

Агджоян А.Т.^{1,2)}, Дамба Л.Д.⁴⁾, Запорожченко В.В.^{2,1)}, Балановский О.П.^{1,2,3)}

¹⁾ ФГБУН Институт общей генетики им. Н.И. Вавилова РАН,
ул. Губкина, д. 3, ГСП-1, Москва, 119991, Россия;

²⁾ ФГБНУ «Медико-генетический научный центр им. академика Н.П. Бочкова»,
ул. Москворечье, д. 1, Москва, 115522, Россия;

³⁾ АНО «Биобанк Северной Евразии», ул. Котляковская, д. 3, Москва, 115201, Россия;

⁴⁾ ГБУ «НИИ медико-социальных проблем и управления Республики Тыва»,
ул. Кечил-оола, д. 2А, Кызыл, 667003, Россия

К ВОПРОСУ О САМОДИЙСКОМ СУБСТРАТЕ В ГЕНОФОНДЕ КОРЕННОГО НАСЕЛЕНИЯ ЮЖНОЙ СИБИРИ: ФИЛОГЕОГРАФИЯ Y-ХРОМОСОМНОЙ ГАПЛОГРУППЫ N-L666

Введение. Целью данной работы являлось изучение разнообразия гаплогруппы **N2** – предполагаемого генетического наследия самодийских племен – у территориальных и родовых групп тувинцев и тофалар, анализ структуры кластеров и датировок их формирования в контексте популяционной истории.

Материалы и методы. Изучено разнообразие Y-хромосомной гаплогруппы **N-L666** (филогеографически почти эквивалентной **N2-P43**) у популяций Южной Сибири. Выборка включила 590 представителей западных, центральных, южных, юго-восточных и северо-восточных (тоджинцы) тувинцев, относящих себя к одной из 21 родоплеменных групп, а также тофалар. Из этой выборки 138 человек относились к гаплогруппе **N-L666** и были изучены по 15 STR-маркерам в двух масштабах: локальном (территориальные и родоплеменные группы тувинцев) и региональном (сравнение с популяциями Южной и Западной Сибири).

Результаты. В генофонде тувинцев и тофалар выявлены два кластера гаплотипов гаплогруппы **N-L666**: специфичный для них кластер А (охватывает 19% и 16% генофондов тувинцев и тофалар, соответственно) и кластер В с широким ареалом от Западной Сибири до Забайкалья (составляет около 30% генофонда тофалар). Популяционное разнообразие и датировки кластера В указывают на его происхождение из прапопуляции – источника гаплогруппы **N-L666** в Сибири, связываемой с самодийским населением. Популяционный состав и датировки кластера А указывают на его формирование в ареале тувинцев и тофалар в течение последней тысячи лет. Появление субкластеров А1, А2, В1 может быть результатом демографического роста в популяциях тувинцев, южных алтайцев и хакасов около 300–450 лет назад.

Распространение одних и тех же гаплотипов, кластеров и субкластеров среди разных территориальных групп и родов тувинцев указывает на единый для них источник гаплогруппы **N-L666**, существовавший в генофонде задолго до обособления этих групп.

Заключение. Специфический кластер гаплогруппы **N-L666** у тувинцев предположительно был основан представителем одного из самодийских племен, многочисленные потомки которого участвовали в формировании тувинцев, тофалар и южных алтайцев в течение последней тысячи лет.

Ключевые слова: коренное население; Южная Сибирь; самодийцы; тувинцы; тоджинцы; тофалары; генофонд; Y-хромосома; гаплогруппа **N-L666**; гаплотип; филогенетический анализ

Введение

Тувинцы и тофалары – тюркоязычные этносы Южной Сибири, в этногенезе которых принимали участие также монголоязычные, самодийские, кетские, тунгусские группы. Ко времени миграции тюркоязычных племен (в конце I тысячелетия н.э.) на территории Тувы проживало кетоязычное и самодийское население, которое прослеживается здесь и несколько веков спустя (в конце XIII в. – начале XIV в.) [Тишков с соавт., 2008]. Наличие самодийских элементов в этническом составе тувинцев и тофалар отмечено исследователями со второй половины XVIII в.: П.С. Палласом, Ю. Клапротом, И. Георги, и М.А. Кастреном, подтверждающим свою точку зрения лингвистическими данными. А.П. Окладников отметил появление предков самодийских племен в таежных районах Среднего Енисея еще в позднепалеолитическое время и в эпоху бронзы [Маннай-Оол, 2004; Тишков с соавт., 2008].

Территория расселения самодийских племен охватывала северные склоны Саянского нагорья: ареал тувинцев, хакасов, тофалар и др. Не позднее I тысячелетия н.э. часть самодийцев достигла приполярных районов Сибири, куда они принесли свой язык и культуру; их современными потомками являются ненцы, энцы, нганасаны и селькупы [Хелимский, 2000]. Оставшееся в Южной Сибири население было ассимилировано тюркоязычными племенами: так, у тувинцев самодийский след сохранился в малочисленных топонимах и единичных названиях родоплеменных групп (*маады*, *иргит*, *чооду*, *соян*, *хаазыт* и др.). Тофалар исследователи ещё в XVIII в. считали «самодийским племенем с чистым самодийским языком, не столь испорченным, как у других саянских самодийцев» (что впоследствии объяснялось вхождением в состав тофалар самодийских родов, передавших свой язык) [Тишков с соавт., 2008].

В генетическом ландшафте Сибири с древними самодийскими популяциями связывают распространение гаплогруппы **N2 (N1b-P43)** Y-хромосомы [Харьков с соавт., 2013]: эта линия встречается с высокой частотой как у приполярных самодийцев, так и в Южной Сибири [Ilumäe et al., 2016; Karafet et al., 2018]. Основная доля разнообразия гаплогруппы **N2** в Сибири приходится на её

ветвь **N2a1-B478** (возраст около 3 тыс. лет, рассчитан на основе данных полного секвенирования Y-хромосомы) [Ilumäe et al., 2016]. По данным других исследователей в генофонде тувинцев гаплогруппа **N2** встречается в среднем от 19% [Derenko et al., 2007] до 24% [Харьков с соавт., 2013]. К сожалению, в этих работах «тонкая структура» гаплогруппы **N2** (разнообразие гаплотипов, их филогенетические связи) для субэтнических групп не рассматривалась, хотя при анализе генофонда Сибири в целом публиковались результаты филогенетического анализа с включением тувинцев [Харьков с соавт., 2013; Derenko et al., 2007; Ilumäe et al., 2016]. Небольшая опубликованная выборка тофалар (N=32), включившая только население наиболее доступного с. Алыгджер и изученная по узкой панели SNP маркеров Y-хромосомы [Derenko et al., 2006], выявила линию **N2** с частотой около 43%.

Поясним филогенетические соотношения упоминаемых ветвей гаплогруппы **N**. Ветвь **N-L666** включает в себя ветвь **N2-P43** и очень редкую ветвь **N1-M128** [Ilumäe et al., 2016]. Поэтому далее паттерн гаплогруппы **N-L666** в Сибири приравнивается к таковому для ранее широко изучавшейся **N2 (N1b-P43)**.

В исследованиях нашего научного коллектива [Дамба с соавт., 2018; 2019а, 2019б; Балановская с соавт., 2019; Жабагин с соавт., 2020] у тувинцев средняя частота гаплогруппы **N-L666** составила около 20%, при этом максимальные значения обнаружены у западных тувинцев (34%), минимальные – у юго-восточных (9%), а в остальных территориальных группах тувинцев частота **N-L666** варьировала от 19% до 28%. Среди наиболее многочисленных родов тувинцев [Жабагин с соавт., 2020] гаплогруппа **N-L666** распространена равномерно и встречается со средней частотой 20% (от 14% до 32%). В нашем исследовании [Балановская с соавт., 2019] выборка (N=45) равномерно охватывает все три изолированные поселения тофалар (Алыгджер, Верхняя Гутара, Нерха), и частота **N-L666** составляет 45%.

Частоты гаплогруппы не отражают её внутреннее разнообразие: за равномерным распространением может скрываться как наличие своего кластера гаплотипов у каждого рода (если

он происходит от одного биологического предка), так и принадлежность представителей разных родовых групп к одному кластеру гаплотипов (разные социальные группы происходят от одной древней генетической линии), или в целом – отсутствие кластеров (каждый гаплотип – отдельная редкая линия).

Основной целью данной работы являлось изучение разнообразия гаплогруппы **N2** – предполагаемого генетического наследия самодийских племен – у территориальных и родовых групп тувинцев и тофалар, анализ структуры кластеров и датировок их формирования в контексте популяционной истории.

Материалы и методы

В данном исследовании популяции тувинцев ($N_{\text{TOT}}=545$), где встречается гаплогруппа Y-хромосомы **N-L666** ($N_{L666}=117$), рассмотрены в двух ракурсах – географических групп и родовых групп. Географические группы включили западных ($N_{\text{TOT}}=89$, $N_{L666}=29$), центральных ($N_{\text{TOT}}=184$, $N_{L666}=41$), южных ($N_{\text{TOT}}=92$, $N_{L666}=19$), юго-восточных ($N_{\text{TOT}}=90$, $N_{L666}=8$) и северо-восточных тувинцев (тоджинцы, $N_{\text{TOT}}=90$, $N_{L666}=20$), частоты гаплогруппы **N-L666** для них ранее опубликованы в [Балановская с соавт., 2019]. Рассмотрены и родовые группы тувинцев (21 группа), к которым указана принадлежность предков индивидов по мужской линии (частоты гаплогруппы **N-L666** для 10 наиболее многочисленных родов тувинцев ранее опубликованы в [Жабагин с соавт., 2020]). Популяция тофалар ($N_{\text{TOT}}=45$, $N_{L666}=21$) по данным о частотах гаплогрупп Y-хромосомы ранее описана в [Балановская с соавт., 2019]. Таким образом, в данной работе суммарно проанализировано 138 образцов ДНК тувинцев и тофалар – носителей гаплогруппы **N-L666**, включая ранее не публиковавшиеся данные по остальным родам, а также STR-гаплотипы.

Суммарная выборка образцов собрана Л.Д. Дамба под руководством Е.В. Балановской согласно правилам Биобанка Северной Евразии [Балановская с соавт., 2016] в экспедиционных обследованиях 2015–2017 гг. (см. популяции, описанные ранее [Дамба с соавт., 2018; 2019а, 2019б; Балановская с соавт., 2019; Жабагин с

соавт., 2020]). В выборки включены только неродственные между собой индивиды, все предки которых на протяжении не менее трех поколений относили себя к данному этносу и родились на данной территории. Сбор образцов сопровождался письменным информированным согласием обследуемых под контролем Этической комиссии ФГБНУ «МГНЦ» (г. Москва).

Материалом для исследования послужила геномная ДНК, выделенная из образцов венозной крови методом фенол-хлороформной экстракции. Принадлежность образцов ДНК к гаплогруппам Y-хромосомы определялась по данным анализа широкого спектра SNP-маркеров аналогично описанному в [Балановская с соавт., 2019].

Фрагментный анализ 17 STR локусов проведен на приборе ABI 3130xl (Applied Biosystems) с использованием набора Y-filer PCR Amplification Kit (Applied Biosystems). Для сравнения привлечены гаплотипы северных и южных алтайцев, хакасов и горных шорцев, изученные в [Балаганская, 2011], а также неопубликованные данные научного коллектива (для тофалар, тоболо-иртышских сибирских татар, бурят и хамниган).

Филогенетические сети построены в программе Network 10.2.0.0. [URL: <http://www.fluxus-engineering.com> (дата обращения – 28.12.2020)] по данным о 15 STR-маркерах (DYS389I, DYS389II, DYS390, DYS456, DYS19, DYS458, DYS437, DYS438, DYS448, GATAH4, DYS391, DYS392, DYS393, DYS439, DYS635) и визуализированы в программе Network Publisher v. 2.1.2.5 [URL: <http://www.fluxus-engineering.com> (дата обращения – 28.12.2020)].

Возраст филогенетических сетей и кластеров оценен по числу мутационных шагов (ρ -статистика) [Forster et al., 1996] с использованием длины поколения 31 год [Fenner, 2005] и двух скоростей мутирования: «генеалогической» и «комбинированной». «Генеалогическая» скорость (0,027 мутаций на локус на поколение, аналогично [Балановская с соавт., 2019]) учитывает накопление мутаций в среднем для разных STR-локусов Y-хромосомы, её значение было получено в исследованиях при прямом подсчете мутаций для Y-хромосомы в известных родословных [Балановский, Запорожченко, 2015].

«Комбинированная» скорость (0.042 мутаций на локус за поколение) представляет собой сумму скоростей мутирования для каждого из проанализированных 15 STR-локусов, данные получены с ресурса YHRD.org (Available at: https://yhrd.org/pages/resources/mutation_rates. Accessed: 25.08.2020), и, таким образом, учитывает скорости мутирования только 15 использованных в анализе STR-маркеров. Значение «генеалогической» скорости – в полтора раза меньше, чем «комбинированной», и, соответственно, датировки – в полтора раза выше. Хотя были рассчитаны оба значения датировок, далее в тексте при описании и обсуждении результатов чаще используется «комбинированная» скорость: во-первых, она учитывает динамику именно для анализируемого сочетания 15 маркеров, во-вторых, для расчётов датировок кластеров внутри гаплогрупп возрастом менее 5-10 тыс. лет (для гаплогруппы **N-L666** он составляет в среднем 10 тыс. лет [URL: <https://www.yfull.com/tree/N-L666/> (дата обращения: 16.04.2021)]) рекомендуется использовать близкое или превосходящее «генеалогическую» скорость значение [Балановский, Запороженко, 2015].

Результаты и обсуждение

Разнообразие гаплогруппы N-L666 у разных территориальных групп и родов тувинцев

Результаты филогенетического анализа гаплогруппы **N-L666** у тувинцев представлены на рисунке 1 в трёх вариантах: на рисунке 1.1 отображена схема-сеть, выделены кластеры и субкластеры, указаны гаплотипы вероятных основателей кластеров и субкластеров; на рисунке 1.2 разными цветами обозначены представители территориальных групп тувинцев (западные, центральные, южные, юго-восточные, северо-восточные – тоджинцы), а на рисунке 1.3 – представители разных родов (21 род, из них 10 наиболее многочисленных показаны отдельными цветами, а остальные 11 – серым).

На филогенетической сети (рис. 1.1) выделяются два кластера, обозначенные как А и В. Кластер А – более обширный (90% образцов и 81% гаплотипов всех тувинцев с гаплогруппой **N-L666**, табл. 1) – имеет «звездчатую» структуру с выраженным основателем (α , гаплотип приведен в табл. 2) и двумя субкластерами в его составе (А1

и А2, их основатели – α_1 и α_2 соответственно). Кластер В (гаплотип основателя β , табл. 2) контрастно отличается и по числу образцов (почти в 10 раз меньше, чем в кластере А; табл. 1), и по числу гаплотипов (в 4 раза меньше, чем в кластере А, табл. 1).

Носители более частых гаплотипов (α , α_1 , α_2 ; на рисунках 1.2 и 1.3 – крупные пестрые круги с разноцветными секциями) встречаются в разных территориальных (рис. 1.2) и родовых (рис. 1.3) группах тувинцев. Внутри выделенных кластеров и субкластеров не отмечается накопления гаплотипов из какой-то одной или территориальной, или родовой группы (на рисунках 1.2 и 1.3 не прослеживается преобладания какого-то одного цвета даже внутри субкластеров). Такой паттерн указывает на распространение одних и тех же (или родственных) гаплотипов из единого источника в разные концы ареала тувинцев – от западных до северо-восточных (тоджинцев). Своеобразие генофонда тувинцев-тоджинцев, показанное по Y-хромосомному генофонду в целом [Харьков с соавт., 2013; Балановская с соавт., 2019; Derenko et al., 2006], не прослеживается в пределах гаплогруппы **N-L666**: у них встречаются гаплотипы, идентичные или близкие к таковым у других территориальных групп (рис. 1.2).

Представленность кластера А у разных территориальных популяций и родов тувинцев указывает на его распространение ещё на заре формирования этих групп. Частота кластера А среди всех носителей гаплогруппы **N-L666** у тувинцев составляет 90%, а в генофонде тувинцев в целом – около 19% (число носителей N=105, табл. 1). Допуская распространение гаплогруппы **N-L666** самодийскими племенами, можно предположить аналогичное происхождение и далёкого предка всех представителей кластера А. В таком случае важно представлять, встречается ли кластер А за пределами этнического ареала тувинцев, особенно у соседних популяций Южной Сибири: тофалар, алтайцев, хакасов, для которых по этнографическим данным прослеживается влияние самодийских племен в этногенезе, а в генофонде значительно выражена гаплогруппа **N-L666** [Балаганская, 2011; Балановская с соавт., 2019; Тишков с соавт., 2008]. Для прояснения этого вопроса далее проведен филогенетический анализ с привлечением и этих популяций.

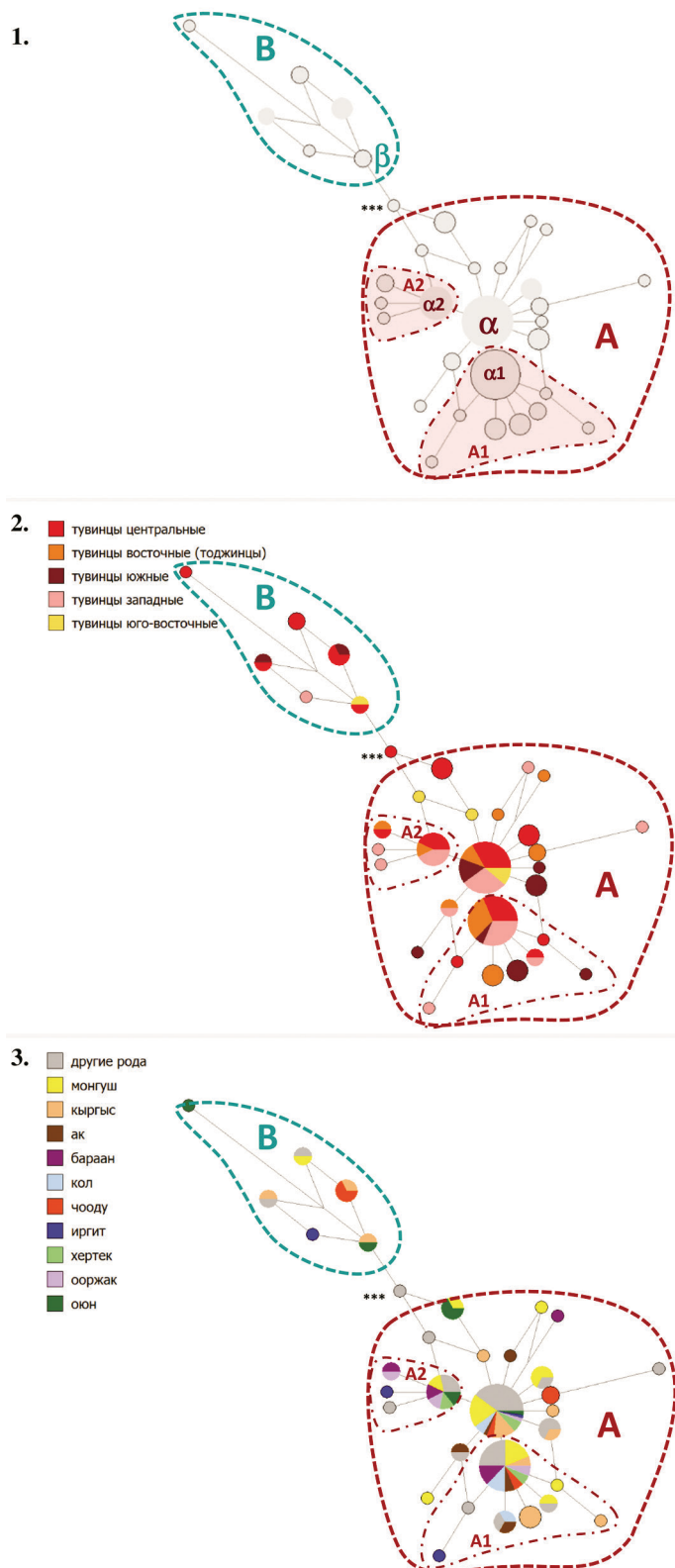


Рисунок 1. Медианная сеть гаплогруппы N-L666 у тувинцев.

(1) – выделены кластеры, субкластеры, гаплотипы основателей; (2) – в легенде обозначена принадлежность носителей гаплотипов к одной из территориальных групп тувинцев; (3) – в легенде обозначена принадлежность носителей гаплотипов к наиболее многочисленным родам тувинцев (всеми цветами, кроме серого) или объединенной выборке из 11 родов (шынаа, ховалыг, саая, долаан, байкара, сат, куулар, шалык, сарыглар, ондар, маады), обозначенной серым

Figure 1. Median network of haplogroup N-L666 in Tuvans. (1) – Clusters, sub-clusters and founder haplotypes are labeled; (2) – In the legend, see the affiliation of the individuals with one of the regional groups of Tuvans; (3) – In the legend, see either the affiliation of the individuals with the most numerous clans of Tuvans (painted by all colours but grey) or the presence in the pooled sample of 11 minor clans (Shyna, Khovalyg, Saaya, Dolaan, Baykara, Sat, Ruular, Shalyk, Saryglar, Ondar, Maady) in grey

Примечания. Каждый круг обозначает отдельный гаплотип (уникальное сочетание 15 аллелей); длина линий, связывающих гаплотипы, пропорциональна числу мутаций (различий между гаплотипами); размер круга соответствует числу носителей гаплотипа (минимальный размер круга показан для единичных гаплотипов).

Notes. Each circle depicts a distinct haplotype (which is a unique combination of 15 alleles); the length of the edge connecting two haplotypes is proportional to the number of changes between them; the size of each circle corresponds to the number of the haplotype-bearing individuals (fixed minimum size is used for the singletons).

Взаимодействие тувинцев с окружающими популяциями

При расширении географии популяций-носителей гаплогруппы **N-L666** на филогенетической сети (рис. 2.1–2.2) оба основных кластера сохраняются, но при этом значительно расширяется только кластер В. В составе кластера В преобладают хакасы (качинцы и сагайцы), что отражается в значении специфичности кластера В (47% – только хакасы, табл. 1), довольно заметны тоболо-иртышские татары (в основном, ясколбинские, или «заболотные» сибирские татары), также встречаются тофалары, северные алтайцы, горные шорцы, реже – гаплотипы южных алтайцев, бурят Забайкалья и хамниган. Тувинцы в кластере В (представители всех групп, кроме тоджинцев) вошли в состав «хакасского» субкластера В1, гаплотипа основателя γ и отходящей от него ветви, содержащей два гаплотипа тофалар (рис. 2.2).

В кластер А включились – кроме тувинцев – гаплотипы только тофалар и южных алтайцев (алтай-кижи, теленгиты); и, таким образом, их распределение на данной филогенетической сети приближено к таковому для тувинцев: тофалары и южные алтайцы встречаются в обоих кластерах. Заметим, что гаплогруппа **N-L666** доминирует в генофонде тофалар (частота достигает 45%), что определяет их генетическое сходство с популяциями не только Южной, но и Западной Сибири, Урала [Балановская с соавт., 2019]. На филогенетической сети (рис. 2.2) большинство образцов тофалар (2/3 от общего числа носителей гаплогруппы **N-L666**) распределились в кластер В, но большинство из них (N=12) – носители двух родственных гаплотипов. Доля носителей только двух этих гаплотипов – более четверти Y-хромосомного генофонда тофалар, что вероятнее всего связано с дрейфом генов в этой изолированной популяции.

Наличие у тувинцев с гаплогруппой **N-L666** двух генетических кластеров, один из которых доминирует в их генофонде (в данном исследовании – кластер А), а второй чаще встречается у соседних популяций Южной и Западной Сибири, показано и ранее: в работе [Харьков, 2012] на филогенетической сети гаплогруппы **N1b (N2-P43)** – основной линии внутри гаплогруппы **N-L666**. Данный результат, полученный независимо разными научными коллективами при изучении разных выборок

тувинцев и популяций сравнения, подчеркивает реальность выявленных закономерностей.

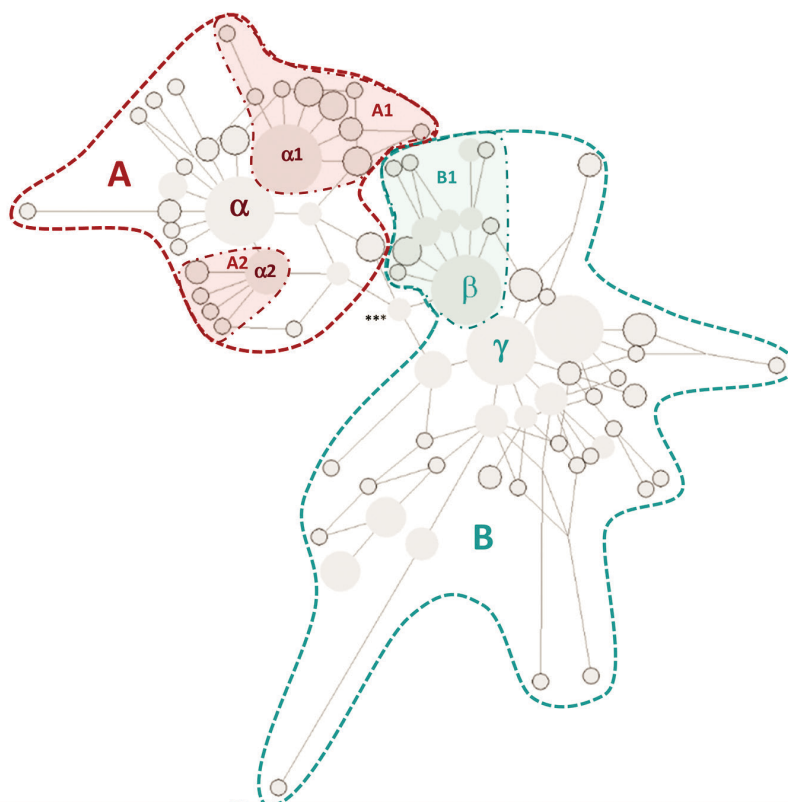
Представленность в кластере А гаплотипов только тувинцев, южных алтайцев и части тофалар указывает на редкость этого варианта у других популяций (хакасы, сибирские татары, северные алтайцы, горные шорцы, буряты, хамнигане). Данный факт, а также высокое популяционное разнообразие кластера В указывает на его происхождение из «предковой» популяции – источника гаплогруппы **N-L666** в Южной Сибири. Предполагается, что носители кластера В расселились достаточно широко, но в ареале тувинцев, тофалар и южных алтайцев одна из линий получила развитие и сформировался специфичный кластер А. Формирование субкластеров А1 и А2 внутри этой линии указывает на демографический рост в ареале указанных популяций. Для датировки данных событий мы рассчитали возраст каждого кластера и субкластера.

Датировки и гипотезы

Датировки формирования кластеров и субкластеров, обозначенных на филогенетических сетях (рис. 1 и 2), представлены в таблице 1. Согласно нашим расчётам, возраст кластера А составляет около 600–700 лет. Возраст кластера В составляет около 1500 лет (рис. 2, табл. 1). При добавлении других популяций в анализ на рисунке 2 изменилась и структура кластера В: стала больше соответствовать «звёздчатой», но в центре кластера при этом выделяется гаплотип γ – другой вероятный основатель (поэтому датировки для кластера В на рисунке 2 приведены при основателях β и γ , табл. 1, 2). Датировка формирования кластера В при основателе γ ниже (в среднем около 1200 лет), чем при основателе β , что связано с положением гаплотипа γ ближе к центру сети (иными словами, для формирования такой модели потребовалось бы меньше времени, чем при основателе β , табл. 1).

Достаточно близкими оказались датировки возникновения субкластеров А1, А2, В1: они составили в среднем 300–450 лет (500–700 лет – при использовании «генеалогической» скорости мутирования; рис. 2, табл. 1). Данный факт может указывать на демографический рост в это время не только в популяциях тувинцев, но также – южных алтайцев и хакасов.

1.



2.

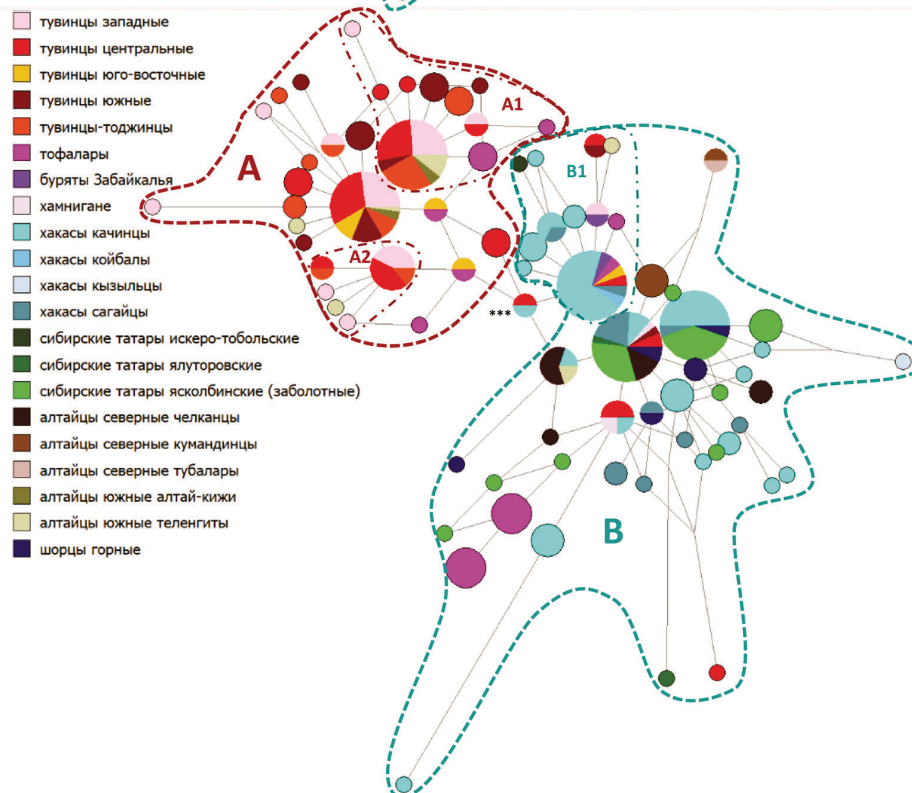


Рисунок 2. Медианная сеть гаплогруппы **N-L666** в популяциях Южной Сибири. (1) – выделены кластеры, субкластеры, гаплотипы основателей; (2) – в легенде обозначена принадлежность носителей гаплотипов к популяциям тувинцев и групп сравнения
 Figure 2. Median network of haplogroup **N-L666** in the populations of Southern Siberia. (1) – Clusters, subclusters and founder haplotypes are labeled; (2) – In the legend, see the affiliation of the individuals with Tuvans and the populations used for comparison

Таблица 1. Возраст кластеров STR-гаплотипов на филогенетических сетях
Table 1. Age of the STR haplotype clusters in the phylogenetic networks

| Рисунок | КЛАСТЕР | СОСТАВ КЛАСТЕРА | Количество образцов в кластере | Количество гаплотипов в кластере | Специфичность кластера (Sp) | Возраст, лет число мутаций ($\rho \pm \sigma$) | Возраст, лет «генеалогическая» скорость) | Возраст, лет «комбинированная» скорость) |
|---------|---------------------------------|-----------------|--------------------------------|----------------------------------|-----------------------------|--|--|--|
| 1 | Вся сеть (основатель ***) | | 117 | 33 | – | 3,33±1,38 | 3800±1600 | 2500±1000 |
| 1 | A | Тувинцы | 105 | 26 | 100% | 0,83±0,31 | 950±350 | 600±230 |
| 1 | A1 | Тувинцы | 28 | 8 | 100% | 0,50±0,20 | 580±230 | 370±150 |
| 1 | A2 | Тувинцы | 11 | 4 | 100% | 0,36±0,22 | 420±260 | 270±170 |
| 1 | B (основатель β) | Тувинцы | 11 | 6 | 100% | 1,45±0,61 | 1700±700 | 1100±450 |
| 2 | Вся сеть (основатель ***) | | 275 | 79 | – | 3,11±0,99 | 3600±1150 | 2300±750 |
| 2 | Вся сеть (основатель β) | | 275 | 79 | – | 3,07±0,91 | 3500±1150 | 2300±700 |
| 2 | Вся сеть (основатель γ) | | 275 | 79 | – | 3,21±1,00 | 3700±1130 | 2400±730 |
| 2 | A | Тувинцы | 120 | 31 | 88% | 0,90±0,33 | 1050±400 | 700±250 |
| 2 | A1 | Тувинцы | 35 | 10 | 80% | 0,54±0,20 | 620±220 | 400±150 |
| 2 | A2 | Тувинцы | 12 | 5 | 100% | 0,42±0,22 | 480±250 | 300±160 |
| 2 | B (основатель β) | Хакасы | 154 | 47 | 48% | 2,07±0,83 | 2400±950 | 1500±600 |
| 2 | B (основатель γ) | Хакасы | 154 | 47 | 48% | 1,54±0,40 | 1800±460 | 1150±300 |
| 2 | B1 (основатель β) | Хакасы | 36 | 11 | 69% | 0,61±0,24 | 700±270 | 450±180 |

Примечания и обозначения: Рис. – номер рисунка, на котором на медианной сети выделен соответствующий кластер. Sp – специфичность кластера, т.е. процент образцов данного кластера, происходящих из популяций, указанных в названии кластера; ρ -статистика (ρ) обозначает среднее число мутаций от основателя до периферии кластера, σ – среднеквадратичное отклонение (погрешность оценки возраста); возраст в годах в таблице указан от современности.

Notes and Legend: Рис. (Figure) – The index of the picture representing a network with the corresponding cluster. Sp – specificity of the cluster, i.e. the percentage of the individuals of the specified populational origin in the cluster; rho-statistics (ρ) is the mean number of mutations from the founder haplotype to the leaf ones, σ – standard deviation for the rho value; the age is specified in years before present.

Что касается времени возникновения самих кластеров A и B (время их разделения от общего предка), то по STR-маркерам нами получена датировка 2,5 тыс. лет назад. По данным полного секвенирования Y-хромосомы [Ilumäe et al., 2016; Valinova et al., 2019] большинство тувинцев (и тувинцы-тоджинцы) с гаплогруппой **N-L666** являются носителями линии **N2a1-B478**, которая сформировалась в близкий период времени (около 3 тыс. лет назад, по полногеномным данным) и встречается в популяциях Южной, Западной и Центральной Сибири (алтайцев, хакасов, ненцев, нганасан, долган).

Отметим близость полученных нами и «полногеномных» датировок в отношении субкластеров A1 и A2: на филогенетическом дереве (в [Ilumäe et al., 2016] внутри линии **N2a1-B478** выделен кластер, включающий только образцы тувинцев, а его датировка составляет в среднем 1,1 тыс. лет.

Заключение

Проведенный филогенетический анализ Y-хромосомной гаплогруппы **N-L666**, распространенной в Западной и Южной Сибири, позволяет сделать ряд выводов. Прежде всего, датировка

Таблица 2. STR гаплотипы-основатели (founder) для кластеров филогенетических сетей гаплогруппы N-L666

Table 2. The founder haplotypes for the phylogenetic networks of haplogroup N-L666

| Рис. | КЛАСТЕР и гаплотип-основатель | DYS389I | DYS389II | DYS390 | DYS456 | DYS19 | DYS458 | DYS437 | DYS438 | DYS448 | GATA H4 | DYS391 | DYS392 | DYS393 | DYS439 | DYS635 |
|------|----------------------------------|---------|----------|--------|--------|-------|--------|--------|--------|--------|---------|--------|--------|--------|--------|--------|
| 1; 2 | Основатель сети (***) | 13 | 16 | 23 | 15 | 14 | 16 | 14 | 10 | 19 | 12 | 11 | 14 | 13 | 10 | 24 |
| 1; 2 | A (основатель α) | 13 | 16 | 23 | 15 | 15 | 16 | 14 | 10 | 19 | 12 | 11 | 14 | 13 | 11 | 25 |
| 1; 2 | A1 (основатель α_1) | 13 | 16 | 23 | 15 | 15 | 16 | 14 | 10 | 19 | 12 | 11 | 14 | 13 | 11 | 26 |
| 1; 2 | A2 (основатель α_2) | 13 | 16 | 23 | 15 | 15 | 16 | 14 | 10 | 19 | 12 | 11 | 14 | 13 | 11 | 24 |
| 1; 2 | B (основатель β) | 13 | 16 | 23 | 15 | 14 | 16 | 14 | 10 | 19 | 12 | 10 | 14 | 13 | 10 | 24 |
| 2 | B (основатель γ) | 13 | 16 | 23 | 15 | 14 | 17 | 14 | 10 | 19 | 12 | 10 | 14 | 13 | 10 | 24 |

Примечания: Рис. – номер рисунка, на котором на медианной сети выделен соответствующий кластер. Основателем филогенетической сети предполагается гаплотип, связывающий кластеры А и Б и обозначенный *** (рис. 1-2). В таблице цветами обозначены значения STR-маркеров: белым – одинаковые с гаплотипом ***, серым – отличающие гаплотипы α , α_1 , α_2 , β и γ от гаплотипа ***, темно-серым – отличие гаплотипа α_1 от α .

Notes: Рис. (Figure) – The index of the picture representing a network with the corresponding cluster. One of the haplotypes connecting clusters A and B is chosen as the founder and named *** (pic. 1-2). STR allelic values are colored in the table: ones equal to the founder *** are in white, those separating the founder *** from α , α_1 , α_2 , β and γ are in light-grey, and ones separating α_1 and α are in dark-grey.

разделения гаплогруппы **N-L666** на два кластера, по результатам анализа как STR-маркеров, так и полного секвенирования, составляет около 3 тысяч лет назад. Иными словами, около 3 тыс. лет назад ещё существовал единый источник гаплогруппы **N2** для всех южносибирских популяций – от сибирских татар до бурят. Затем из него выделился кластер А, характерный для тувинцев и тофалар (встречается также у южных алтайцев). Полученные нами датировки его возраста – около 700 лет с использованием «комбинированной» скорости мутирования, около 1000 лет – «генеалогической» (табл. 1) близки к датировке 1 100 лет, основанной на полном секвенировании [Ilumäe et al., 2016]. Таким образом, на формирование специфической «тувинской» линии гаплогруппы **N-L666** потребовалось около 1 тыс. лет, и очевидно, что её основатель появился в ареале тувинцев не позднее этой даты, возможно, что в составе одного из самодийских племен. Потомки этого основателя оказались достаточно многочисленными, расселились широко в течение последнего тысячелетия и в наши дни обнаружались во всех концах этнического ареала тувинцев, а также среди их восточных и западных соседей – тофалар и южных алтайцев.

Благодарности

Исследование выполнено при финансовой поддержке Государственного задания Министерства науки и высшего образования РФ для Медико-генетического научного центра им. акад. Н.П. Бочкова (работы по генотипированию, филогенетический анализ), Государственного задания Министерства науки и высшего образования РФ для Института общей генетики им. Н.И. Вавилова РАН (интерпретация результатов, написание текста статьи).

Авторы выражают благодарность всем донорам образцов, которые принимали участие в данном исследовании, а также АНО «Биобанк Северной Евразии» за предоставление коллекций ДНК.

Библиография

Балаганская О.А. Полиморфизм Y хромосомы у тюркоязычного населения Алтая, Саян, Тянь-Шаня и Памира в контексте взаимодействия генофондов Западной и Восточной Евразии. Автореф. дисс. ... кандидата биол. наук, 2011, 27 с.

Балановская Е.В., Жабагин М.К., Аджоян А.Т., Чухряева М.И., Маркина Н.В. с соавт. Популяционные биобанки: принципы организации и перспективы

применения в геногеографии и персонализированной медицине // *Генетика*, 2016. Т. 52. № 12. С. 1371–1387. DOI: 10.7868/S001667581612002X.

Балановская Е.В., Дамба Л.Д., Агджоян А.Т., Жабагин М.К., Олькова М.В., с соавт. Генофонд охотников-оленоводов Южной Сибири: тофалары и тоджинцы // *Вестник Московского университета. Серия XXIII. Антропология*, 2019. № 4. С. 67–80. DOI: 10.32521/2074-8132.2019.4.067-080.

Балановский О.П., Запорожченко В.В. Хромосома-летописец: датировки генетики, события истории, соблазн ДНК-генеалогии // *Генетика*, 2016. Т. 52. № 7. С. 708–725. DOI: 0.1134/S1022795416070048.

Дамба Л.Д., Айыжы Е.В., Монгуш Б.Б., Жабагин М.К., Юсупов Ю.М. с соавт. Комплексный подход в изучении родовой структуры тувинцев Республики Тыва на примере родов монгуш и ооржак // *Вестник Тувинского государственного университета*. Вып. 2. Естественные и сельскохозяйственные науки, 2018. № 2 (37). С. 37–44.

Дамба Л.Д., Балановская Е.В., Агджоян А.Т., Короткова Н.А., Олькова М.В. с соавт. Генофонд трех восточных родов тувинцев по данным полиморфизма Y-хромосомы // *Вестник Московского университета. Серия XXIII. Антропология*, 2019а. № 1. С. 74–85. DOI: 10.32521/2074-8132.2019.1.074-085.

Дамба Л.Д., Айыжы Е.В., Монгуш Б.Б., Короткова Н.А., Чернышенко Д.М. с соавт. Краткий обзор некоторых родоплеменных групп тувинцев по данным междисциплинарных исследований // *Вестник Тувинского государственного университета*. Вып. 2. Естественные и сельскохозяйственные науки, 2019б. № 2 (45). С. 19–30. DOI: 10.24411/2077-5326-2019-10003.

Жабагин М.К., Дамба Л.Д., Короткова Н.А., Чернышенко Д.Н., Утриван С.А. с соавт. Анализ родо-

племенной структуры тувинцев по маркерам Y-хромосомы // *Генетика*, 2020. Т. 56. № 6. С. 732–738. DOI: 10.31857/S0016675820060132.

Маннай-Оол М. Х. Тувинцы. Происхождение и формирование этноса. Новосибирск: Наука, 2004. 166 с.

Тишков В.А., Чешко С.В., Миссонова Л.И., Аксянова Г.А., Вайнштейн С.И. и соавт. Тюркские народы Восточной Сибири. М.: Наука, 2008. 422 с.

Хелимский Е.А. Компаративистика, уралистика: Лекции и статьи. М.: Языки русской культуры, 2000. 640 с.

Харьков В.Н. Структура и филогеография генофонда коренного населения Сибири по маркерам Y-хромосомы: Автореф. дисс. ... доктора биол. наук, 2012. 45 с.

Харьков В.Н., Хамина К.В., Медведева О.Ф., Сиимонова К.В., Хитринская И.Ю. с соавт. Структура генофонда тувинцев по маркерам Y-хромосомы // *Генетика*, 2013. Т. 49. № 12. С. 1416–1425. DOI: 10.7868/S0016675813120035.

Сведения об авторах

Агджоян Анастасия Торосовна, к.б.н., ORCID ID:

0000-0002-8776-2934; aagdzhojan@gmail.com;

Дамба Лариса Доржуевна, ORCID ID: 0000-0003-1736-9210; larissa.damba@gmail.com;

Запорожченко Валерий Владимирович, ORCID ID:

0000-0003-4649-1602; valeryz2001@gmail.com;

Балановский Олег Павлович, д.б.н., профессор РАН,

ORCID ID: 0000-0003-4218-6889; balanovsky@inbox.ru.

Поступила в редакцию 06.05.2021,
принята к публикации 17.05.2021.

Agdzhojan A.T.^{1,2)}, Damba L.D.⁴⁾, Zaporozhchenko V.V.^{2,1)}, Balanovsky O.P.^{1,2,3)}

¹⁾ Vavilov Institute of General Genetics, RAS, Gubkina st., 3, GSP-1, Moscow, 119991, Russia;

²⁾ Research Centre for Medical Genetics, Moskvorechie st., 1, Moscow, 115522, Russia;

³⁾ Biobank of Northern Eurasia, Kotlyakovskaya st., 3, Moscow, 115201, Russia;

⁴⁾ Research Institute of Medical and Social Problems and Control of the Healthcare Department, Kechil-ool st., 2A, Kyzyl, 667003, Russia

IN ADDRESSING THE QUESTION ABOUT THE SAMOYEDIC SUBSTRATE IN THE SOUTH SIBERIAN POPULATIONS: THE PHYLOGEOGRAPHY OF Y-CHROMOSOME HAPLOGROUP N-L666

Introduction. The aim of this work was to study the diversity of haplogroup N2 – the alleged genetic heritage of the Samoyedic tribes – in the territorial and clan groups of Tuvans and Tofalars, to analyze the structure of clusters and the dates of their formation in the context of population history.

Materials and methods. We studied the diversity of Y-chromosomal haplogroup N-L666 in Southern Siberia, where this lineage itself is an approximate equivalent to N2-P43. The whole sample included 590 representatives of western, central, southern, southeastern and northeastern (Tojin) Tuvans, who identified themselves as one of 21 tribal groups, as well as Tofalars. The N-L666 subsample consisted of 138 individuals and was studied using 15 STR markers in two scales: local (considering the areal groups and tribal clans of Tuvans) and regional (in comparison with the populations of Southern and Western Siberia).

Results. Two clusters of N-L666 STR haplotypes were identified: cluster A, specific for Tuvans and Tofalars (covering 19% and 16% of their gene pools respectively) and cluster B, widely scattered throughout Siberia from the West to the Transbaikalia (reaching ~30% in Tofalars). The ubiquity and a greater age of the cluster B favor the idea of its origin in the ancestral population – the ultimate source of the haplogroup N-L666 in Siberia – commonly alleged to be Samoyedic by language. On the contrary, the narrow geographic range and a relatively recent age of the cluster A indicate its formation in the area inhabited by Tuvans and Tofalars during the last thousand years. The emergence of subclusters A1, A2, B1 may be the result of demographic growth in the populations of Tuvans, southern Altaians and Khakas about 300-450 years ago.

The spread of the same haplotypes, clusters and subclusters among different regional groups and clans of Tuvans indicates a common source of haplogroup N-L666 for them, which existed in the gene pool long before the separation of the studied populations.

Conclusions. A specific cluster of haplogroup N-L666 in Tuvans was presumably founded by a representative of one of the Samoyedic tribes, whose numerous descendants participated in the formation of the Tuvans, Tofalars and southern Altaians over the last thousand years.

Keywords: indigenous people; Southern Siberia; Samoyedic peoples; Tuvans (Tuvinians); Tojin Tuvans; Tofalars; gene pool; Y-chromosome; haplogroup N-L666; haplotype; phylogenetic analysis

References

- Balaganskaya O.A. *Polimorfizm Y khromosomy u tyurkoyazychnogo naseleniya Altaya, Sayan, Tyan'-Shanya i Pamira v kontekste vzaimodeistviya genofondov Zapadnoi i Vostochnoi Evrazii*. [Y-chromosomal polymorphism in turkic-speaking populations from Altai, Sayan, Tien-Shan and Pamir Mountains in context of the interaction between Western and Eastern Eurasian gene pools] PhD in Biology Thesis. Moscow, 2011. 27 p. (In Russ.).
- Balanovskaya E.V., Zhabagin M.K., Agdzhoyan A.T., Chuhryaeva M.I., Markina N.V. et al. Populyatsionnye biobanki: principy organizatsii i perspektivy primeneniya v genogeografii i personalizirovannoy meditsine [Population biobanks: Organizational models and prospects of application in genegeography and personalized medicine]. *Genetika* [Russian Journal of Genetics], 2016, 52 (12), pp. 1227–1243. (In Russ.). DOI: 10.1134/S1022795416120024.
- Balanovskaya E.V., Damba L.D., Agdzhoyan A.T., Zhabagin M.K., Ol'kova M.V. et al. Genofond okhotnikov-olenevodov Yuzhnoi Sibiri: tofalary i todzhintsy [The Gene Pool of Hunters-Reindeer Herders of Southern Siberia: Tofalars and Todzhintsy]. *Moscow University Anthropology Bulletin* [Vestnik Moskovskogo universiteta. Seriya XXIII. Antropologiya], 2019, 4, pp. 67–80. (In Russ.). DOI: 10.32521/2074-8132.2019.4.067-080.
- Balanovskii O.P., Zaporozhchenko V.V. Khromosomaletopisets: datirovki genetiki, sobytiya istorii, soblazn DNK-genealogii [Chromosome as a chronicle: Genetic dating, historical events, and DNA-genealogic temptation]. *Genetika* [Russian Journal of Genetics], 2016, 52 (7), pp. 708–725. (In Russ.). DOI: 0.1134/S1022795416070048.
- Damba L.D., Aiyzhy E.V., Mongush B.B., Zhabagin M.K., Yusupov YU.M. et al. Kompleksnyi podkhod v izuchenii rodovoi struktury tuvintsev Respubliki Tyva na primere rodov mongush i oorzhak [Complex approach in clan structure of Tuvans by the example of clans Mongush and Oorzhak]. *Vestnik Tuvinskogo gosudarstvennogo universiteta. Vypusk 2. Estestvennye i sel'skokhozyaistvennye nauki*. [Bulletin of Tuva state University. Issue 2. Natural and agricultural Sciences], 2018, 2 (37), pp. 37–44. (In Russ.).
- Damba L.D., Balanovskaya E.V., Agdzhoyan A.T., Korotkova N.A., Ol'kova M.V. et al. Genofond trekh vostochnykh rodov tuvintsev po dannym polimorfizma Y-khromosomy [Gene pool of three eastern Tuvan clans according to Y-chromosome polymorphism] *Moscow University Anthropology Bulletin* [Vestnik Moskovskogo universiteta. Seriya XXIII. Antropologiya], 2019a, 1, pp. 74–85. DOI: 10.32521/2074-8132.2019.1.074-085. (In Russ.).
- Damba L.D., Aiyzhy E.V., Mongush B.B., Korotkova N.A., Chernyshenko D.M. et al. Kratkii obzor nekotorykh rodoplemennykh grupp tuvintsev po dannym mezhdistsiplinarnykh issledovaniy [Summary of some family breeding groups of Tuvans under data of interdisciplinary research]. *Vestnik Tuvinskogo gosudarstvennogo universiteta. Vypusk 2. Estestvennye i sel'skokhozyaistvennye nauki*. [Bulletin of Tuva state University. Issue 2. Natural and agricultural Sciences], 2019b, 2 (45), pp. 19–30. (In Russ.). DOI: 10.24411/2077-5326-2019-10003.
- Zhabagin M.K., Damba L.D., Korotkova N.A., Chernyshenko D.N., Utrivan S.A. et al. Analiz rodoplemennoi struktury tuvintsev po markeram Y-khromosomy [Analysis of Clan Structure of Tuvans by Y-Chromosome Markers]. *Genetika* [Russian Journal of Genetics], 2020, 56 (6), pp. 732–738. (In Russ.). DOI: 10.31857/S0016675820060132.
- Mannai-Ool M. KH. *Tuvintsy. Proiskhozhdenie i formirovanie ehtnosa*. [Tuvans. The origin and formation of the ethnic group] Novosibirsk, Nauka Publ., 2004. 166 p. (In Russ.).
- Tishkov V.A., Cheshko S.V., Missonova L.I., Aksyanova G.A., Vainshtein S.I. i soavt. *Tyurkskie narody Vostochnoi Sibiri* [The Turkic peoples of Eastern Siberia]. Moscow, Nauka Publ., 2008. 422 p. (In Russ.).
- Khelimskii E.A. *Komparativistika, uralistika: Lektsii i stat'i*. [Comparative studies and Uralistics: Lectures and articles.] Moscow, Yazyki russkoi kul'tury Pibl., 2000. 640 p. (In Russ.).
- Kharkov V.N. *Struktura i filogeografiya genofonda korennogo naseleniya Sibiri po markeram Y-hromosomy* [Structure and phylogeography of the gene pool of the indigenous population of Siberia by Y-chromosome

markers] Dr. in Biology Thesis. Tomsk, 2012. 45 p. (In Russ.).

KHar'kov V.N., KHamina K.V., Medvedeva O.F., Simonova K.V., KHitrinskaya I.YU. et al. Struktura genofonda tuvintsev po markeram Y-khromosomy [Gene-Pool Structure of Tuvinians Inferred from Y-Chromosome Marker Data]. *Genetika* [Russian Journal of Genetics], 2013, 49 (12), pp. 1236–1244. DOI: 10.1134/S102279541312003X. (In Russ.).

Balinova N., Post H., Kushniarevich A., Flores R., Karmin M. et al. Y-chromosomal analysis of clan structure of Kalmyks, the only European Mongol people, and their relationship to Oirat-Mongols of Inner Asia. *Eur. J. Hum. Genet.*, 2019, 27, pp. 1466–1474. DOI: 10.1038/s41431-019-0399-0.

Derenko M., Malyarchuk B., Denisova G. et al. Contrasting patterns of Y-chromosome variation in South Siberian population from Baikal and Altai Sayan regions. *Hum. Genet.*, 2006, 118, pp. 591–604. DOI: 10.1007/s00439-005-0076-y.

Derenko M., Malyarchuk B., Denisova G., Wozniak M., Grzybowski T. et al. Y-chromosome haplogroup N dispersals from south Siberia to Europe. *J. Hum. Genet.*, 2007, 52 (9), pp. 763–770. DOI: 10.1007/s10038-007-0179-5.

Illumäe A.M., Reidla M., Chukhryaeva M., Järve M., Post H. et al. Human Y Chromosome Haplogroup N: A Non-trivial Time-Resolved Phylogeography that Cuts across Language Families. *Am. J. Hum. Genet.*, 2016, 99 (1), pp. 163–173. DOI: 10.1016/j.ajhg.2016.05.025.

Fenner J.N. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am. J. Phys. Anthropol.*, 2005, 128, pp. 415–423. DOI: 10.1002/ajpa.20188

Forster P., Harding R., Torroni A., Bandelt H. J. Origin and evolution of Native American mtDNA variation: a reappraisal. *Am. J. Hum. Genet.*, 1996, 59 (4), pp. 935–945.

Karafet T.M., Osipova L.P., Savina O.V., Hallmark B., Hammer M.F. Siberian genetic diversity reveals complex origins of the Samoyedic-speaking populations. *Am. J. Hum. Biol.*, 2018, 30:e23194. DOI: 10.1002/ajhb.23194.

Information about Authors

Agdzhoyan Anastasiya T., PhD of Biology,
ORCID ID: 0000-0002-8776-2934;
aagdzhoyan@gmail.com;

Damba Larissa D., ORCID ID: 0000-0003-1736-9210;
larissa.damba@gmail.com;

Zaporozhchenko Valery V.,
ORCID ID: 0000-0003-4649-1602; valeryz2001@gmail.com;

Balanovsky Oleg P., Doctor of Biology, Professor RAS,
ORCID ID: 0000-0003-4218-6889; balanovsky@inbox.ru.