УДК 575.17

Е. В. Балановская, В. В. Напольских, В. С. Чураков, Н. В. Пислегин, Ю. С. Записецкая, Д. К. Черневский, О. П. Балановский

ГЕНОФОНДЫ УДМУРТОВ И БЕСЕРМЯН В КОНТЕКСТЕ ФИННО-УГОРСКИХ И ДРУГИХ ОКРУЖАЮЩИХ НАРОДОВ: ПОЛНОГЕНОМНЫЕ И ФАРМАКОГЕНЕТИЧЕСКИЕ ДАННЫЕ[1]

Впервые генофонд коренных народов Удмуртии (пять популяций удмуртов и бесермян) исследован по обширной полногеномной панели ДНК-маркеров в широком контексте всех финно-угорских и окружающих народов: суммарно анализируются 728 геномов из 76 популяций.

Анализ предковых компонент ADMIXTURE выявил особую «удмуртскую» предковую компоненту, составляющую почти 100 % геномов всех удмуртов и основную часть геномов бесермян. Второй по значимости предковой компонентой генофонда бесермян оказалась «беломорская», составляющая основу геномов коми-зырян и русских беломорского севера. В целом три независимых метода анализа полногеномных данных (ADMIXTURE, F_{ST}, PCA) обнаруживают наибольшее генетическое сходство удмуртов и бесермян с коми-пермяками, далее по степени близости генофондов следуют марийцы, чуваши, башкиры и татары Поволжья.

Впервые дана характеристика фармакогенетического статуса населения Удмуртии, определены частоты 45 фармакогенетических маркеров. Картографирование показывает, что частоты генетических маркеров, используемых в фармакогенетических рекомендациях, у бесермян и удмуртов близки к частотам у коренного населения Поволжья, Приуралья и Южного Урала, но не более отдаленных регионов: в пределах этой территории можно пользоваться едиными протоколами при развитии персонифицированной медицины.

Публикуемые карты распространения «удмуртских» предковых компонент генофонда народов европейской части России и Урала дают возможность представителям гуманитарных наук самостоятельно анализировать генетические следы взаимодействия удмуртов и бесермян с другими народами.

Ключевые слова: удмурты, бесермяне, этногенез, генофонд, геногеография, полногеномные панели, геном, Удмуртия, фармакогенетика, финно-угры, генетическая история населения, Волго-Уральский регион.

DOI: 10.35634/2224-9443-2022-16-2-328-346

Введение

Популяционная генетика, опирающаяся на анализ современной и древней ДНК, стала признанным историческим источником: палеогенетика служит вспомогательным инструментом, выступая в союзе с анализом генофонда современного населения. Для реконструкции генетической истории населения используются два наиболее эффективных инструмента молекулярной генетики: изучение вариантов Y-хромосомы по очень широкой панели ДНК-маркеров и анализ аутосомного генофонда по обширным полногеномным панелям.

Поскольку у «мужской» Y-хромосомы нет пары и ей не с кем обмениваться фрагментами при кроссинговере, она изменяется лишь благодаря мутациям: они ведут четкую запись генетической истории, позволяя восстановить генеалогическое древо по отцовской линии – «отец мужчины–дед–прадед–прапрадед» и далее до Адама. Исследования Y-хромосомного генофонда удмуртов и бесермян [Tambets et al., 2004; Трофимова и др., 2015; наши неопубликованные данные] указывают на его ярко выраженное своеобразие. Поэтому на первый план выступает необходимость анализа их аутосомного генофонда, что позволит получить «бинокулярное зрение»: возможность объективно сравнить «видение» обеих наиболее информативных генетических систем – Y-хромосомы и аутосомного генофонда.

Формирование удмуртского народа происходило в Предуралье и Прикамье. Удмуртский язык вместе с близкими к нему языками коми восходит к пермскому праязыку. Поздняя (к концу I тыс. н.э.) пермская прародина локализуется в Пермско-Сарапульском Прикамье; распад единства произошёл ме-

[1] Исследование выполнено при финансовой поддержке Государственного задания Министерства науки и высшего образования РФ для Медико-генетического научного центра им. академика Н. П. Бочкова (биоинформатический анализ данных, интерпретация результатов) и гранта РНФ № 21-14-00363 (анализ фармакогенетических маркеров, картографический анализ).

нее тысячи лет назад [Белых, 2009]. Ещё в прапермское время (в I тыс. н.э. и ранее) южные пермяне – предки удмуртов – активно контактировали и с ираноязычными (сарматскими, аланскими) группами Предуралья и Поволжья, и с носителями балтских языков, и с предками венгров. На западе, в бассейне Вятки, можно предполагать присутствие парапермских групп, через посредство которых предки удмуртов находились в контакте с марийско-мерянской общностью [Смирнов, 2013, 2014; Напольских, 2015].

Распад прапермской общности (в начале II тыс. н.э.) связан с появлением в нижнем Прикамье (не позднее IX в.) булгар: в рамках Волжской Булгарии, по-видимому, происходило становление удмуртского народа и языка. Тесные контакты предков удмуртов с тюрками – булгарами и кыпчаками – во многом отличают языки удмуртов и коми. В ходе взаимодействия предков удмуртов с булгарами в XI–XV вв. произошло формирование этнической группы бесермян, обнаруживающих в культуре черты мощного булгаро-чувашского и исламского влияния [Белых, 2009; Напольских, 2015].

Новый этап в этнической истории удмуртов и бесермян, приведший к возникновению двух этнографических групп удмуртов (южной и северной) и полной смене ареала бесермян, связан с развернувшимся во второй половине XV в. противоборством Русского государства и Казанского ханства за контроль над Вятско-Камским междуречьем. При описании очередного конфликта Москвы и Казани в 1469 г. впервые упоминается подвластная казанскому хану *Вотятъцкая земля* (вотяки – старое русское название удмуртов). Судя по сохранившимся грамотам, Иван III начал проводить политику по стимулированию переселения удмуртов и бесермян из «Казанских мест» в формирующийся Каринский стан (район впадения Чепцы в Вятку к востоку от г. Кирова). По переписи 1615 г. в Каринском стане насчитывалось 10 дворов «княжьих» и 4 «тотарских», не менее 32 дворов бесермян и около 270 дворов удмуртов. К ним следует добавить 60 дворов удмуртов соседнего Лужановского стана[2]. Вплоть до реформы Екатерины II (1780 г.), южные удмурты проживали в пределах удмуртских сотен Арской дороги Казанского уезда (правобережье нижней Вятки – районы, известные как Заказанье), а северные, периодически включавшие в свой состав переселенцев с юга, – в Каринском стане. При этом в освоении башкирских земель (рубеж XVI–XVII в.) приняли участие и северные, и южные удмурты, и отдельные представители бесермян [Чураков, 2017; Пислегин, Чураков, 2020; Садиков, 2001]. Из 552 тыс. удмуртов (перепись 2010 г.) южные составляют 3/5. Бесермяне после переписи 1926 г. записывались в число удмуртов, что привело к смене этнической самоидентификации у значительной части бесермян: перепись 2010 г. указывает 2,2 тыс. чел.

По данным антропологии, отмечается высокое сходство между разными группами удмуртов. Их антропологический тип, для которого «несомненным является автохтонное, смешанное и преимущественно европеоидное происхождение в волго-уральском регионе» [Аксянова, 2003, 122], близок к коми-пермякам, южным популяциям коми, марийцам, татарам, башкирам и чувашам, хотя по данным разных авторов наблюдается разная степень сходства с перечисленными народами [Дерябин, 1998; Аксянова, 2003]. Сближение с марийцами может быть связано как с поглощением общего доудмуртского населения, так и с ассимиляцией удмуртов при продвижении марийцев на восток. Сходство с башкирами и северо-западными татарами может быть либо наследием общих прапермских популяций, либо результатом ассимиляции удмуртов в башкиро-татарской среде. Антропологический тип бесермян близок к удмуртам, марийцам, некоторым популяциям татар и чувашей. У бесермян усилен тот расовый компонент, который выделяет удмуртов на фоне пермских финнов и марийцев и сближает удмуртов и бесермян с тюрками Поволжья [Дерябин, 1998; Аксянова, 2012]. По данным краниологии в популяциях пермских народов XVII–XIX вв. выделяются два европеоидных комплекса: к прикамскому относятся коми-пермяки и большая часть удмуртов, к североевропейскому – комизыряне и среднечепецкие удмурты [Широбоков 2014].

Популяции удмуртов изучались в целом ряде работ практически по всему спектру генетических маркеров: по маркерам митохондриальной ДНК [Бермишева и др., 2002; Грошева и др., 2013, 2014], Y-хромосомы [Трофимова и др., 2015], отдельным аутосомным маркерам [Бермишева и др., 2007; Зинченко и др., 2007], а также полногеномным (широкогеномным) панелям [Triska et al., 2017; Tambets et al., 2018; Jeong et al., 2019]. Но удмурты редко выступали основным объектом исследования, чаще они изучались в общем ряду других народов. В результате в литературе имеется не так мало данных о генофонде удмуртов, но не так много выводов из этих данных.

Из всех генетических маркеров для удмуртов наиболее подробно изучена митохондриальная ДНК. Впервые она была исследована специалистами Эстонского биоцентра в ряду народов Волго-

[2] Спустя 300 лет, опираясь на данные переписи 1926 г., численность каринских (чепецких) татар можно оценить в 15 тыс. чел., северных удмуртов – в 200 тыс. чел., бесермян – 10 тыс. чел.

Уральского региона [Бермишева и др., 2002]. Выборка удмуртов происходила из Малопургинского района на юге Удмуртии и Татышлинского района на севере Башкирии (N=101). Суммарная частота восточно-евразийских гаплогрупп у удмуртов в этой работе была определена в 21 %, высказано предположение о значительном действии дрейфа генов в популяции удмуртов. В систематическом исследовании митохондриальной ДНК удмуртов [Грошева и др., 2014] доля восточно-евразийских гаплогрупп в разных популяциях удмуртов варьировала от 16 % до 40 % (в среднем 30 %) – это очень высокое значение для Волго-Уральского региона и максимальное (кроме бесермян) для популяций к западу от Уральских гор [Tambets et al., 2018]. В работе [Грошева и др., 2014] подчеркивается выделение двух групп удмуртских родов. К первой группе – Ватка – авторы относят рода, расселенные и на севере Удмуртии, и на юге (современный Татарстан). Ко второй группе – Калмез – отнесены рода, расселенные посредине ареала, при этом указывается брачная изоляция двух групп родов. Северных удмуртов авторы считают потомками родов Ватка, центральных удмуртов – потомками родов Калмез, а южных удмуртов – результатом смешения родов Калмез и Ватка. Изучен генофонд 10 географических популяций, из них 200 образцов отнесены не только к географическим, но и к родовым группам Ватка, Калмез и Арск: генетическая изменчивость «родовых» выборок оказалась выше, чем «географических». Однако авторская методика вызывает вопросы: родовая принадлежность у удмуртов наследуется по отцовской линии [Чураков, 2003], а анализируемая митохондриальная ДНК – по материнской. Более того, этнографические исследования указывают на отсутствие деления удмуртов на племена: объединение Ватка – это чепецкие удмурты, а объединение Калмез – иллюзорно [Косарева, 2017] , поскольку является отражением так называемой Кильмезской волости, у населения которой сформировалось локальное самосознание.

Той же группой авторов по мтДНК изучена и выборка (N=41) бесермян [Грошева и др., 2013]. Доля восточно-евразийских гаплогрупп у бесермян (46 %) оказалась максимальной среди всех народов Волго-Уральского региона. Преобладает гаплогруппа C, и ее гаплотипы, выявленные у бесермян, распространены в популяциях Южной Сибири. Некоторое генетическое отдаление бесермян от других популяций Волго-Уральского региона обнаруживается по всему спектру гаплогрупп. На основании всех этих данных авторы работы [Грошева и др., 2013] приходят к выводу о смешанном удмуртско-тюркском происхождении бесермян.

Исследование маркеров Y-хромосомы [Трофимова и др., 2015] посвящено не удмуртам, а в целом финно-угорским и тюркоязычным народам Волго-Уральского региона. Уровень филогенетического разрешения невысок: только 12 гаплогрупп. Для удмуртов указано сходство с остальными финно-угорскими популяциями региона (кроме мордвы), а для бесермян – финно-угорская, а не тюркская основа генофонда. При этом авторы отмечают, что без подразделения гаплогруппы N на субтипы этот вывод является предварительным. Ее подразделение было проведено в работе [Ilumae et al., 2016], выполненной большим международным коллективом с нашим участием. Хотя удмурты в тексте статьи не упоминаются, но из приложения и карт следует, что у них преобладает ветвь N3a1, роднящая их с коми-зырянами, и в меньшей степени – с коми-пермяками, чувашами и марийцами.

Для третьей группы генетических маркеров – аутосомных – есть данные [Бермишева и др., 2007] лишь о 10 ДНК-маркерах в шести районах Удмуртии. Основной вывод работы состоит в кластеризации обследованных районов в две группы – южную и северную, которые слабо различаются и по частотам аллелей, и по уровню гетерозиготности. Размах межпопуляционных различий у удмуртов несколько выше, чем у марийцев, но ниже, чем у чувашей и башкир.

Популяции тех же шести районов Удмуртии изучены в работе [Зинченко и др., 2007] по 4 ДНК-маркерам (2 совпадают с маркерами в [Бермишева и др., 2007]), но им дана также генетико-демографическая характеристика. Показано, что индекс эндогамии сельского населения варьирует по районам от 48 % до 76 %. Авторы отмечают (правда, не приводя ссылок) быстрый рост численности удмуртов – в 15 раз за последние 10 поколений – и предполагают, что это обусловило интенсивный дрейф генов и особенности в распространении наследственной патологии.

Наконец, удмурты были изучены и по полногеномным панелям, включающим сотни тысяч аутосомных ДНК-маркеров. В статье [Triska et al., 2017] 30 образцов удмуртов изучены по панели Illumina 660k; в [Tambets et al., 2018] 16 образцов удмуртов генотипированы по той же панели; в статье [Jeong et al., 2019] с нашим участием 10 образцов удмуртов и 6 образцов бесермян изучены по панели Affimetrix Human Origin, также включающей около 600 тысяч маркеров. Однако во всех работах удмурты выступали лишь одной из ряда популяций, и данные о них не обсуждались. В [Tambets et al., 2018] выявлено несколько большее генетическое сходство между уралоязычными популяция-

ми, чем между уралоязычными и другими популяциями тех же географических регионов, но и в этой статье генофонд пермской группы народов не рассматривался.

Таким образом, хотя генетические данные об удмуртах имеются по всем основным генетическим системам, но выводы об их генетической истории или отсутствуют, или основаны на данных о небольших панелях ДНК-маркеров. Для следующего шага в изучении их генофондов необходимо получить углубленные данные об изменчивости Y-хромосомы и провести целенаправленный анализ удмуртов и бесермян по широкогеномным панелям маркеров. Реализации второго из этих подходов и посвящена данная работа.

Ее цель – детальный анализ генофонда популяций удмуртов и бесермян по обширным полногеномным панелям, а основная задача – оценить, какое положение занимают аутосомные генофонды удмуртов и бесермян среди народов финно-угорского мира и окружающих славянских и тюркских народов.

Материалы и методы

Методические и биоинформатические аспекты изучения аутосомного генофонда по полногеномной панели детально описаны в [Балановский и др., 2020], что позволяет здесь ограничиться только основными штрихами.

При изучении единичных ДНК-маркеров выборка должна быть не менее 50 индивидов. При анализе полногеномных панелей, включающих сотни тысяч ДНК-маркеров, адекватная выборка составляет 5-15 индивидов. Однако при этом резко возрастают требования к качеству выборки, поскольку эти единичные образцы должны стать полномочными представителями геномов всей популяции. Если задача исследования – создать не «мгновенный снимок» генофонда популяции, который очень скоро изменится из-за мощных современных миграций, а реконструировать генетическую историю народа, то выборка должна отвечать правилу «трех поколений» [Балановская и др., 2016а]: генеалогия индивидов на протяжении не менее трех поколений должна быть связана только с исследуемой популяцией, а сами популяции должны отражать субэтническую структуру изучаемых народов. Поэтому в наше исследование включены две популяции бесермян – восточных (5 геномов из Балезинского района) и западных (5 геномов из Юкаменского района Удмуртии). Популяции удмуртов (18 геномов) разделены согласно историческому делению на северных (5 геномов из Юкаменского района) и более южных (13 геномов), подразделенных на «восточных» и «центральных». Популяция Шарканского района занимает промежуточное положение между северными и южными удмуртами, поэтому ее выделили в отдельную «восточную» популяцию (5 геномов). Остальные 8 геномов отнесли к «центральной» популяции (геномы из Вавожского, Завьяловского, Увинского районов), включив в нее и геном удмурта из Кукморского района Татарстана (но выделив его как южных удмуртов на графике главных компонент PCA). Итого в суммарную выборку из Удмуртии включены те 28 индивидов, у которых в генеалогической части анкеты указано, что четверо их бабушек и дедов относили себя к данному этносу и родились в данной популяции.

Их основной анализ (методом ADMIXTURE) проведен в обширном контексте 700 геномов из 71 популяции других народов Восточно-Европейской равнины и Урала (табл. 1): 217 геномов, охватывающих все финно-угорские народы России, 11 геномов самодийцев (ненцев), 112 геномов тюркоязычных народов Волго-Уральского региона, 8 геномов балтоязычных литовцев, 352 генома славян.

Генотипирование по полногеномной панели ДНК-маркеров выполнено с помощью биочипа Infinium OmniExome BeadChip Kit (Illumina, США) на приборе iScan (Illumina, США). Первичный анализ и оценка качества проведены в программе GenomeStudio v2011. 1. Для исследованных образцов показатель CallRate составлял не менее 0,99.

Анализ генотипов всех 728 геномов из 76 популяций проведен методом ADMIXTURE [Alexander et al., 2009], который, наряду с анализом главных компонент, является базовым методом описания генетической структуры популяций по полногеномным данным. Метод ADMIXTURE дает количественную оценку вклада разных предковых компонент в каждый индивидуальный геном. Генетический состав каждой предковой компоненты определяется автоматически на основе одной и той же совокупности изученных геномов. Результаты являются объективными и не зависят от исследователя: единственный параметр, который задает исследователь – число предковых компонент k. При k=2 для каждого проанализированного генома выявляется вклад двух предковых компонент; при k=3 для тех же геномов – вклад уже трех компонент; при k=4 – четырех компонент; при k=20

реконструируется вклад двадцати предковых компонент для той же совокупности геномов. Таким образом, при увеличении k программа для одних и тех же геномов рассчитывает все большее число более дробных предковых компонент: при большом числе k выделяет все более детальные предковые компоненты, позволяя охарактеризовать субэтносы, а порой и отдельные популяции. Для одной и той же совокупности геномов анализ проводится для разных k. При каждом k программа ADMIXTURE выдает оценку вклада (0 %<C<100 %) каждой из предковых компонент в каждый индивидуальный геном. Затем, усреднив на каждом уровне k вклад каждой компоненты во все геномы данной популяции, оценивается средний вклад каждой предковой компоненты в ее генофонд. Полученные величины могут быть картографированы. Для реконструкции генетической истории удмуртов и бесермян в работе представлены 4 карты для предковых компонент при k=5, k=6, k=14.

Расчет генетических расстояний F_{ST} проведен попарно между всеми популяциями. Фильтрация массива геномных данных (728 геномов) с параметрами geno = 0.1, maf = 0.01, mind = 0.05, indep-pairwise = 1500, 150, 0.5 проведена с помощью программного обеспечения PLINK [Chang, 2015]; затем файлы из формата plink были конвертированы в формат eigensoft [Patterson, 2006] и далее рассчитана матрица расстояний программой smartpca [Patterson, 2006] с параметром fstonly = YES.

При расчете главных компонент PCA для охвата зарубежных финно-угорских популяций были включены данные [Lazaridis 2014; Jeong 2019; Patterson 2012] по панели Human Origins (550 тысяч ДНК-маркеров). Объединенный массив включил 433 генома, проанализированного по 410 тысячам ДНК-маркеров, присутствующих в обоих исходных массивах (Human Origins и наш массив). Фильтрация полученного объединенного массива проведена с параметрами geno = 0.05, maf = 0.01, mind = 0.1, indep-pairwise = 1500, 150, 0.2 (PLINK [Chang CC, 2015), далее файлы переведены в формат eigensoft программой convertf [Patterson 2006]. Расчет главных компонент проведен программой smartpca [Patterson, 2006], графики построены с использованием библиотек на языке Python3 [Waskom 2021].

Для описания распространения в популяциях удмуртов и бесермян клинически значимых ДНК-маркеров привлечены данные о 45 ключевых фармакогенетических маркерах, детально описанных в [Балановская и др., 2020]. Информация об их распределении в популяциях России и сопредельных стран извлечена в результате биоинформатического анализа нескольких массивов данных, полученных в рамках популяционно-генетических исследований с использованием широкогеномных панелей SNP-маркеров: отобраны 45 маркеров (генов ADME; генов, кодирующих фармакодинамические мишени лекарственных средств; генов, кодирующих компоненты системы гемостаза). Для анализа привлечены данные о 2197 геномах из 137 этнических и субэтнических групп, объединенных для целей анализа в 50 метапопуляций. На основе матрицы встречаемости 45 фармакогенетических ДНК-маркеров в 50 популяциях Северной Евразии рассчитаны генетические расстояния М. Нея от средней частоты в объединенной популяции удмуртов и бесермян частоты каждого ДНК-маркера до всех 49 популяций и созданы 45 карт генетических расстояний от популяции Удмуртии. После усреднения этих 45 карт получена карта средних генетических расстояний от популяции Удмуртии до всех изученных популяций, отражающая особенности фармакогенетического статуса удмуртов и бесермян.

Все карты построены с помощью оригинального картографического пакета GeneGeo [Кошель, 2012] методом средневзвешенной интерполяции с радиусом влияния 700 км и значением весовой функции 3. Детальное описание процесса создания геногеографических карт, их возможностей и чтения дано в [Балановская, Балановский, 2008; Балановский, 2015].

1. Анализ предковых компонент ADMIXTURE

При анализе генофонда коренного населения Удмуртии методом ADMIXTURE охвачен весь финно-угорский мир России – от саамов и ижоры на западе до обских угров на востоке. Анализ ADMIXTURE провели 14 раз, задавая число предковых компонент от k=2 до k=15. В результате получили для каждого из 728 геномов по 119 оценок вклада (contribution, C) каждой из 119 предковых компонент.

Их сравнение показывает, что на всех уровнях от k=4 до k=15 геномы удмуртов отличаются наличием очень высокого вклада (91 % <C<100 %) одной из предковых компонент (табл. 1). У бесермян эта компонента также достигает C=97 % (за исключением уровней k=4 и k=6, где ее вклад C=77 %). Эту предковую компоненту можно смело назвать «удмуртской», поскольку ее вклад в 700 изученных геномов других народов очень мал (табл. 1). Пространственное распределение вклада «удмуртской» предковой компоненты в геномы изученных народов отражено в таблице 1 и на картах: основной вариант при k=5 и k=14 на рис. 1, второй по значимости вариант при k=6 на рис. 2. Рассмотрим их.

Таблица 1

Вклад (в %) «удмуртской» предковой компоненты ADMIXTURE в геномы изученных популяций при k=5, k=6 и k=14

ЭТНОС (популяция)	Число популяций	Число геномов	Число предковых компонент k								
			k=5			**k=6**			**k=14**		
			MEAN	MIN	MAX	**MEAN**	MIN	MAX	**MEAN**	MIN	MAX
Бесермяне западные	1	*5*	**94**	68	100	**73**	52	85	**94**	68	100
Бесермяне восточные	1	*5*	**100**	100	100	**81**	72	85	**100**	100	100
Удмурты северные	1	*5*	**100**	100	100	**98**	92	100	**100**	100	100
Удмурты восточные	1	*5*	**100**	100	100	**100**	99	100	**100**	100	100
Удмурты центральные	1	*8*	**100**	100	100	**99**	94	100	**100**	100	100
Башкиры	7	*41*	**11**	0	25	**21**	0	37	**6**	0	22
Вепсы	1	*10*	**0**	0	2	**0**	0	5	**0**	0	2
Водь	1	*6*	**0**	0	0	**0**	0	0	**0**	0	0
Ижора	2	*14*	**0**	0	3	**0**	0	0	**0**	0	0
Карелы	3	*27*	**0**	0	3	**1**	0	9	**0**	0	0
Коми-зыряне	1	*7*	**21**	17	24	**18**	6	26	**2**	0	12
Коми-пермяки	5	*47*	**54**	37	68	**40**	29	50	**0**	0	15
Литовцы	1	*8*	**0**	0	0	**0**	0	2	**0**	0	0
Манси	1	*21*	**0**	0	8	**0**	0	0	**0**	0	0
Марийцы	2	*22*	**0**	0	0	**97**	65	100	**0**	0	0
Мордва	3	*39*	**3**	0	7	**13**	6	19	**0**	0	3
Ненцы	1	*11*	**1**	0	16	**1**	0	12	**0**	0	0
Русские	22	*275*	**2**	0	27	**1**	0	23	**0**	0	22
Саамы	1	*3*	**0**	0	0	**0**	0	0	**0**	0	0
Татары Поволжья	4	*41*	**10**	1	24	**25**	15	40	**7**	0	18
Украинцы	12	*77*	**1**	0	6	**2**	0	7	**0**	0	0
Ханты	1	*21*	**0**	0	7	**0**	0	3	**0**	0	0
Чуваши	3	*30*	**1**	0	6	**74**	44	96	**0**	0	0
Суммарно	76	*278*									

Обозначения:
MEAN – средний вклад предковой компоненты в генофонд данной популяции;
MIN – минимальное значение предковой компоненты среди изученных геномов данной популяции;
MAX – максимальное значение предковой компоненты среди изученных геномов данной популяции.

1А. Анализ ADMIXTURE на уровне k=5 (пять предковых компонент)

На уровне k=5 (как и на всех уровнях k≥5, за исключением k=6) все геномы и удмуртов, и бесермян имеют 100 % вклад «удмуртской» компоненты (табл. 1, рис. 1А). Лишь в одном геноме бесермянина этот вклад составляет не 100 %, а две трети генома, причем и при других значениях k, варьирует на уровне около двух третей генома. Остальную треть генома этого бесермянина на всех уровнях k составляет «новгородская» предковая компонента (ее детальный анализ см. [Балановская и др., 2021]). Этот пример показывает, что, анализируя индивидуальные геномы, мы можем видеть, какие именно генетические пласты они унаследовали и в какой мере.

На фоне практически полного единообразия удмуртов и бесермян по вкладу «удмуртской» компоненты (около 100 %), другие этносы на уровне k=5 явно делятся на три группы (табл. 1) по ее вкладу (С): в первую («пермскую») группу входят геномы коми-пермяков (С=54 %) и коми-зырян (С=21 %); во вторую («тюркскую») группу – геномы башкир и татар (10 %<С<11 %); в третью группу с отсутствием удмуртского влияния (0<С<3 %) – все остальные изученные популяции (табл. 1). Рассмотрим первые две группы детально.

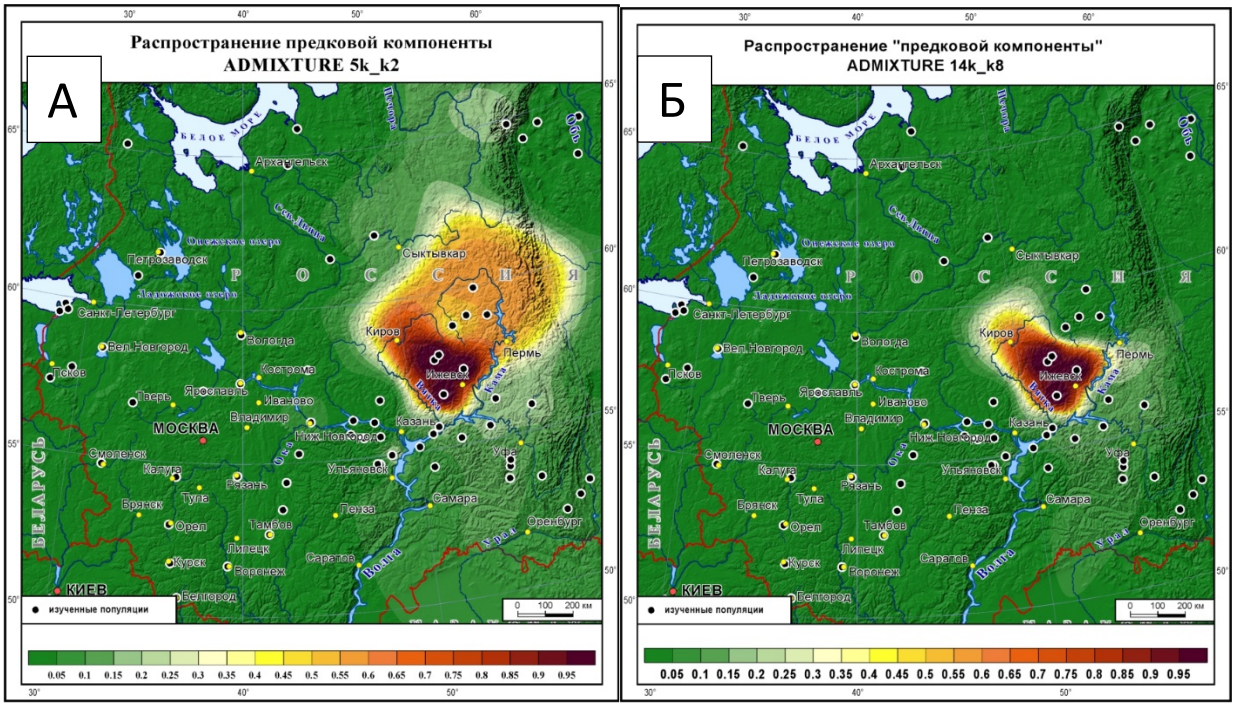

Рис. 1. Распространение условно «удмуртской» предковой компоненты ADMIXTURE:
А – при k=5 предковых компонент; Б – при k=14 предковых компонент

Обозначения: высокие значения предковой компоненты обозначены красно-коричневыми тонами, низкие – зелеными, шкала переходов приведена в легенде под картой; изученные популяции обозначены черными кружками.

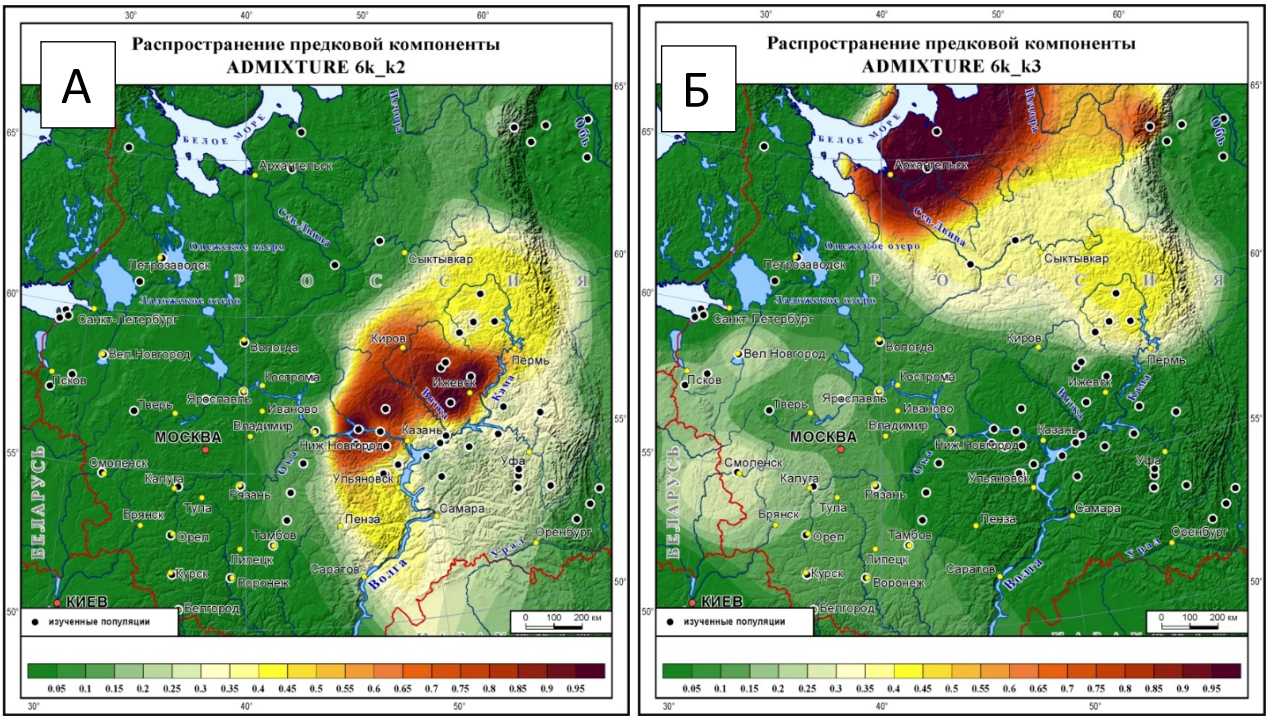

Рис. 2. Распространение предковых компонент ADMIXTURE при k=6:
А – условно «удмуртской», Б – условно «беломорской».

Обозначения: высокие значения предковой компоненты обозначены красно-коричневыми тонами, низкие – зелеными, шкала переходов приведена в легенде под картой; изученные популяции обозначены черными кружками.

«Пермская» группа включает популяции коми-пермяков с вкладом «удмуртской» компоненты около половины генофонда. Причем очень важно, что этот вклад стабилен: ни у одного индивида он не опускается ниже трети и не превышает двух третей генома (табл. 1). Такая стабильность, выявленная для геномов всех пяти популяций коми-пермяков, указывает на общий для них мощный генетический пласт, составляющий половину их генофонда и роднящий их с удмуртами и бесермянами. В популяции коми-зырян этот генетический пласт тоньше – он составляет пятую часть (C=21 %) генофонда, но тоже стабилен, варьируя по индивидуальным геномам от 17 % до 24 % (табл. 1). Поскольку эта популяция одна из самых географически удаленных групп коми-зырян (из Ханты-Мансийского округа), то можно предполагать, что мы охватываем весь диапазон вклада «удмуртской» компоненты (от пятой части до половины генофонда) в геномы народов, говорящих на языках той же пермской лингвистической ветви, что и удмурты.

Вторая («тюркская») группа охватывает многие популяции башкир и татар. Вклад в их геномы «удмуртской» компоненты невелик (в среднем десятая часть генофонда), но тоже стабилен (0 %<C<25 %). При изучении этих популяций по Y-хромосоме постоянно фиксировался значительный вклад дотюркского населения в их генофонды [Балановская и др., 2016б; Юсупов и др., 2017]. Анализ аутосомного генома показывает, что часть этого наследия, полученного от коренного населения Урала, принадлежит народам пермской языковой группы. Ассимиляция южнопермского (древнеудмуртского) населения происходила в ходе формирования кыпчакоязычных народов Поволжья и Приуралья в XIII–XVI вв. В двух районах южнопермский вклад в сложение башкир и татар мог быть наиболее значительным. Во-первых, это правобережье Нижней Камы и нижнее течение Вятки, где формировалось «ядро» удмуртов (и бесермян) еще в булгарское время. Во-вторых, это нижнее и среднее течение р. Белой – район распространения в I тыс. н. э. археологических культур (караабызской, бахмутинской и др.), в носителях которых можно видеть периферийно-южные параперм-ские группы прапермской общности, наиболее близкие к предкам удмуртов.

Рассмотренные на уровне k=5 закономерности являются ведущими, поскольку в основных чертах остаются неизменными почти на всех более высоких уровнях, лишь сужая, как «шагреневая кожа», географический ареал влияния (табл., рис. 1Б). Но в «пермской» группе на ряде уровней k вклад «удмуртской» компоненты полностью исчезает, заменяясь на свою собственную предковую компоненту.

Точно так же и для популяций Удмуртии не на всех уровнях k наблюдается 100 % преобладание «удмуртской» предковой компоненты. Как отмечалось выше, на уровнях k=4 и k=6 проявляются иные, более редкие, но не менее важные закономерности, требующие отдельного рассмотрения.

1Б. Анализ ADMIXTURE на уровне k=6 (шесть предковых компонент)

На уровне k=6 (табл. 1, рис. 2А) вклад «удмуртской» компоненты в генофонд бесермян снижается до 77 %. Анализ других предковых компонент при k=6 показал, что пятую часть геномов бесермян составляет еще один генетический пласт – условно названный «беломорским», поскольку в северных популяциях Архангельской области (Лешуконский и Пинежский районы) он составляет 97 % генофонда (рис. 2Б). Также велика доля «беломорского» генетического пласта у коми-зырян (74 %). Конечно, здесь мы видим генетический вклад не славянского, а субстратного финно-угорского населения европейского севера России, языки которого были либо близки прибалтийско-финско-саамским (но не идентичны им), либо образовывали особые группы в пределах финно-пермской общности. Сегодня мы неплохо можем представить себе языковой ландшафт дорусского севера Восточной Европы благодаря эпохальным исследованиям субстратной топонимии, проделанным А. К. Матвеевым [Матвеев, 2001, 2004, 2007, 2015].

Но если у бесермян на уровне k=6 вклад «удмуртской» компоненты снижается, то у марийцев при переходе с уровня k=5 (C=0 %) на уровень k=6 подскакивает до максимума (табл. 1, рис. 2А). Это согласуется с данными антропологии [Аксянова 2003], фиксирующими большое сходство фенотипа удмуртов и марийцев. Сходная картина обнаруживается и у чувашей: при переходе с k=5 на уровень k=6 вклад «удмуртской» компоненты от 0 % поднимается в среднем до трех четвертей генома чувашей (табл. 1, рис. 2А), достигая у верховых чувашей (вирьял) 82 % их генофонда.

Во вторую группу – с вкладом «удмуртской» компоненты на уровне k=6 меньше половины генома – входят (табл. 1, рис. 2А) коми-пермяки (C= 40 %), башкиры и татары (C≈25 %), коми-зыряне и народы Мордовии (с вкладом от 13 %<C<18 %).

Эти закономерности проявляются при малом числе предковых компонент k=3, 4, 6. При большем числе компонент (k=5, 7–15) и для марийцев, и для чувашей уже характерна «собственная» общая предковая компонента, составляя подавляющую часть их генофондов. Это позволяет выдвинуть

гипотезу, что сходство популяций Удмуртии, марийцев и чувашей отражает наиболее древний генетический пласт, возможно, сформировавшийся до того, как эти популяции приобрели наблюдаемое ныне лингвистическое своеобразие. Более вероятно, что здесь мы имеем дело с генетическим вкладом парапермских групп западной периферии пермской прародины, населявших бассейн Вятки и ассимилированных в начале II тыс. н. э. марийцами, удмуртами (потомками эндопермян из Среднего Прикамья) и чувашами.

2. Положение генофондов удмуртов и бесермян в генетическом пространстве

Рассмотренные результаты метода ADMIXTURE дают детальную реконструкцию генетической истории отдельных геномов. Ее обобщенную картину представляют два других независимых метода – расчет генетических расстояний F_{ST} между популяциями (табл. 2) и анализ расположения геномов в пространстве главных компонент PCA (рис. 3).

Таблица 2

Генетические расстояния F_{ST} от популяций удмуртов и бесермян до генофондов других народов

	Удмурты	Бесермяне
Удмурты	0,000	0,008
Бесермяне	0,008	0,000
Коми-пермяки	0,009	0,009
Башкиры	0,010	0,011
Татары Поволжья	0,011	0,010
Чуваши	0,012	0,012
Марийцы	0,015	0,016
Карелы	0,017	0,016
Русские	0,017	0,014
Мокша, шокша, эрзя	0,018	0,016
Украинцы	0,019	0,017
Вепсы	0,021	0,020
Коми-зыряне	0,021	0,021
Литовцы	0,021	0,020
Ижора	0,022	0,020
Ханты	0,027	0,031
Манси	0,028	0,033
Водь	0,029	0,026
Саамы	0,036	0,035
Ненцы	0,040	0,043

Генетические расстояния между популяциями F_{ST} рассчитаны по тем же данным, что и в предыдущем разделе (табл. 1), но метод F_{ST} предназначен для анализа выборок. Прежде всего отметим высокую гомогенность удмуртов (отмеченную и антропологами [Аксянова, 2003]): средние генетические расстояния между популяциями удмуртов в три раза меньше, чем расстояние от удмуртов до бесермян. При расположении популяций других народов по степени возрастания их генетических расстояний от удмуртов (табл. 2) видно их деление на группы. В группу наиболее близких к генофонду удмуртов и бесермян ($0.009 \geq F_{ST} \leq 0.012$) вошли коми-пермяки и все рассмотренные тюркоязычные популяции: башкиры, татары и чуваши. Следующую ($0.014 \geq F_{ST} \leq 0.019$) крайне разнородную группу составляют марийцы, карелы, популяции Мордовии (мокша, шокша, эрзя), русских и украинцев. Еще более удалена от удмуртов компактная группа с $0.020 \geq F_{ST} \leq 0.022$, в которую, кроме финно-угорских популяций (коми-зыряне, вепсы, ижора), вошли балтоязычные литовцы. Среди наиболее генетически далеких от удмуртов популяций ($0.027 \geq F_{ST} \leq 0.043$) оказались и западные (водь, саамы), и восточные (ханты, манси, ненцы) уралоязычные народы.

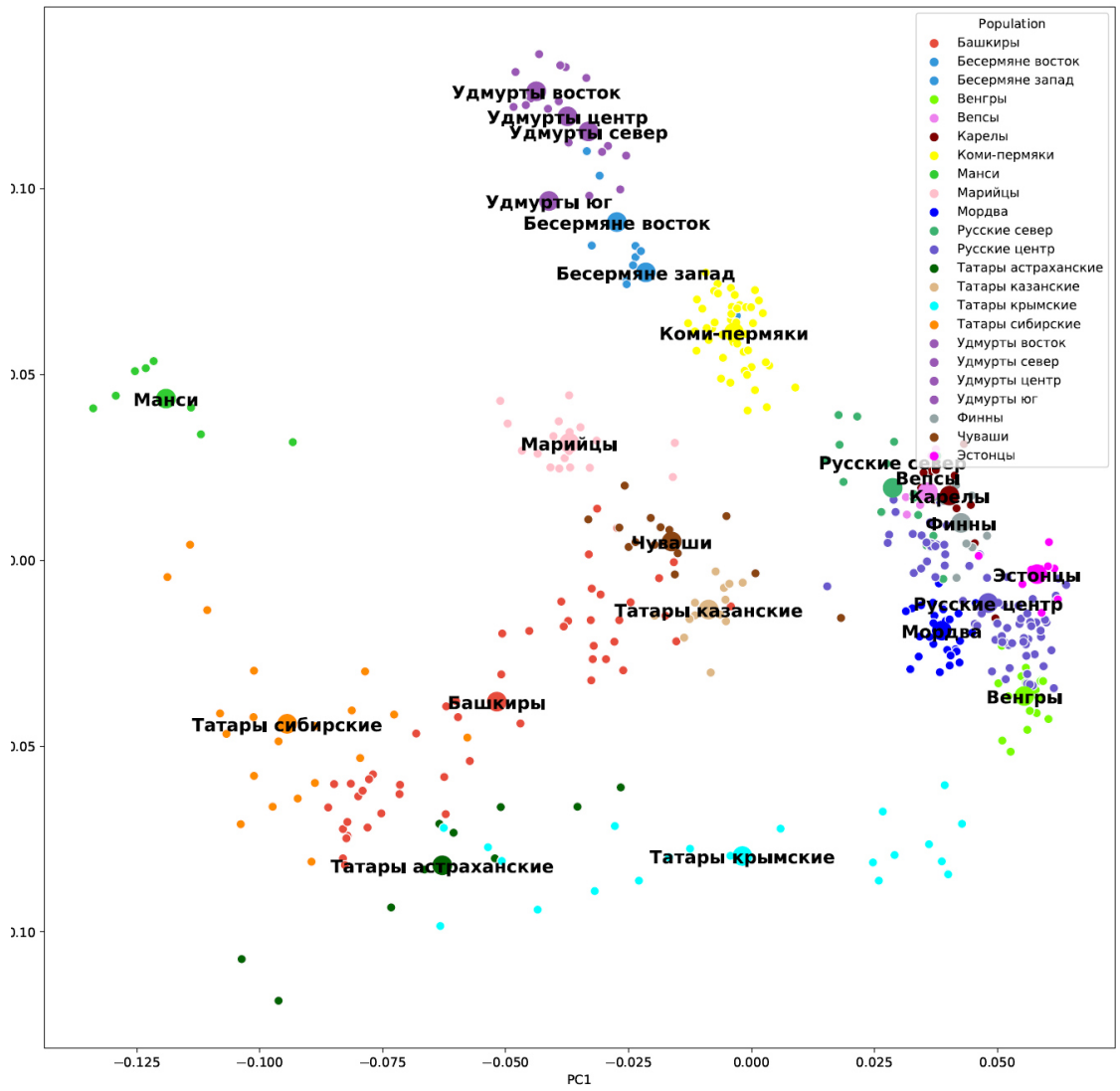

Рис. 3. Положение геномов удмуртов и бесермян в генетическом пространстве 1 и 2 главных компонент финно-угорских и других народов Европы

Обозначения: каждый индивидуальный геном обозначен отдельной точкой; усредненное положение популяции показано большим кругом; цветовой обозначение кругов приведено в легенде.

Положение популяций в пространстве главных компонент (PCA) рассчитывается независимым методом на основе корреляционной матрицы. Для того, чтобы максимально охватить генофонды финно-угорского мира, при анализе главных компонент привлечены также геномы, изученные по другой панели аутосомных маркеров (Human Origin), с целью включения в анализ эстонцев, финнов и венгров. В пространстве первой (PC1) и второй (PC2) главных компонент (рис. 3) наиболее генетически близки к удмуртам и бесермянам коми-пермяки, геномы которых в свою очередь тяготеют к популяциям Русского Севера. Четко видно, что бесермяне (особенно западные) намного ближе к коми-пермякам, чем удмурты. Эти же закономерности мы наблюдали при анализе другими независимыми методами – и ADMIXTURE, и F_{ST}, что подтверждает надежность закономерностей. На втором месте по степени сходства с генофондами бесермян и удмуртов – геномы марийцев. К ним близки геномы чувашей, среди которых на графике располагается и целый ряд геномов башкир. В целом уральские тюрки – чуваши, казанские татары и значительная часть геномов башкир – образуют весьма рыхлое, но общее облако, особенно на фоне огромной дифференциации геномов башкир, перекидывающих генетический мост от татар казанских к татарам сибирским. Западные популяции финно-угров – финны, карелы, эстонцы, венгры – вошли в общий тесный массив геномов вместе с популяциями Мордовии и русскими.

3. Фармакогенетический статус популяций Удмуртии

Развитие персонализированной медицины во многом связано с достижениями фармакогенетики. Фармакогенетические исследования, проводимые во всем мире, выявили ключевые гены, которые учитываются при назначении разных видов лекарств [Core ADME Gene List 2020; pharmgkb.org, 2020]. В мировой практике разработаны стандартные тесты и протоколы по подбору препарата и дозы в зависимости от генотипа пациента [Scott et al. 2013; Johnson et al. 2017; Lima et al. 2020;], но подавляющее число исследований проводится в популяциях Западной Европы, генофонды которых отличаются от большинства генофондов России. Поэтому основные ограничения для внедрения фармакогенетических тестов связаны с межэтническими различиями: доступные на рынке тесты их не учитывают.

Недавно было проведено первое систематическое изучение распространенности 45 значимых фармакогенетических вариантов в популяциях России, на основе которого создан первый фармакогенетический атлас для населения России и сопредельных стран [Балановская и др., 2020]. Он основан на массивах данных, полученных в рамках популяционно-генетических исследований с использованием полногеномных панелей ДНК-маркеров. Конечно, индивидуальные генотипы могут отличаться от средних показателей популяции. Но усредненный прогноз для конкретной российской популяции, учитывающий особенности ее генофонда, лучше подходит для ее членов, чем стандартные рекомендации по выбору препаратов и их дозировок, ориентированные на западноевропейские популяции.

Поскольку данное исследование популяций удмуртов и бесермян выявило яркое своеобразие их генофонда, можно ожидать, что оно нашло отражение и в своеобразии их фармакогенетического статуса. Поэтому провели его анализ, результаты которого продемонстрированы картой генетических расстояний (рис. 4).

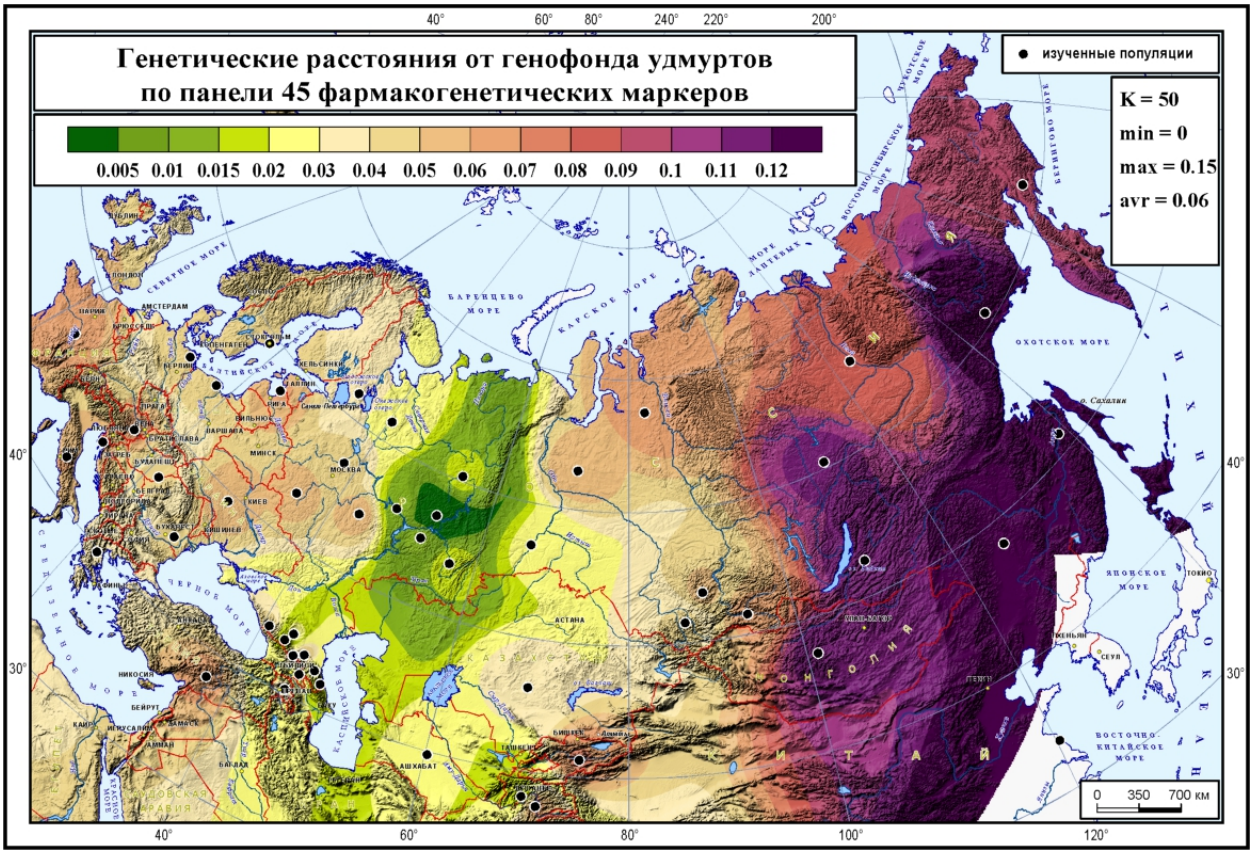

Рис. 4. Карта средних генетических расстояний от популяции коренного населения Удмуртии до популяций Северной Евразии

Обозначения: минимальные генетические расстояния (близость к генофонду Удмуртии) обозначены зелеными тонами; максимальные генетические расстояния (несходство с генофондом Удмуртии) обозначены красно-фиолетовыми тонами, шкала переходов приведена в легенде над картой; изученные популяции обозначены черными кружками.

В табл. 3 приведены частоты 45 фармакогенетических ДНК-маркеров, рассчитанные для объединенной популяции Удмуртии, включившей геномы удмуртов и бесермян. На основе этих частот построены 45 карт генетических расстояний от популяции удмуртов до каждой из 49 метапопуляций, и затем построена карта средних по 45 ДНК-маркерам генетических расстояний, дающая представление о своеобразии фармакогенетического статуса коренного населения Удмуртии. Она показывает, что ареал наиболее «фармакогенетически» близких популяций (отраженный на карте зелеными тонами) охватывает в основном коренное население Поволжья и Урала. И на запад от этого ареала, и особенно на восток тянутся регионы с иным фармакогенетичесим статусом. Эта карта (рис. 4) может служить первым ориентиром, указывающим, что фармакогенетические рекомендации, пригодные для персонифицированной медицины Удмуртии, могут охватывать коренное население Поволжья, Приуралья и Южного Урала, но не более отдаленные регионы.

Таблица 3

Частота фармакогенетических ДНК-маркеров у удмуртов и бесермян

Ген	Маркер	Частота в Удмуртии	Ген	Маркер	Частота в Удмуртии
ABCB1	rs1045642-G	0,47	CYP3A4	rs12721629-A	0,00
	rs4148738-G	0,50		rs2242480-A	0,11
APOE	rs429358-G	0,00		rs4986910-G	0,00
CES1	rs2244613-C	0,39		rs62471956-A	0,02
COMT	rs4680-A	0,63	CYP3A5	rs10264272-A	0,00
CYP1A2	rs12720461-A	0,00		rs28365083-A	0,00
				rs41303343-AA	0,00
	rs2069526-C	0,07		rs55817950-A	0,00
	rs762551-C	0,38		rs776746-A	0,11
CYP2B6	rs28399499-G	0,00	CYP4F2	*3 rs2108622-A	0,27
CYP2C19	rs28399504-G	0,00	DPYD	*2A rs3918290-A	0,00
	rs41291556-G	0,00	Factor II	rs1799963-A	0,00
	rs4244285-A	0,14	Factor V Leiden	rs6025-A	0,05
	rs4986893-A	0,02	IFNL3	rs8099917-C	0,18
	rs56337013-A	0,00	ITGB3	rs5918-G	0,16
	rs6413438-A	0,00	MTFHR	rs1801131-C	0,35
CYP2C9	rs1057910-C	0,09		rs1801133-A	0,18
	rs28371685-A	0,00	PON1	rs662-G	0,28
	rs28371686-G	0,00	SLCO1B1	rs4149056-G	0,20
	rs56165452-G	0,00	TPMT	rs1142345-G	0,02
CYP2D6	rs28371725-A	0,00		rs1800460-A	0,01
	rs5030862-A	0,00	VKORC1	rs9923231-G	0,40
	rs59421388-A	0,00	VKORC1	rs9923231-T (A)	0,60

Заключение и выводы

1. Генофонд народонаселения Удмуртии впервые исследован по обширной полногеномной панели ДНК-маркеров. Пять популяций удмуртов и бесермян (геномы 28 индивидов) рассмотрены в широком контексте популяций европейской части России и Урала. В основной анализ включены все финно-угорские народы России (в анализ главных компонент также и народы зарубежья – венгры, финны, эстонцы). Они рассмотрены в широком контексте окружающих народов – тюркоязычных, славянских, балтоязычных, самодийских. Исследования, проведенные по обширной полногеномной панели и корректным выборкам, дают устойчивые результаты, которые уже существенно не изменятся и могут лишь корректироваться при введении новых выборок из других популяций.

2. При анализе всего народонаселения европейской части России и Урала методом ADMIXTURE, начиная с уровня четырех предковых компонент генофонда, ярко проявляется «удмуртская» компонента. Она составляет почти 100 % геномов всех удмуртов и бесермян. В других популяциях эта компонента встречается лишь у ближайших соседей удмуртов. А именно, при числе предковых компонент k=5 «удмуртская» компонента характерна также для коми-пермяков, в меньшей степени – для коми-зырян, башкир, татар, у остальных народов практически отсутствует. При числе предковых компонент k=6 «удмуртская» предковая компонента является в равной степени и «марийской» (составляя практически весь их генофонд). При этом у бесермян «удмуртская» предковая компонента снижается до трех четвертей генофонда за счет проявления «беломорской» предковой компоненты, которая составляет основную часть геномов коми-зырян и весь генофонд русских беломорского севера.

3. Публикуемые карты распространения «удмуртских» предковых компонент генофонда народов европейской части России и Урала дают возможность представителям гуманитарных наук самостоятельно анализировать генетические следы взаимодействия удмуртов и бесермян с другими народами региона.

4. Сравнение результатов анализа тремя независимыми методами – предковых компонент ADMIXTURE, генетических расстояний F_{ST} и главных компонент PCA – показывает, что по геномным данным удмурты обнаруживают наибольшее сходство с коми-пермяками. Второй круг генетического сходства включает марийцев и уральских тюрков. Хотя при анализе разными методами их вклад оценивается различно, но в наиболее общем виде можно очертить такую последовательность генетического сходства с удмуртами и бесермянами: коми-пермяки, марийцы, чуваши, часть популяций башкир, татары казанские.

5. Эти же закономерности характерны и для генофонда бесермян. Однако с тем отличием, что бесермяне генетически ближе ко всем перечисленным народам, чем популяции удмуртов.

6. Генетическое сходство между удмуртами, бесермянами, коми-пермяками, марийцами и уральскими тюрками особенно ярко проявляется на фоне их генетической отдаленности от остальных финно-угров России (карел, води, ижоры, хантов и манси) и зарубежья (финнов, эстонцев, венгров).

7. Для адаптации опубликованных фармакогенетических алгоритмов к применению в России необходимо получение данных о распространенности клинически значимых маркеров среди российских популяций. Нами дана характеристика фармакогенетического статуса населения Удмуртии. Показано, что фармакогенетические рекомендации, пригодные для персонифицированной медицины Удмуртии, могут охватывать коренное население Поволжья, Приуралья и Южного Урала, но не более отдаленные регионы.

Соблюдение этических стандартов: исследование одобрено Этическим комитетом Медико-генетического научного центра (протокол № 3/1 от 5 сентября 2018 г.), выполнено на образцах, полученных в ходе популяционно-генетического обследования генофонда; все обследуемые подписали добровольное информированное согласие.

Благодарности: Мы благодарим всех доноров образцов, которые принимали участие в данном исследовании, АНО «Биобанк Северной Евразии» за предоставление коллекций ДНК.

СПИСОК ЛИТЕРАТУРЫ

Аксянова Г. А. Антропология бесермян // Этнографическое обозрение. 2012. № 2. С. 100–117.

Аксянова Г. А. К антропологической характеристике южных удмуртов // Этнографическое обозрение. 2003. № 2. С. 110–126.

Балановская Е. В., Жабагин М. К., Балановский О. П. и др. Популяционные биобанки: принципы организации и перспективы применения в геногеографии и персонализированной медицине // Генетика. 2016. № 12. С. 1371–1387.

Балановская Е. В., Агджоян А. Т., Балановский О. П. и др. Татары Евразии: своеобразие генофондов крымских, поволжских и сибирских татар // Вестник Московского университета. Серия XXIII. "Антропология". 2016. №3. С. 72–82.

Балановская Е. В., Черневский Д. К., Балановский О. П. Своеобразие новгородского генофонда в контексте народонаселения Европейской части России // Ученые записки Новгородского государственного университета. 2021. В печати.

Балановский О. П. Генофонд Европы. Москва: Товарищество научных изданий КМК, 2015. 354 с.

Балановский О. П., Записецкая Ю. С., Балановская Е. В. и др. Взаимодействие генофондов русского и финноязычного населения Тверской области: анализ 4 млн SNP-маркеров // Вестник РГМУ. 2020. № 6. С. 24–31.

Белых С. К. Проблема распада прапермской этноязыковой общности. Ижевск: Удмуртский университет, 2009. 150 с.

Бермишева М. А., Гинтер Е. К., Хуснутдинова Э. К. и др. Популяционно-генетическое исследование популяции удмуртов (анализ десяти полиморфных ДНК-локусов ядерного генома) // Генетика. 2007. Т. 43, № 5. С. 688–705.

Грошева А. Н., Шнейдер Ю. В., Морозова И. Ю. и др. Генетическое разнообразие бесермян по данным о полиморфизме митохондриальной ДНК // Генетика. 2013. № 11. С. 1337–1344.

Грошева А.Н., Шнейдер Ю.В., Жукова О.В. и др. Особенности митохондриального генофонда удмуртов в связи с их родоплеменной структурой // Генетика. 2014. Т. 50. № 9. С. 1104–1104.

Дерябин В. Е. К этнической антропологии современных финских и тюркских народов восточноевропейской России // Народы России: от прошлого к настоящему. Антропология. 1998. С. 14–37.

Косарева И. А. Этнографические группы удмуртского народа (опыт выделения). Часть 1 // Известия Общества археологии, истории и этнографии при Казанском университете. 2017. Т. 37. №2. С. 195.

Кошель С. М. Геоинформационные технологии в геногеографии // Современная географическая картография. Москва: Дата+, 2012. С. 158–166.

Матвеев А. К. Субстратная топонимия Русского Севера. IV. Топонимия мерянского. Екатеринбург: Издательство Уральского университета, 2015. 313 с.

Матвеев А. К. Субстратная топонимия Русского Севера. Часть I. Екатеринбург, 2001. 277 с.

Матвеев А. К. Субстратная топонимия Русского Севера. Часть III. Екатеринбург, 2007.

Матвеев А. К. Субстратная топонимия Русского Севера. Часть II. Екатеринбург: Издательство Уральского университета, 2004. 21 с.

Напольских В. В. Очерки по этнической истории. Казань: Казанская недвижимость, 2015. 646 с.

Пислегин Н. В., Чураков В. С. Населенные пункты Удмуртского Прикамья XVII – середины XIX в. Ижевск: Удмуртский федеральный исследовательский центр Уральского отделения Российской академии наук, 2020. 784 с.

Садиков Р. Р. Поселения и жилища закамских удмуртов (материальный и духовный аспекты). Уфа: Гилем, 2001. 181 с.

Смирнов О. В. К вопросу о пермском топонимическом субстрате на территории Марий Эл и в бассейне среднего течения Вятки (в свете этнической интерпретации археологических культур). 1 // Вопросы ономастики. 2013. №2(15). С. 7–59.

Смирнов О. В. К вопросу о пермском топонимическом субстрате на территории Марий Эл и в бассейне среднего течения Вятки (в свете этнической интерпретации археологических культур). 2 // Вопросы ономастики. 2014. № 2 (16). С. 7–33.

Трофимова Н. В., Литвинов С. С Хуснутдинова Э.К. и др. Генетическая характеристика популяций Волго-Уральского региона по данным об изменчивости Y-хромосомы // Генетика, 2015. Т. 51. № 1. С. 122.

Чураков В. С. Каринский стан: история, территория, население // Исторические, философские, политические и юридические науки, культурология и искусствоведение. Вопросы теории и практики. 2017. № 12. Ч. 3. С. 215–219.

Чураков В. С. К критике воршудной теории // Финно-угроведение. 2003. № 2. С 3–18.

Широбоков И. Г. К антропологии пермских народов XVII–XIX вв. (краниологические данные) // Ежегодник финно-угорских исследований. 2014. № 3. С. 80–98.

Юсупов Ю. М., Схаляхо Р. А., Балановский О. П. и др. Генетические портреты семи кланов северо-западных башкир: вклад финно-угорского компонента в генофонд башкир // Вестник Московского университета. Серия XXIII. Антропология. 2017. № 3. С. 94–103.

Alexander D. H., Novembre J., Lange K. Fast model-based estimation of ancestry in unrelated individuals // Genome Research. 2009. V. 19. P. 1655–64.

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. no. 1. P. 13742–015.

Core ADME Gene List [Электронный ресурс]. URL: http://pharmaadme.org/joomla/index.php ?option=com_content&task=view&id=12&Itemid=27 (дата обращения 25.04.2021).

GG-base [Электронный ресурс]. URL: https://www.gg-base.org/ (дата обращения 25.04.2021).

Ilumäe A. M., Balanovska E., Balanovsky O. et al. Human Y Chromosome Haplogroup N: A Non-trivial Time-Resolved Phylogeography that Cuts across Language Families // Am J Hum Genet. 2016. V. 99. P. 163–173.

Jeong C., Balanovsky O., Balanovska E. et al. The genetic history of admixture across inner Eurasia // Nature ecology & evolution. 2019. V. 3. no. 6. P. 966–976.

Johnson J. A., Caudle K. E., Gong L. et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for pharmacogenetics-guided warfarin dosing: 2017 update // Clinical Pharmacology & Therapeutics. 2017. V. 102. no. 3. P. 397–404.

Lazaridis I., Balanovska E., Balanovsky O. et al. Ancient human genomes suggest three ancestral populations for present-day Europeans // Nature. 2014. V. 513. no. 7518. P. 409–413.

Lima J. J., Thomas C. D., Barbarino J. et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for CYP2C19 and proton pump inhibitor dosing // Clinical Pharmacology & Therapeutics. 2020.

Patterson N., Moorjani P. et al. Ancient admixture in human history // Genetics. 2012. Vol. 192, no. 3. P. 1065–1093.

Patterson N., Price A. L., Reich D. Population structure and eigenanalysis // PLoS genet. 2006. Vol. 2, no. 12. P. 190.

PharmGKB [Электронный ресурс]. URL: https://www.pharmgkb.org (дата обращения 25.04.2021).

Scott S. A., Sangkuhl K., Stein C. M. et al. Clinical Pharmacogenetics Implementation Consortium guidelines for CYP2C19 genotype and clopidogrel therapy: 2013 update //Clinical Pharmacology & Therapeutics. 2013. Vol. 94, no. 3. P. 317–323.

Tambets K., Balanovsky O., Balanovska E. et al. Genes reveal traces of common recent demographic history for most of the Uralic-speaking populations // Genome biology. 2018. V. 19. no. 1. P. 1–20.

Tambets K., Balanovsky O., Balanovska E. et al. The western and eastern roots of the Saami—the story of genetic "outliers" told by mitochondrial DNA and Y chromosomes // The American Journal of Human Genetics. 2004. T. 74. Vol. 4. P. 661–682.

Triska P., Chekanov N., Prokhortchouk E.. et al. Between Lake Baikal and the Baltic Sea: genomic history of the gateway to Europe. BMC genetics. 2017. V. 18. no. 1. P. 5–20.

Waskom M. L. Seaborn: statistical data visualization. Journal of Open Source Software. 2021. V. 6. no. 60. P. 3021.

Поступила в редакцию 20.07.2021

Балановская Елена Владимировна,
доктор биологических наук, профессор,
ФГБНУ «Медико-генетический научный центр им. акад. Н. П. Бочкова»,
115478, Россия, г. Москва, ул. Москворечье, д. 1
E-mail: balanovska@mail.ru

Напольских Владимир Владимирович,
член-корреспондент РАН, доктор исторических наук, профессор,
лаборатория античной культуры ШАГИ
Российской академии народного хозяйства и государственной службы при Президенте РФ,
119571, г. Москва, проспект Вернадского, д. 82, 84
E-mail: vovia@udm.ru

Чураков Владимир Сергеевич
кандидат исторических наук, старший научный сотрудник,
Удмуртский институт истории, языка и литературы УдмФИЦ УрО РАН,
426067, Россия, г. Ижевск, ул. Ломоносова, 4
E-mail: vermis3@gmail.com

Пислегин Николай Викторович
кандидат исторических наук, старший научный сотрудник,
Удмуртский институт истории, языка и литературы УдмФИЦ УрО РАН,
426067, Россия, г. Ижевск, ул. Ломоносова,4
E-mail: cpeg@rambler.ru

Записецкая Юлия Сергеевна,
стажер лаборатории популяционной генетики человека,
ФГБНУ «Медико-генетический научный центр им. акад. Н. П. Бочкова»,
115478, Россия, г. Москва, ул. Москворечье, 1
E-mail: zapisetskaya.yus@phystech.edu

Черневский Денис Константинович,
научный сотрудник лаборатории популяционной генетики человека,
ФГБНУ «Медико-генетический научный центр им. акад. Н. П. Бочкова»,
115478, Россия, г. Москва, ул. Москворечье, 1
E-mail: dk_ch@mail.ru

Балановский Олег Павлович,
доктор биологических наук, профессор РАН,
ФГБУН Институт общей генетики им. Н. И. Вавилова РАН,
119991, Россия, Москва, ул. Губкина, 3
E-mail: balanovsky@inbox.ru

E. V. Balanovska, V.V. Napolskikh, V.S. Churakov, N.V. Pislegin, Yu. S. Zapisetskaya, D. K. Chernevskiy, O. P. Balanovsky

GENE POOLS OF UDMURTS AND BESERMYANS IN THE CONTEXT OF FINNO-UGRIC AND OTHER NEIGHBORING ETHNIC GROUPS: THE GENOME-WIDE AND PHARMACOGENETIC DATA

DOI: 10.35634/2224-9443-2022-16-2-328-346

The indigenous populations of Udmurtia (five populations of Udmurts and Besermyans) have been studied for the first time by the genome-wide array. The populations were studied in the context of all Finno-Ugric and other surrounding ethnic groups. The total analyzed dataset included 728 individual genomes from 76 populations.
The ADMIXTURE analysis of ancestral components identified the specific "Udmurt" component comprising nearby 100 % of genomes of all studied Udmurt individuals and the lion's share of genomes of the studied Besermyans. The second component in Besermyan was the "White Sea" one predominating in Komi Zyryans and northernmost populations of ethnic Russians. The three independent methods – ADMIXTURE, F_{ST}, and PCA – demonstrated that Udmurts and Besermyans are genetically closest to Komi-Permyaks, followed by Mari, Chuvash, Bashkirs, and Volga Tatars.
We also characterized the pharmacogenetic status of Udmurtia population by estimating frequencies of 45 pharmacogenetic markers. The cartographic analysis revealed that genetic markers used in pharmacogenetic recommendations were observed in Besermyan and Udmurt populations at frequencies close to the frequencies in indigenous groups from Volga-Ural region but not from the more remote regions. Thus, within the revealed area of genetic similarity the same pharmacogenetic protocols can be applied.
Here we publish the maps of distribution of the "Udmurt" ancestral components in the area of European Russia and Ural. The maps can be used by scholars in humanitarian fields in their own studies to trace genetically the interactions of Udmurts and Besermyans with other ethnic groups.

Keywords: Udmurts, Besermyans, gene pool, gene geography, genome-wide arrays, Udmurtia, pharmacogenetics, Finno-Ugric populations, genetic population history, Volga-Ural region.

Citation: Yearbook of Finno-Ugric Studies, 2022, vol. 16, issue 2. P. 328–346. In Russian.

REFERENCES

Aksyanova G. A. *Antropologiya besermyan* [Anthropology of Besermyan] *Etnograficheskoe obozrenie* [Ethnographic review], 2012, no, 2. P. 100–117. In Russian.

Aksyanova G. A. *K antropologicheskoi kharakteristike yuzhnykh udmurtov* [Anthropological characteristics of the southern Udmurts]. *Etnograficheskoe obozrenie* [Ethnographic review], 2003, no. 2. P. 110–126. In Russian.

Balanovska E. V., Zhabagin M. K., Balanovsky O. P. et al. *Populyatsionnye biobanki: printsipy organizatsii i perspektivy primeneniya v genogeografii i personalizirovannoi meditsine* [Population biobanks: principles of organization and prospects of application in genogeography and personalized medicine] // *Genetika* [Genetics], 2016, no. 12. P. 1371–1387. In Russian.

Balanovska E. V., Agdzhoyan A. T., Balanovsky O. P. et al. *Tatary Evrazii: svoeobrazie genofondov krymskikh, povolzhskikh i sibirskikh tatar* [Tatars of Eurasia: the originality of the gene pools of the Crimean, Volga and Siberian Tatars]. *Vestnik Moskovskogo universiteta. Seriya XXIII. "Antropologiya"* [Bulletin of Moscow University. Series XXIII. "Anthropology"], 2016, no.3. P. 72–82. In Russian.

Balanovska E. V., Chernevskiy D. K., Balanovsky O. P. *Svoeobrazie novgorodskogo genofonda v kontekste narodonaseleniya Evropeiskoi ashki Rossii* [The originality of the Novgorod gene pool in the context of the population of the European part of Russia]. *Uchenye zapiski Novgorodskogo gosudarstvennogo universiteta* [Memoirs of NovSU], 2021. In print. In Russian.

Balanovsky O. P. *Genofond Evropy* [The gene pool of Europe.]. Moscow: Tovarishchestvo nauchnykh izdanii KMK Publ., 2015. 354 p. In Russian.

Balanovsky O. P., Gorin I. O., Balanovska E. V. et al. *Vzaimodeistvie genofondov russkogo i finnoyazychnogo naseleniya Tverskoi oblasti: analiz 4 mln SNP-markerov* [Interaction of gene pools of the Russian and Finnish-speaking population of the Tver region: analysis of 4 million SNP markers]. *Vestnik RGMU* [Bulletin of the Russian State Medical University], 2020, no. 6. P. 24–31. In Russian.

Belykh S. K. *Problema raspada prapermskoi etnoyazykovoi obshchnosti* [The problem of the collapse of the pre-Permian ethno-linguistic community]. Izhevsk: Udmurtskii gosudarstvennyi universitet Publ., 2009. 149 p. In Russian.

Bermisheva M. A., Ginter E. K., Khusnutdinova E. K. et al. *Populyatsionno-geneticheskoe issledovanie populyatsii udmurtov (analiz desyati polimorfnykh DNK-lokusov yadernogo genoma)* [Population genetic study of the Udmurt population (analysis of ten polymorphic DNA loci of the nuclear genome)]. *Genetika* [Genetics], 2007, vol. 43, no. 5. P. 688–705. In Russian.

Grosheva A. N., Shneider Yu. V., Morozova I. Yu., et al. *Geneticheskoe raznoobrazie besermyan po dannym o polimorfizme mitokhondrial'noi DNK* [Genetic diversity of besermyans according to data on mitochondrial DNA polymorphism]. Genetika [Genetics], 2013, no. 11. P. 1337–1344. In Russian.

Grosheva A. N., Shneider Yu.V., Zhukova O. V., et al. *Osobennosti mitokhondrial'nogo genofonda udmurtov v svyazi s ikh rodoplemennoi strukturoi* [Features of the mitochondrial gene pool of Udmurts in connection with their tribal structure]. *Genetika* [Genetics], 2014, vol. 50, no. 9. P. 1104–1104. In Russian.

Deryabin V. E. *K etnicheskoi antropologii sovremennykh finskikh i tyurkskikh narodov vostochnoevropeiskoi Rossii* [On the ethnic anthropology of modern Finnish and Turkic peoples of East European Russia]. *Narody Rossii: ot proshlogo k nastoyashchemu. Antropologiya* [Peoples of Russia: from the past to the present. Anthropology], 1998. P. 14–37. In Russian.

Kosareva I. A. *Etnograficheskie gruppy udmurtskogo naroda (opyt vydeleniya). Chast' 1* [Ethnographic groups of the Udmurt people (isolation experience). Part 1]. *Izvestiya Obshchestva arkheologii, istorii i etnografii pri Kazanskom universitete* [News of the Society of Archeology, History and Ethnography at Kazan University], 2017, vol. 37, no. 2. 195 p. In Russian.

Koshel' S. M. *Geoinformatsionnye tekhnologii v genogeografii* [Geoinformation technologies in genogeography]. *Sovremennaya geograficheskaya kartografiya* [Modern geographical cartography]. Moscow: Data+ Publ., 2012. P. 158–166. In Russian.

Matveev A. K. *Substratnaya toponimiya Russkogo Severa. IV. Toponimiya meryanskogo* [Substratum toponymy of the Russian North. IV. Meryanskiy toponymy]. Yekaterinburg: Izdatel'stvo Ural'skogo universiteta Publ., 2015. 313 p. In Russian.

Matveev A. K. *Substratnaya toponimiya Russkogo Severa. Chast' I* [Substratum toponymy of the Russian North. Part I.]. Yekaterinburg, 2001. 277 p. In Russian.

Matveev A. K. *Substratnaya toponimiya Russkogo Severa. Chast' III* [Substratum toponymy of the Russian North. Part III.]. Yekaterinburg, 2007. In Russian.

Matveev, A. K. *Substratnaya toponimiya Russkogo Severa. Chast' II* [Substratum toponymy of the Russian North. Part II]. Yekaterinburg : Izdatel'stvo Ural'skogo universiteta Publ., 2004. 21 p. In Russian.

Napol'skikh, V. V. *Ocherki po etnicheskoi istorii* [Essays on Ethnic History]. Kazan : Kazanskaya nedvizhimost' Publ., 2015. 646 p. In Russian.

Pislegin N. V., Churakov V. S. *Naselennye punkty Udmurtskogo Prikam'ya XVII – serediny XIX v.* [Settlements of the Udmurt Prikamye of the 17th – mid-19th centuries]. Izhevsk: Udmurtskii federal'nyi issledovatel'skii tsentr Ural'skogo otdeleniya Rossiiskoi akademii nauk Publ., 2020, 784 p. In Russian.

Sadikov R. R. Poseleniya i zhilishcha zakamskikh udmurtov (material'nyi i dukhovnyi aspekty) [Settlements and dwellings of the Trans-Kama Udmurts (material and spiritual aspects)]. Ufa: Gilem Publ, 2001, 181 p. In Russian.

Smirnov O. V. *K voprosu o permskom toponimicheskom substrate na territorii Marii El i v basseine srednego techeniya Vyatki (v svete etnicheskoi interpretatsii arkheologicheskikh kul'tur). 1* [On the question of the Permian toponymic substratum on the territory of Mari El and in the basin of the middle reaches of the Vyatka (in the light of the ethnic interpretation of archaeological cultures). 1]. *Voprosy onomastiki* [Onomastics issues], 2013, no. 2(15). P. 7–59. In Russian.

Smirnov O. V. *K voprosu o permskom toponimicheskom substrate na territorii Marii El i v basseine srednego techeniya Vyatki (v svete etnicheskoi interpretatsii arkheologicheskikh kul'tur). 2* [On the question of the Permian toponymic substratum on the territory of Mari El and in the basin of the middle reaches of the Vyatka (in the light of the ethnic interpretation of archaeological cultures). 2] // *Voprosy onomastiki* [Onomastics issues], 2014, no. 2(16). P. 7–33. In Russian.

Trofimova, N. V., Litvinov, S. S., Khusnutdinova, E. K et al. *Geneticheskaya kharakteristika populyatsii Volgo-Ural'skogo regiona po dannym ob izmenchivosti Y-khromosomy* [Genetic characteristics of the populations of the Volga-Ural region according to the data on the variability of the Y chromosome]. *Genetika* [Genetics], 2015, vol. 51, no. 1. 122 p. In Russian.

Churakov V. S. *Karinskii stan: istoriya, territoriya, naselenie* [Karinsky Stan: history, territory, population]. *Istoricheskie, filosofskie, politicheskie i yuridicheskie nauki, kul'turologiya i iskusstvovedenie. Voprosy teorii i praktiki* [Historical, philosophical, political and legal sciences, cultural studies and art history. Questions of theory and practice], 2017, no. 12. Part. 3. P. 215–219. In Russian.

Churakov V.S. *K kritike vorshudnoi teorii* [To the criticism of the vorshudnaya theory]. *Finno-ugrovedenie* [Finno-Ugric Studies], 2003, no. 2. P. 3–18. In Russian.

Shirobokov I. G. *K antropologii permskikh narodov XVII–XIX vv. (kraniologicheskie dannye)* [On the anthropology of the Permian peoples of the 17th–19th centuries. (craniological data)]. *Ezhegodnik finno-ugorskikh issledovanii* [Yearbook of Finno-Ugric Studies], 2014, no. 3. P. 80–98. In Russian.

Yusupov Yu. M., Skhalyakho R. A., Balanovsky O. P. et al. *Geneticheskie portrety semi klanov severo-zapadnykh ashkir: vklad finno-ugorskogo komponenta v genofond ashkir* [Genetic portraits of the seven clans of the northwestern Bashkirs: the contribution of the Finno-Ugric component to the Bashkir gene pool]. *Vestnik Moskovskogo universiteta. Seriya XXIII. Antropologiya* [Bulletin of Moscow University. Series XXIII. Anthropology], 2017, no. 3. P. 94–103. In Russian.

Alexander D. H., Novembre J., Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Research, 2009, vol. 19. P. 1655–64.

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, vol. 4, no. 1. P. 13742–015.

Core ADME Gene List [internet]. URL: http://pharmaadme.org/joomla/index.php?option=com_content &task=view&id=12&Itemid=27 (accessed 25.04.2021).

GG-base [internet]. URL: https://www.gg-base.org/ (accessed 25.04.2021).

Ilumäe A. M., Balanovska E., Balanovsky O. et al. Human Y Chromosome Haplogroup N: A Non-trivial Time-Resolved Phylogeography that Cuts across Language Families. Am J Hum Genet, 2016, vol. 99. P. 163–173.

Jeong C., Balanovsky O., Balanovska E. et al. The genetic history of admixture across inner Eurasia. Nature ecology & evolution, 2019, vol. 3, no. 6. P. 966–976.

Johnson J. A., Caudle K. E., Gong L. et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for pharmacogenetics-guided warfarin dosing: 2017 update. Clinical Pharmacology & Therapeutics, 2017, vol. 102, no. 3. P. 397–404.

Lazaridis I., Balanovska E., Balanovsky O. et al. Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature, 2014, vol. 513, no. 7518. P. 409–413.

Lima J.J., Thomas C.D., Barbarino J. Et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for CYP2C19 and proton pump inhibitor dosing. Clinical Pharmacology & Therapeutics, 2020.

Patterson N., Moorjani P., Luo Y. et al. Ancient admixture in human history. Genetics. 2012, vol. 192, no. 3. P. 1065–1093.

Patterson N., Price A. L., Reich D. Population structure and eigenanalysis. PloS genet, 2006, vol. 2, no. 12. P. 190.

Pharm GKB [internet]. URL: https://www.pharmgkb.org (accessed 25.04.2021).

Scott S. A., Sangkuhl K., Stein C. M. Et al. Clinical Pharmacogenetics Implementation Consortium guidelines for CYP2C19 genotype and clopidogrel therapy: 2013 update. Clinical Pharmacology & Therapeutics, 2013, vol. 94, no. 3. P. 317–323.

Tambets K., Balanovsky O., Balanovska E. et al. Genes reveal traces of common recent demographic history for most of the Uralic-speaking populations. Genome biology, 2018, vol. 19, no. 1. P. 1–20.

Tambets K., Balanovsky O., Balanovska E. et al. The western and eastern roots of the Saami—the story of genetic "outliers" told by mitochondrial DNA and Y chromosomes. The American Journal of Human Genetics, 2004, vol. 74, no. 4. P. 661–682.

Triska P., Balanovsky O., Prokhortchouk E. et al. Between Lake Baikal and the Baltic Sea: genomic history of the gateway to Europe. BMC genetics, 2017, vol. 18, no. 1. P. 5–20.

Waskom M. L. Seaborn: statistical data visualization. Journal of Open Source Software, 2021, vol. 6, no. 60, 3021 p.

Received 07.20.2021

Balanovska Elena Vladimirovna,
Doctor of Biological Sciences, Professor,
Head of the Laboratory of Human Population Genetics
Research Centre for Medical Genetics,
115478, Russia, Moscow, Moskvorechye st., 1
E-mail: balanovska@mail.ru

Napolskikh Vladimir Vladimirovich,
Corresponding Member of the Russian Academy of Sciences,
Doctor of Historical Sciences, Professor,
The Russian Presidential Academy of National Economy and Public Administration (RANEPA)
119571, Moscow, Vernadsky Avenue, 82, 84
E-mail: vovia@udm.ru

Churakov Vladimir Sergeevich,
Candidate of Historical Sciences, Senior Researcher,
Udmurt institute of History, Language and Literature UdmFRC UB RAS
4, Lomonosova st., Izhevsk, 426004, Russian Federation
E-mail: vermis3@gmail.com

Pislegin Nikolay Viktorovich,
Candidate of Historical Sciences, Associate Professor,
Udmurt institute of History, Language and Literature UdmFRC UB RAS,
4, Lomonosova st., Izhevsk, 426004, Russian Federation
E-mail: cpeg@rambler.ru

Zapisetskaya Julia Sergeevna,
Trainee, Laboratory of Human Population Genetics
Research Centre for Medical Genetics,
115478, Russia, Moscow, Moskvorechye st., 1
E-mail: zapisetskaya.yus@phystech.edu

Chernevskiy Denis Konstantinovich,
Researcher, Laboratory of Human Population Genetics
Research Centre for Medical Genetics,
115478, Russia, Moscow, Moskvorechye st., 1
E-mail: dk_ch@mail.ru

Balanovsky Oleg Pavlovich,
Doctor of Biological Sciences, Professor of the Russian Academy of Sciences,
Head of the Laboratory of Genomic Geography
Vavilov Institute of General Genetics Russian Academy of Sciences,
119991, Russia, Moscow, st. Gubkina, 3
E-mail: balanovsky@inbox.ru