Геногеография: древняя ДНК, современные генофонды и сотрудничество с этнографами

Олег Балановский

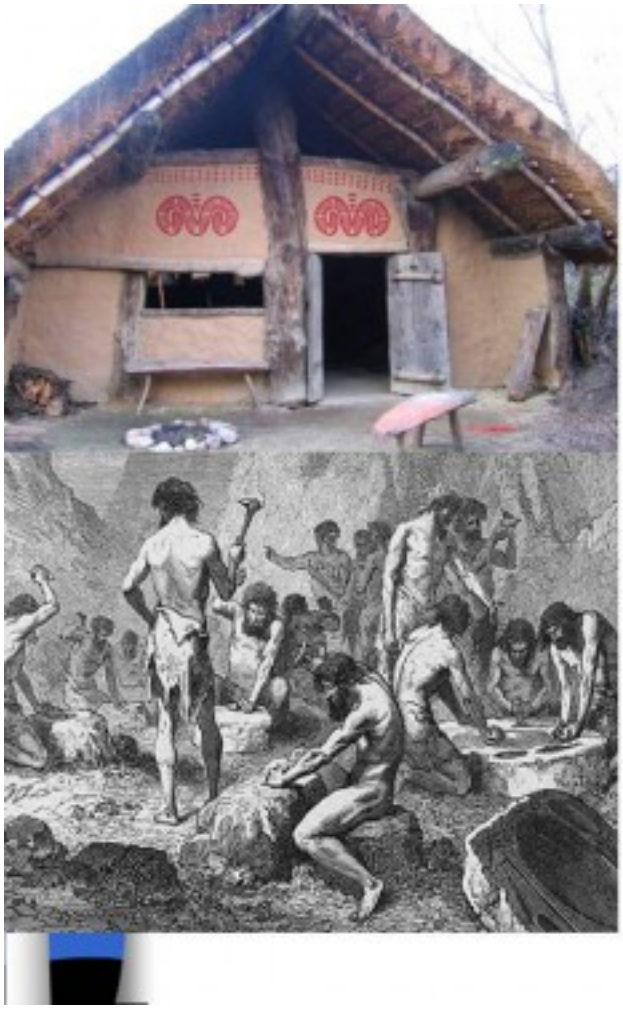
Одной из пограничных областей, в которой особенно ярко проявляются контакты между науками, является анализ древней ДНК, потому что тут работают совместно генетики, археологи, палеоантропологи, а зачастую привлекаются и лингвисты. Поэтому первая часть доклада будет посвящена древней ДНК, во многих исследованиях которой мне довелось участвовать. Анализ современных генофондов, хотя и находится в ведении только одной науки – популяционной генетики – но и эти результаты активно обсуждаются самыми разными специалистами, и это вторая часть доклада. Наконец, сотрудничество генетиков и этнографов – один из ярких примеров междисциплинарного подхода, но оно иногда наталкивается на недопонимание того, в чем совпадают, и в чем различаются этнос (объект этнологов) и популяция (объект генетиков).

Часть I Древняя ДНК

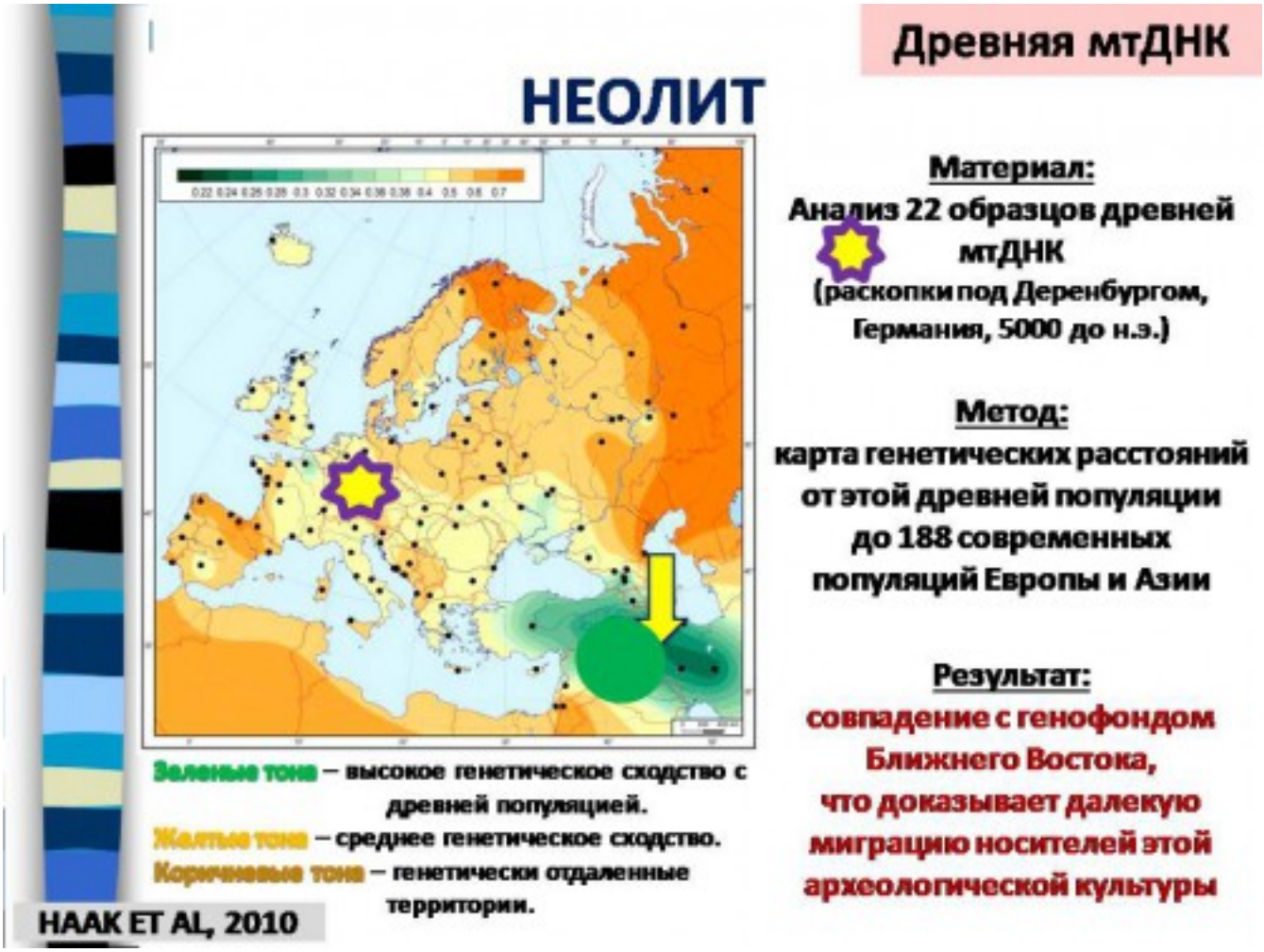
Идея изучения древней ДНК очень простая – взять древнюю косточку и сравнить ее ДНК с ДНК из всех современных популяций. Но технически это сложно: в выдыхаемом нами воздухе содержится больше ДНК, чем во всей этой косточке. И для того, чтобы изучать не свою ДНК, а ту, которая содержалась в этой косточке, нужно надеть скафандры и 90% времени тратить на поддержание стерильности помещения. В мире есть некоторое количество лабораторий, которые занимаются этим достаточно надежно.

Один из вопросов, который привлекает больше всего внимания, это неолитизация Европы. У исследователей сложилась такая точка зрения, что генофонд Европы сформировался в ходе двух событий– во-первых, когда люди туда попали впервые, и, во-вторых, когда туда пришли земледельцы. Ведь появление земледелия (производящего хозяйства) привело к многократному увеличению численности, поэтому вклад этой миграции многократно важнее, чем вклад последующих. И сейчас господствует точка зрения, что генофонд Европы сформировался под влиянием миграции населения с Ближнего Востока. Согласно теории демической диффузии, неолитическая популяция, расселяясь, ассимилирует автохтонное население, формируя градиент в генофонде. И этот градиент подтверждается данными последних исследований. Земледельцы двигались с Ближнего Востока, постепенно смешиваясь с местным населением, и формируя генофонд Европы. Однако данные по митохондриальной ДНК указывают на более древние датировки формирования европейского генофонда, чем неолит, что не согласуется с такой концепцией.

Один из способов разрешить это противоречие, это непосредственно посмотреть генофонд первых неолитических земледельцев Европы и разобраться, местные ли это уроженцы, которые восприняли технологию земледелия, или это мигранты издалека.



На карте ниже показаны результаты сравнения 22 образцов древней мтДНК, найденных на территории Германии, с 188 современными популяциями Европы и Азии. Методом генетических расстояний показано сходство древней мтДНК с генофондом Ближнего Востока. Так что не приходится сомневаться, что была миграция древнего ближневосточного населения, земледельцев. По-видимому, она была не слишком многочисленна, и генетически растворилась в местном населении.



С появлением возможности исследования полных геномов, в том числе древних ДНК это представление получило наглядное графическое подтверждение. На этом рисунке видно, где находится генофонд населения Ближнего Востока, где – геном европейских охотников-собирателей, и что геном современных европейцев находится посередине между ними. Так что выявляется смешение этих групп населения. Но выявляется и нечто третье – это древний евразийский компонент, который тоже вносит свой вклад в генофонд Европы. Статья о взаимодействии этих трех компонентов была опубликована в Nature.

Если обратиться к более древней эпохе – к мезолиту, то была изучена популяция мезолитических охотников-собирателей Оленеостровского могильника на Южном Оленьем Острове в Онежском озере. А зона генетически сходных с ними популяций, как показывает карта, оказалась в Сибири. И это указывает на миграцию, произошедшую в мезолите из Сибири на территорию Европы.

Была изучена жившая примерно в том же регионе, но на 3 тысячи лет позже, популяция с Большого Оленьего острова, которая тоже обнаружила сходство с популяциями Сибири. Это доказывает, что миграции из Сибири были множественными. Интересно, что современное население этих мест не имеет ничего общего с той древней популяцией. То есть, та древняя популяция не оставила ныне живущих потомков, а значит, установить эту миграцию по современному населению мы бы никогда не смогли.

Приведу одну из самых недавних глобальных реконструкций, в которую включены данные и по древней, и по современной ДНК, эта некая схема ветвления человечества. Тут есть разделение на африканцев и неафриканцев, потом отделение некой загадочной ветви «базальных европейцев», последующее разделение населения Северной Евразии и Западной Евразии и объединение некоторых из этих потоков. Здесь важно то, что сейчас мы имеем возможность не реконструировать некоторые отдельные направления миграции, а осознавать всю сложность популяционной системы – как разделение, так и смешение.

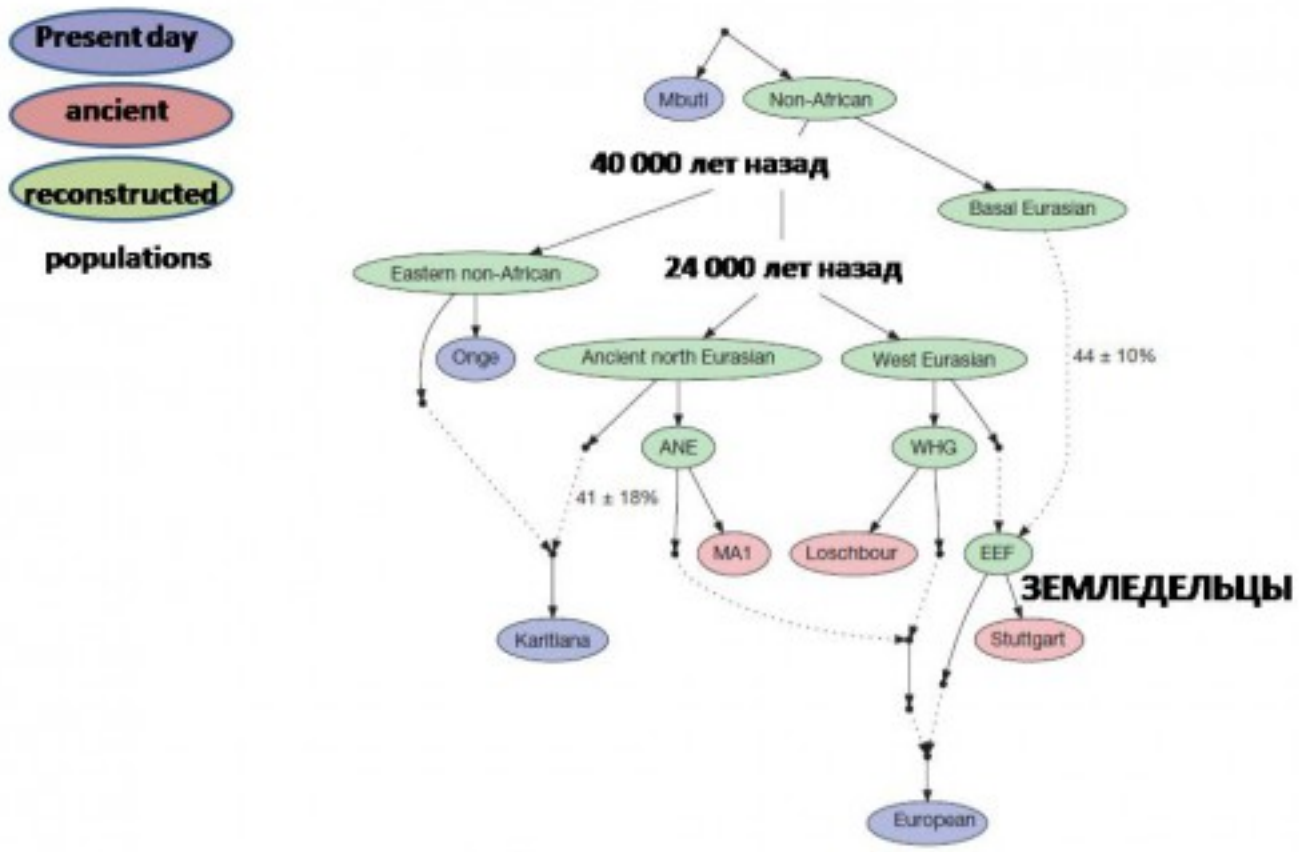


Figure 3 | Modelling of west Eurasian population history. A three-way

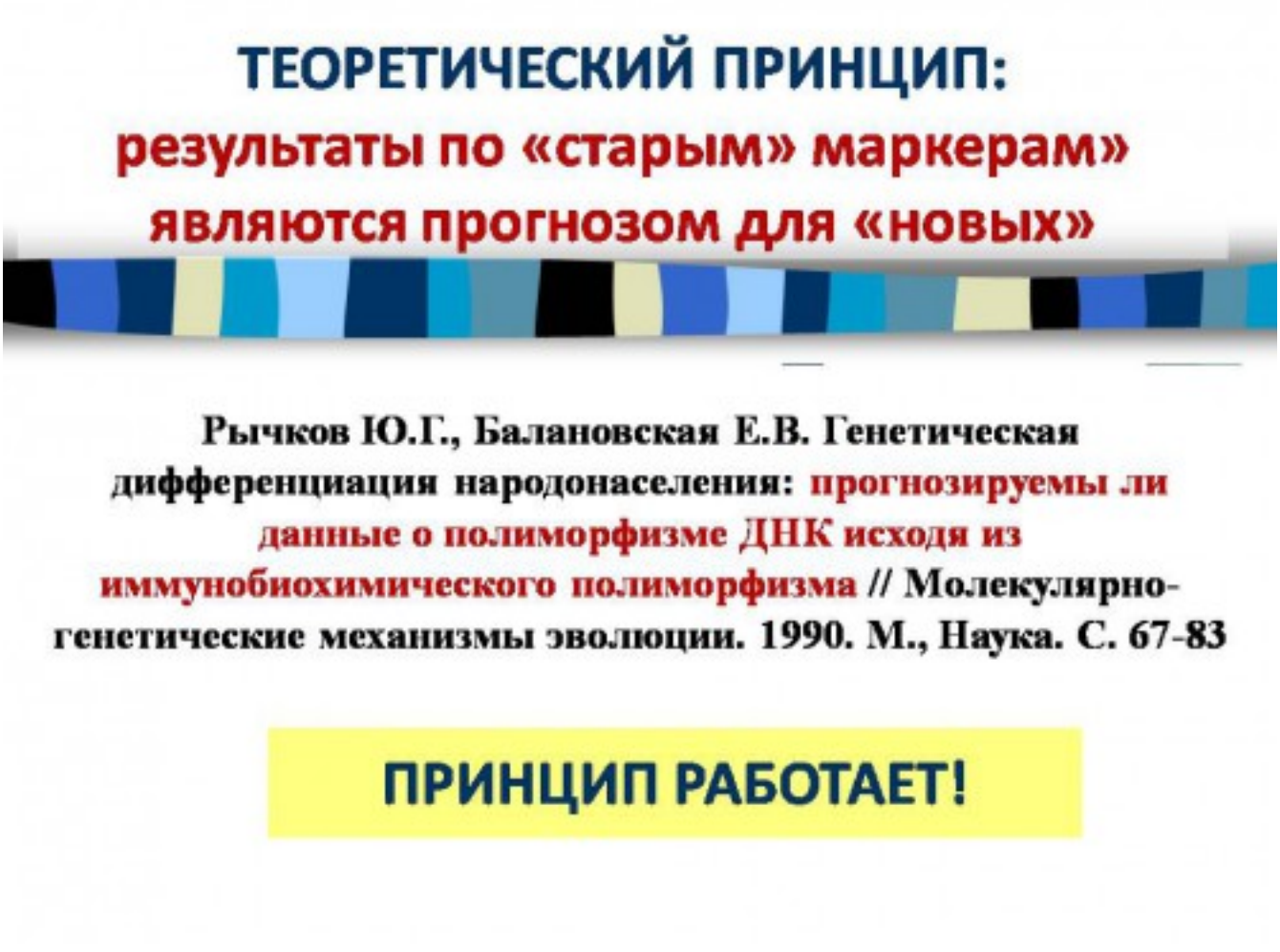
Часть II Современные генофонды

Перейдем от древней ДНК к современной. Наука геногеография – в широком смысле это наука, изучающая генофонды. В узком понимании генетическая география — это карты. Справедливость этого в том, что карта действительно является сердцевиной всей геногеографии. Пространство — это гораздо более важное измерение для изучения генофонда, чем время. В этом году геногеография могла бы праздновать свой столетний юбилей, потому что впервые идея о том, что частоты генов могут быть у людей разными, что на одной территории гены встречаются с одной частотой, а на другой территории могут быть с совсем другой частотой – возникла из практики Первой мировой войны.

Когда стало необходимым массовое переливание крови, выяснилось, что для переливания крови русским, сербам, французам, неграм, индусам и т.д. нужно запасаться разными объемами одних и тех же групп крови. С этого и начались исследования разнообразия человечества по генетическим маркерам. За сотню лет опубликованы десятки тысяч статей по разным типам генетических маркеров. Сначала использовались классические иммунологические маркеры, потом перешли к классическим биохимическим, потом началась эра ДНК. Существует очень много типов генетических маркеров, но сейчас используются три основных типа: Y-хромосома, митохондриальная ДНК и полногеномные маркеры. По времени использование этих пяти типов маркеров почти последовательно сменяет друг друга. Например, сейчас стремительно растут и будут расти исследования полногеномных маркеров.

Я продемонстрирую основные исследования за последние полвека по этим всем маркерам, чтобы показать, как стабильность, то есть постоянно выявляемые одни и те же результаты, так и динамику.

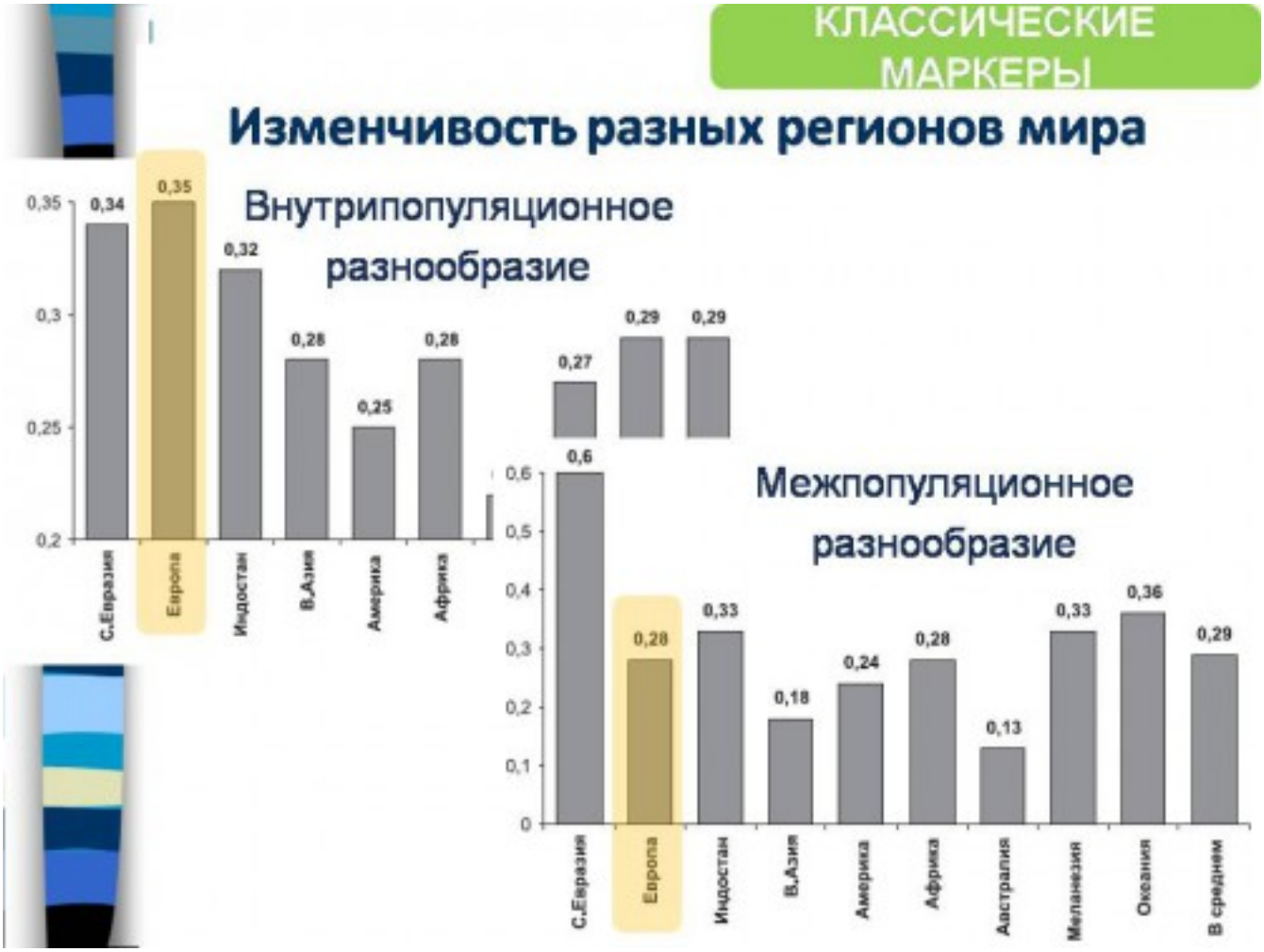
Переход от старых маркеров к ДНК стал неким испытанием, потому что можно было предполагать, что ДНК даст нечто совсем другое. Но на самом деле результаты по старым маркерам стали прогнозом для новых. Этот принцип с тех пор неоднократно подтвердился.



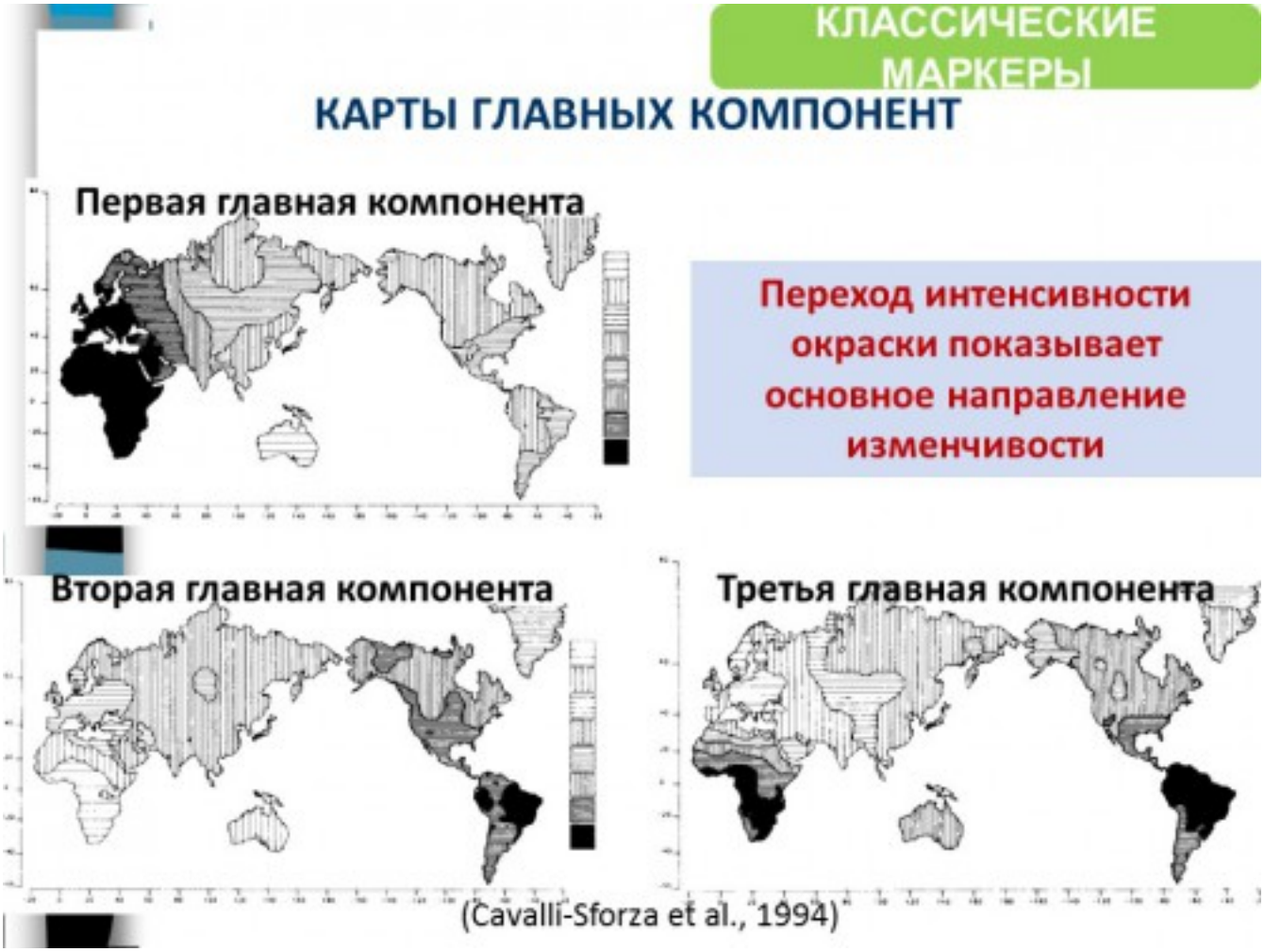
Это позволяет нам ввести полисистемный подход, то есть когда у нас присутствуют разные типы маркеров, разные признаки, мы можем изучить какие-то популяции по разным маркерам и сравнить результаты. Если они окажутся разными, это либо ошибки, либо особенности восприятия разных маркеров, которые указывают на объективные черты генофонда, но которые по-разному преломляются в призмах разных маркеров.

На следующей карте показаны популяции, обследованные в наших экспедициях. Наши ДНК-коллекции – это крупнейший в мире банк ДНК коренного населения для Северной Евразии. С 1998 по 2014 год мы собрали 20400 образцов из 220 коренных популяций. Но в базы данных, которые м ы создаем и которые используются для всех анализов, включены не только наши данные, но и почти все то, что публикуется в мире.

По классическим маркерам можно сравнить внутрипопуляционное и межпопуляционное разнообразие. Например, если посмотреть на Европу, то видно, насколько велики внутрипопуляционные различия и низки различия между популяциями.



Другим методом получены карты главных компонент, на которых переход цвета соответствует важнейшему направлению генетической изменчивости.



Главная компонента не одна – их столько, сколько исходных признаков, но важны только первые две-три из них. Так, по первой главной компоненте мы видим кардинальные различия между населением Африки и населением Австралии, а все остальное – в промежутке. Вторая главная компонента показывает основные различия между населением Африки и населением Южной Америки. Третья – изменчивость между западом и востоком Евразии.

На следующей карте изображено внутрипопуляционное разнообразие, видно, что оно наиболее высоко в западных и южных областях Евразии.

Эти данные (на рисунке выше) получены зарубежной школой геногеографии, а независимо работала отечественная школа Рычкова (рисунок ниже), и результат они получили тот же самый.

По мтДНК были проанализированы более сотни гаплогрупп в более чем 600 популяциях мира. Мы посмотрели корреляции разных признаков друг с другом. Видно, что признаки образуют неожиданное четкие кластеры по их пространственной корреляции друг с другом, и всего выявлено шесть таких главных кластеров. Сто восемнадцать признаков сходятся всего к шести основным паттернам изменчивости! Рассмотрим эти паттерны.

Если суммировать частоты гаплогрупп, входящих в разные кластеры, то видно, что первый, самый многочисленный кластер приурочен к Восточной Азии,

второй кластер – однозначно африканский,

а третий оказывается приурочен к Австралии.

Затем у нас выделяется западноевразийский кластер,

и в отдельный кластер выделилась Сибирь.

Это означает, что по мтДНК северные и южные монголоиды (кластеры 1 и 6) повышаются в ранге до вполне самостоятельных континуумов, сравнимых с африканским и австралийским.

Итак, по мтДНК выявлены те же самые геногеографические провинции мира, которые выявлялись и экстремумами главных компонент.

Говоря про Y-хромосому, выяснилось, что если проанализировать данные теми же методами, то мы получим похожие результаты. И поскольку времени недостаточно, для Y-хромосомы я не буду показывать континуумы гаплогрупп и внутрипопуляционное разнообразие, покажу анализ межпопуляционного разнообразия. Оно может быть большим на стыках контрастных, различающихся генофондов – так же, как на стыках двух материковых плит возникают горы. Примерно то же самое происходит и в генофондах. И выделяются зоны высокого популяционного разнообразия – «разломы» в мировом генофонде.

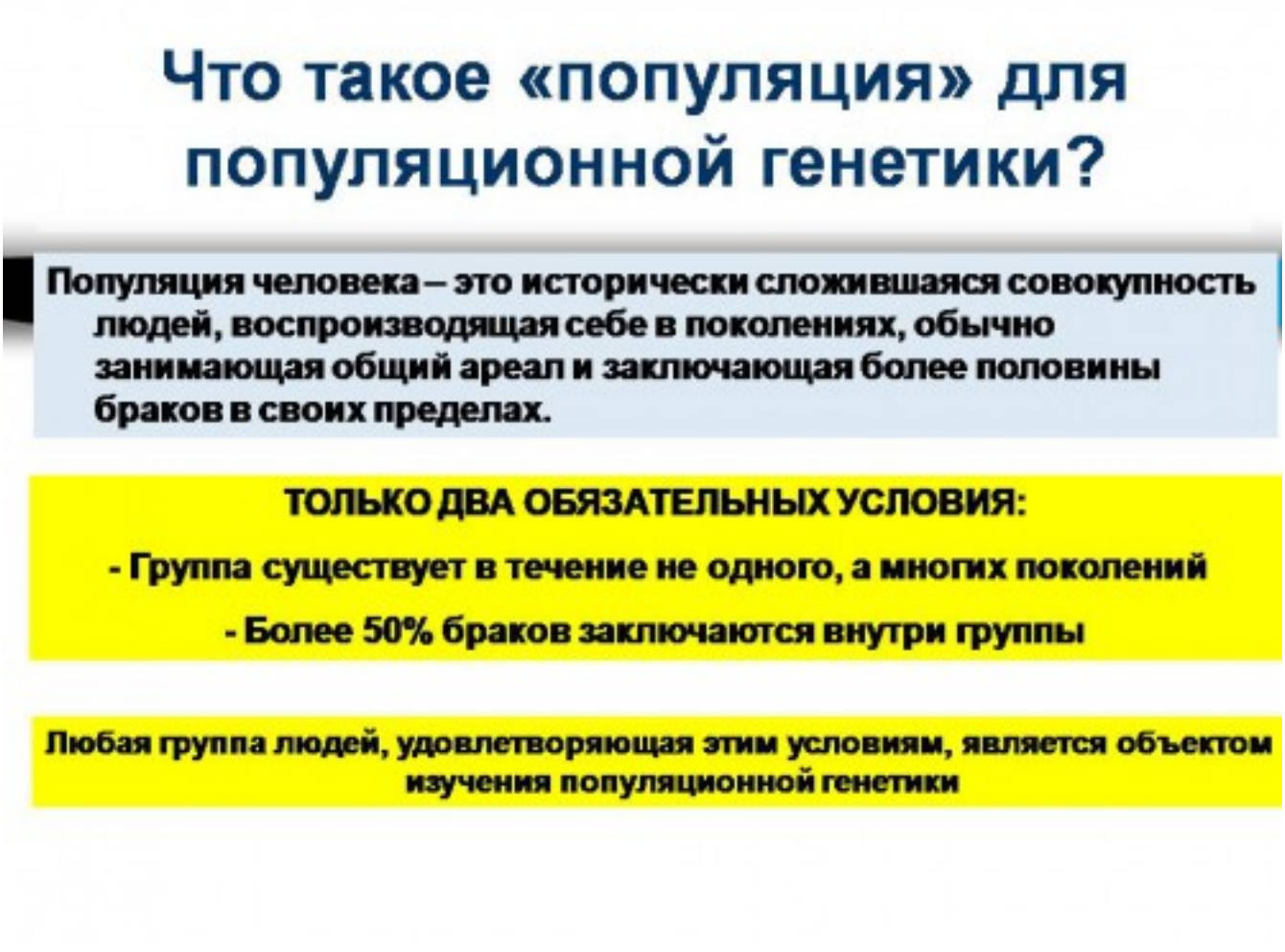
Например, одна из таких зон – это Гималаи, по разные стороны Гималаев очень разное население. А основной разлом проходит наискосок от Кавказа до Средней Сибири и делит Евразию пополам.

По полногеномным маркерам покажу только один результат – если в генетическом пространстве разместить отдельных индивидов, то люди, живущие в определенных регионах, хорошо кластеризуются друг с другом.

Часть III Генетики и этнологи

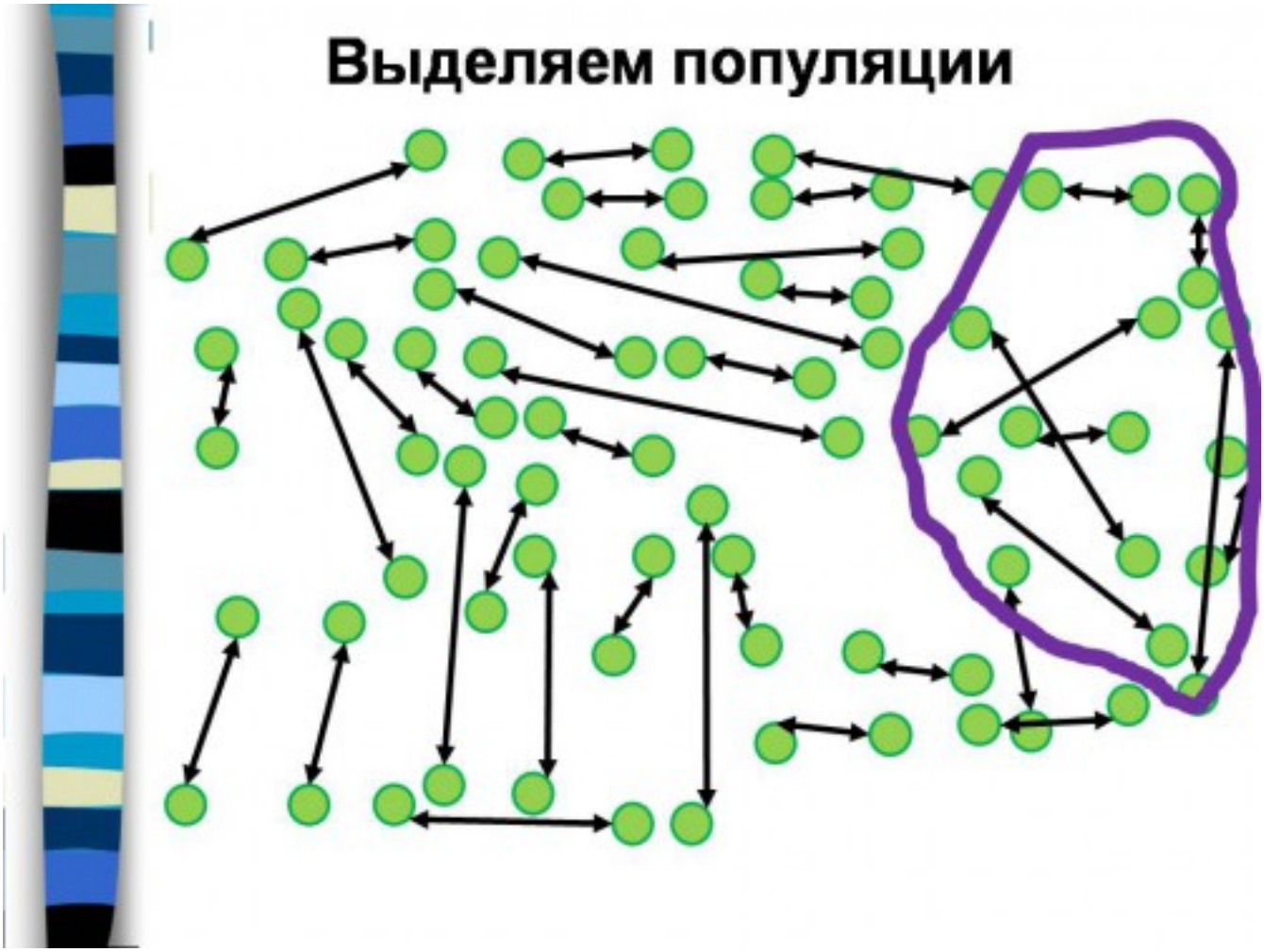
Из всех этих и многих других сведений о структуре генофонда какую информацию можно извлечь об истории популяций и происхождении народов? Обратите внимание, что до сих пор я ни разу не упомянул этносы. Привязка данных геногеографии к этносам является далеко не самым главным в изучении генофонда человечества. Но именно это вызывает наибольшее количество вопросов и наибольшую степень непонимания. Геногеография все время находится между Сциллой и Харибдой. Пучина Харибды – утверждение о том, что не существует никакой связи между генетикой и этносом. Но ужасы Сциллы еще хуже – она утверждает 100% связь между генетикой и этносом. Это то, к чему склоняется ДНК-генеалогия.

Как же все-таки пройти между Сциллой и Харибдой? Для этого надо задуматься, что если генетика имеет дело с биологическими объектами, с популяциями, откуда там вообще возникает этнос – категория социальная? Дело в том, что популяция для популяционной генетики – это любая группа людей, которая удовлетворяет всего-навсего двум условиям. Она должна существовать более чем одно поколение, и более половины браков должно заключаться внутри нее. Любая группа людей, удовлетворяющая этим критериям, является популяцией.

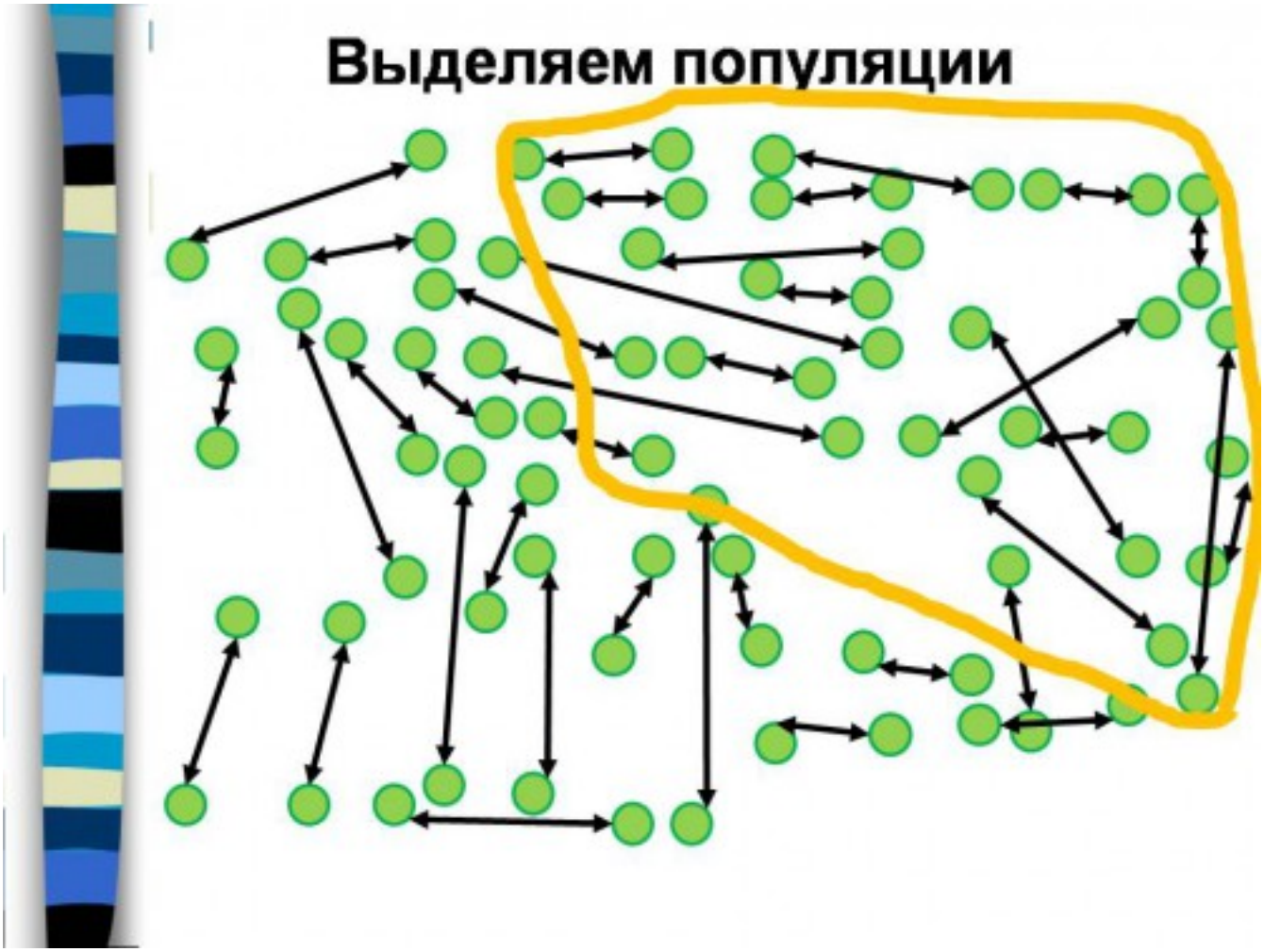


Обозначим людей кружками, а заключаемые ими браки – стрелками. Тогда совокупность людей, обрисованная красной линией, является популяцией по критерию 50% браков, потому что более половины стрелочек, которые начинаются внутри этой линии, внутри нее и заканчиваются и только 4 стрелки выходят за ее пределы.

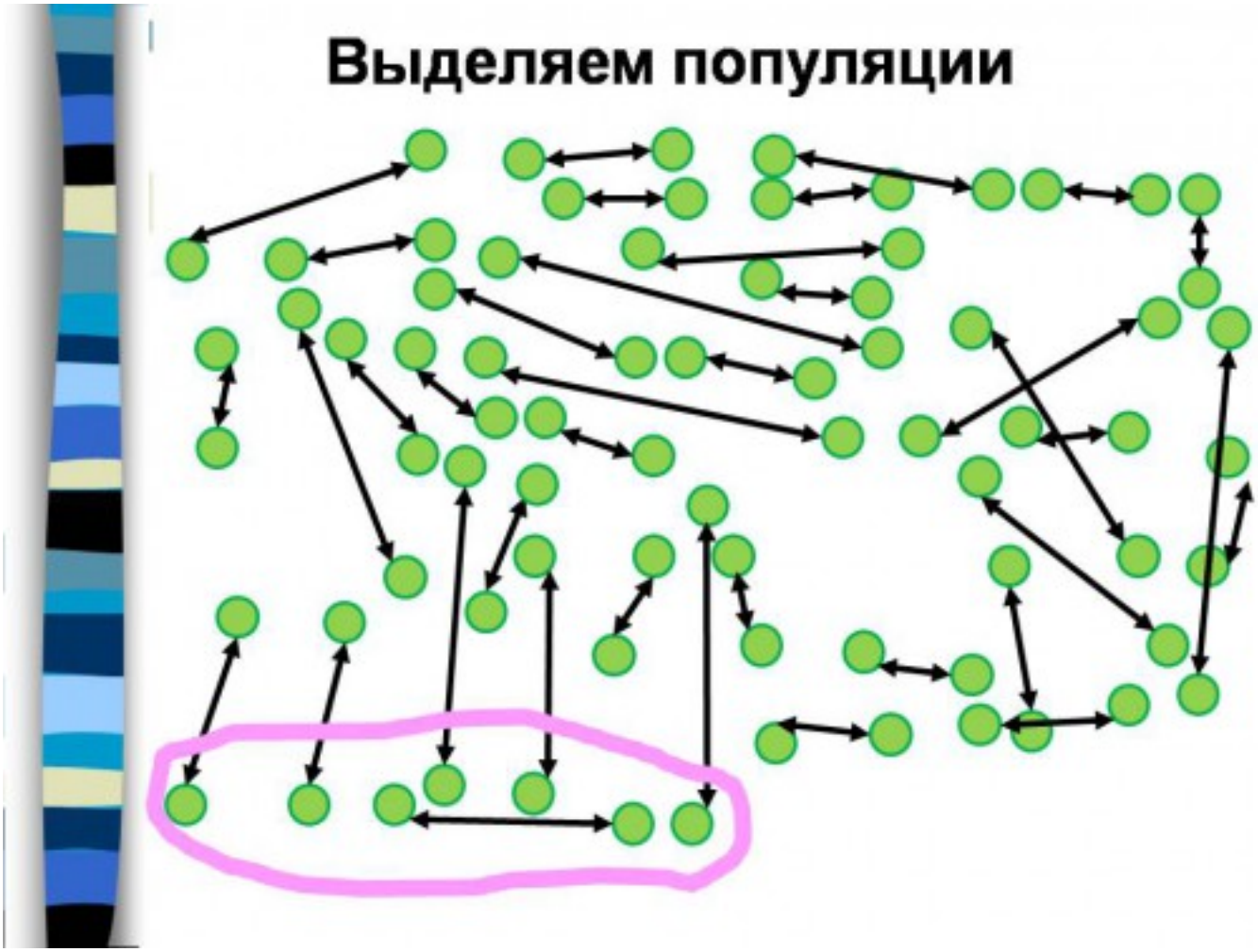
Группа людей, ограниченная фиолетовой линией, тоже является популяцией, потому что опять большинство браков заключены в ее пределах.



И группа людей, ограниченная желтой линией, тоже является популяцией. Она включает часть людей из первой популяции, и часть из второй. На этом примере наглядно видно, что популяции можно выделить разными способами, они могут быть вложенными друг в друга или перекрываться.



А является ли популяцией группа людей, ограниченная розовой линией? Давайте посчитаем: в этой группе 7 человек, но только двое из них заключили брак в пределах этой группы, а остальные 5 (большинство) – за ее пределами.



Когда доступны такие полные данные по структуре браков, выясняется, что, например, в Архангельской области – если взять одну деревню, то внутри нее заключается меньше 50% браков – это не популяция. Но если взять несколько соседних деревень, то уже больше 50% — это уже популяция.

А на юге – если взять весь Краснодарский край, то больше половины браков заключены с уроженцами других регионов. Значит, население всего Краснодарского края не является популяцией.

Когда доступны полные данные по структуре браков

АРХАНГЕЛЬСКАЯ ОБЛАСТЬ: в пределах одной деревни заключается менее 50% браков, но уже группа соседних деревень является популяцией. Население всей области – тем более является популяцией.

КРАСНОДАРСКИЙ КРАЙ – более половины браков заключается с уроженцами других регионов, поэтому населения Краснодарского края не является популяцией.

Популяции всегда перекрываются

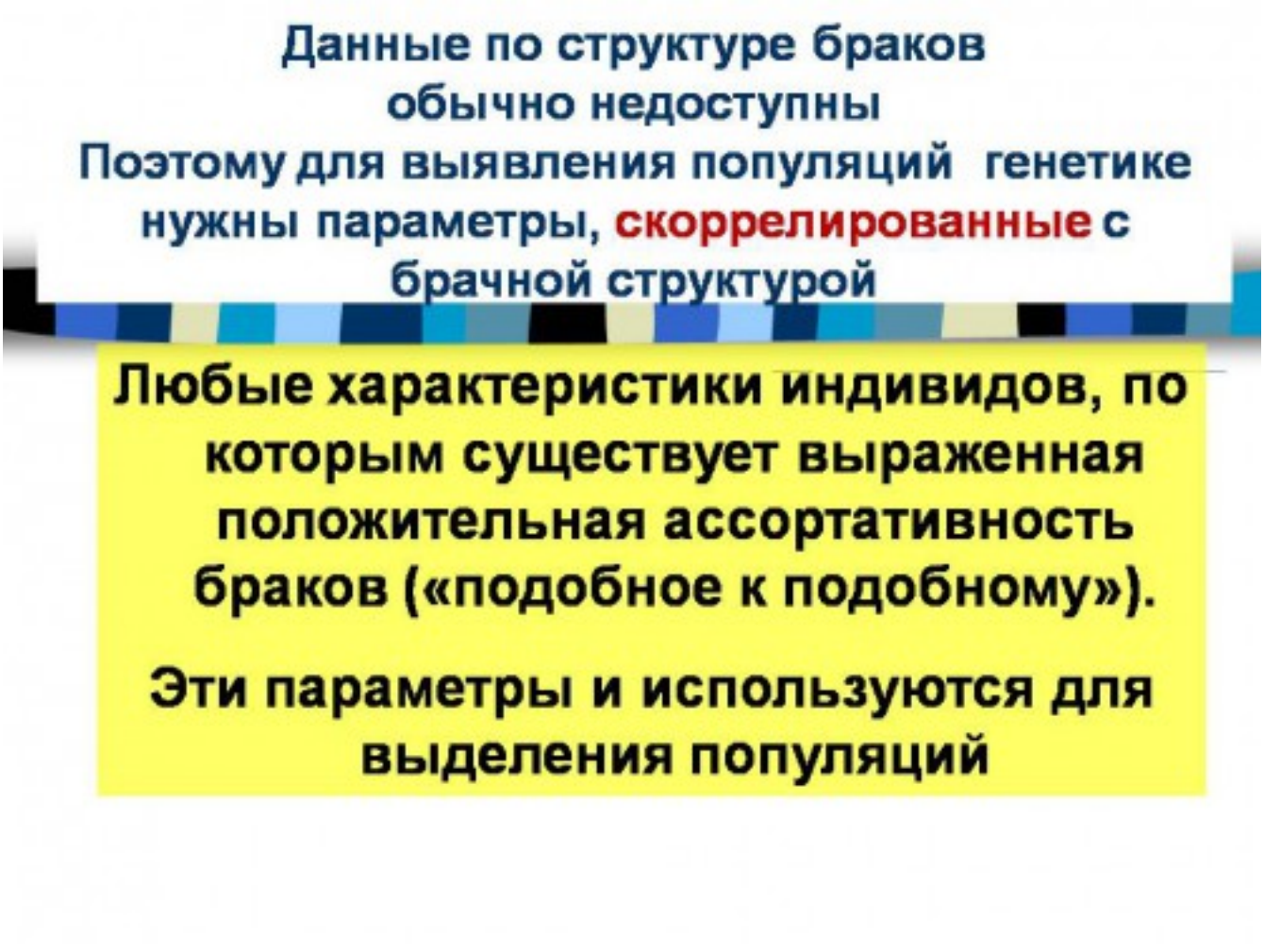
(в Краснодарском крае – есть изолят шапсугов),

их можно выделить по-разному

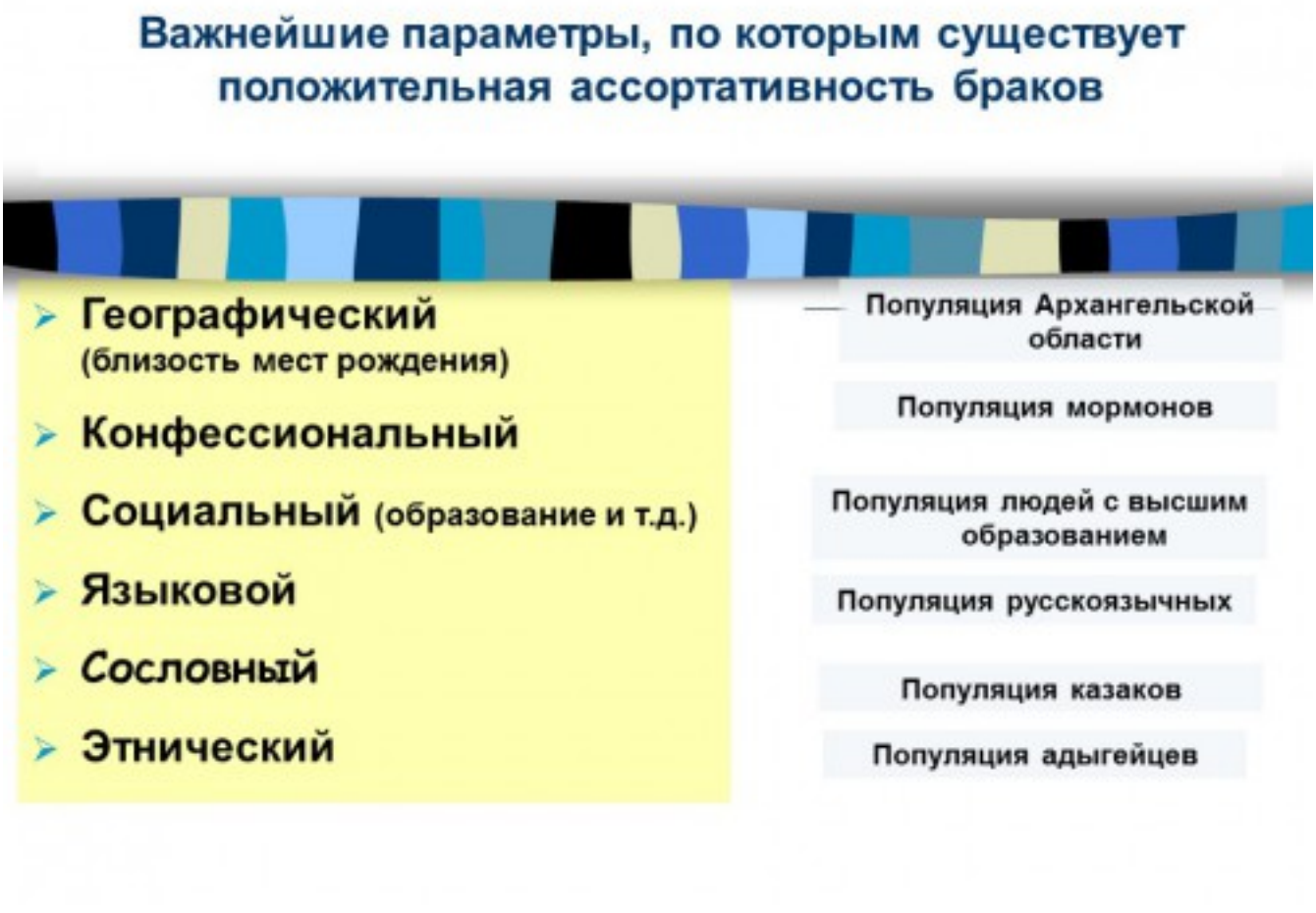
(иерархия популяций - как матрешка в матрешке),

человечество с точки зрения генетики – сложнейшая система популяций.

Популяции входят одна в другую, как матрешки, они перекрываются. Человечество представляет собой сложнейшую систему популяций. И поскольку полные данные по структуре браков обычно недоступны, то популяционным генетикам для выделения популяций нужны какие-то другие параметры, которые с брачной структурой скоррелированы. Поэтому любые признаки, по которым существует брачная ассортативность, являются признаками, очень полезными для популяционных генетиков.



Например, географический признак – он работает почти всегда – как правило, большинство браков заключается с уроженцами не слишком отдаленных регионов. Конфессиональный – например, популяция мормонов США – это классический объект популяционной генетики, на котором сделано много работ. Социальный – популяция людей с высшим образованием. Языковой, например, популяция русскоязычных, сословный – популяция казаков. Наконец, этнический.



Для самых разных популяций максимальная ассортативность браков происходит по принципу географического соседства и по формальному критерию этничности – той этничности, которую люди себе указывают.

Итак, для каждого человека, биологический образец которого анализируется популяционными генетиками, фиксируются как минимум два параметра: его происхождение из одного из известных регионов (из списка) и происхождение от одного из народов (тоже из списка).

Возникает вопрос, а откуда берутся эти списки и кто, вообще говоря, определяет, что такое регион и что такое народ?

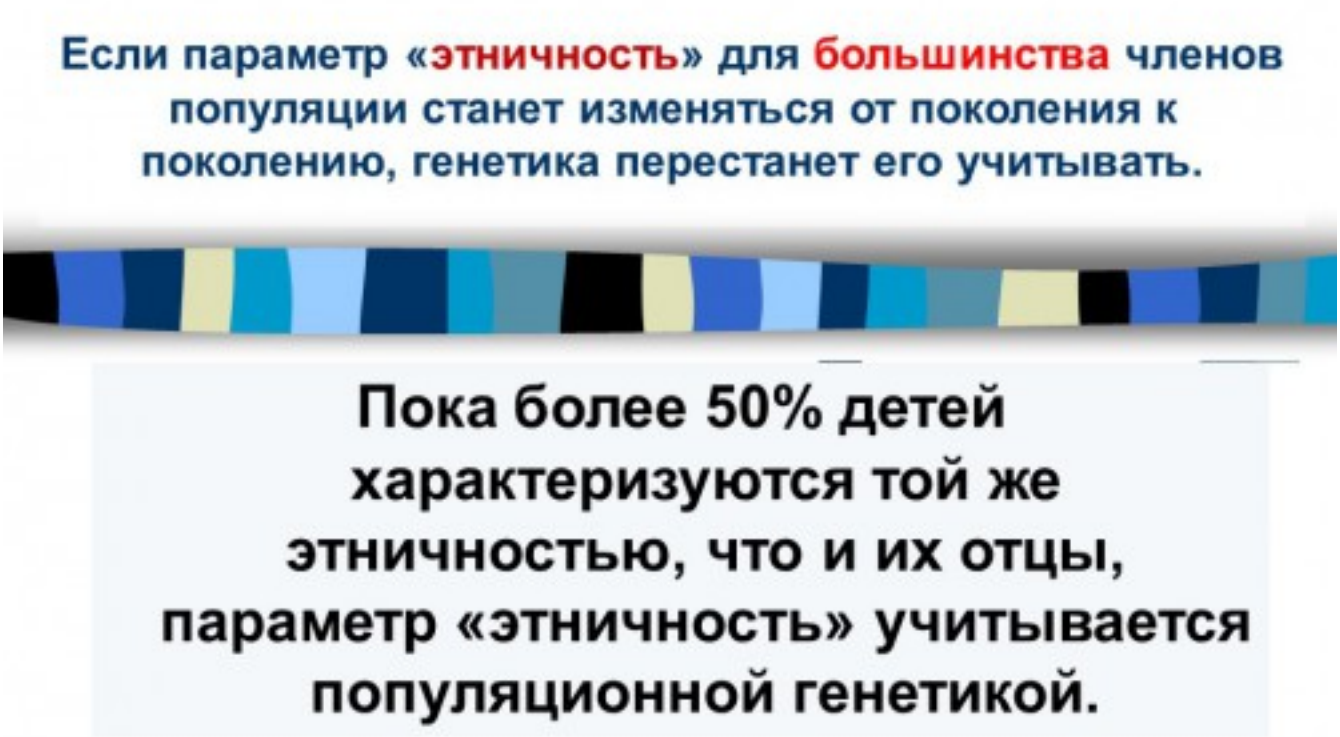
Вопрос, что такое регион, очень простой — его решает правительство, а генетика принимает то, что говорит правительство. В данный момент господствует практика, что географические границы регионов жестко закреплены, заданы на географических картах и считаются неотъемлемыми свойствами регионов. Генетика это принимает. Если же завтра выйдет закон, по которому границы регионов являются воображаемыми, конструируются самими жителями регионов и легко изменчивы, генетика примет и это. Если когда-то регионы будут вообще отменены, а место рождения будет записываться, как GPS-координаты, генетика примет и это. Мы будем работать с такими данными и выделять популяции по принципу близости GPS-координат.

А вопрос, что такое этнос, решает этнология. Генетика принимает то, что говорит этнология. До недавнего времени существовала единственная концепция эссенциализма, которая воспринимает принадлежность к определенному этносу как изначальное и неотъемлемое свойство людей. И мы с этим работаем. Недавно, наряду с эссенциализмом, появилась и концепция конструктивизма, которая рассматривает этнос как воображаемую конструкцию, создаваемую самими людьми и легко изменчивую. И мы это тоже принимаем, и работает с таким пониманием этноса. Если в будущем в этнологии возникнут иные концепции, генетика примет и их.

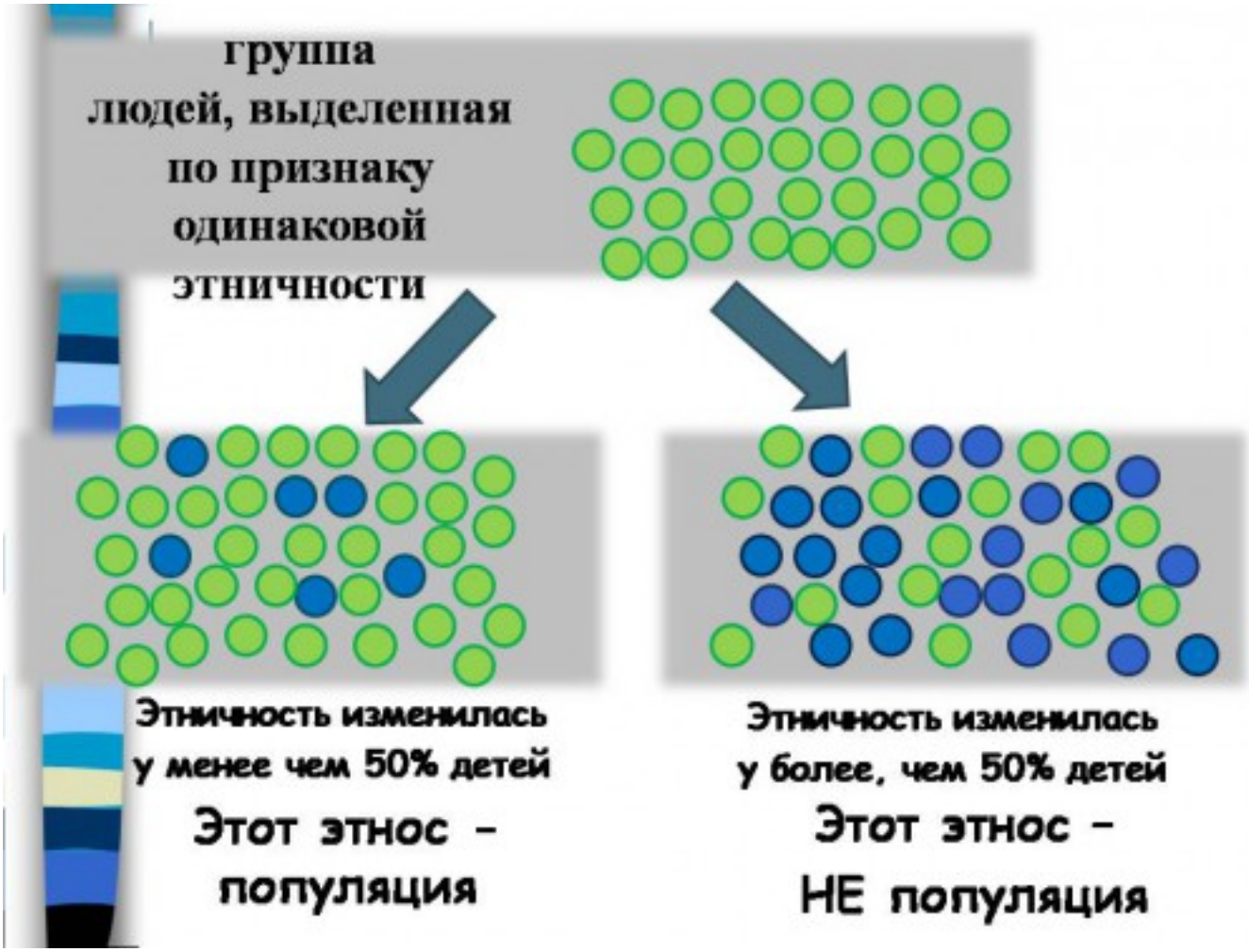
Генетику интересует только одно – по какому признаку ассортативности заключаются браки. Но тут возникает вопрос о стабильности признака в поколениях, потому что если признак нестабилен, то он становится неинформативен для выделения популяций. И вот более 20 лет назад на одном из докладов мы представили пример из нашей экспедиционной практики на Сахалине, когда человек, оба родители которого – нивхи, объявил, что он норвежец.

Вопрос тут число количественный. Если место рождения для **большинства** членов популяции станет изменяться от поколения к поколению, генетика перестанет учитывать место рождения как информативный для генетики показатель. Но пока более 50% детей характеризуются тем же регионом рождения, что и их родители, место рождения учитывается и помогает нам работать.

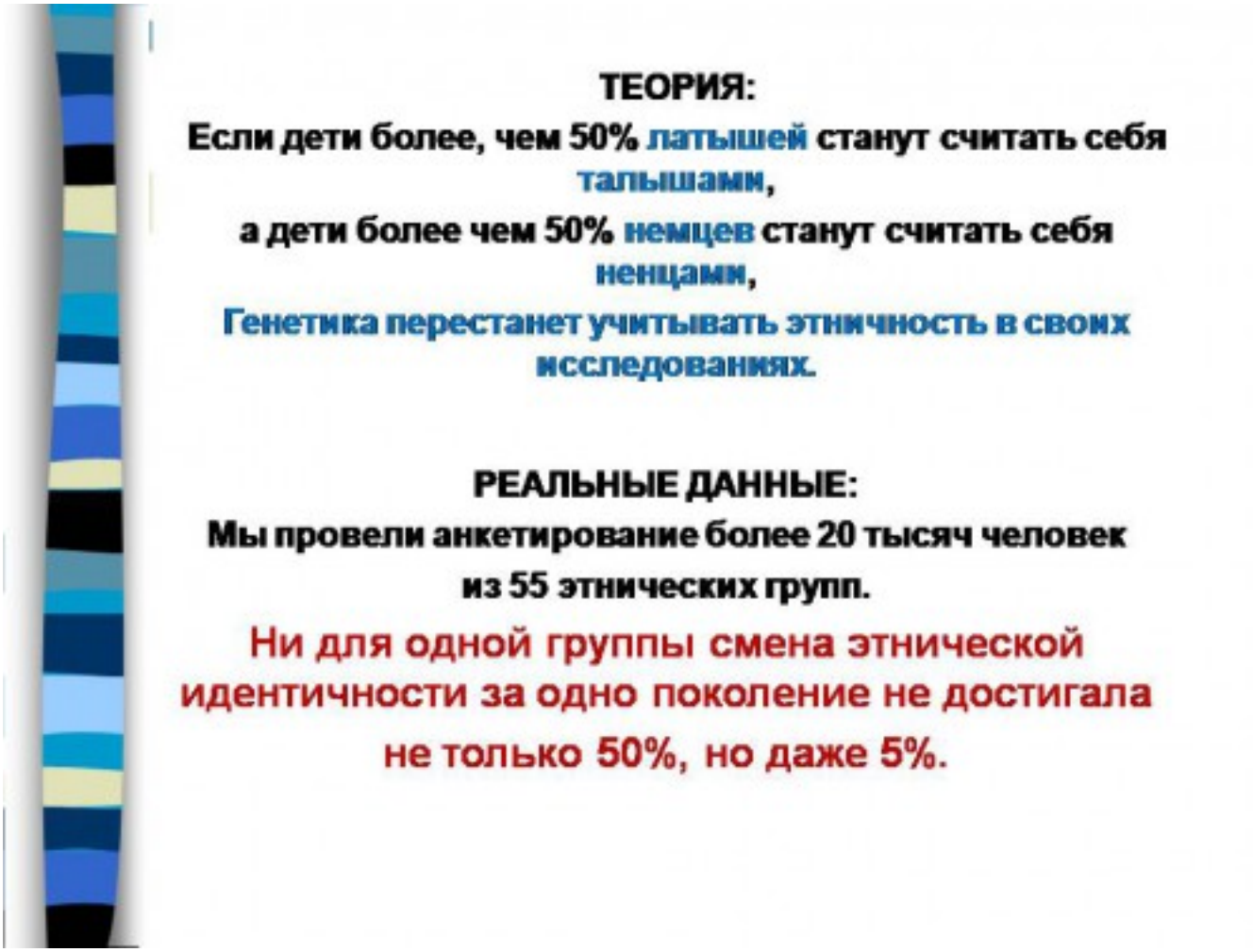
Точно так же, если параметр этничность для **большинства** членов популяции станет изменяться от поколения к поколению, генетика не станет его учитывать.



Если взять группу людей одинаковой этничности, то есть вероятность, что среди их потомков часть будет причислять себя к другому этносу. Но если таких потомков меньше 50%, то признак информативен, этот этнос – популяция. А вот если этничность изменится у более 50% детей, значит, этот этнос не популяция. И тогда признак этничности для генетики не работает.



Реальные данные такие – в ходе наших исследований мы провели анкетирование более 20 тысяч человек из 55 этнических групп. И ни для одной этнической группы смена этнической идентичности за одно поколение не достигала не только 50%, но даже и 5%. Поэтому эти два параметра – этническая самоидентификация и место рождения, пока отражают объективную реальность и остаются важнейшими критериями выделения популяций.



Практическая эффективность такого подхода была выявлена в том случае, когда по Y-хромосомному гаплотипу исполнителя домодедовского теракта удалось определить место его рождения. Этот гаплотип точно совпал с группой, которая была встречена только у одной группы в мире – ингушей Ингушетии. Благодаря этому личность исполнителя была быстро установлена и подтверждена прямой ДНК-идентификацией.

Это не единичный случай – Следственный комитет обращается к нам ежемесячно.

ОТНЮДЬ НЕ ЕДИНИЧНЫЙ СЛУЧАЙ

Ежемесячные запросы от Следственного комитета

Благодарность от МВД

Уважаемому ученому сотруднику,
руководителю группы геномной генографии
Институт Российской академии наук
БАЛАНОВСКОМУ О.П.

БЛАГОДАРСТВЕННОЕ ПИСЬМО

Уважаемый Олег Павлович!

Примите, пожалуйста, искреннюю благодарность за Вашу серьёзную помощь в личном участии в розыске в Ставропольском крае особо опасного преступления. Совместная работа с Вами послужила для Вашего более старшего наглядным примером высокого профессионализма и творческого подхода к порученному делу.

Желаю Вам отличного здоровья, семейного благополучия, отличных и новых успехов в разработке современных криминалистических методов, используемых на практике при установлении личности виновных участников преступного события.

С уважением,

Начальник ГУ МВД России по Ставропольскому краю генерал-майор полиции А.Г. Ондар

г. Ставрополь
октябрь 2012 г.

Сотрудничество с этнографами для генетиков очень важно, и очень многие этнографы отвечают нам взаимностью.

Сотрудничество с этнографами для генетиков очень важно.

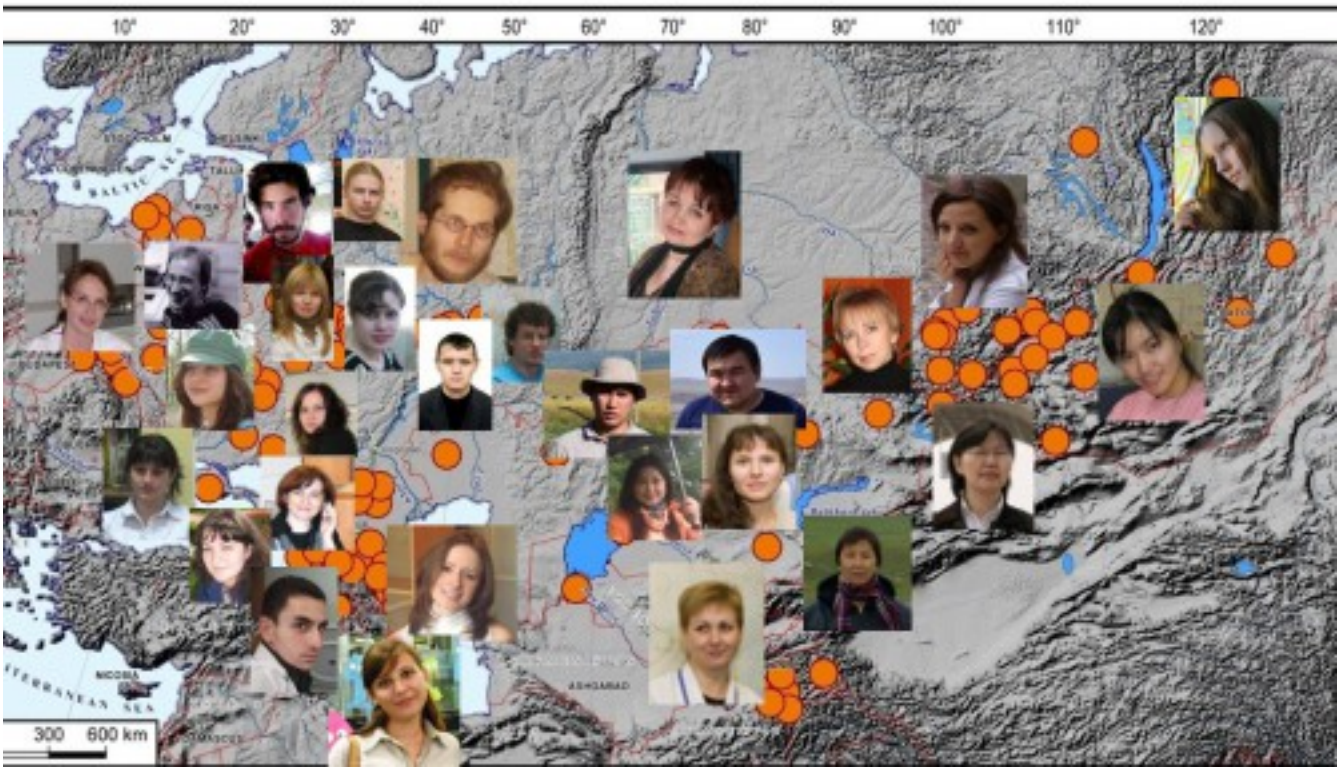
Многие этнографы считают, что это взаимно.

Экспедиция в Бурятию: этнологи и генетики.
работаем

Экспедиция в Бурятию: этнологи и генетики.
Работаем в музее

Среди нашей большой команды разнообразны места рождения.

МЕСТА РОЖДЕНИЯ ЧЛЕНОВ НАШЕЙ КОМАНДЫ



И разнообразна этническая принадлежность.

А вот браков внутри нашей команды всего 4%. Поэтому наш коллектив не является популяцией. Мы являемся частью общей популяции научного сообщества и открыты для взаимодействия с другими коллективами.

БРАКОВ ВНУТРИ НАШЕЙ КОМАНДЫ: 4%

Поэтому наш коллектив не является популяцией…

зато является частью общей «популяции» научного сообщества и открыт для сотрудничества с другими коллективами

СПАСИБО ЗА ВНИМАНИЕ

видеозапись доклада https://www.youtube.com/watch?v=YmcM2MifMeU