

АНАЛИЗ ФАМИЛЬНОГО ЛАНДШАФТА КАРАЧАЕВО-ЧЕРКЕССИИ

Ельчинова Г.И.¹, Макаов А.Х.-М.², Зинченко Р.А.^{1,3}

¹ ФГБНУ Медико-генетический научный центр, Москва 115478, elchinova@med-gen.ru;

² Муниципальное бюджетное лечебно-профилактическое учреждение «Хабезская центральная районная больница», Хабез Карачаево-Черкесской Республики, e-mail: makaov@yandex.ru;

³ Российский национальный исследовательский медицинский университет имени Н.И.Пирогова Министерства здравоохранения Российской Федерации, кафедра молекулярной и клеточной генетики, Москва, 117997, renazinchenko@mail.ru

Исследование выполнено в рамках комплексного медико-популяционно-генетического изучения населения Карачаево-Черкесии. Для анализа популяционно-генетической структуры населения Карачаево-Черкесии для популяции ранга «республика» в работе в качестве квазигенетического маркера использованы фамилии, выбранные тотально из списков избирателей. Проведен кластерный анализ (пошаговая кластеризация) по 1022 частым фамилиям (частота более 0,1%) с использованием среднесвязывающего метода. При относительно ожидаемом порядке кластеризации высокие значения фамильных дистанций оказались необычными. Причина, предположительно, в тухумной организации социума на Северном Кавказе, а не в отсутствии брачных миграций между районами. Сделан вывод о том, что при использовании фамилий в качестве квазигенетического маркера при медико-биологических и популяционно-генетических исследованиях помимо стандартных ограничений (монофилетичность, патроклинное наследование, длительность существования) необходимо еще одно – моноэтничность популяции.

Ключевые слова: Карачаево-Черкесия, фамилия, популяционно-генетические исследования, кластерный анализ, квазигенетический маркер.

ANALYSIS OF SURNAME LANDSCAPE KARACHAY-CHERKESSIA

Elchinova G.I.¹, Makaov A.K.-M.², Zinchenko R.A.^{1,3}

¹ Federal state scientific budgetary Institution «Research Centre for Medical Genetics» Moscow, 115478, e-mail: elchinova@med-gen.ru;

² Municipal Budgetary Health Care setting "Habezskaya central district hospital", Habez Karachai-Cherkess Republic makaov@yandex.ru;

³ Pirogov Russian National Research Medical University, Moscow, 117997, Russia, e-mail: renazinchenko@mail.ru

The study is performed within the integrated medical and population-genetic study of Karachay-Cherkessia people. Chosen totally from voter lists surnames are used as quasigenetic marker for the analysis of population genetic structure of rank "Republic. Cluster analysis (incremental clustering) of surname distances calculated by 1022 frequent surnames (frequency 0.1%) using average linkage clustering is carried out. The order of clustering has been expected, but high values of surname distances are unusual. The reason of that can be in society clan organization in the North Caucasus, but not in the absence of marriage migration between districts. Monoethnicity of the population turned out to be another condition for the use of surnames as quasigenetic marker in biomedical and population-genetic studies along with standard restrictions (monofilament, patrocina inheritance, duration of existence).

Keywords: Karachay-Cherkessia, surname, population-genetical research, cluster analysis, quasigenetic marker

Комплексное медико-популяционно-генетическое изучение населения Карачаево-Черкесской республики (КЧР) проводится сотрудниками лаборатории генетической эпидемиологии Медико-генетического научного центра с 2013 года. Исследование проводится по стандартному протоколу, разработанному в лаборатории более 30 лет назад под руководством Е.К.Гинтера и А.А.Ревазова и успешно применяемому в настоящее время. Единообразие сбора и обработки материала позволяет проводить адекватное сравнение результатов всех исследований по данному направлению.

Протокол включает в себя обследование населения выездной бригадой врачей различных специальностей, ориентированных на диагностику наследственной патологии (генетик, педиатр, офтальмолог, ортопед, невролог, дерматолог, сурдолог). Подтверждающая ДНК-диагностика осуществляется сотрудниками других лабораторий Центра или (в случае полноэкзомного и полногеномного секвенирования) другими институтами. Сбор популяционно-генетической и генетико-демографической информации как биологического (кровь), так и небιологического (паспортные и анкетные данные) характера осуществляется в этих же популяциях, что предполагает комплексную оценку населения регионов через различные системы. Далее составляются базы данных, реестры, регистры.

Ранее подобным образом обследовано население Костромской, Кировской, Архангельской, Брянской, Ростовской областей, республик Марий Эл, Чувашия, Удмуртия, Татарстан, Башкортостан, Адыгея, Краснодарского края с общей численностью обследованного населения более 3 млн. человек. Обследование населения по единому стандарту позволяет проводить различные виды сравнительного анализа, не только выявляя очаги накопления отдельных нозологий в регионах и этносах, но и объясняя механизм локальной аккумуляции заболеваний [8]. Результаты этих исследований отражены в отечественной (журналы Генетика, Медицинская генетика, Педиатрия, Офтальмология, Стоматолог и др.) и зарубежной печати (Science, European Journal of Human Genetics и др.), издан ряд монографий.

Неоднократно было показано, что особенности распространения наследственной патологии, а также характер груза и спектр наследственных болезней находятся в прямой зависимости от популяционно-генетических параметров изучаемого региона. На основании рассчитанных популяционно-генетических характеристик осуществлен успешный прогноз распространенности наследственной патологии на еще необследованные территории региона [7]. Молекулярно-генетические анализы дорогостоящи и трудоемки, а получение популяционно-генетических характеристик из небιологических источников информации позволяет провести экспресс-оценку параметров популяционно-генетической структуры относительно дешево, легко и быстро.

К небιологическим источникам информации относятся частоты фамилий, обычно получаемые из списков избирателей, длины брачных миграций, подсчитанные на основании мест рождения супругов по сведениям ЗАГСа или похозяйственных книг, и демографические анкеты, полученные путем опроса женщин пострепродуктивного возраста (старше 45 лет). Фамилии давно традиционно используются при популяционно-генетических исследованиях в качестве квазигенетического маркера как зарубежными, так и отечественными исследователями. По классическому постулату Мортонa «использование

фамилий в качестве селективно нейтрального маркера имеет информационную ценность, равную лучшей кодоминантной генетической системе» [12]. Использование фамилий в качестве квазигенетического маркера имеет ряд ограничений. В популяции должно практиковаться патроклинное (т.е. по отцу) наследование фамилий, фамилии должны использоваться достаточно долго (не менее 10 поколений) и иметь монофилетичное происхождение (хотя бы априорно).

Карачаево-Черкесия характеризуется полиэтничным составом населения, причем пять этносов – карачаевцы, русские, черкесы, абазины и ногайцы - значительно представлены в Республике. По материалам Всероссийской переписи населения 2010 года [1] карачаевцы составляют 40,67%, русские – 31,40%, черкесы – 11,82%, абазины – 7,73%, ногайцы – 3,28% населения. В Республике наблюдается компактное проживание представителей различных этнических групп: карачаевцы – в Карачаевском, Малокарачаевском, Усть-Джегутинском и Прикубанском районах, абазины – в Абазинском районе, ногайцы – в Ногайском, черкесы – в Хабезском районе. В остальных районах представители этих этносов проживают также компактно, но занимают лишь часть района. Русские проживают преимущественно в городах – Черкесске, Карачаевске, Усть-Джегуте, Теберде. Впервые в наших исследованиях мы работаем в популяции, где широко представлены более трех этнических групп населения.

Фамилии успешно применены нами для описания популяционно-генетической структуры карачаевцев [4], черкесов, ногайцев [5] и абазин [3] по отдельности. Скоррелированность параметров, полученных разными методами из различных источников информации свидетельствует о корректности проведенных исследований и расчетов. Для русских распределение фамилий не рассматривалось и, соответственно, не вычислялись популяционно-генетические характеристики, т.к. русские не являются коренным населением Северного Кавказа и в значительной степени являются мигрантами из разных регионов России и не только РФ, не образуют популяцию в классическом смысле этого определения: «под термином «популяция» понимают совокупность особей определенного вида, в течение достаточно длительного времени (большого числа поколений) населяющих определенное пространство, внутри которого осуществляется та или иная степень панмиксии и нет заметных изоляционных барьеров, которая отделена от соседних таких же совокупностей особей данного вида той или иной степенью давления тех или иных форм изоляции» [10]. Однако вопрос совокупности основных этносов в рамках КЧР до сих пор нами не рассматривался. Изучению фамильных статистик для популяции ранга «республика» посвящено настоящее сообщение. Помимо чисто утилитарных потребностей медико-генетического консультирования популяционно-генетические исследования человека

являются самостоятельным научным направлением и представляют интерес для генетиков, антропологов, историков, краеведов, социологов.

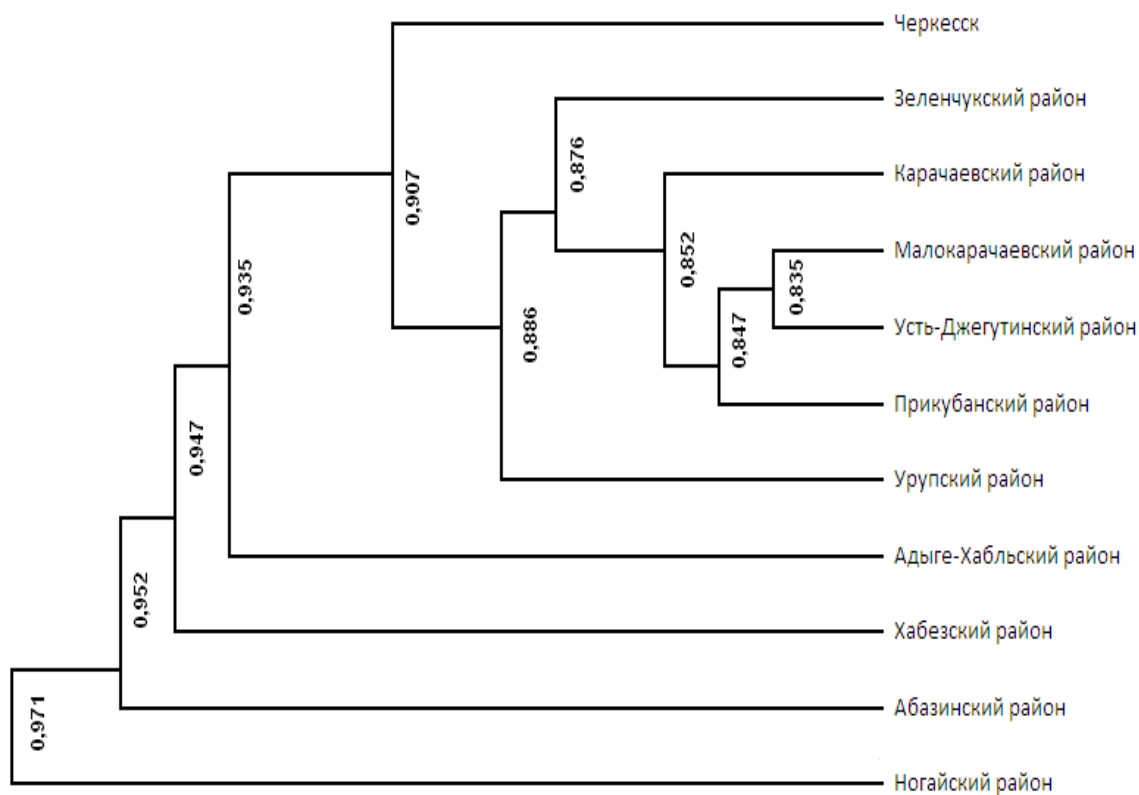


Рис.1 Дендрограмма по матрице фамильных дистанций по частым фамилиям

На рис.1 представлена дендрограмма, составленная по треугольной матрице фамильных дистанций размера [11×11], подсчитанных по частым фамилиям КЧР (хотя бы в одном районе частота фамилии более 0,1%), в анализе участвует 1022 фамилии. Фамильные дистанции рассчитаны по формуле:

$$D_{12} = \frac{2}{\pi} \arccos\left(\sum_k \sqrt{p_{1k} p_{2k}}\right)$$

Фамилии рассматривались как аллели одного локуса, p_{1k} и p_{2k} частоты k -ой фамилии в изучаемых 1-ой и 2-ой популяциях [11]. Пошаговая кластеризация проведена среднесвязывающим методом [2]. Фамилии выбраны из списков избирателей тотально, затем отобраны частые. Использование частых фамилий отсекает фамилии случайных мигрантов и вырождающиеся фамилии, так что в основном мы имеем дело с фамилиями коренного населения популяции. Допускается использование и других частотных критериев. Мы

выбрали именно этот, т.к. (см.формулу) при произведении двух величин менее 0,001 каждая под радикалом получается практически нулевое значение. Использование подобного метода увеличивает масштаб рассматриваемых характеристик, делая более значимыми имеющиеся различия. Поскольку дети обычно носят фамилию одного из родителей (редко – обоих), то мы полагаем, что имеем дело с тотальной выборкой фамилий.

Порядок кластеризации представляется достаточно естественным и ожидаемым – сначала в общий кластер объединяются районы, населенные преимущественно карачаевцами (Карачевский, Малокарачаевский, Усть-Джегутинский, Прикубанский), затем к ним присоединяются районы (Зеленчукский и Урупский) и г.Черкесск с меньшей долей карачаевцев в населении, а затем районы, население которых представлено преимущественно черкесами, абазинами и ногайцами. При этом обращают на себя очень высокие значения дистанций, на которых начинается кластеризация (0,835). Первое объяснение, которое приходит в голову, это отсутствие брачных миграций между районами, по крайней мере, что они минимальны. Однако, невысокие значения индекса эндогамии даже в районах, где титульной нацией являются карачаевцы (0,34 в Прикубанском, 0,47 в Усть-Джегутинском, 0,56 в Карачаевском, 0,63 в Малокарачаевском [6]), противоречат этой гипотезе. Среднее значение дистанционной матрицы составило 0,926. В то же время, ни одно из значений матрицы не достигает единицы, поэтому нельзя утверждать, что между какими-то районами отсутствуют брачные контакты. Видимо, подобная особенность кластеризации является следствием тухумной организации социума Северного Кавказа [9], сохранившейся до настоящего времени. Заметим также, что Черкесск – столица Республики – не является центром кластеризации, что также достаточно нетипично, и это является следствием полиэтничного состава населения КЧР.

Еще одной важнейшей популяционно-генетической характеристикой является значение случайного инбридинга F_{st} , вычисляемого как четверть суммы квадратов частот фамилий [11]. Под инбридингом мы в данном случае понимаем вероятность случайной встречи в зиготе идентичных по происхождению гамет. В свете всего вышесказанного оценкой F_{st} для популяции ранга район следует пользоваться с осторожностью. Поэтому в табл.1 мы приводим значения F_{st} , подсчитанные для популяции ранга район и средневзвешенные значения для сельских поселений района (популяция ранга «сельсовет»). Эти характеристики подсчитаны по всему списку фамилий. Кроме того, мы не исключаем возможности того, что более корректными могут оказаться значения F_{st} , полученные тем же способом по частым фамилиям. В табл. 1 приводятся значения случайного инбридинга, полученные четырьмя различными способами. Который из них окажется более корректным при работе в полиэтничной популяции, покажет дальнейшее сравнение этих значений с

другими популяционно-генетическими характеристиками. Это будет одним из следующих этапов нашей работы по изучению популяционно-генетических параметров населения Карачаево-Черкесии.

Таблица 1

Значения случайного инбридинга F_{st} , полученные различными способами

район	по всему списку фамилий		по частым фамилиям	
	для популяции ранга «район»	средневзвешенное значение для популяции ранга «сельсовет»	для популяции ранга «район»	средневзвешенное значение для популяции ранга «сельсовет»
Усть-Джегутинский	0,0016	0,0024	0,0038	0,0051
Адыге-Хабльский	0,0008	0,0040	0,0019	0,0085
Карачаевский	0,0015	0,0045	0,0027	0,0076
Малокарачаевский	0,0035	0,0056	0,0047	0,0074
Прикубанский	0,0015	0,0033	0,0033	0,0062
Абазинский	0,0018	0,0054	0,0026	0,0078
Урупский	0,0003	0,0011	0,0022	0,0046
Хабезский	0,0015	0,0080	0,0020	0,0105
Ногайский	0,0023	0,0073	0,0051	0,0091
Зеленчукский	0,0008	0,0021	0,0033	0,0063

Заметим лишь, что значения, полученные по всему списку фамилий, априорно представляются более логичными: наиболее низкие значения (как для популяции ранга «район», так и средневзвешенные значения для популяций ранга «сельсовет» в районе) в Урупском районе с преобладанием русского населения, низкие значения в Адыге-Хабльском и Зеленчукском районах с полиэтническим составом при отсутствии титульной нации.

Таким образом, успешно описав при помощи фамилий популяционно-генетическую структуру коренного населения Карачаево-Черкесии (каждый этнос по отдельности), мы не сумели извлечь всю ожидаемую полноту генетической информации при попытке совместного анализа популяционно-генетической структуры совокупности различных этносов. Тем самым мы подтвердили осторожно высказываемую нами ранее гипотезу о том, что при использовании фамилий в качестве квазигенетического маркера при медико-биологических исследованиях помимо стандартных ограничений на их применение (монофилетичность, патроклинное наследование, длительность существования) необходимо еще одно – моноэтничность популяции. Сделать такой вывод окончательно нам позволила именно работа в популяции с полиэтническим составом населения.

Список литературы

1. Всероссийская перепись населения 2010 года // http://www.gks.ru/free_doc/new_site/perepis2010/croc/results2.html (дата обращения 29.09.2015).
2. Дерябин В.Е. Многомерная биометрия для антропологов // М.: МГУ. 1983. - 227 с.
3. Ельчинова Г.И., Шакманов М.М., Ревазова Ю.А., Зинченко Р.А. Популяционно-генетическая характеристика абазин Карачаево-Черкесии (по данным о распределении фамилий) и брачных миграциях // Генетика. – 2015. – Т.51. – № 10. – С. 1194-1190.
4. Ельчинова Г.И., Эльканова Л.А., Зинченко Р.А. Распределение карачаевских фамилий // Вестник Московского университета. Серия XXIII. Антропология. – 2015. – № 1. – С. 111-116.
5. Ельчинова Г.И., Ревазова Ю.А., Зинченко Р.А. Популяционно-генетическая характеристика ногайцев Карачаево-Черкесии (по брачным миграциям и частотам распределения фамилий) // Серия XXIII. Антропология. – 2016. – В печати.
6. Ельчинова Г.И., Шакманов М.М., Ревазова Ю.А. и др. Брачно-миграционная характеристика карачаевцев // Генетика. - 2015. - Т. 51. - № 8. - С. 941-945
7. Ельчинова Г.И., Зинченко Р.А., Кривенцова Н.В., Амелина С.С., Тереховская И.Г., Валькова Т.И., Вальков В.А., Салимов Р.Р. Прогнозирование распространенности наследственной патологии на основании инбридинга и случайной изонимии // Медицинская генетика. – 2007. – Т. 6. – № 11 (65). – С. 29-33.
8. Зинченко Р.А., Ельчинова Г.И., Галкина В.А., Кириллов А.Г., Аbruкова А.В., Петрова Н.В., Тимковская Е.Е., Зинченко С.П., Шокарев Р.А., Морозова А.А., Близнец Е.А., Вассерман Н.Н., Степанова А.А., Поляков А.В., Гинтер Е.К. Дифференциация этнических групп России по генам наследственных болезней // Медицинская генетика. – 2007. – Т.6, № 2. – С. 29-37.
9. Историко-культурные традиции народов Северного Кавказа. Научно-справочное пособие / Под ред. акад. Тишкова В.А. – М.: ИЭА РАН. 2013. – 114 с.
10. Тимофеев-Ресовский Н.В., Яблоков А.В., Глотов Н.В. Очерк учения о популяциях. – М.: Наука. 1973. – 277 с.
11. Cavalli-Sforza L.L., Bodmer W.F. The Genetics of Human populations // San Francisco: Ed. W.H.Freeman and Company. –1971. – 965 p.

12. Morton N.E., Yee S., Harris D.E., Lew R. Bioassay of kinship // Theoretical Population Biology. – 1971. – V. 2. – P. 507-521.

Рецензенты:

Петрова Н.В., д.б.н., в.н.с. лаборатории генетической эпидемиологии, ФГБНУ «Медико-генетический научный центр», г. Москва;

Хлебникова О.В., д.м.н., врач-генетик, в.н.с. лаборатории генетической эпидемиологии ФГБНУ «Медико-генетический научный центр», г. Москва.