

**Issue 2 / Часть 2**

**Articles / Статьи**



*В. Н. Харьков\**, *А. Ю. Урманчиева†*, *Н. А. Колесников‡*, *А. А. Зарубин§*,  
*Л. В. Валихова\*\**, *И. Ю. Хитринская††*, *И. А. Грунтов‡‡*, *И. М. Егоров§§*,  
*В. М. Лемская\*\*\**, *А. В. Дыбо†††*, *В. А. Степанов‡‡‡*

\* Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр РАН; vladimir-kharkov@medgenetics.ru

† Институт лингвистических исследований РАН; urmanna@yandex.ru

‡ Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр РАН; nikita.kolesnikov@medgenetics.ru

§ Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр РАН; aleksei.zarubin@medgenetics.ru

\*\* Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр РАН; larisa.valikhova@medgenetics.ru

†† Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр РАН; khitrinskaya@tnimc.ru

‡‡ Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр РАН; Институт языкознания РАН, Москва; groutov@mail.ru

§§ Институт финно-угроведения/уралистики Университета Людвиг Максимилиана, Мюнхен; i.m.jegorow@gmail.com

\*\*\* ФГБОУ ВО «Томский государственный педагогический университет», Томск; lemskaya@tspu.edu.ru

††† Институт языкознания РАН, Москва; adybo@mail.ru

‡‡‡ Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр РАН; vadim.stepanov@medgenetics.ru

## **Популяционная генетика и лингвистическая компаративистика: филогения и филогеография гаплогруппы N1a2b и лингвистическая предыстория самодийцев<sup>1</sup>**

В работе проведено сопоставление данных этногенетики по гаплогруппе N1a2b Y-хромосомы и лингвистики самодийской языковой группы. Гаплогруппа N1a2b унаследована современными народами от более раннего населения уралоидного антропологического типа, заселявшего территорию Южной Сибири. При детальной реконструкции филогении этой гаплогруппы показана этноспецифичная структура сублиний N1a2b. Лингвистическая предыстория самодийцев реконструируется путем анализа и стратификации лексических заимствований. Релевантные для данной статьи результаты по лингвистической филогенетике получены методом ручного анализа общих инноваций (для самодийских языков) и байесовской лексикостатистики (для тюркских языков). Между результатами генетического и лингвистического анализа обнаружены следующие соответствия. Регион начального распространения гаплогруппы N1a2b в целом соотносится с прародиной самодийцев, установленной по лингвистическим данным. Специфическая близость лексики маторского и ненецких языков находит подтверждение в истории сублинии N1a2b1b1. Эта сублиния представлена у тундровых ненцев (но не у энцев) и у хакасов качинцев, на язык которых как раз и могли перейти маторы. Наличие сублинии N1a2b1b1 у казымских хантов объясняется их вторичными контактами с ненцами, о чем также свидетельствуют ненецкие культурные заимствования. Сублиния N1a2b1b2a1 обнаруживается у тувинцев и монголов, что коррелирует с тувино-монгольскими языковыми контактами. Полученные современные совместные результаты данных популяционной генетики и лингвистики хорошо дополняют друг

---

<sup>1</sup> Исследование выполнено за счет гранта Российского научного фонда № 22-64-00060, <https://rscf.ru/project/22-64-00060/>.

друга и дают достоверное подтверждение языковому родству и генетико-демографическим процессам в популяциях Сибири.

*Ключевые слова:* этногенетика; Y-хромосома; лингвистическая предыстория; самодийские народы.

## Введение

Анализ структуры генофондов популяций человека с использованием диаллельных и микросателлитных маркеров Y-хромосомы является одним из наиболее современных и результативных методов изучения генетического разнообразия популяций человека, позволяет проводить комплексную оценку популяционной истории, выполнять филогенетические реконструкции отдельных линий, филогеографические исследования гаплогрупп, анализ миграционных процессов. Несмотря на бурное развитие методов, позволяющих изучать генофонды популяций на уровне полных геномов, использование маркеров Y-хромосомы по-прежнему является актуальным для решения целого ряда вопросов популяционной генетики (Underhill et al. 2000; Jobling & Tyler-Smith 2003; Karmin et al. 2015; Watanabe et al. 2019; Esteller-Cucala et al. 2023).

Территория Южной и Западной Сибири — уникальный регион с точки зрения разнообразия естественно-географических условий и этнокультурного многообразия проживающего здесь населения. Одним из достаточно древних и относительно широко распространенных на этой территории по данным антропологии, этнологии и лингвистики является самодийский компонент, который связан с одной из основных миграционных волн в Южной, Западной Сибири и Арктике.

Современные народы, говорящие на самодийских языках, занимают значительное пространство: от Белого моря до Таймыра на севере, а также среднее и нижнее Приобье. В настоящее время это лесные и тундровые ненцы, лесные и тундровые энцы, нганасаны и селькупы, разделенные на несколько территориальных групп. Еще относительно недавно некоторые самодийские этносы или субэтносы (камасинцы, маторы, койбалы, карагасы, тайгийцы, кашинцы) проживали в Саянах. Сейчас они полностью растворились в основном тюркском населении (Дульзон 1960). Поскольку эти самодийцы сыграли существенную роль в формировании тюркских этносов саянского региона, генетические и лингвистические данные, относящиеся к популяциям Южной Сибири, необходимо рассматривать в комплексе для самодийских и тюркских народов.

Проблема корректного соотнесения данных генетики и лингвистики, их объединения в комплексный междисциплинарный подход является чрезвычайно актуальной, поскольку именно попытки сопоставления генетических и лингвистических данных зачастую вызывают пристальное внимание специалистов широкого круга направлений, связанных с реконструкцией истории населения различных регионов.

Процессы ассимиляции и популяционной дифференциации также сыграли свою роль в формировании современных народов Сибири. Самодийцы относятся к уральскому антропологическому типу, который по своим характеристикам занимает промежуточное положение между монголоидной и европеоидной расами (Рогинский & Левин 1978). Внутри самодийской группы различия по антропологическому типу сильно варьируются. К настоящему времени исследования, предполагающие выяснение соотношения лингвистической и генетической истории самодийцев, проводились на ограниченном материале и с не очень высокой точностью (Волков 2013; Karafet et al. 2018). Все самодийские языки можно разделить на две группы: северосамодийские (тундровый

и лесной ненецкий, тундровый и лесной энецкий, нганасанский, а также маторский) и южносамодийские (селькупский язык, представленный пятью диалектами, а также камасинский).

Многие сибирские этносы антропологически и генетически довольно гетерогенны (например, алтайцы, тувинцы, хакасы, ненцы и эвенки), а исследованные до настоящего времени выборки не отражают всей многокомпонентности генофонда, характеризуя лишь часть имеющегося разнообразия. Исследование направлено на анализ состава генофондов современного коренного населения Южной и Западной Сибири.

Исследование состава генофондов современных популяций человека по различным генетическим маркерным системам, реконструкция их происхождения, скоррелированные с современными представлениями об истории языков данных этносов, имеют важное значение для комплексного понимания формирования различных этно-территориальных групп. Проблематика, касающаяся анализа структуры и соотношения различных генетических компонент у народов Сибири, генеалогического ветвления их языков, а также межязыковых заимствований, как материальных, так и структурных, несмотря на относительно высокий уровень изученности, имеет ряд невыясненных вопросов. Генетика в этой связи дает важные результаты для междисциплинарного исследования, поскольку развитие современных методов анализа популяционного генофонда позволило вывести этногенетические исследования на совершенно новый уровень.

История этнических групп, населяющих Сибирь, была насыщена различными событиями и претерпела изменения популяционного, языкового и культурного состава за счет миграций различных популяций, относящихся к разным языковым семьям и группам (енисейской, самодийской, угорской, тюркской, монгольской, тунгусо-маньчжурской, юкагирской и др.). 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.

Для маркеров Y-хромосомы показан наибольший уровень генетических отличий между популяциями в сравнении с любыми другими генетическими системами. Это приводит к значительной степени территориальной и межэтнической генетической дифференциации и высокому уровню специфичности гаплотипов. Высокие показатели межпопуляционного и межгруппового разнообразия Y-хромосомы делают ее маркеры наиболее дифференцирующими из всех, применяемых в популяционной генетике и геномике человека, и подчеркивают эффективность анализа гаплотипов Y-хромосомы для выявления различий между популяциями и при изучении миграций (Duranloup et al. 2003; Barbujani et al. 2003). Кроме характеристики генетического разнообразия, дифференциации и состава популяционного генофонда, исследование маркеров Y-хромосомы очень важно для филогенетических реконструкций различных гаплогрупп.

Гаплогруппа N1a2b (ранее обозначавшаяся как N2 и N1b), неравномерно распределенная в популяциях Сибири и Восточной Европы, представляет собой маркер доисторической связи между сибирскими и европейскими популяциями. Она связана с древней общностью населения уралоидного антропологического типа, заселявшего территории к востоку и западу от Уральских гор. Наибольшая частота гаплогруппы N1a2b приходится на территорию Западной Сибири и Алтае-Саянского региона (Zerjal et al. 1997; Karmin et al. 2015; Puumäe et al. 2016; Karafet et al. 2018). Именно эта гаплогруппа доминирует у всех этносов, принадлежащих к группе самодийских языков уральской языковой семьи, в Западной Сибири — 92% у нганасан, 78% у энцев и 74% у тундровых ненцев (Karafet et al. 2002; Kharkov et al. 2021). Ее частота в Южной Сибири, в частности,

у хакасов и тувинцев, также отражает вклад в их современный генофонд самодийских этносов, заселявших эту территорию ранее (см. Таблицу 1).

Распространение N1a2b на территории Европы связано с расселением предков современных финно-угорских народов, поскольку данная гаплогруппа характерна для этих народов Урала и Поволжья, где она составляет более 20% в некоторых популяциях (удмурты, коми) (Tambets et al. 2004; Karmin et al. 2015). Присутствие гаплогруппы N1a2b у татар, башкир и русских связано с включением в состав этих народов финно-угорского генетического субстрата. Западной границей ареала этой гаплогруппы является Финляндия, где она обнаруживается у единичных образцов (Lappalainen et al. 2006; Rootsi et al. 2007). Предположительным местом возникновения гаплогруппы N1a2b является Южная Сибирь.

В диахронической лингвистике существует две группы методов, дающих результаты, релевантные для истории этнических групп и подходящие для непосредственного сравнения с результатами других наук, в первую очередь популяционной генетики, археологии и антропологии. Первая группа методов связана с изучением дивергентных процессов, установлением языкового родства и построением филогенетических деревьев. Вторая группа методов сосредотачивается вокруг изучения конвергентных процессов, т. е. заимствований и эффектов ареальной диффузии.

Филогенетические деревья в сравнительно-историческом языкознании бывают двух типов: качественные и количественные. Качественное дерево строится на основании выявленных в родственных языках/диалектах процессов фонетических, морфологических и лексических изменений. Последовательность прохождения этих процессов может быть упорядочена по времени на основании результатов, наблюдаемых в каждом из языков. Сопоставление последовательностей дает возможность определить, какие из процессов проходили в общем состоянии этих языков, а какие могли идти уже только после разделения общего состояния на отдельные группы и далее на отдельные языки. Такое дерево отражает относительную хронологию языковых изменений, но допускает абсолютную датировку только по внешним по отношению к процедуре данным (письменные памятники, заимствования между языками с письменной традицией, лексика, связанная с датированными технологиями и т. п.). Количественные деревья строятся лексикостатистическим методом, основанным на количественных сходствах фрагментов базового словаря сравниваемых языков. В пределах базисной лексики предполагается относительно равномерная скорость внутренних замен, что допускает абсолютную датировку (Старостин 2007). Лексикостатистические матрицы обрабатываются алгоритмами, практически идентичными с теми, что используются для филогенетического анализа на основе ДНК. Лексикостатистические методы могут быть применены к любым достаточно документированным языкам/диалектам, поскольку основаны на анализе универсальных наборов концептов. Их основное ограничение состоит в том, что они игнорируют различие между синапоморфией и симплезиоморфией.

Взаимодействие этнических групп в прошлом может быть реконструировано на основании заимствований. Заимствоваться могут практически любые элементы языковой системы, но наиболее надежно идентифицируются лексические заимствования. Лексические заимствования могут стратифицироваться с привлечением сведений о топологии филогенетического древа и относительной хронологии фонетических процессов. Количество заимствований и распределение по семантическим группам дает сведения об интенсивности и характере взаимодействий этнических групп. Различаются адстратные и субстратные заимствования. Адстратные заимствования указывают на культурный обмен и торговые связи между этническими группами. Они охватывают небольшое коли-



чество семантических полей и редко проникают в базисную лексику. Субстратные заимствования указывают, что в истории данной популяции имел место языковой сдвиг, т. е. переход с одного языка на другой. Они охватывают многие семантические поля, в том числе проникая и в базисную лексику; могут иметь характерные фонотактические паттерны, контрастирующие с унаследованной от праязыка лексикой. Данные топонимики также часто могут указывать на наличие субстрата (Дульзон 1950: 175–183; Дульзон 1953; Дульзон 1954: 59–94; Дульзон 1973).

С точки зрения генетики при распаде популяции и расширении кластера гаплотипов в одной из дочерних популяций этот кластер будет ограничен только этой дочерней популяцией (при условии, что взаимосвязь между популяциями незначительна). Следовательно, оценка его датировки будет определяться не самим разделением (т. е. расхождением популяции), а событием расширения кластера, которое могло произойти в любой момент эволюции дочерней популяции, во время или после разделения.

С позиции лингвистики инновации, которые обнаруживаются в одном языке/диалекте, но отсутствуют в другом, возникают после физического разделения протопопуляции носителей языка на частично изолированные субпопуляции. Лингвистическая датировка разделения, таким образом, указывает на момент времени, когда языковые различия были закреплены, а не на возникновение этого процесса. Следовательно, начало дивергенции у населения могло иметь место за несколько поколений до времени, определяемого по лингвистическим данным.

Целью настоящего исследования является взаимная верификация результатов, которая становится возможна благодаря тому, что и генетические, и лингвистические исследования позволяют строить филогенетические деревья и датировать точки отделения ветвей (Balanovsky et al. 2011; Mallory et al. 2019). Вычисление временных промежутков по данным лингвистики основано на подсчете расхождений, образовавшихся после разделения языков выборки. Аналогичным образом, временные отрезки по данным генетики Y-хромосомы рассчитываются исходя из количества мутаций, возникших после начала расширения гаплогруппы.

В работе выявлен вклад самодийского генетического субстрата на основе генотипирования различных SNP и STR маркеров Y-хромосомы, филогенетического и филогеографического анализа линии N1a2b Y-хромосомы, маркирующей самодийский генетический субстрат. На основании данных о популяционно-специфичной структуре гаплогрупп Y-хромосомы выполнена детальная характеристика и реконструкция формирования современного генетического разнообразия исследуемых популяций, что позволяет точнее реконструировать историю формирования генофондов современных сибирских народов.

### Экспериментальная часть

**Материалы.** В работе использовались образцы ДНК мужчин из различных популяций. Забор первичного биологического материала (венозной крови) у доноров производили с соблюдением процедуры письменного информированного согласия на проведение исследования. На каждого донора составлялась анкета с краткой родословной, указанием этнической принадлежности и мест рождения предков. Индивида относили к данной этнической группе на основании его собственной этнической идентификации, его родителей и места рождения. Для анализа YSTR-гаплотипов гаплогруппы N1a2 были использованы 592 образца ДНК мужчин принадлежащих к ее основной азиатской



ветви: ненцев (N=256), хантов (N=48), хакасов сагайцев (N=88), хакасов качинцев (N=43), чулымских тюрков (N=21), шорцев (N=11), тувинцев (N=110), якутов (N=2), бурятов (N=1), челканцев (N=3) и калмыков (N=9).

Методы исследования гаплотипов Y-хромосомы. Для изучения состава и структуры гаплогрупп Y-хромосомы в исследование были включены две системы генетических маркеров: диаллельных локусов, представленных SNP, и полиаллельных высоковариабельных микросателлитов (YSTR). С помощью 445 SNP-маркеров определяли принадлежность образцов к различным сублиниям N1a2b. Классификация гаплогрупп дана в соответствии с данными Международного общества генетической генеалогии (ISOGG). Анализ STR-гаплотипов внутри гаплогрупп проводили с использованием 46 микросателлитных маркеров нерекombинирующей части Y-хромосомы (YSTR) (DYS19, 385a, 385b, 388, 389I, 389II, 390, 391, 392, 393, 426, 434, 435, 436, 437, 438, 439, 442, 444, 445, 448, 449, 456, 458, 460, 461, 481, 504, 505, 518, 525, 531, 533, 537, 549, 552, 570, 576, 635, 643, YCAIIa, YCAIIb, GATAH4.1, Y-GATA-A10, GGAAT1B07, DYF387; Kharkov et al. 2020). STR-маркеры генотипировали с помощью капиллярного электрофореза на генетическом анализаторе ABI Prism 3730 и Нанофор-05. Генотипирование SNP-маркеров проводили с помощью ПЦР и последующего анализа фрагментов ДНК с помощью ПДРФ-анализа. Экспериментальные исследования проведены на базе Центра коллективного пользования научно-исследовательским оборудованием «Медицинская геномика» (НИИ медицинской генетики Томского НИМЦ). Построение медианных сетей гаплотипов Y-хромосомы проводили с использованием программы Network v. 10.2.0.0 ([www.fluxus-engineering.com](http://www.fluxus-engineering.com)) по методу медианных сетей Бандельта (Bandelt et al. 1999). Оценку возраста генерации наблюдаемого разнообразия гаплотипов в гаплогруппах проводили методом ASD (Zhivotovsky et al. 2004) на основании средних квадратичных отличий в числе повторов между всеми маркерами. Возраст поколения принимался равным 30 годам, темп мутирования 0,0021 на локус на поколение.

Полногеномное секвенирование мужчин с гаплогруппой N1a2b выполнено методом массового параллельного секвенирования NGS (Next Generation Sequencing) на 84 образцах ДНК, принадлежащих к азиатской линии N1a2b. Секвенирование с покрытием  $\times 30$  проводили на приборе Illumina NovaSeq. Выполнен биоинформатический анализ первичных данных секвенирования геномов, включающий оценку качества полученных данных секвенирования для проанализированных образцов. Анализ данных полногеномного секвенирования образцов показал, что качество сгенерированных прочтений было достаточно высоким. Эти геномные данные были включены в базу данных для ее расширения и дальнейшей биоинформатической обработки общего массива образцов. Анализ и извлечение генотипов из данных секвенирования проводились с использованием Genome Analysis Toolkit 4 (<https://software.broadinstitute.org/gatk>).

Для привязки генотипов образцов к уже описанным SNP и обнаружения новых мутантных позиций было применено два подхода:

1) фильтрация данных с пропущенными в различных образцах позициями, уровнем покрытия и ошибочных аллелей в одной позиции. К сожалению, в результате этого подхода было отсеяно очень большое количество SNP, включенных в список, особенно это касалось многих терминальных SNP. Это удаляло также и много не описанных ранее SNP, которые очень важны для более точной детализации филогении различных линий;

2) после этого были удалены все позиции без мутаций в файлах этих гаплогрупп. Фильтрация данных в автоматическом режиме не проводилась. Далее эти данные в табличном виде были перенесены в Excel и начат отбор мутантных позиций и отсева оши-

бочных вручную на всем наборе образцов мужчин, принадлежащих к этим гаплогруппам. Это позволило оставить практически все привязанные к базе SNP, обозначаемые различными шифрами (A, AM, B, BY, C, CTS, F, FGC, FT, L, M, MF, VL, S, SUR, TY, Y, YP, Z, ZS), а также существенно увеличить количество новых YSNP.

Полученные после отбора данные по позициям YSNP для всех секвенированных образцов были использованы для создания отдельных vcf-файлов с достоверными мутантными позициями. С помощью программ Y-LineageTracker и StarBeast3 были построены филогенетические деревья для N1a2 на уровне отдельных образцов (см. рисунки 1, 2 в онлайн-приложении к настоящей работе). Построение филогенетических деревьев с использованием пакета программ BEAST основано на использовании байесовской статистики (метода Монте-Карло для марковских цепей), вовлекающей сложный математический аппарат. Эти деревья более точно отображают разделение различных ветвей внутри гаплогрупп. Все построенные деревья полностью согласуются с количеством YSNP в отдельных линиях и демонстрируют разделение образцов на этноспецифичные группы.

Подробная филогения гаплогруппы N1a2b построена по результатам NGS-секвенирования полных геномов различных образцов мужчин. Данные по Y-хромосомным SNP получены после биоинформатической обработки данных, привязки описанных ранее мутаций к указанным позициям и отбора новых терминальных SNP для более точного разделения сублиний внутри этой гаплогруппы. Детальное разделение на отдельные ветви показывает специфичность разных линий для всех исследованных этносов, субэтносов и популяций.

Для гаплогруппы N1a2b по всем секвенированным образцам была выполнена точная привязка к детальным этноспецифичным вариантам отдельных сублиний Y-хромосомы, за счет принадлежности к уже описанным мутантным SNP. У многих образцов мужчин из этих гаплогрупп, относящихся к сибирским популяциям, было обнаружено большое количество новых, не описанных ранее SNP. В общем массиве маркеров они составляют около 20%, но для различных образцов их доля в терминальных SNP порой доходит до 100%. Эти результаты позволяют гораздо точнее разделить филогенетическую структуру этой линии, детализировать их происхождение.

**Лингвистические методы.** Лингвистический подход включал в себя исследование заимствования лексики между различными уральскими идиомами (тундровый ненецкий и казымский хантыйский) методами сравнительно-исторического языкознания и контактной лингвистики (Бурлак & Старостин 2005; Campbell 2013), анализ положения маторского языка внутри самодийской подгруппы путем изучения общих инноваций (Campbell 2013), объединяющих маторский язык с северносамодийскими и южносамодийскими языками, анализ языковых контактов тюрков и самодийцев, тюрков и енисейцев, тюрков (тувинцев и чулымцев) и монголов, что позволяет датировать контакты, а также получать данные о древних миграциях носителей этих языков (Erps 2015).

## Результаты и обсуждение

При определении прародины гаплогрупп необходимо использовать несколько критериев. Во-первых, одним из критериев является высокая частота гаплогруппы, указывающая на возможное место ее происхождения. Но максимальная частота не может служить однозначным показателем центра зарождения гаплогруппы, поскольку и на

периферии ареала гаплогруппы популяции могут испытывать сильное влияние дрейфа, что может привести к эффекту основателя по данной линии у конкретного этноса или группы популяций. Во-вторых, наличие предковой по отношению к исследуемой линии и близкородственных сестринских гаплогрупп также указывает на прародину. В большинстве случаев местом возникновения гаплогруппы может быть регион с высокой частотой базальных ветвей этой клады. Третьим критерием является разнообразие YSTR-гаплотипов в пределах изучаемой гаплогруппы. Территорию с наибольшим разнообразием какой-то родственной группы гаплотипов принято рассматривать как вероятную зону их общего происхождения. Там, где гаплогруппа впервые достигла значимой частоты и откуда она затем расширяла свой ареал, должны были накопиться многочисленные мутации и сформироваться множество дочерних гаплотипов. В месте происхождения показатели разнообразия должны быть максимальны. В отдаленных популяциях гаплотипическое разнообразие должно быть мало, поскольку с миграционными потоками они получают не репрезентативную выборку гаплотипов предковой популяции, а лишь случайную часть спектра.

Построенное филогенетическое древо азиатской гаплогруппы N1a2b (рисунок 3 в онлайн-приложении к настоящей работе) показывает разделение образцов на различные ветви в соответствии с теми специфичными сублиниями, которые определяются терминальными SNP. Близость секвенированных образцов мужчин, принадлежащих к этой гаплогруппе, полностью совпадает с их этнической привязкой и данными их анкетирования.

#### Сублинии гаплогруппы N1a2b в популяциях Алтае-Саян и лингвистические данные

Самая древняя ветвь этой линии присутствует в популяциях Алтая и Саян N1a2b2a2a~ (B478, Z35080, Z35081, Z35082, Z35083, Z35084, YP3079 (xZ35088, B168, B169)). К ней принадлежат образцы челканцев, хакасов-сагайцев и шорцев. Большинство из них формируют попарные ветви, с различными терминальными SNP, объединяющие эти образцы в соответствии с их этническим происхождением. Относительно небольшая частота этой сублинии в популяциях северных алтайцев обусловлена тем, что этот базовый субстратный самодийский компонент был достаточно сильно замещён в течение последних полутора тысяч лет тюркскими и монгольскими группами, но его присутствие в генофонде современных популяций южно-сибирских этносов демонстрируют наличие вклада древнего самодийского населения. Присутствие этой линии с этноспецифичными терминальными SNP и отличающимися гаплотипами говорит о их наследовании от достаточно древних общих предков, входивших в состав раннего населения этих территорий.

Близость образцов, принадлежащих к этнически родственным популяциям, может быть сопоставлена с лингвистической классификацией. Генетические образцы, для которых можно предполагать самодийский субстрат, были взяты у носителей тюркских языков.

В Южной Сибири в исторически засвидетельствованный период представлены следующие группы тюркских языков (связанные с затронутыми в исследовании популяциями): саянская (тувинский с диалектами и тофаларский); 3-группа (среднечулымский, мрасский диалект шорского, хакаский с диалектами, а именно кызыльским, качинским, сагайским, шорским); алтайская, делящаяся на южно-алтайскую подгруппу (алтайкижи, теленгитский, телеутский), и северно-алтайскую (кумандинский, челканский, ту-

баларский; возможно, также кондомский диалект шорского). Алтайская группа, по последним данным (Egorov et al. 2022), представляет собой раннее ответвление кыпчакской группы. С антропологической точки зрения традиционная культура алтайцев больше связана со степью, чем культура саянской и 3-групп, так что алтайцы могут быть более новыми пришельцами в Южную Сибирь, чем носители первых двух групп тюркских языков. Датировка появления тюрков в Южной Сибири связывается обычно с алтайским походом Маодуня (шаньюй сюнну, 201 г. до н.э., см. напр., Хелимский 2000). Неплохо датируются, например, контакты тюрков с енисейцами. Если ориентироваться на реконструкцию праенисейской фонетики в Старостин 1982 и праенисейский лексикон, реконструированный в Старостин 1995, а также расширенный и учитывающий новейшую литературу вариант его, опубликованный в интернете по адресу [starlingdb.org](http://starlingdb.org) (ТоВ), то распад праенисейского языка датируется глоттохронологически II в. до н.э., когда выделяются две ветви — «кетская» и «коттская». От «кетской» отделяется в V в. н.э. пумпокольский язык, затем в XII в. н.э. распадаются кетский и югский языки. Коттская группа в VI в. н.э. распадается на коттский и аринский.

Соответственно выделяется три пласта старых тюркских заимствований в енисейские: (1) пратюркские заимствования в праенисейский (не позднее рубежа эры); (2) заимствования, происшедшие не позднее VI в. н.э. (только в «коттской» ветви); (3) заимствования, происшедшие не позднее XII в. н.э. (только в кетско-югской подветви), см. Дыбо 2007. Это предполагает по крайней мере три волны контактов — и, возможно, три волны переселений.

#### Сублинии гаплогруппы N1a2b у ненцев

Вторая специфичная сублиния N1a2b1 (B478, Z35080, Z35081, Z35082, Z35083, Z35084 (xZ35088, B169)), являющаяся чуть менее древней, обнаружена у тундровых ненцев восточной части ЯНАО. К ней принадлежат исключительно образцы ненцев, которые по мужской линии имеют принадлежность к энецким родам, вошедшим в состав тундровых ненцев. Эта линия является наследием более раннего населения тундровой территории, которое появилось там до прихода предков ненцев, и эта более ранняя миграционная волна самодийцев обладала другим составом сублиний и гаплотипов. Ещё одна очень специфическая сублиния N1a2b1b~ обнаружена у двух якутов и одного бурята.

#### Сублинии гаплогруппы N1a2b у тувинцев

Следующая ветвь присутствует у тувинцев, якутского эвенка и калмыков. У тувинцев она разделена на три сублинии, отличающиеся по терминальным SNP и кластерам гаплотипов. С максимальной частотой у тувинцев представлен ее вариант N1a2b1b2b1 (B178, PF3415, Z35147, Z35149, Z35152). К нему относятся два образца южных алтайцев. Её возраст составляет 1442 года (SD=368 лет). Четыре образца тувинцев принадлежат к очень редкой параллельной линии N1a2b1b2a1~ (B228, Z35125, Z35127, Z35128). Ранее она была обнаружена у монголов (Пумäe et al. 2016). К третьей тувинской сублинии (xB175, Z35117, Z35118) относятся 10 образцов. Монголы и тувинцы находились в интенсивных контактах, как минимум, с XIII века. В XIV–XVII веках наблюдалось сильное влияние ойратов (ближайших родственников современных калмыков), позже преобладало влияние халха-монгольского языка. Заимствования шли, главным образом, в одном направлении — из монгольского языка в тувинский. Тувинцы жили, фактически, в ситуации билингвизма, что отразилось в значительном объеме лексики, заимствованной из монгольских языков

(более 1500 лексем по оценке в Khabtagaeva 2006). Интенсивность контактов и старый билингвизм, сохраняющийся и до настоящего времени, подтверждаются проникновением монгольских заимствований в базисную лексику. В различных диалектах тобаской группы монголизмы отмечены в следующих слотах 110-словной версии списка Сводеша: 'все', 'земля', 'зеленый', 'дождь', 'песок', 'семя', 'маленький', 'хвост', 'женщина'. Что касается самодийских контактов языков саянской группы (тувинского и тофаларского), то известно существенное количество тюркизмов, определяемых как саянские, в матор-тайги-карагасской группе, и небольшое количество заимствований из матор-тайги-карагасского в тувинском, больше в тофаларском (см. ниже).

### Сублинии гаплогруппы N1a2b у чулымских тюрков

Следующая специфичная линия выявлена у чулымских тюрков (VL65, Z35095, Z35099, Z35102) и является наиболее близкой к сестринской ей линии N1a2b1b1 у хакасов качинцев. Медианная сеть гаплотипов (см. рисунок 4 в онлайн-приложении) демонстрирует у чулымцев звездообразную филогению с недавним эффектом основателя и доминированием по частоте предкового гаплотипа. Кластер чулымских гаплотипов равноудален от всех качинских сеоков. Возраст этого кластера у чулымцев составил 667 лет (SD=194 года), что достаточно хорошо совпадает с предположительным временем расселения предков чулымцев на север.

При исследовании истории тюркского населения среднего течения р. Чулым А. П. Дульзон обращался прежде всего к данным топонимики (Дульзон 1950), а также номенклатуре родства и свойства (Дульзон 1954) в сравнении с материалом языков других народов Сибири. Лингвистические данные, обработанные А. П. Дульзоном, указывают на наличие субстратов в языке чулымских тюрков — прежде всего енисейского (кетского) и самодийского (селькупского). Распространение тюркского языка в среднем течении р. Чулым, с точки зрения А. П. Дульзона (Дульзон 1973), происходило в три этапа: начиная с VII–VIII вв. с верховьев реки по направлению вниз; в XII–XIII вв. с юга вниз по р. Кие; не позднее XV в. — с юго-запада вверх по р. Чулыму (либо вниз по р. Яе). По данным интерпретации гидронимов среднего Причулымья, тюркизация этой территории произошла путем наслоения на три кетоязычные народности (ассаны, пумпоколы, арины; Дульзон 1973). Помимо этого, на основе комплекса этнолингвистических и археологических данных отмечается «преемственная связь причулымского населения» с селькупам (Дульзон 1953; Дульзон 1954).

Следует отметить наблюдения о чулымско-тюркских монголизмах. В. И. Рассадин (1980) на материале, собранном В. В. Радловым (нижнечулымские диалекты) и А. П. Дульзоном (диалектная принадлежность не указывается), устанавливает пласт монгольской лексики в чулымско-тюркском языке. Вопрос о монгольском субстрате в данном случае не ставился; пласт монголизмов представляет собой единичные лексические единицы (не более 50), проникшие в чулымский язык в XIII–XIV вв., либо, возможно, через более поздние контакты с джунгарскими ойратами (Рассадин 1980).

### Сублинии гаплогруппы N1a2b у хантов и ненцев

Все остальные исследованные образцы разделяются на две крупные субклады. К первой из них (N1a2b1b1-Y68212, Y70717, Y70315, Y70327) относятся ханты и хакасы-качинцы, которые разделились между собой достаточно давно. Эта сублиния хантов близка к хакасам-качинцам (Z35093, Z35097, Z35103; Valikhova et al. 2022). К гаплогруппе N1a2b2b1~

(Z35076) относятся три образца казымских хантов. Медианная сеть гаплотипов (см. рисунок 5 в онлайн-приложении) также демонстрирует у хантов звездообразную филогению с недавним эффектом основателя и преобладанием по частоте предкового гаплогруппы. Возраст этого кластера у хантов составил 858 лет (SD=338 лет).

Вторая сублиния также разделяется на две ветви. Первая N1a2b1b1a~ (B170, B171, Z35091, Z35092, Z35093, Z35095, Z35096, Z35097, Z35099, Z35102, Z35103, Z35105 и др. (xZ35104, Z35106, Z35107)). К ней принадлежат практически подавляющее большинство ненцев фратрии Харючи и многие образцы, относящиеся к фратрии Вануйто. Вторая сублиния N1a2b1b1b (B172, Z35108, Z35109, Z35112 и др.) также присутствует у ненцев и хантов, которая унаследована ими от ненецких предков по мужской линии. К ней принадлежат все ненцы из фратрии Вануйто, относящиеся к родам Вануйто, Пуйко и Юнгат, и хантыйского по происхождению рода Пурунгуи.

Структура этого филогенетического дерева и соотношение возраста отдельных сублиний хорошо согласуется с данными антропологии и лингвистики, которые описывают территорию Саян как исходную территорию расселения самодийских групп.

### Филогения гаплогруппы N1a2b и лингвистическая история самодийцев

На данном филогенетическом дереве с точки зрения лингвистической истории самодийцев наиболее примечательной представляется линия N1a2b1b1. Как можно видеть по схеме, представленной на Рисунке 3 «Разделение сублиний гаплогруппы N1a2b Азиатской ветви по этноспецифичным терминальным SNP», эта линия представлена у ненцев, хантов и хакасов-качинцев. Эта линия обнаружена у казымских хантов, что, вероятно, можно объяснить их контактами с тундровыми ненцами. Лингвистически эти контакты отражаются в достаточно значительном пласте культурной и ландшафтной лексики, заимствованной из тундрового ненецкого в северные диалекты мансийского и хантыйского языков (заимствований в хантыйский при этом значительно больше). Основной корпус этих заимствований был выделен еще В. Штейницем (Steinitz 1959). В статье А. Ю. Урманчиевой (Урманчиева 2021) показано, что эти заимствования сохраняют целый ряд архаичных черт фонетики тундрового ненецкого. Историческая фонетика обско-угорских языков указывает на то, что эти заимствования произошли не позднее XV века.

Приведем заимствования из ненецкого в казымский хантыйский: т. нен. *xabt*<sup>o</sup> ‘кастрированный олень’ > каз. хант. *χəpti*; т. нен. *xabt*<sup>o</sup>*rka* ‘бесплодная важенка, которую дрессируют для езды’ > каз. хант. *χəptərka*; т. нен. *maly*<sup>o</sup>*cyu* ‘малица’ > каз. хант. *mɔl’ši* [иногда: *mɔləp’si*]; т. нен. *nyada* ‘ягель’ > каз. хант. *ńɔta*; т. нен. *pan*<sup>o</sup> ‘нижний край (например, нюка); подол одежды’ > каз. хант. *pəŋ*; т. нен. *parka* ‘парка’ > каз. хант. *pəɾχa*; т. нен. *tati* ‘младшая жена (при многоженстве)’ > каз. хант. *tətti*; т. нен. *syaqmey*<sup>o</sup> ‘менструация’ > каз. хант. *śɔt’ŋa*; т. нен. *ŋaw*<sup>o</sup>*ka* ‘авка’ > каз. хант. *awkka*; т. нен. *mador* ‘группа, толпа, стадо, стая’ > каз. хант. *matər*; т. нен. *syap*<sup>o</sup> ‘две плоские костяшки, скрепленные ремнем на узде оленьей упряжки’ > каз. хант. *śaŋəw*, *-ŋw*; т. нен. *syar* ‘кожа (под волосами)’ > каз. хант. *śar*, обд. хант. *śār* ‘верхний слой шкуры’; т.нен. *nyamsyer* ‘невозмутимый, хладнокровный (о человеке)’ > каз. хант. *ńɔtmʒa*; т. нен. *nyəb*<sup>o</sup>*lyuy*<sup>o</sup> ‘неблуж (осенний теленок); телючья шкура, снятая примерно в сентябре, пригодная для малицы’ > каз. хант. *ńəpɾəw*, *-ŋw*; т. нен. *yədalabcy*<sup>o</sup>*h* ‘скребок для выделки шкур’ > каз. хант. *jətləp’si*; т. нен. *nyəw*<sup>o</sup>*di*<sup>o</sup> ‘теленок годовалой важенки; внебрачный ребенок’ > каз. хант. *ńəwti*; т. нен. *pətna* ‘лоншак, олень-самец по второму году’ > каз. хант. *ńətna*; т. нен. *məlk*<sup>o</sup> ‘комолый’ > каз. хант. *məlχa*; т. нен. *nyulək*<sup>o</sup> ‘мягкий’ > каз. хант. *ńūlak* пух; т. нен. *χuŋkər*<sup>o</sup> ‘выемка между ягодицами (у человека)’ > каз. хант. *šāš-χəŋχart* ‘на крупе идущее от спины вниз углубление, выемка в крестце

(и спине) лошади'; т. нен. *tíncyah* ( $h < \eta$ ) 'аркан для поимки оленей' > каз. хант. *tĩńsaŋ*; т. нен. *túcyah* ( $h < \eta$ ) 'женская сумочка для швейных принадлежностей (из оленьей шкуры или цветного сукна)' > каз. хант. *tűtsaŋ*; т. нен. *syobyua* 'капюшон (малицы или совика)' > каз. хант. *śqta*; т. нен. *тууу<sup>о</sup>ко* 'подночье (нижняя зимняя покрывка чума мехом внутрь)' > каз. хант. *mǝjǝkkǝ*.

Таким образом, представленность линии N1a2b1b1 у казымских хантов можно объяснить вторичными контактами с тундровыми ненцами. Тогда эта линия остается представленной у тундровых ненцев и хакасов-качинцев, причем парадоксальным образом тундровые ненцы на рассматриваемом филогенетическом дереве оказываются теснее связаны с хакасами-качинцами, чем даже с энцами. Но, возможно, лингвистические данные помогут объяснить этот парадокс. Традиционно ближайшими лингвистическими родственниками тундрового ненецкого считались северносамодийские языки: ближайший к тундровому лесной ненецкий, достаточно близкие к обоим ненецким лесной и тундровый энецкий и занимающий достаточно обособленную позицию в северносамодийской подгруппе нганасанский. Но, как показано в работе Helinski 1997, к северносамодийским, несмотря на свое крайне южное положение, относился также маторско-тайгийско-карагасский язык. Е. А. Хелимский проверил всю известную маторскую лексику самодийского происхождения методом сепаратных изоглосс, т. е. он выделял такие лингвистические параллели, которые связывают маторский либо с южносамодийскими (селькупским и/или камасинским), либо с северносамодийскими (ненецким и/или энецким и/или нганасанским). Маторско-южносамодийских лексических изоглосс, по его подсчетам — 38, маторско-северносамодийских — 97 (Helinski 1997). На основании этого Е. А. Хелимский приходит к выводу о том, что маторский относится к северносамодийским языкам, к ненецко-энецкой подгруппе.

В работе Урманчиева 2023 проведено дополнительное исследование маторско-северносамодийских связей. Был несколько изменен метод: не только определялись лексические параллели, связывающие маторский с южно- либо с северносамодийскими, но также оценивалось, должны ли выделенные параллели рассматриваться как сохранение общего архаизма (что может быть случайным явлением) либо как общая инновация (поскольку инновации являются ключевыми для определения подгруппы в лингвистической классификации). В таблице 2 ниже учитываются только такие параллели, которые могут претендовать на статус общих инноваций (из числа которых удалены однозначно вторичные параллели контактного происхождения). Параллели, которые сложно охарактеризовать как архаизмы/инновации и общие архаизмы из подсчетов исключаются.

Таблица 2. Лексические инновации, связывающие маторский с южно- и северносамодийскими языками

	общие инновации
– с камасинским и селькупским	4
– с селькупским	4
– с камасинским	3
– всего с южносамодийскими	11
– с ненецким, энецким и нганасанским	14
– с ненецким и энецким	12
– с энецким	0
– с <b>ненецким</b>	8
– всего с северносамодийскими	34

Как видно из этой таблицы, число параллелей, которые могут трактоваться как общие генетические инновации, у маторского существенно (в три раза) выше с северносамодийскими, чем с южносамодийскими. Е.А. Хелимский отмечает бóльшую близость маторского к энецко-ненецкой подгруппе (Helimski 2022). Но интересно, что внутри этой подгруппы маторский демонстрирует отчетливую близость с ненецким (строка таблицы, выделенная полужирным): иными словами, ненецкий и маторский разошлись позже, чем ненецкий и энецкий. Если теперь мы предположим, что среди предков современных хакасов-качинцев были носители маторского, настолько же в генетическом отношении близкие к ненцам, насколько маторский близок к ненецкому в лингвистическом отношении, это позволит нам объяснить, почему на полученном методами генетики филогенетическом дереве хакасы-качинцы близки к тундровым ненцам. Естественно, это достаточно прямолинейная и упрощенная картина: говоря о «носителях маторского», мы должны иметь в виду скорее какой-то из диалектов маторско-тайгийско-карагасского континуума — возможно, даже тот, о котором не сохранилось вообще никаких сведений. Но принципиально картину можно очертить именно таким образом, как это сделано выше.

Медианные сети YSTR-гаплотипов мужчин, относящихся к этой кладе, демонстрируют значительные этнические различия и специфичность спектра гаплотипов для отдельных этносов и субэтнических групп. Практически все основные этно-специфичные сублинии показывают звездообразную филогению этих гаплотипов, которые демонстрируют эффект основателя и во многих случаях достаточно активный демографический рост предковых групп. Наиболее сильный эффект основателя выявлен у ненцев, что связано с переходом их на тундровое оленеводство и значительным увеличением их численности из поколения в поколение. У хакасов присутствуют две разные сублинии, которые связаны с базовым самодийским населением, разделённым территориально и вошедшим в состав сагайцев и качинцев. Южные сагайцы из Аскизского и Таштыпского районов демонстрирует генетическую близость к шорцам и челканцам. Качинцы и чулымцы также достаточно близки друг другу, но более удалены от хакасов-сагайцев.

Что касается качинцев, следует отметить, что генетический материал забирался только у северных качинцев, носителей так называемого белоюсского говора качинского языка. Именно эта группа демонстрирует в языковом отношении (как, впрочем, и по фамилиям) сильное влияние кызыльского адстрата (а, может быть, и субстрата): наличие заимствований из кызыльского диалекта с переходом *\*ni-* > *i-*, презенса с *odur-* и под. Так что, возможно, эти генетические данные касаются кызыльцев и чулымцев, а не качинцев и чулымцев. Южные сагайцы (прежде всего бельтыры), напротив, и в языковом, и в этнологическом отношении связаны с шорцами, поскольку с XVII в. зафиксированы повторяющиеся миграции шорцев на территорию нынешней южной Хакасии. Что касается языковых контактов с различными самодийскими группами, то лингвистам известны следующие факты (Дыбо 2007).

1. Установлено как минимум 19 достаточно надежных заимствований из праторского в прасамодийский, т. е. отражающихся в северосамодийских языках: ПСам. *\*jetmä-* ‘латать’ < ОТ *\*jama-* ‘латать, штопать’, ПСам. *\*jesəj-* ‘ставить чум’ < ОТ *\*jasa-* ‘строить, устраивать’, ПСам. *\*keru* ‘оса’ < ОТ *\*Кариη* (?) ‘шмель’, ПСам. *\*päjmä* ‘сапог’ < ПТ *\*balmak* ‘сапог, обувь’, ПСам. *\*kän* ‘правитель’ < ПТ *\*kän* ‘правитель’, ПСам. *\*ta(ə)j* ‘ветвь, сук’ < ОТ *\*dal* ‘ветка, ива’, ПСам. *\*pät-* ‘погружаться, заходить (о солнце)’ < ПТ *\*bat-* ‘погружаться, тонуть, заходить (о солнце)’, ПСам. *\*jekä* ‘близнец’ < ПТ *\*(h)ejkiř* ‘близнецы’, ПСам. *\*wekänä/\*wekzrз* ‘осетр’ < ПТ *\*bEkre* ‘вид



- осетровой рыбы', ПСам. \**jokā*- (? \**jok*-) 'sich verirren' < ПТ \**jōk* 'нет', \**jōk-ad*- (отыменной глагол), \**jōk-a-l*- (рефлексив отыменного глагола) 'потеряться, исчезнуть', ПСам. \**tet<sup>2</sup>tā* (? \**tettā*) 'четыре' < ПТ \**dōrt* 'четыре', ПСам. \**juntz* 'лошадь' < ОТ \**junt* 'лошадь, кобыла', ПСам. \**jür* 'сто' < ПТ \**jūr* 'сто', ПСам. \**jür* 'жир' < ПТ \**ūr* 'жир', ПСам. \**ken* 'ножны' < ПТ \**kījn* 'ножны', ПСам. \**kīr* 'седой, светлый, белый' < ОТ \**Kīr* 'серый, седой', ПСам. \**ki*, \**kilz* 'соболь' < ПТ \**kīl* 'соболь', ПСам. \**kāpsā* 'шаманская ложка для окроплений' < ПТ \**Kamič* 'черпак'. Заимствования из пра-самодийского в пратюркский более сомнительны и малочисленны, их насчитывается всего четыре: ОТ \**K(i)aja* 'скала, гора' < ПСам. \**koājā* 'гора, водораздел', ОТ \**kās* 'кора' < ПСам. \**kāsa* 'кора'; ОТ \**tīt* 'лиственница/кедр' < ПСам. \**titeŋ* 'кедр'; ОТ \**būn* (\**būjn*?) < ПСам. \**wēŋz* 'мясной или рыбный суп, бульон'. Это должны быть очень ранние контакты, до распада северных и южных самодийцев, датируемого приблизительно рубежом эры.
2. Относительно ранние, т. е. не из современных тюркских языков Южной Сибири, заимствования в матор-тайги-карагасский (Хелимский 2000); обсуждение см. в Дыбо 2007, с возможной передатировкой в "современные" — 3 vs 2 слова из гипотетического пра-южносамодийского во все тюркские языки Южной Сибири: тув. *kada* 'тетя', 'бабушка, свекровь', *kadaŋ* 'жена, старуха, баба', тоф. *qadaŋ*, качин. *qadaŋ* < ЮСам. < ПСам. \**kātā* 'бабушка'; тув. *dazıl*; шор. *tazıl*; галт. *t.azıl*; туба *tazıl*; куу *tazıl* 'корень' < ЮСам. \*\**tasV* < ПСам. \**tensā* ~ \**censā* (> ПССам. \**tensā*); МТК \**üüd8* 'корова, крупный рогатый скот' < тув. *ud* 'вол, бык'; МТК \**kaj* 'гусь' < ПТ \**Kāŋ* 'гусь'; Камас.-койб. *boi* 'лед' < ОТ \**būz* 'лед'.
  3. Относительно современные заимствования в матор-тайги-карагасский и в камасинский (там, где это можно определить — в матор-тайги-карагасский из языков саянской группы, в основном тофаларского, и языков z-группы, в основном хакасских диалектов, в том числе сагайского; в камасинский — из языков z-группы, хакасских и шорских диалектов, в том числе качинского и кызыльского) — всего больше 500 (Joki 1952; Helimski 1997). Относительно современные заимствования из МТК в тувинский и тофаларский — 3: тув. *ire* 'дед', *irek* 'дед, старик', тоф. *ire* 'дедушка, отец мужа, прадед, предок по муж. линии, медведь' < матор. *irei* 'медведь' < ПСам. \**irāj* 'дед, старик'; тоф. *tarhiš* 'годовалый олененок (неодобрительное название)' < ПСам. \**cərkaj*- 'олененок'; тоф. *čoydī* этноним < матор. *čagadi* 'старый'. Из камасинского в z-группу — 4: хакас. *sidiĭk*, саг. *sedik* 'тяжелый, трудный' < камас. *sed-em* тж. < ПСам. \**seč*- тж. (Helimski 1995); хак. кач. (XIX в.) *мугнэ* 'росомаха' < койб. камас. *мугна* < ПСам. \**wiŋkəncä*; хак. койб. (уст.) *ка* 'лось' < койб. камас. *ка* < ПСам. \**kā*; хак. кач. (XIX в.) *тямь* 'выдра' < койб. камас. *тямь* < ПСам. \**töt*. Ср. еще кальку: кызыл. *харга айы* март (букв. вороны месяц) — камас. *bārik'i* 'первый месяц весны' (*bāri* ворона).

Данных по языковым контактам с самодийцами, выделяющих разные зоны контакта для кызыльцев (или белоюсских качинцев; + чулымцы) и шорцев, пока не обнаруживается. Исторически (по картам) видно, что маторы могли контактировать тесно с койбалями, которые, по материалам Кастрена, несомненно, хакасы и по некоторым параметрам ближе к качинцам. Северные качинцы и кызыльцы в засвидетельствованное время контактируют с камасинцами. Структурные изогlossы с маторским также не показывают разделения «север — юг» для диалектов z-группы; например, специфически маторская депалатализация начального \**ŋ*- происходит в кызыльском, качинском, сагайском и шорском, не затрагивает чулымский; специфически маторский процесс *č*->*s*-/\_*Vs* нахо-

дит свою параллель в кызыльском, но там это совсем новый процесс, идущий на протяжении XX в., и с несколько иными условиями распределения. С другой стороны, например, данные по истории распространения показателей творительного падежа в хакасских диалектах (Дыбо & Мальцева 2021) позволяют предположить миграцию койбалов-качинцев в зону контакта с кызыльцами в течение первой половины XX века. Учитывая отсутствие генетических данных по усть-абаканским, староиюсским качинцам и кызыльцам, следует признать, что как лингвистические, так и генетические данные нуждаются в дальнейшем исследовании.

Генотипирование новых высокоинформативных SNP-маркеров показывает, что в пределах гаплогруппы N1a2b1 Y-хромосомы наблюдается разделение не только на этноспецифические линии, но и на более узкие сублинии и кластеры гаплотипов, вплоть до индивидуального уровня. Большинство этнических генофондов характеризуются преобладанием или даже полным доминированием специфичных SNP по этой гаплогруппе при использовании нескольких сотен YSNP. Это позволит подробно и очень точно описать территориальную и межэтническую генетическую дифференциацию по специфичности различий YSNP и YSTR-гаплотипам.

Присутствие у всех исследованных популяций разных этноспецифичных вариантов гаплогруппы N1a2b подтверждает отсутствие миграций носителей этой гаплогруппы и обмена генами на протяжении последних нескольких сотен лет. Основным фактором разделения этой гаплогруппы на отдельные сублинии явилась генетическая изоляция разных предковых самодийских групп коренного населения и их дальнейший демографический рост.

Значительная частота этой гаплогруппы в Южной Сибири, в частности, у хакасов, тувинцев и северных алтайцев, вероятнее всего отражает вклад в их генофонд самодийских этносов, заселявших эту территорию ранее. Очевидно, в Алтае-Саянском регионе они не были вытеснены или истреблены более поздними мигрантами, а ассимилированы и включены в состав современных тюркоязычных групп. При расселении на север предков нгансан, энцев и ненцев происходило их смешение с древним аборигенным населением и увеличение численности людей в популяциях при смене поколений.

С точки зрения лингвистики наблюдается изменение состава лексики в записях словарных материалов, сделанных на Саянах с XVIII по середину XIX в. В начале этого периода мы видим самодийские словари с вкраплением тюркских заимствований, в конце — тюркские словари с единичными самодийскими заимствованиями (Дыбо & Мальцева 2021).

Местом начала распространения на территории Сибири гаплогруппы N1a2b является регион Алтая и Саян. Территория, соответствующая современному ареалу тувинцев, алтайцев и хакасов, была первичным очагом генерации разнообразия и экспансии численности предковой группы носителей гаплогруппы N1a2b. По мере расселения носителей гаплогруппы сформировались вторичные очаги генерации разнообразия и экспансии численности — на территории Западной Сибири и ЯНАО. По лингвистическим данным прародина самодийцев находилась примерно в этом же регионе. Часть самодийцев мигрировала отсюда на север около рубежа нашей эры. В позднейшее время оставшееся в Саяно-Алтайском регионе самодийское население было ассимилировано сибирскими тюркскими народами.

Расширенный состав терминальных SNP для выявленных гаплогрупп позволил детально охарактеризовать и проанализировать различия в гаплотипической структуре отдельных этноспецифичных сублиний, определить их родство, следы экспансии численности в генофонде исследованных популяций.

## Заключение

В рамках исследования на основе данных о полных геномах и структуре сублиний N1a2b Y-хромосомы выполнено детальное описание самодийского компонента в генофонде современных популяций Южной и Западной Сибири, оценка его вклада и реконструкция формирования современного генетического разнообразия исследуемых популяций, что позволяет точнее охарактеризовать историю генофондов современных сибирских народов. Получены новые результаты сопоставления данных этногенетики и лингвистики.

Гаплогруппа N1a2b маркирует древнюю популяционную общность населения уралоидного антропологического типа, заселявшего территории к востоку от Уральских гор. Распространение этой гаплогруппы на территории Западной и Южной Сибири связано с расселением самодийцев. Наибольшая частота в популяциях северных самодийцев — нганасан, ненцев и энцев служит этому прямым подтверждением. Полученные результаты позволяют по-новому взглянуть на историю формирования ненцев, энцев, хантов, хакасов, шорцев, чулымских тюрков, тувинцев, северных алтайцев, телеутов, сибирских татар и других народов. Детально охарактеризована филогения гаплогруппы N1a2b Y-хромосомы, являющейся основой генофонда многих этносов Сибири.

Филогенетический анализ Y-хромосомных сублиний и гаплотипов доказывает, что центром происхождения и расселения носителей самодийского компонента в Южной и Западной Сибири является территория современного Алтая и Саян. Полученные результаты хорошо согласуются с данными лингвистики, этнологии и антропологии о вкладе самодийского компонента в формирование различных народов Алтая-Саян и историческими ареалами самодийских и других языков уральской языковой семьи.

Обнаружение большого количества SNP-маркеров, которые определяют филогенетически недавние ветви для различных гаплогрупп Y-хромосомы и данные об их распространении в современных популяциях, позволили провести отбор специфичных маркеров для популяций. Использование таких высокоинформативных этноспецифичных SNP-маркеров Y-хромосомы является одним из наиболее перспективных инструментов для анализа региональных и этнических популяционных генофондов на уровне субэтносов, родов и локальных территориальных групп.

Все исследованные выборки характеризуются преобладанием и даже полным доминированием специфичных SNP по гаплогруппе N1a2b. Сопоставление проводится в первую очередь по терминальным SNP для различных сублиний и консервативным, более медленно мутирующим STR-маркерам, которые обладают наибольшей этнодифференцирующей информативностью.

Корректность полученного методами генетики филогенетического дерева, отражающего взаимную близость популяций, была верифицирована путем сопоставления с выводами об этнической истории данных этносов, полученными лингвистическими методами. Это сопоставление позволило подтвердить в нескольких важных аспектах правильность соотнесения гаплогруппы N1a2b с самодийским компонентом. Как минимум некоторые из разделений данной гаплогруппы на сублинии коррелируют с современными представлениями о сложных процессах смены языка, адстратных и субстратных влияний, оставляющих свой след в языке: на основе таких специфических лингвистических маркеров реконструируются различные контактные ситуации, в которые в разные периоды своей истории были вовлечены носители самодийских языков. При этом необходимо отметить, что данное междисциплинарное исследование показало *взаимную* верификацию выводов: для лингвистов столь же важно, что реконструируемые ими про-

цессы находят подтверждение в результатах, полученных методами других наук. Перечислим вкратце наиболее существенные корреляции между данными генетики и лингвистики:

- регион, в котором в Сибири началось распространение гаплогруппы N1a2b, в целом близок к той зоне, которую лингвисты очерчивают как прародину самодийцев, то есть как ту территорию, на которой жили носители реконструируемого самодийского праязыка и на которой начались процессы его диалектной дивергенции;
- наиболее загадочно с лингвистической точки зрения положение маторского языка в группе самодийских. Этот язык, угасший в XIX веке, географически был самым южным из самодийских языков (северные предгорья Саян), однако сравнительно-исторические исследования маторской лексики, проводившиеся в последние десятилетия, показывают, что он принадлежит к группе северносамодийских языков. Е. А. Хелимский (Helimski 1997) делает вывод о том, что маторский относится к ненецко-энецкой подгруппе; в диссертации А. Ю. Урманчиевой (Урманчиева 2023) делается попытка обосновать особую близость внутри этой подгруппы маторского к ненецкому. Однако лингвистический материал по маторскому достаточно скуден: мы располагаем только списками слов (в донаучной записи, которая требует тщательной и не всегда возможной лингвистической интерпретации), материалов по грамматике этого языка до нас практически не дошло, так как его носители были ассимилированы тюркоязычными народами данного региона. В этой ситуации (слишком смелые выводы, сделанные на ограниченном материале) очень важно получить независимое подтверждение лингвистической гипотезы. Прямое подтверждение мы находим в истории сублинии N1a2b1b1: она представлена у тундровых ненцев и у хакасов качинцев (на язык которых как раз и могли перейти маторы), но не у энцев;
- наличие сублинии N1a2b1b1 у казымских хантов объясняется их вторичными контактами с ненцами, что отражается в многочисленных культурных заимствованиях из ненецкого в северные диалекты хантыйского;
- на примере детального разбора фрагмента филогенетического дерева — поиска лингвистических корреляций истории сублинии N1a2b1b1 — мы показали перспективность таких междисциплинарных исследований. Валидность этой методики подтверждена также тем, что сублиния N1a2b1b2a1, обнаруженная у тувинцев и монголов, находит соответствие в истории тувинско-монгольских языковых контактов. Несомненно, столь же многообещающие результаты ожидаются и в отношении истории других этносов, говорящих или говоривших на самодийских языках. Однако это требует существенно более полного охвата популяций Сибири генетическими обследованиями. Построенное на современном этапе генетическими методами филогенетическое дерево заведомо неполно: на нем отсутствует информация по нганасанам, лесным ненцам, селькупам, информация по носителям енисейских языков, поэтому крайне желательным было бы расширение исследовательской базы.

Полученные результаты не противоречат общепринятым версиям этногенеза самодийских народов, но позволяют более детально описать этот процесс. Информация о структуре генофонда популяций Южной и Западной Сибири по Y-хромосомным сублиниям является важным дополнением к существующим лингвистическим, а также антропологическим и археологическим данным о их происхождении, формировании, миграциях, демографических процессах и родственных связях с другими народами.

Онлайн-приложение к статье (доступно по адресу [jолr.ru/jlr22/kharkov.zip](http://jолr.ru/jlr22/kharkov.zip)) содержит следующие иллюстративные материалы:

Рисунок 1. Филогенетическое древо гаплогруппы N1a2 (StarBeast3).

Рисунок 2. Филогенетическое древо гаплогруппы N1a2 (Y-LineageTracker).

Рисунок 3. Разделение сублиний гаплогруппы N1a2b азиатской ветви по этноспецифичным терминальным SNP.

Рисунок 4. Медианная сеть YSTR-гаплотипов гаплогруппы N1a2b1 у чулымцев.

Рисунок 5. Медианная сеть YSTR-гаплотипов гаплогруппы N1a2b1b1 у хантов.

### Список сокращений

арин. — аринский; галт. — горноалтайский (алтай-кижи); каз. — казымский (диалект хантыйского); камас. — камасинский; кач. — качинский (диалект хакасского); койб. — койбальский; кот. — коттский; матор. — маторский; МК — язык(и) словаря Махмуда Кашгарского; МТК — маторско-гайгийско-карагасский; огуз. — огузский; обд. — обдорский (диалект хантыйского); ОТ — общетюркский; ПЕн. — праенисейский; ПСам. — прасамодийский; ПТ — пратюркский; саг. — сагайский (диалект хакасского); т.нен. — тундровый ненецкий; тоф. — тофаларский; тув. — тувинский; хак. — хакасский; хант. — хантыйский; шор. — шорский.

### Литература

- Бурлак, С. А., С. А. Старостин. 2005. *Сравнительно-историческое языкознание*. Москва: Издательский центр «Академия».
- Волков, В. Г. 2013. Древние миграции самодийцев и енисейцев в свете генетических данных. *Томский журнал лингвистических и антропологических исследований* 1(1): 79–96.
- Дульзон, А. П. 1950. Древние смены народов на территории Томской области по данным топонимики. *Ученые записки Томского государственного педагогического института* VI: 175–187.
- Дульзон, А. П. 1953. Поздние археологические памятники и проблема происхождения чулымских татар. *Ученые записки Томского государственного педагогического института* X: 127–133.
- Дульзон, А. П. 1954. Термины родства и свойства в языках Нарымского края и Причулымья. *Ученые записки Томского государственного педагогического института* XI: 59–94.
- Дульзон, А. П. 1960. Этнический состав древнего населения Западной Сибири по данным топонимики. В кн.: *Доклады делегации СССР: XXV Международный конгресс востоковедов*: 1–10. Москва: Изд-во Вост. лит.
- Дульзон, А. П. 1973. Диалекты и говоры тюрков Чулыма. *Советская тюркология* 2: 16–29.
- Дыбо, А. В. 2007. *Лингвистические контакты ранних тюрков. Лексический фонд: пратюркский период*. Москва: Восточная литература.
- Дыбо, А. В., В. С. Мальцева. 2021. Хакасские диалекты: конец XIX в. и начало XXI в. *Oriental Studies* 14(6): 1353–1367.
- Рассадин, В. И. 1980. Монголизмы в чулымско-тюркском языке В кн.: И. Д. Буряев (ред.). *Монголо-бурятские заимствования в сибирских тюркских языках*: 48–51. Москва: Наука.
- Рогинский, Я. Я., М. Г. Левин. 1978. *Антропология: Учебник для студентов университетов. 3-е изд.* Москва: Высшая школа.
- Старостин, С. А. 2007. *Труды по языкознанию*. Москва: Языки славянских культур.
- Старостин, С. А. 1982. Праенисейская реконструкция и внешние связи енисейских языков. В кн.: Е. А. Алексеев (ред.). *Кетский сборник. Антропология, этнография, мифология, лингвистика*: 144–237. Ленинград: Наука.
- Старостин, С. А. 1995. Сравнительный словарь енисейских языков. В кн.: С. А. Старостин (ред.). *Кетский сборник. Лингвистика*: 176–315. Москва: Языки русской культуры.
- Урманчиева, А. Ю. 2021. Фонетические особенности ненецких заимствований в обско-угорских языках. *Урал-алтайские исследования* 40(1): 101–123.
- Урманчиева, А. Ю. 2023. *Реконструкция лингвистического ландшафта Западной Сибири (на материале самодийских языков): диссертация на соискание ученой степени доктора филологических наук*. Санкт-Петербург.
- Хелимский, Е. А. 2000. *Компаративистика, уралоистика. Лекции и статьи*. Москва: Языки русской культуры.

## References

- Balanovsky, O., K. Dibirova, A. Dybo, O. Mudrak, S. Frolova, E. Pocheshkhova, M. Haber, D. Platt, T. Schurr, W. Haak et al. 2011. Parallel Evolution of Genes and Languages in the Caucasus Region. *Molecular Biology and Evolution* 28(10): 2905–2920.
- Bandelt, H.-J., P. Forster, A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16(1): 37–48.
- Barbujani, G. et al. 2003. A recent shift from polygyny to monogamy in humans is suggested by the analysis of worldwide Y-chromosome diversity. *Journal of Molecular Evolution* 57: 85–97.
- Burlak, S. A., S. A. Starostin. 2005. *Sravnitel'no-istoricheskoe jazykoznanije*. Moskva: Izdatel'skij centr «Akademija».
- Campbell, Lyle. 2013. *Historical linguistics: an introduction*. 3rd edition. Edinburgh University Press.
- Dul'zon, A. P. 1950. Drevnije smeny narodov na territorii Tomskoj oblasti po dannym toponimiki. *Uchenyje zapiski Tomskogo gosudarstvennogo pedagogicheskogo instituta* VI: 175–187.
- Dul'zon, A. P. 1953. Pozdnije arxeologicheskije pam'atniki i problema proisxozhdenija chulymskix tatar. *Uchenyje zapiski Tomskogo gosudarstvennogo pedagogicheskogo instituta* X: 127–133.
- Dul'zon, A. P. 1954. Terminy rodstva i svojstva v jazykax Narymskogo kraja i Prichulym'ja. *Uchenyje zapiski Tomskogo gosudarstvennogo pedagogicheskogo instituta* XI: 59–94.
- Dul'zon, A. P. 1960. Etnicheskij sostav drevnego naselenija Zapadnoj Sibiri po dannym toponimiki. In: *Doklady delegacii SSSR: XXV Mezhdunarodnyj kongress vostokovedov*: 1–10. Moskva: Izdatel'stvo vostochnoj literatury.
- Dul'zon, A. P. 1973. Dialekty i govory t'urkov Chulyma. *Sovetskaja t'urkologija* 2: 16–29.
- Dupanloup, I., L. Pereira, G. Bertorelle, F. Calafell, M. J. Prata, A. Amorim, G. A. Barbujani. 2003. A Recent Shift from Polygyny to Monogamy in Humans Is Suggested by the Analysis of Worldwide Y-Chromosome Diversity. *Journal of Molecular Evolution* 57(1): 85–97.
- Dybo, Anna V. 2007. *Lingvisticheskie kontakty rannix t'urkov*. *Leksicheskij fond: pratt'urkskij period*. Moskva: Vostochnaja literatura RAN.
- Dybo, Anna V., Vera S. Mal'ceva 2021. Hakasskije dialekty: konec XIX v. i nachalo XXI v.. *Oriental Studies* 14(6): 1353–1367.
- Egorov, I. M., A. V. Dybo, A. S. Kassian. 2022. Phylogeny of the Turkic Languages Inferred from Basic Vocabulary: Limitations of the Lexicostatistical Methods in an Intensive Contact Situation. *Journal of Language Evolution* 7(1): 16–39.
- Epps, Patience. 2015. Historical linguistics and socio-cultural reconstruction. In: Claire Bowern, Bethwyn Evans (eds.). *The Routledge handbook of historical linguistics*: 579–597. London / New York: Routledge.
- Esteller-Cucala, P., M. Palmada-Flores, L. F. K. Kuderna, C. Fontserre, A. Serres-Armero, M. Dabad, M. Torralvo, A. Faella, L. Ferrández-Peral, L. Llovera et al. 2023. Y chromosome sequence and epigenomic reconstruction across human populations. *Communications Biology* 6(1): 623.
- Helimski, Eugen. 1995. Laut- und Wortgeschichte der Türksprachen. *Turkologica* 26: 75–95.
- Helimski, Eugen. 1997. *Die matorische Sprache*. *Studia uralo-altaica* 41. University of Szeged.
- Helimski, Eugen. 2022. Zur Stellung des Matorischen innerhalb der samojedischen Sprachen. In: Valentin Gusev, Anna Urmanchieva, Aleksandr Anikin (eds.). *Siberica et Uralica: In memoriam Eugen Helimski (Studia uralo-altaica 56)*: 479–495. University of Szeged.
- Helimskij, E. A. 2000. *Komparativistika, uralistika. Lekcii i stat'ji*. Moskva: Jazyki russkoj kul'tury.
- Illumäe, A. M., M. Reidla, M. Chukhryaeva, M. Järve, H. Post, M. Karmin, L. Saag, A. Agdzhoyan, A. Kushniarevich, S. Litvinov et al. 2016. Human Y Chromosome Haplogroup N: A Non-trivial Time-Resolved Phylogeography that Cuts across Language Families. *The American Journal of Human Genetics* 99(1): 163–173.
- ISOGG = *International Society of Genetic Genealogy*. On-line resource: [www.isogg.org](http://www.isogg.org). Accessed on 07.12.2023.
- Jobling, M. A., C. Tyler-Smith. 2003. The human Y chromosome: an evolutionary marker comes of age. *Nature Reviews Genetics* 4: 598–612.
- Joki, Aulis J. 1952. *Die Lehnwörter des Sajansamojedischen*. *Mémoires de la Société Finno-Ougrienne* 103. Helsinki: Suomalais-Ugrilainen Seura.
- Karafet, T. M., L. P. Osipova, M. A. Gubina, O. L. Posukh, S. L. Zegura, M. F. Hammer. 2002. High levels of Y-chromosome differentiation among native Siberian populations and the genetic signature of a boreal hunter-gatherer way of life. *Human Biology* 74(6): 761–789.
- Karafet, T. M., L. P. Osipova, O. V. Savina, B. Hallmark, M. F. Hammer. 2018. Siberian genetic diversity reveals complex origins of the Samoyedic-speaking populations. *American Journal of Human Biology* 30(6): e23194.

- Karmin, M., L. Saag, M. Vicente, M. A. Wilson Sayres, M. Järve, U. G. Talas, S. Rootsi, A. M. Ilumäe, R. Mägi, M. Mitt et al. 2015. A recent bottleneck of Y chromosome diversity coincides with a global change in culture *Genome Research* 25(4): 459–466.
- Khabtagaeva, Baiarma. 2006. The Intensity of Mongolian Influence in the Tuva Language. *Acta Orientalia* 59(1): 85–94.
- Kharkov, V. N., L. M. Novikova, I. Y. Khitrinskaya, V. G. Volkov, V. A. Stepanov, O. V. Shtygasheva, F. A. Luzina. 2020. Gene Pool of Khakass and Shors for Y Chromosome Markers: Common Components and Tribal Genetic Structure. *Russian Journal of Genetics* 56(7): 849–855.
- Kharkov, V. N., L. V. Valikhova, E. L. Yakovleva, V. N. Serebrova, N. A. Kolesnikov, T. I. Petelina, I. Yu. Khitrinskaya, V. A. Stepanov. 2021. Reconstruction of the Origin of the Gydan Nenets Based on Genetic Analysis of Their Tribal Structure Using a New Set of YSTR Markers. *Russian Journal of Genetics* 57(12): 1414–1423.
- Lappalainen, T., S. Koivumäki, E. Salmela, K. Huoponen, P. Sistonen, M. L. Savontaus, P. Lahermo. 2006. Regional differences among the Finns: a Y-chromosomal perspective. *Gene* 376(2): 207–215.
- Mallory, James, Anna Dybo, Oleg Balanovsky. 2019. The Impact of Genetics Research on Archaeology and Linguistics in Eurasia. *Russian Journal of Genetics* 55(12): 1472–1487.
- Rassadin, Valentin I. 1980. Mongolizmy v chulymsko-t'urkskom jazyke. In: I. D. Bur'ajev (ed.). *Mongolo-bur'atskije zaimstvovaniya v sibirskix t'urkskix jazykax*: 48–51. Moskva: Nauka.
- Roginskij, Ja. Ja., M. G. Levin. 1978. *Antropologija: Uchebnik dl'a studentov universitetov. 3-e izd.* Moskva: Vysshaja shkola.
- Rootsi, S., L. A. Zhivotovsky, M. Baldovic, M. Kayser, I. A. Kutuev, R. Khusainova, M. A. Bermisheva, M. Gubina, S. A. Fedorova, A. M. Ilumäe et al. 2007. A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. *European Journal of Human Genetics* 15(2): 204–211.
- Starostin, S. A. 1982. Praenisejskaja rekonstrukcija i vneshnije sv'azi jenisejskix jazykov. In: E. A. Alekseenko (ed.). *Ketskij sbornik. Antropologija, etnografija, mifologija, lingvistika*: 144–237. Leningrad: Nauka.
- Starostin, S. A. 1995. Sravnitel'nyj slovar' jenisejskix jazykov. In: S. A. Starostin (ed.). *Ketskij sbornik. Lingvistika*: 176–315. Moskva: Jazyki russoj kul'tury.
- Starostin, S. A. 2007. *Trudy po jazykoznaniju*. Moskva: Jazyki slav'anskix kul'tur.
- Steinitz, Wolfgang. 1959. Zu den samojedischen Lehnwörtern im Ob-Ugrischen. *Ural-Altische Jahrbücher XXXI*: 426–453.
- Tambets, K., S. Rootsi, T. Kivisild, H. Help, P. Serk, E. L. Loogväli, H. V. Tolk, M. Reidla, E. Metspalu, L. Pliss et al. 2004. 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* 74(4): 661–682.
- ToB = Starostin, S. A. et al. 1998–2005. *The Tower of Babel. An etymological database project*. Online resource available at: [starlingdb.org](http://starlingdb.org).
- Underhill, P. A., P. Shen, A. A. Lin, L. Jin, G. Passarino, W. H. Yang, E. Kauffman, B. Bonnè-Tamir, J. Bertranpetit, P. Francalacci, M. Ibrahim, T. Jenkins, J. R. Kidd, S. Q. Mehdi, M. T. Seielstad, R. S. Wells, A. Piazza, R. W. Davis, M. W. Feldman, L. L. Cavalli-Sforza, P. J. Oefner. 2000. Y chromosome sequence variation and the history of human populations. *Nature Genetics* 26(3): 358–361.
- Urmanchieva, Anna Ju. 2021. Foneticheskie osobennosti nenetskix zaimstvovanij v obsko-ugorskix jazykax. *Uralo-altajskie issledovaniya* 40(1): 101–123.
- Urmanchieva, Anna Ju. 2023. *Rekonstrukcija lingvisticheskogo landshafta Zapadnoj Sibiri (na materiale samodijskix jazykov)*. Doctoral thesis. Saint-Petersburg.
- Valikhova, L. V., V. N. Kharkov, A. A. Zarubin, N. A. Kolesnikov, M. G. Svarovskaya, I. Yu. Khitrinskaya, O. V. Shtygasheva, V. G. Volkov, V. A. Stepanov. 2022. Genetic Interrelation of the Chulym Turks with Khakass and Kets according to Autosomal SNP Data and Y-Chromosome Haplogroups. *Russian Journal of Genetics* 58(10): 1228–1234.
- Volkov, V. G. 2013. Drevnie migracii samodijcev i enisejcev v svete geneticheskix dannyx. *Tomskij zhurnal lingvisticheskix i antropologicheskix issledovanij* 1(1): 79–96.
- Watanabe, Y., I. Naka, S. S. Khor, H. Sawai, Y. Hitomi, K. Tokunaga, J. Ohashi. 2019. Analysis of whole Y-chromosome sequences reveals the Japanese population history in the Jomon period. *Scientific Reports* 9(1): 8556.
- Zerjal, T., B. Dashnyam, A. Pandya, M. Kayser, L. Roewer, F. R. Santos, W. Schiefenhövel, N. Fretwell, M. A. Jobling, S. Harihara et al. 1997. Genetic relationships of Asians and Northern Europeans, revealed by Y-chromosomal DNA analysis. *The American Journal of Human Genetics* 60(5): 1174–1183.
- Zhivotovsky, L. A., P. A. Underhill, C. Cinnioglu, M. Kayser, B. Morar, T. Kivisild, R. Scozzari, F. Cruciani, G. Destro-Bisol, G. Spedini et al. 2004. The Effective Mutation Rate at Y Chromosome Short Tandem Repeats, with Application to Human Population-Divergence Time. *The American Journal of Human Genetics* 74(1): 50–61.

*Vladimir N. Kharkov, Anna Y. Urmanchieva, Nikita A. Kolesnikov, Aleksey A. Zarubin, Larisa V. Valikhova, Irina Yu. Khitrinskaya, Ilya A. Gruntov, Ilya M. Egorov, Valeriya M. Lemskaya, Anna V. Dybo, Vadim A. Stepanov.* Population genetics and comparative linguistics: phylogeny and phylogeography of haplogroup N1a2b and the linguistic prehistory of the Samoyed peoples

In this paper, we carry out a comparison of ethnogenetic data on the Y chromosome haplogroup N1a2b and linguistics of the Samoyed language group. Haplogroup N1a2b was inherited by modern peoples from an earlier population of the Uraloid anthropological type that inhabited the territory of Southern Siberia. A detailed reconstruction of the phylogeny of this haplogroup shows the ethnospecific structure of the N1a2b sublineages. The linguistic prehistory of the Samoyed peoples is reconstructed based on the analysis and stratification of loanwords. The results on linguistic phylogeny that are relevant for the present paper are obtained using manual analysis of common innovations (for the Samoyedic languages) and Bayesian lexicostatistics (for the Turkic languages). The following correspondences between the results of the genetic and linguistic analyses were uncovered. The region of the initial spread of haplogroup N1a2b roughly corresponds to the area identified as the Samoyedic homeland based on linguistic data. Specific proximity of Mator to the Nenets languages finds support in the history of N1a2b1b1 lineage. This lineage is attested among the Tundra Nenets and Khakas Kachins, whose language could be acquired by the former Mator-speaking population. The presence of N1a2b1b1 lineage among Kazym Khanty is explained by their secondary contacts with Nenets, entailed by the Nenets loanwords in the cultural vocabulary of the Northern Khanty dialects. N1a2b1b2a1 sublineage has been found among Mongols and Tuvans, correlating with the history of the Mongol-Tuvanian language contacts. The undertaken detailed analyses of the phylogeny of N1a2b1b1 lineage confirms the credibility of these conclusions. The obtained modern joint results from population genetics and linguistics complement each other fairly well and provide reliable confirmation of linguistic kinship and genetic-demographic processes in the populations of Siberia.

*Keywords:* ethnogenetics; Y-chromosome; linguistic prehistory; Samoyed people.