

*Ultima ratio*

**Вестник Российской Академии  
ДНК-генеалогии**

**Том 4, № 11**

**2011 ноябрь**

**Российская Академия ДНК-генеалогии**

ISSN 1942-7484

**Вестник Российской Академии ДНК-генеалогии.** Научно-публицистическое издание Российской Академии ДНК-генеалогии. Издательство Lulu inc., 2011.

*Авторские права защищены. Ни одна из частей данного издания не может быть воспроизведена, переделана в любой форме и любыми средствами: механическими, электронными, с помощью фотокопирования и т. п. без предварительного письменного разрешения авторов статей.*

*При цитировании ссылка на данное издание обязательна.*

Составитель  
*Российская Академия ДНК-генеалогии*

Оформление издания  
*Anatole A. Klyosov*  
*Павел Шварев*

© Авторские права на статьи принадлежат Российской Академии ДНК-генеалогии, 2011. При перепечатке ссылка обязательна.

© РА-ДНК, 2011

# СОДЕРЖАНИЕ НОМЕРА

Оглавление .....	2062
Гаплотип фараона Тутанхамона (1333 – 1323 до н.э.) и его возможное происхождение. <i>А. А. Клёсов</i> . .....	2063
Гаплотипы гаплогрупп R1a1 и N1 в восточных Гималаях. <i>А. А. Клёсов</i> . . .	2070
Армянские гаплотипы гаплогруппы Т-М184 из Закавказья. <i>А.А. Лабай</i> . . .	2073
Об удивительных и неожиданных языковых совпадениях между лакским и аккадским языками. Топонимика и личные имена лаков (Окончание). <i>Р.А. Омариева</i> . .....	2080
Исследования мтДНК. Предисловие редактора. <i>А.А. Клёсов</i> .....	2088
Исследование мтДНК, популяционная и демографическая генетика, русской сельской популяции Ярославской области в прошлом и настоящем. <i>А. Ф. Назарова</i> . .....	2093
ДНК-генеалогия и письма из Киргизии (хрестоматийные «что делать» и «с чего начать»). <i>Чинара Сейдахматова – Анатолий Клёсов</i> . .....	2100
DISCUSSIONS	
The mutation rate constants in DNA genealogy and related issues. ....	2108
ДНК-ГЕНЕАЛОГИЯ ДЛЯ НАЧИНАЮЩИХ	
ДНК-генеалогия. О чем эта наука, что она определяет и выявляет, и кому она интересна. <i>А.А.Клёсов</i> . .....	2157
Letters from the Readers: PERSONAL CASES	
Part 33, letters 109-111. ....	2196

# Гаплотип фараона Тутанхамона (1333 – 1323 до н.э.) и его возможное происхождение

А.А. Клёсов

<http://aklyosov.home.comcast.net>



Для начала – несколько справочных сведений. Тутанхамон, фараон XVIII династии Древнего Египта, правил в 1333-1323 гг до нашей эры, то есть

примерно 3330 лет назад, если дату привести в вид, пригодный для расчетов в рамках ДНК-генеалогии.

Перед ним в XVIII династии правили фараоны Яхмос I → Тутмос I → Тутмос III → Аменхотеп II → Аменхотеп III → Эхнатон → Хоремхеб → Тутанхамон. Тутанхамон был сыном Эхнатона, и в династии ведет свою родословную от Тутмоса I (1504-1492 гг до н.э.), и, возможно, от Яхмоса I.

Ятмос I, основатель XVIII династии, известный также как Яхмес, Афмос или Амасис I, сын Секенен-Ра II, правил в 1550-1525 до н.э., и жил, соответственно, 3560 лет назад. Это было время войны с гиксосами, на которой погиб его отец.

Перед Ятмосом I правили 18 фараонов, но по особому принципу счета Яхмос начинает 18-ю династию, в которой Тутанхамон был последним, восьмым по счету. Перед XVIII династией были пять фараонов Среднего Царства (Ментухотеп II → Аменемхет I → Сенусерт I → Сенусерт III → Аменемхет III), перед ними семь фараонов Древнего Царства (Джосер → Снофру → Хуфу → Хафра → Менкаура → Сахура → Пиопи II), и перед ними – шесть фараонов раннего царства (Нармер → Менес → Хор Аха → Джер → Ден → Хасехемуи).

17 февраля 2010 года министр культуры Египта Фарук Хосни и генеральный секретарь Высшего совета Египта по древностям Захи Хавасс объявили о результатах исследований 2007-2009 гг, согласно которым были определены (в некоторой степени) генетические особенности Тутанхамона и определена его гаплогруппа. Гаплогруппа объявлена не была, и судя по разным признакам, ее объявлять руководителям Египта не хотелось. Было ясно, что гаплогруппа какая-то не та, которую вождям Египта хотелось бы увидеть. То есть было понятно, что это не типичная египетская гаплогруппа, а какая-то иностранная, и потому фараоны были тоже чужестранные. Это, видимо, травмировало патриотические чувства египетского руководства.

Но утечка информации состоялась. Сначала ходили упорные слухи, что Тутанхамон имел гаплогруппу R1b, то есть «европейскую», в понятиях обывателей, а значит, в понятиях обывателей, предки Тутанхамона прибыли из Европы. Не исключено, что это были англичане или французы, ненавистные еще по воспоминаниям о войне 1956 года. Поэтому гаплогруппу засекретили. Исходя из этого, именно гаплогруппа R1b была наиболее вероятной для фараона. Потом состоялась утечка информации о самом гаплотипе, которой в самом деле принадлежал к гаплогруппе R1b1a2.



Наконец, на сайте iGENEA, швейцарской компании (Цюрих)  
<http://www.igene.com/en/index.php?c=62>

был обнародован и сам 16-маркерный гаплотип

13 24 14 11 11 14 X X 10 13 13 30 -- 16 14 19 10 15 12

Здесь первые 12 аллелей представлены в формате FTDNA, за исключением маркеров DYS426 и DYS388, остальные шесть аллелей - DYS458, 437, 448, GATAN4, DYS 456, 438. Другими словами, это известный Y-файлер, 17-маркерный формат гаплотипов со снятым маркером DYS635.

Понимающему человеку уже ясно, что это не обычный «европейский»

гаплотип R1b1a2, который швейцарская компания могла подsunуть, выдавая якобы за гаплотип фараона. Дело в том, что почти все европейские гаплотипы группы R1b1a2 имеют DYS439=12 (это маркер сразу после X выше), здесь же он 10. Это - редкость, таких - всего примерно 0.5% европейских R1b1a2 гаплотипов.

iGENEA, как коммерческая компания, естественно, сразу начала коммерческую кампанию, вариант сказки про Золушку. А именно, объявила конкурс на примеривание этой хрустальной туфельки за 179 или 399 долларов, что есть цена за определение гаплотипа всем желающим, в обычном или улучшенном варианте тестирования. Было объявлено, что начинается поиск европейских родственников фараона.

При этом на сайте компании появилась завлекательная «информация», от которой вянут уши. Например, что гаплогруппа R1b1a2 появилась 9500 лет назад в районе Черного моря, и вместе с продвижением сельского хозяйства начиная с 7000 лет назад носители этой гаплогруппы двинулись в Европу. Они были якобы индоевропейцами, которые распространились по Европе несколькими волнами вскоре после 7000 лет назад. Сообщается, что в Египте этой гаплогруппы менее 1%, что вызвано отчасти европейскими миграциями в Египет в последние 2000 лет.

Правда, при этом непонятно, откуда эта гаплогруппа появилась в районе Черного моря, хотя можно догадаться, что ноги там растут от работ Марии Гимбутас про нашествие индоевропейцев в Европу примерно в те же времена. Правда, индоевропейцами должны быть носители R1a1, и тогда непонятно, при чем здесь R1b1a2, но чего не бывает. Тем более что iGENEA сообщает, что «индоевропейская курганная культура» распространилась по Европе начиная с 6400 лет назад, и это были R1b1a2. Откуда это iGENEA взяла, на основании каких данных про R1b1a2 в Европе в такие ранние времена - остаётся неизвестным. На самом деле ранее 4500 лет назад, то есть и две тысячи лет спустя, R1b1a2 в Европе не зафиксированы, и только реконструкция их базовых гаплотипов дает датировку 4800 лет назад для их прибытия в Европу и начала движения культуры колоколовидных кубков.

Так откуда появилась гаплогруппа R1b1a2 у египетских фараонов? Без датировок беспредметные гадания только продолжатся. Как мы видим, никто не принимает во внимание сами гаплотипы, их аллели и мутации, а значит - датировки. Не принимают - потому что не умеют по ним датировать. Посмотрим, что можно в этом отношении сделать. Взглянем еще раз на гаплотип Тутанхамона:

13 24 14 11 11 14 X X 10 13 13 30 -- 16 14 19 10 15 12

Он отличается на 6 мутаций от базового гаплотипа гаплогруппы R-M269\* (с возрастом примерно 7000 лет назад)

12 24 14 11 11 14 X X 12 13 13 29 -- 16 15 19 11 15 12

на 7 мутаций от базового гаплотипа гаплогруппы R-L51 (с возрастом примерно 5850 лет назад)

13 25 14 11 11 14 X X 12 13 13 29 -- 17 15 19 11 15 12

на 6 мутаций от базового гаплотипа гаплогруппы R-L11 (с возрастом примерно 5300 лет назад)

13 24 14 11 11 14 X X 12 13 13 29 -- 17 15 19 11 15 12

на 6 мутаций от базового гаплотипа гаплогруппы R-P312 (с возрастом от 3625 до 4100 лет назад по разным данным)

13 24 14 11 11 14 X X 12 13 13 29 -- 17 15 19 11 15 12

и на 8 мутаций от базового гаплотипа гаплогруппы R-U106 (с возрастом 4175 лет назад)

13 23 14 11 11 14 X X 12 13 13 29 -- 17 15 19 11 16 12

6, 7, и 8 мутаций между двумя указанными 16-маркерными гаплотипами транслируются соответственно в

$6/0.0315 = 190 \rightarrow 234$  поколения, то есть 5850 лет;

$7/0.0315 = 222 \rightarrow 284$  поколений, то есть 7100 лет;

$8/0.0315 = 254 \rightarrow 338$  поколений, то есть 8450 лет.

В таком случае общий предок Тутанхамона с базовыми гаплотипами M269\*, L51, L11, P312 и U106 жил, соответственно

8,090 лет назад – с M269\*

8,140 лет назад – с L51

7,210 лет назад – с L11

6,400-6,600 лет назад – с P312

7980 лет назад – с U106



Все эти датировки свидетельствуют о том, что гаплотип Тутанхамона относится скорее к субкладу R1b1a2-M269\*, с общим предком 6000-8000 лет назад. Это, естественно, не европейский гаплотип, где общие предки в подавляющем большинстве имеют «возраст» 4200-4500 лет назад для самых древних европейских субкладов R1b1a2.

Итак, гаплотип R1b1a2 фараона Тутанхамона – не европейский. Продолжаем вопрос – откуда он появился в Египте 3330 лет назад, если не из Европы?

Ответ на этот вопрос я дал несколько лет назад, причем в серии публикаций. После прибытия на Ближний Восток миграционным маршрутом с Русской равнины (7000-6500 лет назад) [а перед тем из Центральной Азии 16 тысяч лет назад] через Кавказ (6000 лет назад) и Анатолию (6000 лет назад), основав Шумер (5500 лет назад и позже), с датировками в Ливане (5200±670 лет назад), носители гаплогруппы R1b1a2 направились на запад, по северо-африканскому побережью, прошли через Египет примерно 5500-5200 лет назад, и в итоге вышли к Атлантике, переправились через Гибралтар и высадились на Пиренеях около 5000 лет назад. Это было начало культуры колоколовидных кубков, которая примерно 4800 лет назад двинулась с Пиреней на север и заселила Европу между 4500 и 3000 лет назад.

Данные по гаплотипу Тутанхамона показывают, что в ходе этого маршрута носители R1b1a2 установили в Египте правящую верхушку, которая положила начало династиям фараонов. Когда это было? По датировкам миграционного маршрута – между 5500 и 5200 лет назад. Действительно, 6000 лет назад они еще были на Кавказе и в Анатолии, 5500-5200 лет назад на Ближнем Востоке, и 4800 лет назад уже на Пиренеях.

И вот сейчас переходим к списку фараонов, который не случайно был приведен в начале этой статьи. Родоначальник всей линии фараонов, Нармер, фараон Раннего Царства, нулевая династия, жил в 32-м веке до н.э., то есть 5200 лет назад. Примечательно, что его имя отсутствует в царских списках, то есть он появился неизвестно откуда, но он был победителем Нижнего Египта и объединил его с Верхним Египтом. Поскольку сам Нармер не был документированным царем, то основателем Древнего Египта исторические источники считают его сына Менеса, который уже имел официальный царский титул. Началась новая эра в истории Египта и новая царская линия, линия фараонов. Как вытекает из вышесказанного, линия гаплогруппы R1b1a2.

Источники насчитывают пять основных теорий происхождения Нармера. На самом деле ни одна из них ничего не говорит о происхождении Нармера. Эти «теории» говорят о том, был ли Менес наследником Нармера, или Менес и Нармер – одно и то же лицо, как долго шло объединение Египта, когда оно было закончено, был ли победитель восстания в Нижнем Египте подлинным объединителем Египта и так далее. Как мы видим, о происхождении Нармера там ровным счетом ничего нет. Но мне не жалко, пусть моя теория будет шестой. Она – о том, что Нармер или его прямые предки, носители гаплогруппы R1b1a2, прибыли с Ближнего Востока, а перед тем – с Кавказа и с Русской равнины, а перед тем – с Урала и Центральной Азии. В свою очередь, потомки этих людей сейчас составляют до 60% населения Западной и Центральной Европы. Они, естественно, не потомки египетских фараонов, они потомки их предков.

# Гаплотипы гаплогрупп R1a1 и N1 в восточных Гималаях

А.А. Клёсов

<http://aklyosov.home.comcast.net>

В конце ноября 2011 года была опубликована статья (Kang et al, 2011), в которой приведены 170 14-маркерных гаплотипов племен Luoba и Deng восточных Гималаев, которых относят к сино-тибетскому населению. 14 маркеров – это первые 12 маркеров панели FTDNA плюс DYS437 и DYS438.

Все 170 гаплотипов разошлись по 14 гаплогруппам и субкладам, из которых по одному гаплотипу относятся к гаплогруппам C, D\*, J, O3\*, Q, R1, по три гаплотипа – к D1, O\*, R\*. Пять гаплотипов принадлежат гаплогруппе R1a1-M17, все – из племени Луоба, как и четыре гаплотипа гаплогрупп R\* и R1.

18 гаплотипов отнесены к гаплогруппе D3a, 40 гаплотипов – к гаплогруппе N1\* (все – племя Луоба, кроме одного гаплотипа), и 92 гаплотипа (54% от всех) – к субкладам O3a3c\* и O3a3c1.

Дерево гаплотипов приведено на рис. 1. Пять гаплотипов гаплогруппы R1a1 имеют всего четыре мутации от базового гаплотипа

13 25 15 10 11 14 13 14 10 13 11 17 – 14 11

что дает  $4/5/0.0215 = 37 \rightarrow 38$  поколений, то есть 950 лет до общего предка. Но сам базовый гаплотип очень необычный, в нем DYS426 = 13, и DYS388 = 14. Это явно не недавний визитер с Русской равнины. С базовым гаплотипом последней данный базовый гаплотип разделяет 5 мутаций, то есть  $5/0.0215 = 233 \rightarrow 302$  поколения, то есть примерно 7550 лет. Если принять, что базовому гаплотипу Русской равнины 4800 лет, то общий предок тибетского R1a1 и европейского R1a1 субкладов жил примерно  $(7550+4800+950)/2 = 6650$  лет назад. Эта цифра знакома и встречается у азиатских гаплотипов R1a1, но не с DYS426=13 и DYS388=14. Это определенно какая-то отдельная ветвь. Правда, азиатские гаплотипы обычно короткие и не содержат, как правило, данные маркеры, но в древних алтайских и тувинских базовых гаплотипах, а также в пакистанских базовых гаплотипах DYS388=12.



от которого во всех трех гаплотипах всего одна мутация (390 лет до общего предка). Но этот базовый гаплотип отличается от базового тибетского гаплотипа на 16 мутаций в 14 маркерах. Это эквивалентно  $16/0.0215 = 744 \rightarrow 1870$  поколений, то есть 46,750 лет между их общими предками, и помещает время жизни ИХ общего предка на 24 тысячи лет назад. Это – несколько ниже, чем ожидалось бы для гаплогруппы R, но для столь малого числа гаплотипов статистика недостаточна.

Поскольку данных в литературе для гаплогруппы N1\* мало, то рассмотрим, что здесь могут дать тибетские гаплотипы. Из рис. 1 видно, что ветвь справа (из 30 гаплотипов) плоская, следовательно, относительно недавняя. Действительно, она содержит 39 мутаций от базового гаплотипа

13 23 14 10 12 12 12 14 12 14 14 17 - 14 11

что дает  $39/30/0.0215 = 60 \rightarrow 64$  поколения, то есть примерно 1600 лет до общего предка. Это – середина 1-го тысячелетия нашей эры.

Вторая ветвь малая, и состоит из пары идентичных гаплотипов гаплогруппы N1\*

13 23 14 10 12 12 12 14 12 13 14 17 - 14 11

то есть отличается всего на одну мутацию от большой ветви, и одного отдельно стоящего гаплотипа

13 23 14 10 13 15 12 14 12 14 14 17 - 14 11

На все три базовых гаплотипа – пять мутаций, что дает  $5/3/0.0215 = 78 \rightarrow 85$  поколений, или 2125 лет дистанции от усредненного возраста всех трех гаплотипов (530 лет). Это суммарно дает 2655 лет до общего предка всех трех ветвей гаплогруппы N1\*. Это, конечно, совершенно не отражает возраст N1\*, и определенно указывает на прохождение бутылочного горлышка тибетской популяции N1\* относительно недавно, в середине 1-го тысячелетия до нашей эры.

## Литература

Kang, L., Lu, Y., Wang, C., Hu, K., Chen, F., Liu, K., Li, S., Jin, L. and Li, H. (2011), Y-chromosome O3 Haplogroup Diversity in Sino-Tibetan Populations Reveals Two Migration Routes into the Eastern Himalayas. *Annals of Human Genetics*. doi: 10.1111/j.1469-1809.2011.00690.x

# Армянские гаплотипы гаплогруппы Т-M184 из Закавказья

А.А. Лабай  
V12189@mail.ru

Абсолютная аксиома – человек не выбирает ни родителей, ни место, ни время рождения. А значит, мужчина не выбирает свою гаплогруппу и гаплотип. Они передаются ему отцом через акт рождения и не меняются в процессе жизни. Эти биологические признаки не определяют ни внешность, ни способности, ни другие, наследуемые от родителей признаки, а несут информацию о предках по мужской линии – принадлежность к роду и о времени жизни предка, прошедшего «бутылочное горлышко» внешних условий и ставшего «патриархом» той или иной группы (популяции) мужчин.

Вот так получилось, что у меня гаплогруппа Т, редкая и загадочная, разбросанная по всему миру и встречающаяся среди многих национальностей и адептов разных религий. Интерес к своим корням привёл меня в ДНК-генеалогию с целью получить как можно больше информации о своём роде.

Первое исследование 49- и 22-маркерных гаплотипов гаплогруппы Т из базы FTDNA провёл А.А.Клёсов (2011а). Полученные результаты позволили сделать следующие выводы:

1. Гаплогруппа Т, как и гаплогруппа В, произошла от «бета»-гаплогруппы, которая возникла 64,000±6,000 лет назад.
2. Возраст общего предка гаплогруппы Т по имеющимся гаплотипам составляет 20,000 лет.
3. Возраст общего предка гаплогруппы Т современной популяции мужчин примерно 10,000 лет назад.
4. Современная популяция гаплогруппы складывается более чем дюжины популяций из разных уголков мира.

Второе исследование было так же проведено А.А.Клёсовым (2011b). На этот раз, расчёты проводились по выборке из 35-ти гаплотипов этнических армян из Закавказья. 17-маркерные гаплотипы были взяты из (Herrera et al., 2011). Всего в статье опубликовано 260 гаплотипов, из них 120

гаплотипов R1b1b2 (46%), 105 гаплотипов J2 (40%) и 35 гаплотипов гаплогруппы T (14%). Это уже даёт представление о порядке распределения основных гаплогрупп среди армян Закавказья.

Экспресс-анализ дерева из 35 гаплотипов, без разделения на ветви, дал время жизни общего предка армянской популяции субклада M184 -  $6450 \pm 770$  лет назад. Проверка по двум ветвям с наиболее выраженными базовыми гаплотипами подтвердила эту дату - примерно 7400 лет назад (Клёсов, 2011b). Так как статья А.А. Клёсова изначально ставила перед собой цель провести только беглый анализ, то подробное рассмотрение дерева не проводилось.

Полученный результат не противоречит данным из первой статьи и позволяет сделать вывод о почтенном возрасте армянского субклада M184. Но учитывая, что и первая и вторая ветвь имели предков на временном интервале не более 1000 лет назад, остаётся открытым вопрос об автохтонности гаплогруппы T в Закавказье.

Проверка времени жизни общего предка всей популяции пермутационным методом показала, что на 35 гаплотипов имеется 20528 пермутаций. Тогда время жизни общего предка будет:

$20528/2/1225/17/0,002 = 246$  поколений, или 6160 лет назад.

При пермутационном (квадратичном) способе поправка на возвратные мутации не вводится (Klyosov, 2009). Если сравнить цифры, полученные разными методами (6160,  $6450 \pm 770$  и 7400 лет), то можно предположить, что в популяции имеются молодые субклады, которые и омолаживают общего предка и надо проводить анализ по каждой популяции отдельно. Для этого было построено дерево из 35 гаплотипов гаплогруппы T субклада M184, которое представлено на рис. 1.

Нумерация гаплотипов взята из первоисточника без изменения и отражает, дополнительно, региональную принадлежность гаплотипов. Видно, что представители армянской нации происходят из четырёх регионов:

1. Ararat Valley (Турция и Армения)
2. Gardman ( Азербайджан)
3. Sasun (Турция)
4. Lake Van (Турция)

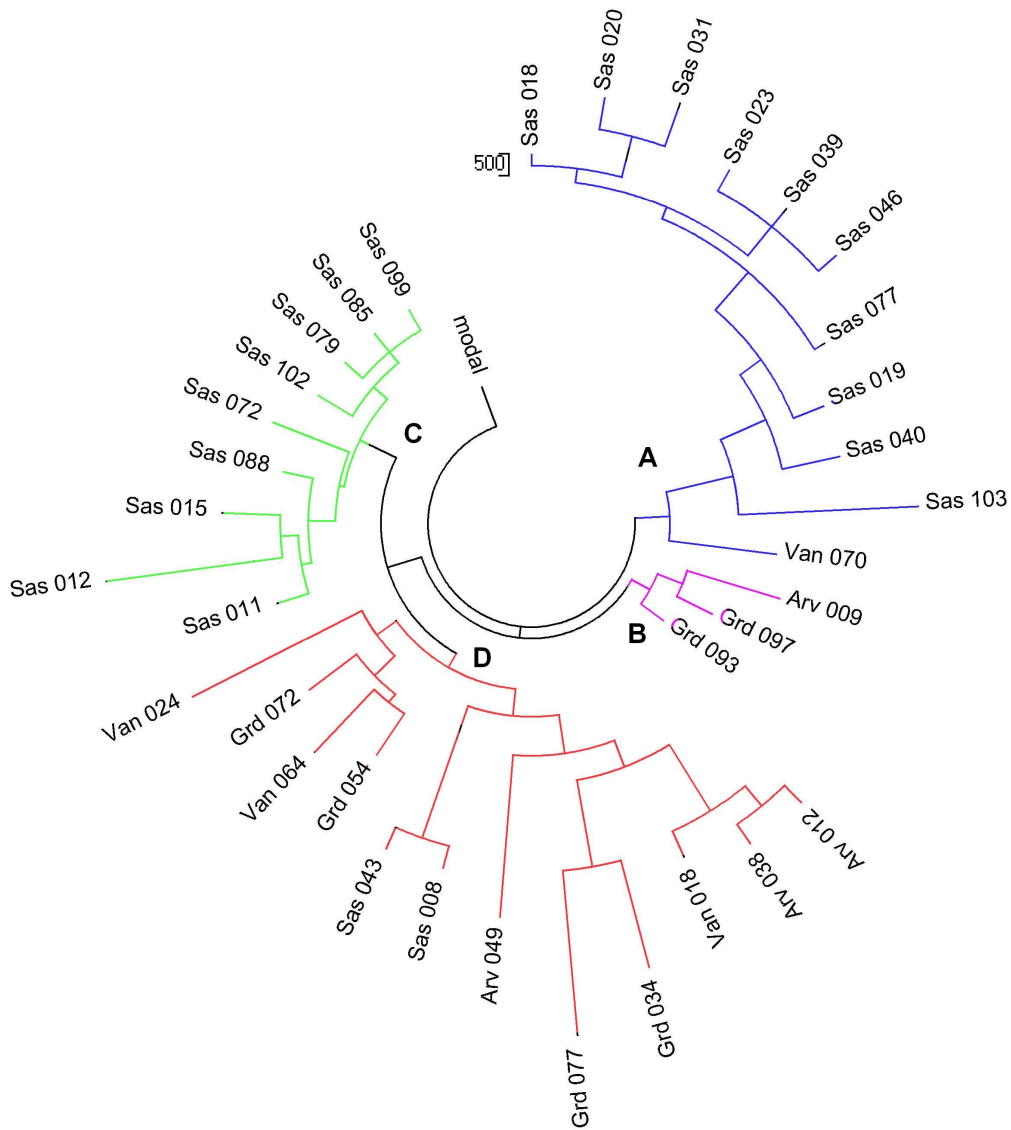
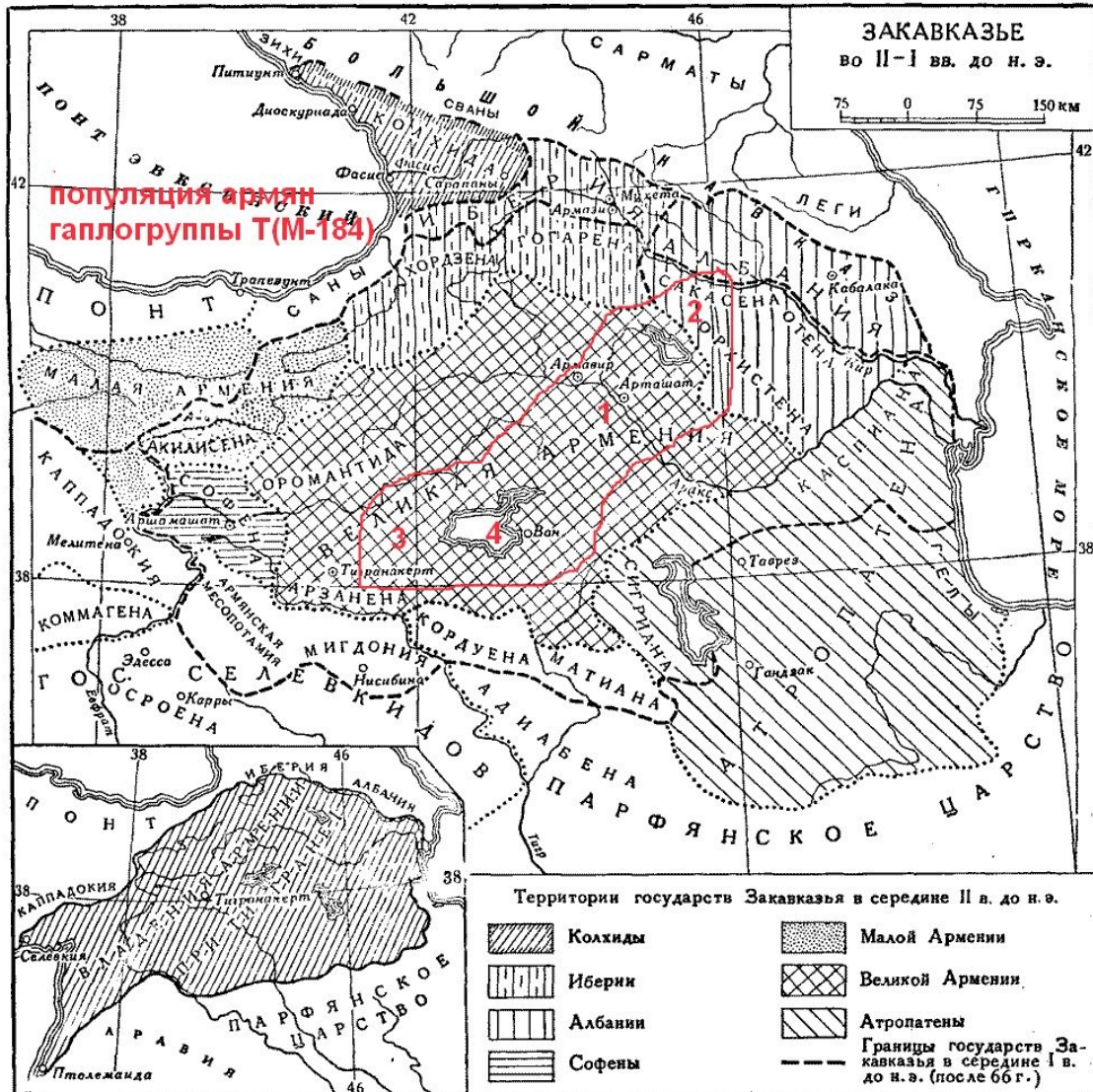


Рис. 1 Дерево 35-ти 17-маркерных гаплотипов армян из гаплогруппы Т (М-184).

Расположение армянских представителей гаплогруппы Т схематично очерчено на карте «Закавказье во II-I вв. до н.э.»





Анализ ветвей проводился линейным методом. Для каждой ветви был определён базовый гаплотип, подсчитаны мутации, определено время жизни общего предка для каждой ветви. Затем была сформирована матрица базовых гаплотипов и проведен расчёт общего предка для всех ветвей.

Первая ветвь обозначена индексом А.

Базовый гаплотип ветви А, дающий наименьшее число мутаций, имеет вид (расположение маркеров в гаплотипе соответствует расположению маркеров в стандартных панелях FTDNA):

14-23-15-10-14-18-X-X-11-14-15-17---17-15-19-11-15-9-21 (А)

Относительно этого гаплотипа в выборке из 11 гаплотипов набирается 37 мутаций. Тогда время жизни общего предка ветви А:

$$37/11/0,034 = 99 \rightarrow 106 \text{ поколений} = 2740 \pm 620 \text{ лет назад}$$

Вторая ветвь обозначена индексом В. Её базовый гаплотип имеет вид :

13-23-14-10-14-16-X-X-11-14-14-17---17-14-19-11-16-9-21 (В)

Относительно этого гаплотипа в выборке из 3 гаплотипов имеется 5 мутаций. Тогда время жизни общего предка ветви В будет:

$$5/3/0,034 = 49 \rightarrow 52 \text{ поколений} = 1290 \pm 610 \text{ лет назад.}$$

Третья ветвь обозначена индексом С. На самом деле эта ветвь состоит из двух ветвей, но разделять её на подветви смысла нет, так как популяция малочисленна. Её базовый гаплотип, дающий наименьшее число мутаций, имеет вид :

13-24-14-10-14-16-X-X-11-13-13-17---16-14-19-11-16-9-21 (С)

Относительно этого гаплотипа в выборке из 9 гаплотипов набирается 21 мутация. Тогда время жизни общего предка ветви С равно:

$$21/9/0,034 = 69 \rightarrow 74 \text{ поколений} = 1840 \pm 490 \text{ лет назад.}$$

Четвёртая ветвь, обозначенная индексом D, также состоит из двух ветвей. Но и её дробить нет смысла. Базовый гаплотип, дающий наименьшее число мутаций, имеет вид :

13-23-15-10-14-16-X-X-12-13-13-16---16-14-19-11-15-9-21 (D)

Относительно этого гаплотипа в выборке из 12 гаплотипов набирается 79 мутаций. Тогда время жизни общего предка ветви D:

$$79/12/0,034 = 194 \rightarrow 239 \text{ поколений} = 5980 \pm 1130 \text{ лет назад.}$$

Матрица базовых гаплотипов всех ветвей:

13-23-14-10-14-16-X-X-11-14-14-17---17-14-19-11-16-9-21 (В) [1290±610 л.н.]  
13-24-14-10-14-16-X-X-11-13-13-17---16-14-19-11-16-9-21 (С) [1840±490 л.н.]  
14-23-15-10-14-18-X-X-11-14-15-17---17-15-19-11-15-9-21 (А) [2740±620 л.н.]  
13-23-15-10-14-16-X-X-12-13-13-16---16-14-19-11-15-9-21 (D) [5980±1130 л.н.]

В четырёх базовых гаплотипах имеется 18 мутаций. Это разводит их общих предков на

$18/4/0,034 = 132 \rightarrow 152$  поколений, то есть 3800 лет назад, плюс усредненное время жизни предков всех четырех базовых гаплотипов (~ 2960 лет назад), то есть примерно 6,760 лет назад.

Эта величина в пределах погрешности равна трем величинам, приведенным выше – 6160,  $6450 \pm 770$ , 7400 лет назад. Таким образом, оказалось, что в данном случае считать по всему дереву или делить на ветви не влияет на конечный результат. Иначе говоря, ветви на дереве достаточно сбалансированы.

Можно утверждать, что самой древней популяцией является ветвь D, охватывающая все четыре географических региона. Ее возраст,  $5980 \pm 1130$  лет, в пределах погрешности совпадает с возрастом всего дерева. Видимо, это и есть предковая популяция гаплогруппы T в Закавказье.

Затем по древности можно отметить ветвь A из региона Sasun и примыкающего района о. Ван. Далее следует ветвь C из региона Sasun. И самая «молодая» ветвь B из региона Gardman и Ararat Valey.

Неудивительно, что представители гаплогруппы T субклада M184 связали свою судьбу с армянской национальностью, т.к. на карте видно, что носители гаплотипов выборки проживают в регионе, который со II по I столетие до н.э. носил название Великой Армении.

Вывод о векторе распространения этого субклада в Закавказье делать преждевременно. Здесь большую помощь оказали бы ископаемые гаплотипы. Также остается открытым вопрос об автохтонности представителей гаплогруппы T в Закавказье. Но можно сделать осторожное предположение об участии представителей гаплогруппы T в жизни Куро-Аракской археологической культуры, и также предполагать о тесной связи с представителями R1b и J2, по крайней мере, в Закавказье, потому что датировки пребывания древних носителей этих гаплотипов примерно совпадают.

## Литература

Клёсов А.А. (2011а) ДНК-генеалогия основных гаплогрупп мужской половины человечества (Часть 2). Вестник Российской Академии ДНК-генеалогии, 4, No. 7, 1367-1394.

Клёсов А.А. (2011b) Гаплогруппы и гаплотипы Армении (гаплогруппы J2, R1b-L23, R1b-M269 и T-M184). Вестник Российской Академии ДНК-генеалогии, 4, No. 10, 1985-1993.

Herrera, K.J., Lowery, R.K., Hadden, L., Calderon, S., Chiou, C., Yepiskoposyan, L., Regueiro, M., Underhill, P.A., Herrera, R.J. (2011) Neolithic patrilineal signals indicate that the Armenian plateau was repopulated by agriculturalists. *Eur. J. Human Genetics*, doi:10.1038/ejhg.2011.192 (16 November 2011).

Klyosov, A.A. (2009) DNA Genealogy, mutation rates, and some historical evidences written in Y-chromosome. I. Basic principles and the method. *J. Genetic Genealogy*, 5, 186-216.

# **Об удивительных и неожиданных языковых совпадениях между лакским и аккадским языками. Топонимика и личные имена лаков**

**Р.А. Омариева**

(Окончание; начало см. в Вестнике том 4, № 8, стр. 1584-1592 и обсуждение стр. 1592-1603; продолжение – в Вестнике том 4, №9, стр. 1776-1782)

Лакцы проживают в высокогорной части Дагестана, в Кулинском и Лакском районах. Третий, еще один район лаков – Новолакский – был создан искусственно, часть лакцев в добровольно–принудительном порядке была переселена на земли репрессированных и отправленных в Среднюю Азию чеченцев, после реабилитации которых во избежание конфликта с соседями и справедливости ради правительством Дагестана было принято решение вновь переселить лаков, выделив им земельные наделы на территории Дагестана и освободить чеченские земли в пользу вернувшихся владельцев.

Территория лаков начинается с узкого живописного Цудахарского ущелья, села горные sarhu-saravallu – шарху-шяраваллу лаков, так мы называем свои села, расположены в долине реки по горным склонам. Многие села опустели и разрушены, массовый отток населения идет последние 15-20 лет - последствия очередной революции в России, якобы бескровной, но на деле унесшей не менее жизней, чем октябрьская революция, если учесть и то обстоятельство, что целые поколения потенциальные просто не появились на свет в стрессовые 90-е годы - годы распада СССР. Я не знаю, как все происходило исторически - лаки изначально проживали на Кавказе, потом мигрировали на территорию Междуречья, или миграция шла наоборот, из Междуречья. Знаю, что Кавказ обитаем был издревле, наскальному солнечному календарю, обнаруженному западной экспедицией, насчитали 8000 лет, да и дагестанскому Дербенту не менее, а может и более 5000 лет по последним данным. Существует странная, на уровне каких-то физических полей, энергетики, привязанность лаков к высокогорью. От длительности пребывания поколениями в горах лаки как-будто физически срослись с горами. Новолакские лаки - переселенцы, о которых я говорила выше, на благодатных землях низменного Дагестана, где зрели абрикосы, виноград, орехи, ностальгировали, скучали по суровому высокогорью и никак не хотели принять прелести равнинной жизни.

Как бы и куда бы лаки ни мигрировали - с Кавказа в Междуречье или обратно, топонимика и личные имена лаков являют собой повторы топонимики и личных имен аккадцев, древних жителей Месопотамии, отраженных в трудах западных ассириологов, в частности - И. Гельба, внесшего существенный вклад в составление Чикаго-Ассирийского словаря. Если взять названия горных лакских сел и попытаться выяснить происхождение названия, без Месопотамских корней выяснить ничего не удастся. Само слово Месопотамия представляет греческую кальку Междуречья, называемого аборигенами BET NAHRAIN. Здесь bet - лак. корень «бит-ан» - оставить, поставить, поселить; nahrain-лак. нех, нехру-река, реки. нехрай-на реках, нехрайн- на реки.

Древние названия лакских сел смысловой расшифровки не имеют, вернее, эта расшифровка утеряна. Существуют всякие домыслы и предположения, никак не подтверждаемые, например, село Вачи Кулинского района произносится Ваччи, Вачч-алу - названо якобы в честь албанского царя Ваче, тогда, спрашивается, где название предыдущее? Албания - это сравнительно недавние события на фоне истории Междуречья. Еще село Цлуцар, цлу-сса- «новое», щар- «село, поселение» - позволяет предположить данную интерпретацию. Но и это название фигурирует среди географических названий Месопотамии. В лакском языке сохранились также названия стран, упоминаемых на аккадском языке, в неизменном виде: Египет - Мисри, mat MISRI у аккадцев, Сирия - Шам, SAMEYA, Персия - Парс, mat PARSU, Йемен- Ямани, mat. YAMINA, - кстати, понятого толкователями САД как Иония.

Географическим названиям Месопотамии предшествуют определяющие слова: kur -территория, если название идет как название территории ( Лак. кьур- поле, территория), mat- страна, если название идет как название страны ( мат - страна у лаков сохранилось только в составе слов, например, «хлуку-мат»- правительство, или в личных именах- Мусли-мат. Ума-мат) sarrî -город, если указывается город ( Лак. щар- село, населенный пункт). Названия топонимов Месопотамии я обозначила заглавной латинской буквой, обозначающей том словаря на эту букву, откуда взят топоним, и далее цифрой, указывающей страницу этого тома.

Горное ущелье, где начинается территория лаков, называется Цлахьар, Цлахьар-алу. Название SAHARATU имеется в САД, S, 36, толкование - «название месяца». Если в текстах будет встречено название SAHAR, можно будет полагать, что это все же географическое название, SAHAR-atu означает тогда из SAHAR.

Село Ц1увк1ул – их два села, одно из них лаки называют Т1аннул Ц1увк1ул. S,356 упоминается LU SUKKAL, Tannulli SUKKAL, о которых говорится temti halki- народ дерева, древесины,- возможно, в связи с занятием поставками древесины или столярным ремеслом.( Т-1, 175 tannu –сущ. деревянный сосуд; Лак. Т1анну-л – деревянный, из дерева). Название LU SUKKAL, LU -нижний, предполагает наличие еще одного SUCCAL, возможно, верхнего, и объясняет наличие двух сел с одинаковыми названиями «Ц1увк1ул» в маленьком Кулинском районе.

Село Суммат1 - S,379, географическое название SUMMATU.

Село ХЪювхъи.- В,274, kur HAVHI.

Село Сук1иях-S, 398, SUQA- географическое название , судя по производным SUQAYA, SUQAITU – (выходец из Suqa).

Село Кули – А-2, 117, KULLAB. Название горы Kullar около озера Урмия отразилось в лакской топонимике как Къуллар-ду, и, видимо, к названию «Кули» отношения не имеет.

По устным преданиям (письменные источники были уничтожены при Советах, как религиозная ересь- см. ниже), существовали села Куркли, Ч1уч1абах-алу, Ч1ютра-Эя, Ччитал Мащи, Атрал Тарац1. Кули было образовано объединением этих сел и название его означает «все, общность». Как ни удивительно, слово R,504,kullatu – «сущ. общность, все количество» как бы подтверждает эту версию, и остановимся пока на ней.

Должна буду сказать, что некоторые версии названий имеют предварительный характер, работа с текстами в САД сулит много интересного, и информация может быть уточнена.

Село Хъусра, вероятно, имеет отношение к названию реки HUSUR- KHOSR- в латинской транскрипции, в Месопотамии, упоминается в В, 209, возможно, было аналогичное поселение и еще встретится в текстах. Село Гъумук- Кумух -упоминается в словаре, G 132 как географическое название на территории Ура GUMUH, или GUDMUH. Также входит в список сирийско-хаттских городов в Анатолии из англоязычной Википедии (статья Месопотамия), - как город KUMMUH. Видимо, к этому Анатолийскому варианту Кумуха относятся описания древних историков, повествующих о богатом и прекрасном городе,- горное село Кумух никак не создает впечатления, что стоит на развалинах древнего города, обычное горное село в долине реки, как и остальные села.

В этом же списке находим QUWE – аналог топонима села Лакского района Къува, GURGUM – аналог названия села Ккурккул, KARCHEMISH – Хъурхъи. Там же BARSIP – аналог названия села Бурши; HILAKKU- аналог

названия Гьалк1расси. НАМАТН- аналог названия Гьамияци. Там же город SAMHAL, аналог дагестанского Шамхала.

Далее перейдем к топонимам из САД.

Село К1унди-ми, Кунди, S-2, 102, mat KUNDI.

Село Къян, Кани-В,331, mat CANAAN.

Село Къуви- Куви-В,209, HUWAWA.

Село К1увур-Кувур- В,209, KUBU.

Село К1амахъал-Камакал-В,209, КАММАКЕНЕ.

Село Виц1хъи. Вицки-В,386, WIZIKU, UIZIKU, KUR UIZIKU.

Село Ххути- Хути- Н,265,hutaru=hattu,-народ хатты и mat НАТТИ.

Село Ххури- Хури К,279, mat HURRI.

Село Ккут1нищи – Кутнищи-К497, QATNA.

Село Варайми –Варими-S-2, 104, mat WARIM.

Село Ури – UR.

Село Куми – KUME.

Что касается наиболее древних и самых известных названий городов Месопотамии, также известных в ареале племен, их названия сохранены также у лаков в родовых именах и топонимике отдельных сел. Например, город Аккад в греческой транскрипции выглядит как Archad, эта транскрипция представляется более правдивой в силу непосредственных контактов древних греков с Месопотамией. Название соответствует названию лакского села Арчутти- ныне уже, похоже, не существующего, нашла в словаре Хайдакова 1962 года издания, а в списки современных лакских сел не входит. В местной топонимике Топоним UR-SA-TU, как название реки, и AR-SA-TI-A -как название поля, находим также в словнике И.Гельба, стр.77.

Многие топонимы имеют повторы в отдельных селах, к ним относится и данный топоним. Например, в Кулинской топонимике есть Кума (Kume) Хъахъия-зунгту, Хъахъия- рат1 (Haha) Арч-алу (Archad) Хана-ц1алу (Hana) Улия –рат1 –речка Ulai в Месопотамии.

Вот несколько топонимических названий лаков из словаря Хайдакова, который, несмотря на наивный перевод названий на русский язык по созвучию слов, хотя бы собрал воедино большую часть лакской топонимики.

Кумар-алу –Kumar,

Кунукли-ялу –Kunuk,

Ккуртта- ялу- Kurta,



Къюни-ялу -Huni,  
 Къуллар-дахалу- Kullar,  
 К1икъубав-алу- Kikuba,  
 Лув- Ц1урттазани -Lu-Surtazani,  
 Лярхъал-алу -Larha,  
 Мяри-ялу -Mari,  
 Палах-алу- Palah,  
 Суси-ялу - Susi,  
 Ттуккул-к1иллу - Tukkul killu,  
 Эя -халу Ea,  
 Ясария -Yasaria,  
 Яшар - Yasar,  
 Т1аннил -Маршрабу-къур -Tannil-Marsrabu-kur,  
 Лясат1и- Lasati,  
 Шамхал-алу -Samhal,  
 Эял-дара - Ea-dara,  
 Сухъа- Suqa,  
 Сарану-халу Saranu,  
 Ляхар- Lahar,  
 Масру -Masru,  
 Уттакки-ялу -Uttakku,  
 Ттуккул-Ламу-рат1 -Tukkul Lamu-rat,  
 Куркулла-ялу -Kurkul,  
 Аькул- алу -Akullu,  
 Аттара -ц1алу- Attara,  
 Арц-алу- Arsi,  
 Арцул-алу Arsu,  
 Яриппа-ялу- Yarippa.

Это маленькая часть лакской топонимики из словаря Хайдакова, по которой видно явное совпадение названий с названиями в Месопотамии и сходство образования топонимов – возможно, многие из них, или все - присутствовали в Месопотамии в названиях территорий и населенных пунктов. Топонимы Kumar, Kunuk, Kullar, Ea, Susi, Mari, Akullu, Arsi, Suqa, Attara (Addaru), Samhal, Kikuba, -однозначно присутствуют в топонимике Месопотамии и отражены в САД и в словнике И.Гельба. Определение остальных топонимов лаков – вошедших в этот маленький список и оставшихся вне поля зрения остальных - просто дело времени. Дефисом я отделила словообразующие части, уточняющие основное слово-название: - алу, -ялу- место. Например, «нех» -река, «нех»-алу, «нех»-аялу- место на реке, «нех»-ац1алу-место у реки, «нех»-ахалу место за рекой.

Название города Урук в Википедии в греческой транскрипции выглядит как Orchoe, Orugeia. Это соотносится с лакским «Оьргъ» - так лакцы называют жителей села златокузнецов Кубачи, вероятно, выходцев в историческом прошлом из Урука.

Этимология топонима «Кубачи»: Географическое название Kubarra -река упоминается в А-2,528. Далее. К,481,находим слово kubahatu- переведено как «топографический термин». Означает это из -за Kuba. Кубачи восходит, видимо, к названию реки Kubarra- как Хосрех Хъусра - к названию реки Khosr. Хосрехцев аналогично называют Хъусра-щи. Также один из самых благородных родов в Кули носит название «Оьргъу». Другие родовые имена связаны названиями древних городов. Род «Кишитлу» в Кули-из города Kish, род «Ккуркку»- из Gurgum, род «Марххащи» из Marhashe, род Nihru - выходцы из Nihru, А-2, 94, Nihru.

Есть родовые имена, связанные с происхождением из разных народов на территории Месопотамии: род «Сюти» - из Сутиев, род «Къуте» -из Кутиев, род «Каччи» -из Касситов, род «Маннан» - из Маннеев. Некоторые лакские топонимы соответствуют в словаре САД словам с определенным толкователями смыслом. Но, поскольку переводы в САД крайне ненадежны, за исключением слов из списков синонимов, хотя и тут есть несоответствия,- упор на эти переводы делать не стоит, пока тема не проработана достаточно.

Относительно личных имен лаков. Буквально все имена, мужские и женские, за исключением позднейших имен, данных детям в честь героев кинофильмов, или заимствованных славянских имен - все имена соответствуют аккадским личным именам, систематизированным И.Гельбом в труде «Glossary of old akkadian».изданном в 1957 году Чикагским университетом. Исключение - имя Дингир: мной это имя встречено только в одном детском стишке в какой-то детской книжке, где говорится о некоем загадочном «Дингир- Дангиричу»-Дингире из Дангира. Запомнилось, потому что непонятно было, кто это, поскольку имя такое у лаков не встречается. Так вот, кроме этого загадочного Дингира, все имена лаков, в частности кулинцев, присутствуют в словнике И.Гельба, как имена аккадцев. Но следует иметь в виду, что эти имена у нас большей частью вышли из употребления, и являются уже редкими, устаревшими, остающимися только на могильных плитах.

Процесс предания забвению этнических имен начался с насаждения ислама, правоверные должны были, например, называть сыновей в честь пророка Мухаммеда, и на сегодняшний день почти каждый мужчина или Магомед, или Могомедов. Имена Х1абилла, Х1апис, Аьбид, Аьллу, Аьлил,

женские Баху, Амму, Умму, Бахттун, К1ук1у, Халун, Бича, Шаме, Гьуме, Х1юри, - вот часть преданных забвению этнических имен, которые носили еще древние аккадцы. Родовые имена кулинцев- Мурачу-Murasu, Ахъушан-Ahusan, Риххи-Rihi, Къап1уда -Naruda, Аьмат1у -Amatu, Къурти -Kurti, Аьк1и-Aqi, Нана-Гада -Nana-Gada, Къурти-База- Kurti-Baza,- которыми мы в детстве обзывали обидевшего нас представителя рода, и тот, оскорбленный, также «обзывался» в ответ, оказались вполне обычными именами людей, которые, возможно, были достойными основателями своих родов. Даже те имена, которые присутствуют у лаков в дополнение к основному имени с оттенком прозвища, и они присутствуют в древнем Аккаде, похоже, также дополняя другое, основное имя - например, имена Абур, Къудур (имя, которое мы знаем по текстам учебников истории как Навухудоносор, в оригинальной записи выглядит как Nabu-Kudurru-Nasur).

Лаки - народ, потерявший знания о своих истоках, впрочем, как и другие дагестанцы. За период менее чем полвека дагестанские этносы становились поголовно безграмотными два-три раза. Будучи осколками племен, они имели языком образования и межплеменного общения арабский и тюркский языки на основе арабской письменности. После революции арабское письмо заменили латиницей, и все стали поголовно безграмотными. Далее перешли на кириллицу, снова все стали безграмотными. Параллельно уничтожались книги на арабском языке в русле борьбы с «религией-опиумом для народа». Так были уничтожены также книги, содержащие сведения о населении, его истории, родословные. Бабушка моя по отцу, Х1юрия, долго хранила огромные с побуревшими от старости страницами фолианты на лакском языке арабской вязью. Иногда брала, читала мне повествования о каких-то войнах и сражениях, много имен, которые мне ничего не говорили, одна из книг- в стихотворной форме. Но я не понимала, слушала только из вежливости - чтобы бабушку не обидеть.

Постоянно в селе появлялись скупщики- то книг, то старинных украшений и шелковых платков, ковров - все отобрали хитрые скупщики у доверчивых кулинцев. Последние годы скупщики охотились за старинной медной утварью. Огромные «саргас»-блюда медные под метр и более метра в диаметре, в них сушат травы и «зиниччат1» для приготовления домашнего пива, меньшие по размеру блюда «аму» и «ланжари», разные котлы с расписными крышами, подойники, также с расписанными стенками, далее- всевозможные сосуды для воды, для еды,- короче вся посуда для применения в доме из меди, луженой оловом, - присутствовала в каждом хозяйстве. Молодежь, считающая это все ненужной рухлядью в доме, все сдала скупщикам. Таким образом, каких то реликвий, письменных

источников или вещей, связанных со своей историей, у лаков не осталось практически ничего.

Полное имущественное разорение причинено было войсками Хромого Тимура - по преданиям, кованые сундуки женщин были полностью выпотрошены и содержимое забрано в качестве трофеев. Сундуки для девушек на выданье – это традиция, сохранившаяся в горных селах до сих пор. Сундук деревянный, украшенный ковкой из металла, готовят для девушки к замужеству заранее. Украшения и подарки, полученные от родных, складываются туда. Это, как правило, серебряные кольца, серьги, браслеты, пояса для женской праздничной одежды, украшения, нашиваемые на одежду, на наволосники, бусы коралловые и из разных камней.

Лакские ювелиры предпочитали работать с серебром, чем с золотом - якобы за особую пластичность серебра, и таких вещей у жителей было достаточно даже после монголо-татарского грабежа. Ценность их была в том, что они поколениями переходили от матери к дочери, и они могли представлять собой образцы ювелирного искусства далекой старины. Вот это тоже большей частью ушло в руки скупщиков.

Все, что у нас осталось для того, чтобы определить нас как оригинальный этнос - это наш язык, оскудевший, заброшенный. Но - пока еще живой. Не будет сел в горах - его не станет, потому что городские лакцы не придают значения обучению своих детей родному языку. Одни - в силу суматошного ритма жизни, обрекающего людей на напряженное физическое выживание, другие - в силу достигнутой при «строительстве капитализма» в России финансовой мощи и появившегося пренебрежения к своим корням, таким же незначительным на их взгляд, как и их бедные родственники. Я же намерена работать с Чикаго-Ассирийским словарем и дальше, по крупинке, выяснять историю народа и языка.

Благодарю редакцию и сотрудников «Вестника» за предоставленную мне возможность высказаться на отличную от направления журнала, но где-то пограничную тему.

Р. Омариева.

# Исследования мтДНК

## Предисловие редактора

Доктор биологических наук А.Ф. Назарова представила статью о мтДНК в относительно небольшой популяции русского (в основном) городка. Скептик может сказать, а что это даёт – городок маленький, на самом деле деревенька в несколько дворов (50 человек населения 15 лет назад, сейчас, наверное, и того нет), что за статистика?

На это отвечаю, что это мог бы быть интересный материал. Никто и не собирается рассматривать эту деревеньку как репрезентативную для Российской Федерации или для русского населения в целом. Но это – маленькая часть большой мозаики, и без таких частей не будет и мозаики. Далее, А.Ф. Назарова – профессиональный популяционный генетик, и интересно посмотреть, как популяционные генетики подают и рассматривают подобный материал. Далее, выявленные и показанные здесь гаплотипы мтДНК – характерные для русского населения, хотя, повторяю, не могут приниматься как полностью отражающие структуру мтДНК русского народа, но никто противоположного и не утверждает. Наконец, это материал по мтДНК, который на страницах Вестника представлен крайне редко. Там есть свои проблемы, часть которых я и постараюсь затронуть в этом Предисловии, поскольку они не затронуты в самой статье.

С мтДНК есть занятная ситуация – они (молекулы митохондриальной ДНК) значительно легче определяются, но в отношении ДНК-генеалогии несравненно хуже изучены. Причина проста – популяционные генетики не умеют работать с кинетикой мутаций, а значит – и с датировками. В этом их никак нельзя обвинять, это просто не их специальность. Их дело – выявить популяцию для изучения, выполнить формальности типа получить юридическое согласие испытуемых в письменном виде, собрать тестируемый материал (кровь, слюну), определить гаплотипы и гаплогруппы тестируемых, сделать свои популяционно-генетические заключения и выводы. Но для того, чтобы делать датировки, они должны как минимум уметь определять константы скоростей мутаций. Они этого не умеют, за исключением нескольких человек в мире, да и те никак не договорятся, как это делать.

Вот пример того, как это делается в популяционной генетике теми, кто не знают, как это делать (ниже будет дан пример тех, кто знают, но ответ далеко не окончательный).

В своей работе «Этногеномика и генетическая история народов Восточной Европы» (Вестник Российской академии наук, 73, 614-621, 2003) , на которую, кстати, дана ссылка в статье А.Ф. Назаровой ниже, Э.К. Хуснутдинова, чл.-корр. Академии наук Башкортостана, зав. отделом Института биохимии и генетики Уфимского научного центра РАН, пишет следующее:

*«По ориентировочным оценкам, возраст дивергенции линий, выявленных у народов Волго-Уральского региона, варьировал от  $273 \pm 57$  тыс. лет для азиатской линии Z до  $22.76 \pm 5.250$  тыс. лет для линии C. Возраст дивергенции самой крупной европейской линии H определен в  $20.036 \pm 4.250$  тыс. лет, что соответствует археологическому времени повторной экспансии населения на территории Урала в постледниковый период.»*

Вот это «по ориентировочным оценкам» и показывает, что автор не знает. Как считали и кто считал, и на основании чего считали – в статье Э. Хуснутдиновой нет. То, что «азиатская линия Z» произошла 273 тысяч лет назад (плюс-минус как указано), довольно занятно, поскольку гаплогруппа Z – это третье «поколение» митохондриальных гаплогрупп  $L \rightarrow M \rightarrow Z$

Но поскольку не указано, как считали, кто считал и на основании чего, то забавным и остается. Особенно потому что в недавней работе (Soares et al, "Correcting for purifying selection: an improved human mitochondrial molecular clock", The American Journal of Human Genetics, 84, 740-759, 2009) сообщается, что возраст гаплогруппы Z составляет 24,300 лет, и дается 95%-ный доверительный интервал между 15,400 и 33,600 лет. Но уж никак не 273 тысячи лет назад.

Кстати, самая древняя гаплогруппа мтДНК , L0, в последней работе датируется 149,700 лет назад (112,200 до 188,000 лет), а последующая гаплогруппа M в Восточной Азии датируется 60,600 лет назад (47,300-74,300 лет), в Южной Азии – 49,400 (39,000-60,200) лет.

Я не знаю, у кого датировки точнее – у Э. Хуснутдиновой или у Педро Соареса, но что-то мне подсказывает, что у второго будут более близкие к реальности. Но разница между 273 тысячи лет и 24 тысячи лет назад как-то плохо укладывается в голову, тем более что в той же своей статье Э. Хуснутдинова пишет, что человек вышел из Африки «около 60-70 тыс лет назад». Кто-нибудь что-нибудь в этом понимает, что написала член-корреспондент Академии наук Башкортостана?

Кем и как определен «возраст дивергенции самой крупной европейской линии H  $20.036 \pm 4.250$  тыс. лет» тоже осталось неизвестным, но это уже ближе к другим оценкам. Например, в той же работе (Soares et al, 2009) эта

датировка дается как 18,600 (14,700-22,600) лет. Wikipedia, правда, дает “over 35,000 years ago”, но цену Википедиям мы знаем. С другой стороны, это показывает полный хаос в популяционной генетике в отношении датировок.

В общем, читатель уже понял, что все эти расчеты популяционными генетиками и жонглирование цифрами, особенно без ссылок и указаний на то, как это делалось, не имеют никакого смысла.

Продолжение цитаты более-менее понятно:

*«Используя данные о числе мутационных замен и скорости накопления мутации для гипервариабельного участка мтДНК, равной одной мутационной замене за 20.18 тыс. лет, мы получили среднее значение времени дивергенции мтДНК для народов Волго-Уральского региона. Оно составляет 49.60 тыс. лет назад, что соответствует периоду расселения человека на европейском континенте в эпоху верхнего палеолита».*

Это более-менее понятно потому, что эта величина, 20180 лет на одну мутацию, действительно принята в известном программном обеспечении Network. Это – скорость мутации в гипервариабельном участке мтДНК (HVS-I), между позициями 16090 и 16365. Но в любом случае из цитаты Э. Хуснутдиновой неясно, что такое «мы получили среднее значение времени дивергенции мтДНК для народов Волго-Уральского региона», поскольку неясно, о каких гаплогруппах речь. У всех же гаплогрупп разный возраст. Как можно усреднять возраст, скажем, отца одного и прадедушки другого, и что это даст? Поэтому «Оно составляет 49.60 тыс. лет назад, что соответствует периоду расселения человека на европейском континенте в эпоху верхнего палеолита» остается «вещью в себе». Что за человек расселялся? Какая у него была гаплогруппа? Впрочем, это типичный подход популяционной генетики – смешать несмешиваемое и вычислить нечто неусредняемое. Типа той же «популяционной скорости Л. Животовского».

Теперь о том, почему я ссылаюсь на работу (Soares et al, 2009) и доверяю ей больше. Да потому, что это наиболее подробное и фундаментальное исследование, в котором проведен анализ предыдущих подходов и результатов. Не факт, что и оно верное, но это, видимо, лучшее из того, что сейчас есть.

И это тоже одна их причин, по которой статья А.Ф. Назаровой оказалась кстати для публикации – чтобы обсудить мтДНК для целей ДНК-генеалогии и для изучения популяций в том ключе, в каком приведено в статье ниже.

Как сообщают Soares et al (2009) одна мутация во всей мтДНК происходит в среднем один раз в 3624 года. Понятно, что при такой медленной скорости мутации много ДНК-генеалогии не получится, только для самых древних линий, или для очень крупных массивов мтДНК, идущих от одного общего предка. Но откуда такие цифры и насколько они обоснованы? По данным (Soares et al, 2009), они получены при изучении 2196 полных геномов мтДНК и откалиброваны по человеку и шимпанзе, а также по древним стоянкам на Канарских островах, в Океании и в Европе до и после ледникового периода. Это дало при проверке время расселения человека в Америке ~ 15 тысяч лет назад, что не противоречит археологическим свидетельствам. Правда, за дату «расселения человека из Африки 55-70 тысяч лет назад» авторы, видимо, принимают время бета-гаплогруппы 64,000±6,000 лет назад, то есть время появления внеафриканского человека, гаплогруппы ВТ. Авторы в 2009 году не знали, что у гаплогрупп ВТ нет африканских снипов (снипов гаплогруппы А), и принимали их за «расселение из Африки». Но датировка, как видим, вполне удовлетворительна, если не абсолютно точна в пределах погрешности. Эта же калибровка дала время расхождения мтДНК современного человека и неандертальца 550,000±54,000 лет назад, что тоже не противоречит массиву оценок в этом отношении.

Правда, затем авторы вводят понятие «синонимных молекулярных часов», которые отклоняются от «линейной скорости мутации» посредством дополнительного коэффициента, что делает скорость мутации зависящей от времени, и таким образом «откорректированной» за счет “purifying selection”. Я так и не понял, идет ли речь о введении поправки на возвратные мутации, что вполне возможно, судя по используемым поправочным формулам, содержащим экспоненциальные компоненты и рассуждениям о «скрытых мутациях» (hidden mutations). Применяя эти коэффициенты, авторы получают из скорости мутации 3624 лет<sup>-1</sup> (одна мутация в 3624 года) скорость мутации одна в 7557 лет при поправочном коэффициенте 47.94% в одном случае, и 7790 лет в другом случае. Например, между современным человеком и шимпанзе наблюдаются 965 «синонимных мутаций», но при введении поправки получается 2108 «всех синонимных мутаций». Тогда средняя скорость мутации получается одна мутация в 6687 лет (2108 мутаций за 14 миллионов лет). Еще одна величина скорости мутации, приводимая авторами (Soares et al, 2009) – одна в 9503 лет. Средняя величина, сообщаемая авторами, равна одна мутация в 7885 лет.

Понятно, что эти упражнения никак не способствуют принятию данных скоростей мутации для ДНК-генеалогических расчетов с использованием мтДНК.



Возвращаемся к статье А.Ф. Назаровой. Она приводит данные, что население деревни Сараево Ярославской области имеет следующие гаплогруппы мтДНК (числа округлены):

- 38% Н
- 19% W
- 19% Т (Т\* и Т1)
- по одной (из 16) preV2, I, U, X.

Вот сокращенная схема образования этих гаплогрупп

$L \rightarrow M + N$  (L и M в деревне не оказалось, только дочерние гаплогруппы N)

$M \rightarrow (C, Z, E, G, Q, D)$  - ни одной в деревне не оказалось

$N \rightarrow I, W, Y, X, R, A, S$

$R \rightarrow R0 + JT + U + P + F + B$

$JT \rightarrow J, T$

$U \rightarrow K$  (в деревне не оказалось)

$R0 \rightarrow H, V$

Для краткой справки приведем примерный возраст соответствующих гаплогрупп, найденных в деревне Сараево (возраст - из статьи Soares et al, 2009):

- U - 54,000 лет
- X - 31,800 лет
- Т\* - 26,800 лет
- I - 26,300 лет
- W - 21,200 лет
- H - 18,600 лет
- Т1 - 14,700 лет

После этого введения можно переходить и к самой статье А.Ф. Назаровой.

# **Исследование мт ДНК, популяционная и демографическая генетика, русской сельской популяции Ярославской области в прошлом и настоящем**

**А. Ф. Назарова**

**Институт проблем экологии и эволюции РАН, Москва  
afnazar@yandex.ru**

## **Введение**

Около 10 лет назад мы начали популяционно-генетическое изучение сельской популяции Ярославской области, которое к настоящему времени дополнилось генетико-демографическими и молекулярно-генетическими исследованиями. В данной работе мы исследовали полиморфизм митохондриальной ДНК (мтДНК) русской сельской популяции Ярославской области, провели тотальный генеалогический анализ (по результатам опроса населения) этой популяции.

## **Материалы и методы**

Материалом для исследования полиморфизма мтДНК послужили пробы волос населения Сараево, собранные А.Ф.Назаровой во время экспедиций в Ярославскую область 2003–2004 г.г. Скрининг полиморфных сайтов, определяющих основные группы типов мтДНК, распространенных в популяциях Евразии (табл.1), проводился с помощью рестрикционного анализа участков мтДНК, амплифицированных в полимеразной цепной реакции. Рестрикционные фрагменты фракционировались электрофоретически в 8% полиакриламидных гелях. Для детекции ДНК использовалась окраска гелей бромистым этидием с последующей визуализацией ДНК в УФ-свете. Полиморфизм учитывался по наличию (+) и отсутствию (-) сайтов рестрикции ДНК. Анализ полиморфизма мтДНК проводился на базе лаборатории генетики Института биологических проблем Севера ДВО РАН (г.Магадан) Б.А.Малярчуком, М.В.Деренко и А.В.Лункиной.

Материалом для исследования демографо-генеалогической ситуации в популяциях Ярославской области послужили данные тотального

генеалогического анализа популяции д. Сараево и демографо-генеалогического анализа популяции с. Заозерье Ростовского района, проведенных А.Ф.Назаровой и С.М.Алхутовым во время экспедиций 1991-1994 г.г., и продолжавшегося во время последующих экспедиций 1995-2000 г.г. и 2001-2004 г.г. А.Ф.Назаровой и М.Г.Кузнецовой. Материалы по популяции Заозерья опубликованы нами ранее (Назарова и др., 1996).

**Таблица 1**

**Полиморфные рестрикционные варианты, определяющие группы типов мтДНК у населения Евразии**

Группы мтДНК	Маркеры
H	-14766 <i>MseI</i> , – 7025 <i>AluI</i>
pre*V1	-14766 <i>MseI</i> , -15904 <i>MseI</i> , -16297 <i>MseI</i> , +4577 <i>NlaIII</i>
pre*V2	-14766 <i>MseI</i> , +15904 <i>MseI</i> , -16297 <i>MseI</i> , +4577 <i>NlaIII</i>
V	-14766 <i>MseI</i> , +15904 <i>MseI</i> , -16297 <i>MseI</i> , -4577 <i>NlaIII</i>
HV*	-14766 <i>MseI</i>
U	+12308 <i>Hinfl</i>
K	+12308 <i>Hinfl</i> , -9052 <i>HaeII</i> , +10394 <i>DdeI</i>
J	-13704 <i>BstNI</i> , +10394 <i>DdeI</i>
T*	+13366 <i>BamHI</i> , +15606 <i>AluI</i>
T1	+13366 <i>BamHI</i> , +15606 <i>AluI</i> , -12629 <i>AvaII</i>
N1	-12498 <i>NlaIII</i>
I	-4529 <i>HaeII</i> , +8249 <i>AvaII</i> , +10032 <i>AluI</i> , +10394 <i>DdeI</i> , -12498 <i>NlaIII</i> , +16389 <i>BamHI</i>
W	+8249 <i>AvaII</i> , -8994 <i>HaeIII</i>
X	+14465 <i>AccI</i>
C	+10394 <i>DdeI</i> , +10397 <i>AluI</i> , -13259 <i>HincII</i> /+13262 <i>AluI</i>
D	+10394 <i>DdeI</i> , +10397 <i>AluI</i> , -5176 <i>AluI</i>
G	+10394 <i>DdeI</i> , +10397 <i>AluI</i> , +4830 <i>HaeII</i> /+4831 <i>HhaI</i>
E	+10394 <i>DdeI</i> , +10397 <i>AluI</i> , -7598 <i>HhaI</i>
M*	+10394 <i>DdeI</i> , +10397 <i>AluI</i>
A	+ 663 <i>HaeIII</i>
B	9 п.н. делеция в регионе V
F	-12406 <i>HpaI</i> / <i>HincII</i>

**Примечание.** Группы мтДНК обозначены в соответствии с классификацией, предложенной в работах (Macaulay et al, 1999; Richards et al, 2000; Malyarchuk et al, 2003). Указаны позиции полиморфных рестрикционных сайтов относительно кембриджской последовательности мтДНК человека (Anderson et al, 1981).

В генеалогическую схему Сараево были включены более 300 человек, в подавляющем большинстве русские. Генеалогический анализ проводили путем опроса жителей деревни. Число изученных поколений по отдельным семьям 5-6, в остальных не менее четырех.

## **Результаты исследований и их обсуждение**

Популяция деревни Сараево Переславского района Ярославской области была исследована нами практически тотально в генеалогическом (по опросу населения) и демографическом аспектах. Основана эта деревня была очень давно, находится она в интересном в историческом отношении регионе. Во времена позднего палеолита (6000-5000 лет до н.э.) на территории будущей Ярославской области обитали племена первобытных охотников и рыболовов. В Бронзовом веке (2000-1000 лет до н.э.) этот регион был освоен племенами скотоводов Фатьяновской археологической культуры (Горюнова, 1961). И только чуть более 1000 лет назад местное финно-угорское население меря стало смешиваться со славянами, пришедшими в этот регион. В дальнейшем, во время татаро-монгольского ига, граница между территорией, контролируемой татарами-монголами, и землями, охраняемыми русскими войсками, проходила по реке Нерли, текущей менее чем в 20 км севернее Сараево. Антропологическая гетерогенность русской популяции (по морфологическим признакам) показана в работах (Алексеева, 1973, 1999).

Для выяснения глубинной этнической природы местного населения, русского по официальной принадлежности, но явно имеющего по фенотипическим признакам и финно-угорские (мерянские) корни, мы исследовали пробы митохондриальной ДНК волос представителей популяции Сараево. Учитывая, что объем популяции Сараево на 1995 год составлял всего 50 человек (речь идет только о постоянном населении), выборка, исследованная в плане полиморфизма митохондриальной ДНК составила 17 человек. В Табл. 2 приведены гаплогруппы мтДНК населения Сараево. Как видно из таблицы 2, практически все исследованные лица из популяции Сараево несут гаплогруппы митохондриальной ДНК, в основном обнаруживаемые у европейского населения. Наиболее частой в

этой выборке является гаплогруппа Н. Гаплогруппы мтДНК, обнаруженные в популяции Сараево, относятся в основном к макрогруппе R: это гаплогруппы Н, pre-V, U, Т. У трех индивидуумов обнаружена гаплогруппа W и у одного – гаплогруппа X, относящиеся к макрогруппе N. Как и в тотальной популяции русских (Малярчук, 2003), в населении Сараево наиболее часты гаплогруппы Н и Т.

В работе по исследованию мтДНК русских ряда областей России Малярчук и Деренко (2007) установили, что наиболее часто у русских встречаются гаплогруппы Н, U, Т и J - те же гаплогруппы, которые широко распространены в генофондах других европейских народов. Примесь монголоидных гаплогрупп у русских незначительна.

Возраст дивергенции самой крупной европейской линии Н определен как 20 тысяч лет (цит. по Хуснутдинова, 2003), что соответствует археологическому времени повторной экспансии населения на территории Урала в послеледниковый период (там же). Используя данные о числе мутационных замен и скорости накопления мутаций для гипервариабельного участка мтДНК, равной одной мутационной замене за 20 тыс. лет, Хуснутдинова (там же) получила среднее время дивергенции мтДНК для народов Волго-Уральского региона. Оно составляет 49 тыс. лет назад, что соответствует периоду расселения человека на европейском континенте в эпоху верхнего палеолита.

Монголоидных гаплогрупп (А, В, С и D) в популяции Сараево мы не обнаружили, что может говорить об отсутствии вклада монголоидных женщин в формирование популяции Сараево. Это вполне объяснимо, поскольку во время татаро-монгольского нашествия и последующего ига на Руси присутствовал контингент татаро-монгольских войск, представленный мужчинами, и генное влияние монголоидов, коль скоро таковое могло быть, должно было идти через ядерный, а не митохондриальный геном.

Основателями современной популяции Сараево являются три или четыре семьи, которые обитали здесь по крайней мере с 18-го века (*гаплогруппы автором не представлены – прим. редактора*). Это семьи Белоусовых, Садовниковых, Муравьевых и Бурловых. Менее часта в этой популяции фамилия Сусловы. Национальный состав – практически полностью русские, однако в последних двух поколениях, ставших почти полностью мигрантами в другие поселения и регионы, встречаются браки с мордовцами, немцами и узбеками. Под номерами 1, 5 и 8 в табл. 2 присутствуют маятниковые мигранты российского происхождения, из них номер 1 предположительно имеет предков из данного региона, а номера 5 и 8 – с юга России: речь идет о предках в давних поколениях. Гаплогруппы

этих лиц оказались соответствующими происхождению их предков: номер 1 – гаплогруппа Н (наиболее часта в выборке Сараево), номер 5 – pre V, номер 8 – U. Последние две гаплогруппы, pre\*V2 и U, наиболее часты в популяциях русских юга России, как показано Малярчуком (2003).

Итак, генетическая история популяции Сараево привела к формированию европеоидной группы населения с гаплогруппами митохондриальной ДНК, характерными для русской популяции в целом. В работе Малярчука (2003) показано, что в общей популяции русских, исследованных в Краснодарском и Ставропольском краях, Белгородской, Орловской, Саратовской и Нижегородской областях обнаруживаются макрогруппы мтДНК R, N и M. В изученной выборке Сараево встречены две из трех макрогрупп, R и N.

**Таблица 2.**

Гаплогруппы митохондриальной ДНК населения д.Сараево Ярославской области

Номера исследованных лиц	Гаплогруппа мтДНК
+ 1	H
2	W
3	I
4	H
+ 5	(pre*V2)
+ 6	H
7	H
+ 8	U
9	X
10	T*
11	T*
12	W
13	H
14	H
15	W
16	T1

На основании генеалогического анализа населения д. Сараево мы составили развернутую родословную, в которую вошли более 390 человек – как из предшествующих поколений, так и ныне живущие. Дети и внуки большинства лиц, живущих сейчас в популяции Сараево, в основном мигрировали в другие сельские населенные пункты и города России. Мы выявили 5 браков между троюродными сибсами в популяции Сараево. Инбредных браков других типов (между двоюродными сибсами, между дядей и племянницей и т.д.) в этой популяции не оказалось. Коэффициент инбридинга, вычисленный нами по генеалогическим данным по методу С.Райта (Spuhler, 1976),  $F$  оказался равным 0,0002018. Если сопоставить это значение с коэффициентом инбридинга других человеческих популяций, например с популяцией северных европеоидов Швеции, где много сельских изолятов, то в Пайоле  $F = 0,0008$ , а в Мунионалусте  $F = 0,0058$  (Малярчук и Деренко, 2007). Ранее нами была показана гетерогенность русской популяции (Назарова, 1994), поэтому возьмем для сравнения также гетерогенную в связи с ее происхождением бразильскую популяцию. В сельской бразильской популяции Парнамирины  $F = 0,01563$ . Меньшее значение коэффициента инбридинга в исследованной нами сельской популяции, видимо, связано с отсутствием в русской деревне браков между двоюродными сибсами, тогда как доля таковых в бразильской популяции равна 19,55%, а в шведских – 0,95% и 6, 8%, соответственно. Итак, можно заключить, что инбридинг в русской сельской популяции Ярославской области незначителен, и меньше такового в ряде зарубежных сельских популяций Европы и Америки.

Автор выражает благодарность Б.А.Малярчуку, М.В.Деренко и А.В.Лункиной (Институт биологических проблем Севера ДВО РАН, г. Магадан) за помощь в проведении данного исследования.

### *Литература.*

Алексеева Т.И. // Этногенез восточных славян по данным антропологии.. – Изд-во МГУ.- 1973.- 329 С.

Алексеева Т.И. Этногенез и этническая история восточных славян по данным антропологии. // В кн.: Восточные славяне. Антропология и этническая история. – М.- Научный мир.-1999.- 307-315.

Горюнова Е.И. // Этническая история Волго-Окского междуречья. М.- 1961.- Изд-во АН СССР.

Малярчук Б.А., Деренко М.В. Структура русского генофонда. Природа, № 4, 2007.

Малярчук Б.А. Изменчивость митохондриального генома человека в аспекте генетической истории славян. // Докт. дисс...- М.- 2003.

Назарова А.Ф. Популяционная генетика русских: генеалогический анализ, частоты генов и генетические расстояния. // Доклады РАН.- 1994.- 339, № 4.- С.563-568.

Назарова А.Ф., Алхутов С.М., Машуров А.М. Популяционно-генетическое исследование русской деревни: генеалогический анализ, инбридинг и индексы потенциального отбора двух сельских популяций Ярославской области. // Доклады РАН.- 1996.- 349, № 1.- с. 133-137.

Хуснутдинова Э.К. Этногеномика и генетическая история народов Восточной Европы. Вестник Российской Академии наук, 2003, т.73, № 7, с. 614-621.

Anderson S., Bankier A.T., Barrell B.G. et al. Sequence and organization of the human mitochondrial genome // Nature. 1981. V. 290. P. 457-465.

Macaulay V., Richards M., Hickey E. et al. The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs // Am. J. Hum. Genet. 1999. V. 64. P. 232-249.

Malyarchuk B.A., Grzybowski T., Derenko M.V. et al. Mitochondrial DNA variability in Bosnians and Slovenians // Ann. Hum. Genet. 2003. V. 67. P. 412-425.

Richards M., Macaulay V., Hickey E. et al. Tracing European founder lineages in the Near Eastern mtDNA pool // Am. J. Hum. Genet. 2000. V. 67. P. 1251-1276.

Spuhler J.F. The maximum opportunity for natural selection in human populations. // Demographic anthropology. – Univ. of New Mexico Press – 1976.- P.185.



# ДНК-генеалогия и письма из Киргизии (Хрестоматийные «что делать» и «с чего начать»)

Чинара Сейдахматова и Анатолий Клёсов

## Краткое предисловие

Автор приведенных ниже писем дала согласие на их публикацию в качестве обсуждения исходных вопросов ДНК-генеалогии. На мой взгляд, это хороший пример отношения заинтересованного человека, особенно в регионе, где вопросы генеалогии традиционно имеют особую значимость.

## Письмо 1

(27 ноября 2011)

Здравствуйте, уважаемый Анатолий Алексеевич. Простите мне мою дерзость, но я не могла удержаться от сумасшествия, в которое я впала, прочитав Вашу статью «Древние арии: кто они были и откуда?» (Примечание – статья была опубликована в Вестнике, ноябрь 2008 г., том 1, №5, с. 908-928, и в книге «Интернет: Заметки научного сотрудника, 2010, Изд. Московского университета). Возможности Вашей науки поражают и обнадёживают.

Меня зовут Чинара Сейдахматова. Я живу в Киргизии, киргизка по происхождению. Последние семь лет занимаюсь изучением родоплеменного устройства киргизов и собираю традиционное знание киргизов по вопросу нашего происхождения. Вы, наверно не знаете, что у киргизов есть живое генеалогическое древо, восходящее почти за 1000 лет поименно, т.е., если я знаю порядка 14 отцов, дальше по общему древу, могу узнать остальных за 1000 лет вперёд. По сей день, сохранена духовная традиция – каждый уважающий себя киргиз из хорошего рода обязан знать семь своих отцов. Благодаря этому сохраняется древо.

У нас очень много мифов о формировании нашего этноса и его древности, но меня беспокоит не древность нашего происхождения, а степень родства наших родов. Дело в том, что есть легенда, в которой говорится, что сорок кочевых родов собрались и назвали киргизами, признав единые духовные законы Тенгрианства. Некоторые знатоки «санжира» (так называется наше древо) говорят, что наши рода не имеют кровного родства, что три-четыре стволых рода («Бугу», «Саяки», «Солто» и ???) объединились и к ним

примкнули остальные. Для нас очень важна степень родства этих родов.

Конечно, очень много мифов об арийскости киргизов, но я никогда не могла понять, а почему хорошо быть арием. Прочитав Вашу статью, я так и не поняла, почему весь мир хочет быть к ним причастен?

Мой вопрос заключается в следующем: могут ли пользоваться Вашей методикой и вести такие же исследования другие группы генетиков? Если могут, то можно ли с Вами связаться для консультаций и на каких условиях? Если не могут, могли бы Вы помочь нам организовать подобные исследования? Если «да», то, что для этого необходимо с нашей стороны и в виде каких ресурсов? Насколько сложное техническое сопровождение такого проекта?

Я занимаюсь широкой общественной деятельностью, обладаю информацией, насколько актуальны сегодня для нас подобные исследования. Они могли бы положить конец раздору в нашей стране. Ведь все мы в конце концов братья. В нашей стране лет десять живёт один американский еврей, который не жалеет собственных денег, на популяризацию идеи братства киргизов с евреями, он утверждает, что киргизы - утерянное двенадцатое колено евреев. Я не антисемитка, меня это мало волнует, возможно, и братья, но это не поможет оздоровлению общества. А вот если мы выявим кровное родство киргизских родов между собой, а затем родство с калмаками (это - наше киргизское произношение), казахами и другими соседями кочевниками, мы сумели бы воспользоваться такой информацией в целях консолидации общества.

С восхищением и благодарностью, Чинара Сейдахматова.

### **Мой ответ:**

Уважаемая Чинара,

Спасибо за добрые слова. То, что Вы пишете, очень интересно и важно.

*>Конечно, очень много мифов об арийскости киргизов, но я никогда не могла понять, а почему хорошо быть арием, прочитав Вашу статью, я так и не поняла, почему весь мир хочет быть к ним причастен?*

Вы не поняли, потому что в статье этого нет и быть не могло, что "весь мир хочет быть к ним причастен". Это, конечно, не так. Скорее всего, арии вызывают интерес, потому что загадке происхождения ариев уже 200 лет, и

только сейчас она уже в значительной степени решена, причем именно благодаря ДНК-генеалогии. Более того, тест на ДНК очень простой, и можно "в одно касание" узнать, принадлежит ли мужчина к потомкам этого загадочного рода R1a1, к которому принадлежали арии. Всегда интересно узнать, что принадлежишь к столь древнему роду, о котором уже известно так много - и датировки, и миграции, и происхождение, и генеалогические ветви и подветви, и можно играть в увлекательную игру, сравнивая гаплотипы (фрагменты Y-хромосомы) длиной до 67, и сейчас уже и до 111 звеньев (маркеров), и узнавая, кто близкий родственник, а кто относительно далекий, и все в пределах одного и того же рода, рода легендарных ариев.

Естественно, это интересно. Но это касается далеко не только ариев, а вообще любых легендарных исторических личностей и племен. Вы думаете, меньше интерес найти себя потомками майя? Или викингов? Или пророка Мухаммада? (правда, у него не было прямых потомков, но близкие родственники были, например, двоюродный брат Али ибн Аби Талиб, и сыновья последнего Хассан и Хуссейн, родоначальники Сейидов). Арабы хотят оказаться Сейидами. Евреи - потомками Аарона (козны). Или людям неинтересно было бы узнать, что они прямые потомки Александра Македонского? Или Юлия Цезаря (хотя детей у него не было)? Такова человеческая натура, нравится это кому или не нравится. Я, например, потомок детей боярских, и потратил массу времени на изучение средневековых (и позже) архивов, чтобы до этого докопаться и выяснить детали. И все началось с того, что я узнал, что тоже принадлежу гаплогруппе R1a1, но задавал себе вопрос - а что было после ариев?

Так что дело не в ариях, хотя это тоже интересно, а в стремлении многих людей узнать больше о своих корнях. У мужчин это, кстати, значительно более выражено, чем у женщин, и это "экспериментальный факт".

Арии интересны и потому, что они представляют яркий пример рода древности, внесшего значительный вклад в современную цивилизацию, распространившего по планете индоевропейские языки, проложившего культурно-исторический мост между Европой и Азией еще тысячелетия назад. Считалось, что они безвозвратно исчезли, сгинули, а сейчас оказалось, что вовсе нет. Прямые потомки рода ариев - среди нас. Кстати, по современным данным (которые, конечно, будут уточняться) у киргизов примерно 63% гаплогруппы R1a1, то есть почти две трети. Остальная часть расходится по гаплогруппе N1b (это - древняя, азиатская, алтайская, сибирская, уральская ветвь), R1b (всего 1.3%), это - предки большинства современных европейцев, а до того - башкир, кавказцев (грузин, армян,

дагестанцев), турок, легендарных шумеров (и потомков их – многих современных ассирийцев), возможно, и части египетских фараонов.

Так что на ариях свет клином далеко не сошелся. Просто это один из ярких примеров загадок древности. А загадок еще много. Кто были, например, енисейские киргизы? Почему древность киргизских R1a1 обрывается примерно 1300 лет назад, то есть в 8-м веке нашей эры, плюс-минус пара столетий? А что было раньше, и вообще что произошло 1300 лет назад с киргизами? Как видите, с историческими ариями у них большой разрыв, в два тысячелетия. Что случилось в середине-конце первого тысячелетия? Геноцид? Переселение? У меня ответа пока нет.

Генетики знают, как определять гаплотипы-гаплогруппы, если они генетики. Знают они и то, с кем нужно связаться для закупки оборудования, хотя у генетиков оно обычно есть. Если нет - это недешево, может достигать нескольких сотен тысяч долларов. Но, повторяю, это их вопрос. Для расчетов и интерпретации результатов тестирования я, конечно, могу принять участие. Если это немногие случаи, то никаких условий у меня нет, и уж во всяком случае финансовых. Если это станет значительной по времени работой, то к условиям можно вернуться.

*>В нашей стране лет десять живёт один американский еврей, который не жалеет собственных денег, на популяризацию идеи братства киргизов с евреями, он утверждает, что киргизы утерянное двенадцатое колено евреев.*

Мне эта безумная "теория", конечно, знакома, но это не так. Никакого основания у нее нет, кроме фантастических предположений, никак не проверяемых. Сказав это, я, конечно, не могу исключить, что среди жителей Киргизии есть евреи, но это совершенно другое дело. Есть и еще один поворот такого предположения, но он окажется неожиданным и для самих евреев. Дело в том, что есть евреи гаплогруппы R1a1, той же самой, что и киргизы гаплогруппы R1a1. Эти евреи, конечно, никакое не "колono" и к "кошерным" евреям не относятся. Это - потомки древних ариев, которые прошли по Ближнему Востоку еще до времен Авраама, когда евреев как таковых еще не было. Другая миграционная ветвь ариев оставила потомков в Киргизии. Общий предок евреев R1a1 и киргизов R1a1 жил 4500 лет назад, скорее всего, на Русской равнине. Так получается при расчете, основанном на структуре гаплотипов R1a1 евреев и киргизов - кстати, и у тех, и других линия обрывается 1300 лет назад, но сами линии евреев и киргизов рода R1a1 очень удалены друг от друга. Поэтому и имеют общего предка 4500 лет назад. Как понимаете, на одно "колono" это никак не тянет, но вот два колono в гаплогруппе R1a1 из евреев и киргизов мы уже имеем. Это помимо

еще двух десятков других "колен" в гаплогруппе R1a1, арийской (условно говоря) гаплогруппе.

*>Я не антисемитка меня это мало волнует, возможно, и братья, но это не поможет оздоровлению общества.*

Да, это был бы необычный ракурс в решении задачи "оздоровлении общества", а именно киргизского общества.

*>А вот если мы выявим кровное родство киргизских родов между собой, а затем родство с калмаками, казахами и другими соседями кочевниками, мы сумели бы воспользоваться такой информацией в целях консолидации общества.*

На мой взгляд, подход идеалистический, но зависит от того, как это разыгрывать. В гражданской войне в России 90 лет назад русские убивали русских, и знание общего родства им не помешало. То же самое в любой гражданской войне. Евреи и арабы вообще к одному роду принадлежат, как и армяне и турки. Однако ненависть порой зашкаливает. Тем не менее, при прочих равных факторах общее родство может помочь восстановить мир, с этим я согласен. Вообще в неустойчивых ситуациях порой достаточно небольшого фактора, чтобы в целом перевесило.

Так что удачи Вам. Всего наилучшего. Да, еще один вопрос. Ваше письмо содержит интересные данные и вообще показывает заинтересованный взгляд неравнодушного человека. У нас есть ежемесячное издание "Вестник Российской Академии ДНК-генеалогии", в котором есть соответствующие разделы - и переписка с читателями, и мнения, и дискуссии. Если Вы не возражаете, я бы поместил Ваше письмо в следующем номере, вместе с моим ответом. Нет возражений?

Спасибо.

### **Продолжение переписки:**

Несказанно рада Вашему ответу, честно говоря, и не надеялась. Конечно, я согласна на Ваше предложение по журналу. В свою очередь, хочу просить Вашего разрешения на расширение круга нашей переписки. Есть интересующиеся, можно я включу их в рассылку? Я правильно поняла, что мы можем получить доступ к журнальной переписке?

Не могли бы Вы рекомендовать какую-нибудь Вашу статью на общее понимание древности человеческого рода и его миграционных схем? Несколько лет назад мне попадалась статья о генетических исследованиях

древнейших этносов, по-моему Масачусетского университета, где давалась схема родственных групп. Меня тогда удивило то, что киргизы были в одной группе со славянами и японцами, а все народы тюркской языковой группы оказались в разных генетических группах. Даже с казахами, которых мы считаем близкими родственниками, мы были в разных группах. Схема была очень выразительной в виде цветка из четырёх крупных лепестков и четырёх более мелких между ними. Сейчас, когда я стала более серьёзно интересоваться этими вопросами, я не знаю где искать ту схему. Возможно в Ваших работах есть об этом?

Очень интересно, где проживают самые древние рода? У киргизов много легенд, как кочевали и где осели наши родственные племена. Было бы здорово получить подтверждение. Например, род "Бугу" (оленей), есть на Алтае, высоко в горах Гималая и Балканах. Некоторые алтайские киргизы (для нас они киргизы, мы считаем себя одним народом с алтайцами) считают, что "Бугу" с калмаками близкие братья, хотя здесь на Тянь-Шане наши киргизы имеют образ калмаков непримиримыми врагами. У нас различают российских калмыков и наших "сарт калмаков", которые и есть тот самый «вражеский образ».

Мне кажется, сегодня веление времени, когда все ищут свои корни, своих собратьев. Есть необъяснимое стремление причастности к большому сообществу. Возможно, это глубинный стадный инстинкт, который спал доселе. Может быть, на таком фоне изменения общественной психологии, наши идеи консолидации, на основании причастности к единым корням, не так уж идеалистичны.

Очень много вопросов, надеюсь получить ответы в Ваших статьях. С благодарностью и почтением, Чинара.

### **Мой ответ:**

Литература обширная, но подходы разные и разная трактовка и на разных уровнях понимания. Если на популярном уровне - то подборка моих статей на эту тему есть на моем сайте <http://aklyosov.home.comcast.net>, там в последней части сайта галерея статей и журнала "Вестник Российской Академии ДНК-генеалогии". Пусть Вас не смущают названия линков на английском языке, там тексты на русском. Там же есть книга "Происхождение славян и других народов - очерки ДНК-генеалогии", она так и не издана, руки не доходят. Еще - моя книга "Интернет - заметки научного сотрудника" (изд-во Московского университета, 2010, 512 стр.),

там часть книги - популярные очерки по тематике ДНК-генеалогии (и родственным вопросам).

Если на более научном уровне - моя обзорная статья в журнале "Биохимия" (Москва), это - старейший академический журнал, 2011 г., том 76, выпуск 5, стр. 636-653; книга "Происхождение человека" (А.А. Клёсов, А.А. Тюняев, изд-во "Белые Альвы", 2010, 1022 стр.).

Если на совсем научном уровне - мои (и с соавторами) статьи на англ. языке в зарубежных академических изданиях, ссылки есть на сайте, но не думаю, что Вам нужно это сейчас. Почему я даю ссылки только на свои работы - потому что у других - другие аспекты смежных наук. Это подробно пояснено в статье в журнале "Биохимия".

То, что Вы написали о некой статье Массачусеттского университета, я бы не стал тратить время. Есть масса всяческих суррогатов. В общем, все они познавательны в той или иной степени, но все полны басен, фантазий, подтасовок и манипуляций.

Что касается проживания древних родов - это только начинает выясняться. Поэтому мало-мальски полного или достоверного ответа Вы нигде не получите. Мы над этим работаем, но нужны данные по ДНК, а их мало (хотя поступают, но больше на Западе, там люди более подвижны).

### **Продолжение переписки:**

Большое спасибо за подборки, меня интересовали именно Ваши статьи, для популярного чтения. Теперь займусь изучением. Скажите, пожалуйста, как к Вам поступают данные ДНК? Мы могли бы этому способствовать, только сначала нам на месте надо сорганизоваться, понять самим и донести сообществу для чего нам это нужно. Мы, киргизы, очень носимся со своей древностью и чистотой крови (я в том числе), но у меня есть подозрения, что степень ассимиляции очень высокая. Ведь миграционный путь наших предков длителен и обширен. В принципе, аргумент в пользу чистоты крови существует: 1) традиционные знания генного влияния со стороны матери и отца, какие качества на каком ребёнке и как отражаются. 2) традиции подбора невест для сына - это целая наука, 3) традиция сохранения древа (обязательное знание семи отцов), 4) мифы имеют очень детализированные описания истории и путей миграции наших предков. В ближайшее время, после изучения материала, мы с коллегами, продумаем дальнейшие шаги и поставим Вас в известность.

С уважением и пожеланием успехов,

**Мой ответ:**

В отношении "поступления ДНК" есть два пути - определять самим или использовать базы данных. Киргизов в них очень мало. Чтобы понять, что это такое - ниже линк на большую базу данных, с сотней тысяч гаплотипов (или уже больше)

<http://www.ysearch.org>

при входе откроется панель, в которой две нижних "кнопки" слева - для введения имени или индекса. Один, например, это WMJ46 , другой E4QS4. Оба - киргизы. Вот они знают, как заказывать анализ ДНК. Знаю и я, конечно, но знают и Ваши генетики.

Желаю успеха.



## DISCUSSIONS

### The mutation rate constants in DNA genealogy and related issues

Arranged by  
**Anatole A. Klyosov**  
<http://aklyosov.home.comcast.net>

A lengthy discussion on the mutation rate constants and related issues took place in the RootsWeb Forum in November 2011. After exchanging some comments under the "TMRCA Calculations" subject, Andrew McEachern has noticed:

**Andrew McEachern:**

>... although Anatole Klyosov's method has been repeatedly criticised no one as yet has proven it incorrect in terms of genealogy. I still use the method and it works for me down to a point where I am comfortable that the analysis resulting from his methodology pretty well follows what we know of history of certain family groups down to about 1000AD.

I have even read on other forums that it is believed that the methodology is quite sound for the last two millennium. And that was written by one of his chief protagonists on this forum.

Why don't we have a TMRCA challenge? We have an extremely well known family group being Clan Donald. They have exhaustive and verifiable genealogies going down to about 1350AD. We know who their common ancestor was and we know exactly when he lived and we know that their family group has three main branches and we even know with reasonable certainty the common ancestor date for each of the main branches.

So get out your TMRCA calcs and show the methodology and working out and give us some dates using all these TMRCA methodologies. Lets see who is near and who is far.

"One of chief protagonists" has immediately responded:

**Kenneth Nordtvedt:**

>...I have indicated in a number of messages that as far as TMRCA estimates are concerned, it is close enough to the variance method to make his estimates ok. Where I have problems with his numbers are the impossibly small sigmas or "confidence intervals" he gives. The true sigmas are substantially larger.

There are serious limitations of TMRCA estimates in both the small and large time limits. For small times the intrinsic sigmas are large as fraction of the estimate. For long times the faster STRs don't behave as simple boxes with a single mutation rate the same for all repeat values and up as well as down.

**MY COMMENT - Anatole Klyosov** (NOTE: Since I consider materials of this Discussion as having a serious educational value, I will comment "off-line" on some of statements placed in the Forum).

Above is a typical comment by K. Nordtvedt. They are never specific, never illustrated by "real life" examples, never exemplified with actual haplotypes and actual calculations. Not once I have challenged K. Nordtvedt with actual "real life" examples and suggested to calculate "true sigmas" from actual experimental data just to provide support to his words, and he never did it. I can only guess that he knew that as soon as he provided actual calculations, I would immediately corner him. So, he decided to be empty-worded but negative.

An interesting detail: while calculated sigma in dozens and hundreds of my TMRCA calculations NEVER goes below 10%+, and it is typically within 11-14% in the best cases and often goes above 15-25% in ordinary cases, K. Nordtvedt listed HIS sigma as 13.3% and then corrected it to 7.7%.

Here as it was going:

KN: "This gives for sigma/G of square root  $(2[G-Gstar]/mNstrG^2)$ . For  $G-Gstar = 10$  generations,  $m = 1/300$ ,  $G = 100$ ,  $Nstr = 50$ , the correct formula gives 13.3 percent. "

His opponent: "I can't get the 13.3 percent from your example. Maybe a typo".

KN: "Typo indeed.... Fortunately when I evaluated the 13.3 percent I had already caught the wayward "2". So you should get the 13.2 now I think.

His opponent: "I get 7.7% now. Maybe your 13.3 % was the mistake?"

KN: "7.7 percent is correct this evening."

So, as we see, my >10% were "impossibly small sigmas... The true sigmas are substantially larger". His 13.3% or even 7.7% sigmas are O.K. Thus passes "scientific" honesty.

**Sasson Margaliot:**

>Ken wrote:

>Where I have problems with his numbers are the impossibly small sigmas or >"confidence intervals" he gives. The true sigmas are substantially larger.

... Anatole would only provide his estimates after \*personally\* verifying that the database has no offending sub-clusters leading to (so called) "phantom" estimates. There is no formula that gives "sigma" for such "phantom-free" case.

**MY COMMENT:**

Sasson is right in his first statement and wrong in the second one. Indeed, before calculations I usually sort out haplotype datasets using specially developed criteria, in order to make sure that the dataset contains haplotypes descended from one (and only one, in terms of DNA genealogy) common ancestor. This significantly reduces a margin of error. Regarding that "there is no formula... for such "phantom-free" case" - it is an incorrect statement. The formula is given in (Klyosov, A.A. (2009) *DNA Genealogy, mutation rates, and some historical evidences written in Y-chromosome. I. Basic principles and the method. J. Genetic Genealogy, 5, 186-216*).

Then **Kenneth Nordtvedt** has announced his next in line "theoretical" utility, with some imaginary features:

>I have added another "Challenge" excel dataset 4408... Someone wanted to see how much or if more STRs might increase the accuracy of tree reconstruction from a sample of final haplotypes. So Challenge4408 has haplotypes of 150 STRs with total mutation rate of about .4 and with individual rates varying from about .01 down to .0001. ... I originally thought people who have software or algorithms for tree reconstruction might want a test bed to see how their tools work in the context of bushy trees of the type which are often the object of reconstruction attempts these days....

**Anatole Klyosov:**

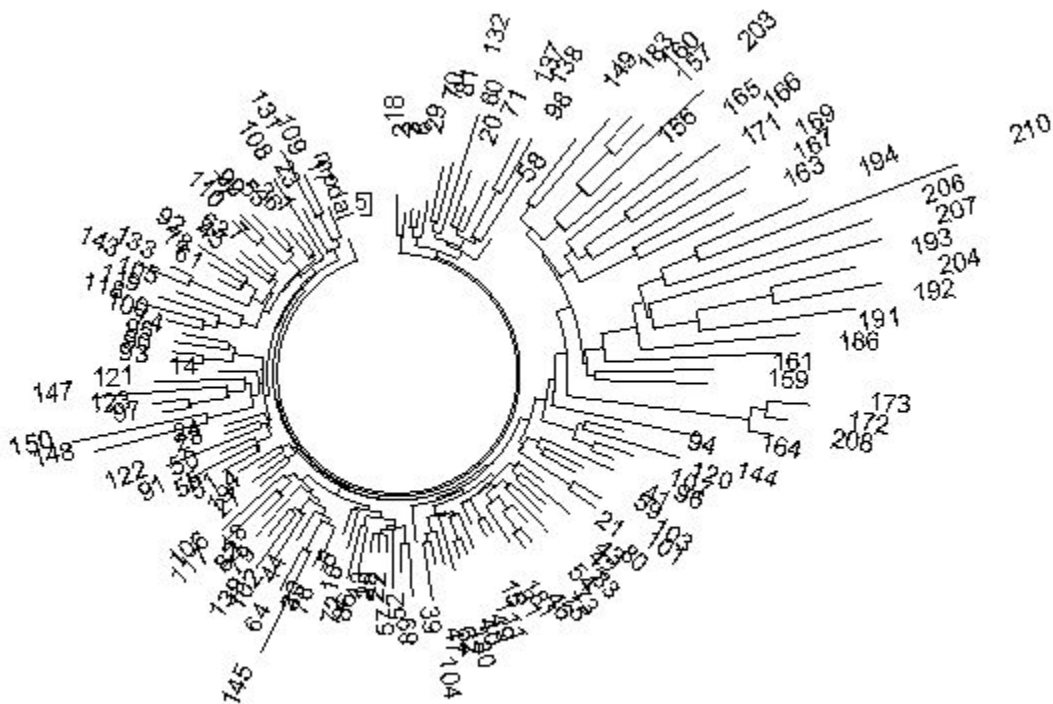
Those "challenges" are solved long ago. It is waste of time to consider "bushy trees" since in reality they typically contain branches of different size and different "ages". Therefore, an analysis of a "bushy tree" in its entirety produces, as a rule, a phantom "common ancestor", and that is why a margin of error must be huge and the whole exercise useless. As a result,

fables on "different mutation rates for different haplogroups", "huge confidence intervals in TMRCA's", etc.

Instead of walking in circles considering "bushy trees" all these years and complaining on "huge confidence intervals", one better take ACTUAL genealogy data, ACTUAL haplotype datasets, and compare actual dates with those resulted from DNA genealogy. This will show what ACTUAL margins of error look like. With "bushy trees", they should be first subdivided on separate branches, and each branch should be analyzed individually.

In fact, it is done on dozens of datasets.

Here is an example. The Donald Clan dataset (red, green, and yellow groups) contains 214 haplotypes (effective this week), with 136 of them of 67 marker haplotypes. A haplotype tree shows that the dataset in fact splits into 14 branches (including mini-branches [mini-lineages] of four-five haplotypes, however, having their distinct position of the tree).



The Donald Clan 67 marker haplotype tree (red, green, and yellow groups) of 136 haplotypes. The whole dataset contains 214 haplotypes, however, 78 of them are determined in 12, 25 or 37 marker format.

In many cases those color groups were mixed up in the original dataset. All 14 branches have been analyzed, TMRCA's were obtained with their confidence intervals. The main (principal) branch contains 39 haplotypes, with the base haplotype as follows:

13 25 15 11 11 14 12 12 10 14 11 31 -- 16 8 10 11 11 23 14 20 31 12 15 15  
 16 - 11 12 19 21 17 16 17 18 34 39 12 11 - 11 8 17 17 8 12 10 8 11 10 12 22  
 22 15 11 12 12 13 8 14 23 21 12 12 11 13 11 11 12 12

No wonder, it is the "modal" haplotype on the original list, since the branch is the largest one. It is the same haplotype we identified earlier with our publication (2010) with Andrew MacEacharn.

Those 39 haplotypes contain 27, 64, 120, and 159 mutations in the first 12, 25, 37, and 67 marker haplotypes. This gives

$27/39/0.020 = 35 \rightarrow 36$  generations, or  $900 \pm 195$  ybp, or  
 $64/39/0.046 = 36 \rightarrow 37$  generations, or  $925 \pm 150$  ybp, or  
 $120/39/0.09 = 34 \rightarrow 35$  generations, or  $875 \pm 120$  ybp, or  
 $159/39/0.12 = 34 \rightarrow 35$  generations, or  $875 \pm 110$  ybp.

Please notice how all the four lines fit each other with their TMRCA. The most reliable is, of course, the 67 marker series, and it gives 2011 minus 875 = 1136 AD. As you might remember, Somerled lived 1100-1164 AD.

(The denominators show the mutation rate constants for the respective 12-, 25-, 37- and 67-marker haplotypes.  $\rightarrow$  indicates a correction for back mutations according to the published (2009) Table).

Another large 19-haplotype branch, with a base haplotype:

13 25 15 11 11 14 12 12 10 14 11 31 -- 16 8 10 11 11 23 14 20 31 12 15 15  
 16 - 11 12 19 21 17 16 17 18 34 38 12 11 - 11 8 17 17 8 12 10 8 11 10 12 22  
 22 16 11 12 12 13 8 14 23 21 12 12 11 13 11 11 12 12

It differs quite distinctly from the Somerled direct DNA-lineage in two markers. All nineteen haplotypes have only 60 mutations which gives  $60/19/0.12 = 26 \rightarrow 27$  gen, =  $675 \pm 110$  ybp. Obviously, they are also derived from Somerled, but through an intermediate common ancestor who had those two mutations compared to Somerled. It might be from John, Lord of the Isles.

Another large, 23-haplotype branch, with a base haplotype:

13 25 15 11 11 14 12 12 10 14 11 31 -- 16 8 10 11 11 23 14 20 31 12 15 15  
16 - 11 12 19 21 17 16 17 19 34 39 12 11 - 11 8 17 17 8 12 10 8 11 10 12 22  
22 15 11 12 12 13 8 14 23 21 12 12 11 13 11 11 12 12

This is recent lineage, 50 mutations in all 23 haplotypes. It gives  $50/23/0.12 = 18$  gen, or  $450 \pm 80$  years to a common ancestor.

And so on. The green and yellow groups (actual branches) were analyzed as well.

**Didier Vernade:**

I have 2 questions; one regarding this example: Is the Mc Donald tree known by paper records or was it derived by some comparative method? Marko Heinila is finding differences according to which methods is used to produce the tree. Second question: what about Busby et al.? You didn't react on the many discussions on that paper. Latest post being by Mike W with his variance analysis (a few days ago).

**Anatole Klyosov:**

The Mc Donald tree was composed by me a couple of days ago from the latest (updated) Donald Clan haplotypes edition, currently 136 of 67 marker R1a1 haplotypes. Many programs exist; however, they do the same thing - they sort out haplotypes (or whatever) based on minimization of the difference between the objects. I do not employ those programs for calculations (they might be different in that regard), I just use their principal function, which in this particular case is to split the dataset to a series of branches. The program which I use does it nicely. The "difference" you are talking about might be rotating the tree, for example, and the left side branch next time (when you change setting parameters or add haplotypes) might come up on the right side of the tree, which I do not care about. There are haplotypes which mutations place them in an unstable position between the branches, and they can switch between branches when you add or subtract haplotypes, however, this does not change the pattern and is reflected in the margin of error of the final calculation.

In other words, the tree can be technically different (see above) but branches are practically the same. That is what counts.

"What about Busby et al". What about them? Their paper is very "heterogeneous", and the part with "mutation rates" is a total disaster. In that regard the paper was doomed. Another part about "no geographical trends in diversity" of R1b1b2-M269" in Europe was doomed as well, since they have used bikini-haplotypes. They were not supposed to see anything meaningful. That

was the main reason that I "did not react", since I prefer to make positive remarks in discussions, not negative, unless I am challenged. However, in that particular case I WAS challenged, and, as a result, the detailed analysis of the Busby at al. paper is here:

[http://aklyosov.home.comcast.net/~aklyosov/Vestnik\\_4\\_09.pdf](http://aklyosov.home.comcast.net/~aklyosov/Vestnik_4_09.pdf)

The content of the journal is mainly in Russian, however, the Busby analysis and the following discussion is in English (p. 1831 and further). Technically, the reference is Proceedings of the Russian Academy of DNA Genealogy, vol. 4, No. 9, pp. 1831-1892 (2011).

**Lindsey:**

How to you count mutations? ie if a member of the set has 14 at DYS 458, but the modal is 16 -- do you consider that one mutation or two? What about 464 -- do you count any mathematical difference on this marker as a single mutation? When several multi-copy markers are involved and RecLOH is suspected, do you count all changes as one mutation?

I'd like to try this method with my family which belongs to a small subgroup with a common ancestor who probably lived about 1600 ybp. The oldest lineage in the group goes back to c. 1550, and several others to c. 1580-1670.

Four or five men in the group were tested on 43 markers--do you have a formula for that resolution? If not, they could be compared at the 25 marker level with TMRCA calculation made accordingly.

**Anatole Klyosov**

1. There are pretty straightforward rules for mutation counting. By the way, I typically use 67 marker haplotypes without removing any markers, because they ALL are valuable source of information. Of course, when 67 (or 111) marker haplotypes are not available, I use whatever available. For example, just four marker haplotypes (DYS464a,b,c,d) give excellent and informative results, very distinct for various datasets. For particularly ancient times (between 10,000 and 150,000 years before present) I use the slowest 22 marker panel, in which one mutation occurs on average once in 4,500 years.

2. The word "modal" does not belong to my vocabulary. It is a very confusing term and poorly (and variably) defined. Any "bushy tree" can have a modal haplotype which would drift with changing size of the branches, adding or removing haplotypes, etc. I use the term "base haplotype" which is either the ancestral haplotype of a given population or determined as close to the ancestral as possible. Every branch on a haplotype tree has its own base haplotype,

because every branch on a tree has its own common ancestor.

3. If the base haplotype has DYS458 =16, then a haplotype with DYS458 = 14 has two mutations in this allele. All mutations counted as one-step mutations, except in multi-copy markers. There you count a number of "copies" over the whole multi-copy set of markers. For example, 19-21 → 21-21 is one mutation. 12-15-15-16 → 15-15-15-16 is one mutation. 12-15-15-16 → 16-16-16-16 is two mutations. 11-14 → 11-11 is one mutation. 11-14 → 14-14 is one mutation.

In fact, it is much ado about nothing. In most cases neglecting those multi-copy mutations will add something within the margin of error. So, it is just a matter of being a perfectionist, which feels good in many cases.

4. I certainly have. For the 43 marker haplotype it is 0.0754 mutation/haplotype/conditional generation of 25 years, or  $0.0754/43 = 0.00175$  mutation/marker/generation. If you are a perfectionist, try a double-check: remove four markers of DYS464a,b,c,d, and verify the result for the 39 marker haplotype (this is what some people use when they are afraid of DYS464) with the mutation rate constant of 0.071 per haplotype or 0.00183 per marker. You can verify it again using the 25 marker haplotypes (as you have suggested) using the mutation rate constant of 0.046 per haplotype or 0.00184 per marker per generation of 25 years. I am reminding that those 25 years have nothing to do with "actual generation length" which is a floating thing. Those 25 years here is just a mathematical value. It is bound to the mutation rate constant.

**M.A. Farrel:**

... I have NO generation of only 25 years back to our common ancestor b.1706; these generations were actually from 30 to 35 years apart because they were all men of the ydna line. Women were younger when they married...

**Anatole Klyosov:**

This is a very timely comment. The thing is that we should not compare actual generations (which can be any from, say, 16 to 45 years in length) with "generations" from DNA genealogy, because those are "conditional generations" for the simplicity of calculations. In fact, those "25 years" are a part of the mutation rate constant. Therefore, we should compare YEARS from the documented genealogy with YEARS of the TMRCA. Not "generations".

**Lindsey:**

Thank you, Anatole--I've just tried your calculation for my cousins, six of whom have 43 marker results. The earliest-known ancestor was born in 1672 or 1673 and we have his 43 marker DNA signature by testing one descendant (A) of one



son and five descendants (Group B) of another son. One man from Group B matched A 43/43. Three men from Group B (two brothers and the son of the older brother) matched A on 42 of 43 markers. The fifth man in Group B matched A on the marker where the three men in B differed, but had a mutation not shared by anyone else at DYS464c. I was concerned that the close relationship of the three men (my first cousins) and/or size of the group might throw the calculation off, but when I ran the numbers I got 8.841 (generations). My two first cousins are the 8th generation in descent from the ancestor born in 1672/3 and the son/nephew is the 9th generation. I would have to look up the lineages of the other three men before I could say exactly how many generations removed they are from the common ancestor, but I'm sure that it would be about 8 or 9 generations.

Since the common ancestor of Group B was born in December, 1761, we know that the four mutations in his line occurred in one of several subsequent generations.

Now I'll see what I can do with the subgroup to which this family belongs, where, I regret to say, genealogical information is rather thin. Nonetheless, we are confident that the largest family in the group descends from an ancestor born c. 1550, and as I mentioned earlier, we have several other lines in the group that go back to c. 1600.

It appears to me from the genetic distances involved within the group that some of these families may descend from an ancestor who immediately predated the first surnames in England (c 1100-1200), that the outliers go back to the subclade founder who lived in the early Dark Ages, and that a handful of families are probably the result of NPE's in the past several hundred years. Some of these NPE's are known, some are suspected, and most look as if they might be offshoots of the largest family in the subgroup.

I asked about counting mutations vs measuring genetic distance, because I was under the impression, perhaps incorrectly, that some preferred the former method. Also, in my Britton project, we have an R1b with 15 at DYS447. This was so unusual that I asked FTDNA to check the result again and when 15 was confirmed in a second test, I was told that the difference between 15 and the typical value of 25 should be counted as one rather than ten.

**Anatole Klyosov:**

Of course, the fit between the calculated 8.8 generations and the actual 8-9 generations looks as a pure coincident (and I myself often prefer to think so), however, those fits happen too often to consider them as just coincidents. There is paper in press in which I (and a co-author) have considered many actual

(documented) renealogies, and plotted actual dates vs. calculated ones. The correlation is rather striking, with a correlation coefficient 0.95 for 67 marker haplotypes.

That is why I am all smiling when I read here that it is too complex, too erroneous, and with too wide confidence intervals. Of course, anything can happen, and I as a chemist, biochemist and a biomedical scientist can testify that any lab observes mishaps, deviations, etc. from theoretical considerations.

Re. your next remark, yes, some preferred whatever they want to prefer. There are many ways to handle haplotypes and their mutations. The way I have chosen I optimized for years.

Finally - yes, 25 → 15 you should consider as one mutational event.

**Mary Alice Farrell:**

Thank you for the necessary clarification, Anatole. I do understand this is a method of estimating rate of mutations over great periods of time...

**Anatole Klyosov:**

No, it is not only "over great periods of time". Generally, it is not a right angle at the problem at all.

When I run my experiments in chemical kinetics (this is science about rates and mechanisms of chemical and biochemical reactions, in a nutshell), and if I take only two measurements along the reaction curve, I get, as a rule, an unreliable quantitative description of the reaction. However, I will not blame chemical kinetics for that unreliability. I just generated too few data. Also, I am not going to say, that chemical kinetics works only "over great periods of time". It would be an incorrect statement. Also, if I mix several compounds in my reaction system and get a mess, I am not going to blame chemical kinetics. It would be me to blame, for sloppiness, ignorance, lack of skill, etc.

The same is in DNA genealogy. It all depends on how many "experimental points" you have (that is, how many haplotypes in the dataset) and how you handle them. This comes with years of experience. The same as in any field of science, and elsewhere. However, there is always somebody who says "dreamers". It is not them who actually do science. There are two tools in science: a dream and its proper execution. Scientists are practical dreamers, converting the dream into reality.

You have to understand whether ALL haplotypes in your dataset comes from ONE (and only one) common ancestor (there are criteria for it), which haplotypes

you compare to each other, whether they belong (or might belong) to different DNA-lineages, whether a timespan is long enough to accumulate some back mutations (typically after 600 years), etc. Also, a small number of mutations results in a large margin of error. This is because in a system with only two mutations there could have been easily 1, 0, 2, 3, 4 mutations. Silly statistics.

**Lindsey:**

What are the criteria for determining whether a group of haplotypes comes from one common ancestor when the dataset consists of a number of families with different surnames but presumed, from STR results, to descend from a common ancestor who lived 1500-1600 years ago?

How do you deal with the possibility of back mutations in lineages older than 600 years?

**Anatole Klyosov:**

1. The criterion is the equality of TMRCA calculated by the linear and the logarithmic methods. If the two produce the same TMRCA, the dataset is "uniform", obeys the first order kinetics, and should descend from one common ancestor.

For example, if the haplotype dataset contains 20 of 25 marker haplotypes, and 8 of them are identical to each other (hence, 8 base haplotypes in the dataset), then  $[\ln(20/8)]/0.046 = 20$  conditional generations, that is 500 years to a common ancestor. As you see, you do not need to count mutations here, in the logarithmic method. In the linear method you count mutations, and, suppose, those 20 of 25 marker haplotypes contain 18 mutations. Then  $18/20/0.046 = 20$  conditional generations, the same 500 years to the common ancestor. You are all set, since you have a right dataset, with one common ancestor, whatever surnames they have. If, however, that 20-marker dataset contains, say, 30 mutations, you have  $30/20/0.046 = 33$  generations. End of the story, you cannot continue with your calculations. You have a bunch of different lineages, at least two, and the dataset has a phantom common ancestor.

Unfortunately, that is exactly what population geneticists do, whether they are known, renown, famous or reputable. All of them. They take a number of haplotypes, count mutations, divide by a number of haplotypes, and voila. They get in fact a senseless, phantom "TMRCA". And then they blame mutations which are "unruled". They should blame themselves.

2. After 23 conditional generations ( $23 \times 25 = 575$  years) you observe less mutations in the dataset, than occurred in reality, since a certain fraction of mutations came back and become "invisible". A table is published which shows a correction factor for any given number of generations (Klyosov, J. Genet. Geneal,

2009). For example, 40 generations should be converted to 42 (40 → 42) [all generations are here of 25 years], that is 1000 → 1050 years. 58 → 62 generations, that is 1450 → 1550 years; 100 → 111 generations, that is 2500 → 2775 years, and so on.

**Lindsey:**

I am analysing DNA results from a group of 30 families. For the purpose of this calculation I excluded two families for which only a single 12 marker haplotype is available. For 28 of them, I have 25 marker haplotypes; for 14 I have 43 marker haplotypes, and for 13 I have 67 marker haplotypes. Some families appear in all three sets--some in only one set. The group contains two outliers, both of which are included in the 25 and 43 marker sets, while only one is included in the 67 marker set. Results as follows:

25 Markers, 59 mutations yielding 45.807 G or 1145.186 ybp

43 Markers, 65 mutations yielding 61.576 G or 1539.408 ybp, without the two outliers--45 mutations yielding 42.629 G or 1065.74 ybp

67 Markers, 96.5 mutations yielding 61.858 G or 1546 ybp without the outlier who had 16 or 19 mutations, depending on whether one counts a decrease from 12 to 8 as a single mutation -- 51.602 G = 1290.064 ybp or 49.679 G = 1241.98 ybp

... The 43 and 67 marker calculations of 1500-1600 years are nearly identical to Ken's estimate for common ancestor of this subgroup. I think he said that his database consists of 67 marker results -- I don't know whether he includes families not tested at that level in any of his calculations, but when they are included, TMRCA drops significantly. I also find the decrease when the two outliers are excluded of particular interest, since one of them is almost certainly not English and there is reason to think that the other may not be English either. The names of the rest suggest that they came from England/Britain.

**Anatole Klyosov:**

It was a good try. The only one thing which was forgotten (or neglected) was a correction for back mutation. So, you get 46, 43 and 50 (conditional) generations to a common ancestor, on average  $46.0 \pm 3.5$  generations, that is only 7.6% variation between the three. It is remarkably small margin of error, considering that you took a different number of haplotypes in each panel of markers, and use different mutation rate constants for each panel. However, such a small margin is rather common for calculations, which shows their robustness. If calculations are done properly, they give pretty robust results.

Now, you need to introduce the correction. The table of corrections for back mutations is published in JOGG, 5(2) 186-216 (2009), Table A in the

Appendix:

<http://www.jogg.info/52/files/Klyosov1.pdf>

For 46 (observed) generations it shows →48 generations, that is 1,200 years to a common ancestor.

However, in those situations I typically take a 67 marker set alone (as the most extended one), and others (25, 43, 37 marker sets) I consider as just supportive sets, provided that the results are about the same. In your case it gave 50 → 53 generations, that is 1325±200 years to a common ancestor. I have calculated this 15%-margin of error based on your actual data.

As I have mentioned earlier not once, "quadratic" methods typically give elevated results due to their sensitivity in real haplotype datasets (some admixtures, slight deviations, multi-copy/recLOH effects, etc.). Hence, the elevated 1500-1600 years quoted above. Therefore, higher error margins for "quadratic" methods, in this case 20-30% (at least).

**Lindsey:**

Yes, I did forget the adjustment -- probably because I was working from a list which didn't contain a table for adjustments.

Anyhow, the calculations of 50 and 53 conditional generations were obtained for the 43 and 67 marker sets by deleting the two most distant haplotypes from the 43 marker set and deleting one of those from the 67 marker set.

However, when both of them are included in the 43 marker set and in the 67 marker set, I get 61.576G (not adjusted for back mutations) for 12 haplotypes compared on 43 markers and 61.858G (not adjusted) for 12 haplotypes compared on 67 markers vs 45.807G (not adjusted) when all 28 haplotypes in the group are compared at the 25 marker level. If both are eliminated from the 25 marker set, I get 40.0698G (not adjusted).

Which of these estimates--46G for the complete set -- or 62G for a partial set--do you consider more reliable?

**Anatole Klyosov:**

I thought you gave enough arguments why those haplotypes cannot belong to the dataset. However, the problem of outliers is typically a delicate problem. Sometimes the solution is rather obvious (as I believe in your case, considering

your arguments), sometimes it is more complicated, and, generally, to do "cherry-picking" is not a scientific way.

That is why I have mentioned here (and not once) criteria which separate outliers more scientifically and on a more objective and neutral ground. Two the most direct ones are: (a) to compose a haplotype tree and see which haplotypes are clearly do not belong to the branch, and (b) to compare the TMRCA's obtained from the linear and the logarithmic methods. If after removing haplotypes in question you get much better fit between the both methods, you should remove them.

I understand that the both ways are difficult to handle by a novice, but life is tough. In other ways, here is a borderline between a work and a play.

Lindsey:

Thank you -- I'm glad you consider my arguments sufficient; however, I will also try the logarithmic method on the group. My thought about the outliers has long been that they share an earlier common ancestor with the rest of the group, and that if he lived about 1600 ybp, then the common ancestor of the English group minus the outliers should be several hundred years later.

But the outliers aren't the only hazard in analysing this group -- several families are probably the result of relatively recent NPE's, and quite a few mutations occur toward the end of the 37 marker set or in markers 38-67, yet only 17 of 30 families have been tested either at the 37 or 67 marker level.

**Anatole Klyosov:**

You are quite right in both of the items. The thing is that all of us share an "earlier" common ancestor, in fact, many of them - in our branch, in our subclade, in our haplogroup, etc., down to a chimpanzee and still down. So, you have to figure out which one you are looking for. Hence, "the most recent" in the TMRCA abbreviation. The same is related to NPEs.

Therefore, we need to resolve haplotype datasets to separate "branches", "clusters", "family" or whatever we call them, each one with its own and only one common ancestor (in terms of DNA genealogy; meaning having one base aka ancestral haplotype; they might have been brothers or father and sons, etc, it would be "one common ancestor" anyway). That is why the logarithmic method is so useful, when coupled with the "linear" method. When the two do not give the same results, it sounds "alarm!". "Something is wrong".

**Keith Britton:**

One facet of your methodology is to pre-examine data by fitting a tree. I would like you to do that, using the datasets. I'd like also to ask Ken Nordtvedt and anyone else interested to do the same. This is not a challenge, and I'm not asking for any results; it's to set the direction from which later questions will be assessed.

I recently happened upon an interesting and instructive website and dataset. I found my analysis (quite possibly defective) rather surprising, and the implications touched other recent threads or "old songs". For John Chandler, I'm aware some markers have high mutation rates, but DYS460? I'll continue appropriate aspects on this thread, initiate another with a different view of the dataset, and then provide the link to the website containing the original material. This also bears on the Younger Tens. That area I will discuss offlist directly with you, Anatole. You and your co-workers are primarily responsible for knowledge in this area, and your journal deserves first look at anything I can contribute.

The dataset is 15 haplotypes. They are, as is common in project datasets, in all FTDNA formats from 12 to 111 markers. For compactness and other reasons they are in delta format (improved since you last saw it Anatole). Basically, that's reducing data to differences from a reference, typically a modal or an Anatole Base.

**Anatole Klyosov:**

I am familiar with the Britton ("the Tens", DYS388=10 R1a1) haplotypes, albeit, apparently, in the outdated version. I have four of the 67 marker haplotypes which you have mentioned, but some of them are not totally complete. Those I have are very similar to each other, and they contain only 4 mutations in their 37 marker format (some alleles are still missing). It results in 11 "conditional" generations, that is ~ 275 years to a common ancestor, however, because of only four mutations (which could have easily been 2, 3, 5 or 6 mutations), the margin of error is  $275 \pm 140$  years.

Obviously, it is not a very useful estimate. Maybe your updated version can help with refining the data.

Regarding DYS460, it is one of the fastest markers. It is comparable with the mutation rate of DYS439 in the first 12 marker panel. The 22 "slow" marker panel which I use for very distant TMRCA calculations includes markers with mutation rates from 0.00001 to 0.00055 per marker per conditional generation (of 25 years). DYS 460 has - according to various estimates (from father-son pairs) -

mutation rates of 0.0040, 0.0058, 0.0038 per actual generation, that is around 10-60 times faster compared to the slowest markers.

**Lancaster-Boon:**

Nevertheless it remains true that whatever the merits of trying to develop a clock out of STR haplotypes, this has diverted a lot of resources away from other ways of using STR haplotypes in ways which would be more useful for genealogy.

And I think this is quite critical given that the people developing clocks are effectively encouraging both companies and their genealogist customers to spend all their money in this direction, and not in other directions - in order to have bigger and bigger data bases filled with "slow moving" markers.

**Anatole Klyosov:**

Nobody sees it as a "clock", literally speaking. It is a rather elongated "cloud" directed to the past. Some people hope to use it in "classical" genealogy, some would like to understand better a nature of mutations, some consider it as an object of chemical kinetics (I am among them, and I am trying to develop chemical kinetics of mutations in haplotypes with as detailed description as possible for potential application in various fields), some are looking forward to develop a tool for anthropology, archaeology, linguistics (I am among them as well). As you see, goals are different, and some of people working on them are delighted with (and are in need to use) the panels of 37-67 and 67-111, and further on - haplotypes.

That is why I protested when you were trying to project your personal goal (probably quite narrow compared with those mentioned above) to the whole field.

Now, I have explained about "developing clocks". Often we do not even need a "clock". I did not need a clock when several years ago I came to a conclusion that bearers of R1b in Europe did not live there 30,000 years ago, as "famous, reputable, heavyweight", etc. founder fathers have insisted, and, as it turned out, without ANY DATA. They just thought (!) so. My simple calculations gave 4500 ybp for the time of appearance of R1b in Central and Western Europe. Now, after several years, it is refined to 4,800 years. As you see, principally there is no difference, clock or no clock. It was a principal conclusion. Some folks remember scream which was here in response to my dating of 4500 ybp for R1b in Europe. DF was one of those main screamers. And now he is dare to stick his neck again with next series of negative words.



So, for some principal issues we do not need "clocks". We solve principal problems. For other issues we need something like a clock, let it be with one hour arrow.

Back to "slow moving" markers. We have enough of them. I have mentioned here my 22 marker panel, in which one mutation occurs once in 4500 years. Even with such slow markers we (with co-authors) have found that some (African) haplotypes have a gap of many mutations from the pack, placing them to >100,000 year times back. So, how "slower" markers do you need?

**Sandy Patterson:**

As you know, M222+ markers are pretty tightly grouped around the modal. As such, it is difficult to identify potential sub-clades....

**Anatole Klyosov:**

Maybe this would be of help: Proceedings of the Russian Academy of DNA Genealogy in its June (2011) issue, vol. 4, No. 6 (in English) lists an article "*Haplotypes of R1b1a2-P312 and related subclades: origin and "ages" of most recent common ancestors*" (pp. 1127-1195).

[http://www.lulu.com/items/volume\\_71/10723000/10723072/2/print/10723072.pdf](http://www.lulu.com/items/volume_71/10723000/10723072/2/print/10723072.pdf)

Fig. 4 on p. 1152 shows a position of L21-M222 on the haplotype tree, and you can see that it contains at least 20 of sub-branches. A brief description of M222 is given on p. 1154.

312 haplotypes of this subclade take almost the whole right-hand side of the tree in Fig. 4. A detailed analysis of this subclade was performed in (Klyosov, 2010b), and the following base haplotype was determined:

13 25 14 11 11 13 12 12 12 13 14 29 - 17 9 10 11 11 25 15 18 30 15 16 16 17 -  
11 11 19 23 17 16 18 17 38 39 12 12 - 11 9 15 16 8 10 10 8 10 10 12 21 23 16 10 12 12  
16 8 12 25 20 13 12 11 13 11 11 12 12 (R1b1a2-L21-M222, 1450±160 ybp)

Analysis of 257 haplotypes of this subclade showed that a common ancestor of this subclade lived 1450±160 ybp. 312 of M222 haplotypes in this dataset (of 2299 haplotypes) contained 98 of 12 marker base haplotypes and 25 of 25 marker base haplotypes. This gives  $\ln(312/98)/0.022 = 53 \rightarrow 56$  generations, that is 1400 ybp with 12 marker haplotypes, and  $\ln(312/25)/0.046 = 55 \rightarrow 58$  generations, that is 1450 ybp with 25 marker haplotypes. These figures are practically identical.

\*\*\*\*\*

I can add that if these figures (1400 and 1450 ybp) are practically identical, it means that the dataset has one common ancestor.

**Terry Barton:**

The Barton DNA project is over 10 years old. There are 96 men in Barton Lineage I - all one genetic family. Forty six are tested at 67 and sixty one at 37 markers. Most of the others are an assortment of 25, 26 & 43 markers. We already know where our fast markers are - and there aren't very many of them. I have a handful of the 96 who still may learn something with an additional marker or panel upgrade - so I agree that can be useful on a carefully selected basis.

The three Barton men tested to 111 found only 2 more mutations and one of them is practically useless to us - as I already knew about it - actually I started it myself. It's only found in me and my son - but not my father nor any of the others who tested it. So, in effect, we found one potentially useful mutation in 3 tests. That's not very powerful - and these three men weren't close - based on what we know from paper trails that don't connect. In 2006, I also tested 10 of my men on 30 odd of Thomas Krahn's markers when he was still at DNA Fingerprint and found only one mutation then. (neither of these last two efforts are on line - but the results up to 76 markers are)

<http://www.worldfamilies.net/surnames/barton/results>

... These days, I've changed from wanting "more markers" to wanting "fast markers". Hope this helps explain why I seek a panel(s) of truly fast markers. (I do recognize that they will come with both back and parallel mutations. I figure that's the price we pay.) Possibly, the "jillions" of coming SNPs will actually provide the branching we seek within our Genetic Lineages. If so - great. But until then, I will continue my quest to learn more through the DNA tool that has been our Lineage definer - ySTRs.

Terry Barton

WorldFamilies.net

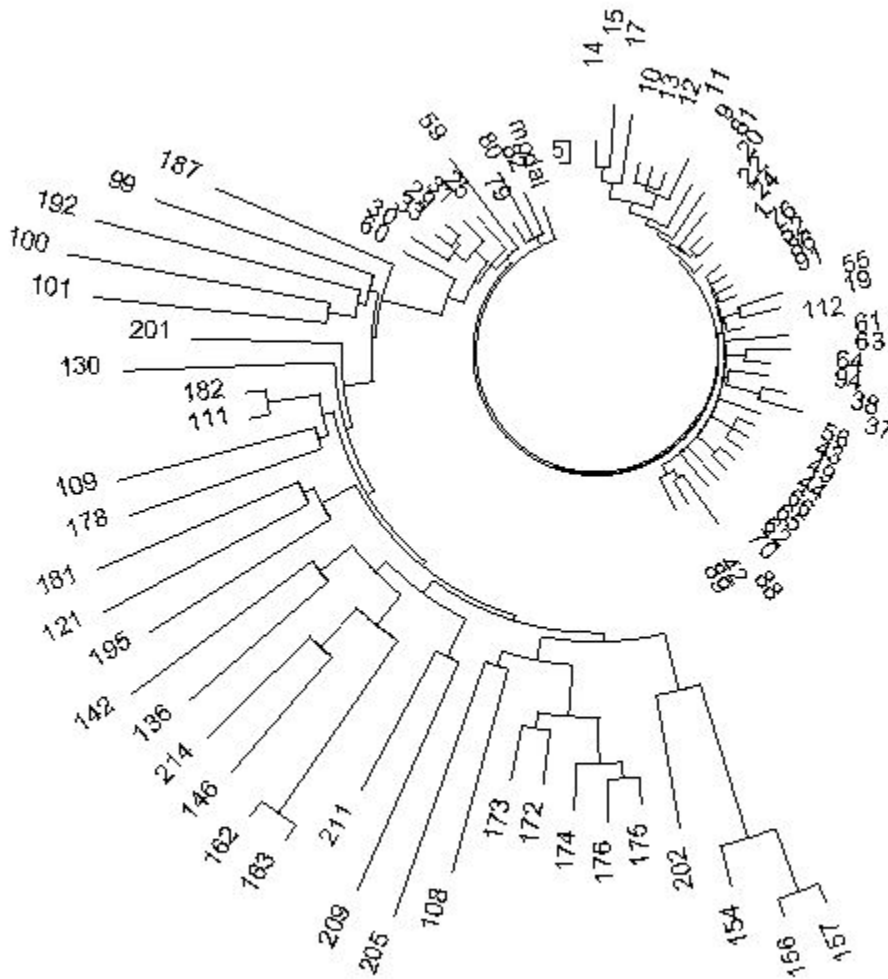
[www.worldfamilies.net](http://www.worldfamilies.net)

**Anatole Klyosov:**

I took a look at your Barton DNA Project dataset and composed a haplotype tree from all 80 of 67 marker haplotypes, including other lineages.

It is a nice tree, with 47 haplotypes forming a close branch. These 47 haplotypes contained 99 mutations from their base haplotype. This gives  $99/47/0.12 = 17.6$  conditional generations, that is 439 years from the common

ancestor. I do not want to round up at this stage, so you can see a reproducibility of the results. It is 1572 AD±62 years.



I also ran four calculations with the logarithmic method.

In the 12 marker panel there were 35 identical haplotypes, that is the base haplotypes. Clearly, this base is the ancestral haplotype in this case. This gives  $[\ln(47/35)]/0.02 = 14.7$  generations.

In the 25 marker panel there were 26 base haplotypes.  $[\ln(47/26)]/0.046 = 12.9$  generations.

In the 37 marker panel there were 9 base haplotypes.  $[\ln(47/9)]/0.090 = 18.4$  generations.

In the 67 marker panel there were 6 base haplotypes.  $[\ln(47/6)]/0.12 = 17.2$  generations.

Please notice that the 67 marker panel gave 17.6 generations and 17.2 generations for the linear and the logarithmic methods, respectively. If we consider all these five calculations, we get an average of 17.6, 14.7, 12.9, 18.4, 17.2 generations, that is  $16.2 \pm 2.3$  generations, with  $\sim 14\%$  variation. It would give you 405 years to a common ancestor, that is approximately 1606B AD $\pm$ 58 years. As you see, the two estimates are overlapping in time.

So, you can take either 17.6 generations (1572 AD) from 47 of 67 marker haplotypes, or 16.2 generations from an averaged logarithmic method, it would not make a large difference. It is about one generation difference anyway. By the way, no correction for back mutations is needed here.

Do you still need fast markers?

**Terry Barton:**

...Dating the MRCA is the least of my interests... Your initial estimate of MRCA at 17.6 generations becomes useful when adjusted to use our actual average years per generation... of 30 and 33 years. We also know that those two paper trail's MRCA could not have been born any later than the 1620s.

I also consider the nearby Barton estate dating to the 1200s and the more than 100 Barton families in area Parish church records in the 1600s.

However, I am unwavering - I still want more fast markers! My interest since the very earliest days of the Barton project is to use yDNA to help reconstruct the Barton Lineage I family - not to simply group them as man sharing a recent MRCA. ...I will continue to be on a quest for more useful ySTR mutations - as they are the best way I have found to identify branching and to partially replace the lost paper trails.

Does that mean we agree or disagree - or does it mean that we are talking past each other?

**Anatole Klyosov:**

Let me repeat what I have found:

*>These 47 haplotypes contained 99 mutations from their base haplotype. This gives ...17.6 conditional generations (25 years each), that is 439 years from the common ancestor... It is 1572 AD $\pm$ 62 years.*

*>I also ran four calculations with the logarithmic method...(It results in) 1606B AD $\pm$ 58 years. As you see, the two estimates are overlapping in time.*

Let me repeat you comment:

*>Your initial estimate of MRCA at 17.6 generations becomes useful when adjusted to use our actual average years per generation... of 30 and 33 years. We also know that those two paper trail's MRCA could not have been born any later than the 1620s.*

What do you think, 1572 and 1606 are “any later than the 1620s”?

Yes, we are talking past each other, at least partly.

Let me explain.

We are seemingly (it looks like that from aside) pursuing different goals, however, your goals are essentially incorporated into mine goals. You apparently did not realize it.

I (and my co-workers and colleagues) pursue three principal goals:

1) To understand and to show others the potential of DNA genealogy in examination, verifications and sorting out the "documented", "classical" genealogy. For this, I work with documented genealogy, such as yours, and many others, and even with "mythical" genealogies, such as Abraham (the Jewish and the Arabic) genealogy (J1 and J2), Sayyd (Sayed, Sharif) genealogy (J1), Rurik genealogy (N1c1 and R1a), etc.

2) For this, I continuously examine and fine-tune the apparatus of DNA genealogy. It is not population genetics at all. It is chemical kinetics considering sequenced DNA fragments, such as haplotypes of Y chromosome, and accompanying SNP mutations. Population genetics considers similar objects plus many others, such as genes, however, DNA genealogy employs very different methodology.

3) The ultimate goal of DNA genealogy is examination, verification and sorting out our knowledge in history, archaeology, anthropology, linguistics, etc. We are not there as yet, however, we are moving in all those directions, studying feedbacks, correcting ourselves and our approaches, including conceptual and computational approaches.

As you see, your interest and mine are supposed to cross in the first item. Let us see why "we are talking past each other", despite a good fit of the calculations.

As you see, my calculations regarding the lineage of 47 individuals in your Project are resulted in  $1572 \pm 62$  years (the linear method) and  $1606 \pm 58$  years (the

logarithmic method). You gave the date as "not later than the 1620s." I am not surprised with such a reasonable fit, I see same things often. However, as you have noticed, it is not of your interest. Fine, it is MY interest. So, I have accomplished my goal in this particular case. However, here we are, as you said, talk past each other. Different interests.

Next item - you misunderstood that "17 conditional generations" in my calculations it is  $17 \times 25 = 425$  years. Period. You cannot "adjust" those conditional generations using 30, 33 years or whatever and get 510 or 561 years. Those 25 years are built into the mutation rate constant, which is 0.12 mutation/haplotype/conditional generation of 25 years. With 33 years per generation I need to change the mutation rate constant from 0.12 to  $0.12 \times 33 / 25 = 0.1584$ , and in that case I will get not 17 but 12.9 generations to a common ancestor, that still the same  $12.9 \times 33 = 425$  years. So, no "adjustments" please. So, I hope, "talking past each other" here is correctable. At least numerically.

Next item - there is no "1200" years from a common ancestor of anything of the haplotype tree of those 47 individuals. It would have been a distinct and a separate branch on the tree in that or next to them branch. There is none with such an "age". Instead, next to that lineage of 47 men (with a common ancestor around 1600 AD) there is a branch with a common ancestor of 2400 ybp (400 BC), which is the ancestral branch to yours of 47 people. The 67 marker base haplotype of the older branch differs by 9.4 mutations from that of the younger (yours) branch, which confirms its ancestral nature, and fits with the both TMRCA.

However, it is apparently is not of interest to you. It is of interest to me, though.

**Lindsey:**

So if the 67 marker base haplotype of my Britton family has a genetic distance of at least 10-12 from the 67 marker base haplotypes of its closest DNA matches (which have different surnames), then the common ancestor would have lived more than 2400 YBP?

**Anatole Klyosov**

Certainly. Here is a small and handy table showing a distance between two 67 marker base haplotypes. The left column shows a number of one-step mutations from the base haplotype, the right column shown a time difference between the two common ancestors (corrected for back mutations). Let's call it "delta". Unless the following two conditions are in place, the difference will be just a crude approximation, a semi-quantitative one.

Condition 1: the two base haplotypes should be verified that they belong to one person (in terms of DNA genealogy), not some phantom stuff.

Condition 2: The two base haplotypes should belong to the same principal subclade in a haplogroup. If they belong to two principally different subclades, the delta is underestimated, and could be underestimated quite significantly.

Assuming that the two base haplotypes are symmetrical (in time) with respect to their common ancestor, the timespan to THEIR common ancestor equals to  $(\text{delta} + \text{TMRCA1} + \text{TMRCA2})/2$ .

I repeat in words: In order to calculate a timespan to the "overall" common ancestor, that is the common ancestor of lineage 1 (having TMRCA1, calculated for the first base haplotype) and lineage 2 (having TMRCA2, calculated for the second common ancestor), you need to sum up all the tree figures (the delta, TMRCA1, and TMRCA2) and divide the sum by 2.

An example: for the Barton 47 haplotypes the TMRCA was 439 years (again, I should have rounded up, but I do not want to confuse folks here), and for the small branch of 5 haplotypes sitting next to it on the tree the TMRCA was 2,400 years. The delta (9.4 mutations) gave  $9.4/0.12 = 78 \rightarrow 85$  conditional generations, that is  $85 \times 25 = 2125$  years. That is the time-difference between the two lineages. The timespan to THEIR common ancestor is  $(2125 + 439 + 2400)/2 = 2,482$  years. Obviously, it is the same common ancestor of 2,400 ybp. Therefore, he is the ancestor of both the 2400 and 439 ybp lineages. The last one, young, just split off several centuries back.

- 1 ----- 210 years
- 2 ----- 420 years
- 3 ----- 650 years
- 4 ----- 850 years
- 5 ----- 1100 years
- 6 ----- 1325 years
- 7 ----- 1550 years
- 8 ----- 1800 years
- 9 ----- 2025 years
- 10 ----- 2275 years
- 11 ----- 2550 years
- 12 ----- 2775 years
- 13 ----- 3025 years
- 14 ----- 3325 years
- 15 ----- 3575 years

**Sandy Paterson:**

My estimate for the sum of the 111-marker FTDNA mutation rates is just over 0.41

Anyone else?

**Anatole Klyosov:**

How come? Based on what?

The mutations rate constant for the 111 marker panel was published in Proceedings of the Russian Academy of DNA Genealogy, vol. 4, No. 7 (July 2011), p. 1535-1539, "*The mutation rate constant for the 111 marker haplotype panel*", by A.A. Klyosov. Illustrations and calculations were given, among others, using a dataset and the respective haplotype tree for I1 haplogroup. The 67 marker panel (with the mutation rate constant 0.12 per haplotype, or  $0.12/67 = 0.00179$  per marker) gave  $3625 \pm 385$  years to a common ancestor, the 111 marker panel (with the mutation rate constant 0.198 per haplotype, or  $0.198/111 = 0.00178$  per marker) gave 3700 years.

Let's verify it here, just for fun. 20 Donald Clan haplotypes (Red Group) are available in the 111 marker format. The first 67 markers have 70 mutations from the base haplotype, that is  $70/20/0.12 = 29$  generations (without a small correction for back mutations), all the 111 markers add 37 mutations, resulting in 107 mutations total.  $107/20/0.198 = 27$  generations.

(By the way, somebody here was talking on a "huge error of margin")

With your 0.41 value for the 111 markers the Donalds would have had a common ancestor only  $107/20/0.41 = 13$  generations ago, that is 325 years ago. Forget about John, Lord of the Isles :-)) (I am reminding again that those "generations" are conditional ones, 25 years of length, you cannot change it without a recalculation of the mutation rate constant).

Another example. A couple of days ago I have calculated here a timespan to a common ancestor of E-V13 subclade based on 193 of 67 marker haplotypes. It gave 115 generations for the 25 marker panel, 117 generations for the 37 marker panel, and 123 generations for the 67 marker panel (without corrections for back mutations).



Now, out of those 193 haplotypes 26 were determined in the 111 marker format. Let's verify the calculations.

The first 25 markers contain 129 mutations, that is  $129/26/0.046 = 108$  generations.

The first 37 markers contain 256 mutations, that is  $256/26/0.090 = 109$  generations.

The first 67 markers contain 355 mutations, that is  $355/26/0.120 = 114$  generations.

The first 111 markers contain 552 mutations, that is  $552/26/0.198 = 107$  generations.

One can see, that even the extreme numbers of generations, 114 and 107 here, differ by only 6%. On average all four TMRCAs gave  $109.5 \pm 3.1$  generations, that is  $\pm 2.8\%$ .

(Again: by the way, somebody here was talking on a "huge error of margin", and somebody else was talking that he does "not believe in those calculations").

**Sandy Patterson:**

>>My estimation for the sum of the mutation rates for FTDNA markers 1-111 is just over 0.41.

>Anatole Klyosov: How come? Based on what?

It's based on observed YHRD mutation rates plus those published by Ballantyne et al 2010 and Burgarella & Vasques 2010.

I found that these three sources gave results for 35 of the 111 markers. I then took observed variances in M222+ and L21+ and, using these, modelled the observed 35 mutation rates. I then used that model to re-estimate all 111 mutation rates.

By memory, I think John Chandler's estimated rates sum to .2243 over 67 markers. I know that he's expressed satisfaction with the rates 1-37 but has been more cautious about markers 38-67. Still, I don't think .2243 is far off.

If we then do a crude sanity check on the estimate of .41, we get  $.2243 \times 111 / 67 = .37$

So I don't think .41 can be that far out.

**Anatole Klyosov:**

Your "sanity check" was based on wrong numbers by John Chandler for the 67 marker panel. It is not 0.2243, but 0.12. So, your "sanity check" should have been

$0.12 \times 111/67 = 0.199$ . In fact, it is almost exactly correct, the right (calibrated) number is 0.198 mutations per 111 marker haplotype per "conditional" generation of 25 years.

Regarding Ballantyne, Burgarella et al I can only exclaim - again?

I gave here a detailed analysis of their "mutation rates". Well, read again, [http://aklyosov.home.comcast.net/~aklyosov/Vestnik\\_4\\_09.pdf](http://aklyosov.home.comcast.net/~aklyosov/Vestnik_4_09.pdf) p. 1831 and further, particularly pp. 1834-1837.

"Mutation rates" by Ballantyne et al (2010) and Burgarella and Vasques (2010) might be good for something else (I frankly do not know for what, but maybe for some forensic things), but NOT for DNA genealogy.

I do not know about your "memory", however, 0.2243 for 67 marker haplotypes is a totally unacceptable value. It is 187% higher than the proven value. In the "correct TMRCA analysis" I have noted that 39 of 67 marker Clan Donald haplotypes have 159 mutations, which resulted in  $875 \pm 110$  years to (supposed) Somerled:  $159/39/0.12 = 34 \rightarrow 35$  (conditional) generations, that is  $875 \pm 110$  ybp. If you apply that 0.2243 figure, you will get that Somerled lived  $159/39/0.2243 = 18$  (conditional) generations, that is 450 ybp. John, Lord of the Isles (died in 1386) would have lived about 350 years ago.

*>So I don't think .41 can be that far out.*

:~))

It is VERY far out.

**Linda:**

*>The first 25 markers contain 129 mutations, that is  $129/26/0.046 = 108$  generations.*

*>The first 37 markers contain 256 mutations, that is  $256/26/0.090 = 109$  generations.*

*>The first 67 markers contain 355 mutations, that is  $355/26/0.120 = 114$  generations.*

*>The first 111 markers contain 552 mutations, that is  $552/26/0.198 = 107$  generations.*

I am into family history research and the use of yDNA evidence for supporting my genealogy research.

Your statements seem contrary to what I thought was "the way it works" . That being that the more STR markers one tested the more near to current time you became in matching with folks of your surname and haplotype.

I suspected a few months ago that all the STR tests were useless (for my

genealogy work) because of the SNP results that seem to be tossing all those STR results around like bouncing balls and your observations above seem to confirm that.

Am I seeing that correctly? And if I am seeing it correctly would you comment on any theory you may have about how yDNA works with genealogy?

**Anatole Klyosov:**

I have no idea how you manage to see in the example above (and in tons of other similar in kind examples) that (1) the more STR markers, the more near to the current time you became, (2) all the STR ... useless, (3) SNP ... toss all those STR results around, and (4) that my observation above seem to confirm that.

If I understood what you have said and rather accurately summarize it above in those four items, then I see it completely opposite, all the four items.

Let's go item by item.

(1) Where you see that "the more STR markers, the more to the current time? Above you see four sets of STR markers, of 25, 37, 67, and 111 markers. The calculated dates are 108, 109, 114, and 107 generations. In fact, the last one is 108 generations, if I take one transition 15-->9 as 6 mutations, but I took it as only one mutation. As you see, I had an opportunity to make the final result "nicer", but I did not want to do it, undercounted 5 mutations and calculated 107, and not 108 generations.

It is not "more to the current time when you move from 25 STR markers to 111 STR markers, it stays the same within just a few percent.

(2) Well, maybe "all the STR are useless" to you, but not to me. By the way, I am not "into family history research", my specialty is time-dependent processes, including mutations in the DNA. So, I am a professional in the area in the strict sense of the word.

(3) SNPs do not "toss" STR results around, SNPs help to separate datasets into more justified groups of haplotypes, particularly when haplotypes are too close to each other. In fact, SNPs do often confirm the structure of a haplotype tree. Sometimes a haplotype tree helps to identify potential SNPs. For example, several years ago our group published an article in which we found two distinct branches on the R1a1 haplotype tree. We named these branches "Central European branch" and "Western Slavic branch", and described their base haplotypes. About a year after it SNP M458 was discovered, which embraced the both branches. Still no separate SNPs which would identify each of the two

branches was discovered, but they will be discovered (may be they are very new L260 and M334, but it is not certain as yet). As you see, SNP did not "toss" those two branches around, it just confirmed the existence of the two branches.

(4) How did you get to a "conclusion" that my calculations and observation confirm your items (1) to (3) above?

>...would you comment on any theory you may have about how yDNA works with genealogy?

Your question covers the whole field of science. Judging from your comments above, you have a VERY vague idea on STR, their mutations, etc. I suggest you to read materials of this Forum. There are many good comments here.

**Ann Turner:**

Linda, I don't believe Anatole understood your perspective. He is looking at the total number of mutations found in a large group of people. The more mutations overall, the more distant the common ancestor. The more markers you test, the more mutations you will find in the group, but it stays in proportion, so the MRCA calculation for the group as a whole gives about the same result whether you test 25 or 111 markers.

You're looking at this from a different perspective. When you are comparing two individuals, a perfect 67-marker match is much harder to achieve than a perfect 12-marker match, so the MRCA will tend to be more recent.

**Anatole Klyosov:**

There are three statements above. One maybe correct, and two are incorrect.

The first, that I do not understand the "perspective" is well may be correct. Some "perspectives" I even refuse to understand, if they base on wrong assumptions.

The second one, that "He is looking at the total number of mutations found in a large group of people" is flat incorrect. I often look at haplotypes of two individuals, and it is O.K., however, error margin in those cases is often too large to be truly informative.

The third one, "a perfect 67-marker match is much harder to achieve than a perfect 12-marker match, so the MRCA will tend to be more recent" - is incorrect. MRCA would not be "more recent", it would be much more jumpy for 12 marker haplotypes. A "perfect match" for two 12 marker haplotypes does not mean much, and one mutation would jump the "TMRCA" from "zero" to

50±50 generations from a "common ancestor", that is to 1250±1250 years ago.

As you see, the TMRCA is not "more recent", it is just uncertain, it is between 0 and 1250±1250 for one mutation. Estimates between two 67 marker haplotypes (better if between two 111 marker haplotypes) are much more certain, not "more recent".

**Ann Turner:**

I agree that the effect of a mutation is "jumpier" for 12 markers, but the MRCA would tend to be more recent for a 66/67 marker match than an 11/12 marker match.

Specifically, the 50% probability is ~ 24 generations for 11/12 vs 5 generations for 66/67 when using .003 for the mutation rate.

**Anatole Klyosov:**

Well, let us see. For two 67 marker haplotypes with just one mutation between them, the time-wise distance between them is  $1/0.12 = 8.3 \pm 8.3$  generations (25 years each). Of course, 8.3 generations is ridiculous, however, let's do it for the sake of the demonstration. It formally (sic!) means that if the two haplotypes belong to the same subclade and to the same lineage, their common ancestor lived only  $100 \pm 100$  years ago. That is, between 1811 and our time, formally (sic!) speaking.

For two 12 marker haplotypes with just one mutation between them, the time-wise distance between them is  $1/0.02 = 50 \pm 50$  generations. It formally means that if the two haplotypes belong to the same subclade and to the same lineage, their common ancestor lived  $1325 \pm 1325$  years ago. That is, between 686 AD and our time, formally speaking.

If you prefer to call the first one as "more recent", fine with me. Could be the same common ancestor, though.

I would prefer to call it not "more recent" but "less certain".

Your last sentence is beyond my understanding. What is 0.003 ("for the mutation rate") anyway?

**Gary Felix:**

Let me explain, I am against the use of STR's to determine the ancient ages of Y DNA lines... Evidently you are not aware of the "surfing effect" or "wave front"... In the mean time perhaps you can explain how one Y DNA line surrounded by

thousands of farmers comes to leave his Y across Europe to the detriment of all but a few lines.

**Anatole Klyosov:**

Yes, I understood it immediately from your comments. Hence, my ironic replies. No problem, you live in a free country. You are entitled to have your opinion. So why you are "against"? Don't you think that some also live in free countries?

As always, ignorance goes hand in hand with categorical statements ("evidently"...). Incidentally, I am familiar with the articles you have quoted on the "surfing effect" or "wave front". Some things in them are reasonable, some I do not share, since I have other data not in agreement with their statements. It is normal for science, isn't it?

For example, Chiaroni et al (2009) paper cites data by Prugnolle et al (2005) regarding "a linear fall of genetic diversity with the geographical distance in each population from the origin in East Africa". First, "origin in East Africa" is not proven, and a number of other geographical points can be indicated with the same "fall in diversity" with distance from them. As you see, "the origin in East Africa" is just postulated there.

Second, "genetic diversity" can be expressed very differently. Alternatives were not even considered there, which is no-no in science. Third, the known phylogenetic tree in Fig. 1 in the paper by Chiaroni et al shows (that we all know, I believe) that haplogroups from D to R/T do not have SNPs of haplogroup A and haplogroup B. So, "high genetic diversity" in Africa is already irrelevant here, at least for non-African males on Earth. This because if that "high diversity in Africa" is related to only haplogroups A and B, and other haplogroups did not descend from either A or B, then who cares about their "high diversity" in Africa, except when African populations of haplogroups A and B are studied?

In other words, that "linear fall of genetic diversity with geographical distance" is not related to male chromosome Y. Or it is a product of imagination. Or it equally falls when moving from elsewhere.

Indeed, if you write down base ("ancestral") haplotypes for all haplogroups from A to T, all in that order, you will not see any sign of this "linear fall" from haplogroups A or B.

Regarding "surfing effect", "wave front" and all those buzzwords, no problem, they are just not from my vocabulary. When I showed you that the 7,000 ybp and 3,400 ybp base haplotypes are differed by 4 mutations, there was no need for me to brag about "surfing effect" or "wave front". It was just a 4-mutation gap

between two base haplotypes which can be described in years, quantitatively. It is that simple.

Regarding your last question, I would appreciate if you rephrase the question. Thank you. By the way, why "farmers"? Maybe they were military guys? Or storytellers? Or glassblowers?

**Gary Felix:**

I figured he would have to be a farmer with that many sons to feed.

I ask, what made this storytellers Y superior to the majority of contemporary farmers who had been populating Europe for many thousands of years prior? Especially since this storyteller's line almost didn't survive the neolithic expansion in the first place.

*>...then who cares about their "high diversity" in Africa, except when African populations of haplogroups A and B are studied?*

Already supplied papers showing that diversity decreases away from its origin. Also if you look at the haplotree you will see the earliest lines are concentrated closest to Africa with the more recent ends of each line away from its origin.

The fact that Africa has A and B shows more diversity than the lines that left Africa.

**Anatole Klyosov:**

:~)) (I can only smile at the above reply )

If the "lines that left Africa" have no African SNPs (neither M91, P97 of haplogroup A, nor M60, M181, P90 of haplogroup B) how do you know that they "left Africa"?

I bet you do not have those M91, P97, M60, M181, P90. How can you claim that your (male) ancestors came out of Africa??

Please no somebody else's "opinions". Give me DATA. FACTS. Your words, not somebody's "thinking".

So, fact No. 1. You do not have said SNPs. If you do not know about yourself, I know hundreds of others (non-Africans) who do not have them.

Your explanation? Haplotypes maybe? Calculations? Oh, yes, you do not believe in them. So, what's left?

**Gary Felix:**

*>If the "lines that left Africa" have no African SNPs (neither M91, P97 of haplogroup A, nor M60, M181, P90 of haplogroup B) how do you know that they "left Africa"?*

Not true haplo D & E have the YAP marker which is African in origin.

**Anatole Klyosov:**

Great. It seems the discussion might be an interesting one. Hopefully.

I would be really happy to learn that the YAP marker has an African origin, and haplogroup DE was originated in Africa (this would make my life easier in understanding the origin of man), however, we are not there. You are too categorical again ("not true", "African in origin"). I wish I know that much as seemingly you do.

First, we do not know whether YAP has an African origin, as far as I am concerned. Sources typically indicate "possible place Africa or Asia", regarding origin of DE.

Second, in their work Underhill P.A. and Kivisild T. (*"Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations"*. *Annu. Rev. Genet.* 41: 539-64, 2007) used an obsolete phylogeny tree. We now know that it is not correct. B did not descend from A, and the others did not descend from B. So, your thinking probably is still based on that obsolete picture, if you are saying that haplogroups D and E are "African in origin".

Third, there are many circumstantial evidences that DE was hardly originated in Africa. There are data that out of 8,000 Nigerians only five individuals have marker DE. Sounds like 0.06% to me. Other DE are found in Asia, some (seldom) in the Middle East. I do not draw any conclusions from it, I am just showing an uncertainty with geographical location.

Fourth, as I remember, in the WTY Project participants in haplogroups A, on the one hand, and D/E, on the other, do not have overlapping SNPs at all.

Fifth, if even you show that YAP has indeed the African origin (how are you going to show it, I wonder), still there is life "above" (or "below") haplogroups D/E, namely in the F-T category. They do not have any know "African" SNPs, do they?



What say you?

*>Already supplied papers showing that diversity decreases away from it's  
>origin. Also if you look at the haplotree you will see the earliest lines  
>are concentrated closest to Africa with the more recent ends of each line  
>away from its origin.*

I have already explained that the first is not an argument. If R1a1 had migrated to South Siberia, it does not mean that it came from Africa. The second is even less an "argument".

*>I ask, what made this storytellers Y superior to the majority of contemporary farmers  
>who had been populating Europe for many thousands of years prior? Especially since  
>this storyteller's line almost didn't survive the neolithic expansion in the first place.*

I do not know specifically. It happened all the times, with all haplogroups, subclades, lineages. One of them die out, other multiply. As Leo Tolstoy wrote, "All happy families are happy the same way, all unhappy families unhappy in their own ways".

Maybe some storytellers used to tell stories better than others. I do not know.

However, I suggest you to toss a coin simulating proliferation of several lineages, with a son be a tail and a daughter be a head. Or conversely, if you prefer that way. You will see that in 15-20 tosses only one lineage will left. What made it superior, using your language

**Sasson Margaliot:**

Underhill P.A. and Kivisild T. in their study of 2007 never said that A and/or B were ANCESTRAL to other groups. What they said was that A and B were the first to branch out: first A, then B.

The next group to branch out, they say, was DE, and then C. That's what folks here believe that happened.

This version of Tree Trunk is repeated everywhere, and is universally accepted - without evidence.

The Y-DNA from Nea was sequenced in Europe, which does not support the "accepted" version of Trunk Tree - but it was withheld from publication.

**Anatole Klyosov:**

When I was writing my comment, I checked the "Haplogroup DE" piece in Wikipedia. There is a diagram there with a reference to the Underhill and Kivisild 2007 paper, in which A is ancestral to B and everything else, including DE, and B is ancestral to everything else, including DE.

It could have been that Wiki distorted the original data (it often does), however, first, it is not relevant here, since DE is hardly descended from either A or B, and, second, the statement from Underhill and Kivisild, which you have reproduced here "What they said was that A and B were the first to branch out: first A, then B. The next group to branch out, they say, was DE, and then C" is also hardly correct. Well, I am trying to be diplomatic. It is almost definitely incorrect.

"Folks" believe, indeed, but that was "folks" normally do. They "believe" and "accept", as soon as it is published. "Peer review" for "folks" are truly magic words.

The current version of the tree (ISOGG-2011) is generally O.K., however, I do not see there the word "Africa" carved in, particularly above (or below, or on side, rather) haplogroup A.

All that fuss about "diversity" in Africa is in fact irrelevant to most of the haplogroups. New York City is more diverse than Newton, Mass (where I live), but it does not mean that Newton descended from New York. Diversity is not necessarily the origin, it is often a mix.

I (with co-workers) also have a different version of the tree: formally the same (or similar) but very different IN KIND.

**MY COMMENT:** The discussion further moved to an estimation of a time span to the most recent common ancestor of current bearers of the E-V13 subclade. The thing is that recent excavations in Spain dated as 7,000 years before present have identified some remains of a bearer of E-V13, which would indicate that the subclade is at least 7,000 years old. A reference was made to my earlier calculations of the TMRCA for E-V13, based on short haplotypes published by Cruciani et al and Battaglia et al, which gave the TMRCA of about 2,600 years before present (ybp). I have explained at the Forum that when we determine the TMRCA, then by definition we calculate a timespan to the most recent common ancestors of bearers of respective haplotypes who live TODAY. If that common ancestor lived 2,600 ybp, it means that either Cruciani et al, Battaglia et al did not have a good representative series of haplotypes (that is they collected haplotypes for their publications unprofessionally, which I seriously doubt), and

they (haplotypes) are skewed to more recent times, or more earlier ancestors of E-V13 subclade (in this particular case) had vanished along with their descendants, and their descendants are not around. The third option, that our calculations cannot distinguish between 7,000 and 3,000 years ago, is totally unrealistic. Of course, ignorant and negative-minded people, who have no idea about calculations their calibrations and verifications, etc., happily celebrated the third option. Examples are given immediately below.

**Dienekes Pontikos:**

If anyone is still in doubt that Anatole Klyosov's method and his self-assurance in his narrow confidence intervals are without any merit... The recent publication of 7,000 year old E-V13 from Neolithic Spain... ..indicates that Klyosov's age estimate for E-V13 is off by a factor of 2.5 Huge confidence intervals indeed!

**David Faux:**

Alas Anatole, not only are Y-STR MRCA analyses for "deep ancestry" on their "way out" due to "poor and erratic performance" (as per Dienekes assessment), but the same can be said for their use for genealogical timeframes... Very futile efforts whether 300 years, 3,000 years, or 30,000 years. It is a tool that is broken and needs to be discarded because it cannot be fixed.

**Gary Felix:**

I am against the use of STR's to determine the ancient ages of Y DNA lines... How are people coming up with these old ages for the V13 V12 node with your utility?... Estimates can vary greatly based on the markers used.

\*\*\*\*\*

**Anatole Klyosov, re. Dienekes Pontikos:**

Funny. I thought that everyone here understands that by measuring TMRCA we identify the timespan to a SURVIVED common ancestor, who might very likely pass a population bottleneck and his DNA was carried on to the present time. By studying excavated bones we measure a timespan to an actual person who might likely died well before that population bottleneck.

Should I really remind this very simple thing on this Forum?

As a bonus: here are some data on CURRENT population of V13, two different datasets.

In one study of 205 of 67 marker haplotypes it was  $3725 \pm 380$  ybp (from the Discussion I have referred to today); in another study 261 of 11 marker

haplotypes were listed (Cruciani et al, 2007), and the calculations gave  $2675 \pm 270$  ybp (the linear method) and  $2450 \pm 250$  ybp (the logarithmic method).

In other words, we compare here apples and oranges with the excavated bones. This is the classical dilemma when one compares direct anthropological data with TMRCA's.

**Didier Vernade :**

I noticed the E-V13 discrepancy but... The SNP test on old DNA may also be suspicious... Only 1 case... and only one SNP.

... I still have doubt on the V13 result... If both groups were V13+ with a bottleneck, one would expect many more SNPs separating the "new" V13 group from other E-M78 subclades.

The test is done on DNA degraded by presumably 7000 years in a cave. Many tests are failing in such conditions. One should keep some caution before considering this one result : One skeleton and only one SNP is indicating E-V13 : so, yes, may be the result is wrong... E-V13 is only known in a few cases in Spain and in France ; it's mostly a marker from the Balkans, Central Europe and East Europe. Spain is not the place one was expecting to find E-V13.

Now, it's possible that the E-V13 result is valid but then I would expect a group in Spain and south western France (next to Spain) to display a large variance (or a different result according to A. Klyosov calculations). The idea could be that there was an old (7000 bp) group E-V13 and the haplotype, mainly from the East collected today are belonging to a subgroup of E-V13. The analysis conducted by A. Klyosov might be right for this subgroup (to be defined) while the whole E-V13 group is a more complex one. This possible explanation of the discrepancy has still to be found, meaning : we need a collection of E-V13 haplotypes from Spain and see if it can corroborate the cave finding. In the absence of such finding I would very much doubt of the result, not confirmed yet by another work, by Marie Lacan et al.

... I consider equally likely that the V13 result was wrong.

...My position is reasonable doubt on DNA analysis when the sample is 7000 years old and the result is sustained by only 1 sample and only 1 SNP.

**Anatole Klyosov:**

Let's assume that the identification of E-V13 in the excavated material was correct. Of course, this alternative - that it is incorrect - stays, however, it is just an alternative. Let's keep it in mind and go further.

So, it is correct. What it would give us?

The whole discussion in this thread broke loose, and totally non-constructively, on just a personal matter. Let's put it aside, and come back to science.

There are two scientific issues to be resolved: (a) whether the calculations for TMRCA for E-V13 are correct, and a common ancestor of the three considered groups of V13 bearers:

- the Cruciani et al [2007] list of 261 of V13 11-marker haplotypes;
- the Battaglia et al [2008] list of 108 of V13 12-marker haplotypes; and
- Steve Goodrich list [2011] of 205 and 165 of V13 67-marker haplotypes.

lived  $2625 \pm 290$ ,  $2650 \pm 390$ ,  $3300 \pm 600$ , and  $3250 \pm 360$  years ago, respectively. In Steve' data the linear method and the logarithmic method gave 3500 and 3300 ybp in the first case, and 3250 and 3250 ybp in the second case.

For the record, the first two short lists produced the base haplotypes as follows (DYS19, 390, 388, 389-1, 389-2, 391, 392, 393, 439, 460, YCAIIa, YCAIIf):

13 XX XX XX XX 10 XX 13 12 9 21 19 , and

13 24 12 13 30 10 11 13 12 9 21 19

(b) Whether E-V13 came from North Africa to Iberia ~ 4,800 ybp on shoulders of R1b1a2 (Bell Beakers), which I have dismissed in the earlier discussion based on rather recent "age" of E-V13 and their presently known location (none in Africa).

The first issue has a firm answer: all known V13 haplotypes are of a rather "recent" origin, as calculated above. They could not possibly be dated by ~ 7,000 years ago. I cannot exclude that some very mutated V13 haplotype can surface some day, however, I am talking of those ~ 700 of V13 haplotypes considered above and their TMRCA's.

Let's verify it differently, since some folks probably did not grasp those multi-marker calculations. Let's minimize the task, and consider some of the slow markers in the listed haplotype datasets. Take DYS393. According the well-known John Chandler table of mutation rate constants, which I have examined and generally confirmed for the fist 12 marker panel, the mutation rate constant for DYS393 equals 0.00076. How many mutations would be accumulated after 7,000 years in E-V13 haplotypes? The answer: 16-17

mutations on average in 100 haplotypes. Verification:  $16/100/0.00076 = 211$  ( 266 generations, that is 6650 ybp;  $17/100/0.00076 = 224$  ( 288 generations, that is 7200 ybp. The correction for back mutations is shown. If someone does not like corrections for back mutations or not trust them, no problem. In that case the amount of mutations will be 21-22 per hundred haplotypes.

How many of mutations are in reality in those listed datasets?

Cruciani: 8 (eight) in 261 haplotypes, that is 3.1 per 100 haplotypes. Not 16-17, not 21-22, but 3.1.

Calculations based on 8 mutations will be too imprecise, however, this is not for calculations. It is to show that the Cruciani 261 haplotype could not be derived from a 7,000 year old common ancestor.

Battaglia: 2 (two) in 108 haplotypes, that is 1.9 per 100 haplotypes.

Goodrich: 3.96 per 100 haplotypes.

So, it could not be a 7,000 "old" ancestor.

O.K., let's take a faster marker, such as DYS19, with the mutation rate constant 0.00151 (John Chandler). After 7,000 years it should produce on average 33 mutations per 100 haplotypes.  $33/100/0.00151 = 219$  ( 280 generations, that is 7,000 years. Without a correction, it should be 42 mutations.  $42/100/0.00151 = 278$  generations, or 6950 years.

Cruciani: 9 (nine) mutations in 261 haplotypes, that is 3.4 per 100 haplotypes.

Battaglia: 14 mutations in 108 haplotypes, that is 13 mutations per 100 haplotypes.

Goodrich: 7.1 mutations per 100 markers. Not 33 mutations, not 42 mutations for 7,000 years.

Conclusion: It is clear that the present day population of E-V13 tested for this subclade is descended from a common ancestor who lived not 7,000, but around 3,000 years ago.

Let's move to item (b), whether E-V13 came from North Africa to Iberia ~ 4,800 ybp on shoulders of R1b1a2 (Bell Beakers), which I have dismissed in the earlier discussion based on rather recent "age" of E-V13 and their presently known location (none in Africa).

The answer is also clear. According to data collected by M. Lacan, out of 215 bearers of E-V13 none lives in North Africa. 75% live in Central

Mediterranean, 25% in Western Mediterranean. E-V13 have not been with the Bell Beakers, because amount of V13 in French is 0, in Irish is 0, in Belgians is 0, in Dutch is 0, in North Western Europe is 0.

So even if E-V13 were in Iberia 7,000 years ago, this would not change a thing in my earlier conclusion, when I excluded E-V13 from those who had followed Bell Beakers. They came not from Africa, but from the east of Europe.

End of the story.

**grandcross:**

Didier wrote: (...)

As for the presumed age of E-V13, here is a sampling of opinion:

E-V13 expansion times observed for Konya, Franchthi Cave and Macedonian Greece, all approximately 9000 years ago. Moreover, E-V13 YSTR-related data from Bulgaria and Macedonia, both with a variances of 0.28, suggest an expansion time of approximately 10 000 years ago...

As to a western Asia-Europe connection, our data suggest that western Asians carrying E-V13 may have reached the Balkans anytime after 17.0 ky ago...

You can argue all you want about the methodology used to arrive at these numbers...

**Anatole Klyosov:**

grandcross, your rather aggressive comments require at least some substance to base them upon. However, all your "substance" is summarized here:

*>As for the presumed age of E-V13, here is a sampling of opinion*

Personally, I could not care less about "opinions". I need DATA and, particularly, HOW they were analyzed. Cruciani, for example, uses infamous Zhivotovsky "population mutation rates", which increase dates by 250 - 300%. Why do I need his "opinion" of the "presumed age"?

*>E-V13 expansion times observed ... all approximately 9000 years ago.*

Zhivotovsky again?

>E-V13 YSTR-related data from Bulgaria and Macedonia... suggest an expansion time of approximately 10 000 years ago.

Zhivotovsky again?

>As to a western Asia-Europe connection, our data suggest that western Asians carrying E-V13 may have reached the Balkans anytime after 17.0 ky ago

Zhivotovsky again?

**Comment off-line (Anatole Klyosov):** "grandcross" did not answer the above questions. Instead of a specific answer, HOW those numbers were obtained, he wrote:

**grandcross:**

The "substance" of my remarks are Lacan's analysis and the opinions of some very credible population geneticists over the years who have studied and published papers relating to E-V13. ... It seems to me that the difference between your opinion and theirs is the existence now of physical evidence E-V13 was in Spain 7,000 years ago. That is a fact you cannot deny...

...So how \*do\* you explain the E-V13 bones found in a Neolithic grave in Spain?

**Anatole Klyosov**

First part of your first statement is not correct. There is a diagram in Fig. S1 of the paper by Lacan et al. which shows that out of 215 bearers of E-V13, 75% are from Central Mediterranean, and other 25% from Western Mediterranean. Table S3 in the same paper shows distribution of V13 among regions and countries, and as I have already cited, in French it is 0, in Irish 0, in Belgians 0, in Dutch 0, etc. I would really appreciate if you, when mentioning Lacan's analysis, be more specific. Or at least truthful.

Regarding opinions "of some credible populations geneticists" I have to repeat that I could not care less about their "opinions", when they (opinions) are not accompanying with data. I have a high opinion of work by Cruciani, however, I have a VERY low opinion on his applications of Zhivotovsky "approach" to TMRCA's. As you see, an "opinion" is not a religion, when you believe and are not suppose to ask questions.

Also, you did not answer my question on HOW those astronomical TMRCA's for V13 were calculated, based on what?



Regarding your second statement,

*> It seems to me that the difference between your opinion and theirs is the existence now of physical evidence E-V13 was in Spain 7,000 years ago. That is a fact you cannot deny...*

well, if you (seemingly) can write, you probably can read. If so, it would be a good idea to read what I have written on E-V13 in Spain 7,000 years ago. As I have described, I do not question the data. I do not have a specific ground to question them. So, I (naturally) accept the existence of V13 in Spain 7,000 ybp. So what, except it is an interesting fact? The Earth is peppered by ancient Y chromosomes which we are not going to see among present day people. What we determine using approaches of DNA genealogy, it is a timespan to the earliest survived common ancestor of the given population, this is all. There are two ways to check whether the dates could be "deeper" - to collect and study more haplotypes, and excavate haplotypes (in order to study them). It is a no-brainer.

As you see, I do not deny it at all. I accept it, though on the back of my mind there is always a thought that it might be erroneous data. It is my profession to think that way.

*>So how \*do\* you explain the E-V13 bones found in a Neolithic grave in Spain?*

Easy. That fellow either lived there 7,000 ybp, or came from elsewhere, or there is a wrong dating, or there is a wrong typing of V13. All is possible. As Didier rightly hinted, one point it is not quite a point. Though, we have it, and should take it into account. As simple as that.

**Lawrence Mayka:**

Nordtvedt's Generations6 spreadsheet estimates the interclade TMRCA between E-V12 and E-V13 (using the 67-marker entries in the E1b1b Project) as 319 generations, or about 9600 years; and the intraclade age as 166 generations, or about 5000 years. This means that the actual V13 SNP appeared sometime between 5000 and 9600 years ago--not even counting confidence intervals on these numbers...

**Anatole Klyosov:**

This is a good example of how a discussions reach a certain stage of absurd. The interclade between V12 and V13 shows a time span to A COMMON ANCESTOR of V12 and V13, that is E-M78. In fact, E-M78 is a common ancestor of V12, V13, V22, V65, and M531. Why do we need an TMRCA for E-M78 in the context of a timespan to V13?

How these "166 generations" did appear out of nowhere? How many haplotypes have been considered, how many mutations in them, were they indeed V13, since only "E1b1 Project" is mentioned? Since when 166 generations is "about 5000 years", how was it calibrated to make it into 5000 years? Was it verified that it was only one lineage (one common ancestor)?

Absurd. This is the way to kill any work. Make it a mess and claim that nothing works.

**Lawrence Mayka:**

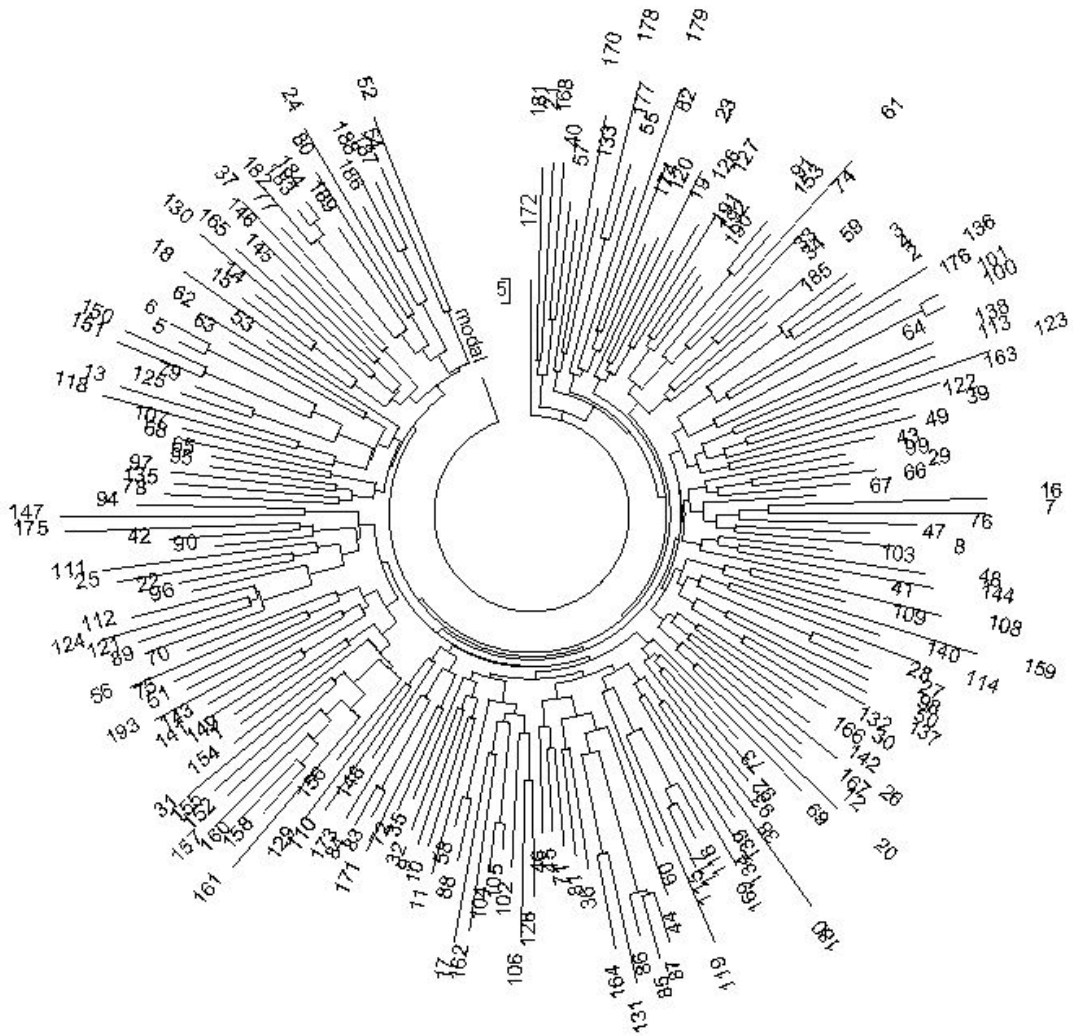
I have sent you an Excel file of the E-V13 haplotypes. You are free to perform your own analysis.

**Anatole Klyosov:**

Thank you Larry, I have received the dataset, composed and sent you the haplotype tree of those 193 of 67 marker haplotypes, along with results of calculations.

As I have expected, your calculations using that "Generation" utility produced significantly elevated TMRCA. I guess I know why. My experience with "quadratic" calculations showed that squares of the distances are extremely sensitive to any aberrations, deviations, recLOH, etc. This particular set contains plenty of multi-copy mutations. How do you think the "Generation" handled them? Right, it squared all of them as there is no tomorrow.

I do all my calculations manually, including all multi-copies. The haplotype tree is nice and almost "ideal" (see below), except a few haplotypes on the very top, which form a separate small branch. They could be some erroneous haplotypes, with mistyped alleles, and the program threw them together. I did not filter them out, they might have distorted things just a little bit. However, in the quadratic method the distortion is likely significant.



**Haplotype tree composed of 193 of 67 marker haplotypes, provided (the haplotypes) by Lawrence Mayka.**

The first 12 markers contain 427 mutations, which gives  $427/193/0.02 = 111 \rightarrow 124$  generations, or  $3100 \pm 340$  ybp

The first 25 markers contain 1020 mutations,  $1020/193/0.046 = 115 \rightarrow 130$  gen, that is  $3250 \pm 340$  ybp

The first 37 markers contain 2033 mutations,  $2033/193/0.090 = 117 \rightarrow 133$  gen, that is  $3325 \pm 340$  ybp

The first 67 markers contain 2857 mutations,  $2857/193/0.12 = 123 \rightarrow 141$  gen, that is  $3525 \pm 360$  ybp

Those 193 haplotypes contain 14 base 12 marker haplotypes, which gives  $[\ln(193/14)/0.02 = 131 \rightarrow 150]$  generations, that is  $\sim 3750 \pm 500$  years. This 6% discrepancy between 3525 and 3750 years (or slightly higher with shorter haplotype series) points at a slight deviation of the dataset from "one common ancestor" rule. However, not much in this particular case.

As you see, all the numbers are within margin of error, and agree with what I have published earlier.

One can see that your conclusion that V13 are "close to 5000 years" and "between 5000 and 9000 years old" are completely off the mark.

**Lawrence Mayka:**

I now see that my calculation had an easy-to-make error: I forgot to exclude haplotypes with null values (other than DYS425) and obvious multi-step jumps. The Generations6 spreadsheet treats these as multiple one-step mutations, inflating the age estimate.

After excluding two haplotypes with a null value at DYS448, and a haplotype with 9 at DYS19 (an obvious multi-step mutation), Generations6 calculates the intraclade age ("Ga") of the E-V13 haplotypes to be 125 generations, or roughly 3750 years according to the 30-year generation span we ordinarily use with this spreadsheet. The coalescence age "Gacoal" is 118 generations, or 3540 years.

I'm sorry for any confusion caused by my earlier post.

**Lawrence Mayka:**

... I found one more null value in the E-V13 dataset, at DYS534. After removing that haplotype, Generations6 now calculates an intraclade age of 113 generations, or 3400 years – squarely within the range calculated by Anatole Klyosov.

**Gary Felix:**

Here is the haplotype of the V13 sample 7K years ago: 456=16, 389I=13, 390=24, 389II=31, 458=16, 19=13, 385ab=16/19, 393=13, 391=10, 439=11, 635=22, 392=11, H4=10, 437=14, 438=10, 448=20

Here is a V13 Mexico project member: 456=17, 389I=13, 390=24, 389II=30, 458=15, 19=13, 385ab=14/18, 393=13, 391=10, 439=12, 635=not tested, 392=11, H4=10, 437=14, 438=10, 448=20.

Throwing out the fast marker that rarely mutates from 3 values, 439 and the multicopy 385ab and we are 3 mutations off on 13 markers over 7K years. Mexico DNA Project Admin.

### **Anatole Klyosov**

Actually, this ancient haplotype tells us an interesting story. In a short 13 marker format (DYS 393, 390, 19, 391, 385a, 385b, 439, 389-1, 392, 389-2, 458, 437, 448) it is as follows:

13 24 13 10 16 19 11 13 11 31 -- 16 14 20

and the base (deduced ancestral) haplotype of present day E-V13 (with a common ancestor of 3200-3500 ybp) is

13 24 13 10 16 18 12 13 11 30 -- 15 14 20

There are 4 mutations between them, which (for the mutation rate constant for this 13 marker haplotype of 0.0305 mutation/haplotype/25 years) separates the 7,000 ybp fellow and the common ancestor of the current V13 by 3750 years. Since a common ancestor of the current V13 lived 3200-3500 years ago, this separation is just about right (3200-3500 + 3750 = 6950-7250 years).

Four important conclusions follow from it:

- 1) The 7,000 ybp excavated fellow was NOT a common ancestor of our contemporary V13 folks. It has a 4-mutation-off haplotype, it is distant by 3750 years, and our contemporary V13 are NOT his descendants.
- 2) However, the time is right, and calculations confirm the 7,000 years haplotype in terms that can be 7,000 year old.
- 3) Our calculations support the excavated haplotype.
- 4) E-V13 went through a population bottleneck before 3500 ybp, therefore our contemporary E-V13 descended from a man who survived that bottleneck and continued the chain of V13 generations from ~ 3500 ybp to the present time.

#### **An addition off-line (Anatole Klyosov):**

"A V13 Mexico project member" referenced to by Mr. Felix, has almost the same haplotype as the base haplotype of present day E-V13, and differs from the latter by only two mutations (DYS385a = 14 instead of DYS385a = 16)

13 24 13 10 14 18 12 13 11 30 -- 15 14 20

and therefore differs from the 7,000 years old haplotype (see above) by 6 mutations. This gives  $6/0.0305 = 197 \rightarrow 244$  generations, that is ~ 6100 years between them. Therefore, THEIR common ancestor lived  $7000+6100/2 = 6550$

years ago, plus-minus about a thousand years. In other words, alleles in the excavated 7000 year old haplotype are properly positioned with respect to the present-day E-V13 haplotypes and their common ancestor 3200-3500 years ago.

**Gary Felix:**

The ancestral haplotype must be very close if not the same as the 7 Kybp V13 because even on range expansion the modality wins out the vast majority of the time. The time of expansion of this marker out of the Balkans appears relatively close to 7K years ago.

Why "a bottleneck" showing one rather than reoccurring bottlenecks? Is this bottleneck showing up on other TMRCA estimates? One V13 survives from 3200-3500 years ago at a time when Europe was saturated with no more new lands to explore. How does this happen?

**Anatole Klyosov:**

Interesting. I have placed next to each other two haplotypes in the same format, one the 7,000 ybp and the other the "ancestral", but actually base (deduced) haplotype of the present day 193 individuals. They obviously differ by four mutations.

A reader looks at them and says: they "must be very close if not the same".

How to explain that kind of an eyesight?

A hint: Some 140,000 years ago in Africa or elsewhere "even on range expansion the modality" was winning "out the vast majority of the time". Everyone was happy and dancing and singing. However, on some reason we do not see their haplotypes today.

An explanation: Ancient haplotypes could have been expanding and winning "out of the vast majority of the time", but then .... boooom, a black plague or something, and we see only an "ancestral" haplotype of a lucky guy who miraculously survived (with a gal next to it), about 4,000 years AFTER the ancestral haplotype of his subclade arose. If his (the lucky guy's) haplotype was shifted by, say, four mutations compared with the more ancient but extinct ancestral haplotype, that is what we see now, the lucky guy haplotype. Of the lucky guy who lived some 3,400 years before present. And then we excavate 7,000 ybp-bones with a haplotype of 4 mutation off, and scratch our head - how come?

The answer: easy. A population bottleneck, or however we call it. Happened tons and tons of times.

**Gary Felix:**

The reasoning must be clarified. The idea that V13 would come down to one individual that lived 2300 to 2500 years ago is extremely unlikely barring positive selection along with extreme circumstances.

If a variant's chances of survival are at the outer fringes of its expansion and most of these populations of V13 go extinct then this could change the modality to a new subpopulation of V13. Of course expansion rate plays a big part in this scenario. This would most likely occur when populations are small and there is still range expansion. Otherwise all things being equal, any variant is outnumbered and likely to die out unless there is an effective small population size involved such as with endogamous populations for instance.

**Anatole Klyosov:**

:~))

The reasoning is always nice to clarify. I would greatly appreciate if you or someone would clarify the reasoning of origin of life on Earth. Or reasoning why we see the TMRCA for I1 ~ 3,400 ybp all across Europe, while it must be "older".

By the way, if tomorrow an excavated bone of 8,000 ybp shows I1, what would you say? That "the reasoning must be clarified. The idea that I1 would come down to one individual that lived 3,400 to 3,600 years ago is extremely unlikely barring positive selection along with extreme circumstances"?

**Gary Felix:**

Your calculation of V13 was shown to be in error by about 4K years accounting for V13's trek across Europe.

**Anatole Klyosov:**

How you or anyone showed that my calculations "shown to be in error"?

Why would not you be honest and say it differently: "I am not qualified in those calculations and know nothing about them".

Why without knowing a thing about it you say "in error"?  
A mystery.

**Sasson Margaliot:**

Anatole,

I do not understand why the field still resists the new approach proposed by you, namely cleansing the data-set from phantom-inducing subsets, which must be considered separately.

Why to insist on applying such or another software algorithm indiscriminately to any data-set, even in extreme cases like J1 Cohens where there is a sub-family that needs to be considered all by itself?

**Anatole Klyosov:**

Dear Sasson,

I have been asking this very question myself for the last four years, since I first brought the logarithmic method here, to this Forum, and listed here all the mutation rate constants for 6, 12, 25, 37, and 67 marker haplotypes. They stay essentially the same since then. In 2009 my paper appeared in JOGG, and Whit Athey wrote me (as the editor) with some surprise that all reviews to it were negative, but he liked it and wanted to publish it. In fact, there were two papers published, one "methodology" and one "practical illustrations". In fact, as I remember, he did not understand that time why the reviews were negative. There were no explanations given. Negative, period. With my 300+ peer review publications in academic field (and a number of books) I have never met anything like that. It showed me that there was something wrong with the "population genetics", or "genetic genealogy" area. It still is, as it is obvious from reading the "academic publications" in 2010-2011. The factual material there is fine and useful, but interpretations are hair-raising matters.

Then, in 2009, my paper of about 10 pages was published in Human Genetics, and it took only 7 days to get is accepted. That was a good sign, as I have only a few papers with such a record. It was about the J1 Cohens which you have mentioned. It was a critique of the approach employed in calculations in "genetic genealogy" by M. Hammer, D. Behar, T. Karafet, L. Zhivotovsky, K. Skorecki in general, and in considering the Cohen DNA genealogy in particular. What they have created (during the 10-year period) was a complete mess. I have explain in detail in the paper a more correct and justified approach, and reconsidered their data. What was then in the field in that regard? Well... nothing. "Population rates" and the flawed "calculations" in academic literature continue. Pick ANY paper published in



2011, it is mind-boggling how far those authors are from understanding what they are doing when trying to interpret patterns of mutations on a time scale. In other words, when they work with datings based on their Y-chromosomal results. There was only one article, by Cruciani et al, in which ~ 140,000 ybp was reasonably well calculated for haplogroup A in Africa (I have obtained the same date in a publication a month earlier, but using STR, not SNP, as in the Cruciani paper). However, Criciani et al have made a wrong conclusion about the African origin of the "root", as I see it. However, it is a different subject.

A few days ago a paper was accepted in an international "academic" journal, *Advances in Anthropology* (by I. Rozhanskii and A. Klyosov), which explicitly describes the methodology and exemplifies it with several thousand haplotypes from various haplogroups, with correlations of calculated data and actual (documented) genealogies, with all basic marker panels, etc. One more step ahead.

Regards,

Anatole Klyosov

# **ДНК-генеалогия для начинающих**

## **ДНК-генеалогия.**

### **О чем эта наука, что она определяет и выявляет, и кому она интересна**

**А.А. Клёсов**

<http://aklyosov.home.comcast.net>

Что за штука такая – ДНК-генеалогия? Что и как она измеряет? Что даёт? Что показывает? Можно ли с ее помощью узнать, например, национальность? Принадлежность к определенному этносу? Выявить, кто и откуда были прямые предки? Как давно жили? В каких исторических событиях принимали, или могли принимать участие?

Вот об этом и поговорим.

Начнем с того, что ДНК-генеалогия – наука по сути историческая, как и сама «классическая» генеалогия. Принципиальная разница в том, что обычная генеалогия опирается на документальные материалы – архивные свидетельства, церковные хроники, семейные воспоминания, в устной или письменной форме, предметы в семейной коллекции, и тому подобные сведения, а ДНК-генеалогия – на беспристрастные и беспристрастные научные данные, свободные от эмоциональной окраски. Эту окраску им придает уже последующее обсуждение, например, в широкой прессе. Там уже разгуливаются фантазии, манипуляции, политические и личные пристрастия. На самом же деле в ДНК-генеалогии есть всего три сугубо научных операции –

(1) Определение практически необратимых мутаций в ДНК, так называемых «снимов», которые тысячелетиями и десятками тысяч лет передаются по наследству, от отца сыновьям и от матери дочерям (у отцов и матерей совершенно разные снимы, так что здесь имеют место две совершенно разные ДНК-генеалогические линии, мужская и женская, каждая из которых ведут из совершенно древних времен до наших современников и их потомков),

(2) Определение более короткоживущих мутаций в фрагментах ДНК, которые продолжают мутировать и дальше, примерно раз в несколько

поколений, или раз в несколько десятков поколений (у разных фрагментов по-разному), и создают картину мутаций, которая у каждого человека своя, и служит ДНК-генеалогическим «паспортом» человека,

(3) Метод расчета, количественной обработки картины мутаций у групп людей, популяций, и превращение этой картины в хронологические, фактически исторические свидетельства. Наиболее обычный подход в ДНК-генеалогии – это расчет исторического времени, когда жил общий предок данной группы людей, племени, популяции, рода. И когда эти расчеты проводятся для разных территорий, регионов, стран, континентов, то зачастую выявляется картина миграций предков наших современников, когда эти миграции происходили и по каким территориям, и где потомки в итоге осели. Часто оказывается, что осели по всему миру, но в каждом регионе в свое время, и можно рассчитать, когда именно. Это можно сопоставлять с данными археологии, лингвистики, этнографии по тем же территориям, и получать более полную, более объемную картину мира и в пространстве, и во времени.

Первые две операции проводятся методами «секвенирования ДНК», то есть определения последовательности нуклеотидов в ДНК. Это – практически полностью автоматизированный метод, доступный любому квалифицированному лаборанту, при наличии, конечно, высокотехнического и дорогостоящего оборудования, ценой обычно в сотни тысяч долларов. Техник проводит секвенирование или Y-хромосомы ДНК (Y-ДНК), то есть мужской половой хромосомы, которая есть только у мужчин, или митохондриальной ДНК (мтДНК), которая есть и у мужчин и у женщин, но от мужчин потомству далее не передается. Сперматозоиды несут только мужские хромосомы. Точнее, у сперматозоидов есть только немного митохондрий, чтобы хвостом покрутить и потому «плыть против течения» (митохондрии дают энергию для этого движения), а потом митохондрии отваливаются, когда цель достигнута. Узнаете мужчин?

А женщина вынашивает ребенка, и тем самым передает ему свою мтДНК, помимо своих хромосомных ДНК, но мужской половой хромосомы среди них нет. Поэтому мать передает свою мтДНК и сыновьям и дочерям, а отец свою Y-ДНК – только сыновьям. И это – те же мтДНК и Y-ДНК, которые тысячелетиями, десятками и сотнями тысяч лет, да и миллионами лет передавались по прямой родословной цепочке от отца к сыну (Y-ДНК) и от матери дочери (мтДНК), только обрастая мутациями в ходе десятков, сотен и тысяч поколений. Потому и стала возможной ДНК-генеалогия, потому что у каждого нашего современника своя картина мутаций, отражающая этот исторический путь Y-ДНК и мтДНК. Если картина мутаций сходная – то это могут быть близкие родственники, значительно различающаяся





рассказывалось в предыдущем выпуске Вестника (октябрь 2011, том 4, №10, стр. 1908-1977).

Так вот, гаплогруппы от А до О (от африканского до китайского, он же юго-восточной Азии, условно говоря), а также S и T, имеют в седьмом маркере на 99% аллель 11, а гаплогруппы, или рода P, Q и R имеют там на 99% аллель 12. R – это большинство западноевропейцев и восточноевропейцев, включая больше половины этнических русских. Восьмой маркер у большинства мужчин планеты имеет аллель 12, но у рода I там в основном 13 и 14, а у евреев и арабов (преимущественно род J) там в основном 15 и 16. Род I – исторически «балканский» род, а также некоторые прибалтийские и скандинавские по происхождению рода. Этот восьмой маркер представляет собой триплет нуклеотидов АТА, аденин – тимин – аденин, так что почти все этнические русские рода R1a (которых в России больше половины) имеют такой прогон в своих ДНК:

АТААТААТААТААТААТААТААТААТААТААТААТААТА

Возвращаясь к теме разговора, 67-маркерные гаплотипы, типа приведенного выше, представляют большое количество информации, поскольку большинство их маркеров могут мутировать в любую сторону, и одна мутация в таком маркере происходит в среднем раз в 8 поколений, то есть примерно раз в 200 лет. В ДНК-генеалогии принято считать в «условных поколениях», 25 лет на каждое. То есть 100 поколений – 2500 лет, середина первого тысячелетия до нашей эры. Можно спорить, что поколение на самом деле бывает короче или длиннее, но спор будет пустой, потому что данных ни у кого все равно нет, чтобы знать продолжительность поколений на тысячелетия вглубь.

А вот с мтДНК в этом отношении не так хорошо, информации значительно меньше. Вот, например, написание моей мтДНК в рамках этой науки:

16519С  
263G, 309.1С, 315.1С, 477С

Мало того, что всего пять «точек опоры», по сравнению с 67 в мужских гаплотипах, так там и мутации происходят раз в три с лишним тысячи лет. Так что с генеалогическими расчетами не разбежишься. Еще одна особенность мтДНК – это то, что они весьма хаотически разбросаны по территориям, например, по всей Евразии. Дело в том, что по исторической традиции – где узаконенной, а где так сложилось – жен и наложниц умыкали где получится. Помимо того, если даже не умыкали, то жены обычно приходили в род, племя мужа, а не наоборот. Это и привело к

фактическому рассеянию мтДНК по территориям, так что о направленных миграциях мтДНК, в отличие от Y-ДНК, говорить не приходится. Бывают, впрочем, случаи, когда мтДНК образуют свои «оазисы», но это случаи довольно редкие и экзотические, как правило, на изолированных территориях, с затрудненными контактами с внешним миром на протяжении тысячелетий, например, среди некоторых племен американских индейцев. Тогда это очень интересно и информативно, но это скорее исключение, чем правило. И опять, с одной мутацией на несколько тысяч лет не до ДНК-генеалогии. Обычно при этом просто фиксируется территория, на которой такой «оазис» мтДНК найден. Этим занимается обычно наука популяционная генетика, не ДНК-генеалогия.

Но поскольку мы рассказали про Y-ДНК, то расскажем и о мтДНК более подробно. Помимо ДНК хромосом (и Y-хромосомы среди них), которые хранятся в ядре каждой клетки, молекулы ДНК находятся также в митохондриях. Митохондрии – это маленькие образования, плавающие во внутриклеточной жидкости, в цитоплазме. Их – от нескольких сотен до нескольких тысяч, даже до десятков тысяч на каждую клетку. И в каждой – короткая молекула ДНК, в виде несимметричного кольца. Длина ее – всего 16 с половиной тысяч нуклеотидов. Сравните с мужской хромосомой Y, которая в три тысячи раз длиннее, 50 миллионов нуклеотидов.

Митохондриальная ДНК (мтДНК) состоит из двух частей – выпирающая в сторону петля, и остаток кольца. Оказалось, эта петля является носителем генеалогической информации, но совершенно по другому, чем у мужчин. В мтДНК нет таких тандемных повторов, как в Y-хромосоме. Там нет подобных маркеров, о которых шла речь выше. Но мутации – есть. Время от времени, причем намного реже, чем у мужчин, считывающий фермент ошибается и вместо одного нуклеотида вставляет другой. Или вообще вставляет лишний. Поэтому мутации записываются, например, так – 1651C. Поскольку известно, что в «стандартной» мтДНК нуклеотид под номером 1651 – тимин (Т), сразу ясно, что в этом положении тимин заменен на С (цитозин). Или запись такая: 315.1C. Это значит, что после 315-го нуклеотида в «стандартную» цепь вставлен один лишний цитозин.

Иначе говоря, у мужчин ДНК-генеалогия основана на изменении числа повторов определенных маркеров в хромосомной ДНК, а у женщин – на разовых нарушениях одиночных нуклеотидов в митохондриальной ДНК. То есть совершенно другой принцип. Маркеров как таковых у женщин нет, вся петля ДНК – один сплошной маркер. А сравнивают – со «стандартной» мтДНК.

Поскольку мтДНК в основном некодирующая, то эти мутации в петле ни к чему жизненно важному не приводят. Просто запись в генетической книге учета.

А что такое «стандартная» мтДНК? А это та, которую изучили первой. Поэтому и приняли за стандарт. В 1981 году генетики в Кембридже, заречном пригороде Бостона, были готовы провести первое определение последовательности, или «первичной структуры» мтДНК. Нужна была любая плацента, клетки которой крайне богаты митохондриями. А в соседнем госпитале как раз рожала женщина. Взяли ее плаценту, выделили митохондрии, отсюда – мтДНК, и провели полный анализ ее последовательности. Поскольку это была первая последовательность мтДНК – её и взяли за международный стандарт. И мутации в последовательности отсчитывают от нее.

Впоследствии оказалось, что эта стандартная мтДНК присуща именно европейскому типу, и мтДНК женщин с «европейскими корнями» отклоняются от нее всего на несколько позиций, или вообще не отклоняются. Фамилия этой кембриджской женщины осталась нераскрытой. А последовательность ее мтДНК в генетике называют «кембриджской стандартной последовательностью».

Здесь – важное отступление, о котором я рассказал выше. Мужчины получают свои митохондрии от мамы, но своим сыновьям не передают. Поэтому митохондриальная ДНК на каждом мужчине терминируется. Нет девочек в роду – мтДНК терминировалась на мальчиках, связь этой линии с праmaterью («митохондриальная Ева») потерялась. Нет мальчиков в роду – терминировалась Y-хромосома, потерялась генеалогическая связь с праотцом («хромосомный Адам»). Нет детей – полная терминация генеалогической информации от отца с матерью. Но каждый мужчина имеет мтДНК, и ее анализ дает такую же генеалогическую информацию, как и анализ мтДНК его матери или сестры. А, повторяю, наличие такой информации – генеалогический «прострел» к прародителям – десятки, а то и больше сотни тысяч лет назад, как к «Еве» (митДНК), так и к «Адаму» (Y-хромосома), плюс информационное богатство всех мутаций на историческом пути.

Здесь надо заметить, что «митохондриальная Ева» – это вовсе не первая женщина, не библейская, а ближайшая по времени прародительница всех женщин на Земле. Та, к которой сходятся генеалогические нити от всех живущих на планете. У нее, естественно, были родители, но у «Евы» еще один ограничительный признак – у нее должно было быть по меньшей мере две дочери, чтобы от Евы и пошел тот генеалогический «разбег», в



итоге породивший все человечество. Подруги Евы не стали прародителями человеческого рода, а также не стали те, кто жили вокруг, или в отдалении, тысячи и десятки тысяч, а то и сотни тысяч лет до «Евы». Их потомство не оставило генеалогических следов в живущих в настоящее время на Земле.

Поскольку мтДНК – кольцо, то неважно, откуда начинать считать. Договорились считать от середины выпирающей петли как наиболее заметного образования. А поскольку в «стандартной» мтДНК всего 16569 нуклеотидов, то моя мутация (точнее, моих бабушек) номер 16519С произошла недалеко от точки отсчета, на 50 нуклеотидов ниже начала. Там, от номера 16001 и до 16569, находится область, относительно богатая мутациями. Ее называют областью низкого разрешения, или HVR1 (hypervariable region 1). Вторая, область высокого разрешения, HVR2, находится сразу за первой, с нуклеотида 1 до 580. Собственно, часто только там, в этих двух областях, и проводится генеалогический анализ митохондриальной ДНК.

Итак, верхняя строка в «моих» мутациях относится к HVR1, нижняя – к HVR2. Более детально это означает – одна мутация с заменой тимина в позиции 16519 на цитозин, и четыре мутации в позициях 263 (аденин на гуанин), 309.1 (вставка лишнего цитозина), 315.1 (еще одна вставка цитозина), и 477 (замена тимина на цитозин).

Строго говоря, и мутациями-то эти изменения назвать нельзя. Это – просто отклонения от «стандартной», условно выбранной последовательности. То есть вполне может быть, что это не у меня мутация в виде замены тимина на цитозин в позиции 16519, а у той женщины, чья мтДНК принята за стандарт, древний цитозин в этой позиции мутировал на тимин. Все в мире относительно. Но для целей генеалогического анализа это не так важно.

Вот несколько примеров, что такой анализ дает. Жена последнего русского царя Николая II, Алиса Гессен-Дармштадтская, внучка английской королевы Виктории, имела такой гаплотип:

16111T, 16357C  
263G, 315.1C

У нее с королевой Викторией точно такие же мутации в зоне высокого разрешения, как и у моих бабушек. И гаплогруппа та же, H, но это уже индексы митохондриальных родов, или гаплогрупп. Так что европейская общая бабушка у нас с ними одна. Кстати, принц Филип, дюк Эдинбургский и муж ныне здравствующей королевы Елизаветы II, имеет

точно такой же гаплотип, как и Алиса. Что не удивительно, они – племянники.

Поехали дальше. Как насчет Марии Антуанетты? Королевы Франции, обезглавленной в 1793 году в возрасте 37 лет? Жены Людовика XVI? Вот ее гаплотип:

16519С  
152С, 194Т, 263G, 315.1С

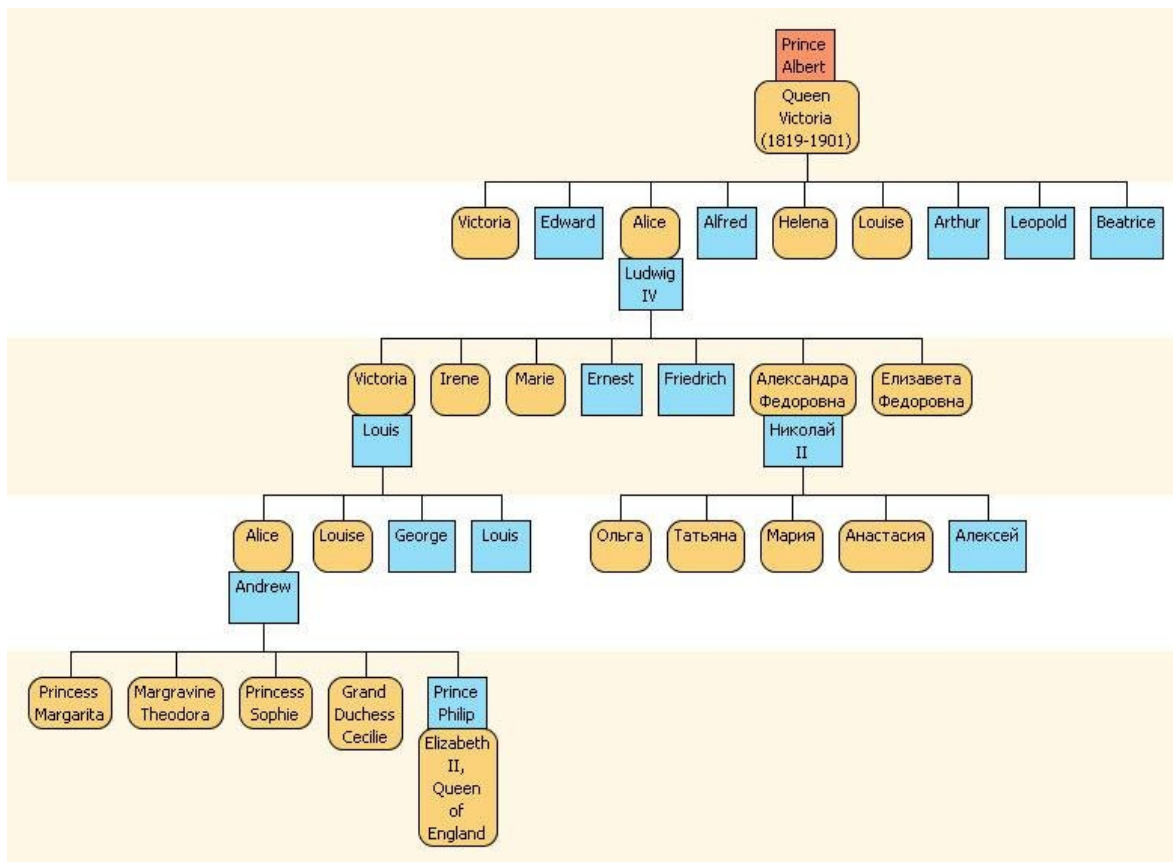
Ничего не напоминает? Ну как же, гаплотип моих бабушек по материнской линии:

16519С  
263G, 309.1С, 315. 1С, 477С

Та же «входная» мутация в подгруппу гаплогруппы Н, тот же вариант петли митДНК низкого разрешения, те же две мутации из четырех в области высокого разрешения. Даже мутация 315.1С та же, вставка лишнего цитозина в том же месте. Определенно общая бабушка у Марии Антуанетты с моими, причем уже в Европе, относительно недавно. А как недавно? Да, видимо, несколько тысяч лет назад. А «недавно» - потому что в мтДНК много родов, или гаплогрупп, и то, что мы с королевой Викторией или Марией Антуанеттой входим в одну гаплогруппу, уже означает относительно близкое родство. Принципиальное, если угодно. Один род.

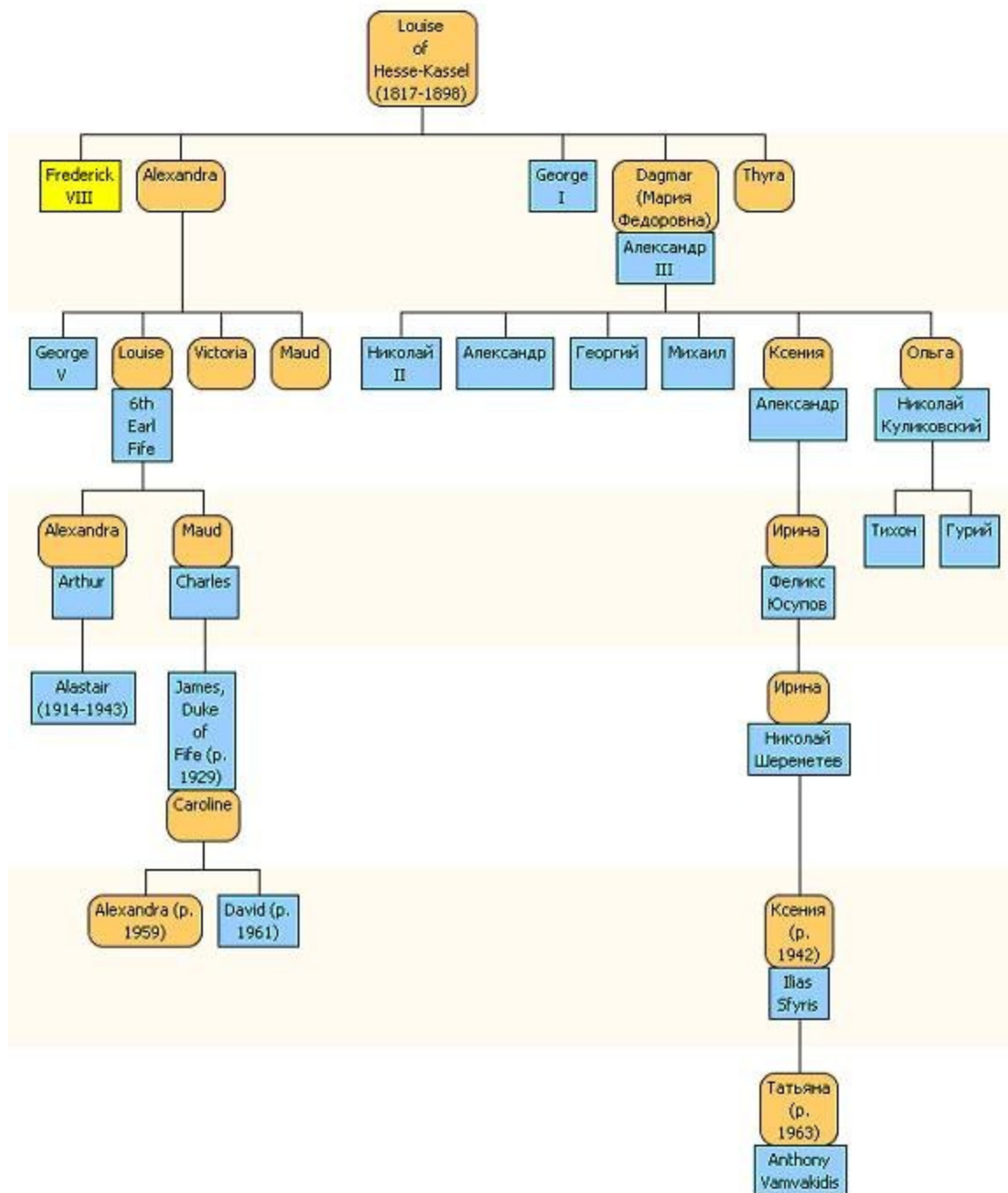
Раз уж мы погрузились в рассмотрение мтДНК, перед тем, как вернуться к Y-ДНК, посмотрим на генеалогическое древо английской королевы Виктории, которое доходит до семьи Николая Романова, то есть дочерей Николая II и Александры, погибших с родителями в 1918-м, а также до принца Филипа по параллельной линии. Как мы уже рассказывали, мтДНК передается в поколениях исключительно по женской линии. Потомству мтДНК передает только мать, из своей яйцеклетки (женской половой клетки). Мужчина мтДНК не передает. Не мужское это дело.

Вот как выглядит женская линия королевы Виктории до Александры Федоровны Романовой, она же Алиса (Аликс) Гессен-Дармштадская, внучка английской королевы:



На этом рисунке показано, как мтДНК переходит по желтым, женским линиям от королевы Виктории через ее дочь, великую княгиню Алис, и далее расходится по нескольким параллельным потокам, из которых здесь показаны два – до герцога Эдинбургского Филипа, мужа нынешней королевы Великобритании Елизаветы II, и его четырех сестер, с одной стороны, и до Александры Федоровны Романовой и ее детей – княжен Ольги, Татьяны, Марии, Анастасии и царевича Алексея. Таким образом, все мужчины и женщины, показанные на этой генеалогической схеме (кроме супругов, пришедших «со стороны»), имеют одну и ту же мтДНК. Кстати, её для тестирования и определения картины мутаций в мтДНК членов царской семьи во время расследования, чьи же останки найдены у Екатеринбурга, предоставили несколько членов монарших семей Европы, для схемы выше – принц Филип.

Но по этой схеме Николай II выпадает из рассмотрения. Он – муж Александры Федоровны, пришел «со стороны». Поэтому посмотрим на другую схему, фрагмент генеалогического дерева ветви европейских монархов, идущей по женской линии от Луизы фон Гессен-Кассель, бабушки Николая Второго. Мать самой Луизы была Шарлотта Датская, бабушка – София Фредерика, супруга принца Датского и Норвежского.



Эта ветвь, естественно, неполная, иначе в ней должны были быть отражены десятки, если не сотни человек. Суть в том, что здесь показаны две отдельные ветви, потомки которой предоставили свою мтДНК для идентификации мтДНК Николая II, и для сопоставления с ДНК других останков из могилы при известном расследовании. ДНК для анализа предоставили три человека – (а) принц Филип, муж ныне здравствующей английской королевы Елизаветы, (б) Ксения Николаевна, потомок родной

сестры Николая Второго, Ксении Александровны, и английский герцог Джеймс Карнеги, потомок королевского рода по материнской линии.

Если мы проследим по женской линии, то есть по желтым фигуркам от королевы Виктории вниз по поколениям (первая генеалогическая схема выше), мы увидим, что принц Филип должен иметь ту же мтДНК, что и дети Александры Федоровны, включая Анастасию и царевича Алексея. Как мы помним, мтДНК передается мужчинам в первом поколении, а затем на них обрывается. А вот у Николая Второго мтДНК будет уже другая, он в первой генеалогической схеме «со стороны», его мтДНК - по его матери и бабушек по материнской линии.

Зато согласно второй схеме, мтДНК Николая Второго будет та же, что у его бабушки Луизы по материнской линии, и та же ДНК будет и у его сестры Ксении, и у ее правнучки Ксении, что родилась в 1942 году, и у герцога Джеймса Карнеги, который родился в 1929 году и до сих пор здравствует. И вот именно таким образом было показано, что останки, найденные в могилах, принадлежат именно царской семье. Как именно было показано, описано в Вестнике за март 2009 года, том 2, № 3, стр. 452-481.

Возвращаемся к мужской ДНК-генеалогии, и поговорим о том, в каком виде представляют и рассматривают данные, те самые гаплотипы, фрагменты ДНК, о которых речь шла выше. Например, мой 67-маркерный гаплотип, приведенный выше, или даже мой 111-маркерный гаплотип:

13 24 16 11 11 15 12 12 10 13 11 30 - 16 9 10 11 11 24 14 20 34 15 15 16 16 - 11 11 19  
23 15 16 17 21 36 41 12 11 - 11 9 17 17 8 11 10 8 10 10 12 22 22 15 10 12 12 13 8 15 23  
21 12 13 11 13 11 11 12 13 - 31 15 9 15 12 25 27 19 12 12 12 12 10 9 12 11 10 11 12 30  
12 14 25 13 9 10 18 15 20 12 24 15 12 15 24 12 23 19 11 15 17 9 11 11

Итак, после того, как гаплотипы определили таким же образом, то есть определяя последовательность нуклеотидов в заранее определенных и согласованных участках ДНК, их помещают в базы данных, обычно открытых для людей интересующихся. Вот пример такой базы данных, в которой больше 100 тысяч гаплотипов, распределенным по разным родам-гаплогруппам по всему миру: <http://www.ysearch.org>. Есть и более узкие базы данных, как, например, <http://dna-project.clan-donald-usa.org/tables.htm>, в которой размещены гаплогруппы обширного генеалогического сообщества МакДоналдов. Они тоже распределены по гаплогруппам, что уже означает, что они принадлежат разным родам, которые друг другу далеко не родственники. 214 человек из них принадлежат роду R1a1, тому самому, к какому принадлежат большинство русских славян, и о чем речь пойдет ниже. Как древний шотландский род



**и шестом. Опытный глаз сразу увидит, сколько мутаций, где и какие.**

оказался праславянского происхождения – это вопрос исторически интересный. Древние шотландские легенды повествуют, что шотландцы произошли от скифов из причерноморских степей, которые в непамятные времена прибыли на Британские острова. Как видим, ДНК-генеалогия этому не противоречит, общие предки у этих шотландцев и у русских рода R1a1 действительно одни и те же. Из этих 214 человек более сотни произошли от одного общего предка по имени Джон Лорд Островов (John Lord of the Isles), который умер в 1386 году. Принимая условное поколение за 25 лет, получим, что лорд Джон жил примерно 26 поколений назад.

Если выписать из базы данных первые сто гаплотипов МакДоналдов рода R1a1 и построить ДНК-генеалогическое дерево с помощью профессиональной компьютерной программы, то для первых шести маркеров дерево будет иметь следующий вид (Рис. 1).

За прошедшие 26 условных поколений, или 650 лет, 75 гаплотипов из сотни его потомков остались нетронутыми в первых 6 маркерах, а в остальных 25 гаплотипах оказались 32 мутации. В ДНК-генеалогии известно, что на первых 6 маркерах средняя скорость мутаций составляет 0.0017 мутаций на маркер на поколение в 25 лет, или 0.0103 мутаций на весь гаплотип на поколение. Поэтому считаем:  $32/100/6 = 0.0533$ , то есть за время, прошедшее со времен лорда Джона, накопилось в среднем 0.0533 мутации на маркер. Делим на 0.0017, и получаем, что общий предок всех ста человек жил примерно 31 поколений назад, то есть примерно 775 лет назад. Но поскольку 32 мутации – это статистически небольшая величина, то погрешность такого расчета составит 20.31%, если формально применить математическую статистику, что, конечно, с такой точностью здесь не имеет смысла. Получается  $31 \pm 6$  поколений, или  $775 \pm 157$  лет назад. Неплохо, поскольку мы знаем, что на самом деле Джон умер 625 лет назад.

На самом деле придирайтесь к точным годам здесь не стоит. Мутации – дело статистическое, поэтому пара поколений туда-сюда ничего по сути не изменит. Из этого примера ясно и то, что с 32 мутациями в серии данных далеко не уехать, желательнее применять не 6-маркерные гаплотипы, в 67-маркерные или даже 111-маркерные, точность расчетов сразу резко возрастает.

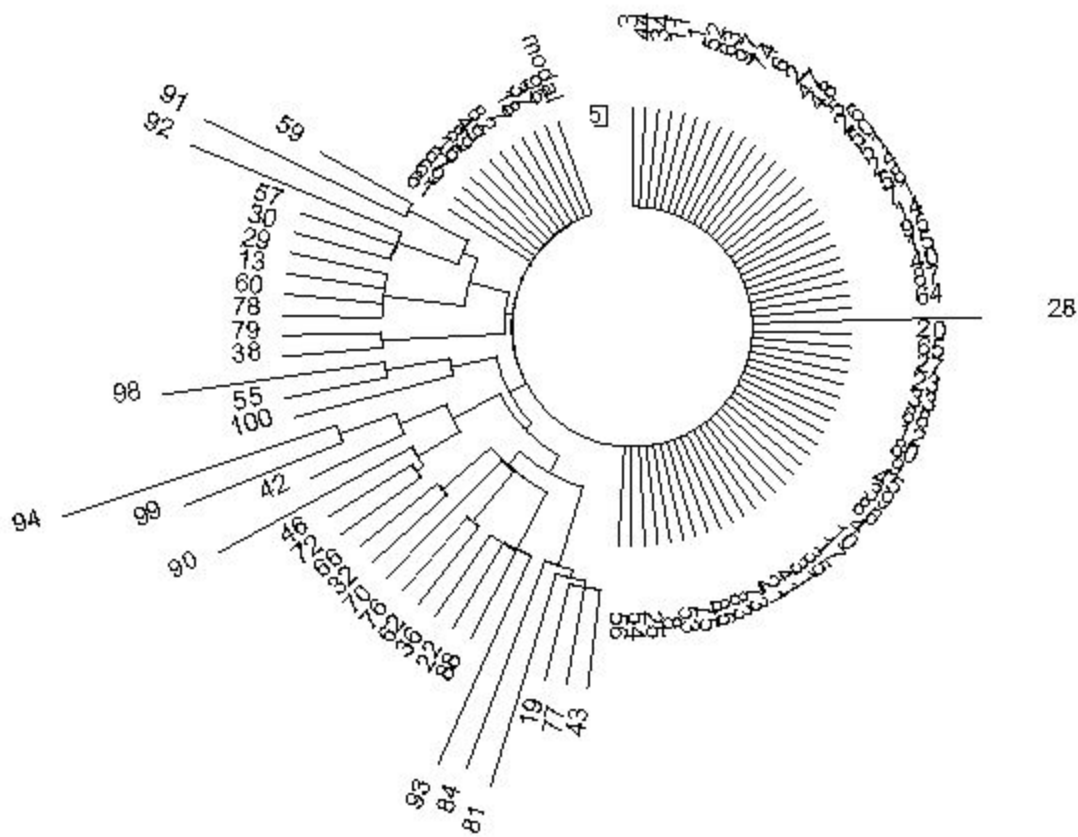
Так что расчеты в ДНК-генеалогии в простейшем случае сводятся к тому, что считается число мутаций во всей серии гаплотипов, потом рассчитывается, сколько мутаций в среднем приходится на один маркер, и делится на константу скорости мутации, которая несколько различается

для гаплотипов разной длины, но не очень значительно. Для 6-маркерных гаплотипов приведенного формата это 0.0017 (мутаций на маркер на поколение), для 12-маркерных 0.0017, для 17-маркерных 0.0020, для 25-маркерных 0.0018, для 67-маркерных 0.0018, для 111-маркерных 0.0018. Эти и многие другие значения констант скоростей мутаций сведены в таблицы и опубликованы в научной печати.

Если еще один простой способ примерного подсчета времени жизни общего предка, при котором и мутаций не нужно считать. Нужно просто взять натуральный логарифм от отношения всех гаплотипов в серии к числу нетронутых (базовых) гаплотипов, и разделить на ту же константу скорости мутации, но уже для всего гаплотипа. В нашем случае это будет  $[\ln(100/75)]/0.0103 = 28 \pm 4$  поколений, то есть 700 $\pm$ 100 лет до общего предка. Так что в одном случае мы получили 31 $\pm$ 6 поколение, в другом 28 $\pm$ 4 поколения, а на самом деле примерно 26 поколений. Поскольку здесь мы знаем довольно точную дату из документальной генеалогии, то можем критически оценить погрешность расчета. Но обычно дата неизвестна, поэтому любая, даже приблизительная датировка приветствуется.

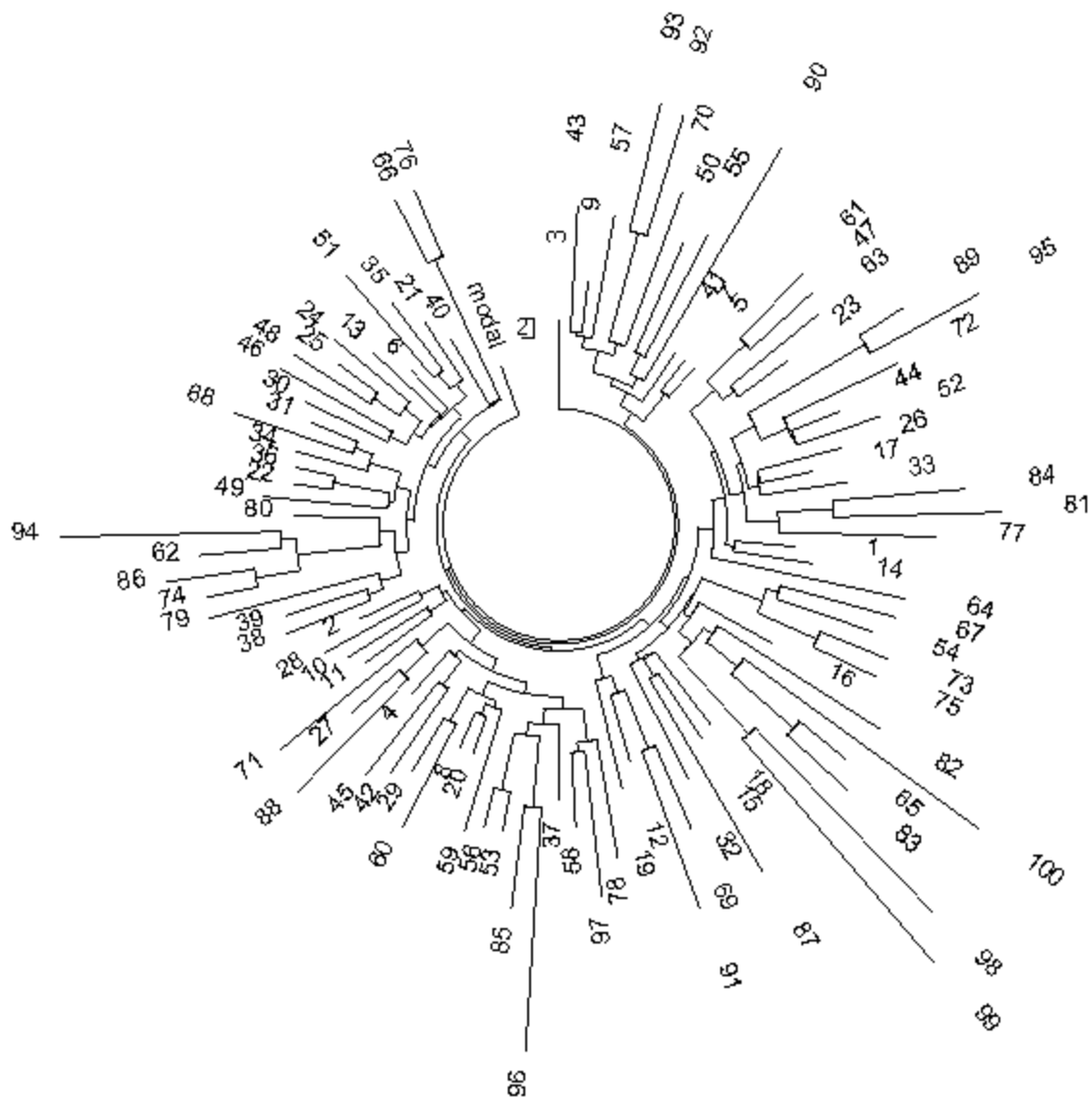
Если мы продвинемся до 12-маркерных гаплотипов, дерево несколько усложнится, поскольку возможностей для мутаций в гаплотипах окажется больше. Вот как выглядит в этом случае ДНК-генеалогическое дерево МакДоналдов:





**Рис. 2. Дерево из ста 12-маркерных гаплотипов семейства МакДоналдов гаплогруппы R1a1. Базовый гаплотип 13 25 15 11 11 14 12 12 10 14 11 31**

Наконец, ниже приведено дерево 67-маркерных гаплотипов МакДоналдов. Базовых гаплотипов в нем уже нет, поскольку в 67-маркерных гаплотипах одна мутация происходит в среднем в ходе 8 поколений. Ясно, что за 26 поколений вероятность того, чтобы хоть один гаплотип не мутировал, весьма мала. Все дерево содержит 308 мутаций, что дает  $308/100/0.12 = 26$  поколений до общего предка, что и совпадает с «возрастом» общего предка всей серии гаплотипов. Как видно, 67-маркерные гаплотипы оказались наиболее точными, что и следовало ожидать.



**Рис. 3. Дерево из ста 67-маркерных гаплотипов семейства МакДоналдов гаплогруппы R1a1. Базовых гаплотипов на дереве уже нет, все мутировали.**

Теперь обратимся к более далеким историческим временам. Поговорим о славянах, об этнических русских. Когда они появились, в какие времена, кому родственны по прямым мужским предкам. О наших корнях, иначе говоря.

Общие разговоры закончились, хотя они были необходимы для введения в новую науку. Теперь дадим конкретный пример, чтобы показать характер

исходных данных, суть расчетов, и вид получаемых результатов.

Для изучения ДНК-генеалогии этнических русских отбирались мужчины, которые считали своим родным языком русский, предки которых как минимум три поколения жили в одной из 12 областей Европейской части Российской Федерации и все эти как минимум три поколения говорили на русском языке, и среди которых не было родственников как минимум в трех поколениях. Оказалось, что среди этнических русских имеются три основных по численности рода, назовем их условно восточно-славянский, западно-славянский и финно-угорский. Условно – потому что это лингвистическая классификация, но она в целом, и не без причин, перекликается с классификацией ДНК-генеалогии. А почему бы ей и не перекликаться, поскольку рода поначалу говорили на своих языках? Рода обычно возникают на разных территориях, среди разных людей, потому и языки поначалу разные.

В классификации ДНК-генеалогии эти рода различаются по мутациям: все без исключения «восточные славяне» имеют снип M198, мутация цитозина в тимин в определенном участке Y-хромосомы, все «западные славяне» - снип M170, мутация аденина в цитозин в другом участке, все «угро-финны» - снип M46, мутация цитозина в тимин еще в другом участке. Эти мутации-снипы все пронумерованы, и их в наших ДНК много миллионов. Для ДНК-генеалогии сейчас используют 20 мутаций-снипов для главных родов, а всего, для всевозможных подчиненных родов, уже около тысячи снипов. Род «восточных славян» называется R1a1, «западных славян» I, «угро-финнов» - N1c. Это и есть под-рода основных родов R, I, N. Эти рода разошлись от своих общих предков десятки тысяч лет назад, и вот опять для миллионов людей сошлись в Российской Федерации.

Когда эти рода имели у этнических русских общего предка? Где эти предки жили? Когда пришли на земли, ставшие намного позже Россией?

Это – вопросы ДНК-генеалогии.

Всего для данного исследования было отобрано 545 этнических русских.

У-у-у, скажет скептик, 545 человек из 140 миллионов в Федерации, из которых около 70 миллионов мужчин? Ну, для начала, не все 70 миллионов этнические русские, не у всех русский – родной язык, не все живут поколениями на одной территории. Но, безусловно, приведенные выше критерии прошли бы миллионы. Так что вопрос скептиков совершенно понятен. Любой новичок в ДНК-генеалогии начинает с этого вопроса. На него есть два ответа – один полусерьезный, другой – серьезный. Первый –

чтобы убедиться, что море соленое, вовсе не обязательно пробовать воду в каждой бухточке и с каждого берега. Скептикам рекомендуется попробовать, и интересно, на какой попытке они признают, что вовсе необязательно пробовать воду тысячи и десятки тысяч раз, а то и миллионы раз. Дело в том, что если хорошо перемешано, то разницы нет.

То же и с этническими русскими. За тысячелетия перемешаны так, что вряд ли можно найти места, где люди так и живут компактно с прошлой эры. Войны, призывы в армию, переезды из деревни в города и наоборот, переезды в институты для обучения, на стройки, насильственные перемещения, в том числе ссылки, заключения, Сибирь, Дальний Восток, в общем, Россия в этом отношении одна из самых мобильных стран. А серьезный ответ в том, что это специально изучалось. То есть гаплотипы отбирались из разных регионов, от Калининградской области до Камчатки и Дальнего Востока, от Сочи до Норильска, и всегда это был набор по сути одних и тех же гаплотипов, с одним и тем же суммарным числом мутаций, и, соответственно, одним и тем же «возрастом» общих предков восточных славян, западных славян и финно-угров, которых уместно назвать северными славянами. Родной язык-то у них русский, славянский. Количественные данные будут приведены ниже.

Оказалось, что половина всех этнических русских, а на юге – в Курской, Белгородской, Орловской областях, более 60%, то есть практически две трети относятся к роду R1a1. Это – тот же род, что и западноевропейский род R1b1 (многие англичане, французы, немцы, голландцы, скандинавы), которые составляют примерно 60% западной и центральной Европы. Как видно, это тот же род на двух уровнях родства – R1. Только на третьем уровне идет разделение, на R1a и R1b. Оба рода – европеиды, но это мы и так знаем. К роду I относятся 21% этнических русских, к «финно-угорскому», а на самом деле к уральскому роду N1c1 – 14%. О них – ниже. Среди этнических русских есть и численно менее представительные гаплогруппы (рода) R1b (5%), J2 (3%), E (3%), G (2%), K (2%), F (1%), C (0.4%) [цифры округлены, проценты – от общей численности этнических русских в РФ]. Для примерной ориентации можно сказать, что R1b – это потомки башкир, чувашей, казахов, тувинцев, уйгур, а также западноевропейцев, как показано выше. Что их связывает?

А связывает то, что предки первых и пришли в Европу примерно пять тысяч лет назад, и стали западноевропейцами. Общий предок рода R1b появился в Центральной Азии примерно 16 тысяч лет назад, и был европеидом. Об этом тоже можно поговорить отдельно. J2 – это потомки

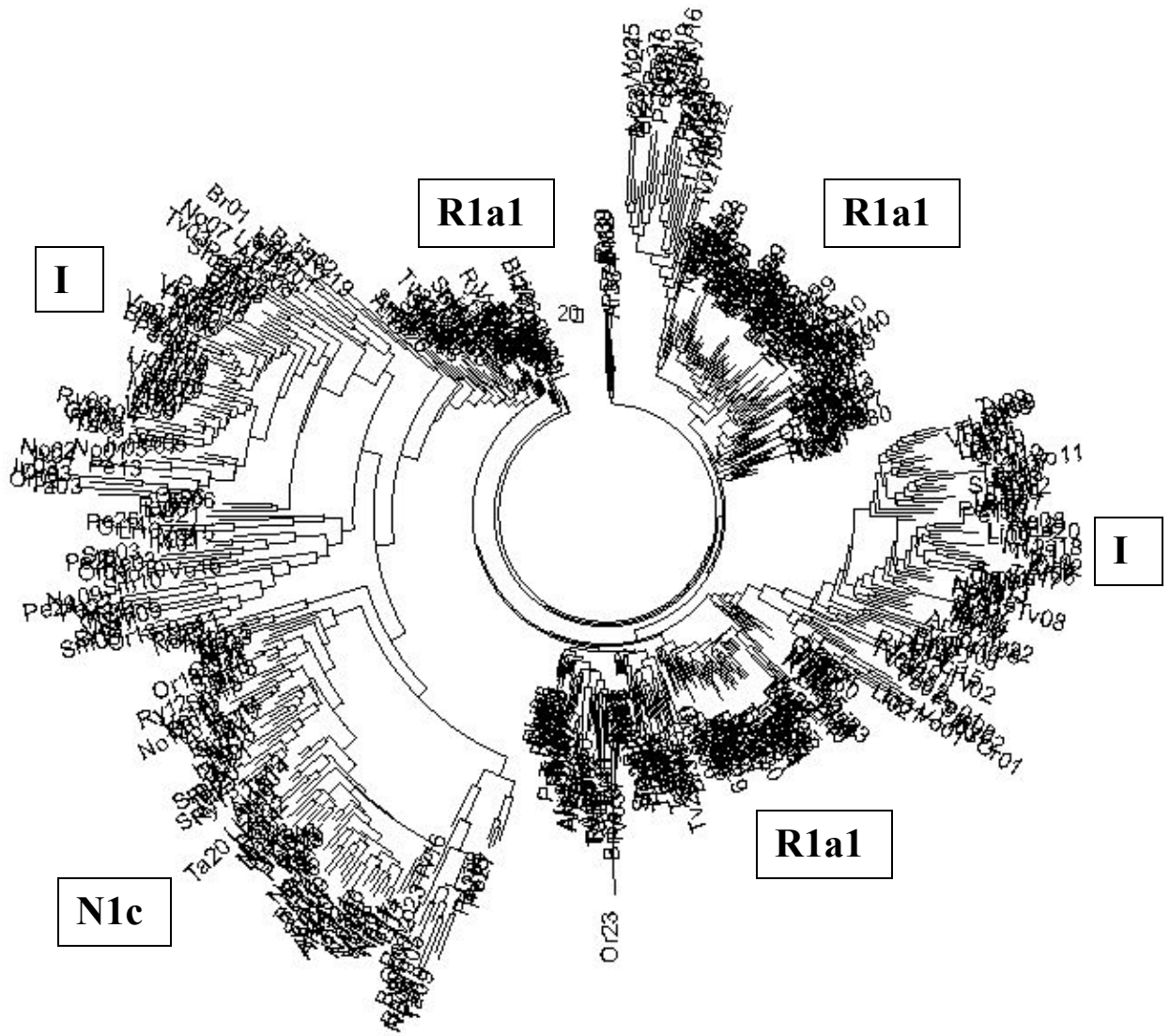


Рис. 4. Дерево из 545 17-маркерных гаплотипов по двенадцати областям Российской Федерации (Архангельская, Брянская, Вологодская, Ивановская, Липецкая, Пензенская, Новгородская, Орловская, Смоленская, Рязанская, Тамбовская, Тверская). Показаны ветви наиболее представленных (в численном отношении) гаплогрупп (родов) в Российской Федерации - R1a1 (47% в среднем, до 62% по территориям), N1c (14% в среднем, более представлена на севере РФ, от Архангельской до Псковской области включительно), I (22%, более представлена в Прибалтике и в западных областях РФ).

средиземноморцев, а также древних жителей Месопотамии, Кавказа, Ближнего Востока, часть евреев. Самый близкий к нам ареал – Кавказ, многие армяне, дагестанцы, и вообще многие горцы. Е – это Греция, Северная Африка, Ближний Восток, многие евреи. G – Кавказ (особенно Осетия), древний Иран, Турция. С – монголы, полинезийцы. Насчет полинезийцев – на Русской равнине это менее вероятно, чем монголы, но в любом случае среди этнических русских их практически нет, 0.4%. Так что если монголы и были в наших краях, то особого следа не оставили. Или молва, возвращенная на исторических науках, это сильно преувеличивает. Как и сами исторические науки.

Заметим, что отсюда уже ясно, что национальность с помощью гаплотипов-гаплогрупп не определить. Национальности появились недавно, а гаплогруппы-рода – десятки тысяч лет назад. С тех пор и разошлись, объявившись сейчас по разным национальностям, гражданствам, партийностям. Есть, конечно, более концентрированные места проживания людей определенных гаплогрупп, но все равно 100%-ной корреляции с национальностью нет.

Вернемся к вопросу, как анализируются такие серии гаплотипов. Это и есть центральное место ДНК-генеалогии. В рассматриваемой серии гаплотипов выделяется базовый гаплотип, он же предковый, от которого мутации расходятся, как круги по воде. Так и должно быть. Чем дальше потомок отошел от своего прямого предка, тем более мутировал его гаплотип, то есть тем больше в нем мутаций. Это – закон природы. Старые горы более «изношены», чем молодые, на них больше трещин, они больше разваливаются на камни. На старых людях больше морщин, чем на молодых, это тоже определенные «мутации» кожи, да и всего организма. Эволюция в живой природе идет по пути изменений, больше времени прошло – больше эволюционных изменений. Так и мутации в гаплотипах. В 67-маркерных гаплотипах, описанных выше, очередная мутация появляется, как уже сообщалось, каждые примерно 8 поколений. В среднем 10 мутаций на каждый такой гаплотип в серии, то есть 500 мутаций в серии из 50 таких гаплотипов, показывают, что общий предок такой серии гаплотипов жил более 2000 лет назад, в прошлой эре. Это – когда  $10/67 = 0.149$  мутаций на маркер в среднем. Вот уже и первое простое сведение из ДНК-генеалогии: если среднее число мутаций на маркер больше 0.149, то общий предок серии гаплотипов (то есть группы людей, предоставивших свои гаплотипы) жил примерно 2275 лет назад. Это уже подсчитано по правилам ДНК-генеалогии, с учетом «тонких факторов», на которых здесь нет возможности, да и нужды останавливаться. А если в среднем 15 мутаций на 67-маркерный гаплотип, то есть 0.224 мутации в среднем на маркер, то общий предок жил примерно 3575 лет назад. Естественно, при

этом каждый раз рассчитывается погрешность полученных данных, их доверительный интервал. Но об этом здесь вряд ли стоит говорить в деталях.

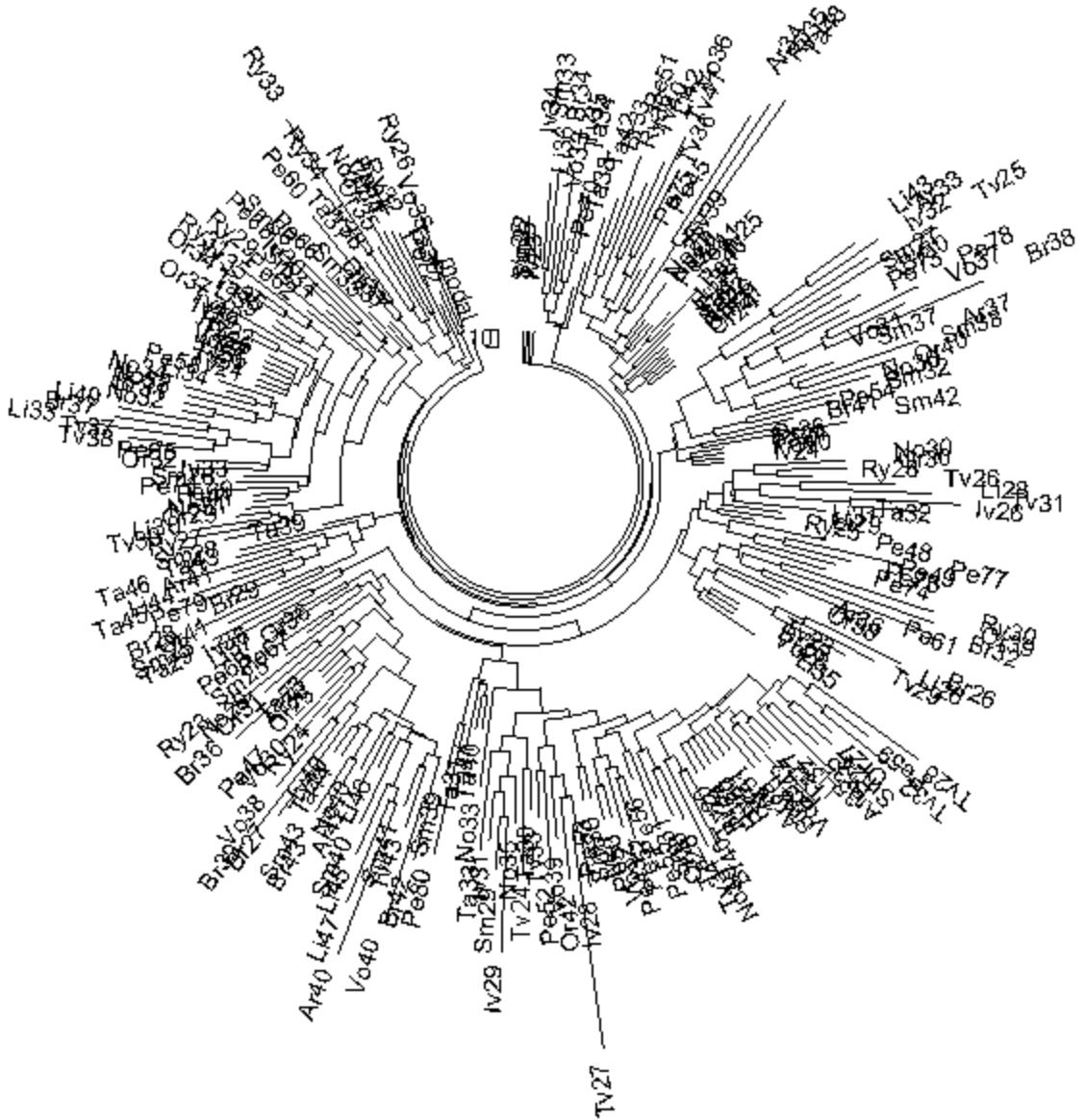


Рис. 5. Дерево из 255 17-маркерных гаплотипов гаплогруппы R1a1 по двенадцати областям Российской Федерации (см. рис. 4).

Так вот, оказалось, что все 255 гаплотипов этнических русских гаплогруппы R1a1 в данной серии имеют общего предка, который жил примерно 4475 лет назад. Середина 3-го тысячелетия до нашей эры. Эти данные рассчитывались по более коротким, 17-маркерным гаплотипам, как указано в подписи к рис. 5. Будет ли разница, если считать по более детальным, 67-маркерным гаплотипам?

Для ответа на этот вопрос взяли еще одну серию, из 148 67-маркерных гаплотипов, которые уже собирали не по 12 областям РФ, а по всей стране, до Тихого океана. Туда попали русские (92 чел), украинцы (20 чел), белорусы (12), литовцы (11), казахи (3), армяне (3), эстонцы (2), азербайджанцы (2), киргиз, латыш и молдаванин (по 1). Все они гаплогруппу R1a1, так что было логично рассудить, что в те далекие времена не было ни эстонцев, ни латышей, ни прочих национальностей, и уже только намного позже потомки R1a1 разошлись по своим странам.

Проверка серии гаплотипов специальными критериями ДНК-генеалогии показала, что общий предок у всех 148 представителей этой серии был действительно один. Гаплотип у него был такой:

13 25 16 11 11 14 12 12 10 13 11 30 -- 15 9 10 11 11 24 14 20 32 12 15 15 16 - 11 11 19  
23 16 16 18 19 34 39 13 11 - 11 8 17 17 8 12 10 8 11 10 12 22 22 15 10 12 12 13 8 14 23  
21 12 12 11 13 11 11 12 13

Это – точно такой же, какой был найден ранее у 255 человек, общий предок которых жил 4475 лет назад, на соответствующих 17 маркерах. Но здесь было решено поиграть с маркерами, чтобы глубже понять надежность расчетов. Было найдено, что если ограничиться только первыми 25 маркерами, то во всех 148 гаплотипах имеется 1037 мутаций, что дает среднюю величину  $0.280 \pm 0.017$  мутаций на маркер, то есть 4500 лет до общего предка. Такое практически абсолютное совпадение (отклонение на 25 лет, то есть на полпроцента) в данном случае, конечно, случайно, но в целом подобные совпадения, в пределах погрешности с 95%-ной достоверностью, скорее правило, чем исключение.



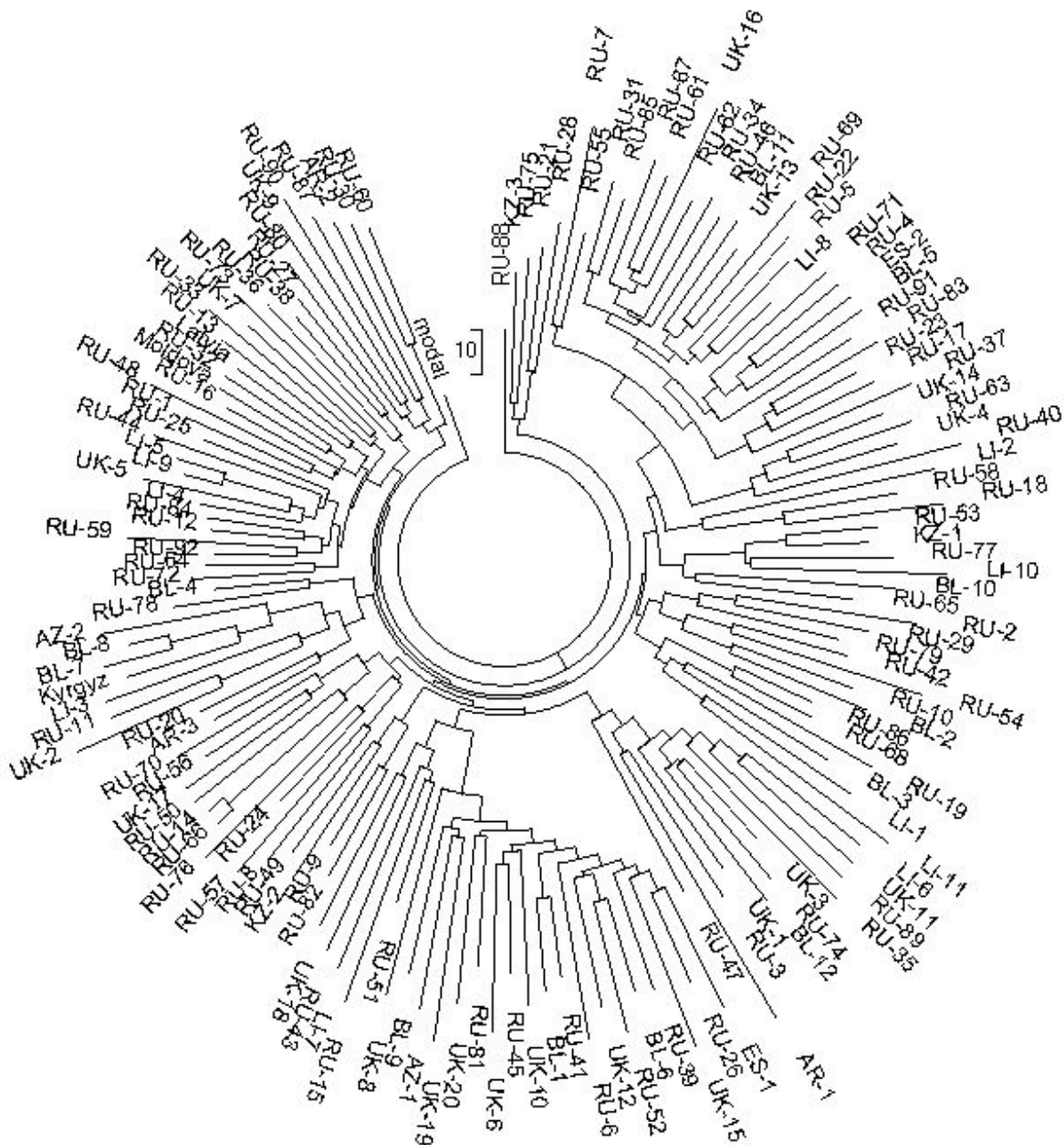


Рис. 6. Дерево из 148 67-маркерных гаплотипов гаплогруппы R1a1 постсоветского пространства. Гаплотипы взяты из базы данных YSearch. Индексы гаплотипов RU, UK, BL, LI, KZ, AR, ES, AZ соответствуют русским (92 чел), украинцам (20), белорусам (12), литовцам (11), казахам (3), армянам (3), эстонцам (2), азербайджанцам (2); Kyrgyz, Latvia, Moldova (по одному гаплотипу). Количество гаплотипов на дереве отражают два фактора - относительная численность данной гаплогруппы в популяции, и число людей, решивших провести соответствующий анализ своей ДНК.

Далее, выйдем за пределы 25-маркерных гаплотипов, и посчитаем мутации для всех 148 гаплотипов в первых 37 маркерах. Их там 2023 мутации, но и константа скорости для 37-маркерных гаплотипов выше, чем для 25- или 67-маркерных. Расчет 4475 лет до общего предка. Совпадение абсолютное. Наконец, во всех 67 маркерах было 2748 мутаций, то есть 0.277 мутаций на маркер, или 4575 лет до общего предка. Совпадение опять почти абсолютное. Средний «возраст» общего предка для приведенных 148 гаплотипов  $4500 \pm 60$  лет. Это, конечно, формальный расчет в отношении погрешности, но суть ясна.

И вот теперь стоит привести результаты расчетов возраста общего предка гаплогруппы R1a1 на Русской равнине, которые велись по мере поступления гаплотипов с июня 2008 года до настоящего времени. В то время в нашем распоряжении было только 26 25-маркерных гаплотипов в которых было всего 148 мутаций, что дало 4400 лет до общего предка. К ноябрю того же года поступило уже 44 гаплотипа, которые дали 4825 лет до общего предка. Сейчас мы уже знаем причину такого завышения, но тогда не знали. Мы споткнулись о так называемые палиндромные мутации, которые тогда в мире не были известны. В начале 2009 года у нас было уже 58 25-маркерных гаплотипов, что дало 4725 лет до общего предка. Далее серии гаплотипов включали 98, 110 и 148 гаплотипов, и все они давали примерно одно и то же время до общего предка, а именно примерно 4500 лет. Те 255 гаплотипов, что описаны выше, уже не имели палиндромных мутаций, и дали 4475 лет до общего предка, как показано выше. В целом все эти расчеты за последние три года дают среднее время до общего предка гаплогруппы R1a1 на пост-советском пространстве, большинство из потомков которого в выборке – этнические русские и их «родственники» по роду, равное  $4600 \pm 150$  лет.

Отсюда следует целый ряд выводов. Один – что расчеты практически не зависят от выборки гаплотипов, в каком бы месте России они не брались, надо только, чтобы они захватывали большие территории. Далее, что наши предки, имя им праславяне, жили на наших территориях не менее 4500 лет назад. Более того, оказалось, что те относительно немногие вкрапления гаплогруппы R1a1, которые есть на Британских островах, во Франции, Испании, Италии, Греции, северо-западной Европе, они все ведут свою родословную с Русской равнины, от праславянских предков. Гаплотипы все практически те же. Есть только одно исключение – на Балканах и в нескольких регионах Европы обнаружены значительно более древние гаплотипы, с необычными мутациями, идущие от общего предка примерно 9-11 тысяч лет назад. А сама гаплогруппа R1a1 появилась в Южной Сибири, в Алтайском регионе, примерно 20 тысяч лет назад, и ее носители,

европеоиды, мигрировали по южной части Евразии в Европу, прибыв на Балканы примерно 7-10 тысяч лет назад, по разным оценкам.

Раскопки показали, что на юге Германии носители гаплогруппы R1a1 обитали 4600 лет назад, причем гаплотипы точно такие же, как и на Русской равнине 4500 лет назад. Поскольку и там и там погрешность определения не меньше 200 лет, то нельзя сказать, кто куда пришел и откуда вышел. В сочетании с археологическими и лингвистическими данными можно представить картину, что носители гаплогруппы R1a1 прибыли с Балкан и в Германию, и на Русскую равнину примерно 4800 лет назад, и это и были легендарные арии. У лингвистов и археологов их называют арии, «индоарии», «иранцы», «индоевропейцы», хотя в те времена у них ни Ирана, ни Индии и в проекте не было. Просто тысячелетия спустя лингвисты придумали классификационную систему арийских языков, так у них и прижилось. А потом арии стали чуть ли не бранным словом, из-за нацистских псевдонаучных теорий, по которым арии были зачислены в отдельную расу со статусом «сверхчеловеков». Всё это оказалось неверным, но «осадок остался». Сейчас оказалось, что арии – это никакая ни раса, и никаких «сверхчеловеческих» качеств у них не было. Это были наши предки, гаплогруппа R1a1, и это было показано, как описано в следующих нескольких абзацах этого рассказа.

В процессе наших исследований было обнаружено, что гаплотипы большинства индийцев гаплогруппы R1a1, которых в Индии насчитывается не менее 100 миллионов человек, практически идентичны гаплотипам этнических русских той же гаплогруппы. Иначе говоря, половина русских и от четверти до трети индийцев – это потомки одного и того же общего предка, который у индийцев жил примерно 4050 лет назад, и тогда же у иранцев той же гаплогруппы R1a1. Базовый гаплотип, приведенный выше – одинаковый у русских, индийцев и иранцев рода R1a1, но общий предок на Русской равнине – старше примерно на 400-800 лет. Это влечет за собой решение интереснейшей исторической задачи, которой уже не менее двухсот лет. Эта загадка относилась к роду ариев, под этим названием они вошли в индийские веды. До самого недавнего времени наука не знала, как их обозначить в «научных терминах». Какой объективный, измеряемый параметр их объединяет? Собственно, вопрос так и не ставился. Некие «арии», пришельцы в Индостан с севера, знают снег, холода, им знакомы береза, ясень, бук, им знакомы волки, медведи, знакома лошадь. Сейчас стало известно – это люди рода R1a1, потомками которых являются, в частности и в особенности, восточные славяне – русские, украинцы, белорусы. В особенности – потому что среди них, восточных славян, доля гаплогруппы R1a1 самая значительная, и времена жизни общих предков предшествуют переходу носителей гаплогруппы R1a1 в Индостан и на

Иранское плато, что произошло, по историческим данным, примерно 3500 лет назад, в середине 2-го тысячелетия до н.э.

В качестве дополнительной иллюстрации этим исследованиям совместной истории предков на Русской равнине и в Индостане, опять приведем 67-маркерный гаплотип автора этой статьи, славянина гаплогруппы R1a1,

13 24 16 11 11 15 12 12 10 13 11 30 – 16 9 10 11 11 24 14 20 34 15 15 16 16 – 11 11 19  
23 15 16 17 21 36 41 12 11 – 11 9 17 17 8 11 10 8 10 10 12 22 22 15 10 12 12 13 8 15 23  
21 12 13 11 13 11 11 12 13

и три типичных 67-маркерных гаплотипа индийцев-«индоевропейцев», совершенно неупорядоченно взятых с индийского сайта FTDNA. Мутационные различия между ними выделены:

13 24 **17 10** 11 14 12 12 10 13 11 **32** – 16 9 10 11 11 24 14 20 **31** 12 15 15 16 – 11 **10** 19  
23 **16 16 17 20 33 34 13** 11 – 11 8 17 17 8 **11** 10 8 11 10 12 22 22 15 10 12 12 13 8 14 23  
21 13 13 11 13 11 11 12 **13**

13 24 16 11 11 14 12 12 10 13 11 **31** -- 16 9 10 11 11 24 14 20 **33** 12 15 15 16 – **10** 12 19  
23 15 17 18 **18 35 41 15** 11 – 11 8 17 17 8 12 10 8 11 10 12 22 22 15 10 12 12 13 8 **13** 23  
21 **12 12** 11 13 **10** 11 12 12

13 **23** 16 11 **12 15** 12 12 10 13 11 **30** – – 9 10 11 11 24 14 20 **30** 12 **16 16** 16 – 11 12 19  
23 15 16 18 **21 35 39 12** 11 – 11 8 17 17 8 12 10 8 11 10 12 22 22 **16** 10 12 12 13 8 14 **24**  
**22** 13 13 11 13 11 11 12 12

Степень сходства между гаплотипами видна сразу. Заметим, что число мутаций между индийскими гаплотипами попарно равно 27-30, и между славянским (по определению) гаплотипом автора и каждым из индийских – тоже попарно – мутационная разница равна 25-30. Иначе говоря, гаплотип автора ближе к индийцам, чем они сами между собой. На самом деле это различие в пределах погрешности, и различия практически равны друг другу.

Для примера - типичный базовый западноевропейский гаплотип (гаплогруппы R1b), который (и его производные) имеют примерно 60% западно- и центрально-европейцев, и до 90% жителей Британских островов, имеет вид:

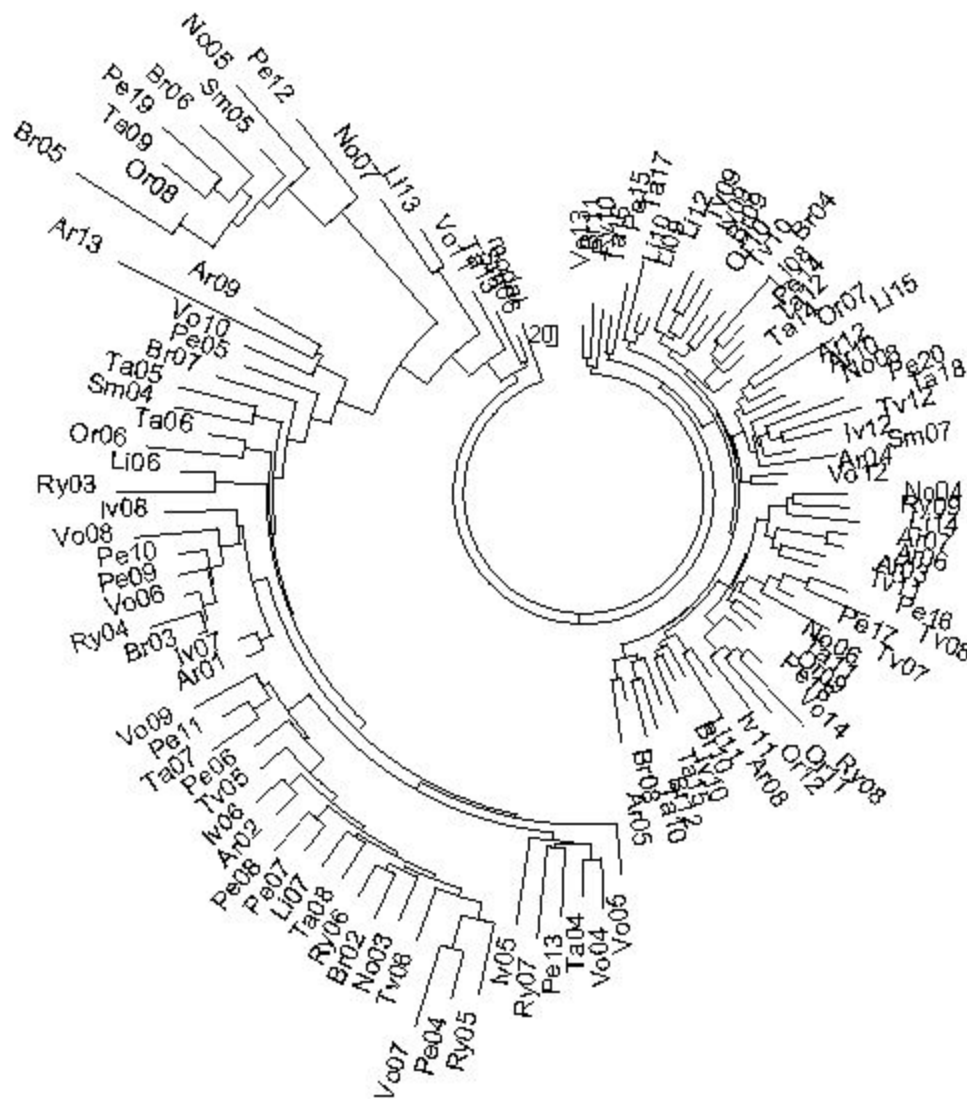
13 24 14 11 11 14 12 12 12 13 13 29 – 17 9 10 11 11 25 15 19 29 15 15 17 17 -- 11 11  
19 23 15 15 18 17 36 38 12 12 – 11 9 15 16 8 10 10 8 10 10 12 23 23 16 10 12 12 15 8 12  
22 20 13 12 11 13 11 11 12 12

Число мутаций между ним и индийскими гаплотипами (и гаплотипами этнических русских) приближается к 50-ти, что не удивительно – их предков разделяют не менее тридцати тысяч лет. В Индии и Иране гаплотипов гаплогруппы R1b очень мало, почти нет. Среди ариев 3500 лет назад предков современных западноевропейцев, похоже, не было. Стоит отметить, что в высших индийских кастах в настоящее время до 72% носителей гаплогруппы R1a1, в основном, среди браминов. В то же время среди 367 браминов не было найдено ни одной гаплогруппы R1b.

Пройдемся по результатам подобных же подходов к анализу гаплотипов этнических русских гаплогруппы I («западные славяне») и N1c («финно-угры»). Среди 545 гаплотипов этнических русских 255, как было отмечено выше, были носители гаплогруппы R1a1 (47%), 117 гаплотипов группы I1/I2 (21%), и 76 гаплотипов финно-угров (14%). Это – в сумме 82% от всей общей выборки.

Результаты детального исследования показали, что род I возник не менее 30 тысяч лет назад, причем это было скорее всего на Восточно-Европейской (Русской) равнине, и мигрировал в Европу. Видимо, предки рода I и R жили в те времена на Русской равнине вместе, чем и можно объяснить то, что и те, и другие европеоиды. Поскольку гаплогруппа I оказалась больше 20 тысяч лет назад в Европе, а гаплогруппа R – в Южной Сибири, и обе были европеоидны, то невозможно себе представить, чтобы европеоиды возникли независимо друг от друга в разных концах Евразии. Это – исключительно сложный каскад антропологических признаков, который просто не мог дважды и независимо воспроизвестись. Но в Азии 20-30 тысяч лет назад не было I, а в Европе в те времена не было R. Значит, разошлись в разные стороны. Тем более что у антропологов есть свидетельства о приходе древних людей в Южную Сибирь с запада, не менее 30-40 тысяч лет назад.

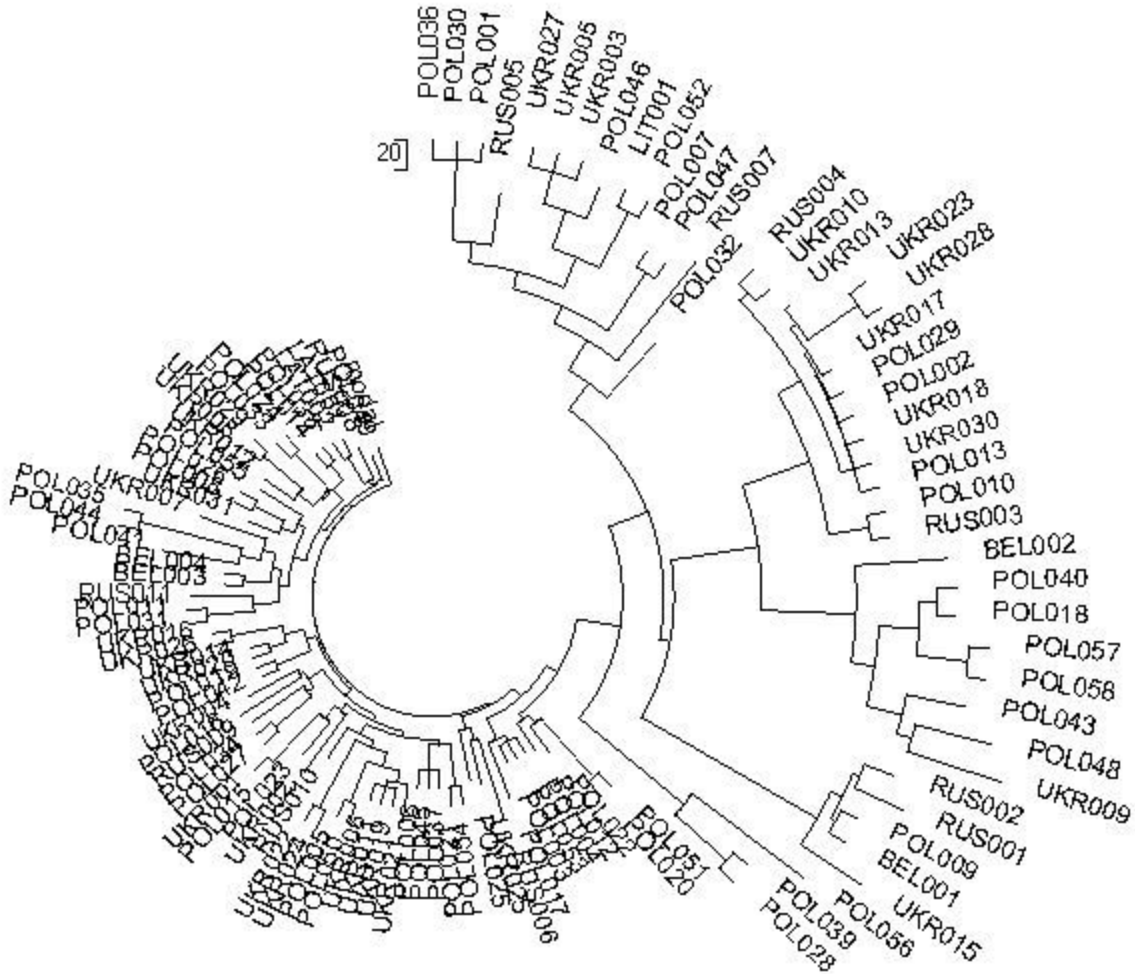
Вернемся к этническим русским рода I, 117 гаплотипов которых приведены на рис. 7. Их дерево впечатляюще разошлось на две ветви. Справа – молодая ветвь гаплогруппы I2 (примерно 3000 лет до общего предка), слева – ветвь гаплогруппы I1 (3650 лет до общего предка (средняя часть ветви), и древней гаплогруппы I2 (10500 лет до общего предка). Более точные датировки указаны в подписи к рисунку. Эти гаплогруппы, I1 и I2, разошлись от родительской гаплогруппы I не менее 15 тысяч лет назад, и эти ветви каждая имели свою историю и своих предков, живших в разные исторические эпохи.



**Рис. 7.** Дерево из 117 17-маркерных гаплотипов этнических русских сводной гаплогруппы I по двенадцати областям Российской Федерации. Справа - молодая ветвь гаплогруппы I2 (3000±380 лет до общего предка), слева - ветвь гаплогруппы I1 (3650±800 лет до общего предка (средняя часть ветви), и древней гаплогруппы I2 (10500±1100 лет до общего предка)

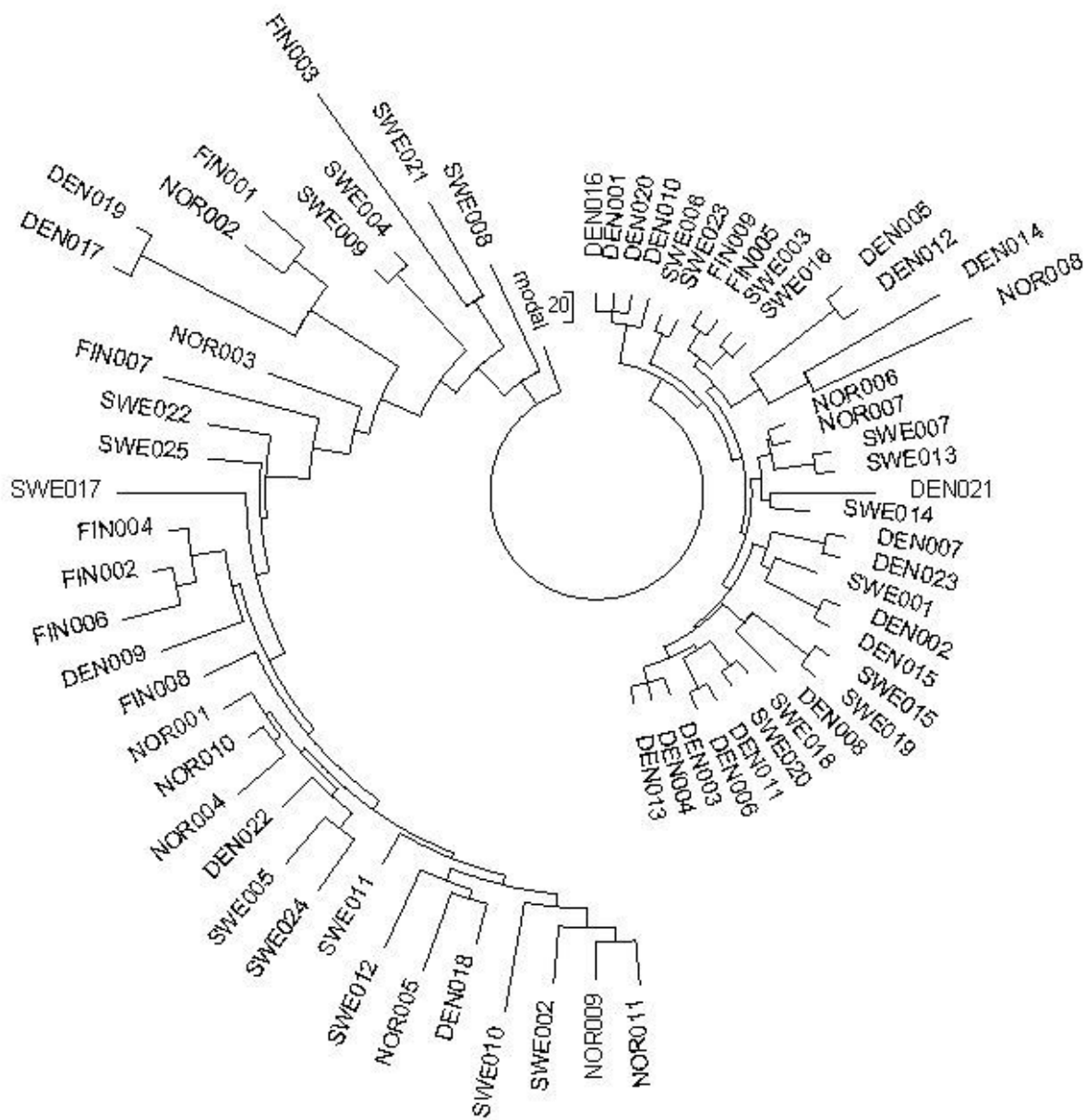
Мы уже знаем, что среднее число мутаций на маркер является показателем, как давно жил общий предок серии гаплотипов. Помните, для протославянских гаплотипов (род R1a1) на Русской равнине с общим предком, который жил 4500 лет назад, среднее число мутаций на маркер было 0.277-0.280. А у протославянских же (род I2) гаплотипов - 0.597

мутаций на маркер. Это и дает примерно 10500 лет до общего предка. Только-только начали сходиться ледники Великого оледенения.



**Рис. 8. Дерево 25-маркерных гаплотипов Восточной Европы гаплогруппы I2. Дерево состоит из 96 гаплотипов, построено по данным базы данных YSearch. Индексы гаплотипов соответствуют России, Украине, Белоруссии, Польше, Литве, Латвии (последняя - только в «молодой» ветви).**

Теперь используем совершенно другую серию гаплотипов рода I2, 25-маркерных, взятых из базы данных YSearch, про которую шла речь выше. Как видно, принципиальная картина (рис. 8) та же, только дерево развернуто в другую сторону, что не имеет никакого значения ни по сути, ни для расчетов. Здесь уже - не только этнические русские, а Восточная



**Рис. 9. Дерево 25-маркерных скандинавских гаплотипов гаплогруппы I2, построено по данным базы данных YSearch. Дерево состоит из 68 гаплотипов. Индексы гаплотипов соответствуют Швеции, Норвегии, Дании, Финляндии.**

Европа - Россия, Украина, Белоруссия, Польша, Литва, Латвия. Здесь возраст молодой ветви -  $2650 \pm 320$  лет, что несколько «моложе», чем для этнических русских ( $3000 \pm 380$  лет), но совпадает в пределах погрешности расчетов. Возраст древней ветви -  $10800 \pm 1200$  лет, что практически



совпадает с возрастом ветви I2 для этнических русских (10500±1100 лет). То есть род I2 в Восточной Европе – это один и тот же род, с одними общими предками, независимо от нынешних государственных границ.

Базовый (предковый) гаплотип один и тот же, и в 25-маркерном формате он имеет вид

13 24 16 11 14 15 11 13 13 13 11 31 – 17 8 10 11 11 25 15 20 32 12 14 15 15

А вот в Скандинавии общие предки данной гаплогруппы совершенно другие, хотя возраст ветвей тот же. Дерево гаплотипов гаплогруппы I2 Скандинавии приведено на рис. 9. Правая ветвь имеет общего предка, который жил 3025±470 лет назад, что в пределах погрешности совпадает с временем жизни общего предка этнических русских «молодой» ветви (3000±380 лет назад), и восточноевропейцев в целом (2650±320 лет назад). Но базовый гаплотип этой скандинавской ветви совершенно другой

13 22 14 10 13 14 11 14 11 12 11 28 – 15 8 9 8 11 23 16 20 28 12 14 15 15

Он отличается от восточноевропейского (включая и этнических русских) на 26 мутаций на 25 маркерах (!), что разводит их общих предков на 29 тысяч лет. Это означает, что обнаруживаемая в гаплотипах молодой ветви гаплогруппы I2 картина мутаций – только «верхушка айсберга», то, что биологи называют «прохождением бутылочного горлышка популяции». От древнего общего предка остались только (чудом) выжившие потомки, которые тысячи лет назад дали побеги новой «молодой» ветви в Скандинавии, другой ветви в Восточной Европе, и только по ним, обрывкам древней популяции, можно выявить, что сама ДНК-линия гаплогруппы I2 ведет от предка той же гаплогруппы, который жил более 20 тысяч лет назад, видимо, в Европе. Это – времена неандертальцев.

Другим обрывком той же древней ветви является ветвь I2 возрастом примерно 10 тысяч лет (10500±1100 лет у этнических русских и 10800±1200 в Восточной Европе в целом). В Скандинавии эта древняя ветвь (левая ветвь на рис. 9) тоже совершенно другая, чем у восточноевропейских I2. Их базовые гаплотипы соответственно равны:

13 23 14 10 14 15 11 14 11 12 11 28 – 16 8 9 8 11 24 16 20 28 12 14 14 15

13 23 15 10 15 15 11 13 11 13 12 29 – 16 8 9 11 11 24 14 20 27 12 14 15 16

В первой, скандинавской ветви, в среднем 0.509±0.065 мутаций на маркер, что дает 9575±1140 лет до общего предка. Во второй, восточноевропейской

ветви –  $0.556 \pm 0.045$  мутаций на маркер, что дает  $10800 \pm 1200$  лет до общего предка.

Между их базовыми гаплотипами – 13 мутаций, что помещает их общего предка примерно на 15 тысяч лет назад. Это всё – тоже «обрывки» древней линии гаплогруппы I2, а до того – I, древнейшей гаплогруппы в Европе.

Следует отметить, что по виду «молодая» ветвь гаплогруппы I1 на рис. 7 действительно оказывается самой «молодой» в сводной гаплогруппе I. Ее общий предок жил  $3650 \pm 800$  лет назад, и имел (базовый) гаплотип в 25-маркерном формате

13 22 14 10 13 14 11 14 11 12 11 28 – 15 8 9 8 11 23 16 20 28 12 14 15 16

Помимо того, есть небольшая серия гаплотипов рода I1, все европейские (Франция, Англия, Швейцария, Греция, Германия и Польша), которые разбегаются по мутациям настолько, что необходимо примерно 21400 лет для такого разбега. Это и есть минимальный возраст общего предка гаплогруппы I1 в Европе. Общий же предок гаплогрупп I1 и I2, то есть общий предок гаплогруппы I, уходит вглубь более чем на 30 тысяч лет назад, и, вероятно, более 40 тысяч лет назад.

Перейдем к этническим русским «финно-угорского» происхождения, гаплогруппа N1c. Здесь надо подчеркнуть, что это – такие же этнические русские, как и упоминались выше, и отбирались для тестирования по тем же критериям, что и остальные. Это – славяне, как и «западные» и «восточные». По аналогии их можно назвать «северными славянами». Основная разница в происхождении древних предков в том, что древние носители гаплогруппы N1c мигрировали тысячелетия назад из Южно-Сибирского региона, видимо, с Алтая, и их язык был реконструирован как принадлежащий алтайской языковой семье, и далее уральской группе языков. Гипотеза об урало-алтайской семье языков существует с 18-го века. Так что носители гаплогруппы N1c скорее уральцы, чем угро-финны.

На рис. 10 приведено дерево гаплотипов этнических русских гаплогруппы N1c. Его анализ показал, что все гаплотипы тоже происходят от одного общего предка, который жил  $3525 \pm 400$  лет назад. Это – середина 2-го тысячелетия до н.э. В это время арии (гаплогруппа R1a1) уже уходили в Индостан и на Иранское плато.



получает сведения о мутациях, накопившихся в определённых участках ДНК и может сравнивать характер (рисунок) этих мутаций у конкретных людей или их коллективов, популяций, этнических групп, народов. Это позволяет получать сведения о передвижениях предков современных (и ископаемых) носителей анализируемых молекул ДНК, опять же в пространстве и во времени, вплоть до времён 50 – 80 тысяч лет назад и на любой территории. На самом деле, вплоть до любых времён жизни предков человека, до сотен тысяч и миллионов лет назад, но наука на таких временах пока имеет слишком малую базу.

Помимо того, любая человеческая популяция в определённые моменты времени проходила то, что генетики называют «бутылочным горлышком» популяции. Это означает, что популяция, будь то род, племя, или просто группа родственников, сокращается в размере настолько, что или прекращает свое существование, генеалогическая линия прерывается (не прошли «бутылочное горлышко»), или сокращается до нескольких, или буквально до одного человека, потомство которого в итоге выживает и увеличивается в числе. В таких случаях этот человек и оказывается «общим предком» выжившей популяции. Методы расчёта времени жизни общего предка данной, выжившей популяции ведут именно к тому человеку, картина мутаций в ДНК его потомков сводится именно к нему, потому что «разбег» мутаций в ДНК потомков отсчитывается именно от этого, выжившего предка. Он становится, в рамках понятий ДНК-генеалогии, общим предком данной популяции.

«Буквально до одного человека» здесь – это в понятии ДНК-генеалогии, то есть скорее «до одного гаплотипа». Людей, оставшихся в данной популяции, могло быть несколько, это могли быть отец и сын, или братья, группа родственников, у которых гаплотип мог и не различаться, или различаться несколькими мутациями, но в итоге выжил только один гаплотип, и именно от него пошёл «разбег мутаций» у потомков.

Методы ДНК-генеалогии позволяют узнать, когда жил общий предок, и, значит, – когда имело место «бутылочное горлышко» популяции на абсолютной шкале времени. При этом при наличии умеренной статистики, а именно при рассмотрении всего нескольких десятков или (лучше) сотен, а иногда и тысяч образцов ДНК потомков, можно идентифицировать времена жизни общих предков с точностью, близкой к 10% при 95%-ной достоверности полученных абсолютных значений времени. Так, если общий предок данной популяции жил 5000 лет назад, то при наличии (путём тестирования) сотни гаплотипов, то есть определённых фрагментов Y-хромосомы современников, потомков данного общего предка, время жизни общего предка определится с точностью  $5000 \pm 530$  лет до настоящего

времени (для сотни 25-маркерных гаплотипов). Это означает, что время жизни общего предка данной популяции попадает в указанный интервал, между 4470 и 5530 лет назад, с надёжностью 95%.

Такая точность вызвана тем, что в сотне 25-маркерных гаплотипов (то есть когда в каждом гаплотипе имеется по 25 нуклеотидных маркеров, и каждый мутирует с определённой средней скоростью на протяжении этих тысяч лет) в распоряжении исследователя есть 2500 «экспериментальных точек», в которых на протяжении  $5000 \pm 530$  лет произойдёт, как исследователь знает, примерно 770 мутаций. Эта величина жестко завязана на среднюю скорость мутаций, установленную и калиброванную. Если мутаций меньше, то общий предок жил более недавно, и опять можно вполне надёжно установить, когда он жил, как показано на ряде примеров в настоящей статье.

Выше было показано, что общий предок этнических русских гаплогруппы I1 жил  $3650 \pm 800$  лет назад. Это было определено всего по паре десятков 17-маркерных гаплотипов. Для сравнения, приведем результаты массивного исследования, в котором рассматривалась выборка из 857 25-маркерных гаплотипов Англии, тот же род I1, то есть 21425 «экспериментальных точек». Эти гаплотипы содержали 4868 мутаций от базового гаплотипа. Такое число мутаций позволяет определять среднее число мутаций на маркер (а именно это среднее число фактически используется для расчёта времени жизни общего предка, как было показано выше) с точностью до 2 – 3% с 95%-ной надёжностью. Было найдено, что общий предок для 857 англичан в данной выборке жил  $3425 \pm 350$  лет назад, причем базовый, то есть предковый гаплотип был совершенно такой же, идентичный. Это – очередная тайна ДНК-генеалогии, кто же от кого произошел три с половиной тысячи лет назад. Кто принес предковый гаплотип, который оказался одним и тем же у русских и англичан? И такие тайны в ДНК-генеалогии – на каждом шагу.

Данные здесь примеры показывают, что ДНК-генеалогия – это не просто нарождающаяся наука с зыбким фундаментом (что характерно для новых областей наук). За последние несколько лет ДНК-генеалогия практически закончила формировать расчётный базис, платформу, и временные расчёты проводятся теперь с достаточной надёжностью. Было экспериментально показано, с использованием геномов многих популяций людей, а также геномов шимпанзе, что мутации представляют собой действительно «молекулярные часы», скорость которых в ДНК неизменна на протяжении по меньшей мере последних двух миллионов лет.

Было экспериментально показано на тысячах пар «отец-сын», что мутации в гаплотипах действительно равновероятны по «направлению», и тандемные, повторяющиеся блоки нуклеотидов, называемые в ДНК-генеалогии маркерами, могут укорачиваться или удлиняться на блок (то есть менять чисто аллелей) с одинаковой вероятностью.

Сложнее с географической привязкой, с выявлением того, где именно жил общий предок, так как время его жизни (то есть сколько поколений или лет назад он жил) о географии не говорит. Здесь показателен приведенный выше пример про англичан и русских одного и того же рода – П. Для выявления того, где, на каких территориях жили общие предки популяции, приходится привлекать независимые данные археологии, антропологии, лингвистики, понимая частую условность их сведений, поскольку о том, что те ископаемые предки действительно дожили в потомках до наших дней, приходится только гадать.

Именно потому союз антропологии, археологии, лингвистики с ДНК-генеалогией так важен. Как уже отмечалось выше, ДНК-генеалогия предоставляет в их распоряжение жёсткую привязку в виде «метки» рода, определённую и однозначно определяемую мутацию, снип, в Y-хромосоме ДНК, которая (мутация) всегда сопровождает каждого члена рода. Эта мутация не ассимилируется в популяциях, как ассимилируются языки, культуры, религии, физические черты, антропологические показатели.

Эта мутация, снип, одна и та же в смешанных популяциях, позволяет отличить члена рода через тысячи и десятки тысяч лет. Она позволяет проследить миграции родов и отдельных представителей рода. Она позволяет понять, останки представителей каких родов находятся в археологических раскопах, и как археологические культуры связаны друг с другом – не только через материальные, культуурообразующие носители, но и через людей, через конкретные рода, понять генезис, динамику археологических культур, добавить важнейшую компоненту к динамике человеческих популяций и их материальных носителей.

Похоже, что в ряде случаев потомки обитателей древних стоянок 50 – 45 тысяч лет назад действительно дожили до настоящего времени. Но это может быть верифицировано только путём анализа ДНК ископаемых останков. Технически это возможно – определением нуклеотидной последовательности фрагментов костной ДНК, глубоко упрятанной в костях и частично пережившей тысячелетия. Эти исследования пока немногочисленны, фактически единичны, эта работа очень трудна и очень дорогостояща.

Следует ещё раз подчеркнуть – ДНК-генеалогия наших современников «докапывается» только до «бутылочных горлышек» популяций, родов, племён, генеалогических линий, связывающих современников с их предками. «Бутылочное горлышко» – это далеко не обязательно результат мора, эпидемий, войн, природных катаклизмов в прошлом. Хотя все эти факторы непременно имели место и оказали влияние на состав современных популяций. Трудно представить, насколько чума середины 14-го века, которая выкосила четверть европейцев, терминировала генеалогические линии, гаплотипы и, возможно, целые гаплогруппы, рода. Много генеалогических линий начинаются именно в середине 14-го века. Это – выжившие люди, поведшие линию популяции опять сначала, ставшие «общими предками» многочисленных групп наших современников.

Геноцид – худший враг ДНК-генеалогии, не говоря обо всём человечестве. Сколько генеалогических линий прервала резня в Армении второго десятилетия прошлого века, геноцид евреев, цыган, других национальных, религиозных групп и популяций – об этом можно только догадываться. Сколько генеалогических линий прервалась в мировых войнах 20-го века, прерывается сейчас... Любая война вносит необратимый вклад в уничтожение генетического «материала». Галльские войны Юлия Цезаря, по данным Плутарха, привели к гибели более миллиона жителей Центральной Европы, и ещё миллион был угнан в рабство. Опять, можно только гадать, как это изменило ландшафт гаплотипов и гаплогрупп в Европе.

Но не только войны и эпидемии создают «бутылочные горлышки», но и переезд, переход, миграция носителя гаплогруппы и гаплотипа на новое место. Если миграция массовая или даже нескольких человек, то эти люди «переносят» мутации своего общего предка на новое место, и, с точки зрения ДНК-генеалогии, общий предок мигрантов как был тысячелетия назад, так и остался. Так, например, общий предок англичан в Англии и в США – один и тот же. Как и русских, восточных славян в США, например. Это вовсе не означает, что восточные славяне жили также и в Северной Америке 4500 лет назад, как на территории современной России. Тем не менее, подобные заключения – типичная ошибка начинающих в области ДНК-генеалогии. Так, Исландия была заселена только в 9-м веке нашей эры, а общий предок исландцев практически по всем родам – тысячелетия назад, как и в континентальной Европе.

С другой стороны – общий предок гаплогруппы R1b1 по расчётам ДНК-генеалогии (по мутациям в гаплотипах) жил в Центральной (Средней) Азии 16 тысяч лет назад; на территории современной России, среди этнических

русских – 6775 лет назад; на Ближнем Востоке (в частности, среди евреев, а также в Ливане) 5500 – 5200 лет назад, на Пиренеях – 4800 лет назад, в Ирландии – 3800 – 3400 лет назад. Вот такой шлейф даёт обоснованное представление о временах и направлении миграций рода R1b1, более того – даёт основания связать этот род с определенными археологическими культурами (иногда – с серией культур, горизонтом, культурно-исторической общностью).

Отсюда и новый термин – молекулярная история, то есть создание исторических реконструкций, исходя из молекулярных характеристик ДНК потомков, а порой и (ископаемых) предков. Поскольку далёкие предки, передвигаясь, несли в новые края языки, то, прослеживая миграции предков, происходившие сотни, тысячи и десятки тысяч лет назад, можно получать сведения о миграции языков во времена столь глубокой древности. Сопоставление этих реконструкций с данными лингвистики, полученными принципиально другими методами, может позволить получать более обоснованные сведения в области языкознания, проверять существующие концепции и приходить к новым, совершенно неожиданным концепциям и идеям.

Так ДНК-генеалогия повернулась своей неожиданной и непредсказуемой ранее гранью к наукам гуманитарным, историческим, лингвистическим. Хотелось бы – чтобы и в отношении установления и укрепления мира на планете. Все мы, независимо от расовых и религиозных отношений, потомки одних и тех же, не очень удаленных общих предков. Анализ гаплотипов разных популяций показывает это опять и опять, что все люди в конечном итоге – родственники. Более того, сейчас уже можно довольно надежно рассчитывать степень этого родства, и ДНК-генеалогия в этом играет важную роль.



**Обращения читателей и персональные случаи ДНК-генеалогии**

**LETTERS FROM THE READERS: PERSONAL CASES**

**Part 33**

**Anatole A. Klyosov**

Newton, Massachusetts 02459, U.S.A.  
<http://aklyosov.home.comcast.net>

**LETTER 109**

I am from Kuwait. May I know what is the TMRCA for those J2 haplotypes which I am sending along with this letter?

**MY RESPONSE:**

Please explain what do you want to know from those haplotypes, and in which form you would like to obtain an answer. Also, why did you chose THOSE haplotypes, what was the main idea of that selection?

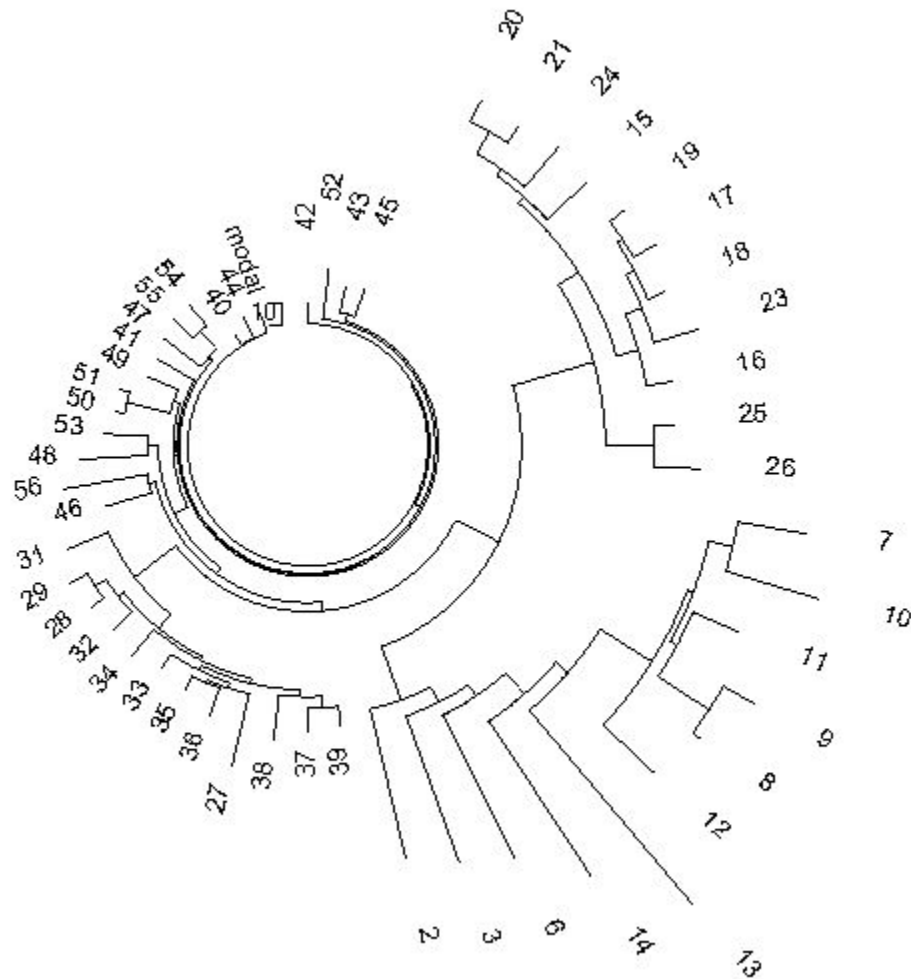
What do you expect an TMRCA would be? I just want to know at which level I should answer your question.

**CONTINUATION:**

I choosed those haplotypes due to all of them have participated in the Levant project and as I compared those haplotypes together they appeared to have the TMRCA with me in 2600-3300 year timeframe. So I would like to verify it and also to know in what time the branches split from each other. And why I choosed those because they were nearest to me as I compared and I want to know relationship between them, and split of branches, and direction of immigration in ancient times. But, first of all, I wanted to see my "family".

MY RESPONSE:

O.K., understood. Attached is a tree composed of all those 51 of 67 marker haplotypes (the tree of 56 of 37 marker haplotypes is essentially the same). The numbering is according to your list (you are No. 11).



Your "family" is on the lower right-hand side, a flat branch with haplotypes 7 to 12. They are (the names are omitted in this Proceedings):

- 7 - Al-S.. (Kuwait)
- 8 - M... (Kuwait)
- 9 - Al S... (Saudi Arabia)
- 10 - Al-... (UAE)
- 11 - you
- 12 - K... (UAE)

Now you know your actual origin, this is all your family. All are from J2 haplogroup.

All the six individuals have 53 (actual) mutations in their haplotypes, which gives  $53/6/0.12 = 74 \rightarrow 80$  conditional generations (25 years in each), that is  $2000 \pm 340$  years to a common ancestor of all the six individuals. This is the beginning of the Common Era.

Congratulations, and best regards,

CONTINUATION:

Thank you very much. May I know all of those haplotypes, plus subclades J2a4h + J2a4 - have common ancestor how many years ago?

MY RESPONSE:

It goes to the beginning of times for haplogroup J2, that is 15-20 thousand years ago. For example, a common ancestor of only two branches, of your flat branch and the next one on the top (counter clockwise) lived 9600 years ago. A common ancestor between your branch and a flat branch on the left lower side (haplotypes 28-39) lived about 10,000 years ago, but it had a different common ancestor, because a common ancestor between that last branch and that on the right top lived 6900 years ago. That is, a common ancestor of only those three branches lived ~ 15,000 years go.

## LETTER 110

You have considered my case in Letter 55. I live in Northern France, my ancestors (since at least the 13<sup>th</sup> century) lived in Antwerp, and my haplogroup is R1a1-M17, which is quite rare among Flanders. You have suggested that my haplotype belongs to the "Old Scandinavian branch" based on a position of my branch on the R1a1 haplotype tree (Letter 55, Proceedings, 2011 June, vol. 4, No. 6, p. 1349).

I would like to know more about this branch, and it has already helped me thank by your numerous and edifying work on our haplogroup. R1a.org was as decisive.

I read that "Western Eurasian branch" belongs to the family of lineage which are positive to SNP - Z280 and that "Old Scandinavian branch" is characterized by

SNP - Z284. My 37 markers will eventually be enough to define my branch. Do you think that research of this SNP - Z284 should confirm your verdict?

MY RESPONSE:

The both questions you have addressed are still uncertain. First, as I have explained earlier, your 37 markers are not good enough to make a definite conclusion to which branch you belong. The Old Scandinavian branch looks better in your case, however, your haplotype has 30 markers missing (compared to the 67 marker haplotypes), and those 30 missing alleles can change the story. As you see, you have close to only half of the markers, that is why ANY assignments to branches are only temporary. Some people have only few mutations from the ancestral haplotypes of the branch, and in those cases it is rather easy to place them on the branch where they belong. Some people, and you among them, have more mutations, and you are between several branches with your mutations in the 37 markers. So, based on only 37 markers, your question will never be solved completely. Only with some probability.

Regarding the SNPs there is also some uncertainty. They are very new SNP and the work still continues with their assignments to the branches. Positions of Z280 and Z284 are still uncertain, however, it looks like Z280 indeed defines the Eurasian branch, and Z284 embraces both the Old and Young Scandinavian branches. Probably in a couple of months this will be more clarified. It would be a good idea for you to test the both SNPs, 280 and 284, even if you have the 67 marker haplotype (which you do not have as yet).

## **LETTER 111**

I live in the Netherlands. Recently I have read your article "Origin of the Jews and the Arabs: Date of their Most Recent Common Ancestor is Written in their Y-Chromosomes - However, There were Two of Them" and I saw that you have described there the same haplogroup - J2a4b1 - to which I belong to. So, here is my very recent story.

67 marker haplotype testing helped me to identify a family member unknown to me before, and with the same last name. The archives confirmed that we are distant relatives with a common ancestor who lived back in 1580. In principle, we can come back in the archives even until 1375. However, a mystery for us is that our 67 marker haplotypes, of both of us, are identical. This is after 430 years! It seems that we both are have the oldest lineage in which two 67 marker haplotypes are identical. How can you explain it?

MY RESPONSE:

Yes, indeed, your case is rare but not impossible. To clarify it, I have contacted the FTDNA company, which ran your tests, and they confirmed that the the first 67 markers you and your relative have identical alleles, however, in DYS635 (which was not reported to you since it is not a part of the 67 marker haplotype) your alleles differ by one mutation. You have there 20, and your relative has 21.

Now, let me explain some statistics. On average, one mutation in a 67 marker haplotype happens once in 208 years. This calculations is based on the mutation rate constant which is equal to 0.12 mutation per haplotype per 25 years (a "conditional generation"). However, this is on average. The "average" has in fact a rather wide distribution. In mathematical statistics, we say that "with 68% probability (that is "one sigma" in mathematics) one mutation happens between  $208 \pm 208$  years, that is between now and 416 years ago".

In other words, if to collect a huge amount of people in 67 markers of which happened just one mutation, then 68% of those people have it happens between 1595 AD and this time (meaning their generation).

In your case it was between 1580 and now. As you see, you are pretty close, and this is only with 68% of those with one mutation in 67 marker haplotypes.

Granted, in your case one mutation happened not in 67 marker haplotypes, but in a more extended haplotype including DYS635). So, technically, you do not fit to those 68% of all those people.

When we move to 95% probability ("two sigma"), that is a timespan when 95% of all the mutated 67 marker haplotypes fit, the time interval is increased by two, that is, it becomes  $208 \pm 416$ , an stretches back by 624 years, to 1387. Now you clearly fit this time interval.

This confirms my statement above, that your case is rare, but not impossible.

Best regards