Структура генофонда Европы в зеркале митохондриальной ДНК

Олег Балановский

Фрагмент из книги «Генофонд Европы»

Уже около двадцати лет митохондриальная ДНК является одной из наиболее востребованных генетических систем в популяционно-генетических исследованиях. Вместе с Y-хромосомой мтДНК образовала своеобразный дуэт, и эти две нерекомбинирующие генетические системы с однородительским типом наследования стали наиболее популярными маркерами в исследовании популяций человека. Даже сейчас, после внедрения широкогеномных маркеров и полногеномного секвенирования, число выходящих статей по изменчивости мтДНК и Y-хромосомы в несколько раз превышает число полногеномных исследований, хотя и уступает им по импакт-фактору журналов. В целом можно уверенно констатировать, что появление широкогеномных маркеров добавило третью систему, но не заменило первые две.

Поскольку митохондриальная ДНК и Y-хромосома настолько параллельны друг другу – и по методам анализа, и по объему накопленных данных – в книге анализ этих двух систем проведен по возможности аналогично, чтобы было легче сопоставить результаты: и общие закономерности, выявляемые обеими системами, и частные факты, обнаруженные только по одной из них. Поэтому план изложения исследований митохондриального генофонда Европы такой же, как и для Y-хромосомного генофонда. Сначала мы кратко рассмотрим итоги предшествующих исследований. Затем проанализируем генетические взаимоотношения популяций, причем, как и для Y-хромосомы, рассмотрим закономерности в трех приближениях. (Правда, поскольку межпопуляционная вариабельность мтДНК в Европе куда меньше, ее изменчивость на внутриэтническом уровне, где она совсем мала, не рассматривается). Потом изучим роль географического и лингвистического факторов в структурировании генофонда, и рассчитаем генетическую гетерогенность основных лингвистических групп населения Европы.

Однотипность плана анализа мтДНК и Y-хромосомы придется нарушить лишь в двух частных пунктах: для Y-хромосомы мы не стали анализировать гаплотипическое разнообразие, а для мтДНК не будем рассматривать карты отдельных гаплогрупп. Это вызвано тем, что внутрипопуляционное гаплотипическое разнообразие Y-хромосомы, оцениваемое по наиболее популярной панели из 17 STR маркеров, практически во всех популяциях Европы близко к максимуму, и потому различия между популяциями невелики. Что же касается карт отдельных гаплогрупп мтДНК, то для них географические закономерности (например, более высокая частота гаплогруппы H на севере Европы) известны, но маловыразительны. Причина — уже упоминавшаяся гомогенность Европы по частотам гаплогрупп мтДНК. Поэтому геногеография гаплогрупп мтДНК в Европе специально не рассматривается (в отличие от анализа Y-хромосомы, четким геногеографическим паттернам отдельных гаплогрупп которой посвящен раздел 2.2). Если же читатель все равно интересуется распространением гаплогрупп мтДНК, он найдет эти карты в главе 8, где они построены в масштабе всей Евразии, и естественно включают Европу как составную часть.

Отметим, что в этой главе анализируется панель данных по гипервариабельному сегменту мтДНК и SNP маркерам кодирующей части, а данные полного секвенирования мтДНК не рассматриваются. Дело в том, что геногеография оперирует выборками, а не индивидуальными образцами. И если полное секвенирование нескольких десятков образцов какой-либо гаплогруппы выявляет новые субгаплогруппы, эта информация касается лишь молекулярной, а не популяционной составляющей анализа мтДНК: мы знаем, что эти субгаплогруппы существуют, но мы не знаем, с какой частотой они встречаются в каких популяциях. Поэтому и по Y-хромосоме (глава 2), и по мтДНК (данная глава) анализируются лишь те гаплогруппы, по которым уже накопились данные об их частотах в разных популяциях, а новейшие данные о гаплогруппах, выявленных при полногеномном секвенировании, рассматриваются отдельно, в одном из разделов главы 8.

3.1. Предшествовавшие исследования изменчивости мтДНК в Европе

Авторитет митохондриальной ДНК (мтДНК) в популяционной генетике человека утвердился после яркой иллюстрации африканских корней митохондриальных генофондов населения всех континентов. Это явилось чрезвычайно убедительным аргументом в пользу моноцентристской концепции происхождения человека современного вида и африканской прародины. Меньший географический масштаб, но ничуть не меньший резонанс, имело и крупнейшее исследование митохондриального генофонда Европы [Richards et al., 2000]. В популяционной генетике человека Европа всегда являлась наиболее изученным континентом, поскольку многочисленные европейские лаборатории концентрируются в первую очередь на изучении населения своих собственных стран, и лишь во вторую – населения других частей света.

НЕОЛИТ ИЛИ ПАЛЕОЛИТ?

Митохондриальная ДНК не стала исключением, и, начиная с 90-х годов XX века, стали стремительно накапливаться данные о полиморфизме мтДНК европейских популяций. Эти исследования привели к результатам, которые можно рассматривать как смену «неолитической» парадигмы в трактовке формирования генофонда Европы на «палеолитическую». В 1970-х-90-х годах

преобладала теория формирования европейского генофонда в результате миграции популяций неолитических земледельцев из Передней Азии в Европу [Ammerman, Cavalli-Sforza, 1984]. Но на рубеже XXI века большинство специалистов склонилось на сторону теории неолитизации путем заимствования неолитической культуры без миграции самих земледельцев, что предполагает значительно более древнее (палеолитическое) время формирования европейского генофонда [Richards et al., 1996, 2000]. Тем самым была почти отвергнута знаменитая теория «демической диффузии» [Ammerman, Cavalli-Sforza, 1984; Cavalli-Sforza et al., 1994], возводящая (по данным об аутосомных классических маркерах) формирование основных черт генофонда Европы ко времени расселения неолитических земледельцев, пришедших с Ближнего Востока и ассимилировавших аборигенные палеолитические популяции охотников-собирателей. Взамен была реабилитирована теория «культурной диффузии», рассматривающая неолитизацию Европы как распространение земледелия путем культурных заимствований без миграции самого населения. Правда, в последующие годы однозначность такого вывода несколько поколебалась, о чем мы будем говорить в главе 9.

Ключевой аргумент в работе [Richards et al., 2000] состоял в выявлении характерных для Европы гаплогрупп и кластеров гаплотипов мтДНК и расчете их возраста: древность большинства из них далеко превышала датировки неолита и даже мезолита, уводя в верхний палеолит. Итак, благодаря исследованию [Richards et al., 2000] не только (временно) возобладала теория палеолитического времени формирования генофонда Европы с лишь незначительным вливанием ближневосточного генетического компонента в неолите, но и изучение митохондриальной ДНК приобрело исключительный авторитет в решении проблем истории формирования генофондов.

ГОМОГЕННОСТЬ МИТОХОНДРИАЛЬНОГО ГЕНОФОНДА ЕВРОПЫ

Однако основной успех исследования Richards с соавторами [Richards et al., 2000] заключался в сравнении европейского митохондриального генофонда с генофондом внешнего по отношению к Европе региона Ближнего Востока, а вопрос о внутренней структуре европейского генофонда не ставился. Во многом это было связано с устоявшимся к тому времени взглядом на европейский митохондриальный генофонд как на исключительно гомогенный, однородный по всему европейскому ареалу: исследования показали, что в самых разных частях Европы распространен один и тот же набор гаплогрупп, причем с очень близкими частотами [Simoni et al., 2000].

Последующее накопление данных лишь отчасти внесло коррективы в такую трактовку: было показано, что при увеличении суммарной выборки (по всем изученным европейским популяциям) до 3113 человек и при объединении данных по крупным регионам Европы удается выявить некоторую степень структурированности по географическому принципу: популяции Средиземноморья, Центральной и Северной Европы входят в разные кластеры на графике генетических взаимоотношений населения этих регионов [Richards et al., 2002]. И все же вывод об гомогенности митохондриального генофонда Европы и очень слабых различиях региональных популяций остался в силе и до сих пор часто цитируется.

РЕГИОНАЛЬНЫЕ ИССЛЕДОВАНИЯ

Возможно, именно сложившейся уверенностью в бесперспективности поиска внутренней структуры в европейском митохондриальном генофонде (а также интенсивной обобщающей работой в предшествовавший период, 1995-2002 гг.) объясняется отсутствие новых попыток такого обобщающего анализа за десять с лишним лет, прошедших после работы [Richards et al., 2002].

Вместо этого большинство исследователей сосредоточились на получении данных по отдельным странам и популяциям Европы, и было опубликовано более ста подобных работ. Перечислим лишь важнейшие из них, опубликованные в период 1995-2011 годов и опирающиеся на выборки более, чем 150 образцов мтДНК [Sajantila et al., 1995; Dupuy et al., 1996; Richards et al., 1996, 2000; Opdal et al., 1998; Pfeiffer et al., 1999; Dimo-Simonin et al., 2000; Helgason et al., 2001; Larruga et al., 2001; Meinila et al., 2001; Nasidze et al., 2001; Bermisheva et al., 2002; Malyarchuk et al., 2002, 2004, 2006, 2010; Gonzalez et al., 2003; Maca-Meyer et al., 2003; Poetsch et al., 2003; Dubut et al., 2004; McEvoy et al., 2004; Pereira et al., 2004; Babalini et al., 2005; Goodacre et al., 2005; Alvarez et al., 2006; Bosch et al., 2006; Falchi et al., 2006; Picornell et al., 2006; Tonks et al., 2006; Achilli et al., 2007; Grzybowski et al., 2007; Brandstaetter et al., 2007, 2008; Hedman et al., 2007; Irwin et al., 2007, 2008; Richard et al., 2007; Tetzlaff et al., 2007; Zimmermann et al., 2007; Lappalainen et al., 2008; Lehocky et al., 2008; Martinez et al., 2008; Turchi et al., 2008; Alvarez-Iglesias et al., 2009; Santos et al., 2003; Alvarez et al., 2010; Garcia et al., 2010; Mikkelsen et al., 2010; Pereira et al., 2010; Tillmar et al., 2010; Morozova et al., 2011; Karachanak et al., 2011].

ИССЛЕДОВАНИЯ В РОССИИ

Реальность бума митохондриальных работ иллюстрирует тот факт, что только в России эти исследования были развернуты целым рядом коллективов: Института биологических проблем Севера в Магадане [Derenko et al., 2003, 2007; Malyarchuk et al., 2001, 2002, 2003, 2004, 2006, 2008, 2010, 2012], Института цитологии и генетики в Новосибирске [Гольцова и др., 2005; Derbeneva et al., 2002a,b; Сукерник и др., 2010; Starikovskaya et al., 1998, 2005; Rubinstein et al., 2008; Volodko et al., 2008], Уфимского научного центра в Уфе [Бермишева и др., 2002; 2004], Якутского научного центра в Якутске [Fedorova et al., 2003, 2013], Южного федерального университета в Ростове-на-Дону [Kornienko et al., 2004], Института общей генетики РАН

[Orekhov et al., 1999; Орехов, 2002; Наумова и др., 2008; Bulayeva et al., 2003; Morozova et al., 2011], Института молекулярной генетики РАН [Belyaeva et al., 2003] Медико-генетического научного центра РАМН в Москве [Балановский и др., 2011] и другими научными коллективами страны. Среди них особого упоминания заслуживают работы Б.А. Малярчука, который из всех российских исследователей опубликовал наибольшее число работ по изменчивости мтДНК в Европе [Malyarchuk et al., 2002, 2003, 2004, 2006, 2008, 2010, 2012], во многих из которых проводится изучение различных аспектов митохондриального генофонда как популяций Сибири, так и восточных, западных и южных славян. Огромную роль в проведении исследований российских коллективов сыграл Эстонский биоцентр и его директор Р. Виллемс. В годы, когда секвенирование в России было технически и/или финансово недоступно, молодые специалисты (включая и автора этой книги) из практически всех перечисленных институтов России, так же, как и специалисты из многих других стран Европы, толпами съезжались в Эстонский биоцентр со своими образцами, проводя там и технический этап секвенирования, и обучаясь методам анализа митохондриальных данных.

ЦЕЛЬ — АНАЛИЗ В МАСШТАБЕ ВСЕЙ ЕВРОПЫ

Но, несмотря на такое обилие сильных региональных работ, обобщающие исследования в масштабе всей Европы за последнее десятилетие мне неизвестны. Свою роль в этом сыграли и объективные трудности составления сколько-нибудь полной сводки о полиморфизме мтДНК в Европе в условиях растущей лавины информации по этой тематике, публикуемой во множестве генетических, судебно-медицинских и даже археологических изданий. В связи с этим создававшиеся общедоступные базы данных по мтДНК, аккумулирующие эту информацию, либо переключились более на молекулярную, чем на популяционную сторону изменчивости мтДНК (например, популярные интернет-ресурсы www.mitomap.org и www.phylotree.org), либо рост базы данных прекратился или замедлился после нескольких обновлений (база данных HVRbase, достигшая объема 13 873 образцов мтДНК, но затем угасшая, EMPOP с текущим объемом 34 717 образцов). Из целого ряда предпринятых попыток создания баз данных о популяционном полиморфизме мтДНК, насколько нам известно, лишь база данных, развиваемая нашим коллективом («MURKA»), выдержала целый ряд последовательных обновлений и к настоящему времени содержит сведения о 132 600 образцов мтДНК в 2 101 популяциях мира (табл. 1.3). Применительно к генофонду Европы и соседних территорий в базе данных содержатся 31 935 образцов мтДНК, что более чем в 10 раз превышает объем данных, использованный в последнем обобщающем исследовании полиморфизма мтДНК в Европе [Richards et al., 2002].

Именно последовательная работа по созданию максимально полной базы данных по изменчивости мтДНК («MURKA», раздел 1.2.) является первой предпосылкой для новой попытки обобщающего анализа европейского митохондриального генофонда, предпринятой в данной монографии. Второй предпосылкой стало наше исследование полиморфизма Y-хромосомы в Европе, тоже основанное на обширной базе данных и включившее картографический и статистический анализ этих данных (это исследование было начато в работе [Balanovsky et al., 2008] и завершено в главе 2 данной книги). Естественным продолжением этой работы стало проведение аналогичного анализа данных по митохондриальной ДНК: было решено предпринять попытку выявить основную структуру европейского генофонда по мтДНК и сравнить основные черты генофонда Европы, реконструируемые благодаря дуэту обеих информативных «однородительских» систем.

3.2. Генетические взаимоотношения популяций: три масштаба

ФОРМИРОВАНИЕ МАССИВА ДАННЫХ.

Для выявления генетических взаимоотношений популяций Европы и смежных регионов использован ставший традиционным метод многомерного шкалирования. Однако особенностью проведенного анализа является большое число тщательно отобранных популяций.

Первым критерием отбора был размер выборки. В базе данных MURKA по Европе имелись данные по 336 популяциям, но многие из них были изучены по небольшой выборке. Для обеспечения репрезентативности мы установили минимальный объем выборки N=70. При выборках меньшего объема по возможности объединялись соседние популяции одного народа так, чтобы в сумме составить выборку более 70 образцов. Если такой возможности не было, малые выборки исключались из анализа.

Вторым критерием отбора была надежность определения гаплогрупп. Мы стремились использовать только те выборки из нашей базы данных, в которых принадлежность образцов к данной гаплогруппе была подтверждена анализом SNP маркеров кодирующей части мтДНК (как правило, методом ПДРФ-анализа). Если для данного этноса или региона такие выборки отсутствовали, в анализ включались выборки, где ПДРФ данных не было, но гаплогруппы можно было надежно (с высокой вероятностью) определить по ГВС1. Пороговым уровнем было наличие не более 5% образцов, для которых надежное определение гаплогруппы оказывалось невозможно. Выборки, в которых доля таких образцов превышала 5%, не включались в анализ.

Третьим критерием было наличие четкой географической привязки в описании выборки. Поэтому, если для выборки было известно ее происхождение лишь с точностью до страны, но по этой стране имелись и выборки с региональной привязкой внутри страны, первую выборку приходилось исключать.

В итоге были сформированы 118 выборок, характеризующих все части Европы и смежных регионов, причем каждая представлена не менее чем 70 образцами. Средний объем выборок составляет N=200, суммарно 23 500 образцов. Таким образом, даже после того как две трети изученных к настоящему времени европейских популяций были отсеяны или объединены, суммарный объем анализируемых образцов оказался все-таки на порядок больше, чем в предыдущих обобщающих исследованиях европейского митохондриального генофонда [Richards et al., 2000, 2002]. Но еще более важно для получения надежных результатов то, что мы опираемся на большой объем выборки из популяции (N=200). Карта на рис. 3.1. показывает географическое положение популяций Европы, изученных по полиморфизму мтДНК.

Следующей задачей было формирование набора гаплогрупп. Из двух критериев полиморфизма, принятых в популяционной генетике (однопроцентный и пятипроцентный) мы применили «мягкий». Поэтому рассматривали только гаплогруппы, средняя частота которых в европейских популяциях составляет хотя бы один процент. Для этого из большого числа гаплогрупп часть была объединена (с учетом их филогенетического родства). Например, были объединены **V** и **preV**, также объединялись разные варианты **U2**, и т.д. Часть редких гаплогрупп, отдаленных от других гаплогрупп филогенетически или с отдаленным от Европы ареалом их преимущественного распространения, была объединена в сборную группу «другие» (**other**), причем средняя частота этой сборной парагруппы составила только 2%. В итоге в анализ были включены следующие 35 гаплогрупп: **A, B, C, D, F, G, Z, L, M1, preHV, H, HV, J, T*, T1, T2, I, N1a, N1b, R, K, V, W, X, U1a, U1b, U2, U3, U4, U5a, U5b, U6, U7, U8, other**.



Рис. 3.1. Популяции Европы и смежных регионов, изученные по мтДНК.

В результате была сформирована таблица из 118 строк-популяций и 35 столбцов-гаплогрупп, которая и послужила основой

для всех видов анализа, проведенных на уровне частот гаплогрупп мтДНК. Данные на уровне отдельных гаплотипов используются только для анализа внутрипуляционного гаплотипического разнообразия.

Для рассмотрения основных закономерностей европейского генофонда анализ был проведен на этническом уровне. Для этого данные по популяциям одного народа были усреднены и анализировались частоты для получившихся среднеэтнических популяций. Число анализируемых популяций при этом сократилось до 62 народов. Они включили следующие этнические популяции Европы и прилегающих регионов Северной Африки, Ближнего Востока и Западной Сибири (приводятся в соответствии с англоязычными обозначениями базы данных MURKA): Lithuanians, Estonians, Finns, Karelians, Komi, Mari, Mordvinians, Saami, Udmurts, Czech, Poles, Byelorussians, Russians, Slovaks, Ukrainians, Bashkirs, Chuvash, Tatars, Albanians, Greeks, Aromuns, Romanians, Slovenians, Bosnians, Bulgarians, Croatians, Hungarians, Basques, Irish, Scottish, Austrians, English, Germans, Icelanders, Norway, Swedes, Swiss, French, Italians, Portugals, Sardinians, Sicilians, Spaniards, Turkey, Armenians, Georgians, Kabardinians, Ossets, Azeris, Nogays, Kurds, Iranians, Syrians, Iraq, Jordanians, Palestinians, Saudi Arabia, Morocco, Kazakhs, Mansi, Nenets, Nganasans.

АНАЛИЗ ПОПУЛЯЦИЙ ЕВРОПЫ И БЛИЖНЕГО ВОСТОКА.

Результаты анализа митохондриальных генофондов этих 62 этнических популяций представлены на рис. 3.2. На рисунке 3.2А показаны генетические взаимоотношения этих популяций (их положение в генетическом пространстве), а на рисунке 3.2Б – локализация тех же популяций на карте (их положение в географическом пространстве).

В генетическом пространстве четко выделяются четыре кластера. Первый кластер включил только саамов, что не удивительно, учитывая их генетическое своеобразие [Cavalli-Sforza et al., 1994], в том числе показанное и по мтДНК [Tambets et al., 2004]. Второй кластер включил те популяции восточных рубежей Европы, у которых повышена частота восточно-евразийских гаплогрупп. Третий кластер включил популяции Передней Азии и Кавказа. Наконец, все остальные популяции с основной территории Европы (от Волги до Пиренейского полуострова) помещаются в четвертый «пан-европейский» кластер. Его небольшие размеры на генетическом графике (при огромных размерах на географической карте) свидетельствуют о низкой межпопуляционной изменчивости внутри этого кластера. Эти результаты подтверждают генетическую гомогенность Европы (кроме Приуралья) по мтДНК, и указывают на своеобразие генофондов Приуралья и Ближнего Востока.

Результаты, представленные на рис. 3.2А, таким образом, подтверждают на новом витке исследований и резко возросшем объеме информации два явления: гомогенности митохондриального генофонда Европы и эффективности выделения двух основных вариантов митохондриального генофонда Евразии – западно-евразийских и восточно-евразийских гаплогрупп, поскольку именно наличие восточно-евразийских гаплогрупп обуславливает своеобразие Приуралья.

Но попробуем узнать о Европе новое, — и для этого рассмотрим этот европейский кластер как бы под микроскопом, удалив из анализа внеевропейские популяции и «генетических чужаков» Европы: пограничную полосу красных кружков восточных рубежей Европы и зеленые популяции Ближнего Востока. То есть рассмотрим только одну Европу – только популяции, обозначенные фиолетовым цветом на рис. 3.2А и 3.2Б.

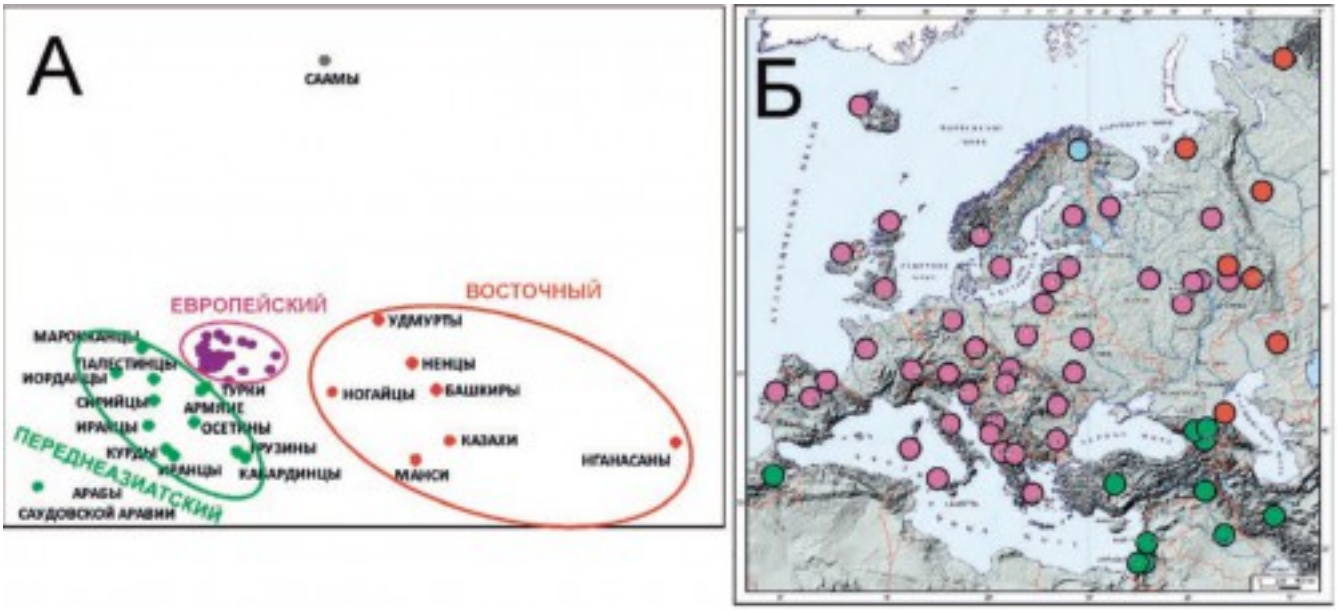


Рис. 3.2. Генетические взаимоотношения европейских и ближневосточных популяций по данным об изменчивости мтДНК.
А. График многомерного шкалирования (стресс=0,085). Популяции разных кластеров показаны разным цветом.
Б. Географическое положение изученных популяций. Цвет каждой популяции такой же, как на графике А.

АНАЛИЗ ТОЛЬКО ЕВРОПЕЙСКИХ ПОПУЛЯЦИЙ.

При анализе только популяций, вошедших в «европейский» (фиолетовый) кластер на рис. 3.2, получаем новый график и новую карту (рис. 3.3). График снова показывает положение популяций в генетическом пространстве и вновь популяции, попавшие в один кластер, обозначены одним цветом (3.3 А), а карта (3.3 Б) отображает географическое положение тех же популяций с теми же цветовыми обозначениями. Мы обнаруживаем новый генетический кластер со значительным разнообразием популяций внутри его (желтый цвет). Обратим внимание, что эти популяции (желтого цвета) занимают большую область только в генетическом пространстве (3.3А). На географической же карте (3.3Б) желтые кружки близки друг к другу, занимая не столь уж большую область на северо-востоке Европы. Таким образом, высокое генетическое разнообразие (большие генетические различия между популяциями, проявляющиеся в большом размере генетического кластера) умещается на географически небольшой территории северо-востока Европы. Рассмотрение того, какие популяции входят в этот кластер, позволяет условно назвать его «финно-угорским». Действительно, в него входят финноязычные народы (коми, мордва, мари, финны, карелы). К этому кластеру приближаются также тюркоязычные чуваши и татары, для которых по антропологическим данным предполагается значительный финно-угорский субстрат. Отметим, что другие тюркские народы Европы (например, башкиры и ногайцы) в этот кластер не входят.

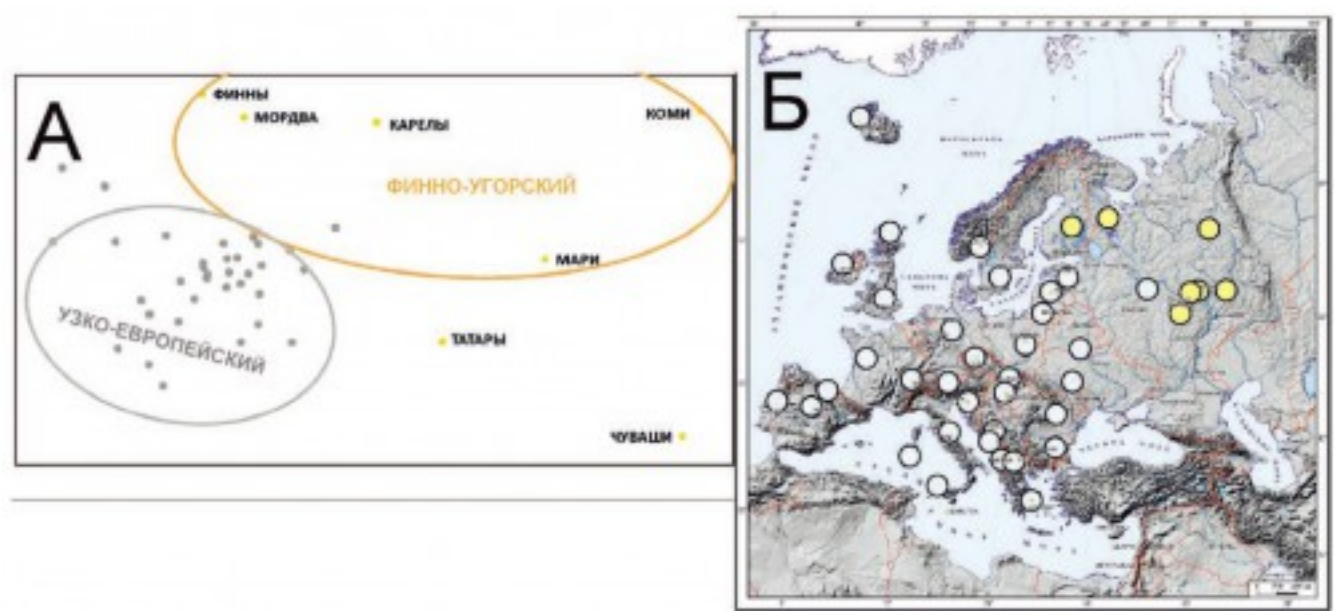


Рис. 3.3. Генетические взаимоотношения популяций, вошедших в «пан-европейский» кластер на рис. 3.2.
А. График многомерного шкалирования (стресс=0,138). Популяции разные кластеров показаны разным цветом.
Б. Географическое положение изученных популяций. Цвет каждой популяции такой же, как на графике А.

Интерпретация высокого межпопуляционного разнообразия в «финно-угорском» кластере будет дана ниже, при рассмотрении гаплотипического разнообразия мтДНК. Пока же отметим, что, хотя этот кластер популяций географически располагается на востоке Европы вблизи границы с Азией, его генетическое своеобразие нельзя объяснять азиатскими влияниями. Ведь миграции из Азии, отразившиеся в высокой частоте восточно-евразийских гаплогрупп, уже сформировали кластер, проявившийся при предыдущем генетическом масштабе (рис. 3.2 А), в котором финноязычные популяции были генетически неотличимы от остальной Европы. Мы видели, что лишь при повышении разрешающей способности анализа и переходе от генетического масштаба рис. 3.2 к генетическому масштабу рис. 3.3. (при рассмотрении только изменчивости генофонда только самой Европы, без восточно-евразийского влияния) проявилось своеобразие финноязычных народов. Таким образом, финноязычные народы (с примкнувшими к ним чувашами и татарами) следует признать по митохондриальному генофонду европейцами, но четкое разделение генофонда Европы на финноязычную и остальную Европу оказывается новым и важным результатом нашего анализа.

Отметим, что все остальные популяции Европы снова оказались генетически почти неразличимы, опять формируя гомогенный кластер, который теперь можно назвать «узко-европейским». Это снова демонстрирует генетическую гомогенность Европы – но на этот раз уже не всего европейского генофонда, а только «не-финноязычной» его части**.**

АНАЛИЗ ЕВРОПЕЙСКОГО ГЕНОФОНДА БЕЗ ФИННОЯЗЫЧНЫХ ПОПУЛЯЦИЙ.

В попытке выявить, присутствует ли все же генетическая структурированность в этом «узко-европейском» кластере, мы вновь повторили тот же прием смены генетического масштаба. Для этого исключили из рассмотрения генетически отличающиеся популяции («финно-угорского» кластера) и проанализировали остальные популяции Европы (только «узко-европейского»

кластера).

Результаты этого анализа показаны на рис. 3.4. Только при этом, третьем по счету, «увеличении» нашего популяционного микроскопа, мы, наконец, обнаруживаем различия между этносами внутри Европы. Хотя изученные популяции занимают почти все пространство графика и четкие кластеры поэтому не образуются, все же по сходству популяций в генетическом пространстве можно выделить несколько групп популяций. Более строго их было бы назвать «зонами», поскольку термин «кластер» предполагает отграниченность от соседних групп, а в нашем случае одна зона тесно соседствует с другой. Но для сохранения единства терминологии мы будем называть выделенные группы генетически сходных популяций по-прежнему «кластерами».

В один кластер (рис. 3.4 А) вошли испанцы, португальцы, французы, сардинцы, итальянцы. Все это народы романской языковой группы индоевропейской лингвистической семьи. В другой кластер вошли латыши и литовцы, то есть народы балтской языковой группы. В третий — соседний — кластер попали русские, словенцы, поляки, чехи, словаки, боснийцы, хорваты, украинцы, болгары. Все это – народы славянской языковой группы. Народы германской языковой группы – немцы, шведы, норвежцы, швейцарцы, австрийцы, англичане, исландцы – тоже образуют свой собственный кластер. Учитывая, что население Шотландии лишь исторически недавно сменило кельтский язык на язык германской группы, их генетическое сходство с ирландцами обрисовывает общность митохондриального генофонда и кельтской группы тоже. Албанцы — единственные представители албанской языковой группы – располагаются примерно в центре графика. Баски – единственный представленный на этом графике народ, который не относится к индоевропейской семье, соответственно занимает обособленное положение.

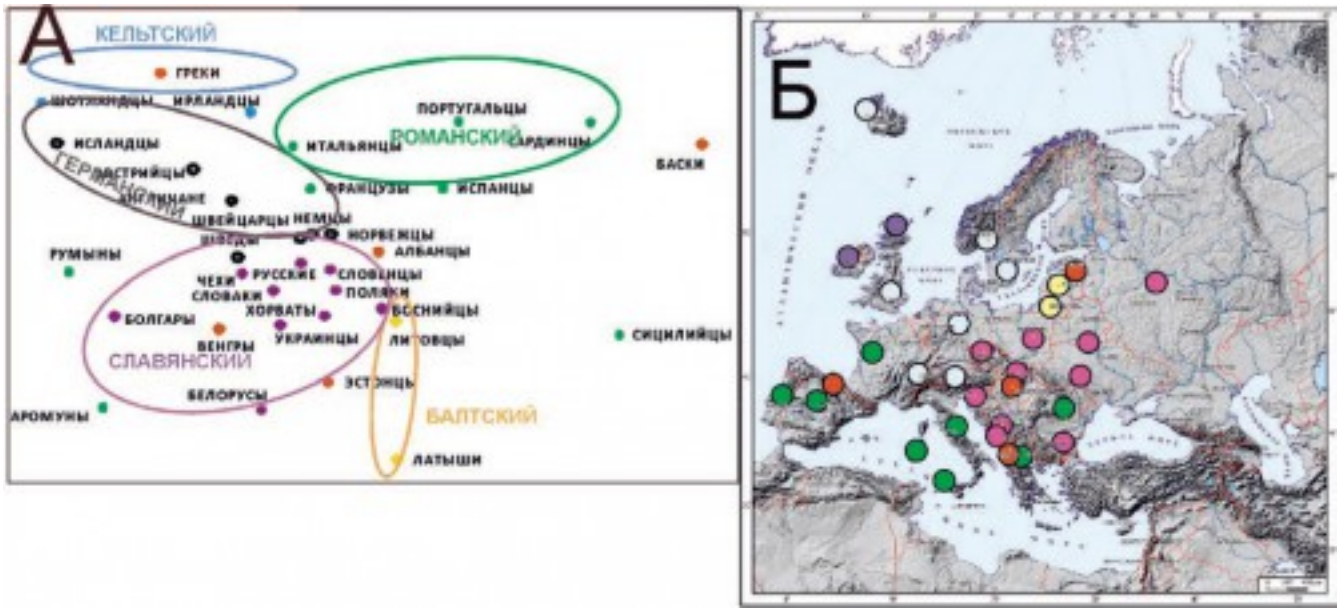


Рис. 3.4. Генетические взаимоотношения популяций, вошедших в «узко-европейский» кластер на рис. 3.3.
А. График многомерного шкалирования (стресс=0,19). Популяции разных кластеров показаны разным цветом.
Б. Географическое положение изученных популяций. Цвет каждой популяции такой же, как на графике А.

Такой впервые полученный результат позволяет нам сделать два важных вывода.

1) Европа не гомогенна по мтДНК: в ее пределах есть генетические кластеры.

2) **Эти кластеры образованы народами, сходными по языку, то есть митохондриальный генофонд Европы упорядочен по лингвистическому принципу**.

Этот вывод подтверждается и тем, что при предыдущем генетическом масштабе формирование кластеров также следовало лингвистическому принципу, поскольку отделились финно-язычные народы (уральская лингвистическая семья), а остались в кластере только народы, говорящие на языках индоевропейской лингвистической семьи, и баски – осколок мира, предшествовавшего распространению индоевропейских языков.

Из общей закономерности, выявленной в ходе анализа митохондриального генофонда Европы (популяции одной языковой группы характеризуются сходными митохондриальными генофондами, рис. 3.3 и 3.4), обнаруживаются всего лишь три исключения. Они заданы наложением географического фактора на лингвистический. Так, романоязычные румыны и аромуны, географически оторванные от большинства романских народов, и генетически располагаются особняком. Эстонцы, единственные из финно-угров, расположились между славянским и балтским кластером — то есть между своими географическими соседями. Итак, на графике можно увидеть как роль сходства языков, так и важность фактора

географического соседства. Поэтому необходимо обратиться к решению вопроса: так что же все-таки определяет генетические дистанции между популяциями Европы — лингвистика или география? Этот вопрос рассматривается в следующем разделе.

Отметим, что успех проведенного анализа, по сути, впервые выявившего структурированность европейского генофонда по мтДНК, обусловлен, на наш взгляд, прежде всего обширностью созданной базы данных и использованием в анализе лишь больших выборок. Если бы в анализ были включены выборки малого объема, то статистический «шум» не позволил бы разглядеть закономерности в структуре генофонда. Вторым фактором успешности проведенного анализа явился примененный метод последовательного «увеличения» разрешающей способности через последовательное выявление и исключение из анализа генетически удаленных популяций.