

## КОМПОНЕНТНЫЙ СОСТАВ ГЕНОФОНДА КОРЕННЫХ ЭТНОСОВ ДАГЕСТАНА

© 2024 г. М. О. Раджабов<sup>1</sup>, \*, В. Н. Харьков<sup>2</sup>, Н. А. Колесников<sup>2</sup>,  
А. А. Зарубин<sup>2</sup>, И. Ю. Хитринская<sup>2</sup>, М. А. Джаубермезов<sup>3,4</sup>, О. М. Раджабов<sup>5</sup>,  
Э. К. Хуснутдинова<sup>3,4</sup>, В. А. Степанов<sup>2</sup>

<sup>1</sup>ФГБУН Дагестанский федеральный исследовательский центр РАН, Махачкала, 367000 Россия

<sup>2</sup>НИИ медицинской генетики ФГБНУ Томский национальный исследовательский центр РАН, Томск, 634000 Россия

<sup>3</sup>Институт биохимии и генетики ОСП ФГБНУ УФИЦ РАН, Уфа, 450054 Россия

<sup>4</sup>ФГБОУ ВО Уфимский университет науки и технологий, Уфа, 450076 Россия

<sup>5</sup>Научный центр генетики и наук о жизни Научно-технологического университета “Сириус”,  
федеральная территория “Сириус”, 354340 Россия

\*e-mail: radjabov\_m@mail.ru

Поступила в редакцию 22.03.2024 г.

После доработки 14.06.2024 г.

Принята к публикации 28.10.2024 г.

Проведен компонентный анализ состава генофондов коренных этносов Дагестана, а также других этносов Северного Кавказа и Центральной Азии на основе широкогеномного генотипирования SNP-маркеров на высокоплотных ДНК-микрочипах Infinium Multi-Ethnic Global-8. Анализ массива генотипических данных этнических популяций упомянутых выше регионов реализован в алгоритме Admixture (K = 12). Несмотря на то, что в этно-лингвистическом плане мегрелы относятся к народности картвельской языковой семьи, абхазы, адыги и черкесы принадлежат к группе этносов абхазо-адыгских языков, осетины, являются ираноязычным этносом, а карачаевцы и балкарцы представляют тюркскую группу народов Северного Кавказа, все перечисленные этносы имеют спектр одних и тех же предковых компонент, к тому же в сопоставимых пропорциях. Ингуши и чеченцы, будучи представителями нахской группы нахско-дагестанской языковой семьи, демонстрируют одну общую предковую компоненту с высокой степенью представленности, подчеркивающую обособленную генетическую историю формирования их генофонда. Караногайцы Дагестана и ногайцы из Карачаево-Черкесии по спектру и частоте представленности предковых компонент указывают на общность с центрально-азиатскими этносами, такими как казахи и узбеки. Этносты Дагестана демонстрируют уникальную картину разнообразия как состава, так и частоты встречаемости предковых компонент в структуре их аутосомного генофонда. Это подчеркивает не только обособленную от других народов Северного Кавказа историю их формирования, но и длительно изолированную существенную генетическую историю каждого коренного этноса. Исследование находится в рамках фундаментальных направлений популяционной и эволюционной генетики человека – выявления этноспецифических особенностей структуры геномов и популяционных генофондов современных популяций человека, реконструкции древних миграций и экспансии численности населения, а также эволюции генофондов и механизмов генетической адаптации.

**Ключевые слова:** генофонд, коренной этнос, Дагестан, предковые компоненты, ДНК-микрочип, алгоритм Admixture

**DOI:** 10.56304/S0234275824060127

### ВВЕДЕНИЕ

Исследования структур генофондов этносов и генетического разнообразия в популяциях на основе анализа геномов отдельных индивидов создают возможность решения ряда задач, в том числе: фундаментальных вопросов этногенеза народов методами эволюционной генетики и популяционной геномики, формирования основ фундаментальной и прикладной медицины путем изучения

генетики мультифакториальных и наследственных заболеваний в этнических популяциях, а также фармакогенетики, создания референсного банка геномов для идентификации биологических образцов в судебно-медицинской экспертизе и ДНК-идентификации [1].

Республика Дагестан это уникальный субъект Российской Федерации в составе Северо-Кавказского федерального округа как в плане этно-

культурного и лингвистического многообразия ее населения (более 30 автохтонных коренных этносов), так и в плане генетического разнообразия коренных этносов, демографических процессов и микроэволюционных событий, обуславливающих формирование структур их генофондов, основными из которых выступают система популяционных изолятов, эффект основателя и генетический дрейф в локальных популяциях, высокий показатель коэффициента инбридинга в масштабе всего мира [2–4].

Дагестан, занимая Северо-Восточную часть Кавказа, мог выступать с одной стороны местом, обеспечивающим генетические контакты между народами Европейской и Азиатской частей Евразийского материка, а с другой как горная система, выступающая барьером для миграций, мог стать резервуаром, сохранившим до сегодняшних дней древнейшие генетические пласты в генофонде его автохтонного населения. Обнаружение стоянки древнего человека раннеплейстоценового времени (до 2 млн лет) на территории Центрального Дагестана (Акушинский район, стоянка Айникаб I), однозначно свидетельствует о привлекательности данной территории для заселения человеком даже с раннего палеолита [5].

Учитывая природный ландшафт (горная система), обуславливающий систему популяционных изолятов, географическое положение (от отрогов Большого Кавказского хребта до западного берега Каспийского моря, включая узкую прибрежную полосу, называемую Кавказскими воротами), этно-лингвистическое многообразие населения, Дагестан выступает одной из ключевых историко-этнографических провинций современной России для проведения геномных исследований.

В лингвистическом отношении коренные этносы Дагестана формально представлены двумя языковыми семьями: дагестанской ветвью нахско-дагестанской, а также кипчакской группой тюркских языков алтайской языковой семьи [6].

В первую языковую семью входят:

а) языки аваро-андо-цезские, которые включают собственно аварский язык, андийские (андийский, ботлихский, годоберинский, каратинский, ахвахский, тиндинский, багвалинский, чамалинский) и цезские (цезский, хваршинский, гинухский, бежтинский, гунзибский);

б) даргинский язык (группа языков, включающая, в частности, урахинский, акушинский, кайтагский, кадарский, губденский, сирагинский, кубачинский и, возможно, другие идиомы, квалифицируемые обычно в качестве диалектов);

в) лакский язык;

г) лезгинские языки (лезгинский, табасаранский, агульский, рутульский, цахурский, арчинский, крызский, будухский, удинский);

д) хиналугский.

Здесь следует уточнить, что местом компактного проживания кызов, будухцев, удинов и хиналугов – лезгиноязычной группы этносов, выступает территория современного Азербайджана, а удины, помимо Азербайджана, живут также в России, Украине, Казахстане, Грузии, Армении и в других странах.

Нахская ветвь данной языковой семьи в Дагестане представлена чеченцами-аккинцами, которые сегодня проживают на территориях Новолакского, Хасавюртовского, Бабаюртовского и Казбековского районов современного центрального Дагестана. Вторая языковая семья представлена кумыками, ногайцами и “азербайджанцами” Дагестана.

Популяция является основной эволюционной единицей. Именно особенности структуры локальных популяций определяют состав генофонда этносов и региональных групп населения. Структура генофонда во многом зависит от особенностей эволюционной истории тех или иных популяций и имеет популяционно-специфический характер. Изучение изменчивости генетических маркеров на уровне отдельных популяций, этносов, этнических групп с целью выявления структуры генофондов локальных популяций и филогенетических взаимоотношений между ними является одной из основных задач этногенетики. До недавнего времени большинство этногеномных исследований были направлены на воссоздание картины расселения человека по различным континентам в глобальном масштабе. В настоящее время все большую актуальность приобретают работы, связанные с детализацией эволюционной и демографической истории отдельных регионов или этнорасовых групп [12–14]. Целью исследования является анализ генофондов коренных народов Дагестана, Северного Кавказа и оценка вклада предковых групп населения у разных популяций, а также роль изоляции субэтнических групп от контактов с другими народами.

## УСЛОВИЯ ЭКСПЕРИМЕНТА

Появившиеся биоинформатические методы анализа, а также развитие вычислительной техники дают возможность интерпретировать большие объемы экспериментальных данных и позволяют более детально анализировать картину:

- – структуры генофондов и взаимодействий популяций;

- – генетико-демографических процессов, которые имели место в прошлом, и достаточно точно датировать по времени эти события;
- – действия естественного отбора;
- – реконструкции миграций.

Современные методы исследования генетической структуры основаны на анализе сотен тысяч SNP-маркеров (Single Nucleotide Polymorphism) по всему геному.

Используемый алгоритм кластеризации Admixture [15] метод является одним из наиболее распространенных для анализа аутосомных маркеров. Он позволяет определить генетическую структуру популяций человека, установить степень генетического разнообразия и уровень дифференциации этносов, дает возможность реконструировать генетическую историю, миграции и расселение современного человека.

Программа Admixture используется для идентификации потенциальных предковых популяций, которые внесли свои компоненты в геномы современных народов. Полногеномные данные и высокоплотные панели SNP позволяют на уровне индивидуальных геномов и популяций оценить вклад элементов разного генетического происхождения. Этот метод позволяет анализировать одни и те же данные на разных иерархических уровнях – от индивидуального генома до генофонда человечества.

На основе широкогеномного генотипирования SNP имеется возможность детально изучить паттерны гаплотипического разнообразия, маркирующие различные субстратные и суперстратные пласты генофонда популяций, степень метисации с пришлым населением на различных уровнях – от индивидуального до родового и этнического, а также провести детальный анализ эволюционной и демографической истории этносов.

Алгоритм Admixture позволяет провести контроль качества, оценивая схожесть рисунка компонент у представителей одной популяции, оценить близость популяций по порядку и особенности комбинаций составляющих при постепенном их увеличении, выявить метисацию отдельных индивидов для дальнейшей фильтрации данных.

При биоинформатической обработке полученных данных из анализа были исключены SNP с более чем 10% отсутствующими генотипами. Были исключены все позиции делеций и инсерций. Данные были предварительно отфильтрованы по минимальной частоте редкого аллеля (MAF, minor allele frequency >0.01). В итоге, после фильтрации исходных данных по 1779819 маркерам, в финальный массив данных вошло 886889 аутосомных SNP. Были исключены образцы, имеющие более 5% пропущенных SNP, а также образцы ме-

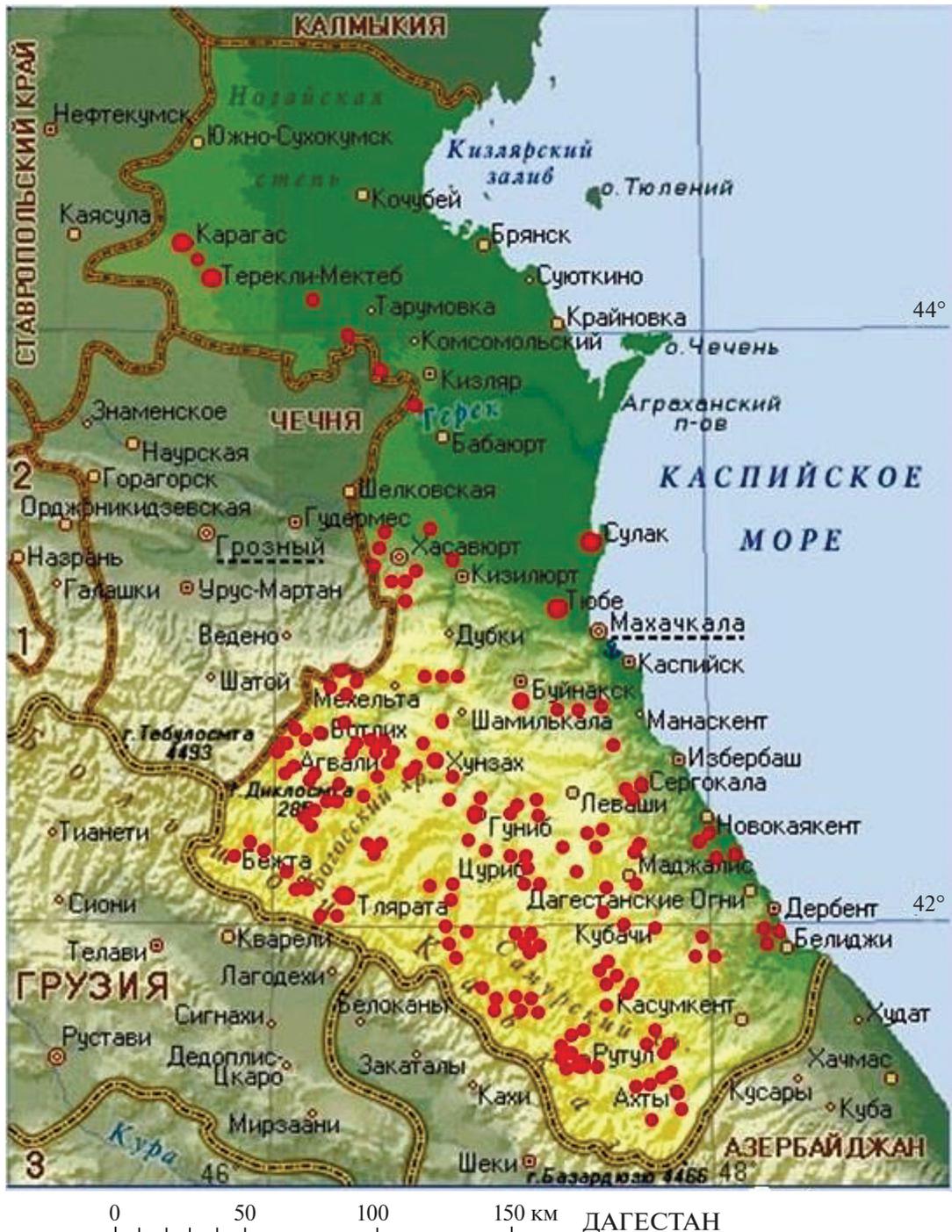
тисов первого, второго и третьего поколения, выявленных с помощью методов определения общих IBD (Identical by descent), построения t-SNE (t-distributed Stochastic Neighbor Embedding) и Admixture. Данная фильтрация позволяет повысить качество выборки увеличив степень точности за счет исключения близкородственных индивидов.

Проведение исследования одобрено Комитетом по биомедицинской этике НИИ медицинской генетики Томского НИМЦ (№ 4 от 21 ноября, 2017). Экспертное заключение о возможности открытого опубликования данной статьи выдана комиссией ФГБОУ ВО «Дагестанский государственный университет» от 06 октября 2023 г.

В ходе экспедиций, организованных Дагестанским федеральным исследовательским центром РАН (г. Махачкала) в период с 2001 по 2022 годы под руководством к. б. н. М.О. Раджабова, и д. б. н. В.Н. Харькова (НИИ медицинской генетики Томского национального медицинского центра РАН), а также в рамках договоров о совместных научных исследованиях с другими российскими научными коллективами был набран биологический материал для анализа. В 72 локальных популяциях (аулах) собраны 518 образцов периферической венозной крови, принадлежащей представителям 37 этносов Северного Кавказа и Центральноазиатского региона, из которых 23 относились к коренным жителям Дагестана (рис. 1). При заборе образцов крови соблюдали процедуру информированного согласия доноров. На каждого донора составляли анкету с краткой родословной, указанием его национальности и мест рождения предков. Собранные образцы относили к различным этническим группам на основании национальности донора и его родителей.

Образцы крови отбирали в вакуумные пробирки Vacuette с КЗЭДТА (Greiner Bio-One, Австрия). После этого выделение геномной ДНК из лейкоцитов периферической крови проводили с помощью метода фенол-хромового экстракта [16]. Затем была проведена проверка концентрации и качества образцов ДНК с использованием флуориметра Qubit 4 (Thermo Fisher Scientific, США) и набора реагентов Qubit™ 1X dsDNA High Sensitivity (Invitrogen, США) и выполнено генотипирование образцов на биочипах.

Полногеномное генотипирование ДНК-маркеров было проведено с использованием платформы Infinium Multi-Ethnic Global BeadChip (Illumina). Данная платформа включает свыше 1.7 млн. отобранных экспертным анализом SNP-маркеров, что позволяет проводить идентификацию генетической ассоциации с широко распространенными и редкими фенотипами, с учетом генетической архитектуры популяций. Для кластеризации масси-



**Рис. 1** Локализация исследованных этнических популяций Дагестана. Красными точками на рисунке показаны экспедиционно обследованные нами более 200 аулов Дагестана.

**Fig. 1.** Localization of the study of ethnic populations of Dagestan. The red dots in the figure show more than 200 villages of Dagestan that we surveyed during the expedition.

ва генотипов и оценки качества генотипирования мы прибегли к протоколу, опубликованному в статье Guo Y. [17] с использованием GenomeStudio (модуль генотипирования v. 2.0.3), программного пакета, который Illumina разработала для анализа полученных геномных данных.

Экспериментальная часть была проведена на базе Центра коллективного пользования НИИ медицинской генетики Томского НИМЦ. Результаты широкомасштабного генотипирования включены в базу данных лаборатории эволюционной генетики НИИ медицинской генетики Томского

НИМЦ. В настоящей статье приводятся пилотные данные по компонентному составу генофондов населения Дагестана. Более детальный биоинформатический анализ данных в настоящее время осуществляется в указанной лаборатории с привлечением данных по более широкому спектру популяций других регионов.

Общий объем выборки неродственных между собой до четвертого поколения жителей автохтонного генезиса составляет около 7 тыс человек. Из них более 90% – совершеннолетние мужчины, поскольку на начальном этапе целью наших исследований было изучение структуры генофонда коренных этнических групп по системам однополовых маркеров генов (Y-хромосома и мтДНК (митохондриальная ДНК)), а как известно, представители мужского пола в своем кариотипе имеют и Y-хромосому, наследуемую по отцовской линии и мтДНК, наследуемую по матери, тогда как представители женского пола имеют лишь мтДНК.

В данном исследовании представлен, как и указывалось выше, материал по 23 коренным этносам Дагестана ( $N = 518$ ), который включает: аварцев ( $N = 24$ ), даргинцев ( $N = 28$ ), лакцев ( $N = 24$ ), табасаранов ( $N = 21$ ), лезгинов ( $N = 28$ ), агулов ( $N = 24$ ), рутульцев ( $N = 22$ ), цахуров ( $N = 24$ ), арчинцев ( $N = 24$ ), андийцев ( $N = 17$ ), ахвахцев ( $N = 24$ ), тиндинцев ( $N = 18$ ), каратинцев ( $N = 24$ ), багулалов ( $N = 23$ ), ботлихцев ( $N = 16$ ), хваршинов ( $N = 15$ ), цезов ( $N = 24$ ), бежтинцев ( $N = 22$ ), чамалинцев ( $N = 24$ ), гунзибцев ( $N = 17$ ), гинухцев ( $N = 19$ ), кумыков ( $N = 37$ ) и караногайцев ( $N = 19$ ).

Коренные популяции Кавказского региона (за исключением Дагестана) ( $N = 294$ ) включают: мегрелов ( $N = 28$ ), абхазов ( $N = 30$ ), адыгейцев ( $N = 10$ ), черкесов ( $N = 30$ ), карачаевцев ( $N = 22$ ), балкарцев ( $N = 50$ ), осетин (С – северные  $N = 30$ , Ю – юж-

ные  $N = 17$ ), ингушей ( $N = 30$ ), чеченцев ( $N = 27$ ) и ногайцев ( $N = 20$ ). Материал выборок Кавказского региона предоставлен сотрудниками Уфимского федерального исследовательского центра Российской академии наук (д. б. н. Э.К. Хуснутдиновой, к. б. н. С.С. Литвиновым).

Популяции Средней Азии ( $N = 77$ ) представлены этническими узбеками ( $N = 22$ ) и казахами (М – младший жуз  $N = 29$ , С – старший жуз  $N = 26$ ).

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

По результатам анализа массива данных с помощью алгоритмов Admixture установлено, что гипотеза о кавказской общности, выдвигаемая исследователями гуманитарных направлений, не совсем корректна по отношению к определенным этническим группам Северного Кавказа. Приходится констатировать, что на фоне гомогенности как по составу предполагаемых предковых компонент, так и долям их представленности в генофонде этносов Северного Кавказа и Центральной Азии у коренных жителей Дагестана компонентный состав и пропорции их представленности в генофонде популяций гетерогенны.

Неожиданным для нас оказалось, что субэтнические группы осетин – дигоцев и иронцев, будучи ираноязычным этносом, имеют схожие генетические профили с тюркоязычными этносами Северного Кавказа (карачаевцами и балкарцами), а также с абхазо-адыгской группой (адыги, абхазы и черкесы).

Кроме этого, мегрелы – субэтническая группа грузин, говорящая на языке, который входит в картвельскую языковую семью, также имеют идентичную с упомянутыми выше этносами Северного Кавказа генетическую платформу (рис. 2).

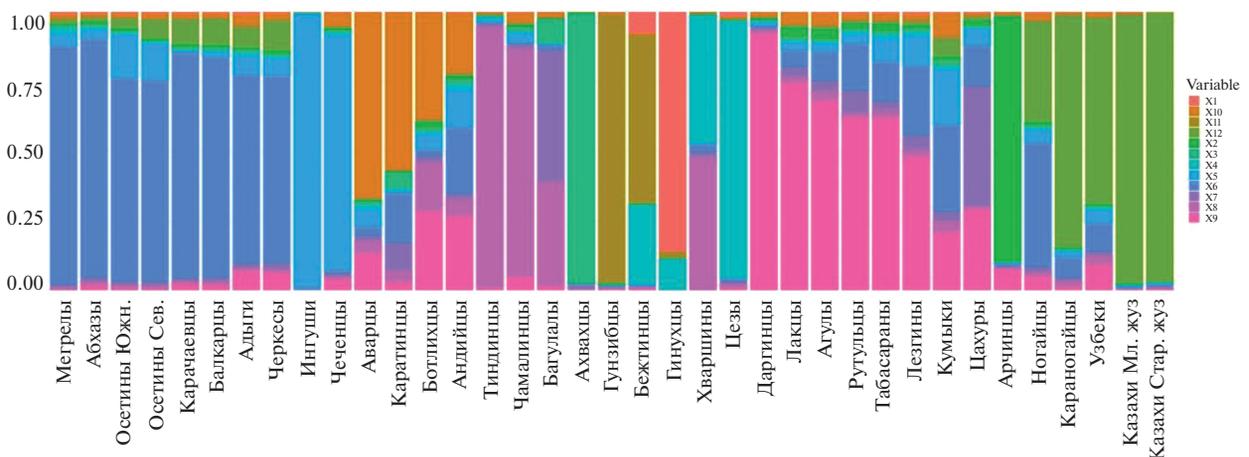


Рис. 2. Упорядоченная картина предковых компонент (K), рассчитанная по алгоритму Admixture ( $K = 12$ ).

Fig. 2. Ordered picture of ancestral components (K), calculated using the Admixture algorithm ( $K = 12$ ).

Интересен также факт, что чеченцы с ингушами (нахская ветвь нахско-дагестанской языковой семьи) по параметрам, формирующим их генофонд, стоят особняком как от дагестанских коренных, так и от других этносов Северного Кавказа. Следует также подчеркнуть, что внутри себя обе эти этнические группы монолитны, то есть одна предковая компонента, которая у ингушей встречается с частотой почти 100% представлена и у чеченцев на уровне 85% (табл. 1). Таким образом, говорить о разных истоках генезиса и обособленной генетической истории их формирования, как этносов, не приходится.

Тюркоязычные караногайцы Дагестана, в большей степени и ногайцы из Карачаево-Черкесской Республики, в меньшей, несмотря на географическую отдаленность демонстрируют идентичную с этносами Центральной Азии картину генофонда – с узбеками и казахами, где превалирует азиатская тюркская предковая компонента. То, что в генофонде карачаево-черкесских ногайцев ощутимо проявляется предковая компонента, свойственная для Северо-Западного Кавказа (абхазы, адыги, балкарцы, карачаевцы, мегрелы, осетины, черкесы), возможно, объясняется попаданием в изученную нами выборку потомков смешанных браков.

Что касается структуры генофонда коренных жителей Дагестана, наблюдается контрастная картина наличия отдельных этноспецифичных предковых компонент для множества этносов, в основном, для малочисленных коренных андоцезских подгрупп, проживающих в высокогорной зоне Республики. Выделение таких этноспецифичных компонент главным образом связано с эффектами основателя и сдвигом частот аллелей для этих популяций. Причиной этого служит их изоляция от контактов с другими группами людей, отсутствие браков между разными удаленными поселками и высокий уровень инбридинга в этих изолированных населенных пунктах. Это приводит к значительной специфичности частот аллелей и блоков сцепления на хромосомах в этих выборках.

В качестве наиболее ярких примеров специфичных микроэволюционных событий в популяциях выступают выборки ахвахцев, арчинцев, цезов (дидойцев), гунзибцев, гинухцев, тиндинцев, чамалинцев и даргинцев. В указанных этнических популяциях на высоком уровне 86–98% представлена доминирующая составляющая, при этом у каждой из перечисленных, своя отдельная обособленная компонента, кроме чамалинцев и тиндинцев, у которых предковая одна и та же.

Следует упомянуть, что чамалинцы и тиндинцы относятся к одной языковой подгруппе – андийской. Генофонды других этносов этой же языковой подгруппы: андийцев, ботлихцев, каратинцев имеют многокомпонентные, но в той или иной степени равные по долям структуры генофондов. Обращает внимание и то, что у последних встречается та же компонента, которая специфична и представлена на уровне 70% для самого многочисленного автохтонного этноса Дагестана – аварцев (численность их чуть более 1 млн. чел.). Ахвахцы из андийской группы позиционируются как обособленный монокомпонентный этнос, в противовес багулалам, которые представлены тремя доминирующими предковыми компонентами, одна из которых встречается, что и логично, также у некоторых других в андийской подгруппе – тиндинцев и чамалинцев, что подтверждает их генетическое родство.

Наиболее гетерогенной структурой генофонда, в плане этноспецифичности предковых компонент и по доле их представленности в генофонде, отличается цезская-дидойская языковая подгруппа этносов, которая относится к аваро-андоцезской группе дагестанской ветви языков. Дидойцы (цезы), гинухцы и гунзибцы представлены собственными специфичными составляющими, при этом их доля чрезвычайно специфична и находится в интервале 86–97%. Генофонд бежтинцев вообрал в себя три компонента, одна из которых встречается у гунзибцев в подавляющей степени (97%), другая у цезов-дидойцев на таком же уровне, и третья встречающаяся у гинухцев. Из этой языковой подгруппы только хваршины содержат в своей структуре, помимо компоненты, свойственной для цезов-дидойцев, бежтинцев и гинухцев, компоненту, маркирующую генофонды этносов андийской подгруппы языков – тиндинцев, чамалинцев и багулалов.

По результатам Admixture прорисовывается картина разделения внутри дидойской группы на западноцезскую и восточноцезскую подгруппы по наблюдаемым предковым компонентам: хваршины и цезы-дидойцы, с одной стороны, гунзибцы и бежтинцы, с другой, при этом у гинухцев, доля общей для тех и других компоненты 10% и своей собственной специфичной – 90%.

Разделение протоцезской группы на протозападноцезскую и протовосточную по лексикостатистическими подсчетам датируется V в. до н.э. Разделение протозападноцезской группы на гинухскую, протохваршинскую и цезскую датируется III в. н.э., а разделение протовосточноцезской на бежтинцев и гунзибцев IX в. н.э., в это же

**Таблица 1.** Доля предковых компонент ( $K = 12$ ) в популяциях  
**Table 1.** The proportion of ancestral components ( $K = 12$ ) in populations

Популяция	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12
Мегрелы	0.01	0.01	0.01	0.02	0.05	0.86	0.01	0.00	0.00	0.01	0.01	0.01
Абхазы	0.01	0.01	0.01	0.01	0.03	0.86	0.01	0.01	0.02	0.01	0.01	0.01
Осетины сев.	0.01	0.01	0.00	0.01	0.13	0.73	0.01	0.01	0.01	0.02	0.00	0.07
Осетины южн.	0.01	0.01	0.01	0.01	0.16	0.73	0.01	0.01	0.01	0.01	0.01	0.03
Карачаевцы	0.01	0.01	0.01	0.00	0.01	0.81	0.00	0.01	0.02	0.02	0.01	0.09
Балкарцы	0.01	0.01	0.01	0.00	0.02	0.79	0.01	0.01	0.02	0.01	0.01	0.09
Адыги	0.01	0.01	0.01	0.01	0.07	0.68	0.01	0.01	0.07	0.04	0.01	0.07
Черкесы	0.01	0.01	0.01	0.01	0.06	0.68	0.02	0.01	0.06	0.02	0.01	0.11
Ингуши	0.00	0.00	0.00	0.00	0.98	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Чеченцы	0.00	0.01	0.01	0.01	0.85	0.01	0.00	0.01	0.04	0.05	0.00	0.00
Аварцы	0.00	0.01	0.01	0.02	0.06	0.03	0.01	0.04	0.13	0.67	0.00	0.00
Каратинцы	0.00	0.00	0.06	0.01	0.01	0.18	0.10	0.04	0.03	0.57	0.00	0.00
Ботлихцы	0.00	0.02	0.01	0.01	0.06	0.02	0.02	0.18	0.28	0.38	0.00	0.00
Андийцы	0.01	0.01	0.01	0.02	0.14	0.24	0.02	0.06	0.27	0.21	0.01	0.00
Тиндинцы	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.94	0.01	0.00	0.01	0.00
Чамалалы	0.01	0.01	0.01	0.01	0.04	0.00	0.02	0.87	0.05	0.03	0.00	0.00
Багулалы	0.00	0.00	0.09	0.00	0.01	0.02	0.47	0.38	0.02	0.02	0.00	0.00
Ахвахцы	0.00	0.00	0.97	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00
Гунзибцы	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.01	0.97	0.00
Бежтинцы	0.08	0.00	0.00	0.29	0.00	0.00	0.00	0.00	0.01	0.00	0.61	0.00
Гинухцы	0.86	0.00	0.00	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00
Хваршины	0.00	0.00	0.00	0.46	0.00	0.03	0.01	0.48	0.00	0.01	0.00	0.00
Цезы	0.02	0.00	0.00	0.93	0.00	0.01	0.00	0.02	0.01	0.00	0.01	0.00
Даргинцы	0.00	0.01	0.01	0.00	0.02	0.00	0.01	0.01	0.92	0.02	0.00	0.00
Лакцы	0.00	0.03	0.01	0.01	0.03	0.06	0.03	0.02	0.75	0.04	0.00	0.01
Агулы	0.01	0.04	0.01	0.01	0.03	0.10	0.03	0.04	0.68	0.04	0.01	0.01
Рутульцы	0.01	0.02	0.01	0.01	0.03	0.17	0.08	0.02	0.62	0.02	0.01	0.00
Табасараны	0.01	0.03	0.01	0.01	0.09	0.14	0.03	0.03	0.62	0.02	0.01	0.00
Лезгины	0.01	0.01	0.01	0.01	0.11	0.25	0.05	0.03	0.48	0.02	0.01	0.01
Кумыки	0.01	0.01	0.02	0.01	0.20	0.31	0.03	0.04	0.21	0.08	0.01	0.06
Цахуры	0.01	0.02	0.00	0.01	0.06	0.15	0.43	0.01	0.29	0.01	0.01	0.01
Арчинцы	0.00	0.88	0.00	0.00	0.00	0.01	0.01	0.00	0.08	0.01	0.00	0.00
Ногайцы	0.01	0.01	0.01	0.01	0.05	0.45	0.01	0.02	0.05	0.02	0.00	0.36
Караногайцы	0.00	0.01	0.00	0.00	0.02	0.07	0.01	0.02	0.01	0.01	0.00	0.83
Узбеки	0.00	0.01	0.01	0.01	0.05	0.10	0.01	0.03	0.09	0.01	0.01	0.67
Казахи ст.жуз	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.01	0.00	0.00	0.97
Казахи мл.жуз	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.96

*Примечание.* Зеленым цветом выделены доминирующие в этносах предковые компоненты.  
*Note.* The ancestral components dominant in ethnic groups are highlighted in green.

время протохваршинская разделяется на хваршинскую и инхкваринскую [7].

Подавляющее доминирование одной предковой компоненты (97%) в структуре генофонда

даргинского автохтонного этноса Дагестана (626.6 тыс. человек), второго по численности населения после аварцев, представленного локальными популяциями, каждая из которых имеет

собственный диалект или говор в языке, показалось нам несколько неожиданным. Такая картина, вероятнее всего результат смещенной выборки. Учитывая относительно небольшое количество проанализированных образцов от каждого этноса (20–30 индивидов), в выборку для анализа могли попасть представители одного или нескольких, но родственных по происхождению аулов.

Следует сосредоточить внимание на преобладании общих и наиболее выраженных генетических компонент у таких коренных этнических популяций Дагестана как даргинцы, лакцы, агулы, рутульцы, табасараны, лезгины, объединяя их в единую центрально- и южнодагестанскую группу.

Сближение лакцев с даргинцами по компонентному составу и частоте представленности их в генофондах было предсказуемо, поскольку и по лингвистическим характеристикам эти этносы являются близкими между собой. При этом то, что представители лезгинской языковой подгруппы (собственно лезгины, агульцы, рутульцы, табасараны) окажутся ближе, чем аварцы и группа андо-цезских языков к даргинско-лакской паре, для нас было также неожиданно. Указанные этносы имеют схожие картины генетических профилей, в которых доминирует генетическая компонента, встречающаяся также у других народов Северного Кавказа, за исключением ингушей, у которых ее 1–5%. Цахуры привлекают внимание тем, что в структуре их генофонда, как и у багвалинцев (андийская подгруппа), в достаточно высокой степени (35%) представлена компонента, которая встречается в других этносах в долях не более 3%. В остальном, по компонентному составу и их долям цахуры вписываются в структуру, свойственную лезгинской языковой группе. Выбывающим из этого порядка оказались арчинцы – малочисленный этнос лезгинской языковой группы, местом проживания которого является анклав внутри аварского этноса Чародинского района Дагестана. Отдельная специфичная компонента доминирует у них в генофонде на уровне 93%.

Промежуточное положение между группами этносов Северо-Западного Кавказа и дагестанскими популяциями, занимают тюркоязычные кумыки Дагестана, у них большие доли компонент этносов нахской группы и северо-западнокавказских. При этом, перечень генетических компонент в составе генофонда тюркоязычных кумыков тот же, что и у даргино-лакской, а также группы лезгиноязычных этносов. На наш взгляд, такое положение занимают кумыки в силу того, что данным исследованием охвачены так называемые северные кумыки, которые проживают в Северо-Восточной части Дагестана – в селениях:

Эндирей (Андрейаул), Костек и Аксай Хасавюртовского района. История основания этих аулов датируются XVI–XVII веками нашей эры, относительно “молодыми” по сравнению с поселениями внутреннегорного и высокогорного Дагестана, возникновение которых согласно историческим источникам относится к эпохам железа, бронзы и даже мезолита. Более того, не исключено, что основателями этих поселений были представители других этносов (не дагестанских) кавказской историко-этнографической общности.

## ЗАКЛЮЧЕНИЕ

Народы Северо-Западного Кавказа (мегрелы, абхазы, адыги, черкесы, осетины, карачаевцы, балкарцы) несмотря на их принадлежность к различным языковым семьям, в генетическом отношении составляют родственную между собой группу этносов с набором одних и тех же предковых субстратных и суперстратных компонент, представленных в их генофондах практически в равных долях.

Этносы нахской ветви нахско-дагестанской языковой семьи (чеченцы и ингуши) отчетливо проявляют специфичный для них генетический пласт, который можно интерпретировать как “нахский”, достигающий почти 98% у ингушей и более 95% у чеченцев. Таким образом, данная этническая общность является обособленной в генетическом плане группой этносов Северного Кавказа.

Коренные этносы Дагестана демонстрируют пеструю картину генетического разнообразия по компонентному составу и по соотношениям их представленности в генофондах. Для таких малочисленных этносов Дагестана как гунзибцы, дидойцы, гинухцы, ахвахцы и арчинцы, обнаружено подавляющее (около 100%) доминирование отдельных, свойственных только для них этноспецифичных компонент, что объясняется довольно сильным эффектом основателя и сдвигом частот аллелей множества SNP для этих популяций. Все это свидетельствует о глубокой древности разделения этих народов и истории их изоляции во времени и пространстве у коренных этносов Дагестана.

Наличие в структуре генофонда северных кумыков, проживающих в Хасавюртовском районе Дагестана (Костек, Эндирей, Аксай) ощутимых долей предковых компонент “Северо-Западного Кавказа” (25%) и “нахского” (20%) дает нам основание сделать несколько предположений:

– предполагать, что этногенез тюркоязычного этноса Дагестана – кумыков происходил в результате консолидации автохтонной основы (кумыки Каякентского, Карабудахкентского и Буйнакского районов), которая формировалась син-

хронно с формированием даргинского, лакского и других аборигенных этносов и населения, обосновавшегося в XVI–XVII в. на северо-восточной части Дагестана;

– не исключать, что языком общения для этой разноплеменной группы обосновавшейся на северо-восточной части Дагестана был карачаевско-балкарский, с которым находясь в контактной зоне впоследствии произошла ассимиляция автохтонного языка кумыков.

Ногайцы и караногайцы Дагестана с преобладающей компонентой, свойственной для этносов Центральной Азии ожидаемо объединились в общий генетический кластер с казахами и узбеками.

#### ФИНАНСИРОВАНИЕ,

Работа выполнена в рамках государственного задания лаборатории эволюционной генетики НИИ медицинской генетики Томского НИМЦ.

#### СПИСОК ЛИТЕРАТУРЫ

1. Prahlow J.A., Cameron T., Arendt A., Cornelis K., Bontrager A., Suth M.S., Black L., Tobey R., Pollock S., Cotter K., Gabrielse J. DNA testing in homicide investigations. *Med. Sci. Law*, 2017, 57(4), 179–191. <https://doi.org/10.1177/0025802417721790>
2. Колесников Н.А., Харьков В.Н., Зарубин А.А., Степанов В.А. Особенности распределения регионов высокой гомозиготности в геномах представителей коренного населения Северной Евразии. *Генетика*, 2019, 55(10), 1231–1236. <https://doi.org/10.1134/S0016675819100072>
3. Колесников Н.А., Харьков В.Н., Раджабов М.О., Зарубин А.А., Литвинов С.С., Екомасова Н.В., Хуснутдинова Э.К., Степанов В.А. Различия по уровню инбридинга в популяциях Дагестана: анализ регионов высокой гомозиготности. *Медицинская генетика*. 2020, 19(7), 21–22. <https://doi.org/10.25557/2073-7998.2020.07.21-22>
4. Колесников Н.А., Харьков В.Н., Зарубин А.А., Раджабов М.О., Воевода М.И., Губина, Э.К. Хуснутдинова, С.С. Литвинов, Н.В. Екомасова, О.В. Штыгашева, Н.Р. Максимова М.А., Сухомясова А.Л., Степанов В.А. Особенности геномного распределения регионов высокой гомозиготности у коренного населения Северной Евразии на индивидуальном и популяционном уровнях на основе анализа SNP высокой плотности. *Генетика*, 2021, 57(11), 1261–1276. <https://doi.org/10.31857/S0016675821110059>
5. Таймазов А.И. Айникаб I: стоянка олдована на Северо-Восточном Кавказе. отв. ред. Х. А. Амиранов. Дагестанский федеральный исследовательский центр Российской академии наук, Институт истории, археологии и этнографии. Махачкала, Мавраевъ, 2021. 288 с.
6. Ярцева В.Н. (Ред.). Языки мира: Кавказские языки. М.: Academia, 1998. 480 с.
7. Шихсаидов А.Р. Распространение ислама в Дагестане. *Исламоведение*, 2010, 1, 31–50.
8. Macoudi. *Les Prairies d'or*. Texte et traduction par C. Barbier de Meynard et Pavet de Courteille. 1863. V 3.
9. Минорский В.Ф. История Ширвана и Дербенда. X–XI вв. Отв. ред. А.А. Али-Заде. М.: Издательство восточной литературы, 1963. 266 с.
10. Гаджиева С.Ш. Дагестанские терекеменцы. XIX–начало XX в.: Историко-этнографическое исследование. Наука, 1990. 216 с. ISBN 5-02-016761-4.
11. Кусаева З.К., Сатцаев Э.Б., Таказов Ф.М. Джалган и джалганцы: синтез персидского и аланского компонентов (по материалам фольклорно-этнографической и лингвистической экспедиции 14–18 июня 2016 г.). *Известия СОИГСИ*, 2016, 21(60), 96–102.
12. Balanovsky O., Dibirova K., Dybo A., Mudrak O., Frolova S., Pocheshkhova E., Haber M., Platt D., Schurr T., Haak W., Kuznetsova M., Radzhabov M., Balaganskaya O., Romanov A., Zakharova T., Hernanz D.F.S., Zalloua P., Koshel S., Ruhlen M., Renfrew C., Wells R.S., Tyler-Smith C., Balanovska E., *Genographic Consortium* Parallel evolution of genes and languages in the Caucasus region. *Mol. Biol. Evol.*, 2011, 28(10), 2905–2920. <https://doi.org/10.1093/molbev/msr126>
13. Biagini S.A., Solé-Morata N., Matisoo-Smith E., Zalloua P., Comas D., Calafell F. People from Ibiza: an unexpected isolate in the Western Mediterranean. *Eur. J. Hum. Genet.*, 2019, 27(6), 941–951.
14. Gopalan S., Berl R.E.W., Myrick J.W., Garfield Z.H., Reynolds A.W., Bafens B.K., Belbin G., Mastoras M., Williams C., Michelle Daya M., Negash A.N., Feldman M.F., Barry S., Hewlett B.S., Henn B.M. Hunter-gatherer genomes reveal diverse demographic trajectories during the rise of farming in Eastern Africa. *Curr. Biol.*, 2022, 32(8), 1852–1860. <https://doi.org/10.1016/j.cub.2022.02.050>
15. Alexander D.H., Novembre J., Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.*, 2009, 19(9), 1655–1664.
16. Маниатис Т., Фрич Э., Сэмбрук Дж. Методы генетической инженерии. Молекулярное клонирование. Москва: Изд-во “Мир”, 1984, 480 с.
17. Guo Y., He J., Zhao S., Wu H., Zhong X., Sheng Q., Samuels D.C., Shyr Y., Long J. Illumina human exome genotyping array clustering and quality control. *Nat. Protoc.*, 2014, 9(11), 2643–2662.

**Component Composition of the Gene Pool Indigenous Ethnos of Dagestan**

**M. O. Radzhabov<sup>a, #</sup>, V. N. Kharkov<sup>b</sup>, N. A. Kolesnikov<sup>b</sup>, A. A. Zarubin<sup>b</sup>, I. Yu. Khitrinskaya<sup>b</sup>,  
M. A. Dzhaubermezov<sup>c, d</sup>, O. M. Radzhabov<sup>e</sup>, E. K. Khusnutdinova<sup>c, d</sup>, and V. A. Stepanov<sup>b</sup>**

<sup>a</sup>*Dagestan Federal Research Center, Mahachkala, 367000 Russia*

<sup>b</sup>*Research Institute of Medical Genetics, Tomsk National Research Medical Center,  
Russian Academy of Sciences, Tomsk, 634050 Russia*

<sup>c</sup>*Institute of Biochemistry and Genetics, Ufa Federal Research Center,  
Russian Academy of Sciences, Ufa, 450054 Russia*

<sup>d</sup>*Ufa University of Science and Technology, Ufa, 450076 Russia*

<sup>e</sup>*Science Centre for Genetics and Life Sciences of Sirius University of Science and Technology,  
Sirius Federal Territory, Sirius Federal Territory, 354340 Russia*

<sup>#</sup>*e-mail: radjabov\_m@mail.ru*

**Abstract**—A component analysis of the gene pool composition of the indigenous ethnic groups of Dagestan, as well as other ethnic groups of the North Caucasus and Central Asia, was carried out based on wide-genome genotyping of SNP markers on high-density DNA microarrays Infinium Multi-Ethnic Global-8. The analysis of the genotypic data array of ethnic populations of the above-mentioned regions was implemented in the Admixture algorithm ( $K = 12$ ). Despite the fact that in ethno-linguistic terms, the Mingrelians belong to the people of the Kartvelian language family, the Abkhazians, Adyghe and Circassians belong to the group of ethnic groups of the Abkhaz-Adyghe languages, the Ossetians are an Iranian-speaking ethnic group, and the Karachays and Balkars represent the Turkic group of peoples of the North Caucasus, all of the listed ethnic groups have a spectrum of the same ancestral components, moreover, in comparable proportions. The Ingush and Chechens, being representatives of the Nakh group of the Nakh-Dagestani language family, demonstrate one common ancestral component with a high degree of representation, emphasizing the isolated genetic history of the formation of their gene pool. The Karanogays of Dagestan and the Nogays of Karachay-Cherkessia, in terms of the spectrum and frequency of representation of ancestral components, indicate a commonality with Central Asian ethnic groups, such as the Kazakhs and Uzbeks. The ethnic groups of Dagestan demonstrate a unique picture of diversity in both the composition and frequency of occurrence of ancestral components in the structure of their autosomal gene pool. This emphasizes not only the history of their formation, isolated from other peoples of the North Caucasus, but also the long-isolated genetic history of each indigenous ethnic group. The study is within the framework of fundamental areas of human population and evolutionary genetics — identifying ethnospecific features of the structure of genomes and population gene pools of modern human populations, reconstructing ancient migrations and population expansion, as well as the evolution of gene pools and mechanisms of genetic adaptation.

**Keywords:** gene pool, indigenous ethnic group, Dagestan, ancestral components, DNA microarray, Admixture algorithm