

Перевод на английский язык <https://vavilov.elpub.ru/jour>

## Идентичные по происхождению блоки в геномах коренного населения Сибири демонстрируют генетические связи между популяциями

Н.А. Колесников , В.Н. Харьков, К.В. Вагайцева, А.А. Зарубин, В.А. Степанов

Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, Россия  
 [nikita.kolesnikov@medgenetics.ru](mailto:nikita.kolesnikov@medgenetics.ru)

**Аннотация.** Генофонд коренного населения Сибири представляет собой уникальную систему с точки зрения исследования популяционно- и эволюционно-генетических процессов, анализа генетического разнообразия и реконструкции генетической истории популяций. Высокое этническое разнообразие является особенностью Сибири как одного из регионов периферийного расселения современного человека. Огромные пространства этого региона и малочисленность аборигенного населения способствовали формированию значительной территориальной и генетической подразделенности. На территории сибирской историко-этнографической провинции расселены около 40 коренных народностей. Проведено масштабное популяционное исследование генофонда коренных народов Сибири на уровне высокоплотного ДНК-микрочипа Infinium Multi-Ethnic Global-8, позволяющее заполнить существенный пробел в геногеографической картине населения Евразии. Для этого были отобраны и проанализированы фрагменты ДНК, унаследованные без рекомбинации каждой парой индивидов от их недавнего общего предка, т.е. сегменты (блоки), идентичные по происхождению (IBD). Распределение блоков IBD в популяциях Сибири хорошо согласуется с географической близостью популяций и их языковой принадлежностью. Чукчи, коряки и нивхи среди сибирских популяций формируют отдельный от основной группы Сибири кластер, причем чукчи и коряки являются более близкородственными. Образуются отдельные субкластеры эвенков и якутов, кетов и чулымцев, тувинцев и алтайцев внутри сибирского кластера. Анализ SNP, которые попадали в большее количество IBD-сегментов анализируемых популяций, позволил составить список из 5358 генов. По результатам расчета, обогащенные этими генами биологические процессы связаны с обнаружением химического раздражителя, участвующего в сенсорном восприятии запаха. Обогащенные найденными генами молекулярные пути связаны с метаболизмом линолевой, арахидоновой, тирозиновой кислот и путем обонятельной трансдукции. При этом анализ литературных данных показал, что некоторые из отобранных генов, которые встречались в большем количестве блоков IBD сразу в нескольких популяциях, могут играть роль в адаптации человека к факторам окружающей среды.

Ключевые слова: IBD; популяции человека; сибирские популяции.

**Для цитирования:** Колесников Н.А., Харьков В.Н., Вагайцева К.В., Зарубин А.А., Степанов В.А. Идентичные по происхождению блоки в геномах коренного населения Сибири демонстрируют генетические связи между популяциями. *Вавиловский журнал генетики и селекции*. 2023;27(1):55-62. DOI 10.18699/VJGB-23-08

## Blocks identical by descent in the genomes of the indigenous population of Siberia demonstrate genetic links between populations

N.A. Kolesnikov , V.N. Kharkov, K.V. Vagaitseva, A.A. Zarubin, V.A. Stepanov

Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia  
 [nikita.kolesnikov@medgenetics.ru](mailto:nikita.kolesnikov@medgenetics.ru)

**Abstract.** The gene pool of the indigenous population of Siberia is a unique system for studying population and evolutionary genetic processes, analyzing genetic diversity, and reconstructing the genetic history of populations. High ethnic diversity is a feature of Siberia, as one of the regions of the peripheral settlement of modern human. The vast expanses of this region and the small number of aboriginal populations contributed to the formation of significant territorial and genetic subdivision. About 40 indigenous peoples are settled on the territory of the Siberian historical and ethnographic province. Within the framework of this work, a large-scale population study of the gene pool of the indigenous peoples of Siberia was carried out for the first time at the level of high-density biochips. This makes it possible to fill in a significant gap in the genogeographic picture of the Eurasian population. For this, DNA fragments were analyzed, which had been inherited without recombination by each pair of individuals from their recent common ancestor, that is, segments (blocks) identical by descent (IBD). The distribution of IBD blocks in the populations of Siberia is in good agreement with the geographical proximity of the populations and their linguistic affiliation. Among the Siberian populations, the Chukchi, Koryaks, and

Nivkhs form a separate cluster from the main Siberian group, with the Chukchi and Koryaks being more closely related. Separate subclusters of Evenks and Yakuts, Kets and Chulymy are formed within the Siberian cluster. Analysis of SNPs that fell into more IBD segments of the analyzed populations made it possible to compile a list of 5358 genes. According to the calculation results, biological processes enriched with these genes are associated with the detection of a chemical stimulus involved in the sensory perception of smell. Enriched for the genes found, molecular pathways are associated with the metabolism of linoleic, arachidonic, tyrosic acids and by olfactory transduction. At the same time, an analysis of the literature data showed that some of the selected genes, which were found in a larger number of IBD blocks in several populations at once, can play a role in genetic adaptation to environmental factors.

Key words: IBD; human populations; Siberian populations.

**For citation:** Kolesnikov N.A., Kharkov V.N., Vagaitseva K.V., Zarubin A.A., Stepanov V.A. Blocks identical by descent in the genomes of the indigenous population of Siberia demonstrate genetic links between populations. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2023;27(1):55-62. DOI 10.18699/VJGB-23-08

## Введение

Генетико-демографические процессы в популяциях, колебания численности, метисационные события, миграции и естественный отбор оказывают влияние на структурированность генетического разнообразия в геномах отдельных индивидов и популяциях в целом. В частности, генетико-демографические процессы приводят к тому, что формируются общие по происхождению блоки сцепления (identity by descent, IBD). Сегментом, имеющим идентичные нуклеотидные последовательности, является IBD у двух или более лиц, если они унаследовали его от общего предка без рекомбинации, т. е. у этих людей сегмент имеет общее происхождение. Ожидаемая длина блока IBD зависит от количества поколений с момента появления последнего общего предка в локусе сегмента (Browning S.R., Browning B.L., 2010; Palamara et al., 2012).

Сегменты IBD можно использовать для оценки демографической истории популяций, включая эффект бутылочного горлышка и потоки генов в популяциях (Gusev et al., 2012). В недавних исследованиях были показаны различия в распределении IBD между африканскими, азиатскими и европейскими популяциями, а также сегменты IBD, общие с древними геномами, такими как геномы неандертальцев или денисовцев (Hochreiter, 2013).

Близкие биологические родственники имеют идентичные друг другу достаточно длинные фрагменты ДНК и, соответственно, в большинстве хромосом находятся идентичные по происхождению блоки значительной длины ( $> 66.7$  сМ), в результате чего ожидаемая длина общих IBD составляет около 1700 сМ. Согласно расчетам, двоюродные и троюродные родственники, вероятно, будут иметь несколько регионов (более 2.5 сегментов) из-за наличия недавних предков, определяющих их родство. Предполагается, что для двоюродных родственников каждый IBD будет иметь общую длину больше 62 сМ и для троюродных – 25 сМ. Дальние родственники, относящиеся к четвероюродным или более дальним кузенам, с большой вероятностью несут один или несколько регионов от своего ближайшего общего предка. Такие пары включают в себя подавляющее большинство людей в конкретной популяции и обычно называются «неродственными», поскольку ожидается, что доля и число IBD по всему геному будут между ними относительно небольшими (Gusev et al., 2012).

Сегменты IBD могут помочь и в обнаружении сигналов естественного отбора в геноме человека. Поиск регионов с избытком сегментов IBD позволяет идентифицировать

регионы в геноме человека, которые находятся под очень недавним и сильным отбором, поскольку отбор обычно увеличивает число сегментов IBD среди людей в популяции (Albrechtsen et al., 2010; Han, Abney, 2011).

Относительно популяций коренных этносов Сибири было высказано предположение, что крупномасштабное рассредоточение и смешение популяций, вероятно, сыграли существенную роль в их предыстории и структуре, что может объяснить необычно высокую долю IBD между популяциями (Pugach et al., 2016).

Цель настоящего исследования – анализ структуры генофонда популяций коренного населения Сибири на основании выявления идентичных по происхождению блоков сцепления, их внутри- и межпопуляционного распределения.

## Материалы и методы

Данные широкогеномного генотипирования получены с использованием микрочипов Infinium Multi-Ethnic Global-8 (Illumina), включающих свыше 1.7 млн маркеров. В анализе не было образцов, имеющих более 5 % пропущенных позиций, а также SNP с более чем 10 % отсутствующими генотипами. Данные были предварительно отфильтрованы по минимальной частоте редкого аллеля (MAF, minor allele frequency  $> 0.01$ ). В итоге после фильтрации исходных данных в финальный массив данных вошло 886 889 аутомомных SNP.

Популяции коренного населения Сибири ( $N = 477$ ) представлены: алтайцами (Б – с. Бешпельтир Чемальского муниципального района,  $N = 24$  и К – с. Кулада Онгудайского района,  $N = 25$ ), бурятами (А – пос. Агинское Агинского района,  $N = 23$  и К – с. Курумкан Курумканского района,  $N = 28$ ), калмыками ( $N = 29$ ), кетами ( $N = 15$ ), коряками ( $N = 20$ ; материал коряков собран в Корякском автономном округе Камчатской области), чукчами ( $N = 25$ ; популяционная выборка чукчей (образцы крови собраны в пос. Лорино, Сиреники, Янарыкот и Новое Чаплино Чукотского автономного округа) относится к береговой группе), нивхами ( $N = 13$ ), татарами (Т – г. Томск,  $N = 20$ ), тувинцами ( $N = 28$ ), удэгейцами ( $N = 15$ ), хантами (К – с. Казым Белоярского района,  $N = 30$  и Р – д. Русскинская Сургутского района,  $N = 26$ ), хакасами (Т – сагайцами Таштыпского района,  $N = 29$  и Ш – качинцами Широинского района,  $N = 26$ ), чулымцами ( $N = 22$ ), эвенками (З – забайкальские (пос. Чара Каларского района, с. Моклакан и с. Тупик Тунгиро-Олёкминского района),  $N = 25$  и Я – якутские,  $N = 28$ ) и якутами ( $N = 26$ ).

**Таблица 1.** Характеристики изученных популяционных выборок коренного населения Сибири

Популяция	Локализация	Объем выборки	Языковая принадлежность (семья/группа)	Антропологический тип
Алтайцы (Б)	Республика Алтай	24	Алтайская/тюркская	Монголоидный (центральноазиатский)
Алтайцы (К)		25		
Буряты (А)	Забайкальский край	23	Алтайская/монгольская	
Буряты (К)	Республика Бурятия	28		
Чукчи	Чукотский автономный округ	25	Чукотско-камчатские языки	Монголоидный (арктический)
Чулымцы	Томская область	22	Алтайская/тюркская	Уральский, монголоидный (южносибирский)
Эвенки (Я)	Якутия	28	Алтайская/тунгусо-маньчжурские языки	Монголоидный (байкальский)
Эвенки (З)	Забайкальский край	25		
Калмыки	Республика Калмыкия	29	Алтайская/монгольская	Монголоидный (центральноазиатский)
Кеты	Красноярский край	15	Енисейская	Енисейский
Хакасы (Ш)	Республика Хакасия	26	Алтайская/тюркская	Уральский, монголоидный (южносибирский)
Хакасы (Т)		29		
Ханты (К)	ХМАО – Югра	30	Уральская/обско-угорская	Уральский
Ханты (Р)		26		
Коряки	Камчатский край	20	Чукотско-камчатские языки	Монголоидный (арктический)
Нивхи	Хабаровский край	13	Палеоазиатский язык	Монголоидный (сахалиноамурский)
Татары (Т)	Томская область	20	Алтайская/тюркская	Уральский, монголоидный (южносибирский)
Тувинцы	Республика Тыва	28		Монголоидный (центральноазиатский)
Удэгейцы	Приморский край	15	Алтайская/тунгусо-маньчжурские языки	Монголоидный (байкальский)
Якуты	Якутия	26	Алтайская/тюркская	Монголоидный (центральноазиатский)

Материал депонирован в биоресурсной коллекции «Био-банк населения Северной Евразии». Характеристика изученных популяций представлена в табл. 1.

Фазирование генотипов проведено с помощью программного обеспечения Beagle 4.1 (Browning S.R., Browning B.L., 2007). Для анализа идентичных по происхождению блоков генома использовали алгоритм Refined IBD, версии refined-ibd.16May19.ad5.jar (Browning B.L., Browning S.R., 2013). Для сравнения популяций были получены суммы средних длин сегментов IBD между парами индивидов 1.5–1.999 сМ, 2–3.999, 4–7.999, 8–15.999 и > 16 сМ (далее в статье данные диапазоны будут записаны как 1.5–2 сМ, 2–4, 4–8 и 8–16 сМ). Тепловая карта с дендрограммой на основании логарифма от суммы средних длин сегментов IBD между парами индивидов для IBD построена с использованием пакета heatmap.2 в программной среде R.

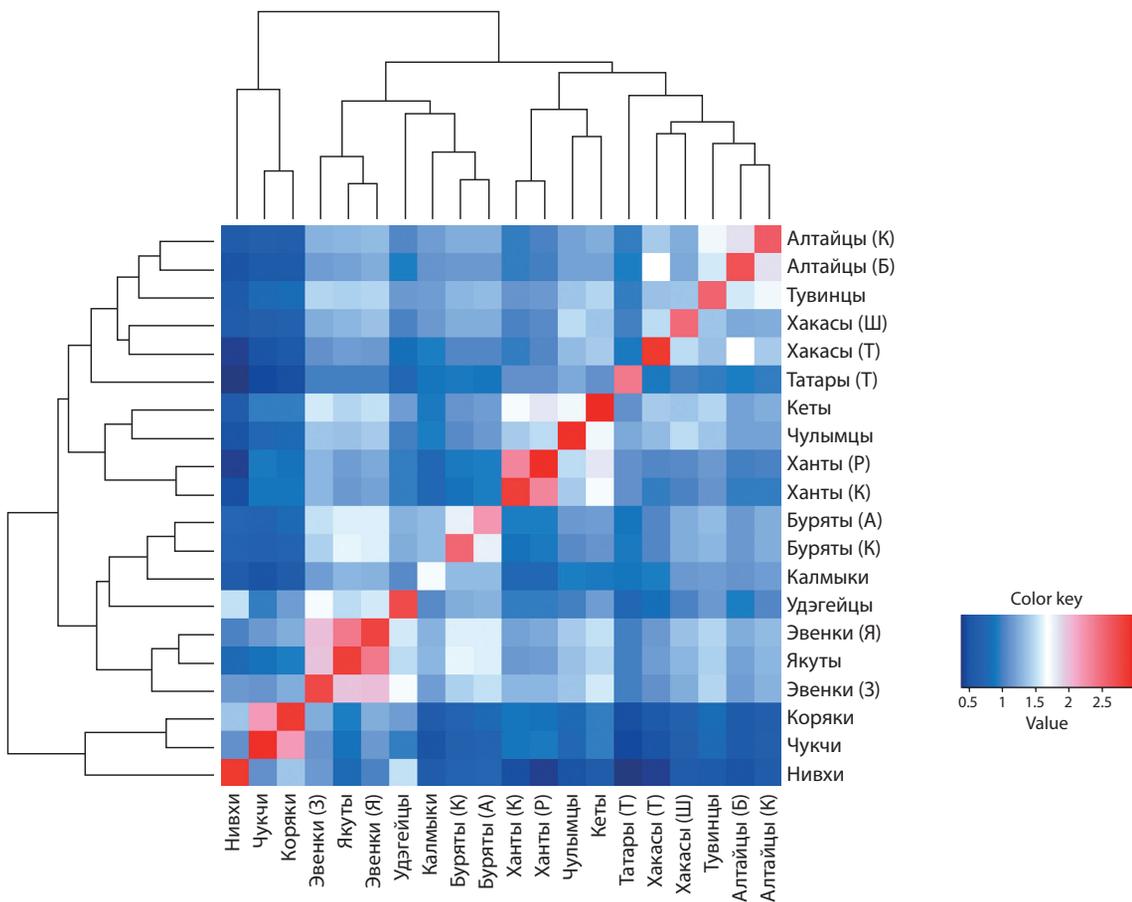
Нами также были идентифицированы SNP, которые попадали в большее количество IBD сегментов анализируемых популяций (частота попадания SNP в IBD выше 99-го квантиля распределения частот), определена принадлежность этих SNP к генам и проведена оценка биологической значимости полученного списка генов. Для анализа использовали веб-ресурс WebGestalt (WEB-based Gene Set Analysis Toolkit), в частности анализ KEGG-путей и генных онтологий (Gene Ontology) методом ORA (over-representation analysis).

## Результаты и обсуждение

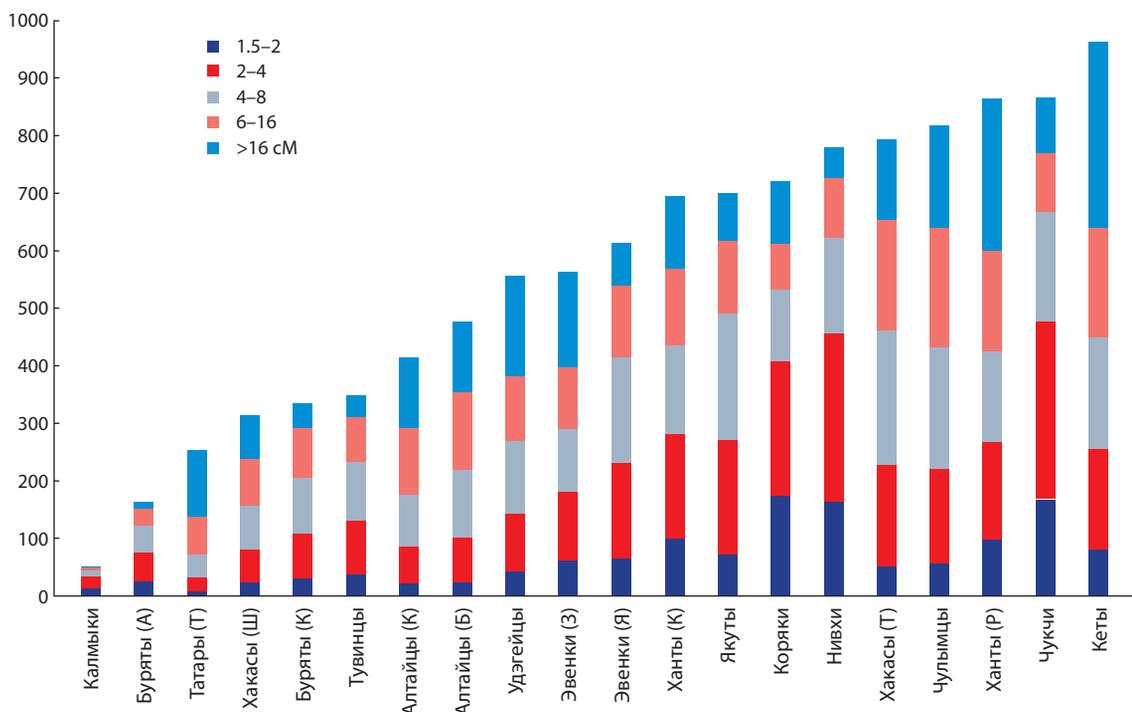
Для более детального анализа генетического родства сибирских популяций и выяснения, в какой степени их генетическая структура может быть объяснена недавними локальными миграциями, мы выделили и проанализировали фрагменты ДНК, которые были унаследованы без рекомбинации каждой парой индивидов от их недавнего общего предка, т. е. сегменты (блоки), идентичные по происхождению (IBD).

Для популяций, населяющих территорию Сибири, характерна уникальная генетико-демографическая история, которая находит отражение в том числе и в распределении блоков IBD как внутри популяции, так и между ними. Нами рассчитаны суммы средних длин сегментов IBD между парами индивидов и на их результатах построена тепловая карта с дендрограммой на основании логарифма от суммы средних длин сегментов IBD (рис. 1).

Количество общих сегментов у представителей разных популяций согласуется с географией их проживания, так как на проживающие рядом народы могут влиять общие генетико-демографические процессы. Анализ тепловой карты демонстрирует кластеризацию популяций сибирской группы, связывающую народности по месту происхождения. Чукчи, коряки и нивхи формируют отдельный от основной группы сибирских популяций кластер, причем чукчи и коряки являются более близкородственными. Формируются отдельные субкластеры эвенков и якутов,



**Рис. 1.** Тепловая карта с дендрограммой на основании логорифма от суммы средних длин сегментов IBD (> 1.5 cM) между парами индивидов.



**Рис. 2.** Диаграмма суммы средних длин сегментов IBD между парами индивидов в исследуемых популяциях для IBD разных размеров (1.5–2, 2–4, 4–8, 8–16, >16 cM).

кетов и чулымцев, тувинцев и алтайцев внутри сибирского кластера.

При градации сегментов IBD по длине в целом тенденция сохраняется, но появляются некоторые отличия, точнее характеризующие недавние взаимоотношения народов. Для более длинных IBD кластеры больше соответствуют нынешнему географическому расположению популяций, отражая недавний обмен общими регионами. По трем диапазонам длин с размерами IBD (1.52–2, 2–4 и 4–8 сМ) популяции лучше разбиваются на близко географически расположенные пары: коряки-чукчи, якуты-эвенки. Кеты практически в равной степени имеют общие блоки IBD с чулымцами (18.7–27.2–7.7 сМ) и хантами (23.4–24.4–4.8 сМ для хантов (К) и 25.9–30.1–7.9 сМ для хантов (Р)). При этом среди хантов большее значение общих блоков IBD наблюдается у русскиных хантов, что соответствует их более близкому географическому расположению по сравнению с хантами Белоярского района.

Для более удаленных от кетов хакасов наблюдается меньшее значение общих блоков IBD практически в два-три раза по сравнению с хантами (9.8–10.4–2.5 сМ для хакасов (Ш) и 9.1–10.6–3.8 сМ для хакасов (Т)). Несмотря на то что эвенки, тувинцы и якуты еще более удалены, для тувинцев (12.3–12.8–2.1 сМ) и якутов (12.8–12.5–2.1 сМ) наблюдаются схожие значения общих блоков IBD. Для забайкальских эвенков (16.4–16.9–3.0 сМ) нужно отметить большие значения блоков IBD, чем с якутскими эвенками (14.7–14.0–2.6 сМ).

За некоторыми исключениями, общие сегменты IBD между сибирскими популяциями лучше объясняются географической близостью популяций, а не их языковой принадлежностью. Так, например, якутские эвенки, проживающие на территории Якутии, имеют больше общих IBD с якутами (252.7 сМ), чем с забайкальскими эвенками (102.5 сМ). При этом сумма средних длин сегментов IBD между парами индивидов между двумя популяциями эвенков сопоставима с суммой между забайкальскими эвенками и якутами (95.2 сМ).

С точки зрения анализа IBD внутри популяций в целом индивиды из популяций Крайнего Севера и Дальнего Востока (коряки, чукчи, нивхи) имеют больше общих IBD с образцами из той же группы, чем индивиды из южносибирских популяций, таких как алтайцы и тувинцы. При этом у чукчей, коряков и нивхов наибольший вклад вносят короткие IBD-фрагменты 1.5–4 сМ (55–57–59 %), что может говорить о бутылочном горлышке в прошлом во время миграций на север и северо-восток и/или сильной изоляции от остальных популяций, населяющих территорию Сибири. При этом у чулымцев, хантов (Р), населяющих центральную часть Сибири и имеющих наибольшую сумму средних длин сегментов IBD между парами индивидов, самый большой вклад вносят IBD длиной более 8 сМ (47–51 %), что говорит о недавнем значительном инбридинге внутри популяции.

Наиболее генетически гетерогенные сибирские популяции, имеющие минимальные значения по коэффициенту геномного инбридинга ( $F_{ROH}$ ) (Колесников и др., 2021), также имеют минимальные значения по сумме средних

длин сегментов IBD. Особенно выражено это у популяций калмыков, агинских бурят и томских татар (рис. 2).

В популяциях бурятов нет значимых различий по средней суммарной длине между парами индивидов с другими популяциями, но есть существенная разница в распределении IBD внутри популяций. Так, у бурятов (К) значительно больше средняя суммарная длина между парами индивидов внутри популяции (335.4 сМ), по сравнению с бурятами (А) (1635 сМ), во многом за счет средних и длинных IBD. У агинских бурят гораздо больше доля коротких IBD (16–30–28–19–7 %), чем у курумканских (9–24–29–25–13 %). Несмотря на то что для бурятов, имеющих большую общую численность, отмечено практически удвоение численности с 237 тыс. человек в 1926 г. до 461 тыс. в 2010 г., данное различие между популяциями можно объяснить резким ростом численности населения пос. Агинское с 451 человека в 1908 г. до 4556 человек в 1939 г. и 15 тыс. – в 2010 г., и стабильностью населения пос. Курумкан с численностью 5617 человек в 1979 г. и 5465 – в 2010 г., находящегося в одном из отдаленных районов Бурятии.

Похожая динамика наблюдается у популяций чукчей (20–36–22–12–11 %), коряков (24–33–17–11–15 %), нивхов (21–38–21–13–7 %) и калмыков (29–39–22–7–3 %), за более резким уменьшением доли длинных IBD. Для чукчей, нивхов и коряков, имеющих небольшую численность (до 13 тыс.), характерно отсутствие резких колебаний численности популяций за последние 100 лет с ростом на 29–9–6 % соответственно, что также могло способствовать сокращению общих длинных IBD-сегментов при условии незначительного количества близкородственных браков внутри популяции.

Всего для 20 популяций получено 189314 SNP, попадающих в наибольшее количество IBD-сегментов анализируемых популяций (частота попадания SNP в IBD выше 99-го квантиля распределения частот). Из них 88530 SNP находятся в межгенных участках, остальные локализованы в области 5358 генов. В табл. 2 представлены гены, которые были показаны в четырех и более популяциях.

Из списка 5358 генов, наиболее часто попадающих в блоках IBD, 1694, по данным WebGestalt, оказались аннотированными по базе KEGG. В результате анализа с учетом поправки Бенджамини–Хохберга ( $FDR = 0.05$ ) были выявлены обогащенные этими генами молекулярные KEGG-пути: путь метаболизма линолевой кислоты hsa00591 ( $FDR = 0.0051$ ), включающий 17 генов (*CYP2C8*, *CYP2C9*, *PLA2G1B*, *PLB1*, *CYP1A2*, *CYP2C19*, *CYP3A4*, *PLA2G10*, *PLA2G2A*, *PLA2G2C*, *PLA2G2D*, *PLA2G2E*, *PLA2G2F*, *PLA2G4A*, *PLA2G4C*, *PLA2G5*, *PLA2G6*); путь метаболизма арахидоновой кислоты hsa00590 ( $FDR = 0.0240$ ), включающий 27 генов (*CYP2C8*, *CYP2C9*, *PLA2G1B*, *PLB1*, *ALOX12*, *ALOX12B*, *ALOX15B*, *ALOX5*, *CYP2B6*, *CYP2C19*, *GPX1*, *GPX3*, *GPX7*, *PLA2G10*, *PLA2G2A*, *PLA2G2C*, *PLA2G2D*, *PLA2G2E*, *PLA2G2F*, *PLA2G4A*, *PLA2G4C*, *PLA2G5*, *PLA2G6*, *PTGIS*, *PTGS1*, *PTGS2*, *TBXAS1*); путь метаболизма тирозина hsa00350 ( $FDR = 0.0240$ ), включающий 18 генов (*ADH1A*, *ADH1B*, *ADH1C*, *ADH4*, *ADH5*, *ADH6*, *ADH7*, *ALDH3B1*, *ALDH3B2*, *AOC2*, *AOC3*, *DDC*, *GOT1*, *HPD*, *IL4I1*, *PNMT*, *TYR*,

**Таблица 2.** Гены, попадающие в регионы с максимальным количеством блоков IBD более чем в четырех популяциях

Ген	Алтайцы (Б)	Алтайцы (К)	Буряты (А)	Буряты (К)	Чулымцы	Эвенки (Я)	Эвенки (З)	Калмыки	Кеты	Хакасы (Ш)	Хакасы (Т)	Ханты (К)	Ханты (Р)	Коряки	Нивхи	Татары (Т)	Тувинцы	Удгейцы	Якуты
AAGAB	+	+	+											+			+		
GSE1						+							+	+	+				+
IQCH	+	+	+											+			+		
IQCH-AS1	+	+	+											+			+		
SMAD3	+	+	+											+			+		
ANGPTL1	+						+					+	+						
RALGPS2	+						+					+	+						
ARFGAP1							+	+							+			+	
C15orf61		+	+											+			+		
C16orf74						+			+				+		+				
CALML4		+	+											+			+		
CHRNA4							+	+							+			+	
CLN6		+	+											+			+		
COL20A1							+	+							+			+	
FEM1B		+	+											+			+		
FLJ16779							+	+							+			+	
ITGA11		+	+											+			+		
LINC02206		+	+											+			+		
LOC100130587							+	+							+			+	
LOC101929076		+	+											+			+		
LOC102723493		+	+											+			+		
MAP2K5		+	+											+			+		
MIR99AHG							+	+						+	+				
NKAIN4							+	+							+			+	
PIAS1		+	+											+			+		
RALGPS2	+						+					+	+						
SKOR1		+	+											+			+		

*TYRP1*); путь обонятельной трансдукции hsa04740 (FDR = 4.55E-08), включающий 159 генов.

Метаболическое превращение полиненасыщенных жирных кислот (ПНЖК), таких как линолевая кислота, в биологически активные ПНЖК с длинной цепью (> 20 атомов углерода, ДЦ-ПНЖК), таких как арахидоновая кислота, необходимо для правильного обмена веществ. ДЦ-ПНЖК и их метаболиты являются важными структурными и сигнальными компонентами для многочисленных биологических систем, включая развитие и функционирование мозга, врожденный иммунитет и энергетический гомеостаз (Marszalek, Lodish, 2005; Calder, 2013). Существуют также пищевые источники предварительно сформированных ДЦ-ПНЖК в яйцах и некоторых видах мяса, содержащих арахидоновую кислоту (Horrocks, Yeo, 1999; Howe et al., 2006; Chilton et al., 2014). Найденные закономерности

в распределении IBD, содержащих гены, участвующие в метаболизме жирных кислот, могут указывать на недавний направленный отбор, связанный с адаптацией к особенностям питания в холодном климате, или отражать влияние западной диеты (Chilton et al., 2014).

Положительный отбор в генах, влияющих как на уровень ДЦ-ПНЖК, так и на метаболическую эффективность, посредством которой образуются ДЦ-ПНЖК в популяциях пигмеев на острове Флорес (Tucci et al., 2018), гренландских инуитов (Fumagalli et al., 2015) и коренных американцев (Amorim et al., 2017; Harris et al., 2019), предположительно также связан с особенностями питания в условиях холодного климата, хотя точное давление отбора неизвестно (Fumagalli et al., 2015). В качестве примера можно привести схожесть генотипов в генах, входящих в систему обоняния, которая может играть роль в фор-

мировании или поддержании социальных связей между индивидами внутри популяции (Christakis, Fowler, 2014). Для таких генотипов выявлены более высокие показатели положительного отбора (Fu et al., 2012).

Анализ геномных онтологий (3511 генов, по данным WebGestalt, оказались аннотированными по базе геномной онтологии) показал девять статистически значимых биологических процессов (с учетом поправки Бенджамини–Хохберга (FDR = 0.05)), связанных с обнаружением химического раздражителя, участвующего в сенсорном восприятии запаха (GO:0050911, GO:0007608, GO:0050907, GO:0050906, GO:0051606, GO:0009593, GO:0007600, GO:0007606, GO:0050877).

Анализ литературы также выявил, что ряд генов, попадающих в блоки IBD, может играть существенную роль в образовании онкологии и влиять на лечение, например ген *AAGAB*, попадающий в блоки IBD у пяти популяций. Ген *AAGAB* состоит из 10 экзонов, кодирующих белок из 315 аминокислот (aa), белок AAGAB (связывающий альфа- и гамма-адаптин). Он широко экспрессируется и взаимодействует с гамма-адаптинным и альфа-адаптинным адаптерными белковыми комплексами, AP1 и AP2, а также участвует в переносе мембран и играет роль в эндоцитозе и сортировке белков. Гетерозиготные мутации *AAGAB* вызывают точечную ладонно-подошвенную кератодермию типа 1 (PPKP1) – кожное заболевание, характеризующееся точечным гиперкератозом ладоней и подошв (Kiritsi et al., 2013). Ген *AAGAB* является многообещающим биомаркером ответа и исхода химиотерапии во время лечения рака молочной железы. Однако точная роль *AAGAB* в развитии рака молочной железы в настоящее время неясна и потенциально требует дальнейшего изучения (Bownes et al., 2019).

Другой ген, *GSE1*, который попадает у пяти популяций в блоки IBD, может функционировать как онкоген при раке молочной железы, желудка и предстательной железы, и также может быть важен при лечении пациентов с раком предстательной железы (Vamodu et al., 2021).

Ген *IQCH-AS1*, кодирующий антисмысловую РНК IQCH 1 (IQCH-AS1), коррелирует с выживаемостью и диагностикой онкологических больных, однако его роль в развитии рака щитовидной железы и химиочувствительности доксорубина остается неясной (Fei et al., 2022).

Участие гена *SMAD3* в регуляции генов, важных для развития клеток, таких как дифференцировка, рост и гибель, подразумевает, что изменение или подавление ее активности может привести к образованию или развитию рака.

Некоторые гены, представленные в табл. 1, могут играть роль в адаптации человека к факторам окружающей среды. Например, ген *IQCH* может быть регулирующим в сперматогенезе (Yin et al., 2005), а также ассоциирован у монголов с ростом взрослого человека (Kimura et al., 2008).

Ген *CHRNA4* кодирует субкомпонент никотиновых рецепторов  $\alpha 4\beta 2$  в человеческом мозге. Показано, что люди, обладающие определенными генотипами гена *CHRNA4*, в задачах визуального поиска лучше справлялись с отслеживанием и идентификацией нескольких объектов (Espeseth et al., 2010). Полиморфизмы в гене *CHRNA4*, по-видимому, также вносят свой вклад в формирование

личности, влияя на степень чувствительности развития как к нормальным, так и неблагоприятным условиям окружающей среды (Grazioplene et al., 2013).

Ген *COL20A1*, кодирующий коллаген типа XX альфа 1, отмечен в ряде генов, имеющих несинонимичные изменения с высокой частотой у современного человека в сравнении с архаичными гоминидами, что, вероятно, способствовало развитию уникальных человеческих черт, и является интересным объектом для изучения (Kuhlwilm, Voesckx, 2019).

## Закключение

Нами получена новая информация о структуре и составе генофондов коренных народов Сибири, их генетических взаимоотношений и генетико-демографических процессах на основании анализа распределения идентичных по происхождению блоков сцепления. Результаты исследования показывают кластеризацию сибирских популяций, связывающую народы по месту происхождения и отражающую общее происхождение и большую степень родства. Для популяций, населяющих территорию Сибири, характерна уникальная генетико-демографическая история, которая отражается в распределении блоков IBD как внутри популяции, так и между ними. Анализ блоков IBD значительно дополняет картину в изучении образования и взаимодействия этносов, но не дает однозначных ответов для популяций, развивающихся в условиях сложного этногенеза. За некоторыми исключениями, общие IBD в Сибири лучше объясняются географической близостью популяций, а не их языковой принадлежностью.

Анализ SNP, которые попадали в большее количество IBD-сегментов анализируемых популяций, позволил составить список из 5358 генов. По результатам расчетов, обогащенные этими генами биологические процессы связаны с обнаружением химического раздражителя, участвующего в сенсорном восприятии запаха. Обогащенные найденными генами молекулярные пути связаны с метаболизмом жирных кислот и путем обонятельной трансдукции. При этом анализ литературных данных показал, что некоторые из отобранных генов, встречающиеся в большем количестве блоков IBD сразу в нескольких популяциях, могут играть роль в адаптации человека к факторам окружающей среды и являются перспективными мишенями для дальнейшего изучения.

## Список литературы / References

- Колесников Н.А., Харьков В.Н., Зарубин А.А., Раджабов М.О., Воевода М.И., Губина М.А., Хуснутдинова Э.К., Литвинов С.С., Екомасова Н.В., Штыгашева О.В., Максимова Н.Р., Сухомьясова 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):1261-1275. DOI 10.31857/S0016675821110059.
- [Kolesnikov N.A., Kharkov V.N., Zarubin A.A., Radzhabov M.O., Voevoda M.I., Gubina M.A., Khusnutdinova E.K., Litvinov S.S., Ekomasova N.V., Shtygashева O.V., Maksimova N.R., Sukhomyasova A.L., Stepanov V.A. Features of the genomic distribution of runs of homozygosity in the indigenous population of Northern Eurasia at the individual and population levels based on high density SNP analysis. *Russ. J. Genet.* 2021;57(11):1271-1284. DOI 10.1134/S1022795421110053.]

- Albrechtsen A., Moltke I., Nielsen R. Natural selection and the distribution of identity-by-descent in the human genome. *Genetics*. 2010; 186(1):295-308. DOI 10.1534/genetics.110.113977.
- Amorim C.E.G., Nunes K., Meyer D., Comas D., Bortolini M.C., Salzano F.M., Hünemeier T. Genetic signature of natural selection in first Americans. *Proc. Natl. Acad. Sci. USA*. 2017;114(9):2195-2199. DOI 10.1073/pnas.1620541114.
- Bamodu O.A., Wang Y.H., Ho C.H., Hu S.W., Lin C.D., Tzou K.Y., Wu W.L., Chen K.C., Wu C.C. Genetic suppressor element 1 (GSE1) promotes the oncogenic and recurrent phenotypes of castration-resistant prostate cancer by targeting tumor-associated calcium signal transducer 2 (TACSTD2). *Cancers (Basel)*. 2021;13(16):3959. DOI 10.3390/cancers13163959.
- Bownes R.J., Turnbull A.K., Martinez-Perez C., Cameron D.A., Sims A.H., Oikonomidou O. On-treatment biomarkers can improve prediction of response to neoadjuvant chemotherapy in breast cancer. *Breast Cancer Res.* 2019;21(1):73. DOI 10.1186/s13058-019-1159-3.
- Browning B.L., Browning S.R. Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics*. 2013; 194(2):459-471. DOI 10.1534/genetics.113.150029.
- Browning S.R., Browning B.L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* 2007; 81(5):1084-1097. DOI 10.1086/521987.
- Browning S.R., Browning B.L. High-resolution detection of identity by descent in unrelated individuals. *Am. J. Hum. Genet.* 2010;86(4): 526-539. DOI 10.1016/j.ajhg.2010.02.021.
- Calder P.C. Long chain fatty acids and gene expression in inflammation and immunity. *Curr. Opin. Clin. Nutr. Metab. Care*. 2013;16(4): 425-433. DOI 10.1097/MCO.0b013e3283620616.
- Chilton F.H., Murphy R.C., Wilson B.A., Sergeant S., Ainsworth H., Seeds M.C., Mathias R.A. Diet-gene interactions and PUFA metabolism: A potential contributor to health disparities and human diseases. *Nutrients*. 2014;6(5):1993-2022. DOI 10.3390/nu6051993.
- Christakis N.A., Fowler J.H. Friendship and natural selection. *Proc. Natl. Acad. Sci. USA*. 2014;111(Suppl. 3):10796-10801. DOI 10.1073/pnas.1400825111.
- Espeseth T., Sneve M.H., Rootwelt H., Laeng B. Nicotinic receptor gene *CHRNA4* interacts with processing load in attention. *PLoS One*. 2010;5(12):e14407. DOI 10.1371/journal.pone.0014407.
- Fei Y., Li Y., Chen F. LncRNA-IQCH-AS1 sensitizes thyroid cancer cells to doxorubicin via modulating the miR-196a-5p/PPP2R1B signalling pathway. *J. Chemother.* 2022;1-9. DOI 10.1080/1120009X.2022.2082348.
- Fu F., Nowak M.A., Christakis N.A., Fowler J.H. The evolution of homophily. *Sci. Rep.* 2012;2:845. DOI 10.1038/srep00845.
- Fumagalli M., Moltke I., Grarup N., Racimo F., Bjerregaard P., Jørgensen M.E., Korneliusen T.S., Gerbault P., Skotte L., Linneberg A., Christensen C., Brandslund I., Jørgensen T., Huerta-Sánchez E., Schmidt E.B., Pedersen O., Hansen T., Albrechtsen A., Nielsen R. Greenlandic Inuit show genetic signatures of diet and climate adaptation. *Science*. 2015;349(6254):1343-1347. DOI 10.1126/science.aab2319.
- Grazioplene R.G., Deyoung C.G., Rogosch F.A., Cicchetti D. A novel differential susceptibility gene: *CHRNA4* and moderation of the effect of maltreatment on child personality. *J. Child Psychol. Psychiatry*. 2013;54(8):872-880. DOI 10.1111/jcpp.12031.
- Gusev A., Palamara P.F., Aponte G., Zhuang Z., Darvasi A., Gregersen P., Pe'er I. The architecture of long-range haplotypes shared within and across populations. *Mol. Biol. Evol.* 2012;29(2):473-486. DOI 10.1093/molbev/msr133.
- Han L., Abney M. Identity by descent estimation with dense genome-wide genotype data. *Genet. Epidemiol.* 2011;35(6):557-567. DOI 10.1002/gepi.20606.
- Harris D.N., Ruczinski I., Yanek L.R., Becker L.C., Becker D.M., Guio H., Cui T., Chilton F.H., Mathias R.A., O'Connor T.D. Evolution of hominin polyunsaturated fatty acid metabolism: from Africa to the New World. *Genome Biol. Evol.* 2019;11(5):1417-1430. DOI 10.1093/gbe/evz071.
- Hochreiter S. HapFABIA: Identification of very short segments of identity by descent characterized by rare variants in large sequencing data. *Nucleic Acids Res.* 2013;41(22):e202. DOI 10.1093/nar/gkt1013.
- Horrocks L.A., Yeo Y.K. Health benefits of docosahexaenoic acid (DHA). *Pharmacol. Res.* 1999;40(3):211-225. DOI 10.1006/phrs.1999.0495.
- Howe P., Meyer B., Record S., Baghurst K. Dietary intake of long-chain  $\omega$ -3 polyunsaturated fatty acids: contribution of meat sources. *Nutrition*. 2006;22(1):47-53. DOI 10.1016/j.nut.2005.05.009.
- Kimura T., Kobayashi T., Munkhbat B., Oyungereel G., Bilegtsaikhan T., Anar D., Jambaldorj J., Munkhsaikhan S., Munkhtuvshin N., Hayaishi H., Oka A., Inoue I., Inoko H. Genome-wide association analysis with selective genotyping identifies candidate loci for adult height at 8q21.13 and 15q22.33-q23 in Mongolians. *Hum. Genet.* 2008; 123(6):655-660. DOI 10.1007/s00439-008-0512-x.
- Kiritits D., Chmel N., Arnold A.W., Jakob T., Bruckner-Tuderman L., Has C. Novel and recurrent *AAGAB* mutations: clinical variability and molecular consequences. *J. Invest. Dermatol.* 2013;133(10): 2483-2486. DOI 10.1038/jid.2013.171.
- Kuhlwil M., Boeckx C. A catalog of single nucleotide changes distinguishing modern humans from archaic hominins. *Sci. Rep.* 2019; 9(1):8463. DOI 10.1038/s41598-019-44877-x.
- Marszalek J.R., Lodish H.F. Docosahexaenoic acid, fatty acid-interacting proteins, and neuronal function: breastmilk and fish are good for you. *Annu. Rev. Cell Dev. Biol.* 2005;21:633-657. DOI 10.1146/annurev.cellbio.21.122303.120624.
- Palamara P.F., Lencz T., Darvasi A., Pe'er I. Length distributions of identity by descent reveal fine-scale demographic history. *Am. J. Hum. Genet.* 2012;91(5):809-822. DOI 10.1016/j.ajhg.2012.08.030.
- Pugach I., Matveev R., Spitsyn V., Makarov S., Novgorodov I., Osakovsky V., Stoneking M., Pakendorf B. The complex admixture history and recent southern origins of Siberian populations. *Mol. Biol. Evol.* 2016;33(7):1777-1795. DOI 10.1093/molbev/msw055.
- Tucci S., Vohr S.H., McCoy R.C., Vernot B., Robinson M.R., Barbieri C., Nelson B.J., Fu W., Purnomo G.A., Sudoyo H., Eichler E.E., Barbujani G., Visscher P.M., Akey J.M., Green R.E. Evolutionary history and adaptation of a human pygmy population of Flores Island, Indonesia. *Science*. 2018;361(6401):511-516. DOI 10.1126/science.aar8486.
- Yin L.L., Li J.M., Zhou Z.M., Sha J.H. Identification of a novel testis-specific gene and its potential roles in testis development/spermatogenesis. *Asian J. Androl.* 2005;7(2):127-137. DOI 10.1111/j.1745-7262.2005.00041.x.

#### ORCID ID

N.A. Kolesnikov [orcid.org/0000-0001-8855-577X](https://orcid.org/0000-0001-8855-577X)  
V.N. Kharkov [orcid.org/0000-0002-1679-2212](https://orcid.org/0000-0002-1679-2212)  
K.V. Vagaitseva [orcid.org/0000-0003-4877-9749](https://orcid.org/0000-0003-4877-9749)  
A.A. Zarubin [orcid.org/0000-0001-6568-6339](https://orcid.org/0000-0001-6568-6339)  
V.A. Stepanov [orcid.org/0000-0002-5166-331X](https://orcid.org/0000-0002-5166-331X)

**Благодарности.** Исследование выполнено за счет гранта Российского научного фонда № 22-64-00060 (<https://rscf.ru/project/22-64-00060/>).

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 21.10.2022. После доработки 17.01.2023. Принята к публикации 24.01.2023.