

¹ НИИ медицинской генетики, Томский национальный исследовательский медицинский центр РАН, г. Томск, vladimir.kharkov@medgenetics.ru

² Дагестанский государственный медицинский университет

³ Институт биохимии и генетики ОСП ФГБНУ УФИЦ РАН

Работа выполнена в рамках Научно-технической программы Союзного государства («ДНК-идентификация») Государственный контракт № 011-17 от 26.09.2017, а также поддержана грантом РФФИ №18-29-13045 «Популяционная геномика и транскриптомика человека: поиск сигналов не-нейтральной эволюции».

Введение:

Изучение генетической структуры современных популяций человека позволяет детально раскрыть ряд вопросов, связанных с формированием их генофонда и этногенезом. Проблематика, касающаяся анализа состава и соотношения различных генетических компонент у народов Кавказа имеет ряд невыясненных вопросов. Современные методы, применяемые в молекулярно-генетических исследованиях, и новые биоинформационные подходы позволяют достоверно выявлять различные предковые генетические компоненты в составе генофонда различных народов. Одним из наиболее широко распространенных на территории России и сопредельных стран является тюркский генетический компонент, маркирующий довольно позднюю волну миграций различных тюркских народов. Республика Дагестан – уникальный район не только России, но и всего мира с точки зрения разнообразия естественно-географических условий, этнокультурного многообразия проживающего здесь населения и историко-генетических проблем. В настоящее время в республике насчитывается около 26 коренных народностей, говорящих на различных языках, в том числе и тюркской группы алтайской языковой семьи.

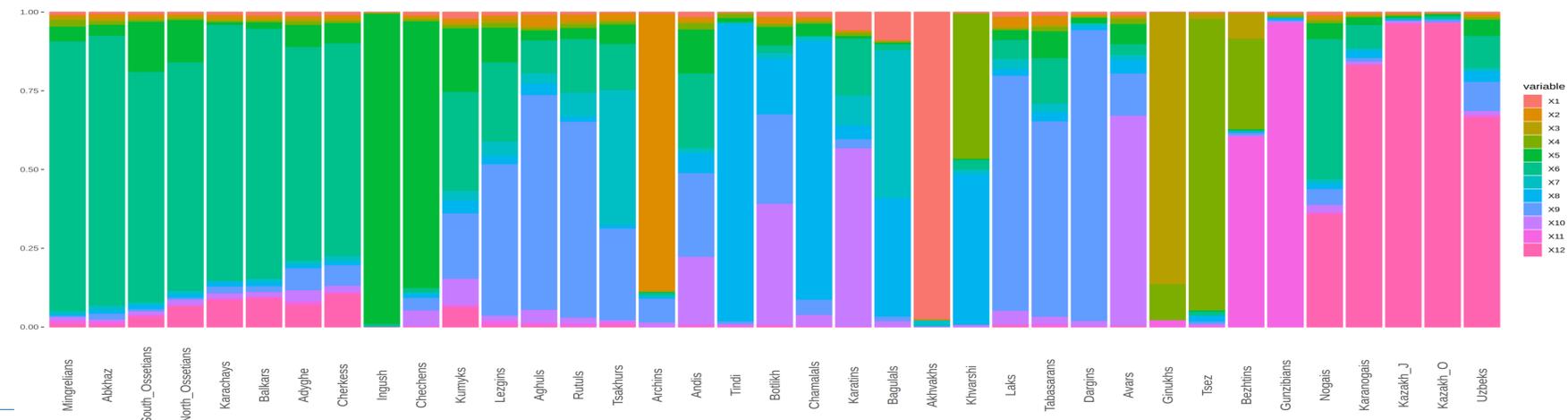
Методы:

Использовались данные генотипов по 1677114 аутосомным SNP (биочип Illumina Multi-Ethnic Global-8) 965 образцов из различных популяций Дагестана, Северного Кавказа и Закавказья (39 выборок) для анализа компонентного состава их генофонда и данные генотипирования более 3000 Y-хромосомных SNP и 44 YSTR у более 3500 образцов мужчин, представляющих коренное население Дагестана, Северного Кавказа, Волго-Уральского региона, Сибири и Средней Азии (более 60 выборок) для сравнительного анализа Y-хромосомных гаплогрупп, маркирующих тюркский генетический компонент. Для выявления компонентов и количества примесей у отдельных индивидов и популяций была использована методика NGS-admix и программа ADMIXTURE. При анализе данных аутосомных генотипов в качестве группы сравнения с наибольшей долей тюркского компонента были использованы различные выборки казахов и узбеков.

Популяции

Казахи	97
Караногайцы	84
Узбеки	67
Ногайцы	36
Черкесы	10
Балкарцы	9,3
Карачаевцы	8,7
Кумыки	7,5
Адыги	7
Осетины С	7
Осетины Ю	3
Ингуши	3
Чеченцы	3
Цахуры	2
Агулы	2
Лезгины	2
Мегрелы	1
Лакцы	1
Аварцы	1
Табасараны	1
Рутульцы	0,5
Абхазы	0,5
Андийцы	0,2
Ботлихцы	0
Даргинцы	0
Чамалинцы	0
Тиндинцы	0
Багулалы	0
Ахвахцы	0
Арчинцы	0
Бежтинцы	0
Гинухцы	0
Гунзибцы	0
Каратинцы	0
Хваршины	0
Цезы	0

Доля Тюркского компонента (%)



Результаты:

Анализ данных по аутосомным SNP в кавказских популяциях показал, что общая картина распределения компонент хорошо согласуется с географическим местоположением исследуемых популяций и привязкой к конкретному региону. При относительно небольшом числе компонент (K=3-4) большинство популяций характеризуются преобладанием отдельного компонента. Начиная от 6 компонент и более, в некоторых популяциях Дагестана выделяются отдельные этноспецифичные компоненты (ахвахцы, гинухцы, бежтинцы, гунзибцы). Начиная от K=3 в общем массиве анализируемых популяций Кавказа отчетливо выделяется компонент, практически полностью доминирующий у казахов (97%) и узбеков (67%). Именно этот генетический компонент маркирует вклад в генофонд населения Кавказа тюркских мигрантов (Компонент 12 на Рисунке). Из популяций Дагестана он полностью преобладает у тюркоязычных караногайцев (84%) и вторым по частоте у ногайцев (36%). У тюркоязычных карачаевцев и балкарцев на долю этого компонента приходится по 8,7% и 9,3%. В генофонде тюркоязычных кумыков он составляет лишь 7,5%. Этот компонент с очень небольшой частотой выявлен у большинства равнинных народов Дагестана – аварцев (от 1 до 2%), лакцев (1%), и табасаран (1%), а также народов лезгинской группы (лезгинов (1%), агулов (1%), рутульцев (1%) и цахуров (2%)). Практически все популяции горных районов Дагестана демонстрируют полное отсутствие тюркского компонента. Таким образом, очень небольшие частоты тюркского компонента в отдельных равнинных популяциях не являются статистической погрешностью, а отражают очень небольшой, но реальный вклад предковых групп в генофонд современных этносов.

Анализ распределения частот гаплогрупп Y-хромосомы обнаружил наличие у всех тюркоязычных народов Кавказа различных азиатских по происхождению линий. В генофонде караногайцев сочетаются несколько линий: R1b1a1a1-M73-M478, C2-Z1453, N1a2b-P43, O2a2b-P164-CTS4723. У кумыков, аварцев, лезгин, агулов, рутульцев и тиндинцев обнаружены только образцы, принадлежащие к азиатским гаплогруппам R1a1a1b2-Z93 и R1a1a1b2a2b-Z93-S4576. У кумыков, также как и у караногайцев, имеется гаплогруппа N1a2b-P43. У рутульцев и даргинцев выявлена гаплогруппа O2a2b-P164-CTS4723.

Выводы:

Тюркоязычные популяции отличаются неоднородностью по доле тюркского по происхождению генетического компонента в составе их генофондов. Наибольшая его доля характерна для караногайцев, минимальная для кумыков. Различия по спектру азиатских гаплогрупп Y-хромосомы свидетельствуют о различных по времени миграциях тюркоязычных групп, которые внесли свой вклад в генофонд современного населения Дагестана и Северного Кавказа.