Генетические взаимоотношения популяций Европы: три уровня

Олег Балановский

Фрагмент из книги "Генофонд Европы"

Проведя картографический анализ отдельных гаплогрупп Y-хромосомы и убедившись в четкой географической приуроченности большинства из них, перейдем к обобщенному (по всем гаплогруппам совместно) статистическому анализу генофонда Европы.

МАССИВЫ ДАННЫХ ДЛЯ СТАТИСТИЧЕСКОГО АНАЛИЗА

Массив данных для статистического анализа, как и массив данных для карт в предыдущем разделе, сформирован на основе нашей базы данных Y-base (раздел 1.1). Как и для картографирования, в массив данных для статистического анализа отбирались только выборки достаточного объема (по возможности более 70 образцов), а меньшие выборки или исключались, или – в пределах одного народа – объединялись друг с другом. Отличие статистического анализа от картографического состоит в том, что каждая популяция должна быть изучена по каждой гаплогруппе. Однако в большинстве статей, особенно не самых последних лет, использовались ограниченные наборы гаплогрупп. Поэтому, если требовать, чтобы каждая выборка была изучена по каждой гаплогруппе, то резко сократится или набор гаплогрупп (менее чем до 10) или набор популяций.

Чтобы избежать и того, и другого, данные по всем выборкам в пределах одного народа были усреднены и в анализе использовались эти среднеэтнические частоты. В таком случае, если по данному народу изучено 10 популяций (выборок), то достаточно, чтобы хотя бы одна из этих популяций была изучена по данной гаплогруппе, и тогда частота этой гаплогруппы в этой популяции будет присвоена всему народу. Если же гаплогруппа была изучена в нескольких региональных популяциях народа, то еще лучше — эти значения усредняются и тогда присвоенное народу значение частоты гаплогруппы получается более надежным. При таком подходе, ориентированном не на локальные популяции, а на народы (этнический уровень популяционной системы) число пустых ячеек в таблице частот гаплогрупп резко сокращается и удается охарактеризовать большинство народов Европы по большинству гаплогрупп.

Правда, какое-то число пустых ячеек все равно остается (для какого-либо народа вообще нет информации о частоте у него какой-либо гаплогруппы). В таких случаях сохраняется проблема выбора – исключать из анализа такой народ или исключать такую гаплогруппу. Анализ был проведен в обоих вариантах и были созданы два массива данных: массив частот 17 гаплогрупп у 62 народов (меньше гаплогрупп, но больше народов) и массив частот 40 гаплогрупп у 40 народов (больше гаплогрупп, но меньше народов). Это увеличение числа гаплогрупп достигнуто за счет подразделения трех основных гаплогрупп Европа (R1b, R1a и N1c) на их ветви.

При описанном анализе на этническом уровне мы избегаем и искажений частот гаплогрупп, и потери числа гаплогрупп или числа народов. Но кое-что все-таки теряется: при таком подходе из поля зрения выпадает изменчивость на внутриэтническом уровне, то есть никак не могут проявиться различия между популяциями, относящимися к одному народу. Чтобы изучить изменчивость и на этом – внутриэтническом – уровне, был создан третий массив данных. В него вошли данные только по тем народам, для которых имелись данные хотя бы о двух разных популяциях этого народа. Число гаплогрупп составило 15 (две гаплогруппы из набора 17 гаплогрупп были исключены из рассмотрения), число народов составило 20, а число популяций этих народов было большим – 120 популяций.

Наконец, поскольку ориентироваться в большом числе точек на графике бывает нелегко, был создан еще один упрощенный массив данных – региональный. Были выделены 15 регионов Европы, частоты гаплогрупп в которых были рассчитаны усреднением частот у народов, относящихся к этим регионам.

В итоге были получены массивы данных, характеризующие генофонд Европы на трех иерархических уровнях: региональном, этническом и субэтническом. Причем этнический уровень охарактеризован двумя массивами данных: по 17 и по 40 гаплогруппам. Таблица 2.4. суммирует важнейшую информацию о всех этих массивах данных. А таблица 2.5. показывает для каждого народа, в каких массивах он представлен.

МАССИВЫ ДАННЫХ ДЛЯ СТАТИСТИЧЕСКОГО АНАЛИЗА

Таблица 2.4. Массивы данных, сформированные для статистического анализа генофонда Европы по маркерам Y-хромосомы.

Название массива данных	Число анализируемых популяций	Число гаплогрупп	Перечень гаплогрупп
Региональный	**15 регионов Европы**	**17**	**C-M130, E-M35, G-M201, I-M253, I-P37.2, I-M223, I-M170(xM253,P37.2,M223), J-M304, N-LLY22g, R-M198(xM458), R-M458, R-M343, L-M11, Q-M242, R-M124, T-M70, other (A, D, E, F, H, I, K, O)**
Этнический 17	**62 народа Европы**	**17**	**Те же гаплогруппы**
Этнический 40	**40 народов Европы**	**40**	**E-M35, G-M201, I-M253, I-P37.2, I-M223, J-M304, C-M130, I-M170(xM253,P37.2,M223), L-M11, Q-M242, R-M124, T-M70, R-M198(xM458,M558,Z284,Z93), R-M458, R-M558, R-Z284, R-Z93, R-M343(xM73xM269), R-V88, R-M73, R-M269(xL23), R-L23(xM412), R-M412(xL11), R-L11(xU106xS116), R-U106(xU198), R-U198, R-S116(xM529xU152), R-M529(xM222), R-M222, R-U152, N-M178(xP298), N-L392(xCTS10760,Z1936,F4205,B202), N-Z1936, N-CTS10760, N-F4205, N-B202, N-P298, N-M2118, N-B211, other (A, D, E, F, H, I, K, O)**
Субэтнический	**120 популяций двадцати народов**	**15**	**C-M130, E-M35, G-M201, I-M253, I-P37.2,I-M223, J-M304, L-M11, N-LLY22g, Q-M242, R-M198, R-M343, R-M124, T-M70, other (A, D, E, F, H, I, K, O)**

Таблица 2.5. Представленность народов Европы в анализируемых массивах данных.

Народ	Представлен в массиве «Этнический 17»	Представлен в массиве «Этнический 40»	Представлен в массиве «Субэтнический» (в скобках – число популяций)	К какому региону относится в массиве «Региональный».
Абазины	**+**	**+**	**-**	**Западный Кавказ**
Абхазы	**+**	**+**	**-**	**Закавказье**
Аварцы	**+**	**+**	**-**	**Восточный Кавказ**
Адыги	**+**	**+**	**-**	**Западный Кавказ**
Албанцы	**+**	**-**	**+ (3)**	**Балканский п-в**
Англичане	**+**	**+**	**-**	**Британские о-ва**
Армяне	**+**	**+**	**-**	**Анатолия**
Балкарцы	**+**	**+**	**-**	**Западный Кавказ**

Баски	**+**	**-**	**-**	**Пиренейский п-в**
Башкиры	**+**	**-**	**+ (5)**	**Степи Европы**
Белорусы	**+**	**+**	**+ (8)**	**Восточно-Европейская равнина**
Бельгийцы	**-**	**-**	**+ (5)**	**Западная Балтика**
Болгары	**+**	**-**	**+ (9)**	**Балканский п-в**
Боснийцы	**+**	**+**	**+ (2)**	**Балканский п-в**
Венгры	**+**	**+**	**-**	**Степи Европы**
Галисийцы	**+**	**-**	**-**	**Пиренейский п-в**
Герцеговинцы	**+**	**+**	**-**	**Балканский п-в**
Греки	**+**	**+**	**+ (5)**	**Балканский п-в**
Грузины	**+**	**-**	**-**	**Закавказье**
Даргинцы	**+**	**+**	**-**	**Восточный Кавказ**
Датчане	**+**	**+**	**-**	**Западная Балтика**
Ингуши	**+**	**+**	**-**	**Восточный Кавказ**
Ирландцы	**+**	**+**	**-**	**Британские о-ва**
Испанцы	**+**	**+**	**-**	**Пиренейский п-в**
Итальянцы	**+**	**+**	**-**	**Аппенинский п-в**
Караногайцы	**+**	**+**	**-**	**Степи Европы**
Карачаевцы	**+**	**+**	**-**	**Западный Кавказ**
Карелы	**+**	**+**	**-**	**Северо-Восточная Европа**
Каталонцы	**+**	**-**	**-**	**Пиренейский п-в**
Коми	**+**	**+**	**+ (2)**	**Северо-Восточная Европа**
Кубачинцы	**+**	**+**	**-**	**Восточный Кавказ**
Кумыки	**+**	**+**	**-**	**Восточный Кавказ**
Латыши	**+**	**-**	**-**	**Восточная Балтика**
Литовцы	**+**	**-**	**+ (6)**	**Восточная Балтика**
Македонцы	**+**	**+**	**+ (2)**	**Балканский п-в**
Мокша и эрзя (мордва)	**+**	**-**	**+ (2)**	**Правобережье Волги**
Немцы	**+**	**+**	**-**	**Западная Балтика**
Ногайцы	**+**	**-**	**-**	**Степи Европы**
Осетины	**+**	**+**	**-**	**Западный Кавказ**
Поляки	**+**	**+**	**+ (6)**	**Восточно-Европейская равнина**
Португальцы	**+**	**+**	**-**	**Пиренейский п-в**
Румыны	**+**	**+**	**-**	**Балканский п-в**
Русские	**+**	**+**	**+ (20)**	**Восточно-Европейская равнина**
Саамы	**+**	**-**	**-**	**Северо-Восточная Европа**
Сардинцы	**+**	**-**	**-**	**Аппенинский п-в**
Сербы	**+**	**-**	**+ (4)**	**Балканский п-в**
Словаки	**+**	**-**	**-**	**Восточно-Европейская равнина**
Словенцы	**+**	**+**	**-**	**Балканский п-в**
Сорбы	**+**	**-**	**-**	**Западная Балтика**
Татары казанские	**+**	**+**	**-**	**Восточно-Европейская равнина**
Татары крымские	**+**	**-**	**-**	**Степи Европы**
Турки	**+**	**+**	**+ (6)**	**Анатолия**
Украинцы	**+**	**+**	**+ (13)**	**Восточно-Европейская равнина**
Финны	**+**	**-**	**+ (7)**	**Северо-Восточная Европа**
Французы	**+**	**-**	**-**	**Франция**
Хорваты	**+**	**+**	**+ (3)**	**Балканский п-в**
Черкесы	**+**	**-**	**-**	**Западный Кавказ**
Чехи	**+**	**+**	**+ (5)**	**Восточно-Европейская равнина**

Чеченцы	**+**	**-**	**-**	**Восточный Кавказ**
Чуваши	**+**	**+**	**-**	**Правобережье Волги**
Шведы	**+**	**+**	**+ (7)**	**Западная Балтика**
Эстонцы	**+**	**+**	**-**	**Восточная Балтика**

РЕГИОНАЛЬНЫЙ УРОВЕНЬ

Рассмотрение на региональном уровне является предварительным – мы сначала посмотрим на Европу как бы из космоса, чтобы уловить скелет, основной паттерн в структуре ее генофонда, а потом уже перейдем к главному анализу — на этническом уровне. Генетические взаимоотношение регионов Европы представлены на рис. 2.38.

Прежде всего, хорошо видны отличия генофонда Кавказа: Западный Кавказ и Восточный Кавказ при взгляде «из космоса» генетически близки и друг к другу, и к народам Закавказья и Анатолии. (Последние два, строго говоря, не относятся к Европе, но включены в анализ для полноты картины). Зато все эти четыре региональных генофонда расположены далеко от генофондов собственно Европы.

Европейские же генофонды делятся на два кластера – Западной Европы и Восточной. В пределах Западной Европы генофонды Средиземноморья (Аппенинский и Пиренейский полуострова) находятся хотя и не близко к анатолийско-кавказскому кругу, но все же ближе к ним, чем генофонды Франции и Британских островов. А эти генофонды, географически более северные, и в генетическом пространстве находятся выше («севернее») средиземноморских.

Мостом между западно-европейскими и восточно-европейскими генофондами служит население Балтики, причем Западная Балтика генетически ближе к Западной Европе, а Восточная Балтика – к Восточной Европе. Географически так и должно быть, поскольку «Западная Балтика» в нашем понимании (табл. 2.5) включает и Швецию, и Данию, и Германию, то есть значительную часть пограничья Западной и Восточной Европы. Правда, регион «Восточно-Европейская равнина» также граничит с Западной Европой. Но ее генофонд на графике и расположился очень близко к генофонду Восточной Балтики.

Положение восточноевропейских генофондов относительно друг друга также вторит географии: популяции Северо-Восточной Европы и в генетическом пространстве (рис. 2.38 А) располагаются справа и выше («северо-восточнее») от популяций Восточно-европейской равнины. А популяции Правобережья Волги (табл. 2.5), которые географически южнее большинства популяций Восточно-Европейской равнины, и на генетическом графике располагаются ниже («южнее») ее (рис. 2.38 А). Популяции восточноевропейских степей, географически расположенные на самом юге Восточной Европы и примыкающие к Кавказу, и на генетическом графике оказываются «южнее» остальных восточно-европейских регионов, и оказываются единственным регионом Европы, генетически приближенным к Кавказу.

Итак, генетические взаимоотношения регионов Европы отражают соотношения их географического положения друг относительно друга. Учитывая, как много только исторически документированных миграций (и, вероятно, еще больше не документированных) происходило в Европе за все время истории ее населения и во всех мыслимых направлениях (что и сделало показатель генетических различий популяций Европы наименьшим среди таких показателей для всех регионов мира [Балановская, 1998]), неудивительно, что по географической близости или удаленности регионов друг от друга можно хорошо предсказывать, соответственно, обилие или немногочисленность миграций между ними и, соответственно, близость или удаленность их генофондов. На самом деле картина интереснее простого отражения географии в генофонде: хотя взаимное положение регионов следует географии, но вот генетические расстояния между регионами отнюдь не прямо пропорциональны географическим. Особенно хорошо это видно в удаленности кластеров Восточной и Западной Европы (географически большинство входящих в них регионов не настолько удалены друг от друга), а также в генетической удаленности Кавказа, хотя географически он близок восточноевропейским регионам.

Кроме общего правила согласованности генетического и географического положения, из рисунка 2.38 следуют и другие исключения. Наиболее яркое из них — совпадение генофонда Британских островов и Балкан: их генофонды не должны бы быть похожи. Они действительно не похожи, что хорошо видно по частотам гаплогрупп и их картам. И анализ главных компонент тоже хорошо уловил эти различия, но отразил их по третьей главной компоненте. По ней генофонды Британских островов и Балканского полуострова характеризуются противоположными, полярными значениями, но при проекции многомерного графика на плоскость первой и второй главных компонент эти два региона наложились друг на друга. Несмотря на это, рассмотрение первых двух главных компонент весьма информативно, поскольку вместе они описывают почти половину суммарной изменчивости, что редко когда достигается при анализе более 10 независимых признаков (в нашем случае 17 гаплогрупп).

График главных компонент, отражающий взаимоотношения генофондов 15 регионов Европы. Использованы данные о частотах 17 гаплогрупп Y-хромосомы. Первая главная компонента описывает 29%, а вторая — 17% общей дисперсии.

Рис. 2.38. **Структура генофонда Европы на региональном уровне.**

ЭТНИЧЕСКИЙ УРОВЕНЬ: 17 ГАПЛОГРУПП, МНОГОМЕРНОЕ ШКАЛИРОВАНИЕ.

ЭТНИЧЕСКИЙ УРОВЕНЬ.

Этнический уровень является основным — и наиболее устойчивым, и наиболее хорошо изученным – среди всех уровней иерархической структуры популяций. Это показано и экспериментально [Балановская, 1990] и интуитивно признается большинством исследователей, потому что любую изученную ими выборку называют прежде всего по названию народа, к которому она относится. (Правда, в Европе могут называть как народ, так и страну, но большинство стран Европы, за исключением разве что России и Великобритании, если и не мононациональны, то близки к этому — в отношении их коренного населения). Этнический уровень является основным и для нашего анализа структуры генофонда Европы, поэтому его мы проанализируем трижды: методом многомерного шкалирования на основе массива данных «Этнический 17» (табл. 2.4), методом главным компонент на основе того же массива данных и методов многомерного шкалирования на основе

массива данных «Этнический 40» (табл. 2.4).

ЧЕТЫРЕ ЗОНЫ НА ГРАФИКЕ.

Генетические взаимоотношения народов Европы при анализе методом многомерного шкалирования показаны на рис. 2.39. При взгляде на этот график бросается в глаза его сложность – почти все поле заполнено многочисленными популяциями и поэтому их группировка в кластеры неоднозначна. Но генетические взаимоотношения между почти всеми народами, существующими в Европе, и не должны быть простыми при их проецировании на двумерный график. Попробуем разобраться в полученной картине. Прежде всего поделим график на четыре зоны – в соответствии с географией Европы и прилегающих регионов. В левой нижней зоне окажутся генофонды Средиземноморья (в соответствии с их географическим положением на юго-западе Европы). В правой нижней зоне окажутся генофонды Передней Азии (включая и Кавказ), что соответствует их географическому положению на юго-восток от Европы. В правой верхней зоне окажутся генофонды Восточной Европы, а в левой верхней – генофонды Западной Европы, опять-таки в соответствии с географией. То есть первый вывод, который мы можем сделать и который облегчит дальнейшее рассмотрение графика – это то, что согласованность положения регионов Европы в географическом и генетическом пространстве сохраняется (как и должно быть) и после того, как каждый регион представлен уже не одной точкой (рис. 2.38), а множеством точек входящих в этот регион народов (рис. 2.39). Рассмотрим поочередно каждую из этих четырех зон.

ЗОНА 1: СРЕДИЗЕМНОМОРЬЕ.

Средиземноморье (южная часть Европы) географически представляет собой цепь из трех больших полуостровов и ряда больших и малых островов. Полуострова, перечисляя с запада на восток – Пиренейский, Аппенинский, Балканский. Далее к востоку находится полуостров Малая Азия (Анатолия), но он формально уже не относится к Европе. И на генетическом графике мы видим ту же цепь генофондов народов, населяющих эти полуострова: слева расположились популяции Пиренейского полуострова; правее (впрочем, без резкого перехода) расположился генофонд Аппенинского полуострова и соседнего с ним острова Сардиния; еще правее – генофонды многочисленных народов Балканского полуострова. Часть балканских генофондов уходит далеко за пределы зоны Средиземноморья, вклиниваясь в скопления этнических точек не только Восточной Европы (что соответствует отсутствию четкой географической границы между Балканами и Восточной Европой), но и в зону графика, занятую популяциями Западной Европы, что выглядит странно. Но в целом, хотя идеального соответствия географического и генетического положения внутри регионов и не прослеживается (чтобы его достичь, каталонцы должны были бы поменяться местами с галисийцами, испанцы – с португальцами, а сардинцы – с итальянцами), то, что популяции одного региона кластеризуются совместно, а расположение этих регионов формирует такую же цепь, как и в географическом пространстве, показывает явный параллелизм между структурированностью генофонда Европы и ее географической структурированностью.

ЗОНА 2 – ПЕРЕДНЯЯ АЗИЯ.

Этот обширный регион в нашем анализе представлен только двумя своими частями, пограничными с Европой – Анатолией и Кавказом. Генофонд Анатолии прекрасно вписывается в цепь полуостровов Средиземноморья, примыкая к генофондам народов Балкан, с которыми Анатолия соседствует и географически. Анатолия в анализе представлена не только турками, но и армянами (поскольку восточные области Анатолии вплоть до геноцида армян в начале 20 века были населены преимущественно армянами). На графике оба этих народа Анатолии кластеризуются совместно, отражая, по всей видимости, общий древний анатолийский субстрат в их генофондах или даже прямую ассимиляцию части армян турками. Кавказ в анализируемых данных представлен подробно, и четко делится на кластеры Западного Кавказа, Восточного Кавказа и Закавказья. Как и на графике регионов (рис. 2.38), из всех европейских популяций Кавказ приближен только к популяциям степей Восточной Европы. Причем этот мостик переброшен через тюркоязычные народы Западного Кавказа (карачаевцев и балкарцев). Конечно, одного лишь положения этих народов на паневропейском графике недостаточно, чтобы делать выводы об их этногенезе (генофонду Кавказа посвящена отдельная глава 7), но все же можно выдвинуть гипотезу о наличии в их генофондах не только ассимилированного тюрками автохтонного компонента Кавказа, но и компонента, пришедшем на Кавказ из степей вместе с носителями тюркских языков.

ЗОНА 3 – ЗАПАДНАЯ ЕВРОПА.

Поскольку Средиземноморье мы рассматриваем отдельно, на долю Западной Европы осталось не так много народов. На

графике присутствуют народы Британских островов, французы, немцы, датчане и шведы. (Еще несколько народов, например, голландцы и норвежцы, к сожалению, надежно не изучены по выбранной панели гаплогрупп). Популяции Британских островов, в четком соответствии с их географическим положением формируют островной кластер в самой левой («западной») части графика. Французы, географически принадлежащие частью к Средиземноморью, а частью граничащие с англичанами и немцами, на генетическом графике тоже находятся между двумя этими популяциями, но решительно приближаются к пиренейскому кластеру Средиземноморской зоны, отдаляясь от немцев. Немцы, напротив, не преподносят никаких сюрпризов, располагаясь на графике в точном соответствии с их географическим положением. К «северу» от них, опять-таки в соответствии с географией, располагаются датчане, а еще «северо-восточнее» (выше и правее на графике), как можно было бы ожидать – шведы. В целом, генетические взаимоотношения западноевропейских народов с немецкой пунктуальностью воспроизводят их географическое положение, и эту монотонную картину нарушает лишь пара неугомонных балканских народов – боснийцы и герцеговинцы – далеко отлетевшие на графике от остальных популяций Балкан в западно-европейскую зону и занявшие место, которое должно бы – по географии – принадлежать отсутствующим в нашем анализе норвежцам. (Это исключение уже упоминалось при описании зоны Средиземноморья, и к нему мы еще раз вернемся ниже).

ЗОНА 4 – ВОСТОЧНАЯ ЕВРОПА.

Среди восточноевропейских народов наиболее многочисленной группой являются народы Восточно-Европейской равнины. Они занимают и самую большую территорию, значительно превышающую площадь остальных выделенных в нашем анализе географических регионов. Но на генетическом графике все они сбились в компактный кластер, отражая генетическую общность населения огромной территории– от Польши до Урала. Правда, внутри кластера, как мы видели и для кластеров Средиземноморья, географические закономерности не прослеживаются, например, татары расположились немного «западнее» русских. В эту группу популяций Восточно-Европейской равнины вошли, прежде всего, западные и восточные славяне, а также татары. Об этой генетической общности мы еще будем подробно говорить в главе 6, посвященной балто-славянским народам. Пока же отметим единственное исключение – чехов, которые, в противовес их географической, лингвистической и исторической близости к словакам и полякам, на графике отлетели далеко на «запад» и расположились вблизи генофондов поляков и немцев. Второе исключение – северные русские, которые демонстрируют большие генетические отличия от прочих славян и близость к популяциям Восточной Балтики и Северо-Восточной Европы, но о своеобразии северных русских уже написано достаточно [Балановская, Балановский, 2007; Balanovsky et al., 2008; Балановская и др., 2011].

Следующий регион, географически расположенный к северу от Восточно-Европейской равнины — Восточная Балтика (в данном анализе соответствует обычному понятию Прибалтики и включает латышей, литовцев и эстонцев) – и на генетическом графике расположен к «северу» от популяций Восточно-Европейской равнины. Близкое положение этих популяций на генетическом графике соответствует небольшим географическим расстояниям между ними. Эти популяции Прибалтики служат мостом от популяций Восточно-Европейской равнины к популяциям Северо-Восточной Европы – опять-таки в соответствии с их географически промежуточным положением. Эти популяции северо-востока Европы (карелы, финны, саами и коми) не преподносят никаких сюрпризов, кроме разве что какой-то слишком уж выраженной близости своих генофондов, не соответствующей ни большим географическим расстояниям между ними, ни хорошо известному для них действию дрейфа генов, который должен был бы подальше разнести их генофонды друг относительно друга. Впрочем, при анализе Y-хромосомы слишком много зависит от уровня филогенетического разрешения отдельных гаплогрупп: поскольку во всех этих генофондах преобладает гаплогруппа N1c-M178, а в этом анализе она еще не подразделена на недавно обнаруженные субтипы [Karmin et al., 2015], то большим межпопуляционным различиям проявиться труднее. Следующий регион – Правобережье Волги – объединяет мордву-мокшу, мордву-эрзю и чувашей. Конечно, географически их ареалы расположены на Восточно-Европейской равнине, но их особое положение на графике вкупе с тесным географическим соседством друг с другом побудило меня рассматривать их как отдельный регион. Они и географически, и генетически располагаются между другими популяциями Восточно-Европейской равнины и популяциями степей Европы. А среди их соотношений друг с другом стоит отметить разве что большие генетические различия между мокшей и эрзей.

Последний из восточно-европейских регионов – степи Европы – выделен не столько по географическому, сколько по «историческому» признаку. Хотя степная зона Европы и представляет собой единую полосу от Урала до Среднедунайской низменности, но тогда в нее надо было бы включать и украинцев (которые обозначены мной как представители народов Восточно-Европейской равнины отчасти из-за их большого генетического сходства с другими славянскими популяциями этой равнины). Все же стоит обратить внимание, что большинство включенных в анализ степных народов Европы (караногайцы, ногайцы, крымские татары, венгры) формируют единую, хотя и вытянутую, зону на графике (только башкиры занимают совсем другое место на графике). Это может служить основанием для гипотезы о наличии общего «степного» пласта в их генофондах, которую, как и любую гипотезу, надо далее проверять.

Итак, планомерно пройдя по всему графику, мы каждый раз выявляли хорошее согласие между, с одной стороны, географическим положением регионов и народов друг относительно друга и, с другой стороны, их генетическими соотношениями. Из этого правила — «генетика следует за географией» — есть, конечно, и несколько исключений. Наиболее яркое из них – уже дважды упоминавшееся положение боснийцев и герцеговинцев вблизи западноевропейских популяций на

графике. Но судя по самой матрице генетических расстояний, рассчитанных по частотам гаплогрупп, пять народов, наиболее генетически сходных с боснийцами, это герцеговинцы, румыны, сербы, хорваты, македонцы; а пять народов, наиболее сходных с герцеговинцами, это боснийцы, румыны, хорваты, сербы, македонцы. То есть и боснийцам, и герцеговинцам прежде всего генетически близки другие народы Балкан, а вовсе не Западной Европы. Потому и на графике боснийцы и герцеговинцы должны бы занять место в балканском кластере, но при визуализации матрицы в виде двумерного графика были помещены далеко в стороне. В том-то и состоит основной недостаток метода многомерного шкалирования, что, не теряя информации (используются генетические расстояния, в расчете которых принимали участие все гаплогруппы), график шкалирования теряет в точности. Ведь многомерное пространство (сколько точек на графике, столько измерений) никогда не проецируется на двумерную плоскость идеально, и геометрические расстояния между точками воспроизводят генетические расстояния между популяциями лишь с наибольшей возможной степенью. Показатели стресса и аллиенации как раз и измеряют степень искажения, внесенную при визуализации на двумерном графике. В нашем случае они высоки (0.19 и 0.20), поэтому скорее удивительно, что популяций с искажением их положения так немного (явные случаи такого искажения только боснийцы и герцеговинцы), чем то, что популяции с искажениями вообще есть.

Поэтому попробуем проанализировать тот же самый массив данных другим методом – главных компонент, который, в отличие от многомерного шкалирования, не теряет в точности, хотя и теряет значительную часть информации: первые две компоненты описывают лишь часть общей изменчивости использованных для анализа гаплогрупп.

Использованы данные о частотах 17 гаплогрупп Y-хромосомы у 63 народов Европы (массив данных «Этнический 17», табл. 2.4. и 2.5.). Для каждого народа цвет значков обозначает регион, а форма значка — лингвистическую группу, цвет надписей народов и регионов соответствует цвету значка, цвета и формы стандартизованы для всех рисунков серии (рис. 2.39. – 2.41.). Приведен график многомерного шкалирования, основанный на матрице генетических расстояний Нея. Уровень стресса составляет 0.19, алиенации – 0.20.

Рис. 2.39. **Структура генофонда Европы на этническом уровне: многомерное шкалирование по данным о 17 гаплогруппах.**

ЭТНИЧЕСКИЙ УРОВЕНЬ: 17 ГАПЛОГРУПП, ГЛАВНЫЕ КОМПОНЕНТЫ.

Генетические взаимоотношения народов Европы, выявленные методом главных компонент, показаны на рис. 2.40. По идее, этот график должен бы совпасть с предыдущим (рис. 2.39), потому что исходные данные для обоих видов анализа идентичные (массив «Этнический 17», способ визуализации тоже идентичен (расположение точек-популяций на двумерном графике) и отличается только сам статистический метод визуализации – многомерное шкалирование на рис. 2.39 и главные компоненты на 2.40.

В первом приближении графики действительно похожи, поскольку на рис. 2.40. выделяются те же самые четыре зоны, соответствующие четырем географическим частям Европы и ближайших к ней частей Передней Азии: переднеазиатская зона в правой нижней части графика (юго-восток изучаемого региона), средиземноморская зона в левой нижней части графика («юго-запад»), западноевропейская зона в левой верхней части графика («северо-запад») и восточноевропейская зона в правой верхней части графика («северо-восток»). Но положение регионов и тем более народов в пределах этих зон более хаотично, чем на рис. 2.39. Рассмотрим зоны поочередно.

ЗОНА 1 – СРЕДИЗЕМНОМОРЬЕ.

Как и на предыдущем графике, можно выделить цепь региональных кластеров, соответствующую полуостровам Средиземноморья. Но народы Пиренейского и Аппенинского полуостровов перемешаны друг с другом, хотя их совместное облако почти не перекрывается с облаком народов Балканского полуострова. Также совершенно не наблюдается близости генофондов Балканского полуострова и Анатолии (рис. 2.40), которая столь явно проявляется и на географической карте, и при анализе тех же данных методом многомерного шкалирования (рис. 2.39).

ЗОНА 2 – ПЕРЕДНЯЯ АЗИЯ.

Народы Анатолии (турки и армяне), как и на предыдущем графике, в целом близки к генофондам Северного Кавказа и Закавказья, хотя парадоксально, что армяне занимают на графике обособленное положение, а турки попадают в само облако кавказских популяций. В пределах Кавказа географическая структурированность почти не проявляется: народы Восточного Кавказа, Западного Кавказа и Закавказья не только не формируют отделенных друг от друга кластеров, но даже не прослеживаются закономерности в их распределении внутри кавказского облака. Зато кардинальные отличия генофонда Северного Кавказа от остальных популяций, входящих в географическое понятие Европы, и на этом графике видны очень четко (рис. 2.40).

ЗОНА 3 – ЗАПАДНАЯ ЕВРОПА.

В этой зоне, в отличие от Кавказа, закономерности могут быть прослежены, хотя они и выражены очень слабо: французы снова приближены к генофондам Средиземноморья, а оба изученных народа Британских островов (англичане и ирландцы) расположены на графике вблизи друг друга и если их выделить в отдельный кластер, то он будет непосредственно примыкать, но все же не перекрываться с народами Западной Балтики – шведами, датчанами и немцами. Правда, британский кластер тесно граничит и с народами Балкан (что не имеет ни географического, ни исторического смысла), но, как и при анализе на уровне регионов, Балканы и Британские острова, почти совпадая по значениям первых двух главных компонент, принимают противоположные значения по третьей главной компоненте. То есть различия их генофондов несомненны, и выявляются они в том числе и методом главных компонент, пусть и не на основном графике.

ЗОНА 4 – ВОСТОЧНАЯ ЕВРОПА.

Как и на предыдущем графике, генофонды Восточно-Европейской равнины, в соответствии с географией, граничат и с

облаком западно-европейских точек, и с балканскими генофондами. Как и на предыдущем графике, народы этого региона, несмотря на обширность Восточно-европейской равнины, демонстрируют выраженное сходство друг с другом. Характерно, что только чехи опять смещены на «запад», демонстрируя такое генетическое сходство с западноевропейскими народами, которое не проявляют другие славянские и неславянские народы Восточно-Европейской равнины. Опять-таки, повторяя предыдущий график и географическую карту, на «северо-восток» от Восточно-Европейской равнины располагаются генофонды Восточной Балтики (Прибалтики), а еще далее – генофонды народов Северо-Восточной Европы (финны, саамы, карелы, коми). И северные русские опять приближены к этим генофондам и к балтам, тогда как южные и центральные русские располагаются в середине популяций Восточно-Европейской равнины, большинство которых составляют славянские народы. Завершая аналогию с предыдущим графиком, отметим, что народы Правобережья Волги снова расположены между остальными популяциями Восточно-Европейской равнины и популяциями степей Евразии, и в пределах этого региона мокша и эрзя снова генетически далеки друг от друга. Но вся эта согласованность с географией выражена куда слабее, чем на графике многомерного шкалирования (рис. 2.39) – четких кластеров нет и говорить можно лишь о тенденциях на графике, и в пределах каждого региона положение большинства народов хаотично. Пожалуй, единственное, в чем график главных компонент лучше отражает географическое положение (и, можно полагать, региональную структуру генофонда), чем график многомерного шкалирования, это положение башкир, которые на рис. 2.40 присоединились к другим популяциям степей Европы. Эти степные популяции вновь являются единственным мостиком, приближающим европейские генофонды к кавказским, и их приближенность друг к другу на графике снова позволяет задуматься о наличии у них общего «степного» компонента в генофонде. Хотя выражена эта тенденция куда слабее, чем на графике многомерного шкалирования – крымские татары отчего-то попали в середину кавказского облака, а венгры (что, напротив, вполне разумно) присоединились к балканским популяциям, удалившись от прочих степняков.

Массив данных и обозначения идентичны рис. 2.39: использованы массив данных «Этнический 17» (табл. 2.4 и 2.5), цвет значков и надписей обозначает регион, а форма значка — лингвистическую группу, как и для всех рисунков серии (рис. 2.39. – 2.41.). Приведен график главных компонент. Первая главная компонента описывает 20 %, а вторая — 13 % общей дисперсии.

Рис. 2.40. **Структура генофонда Европы на этническом уровне: анализ главных компонент по данным о 17 гаплогруппах.**

СРАВНЕНИЕ РЕЗУЛЬТАТОВ МНОГОМЕРНОГО ШКАЛИРОВАНИЯ И ГЛАВНЫХ КОМПОНЕНТ.

Итак, два метода анализа одних и тех же данных (рис. 2.39 и 2.40) дали похожие, но отнюдь не идентичные результаты. Поэтому те черты, которые проявились на обоих графиках, можно счесть наиболее устойчивыми и объективными. Перечислим их:

— положение народов на генетическом графике соответствует их географическому положению на географической карте (с точностью до положения региона Европы эта закономерность прослеживается очень хорошо, с точностью до положения отдельных народов прослеживается не всегда – это часто видно на графике шкалирования и реже на графике главных компонент);

— генофонд Северного Кавказа отличается от генофонда Европы и примыкает к генофондам Передней Азии;

— генофонд чехов, в отличие от других славянских народов, приближен к западноевропейским генофондам;

— генофонд северных русских является вторым исключением из общей тенденции большого сходства всех западно-славянских и восточно-славянских генофондов, потому что приближен к финно-угорским народам северо-востока Европы (и отчасти к балтским народам);

— изо всех популяций Европы только народы степной зоны приближены к генофонду Кавказа и, несмотря на большую протяженность степей (от Урала до Венгрии) демонстрируют некоторое генетическое сходство друг с другом;

— остальные народы Восточно-Европейской равнины, несмотря на обширность этого региона, генетически близки друг другу.

Основное отличие между двумя графиками — то, что на графике многомерного шкалирования (рис. 2.39) географические закономерности видны очень четко, а на графике главных компонент (рис. 2.40) они смазаны. Конечно, когда два разных метода показывают разные паттерны, в общем случае неизвестно, который из них более правильный. Но в данном случае предпочтение можно обоснованно отдать паттерну многомерного шкалирования. Ведь, во-первых, если бы график многомерного шкалирования отличался бы лишь в силу технических причин, то популяции никак не могли бы сами «догадаться» о своем географическом положении – ведь в анализ закладывались лишь частоты гаплогрупп, а география в исходных данных не присутствовала. Значит, высокая степень структурированности в генофонде присутствует объективно, но лучше выявляется методом многомерного шкалирования, чем главных компонент. Во-вторых, к тому же выводу приводит и рассмотрение отдельных регионов, в частности, Кавказа. Четкое деление его генофонда в соответствии с географическим делением на Западный и Восточный Кавказ несомненно (глава 7, а также [Balanovsky et al., 2011; Yunusbaev et al., 2011]) и хорошо выявляется при многомерном шкалировании (рис. 2.39), но не проявляется на графике главных компонент (рис. 2.40). Вообще, хотя теоретически эти два метода равноценны, на практике, по моему опыту, многомерное шкалирование лучше выявляет реально существующую структуру генофонда (не обязательно эта структура географическая, но все равно шкалирование ее обычно выявляет лучше). Потому-то в работах нашего коллектива куда чаще используются графики многомерного шкалирования, чем главных компонент.

ЭТНИЧЕСКИЙ УРОВЕНЬ: 40 ГАПЛОГРУПП, МНОГОМЕРНОЕ ШКАЛИРОВАНИЕ.

В этом виде анализа использовано большое число гаплогрупп, хотя при этом и пришлось исключить из анализа народы, по которым не было столь подробных данных. Основное отличие массивов данных «Этнический 17» и «Этнический 40» — то, что основные западно-европейские гаплогруппы **R1b, R1a** и **N1c** подразделены на субтипы. (Снова и снова повторим, что, хотя этих субтипов сейчас известно очень много, для изучения структуры генофонда и истории популяций полезны лишь те их них, для которых уже проведен скрининг их распространенности в популяциях и определены частоты их встречаемости хотя бы в двух-трех десятках популяций Европы.) Поэтому график шкалирования по 40 гаплогруппам потенциально может точнее выявить особенности генетических взаимоотношений народов Европы.

Рассматривая этот график (рис 2.41) мы снова видим те же 4 зоны (Средиземноморье, Передняя Азия, Западная Европа,

Восточная Европа), а в пределах зон обнаруживаем те же, географические в своей основе, соотношения между генофондами регионов Европы. Чтобы не повторять в третий раз описание тех же закономерностей, обратим внимание лишь на две главные черты.

Первая особенность этого графика – то, что генофонд Балканского кластера переехал из средиземноморской зоны в восточноевропейскую. Это означает, что после разделения основных гаплогрупп на субтипы (что приближает к реальной структуре линий Y-хромосомы) генофонд Балкан проявляет большее сходство с генофондом Восточной Европы, чем Средиземноморья или Западной Европы. Исключением стали лишь греки, генетически приблизившиеся к своим соседям по Средиземноморью – народам Аппенинского полуострова и Анатолии.

Вторая черта – то, что подразделение в каком-либо регионе его мажорной гаплогруппы на субтипы не приводит автоматически к резкому возрастанию генетических отличий народов этого региона друг от друга. Так, несмотря на подразделение гаплогруппы **R1b**, англичане, датчане, ирландцы и шведы остаются на графике вблизи друг от друга, хотя и более четко отделяются от народов Средиземноморья (испанцев, португальцев, итальянцев) и опять примкнувших к этому кругу французов. Точно так же, несмотря на подразделение гаплогруппы **R1a**, народы Восточно-Европейской равнины остаются довольно компактным кластером, а подразделение гаплогруппы **N1c** заметно не увеличивает различия между карелами и коми. Можно думать, что такая стабильность положения народов в генетическом пространстве отчасти задается прочими гаплогруппами (кроме мажорной), данные по которым не меняются, а отчасти определяется тем, что даже при подразделении гаплогруппы на субтипы, ареалы субтипов значительно перекрываются если не по зонам максимумов, то по зонам их умеренных частот.

В целом, увеличение числа гаплогрупп с 17 до 40, при неизбежном уменьшении числа рассматриваемых народов, не увеличило информативность графика. Поэтому, если из серии рассмотренных графиков (рис. 2.39 – 2.41) выбирать какой-то один, я бы однозначно остановился на многомерном шкалировании по 17 гаплогруппам (рис. 2.39). Впрочем, вдумчивым читателям, которые будут использовать материалы этой главы для собственных исследований или собственных выводов, не стоит ограничиваться каким-то одним графиком – исследования будут куда надежнее, а выводы обоснованнее при учете и различных видов статистического анализа, и карт распространения гаплогрупп, и роли лингвистического фактора, о котором скоро пойдет речь. Не стоит забывать и о том, что каждый народ, представленный на графике одной точкой, в действительности состоит из многих локальных популяций, которые могут значительно отличаться друг от друга.

Использованы данные о частотах 40 гаплогрупп Y-хромосомы у 40 народов Европы (массив данных «Этнический 40», табл. 2.4 и 2.5). Как и на других рисунках серии (рис. 2.39 – 2.41) для каждого народа цвет значка и надписей обозначает регион, а форма значка — лингвистическую группу. Приведен график многомерного шкалирования, основанный на матрице генетических расстояний Нея. Уровень стресса составляет 0.16, алиенации – 0.17.

Рис. 2.41. **Структура генофонда Европы на этническом уровне: многомерное шкалирование по данным о 40 гаплогруппах.**

СУБЭТНИЧЕСКИЙ УРОВЕНЬ: ЭТНИЧЕСКИЕ ОБЛАКА.

ИДЕЯ АНАЛИЗА

Поэтому рассмотрим теперь изменчивость генофонда Европы на внутриэтническом уровне. В этот вид анализа включены те европейские народы, для которых имелись данные хотя бы по двум популяциям. Таких народов набралось двадцать, число изученных популяций для разных народов варьирует от 2 до 20 (особенно подробно изучены восточнославянские народы), а суммарное число популяций — 120 (табл. 2.4 и 2.5.). Этот анализ, хотя и ограничен по числу народов, показывает структуру генофонда Европы настолько детально, насколько позволяет внушительный современный объем данных, накопленный за 15

лет интенсивных исследований. Ведь в этом анализе мы имеем дело уже не с усредненными этническими частотами, а со всем многообразием конкретных популяций, населяющих различные уголки Европы.

При рассмотрении этих графиков мы не будем уже специально обращать внимание на взаимоотношения народов друг с другом – об этом выше сказано уже достаточно, да 20 народов и маловато для детальных выводов. Вместо этого мы сосредоточимся на вопросе о том, насколько сходны популяции одного народа друг с другом и насколько они перекрываются с популяциями других народов. То есть речь пойдет о том, насколько важен этнический уровень в структурировании генофонда Европы, то есть насколько принадлежность популяций к одному народу задает сходство их генофондов и, как следствие, насколько обоснован был использованный выше прием усреднения данных по локальным популяциям в среднеэтнические характеристики.

Ранее этот вопрос об уровне внутриэтнической изменчивости в Европе по Y-хромосоме специально не рассматривался (за исключением моей собственной работы [Balanovsky et al., 2008]). Даже исследования, в которых какой-либо народ Европы изучен не в одной, а в нескольких региональных популяциях, немногочисленны. Среди таких исследований следует особо выделить работу [Kayser et al., 2005], в которой изучены популяции двух народов — немцев и поляков. Правда, авторы не анализируют и тем более не сравнивают степень внутренней гетерогенности этих народов, сконцентрировавшись на выявлении генетической границы между ними. Ряд популяций поляков (и их субэтнической группы кашубов) изучен и в статье [Rebala et al., 2012]. В работе [Cinnioğlu et al., 2004] изучены паттерны географической изменчивости гаплогрупп Y-хромосомы в популяциях Турции. В работе Б.А. Малярчука [Malyarchuk et al., 2004] исследована изменчивость в пределах русских популяций (главным образом центральных и южных). В исследованиях [Kasperaviciute et al., 2004; Karlsson et al., 2006; Luca et al., 2007; Lappalainen et al., 2008; Mirabal et al., 2009; Лобов, 2009; Larmuseau et al., 2010; Karachanak et al., 2013; Kushnarevich et al., 2013] изучены популяции, соответственно, литовцев, шведов, чехов, финнов, коми, башкир, бельгийцев, болгар, белорусов. В наших собственных исследованиях были очень подробно изучены русские, украинские и белорусские популяции [Balanovsky et al., 2008; Kushnarevich et al., 2015], а также популяции мордвы [наши неопубликованные данные]. В данном анализе, кроме перечисленных источников, использовались и данные разных авторов по финнам, грекам, албанцам, сербам, хорватам, боснийцам, македонцам (для этих народов, хотя и не проводилось специальных исследований ряда популяций, но данные по разным популяциям одного и того же народа появлялись в разных статьях, и попали в нашу базу данных Y-base).

Анализ генофонда Европы на субэтническом уровне я проводил дважды – в 2007 году (опубликовано в [Balanovsky et al., 2008]), и в 2015 году (впервые публикуется в этой книге). Два этих анализа методически однотипны, но различаются объемом данных – за 8 лет значительно увеличилось и число народов, по которым есть данные о нескольких субпопуляциях (9 народов в 2007 году и 20 народов в 2015 году) и число гаплогрупп, по которым есть данные обо всех субпопуляциях всех этих народов (7 гаплогрупп в 2007 году и 15 гаплогрупп в 2015 году). Тем интереснее сравнить полученные закономерности.

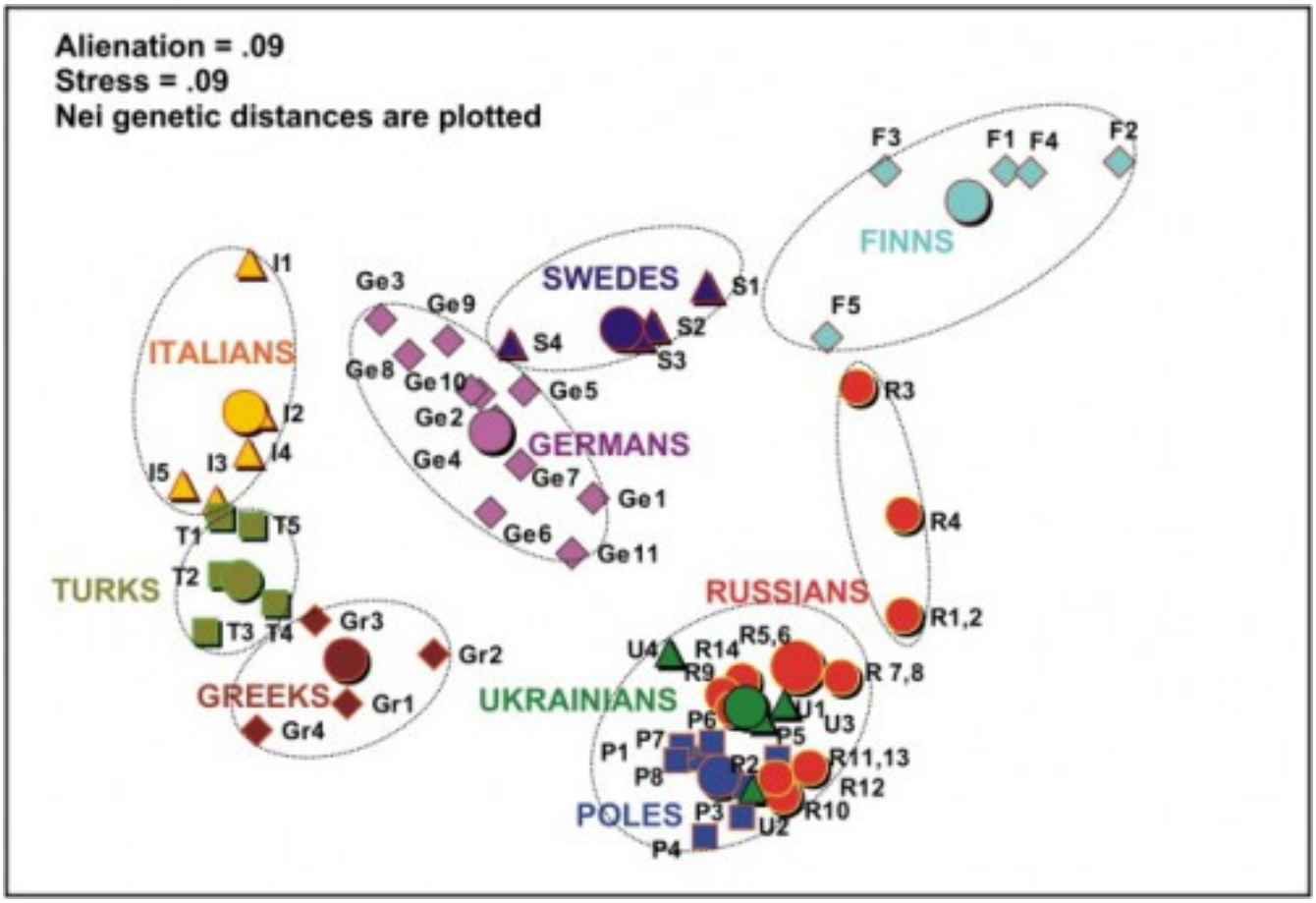
АНАЛИЗ 2007 ГОДА (9 НАРОДОВ, 7 ГАПЛОГРУПП).

Результаты анализа 2007 года представлены на рис. 2.42. Этот график демонстрирует удивительно четкую картину: почти во всех случаях популяции, относящиеся к одному народу, кластеризуются совместно. Причем образующиеся «этнические облака» (скопления точек, соответствующих популяциям одного народа) почти не перекрываются друг с другом. Этот результат имеет принципиальное значение, поскольку он указывает на четкие различия между генофондами народов Европы по маркерам Y-хромосомы.

Два своеобразных исключения касаются славянских групп. Во-первых, северные русские популяции не присоединяются к южным и центральным, а образуют отдельный кластер, находящийся между популяциями шведов, финнов и немцев. Во-вторых, остальные популяции западных и восточных славян — поляков, южных русских и украинцев — формируют не три отдельных, а единое облако, демонстрируя общность своих генофондов. Здесь опять проявляется ведущая роль лингвистического, а не географического фактора: славянские популяции, занимающие половину территории Европы, оказываются не различными (как должно бы быть в случае доминирования географического фактора), а генетически сходными, в согласии с лингвистическими данными о близком родстве восточнославянских и западнославянских языков. Хотя эти данные сами по себе еще не отвергают вывод Россера с соавторами [Rosser et al., 2000] о географической (а не лингвистической) структурированности генофонда Европы, эти результаты указывают, что в отдельных регионах Европы географически различные, но лингвистически сходные популяции оказываются сходными и генетически, указывая на важность и лингвистического фактора в формировании структуры генофонда.

Из полученных результатов можно сделать два вывода. Первый вывод методологический: Y-хромосома является надежной генетической системой с высокой разрешающей способностью. В противном случае генетические различия между этносами не были бы обнаружены или же разные популяции одного народа вели бы себя по-разному, занимая случайные места на графике.

Второй вывод – феноменологический: в структуре генофонда проявляется не только географическая, но и лингвистическая структурированность. (В данном случае этнический фактор совпадает с лингвистическим, поскольку этническая принадлежность популяции определяется в основном по языку). К анализу роли географии и лингвистики в структурировании славянских генофондов мы еще вернемся в главе 6.



Приведен график многомерного шкалирования. Популяции, относящиеся к одному народу, показаны однотипными значками одного цвета.

Рис. 2.42. **Структура генофонда Европы на субэтническом уровне (анализ 2007 года, воспроизводится по [Balanovsky et al., 2008].**

АНАЛИЗ 2015 ГОДА (20 НАРОДОВ, 15 ГАПЛОГРУПП).

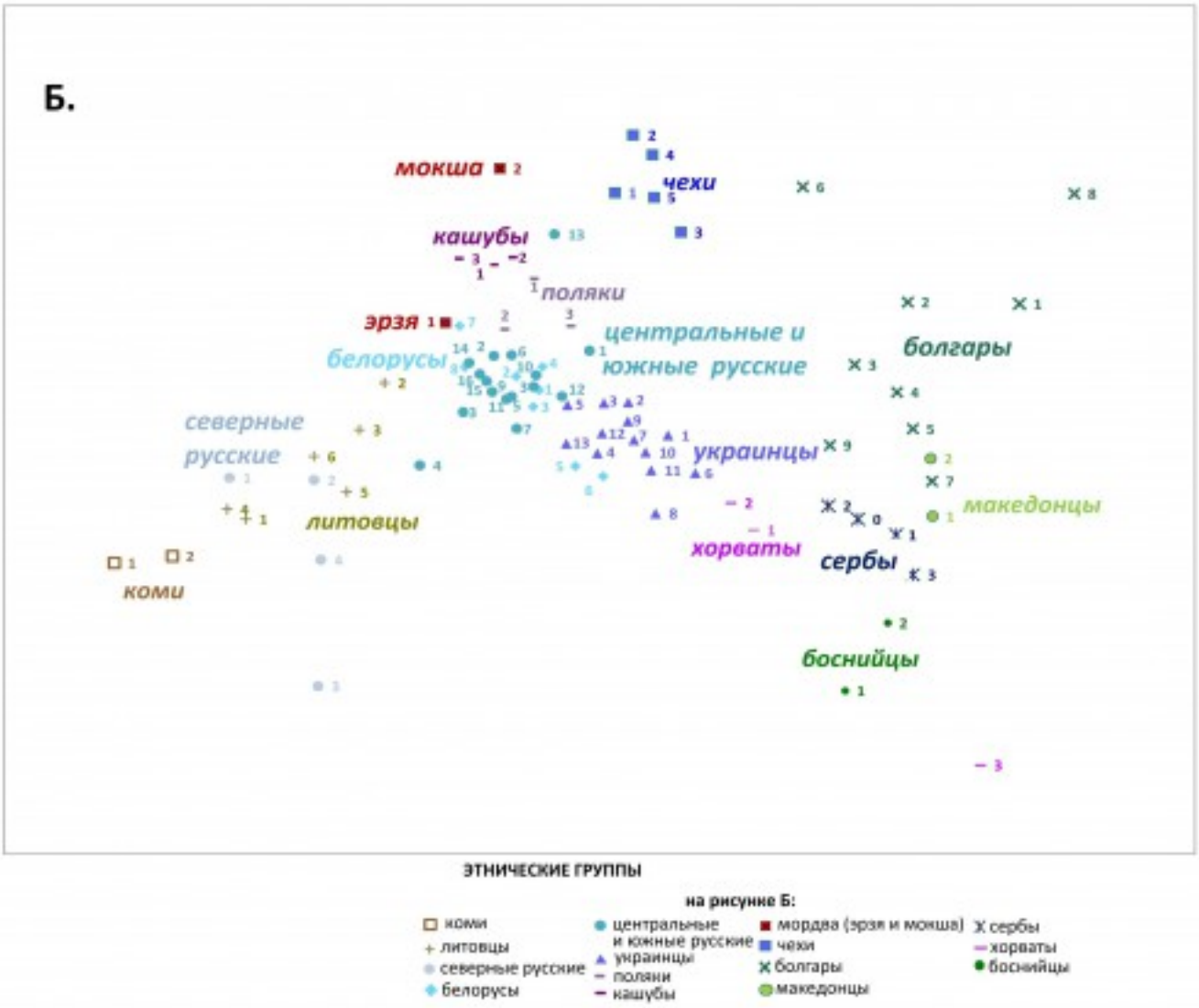
В анализ 2015 года удалось включить в два раза больше и гаплогрупп, и народов, и их субпопуляций. Генетические взаимоотношения этих ста двадцати европейских популяций показаны на рис. 2.43. Популяций так много, что даже современная версия программы Statistica не выводит их на график одновременно; поэтому на рис. 2.43 А выведены популяции всех народов, но восточнославянские народы представлены лишь некоторыми из изученных популяций, а на рис. 2.43 Б выведены лишь народы центрального кластера, проявившегося на рис. 2.43. А, но восточнославянские популяции выведены уже все.

Несмотря на новые данные, закономерность осталась старой – наличие четко выраженных этнических облаков. Действительно, разные региональные популяции одного народа группируются совместно, и эти этнические облака или отделены от этнических облаков других народов или граничат с ними, но за редким исключением не перекрываются. Рассматривая сначала график 2.43 А, видим отграниченные друг от друга этнические облака финнов, шведов, бельгийцев и граничащие друг с другом облака греков, турок и албанцев. При этом облака албанцев и турок четкие и довольно компактные, а греческие популяции генетически очень разнообразны – греки Малой Азии и Крита близки к туркам (что неудивительно, поскольку турки Малой Азии должны были в значительной степени ассимилировать греческое население, особенно на побережье).

Переходя к популяциям остальных народов (рис 2.43 Б) снова видим ту же закономерность: для коми, литовцев, поляков (и кашубов), чехов, украинцев, боснийцев, хорватов и сербов характерны четкие этнические облака: многочисленные популяции каждого из этих народов сбиваются друг с другом в кучки. С остальными народами – мордвой, русскими, белорусами, македонцами и болгарами картина чуть сложнее. Впрочем, с болгарами и македонцами особенность только в том, что облако болгар оказывается довольно рыхлым, показывая большие межпопуляционные различия, но с другими облаками оно почти не перекрывается – лишь немного с облаком македонцев. Особенность мордвы в том, что популяции мокши и эрзи не кластеризуются друг с другом, хотя размах генетических различий между ними на этом графике оказывается не так уж и велик. Особенность русских, проявлявшаяся и на предыдущем графике (рис. 2.42) — это генетическая обособленность северных популяций, которые на этом графике (рис. 2.43). частично перекрываются с этническим облаком литовцев. С белорусами еще интереснее – они четко делятся на южных (Полесье) и северных (вся остальная часть Белоруссии, то есть центр и север); южные белорусы не формируют своего облака, а попадают в облако украинцев, аналогично северные белорусы попадают в самую середину облака русских популяций. Таким образом, увеличив число гаплогрупп и число изученных локальных популяций, удалось увидеть внутреннюю структуру поляко-русско-украинской генетической общности, которая была гомогенной при анализе 2007 года, основанном на меньшем объеме данных: теперь популяции каждого народа занимают свою собственную зону в пределах этой общности. Учитывая, что украинское и русское облака, хотя не перекрываются друг с другом, но тесно соприкасаются, генетические взаимоотношения популяций восточных славян на графике многомерного шкалирования можно описать как генетическую общность, состоящую из русской и украинской половин, причем белорусские популяции распределены почти поровну между обеими половинами.

Использованы данные о частотах 15 гаплогрупп Y-хромосомы (массив данных «Субэтнический», табл. 2.4. и 2.5.). В связи с большим генетическим своеобразием башкир их популяции не выведены на график, соответственно, проанализированы 115 популяций 19 народов. Цвет (а на рис. Б – и форма) значка специфичны для каждого народа. Эти цвета не соответствуют цветам на предыдущих графиках (рис. 2.39 – 2.41) поскольку на них цвет означал регион Европы, а здесь цвет обозначает народ. А. График многомерного шкалирования, основанный на матрице генетических расстояний Нея и включающий все популяции, кроме некоторых восточнославянских. Уровень стресса составляет 0.12, алиенации – тоже 0.12. Б. График многомерного шкалирования, основанный на матрице генетических расстояний Нея и включающий все популяции всех народов, показанных выколотыми кружками на рис. А. На этот график выведены все изученные восточнославянские

популяции. Уровень стресса составляет 0.10, алиенации – тоже 0.10.



Использованы данные о частотах 15 гаплогрупп Y-хромосомы (массив данных «Субэтнический», табл. 2.4. и 2.5.). В связи с большим генетическим своеобразием башкир их популяции не выведены на график, соответственно, проанализированы 115 популяций 19 народов. Цвет (а на рис. Б – и форма) значка специфичны для каждого народа. Эти цвета не соответствуют цветам на предыдущих графиках (рис. 2.39 – 2.41) поскольку на них цвет означал регион Европы, а здесь цвет обозначает народ. А. График многомерного шкалирования, основанный на матрице генетических расстояний Нея и включающий все популяции, кроме некоторых восточнославянских. Уровень стресса составляет 0.12, алиенации – тоже 0.12. Б. График многомерного шкалирования, основанный на матрице генетических расстояний Нея и включающий все популяции всех народов, показанных выколотыми кружками на рис. А. На этот график выведены все изученные восточнославянские популяции. Уровень стресса составляет 0.10, алиенации – тоже 0.10.

Рис. 2.43 (А, Б). **Структура генофонда Европы на субэтническом уровне.**

Конкретные популяции в пределах каждого народа обозначены цифрами.

На рисунке А:

Финны : 1 — Хяме, 2 — Северной Остроботнии, 3 — Северной Савонии, 4 — Сатакунты, 5 — Южной Карелии, 6 — Южной Остроботнии, 7 — Юго-Западной Финляндии.

Шведы: 1 — Блекинге и Кристианстада, 2 — Готланда, 3 — Эстергётланда и Йёнчёпинга, 4 — Скараборга, 5 — Уппсалы, 6

— Вэрмланда, 7 — Вестерботтена.

Бельгийцы: 1 — Антверпена, 2 — Кампина, 3 — Мехелена, 4 — Северного Брабанта, 5 — Фландрии.

Албанцы: 1 — Албании, 2 — Македонии, 3 — Косово.

Греки: 1 — Малой Азии, 2 — Фессалии, 3 — Крита, 4 — Македонии, 5 — Пелопоннеса.

Турки: 1 — центральные, 2 — восточные, 3 — Стамбула, 4 — северо-восточные, 5 — северо-западные, 6 — юго-восточные.

На рисунке Б:

Коми: 1- ижемские, 2 — прилужские.

Мордва: 1 — эрзя, 2 — мокша.

Литовцы: 1 – аукшайты восточные, 2 – аукшайты южные, 3 – аукшайты западные, 4 — жемяйты северные, 5 – жемяйты южные, 6 — жемяйты западные.

Белорусы: 1 – юго-западные (Брестская обл.), 2 — северные (Витебская обл.), 3 — центральные, 4 — восточные, 5 — юго-восточные, 6 — юго-западные, 7 — западные, 8 — северные.

Северные русские: 1 — Архангельская обл., Лешуковский р-н, 2 — Архангельская обл., Пинежский р-н, 3 — Архангельская обл., Красноборский и Ленский р-ны, 4 — Вологодская обл.

Центральные и южные русские: 1 — Костромская обл., Мантуровский и Межевской р-ны, 2 — Тверская обл., Кашинский р-н, 3 — Псковская обл., Порховский и Дедовичский р-ны, 4 — Псковская обл., Островский р-н, 5 — Смоленская обл., Рославльский и Ершичский р-ны, 6 — Орловская обл., Ливненский р-н, 7 — Курская обл., Пристенский р-н, 8 — Воронежская обл., Репьевский р-н, 9 — Белгородская обл., 10 — донские казаки (нижний Дон), 11 — донские казаки (верхний Дон), 12 — кубанские казаки, 13 — терские казаки, 14 — Рязанская обл., Сапожковский р-н, 15 — Рязанская обл., Сараевский р-н, 16 — Рязанская обл. (из разных пунктов).

Украинцы: 1 — западные (Закарпатская обл.), 2 — западные (Ивано-Франковская обл.), 3 — западные (Львовская обл.), 4 — центральные (Черкасская обл.), 5 — западные (Хмельницкая обл.), 6 — западные (Черновецкая обл.), 7 — западные (Ровненская обл.), 8 — северные (Житомирская обл.), 9 — северные (Черниговская обл.), 10 — северные (Сумская обл.), 11 — восточные (Харьковская обл.), 12 — Белгородской обл., 13 — запорожские казаки (Запорожская обл.).

Поляки: 1 — Kociewie, 2 — курпы, 3 — из Вроцлава.

Кашубы: 1 — центральные, 2 — северные, 3 — южные.

Чехи: 1 — Брно, 2 — Йиндржихув-Градец, 3 — Клатови, 4 — Писек, 5 — Тршебич.

Болгары: 1 — Бургас, 2 — Ловеч, 3 — Монтана (Болгария), 4 — Пловдив, 5 — Софийская обл., 6 — Хасково, 7 — Варна, 8 — Разград, 9 — София.

Македонцы: 1 — из Скопье, 2 — из разных регионов Македонии.

Сербы: 1 — из Приштины (Косово), 2 — из разных регионов Сербии, 3 — из Боснии.

Хорваты: 1 — из Сплита, 2 — из разных регионов материковой Хорватии, 3 — из Боснии.

Боснийцы: 1 — центр и юг Боснии и Герцеговины (Зеница, Мостар, Широки-Бриег), 2 — из разных регионов Боснии и Герцеговины.