Расселение 50 казахских родов свидетельствует о генетической истории казахов

Для исследования структуры генофонда казахов использована родовая принадлежность, которая рассматривается как квазигенетический маркер. В работе представлен анализ большого массива данных о 50 казахских родах у 4,2 млн. человек в начале XX века. Эти данные сопоставлены с разнообразием Y-хромосомы в современных казахских популяциях. Частота родов и частота гаплогрупп Y-хромосомы (квазигенетические и генетические маркеры) достоверно коррелируют друг с другом, как показано на примере геногеографических карт для родоплеменной группы Аргын. Анализ межпопуляционной изменчивости по квазигенетическим маркерам при разных группировках популяций привел авторов к выводу, что окончательная структура казахского генофонда, скорее всего, сложилась в эпоху казахского ханства.

Статья, опубликованная в «Вавиловском журнале генетики и селекции», стала результатом комплексного исследования генофонда казахов, проведенного казахскими и российскими генетиками (первый и второй авторы — Максат Жабагин, Национальный центр биотехнологии и Назарбаев университет, Астана, Казахстан; Олег Балановский, Институт общей генетики РАН, Москва). Особенность этой работы в том, что для изучения генофонда были использованы и генетические данные, и квазигенетические — данные о родовой принадлежности.

Род (клан) – это группа людей, возводящих свое происхождение к общему предку. Это понятие, прежде всего, социальное: генеалогическая цепочка рода может соответствовать биологическому родству, а может и не соответствовать. Принадлежность к роду наследуется от отца к сыну, то есть в случае биологического родства передается так же, как и Y-хромосома. Как показали последние исследования этого научного коллектива, у казахов большинство родов действительно состоит в основном из людей, имеющих общего предка по мужской линии. Поэтому род можно рассматривать как квазигенетический маркер для изучения структуры генофонда.

Опираясь на этот факт, авторы работы изучили огромный массив данных о расселении казахских родов в разных регионах Казахстана. Информацию по родам ученые взяли из свода статистических сведений по расселению казахских родов, включающего 4,2 млн человек и основанного на результатах переписи населения Российской империи в конце XIX и начале XX века. Эти уникальные данные авторы представили в виде электронных таблиц и карт расселения родов.

Всего было изучено расселение 50 казахских родов, определены частоты каждого рода в 33 уездах восьми исторических областей Центральной Азии и в шести соседних губерниях Российской империи. Эти данные были сопоставлены с прямыми генетическими данными о распределении гаплогрупп Y-хромосомы в современных казахских популяциях на примере одной из крупнейших родоплеменных групп казахов — Аргын.

Что можно узнать из данных о расселении родов

Поскольку данные о родах рассматривались как квазигенетические маркеры, исследователи, согласно сложившейся традиции, подходили к их анализу так же, как к анализу прямых генетических данных. Каждый род рассматривался как генетический вариант (аллель), наследующийся по однородительскому (отцовскому) типу. По архивным статистическим данным в каждом уезде они рассчитывали частоту встречаемости каждого рода, как обычно популяционные генетики рассчитывают частоту встречаемости гаплогруппы. Квазигенетические маркеры, по аналогии с генетическими, использовались для оценки разнообразия в популяциях (в уездах), индекса миграции, квазигенетических расстояний между популяциями и для оценки случайного инбридинга.

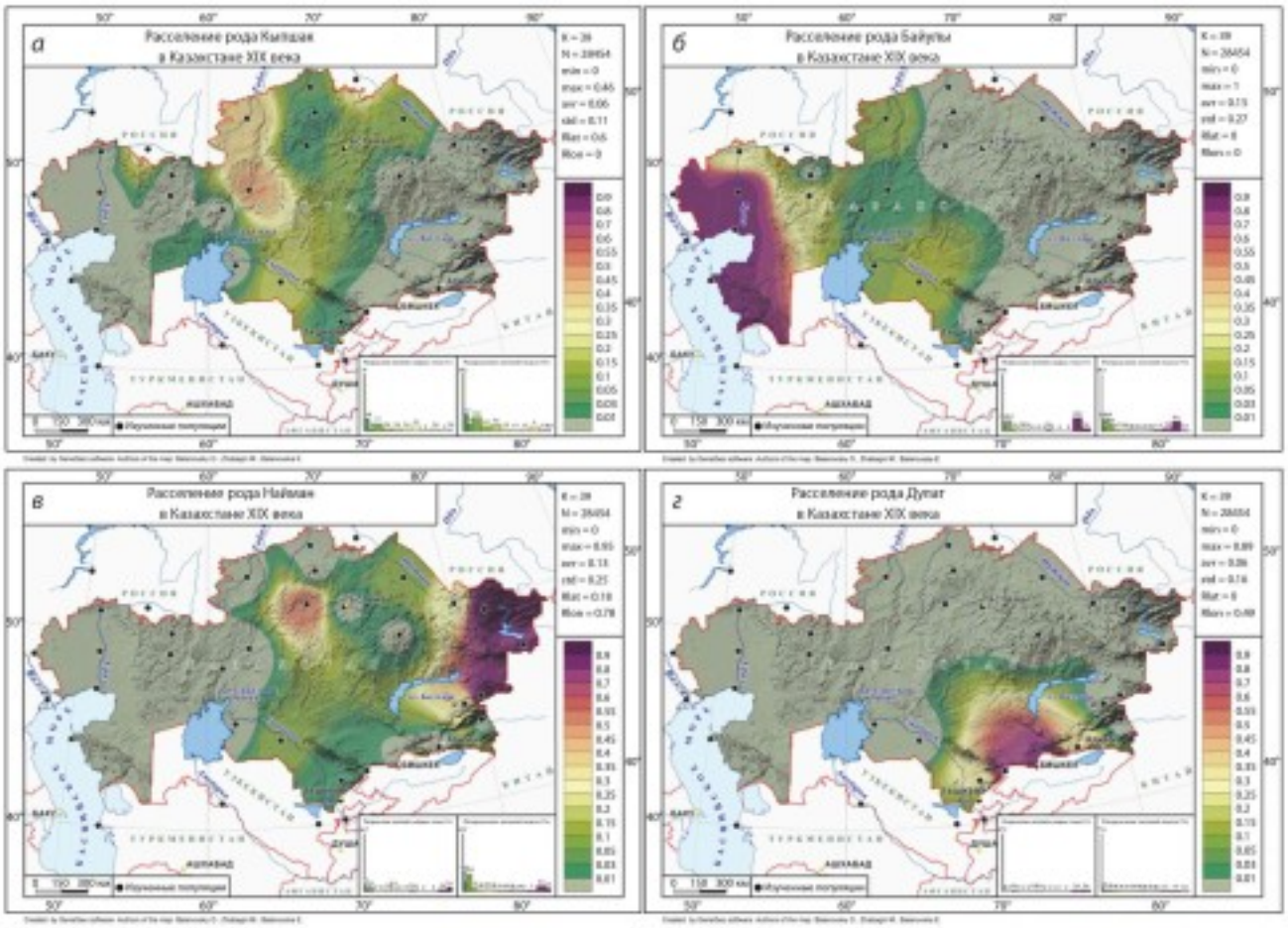
Инбридинг возникает при большой доле близкородственных браков и ведет к увеличению степени гомозиготности (когда на обеих хромосомах индивида расположены одинаковые генетические варианты). Анализируя информацию по родам, авторы для оценки инбридинга использовали показатель изонимии. Чаще всего изонимию определяют по фамилиям – когда оба партнера, вступающие в брак, имеют одинаковые фамилии. В данном случае вычисляли изонимию по вероятности заключения брака между членами одного и того же рода (женщины наследуют род отца).

Самые высокие показатели изонимии и инбридинга авторы отметили на юго-западе (в Закаспийской области, которая отличается и наименьшим числом родов) и на северо-востоке (в Семипалатинской области) региона. Самые низкие показатели изонимии и инбридинга оказались на северо-западе Казахстана (Уральская область) и на смежных территориях России – в Зауралье и нижнем Поволжье. Авторы считают, что эти наблюдения отражают следы исторического заселения

казахами Заволжья в связи с образованием Букеевской Орды в XVIII веке и ростом роли Оренбурга как административного центра Центральной Азии. Они также подчеркивают, что обнаруженный максимальный инбридинг в юго-западных и северо-восточных районах Казахстана позволяет прогнозировать в этих регионах высокий генетический груз (высокую суммарную долю неблагоприятных мутаций).

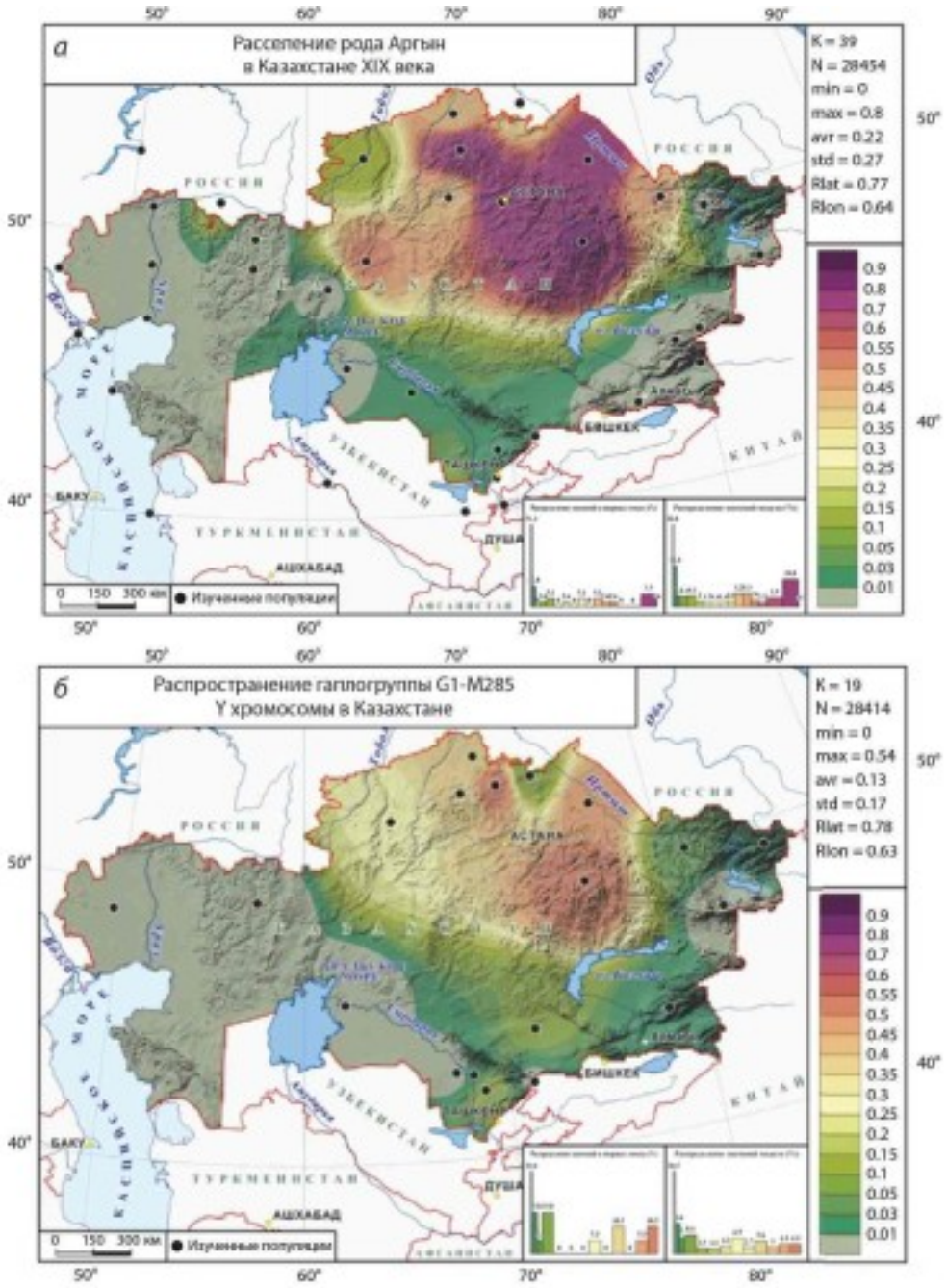
Геногеографические карты по родам и гаплогруппам

По частотам распространения 50 родов авторы построили серию карт. На рисунке представлены четыре таких карты для родов Кыпшак, Байулы, Найман, Дулат. Все карты можно посмотреть в Приложении к статье (ссылка на приложение). На них соответствующим цветом обозначена доля данного рода среди населения региона: цвета фиолетового спектра соответствуют большей доле, цвета зеленого спектра – меньшей доле.



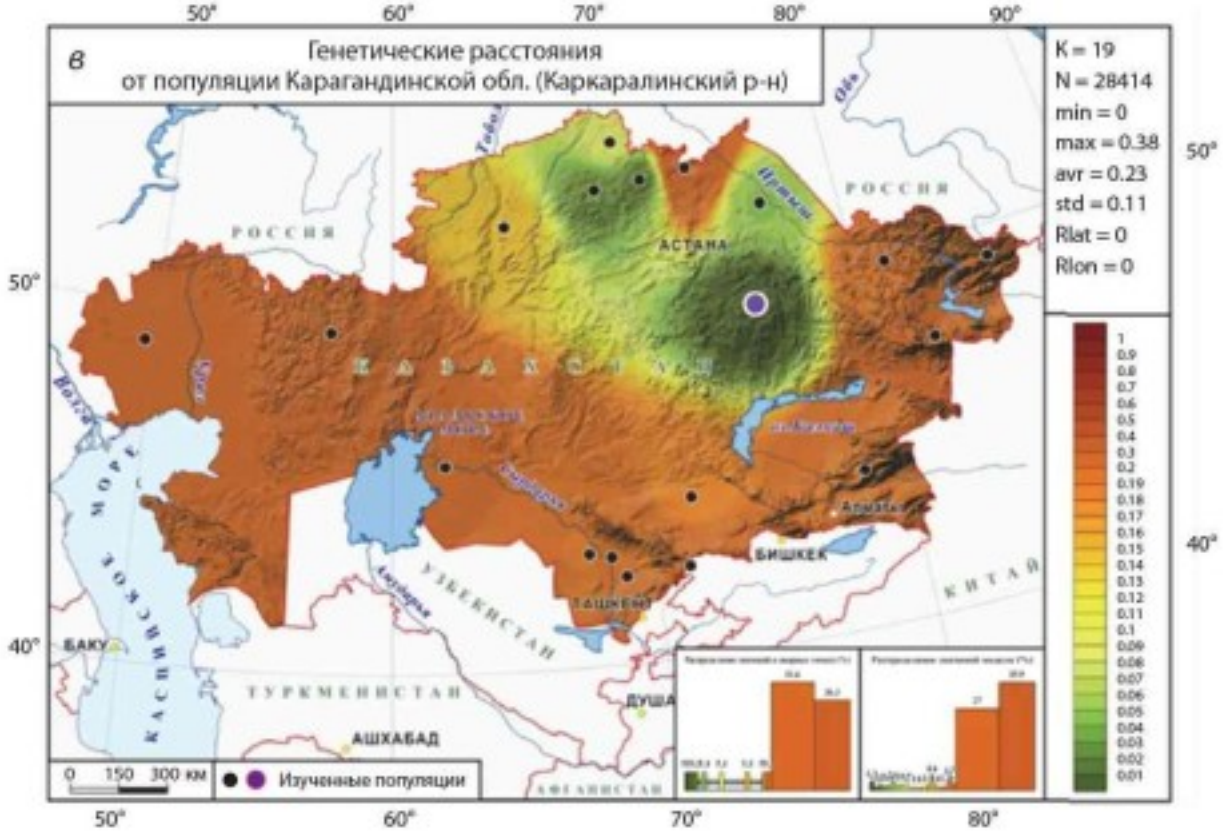
Расселение некоторых казахских родов: а – Кыпшак; б – Байдулы; в – Найман; г – Дулат. Точками обозначены изученные популяции казахов. Цветовая шкала представляет значения распределения признака – доля рода от населения в популяции (от темно-фиолетового – 100% до бледно-зеленого – 1%).

Важно было сравнить карты родов (квазигенетических маркеров) и гаплогрупп (генетических маркеров). На следующем рисунке сопоставлены карты расселения рода Аргын и частоты отдельного генетического маркера (гаплогруппы *G1-M285* Y-хромосомы). Хорошо видно, что эти карты очень близки. Области высокой частоты гаплогруппы G1-M285 совпадают с областями высокой доли рода Аргын в населении.



Геногеография рода Аргын; а – расселение рода (в каждом регионе картографирована частота представителей рода Аргын от общего населения данного региона); б – частота гаплогруппы G1.

Например, в Каркаралинском районе доля рода Аргын особенно высока, и это отражено в особенностях генофонда популяции этого района. На следующем рисунке представлена карта, построенная по генетическим расстояниям от этой популяции до соседних популяций. Зеленым цветом на ней обозначены зоны генетического сходства. Как видно, генетически близкими к популяции Каркаралинского района оказываются те популяции, где 1) высока доля рода Аргын среди населения и 2) высока частота Y-хромосомной гаплогруппы G1-M285. Таким образом, между генетическими и квазигенетическими данными наблюдается четкий параллелизм. Это проявилось и в том, что, когда авторы вычислили квазигенетические расстояния между популяциями, выяснилось, что они с высокой степенью коррелируют с генетическими расстояниями между ними.



Карта, построенная по генетическим расстояниям от популяции Каркалинского района Казахстана. Зеленый цвет обозначает минимальные генетические расстояния, то есть зоны генетического сходства.

Почти полное совпадение квазигенетических и генетических карт говорит о том, что данные о расселении родов могут быть эффективно использованы для прогноза структуры генофонда в тех популяциях, для которых еще не получены прямые генетические данные.

Но вспомним, что квазигенетические и генетические карты различаются во времени: первые получены на материалах начала XX века, а вторые представляют современное население. Из этого генетики заключают, что население большинства регионов, вероятно, сохраняло генетическую преемственность с тех времен до наших дней. Что, по словам авторов, указывает на низкую миграционную активность казахского населения и позволяет задать вопрос – когда же сформировалась современная структура генофонда казахов?

Роль социально-территориальных объединений в структуре генофонда казахов

Чтобы ответить на этот вопрос, авторы исследовали роль различных социально-территориальных объединений казахов в формировании структуры генофонда. Они рассмотрели несколько типов таких объединений.

Первый тип – три государственных образования в эпоху раннего феодализма (750–1050 гг.): Кимакское государство, Карлукское государство, Огузское государство. Предполагается, что они могли стать исходным фактором формирования современной структуры генофонда.

Второй тип объединения – улусы — территориально-административное наследие империи Чингисхана – улусы Джучи, Угэдея и Чагатая, существовавшие на территории расселения исследуемых популяций с XIII по XV век. Возможно, в какой-то момент улусы стали играть бóльшую роль в структуре генофонда, чем государственные образования, предположили авторы.

Третий тип группировки популяций – жузы – социально-территориальные подразделения казахского ханства (1465–1847). Старший, Средний и Младший жузы, существуют и поныне, поэтому именно они могли стать новым основным фактором в структурировании генофонда.

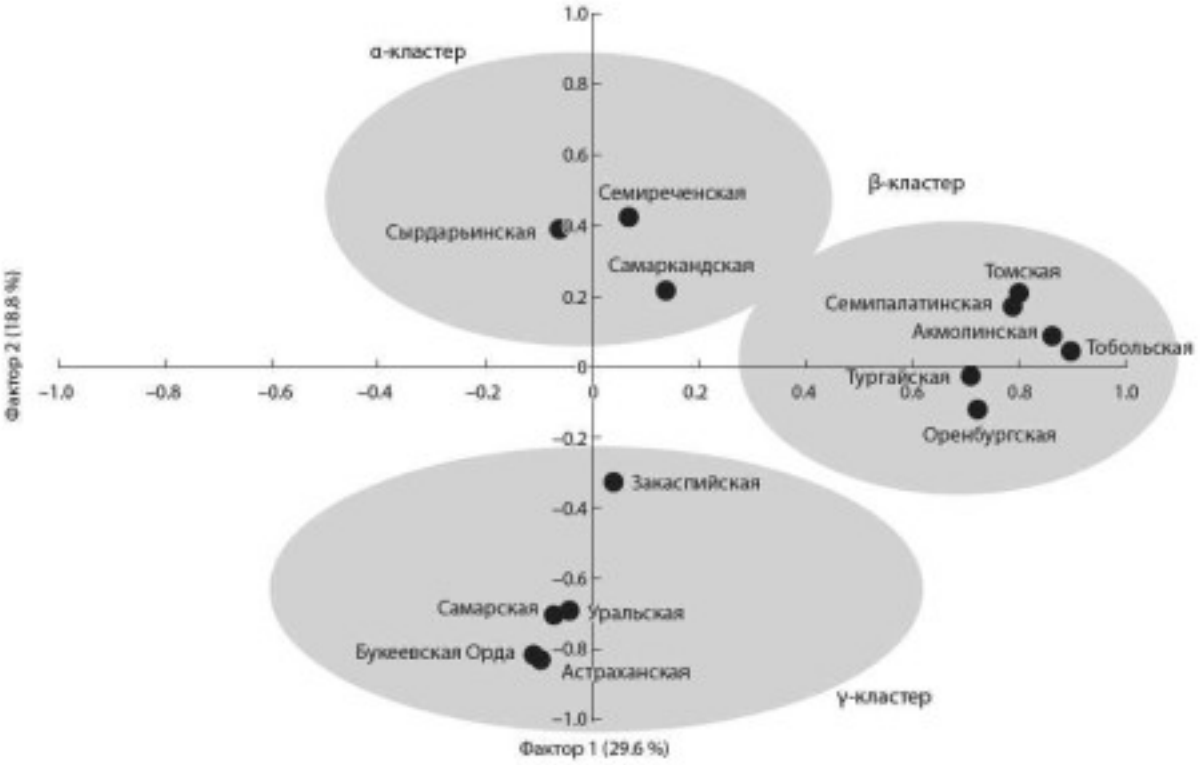
Четвертый тип группировки — губернии, отражает территориально-административное деление Российской империи в

Центральной Азии. Одной из целей такого деления был перевод кочевников к оседлости. Губернии могли также стать значимым фактором в дальнейшем переструктурировании генофонда.

Наконец, пятый тип группировки отражает современное распределение популяций между границами государств (Казахстана, России и стран Средней Азии).

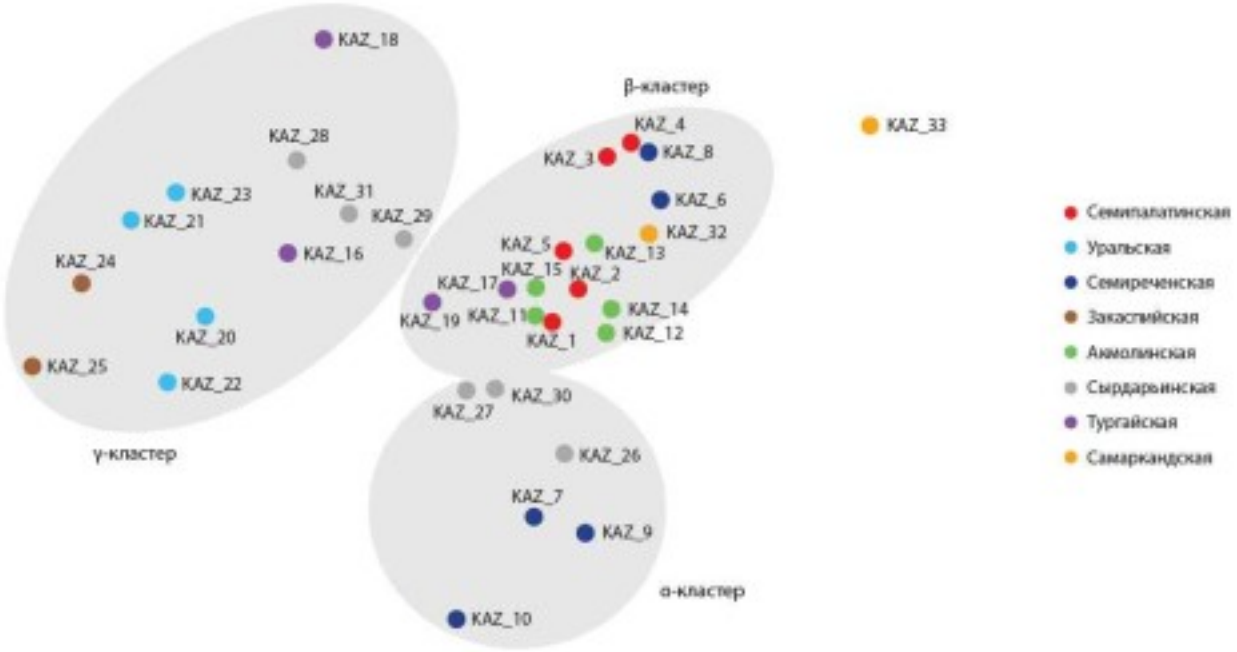
По всем этим группировкам популяций авторы рассчитали межгрупповую изменчивость. Максимальное значение межгрупповой изменчивости (12.3 %) было получено при объединении популяций в жузы. Это говорит о том, что именно во времена казахского ханства, не позднее 550 лет тому назад, и произошло окончательное структурирование казахского генофонда. При группировке по более ранним географическим подразделениям – улусам, изменчивость составила 10.7 %. Это говорит о том, что уже в эпоху Монгольской империи (Золотая Орда) пространственная структура современного генофонда казахов начала формироваться. Как пишут авторы, полученные результаты склоняют чашу весов в пользу гипотезы, что жузы возникли после распада Золотой Орды (альтернативная гипотеза предполагает, что казахи подразделились на жузы еще в X–XII веках, до объединения Чингисханом тюрков и монголов в одну империю). Наконец, наименьшую (3.6 %) межгрупповую изменчивость авторы получили при пятой (современной) группировке популяций.

По данным квазигенетических маркеров (то есть, частотам родов) генетики провели анализ главных компонент казахских популяций, сгруппированных на верхнем уровне административного деления, то есть по областям. На полученном графике областные популяции объединяются в три кластера, соответствующие социально-территориальным объединениям – трем жузам казахов: α-кластер отражает старший жуз; β-кластер – средний; γ-кластер – младший жуз (см рисунок).



Положение популяций казахов верхнего (областного) административного уровня в пространстве главных компонент по данным квазигенетических маркеров.

Когда же исследователи взяли популяции, сгруппированные на нижнем уровне административного деления (уезды) и провели по частотам родов анализ многомерного шкалирования, они получили на графике те же самые кластеры α, β и γ, соответствующие жузам (см. рисунок).



Положение субпопуляций казахов нижнего административного уровня (уезды) на графике многомерного шкалирования, по данным квазигенетических маркеров. Каждый кружок – популяция уезда. Цвет кружков обозначает принадлежность уездов к областям.

Каждый кружок на графике обозначает популяцию уезда, цвет кружка соответствует той или иной области. Авторы отмечают, что в пределах β-кластера обнаруживается сближение юго-восточных популяций KAZ_8, KAZ_6, KAZ_3, KAZ_4 (Семипалатинской и Семиреченской областей) и их связь с KAZ_32 (Самаркандской области), вероятно, связанное с миграцией казахов в Самарканд. Наблюдается также четкое разделение между популяциями вдоль Сырдарьи – низовые (KAZ_28, KAZ_29, KAZ_31) и верховые (KAZ_26, KAZ_27, KAZ_30) уезды Сырдарьинской области.

Таким образом, квазигенетические маркеры (частоты родов) при проведении анализа, стандартного для популяционной генетики, повели себя так, как генетические маркеры (частоты гаплогрупп Y-хромосомы). Это доказывается и поразительным совпадением карт, построенным по частотам родов и частотам гаплогрупп, и корреляцией генетических и квазигенетических расстояний, а также тем, что сгруппированные по разным типам административного деления популяции объединяются в одни и те же кластеры, соответствующие жузам. А именно эти социально-территориальные объединения, по предположению авторов, оказали наибольшее влияние на структуру генофонда казахов.

Итак, работа показывает, что принадлежность к роду – это адекватный квазигенетический маркер, который можно с успехом использовать для исследования структуры генофонда. С помощью этого метода ученые нашли, что высокие межпопуляционные различия проявляются больше всего между регионами, относившимся к разным социально-государственным образованиям в позднемонгольский период и период Казахского ханства. Они делают вывод, что именно эти периоды истории были ключевыми для формирования структуры генофонда. Этим же методом обнаружен более высокий уровень инбридинга в популяциях юго-запада и северо-востока Казахстана, на что следует обратить внимание при медико-генетическом обследовании населения.

Источник:

Жабагин М.К., Балановский О.П., Сабитов Ж.М., Темиргалиев А.З., Агджоян А.Т., Кошель С.М., Раманкулов Е.М., Балановская Е.В. Реконструкция структуры генофонда казахов по данным об их родорасселении // Вавиловский журнал генетики и селекции. 2018. Т.22. №7. С. 895-904 DOI 10.18699/VJ18.431

Приложения к статье с таблицами и картами можно скачать здесь:

Zhabagin et al., 2018b_full_suppl