Человек, который попытался создать каталог человечества

Луиджи Лука Кавалли-Сфорца воплотил идею Дарвина построить древо человечества

Публикуем перевод статьи палеоантрополога Джона Хокса, опубликованную на сайте Medium, о великом итальянском ученом, лидере в области популяционной генетики, впервые начавшим исследовать генетическое разнообразие человечества, Луиджи Кавалли-Сфорца, который скончался 31 августа.

Джон Хокс (John Hawks), палеоантрополог

Луиджи Лука Кавалли-Сфорца, многим поколениям генетиков известный просто как «Лука», умер на прошлой неделе в возрасте 96 лет. Больше, чем кто-либо еще из генетиков человека, Кавалли-Сфорца верил в то, что, соединив знания о генах и культуре, мы можем проследить происхождение человечества. Он автор идей и моделей, объединящих две этих разных области науки. Его работы сформировали фундамент наших сегодняшних знаний о геномном разнообразии человечества по всему миру.

В 1991 году Кавалли-Сфорца написал эссе для журнала *Scientific American,* в котором он сформулировал поставленную перед собой научную задачу. Он вспоминает время, когда в молодости работал в Кембридже, в лаборатории Роналда А.Фишера, одного из основателей современной эволюционной теории.

«Я стал думать о проекте, настолько амбициозном, что он мог показаться безумным: реконструкция того, где возникли человеческие популяции и какими путями они распространялись по миру».

От его первоначальной работы с микроорганизмами эта идея была бесконечно далека. Но Кавалли-Сфорца выбрал удачный момент, чтобы перейти в область генетики человека. В течение 1950-х нарождалась область науки, которая накапливала данные о том, как наследуются разнообразные вариации человечества. Это открывало новые возможности для понимания эволюции популяций.

Антропологи изучали разнообразие человечества, описывая изменчивость его морфологических признаков, таких как форма черепа. Генетики нуждались в более простых системах для того, чтобы начать изучать генетическую изменчивость. Некоторые из ранее известных примеров менделевского наследования относились к генетическим нарушениям, но они были довольно редки, что делало эти примеры не слишком информативными для изучения генетического разнообразия человечества в целом.

Но есть признаки, невидимые глазом, такие как группы крови, которые также следуют законам менделевского наследования. После Второй мировой войны генетики стали развивать методы тестирования населения по этим признакам, позволяющие сравнивать географически далекие популяции. Прежде всего, это были группы крови, затем тест на восприятие горького вкуса вещества фенилтиокарбамида, а позже стали исследовать белки методом электрофореза.

Эти признаки относятся к категории «классических маркеров». Все они передаются в поколениях по менделевским законам

наследования, что дает возможность генетикам использовать математические методы для изучения того, как их частота меняется со временем. Генетики объездили все четыре стороны света, чтобы построить карты частот групп крови и других классических маркеров в популяциях. Никто при этом не знал, как давно возникли группы крови или как давно появились различия между популяциями. Но можно было наблюдать значительные различия: например, в некоторых популяциях практически не было группы крови B, в то время как в остальных она присутствовала в значительной доле. Вплоть до 1980-х классические маркеры оставались основным доказательством генетического разнообразия человечества.

* * *

Кавалли-Сфорца начал работу со своей родной Италии, путешествуя по деревням в долине Пармы и собирая образцы крови. Он пытался понять, как инбридинг в маленьких городках и деревнях был связан с частотой групп крови. С несколькими сотрудниками он изучал церковные книги записей браков и рождений, чтобы отследить миграции людей между деревнями и количество детей у них. Проследив эти множественные линии, он смог показать, что близкородственные браки стали основной причиной генетических различий между маленькими городками. Так он получил самые первые доказательства того, что люди могут быть подвержены генетическому дрейфу, случайным изменениям в частоте генов, которые случаются в маленьких популяциях.

Кавалли-Сфорца осознал, что раз генетический дрейф может объяснить частоту генов в маленьких итальянских городках, он мог влиять и на человечество в целом на протяжении всей его истории. Генетический дрейф был силой, которая в течение долгого времени медленно отодвигала популяции друг от друга, накапливая различия в частоте генов.

Это был период, когда Кавалли-Сфорца начал сотрудничество со специалистом в статистической генетике А.В.Ф.Эдвардсом, разрабатывая пути реконструкции эволюционных деревьев на основе частоты генов. Статистические методы используют изменение расстояния, которое вычисляется из частот некоторых генов в разных популяциях, и они помогли создать новую картину происхождения человека.

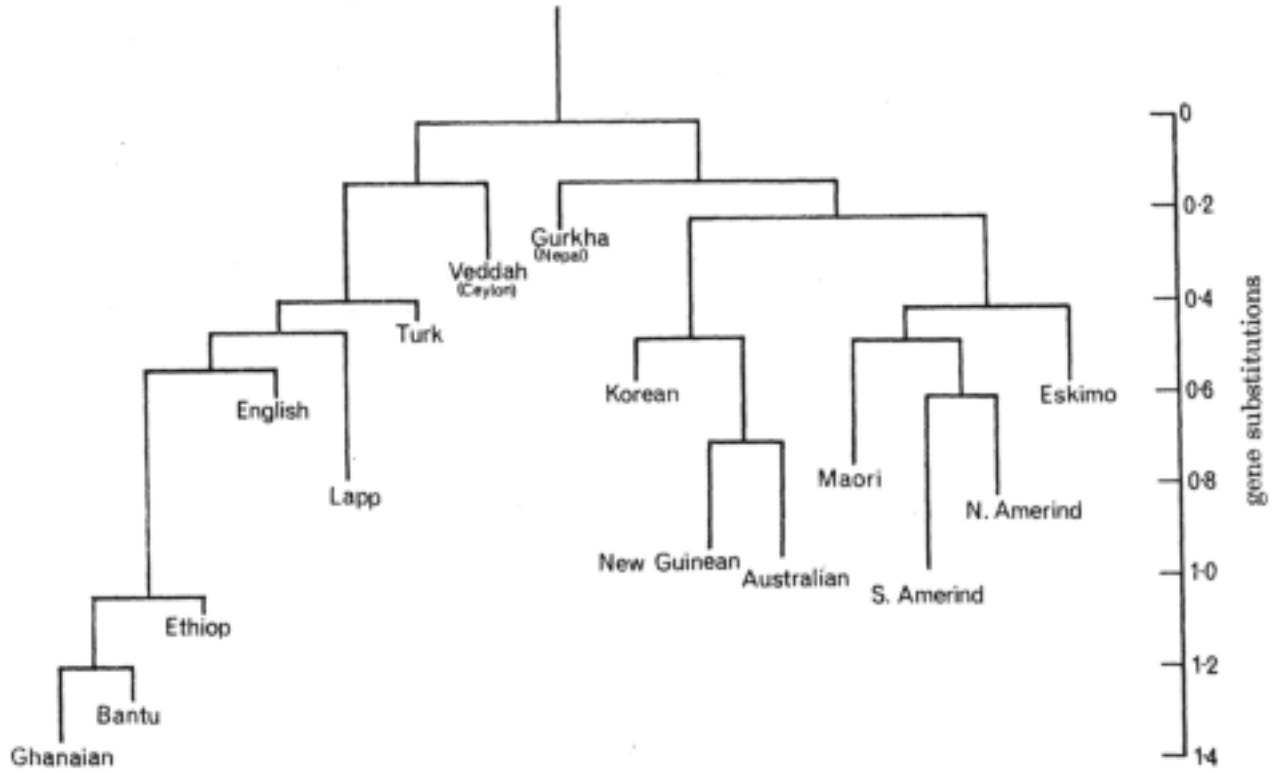


FIGURE 80. Reconstruction of human evolution using a sample of fifteen populations tested for five loci (ABO, MNS, Rh, Di, Fy).

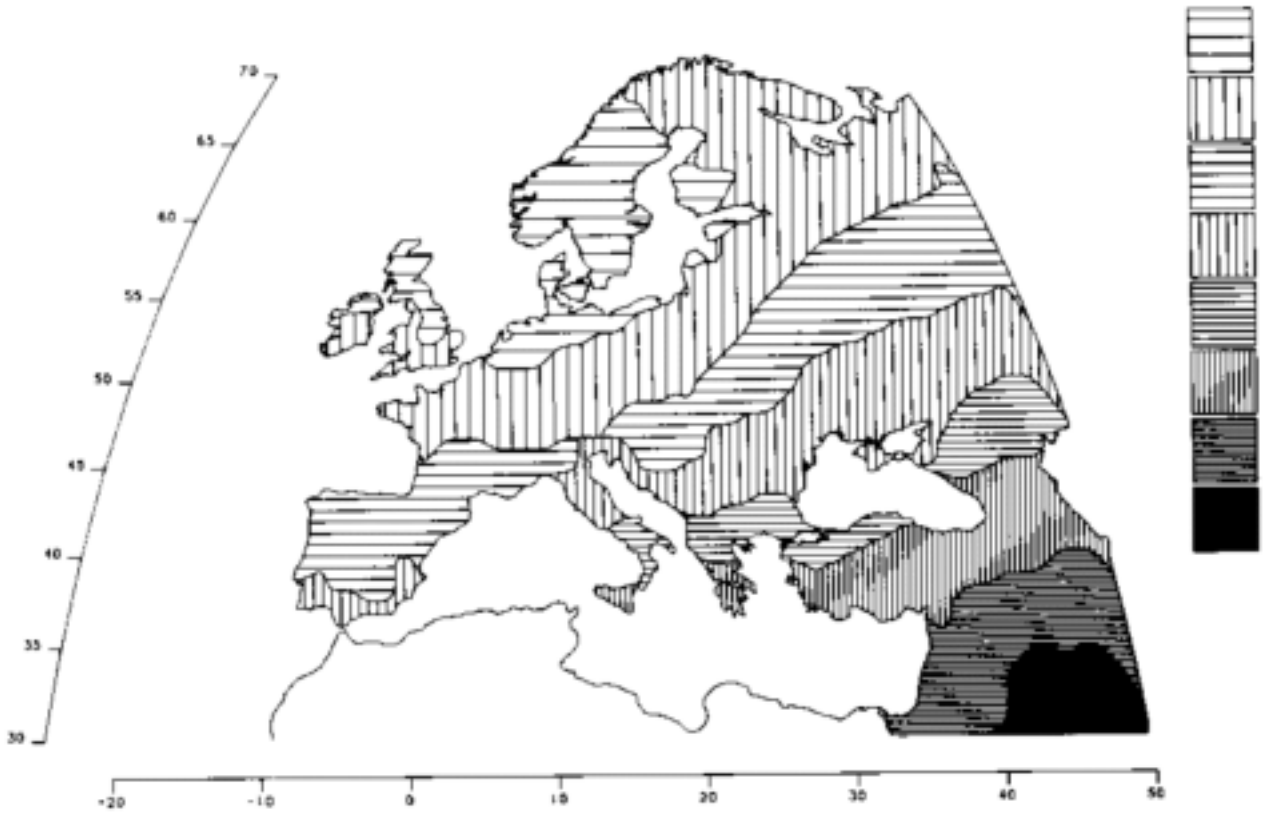
Из работы: Cavalli-Sforza 1966, “Population structure and human evolution.”

На этом дереве на одной большой ветви находятся популяции американских индейцев, Азии и Океании, на другой ветви – Европы и Африки. Оно сильно отличается от наших сегодняшних представлений: сегодня мы рассматриваем африканские популяции как наиболее разнообразную часть человечества, а не просто маленькую веточку. Важно отметить, почему первое

дерево Кавалли-Сфорцы оказалось неверным. Группы крови были открыты и изучены у населения европейского происхождения, это означало, что изменчивость в африканских популяциях не могла быть правильно изучена на основании этих же признаков. Сегодня, используя секвенирование геномов, мы можем убедительно показать, что самая древняя ветвь на дереве человечества – это Африка.

Но что более важно, дерево иллюстрирует методическое ограничение классических маркеров. Частоты нескольких генов не могут дать достаточно информации, чтобы сказать, когда и как много смешений могло происходить между популяциями. Кавалли-Сфорца, основываясь на своей работе в долине Пармы и позже с пигмеями в Центральной Африке, предположил, что миграции и смешения были редкими событиями. В его модели генетический дрейф, но не потоки генов, был основной движущей силой эволюции человечества.

Дерево для наглядного представления генетического разнообразия было мощным инструментом. Но Кавалли-Сфорца и Эдвардс сделали следующий шаг. Они использовали первые появившиеся компьютеры, чтобы математически обработать генетические различия между популяциями и перевести их в главные компоненты, которые отражают обычные корреляции среди частот генов.



Анализ главных компонент генетических вариаций по Еропе, по Cavalli-Sforza 1997 Proc. Nat. Acad. Sci USA.

С этим подходом они смогли не только показать, как изменяются частоты генов на карте, но и показать, как изменяются обычные корреляции частот многих генов. В представлении Кавалли-Сфорца эти карты показывают исторические силы, которые создали разнообразие человечества. Градиент по всем классическим маркерам может показать возможные пути миграций в прошлом. К сожалению, карты не могут показать, как и почему эти миграции совершались.

«Результаты анализа главных компонент выглядели очень интересными, но не было ничего, с чем можно было бы их сравнить, поскольку вопросы, на которые они могли бы ответить, никогда не были заданы».

Он решил искать новые источники данных, которые могли бы придать больший смысл генетическим расстояниям. Вместе с археологом А.Дж. Аммерманом Кавалли-Сфорца обратил свое внимание на неолит. Это была эпоха больших изменений: время, когда земледелие стало распространяться с Ближнего Востока в Европу. Если выбрать из событий прошлого такие, которые достаточно сильно для того, чтобы сформировать частоты генов по всей Европе, то неолит, похоже, был самым сильным из всех.

К началу 1970-х европейская археология вступила в эпоху радиоуглеродной революции в датировках. Этим методом были датированы признаки неолитического перехода в различных регионах Европы, от Греции до Ирландии. И они нарисовали удивительную картину: оказалось, что неолит распространялся по Европе медленно, со скоростью около одного километра в год, от юго-востока на северо-запад.

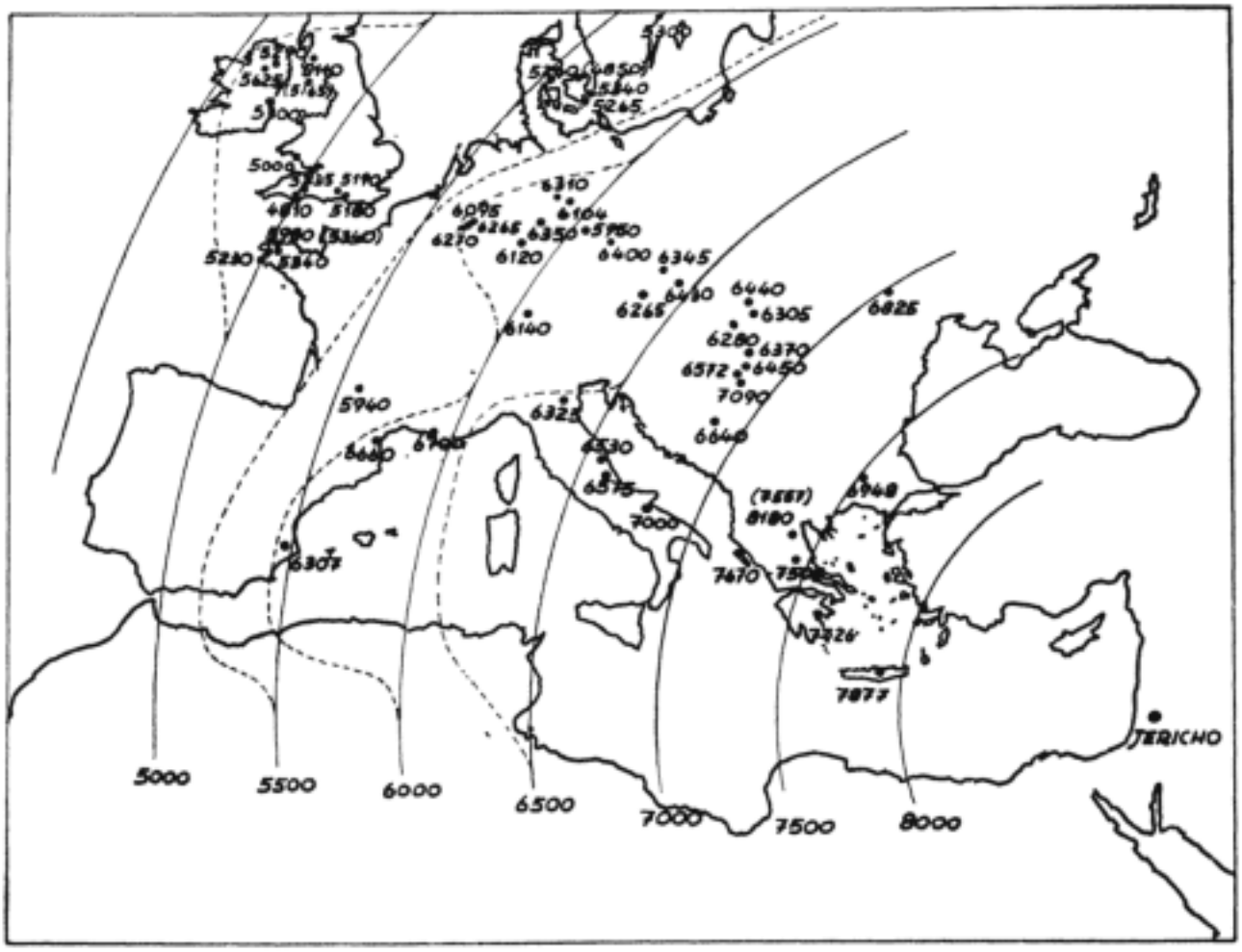


FIGURE 6. Map showing the spread of early farming in Europe. Dates are in years B.P. (see table 1). Arcs indicate the expected position of the spread at 500 year intervals based on the analysis shown in fig. 2. The broken curved lines represent an attempt to take into account some regional variation in the rate of spread.

Карта, показывающая распространение земледелия в Европе. Датировки даны в годах от настоящего времени. Дуги указывают на ожидаемую позицию распространения земледелия по интервалам в 500 лет. Пунктирные линии обозначают попытки связать некоторые региональные вариации с этой скоростью.

Кавалли-Сфорца эта картина привела к такому выводу: не было толп мигрирующих земледельцев, колонизировавших Европу. Вместо этого, земледелие распространялось пос тепенно, по мере того, как популяции росли в численности и передавали свои навыки населению соседних регионов. Но это была не просто «диффузия идей», скорее, это была диффузия культуры совместно с генами. Аммерман и Кавалли-Сфорца назвали этот процесс демической диффузией.

«Эволюционные процессы в основных чертах сходны, независимо от того, какой объект эволюционирует».

В течение 1980-х и 1990-х демическая диффузия стала доминирующей моделью демографических изменений в неолите. Идея подводила основу под теорию Колина Ренфрю, что индоевропейские языки также распространялись в Европе в неолите, замещая языки, бытовавшие с мезолита и более ранних времен. Было похоже, что процессы культурных изменений и распространения культур должны быть связаны с ростом численности групп людей и их продвижением, но для убедительности генетикам надо было заполнить брешь в имеющихся данных. Кавалли-Сфорца работал все больше, чтобы понять, как были связаны культурные и биологические изменения, установив долговременное сотрудничество с Марком

Фельдманом, специалистом по эволюции культур.

Сегодня из данных по древней ДНК мы знаем, что детали этой картины неолита от Кавалли-Сфорца были неверны. Распространение земледелия было важно, но не только оно. Более поздние миграции людей, некоторые из них очень быстрые, изменили генетический ландшафт Европы. Сегодня специалисты склоняются к тому, что индоевропейские языки завоевали Европу в течение бронзового века, и что ранние земледельцы были генетически более всего похожи на современных жителей Сардинии, лигнвистического изолята, который Кавалли-Сфорца хорошо знал.

Эти факты, появившиеся из анализа древней ДНК, стали для большинства генетиков сюрпризом. Синтетический взгляд, продвигаемый Кавалли-Сфорца, был очень притягателен, связывая экономические, демографические и генетические изменения, но для пересмотра взглядов нужны были новые данные.

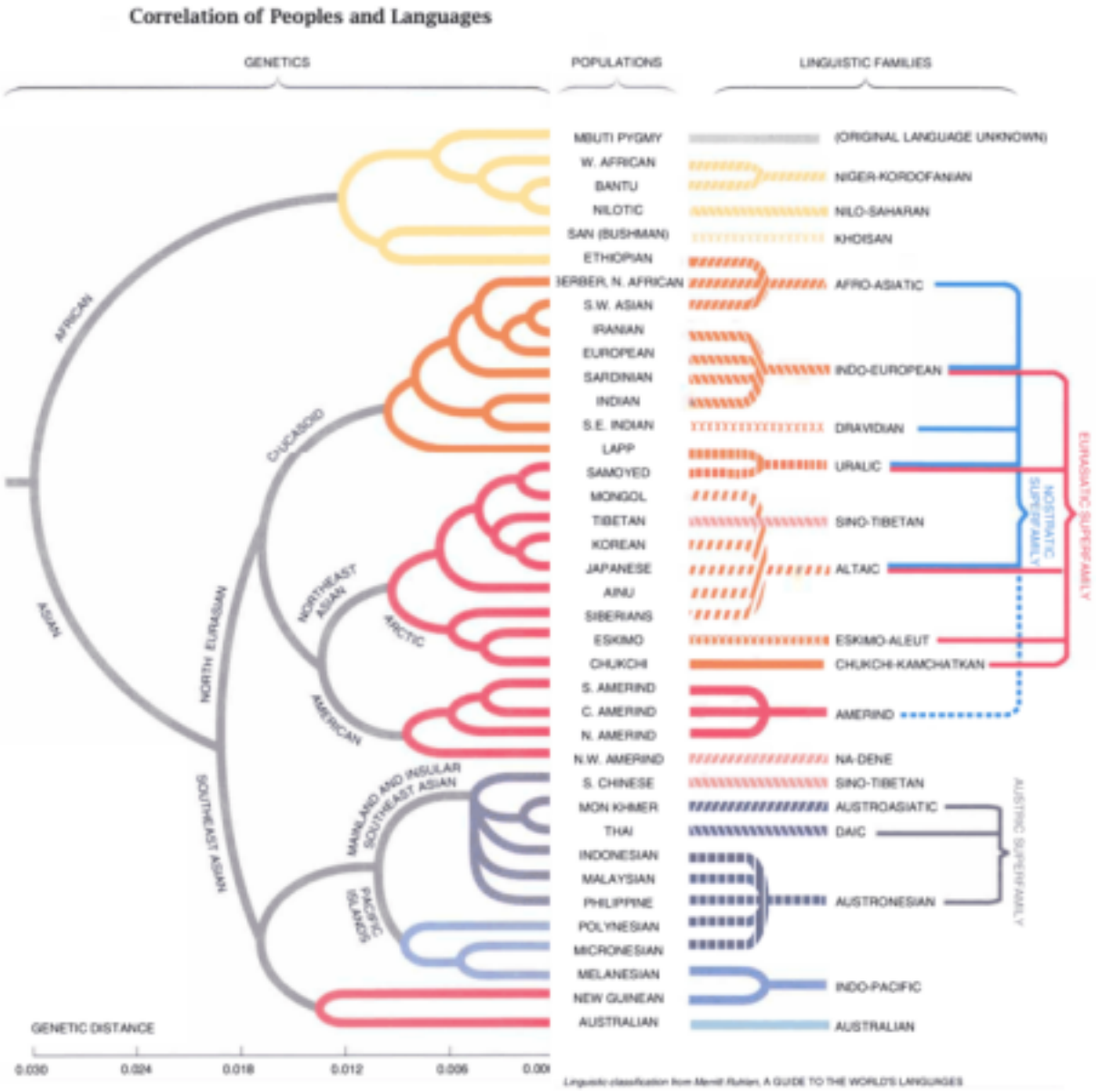
* * *

В течение 1980-х Кавалли-Сфорца работал, чтобы синтезировать растущую базу данных по генетическому разнообразию по всему миру. Он сделал шаг в сторону сенсации вокруг митохондриальной Евы, указав, что, в то время как новые подходы, основанные на секвенировании ДНК, могут проследить историю одного гена с большой точностью, они не способны комбинировать информацию от многих генов. Он стал работать над строительством дерева, которое бы могло суммировать родственные связи между целыми популяциями, а не отдельными индивидами.

Чтобы сделать это, он вступил в коллаборацию с группой лингвистов, которые верили, что они могут построить универсальное дерево человеческих языков. Интеллектуальным лидером этой группы лингвистов был Джозеф Гринберг. Он отрицал подход, которому следовали большинство лингвистов, работая над историей изменений звуков и формирования грамматических конструкций. Гринберг вместо этого использовал намного более простой подход – сравнение общих слов. С его точки зрения этот подход давал возможность определения более глубоких и более древних связей между языками. Так, он предположил, что все языки американских индейцев могут быть сгруппированы в три больших семьи вместо десятка, как считали остальные лингвисты.

Коллега Гринберга Меррит Рулен сделал шаг дальше, предположив, что каждый язык в мире может быть помещен на ветвь общего дерева. Эта идея резонировала со взглядами Кавалли-Сфорца. Если люди действительно произошли от маленькой популяции в течение последних сотен тысяч лет, они должны все происходить от популяции, которая говорила на одном языке. Так же как дерево генов связывает все популяции, так и дерево языков.

И если Кавалли-Сфорца был прав относительно культурных и генетических изменений, эти деревья языков и генов должны совпасть. Все, что осталось, это вырастить их.



Дерево генов и языков. По Cavalli-Sforza 1991, Scientific American.

«Было очень интересно обнаружить, что мы подтвердили гипотезу, восходящую к идеям Чарльза Дарвина — если создать дерево генетической эволюции, это позволило бы проследить эволюцию языков».

Если следовать этому подходу, люди, которые говорят на сходных языках, должны иметь больше генетического сходства, чем люди, которые говорят на языках разных языковых семей. Но из этой закономерности есть много исключений. Некоторые выраженные сходства между генетическим и лингвистическим деревьями могут просто отражать тот факт, что генетически сходные люди живут в непосредственной близости друг от друга и поэтому с большей вероятностью обмениваются языками. С другой стороны, генетическое сходство людей на севере и юге Индии, несмотря на языки разных языковых семей может отражать тысячелетний процесс смешивания между людьми разного происхождения.

Короче говоря, язык и гены отражают ***разную*** историю. Каждый из них подвергается влиянию своих собственных эволюционных сил. Для языка характерен горизонтальный перенос, когда люди перенимают новые языки и под влиянием соседей изменяют свой собственный. Но выявить связь между генами и языками — чрезвычайно соблазнительно. Занимаясь этим, Кавалли-Сфорца разработал каталог частот генов и строил все более и более синтетические карты. Он писал популярные книги и статьи с продвижением своей идеи. Между тем, традиционные лингвисты продолжали отвергать идею о существовании единого «праязыка» для всего человечества, а археологи разделились во взглядах на человеческие миграции.

Вместе с Паоло Меноцци и Альберто Пьяцца в 1994 году Кавалли-Сфорца опубликовал большой труд «История и география

человеческих генов». Эта книга – такая толстая, как телефонный справочник Бостона – включала огромное число графиков и целые страницы таблиц первичных данных о частотах генов. Она могла служить не только для суммирования и обобщения данных в этой области науки, но и как источник для дальнейших исследований.

В то же время к середине 1990-х генетика человека уверенно продвигалась от классических маркеров к микросателлитам и секвенированию ДНК. Между тем, интернет сделал возможным распределенные базы данных, и это означало, что монография с таблицами частот генов уже не была необходимой для новых исследований. Тем не менее, эта книга послужила масштабным утверждением эры генетики человека, и большинство ученых в этой области в середине 1990-х рассматривали ее как чрезвычайно ценный источник знаний.

* * *

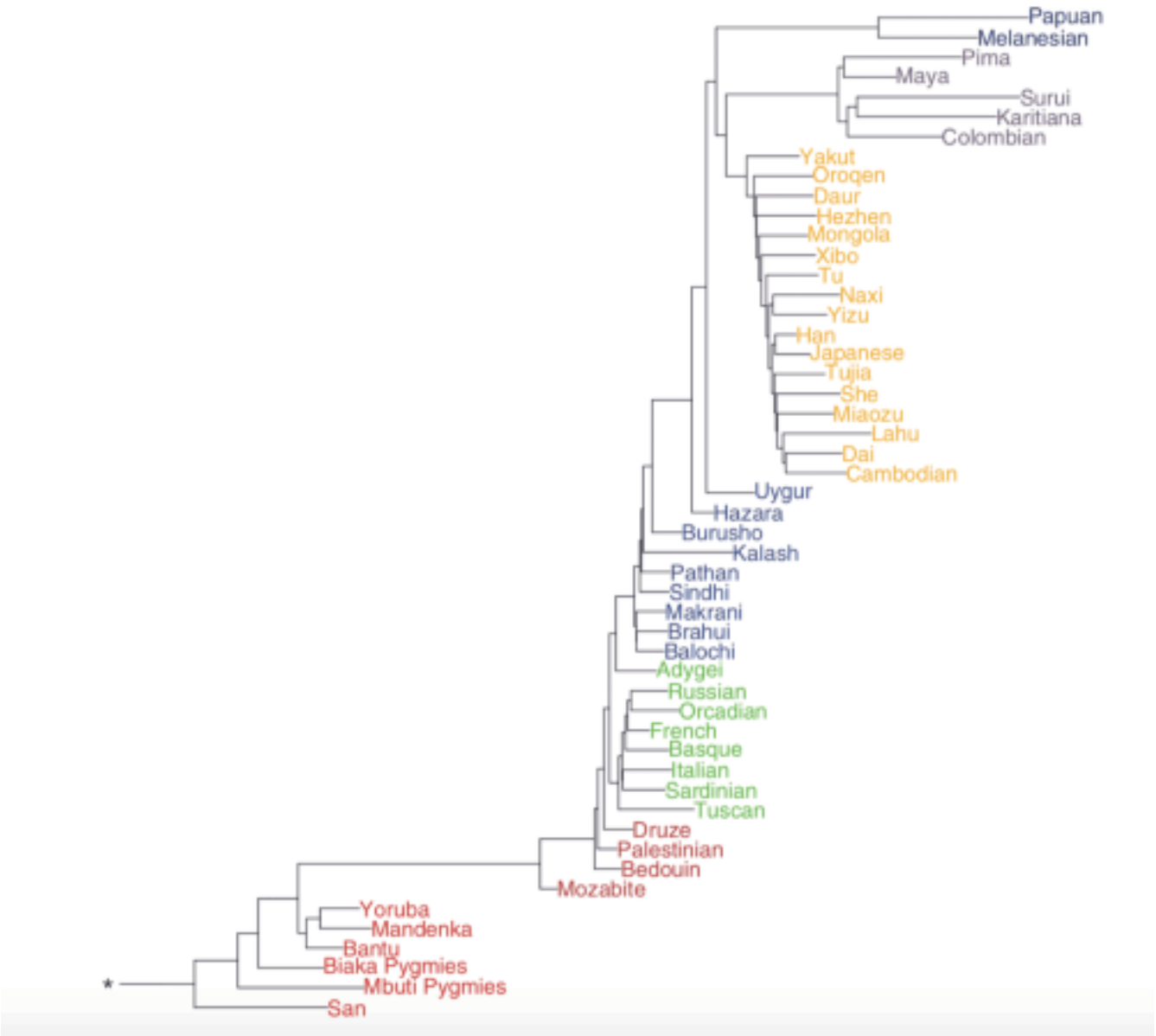
Я впервые встретил Кавалли-Сфорца на симпозиуме по эволюции человека, который он организовал в лаборатории Колд Спринг Харбор в 1997 году. Тогда ему было уже около 75 лет, его внушительная седовласая фигура заметно доминировала над окружающими. Мир изменился – на этом симпозиуме я впервые увидел демонстрацию технологии чипов для параллельного генотипирования десятков локусов. Но Кавалли-Сфорца все равно оставался в центре внимания, его лаборатория была на переднем фронте применения новых технологий.

В это время он организовывал поддержку своей инициативы собрать образцы ДНК малых коренных популяций по всему миру. “Human Genome Diversity Project”, начавшийся в Стэнфорде в начале 1990-х, ставил перед собой задачу собрать генетические образцы от более чем 1000 человек из малых популяций по всему миру. Ученые должны были использовать новую технологию, чтобы трансформировать эти образцы в бессмертные клеточные линии, которые распределить по лабораториям для исследования.

Кавалли-Сфорца и другие видели Human Genome Diversity Project (HGDP) как необходимое доолнение к Human Genome Project, который к тому времени потратил 2,7 млрд долларов на создание чернового варианта последовательностей всех 23 пар хромосом человека. Human Genome Project включал общее представление о геномных вариациях – в разных лабораториях исследовались геномы индивидов европейского, африканского и азиатского происхождения, итоговый черновик генома представлял собой мозаику из геномов разного происхождения. Но в эту схему не были включены сотни популяций по всему миру, и она не охватывала всю широту геномного разнообразия. Кавалли-Сфорца и другие лидеры проекта HGDP ставили задачу собрать образцы ДНК из лингвистических изолятов, надеясь обнаружить их уникальные генетические вариации.

Некоторые исследователи считали HGDP дорогим и нецелесообразным проектом или хуже, наносящим вред сбору ДНК коренных народов для медицинских целей. Спонсоры проекта, которые вначале поддержали идею, решили, что больше не могут его финансировать. Национальный институт здоровья продвигал идею сбора базы данных генетических вариаций, сначала в рамках проекта International HapMap, а потом – в рамках проекта 1000 Genomes Project, но эти проекты фокусировались на больших популяциях, но не на малых популяциях коренных народов.

HGDP включал коллаборацию по всему миру в области сбора образцов и клеточных линий. Centre d’Etude du Polymorphisme Humain-Fondation Jean Dausset (CEPH) в Париже сделал шаг к исследователям в направлении сохранения и распределения клеточных линий, и, в конце концов, сотни лабораторий применили данные для решения проблем происхождения и разнообразия человечества.



Филогенетическое дерево популяций, основанное на данных HGDP-CEPH, по статье Li et al., 2008, Science.

Взяв за основу аналитический подход, разработанный Кавалли-Сфорца, специалисты построили дерево родственных связей между популяциями – и теперь уже правильно поместив африканские популяции у основания дерева, а все неафриканские популяции – на единой ветви, отражающей их общую историю. Они вычислили главные компоненты из данных по частоте генов, получив картину, показывающую разнообразие человечества в пространстве по сотням тысяч генетических маркеров. Тесная корреляция между генетическими вариациями и географией признана следствием генетического дрейфа, отбора и миграций, происходящих в истории популяций.

В течение 1970-х, 1980-х и 1990-х возникали новые подходы к исследованию вариаций ДНК, оставив в прошлом классические маркеры. Но примечательно, что никто принципиально не улучшил основной аналитический подход, введенный Кавалли-Сфорца и Эдвардсом, до тех пор, пока не началась эра полногеномных исследований в начале 2000-х.

Последние 15 лет бурно развивались технологии исследований, которые привели к возрождению популяционной генетической теории и к новому пониманию эволюции человечества. Новые статистические подходы, такие как STRUCTURE и D4, дали возможность генетикам совершить прорыв от бинарных деревьев с отдельными ветвями к деревьям, которые включают связи между разными ветвями и циклы. В последнее десятилетие Кавалли-Сфорца был в курсе этих новых подходов, правда иногда он отмечал, что кажущееся новым прорывом может оказаться и тупиком. В интервью 2011 года он высказал свое восхищение методом STRUCTURE и в то же время сформулировал свои мысли о том, как эффективно использовать генетическую информацию для целей медицины. В то же время, он иронизировал в адрес ставшего широко известным Джона Новембре, который находил слабые стороны классического подхода главных компонентов.

Я был особенно поражен комментарием, который показал, что Кавалли-Сфорца думал о моей собственной работе, а также о

многих других, которые демонстрировали масштабы человеческой эволюции в недавнем прошлом.

«Меня особенно интересует один вопрос: возможно, может быть, даже вероятно, что важная часть генетической медицинской изменчивости, которую мы нашли у нашего вида, возникла в течение последних 10 тысяч лет. Такая вариабельность/адаптация может быть результатом интенсивного культурного развития, которое привело к земледельческо-скотоводческой экономике и образу жизни, который вносит большие различия в жизнь разных слоев человечества по всему миру».

В начале моей карьеры я во многом не соглашался со взглядами Кавалли-Сфорца. Его представления о происхождении человечества не укладывались в границы статистических моделей и данных, доступных в то время. Он отбрасывал возражения, не отвечая на них. Так, он отвергал даже казавшиеся разумными возражения против проекта HGDP.

Сейчас я вижу гораздо яснее, как генетики человека в 1990-х находились на перепутье, нуждаясь в том, чтобы включать эволюцию в свои модели генетической изменчивости, но при этом были неспособны справиться с наблюдениями антропологов. Перечитывая работы Кавалли-Сфорца, я вижу, насколько прозорливым он был, и не только в предугадывании будущих успехов, но и в том, что какие-то методы не оправдали себя.

Сегодня многие генетики в своей работе основываются на открытиях Кавалли-Сфорца – комбинируют базу данных генетических вариаций в прошлом и настоящем, строят деревья с все большим разрешением, взаимодействуют с данным археологии. Некоторые даже пытаются повторить его работы в полногеномном масштабе, мечтая создать атлас вариаций человеческого генома. Это монументальная задача, и, как ярко показала книга Кавалли-Сфорца, возможно, уже в момент публикации этот труд может устареть. Выход в том, чтобы создавать не монументальный том, но разрабатывать метод, идеи, связывающие различные виды данных о нашем прошлом, для понимания, откуда мы пришли и как мы связаны друг с другом.

* * *

Почитать больше:

Кавалли-Сфорца уделял большое внимание публицистике и популяризации, так что некоторые из его трудов доступны для непрофессионального читателя. Я особенно рекомендую Гены, люди и языки и Великое переселение человечества, написанную им вместе с сыном Франческо.

Абстракт его книги был опубликован в журнале *Scientific American в 1991 году под заглавием* “Genes, Peoples, and Languages,” и оттуда же взято приведенное выше дерево генов и языков.

Я также рекомендую прочитать интервью 2010 года Кавалли-Сфорца и Франца Манни, опубликованное в журнале *Human Biology.* “Interview with Luigi Luca Cavalli-Sforza: Past Research and Directions for Future”

А.В.Ф.Эдвардс в 2009 году опубликовал замечательные воспоминания об истории своей работы по статистике с Кавалли-Сфорца “Statistical methods for evolutionary trees,” в которой описал историю их сотрудничества и противостояния с Мотоо Кимура.

Джон Хокс (John Hawks)

Палеоантрополог, Университет Висконсин-Мэдисон

Автор популярного блога по палеоантропологии

перевод Надежды Маркиной

Источник публикации статьи: https://medium.com/@johnhawks/the-man-who-tried-to-catalog-humanity-b433c3f31872