памира ($N = 202$) представлена таджиками [20], а Тибета ($N = 268$) – тибетцами [19], которые предположительно отделились от китайцев-ханьцев около 5 тыс. лет тому назад [22].

Спектры нуклеотидных замен мтДНК реконструировали относительно L-цепей предковых последовательностей, выявленных с помощью филогенетического анализа данных об изменчи-

вости митохондриальных геномов. Исследуемые нуклеотидные замены относятся к генеративным мутациям, порядок появления которых можно проследить в филогенетическом дереве мтДНК человека в направлении от предков к потомкам и реконструировать, тем самым, спектры нуклеотидных замен для анализируемых наборов митохондриальных геномов. Для филогенетического анализа использовали метод максимальной экономии, реализованный в пакете компьютерных программ mtPhyl v4.015 (<https://sites.google.com/site/mtphyl/home>). Построения филогенетических деревьев митохондриальных геномов проводили отдельно для каждой группы населения. Статистическую значимость различий между частотами нуклеотидных замен в популяциях оценивали с помощью точного теста Фишера.

В анализе распределения нуклеотидных замен в филогенетических деревьях учитывали все замены, без разделения их на приватные замены, находящиеся в концевых ветвях дерева, и замены, определяющие гаплогруппы мтДНК и находящиеся в стволах филогенетического дерева. Проведенный ранее анализ митохондриальных геномов показал отсутствие различий в распределении приватных и гаплогруппо-специфичных замен в различных региональных группах населения [17].

Для сравнительного анализа использованы реконструированные ранее спектры нуклеотидных замен целых митохондриальных геномов населения Евразии, проживающего на высоте примерно до 1 км: северо-восточной части Сибири (эскимосы, алеуты, чукчи, коряки и юкагиры; $N = 336$); южной части Сибири и прилегающих территорий Северо-Восточного Китая (буряты, баргуты и хамнигане; $N = 430$); Западной Азии (персы, кашкайцы и ливанцы; $N = 340$) [17]. Для определения высоты проживания использовали калькулятор высоты карты (<https://www.calcmaps.com/ru/map-elevation/>).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Результаты анализа спектров нуклеотидных замен митохондриальных геномов населения высокогорных районов Памира и Тибета в сравнении с другими регионами Евразии показали примерно одинаковую распространенность транзиций в различных популяциях (табл. 1). Во всех спектрах мтДНК преобладают пиримидиновые транзиции, и из них – замены $T \rightarrow C$, составляющие среди всех нуклеотидных замен примерно 33%. Вторыми по распространенности в большинстве региональных групп следуют замены $A \rightarrow G$, а в памирской и северо-восточноазиатской группах – замены $G \rightarrow A$. Тем не менее, межгрупповые различия по частоте транзиций в спектрах мтДНК статистически не значимы ($P > 0.1$, точный тест Фишера). При сопоставлении популяций в двух категориях (высо-

Таблица 1. Спектры нуклеотидных замен мтДНК (по L-цепи) в различных популяциях человека

Регион	C → T	T → C	G → A	A → G	tv	Всего нуклеотидных замен
Тибет (<i>N</i> = 268)	18.2 (196)	33.0 (356)	21.2 (229)	24.1 (260)	3.5 (38)	1079
Памир (<i>N</i> = 202)	19.9 (178)	34.5 (308)	23.0 (206)	19.1 (171)	3.5 (31)	894
Западная Азия (<i>N</i> = 340)	20.3 (388)	33.8 (647)	20.4 (391)	21.5 (412)	4.1 (78)	1916
Южная Сибирь (<i>N</i> = 430)	20.2 (252)	32.3 (404)	21.1 (264)	22.3 (278)	4.1 (51)	1249
Северо-Восточная Сибирь (<i>N</i> = 336)	19.0 (60)	33.5 (106)	24.4 (77)	19.6 (62)	3.5 (11)	316

Примечание. *N* — размер выборки; tv — трансверсии. Частоты нуклеотидных замен приводятся в процентах, в скобках указано количество нуклеотидных замен.

Таблица 2. Распределение трансверсий в мутационных спектрах мтДНК (по L-цепи) в различных популяциях человека

Трансверсии	Памир	Тибет	Южная Сибирь	Западная Азия	Северо-Восточная Сибирь	Все регионы
C → A	29.0 (9)	18.4 (7)	35.3 (18)	32.0 (25)	27.2 (3)	29.6 (62)
A → C	12.9 (4)	29.0 (11)	7.8 (4)	12.8 (10)	18.2 (2)	14.8 (31)
A → T	19.3 (6)	18.4 (7)	15.7 (8)	16.7 (13)	18.2 (2)	17.2 (36)
T → A	9.7 (3)	13.2 (5)	13.7 (7)	9.0 (7)	9.1 (1)	11.0 (23)
G → C	12.9 (4)	13.2 (5)	11.8 (6)	7.7 (6)	18.2 (2)	11.0 (23)
C → G	6.5 (2)	2.6 (1)	3.9 (2)	3.8 (3)	0	3.8 (8)
T → G	6.5 (2)	2.6 (1)	5.9 (3)	7.7 (6)	9.1 (1)	6.2 (13)
G → T	3.2 (1)	2.6 (1)	5.9 (3)	10.3 (8)	0	6.2 (13)

Примечание. Обозначения как в табл. 1.

когорные и не-горные) статистически значимые различия между ними также отсутствуют.

Соотношение транзиций к трансверсиям в памирской и тибетской группах составляет 27.4 и 27.8 соответственно. В других региональных группах величина этого показателя варьирует от 23.5 в южносибирском спектре мтДНК до 27.7 в северо-восточноазиатском. Из трансверсий во всех исследованных группах населения Евразии преобладают замены C → A, кроме тибетской, где чаще замены A → C (табл. 2). Трансверсии C → A составляют около 30% из всех трансверсий (или 1.1% из всех нуклеотидных замен). За ними следуют замены A → T и A → C (17.2 и 14.8% соответственно для всех трансверсий). Распределения трансверсий в различных региональных группах не различаются. Таким образом, полученные результаты свидетельствуют о том, что спектры нуклеотидных замен мтДНК (реконструированные по L-цепи) у населения разных географиче-

ских зон Евразии, включая высокогорные районы, практически не различаются.

Одним из основных последствий влияния окислительно-восстановительных реакций, активно протекающих в митохондриях, на мутационный спектр мтДНК предположительно являются трансверсии C → A (или G → T, если вести учет по H-цепи). Этот тип замен ожидается, когда гуанин в C:G парах окисляется до 8-оксигуанина [10]. Как показала настоящая работа, замены C → A действительно являются очень распространенными среди трансверсий, однако их частота составляет всего лишь примерно 1% от всех выявленных нуклеотидных замен. Кроме этого, различия по частоте замен C → A между высокогорными и не-горными популяциями статистически не значимы.

Наиболее частым типом нуклеотидных замен в L-спектрах мтДНК населения различных географических зон являются транзиции T → C ([17] и

настоящая работа). Ранее отмечалось, что преобладание транзиций $T \rightarrow C$ в мутационных спектрах мтДНК характерно для млекопитающих, и частота этой транзиции положительно коррелирует с продолжительностью жизни [9]. Показано также, что частота замен $T \rightarrow C$ на L-цепи зависит от времени нахождения H-цепи мтДНК в одноцепочечном состоянии во время репликации митохондриального генома [9]. Таким образом, происхождение замен $T \rightarrow C$, регистрируемых на L-цепи, по всей видимости, связано, в основном, с заменами $A \rightarrow G$, произошедшими на H-цепи. Одним из основных механизмов мутагенеза в этом случае считается гидролитическое дезаминирование аденина в одноцепочечных участках H-цепи [7, 9, 23]. Спонтанное превращение аденина в гипоксантин является потенциально мутагенным повреждением ДНК, приводящим к транзициям $A \rightarrow G$ [24]. Кроме этого, к заменам $A \rightarrow G$ могут привести и другие реакции: повреждения аденина гидроксильными радикалами [25], появление под воздействием АФК этенопроизводных аденина и других азотистых оснований [26]. Между тем, транзиции $T \rightarrow C$ могут генерироваться в результате воздействия гидроксильных радикалов на тиминовые основания с образованием тиминового гликоля, однако мутагенный эффект последнего считается низким [27].

Результаты исследований спектров нуклеотидных замен мтДНК человека показали, что важнейшим источником мутаций в митохондриальном геноме является дислокационный мутагенез, приводящий к появлению неспаренных оснований на участках смещения праймерной или матричной цепи в процессе репликации мтДНК [13, 28, 29]. Этот механизм объясняет происхождение примерно 20% варибельных позиций в генах мтДНК [29]. Анализ гиперварибельных участков главной некодирующей области мтДНК показал, что максимальное число дислокационных мутаций в H-цепи (43%) приходится на замены $A \rightarrow G$, в то время как на L-цепи происходят, в основном, замены $T \rightarrow C$ (79%) [28]. Таким образом, смещение цепей во время репликации мтДНК также может приводить к появлению существенного числа транзиций $T \rightarrow C$, регистрируемых в L-спектрах мтДНК.

Отсутствие различий между спектрами нуклеотидных замен мтДНК в высокогорных и негорных популяциях человека является важным результатом настоящей работы. Учитывая факты, свидетельствующие о том, что продукция АФК в митохондриях существенно возрастает в условиях гипоксии [1], ожидалось, что спектры нуклеотидных замен мтДНК в норме и при гипоксии будут различаться — особенно по частоте транзиций $T \rightarrow C$

(по L-цепи), которые предположительно более всего чувствительны к уровню окислительного стресса в митохондриях [9]. Однако сходство в распределении частот транзиций $T \rightarrow C$ и остальных нуклеотидных замен в различных региональных группах и на различной высоте проживания популяций человека свидетельствует о том, что митохондриальные геномы достаточно хорошо защищены от воздействия АФК. Этот вывод распространяется только на генеративные мутации митохондриального генома, поскольку в настоящей работе проводился филогенетический анализ нуклеотидных замен, наследуемых по материнской линии из поколения в поколение в широком временном интервале, вполне достаточном для визуализации генетических эффектов, связанных с адаптацией к различным условиям проживания популяций. Исследование спектров соматических мутаций мтДНК из различных биологических тканей и на протяжении жизни одного индивидуума, очевидно, может привести к несколько другим выводам и к большему разнообразию траекторий мутационных процессов [8, 9].

Исследование выполнено за счет гранта Российского научного фонда № 22-24-00264, <https://rscf.ru/project/22-24-00264/>.

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

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

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. *Burtscher J., Mallet R.T., Pialoux V. et al.* Adaptive responses to hypoxia and/or hyperoxia in humans // *Antioxid. Redox Signal.* 2022. V. 37. P. 887–912. <https://doi.org/10.1089/ars.2021.0280>
2. *Chandel N.S., McClintock D.S., Feliciano C.E. et al.* Reactive oxygen species generated at mitochondrial complex III stabilize hypoxia-inducible factor-1 α during hypoxia: a mechanism of O₂ sensing // *J. Biol. Chem.* 2000. V. 275. P. 25130–25138. <https://doi.org/10.1074/jbc.M001914200>
3. *Paddenbergh R., Ishaq B., Goldenberg A. et al.* Essential role of complex II of the respiratory chain in hypoxia-induced ROS generation in the pulmonary vasculature // *Am. J. Physiol. Lung Cell. Mol. Physiol.* 2003. V. 284. P. L710–L719. <https://doi.org/10.1152/ajplung.00149.2002>

4. Guzy R.D., Schumacker P.T. Oxygen sensing by mitochondria at complex III: The paradox of increased reactive oxygen species during hypoxia // *Exp. Physiol.* 2006. V. 91. P. 807–819.
<https://doi.org/10.1113/expphysiol.2006.033506>
5. Mallet R.T., Burtcher J., Pialoux V. et al. Molecular mechanisms of high-altitude acclimatization // *Int. J. Mol. Sci.* 2023. V. 24.
<https://doi.org/10.3390/ijms24021698>
6. Liu X., Hajnóczky G. Altered fusion dynamics underlie unique morphological changes in mitochondria during hypoxia-reoxygenation stress // *Cell Death Differ.* 2011. V. 18. P. 1561–1572.
<https://doi.org/10.1038/cdd.2011.13>
7. Alexeyev M.F. Is there more to aging than mitochondrial DNA and reactive oxygen species? // *FEBS J.* 2009. V. 276. P. 5768–5787.
<https://doi.org/10.1111/j.1742-4658.2009.07269.x>
8. Zsurka G., Peeva V., Kotlyar A., Kunz W.S. Is there still any role for oxidative stress in mitochondrial DNA-dependent aging? // *Genes (Basel)*. 2018. V. 9. P. 175.
<https://doi.org/10.3390/genes9040175>
9. Mikhailova A.G., Mikhailova A.A., Ushakova K. et al. A mitochondria-specific mutational signature of aging: increased rate of A > G substitutions on the heavy strand // *Nucl. Ac. Res.* 2022. V. 50. P. 10264–10277.
<https://doi.org/10.1093/nar/gkac779>
10. Richter C., Park J.-W., Ames B.N. Normal oxidative damage to mitochondrial and nuclear DNA is extensive // *Proc. Natl Acad. Sci. USA*. 1988. V. 85. P. 6465–6467.
<https://doi.org/10.1073/pnas.85.17.6465>
11. Cheng K.C., Cahill D.S., Kasai H. et al. 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G > T and A > C substitutions // *J. Biol. Chem.* 1992. V. 267. P. 166–172.
12. Kang D., Hamasaki N. Maintenance of mitochondrial DNA integrity: Repair and degradation // *Curr. Genet.* 2002. V. 41. P. 311–322.
<https://doi.org/10.1007/s00294-002-0312-0>
13. Malyarchuk B.A., Rogozin I.B., Berikov V.B., Derenko M.V. Analysis of phylogenetically reconstructed mutational spectra in human mitochondrial DNA control region // *Hum. Genet.* 2002. V. 111. P. 46–53.
<https://doi.org/10.1007/s00439-002-0740-4>
14. Малыарчук Б.А. Анализ распределения нуклеотидных замен в генах митохондриальной ДНК человека // *Генетика*. 2005. Т. 41. № 1. С. 93–99.
15. Kivisild T., Shen P., Wall D.P. et al. The role of selection in the evolution of human mitochondrial genomes // *Genetics*. 2006. V. 172. P. 373–387.
<https://doi.org/10.1534/genetics.105.043901>
16. Pereira L., Freitas F., Fernandes V. et al. The diversity present in 5140 human mitochondrial genomes // *Am. J. Hum. Genet.* 2009. V. 84. P. 628–640.
<https://doi.org/10.1016/j.ajhg.2009.04.013>
17. Малыарчук Б.А. Сравнительный анализ мутационных спектров митохондриальных геномов в популяциях человека // *Мол. биология*. 2023. Т. 57. № 5. С. 792–796.
<https://doi.org/10.31857/S0026898423050117>
18. Kang L., Zheng H.X., Chen F. et al. mtDNA lineage expansions in Sherpa population suggest adaptive evolution in Tibetan highlands // *Mol. Biol. Evol.* 2013. V. 30. P. 2579–2587.
<https://doi.org/10.1093/molbev/mst147>
19. Kang L., Zheng H.X., Zhang M. et al. MtDNA analysis reveals enriched pathogenic mutations in Tibetan highlanders // *Sci. Rep.* 2016. V. 6.
<https://doi.org/10.1038/srep31083>
20. Peng M.S., Xu W., Song J.J. et al. Mitochondrial genomes uncover the maternal history of the Pamir populations // *Eur. J. Hum. Genet.* 2018. V. 26. P. 124–136.
<https://doi.org/10.1038/s41431-017-0028-8>
21. Basnet R., Rai N., Tamang R. et al. The matrilineal ancestry of Nepali populations // *Hum. Genet.* 2023. V. 142. P. 167–180.
<https://doi.org/10.1007/s00439-022-02488-z>
22. Jeong C., Alkorta-Aranburu G., Basnyat B. et al. Admixture facilitates genetic adaptations to high altitude in Tibet // *Nat. Commun.* 2014. V. 5.
<https://doi.org/10.1038/ncomms4281>
23. Корниенко И.В., Малырчук Б.А. Анализ механизмов возникновения мутаций в митохондриальной ДНК человека // *Мол. биология*. 2005. Т. 39. № 5. С. 869–877.
24. Hill-Perkins M., Jones M.D., Karran P. Site-specific mutagenesis *in vivo* by single methylated or deaminated purine bases // *Mut. Res.* 1986. V. 162. P. 153–163.
[https://doi.org/10.1016/0027-5107\(86\)90081-3](https://doi.org/10.1016/0027-5107(86)90081-3)
25. Ide H., Yamaoka T., Kimura Y. Replication of DNA templates containing the alpha-anomer of deoxyadenosine, a major adenine lesion produced by hydroxyl radicals // *Biochemistry*. 1984. V. 33. P. 7127–7133.
<https://doi.org/10.1021/bi00189a016>
26. Pandya G.A., Moriya M. 1,N6-ethenodeoxyadenosine, a DNA adduct highly mutagenic in mammalian cells // *Biochemistry*. 1996. V. 35. P. 11487–11492.
<https://doi.org/10.1021/bi960170h>
27. Basu A.K., Loechler E.L., Leadon S.A., Essigmann J.M. Genetic effects of thymine glycol: site-specific mutagenesis and molecular modeling studies // *Proc. Natl Acad. Sci. USA*. 1989. V. 86. P. 7677–7681.
<https://doi.org/10.1073/pnas.86.20.7677>
28. Malyarchuk B.A., Rogozin I.B. Mutagenesis by transient misalignment in human mitochondrial DNA control region // *Ann. Hum. Genet.* 2004. V. 68. P. 324–339.
<https://doi.org/10.1046/j.1529-8817.2004.00099.x>
29. Малыарчук Б.А. Роль контекста в возникновении мутаций в генах митохондриальной ДНК человека // *Генетика*. 2005. Т. 41. № 3. С. 385–390.

Characterization of the Spectrum of Mitochondrial DNA Nucleotide Substitutions in Human Populations in High Altitude Environments

B. A. Malyarchuk^{a, *}

^a*Institute of Biological Problems of the North, Far East Branch of the Russian Academy of Sciences, Magadan, 685000 Russia*

^{*}*e-mail: malyarchuk@ibpn.ru*

Using phylogenetic analysis of the nucleotide sequences of whole mitochondrial genomes (mtDNA), the spectra of germinal nucleotide substitutions (on the L-chain of mtDNA) were reconstructed in highland populations of the Pamirs and Tibet in comparison with regional indigenous groups of West Asia, Northeast Siberia, and South Siberia. No differences were found in the distribution of nucleotide substitution frequencies in the mtDNA spectra depending on the population distribution by altitude. Pyrimidine transitions dominate in all mtDNA spectra, and T → C substitutions are the most frequent among them. Next in frequency in most regional groups are A → G substitutions, but in the Pamir and northeast Asian groups G → A substitutions are prevalent. Of the transversions in all populations studied C → A replacements were found to be predominant, except for the Tibetan one, where A → C substitutions are more frequent. The lack of differences in the distribution of mtDNA mutations in high-altitude and non-highland populations indicates that the structure of mtDNA spectra in human populations is independent of the oxidative stress intensity in mitochondria.

Keywords: mitochondrial genome, spectrum of nucleotide substitutions, germinal mutations, human populations, hypoxia.

ОСОБЕННОСТИ РАСПРОСТРАНЕНИЯ ГАПЛОГРУПП Y-ХРОМОСОМЫ В НАСЕЛЕНИИ САНКТ-ПЕТЕРБУРГА В СВЯЗИ С ПРОБЛЕМОЙ СОЗДАНИЯ РЕФЕРЕНТНЫХ БАЗ ДАННЫХ

© 2023 г. И. Г. Удина^{1, *}, А. С. Грачева^{1, 2}, С. А. Боринская¹, О. Л. Курбатова¹

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

²Научно-исследовательский институт общей реаниматологии имени В.А. Неговского
Федерального научно-клинического центра реаниматологии и реабилитологии, Москва, 107031 Россия

*e-mail: irina_udina@mail.ru

Поступила в редакцию 29.05.2023 г.

После доработки 14.06.2023 г.

Принята к публикации 15.06.2023 г.

В выборке из мужского населения Санкт-Петербурга определены гаплогруппы Y-хромосомы с помощью генотипирования по 18 STR (*DYS389I*, *DYS389II*, *DYS390*, *DYS19*, *DYS385A*, *DYS385B*, *DYS456*, *DYS437*, *DYS438*, *DYS447*, *DYS448*, *DYS449*, *DYS391*, *DYS392*, *DYS393*, *DYS439*, *DYS635* и *DYS576*) и собраны генетико-демографические данные путем анкетирования. Распределение гаплогрупп Y-хромосомы у жителей Санкт-Петербурга в целом соответствует русскому генофонду, с наиболее распространенными гаплогруппами *R1a*, *R1b*, *E1b1b1*, *N*, *T*, *I1*, *I2*, *J1* и *J2* и с превалированием гаплогруппы *R1a*. Отмечено присутствие “южных по происхождению” гаплогрупп (*C3*, *G2a*, *G2c*, *J1*, *J2*, *L*, *O2*, *O3*, *Q*, *R2* и *T*), попадающих в мегаполис с потоком мигрантов, с частотой 16% (в Москве – 18%). Сравнительный анализ распределений частот гаплогрупп Y-хромосомы у жителей Санкт-Петербурга и Москвы выявил статистически значимые различия по частоте гаплогруппы *E1b1b1*, а также различия по соотношению гаплогрупп *I1* и *I2*, обусловленные географическим положением мегаполисов. Установлены достоверные различия по частоте “южных по происхождению” гаплогрупп между исходной выборкой жителей Санкт-Петербурга (16%) и сформированной на основе данных анкетирования выборкой мужчин, не имевших предков другой этнической принадлежности по мужской линии в двух предшествующих поколениях (4%). Полученный результат подтверждает, что спектр “южных по происхождению” гаплогрупп проникает в генофонд населения мегаполиса с потоками мигрантов и свидетельствует о необходимости генетико-демографического анкетирования при формировании референтных баз данных для мегаполиса, а также для их своевременного обновления, обусловленного изменениями в генофонде под действием миграции.

Ключевые слова: Санкт-Петербург, Москва, население, мегаполис, миграция, генофонд, гаплогруппы Y-хромосомы, референтные базы данных.

DOI: 10.31857/S0016675823110164, **EDN:** QYBTDE

Маркеры Y-хромосомы широко используются в криминалистике для целей ДНК-идентификации личности, в том числе в решении таких важных для социума проблем, как установления отцовства, а также для исторических реконструкций и анализа этнодемографических процессов [1]. Специфические особенности географического распространения гаплогрупп Y-хромосомы позволяют эффективно маркировать и датировать миграционные потоки, что особенно востребовано при изучении миграции в мегаполисы на территории Российской Федерации, и рассматривать вклад определенных потоков генов в генофонд населения мегаполисов.

Изучение гаплогрупп Y-хромосомы позволяет анализировать генетическую структуру популяций

смешанного типа, исторически образованных в результате смешения различных по происхождению потоков населения. В населении США с помощью маркеров Y-хромосомы выявлен уровень смешения африканских, европейских, испаноязычных и азиатских (китайцев и корейцев) групп населения, а также коренного населения Америки – индейцев [2]. Маркеры Y-хромосомы весьма эффективно используются для установления происхождения жителей при смешении различных по происхождению потоков населения, например в истории населения Ирана [3] и Аргентины [4]. Как пример популяций смешанного типа в отношении этнических, антропологических и генетических признаков, можно привести население

мегаполисов, отчасти соответствующее модели “плавильного котла”.

Особенности генетико-демографических процессов в мегаполисах (высокая интенсивность миграционных процессов, полиэтнический состав населения, его смешанный характер, большая доля межнациональных браков) обуславливают повышенную сложность создания референтных баз данных для населения мегаполисов [5, 6]. Основным фактором динамики генофонда популяции мегаполиса является миграция [7], при этом наблюдаются гендерные различия в интенсивности миграционных процессов и этнорегиональном составе мигрантов, а следовательно и различия в характере динамики генофонда населения мегаполисов по однородительским ДНК-маркерам. Динамика генофонда населения мегаполиса под воздействием миграционных процессов обуславливает необходимость постоянного обновления и актуализации популяционных баз данных по генетическим маркерам для целей ДНК-идентификации, включая однородительские маркеры. Базы данных по маркерам Y-хромосомы интенсивно разрабатываются во всем мире [8].

Население Санкт-Петербурга – второго по численности населения мегаполиса Российской Федерации (5.6 млн чел. на 2023 г.) – ранее не было изучено в отношении маркеров Y-хромосомы, но было охарактеризовано в отношении генетико-демографических характеристик [9]. Для Санкт-Петербурга отмечены гендерные различия в интенсивности и этнорегиональном составе потока мигрантов [8]. Было показано, что области миграционного притяжения Москвы и Санкт-Петербурга заметно различаются, что говорит о том, что для двух мегаполисов нужно создавать отдельные базы данных [9, 10]. В этой связи целью нашего исследования является рассмотрение особенностей профиля частот гаплогрупп Y-хромосомы в населении Санкт-Петербурга и проведение сравнительного анализа с ранее изученным населением Москвы.

МАТЕРИАЛЫ И МЕТОДЫ

В 2012 г. Высшей школой экономики в сотрудничестве с ИОГен РАН в рамках проекта “Всемирное исследование ценностей” [11] проведено выборочное исследование жителей Санкт-Петербурга, включающее генетико-демографическое анкетирование ($N = 1037$) и сбор биологических образцов. Данные анкет содержали вопросы о национальности (по самоидентификации), месте и дате рождения анкетированного и его предков в двух предшествующих поколениях. Генетико-демографические параметры выборки представлены в [9]. Данные о молекулярных маркерах в изученной выборке из Санкт-Петербурга представлены в [12, 13].

В настоящее исследование включена только мужская выборка жителей Санкт-Петербурга, от которых были получены биологические образцы ($N = 150$). ДНК из образцов слюны выделена стандартным методом с применением набора реактивов фирмы “Изоген” (Москва). Проведено генотипирование образцов ДНК по 18 STR Y-хромосомы (*DYS389I*, *DYS389II*, *DYS390*, *DYS19*, *DYS385A*, *DYS385B*, *DYS456*, *DYS437*, *DYS438*, *DYS447*, *DYS448*, *DYS449*, *DYS391*, *DYS392*, *DYS393*, *DYS439*, *DYS635* и *DYS576*) на базе ООО “Гордиз” (Москва) с последующим определением принадлежности установленного генотипа к гаплогруппе Y-хромосомы с помощью Интернет-предиктора [14]. Статистический анализ проведен в программе “Statistica”.

От каждого лица, включенного в выборку, получено информированное согласие на использование биологического образца (слюны) и генетико-демографических данных в анонимном виде в популяционном исследовании.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Особенности распределения гаплогрупп Y-хромосомы

Частоты гаплогрупп Y-хромосомы в выборке мужчин из Санкт-Петербурга ($N = 150$) и ранее полученные сравнительные данные для Москвы [15] представлены в табл. 1. В выборке из Санкт-Петербурга наиболее распространены следующие гаплогруппы Y-хромосомы: *R1a*, *R1b*, *E1b1b1*, *N*, *T*, *I1*, *I2*, *J1* и *J2*, что в целом соответствует распределению гаплогрупп у русского населения [16]. Проведен сравнительный статистический анализ распределений частот гаплогрупп Y-хромосомы у жителей Санкт-Петербурга и Москвы: установлено статистически значимое превышение частоты гаплогруппы *E1b1b1* в Петербурге ($G = 8.8897$, $d.f. = 1$, $p < 0.01$). Гаплогруппа *E1b1b1* имеет ближневосточное происхождение, но может проникать в генофонд мегаполиса с запада в связи с особенностями распространения этой гаплогруппы в Европе [17, 18]. Наблюдаются различия между населением Санкт-Петербурга и Москвы по соотношению гаплогрупп *I1* и *I2* (табл. 2), которые вполне согласуются с географическим положением мегаполисов, различиями областей их миграционного притяжения [9, 10] и особенностями распространения рассматриваемых гаплогрупп в Европе (с максимальной представленностью гаплогруппы *I1* на Скандинавском полуострове) [18, 19]. Различия в частотах этих гаплогрупп не достигают порога достоверности, но по соотношению их частот население двух мегаполисов очевидно различается.

Таблица 1. Частоты гаплогрупп Y-хромосомы в изученных выборках из населения Санкт-Петербурга и сравнительные данные по населению Москвы

Гаплогруппа	Санкт-Петербург						Москва, мужчины***	
	мужчины		мужчины с русским отцом**		мужчины с русскими отцом и дедом**			
	<i>N</i>	<i>f</i>	<i>N</i>	<i>f</i>	<i>N</i>	<i>f</i>	<i>N</i>	<i>f</i>
<i>C3*</i>	2	0.013	1	0.009	0	0	7	0.012
<i>E1b1b1</i>	15	0.100	13	0.118	9	0.125	20	0.035
<i>G2a*</i>	3	0.020	3	0.027	0	0	10	0.018
<i>G2c*</i>	1	0.007	0	0	0	0	1	0.002
<i>H*</i>	0	0	0	0	0	0	1	0.002
<i>I1</i>	14	0.093	12	0.109	7	0.097	31	0.054
<i>I2a</i>	2	0.013	1	0.009	0	0	1	0.002
<i>I2a(xI2a1)</i>	9	0.060	7	0.064	5	0.069	57	0.100
<i>I2a1</i>	0	0	0	0	0	0	2	0.004
<i>I2b(xI2b1)</i>	0	0	0	0	0	0	1	0.002
<i>I2b1</i>	0	0	0	0	0	0	4	0.007
<i>J1*</i>	3	0.020	0	0	0	0	18	0.032
<i>J2a*</i>	8	0.053	3	0.027	2	0.028	20	0.035
<i>J2a1b*</i>	1	0.007	1	0.009	0	0	5	0.009
<i>J2ab1*</i>	0	0	0	0	0	0	1	0.002
<i>J2a1bh*</i>	0	0	0	0	0	0	1	0.002
<i>J2a4(xbh)*</i>	0	0	0	0	0	0	3	0.005
<i>J2b*</i>	3	0.020	2	0.018	1	0.014	7	0.012
<i>J2a1xJ2a1-bh*</i>	0	0	0	0	0	0	1	0.002
<i>L*</i>	1	0.007	1	0.009	0	0	2	0.004
<i>N</i>	17	0.113	15	0.136	10	0.139	54	0.095
<i>R1a</i>	62	0.414	47	0.429	35	0.486	260	0.457
<i>R1b</i>	7	0.047	4	0.036	3	0.042	36	0.063
<i>R2*</i>	0	0	0	0	0	0	4	0.007
<i>T*</i>	0	0	0	0	0	0	8	0.014
<i>O2*</i>	0	0	0	0	0	0	1	0.002
<i>O3*</i>	0	0	0	0	0	0	5	0.009
<i>Q*</i>	2	0.013	0	0	0	0	8	0.014
Всего	150	1	110	1	72	1	569	1

Примечание. * — “южные по происхождению” гаплогруппы; *N* — число лиц с данной гаплогруппой, а *f* — численная доля гаплогруппы в выборке. ** — по данным анкетирования. *** — приведены данные для москвичей [15].

“Южные по происхождению” гаплогруппы

Проведение статистического анализа распределения частот гаплогрупп Y-хромосомы в населении Москвы и литературных данных о географическом распространении гаплогрупп Y-хромосомы в населении России и сопредельных стран ранее [14] позволило установить спектр гаплогрупп, которые попадают в мегаполис с потоками

мигрантов, в основном с юга. Этот набор гаплогрупп Y-хромосомы (*C3*, *G2a*, *G2c*, *J1*, *J2*, *L*, *O2*, *O3*, *Q*, *R2* и *T*) обозначен как “южные по происхождению” гаплогруппы [14]. У жителей Санкт-Петербурга “южные по происхождению” гаплогруппы присутствуют с частотой 16.0% (в Москве — 18.1% [15]). Следует отметить, что в Санкт-Петербурге выявлены не все такие гаплогруппы, что,

Таблица 2. Соотношение частот гаплогрупп *I1* и *I2* в изученных выборках Санкт-Петербурга

Гаплогруппы	Санкт-Петербург						Москва, мужчины*** <i>N</i> = 569	
	мужчины <i>N</i> = 150		мужчины с русским отцом <i>N</i> = 110**		мужчины с русскими отцом и дедом** <i>N</i> = 72			
	<i>N</i>	<i>f</i>	<i>N</i>	<i>f</i>	<i>N</i>	<i>f</i>	<i>N</i>	<i>f</i>
<i>I1</i>	14	0.093	12	0.109	7	0.097	31	0.054
<i>I2</i>	11	0.073	8	0.073	5	0.069	65	0.114
Всего	25	0.175	20	0.182	12	0.166	96	0.168
Соотношение гаплогрупп <i>I1</i> и <i>I2</i>								
<i>I1/I2</i>	1.273		1.500		1.40		<i>0.477</i>	
<i>I2/I1</i>	0.786		0.667		0.741		2.097	

Примечание. Жирным шрифтом выделены максимальные значения, курсивом — минимальные; другие обозначения см. табл. 1.

Таблица 3. Частоты крупных гаплогрупп Y-хромосомы в изученных выборках из населения г. Санкт-Петербурга

Гаплогруппа	Санкт-Петербург						Москва, мужчины***	
	мужчины		мужчины с русским отцом**		мужчины с русскими отцом и дедом**			
	<i>N</i>	<i>f</i>	<i>N</i>	<i>f</i>	<i>N</i>	<i>f</i>	<i>N</i>	<i>f</i>
<i>E1b1b1</i>	15	0.100	13	0.118	9	0.125	20	0.035
<i>I1</i>	14	0.093	12	0.109	7	0.097	31	0.054
<i>I2</i>	11	0.073	8	0.073	5	0.069	65	0.114
<i>J1</i>	3	0.020	0	0	0	0	18	0.032
<i>J2</i>	12	0.080	6	0.055	3	0.042	38	0.067
<i>N</i>	17	0.113	15	0.136	10	0.139	54	0.095
<i>R1a</i>	62	0.414	47	0.428	35	0.486	260	0.457
<i>R1b</i>	7	0.047	4	0.036	3	0.042	36	0.063
<i>RH*</i>	9	0.060	5	0.045	0	0	47	0.083
Всего	150	1.000	110	1.000	72	1.000	569	1

Примечание. * *RH* — объединенная группа редких гаплогрупп, частота каждой из которых в суммарной выборке менее 2%; другие обозначения см. табл. 1.

возможно отчасти связано с меньшим размером изученной выборки по сравнению с выборкой из населения Москвы.

Генетико-демографический подход при создании референтных баз данных для мегаполисов

На основе данных анкетирования сформированы две выборки русских по самоидентификации мужчин, не имевших предков другой этнической принадлежности по мужской линии: 1) в одном поколении (с русскими отцами) или 2) в двух предыдущих поколениях (с русскими отцами и дедами) (табл. 1 и 3). Очевидно, что при анкетировании может быть получена неполная информация.

В исходной выборке представлено 16 гаплогрупп, в выборке с русскими предками в одном поколении — 13, а в двух поколениях — всего 8. Примечательно, что в выборке мужчин с русскими предками в двух поколениях наблюдается исчезновение практически только “южных по происхождению” гаплогрупп (*C3*, *G2a*, *G2c*, *J1*, *J2a1b*, *L* и *Q*). С одной стороны, такой результат представляется закономерным, а с другой стороны — свидетельствует о том, что процесс этнического смешения в мегаполисе имеет недавнее происхождение.

Установлены достоверные различия по частоте “южных по происхождению” гаплогрупп, приносимых в генофонд населения мегаполиса ми-

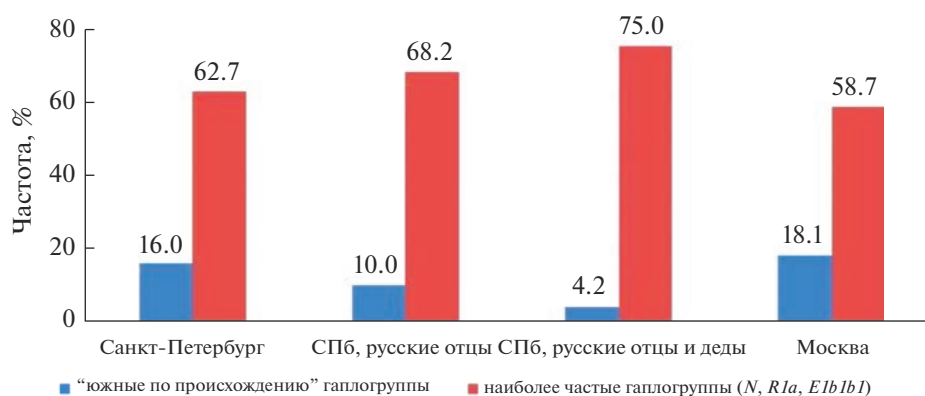


Рис. 1. Суммарная частота «южных по происхождению» гаплогрупп Y-хромосомы и и суммарная частота трех наиболее частых гаплогрупп в выборках населения Санкт-Петербурга и Москвы (по данным табл. 1).

грантами, между исходной выборкой жителей Санкт-Петербурга (16.0%) и выборкой мужчин с русскими предками в двух предшествующих поколениях (4.2%): ($G = 7.3321$, $d.f. = 1$, $p < 0.01$).

На рис. 1 показаны совокупные частоты «южных по происхождению» гаплогрупп и трех самых распространенных гаплогрупп в Санкт-Петербурге (с частотой равной или выше 10%) в выборках с различными генетико-демографическими характеристиками. Следует отметить, что чем меньше присутствие «южных по происхождению» гаплогрупп, тем выше суммарная представленность трех наиболее частых гаплогрупп (*E1b1b1*, *N* и *R1a*). Аналогичный достоверный эффект наблюдали в поколениях населения Москвы для распространенной гаплогруппы *N* [15].

В табл. 3 для наглядности представлено распределение крупных гаплогрупп для сравнительного анализа. По суммарному профилю частот гаплогрупп достоверных различий между группами не выявлено. В этой таблице «южные по происхождению» гаплогруппы объединены в группу редких генотипов *RH* (за исключением *J1* и *J2*). Установленные достоверные различия по частоте гаплогруппы *E1b1b1* и соотношению гаплогрупп *I1* и *I2*, а также специфические особенности распространения наиболее частых гаплогрупп в двух мегаполисах подтверждают необходимость создания отдельных референтных баз данных для рассматриваемых мегаполисов, которая уже определена различной областью миграционного притяжения двух мегаполисов по генетико-демографическим данным (см. выше).

Полученные результаты подтверждают спектр ранее выявленных «южных по происхождению» гаплогрупп [15] как проникающих в генофонд населения мегаполиса с потоками мигрантов и свидетельствуют о необходимости генетико-демографического анкетирования при формировании референтных баз данных для мегаполиса. Также

необходимо своевременное обновление баз данных в связи с изменениями в генофонде под действием миграции. Ранее для жителей Москвы и Санкт-Петербурга нами были разработаны базы данных, содержащие одновременно молекулярно-генетические данные (включая гаплогруппы Y-хромосомы) и генетико-демографические характеристики [20, 21].

Очевидно, что частоты маркеров Y-хромосомы в мегаполисах могут существенно изменяться за несколько поколений, а также могут заметно различаться в одновременно живущих возрастных группах в связи с высокой интенсивностью миграционных процессов. Для населения Санкт-Петербурга оценка коэффициента миграции составила 0.34, что говорит о том, что генофонд популяции практически полностью обновляется за восемь поколений; частотный профиль гаплогрупп Y-хромосомы находится в хорошем соответствии с этнорегиональным составом мигрантов [9]. Для населения Москвы ранее показано увеличение в поколениях частот «южных по происхождению» гаплогрупп Y-хромосомы, привносимых мигрантами из регионов их происхождения [15].

Таким образом, для повышения надежности ДНК-идентификации в криминалистике, референтные базы данных необходимо постоянно обновлять, вследствие интенсивного притока мигрантов, который приводит к изменениям этнического состава населения мегаполиса в поколениях.

Исследование проведено в рамках темы государственного задания ИОГен РАН «Исследования полиморфизма на клеточном, организменном и популяционном уровне как основа создания генетических технологий» (№ 0092-2022-0001).

Авторы благодарят коллег, проводивших сбор биологических образцов и анкетирование жителей Санкт-Петербурга, и самих петербуржцев, принявших участие в исследовании, за ценное сотрудничество.

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.

Исследование одобрено Этическим комитетом ИОГен РАН. От каждого из включенных в исследование участников было получено информированное добровольное согласие на использование биологических образцов и генетико-демографических данных, полученных в результате анкетирования. Результаты исследования представлены и хранятся в анонимной (обезличенной) форме.

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Юсупов Ю.М., Балановская Е.В., Сабитов Ж.М., Балановский О.П. Комплексные исследования этногенеза: союз геногеографии и этнографии // Вестник антропологии. 2017. Т. 38. № 2. С. 28–35.
2. Hammer M.F., Chamberlain V.F., Kearney V.F. et al. Population structure of Y chromosome SNP haplogroups in the United States and forensic implications for constructing Y chromosome STR databases // Forensic Sci. Int. 2006. V. 1. № 164(1). P. 45–55. <https://doi.org/10.1016/j.forsciint.2005.11.013>
3. Grugni V., Battaglia V., Hooshiar Kashani B. et al. Ancient migratory events in the Middle East: new clues from the Y chromosome variation of modern Iranians // PLoS One. 2012. V. 7(7). <https://doi.org/10.1371/journal.pone.0041252>
4. Caputo M., Sala A., Corach D. Demand for larger Y-STR reference databases in ethnic melting-pot countries: Argentina as a test case // Int. J. Legal Med. 2019. V. 133 № 5. P. 1309–1320. <https://doi.org/10.1007/s00414-019-02012-5>
5. Курбатова О.Л., Победоносцева Е.Ю., Веремейчик В.М. и др. Особенности генетико-демографических процессов в населении трех мегаполисов в связи с проблемой создания генетических баз данных // Генетика. 2013. Т. 49. № 4. С. 513–522. <https://doi.org/10.7868/S0016675813040085>
6. Цыбовский И.С., Веремейчик В.М., Котова С.А. и др. Создание судебной референтной базы данных по 18 аутосомным STR для ДНК-идентификации в Республике Беларусь // Генетика. 2017. Т. 53. № 2. С. 249–258. <https://doi.org/10.7868/S0016675817020138>
7. Курбатова О.Л., Янковский Н.К. Миграция – основной фактор популяционной динамики городского населения России // Генетика. 2016. Т. 52. № 7. С. 831–851. <https://doi.org/10.7868/S0016675816070067>
8. de Knijff P. On the forensic use of Y chromosome polymorphisms // Genes (Basel). 2022. V. 17. № 13(5). <https://doi.org/10.3390/genes13050898>
9. Курбатова О.Л., Удина И.Г., Грачева А.С. и др. Генетико-демографические параметры населения г. Санкт-Петербурга. Миграционные процессы // Генетика. 2019. Т. 55. № 9. С. 1071–1082. <https://doi.org/10.1134/S001667581909008X>
10. Курбатова О.Л., Грачева А.С., Победоносцева Е.Ю., Удина И.Г. Генетико-демографические параметры населения г. Москвы. Миграционные процессы // Генетика. 2021. Т. 56. № 12. С. 438–1449. <https://doi.org/31857/S0016675821120080>
11. Понарин Э.Д., Алмакаева А.М., Немировская А.В., Ильченко Н.А. Обзор российских регионов ЛССИ. Москва: ВШЭ, 2012. 121 с. // Свидетельство о регистрации базы данных 2021620200, 02.02.2021. Заявка № 2020622484 от 02.12.2020.
12. Гуреев А.С., Ананьева Е.Д., Рубанович А.В. и др. Ассоциация UVNTR-аллелей гена MAOA с субъективной оценкой благополучия у мужчин // Генетика. 2018. Т. 54. № 5. С. 556–562. <https://doi.org/10.7868/S0016675818050065>
13. Inglehart R.F., Borinskaya S., Cotter A. et al. Genetic factors, cultural predispositions, happiness and gender equality // J. of Res. in Gender Studies. 2014. V. 4(1). P. 32–100.
14. Whit Athey's Haplogroup Predictor. [Электронный ресурс]. Режим доступа: <http://www.hprg.com/hapest6/hapest6G2a/hapest6.htm?order=orig>
15. Удина И.Г., Грачева А.С., Курбатова О.Л. Частоты гаплогрупп Y-хромосомы и процессы миграции в трех поколениях жителей Москвы // Генетика. 2022. Т. 58. № 11. С. 1325–1333. <https://doi.org/10.31857/S001667582110121>
16. Балановская Е.В., Балановский О.П. Русский генофонд на Русской равнине. М.: Луч, 2007. 416 с.
17. Ближневосточные гаплогруппы J1, J2, E1b1b1, G2a, T и др. Описание и связь с археологическими культурами. [Электронный ресурс]. Режим доступа: http://haplogroup.narod.ru/eur_hap2.html
18. Карта географического распространения гаплогрупп Y-хромосомы на территории Евразии. [Электронный ресурс]. Режим доступа: <https://pp.vk.me/bu05HtNeinmUcGTLGwGE0Asv1Z0Xa-wlcEtYHng/nl-BfaaxxKUQ.jpg>
19. Гаплогруппы коренных древних европейцев, гаплогруппа I, II и I2 у народов Европы. [Электронный ресурс]. Режим доступа: <http://haplogroup.narod.ru/ineurope.html>
20. Курбатова О.Л., Победоносцева Е.Ю., Удина И.Г., Грачева А.С. МЕГАПОЛИС–ДНК–1Р. Свидетельство о регистрации базы данных 2021620596, 29.03.2021. Заявка № 2021620445 от 19.03.2021.
21. Курбатова О.Л., Победоносцева Е.Ю., Удина И.Г. и др. МЕГАПОЛИС–ДНК–2Р // Свидетельство о регистрации базы данных 2021620610, 30.03.2021. Заявка № 2021620452 от 19.03.2021.

Distribution Peculiarities of Y-Chromosome Haplogroups in the Population of St. Petersburg in Connection with the Problem of Creation Reference Data Bases

I. G. Udina^{a, *}, A. S. Gracheva^{a, b}, S. A. Borinskaya^a, and O. L. Kurbatova^a

^a*Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, 119991 Russia*

^b*Negovsky Research Institute of General Reanimatology, Federal Research and Clinical Center of Intensive Care Medicine and Rehabilitology, Moscow, 107031 Russia*

*e-mail: irina_udina@mail.ru

In the sample of male residents of St. Petersburg, Y-chromosome haplogroups were determined by genotyping 18 STR Y-chromosome (*DYS389I*, *DYS389II*, *DYS390*, *DYS19*, *DYS385A*, *DYS385B*, *DYS456*, *DYS437*, *DYS438*, *DYS447*, *DYS448*, *DYS449*, *DYS391*, *DYS392*, *DYS393*, *DYS439*, *DYS635* and *DYS576*) and data on genetic demography were collected by means of a questionnaire. The distribution of Y-chromosome haplogroups in St. Petersburg residents generally corresponds to the published data on Russian gene pool, with the most frequent haplogroups *R1a*, *R1b*, *E1b1b1*, *N*, *T*, *I1*, *I2*, *J1* and *J2*, and with the predominance of haplogroup *R1a*. The presence of “southern by origin” haplogroups (*C3*, *G2a*, *G2c*, *J1*, *J2*, *L*, *O2*, *O3*, *Q*, *R2* and *T*) entering the megalopolis with a flow of migrants, with a total frequency of 16% (in Moscow – 18.1%) was noted. A comparative analysis of the frequency distributions of Y-chromosome haplogroups in residents of St. Petersburg and Moscow revealed statistically significant differences in the frequency of haplogroup *E1b1b1*, and differences in the ratio of *I1* and *I2*, determined by geographic position. Based on the survey data, a sample of Russian men who had no ancestors of another ethnicity in the male line in the two previous generations was formed. Significant differences in the frequency of “southern-origin” haplogroups were established between the initial sample of residents of St. Petersburg (16%) and the sample of men with Russian ancestors in two previous generations (4.1%). The obtained result confirms the spectrum of haplogroups of “southern origin” as penetrating into the gene pool of the population of a megalopolis with migrant flows and indicates the need for genetic and demographic questionnaires when forming reference databases for a megalopolis, as well as for their timely updating due to changes in the gene pool under the influence of migration.

Keywords: St. Petersburg, Moscow, population, gene pool, megalopolis, Y-chromosome haplogroups, migration, reference databases.

ОПИСАНИЕ ДИВЕРГЕНЦИИ СУБПОПУЛЯЦИЙ В ИЕРАРХИЧЕСКОЙ
СИСТЕМЕ ПРИ АНАЛИЗЕ ИЗОНИМИИ.
II. ВЕРОЯТНОСТИ НЕИЗОНИМНЫХ ВСТРЕЧ© 2023 г. В. П. Пасеков¹, *¹Федеральный исследовательский центр “Информатика и управление” Российской академии наук,
Москва, 119991 Россия

*e-mail: pass40@mail.ru

Поступила в редакцию 01.05.2023 г.

После доработки 22.05.2023 г.

Принята к публикации 25.05.2023 г.

Рассматриваются типичные метапопуляции человека с иерархической подразделенностью на части (субпопуляции), соответствующие классификации на основе административно-территориального деления (скажем, село, сельсовет, район, область и так далее) или генеалогического подхода, базирующегося на этногенезе, а также на других принципах биологической классификации. Целью настоящей работы является анализ общих свойств распределения концентрации фамилии по субпопуляциям при их иерархической структуре. Внимание концентрируется на описании семейной дивергенции субпопуляций, в качестве показателя которой рассматривается общая вероятность встреч (H_s) лиц с разными фамилиями, понимаемая как вероятность встречи в выбранной наугад субпопуляции, единицы наблюдения, из рассматриваемой метапопуляции с иерархической структурой. Данная вероятность является семейным аналогом концентрации гетерозигот в метапопуляции со случайным скрещиванием в ее субпопуляциях, единицах наблюдения. Получено разложение H_s по уровням иерархии, обобщающее эффект Валунда в популяционной генетике. Общая вероятность неизонимных встреч в иерархически подразделенной метапопуляции меньше вероятности случайных встреч в ней на сумму средних внутригрупповых дисперсий концентрации фамилии, соответствующих отдельным уровням. Такие свойства являются чисто статистическими характеристиками иерархической структуры, а не особенностью конкретной популяционной системы и не выводятся из закономерностей той или иной модели микроэволюции. Они вычислительно формулируются одинаково для любой иерархической системы, хотя в общем случае не совпадают количественно. Полученные результаты относятся к сельским и городским иерархическим метапопуляциям как отдельным компонентам всего населения.

Ключевые слова: иерархическая структура популяций, метапопуляции, концентрации фамилии в субпопуляциях человека, характеристики неоднородности субпопуляций, разложение вероятности встреч носителей разных фамилий по уровням иерархии.

DOI: 10.31857/S0016675823110097, EDN: TNBXXK

Случайный генный дрейф [1] приводит к генетической дивергенции популяций с общим происхождением. Аналогично в результате случайного дрейфа фамилий [2, 3] происходит семейная дивергенция таких популяций. Оба этих процесса протекают синхронно в одной и той же популяции с ограниченной численностью. Семейный состав следующего поколения с численностью N мужской составляющей можно упрощенно рассматривать как результат случайной выборки с возвращением размера N из совокупности фамилий мужчин родительского поколения. Генетический состав нового поколения диплоидной популяции при отсутствии систематических давлений является результатом выборки $4N$ гамет роди-

тельского поколения ($2N$ размер популяции с учетом женщин и $4N$ размер выборки гамет, формирующих $2N$ диплоидных потомков — по две гаметы на потомка). В результате имеется полная аналогия модели семейного дрейфа с генетической моделью диплоидной популяции со случайным скрещиванием, рассматриваемой в отношении одного аутосомного локуса с множественными аллелями. Роль фамилий играют аллели. Аналогом случайной встречи пары индивидуумов с разными фамилиями (следовательно, с разными родоначальниками) является появление гетерозиготного генотипа при случайном скрещивании, а аналогом пар однофамильцев будут гомозиготы.

Очевидно, что закономерности динамики фамильного и генетического состава популяции имеют один и тот же характер выборочного дрейфа. Когда закономерности данных процессов сходны, то по наблюдению за результатами одного из них можно делать выводы о результатах протекания другого. Однако оба процесса не идентичны. Для одной и той же диплоидной популяции случайный процесс выборочных колебаний концентраций фамилий в ряду поколений интенсивнее генного дрейфа (соответствует вчетверо меньшей выборке фамилий, чем выборка гамет). Подробнее об этом говорится в [2, 3].

Популяционно-генетический анализ оперирует с данными по генетической структуре популяций. Получение данных по фамильной структуре менее трудоемко и менее затратно. Поэтому желательно проанализировать пути экстраполяции результатов дивергенции популяций по фамильным данным в генетические выводы. При этом важно уметь описывать свойства дивергенции в типичных для населения ситуациях, и мы начнем с такого описания.

Для метапопуляций человека типична иерархическая подразделенность на части (субпопуляции), соответствующие классификации на основе административно-территориального деления (скажем, село, сельсовет, район, область и так далее) или генеалогического подхода, базирующегося на этногенезе, а также на других принципах биологической классификации. В предыдущей части [4] данной работы в качестве характеристики фамильной дивергенции в иерархически подразделенной метапопуляции служила дисперсия распределения фамилии по субпопуляциям. Теперь обратимся к другой характеристике — общей (полной) вероятности случайной встречи (столкновения) индивидуума с фиксированной фамилией с носителем какой-либо другой, которую понимаем как вероятность встречи в выбранной наугад субпопуляции, единице наблюдения, из рассматриваемой метапопуляции с иерархической структурой.

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. Повторим, что вероятность рассматриваемой встречи аналогична доле (концентрации) гетерозигот в популяционной генетике. Это позволяет при соответствующих предположениях оценивать такой важный генетический показатель как коэффициент инбридинга популяции [5].

Итак, объектом нашего анализа служит реальная или теоретическая метапопуляция s с иерархической подразделенностью на субпопуляции $\{s_i\}$ по уровням иерархии. При подразделенности системы *отношение подчинения означает вхождение подсистемы низкого уровня в качестве составной части в соответствующую подсистему (группу) более высокого ранга*. Например, субпопуляции (скажем, сёла) группируются в сельсоветы, районы и т.д. с уровнями иерархии 2, 3 ... соответственно, и множество субпопуляций на каждом из этих уровней представляет собой разбиение рассматриваемой метапопуляции. Здесь выполняется иерархическое подчинение одной субпопуляции другой, т.е. вхождение первой в качестве части во вторую с более высоким уровнем иерархии.

Дадим несколько слов 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 (V_s и F_s соответственно). Символ тождества ($\equiv$) у нас используется в смысле равенства по определению, а символ $\blacktriangleleft$ обозначает конец доказательства.

Напомним, что идентификацию и положение субпопуляции в иерархической структуре можно задавать с помощью мультиномеров, как это было сделано нами ранее в предыдущей части [4]. Пусть номер конкретного села (первый уровень) обозначается как s_1 ; номер сельсовета (второй уровень), куда входит это село, как s_2 ; номер района (третий уровень), включающего указанные сельсовет и село, как s_3 ; и т.д. Тогда мультиномер $\mathbf{s}_1 \equiv (s_1, s_2, s_3, \dots)$ однозначно определяет рассматриваемое село среди прочих сел первого уровня, вектор $\mathbf{s}_2 \equiv (s_2, s_3, s_4, \dots)$ — идентификатор сельсовета, вектор $\mathbf{s}_i \equiv (s_i, s_{i+1}, \dots)$ идентифицирует субпопуляцию i -го уровня среди прочих таких же субпопуляций внутри соответствующей группы следующего уровня $i + 1$. Это аналогично почтовому адресу, в котором административные составляющие заменены числами.

В результате субпопуляция s_1 входит в надлежащую субпопуляцию s_2 , а s_i входит в s_{i+1} и т.д., т.е. субпопуляция некоторого уровня иерархии включает в себя в качестве своей части соответствующие субпопуляции более низкого уровня. Между объектами и их идентификаторами имеется взаимно однозначное соответствие, и мы иногда будем писать идентификатор вместо названия объекта (села, сельсовета и т.д.). Кроме того, повторим, что множество субпопуляций на каждом отдельно выбранном уровне иерархии представ-

ляет собой разбиение всей метапопуляции, т.е. составляет ее целиком.

Данный способ нумерации приложим также к концентрации x интересующей фамилии, которую в селе s_1 будем обозначать как $x(s_1)$, а концентрацию фамилии в субпопуляции i -го уровня как $x(s_i)$. Когда некоторые из номеров $\{s_i\}$ рассматриваются как случайные величины (например, при выборе субпопуляции наугад), а другие как фиксированные, то для наглядности будем писать фиксированные величины после вертикальной черты. В частности, будем рассматривать $x(s_{i-1}|s_i)$ как *случайную величину*, значениями которой являются концентрации фамилии, получаемые при выборе наугад субпопуляций ($i - 1$)-го уровня из фиксированной группы s_i . *Первый аргумент s_{i-1} у $x(s_{i-1}|s_i)$ указывает на случайно выбираемую субпопуляцию уровня $i - 1$, а второй (s_i) — на содержащую ее фиксированную группу уровня i .*

Таким образом, концентрация фамилии в фиксированной субпопуляции s_i обозначается как $x(s_i)$, а $x(s_{i-1}|s_i)$ дает случайную величину, принимающую значения концентраций фамилии в субпопуляциях уровня $i - 1$, выбираемых наугад из s_i . Обозначим математическое ожидание (среднее значение) случайной величины $x(s_{i-1}|s_i)$ как $\bar{x}(s_i)$, а ее дисперсию как $Vs(x(s_{i-1}|s_i))$.

Цель настоящей работы — продолжение анализа в [4] общих свойств распределения концентрации фамилии по субпопуляциям при их иерархической подразделенности. Они являются чисто статистическими характеристиками иерархической структуры, а не особенностью конкретной популяционной системы. Анализируемые свойства присущи любой иерархически подразделенной метапопуляции и не выводятся из закономерностей той или иной модели микроэволюции. Они вычислительно формулируются одинаково для любой иерархической системы, хотя в общем случае не совпадают количественно. Теоретически это может позволить выделить специфические особенности исследуемого материала. По-прежнему внимание концентрируется на изучении дивергенции субпопуляций, в качестве показателя которой теперь рассматривается не дисперсия концентрации фамилии в субпопуляциях [4], а вероятности встреч пары лиц с неизонимными фамилиями. Обратим внимание на то, что *в рассматриваемых вероятностях учитывается порядок фамилий в паре*, что не вполне традиционно.

Структура изложения материала следующая. После напоминания системы идентификации субпопуляций в метапопуляции с иерархической структурой рассматриваются вероятности встреч носителей разных фамилий в качестве показателей дивергенции субпопуляций на разных уровнях иерархии. Далее подробно рассматривается из-

менчивость субпопуляций в случае многоуровневой иерархической структуры метапопуляции. Выведено разложение по уровням иерархии общей (полной) вероятности неизонимных встреч в выбранной наугад субпопуляции, единице наблюдения. В найденном разложении каждому отдельному уровню иерархии соответствует неотрицательный компонент, равный соответствующей средней внутригрупповой дисперсии концентрации фамилии.

ПОДРАЗДЕЛЕННЫЕ МЕТАПОПУЛЯЦИИ

В случае изучения популяций человека классификация и группировка данных часто производятся на основе административно-территориального деления, имеющего иерархический характер (скажем, село, сельсовет, район, область и др.), генеалогического подхода на основе этногенеза, лингвистических данных и пр. Получаемая группировка субпопуляций приближенно является иерархической. Иерархическая структура метапопуляции отражается на ее свойствах, в частности на распределении фамилий в популяциях человека, где типична опора на официальные данные иерархического характера, сбор и обработку материалов в соответствии с ними. Настоящая статья мотивирована анализом фамильных данных с ориентацией на популяционную генетику. Использование фамилий для получения выводов о генетической структуре популяций основывается на существующих параллелях в передаче на популяционном уровне потомкам фамилий и аутосомных аллелей (см., например, [2, 3]). Плодотворность такого использования продемонстрирована в ряде работ [6–8] (изонимные браки) [9] (фундаментальная монография), в том числе в исследованиях популяций России [10] (медико-генетические аспекты).

Повторим, что у нас объектами являются субпопуляции, их совокупность образует метапопуляцию (иерархическую систему), подчинение означает принадлежность одной субпопуляции другой в качестве ее части. При этом у самой метапопуляции наивысший уровень иерархии, и все субпопуляции являются ее частями. В роли числового признака субпопуляции может рассматриваться любая числовая характеристика, но мы ограничимся преимущественно дискретными признаками.

На популяционном уровне в качестве признака субпопуляции рассматривается среднее значение соответствующей характеристики для множества входящих в нее субпопуляций более низкого уровня. Вычисление средних значений основано на суммировании, и желательно выяснить результат разбиений метапопуляции или субпопуляции на более мелкие группы для соответствующего среднего значения признака. При его вычислении суммирование происходит по входящим в объект базовым единицам самого низкого первого уровня,

которые у нас считаются заданными (служат начальными данными). Варьирование характера разбиения приводит лишь к перегруппировке слагаемых, а результат остается неизменным. Таким образом, среднее значение признака рассматриваемого объекта не зависит от характера разбиения последнего, что подробнее рассматривается далее.

Важным примером характеристики субпопуляции, состоящей из индивидуумов с дискретным признаком T , является средняя величина T , принимающего значения 1 и 0 в зависимости от его категории (1 для интересующей категории и 0 в противном случае; скажем, интересующей категорией может быть фамилия, аллель, гетерозигота и т.п.). Покажем, что среднее значение T в произвольной субпопуляции s_i представляет собой не что иное как концентрацию в ней носителей интересующей категории. Эта концентрация также не зависит от характера подразделенности s_i .

Замечание 1. Пусть субпопуляция s_i состоит из индивидуумов с дискретным признаком T , кодируемым единицей при интересующей категории и кодируемым нулем в противном случае.

Тогда среднее значение T в s_i равно концентрации носителей рассматриваемой категории в данной субпопуляции (доле ее носителей; иначе говоря, вероятности случайно встретить такого носителя в субпопуляции при одинаковых шансах встретить любого индивидуума).

Доказательство. Среднее значение T находится как сумма значений T (по условию это 1 и 0) с весами, равными вероятностям, с которыми T принимает данные значения, т.е. равно

$$1 \times Pr\{T = 1\} + 0 \times Pr\{T = 0\} = Pr\{T = 1\},$$

где $Pr\{T = \dots\}$ обозначает вероятность наблюдения соответствующего значения T . Вероятность $Pr\{T = 1\}$ интерпретируется как вероятность при выборе наугад из s_i индивидуума получить носителя интересующей категории. При одинаковых шансах выбрать любого индивидуума вероятность $Pr\{T = 1\}$ равна доле (концентрации) носителей этой категории в s_i . ◀

Далее рассматривается только популяционный уровень организации, т.е. сами носители остаются за кадром, а анализируемыми объектами являются субпопуляции, характеризующие средними значениями (концентрациями) интересующих показателей, например фамилии. Для субпопуляций самого низкого первого уровня иерархии значения концентрации фамилии полагаем заданными исходно (рассматриваем их как начальные данные).

Обратим внимание на то, что даже когда предметом изучения является единственная иерархически подразделенная метапопуляция, выводы относительно ее свойств можно делать в вероятностных терминах в результате трактовки харак-

теристик субпопуляции как результатов ее случайного выбора из соответствующего множества субпопуляций, частей разбиения метапопуляции. При этом поскольку концентрация является средним значением для T , то к ней приложимы известные свойства средних значений (математических ожиданий).

Рассмотрим метапопуляцию s_j . Покажем, что концентрация $T(s_j)$ носителей интересующей категории дискретного признака T в s_j равна средней величине для его концентраций в составляющих разбиение s_j субпопуляциях $\{s_i\}$ и не зависит от разбиения.

Замечание 2. Пусть метапопуляция s_j с общей численностью $N(s_j)$ разбита на какие-либо группы $\{s_i\}$ с численностями $\{N(s_i)\}$. Положим численность носителей рассматриваемой категории T в метапопуляции s_j обозначена как $T_N(s_j)$.

Тогда концентрация $T(s_j) \equiv T_N(s_j)/N(s_j)$ носителей данной категории T в метапопуляции s_j , произвольно разбитой на группы $\{s_i\}$, находится как

$$\begin{aligned} T(s_j) &= \sum_{s_i \in s_j} T(s_i) \frac{N(s_i)}{N(s_j)} \equiv \\ &\equiv \sum_{s_i \in s_j} T(s_i) Pr(s_i | s_j) = E_{s_i} \{T(s_i) | s_j\}, \end{aligned}$$

$$0 \leq Pr(s_i | s_j) \equiv \frac{N(s_i)}{N(s_j)} \leq 1, \quad \sum_{s_i} Pr(s_i | s_j) = 1,$$

т.е. как математическое ожидание случайной величины $T(s_i | s_j)$, концентрации T в s_i при выборе наугад s_i из s_j с вероятностью $Pr(s_i | s_j) \equiv N(s_i)/N(s_j)$. Здесь E является символом операции получения математического ожидания, нижний индекс у E указывает на используемую при усреднении переменную.

Значение $T(s_j)$ не зависит от разбиения метапопуляции s_j , и $T(s_j)$ не меняется при его изменении.

Доказательство. Концентрация $T(s_j)$ рассматриваемых носителей в разбитой на группы $\{s_i\}$ метапопуляции s_j определяется как отношение их численности $T_N(s_j)$ в s_j , складывающейся из численностей в подгруппах $\{T_N(s_i)\}$, к общей численности $N(s_j)$ метапопуляции s_j , складывающейся из численностей ее подгрупп $\{N(s_i)\}$. Для субпопуляций s_i и $s_j = \{s_i\}$ имеем

$$\begin{aligned} T_N(s_i) &\equiv T(s_i) N(s_i), \\ T_N(s_j) &= \sum_{s_i \in s_j} T_N(s_i) = T(s_i) N(s_i), \end{aligned}$$

$$T(s_j) \equiv T_N(s_j)/N(s_j) = \sum_{s_i \in s_j} T(s_i) \frac{N(s_i)}{N(s_j)}.$$

Следовательно, концентрация $T(s_j)$ в s_j равняется среднему взвешенному значению для концентраций $\{T(s_i)\}$ в составляющих разбиение s_j субпопуля-

циях $\{s_j\}$ с весами, равными относительным численностям субпопуляций $\{N(s_j)/N(s_j)\}$. Очевидно, эти веса в сумме по $\{s_j\}$ дают единицу и неотрицательны. Их значения $\{N(s_j)/N(s_j)\}$ можно интерпретировать как вероятности $\{Pr(s_j|s_j)\}$ попасть на индивидуума из соответствующей субпопуляции s_j при выборе наугад индивидуума из s_j (как вероятности выбора субпопуляции s_i из s_j при одинаковых шансах каждого быть выбранным).

Таким образом, формула для $T(s_j)$ показывает, что $T(s_j)$ является математическим ожиданием для $T(s_i|s_j)$ при выборе наугад соответствующей субпопуляции s_i из s_j :

$$\begin{aligned} T(s_j) &\equiv \sum_{s_i \in s_j} T(s_i) \frac{N(s_i)}{N(s_j)} = \\ &= \sum_{s_i \in s_j} T(s_i) Pr(s_i|s_j) = E_{s_i} \{T(s_i)|s_j\}. \end{aligned}$$

Очевидно, приведенный вывод верен для любого разбиения s_j на субпопуляции $\{s_i\}$. При этом ее общая численность $N(s_j)$ и общее количество носителей $T_N(s_j)$ в ней не изменятся в результате выбора другого разбиения. Следовательно, их отношение (концентрация носителей) не зависит от разбиения s_j и находится по такой же формуле. ◀

Типичным признаком при изучении изонимии служит концентрация x какой-либо фамилии (точнее, носителей данной фамилии). Как и для признака T , среднее значение $x(s_j)$ для концентрации фамилии в некоторой метапопуляции s_j равно средней величине для концентраций в составляющих ее разбиение субпопуляциях. Скажем, если s_j разбита на $\{s_i\}$, то $x(s_j) = E_{s_i} \{x(s_i)|s_j\}$, а если рассматривается разбиение на $\{s_i\}$, то $x(s_j) = E_{s_i} \{x(s_i)|s_j\}$.

Это свойство средних значений признаков в подразделенных популяциях рассматривается далее с более общих позиций как выражение полного математического ожидания.

СВОЙСТВА СРЕДНИХ ЗНАЧЕНИЙ В ИЕРАРХИЧЕСКОЙ СИСТЕМЕ

Пусть дана какая-либо метапопуляция s с иерархической структурой подразделенности. На первом уровне s разбита на субпопуляции $\{s_1\}$, а на i -м уровне на $\{s_j\}$. Понятно, что каждая из субпопуляций s_i в свою очередь, подразделена на соответствующие подмножества субпопуляций первого уровня. Положим, что рассматривается случайная переменная величина T (ею может быть концентрация носителей какой-нибудь категории признака T). Пусть значения концентрации $\{T(s_1)\}$ заданы для всех субпопуляций первого уровня иерархии $\{s_1\}$, а значения $T(s_j)$ на более высоких уровнях находятся по определению как

$$\begin{aligned} T(s_i) &\equiv E_{s_1} \{T(s_1)|s_i\} = \sum_{s_1} T(s_1) Pr(s_1|s_i), \\ i &= 2, 3, \dots; \quad T(s) \equiv E_{s_1} \{T(s_1)|s\}. \end{aligned} \quad (1)$$

Таким образом, на каждом из уровней иерархии i значение $T(s_i)$ является средним значением для $T(s_1)$ при выборе наугад s_1 из s_i .

Рассмотрим некоторые из свойств математического ожидания для иерархических систем, формулируемые для произвольной случайной величины T . Напомним широко используемую далее формулу полного математического ожидания (частным случаем которой является полученная выше формула для $T(s)$ в предыдущем замечании):

$$E\{T\} = E_A \{E_T \{T|A\}\} = E_i E_T \{T|A_i\} Pr(A_i).$$

Здесь случайное событие A принадлежит множеству событий $\{A_i\}$, составляющих полную систему несовместимых случайных событий, реализующихся с вероятностями $\{Pr(A_i)\}$ и таких, что обязательно происходит одно из них; $E_T \{T|A_i\}$ означает условное (условие пишем после вертикальной черты) математическое ожидание для случайной величины T (нижний индекс у E указывает на переменную, которая служит для усреднения) при условии реализации соответствующего случайного события A_i .

У нас при выборе наугад субпопуляции s_i из s_j роль полной системы случайных событий играет множество субпопуляций $\{s_i\}$, представляющих собой разбиение рассматриваемой популяции s_j . Аналогично при случайном выборе s_1 из s_j полную систему образуют субпопуляции $\{s_1\}$. Тогда для случайной величины T формула полного математического ожидания выглядит следующим образом:

$$\begin{aligned} E\{T(s_j)\} &\equiv E_{s_1} \{T(s_1)|s_j\} = \\ &= E_{s_1} \{E_{s_i} \{T(s_i)|s_j\}\} = E_{s_i} \{T(s_i)|s_j\}, \end{aligned} \quad (2)$$

т.е. среднее значение T у некоторой группы s_j является средним для математических ожиданий $E_{s_1} \{T(s_1)|s_j\} \equiv T(s_i)$ у разбивающих ее подгрупп $\{s_i\}$. Это согласуется с полученными ранее формулами для среднего значения случайной величины T или для концентрации x в подразделенной популяции.

В частности, при $j = i + 1$ получаем рекуррентное уравнение

$$T(s_{i+1}) = E_{s_i} \{T(s_i)|s_{i+1}\}. \quad (3)$$

Если в (2) вместо s_j рассматривать всю метапопуляцию s , разбитую на $\{s_j\}$, по формуле полного математического ожидания и определению (1)

$$\begin{aligned} T(s) &= E_{s_1} \{T(s_1)|s\} = E_{s_1} \{E_{s_i} \{T(s_i)|s\}\} = \\ &= \sum_{s_i} T(s_i) Pr(s_i|s), \quad i = 1, 2, \dots, \end{aligned} \quad (4)$$

т.е. $E_{s_1} \{E_{s_i} \{T(s_i)|s\}\} = E_{s_1} \{T(s_1)|s\}$.

Мы имеем два выражения для одного и того же значения $T(s)$ — по определению (1) и по формуле (4), т.е.

$$T(s) \equiv E_{s_i} \{T(s_i)|s\} = E_{s_i} \{T(s_i)|s_i\}, \quad (5)$$

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

Эта формула показывает в явном виде равенство полного (общего) среднего значения $T(s)$ средней величине для математических ожиданий отдельных частей $\{T(s_i)\}$ одного и того же уровня иерархии i , какими бы ни были эти части и выбранный уровень (i принимает значения, соответствующие отдельным уровням иерархии метапопуляции s). Поскольку разбиение s на $\{s_j\}$ может быть любым, то $T(s)$ не зависит от характера подразделенности s .

Аналогичная формула верна для $T(s_j)$, когда s_j разбита на субпопуляции $\{s_i\}$. Если s состоит из $\{s_j\}$, а s_j из $\{s_i\}$, то

$$E_{s_j} \{E_{s_i} \{T(s_i)|s_j\}|s\} = E_{s_i} \{T(s_i)|s\}. \quad (6)$$

Повторим, что здесь случайная величина T принимает значения $\{T(s_i)\}$ с вероятностями $\{Pr(s_i|s_j)\}$ выбора наугад соответствующей субпопуляции s_i на уровне иерархии i из субпопуляции s_j на уровне j . Среднее значение $E_{s_i} \{T(s_i)|s_j\}$ для $T(s_i|s_j)$ обозначается как $T(s_j)$ и характеризует составную часть s_j метапопуляции s . Это среднее значение само является случайной величиной $T(s_j|s)$, если s_j наугад выбирается из фиксированной метапопуляции $s \equiv \{s_j\}$. Согласно формуле полного математического ожидания полное (общее) среднее значение рассматриваемой случайной величины $T(s_i|s_j)$ для всей метапопуляции $s \equiv \{s_j\}$ равно математическому ожиданию для средних значений в ее отдельных частях $\{s_j\}$, что формально отображается формулами (5)–(6).

Признак T определен для субпопуляций на любом уровне иерархии. При этом его базовыми значениями являются значения $\{T(s_i)\}$, заданные для субпопуляций $\{s_i\}$ первого уровня. Через них выражаются величины T для субпопуляций последующих уровней как средние значения в соответствующих группировках базовых значений. Поэтому обозначение $T(s_i)$ подразумевает вычисление при условии значений $\{T(s_i)\}$, а средние величины T для ряда субпопуляций одного и того же уровня не зависят от выбранного уровня и совпадают с величиной для всей метапопуляции s :

$$E_{s_i} \{T(s_i)|s\} = E_{s_i} \{T(s_i)|s_i\} =$$

$$= \dots E_{s_j} \{T(s_j)|s\} = \dots = T(s),$$

что говорит о независимости значения $T(s)$ от разбиений s , в том числе неиерархических.

ПРИМЕРЫ СЛУЧАЙНЫХ ВЕЛИЧИН ПРИ ИЗУЧЕНИИ ИЗОНИМИИ

Далее в качестве случайных величин будем рассматривать такие признаки субпопуляции как концентрация фамилии x и вероятность Hs неизоимных встреч (столкновений) в субпопуляции. Проанализируем также свойства дисперсии $V_{s_{in}}(x(s_i|s_j))$, т.е. дисперсии стоящего в скобках аргумента (случайной величины $x(s_i|s_j)$).

КОНЦЕНТРАЦИЯ ФАМИЛИИ

Рассмотрим подразделенную субпопуляцию s_2 . Положим, что в качестве T рассматривается концентрация x носителей интересующей фамилии, $x(s_2)$ обозначает концентрацию фамилии во всей фиксированной подразделенной субпопуляции s_2 . Она находится по определению (1) как среднее значение для концентраций $\{x(s_i|s_2)\}$ в субпопуляциях единицах наблюдения первого уровня иерархии $\{s_i\}$ внутри s_2 :

$$x(s_2) \equiv E_{s_i} \{x(s_i)|s_2\} = \sum_{s_i} x(s_i) Pr(s_i|s_2).$$

Среднее значение концентрации фамилии $x(s_3)$ для распределения $\{x(s_i|s_3)\}$ в фиксированной субпопуляции s_3 находится по определению (1) и по формуле полного математического ожидания (4) как

$$x(s_3) \equiv E_{s_i} \{x(s_i|s_3)\} = E_{s_2} \{x(s_2)|s_3\} =$$

$$= E_{s_2} \{E_{s_i} \{x(s_i)|s_2\}|s_3\} = \sum_{s_2} x(s_2) Pr(s_2|s_3).$$

Согласно (4) для произвольного уровня иерархии j (при $j > i$, когда s_i входит в s_j как составная часть, элемент разбиения s_j)

$$x(s_j) \equiv E_{s_i} \{x(s_i)|s_j\} = \sum_{s_i} x(s_i) Pr(s_i|s_j) =$$

$$= \sum_{s_i} x(s_i) Pr(s_i|s_j), \quad s_j = \{s_i\}.$$

Если здесь вместо s_j взять всю метапопуляцию s , то ясно, что концентрация фамилии $x(s)$ не зависит от использования при ее подсчетах уровня разбиения s (как мы видели для произвольной случайной величины T в (5)).

Для иерархически подразделенной метапопуляции s мы имеем при заданных значениях $\{x(s_i)\}$:

$$x(s_i) \equiv E_{s_i} \{x(s_i)|s_i\} =$$

$$= \sum_{s_i} x(s_i) Pr(s_i|s_i), \quad s_i = \{s_i\};$$

$$\begin{aligned}
x(s_j) &\equiv E_{s_1} \{x(s_1)|s_j\} = E_{s_1} \{x(s_i)|s_j\} = \\
&= E_{s_1} \{E_{s_i} \{x(s_1)|s_i\}|s_j\} = \sum_{s_i} x(s_i)Pr(s_i|s_j), \\
s_j &= \{s_i\}, \quad s_i = \{s_1\}; \\
x(s) &\equiv E_{s_1} \{x(s_1)|s\} = \\
&= E_{s_1} \{x(s_i)|s\} = \sum_{s_i} x(s_i)Pr(s_i|s).
\end{aligned}$$

При $i = j - 1$ получаем еще одну запись рекуррентного уравнения $x(s_j) \equiv E_{s_{j-1}} \{x(s_{j-1})|s_j\}$, которую можно исследовать с целью получения выражения для $x(s_j)$ на произвольном уровне иерархии j в зависимости от начальных значений $\{x(s_1)\}$.

КЛАССИФИКАЦИЯ ДИСПЕРСИИ КОНЦЕНТРАЦИИ

Рассмотрим внутригрупповую дисперсию $V_{s_{in}}(s_2) \equiv V_s(x(s_1)|s_2)$ концентрации $x(s_1)$ по субпопуляциям $\{s_1\}$ внутри s_2 , т.е. дисперсию распределения $\{x(s_1)\}$ значений x у субпопуляций $\{s_1\}$, являющихся единицами наблюдения и содержащихся в s_2 . Она является случайной при выборе наугад s_2 из s_3 . Признак $V_{s_{in}}$ определен для субпопуляций с уровнем не ниже второго (говорить о внутригрупповой дисперсии для s_1 не имеет смысла, так как в ней нет субпопуляций, различия между которыми характеризуются дисперсией). Формально можно определить значение $V_s(x(s_n)|s_n)$ равным нулю.

Начальные значения для случайной величины x заданы для субпопуляций первого уровня иерархии как $\{x(s_1)\}$. Эти значения $\{x(s_1)\}$ внутри s_2 определяют начальные значения для $V_s(x(s_1)|s_2) = V_{s_{in}}(x(s_1)|s_2)$, внутригрупповой дисперсии $V_{s_{in}}$ для s_2 . Повторим, что данные значения случайны, если s_2 наугад выбирается из фиксированной субпопуляции s_3 .

Напомним классификацию дисперсий в случае метапопуляции s_3 с тремя уровнями иерархии. Здесь выделяют три типа дисперсий. Под *общей (полной) дисперсией* $V_{s_{tot}}(x(s_1)|s_3)$ метапопуляции понимают дисперсию распределения концентрации $\{x(s_1)\}$ признака по субпопуляциям первого уровня иерархии, содержащимся в s_3 . Рассматриваемая метапопуляция s_3 разбита на группы (субпопуляции) $\{s_2\}$ и дисперсия $V_{s_{betw}}(x(s_2)|s_3)$ распределения концентрации $\{x(s_2)\}$ по этим группам называется *межгрупповой дисперсией* в метапопуляции. Наконец, каждая группа состоит из соответствующих субпопуляций $\{s_1\}$ с концентрациями $\{x(s_1)\}$. Дисперсия $V_s(x(s_1)|s_2)$ распределения концентраций $\{x(s_1)\}$ внутри s_2 называется *внутригрупповой дисперсией* $V_{s_{in}}(x(s_1)|s_2)$ группы s_2 . Метапопуляция s_3 в

целом характеризуется *средней внутригрупповой дисперсией* $E_{s_2} \{V_{s_{in}}(x(s_2)|s_3)\}$.

Ремарка 3. Когда у метапопуляции количество уровней иерархии более трех, приведенная классификация дисперсии становится расплывчатой. Отметим, что дисперсия V_s полностью определяется ее аргументом, а индексы *tot*, *betw*, *in* служат только для облегчения ориентации в отношении аспекта рассмотрения данной дисперсии. Без них свободно можно было бы обойтись без ущерба для логики изложения.

При многоуровневой иерархии метапопуляцию s по-прежнему можно охарактеризовать, скажем, в отношении концентрации x некоторого признака *полной дисперсией* $V_{s_{tot}}(x(s_1)|s)$, являющейся дисперсией распределения признака по субпопуляциям самого низкого, первого, уровня иерархии, содержащимся в s .

Еще одной характеристикой субпопуляции s является *межгрупповая дисперсия* s на уровне i , т.е. дисперсия $V_s(x(s_i)|s)$ распределения концентрации фамилии по группам уровня i , на которые разбита исходная субпопуляция s . Для каждого из промежуточных уровней u s будет своя межгрупповая дисперсия.

Другой характеристикой s служит *внутригрупповая дисперсия* $V_{s_{in}}(s_i|s_j)$ субпопуляции s_j на уровне i , т.е. дисперсия распределения $\{x(s_i|s_j)\}$ значений x у субпопуляций s_i в s_j . Внутригрупповую дисперсию можно интерпретировать когда i равно единице как полную дисперсию s_j , т.е. когда в s_j субпопуляции $\{s_i\}$ являются единицей наблюдения.

Напомним, что внутригрупповая дисперсия носит случайный характер при выборе наугад s_j из s . Тогда вся субпопуляция s характеризуется математическим ожиданием для $\{V_{s_{in}}(s_i|s)\}$, иначе говоря, *средней внутригрупповой дисперсией на уровне i для субпопуляций j -го уровня иерархии в s :*

$$E_{s_j} \{V_{s_{in}}(s_i|s_j)|s\} = \sum_{s_j} V_{s_{in}}(s_i|s_j)Pr(s_j|s).$$

Вспомним, что при анализе концентрации фамилии x ее среднее значение в субпопуляции (метапопуляции) высокого уровня, скажем, в s_k выражается согласно (5) через значения u входящих в нее субпопуляций уровнем ниже как $x(s_k) = E_{s_j} \{x(s_j)|s_k\}$. Казалось бы, если в (5) положить $T(s_j) = V_{s_{in}}(s_i|s_j)$, то аналогично $V_{s_{in}}(s_i|s_k) = E_{s_j} \{V_{s_{in}}(s_i|s_j)|s_k\}$. Однако это равенство неверно, что мы подробно проанализируем далее. Суть в том, что перетасовка данных при группировке субпопуляций $\{s_j\}$ не приводит к изменению результатов для концентрации фамилии как показано ранее, в отличие от дисперсии концентрации.

ВЕРОЯТНОСТИ ВСТРЕЧ (СТОЛКНОВЕНИЙ) В ПОДРАЗДЕЛЕННОЙ МЕТАПОПУЛЯЦИИ

Рассмотрим концентрацию (вероятность) $Hs(s)$ встреч (столкновений) индивидуума с данной фамилией с носителем какой-либо другой в метапопуляции s , определяемую как среднее значение Hs для групп одного уровня иерархии. Согласно свойствам средних значений (5) результат усреднения не зависит от того, какое при этом используется разбиение s на группы на данном уровне и от самого уровня. Поскольку концепция встреч Hs скорее умозрительная, чем дающая легко доступный метод определения Hs в произвольной популяции (за исключением случая брачных пар), то возникает проблема как задать Hs на базовом уровне.

Здесь приходится прибегать к косвенным методам, основанным на дополнительных предположениях. Таким предположением является допущение о случайном характере встреч в группах $\{s_i\}$. Оно разумно, когда разбиение метапопуляции состоит из территориально малых групп (базовых субпопуляций $\{s_i\}$ самого низкого уровня иерархии), в каждой из которых существованием каких-либо препятствий для встреч (столкновений) индивидуумов и нарушениями внешних границ при встречах можно пренебречь.

При случайной встрече двух индивидуумов комбинирование ее участников в пары происходит независимо. В результате независимости участников случайных столкновений вероятность встречи равна произведению вероятностей встретить каждого из участников по отдельности, равных их концентрациям. Так, вероятность $Hs(x(s_i))$ встречи индивидуума с интересующей фамилией (с концентрацией $x(s_i)$ ее носителей в s_i) и индивидуума с какой-либо иной с учетом порядка участников, очевидно, находится как $Hs_0(x(s_i)) \equiv x(s_i)(1 - x(s_i))$. Эта формула аналогична выражению для концентрации гетерозигот по закону Харди–Вайнберга в популяционной генетике при учете порядка аллелей в гетерозиготе, поскольку в обоих случаях наблюдаем случайные объединения в пары.

Метапопуляция, состоящая из неподразделенных субпопуляций, характеризуется средней величиной для значений Hs по субпопуляциям. Например, общая (полная) вероятность $Hs(x(s_1|s_2))$ встречи индивидуумов с разными фамилиями в *подразделенной* метапопуляции s_2 (скажем, в сельсовете s_2 , подразделенном на села) определяется как среднее значение $Hs(x(s_i))$ для входящих в нее субпопуляций (сел $\{s_i\}$), под которой понимается вероятность при выборе наугад субпопуляции s_i из рассматриваемой метапопуляции. Тем самым подразумевается, что *встречи происходят только внутри отдельных сел* по аналогии с рядом моделей популяционной генетики, в которых случай-

ное скрещивание идет внутри отдельной элементарной популяции после формирования ее генетического состава к моменту скрещивания с учетом всех факторов динамики. Чтобы подчеркнуть, что когда речь идет о средней вероятности встреч в метапопуляции, а не в отдельной неподразделенной субпопуляции, то среднее значение вероятности встреч называем *общей (полной) вероятностью* $Hs(x(s_2))$.

Очевидно, при произвольном разбиении метапопуляции s на $\{s_i\}$ у нас нет аргументов в пользу случайного характера встреч в каждой из субпопуляций $\{s_i\}$ и соответственно в пользу вероятности встреч в виде $Hs(s_i) = Hs_0(s_i) \equiv x(s_i)(1 - x(s_i))$. При разбиении метапопуляции s на $\{s_i\}$, когда в $\{s_i\}$ встречи случайны,

$$\begin{aligned} Hs(s) &= \sum_{s_i} Hs_0(s_i) Pr(s_i|s) = \\ &= \sum_{s_i} x(s_i)(1 - x(s_i)) Pr(s_i|s). \end{aligned}$$

Здесь и далее для краткости мы пишем $Hs(s)$ вместо $Hs(x(s))$.

Повторим, что роль вероятности $Hs(s_i) = Hs(x(s_i))$ случайной встречи рассматриваемых индивидуумов с учетом порядка их фамилий в неподразделенной субпопуляции s_i с концентрацией интересующей фамилии $x(s_i)$ играет $x(s_i)(1 - x(s_i))$, а вероятность встречи $Hs(s_2)$ в метапопуляции $s_2 = \{s_i\}$ определяется как среднее значение $Hs(s_i)$ при выборе наугад s_i из метапопуляции s_2 . Таким образом, полной вероятностью встречи $Hs(s_2)$ в метапопуляции s_2 будет

$$Hs(s_2) \equiv E_{s_i} \{Hs(s_i)|s_2\} = E_{s_i} \{x(s_i)(1 - x(s_i))|s_2\}.$$

Напомним, что $Pr(s_i|s_2)$ обозначает вероятность случайного выбора села s_i из фиксированного сельсовета s_2 ; E_{s_i} — символ операции получения математического ожидания (среднего значения) случайной величины, стоящей в фигурных скобках; нижний индекс s_i у E_{s_i} указывает на служащую для усреднения переменную s_i . По определению вероятность $Hs(s_i)$ рассматриваемой встречи для произвольного уровня иерархии i задается формулой

$$\begin{aligned} Hs(s_i) &\equiv E_{s_i} \{Hs(s_i)|s_i\} = \sum_{s_i} Hs(s_i) Pr(s_i|s_i) = \\ &= \sum_{s_i} x(s_i)(1 - x(s_i)) Pr(s_i|s_i). \end{aligned}$$

На примере вероятности (концентрации) Hs мы видим, что в качестве характеристики субпопуляций может служить не только концентрация фамилии x , но и функция Hs от x . В каждой субпопуляции s_i значение Hs находится как функция $x(1 - x)$ от концентрации x в s_i . Казалось бы есте-

ственно предположить, что среднее значение Hs по субпопуляциям s_i , составляющим s_j , т.е. $Hs(x(s_i))$, находится как такая же функция $x(1-x)$ теперь от значения x в s_j . Однако $Hs(x(s_i))$ не равно $x(s_j)(1-x(s_j))$, что на первый взгляд может показаться парадоксальным.

ВЕРОЯТНОСТЬ ВСТРЕЧ В ПОДРАЗДЕЛЕННОЙ МЕТАПОПУЛЯЦИИ И КОЭФФИЦИЕНТ ФАМИЛЬНОГО ИНБРИДИНГА

Обратимся к анализу вероятности (доли, концентрации) встреч Hs , являющейся аналогом концентрации гетерозигот в популяционной генетике. Как уже говорилось, эту концентрацию можно уподобить вероятности соответствующего столкновения в некоторой совокупности частиц. Напомним, что здесь подразумевается отыскание $Hs(s_i)$ в подразделенной популяции при условии заданных значений вероятности столкновений в субпопуляциях первого уровня.

На первом уровне иерархии вероятность встречи $Hs(s_1)$ в субпопуляции s_1 двух индивидуумов (с заданной фамилией с концентрацией $x(s_1)$ и какой-либо другой с учетом их порядка) находится в предположении случайных столкновений как $Hs(s_1) \equiv Hs_0(s_1) = x(s_1)(1-x(s_1))$ аналогично закону Харди–Вайнберга. Для субпопуляций более высокого уровня i используется обозначение $Hs(s_i)$, в котором зависимость от значений $Hs(s_1)$ неясна. Концентрация $Hs(s_i)$ находится как среднее значение для $\{Hs(s_1)\}$ в субпопуляциях $\{s_1\}$, составляющих разбиение s_i . Когда $Hs(s_1) = x(s_1)(1-x(s_1))(1-F(s_1))$, для вероятности $Hs(s_i)$ используется обозначение $Hs(s_i|Fs_1)$, явным образом указывающее на вид $Hs(s_1)$.

По определению значение Hs для каждой субпопуляции уровня иерархии i выше единицы равно средней величине для значений Hs у всех входящих в s_i субпопуляций на первом уровне (или на одном из других предшествующих i уровней) согласно (5), а на первом уровне в предположении случайного характера встреч равно $x(s_1)(1-x(s_1))$:

$$\begin{aligned} Hs(s_1) &\equiv Hs_0(s_1) \equiv x(s_1)(1-x(s_1)), \\ Hs(s_i) &\equiv E_{s_1} \{Hs_0(s_1)|s_i\} = \\ &= E_{s_1} \{x(s_1)(1-x(s_1))|s_i\} = \\ &= \sum_{s_1} x(s_1)(1-x(s_1)) Pr(s_1|s_i), \quad i = 2, 3, \dots \end{aligned} \quad (7)$$

Иначе говоря, полная вероятность случайной встречи (концентрация $Hs(s_i)$) в s_i находится как среднее значение $Hs(s_1)$ при выборе наугад субпопуляции s_1 из s_i . По формуле полного математического ожидания (5)

$$Hs(s_j) \equiv E_{s_2} \{Hs(s_2)|s_j\} = E_{s_i} \{Hs(s_i)|s_j\}, \quad j > i > 1.$$

В частности, значения Hs на соседних уровнях иерархии связывает рекуррентное соотношение

$$\begin{aligned} Hs(s_{i+1}) &\equiv E_{s_i} \{Hs(s_i)|s_{i+1}\}, \text{ или} \\ Hs(s_i) &\equiv E_{s_{i-1}} \{Hs(s_{i-1})|s_i\}; \\ Hs(s_1) &\equiv x(s_1)(1-x(s_1)). \end{aligned} \quad (8)$$

Отсюда вытекает согласующаяся с (6) запись формулы полного математического ожидания для Hs

$$E_{s_i} \{Hs(s_i)|s_{i+1}\} = E_{s_i} \{E_{s_{i-1}} \{Hs(s_{i-1})|s_i\}|s_{i+1}\}. \quad (9)$$

Повторим, что (8) означает, что полная вероятность рассматриваемых встреч в субпопуляции s_i является средней величиной для вероятностей встреч в субпопуляциях $\{s_{i-1}\}$, вместе составляющих разбиение s_i . Выясним зависимость вероятности подобной встречи от фамильной дивергенции между субпопуляциями, рассмотренной нами ранее в первой части [4] в терминах дисперсии концентрации рассматриваемой фамилии.

Результат 4 (фамильный аналог эффекта Валунда). Пусть дана метапопуляция s , разбитая на субпопуляции $\{s_1\}$ с концентрациями $\{x(s_1)\}$ рассматриваемой фамилии в них и с вероятностями рассматриваемых неизонимных встреч $\{x(s_1)(1-x(s_1))\}$.

Тогда общая (полная) вероятность $Hs(s)$ случайной встречи двух индивидуумов (с фиксированной фамилией и с какой-либо иной при учете порядка фамилий) в наугад выбранной из s субпопуляции s_1 имеет вид

$$\begin{aligned} Hs(s) &\equiv E_{s_1} \{Hs(s_1)|s\} = \\ &= \sum_{s_1} x(s_1)(1-x(s_1)) Pr(s_1|s) = \\ &= x(s)(1-x(s)) - V_{sbetw}(x(s_1|s)), \\ V_{sbetw}(x(s_1|s)) &\equiv V_s(x(s_1|s)) = \\ &= \sum_{s_1} (x(s_1) - x(s))^2 Pr(s_1|s). \end{aligned} \quad (10)$$

Здесь $V_{sbetw}(x(s_1|s))$ обозначает межгрупповую дисперсию концентраций $\{x(s_1|s)\}$ рассматриваемой фамилии в субпопуляциях (группах) $\{s_1\}$, составляющих разбиение метапопуляции s ; $Pr(s_1|s)$ дает вероятность случайного выбора субпопуляции s_1 из метапопуляции s .

Иным образом общую (полную) вероятность $Hs(s)$ можно представить как

$$\begin{aligned} Hs(s) &= x(s)(1-x(s)) \left(1 - \frac{V_{sbetw}(x(s_1|s))}{x(s)(1-x(s))} \right) = \\ &= x(s)(1-x(s))(1-Fs(s)), \\ Fs(s) &\equiv \frac{V_{sbetw}(x(s_1|s))}{x(s)(1-x(s))} = \frac{V_{sbetw}(x(s_1|s))}{Hs_0(s)} \geq 0, \end{aligned} \quad (11)$$

где $Fs(s)$ является фамильным аналогом коэффициента инбридинга S . Райта в популяционной генетике.

Доказательство. Пусть неподразделенная субпопуляция s_1 рассматривается как единица наблюдения. Например, в сельском населении это может быть село, а метапопуляция s соответствует, скажем, сельсовету. Дисперсия $V_{s_{betw}}(x(s_1|s))$ относится к распределению концентрации заданной фамилии по селам $\{s_1\}$ внутри фиксированного сельсовета s . Данная дисперсия $V_{s_{betw}}(x(s_1|s))$ характеризует фамильную дивергенцию сел внутри сельсовета.

Напомним, что для любой случайной величины x выполняется равенство $E\{x^2\} = (E\{x\})^2 + V(x)$, так как $V(x) \equiv E\{x^2\} - (E\{x\})^2$. Для дальнейшего эту формулу с учетом равенства $E_{s_i}\{x(s_i)|s_{i+1}\} = x(s_{i+1})$ удобно представить как

$$E_{s_i}\{x^2(s_i)|s_{i+1}\} = (E_{s_i}\{x(s_i)|s_{i+1}\})^2 + V_{s_i}(x(s_i)|s_{i+1})) = x^2(s_{i+1}) + V_{s_i}(x(s_i)|s_{i+1})). \quad (12)$$

Как уже говорилось, для субпопуляции s_1 первого уровня иерархии вероятность случайной встречи рассматриваемой пары индивидумов с учетом их порядка равна $x(s_1)(1 - x(s_1))$ подобно закону Харди–Вайнберга в популяционной генетике. Значение $Hs(s)$ для s по определению находится как среднее значение $Hs(s_1)$ при случайном выборе s_1 из s , т.е.

$$\begin{aligned} Hs(s) &= E_{s_1}\{Hs(s_1)|s\} = E_{s_1}\{x(s_1)(1 - x(s_1))|s\} = \\ &= E_{s_1}\{x(s_1)|s\} - E_{s_1}\{x^2(s_1)|s\} = \\ &= x(s) - (V_{s_1}(x(s_1)|s) + (E_{s_1}\{x(s_1)|s\})^2) = \\ &= x(s)(1 - x(s)) - V_{s_1}(x(s_1)|s) \end{aligned}$$

согласно (12). Очевидно, данное выражение $Hs(s)$ можно переписать с использованием коэффициента фамильного инбридинга как $x(s)(1 - x(s))(1 - Fs(s))$, т.е. формула (11) верна. ◀

Согласно доказанному результату в метапопуляции s , разбитой на несколько неподразделенных субпопуляций $\{s_1\}$, в среднем вероятность указанной встречи в субпопуляции s_1 из s меньше на величину межгрупповой дисперсии $V_{s_1}(x(s_1|s))$, чем вероятность независимых сочетаний фамилий двух индивидумов во всей метапопуляции s . Эта формула является практически полным аналогом эффекта Валунда в популяционной генетике (см., например, [1]). Чем больше фамильная дивергенция субпопуляций $\{s_1\}$, тем меньше общая вероятность $Hs(s)$.

Изложенное показывает важность дисперсии распределения концентрации рассматриваемой фамилии по субпопуляциям для анализа их дивергенции. Это вызывает интерес к изучению

компонентов дисперсии концентрации. К данному случаю приложимо правило сложения дисперсий, которое в нашей ситуации выглядит следующим образом.

Замечание 5 (правило сложения дисперсий). Пусть метапопуляция s разбита на субпопуляции $\{s_2\}$, каждая из которых состоит из соответствующих неподразделенных групп $\{s_1\}$, являющихся единицами наблюдения с концентрациями $\{x(s_1)\}$ рассматриваемой фамилии в них.

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

$$V_{s_{tot}}(x(s_1|s)) = V_{s_{betw}}(x(s_2|s)) + E_{s_2}\{V_{s_{in}}(x(s_1|s_2))|s\}. \quad (13)$$

Доказательство. По определению полная (общая) дисперсия $V_{s_{tot}}(x(s_1|s))$ популяции s равна $E_{s_1}\{(x(s_1) - E_{s_1}\{x(s_1)|s\})^2\}$. Вспомним, что $E_{s_2}\{x(s_2|s)\} = x(s)$, $E_{s_1}\{x(s_1|s_2)\} = x(s_2)$, т.е. $V_{s_{tot}}(x(s_1|s)) = E_{s_1}\{(x(s_1) - x(s))^2\}$. Представим $(x(s_1) - x(s))^2$ в виде

$$\begin{aligned} (x(s_1) - x(s_2) + x(s_2) - x(s))^2 &= \\ &= (x(s_1) - x(s_2))^2 + (x(s_2) - x(s))^2 + \\ &+ 2(x(s_1) - x(s_2))(x(s_2) - x(s)). \end{aligned}$$

Математическое ожидание правой части равно сумме средних значений ее отдельных слагаемых, которые находим с применением формулы полного математического ожидания (2). При этом учтем, что за знак E можно выносить независимые от усредняемой величины сомножители.

$$\begin{aligned} E_{s_1}\{(x(s_1) - x(s_2))^2|s\} &= \\ &= E_{s_2}\{E_{s_1}\{(x(s_1) - x(s_2))^2|s_2\}|s\} = \\ &= E_{s_2}\{V_{s_{in}}(x(s_1|s_2))|s\}, \end{aligned}$$

т.е. первое слагаемое равно средней внутригрупповой дисперсии.

Для второго слагаемого находим

$$\begin{aligned} E_{s_1}\{(x(s_2) - x(s))^2|s\} &= \\ &= E_{s_2}\{E_{s_1}\{(x(s_2) - x(s))^2|s_2\}|s\} = \\ &= E_{s_2}\{(x(s_2) - x(s))^2 E_{s_1}\{1|s_2\}|s\} = V_{s_{betw}}(x(s_2|s)), \end{aligned}$$

т.е. получаем межгрупповую дисперсию.

Наконец, покажем, что среднее значение третьего слагаемого равно нулю:

$$E_{s_1} \{2(x(s_1) - x(s_2))(x(s_2) - x(s)) | s\} = \\ = E_{s_2} \{2(x(s_2) - x(s)) E_{s_1} \{(x(s_1) - x(s_2)) | s_2\} | s\} = 0,$$

так как среднее отклонение от средней величины $E_{s_1} \{(x(s_1) - x(s_2)) | s_2\}$ всегда равняется нулю.

В итоге получаем требуемое выражение для полной дисперсии

$$V_{s_{tot}}(x(s_1 | s)) = V_{s_{berw}}(x(s_2 | s)) + \\ + E_{s_2} \{V_{s_{in}}(x(s_1 | s_2)) | s\}. \blacktriangleleft$$

Данное правило приложимо к любой совокупности, разбитой на группы ее элементов, характеризующихся числовым признаком. В частном случае метапопуляции с произвольным количеством уровней иерархии общая дисперсия концентрации фамилии в ней равна сумме межгрупповой дисперсии распределения концентрации по субпопуляциям (группам) одного и того же произвольно выбранного уровня k плюс средняя внутригрупповая дисперсия концентраций в них. Скажем, в разбитой на субпопуляции $\{s_1\}$ метапопуляции $s = s_m$ с m уровнями иерархии правило сложения дисперсий выполняется, когда вместо субпопуляций $\{s_2\}$ рассматриваются $\{s_k\}$:

$$V_{s_{tot}}(x(s_1 | s_m)) = V_{s_{berw}}(x(s_k | s_m)) + \\ + E_{s_k} \{V_{s_{in}}(x(s_1 | s_k)) | s_m\}, \quad 2 < k < m. \quad (14)$$

Выше мы считали, что встречи индивидуумов в субпопуляциях $\{s_1\}$ низшего уровня иерархии (единицах наблюдения) случайны. Благодаря этому $Hs(s_1) = Hs_0(s_1) = x(s_1)(1 - x(s_1))$. В реальных исследованиях данное условие может не выполняться (скажем, при существовании скрытой подразделенности в s_1 или из-за иных причин), и желательно рассмотреть такую ситуацию, поскольку для нее приведенный аналог результата Валунда неверен.

В абстрактном случае, скажем, в совокупности сталкивающихся частиц нескольких типов предположение о случайной природе их столкновений не гарантировано. Распространенным выражением для вероятности Hs неслучайных столкновений частиц рассматриваемого типа с концентрацией x с иными типами является используемая в популяционной генетике вероятность столкновений вида

$$Hs = x(1 - x)(1 - Fs) = Hs_0(1 - Fs), \\ 0 \leq Hs \leq 1, \quad 0 \leq |Fs| \leq 1.$$

При $Fs > 0$ происходит своего рода отталкивание, и разноименные частицы сталкиваются реже, чем чисто случайно. Обратим внимание, что здесь неявно предполагается независимость Fs от типа сталкивающихся частиц (в противном случае рассматриваемая вероятность столкновения

зависит от концентраций всех типов). В случае только двух типов (рассматриваемого и искусственного объединения всех прочих типов) коэффициент Fs имеет смысл корреляции между сталкивающимися типами.

Следствие 6. Рассмотрим метапопуляцию s_2 , разбитую на субпопуляции $\{s_1\}$ с концентрациями $\{x(s_1)\}$ носителей рассматриваемой фамилии в них. Пусть вероятность $Hs(s_1)$ их встреч в s_1 с носителями иной фамилии неслучайна и формулируется как

$$Hs(s_1) = x(s_1)(1 - x(s_1))(1 - Fs(s_1)) = \\ = Hs_0(s_1)(1 - Fs(s_1)), \quad s_2 = \{s_1\},$$

где $Hs_0(s_1)$ обозначает вероятность $x(s_1)(1 - x(s_1))$ чисто случайной встречи в субпопуляции s_1 на первом уровне иерархии.

Тогда общая (полная) вероятность $Hs(s_2 | Fs_1)$ рассматриваемой встречи в наугад выбранной из s_2 субпопуляции s_1 при независимости концентрации x рассматриваемой фамилии и коэффициента фамильного инбридинга $Fs(s_1)$ в субпопуляциях $\{s_1\}$ имеет вид

$$Hs(s_2 | Fs_1) = (x(s_2)(1 - x(s_2)) - \\ - V_{s_{berw}}(x(s_1 | s_2)))(1 - E_{s_1} \{Fs(s_1) | s_2\}) \\ V_{s_{berw}}(x(s_1 | s_2)) = V_s(x(s_1 | s_2)) = \\ = \sum_{s_1} (x(s_1) - x(s_2))^2 Pr(s_1 | s_2). \quad (15)$$

Здесь $V_{s_{berw}}(x(s_1 | s_2))$ обозначает межгрупповую дисперсию распределения концентрации $\{x(s_1 | s_2)\}$ рассматриваемой фамилии по субпопуляциям (группам) $\{s_1\}$, составляющим разбиение метапопуляции s_2 ; $Pr(s_1 | s_2)$ дает вероятность случайного выбора субпопуляции s_1 из метапопуляции s_2 .

Иным образом общую (полную) вероятность $Hs(s_2 | Fs_1)$ можно представить как

$$Hs(s_2 | Fs_1) = Hs_0(s_2) \left(1 - \frac{V_{s_{berw}}(x(s_1 | s_2))}{x(s_2)(1 - x(s_2))}\right) \times \\ \times (1 - E_{s_1} \{Fs(s_1) | s_2\}) = Hs_0(s_2)(1 - Fs(s_2)) \times \\ \times (1 - E_{s_1} \{Fs(s_1) | s_2\}), \\ Fs(s_2) \equiv \frac{V_{s_{berw}}(x(s_1 | s_2))}{x(s_2)(1 - x(s_2))} \geq 0, \quad (16)$$

где $Fs(s_2)$ является фамильным аналогом случайного коэффициента инбридинга F_{ST} , а средний коэффициент фамильного инбридинга $E_{s_1} \{Fs(s_1) | s_2\}$ для субпопуляций $\{s_1\}$ в s_2 служит аналогом коэффициента F_{IS} неслучайного инбридинга С. Райта в популяционной генетике.

Доказательство. Общая вероятность $Hs(s_2)$ рассматриваемых встреч в метапопуляции s_2 по определению находится как математическое ожидание

для значений $Hs(s_1)$ в субпопуляциях $\{s_1\}$, составляющих разбиение s_2 :

$$Hs(s_2) \equiv E_{s_1} \{Hs(s_1)|s_2\} = \\ = E_{s_1} \{x(s_1)(1-x(s_1))(1-Fs(s_1))|s_2\}.$$

При независимости x и Fs данное выражение равно произведению математических ожиданий сомножителей и

$$Hs(s_2) = E_{s_1} \{x(s_1)(1-x(s_1))|s_2\} \times \\ \times E_{s_1} \{1-Fs(s_1)|s_2\} = x(s_2)(1-x(s_2)) \times \\ \times (1-Fs(s_2))(1-E_{s_1} \{Fs(s_1)|s_2\}).$$

Здесь коэффициент фамильного инбридинга $Fs(s_2)$ характеризует подразделенность метапопуляции s_2 (фамильную дивергенцию субпопуляций s_1 внутри s_2) и по определению равен отношению дисперсии распределения концентрации фамилии по субпопуляциям $\{s_1\}$ к вероятности случайных встреч $Hs_0(s_2) \equiv x(s_2)(1-x(s_2))$ в s_2 :

$$Fs(s_2) \equiv \frac{Vs_{betw}(x(s_1)|s_2))}{x(s_2)(1-x(s_2))} = \frac{Vs_{betw}(x(s_1)|s_2))}{Hs_0(s_2)}, \\ Hs_0(s_2) \equiv x(s_2)(1-x(s_2)).$$

Он соответствует коэффициенту случайного инбридинга F_{ST} в популяционной генетике, отражающему генетическую дивергенцию между субпопуляциями.

Коэффициент $E_{s_1} \{Fs(s_1)|s_2\}$ является аналогом неслучайного коэффициента инбридинга F_{IS} в генетике популяций и отражает усредненное нарушение случайного характера встреч (закона Харди–Вайнберга) внутри субпопуляций, служащих единицей наблюдения. ◀

Повторим, что здесь неявно предполагается независимость $Fs(s_1)$ и концентрации x рассматриваемой фамилии. Коэффициенты $\{Fs(s_1)\}$ для субпопуляций первого уровня заданы изначально и по ним находятся средние коэффициенты неслучайного фамильного инбридинга для субпопуляций более высоких уровней. Для единообразия записи вероятности встреч метапопуляции s_2 можно использовать коэффициент общего инбридинга F_{IT} в s_2 . В популяционной генетике он учитывает одновременно влияние как дивергенции между субпопуляциями, так и нарушения закона Харди–Вайнберга внутри них на гетерозиготность H метапопуляции s_2 (и на корреляцию между гаминами генотипа):

$$H(s_2) = x(s_2)(1-x(s_2))(1-F_{IT}), \\ 1-F_{IT} \equiv (1-F_{ST})(1-F_{IS}).$$

РАЗЛОЖЕНИЕ ПО УРОВНЯМ ИЕРАРХИИ ВЕРОЯТНОСТИ ВСТРЕЧ ПРИ ИЕРАРХИЧЕСКОЙ ПОДРАЗДЕЛЕННОСТИ МЕТАПОПУЛЯЦИИ

Рассмотренный случай подразделенности метапопуляции на неподразделенные группы первого уровня можно считать простейшим случаем иерархической структуры. Теперь обратимся к подразделенной метапопуляции s_m с произвольным количеством уровней иерархии m . Если $s = s_m$, то дисперсию распределения концентрации по субпопуляциям $\{s_k\}$ на уровне k в метапопуляции s_m обозначим как $Vs_{betw}(x(s_k)|s_m))$, $k < m$. В частности, $Vs(x(s_1)|s_m)) = Vs_{betw}(x(s_1)|s_m))$, причем $Vs_{betw}(x(s_1)|s_m))$ одновременно является общей дисперсией $Vs_{tot}(x(s_1)|s_m))$ для распределения концентрации фамилии по всем субпопуляциям, служащим единицей наблюдения.

Проанализируем выражение для вероятности встреч Hs в метапопуляции в зависимости от количества ее уровней иерархии m . Дадим решение этой задачи при независимости x и Fs путем последовательного увеличения m и индукции.

В предшествующем следствии мы нашли, что при фиксированной метапопуляции $s = s_2$

$$Hs(s_2) \equiv E_{s_1} \{Hs(s_1)|s_2\} = \\ = (x(s_2)(1-x(s_2)) - Vs(x(s_1)|s_2))(1-E_{s_1} \{Fs(s_1)|s_2\}).$$

Теперь зафиксируем s_3 , подставим в рекуррентное уравнение (8), определяющее $Hs(s_3)$, найденное значение $Hs(s_2)$, учтем эффект Валунда (11) и заменим математическое ожидание произведения произведением математических ожиданий сомножителей в силу предполагаемой независимости x и Fs :

$$Hs(s_3|Fs_1) \equiv E_{s_2} \{Hs(s_2)|s_3\} = \\ = E_{s_2} \{(x(s_2)(1-x(s_2)) - Vs(x(s_1)|s_2))|s_3\} \\ (1-E_{s_2} \{E_{s_1} \{Fs(s_1)|s_2\}|s_3\}) = \\ = (x(s_3)(1-x(s_3)) - Vs(x(s_2)|s_3)) - \\ - E_{s_2} \{Vs(x(s_1)|s_2)|s_3\}(1-E_{s_1} \{Fs(s_1)|s_3\}).$$

Ремарка 7. Заметим, что $Vs(x(s_2)|s_3))$ можно записать как $E_{s_3} \{Vs(x(s_2)|s_3)|s_3\}$, поскольку метапопуляция s_3 единственна. Поэтому

$$E_{s_2} \{Hs(x(s_2)|s_3)\} = \left(x(s_3)(1-x(s_3)) - \sum_{i=1}^2 E_{s_{i+1}} \{Vs(x(s_i)|s_{i+1})|s_3\} \right) (1-E_{s_1} \{Fs(x(s_1)|s_3)\}).$$

Найденное выражение для $Hs(s_3)$ и ремарка подсказывают, что для субпопуляций j -го уровня

$$Hs(s_j) = \left(x(s_j)(1 - x(s_j)) - \sum_{i=1}^{j-1} E_{s_{i+1}} \{Vs(x(s_i|s_{i+1}))|s_j\} \right) (1 - E_{s_1} \{Fs(s_1)|s_j\}).$$

С другой стороны, поскольку $Hs(s_j) = (x(s_j)(1 - x(s_j)) - Vs(x(s_1|s_j)))(1 - E_{s_1} \{Fs(s_1)|s_j\})$ согласно фамильному аналогу эффекта Валунда (11), где роль метапопуляции s играет s_j , то одновременно мы видим разложение дисперсии $Vs(x(s_1|s_j))$ на сумму членов $E_{s_{i+1}} \{Vs(x(s_i|s_{i+1}))|s_j\}$.

Результат 8. Рассмотрим иерархически подразделенную метапопуляцию s_m с уровнями иерархии $i = 1, 2, \dots, m$. Положим, что в ее субпопуляциях $\{s_i\}$ концентрации интересующей фамилии равны $\{x(s_i)\}$ и встречи индивидуумов происходят случайно с вероятностями $\{Hs(s_i) = Hs_0(s_i) \equiv x(s_i)(1 - x(s_i))\}$.

Пусть из s_m наугад выбрана субпопуляция s_1 первого уровня. Тогда вероятность $Hs(s_m)$ встречи в s_1 пары индивидуумов с интересующей фамилией и с какой-нибудь иной с учетом их порядка (полная вероятность встречи) допускает следующее разложение:

$$\begin{aligned} Hs(s_m) &\equiv E_{s_1} \{Hs(s_1)|s_m\} = \\ &= x(s_m)(1 - x(s_m)) - Vs_{betw}(x(s_1|s_m)) = \\ &= Hs_0(s_m) - \sum_{i=1}^{m-1} E_{s_{i+1}} \{Vs_{in}(x(s_i|s_{i+1}))|s_m\}. \end{aligned} \quad (17)$$

Таким образом, общая (полная) вероятность $Hs(s_m)$ равна разности между значением вероятности $Hs_0(s_m) \equiv x(s_m)(1 - x(s_m))$ случайной встречи такой пары во всей метапопуляции s_m и разложением межгрупповой дисперсии $Vs_{betw}(x(s_1|s_m))$ на средние внутригрупповые дисперсии $E_{s_{i+1}} \{Vs_{in}(x(s_i|s_{i+1}))|s_m\}$, соответствующие отдельным уровням иерархии i .

Доказательство. Воспользуемся индукцией. Выше мы видели, что разложение (17) для $Hs(s_m)$ верно при некотором m . Покажем, что тогда оно справедливо для $Hs(s_{m+1})$. Согласно рекуррентному уравнению (8)

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

Рассмотрим отдельное слагаемое в сумме по i . Согласно (6), где роль i играет $i + 1$, а $T(s_{i+1}) = Vs_{in}(x(s_i|s_{i+1}))$, имеем

$$\begin{aligned} E_{s_m} \{E_{s_{i+1}} \{Vs_{in}(x(s_i|s_{i+1}))|s_m\}|s_{m+1}\} &= \\ &= E_{s_{i+1}} \{Vs_{in}(x(s_i|s_{i+1}))|s_{m+1}\}. \end{aligned}$$

Поскольку в соответствии со сделанной ремаркой $Vs_{in}(x(s_m|s_{m+1})) = E_{s_{m+1}} \{Vs_{in}(x(s_m|s_{m+1}))|s_{m+1}\}$, этот член можно включить в рассматриваемую сумму, верхний предел которой увеличится на единицу и станет равен $(m - 1) + 1 = m$:

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

Полученное выражение совпадает с (17), если вместо m поставить $m + 1$. ◀

Из (17) отчетливо видно, что вероятность Hs для уровня выше первого меньше значения, соответствующего случайным встречам, и чем больше сумма средних внутригрупповых дисперсий, тем больше увеличивается указанный разрыв. Полученное разложение является некоторого рода обобщением эффекта Валунда [10]. Заметим, что здесь не требуется независимость концентрации фамилии и коэффициента фамильного инбридинга.

Следствие 9. В иерархически подразделенной метапопуляции s_m с m уровнями иерархии полная дисперсия концентрации фамилии $Vs_{tot}(x(s_1|s_m))$ допускает разложение

$$\begin{aligned} Vs_{tot}(x(s_1|s_m)) &= \sum_{i=1}^{m-1} E_{s_{i+1}} \{Vs(x(s_i|s_{i+1}))|s_m\} = \\ &= \sum_{i=1}^{m-1} (Vs(x(s_i|s_m)) - Vs(x(s_{i+1}|s_m))). \end{aligned} \quad (18)$$

Доказательство первого равенства содержится в предыдущем результате, т.е. там попутно получено разложение дисперсии, найденное нами ранее в [4] более длинным способом.

Второе равенство вытекает из правила сложения дисперсий (14), согласно которому $E_{s_k} \{Vs_{in}(x(s_n|s_k))|s_m\} = Vs_{tot}(x(s_n|s_m)) - Vs_{betw}(x(s_k|s_m)) \geq 0$. В нашем случае при $n = i$, $k = i + 1$

$$\begin{aligned} E_{s_{i+1}} \{Vs_{in}(x(s_i|s_{i+1}))|s_m\} &= Vs(x(s_i|s_m)) - \\ &- Vs(x(s_{i+1}|s_m)), \quad n < k < m. \end{aligned}$$

Таким образом, полная дисперсия $Vs(x(s_1|s_m))$ равна сумме приращений межгрупповой дисперсии при переходе от уровня $i + 1$ к предыдущему i .

Подобно тому, как пройденный путь складывается из приращений по шагам, так и при суммировании приращений дисперсии по i получим полную дисперсию $Vs(x(s_i|s_m))$, поскольку все члены в сумме сокращаются, кроме $Vs(x(s_i|s_m))$, а $Vs(x(s_m|s_m))$ разумно определить как нуль. ◀

Так как левая часть $E_{s_{i+1}}\{Vs(x(s_i|s_{i+1}))|s_m\}$ последней формулы неотрицательна (как среднее значение неотрицательной дисперсии), то при переходе к более высокому уровню $i + 1$ межгрупповая дисперсия (дисперсия концентрации фамилии в субпопуляциях одного и того же уровня) не увеличивается. Это интуитивно ожидается, поскольку при группировке данных разброс значений не возрастает, и дисперсия средних значений для сгруппированных данных не может быть больше дисперсии для исходных данных.

ЗАКЛЮЧЕНИЕ

Пройдемся по основным линиям данной работы — ее мотивации, подходу к анализу, результатам. Интерес к семейной структуре мотивирован вниманием к генеалогии, родственным связям и др. аспектам, так или иначе связанным с генетикой населения. Настоящая статья ориентирована на популяционно-генетическую сторону изучения распределения фамилий, и эта ориентация приводит к использованию и обобщениям традиционных методов анализа генетической структуры по данным о распределении концентраций маркеров в ряде популяций.

Выбор такого подхода обусловлен тем, что популярные модели процессов выборочного генного дрейфа Райта—Фишера и дрейфа фамилий в популяциях ограниченной численности при стандартных предположениях совпадают (отличаются лишь интенсивностью, см. [2, 3]). Следовательно, это должно позволить при соответствующих поправках экстраполировать результаты одного из них на ожидаемые результаты для другого. Природа процессов дрейфа с динамикой генетического и семейного состава популяции обусловлена случайными колебаниями количества потомков у отдельных индивидуумов.

Если у некоторого родителя-мужчины больше сыновей, чем в среднем, то в следующем поколении будет больше носителей его фамилии и его генов. На популяционном уровне это приводит к случайным колебаниям семейного и генетического составов популяции от поколения к поколению, представляющим сущность процессов дрейфа. Оба процесса приводят к дивергенции родственных популяций (генетической и семейной). Очевидно, требуется развивать методы, характеризующие дивергенцию популяций (неоднородность системы популяций в целом) для типичных для человека популяционных структур.

Данные по фамилиям часто доступны в иерархически структурированном виде, например в соответствии с административно-территориальной подразделенностью. В качестве показателей дивергенции служат межгрупповая дисперсия распределения концентрации (кодминантного аллеля, фамилии) в группах (субпопуляциях), нарушения закона Харди—Вайнберга с эффектом Валунда, коэффициент случайного инбридинга. Рассмотрены семейные аналоги этих показателей и даны их обобщения для иерархически подразделенной метапопуляции с произвольным количеством уровней иерархии. Данные обобщения вычислительно универсальны в том смысле, что не зависят от характера разбиений метапопуляции на каждом из уровней, от типа микроэволюционного процесса, приведшего к текущему состоянию метапопуляции, от существования миграций и пр. Неоднородность (дивергенция) в иерархических системах характеризуется алгоритмически одинаковыми показателями, различающимися между системами количественно.

В настоящей статье в качестве основного показателя семейной дивергенции субпопуляций рассматривается общая вероятность встреч (столкновений) Hs лиц с разными фамилиями (аналог концентрации гетерозигот), понимаемая как вероятность встречи в выбранной наугад субпопуляции, единицы наблюдения, из рассматриваемой метапопуляции с иерархической структурой субпопуляций. Получено разложение Hs по уровням иерархии, обобщающее эффект Валунда в популяционной генетике. Общая вероятность неизонимных встреч в иерархически подразделенной метапопуляции меньше вероятности случайных встреч в ней на сумму средних внутригрупповых дисперсий концентрации фамилии, соответствующих отдельным уровням.

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

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

СПИСОК ЛИТЕРАТУРЫ

1. Ли Ч. Введение в популяционную генетику. М.: Мир, 1978. 555 с.
2. 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>
3. 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>
4. Пасеков В.П. Описание дивергенции субпопуляций в иерархической системе при анализе изонимии

- мии. I. Дисперсия как показатель дивергенции // Генетика. 2022. Т. 58. № 6. С. 713–727.
5. *Wright S.* The interpretation of population structure by *F* statistics with special regard to systems of mating // *Evolution*. 1965. V. 19. P. 395–420.
 6. 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.
 7. Ревазов А.А., Парадеева Г.М., Русакова Г.И. Пригодность русских фамилий в качестве квазигенетического маркера // *Генетика*. 1986. Т. 22. № 4. С. 699–703.
 8. *Crow J.F., Mange A.P.* Measurement of inbreeding from the frequency of marriages between persons of the same surname // *Social Biology*. 1982. V. 29. № 1/2. P. 101–105.
 9. *Lasker W.G.* Surnames and Genetic Structure. Cambridge: Cambr. Univ. Press, 1985. 148 p.
 10. Сорокина И.Н., Чурносов М.И., Балтуцкая И.В. и др. Антропогенетическое изучение населения центральной России. М.: Изд-во РАН, 2014. 336 с.

Description of Divergence of Subpopulations in the Hierarchical System under the Analysis of Isonymy. II. Probabilities of Non-Isonymic Encounters

V. P. Passekov^a, *

^a*Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Moscow, 119991 Russia*

*e-mail: pass40@mail.ru

Metapopulations with typical for human populations hierarchical subdivision into parts (subpopulations) are considered, corresponding to classifications based on administrative-territorial division (say, village, village council, district, region, and so on) or on a genealogical approach based on ethnogenesis, as well as on other principles of biological classification. The purpose of this work is to analyze the general properties of the distribution of the surname concentration over subpopulations in their hierarchical structure. Attention is focused on the description of the surname divergence of subpopulations, as an indicator of which the total probability of two person with different surnames encounter is considered, which is understood as the probability of the encounter in a randomly chosen subpopulation, a unit of observation, from metapopulation with a hierarchical structure. Its decomposition by hierarchy levels is obtained, which generalizes the Wahlund effect in population genetics. The total probability of non-isonymic encounters in a hierarchically subdivided metapopulation is less than the probability of random encounters in it by the sum of the average surname concentration intragroup variances corresponding to separate levels. Such properties are purely statistical characteristics of the hierarchical structure, and not a feature of a particular population system, and are not derived from the regularities of one or another model of microevolution. They are computationally formulated in the same way for any hierarchical system, although in the general case they do not coincide quantitatively. The results obtained refer to rural and urban hierarchical metapopulations as separate components of the entire population.

Keywords: hierarchical structure of populations, metapopulations, surname concentrations in human subpopulations, characteristics of heterogeneity of subpopulations, decomposition of the probability of encounter carriers of different surnames by hierarchy levels.

ПОИСК ГЕНОВ-КАНДИДАТОВ ЖИВОЙ МАССЫ У ОВЕЦ ЮЖНОЙ МЯСНОЙ ПОРОДЫ

© 2023 г. В. С. Шевцова^{1, 2, *}, А. Я. Куликова³, Л. В. Гетманцева⁴, А. В. Усатов¹

¹Южный федеральный университет, Ростов-на-Дону, 344000 Россия

²Южный научный центр Российской академии наук, Ростов-на-Дону, 344006 Россия

³Краснодарский научный центр по зоотехнии и ветеринарии, Краснодар, 350055 Россия

⁴Донской государственный аграрный университет, Ростовская обл., пос. Персиановский, 346493 Россия

*e-mail: barbaragen4@mail.ru

Поступила в редакцию 28.06.2023 г.

После доработки 30.06.2023 г.

Принята к публикации 04.07.2023 г.

В работе представлены результаты поиска генов-кандидатов живой массы у овец южной мясной породы. Выборка из 48 овцематок, живая масса которых измерялась в возрасте старше 24 мес., была генотипирована на чипе средней плотности Ovine SNP 50 GenotypingBeadChip. Расчеты, выполненные с помощью метода F_{st} , позволили выявить 55 полиморфизмов, 10 из которых отобраны для анализа влияния на исследуемый показатель. Все отобранные гены-кандидаты продемонстрировали связь с признаком живой массы. Особого внимания заслуживает ген *NDFI1*, в котором детектировано два генетических варианта.

Ключевые слова: овцы, полиморфизмы, живая масса, индексы фиксации.

DOI: 10.31857/S0016675823110139, **EDN:** NNLDIN

Сохранению отечественных пород сельскохозяйственных животных в значительной мере помогает изучение их генетических особенностей. Основанием для этого может служить их хозяйственная ценность. Одной из весьма перспективных отечественных пород мясного направления продуктивности является южная мясная, полученная путем ряда сложных скрещиваний овцематок отечественных пород в типе корриделей с баранами тексель и впоследствии с восточно-фризской молочной [1]. Одним из ключевых в мясном овцеводстве является показатель живой массы. Будучи количественным признаком, живая масса находится под контролем большого числа генов. Кроме того, рядом авторов показана породоспецифичность и зависимость от этапа онтогенеза генных комплексов, вовлекаемых в ее формирование [2].

Цель исследования – выявление на основе результатов генотипирования на чипах средней плотности генетических вариантов, обуславливающих у представителей южной мясной породы высокие показатели живой массы. На чипе средней плотности были генотипированы 48 овцематок, взвешивание которых производилось в возрасте 24 мес. и старше. Фильтрация результатов генотипирования была проведена по методике, рекомендованной в работе [3]. Методом F_{st} (один

из индексов фиксации в статистике Райта) были определены значимые генетические варианты, участвующие в формировании признака живой массы. Для расчетов животные были ранжированы по признаку живой массы и отобраны крайние значения (максимальные и минимальные). Овцематки с живой массой 55–59 кг составили субпопуляцию (шесть голов) с минимальным показателем. Живая масса во второй субпопуляции превышала 80 кг (девять голов). Значимыми вариантами считали те, у которых значения F_{st} превысили уровень квантиля 0.999 (0.1% выбросов). Результаты позволили выявить 55 полиморфизмов, связанных с показателем живой массы у овец южной мясной породы.

Паттерн распределения по аутосомам выявленных генетических вариантов отличается от ранее опубликованных данных по локусам, ассоциированным с живой массой: *OAR9* – 11 SNPs, *OAR5* – 10 SNPs, *OAR2* – 6 SNPs, *OAR3* и *OAR6* – по 5 SNPs, *OAR1* и *OAR17* – по 3 SNPs, *OAR11*, *OAR12* и *OAR21* – по 2 SNPs, по одному SNP – на *OAR7*, *OAR13*, *OAR15*, *OAR23* и *OAR24*. Аннотирование, проведенное в геномном браузере Ensembl, показало, что 30% SNPs попали в интроны ряда генов: 13% – в 3'-область, 36% – в 5'-область генов, 7% составили межгенные участки и для 13% вари-

антов не обнаружено информации в базе данных Ensembl.

Выявленные полиморфизмы позволили отобрать десять генов-кандидатов на основании значения $F_{st} \geq 0.55$ и локализации в пределах определенного гена (*LY6D*, *FAM114A1*, *U6*, *NDFIP1*, *LRRC1* и 5 генов *ENSOARG*, белковые продукты которых еще не изучены). Эффект каждого гена-кандидата на показатель живой массы у южной мясной породы был исследован по всей выборке: по трем генотипам каждого гена рассчитаны средняя живая масса и стандартное отклонение, а также определены уровни значимости различий между средними (p) путем попарного сравнения масс у представителей разных генотипов.

Было установлено, что все отобранные гены-кандидаты влияют на исследуемый показатель живой массы. Наименьший уровень статистической значимости различий между средними значениями живой массы демонстрирует ген *FAM114A1* ($p = 0.012$), для которого разница в живой массе между овцами с генотипами *GG* и *AA* составила 9.85 кг, однако и он находится в диапазоне достоверных различий $p < 0.05$. Предпочтительными генотипами, дающими носителям преимущество в живой массе 7.47–12.5 кг, по генам *ENSOARG00020003076*, *LY6D* и *ENSOARG00020017606* являются гомозиготы *AA*. По генам *FAM114A1*, *ENSOARG00020014128*, *ENSOARG00020002577*, *U6*, *NDFIP1* (два полиморфизма) и *LRRC1* носители генотипов *GG* имеют среднюю живую массу больше на 8.67–12.31 кг по сравнению с остальными генотипами ($p = 0.00026$ – 0.0062).

Функционал белковых продуктов генов-кандидатов весьма разнообразен: участие в развитии и функционировании нервной системы (*LRRC1*, *SEZ6L*, *FAM114A1*), формирование и функцио-

нирование иммунитета (*LY6D*, *NDFIP1*), участие в сплайсинге (*мяРНКУ6*).

Гены-кандидаты живой массы, выявленные в ходе исследования, могут быть рекомендованы, в качестве маркеров для селекционной работы с овцами южной мясной породы. Ген *NDFIP1*, в котором выявлено два генетических варианта, дающих гомозиготам преимущество в живой массе ≈ 9 – 12 кг заслуживает дальнейшего внимательного изучения. Грамотная селекция, учитывающая представленные в данной работе результаты, позволит не только повысить рентабельность мясной отрасли овцеводства, но и сохранить южную мясную породу овец, хорошо адаптированную к условиям юга России.

Исследование выполнено в рамках работ по гранту РФФИ № 20-316-90048.

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. Ульянов А.Н., Куликова А.Я., Баша С.Н. и др. Южная мясная порода овец // Овцы, козы, шерстяное дело. 2010. № 2. С. 65–69.
2. Gholizadeh M., Rahimi-Mianji G., Nejati-Javaremi A. Genome-wide association study of body weight traits in Baluchi sheep // J. of Genetics. 2015. V. 94. P. 143–146. <https://doi.org/10.1007/s12041-015-0469-1>
3. Marees A.T., de Kluiver H., Stringer S. et al. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis // Int. J. of Meth. in Psych. Res. 2018. V. 27. № 2. <https://doi.org/10.1002/mp.1608>

Search for Candidate Genes for Live Weight in Southern Meat Breed Sheep

V. S. Shevtsova^{a, b, *}, A. Ya. Kulikova^c, L. V. Getmantseva^d, and A. V. Usatov^a

^aSouthern Federal University, Rostov-on-Don, 344000 Russia

^bSouthern scientific center, Rostov-on-Don, 344006 Russia

^cScientific Krasnodar Centre for Animal Husbandry and Veterinary Medicine, Krasnodar, 350055 Russia

^dDon State Agrarian University, Rostov oblast, Persianovski set., 346493 Russia

*e-mail: barbaragen4@mail.ru

The paper presents the results of candidate genes search for live weight in southern meat sheep breed. The sample of 48 ewes weighed in the age of 24 months and elder was genotyped on Ovine SNP 50 Genotyping BeadChip array. Using a fixation in dices method (F_{st}), we detected 55 polymorphisms associated with live weight. 10 SNP with the highest F_{st} value were selected to check the effect on the trait under investigation. All the candidate genes showed the association with live weight trait. It's interesting to note that in the *NDFIP1* gene there were detected two significant SNPs with a high F_{st} value.

Keywords: sheep, polymorphisms, live weight, fixation indices.