

УДК 9.903.07

А.А. Клёсов

Академия ДНК-генеалогии, г. Москва, Россия и г. Ньютон, шт. Массачусетс, США

КАК ГЕНОМНЫЙ АНАЛИЗ «ВЫЯВЛЯЕТ ЛИЧНОЕ ПРОИСХОЖДЕНИЕ»

Опубликовано в электронном журнале «Перформат» 3 апреля 2018 г. (<http://pereformat.ru/klyosov/>). Печатается с разрешения автора (<http://pereformat.ru/2018/04/genomnyj-analiz/>).

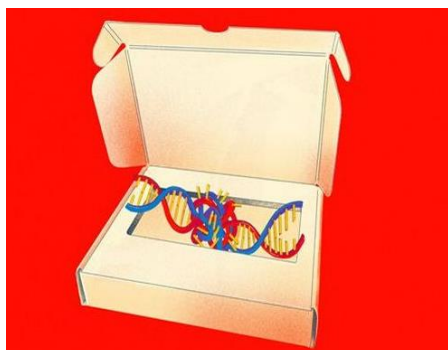
Клёсов Анатолий Алексеевич - доктор химических наук, профессор, советский и американский биохимик, специалист в области полимерных композиционных материалов, биомедицины, ферментативного катализа, лауреат премии Ленинского комсомола (1978) и Государственной премии СССР по науке и технике (1984). Профессор биохимии Гарвардского университета, живёт в Ньютоне (шт. Массачусетс, США). E-mail: aklyosov@comcast.net.



Klyosov Anatoliy Alexeyevich - Doctor of chemical sciences, Professor, biological chemist, expert in the field of polymer composite materials, biomedicine, enzymatic catalysis, Professor of Biochemistry of Harvard University, lives in Newton (Massachusetts, USA); e-mail: aklyosov@comcast.net.

Статья познавательная, с «развлекательной» компонентой, в том смысле, что не засушивает материал. Она о том, как массово дурачат людей с помощью «геномного анализа происхождения», и никакой ответственности, в отличие от Мавроди с его пирамидами, для которого веревочка недолго вилась. А здесь – принципиально новый способ «научного мошенничества». А.А. Клёсов.

На днях в англоязычном сетевом ресурсе «Генетика» вышла статья под названием «**Как геномный тест исковеркал мою семейную родословную, и, вероятно, вашу тоже**». Автор, проживающая в США, описывает, как разослала образец своей ДНК в несколько коммерческих тестирующих «геномных» организаций и что из этого вышло. Об этой истории мы расскажем ниже. Для меня, например, эта история совсем не показалась удивительной. Уже немало лет [я пишу о том же самом на «Перформате»](#) и рассказываю, насколько фантазийными являются «геномные истории», как личные, так и этнические, популяционные, или рассказы о том, кто куда направлялся в древние времена, то есть истории миграционные.

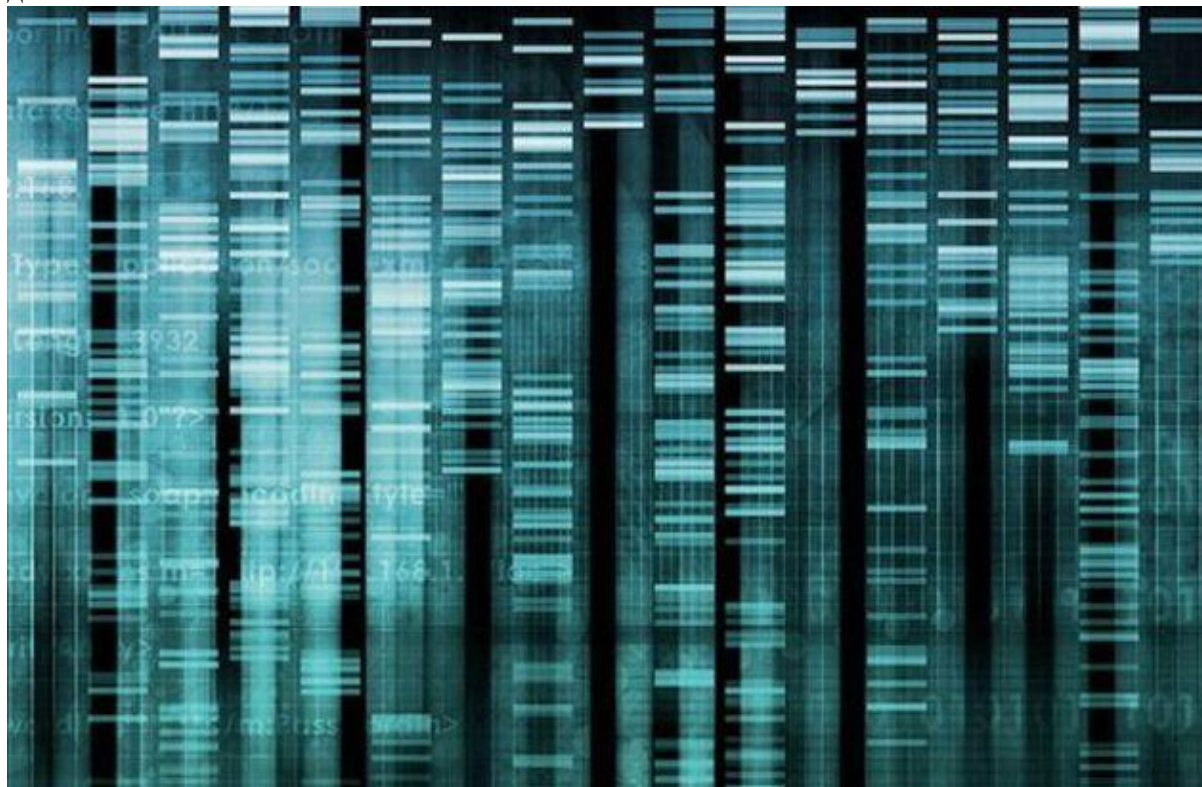


Художник Chelsea Beck/GMG.

Сама методология этих охотнич..., пардон, геномных рассказов является порочной, а если совсем мягко, то фантазийной. Эту методологию следует подразделить на три части, и мои нелестные слова выше относятся в основном к последней, третьей части, именно она является фантазийной, но заказчик полу-

чает только выводы из этой третьей, беллетристической, художественной части. Напомню, что художественная литература в английском языке называется fiction, то есть фикция. Про первую, научную часть, и про вторую, компьютерную часть, заказчик никаких сведений не получает и о них, как правило, понятия не имеет.

Итак, что же представляет собой первая часть геномных исследований, направленных на выяснение истории древнего мира, или персонального происхождения того или иного заказчика, который сдал слюну на ДНК (геномный) анализ, оплатил исследование и получил результаты? К первой части особых претензий нет, там задействовано сложное и дорогостоящее оборудование, там применяют дорогостоящие и хитроумные реактивы, в которых сложные молекулы помечены флюоресцентными метками, там «молекулярными ножничками» разрезают молекулы ДНК, состоящие из десятков и сотен миллионов химических соединений, или звеньев, или нуклеотидов (все три понятия здесь – в данном смысле синонимы), а если исследование «полногеномное», то и из миллиардов звеньев, и полученные фрагменты ДНК, большие и маленькие, разгоняют на «фракции», примерно как в самогонном аппарате. Только в самогонном аппарате главная задача - выделить одну, самую желанную фракцию, а в геномном анализе ДНК таких фракций астрономическое количество, и все желанные. Вот – пример картины таких фракций, каждая из которых (светлые полосы) представляет собой цепочки нуклеотидов.



Эти цепочки нуклеотидов отличаются друг от друга протяженностью, нуклеотидным составом (они все состоят из набора всего четырех нуклеотидов), последовательностью нуклеотидов в цепочках, а также мутациями, когда при копировании этих цепочек, которое выполняется специализированными ферментными комплексами и прочими «ассистирующими» биологическими соединениями, происходят случайные замены одного нуклеотида на другой, или случайные выпадения нуклеотидов, и прочие сбои в биологической машине копирования и исправления допущенных ошибок.

До картинки, приведенной выше, никаких претензий ко всем описанным последовательностям лабораторных операций нет. Естественно, техника совершенствуется, разрешение таких картинок улучшается, «молекулярные ножнички» становятся все

лучше и лучше, и так далее, но так развивается наука. Нельзя в науке критиковать исследования только потому, что в будущем будет еще лучше. Можно, впрочем, критиковать тогда, когда лучше стало уже сейчас, а авторы все еще используют устаревшую технологию, и получают «морально устаревшие» сведения, когда их коллеги уже ушли далеко вперед. И тем более надо непременно критиковать, когда авторы, понимая, что их сведения далеко не полные, а то и с хорошей вероятностью ошибочные, когда авторы выхватывают из многих возможных интерпретаций только одну, потому что им так хочется по некоторым причинам, далеким от науки, но при этом объявляют, что их выводы продвигают науку и являются единственно верными. Часто авторы это не объявляют в прямом виде, но так подают свой материал и выводы, что якобы никаких других объяснений нет. А другие объяснения – «псевдонаучные». Это – прямое указание на то, что авторы – жулики от науки. Начинаешь разбираться – а там других объяснений масса, но авторы предпочли их не упоминать. Это что же, получается, что у самих авторов «псевдонаука»? Нет, это у них просто не наука. Наука – одна, либо она есть, либо ее нет. Как и осетрины «второй свежести».

Повторим еще раз – до картинки, приведенной выше, никаких претензий ко всем описанным последовательностям лабораторных операций нет. Как правило, это – хороший научный уровень. Правда, здесь – натяжка. Уровень-то хороший, а вот науки уже давно нет. Эту работу делают техники-лаборанты, пусть и на сложном оборудовании, и обрабатывают компьютеры по заданным программам. Полета мысли, так сказать, что и есть неотъемлемая часть науки, там нет.

Нет полета мысли и во второй части исследования, поскольку она выполняется компьютером по прописям, которые в тот компьютер закладываются. Главная задача этой, второй части – рассортировать фрагменты ДНК от разных людей так, чтобы отличить закономерные сходства фрагментов от случайных. Закономерные сходства наблюдаются тогда, когда фрагменты ДНК этих людей происходят от недавнего общего предка, где понятие «недавний» определяется поставленной задачей. Такие совпадения называют «идентичные по происхождению» (ИПП). В других случаях совпадения являются случайными по статистическим причинам. Такие совпадения называют «идентичные по сходству» (ИПС). Можно было бы назвать «идентичными по факту», или «одинаковыми благодаря статистике», то есть ОБС, извините за ироническую игру слов, имеющих особое значение в русском языке. В категорию ИПС попадают также фрагменты ДНК, которые сходны не случайно, а отвечают за важные функции жизнедеятельности человека, которые сформировались и устоялись в результате эволюции человека. В категорию ИПС попадают также фрагменты ДНК, которые веками циркулируют в рамках определенного этноса, особенно если этнос относительно замкнутый (как, например, в еврейских сообществах прошлого).

Суть геномного анализа состоит в том, чтобы сопоставить ДНК от разных людей (если речь о геномном анализе человека), при том, что ДНК людей похожи буквально на 99 %, хотя при этом ДНК испещрены мутациями, о которых шла речь выше, то есть заменами одних нуклеотидов на другие, и таких замен в ДНК миллионы. Часть из них обратимые, часть – необратимые. В итоге сравниваются доли процентов различий на фоне более чем 99 % сходства. Но эти сравнения невозможно проводить, пользуясь целыми, интактными молекулами ДНК, они производятся после расщепления ДНК на фрагменты разного размера, как показано на фото выше. Поэтому во второй части исследования компьютер ищет «матчи», то есть совпадения. «Матчи», как пояснено выше, имеют разную природу. Одни похожи потому, что происходят от одного, сравнительно недавнего общего предка (ИПП), другие похожи по другим причинам (ИПС), не имеющим никакого отношения к общему происхождению. Или имеющим лишь косвенное отношение, например, потому что отражают близкое этническое происхождение, но по комбинации происхождения по мужской и женской линии, где общего предка нет в принципе.

В итоге компьютер, подчиняясь определенным программам, которых в интерпретациях ДНК используют множество, сортирует фрагменты по разным критериям одновременно – по числу снипов (определенных и заранее классифицированных мутациях в ДНК), по размерам фрагментов, по степени сходства, и ни в одном варианте нет однозначных критериев. Фрагменты бывают «крупные», «менее крупные», «более средние», «менее средние», «умеренно малые», «малые», а компьютер сам решает, кто кому родственник «по наследственности», а кто «просто похож». Попгенетики это не проверяют, «санитар сказал – в морг, значит, в морг».

Всю эту кашу компьютер сортирует и делает выводы, которые по понятным причинам могут быть совершенно далекими от реальности или частично пересекаться с реальностью, причем степень этой «частичности» может варьироваться в любых пределах. Понятно, что нужны независимые критерии для перекрестной проверки выводов, но популяционные генетики так не работают. Обычно, что получилось, то и получилось. Наиболее очевидным независимым критерием была бы гаплогруппа и субклады, которые тут же вводили бы запреты на происхождение ДНК одного субклада из другого, но популяционные генетики так тоже не работают. У них критерии простые – совпали относительно протяженные сегменты (фрагменты) ДНК – значит, вероятность ИПП выше, значит, общий предок жил не так давно. Совпадают только малые сегменты – значит, преобладает ИПС. Замечаете «качественный» (не количественный) принцип анализа данных? Размеры сегментов измеряют в сантиморганах (сМ), где один сантиморган равен примерно миллиону пар нуклеотидов, или такому расстоянию, на котором степень рекомбинации генов равна примерно 1 %. Есть и ориентировочные формулы, согласно которым если в двух ДНК в среднем совпадают сегменты размером в 11 сМ, то там на 99 % ИПП; если совпадают сегменты в 9 сМ, то вероятности ИПП и ИПС – 50 % на 50 %, а если совпадают только сегменты в 4 сМ, то там на 99 % ИПС (случайные совпадения), и всего на 1 % ИПП.

Понятно, что эти критерии весьма условны, как весьма условны и выводы. И вот с данными сортировки, проведенными компьютером, поскольку вручную такую сортировку делать невозможно, там размеры файлов в гигабайты и терабайты, популяционные генетики включают свой «полет мысли», что должно бы означать начало научной стадии в исследовании. Это и есть третья часть исследования – интерпретационная, превращение тех самых ИПП, ИПС и прочих «сходств и различий» в мутационной картине ДНК из разных источников (конкретных людей или усредненных групп людей, или целых популяций) в выводы – кто от кого, где и когда произошел, куда и когда передвигались группы людей в древности, и так далее, что захотят включить попгенетики в свои «интерпретации».

И вот здесь начинаются настоящие проблемы. Мало того, что критерии условны, как описывалось выше, они часто приводят к полной каше, а выводы надо делать. Вообще говоря, «надо делать» – такого не бывает, никто не заставляет, но выводы делать хочется, потому что надо убеждать мир, что наука на марше. Что клиент должен платить деньги. Если быть честными, то надо показывать, что то, что компьютер выдал, можно объяснить так, а можно и этак. А можно и по-другому, и многими вариантами. Но поступать честно – это признать, что методология еще совсем недоработана. А хочется публикаций, грантов, научного признания, денег. Причем, хочется, чтобы это было быстро. «Мы хотим сегодня, мы хотим сейчас». Итог – как правило, «осетрина второй свежести». Причем, это часто получается «карасем третьей свежести», но в статью и в заявки на гранты идет, несомненно, утверждение, что это свежайшая осетрина. Царская рыба, то есть соответствующее исследование такой же категории.

Остается только удивляться, что иногда выводы разумны, но «разумность» здесь «по понятиям», проверить эту разумность, как правило, не получается. А когда приме-

няется проверка с гаплогруппами-субкладами, вся «разумность», как правило, рассыпается. Поэтому популяционники такой проверки не делают, ни к чему это, а то рассыпется, а это в их планы не входит. Но еще чаще бывает, что компьютер в целом описал ситуацию правильно, например, что геномы древнего байкальского мальчика (с датировкой 24 тысячи лет назад) и современных американских индейцев совпадают на 14-38 % (так – в исходной публикации), и нам (но не авторам статьи) ясно, что это потому, что и у тех, и у других один общий предок – гаплогруппа R (этого в исходной публикации не было). Но популяционники делают свою интерпретацию – что популяция того байкальского мальчика перешла в Америку и стала предком американских индейцев. Это в корне неверно, но в арсенале популяционники слова «неверно» нет. Есть «так получилось, значит, так оно и есть».

Подобные же абсурдные интерпретации делались во множестве случаев. Поскольку авторы соответствующих работ прекрасно понимают, что все их интерпретации построены на песке, то они излагают свои результаты и выводы в крайне вязком виде, практически никогда не напрямую, но в широкую прессу, в СМИ это уже идет в виде, вполне категоричном. Ясно, что подобные «выводы» откровенно дискредитируют науку. Более того, при подобном «анализе» авторы обычно «подтверждают» то, что уже устоялось в науке, даже когда устоялось неверно. Например, что ямная культура (гаплогруппа R1b) принесла индоевропейские языки в Европу. Или что срубная культура (R1a) якобы вышла из ямной (R1b), или что из ямной якобы вышла культура шнуровой керамики (R1a). Это противоречит многим наблюдениям и фактам, но на Западе устоялось (русские историки и лингвисты относятся к этому, как правило, скептически или отрицательно). А раз устоялось, то генетические исследования «подтвердили». А подтвердили, потому что современная популяционная генетика ничего не оспаривает, если что известно или принято – она всегда подтверждает, только чтобы не подумали, что у них что-то не так. Многие примеры такого «подхода» популяционников приведены на «Переформате», не будем повторять. Там же многократно повторялся вывод – что данные «геномного анализа» в популяционной генетике надо непременно проверять перекрестно. Если это пока невозможно, стоит воздержаться от слепого их принятия. Наиболее надежная проверка – с помощью подходов ДНК-генеалогии, сопоставлением гаплогрупп, субкладов, гаплотипов, расчетов датировок.

Поскольку геномные популяционники такой перекрестной проверки не делают, и потому их результаты и выводы часто не являются убедительными для специалистов, то обстановка постепенно накаляется. Первый звоночек – две редакционные статьи, вышедшие 29 марта 2018 года в журнале Nature (стр. 559 и 573-576). Первая имеет характерное название – «Use and abuse of ancient DNA», то есть примерно «Древние ДНК в пользу и во вред», вторая – «Битва за общую территорию». Битва – это между археологами-историками, с одной стороны, и генетиками, с другой. Авторы осторожно пишут, что генетики «упрощают», на самом же деле проблема не в упрощении, генетики слишком часто **искажают**. Авторы критических статей в Nature опять приводят в пример якобы образование культуры шнуровой керамики из ямной по данным генетиков, сообщая, что археологи «в шоке». Мы об этом на «Переформате» уже три года пишем, а на Западе только подходят к пониманию проблемы с «геномными популяционниками». Но до понимания того, что дело не в «упрощении», а в фундаментальной проблеме того, как геномные данные анализируются и интерпретируются, на Западе еще, видимо, далеко. Но уже пишут о «disruptive influence», то есть разрушительном влиянии, которое генетики оказывают на изучение истории человека, о том, что это «irritating», то есть действует раздражающе. При этом речь идет не о прямых данных, данные есть данные, а о безответственных интерпретациях генетиков. О чем мы и пишем последние несколько лет.

Переходим теперь к «геномному анализу» родословных человека, как правило, заказчиков, направляющих образцы своей ДНК в коммерческую лабораторию. Все те

проблемы, описанные выше, здесь присутствуют, но поскольку это не академические «геномные» исследования, которые популяристы обычно пишут вязким языком, и ответа по сути не дают (см. многие примеры того на «Переформате»), то с персональными, частными, оплачиваемыми заказами так дело не пойдет. Там нужно давать выводы рубленным, конкретным языком. Если и включать типа «вероятно», то так, чтобы заказчик всё равно поверил – «вы похожи на...», и тут же «вы очень похожи на...». Проверить-то никто все равно не сможет, где там жил прапрадед многие тысячи лет назад. Что написали, то и хорошо. Более того, заказчик сам часто подыгрывает тому, что ему выдали. Написали ему, что его корни в Италии, заказчик хлопает себя по лбу – так вот почему я пиццу люблю! Вообще вера «в науку» безгранична, и этим популяристы широко пользуются.

Мне прислали письмо, и дама, автор письма, была в восторге, как всё чудесно совпадает с ее родословной, так все здорово! Нашли, что у нее полпроцента полинезийской ДНК. И она пишет, что подтвердилась семейная легенда, что ее прадедушка был моряком и ходил в Полинезию. То, что в лучшем (для прадедушки) случае это его ДНК должна была остаться в Полинезии, а не он привез бы в Россию (что привез-то??), ей в голову не пришло. Вера в торжество науки всё затмила, включая здравый смысл. Я спросил даму – а что, он жену из Полинезии привез? – Нет, конечно, рязанские мы.

Ну вот, мы уже вплотную подошли к сути интересной истории, которую поведала американка, и что мы упомянули выше. Для начала – сухие сведения. Ее прадед по сведениям был сирийцем, прабабка родилась в Пенсильвании, происхождение неизвестно, но по воспоминаниям родственников она говорила на арабском языке. Дед унаследовал от отца (или от обоих родителей) темные глаза и темные, жесткие волосы и женился на блондинке-американке норвежского происхождения. У них было двое детей – отец нашей героини, и его сестра, тетя нашей героини. Как мы видим, у них должно быть выражено сирийское и норвежское происхождение. Тетя отправила образец своей ДНК на анализ в американскую компанию **AncestryDNA**, и оттуда ей сообщили, что тетя на 16 % имеет ближневосточное происхождение, на 15 % имеет кавказское происхождение, и на 30 % она - греко-итальянка. То, что тетя любила макароны, было воспринято, как подтверждение отчасти итальянского происхождения, хотя тетя перелопатила всех своих родственников, но не нашла никого ни с итальянским происхождением, ни с греческим.

Это озадачило героиню нашего повествования, и она решила взять реконструкцию семейного генеалогического дерева в свои руки. Направила образец своей слюны в три компании – AncestryDNA, 23andMe, и National Geographic и была удивлена, получив разные ответы из каждой компании. Первой была AncestryDNA:

Скандинавия	32%
Южная Европа	13%
Ирландия/Шотландия/Уэльс	11%
Великобритания	11%
Ближний Восток	11%
Кавказ	10%

Героиня опять была несколько удивлена своими 11 % происхождения с Ближнего Востока, хотя у сестры ее матери, то есть тети, его нашли равным 16 %. По соображениям героини, у нее ожидалась бы половина тетиного ближневосточного вклада, но что есть, то есть. Греко-итальянское происхождение (у тети 30 %) в новом варианте описания стало, наверное, «южно-европейским», и оказалось 13 %. Тетино 15 %-ное «кавказское» происхождение, которое по соображениям героини, должно было уполовиниться при переходе к самой героине, оказалось выраженным несколько больше. Поскольку ни одна тестирующая компания не привела погрешности в полученных числах, этот вопрос остался невыясненным. Аналогично, 22 % генеалогии с Британских островов у героини осталось вообще невыясненным у тети.

Героиня опять была несколько удивлена своими 11 % происхождения с Ближнего Востока, хотя у сестры ее матери, то есть тети, его нашли равным 16 %. По соображениям героини, у нее ожидалась бы половина тетиного ближневосточного вклада, но что есть, то есть. Греко-итальянское происхождение (у тети 30 %) в новом варианте описания стало, наверное, «южно-европейским», и оказалось 13 %. Тетино 15 %-ное «кавказское» происхождение, которое по соображениям героини, должно было уполовиниться при переходе к самой героине, оказалось выраженным несколько больше. Поскольку ни одна тестирующая компания не привела погрешности в полученных числах, этот вопрос остался невыясненным. Аналогично, 22 % генеалогии с Британских островов у героини осталось вообще невыясненным у тети.

Наша героиня пришла к промежуточному заключению, что ситуация с «геномной генеалогией» у нее и тети остается запутанной, и обратилась со своим образцом ДНК в компанию **National Geographic**. Результаты:

Северо-западная Европа	28%
Юго-западная Европа	23%
Малая Азия	16%
Северо-восточная Европа	10%
Еврейская диаспора	9%
Восточная Европа	8%
Юго-западная Азия/Персидский залив	6%

Появилась Малая Азия с немалыми 16 % плюс Персидский залив, и что занятно, появилась еврейская генеалогия, с ее 9 %. Пришлось обратиться в компанию **23andMe**:

Скандинавия	3.1%
Британия и Ирландия	9.8%
Франция и Германия	7.8%
Остальная Северо-западная Европа	41.9%
Южная Европа	21.6%
Италия	8.4%
Остальная южная Европа	13.2%
Прочая Европа	9.3%
Ближний Восток	3.3%
Северная Африка	2.2%
Еврей-ашкенази	<0.1%
Без отнесения	0.9%

эти данные просто неверны. Зато она теперь только на 3,3 % с Ближнего Востока, несмотря на прадеда-сирийца, и на 2,2 % - северная африканка, и ни много ни мало на 62,6 % - «северо-западная европейка» (включая Центральную и Западную Европу). И никакой Восточной Европы, несмотря на тесты других компаний.

Наша героиня запросила базовый файл данных из 23andMe и отправила его в компанию **GenCove**, которую недавно основали генетики. Хотя они прокрутили те же самые данные, что и 23andMe, они сообщили, что героиня на 8 % происходит из Индостана. У 23andMe ничего подобного не было. В целом данные такие:

Юго-западная Европа	40%
Северная и Центральная Европа	20%
Восточное Средиземноморье	16%
Центральная Индия	8%
Ближний Восток	5%
Анатолия, Кавказ, Иранское плато	2%

Вывод, который сделала наша героиня, логичен и очевиден: четыре «геномные» компании – четыре разных результата, в том числе и те, которые принципиально противоречат ее семейной родословной, в которой у нее, как она пишет, нет оснований сомневаться. В чем причина, размышляет наша героиня? Видимо, в том, как она предполагает, что генетика – наука сравнений. Ее ДНК сравнивали с ДНК других людей и на этом основании приходили к разным выводам. С чем сравнивали, относительно того и получали.

Для того, чтобы проверить свою догадку о причинах разногласий, наша героиня обратилась к ряду специалистов-генетиков. Первым был генетик, автор книги [«Краткая](#)

Героиню поразило, что в дело пошла Восточная Европа, которая в предыдущих тестах и не упоминалась, причем ее суммарное генеалогическое наследство потянуло на 18 %. Интересно, как его могла пропустить компания AncestryDNA? Появилась

В свете столь огромных нестыковок с другими определениями, десятые доли процента вызывают умильную улыбку. За кого, интересно, держат заказчиков-покупателей тестов? Как сообщает героиня нашего рассказа, она с удивлением нашла, что теперь она только на 3% скандинавка, в отличие от других данных и несмотря на отца-норвежца (как она пишет – very Norwegian father), и она уже понимает, что

история каждого, кто когда-либо жил». Он популярно объяснил ей, что она фундаментально не понимает, что именно делают компании по анализу представленных образцов ДНК. Они ведь не изучают наследственные – в прямом смысле – ДНК, они сравнивают вашу ДНК с ДНК современников. С кем более похожа – из тех, кто живут сегодня (!) – значит, и ваша ДНК от них. Просто по определению. Так, ваша тетька оказалась на 30 % итальянкой, и это просто потому, что ДНК в ее образце имела много общего с сегодняшними итальянцами. А именно, 30 % «общности», как показал компьютер.

От себя (А.А.К) добавим, что автор книги изрядно упростил свое объяснение. ДНК тети не показала 30 % сходства с итальянцами. Это ее ДНК, изрубленная в капусту, показала, «по мнению компьютера», 30 % сходства с ДНК современных итальянцев, тоже изрубленных (ДНК, не итальянцев) в капусту, причем компьютер сравнивал, как объяснено выше, всё многообразие фрагментов ДНК, от мелких до больших, все те ИПП и ИПС, и эта «капуста», которая сравнивалась, была «капустой» ДНК от более чем шести миллионов человек в базах данных, как объяснил автор книги. Поэтому и оказываются рядом, в одной таблице, и сирийцы, и норвежцы, и индийцы, и северные африканцы, и с каждым из которых есть какое-то «сходство». И это «сходство» дисциплинированный компьютер выдает в процентах и добавляет доли процента, потому что он дисциплинированный. Но когда сотрудники той коммерческой компании переносят эти доли процента в финальный сертификат, то это не от дисциплинированности, а от откровенной глупости или в лучшем случае от полной расхлябанности, а скорее всего, от склонности к мошенничеству. Там точность, как мы видим, в лучшем случае на уровне 15-20 %, а вовсе не долей процента. Указывая доли процента, сотрудники компании дают заказчику понять, с какой точностью они работают. Это по факту означает «верьте нам». Типа «мы, гусские, друг друга не обманываем».

Последующее расследование героиней нашего рассказа показало, что сами «геномные» компании прекрасно знают цену своих определений. Представитель AncestryDNA поделился с ней такими словами – *«Наши оценки являются, фактически, оценками. Эти оценки переменны и зависят от того, какой метод расчета используется, какая референсная панель берется за основу, какие еще образцы заказчиков используются при анализе»*. На самом деле любой «геномный» исследователь это прекрасно знает. Просто об этом не принято говорить. Более того, *«получаемые оценки зависят от того, сколько геномов той или иной этничности или с той или иной территории есть в «референсной панели»*. Если там мало, скажем, геномов с Ближнего Востока, то даже для ближневосточных заказчиков их «процентная оценка» может оказаться заниженной».

Как мы видим, представители компаний кивают на свои базы данных, типа: ну что мы можем сделать, когда вашей этничности там мало, и оценки получаются перекошенными. Но они никак не хотят признаться, что сама методология «рубить в капусту и сравнивать» является системно проблемной, как было объяснено выше. Так, у того же автора упомянутой книги компания 23andMe нашла небольшую долю ДНК американских индейцев. Его оценка – что этого просто не может быть, и что это просто статистический шум. Но множество людей принимают такое за правду и начинают выдумывать истории своего «происхождения», в соответствии с результатами геномного тестирования.

Героиня обратилась за разъяснениями таких разногласий в компанию 23andMe, и вот что там ей сказали: *«Различные геномные компании используют разные базы данных, различные алгоритмы для расчетов, потому и расхождения в результатах расчетов»*.

Британский генеалог объяснила нашей героине, что тесты AncestryDNA и 23andMe больше настроены на американскую аудиторию, и потому, например, англичане получают «менее полезные результаты», чем американцы, не говоря о людях с Ближнего Востока. Она привела в пример своего отца, все предки которого, насколько

смогли докопаться, происходят из Англии, но геномный тест показал у него только 8% «британского». Как ни странно, американцы часто показывают намного больший процент «британского», чем сами англичане.

Приятель нашей героини, мексиканского происхождения, по геномному тесту 23andMe оказался арабом из Северной Африки. Он сообщил о своем недоумении в компанию, они «upgraded» (то есть, улучшили) тест, и он оказался евреем из Южной Европы. Канадский специалист по этносам рассказал героине, что первый геномный тест показал, что он «100 %-ный ирландец», и по мере совершенствования теста он постепенно становился меньше ирландцем.

Президент компании GenCove, в которой пересчитывали результаты 23andMe и получили совсем другие показатели «происхождения», был озадачен, когда об этом узнал. Он сказал, что ожидал некоторые расхождения, вызванные разными «алгоритмами расчетов», но не настолько. В итоге президент «подработал» алгоритм своей компании и обработал данные еще раз. Расхождение оказалось еще больше. Ответ президента компании нашей героине: *«Честно говоря, я в замешательстве»*.

Еще один специалист, гарвардский генетик, сообщил в разговоре с нашей героиней, что сама суть «геномного тестирования» приводит к разнообразию результатов. Это не означает, что одни результаты правильные, а другие нет. «Правильные» результаты там вообще невозможно получить. Но заказчики тут же придумывают свои якобы семейные истории, которые «поддерживаются» результатами «геномного анализа». На самом деле, по словам гарвардского генетика, эти результаты работают на уровне континентов – ваше происхождение из Европы, ваше из Африки, а ваше из Азии. Это более-менее разумно. Можно также идентифицировать ваших близких родственников. Все остальное – *«примите с хорошей дозой соли»*.

Зная всё это, «геномные компании» через умело организованные рекламные сервисы подогревают интерес потенциальных покупателей, [выпускают постановочные видео](#), в которых актеры, разыгрывая изумление результатами тестов, «находят» экзотических родственников и «узнают» свою категорическую и совершенно безошибочную родословную. Никаких сомнений и других вариантов. Это видео посмотрели уже более 17 миллионов человек. Еще недавняя видео-реклама компании AncestryDNA – герой которой был «100-процентным» немцем, ходил в традиционных кожаных штанах, и не выходя из них, заказал «геномный» тест в Ancestry. Тест показал, что он – шотландец. Герой тут же настолько переформатировался, что купил килт, шотландскую мужскую юбку.

По размышлению, наша героиня пришла к неутешительному (для «геномного анализа») выводу, что она перестала понимать, что такое «происхождение» в «геномном смысле». Дело в том, что в отличие от ДНК-генеалогии по Y-хромосоме, то есть по отцовской линии, в «геномном анализе» рассматривается геномный вклад как по мужской, так и по женской линии, причем с каждым поколением эти линии расходятся в геометрической прогрессии, захватывая порой всё новые и новые этносы. Это что, и есть отражение разных процентных составов от разных этносов и территорий? Или эти разные процентные составы – просто результат усреднения компьютером разных «похожестей», не имеющих к этносам большого отношения? Например, в роду нашей героини были как минимум сирийская линия, норвежская и английская, так какое у нее «происхождение»? Как отметил в разговоре с ней генетик, автор упомянутой выше книги, что если тест заходит на глубину 500 лет и глубже, то ее предки по мужской и женской линиям должны были жить по всей Европе. Так какое же у нее «геномное происхождение»? Как сказал ей еще один генетик – *«Если вы знаете ген, это вовсе не означает, что вы знаете генетическую линию или происхождение»*.

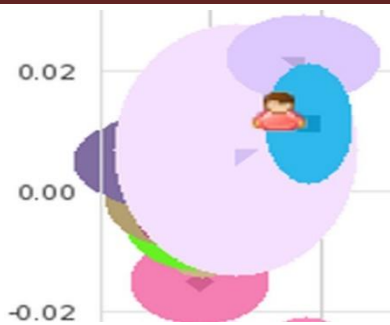
В последнее время подобные «коммерческие геномные компании» появились и в России. Принципы и суть их «этнических исследований» являются такими же, неважно – персональных исследований или целых народов. Какое-то время назад в научной среде

разразился громкий шум в отношении публичного интервью руководителя компании «Генотек» Валерия Ильинского. По его информации, *«среднестатистический житель России 16 процентов генома унаследовал от коренных жителей Центральной России, а все остальные участки являются мозаикой, составленной из фрагментов геномов, характерных для обитателей других регионов»*. В изложении популярных изданий это превратилось в следующее: *«Исследователи Genotek установили, что современные россияне являются коренными русскими только на 16%»*. Кто такие «коренные русские», и по какому критерию и как их находили, руководитель компании «Генотек» не пояснил, хотя потом в социальных сетях компания поделилась, что *«под “коренными русскими” понимались участки, характерные для населения именно Центральной России, не сибиряков и не кавказцев»*. Эта «методология» привела «Генотек» к выводу, что россияне на 10,8 % являются потомками британцев (на 6 % - англичан и на 4,8 % - шотландцев), на 6,3% - потомками венгров, на 1,3 % - пакистанцев. Опять, обратите внимание на точность определений!

На вопрос о том, откуда у русских в предках шотландцы, компания завела известную популяционную песню, уже знакомую нам из того, что описано выше в этой статье: *«Речь идет об одинаковых геномных локусах, присутствующих как в генетическом материале испытуемых, так и в ДНК шотландской популяции. Далее происходит примерно следующее: выделяем некую древнюю популяцию (назовем ее популяцией X). Сравниваем количество участков, характерных для древней популяции X, в геноме современных русских и геноме современных шотландцев. Видим, что геном современных шотландцев обогащен участками, происходящими от популяции X; принимаем популяцию X за “древних шотландцев” и на основе этого делаем заключение о доле “шотландской крови” в геноме испытуемого»*. Занятна здесь не только популяционная «методология», вся в методологических дырах, которые мы разбирали выше на примере нашей американской героини, но и то, что сходство фрагментов ДНК у русских и шотландцев непременно постулируется как то, что русские «являются потомками шотландцев», но никогда не наоборот, хотя «сходство» направления не имеет, оно симметрично.

Более того, не рассматривается, что и те русские, и те шотландцы, у которых «сходство в ДНК», образовались от одного древнего общего предка, а не друг от друга. В ДНК-генеалогии это давно известно. Например, в Шотландии и России есть много носителей одной и той же гаплогруппы R1a, но это вовсе не означает, что «те русские произошли от шотландцев», пользуясь «методологией» популяционных генетиков, как не означает и наоборот. Если понимать предмет, то легко обратить внимание на то, что у тех шотландцев гаплогруппа на самом деле записывается в их большинстве как (опуская несколько промежуточных субкладов) R1a-Z645-Z284 (так называемый «скандинавский субклад»), а у тех русских в их большинстве как R1a-Z645-Z280 (так называемый «субклад Русской равнины»), и одно из другого не образуется. Но и те, и другие образовались от одного общего предка, который имел субклад R1a-Z645. Если бы геномные популяционисты это знали и понимали, они осознавали бы, что «сходство» не означает происхождение одного от другого. Я могу быть похожим на своего дядю, но от дяди не произошел. К сожалению, популяционисты это не понимают, потому что вся их «методология» построена на «сходстве», как и их основной «метод принципиальных компонент», который из понимающих не высмеивал только ленивый.

Именно этот подход является основным в работе российских коммерческих компаний, занимающихся «персональными этническими идентификациями». Приведем пример. В [Академию ДНК-генеалогии](#) недавно обратился клиент одной из таких компаний (а такие обращения к нам нередки), в состоянии полного недоумения. Московская компания по его ДНК-тесту определила, что он по происхождению «мордовец».



Компания пришла к такому выводу на основании того же «метода принципиальных компонент». Приведем фрагмент диаграммы. Фигурка – это клиент, который, разумеется, заплатил компании деньги за эту диаграмму, которая и легла в основу заключения о его «происхождении». Большой светло-сиреневый овал – это «русские». Синий – мордва. Темно-сиреневый выше – это карелы. Приводить расшифровку других цветов, выглядывающих

из-под «русских», большого смысла не имеет, там еще тот винегрет. И вот теперь поставьте себя на место того «клиента»: вы заплатили деньги и в ответ получаете, что вы похожи на «русского», но еще больше на «мордовца», там совпадение 97 %. Но еще есть запасные варианты: с вероятностью 93 % вы украинец, 91 % – поляк, 80 % – татарин, 73 % – румын, 49 % – еврей-ашкенази, а 43 % – еврей-сефард. И так далее. Так за что платили деньги? А, собственно, что вы хотели увидеть, идя на такой широкий «этнический» захват? Что вы не австралийский абориген? Не папуас? Ну вот, увидели. Мордовец, гуляйте. И вот пишет нам новоявленный «мордовец», что, мол, как это так? У меня все в роду русские, о мордовцах в родне и не слыхивал, так что помогите, давайте я у вас получу свой гаплотип и гаплогруппу-субклад, без всяких «этносов», свой этнос я и сам знаю. Русский я, и был всегда. Прав он, что и говорить.

В завершение надо подчеркнуть, что персональная ДНК-генеалогия базируется совершенно на других принципах. Тесты проводятся не на смешанном геноме мужчин и женщин, а на рассмотрении либо Y-хромосомы (мужская, отцовская линия), либо мтДНК (женская, материнская линия). Никаких разветвлений с каждым поколением там нет, одна линия, мужская или женская, уходит в глубь времен на тысячи и десятки тысяч лет, и при ее идентификации рассматривают не кашу, которую сортирует по своему усмотрению компьютер, а совершенно четко определенные гаплотип и гаплогруппа-субклад. Их можно изобразить в виде четкой последовательности чисел и нескольких индексов на листе бумаги (на самом деле на осьмушке листа), и анализ без труда можно провести вручную, не прикрываясь гига- и терабайтовыми компьютерными программами с массой условий, приближений и допущений, что там «по наследству» и что «по похожести».

Результатом персональной интерпретации получаемых данных является не «этничность» в процентах, которой грош цена и которая не воспроизводится от лаборатории к лаборатории, а четко определенные индивидуальные показатели, которые относят человека к определенному древнему роду (гаплогруппе) и племени (субклад) и описывают его гаплотип, по которому, сопоставляя с другими столь же четко определенными гаплотипами, можно проводить исторические отнесения – к археологическим культурам прошлого, к древним миграциям человека, порой и к древним событиям, в которых могли принимать участие предки данного человека. И не нужно при этом привлекать «геномный анализ», который в огромной степени есть фикция, подтасовка или, в лучшем случае, неопределенность. И в осознании этого – главный урок «геномной популяционной генетики».