

БИОЛОГИЧЕСКАЯ ХИМИЯ КАК ОСНОВА ДНК-ГЕНЕАЛОГИИ И ЗАРОЖДЕНИЕ «МОЛЕКУЛЯРНОЙ ИСТОРИИ»

Обзор

© 2011 г. А.А. Клёсов

*MIR International, Inc., 36 Walsh Road, Newton,
Massachusetts 02459, USA; E-mail: aklyosov@comcast.net*

Поступила в редакцию 17.11.10

После доработки 17.12.10

Изложены основы новой научной дисциплины, формирующейся в последние несколько лет, — ДНК-генеалогии. Методологической базой новой дисциплины является использование аппарата химической (биологической) кинетики для анализа картины неупорядоченных мутаций в определенных нерекомбинантных участках ДНК. Цель анализа — перевод динамической картины мутаций в хронологические, временные показатели, датировки времен жизни общих предков популяций, древних родов и племен. Приведены примеры определения времен жизни древних предков этнических русских: восточных славян (род R1a1), западных славян (род I1 и I2), северных, или уральских славян (род N1c), которые жили — по «записям» в их Y-хромосоме — примерно 4600 лет назад (предки современных русских рода R1a1), 3650 лет назад (рода I1), 3000 и 10 500 лет назад (рода I2, две основные ветви), 3525 лет назад (рода N1c) (доверительные интервалы датировок приведены в статье). Полученные значения сопоставлены с соответствующими датировками общих предков «индоевропейского» рода R1a1 в Индии, потомки которых составляют большинство (до 72%) в высших кастах Индии (датировка общего предка 4050 лет назад), и показано, что гаплотипы этнических русских рода R1a1 (до 62% общего населения РФ) и индийцев рода R1a1 (>100 млн. человек в Индии) практически совпадают вплоть до протяженных гаплотипов (67 маркеров). Это фактически решает двухсотлетнюю загадку, кем на самом деле были арии, пришедшие в Индостан примерно 3500 лет назад. Проведено также сопоставление гаплотипов и датировок древних предков этнических русских групп I1 и I2 с гаплотипами и датировками общих предков этих групп в Восточной и Западной Европе и в Скандинавии и показано, что излагаемый подход фактически закладывает основы «молекулярной истории», в которой функцию основного инструментария выполняет высокотехнический анализ молекул ДНК современников и ископаемых останков.

КЛЮЧЕВЫЕ СЛОВА: ДНК-генеалогия, гаплотип, гаплогруппа, мутации, Y-хромосома.

В данной статье описывается новая наука, которая только создается на основе стремительного накопления экспериментального материала. Ежедневно в базы данных поступают десятки новых «экспериментальных точек», которые по принципу обратной связи корректируют методологию новой науки, что приводит к уточнению методов расчета.

Имя этой науки — ДНК-генеалогия. Ее экспериментальные данные — это картина мутаций в нерекомбинантных участках мужской половой хромосомы (на самом деле и в митохондриальной ДНК, но в этой статье речь пойдет только об Y-хромосоме, иначе объем статьи выйдет за пределы допустимого), причем картина мутаций в Y-хромосомах как отдельных людей, так и их

групп, популяций. Методология новой науки — перевод динамической картины мутаций в хронологические показатели, во времена жизни общих предков популяций, а на самом деле — общих предков древних родов и племен. Фактически расчетным путем определяются времена, когда в древности жили эти роды и племена.

Мутации, рассматриваемые в ДНК-генеалогии, — это или одиночные (как правило) замены нуклеотидов в ДНК, например аденина на цитозин или цитозина на тимин, или вставки нуклеотидов, или делеции, или более сложные мутации, при которых ошибка копирующего фермента приводит к переносу целого блока нуклеотидов, tandemного, как его иногда называют. Этим обуславливается или удлинение серии та-

ких блоков на один (редко сразу на два-три блока), или их укорачивание. Первые мутации — SNP (single nucleotide polymorphism), «снипы», обычно очень стабильны. Для ДНК-генеалогии отбирают только такие, которые случаются 1 раз (максимум 2 раза) за историю человечества. Поэтому они являются маркерами человеческих родов. Мутации второго типа — STR (short tandem repeat) — значительно более быстрые. Они происходят в определенных локусах ДНК раз за несколько десятков или сотен поколений. Поэтому гаплотипы ДНК выбирают так, чтобы в них таких локусов было как можно больше (но все-таки чтобы оставаться в рамках практичности), и в ранних работах использовались 6-маркерные гаплотипы, потом 12-маркерные, затем 25- и 37-маркерные, а сейчас работа рутинно ведется с 67-маркерными гаплотипами (правда, в академических публикациях это обычно от 8 до 17 маркерных гаплотипов). В 67-маркерных гаплотипах, например, одна мутация происходит в среднем за период жизни восьми поколений.

Поскольку уже показано и доказано, что эти мутации в локусах происходят в основном (или исключительно) неупорядоченно, то к ним стало возможно применять правила и подходы химической (или биологической) кинетики (кому какой термин больше нравится, суть одна). Каждому маркеру свойственна определенная константа скорости мутации, и каждому гаплотипу (состоящему из набора маркеров) тоже свойственна определенная средняя константа скорости мутации [1, 2]. Поэтому чем древнее общий предок популяции, тем больше мутаций накапливается в гаплотипах его потомков по сравнению с ним, тем больше поколений (и лет) отделяет современных потомков от их общего предка, и это число поколений (и лет) рассчитывается методами ДНК-генеалогии.

Сразу приведем конкретный пример, чтобы показать характер исходных данных, суть расчетов и вид получаемых результатов.

Среди этнических русских (а для их изучения в работе [3] отбирались мужчины, которые считали своим родным языком русский, их предки (как минимум три поколения) жили в одной из 12 областей европейской части Российской Федерации, все (эти как минимум три поколения) говорили на русском языке, и среди них не было родственников как минимум в трех поколениях) имеются три основных по численности рода. Назовем их условно (потому что это лингвистическая классификация) восточно-славянским, западно-славянским и финноугорским. В классификации ДНК-генеалогии эти роды различаются по мутациям: все «восточные

славяне» имеют снип M198 (rs2020857), мутация цитозина в тимин у их общего предка произошла в определенном участке Y-хромосомы; все «западные славяне» — снип M170 (rs2032597), мутация аденина в цитозин у их общего предка произошла в другом участке; все «финноугры» — снип M46 (Tat, rs34442126), мутация цитозина в тимин произошла также в другом участке. Род «восточных славян» называется R1a1, «западных славян» — I, «финноугров» — N1c. Эти роды разошлись от своих общих предков (десятки тысяч лет назад) и опять для миллионов людей сошлись в Российской Федерации.

Когда эти роды имели у этнических русских общего предка? Где эти предки жили? Когда пришли на земли, ставшие намного позже Россией? Это — вопросы ДНК-генеалогии.

В работе [3] приведены списки 17-маркерных гаплотипов 545 этнических русских, отобранных по описанному выше принципу. По этим данным нами было построено общее дерево гаплотипов. Программа, строящая дерево (см. работы [1, 2]), сортирует гаплотипы по принципу похожеści картины мутаций и располагает их по ветвям согласно предполагаемой последовательности мутаций. Наиболее мутированные (по сравнению с другими) гаплотипы уходят вдаль от дерева, наименее мутированные располагаются вблизи «ствола». Исходные, немутированные, т.е. «предковые», гаплотипы (их называют базовыми) «сидят» на стволе дерева. Их, как правило, остается мало в древних сериях гаплотипов, но много в «молодых» сериях, когда гаплотипы еще не успели мутировать, например когда общий предок жил всего 1–2 тысячи лет назад, т.е. в новой эре.

На рис. 1 показано дерево гаплотипов всех 545 этнических русских. Хотя 17-маркерные гаплотипы не обладают хорошей разрешающей способностью, свойственной более протяженным (25-, 37-, 67-маркерным) гаплотипам, видно, что основные гаплогруппы на дереве в целом разрешены; остальные небольшие, как правило, смешанные ветви объединяют гаплотипы этнических русских из численно менее представительных гаплогрупп (родов) R1b (5%), J2 (3%), E (3%), G (2%), K (2%), F (1%), C (0,4%) (числа округлены, 1%—сотая доля общей численности этнических русских в РФ).

Столь сложные деревья в основном иллюстративные. Для их количественного анализа выделяют отдельные ветви, тем более когда они подтверждены анализом снипов (т.е. заранее известно, к какому роду относится любой гаплотип дерева), как в данном случае. На рис. 2 приведено дерево гаплотипов гаплогруппы R1a1, основной по количеству среди этнических

гаплоти́па»: 16 11 14 13 30 25 11 11 13 14 11 10 20
16 15 23 11.

«Базовый» — это фактически синоним понятия «предковый», но отражающий некоторую схематичность понятия, его расчетный характер. Эта запись и отражает «тандемные мутации». В первом маркере, или локусе (DYS19, где **DYS** — это **DNA Y Segment**, локус номер 19), определенная последовательность нуклеотидов

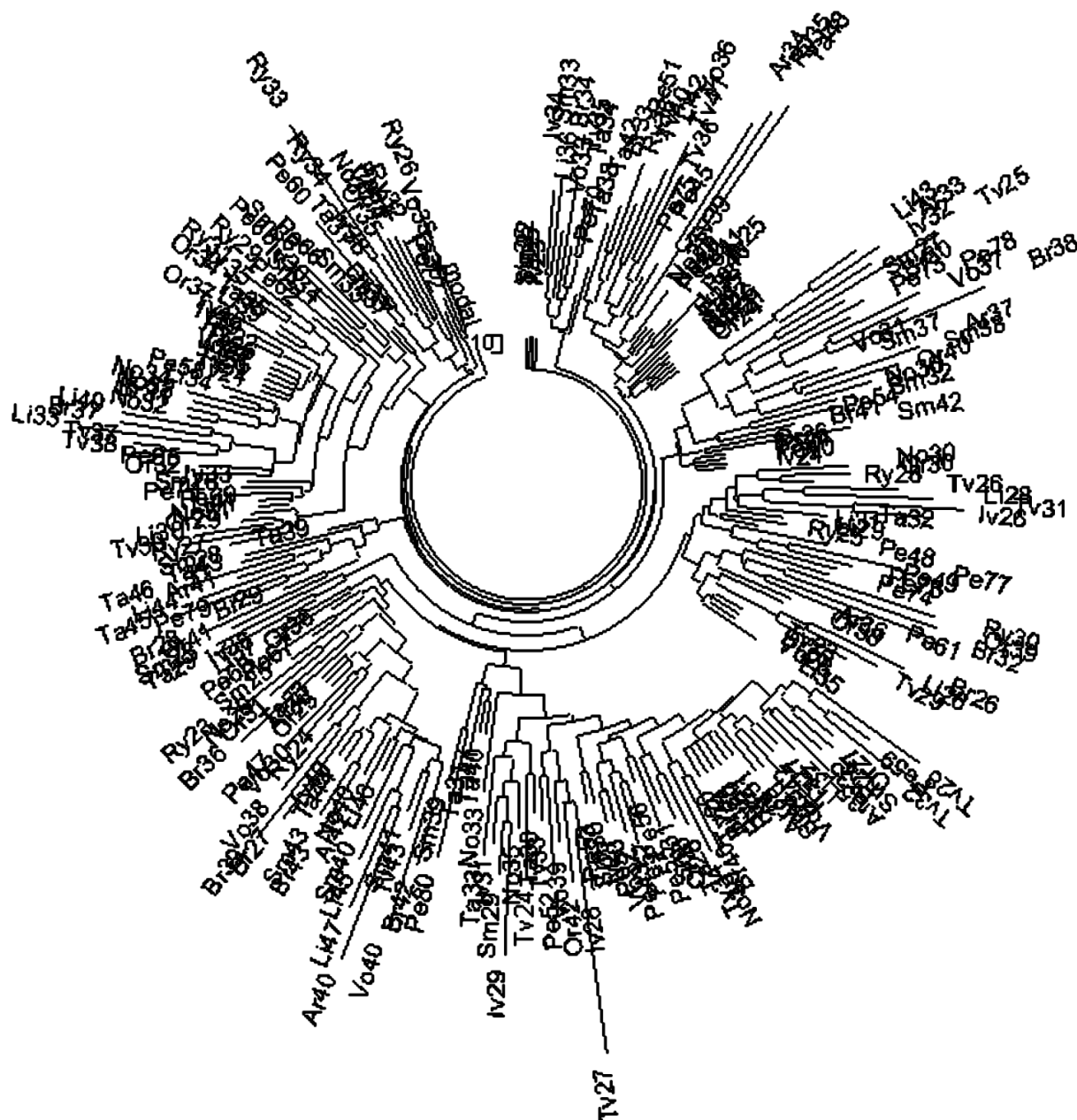


Рис. 2. Дерево из 255 17-маркерных гаплотипов гаплогруппы R1a1 по 12 областям Российской Федерации (см. подпись к рис. 1), построено по данным работы [3]

повторяется 16 раз. Во втором локусе другая последовательность нуклеотидов повторяется 11 раз, в третьем — 14 раз, и т.д.

Оказалось, что все 255 гаплотипов, т.е. все 4335 маркеров (255×17) имеют 1320 мутаций по отношению к базовому гаплотипу гаплогруппы R1a1 этнических русских, т.е. в среднем $0,304 \pm \pm 0,017$ мутации на маркер с 95%-ной достоверностью.

Поскольку средняя скорость мутаций в 17-маркерных гаплотипах равна 0,002 мутации на маркер на «поколение» (как скорость мутаций, так и поколение калибровались в расчете 25 лет на условное «поколение», и при желании константу скорости мутаций можно всегда перекалибровать при другом числе лет на поколение) [1], то общий предок этнических русских гаплогруппы R1a1 (восточные славяне) жил 152 ± 16 поколений назад (без поправки на возвратные мутации) или 179 ± 19 поколений с поправкой [1], т.е. 4475 ± 460 лет назад (середина третьего тысячелетия до н.э.).

Методы расчетов, значения скорости мутаций, калибровка, таблицы возвратных мутаций, способы расчета погрешностей приведены в работе [1]. Проверим полученное значение $4475 \pm \pm 460$ лет назад с помощью совершенно другого набора гаплотипов, а именно 148 67-маркерных гаплотипов гаплогруппы R1a1 постсоветского пространства. Это уже не «академическая», а коммерческая выборка. В академических работах 67-маркерных гаплотипов не бывает, за исключением работ автора настоящей статьи. Причины две: дорого, и специалистам не известно, как их количественно анализировать. Приведенные 148 гаплотипов — это все гаплотипы гаплогруппы R1a1, которые оказались в базе данных YSearch на середину ноября 2010 г. Они были определены по заказу представителей национальных групп, указанных в подписи к рис. 3.

Все эти гаплотипы, независимо от того, к какой национальности относят себя их обладатели, принадлежат к одному роду — R1a1, которому, как отмечено выше, как минимум 4500 лет только среди этнических русских. Естественно, за эти тысячелетия род распространился, охватив представителей многих национальностей. Действительно, гаплотипы всех перечисленных выше представителей разных национальностей практически равномерно перемешаны на дереве и не образуют отдельных выраженных ветвей. Проверим, насколько этот «охват» соотносится с «возрастом» этнических русских гаплогруппы R1a1, восточных славян.

Базовый 67-маркерный гаплотип серии из 148 гаплотипов в соответствующем формате записывается следующим образом: 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 приведен выше, для 16 из соответствующих 17 маркеров, аллели которых выделены (предпоследнего маркера для 17-маркерного гаплотипа нет в 67-маркерной панели): **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.

Обычно для столь протяженных гаплотипов подсчитываются мутации только для первых 25 маркеров, их и так получается много. Кроме того, для первых 25 маркеров система наиболее отработана. Константа скорости мутации для 25-маркерных гаплотипов равна 0,0183 мутации на маркер на условное «поколение» (25 лет) [1]. Как видно, она более «медленная» по сравнению с таковой для 17-маркерных гаплотипов (0,0020 мутации на маркер на «поколение»). В первых 25 маркерах во всех 148 гаплотипах имеется 1037 мутаций, что обуславливает среднее значение $0,280 \pm 0,017$ мутации на маркер или 153 ± 16 поколений без поправки на возвратные мутации, или 180 ± 19 поколений с поправкой, т.е. 4500 ± 470 лет до общего предка. Для 17-маркерных гаплотипов это было $4475 \pm \pm 460$ лет. Такое практически абсолютное совпадение в данном случае, конечно, случайно, но в целом подобные совпадения в пределах погрешности с 95%-ной достоверностью скорее правило, чем исключение.

Попытаемся выйти за пределы 25-маркерных гаплотипов и посчитаем мутации для всех 148 гаплотипов в первых 37 маркерах. Их там 2023 мутации. Константа скорости для 37-маркерных гаплотипов равна 0,00243 на маркер на поколение [1] (как пояснено выше, это фактически не поколение, а промежуток времени (25 лет), т.е. «условное поколение»; истинное время жизни поколения — величина «плавающая» и зависящая от многих факторов: культурных особенностей народов, исторических обстоятельств и т.п.). Получаем $0,369 \pm 0,016$ мутации на маркер или 152 ± 16 поколений без поправки, или 179 ± 18 поколений с поправкой, т.е. $4475 \pm \pm 460$ лет до общего предка. Совпадение абсолютное. Наконец, во всех 67 маркерах было 2748 мутаций, что соответствует $2748/148/0,12 = 155 \pm \pm 16$ поколений без поправки или 183 ± 19 поколений с поправкой, т.е. 4575 ± 470 лет до общего предка (константа скорости мутации для 67-маркерных гаплотипов равна 0,12 мутации на гаплотип или 0,00179 мутации на маркер на поко-

ление). Совпадение опять почти абсолютное. Средний «возраст» общего предка для приведенных 148 гаплотипов 4500 ± 60 лет. Это, конечно, формальный расчет погрешности, но суть ясна.

Как видно, серии гаплотипов, особенно их несколько сотен, образуют «робастную» систему, устойчивую к статистическим вариациям (таблица). Как можно рассчитать с помощью таблицы, среднее время до общего предка гап-

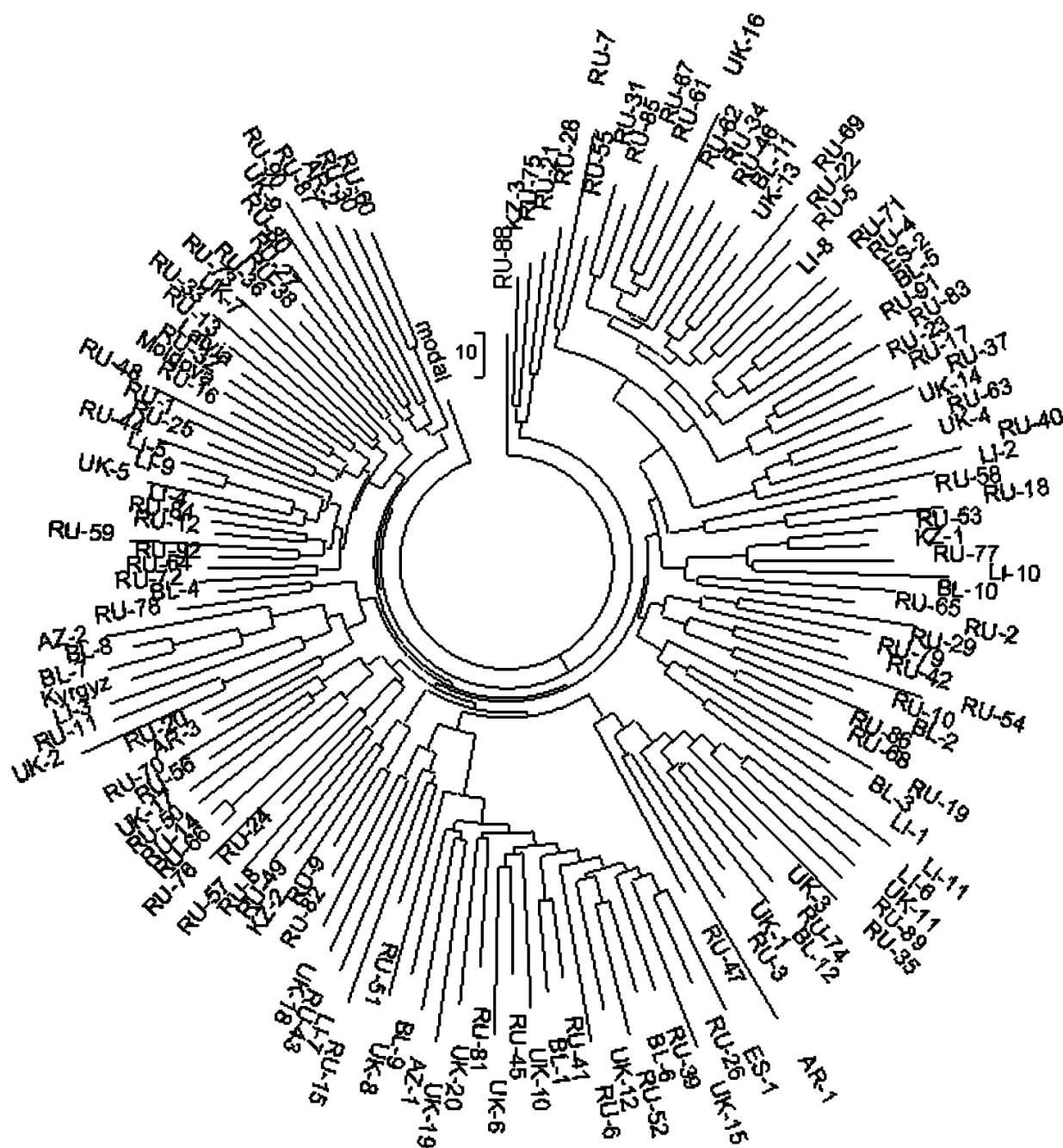


Рис. 3. Дерево из 148 67-маркерных гаплотипов гаплогруппы R1a1 постсоветского пространства. Гаплотипы взяты из базы данных YSearch. Индексы гаплотипов RU, UK, BL, LI, KZ, AR, ES, AZ соответствуют русским (92 чел), украинцам (20), белорусам (12), литовцам (11), казахам (3), армянам (3), эстонцам (2), 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

История определений времени до общего предка гаплогруппы R1a1 на постсоветском пространстве по разным сериям гаплотипов

Дата	Число гаплотипов	Общее число мутаций	Время до общего предка, годы	Источник
Июнь 2008	26	178	4400 ± 550	[5]
Ноябрь 2008	44	326	4825 ± 550	[6]
Январь 2009	58	423	4725 ± 520	[1]
Февраль 2009	255	1320	4475 ± 460	[4]
Март 2009	98	711	4700 ± 500	[1]
Июнь 2009	110	804	4750 ± 500	[1]
Ноябрь 2010	148	1037	4500 ± 470	
		2023	4475 ± 460	
		2748	4575 ± 470	

логруппы R1a1 на постсоветском пространстве, большинство из потомков которого в выборке — этнические русские и их «родственники» по роду, составляет 4600 ± 150 лет. Это заметно меньше той погрешности, которая оценивалась по числу мутаций и по предполагаемой погрешности в определении константы скорости мутаций в гаплотипах.

В процессе этих исследований было обнаружено, что гаплотипы большинства индийцев гаплогруппы R1a1, которых в Индии насчитывается не менее 100 млн., практически идентичны гаплотипам этнических русских той же гаплогруппы [5, 7]. 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 [8]. Базовый гаплотип, приведенный выше, одинаковый у русских, индийцев и иранцев рода R1a1, но общий предок на Русской равнине у этнических русских старше примерно на 500 лет [7]. Это влечет за собой решение интереснейшей исторической задачи, которой уже не менее 200 лет. Однако мы не будем на нем останавливаться, предоставив возможность сделать это самим читателям. Отметим только, что 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, измеримый 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 второго тысячелетия до н.э.

В качестве дополнительной иллюстрации этих исследований совместной истории предков на Русской равнине и в Индостане 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 16 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. Отметим, 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 индийским гаплотипам, чем они сами между собой. На самом деле эти различия практически равны друг другу и находятся в пределах погрешности.

Для примера, типичный базовый западноевропейский гаплотип (гаплогруппы R1b), который (и его производные) имеют ~60% западно- и центральноевропейцев и до 90% жителей Британских островов, имеет вид [9] 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, что не удивительно — их предков разделяют не менее 30 000 лет. В Индии и Иране гаплотипов гаплогруппы R1b очень мало, почти нет. Среди ариев 3500 лет назад предков современных западноевропейцев, похоже, не было. Отметим, что в высших индийских кастах в настоящее время имеется до 72% носителей гаплогруппы R1a1 [10].

Подобные же подходы к анализу гаплотипов этнических русских гаплогруппы I («западные славяне») и N1c («финноугры») позволили построить деревья гаплотипов, которые рассмотрены ниже. Анализ этих деревьев проведен в работах [11, 12]. В работе [3], в которой приведены гаплотипы, их анализ не проводился. Этот анализ вообще практически никем в литературе не проводился, не считая крайне приблизительных оценок в рамках подходов популяционной генетики, когда ставятся совершенно другие задачи и используется другая методология. Популяционная генетика обычно занимается сравнительным анализом популяций путем изучения частот аллелей и их изменения под влиянием эволюционных процессов, объясняя таким образом адаптацию и специализацию в популяциях, и в итоге формулируя закономерности и законы перехода от набора генотипов к серии фенотипов в популяции. Напротив, ДНК-генеалогия — это

по сути наука историческая, создаваемая на базе химической и биологической кинетики и секвенирования ДНК. В ней ярко и количественно выражена временная компонента. Иными словами, ДНК-генеалогия — это продукт слияния определения последовательностей определенных фрагментов ДНК (гаплотипов) и методов химической кинетики, количественно анализирующих динамику изменения этих последовательностей в популяциях. Генетики как таковой в ДНК-генеалогии нет, она рассматривает только негенные области ДНК.

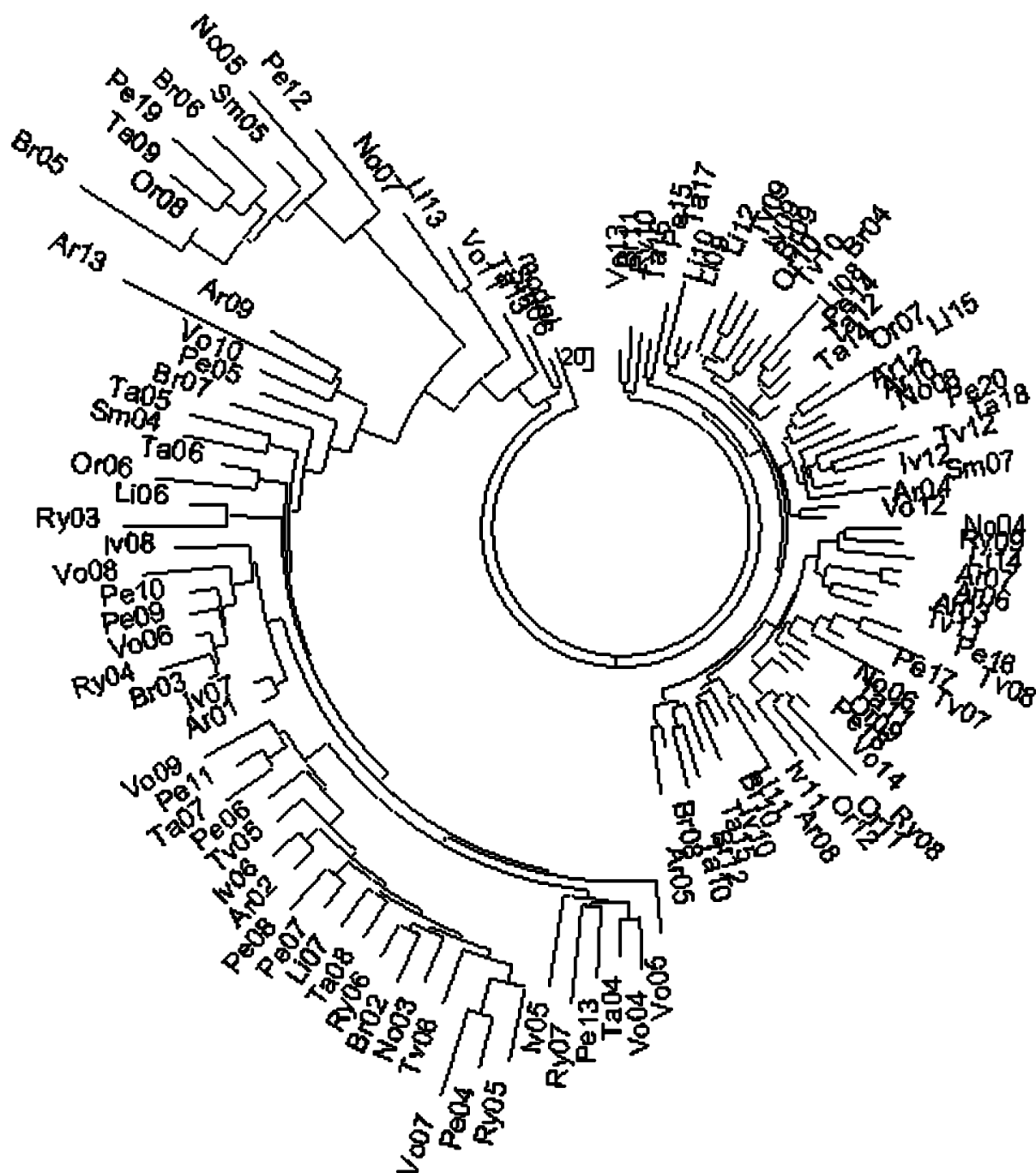
Фактически динамика изменения мутаций в гаплотипах аналогична динамике в системах параллельных и последовательных химических или биологических реакций, причем реакций обратимых (потому что мутации в ДНК от предкового гаплотипа могут происходить «в обе стороны»). Это образует довольно сложную систему, особенно в интервалах времен в тысячи и десятки тысяч лет. Для их анализа и нужны подходы химической кинетики, которые не применяются в популяционной генетике. Эти подходы включают в себя логарифмические зависимости убывания исходных, предковых гаплотипов во времени (происходящие по экспоненциальному закону в каждой ветви), учет накопления мутаций в гаплотипах, учет возвратных мутаций, которые накапливаются во времени, учет симметрии мутаций (увеличения или уменьшения числа аллелей) и т.д. [1, 13–15].

Отметим, что если в исходной цитированной работе [3] в общую гаплогруппу I были зачислены 117 гаплотипов, то дерево гаплотипов ясно показывает наличие совершенно разных ветвей, каждая со своей историей и со своими предками, жившими в разные исторические эпохи (рис. 4). Они действительно разошлись от одного рода I, причем, судя по картине мутаций, это произошло не менее 15 000 лет назад, но образовали свои роды, I1 и I2, которые в свою очередь дали свои ДНК-генеалогические ветви. Правая ветвь гаплотипов этнических русских принадлежит к гаплогруппе I2 с возрастом общего предка 3000 ± 380 лет, причем предок имел гаплотип 16 14 15 13 31 24 11 11 13 15 10 13 20 15 17 23 10 в формате, использованном выше для описания гаплогруппы R1a1 (так называемый формат Y-файлера). От базового гаплотипа R1a1 этнических русских, приведенного выше, он отличается на 14 мутаций, что на самом деле соответствует «латеральной» разнице во временах жизни общих предков в десятки тысяч лет.

Древняя ветвь гаплогруппы I2 этнических русских на рис. 4 (это наиболее «распушенные» и удаленные от ствола серии гаплотипов на ветви слева) имеет общего предка возрастом $10\,500 \pm$

1100 лет. В ней — 203 мутации на 20 гаплотипов, т.е. в среднем $0,597 \pm 0,084$ мутации на маркер (ср. с $0,304 \pm 0,017$ мутации на маркер для гаплогруппы R1a1). Ясно, что эта ветвь значительно более древняя, причем можно рассчитать, насколько более древняя.

Опять сравним результат, полученный с применением 17-маркерных гаплотипов, с результатами расчетов на примере более протяженных коммерческих гаплотипов той же гаплогруппы I2. На рис. 5 приведено дерево гаплотипов Восточной Европы. Как видно, конфигурация дерева



та же, что и у этнических русских (рис. 4), только развернута в другую сторону, что не имеет ни малейшего значения ни по сути, ни для анализа данных. Возраст молодой ветви 2650 ± 320 лет несколько меньше возраста этнических русских (3000 ± 380 лет), но совпадает с ним в пределах погрешности расчетов. Возраст древней ветви $10\,800 \pm 1200$ лет, что практически совпадает с возрастом ветви I2 этнических русских ($10\,500 \pm 1100$ лет), т.е. род I2 в Восточной Европе один и тот же, с одними общими предками, независимо от современных 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. Он точно такой же, как в 17-маркерном формате (выделены те маркеры из 17-маркерного формата, которые пересекаются с 25-маркерным форматом): **13 24 16 11 14 15 11 13 13 11 31 – 17 8 10 11 11 25 15 20 32 12 14 15 15.**

В Скандинавии общие предки данной гаплогруппы совершенно другие, хотя возраст ветвей тот же. Дерево гаплотипов гаплогруппы I2 Скандинавии приведено на рис. 6. 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

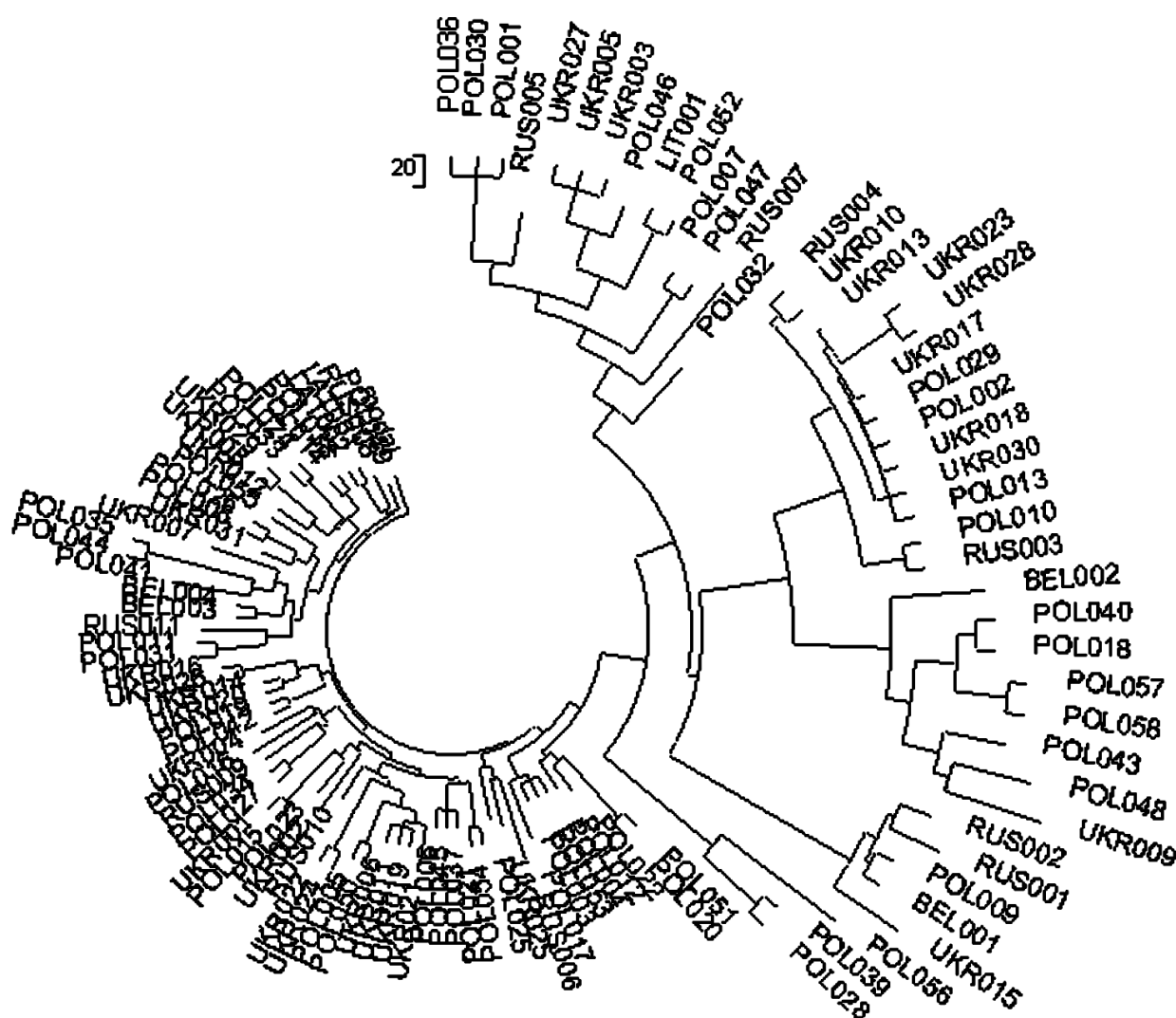


Рис. 5. Дерево 25-маркерных гаплотипов Восточной Европы гаплогруппы I2. Дерево состоит из 96 гаплотипов, построенно с помощью базы данных YSearch. Индексы гаплотипов соответствуют России, Украине, Белоруссии, Польше, Литве, Латвии (последняя — только в «молодой» ветви)

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 восточно-европейского (включая и этнических русских) на 26 мутаций на 25 маркерах (!), что разводит их общих предков на 29 000 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 20 000 лет назад, видимо, в Европе. Это – времена неандертальцев.

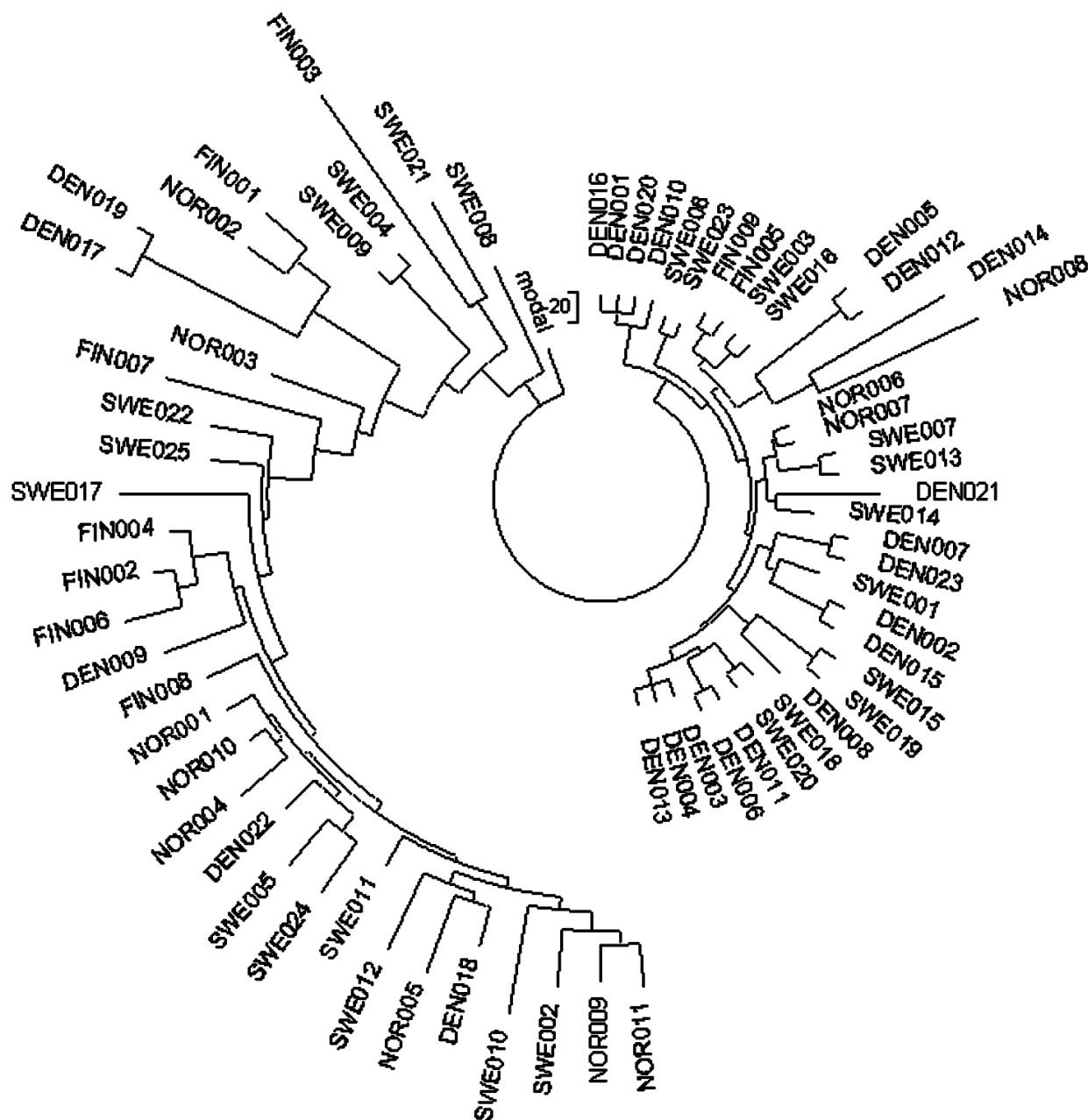


Рис. 6. Дерево 25-маркерных скандинавских гаплотипов гаплогруппы I2, построено с помощью базы данных YSearch. Оно состоит из 68 гаплотипов. Индексы гаплотипов соответствуют Швеции, Норвегии, Дании, Финляндии

Другим обрывком той же древней ветви является ветвь I2 возрастом примерно 10 000 лет ($10\,500 \pm 1100$ лет у этнических русских и $10\,800 \pm 1200$ в Восточной Европе в целом). В Скандинавии эта древняя ветвь (левая ветвь на рис. 6) тоже совершенно другая, чем у восточно-европейских 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 и 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 $0,509 \pm 0,065$ мутации на маркер, что соответствует 9575 ± 1140 лет до общего предка. Во второй, восточно-европейской ветви — $0,556 \pm 0,045$ мутации на маркер, что соответствует $10\,800 \pm 1200$ лет до общего предка. Между их базовыми гаплотипами — 13 мутаций, что помещает их общего предка примерно на 15 000 лет назад. Это все 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. 4 действительно оказывается самой «молодой» в сводной гаплогруппе I. Ее базовый гаплотип в формате Y-файлера 14 14 14 12 29 22 10 11 13 16 10 11 20 14 15 22 10, и общий предок ветви жил 3650 ± 800 лет назад. Практически этот же базовый гаплотип (в 25-маркерном формате, одинаковые аллели выделены, а неодинаковые отличаются всего на доли единицы) **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** был найден для западно-, центрально- и восточно-европейской серий гаплотипов и отдельно Германии [16] с «возрастом» общего предка соответственно 3425 ± 350 , 3425 ± 350 , 3225 ± 360 и 3225 ± 330 лет. При этом для анализа и расчетов использовались довольно большие серии гаплотипов (857, 284, 74 и 276 25-маркерных гаплотипов). Но есть небольшая серия гаплотипов, все европейские (Франция, Англия, Швейцария, Греция, Германия и Польша), которые разбегаются по мутациям настолько, что необходимо 21 400 лет для такого разбега. Это и есть минимальный возраст общего предка гаплогруппы I1 в Европе. Общий же предок гаплогрупп I1 и I2, т.е. общий предок гаплогруппы I, уходит вглубь более чем на 30 000 лет назад, а возможно и более чем на 40 000 лет назад.

Перейдем к этническим русским «финно-угорского» происхождения (гаплогруппа N1c). Надо подчеркнуть, что это такие же этнические русские, которые упоминались выше и отбирались для тестирования по тем же критериям, что и остальные. Это — славяне, как «западные», так и «восточные». По аналогии их можно назвать «северными славянами». Основная разница в

происхождении древних предков состоит в том, что древние носители гаплогруппы N1c мигрировали тысячелетия назад из Южно-Сибирского региона (видимо, с Алтая), и их язык был реконструирован как принадлежащий к алтайской языковой семье, и далее к уральской группе языков. Гипотеза об урало-алтайской семье языков существует с XVIII в., так что носители гаплогруппы N1c скорее уральцы, чем финноугры.

На рис. 7 приведено дерево гаплотипов этнических русских гаплогруппы N1c. Его анализ показал, что все гаплотипы происходят от одного общего предка с базовым гаплотипом (в формате 17-маркерных гаплотипов, как отмечено выше) 14 11 13 14 30 23 11 14 14 14 10 10 19 14 17 22 11. Данный гаплотип отличается на 18 мутаций от базового гаплотипа R1a1 и на 23 мутации от базового гаплотипа I2. Это свидетельствует о том, что общие предки перечисленных родов жили на десятки тысячелетий «врозь» как территориально, так и по времени, и о том, что филогенетически гаплогруппы N1c и R1a действительно ближе, чем N1c и I2.

Все 76 гаплотипов на рис. 7 содержат 317 мутаций от показанного базового гаплотипа [12], что свидетельствует в среднем об относительно небольшом числе ($0,245 \pm 0,028$) мутаций на маркер и соответствует 3525 ± 400 лет до общего предка этнических русских гаплогруппы N1c. Это — середина второго тысячелетия до н.э. В это время арии (гаплогруппа R1a1) уже уходили в Индостан и на Иранское плато. Впоследствии их называли «индоариями», «авестийскими ариями», «индоиранцами».

Попытаемся осмыслить то, что описано выше. На самом деле автор имеет в своем распоряжении сотни примеров ДНК-генеалогических деревьев, которые он строил и изучал в последние годы с помощью подходов биохимии, молекулярной биологии, химической и биологической кинетики. Работа в Гарвардском университете и биомедицинской компании (профессор и главный научный сотрудник) позволила проводить это исследование наряду с главной профессиональной задачей — созданием новых подходов к лечению раковых заболеваний и прочих острых воспалительных патологий [17, 18]. Настоящая статья только набрасывает контуры роли, которую ДНК-генеалогия может играть в понимании происхождения, динамики и миграции человеческих родов, племен, генеалогических линий, а также языков человечества — в пространстве и во времени. В отличие от антропологии, истории, археологии, которые оперируют памятниками прошлого, ДНК-генеалогия имеет дело с молекулами, извлекаемыми из нашего организма. В результате серии отработанных

ных физико-химических операций, доступных грамотному технику-лаборанту и почти полностью автоматизированных, исследователь получает сведения о мутациях, накопившихся в определенных 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 времен 50 000–80 000 лет назад (на самом деле вплоть до любых времен жизни

предков человека, до сотен тысяч и миллионов лет назад, но наука о таких временах пока имеет слишком малую базу данных) и на любой территории.

Помимо этого любая человеческая популяция в определенные 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 уменьшается до

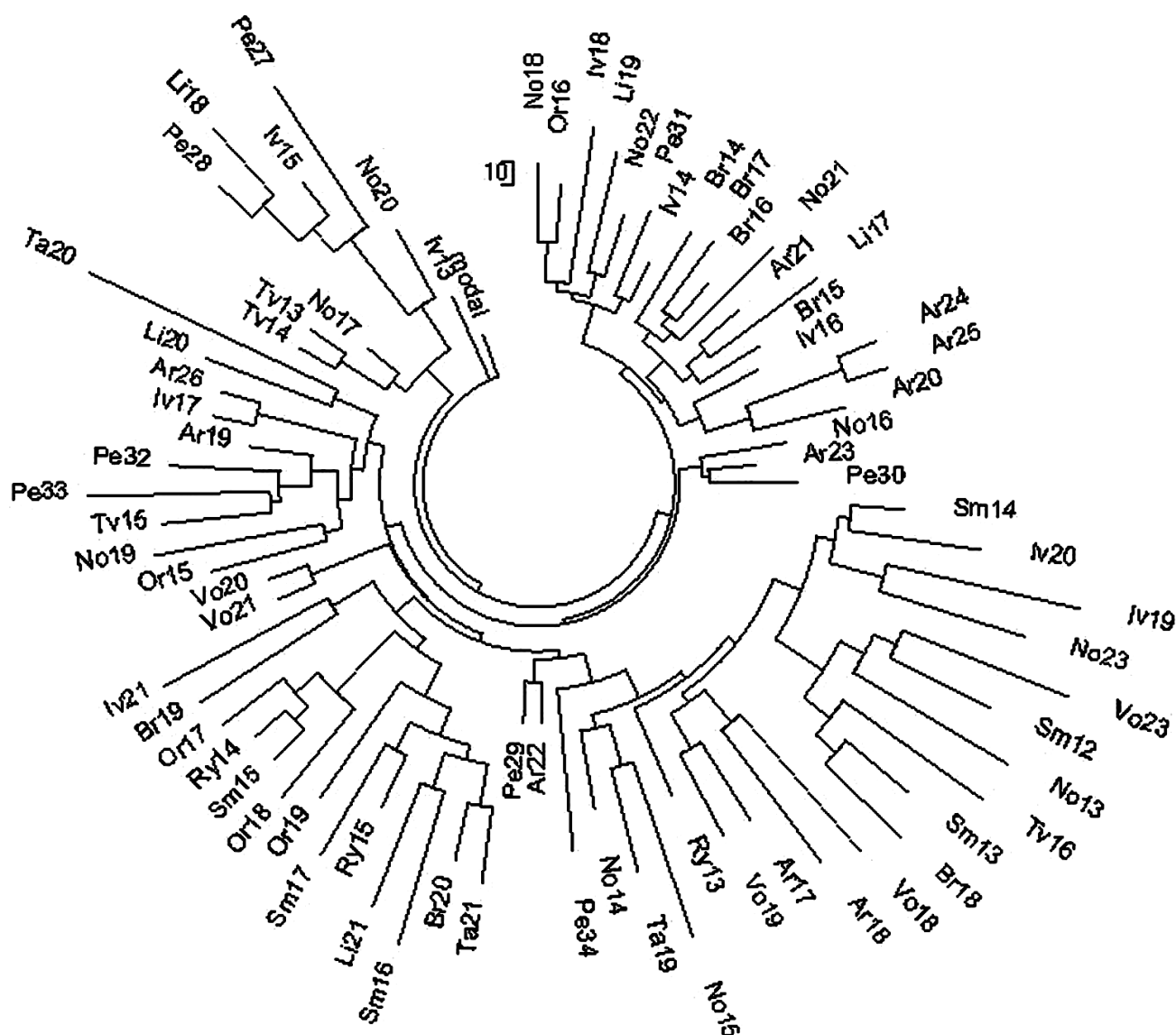


Рис. 7. Дерево из 76 17-маркерных гаплотипов этнических русских гаплогруппы N1c по 12 областям Российской Федерации, построено по данным работы [3]. Индексы при гаплотипах соответствуют сокращенным названиям 12 областей Российской Федерации (см. подпись к рис. 1)

нескольких или до одного человека, потомство которого в итоге выживает и увеличивается в числе. В таких случаях этот человек и оказывается «общим предком» выжившей популяции. Методы расчета 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.

В понятиях ДНК-генеалогии «до одного человека» скорее означает «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 пошел «разбег мутаций» у потомков.

Методы ДНК-генеалогии позволяют узнать, когда жил общий предок и, значит, когда имелось «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 ~10% при 95%-ной достоверности полученных абсолютных значений времени. Так, если общий предок данной популяции жил 5000 лет назад, то при наличии (путем тестирования) сотни гаплотипов, т.е. определенных фрагментов Y-хромосомы современников, потомков данного общего предка, время жизни общего предка определится с точностью 5000 ± 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 надежностью 95%.

Такая точность вызвана тем, что в сотне 25-маркерных гаплотипов (т.е. когда в каждом гаплотипе имеется по 25 нуклеотидных маркеров и каждый мутирует с определенной средней скоростью на протяжении этих тысяч лет) в распоряжении исследователя есть 2500 «экспериментальных точек», в которых на протяжении 5000 ± 530 лет произойдет, как исследователь знает, ~770 мутаций. Это значение жестко связано со средней скоростью мутаций, установленной и калиброванной [1]. Если мутаций в подобной серии из 100 гаплотипов меньше 770, то общий предок жил <5000 лет назад, и опять можно вполне надежно установить, когда он

жил. Например, при рассмотрении 750 19-маркерных гаплотипов басков (Пиренеи), рода (гаплогруппы) R1b1b2, оказалось, что все они содержат 2796 мутаций от гаплотипа, равноудаленного от всех (это и есть предковый гаплотип, так как мутации происходят в подавляющем большинстве совершенно неупорядоченно, со степенью симметрии 0,5 в любую сторону, «вверх» или «вниз» по значению аллелей). Это означает, что общий предок всех 750 басков жил 3625 ± 370 лет назад. На самом деле носители той же гаплогруппы прибыли на Пиренеи ~4800 лет назад и пронесли в континентальную Европу археологическую культуру колоколовидных кубков. По-видимому, пиренейские баски прошли «бутылочное горлышко» популяции, поскольку общий предок их современной популяции на тысячу лет «моложе» времен прибытия носителей гаплогруппы R1b1b2 на Пиренеи.

Можно вообще не обращать внимание на число мутаций и просто посчитать, сколько среди этих 750 гаплотипов есть «равноудаленных», причем идентичных друг другу. Это и есть базовые, или предковые, гаплотипы, и их число задается законами химической кинетики, т.е. физической химии. Примерно такими же законами задается скорость радиоактивного распада, хотя после превращения стабильный изотоп не может стать радиоактивным нуклидом; эти процессы необратимы. Мутация же может вернуться обратно, потому что мутации, как мы отмечали, полностью не упорядочены. Иначе говоря, мутированному гаплотипу все равно в какую сторону мутировать в следующий раз, мутации обратимы. Это усложняет расчеты по сравнению с расчетами простых процессов радиоактивного распада, но проблема с успехом решается.

Среди отмеченных 750 гаплотипов басков 16 оказались базовыми, т.е. за прошедшие 3625 ± 370 лет в полном соответствии с теорией вероятности 16 гаплотипов из 750 так и не успели мутировать, они находятся «на хвосте» кривой вероятностного распределения мутаций. Если считать по законам радиоактивного распада (как говорят физико-химики, по законам кинетики первого порядка), то из 750 гаплотипов 16 останутся неизменными за $\ln(750/16)/0,0285 = 135$ поколений, прошедших со времени жизни общего предка, т.е. за 3375 лет (продолжительность жизни условного «поколения» в 25 лет заложена в константу скорости мутации 0,0285 мутации на гаплотип на поколение, как пояснялось выше). Но при этом расчете не учитывается обратимость мутаций. Учет этой обратимости опять в строгом соответствии с теорией вероятности сдвигает число поколений со 135 до

156, т.е. до 3900 лет до общего предка. Для введения подобных поправок есть детальные таблицы расчетов [1, 13, 14]. Как видно, 3900 лет попадает в интервал 95%-ной надежности расчетов 3625 ± 370 лет до общего предка.

Очевидно, что считать мутации или считать базовые гаплотипы (если от времени жизни общего предка прошло не так много лет и базовых гаплотипов сохранилось достаточно много для проведения расчетов) не имеет большого значения или большой разницы. Значения получаются примерно одинаковые, если расчет ведется правильно и используются надежные константы скоростей мутации, подтвержденные экспериментально, т.е. с учетом известных генеалогий или известных исторических событий. В недавней работе [1] приведены значения констант скоростей мутаций для 22 типичных гаплотипов, применяемых в ДНК-генеалогии, от 5-маркерного до 67-маркерного. В последнее время эта таблица расширена уже до 32 типичных гаплотипов.

Чтобы показать, насколько массивными бывают расчеты и насколько точными получаются данные, приведем еще один пример — выборку из 857 25-маркерных гаплотипов Англии (род I1), т.е. 21 425 «экспериментальных точек» [1, 7]. Эти гаплотипы содержали 4868 мутаций от базового (предкового) гаплотипа. Название «базовый» принято потому, что во многих случаях нет полной уверенности, особенно с очень древними гаплотипами, что «равноудаленный» — это непременно предковый гаплотип. Это может быть хорошее приближение, но не более того. Назвать его «предковым» было бы натяжкой, поэтому называли «базовым».

Такое число мутаций позволяет определять среднее число мутаций на маркер (а именно это среднее число фактически используется для расчета времени жизни общего предка) с точностью 2–3% с 95%-ной надежностью. Это потому, что для несимметричных мутаций точность расчетов (среднеквадратичное отклонение) среднего значения мутаций на маркер задается обратной величиной квадратного корня из числа мутаций, а для симметричных отличается лишь немногим [1]. Для 4868 мутаций это отклонение с 95%-ной надежностью составляет 2,87%, но поскольку точность средних скоростей мутаций в гаплотипах обычно не ниже 5%, то для 95%-ной надежности расчетов времени жизни общего предка предельной точностью может быть только 5%, т.е. 10%-ный интервал.

Поэтому для английских гаплотипов определено, что общий предок всех 857 человек в данной выборке жил 3425 ± 350 лет назад (см. выше, обсуждение гаплотипов гаплогруппы I). Это —

относительно недавнее время, относительно молодые европейские популяции, хотя для некоторых европейских родов общие предки по данным ДНК-генеалогии жили 16 000–23 000 лет назад. Для ряда азиатских родов общий предок рода жил 16 000–21 000, для африканских родов — 28 000–37 000 лет назад. Здесь имеются в виду предки, потомки которых выжили до настоящего времени.

Эти примеры приведены для того, чтобы читатель осознал, 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 расчетный базис, платформу, и временные расчеты проводятся теперь с достаточной надежностью. С использованием геномов многих популяций людей, а также геномов шимпанзе было экспериментально показано, что мутации представляют собой действительно «молекулярные часы»; скорость мутаций в ДНК неизменна на протяжении по меньшей мере последних 2 млн. лет [19].

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 (т.е. менять число аллелей) с одинаковой вероятностью. На этих системах было показано, что «двойные» или «многошаговые» мутации в гаплотипах (маркерах) происходят редко, на уровне единиц процентов, и практически не влияют на расчетные значения времен до общих предков популяций или серий гаплотипов [20].

Сложнее с географической привязкой, с выявлением того, где именно жил общий предок, так как время его жизни (т.е. сколько поколений или лет назад он жил) на географию не указывает. Для выяснения того, где, на каких территориях жили общие предки популяции, приходится привлекать независимые данные археологии, антропологии, лингвистики, понимая частую условность этих сведений. «Условность» потому, что неизвестно, дожили ли гаплотипы тех ископаемых людей до наших времен.

Именно поэтому союз антропологии, археологии, лингвистики с ДНК-генеалогией так важен. Как уже отмечалось выше, ДНК-генеалогия предоставляет в их распоряжение жесткую привязку в виде «метки» рода, определенную и однозначно определяемую мутацию, снп, в Y-хромосоме ДНК, которая (мутация) всегда сопровождает каждого члена рода. Эта мутация не ассимилируется в популяциях, как ассимили-

руются 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, понять, останки представителей каких родов находятся в археологических раскопах и как археологические культуры связаны одна с другой (не только через материальные, культуuroобразующие носители, но и через людей, через конкретные роды), 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 (50 000–45 000 лет назад) действительно дожили до настоящего времени. Но это может быть верифицировано только путем анализа ДНК ископаемых останков. Технически это можно сделать путем определения нуклеотидной последовательности фрагментов костной ДНК, глубоко упрятанной в костях и частично пережившей тысячелетия. Эти исследования пока немногочисленны, фактически единичны, очень трудны и очень дорогостоящи.

Следует еще раз подчеркнуть, что ДНК-генеалогия наших современников «докапывается» только до «бутылочных горлышек» популяций, родов, племен, 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 оказали влияние на состав современных популяций. Трудно представить себе, насколько чума середины XIV в., которая выкосила четверть европейцев, терминировала генеалогические линии, гаплотипы и, возможно, целые гаплогруппы, роды. Много генеалогических линий начинаются именно в середине XIV в. Это — выжившие люди, которые повели линию популяции сначала, ставшие «общими предками» многочисленных групп наших современников.

Геноцид — худший враг ДНК-генеалогии, не говоря о самих носителях гаплотипов. Сколько генеалогических линий прервали резня в Армении во втором десятилетии XX в., геноцид евреев, цыган, других национальных, религиозных групп и популяций — об этом можно только догадываться. Сколько генеалогических линий прервалось в мировых войнах XX в., 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 еще миллион был угнан в рабство. Можно только гадать, как это изменило ландшафт гаплотипов и гаплогрупп в Европе.

Но «бутылочные горлышки» обуславливаются еще и переездом, переходом, миграцией 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 каким был тысячелетия назад, таким и остался. Так, например, общий предок англичан в Англии и в США один и тот же, как и общий предок (но уже другой) русских, восточных славян в США, например. Это совсем не означает, что восточные славяне жили так же 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 IX в. н.э., а общий предок исландцев практически по всем родам — тысячелетия назад, как и в континентальной Европе.

Общий предок гаплогруппы R1b1 по расчетам ДНК-генеалогии (по мутациям в гаплотипах) жил в Центральной (Средней) Азии 16 000 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 дает обоснованное представление о временах и направлении миграций рода R1b1, более того оно позволяет связать этот род с определенными археологическими культурами (иногда с серией культур, горизонтом, культурно-исторической общностью).

Отсюда и новый термин — «молекулярная история», т.е. создание исторических реконструкций исходя из молекулярных характеристик ДНК потомков, а порой и (ископаемых) предков. Поскольку далекие 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 биохимия повернулась своей неожиданной и непредсказуемой ранее гранью к наукам гуманитарным, историческим, лингвистическим

ким (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 общих предков. Анализ гаплотипов разных популяций показывает, что все люди в конечном итоге — родственники. Более того, сейчас уже можно довольно надежно рассчитывать степень этого родства, и биохимия в этом играет важную роль.

В заключение следует подчеркнуть, что автор совсем не приписывает себе роль первопроедца в новой области науки. Скромная задача автора состояла в привнесении количественных подходов в анализ картины мутаций в гаплотипах, в количественном, по возможности, анализе деревьев гаплотипов, расчете хронологических показателей исторических событий по мутациям в гаплотипах их участников. Естественно, в задачи данной статьи не входит исчерпывающий обзор истории создания и развития ДНК-генеалогии, хотя справедливости ради следует отметить, что ДНК-генеалогия вырастает из дисциплины «популяционная генетика», которая была заложена трудами многих специалистов. В приложении популяционной генетики к ДНК человека в 1990-х гг. большую роль сыгра-

ли (хотя в данном списке, безусловно, отражаются личные предпочтения автора) Л. Кавалли-Сфорца, М. Фельдман, Д. Голдстин, М. Хаммер, М. Джоблинг, М. Кейзер, П. Де Ниф, А. Небел, М. Ней, А. Оппенхейм, О. Семино, М. Стоункинг, М. Томас, П. Андерхилл, Б. Уолш, Р. Уэллс, Л. Животовский, Т. Карафет и многие другие, которых здесь отметить просто невозможно (см., например, работы [21–36]).

Этими и десятками других исследователей были проведены тщательные отнесения снипов к принципиальным родам человечества и их подчиненным родам, а также к отдельным генеалогическим линиям рода, имеющим свои мутационные метки в ДНК, называемым субкладами. Снипы тщательно идентифицируют и выбирают для классификации, чтобы роды (в терминах ДНК-генеалогии) были максимально четко дифференцированы. Можно упомянуть, что всего на Земле насчитывают 20 родов, именуемых по буквам латинского алфавита (от А до Т), хотя снипов, применяемых в ДНК-генеалогии, известно уже несколько сотен, и их количество приближается к тысяче. Всего же снипов миллионы. Но так как объем данной статьи ограничен, на этом придется остановиться.

СПИСОК ЛИТЕРАТУРЫ

1. Klyosov, A.A. (2009) *J. Genetic Genealogy*, **5**, 186–216.
2. Klyosov, A.A. (2009) *Human Genetics*, **126**, 719–724.
3. Roewer, L., Willuweit, S., Kruger, C., Nagy, M., Rychkov, S., Morozowa, I., Naumova, O., Schneider, Y., Zhukova, O., Stoneking, M., and Nasidze, I. (2008) *Int. Legal Medicine*, **122**, 219–223.
4. Клёсов А.А. (2009) *Вестн. Российской академии ДНК-генеалогии*, **2**, 232–251.
5. Клёсов А.А. (2008) *Вестн. Российской академии ДНК-генеалогии*, **1**, 400–477.
6. Клёсов А.А. (2008) *Вестн. Российской академии ДНК-генеалогии*, **1**, 947–957.
7. Klyosov, A.A. (2009) *J. Genetic Genealogy*, **5**, 217–256.
8. Клёсов А.А. (2009) *Вестн. Российской академии ДНК-генеалогии*, **2**, 1217–1229.
9. Клёсов, А.А. (2010) *Вестн. Российской академии ДНК-генеалогии*, **3**, 249–299.
10. Sharma, S., Rai, E., Sharma, P., Jena, M., Singh, S., Darvishi, K., Bhat, A.K., Bhanwer, A.J.S., Tiwari, P.K., and Bamezai, R.N.K. (2009) *J. Human Genetics*, **54**, 47–55.
11. Клёсов А.А. (2009) *Вестн. Российской академии ДНК-генеалогии*, **2**, 801–815.
12. Клёсов А.А. (2009) *Вестн. Российской академии ДНК-генеалогии*, **2**, 370–389.
13. Клёсов А.А. (2008) *Вестн. Российской академии ДНК-генеалогии*, **1**, 252–348.
14. Клёсов А.А. (2008) *Вестн. Российской академии ДНК-генеалогии*, **1**, 812–835.
15. Адамов Д.С., Клёсов А.А. (2009) *Вестн. Российской академии ДНК-генеалогии*, **2**, 422–442.
16. Клёсов А.А. (2010) *Вестн. Российской академии ДНК-генеалогии*, **3**, 96–158.
17. Klyosov, A.A., Witczak, Z.J., and Platt, D. (2006) *Carbohydrate Drug Design*, Cambridge University Press, Cambridge, 323 pp.
18. Klyosov, A.A., Witczak, Z.J., and Platt, D. (2008) *Galectins*, John Wiley and Sons, 279 pp.
19. Sun, J.X., Millikin, J.C., Patterson, N., and Reich, D.E. (2009) *Mol. Biol. Evol.*, **26**, 1017–1027.
20. Клёсов А.А. (2010) *Вестн. Российской академии ДНК-генеалогии*, **3**, 1853–1860.
21. Cordaux, R., Bentley, G., Aunger, R., Sirajuddin, S.M., and Stoneking, M.J. (2004) *Forensic Sci.*, **49**, 1–2.
22. Goldstein, D.B., Linares, A.R., Cavalli-Sforza, L.L., and Feldman, M.W. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 6723–6727.
23. Hammer, M.F., Redd, A.J., Wood, E.T., Bonner, M.R., Jarjanazi, H., Karafet, T., Santachiara-Benerecetti, S., Oppenheim, A., Jobling, M.A., Jenkins, T., Ostrer, H., and Bonne-Tamir, B. (2000) *Proc. Natl. Acad. Sci. USA*, **97**, 6769–6774.
24. Heyer, E., Puymirat, J., Dieltjes, P., Bakker, E., and de Knijff, P. (1997) *Human Mol. Genetics*, **6**, 799–803.
25. Jobling, M.A., and Tyler-Smith, C. (1995) *Trends Genet.*, **11**, 449–456.
26. Karafet, T.M., Zegura, S.L., Posukh, O., Osipova, L., Bergen, A., Long, J., Goldman, D., Klitz, W., Harihara, S.,

- de Knijff, P., Wiebe, V., Griffiths, R.C., Templeton, A.R., and Hammer, M.F. (1999) *Am. J. Hum. Genet.*, **64**, 817–831.
27. Kayser, M., Roewer, L., Hedman, M., Henke, L., Hemke, J., Brauer, S., Kruger, C., Krawczak, M., Nagy, M., Dobosz, T., Szibor, R., de Knijff, P., Stoneking, M., and Sajantila, A. (2000) *Am. J. Hum. Genet.*, **66**, 1580–1588.
28. Nebel, A., Filon, D., Weiss, D.A., Weale, M., Faerman, M., Oppenheim, A., and Thomas, M. (2000) *Hum. Genet.*, **107**, 630–641.
29. Nebel, A., Filon, D., Brinkmann, B., Majumder, P.P., Faerman, M., and Oppenheim, A. (2001) *Am. J. Hum. Genet.*, **69**, 1095–1112.
30. Nei, M. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 6720–6722.
31. Semino, O., Passarino, G., Oefner, P.J., Lin, A.A., Arbuzova, S., Beckman, L.E., De Benedictis, G., Francalacci, P., Kouvatsi, A., Limborska, S., Marcikiae, M., Mika, A., Mika, B., Primorac, D., Santachiara-Benerecetti, A.S., Cavalli-Sforza, L.L., Underhill, P.A. (2000) *Science*, **290**, 1155–1159.
32. Takezaki, N., and Nei, M. (1996) *Genetics*, **144**, 389–399.
33. Underhill, P.A., Shen, P., Lin, A.A., Jin, L., Passarino, G., Yang, W.H., Kauffman, E., Bonne-Tamir, B., Bertranpetit, J., Francalacci, P., Ibrahim, M., Jenkins, T., Kidd, J.R., Mehdi, S.Q., Seielstad, M.T., Wells, R.S., Piazza, A., Davis, R.W., Feldman, M.W., Cavalli-Sforza, L.L., and Oefner, P.J. (2000) *Nature Genetics*, **26**, 358–361.
34. Walsh, B. (2001) *Genetics*, **158**, 897–912.
35. Wells, R.S., Yuldasheva, N., Ruzibakiev, R., Underhill, P.A., Evseeva, I., Blue-Smith, L., Jin, L., Su, B., Pitchappan, R., Shanmugalaksmi, S., Balakrishnan, K., Read, M., Pearson, N.M., Zerjal, T., Webster, M.T., Zholoshvili, I., Jamarjashvili, E., Gambarov, S., Nikbin, B., Dostiev, A., Aknazarov, O., Zallous, P., Tsoy, I., Kitaev, M., Mirrakhimov, M., Chariev, A., and Bodmer, W.F. (2001) *Proc. Natl. Acad. Sci. USA*, **98**, 10244–10249.
36. Zhivotovsky, L.A., and Feldman, M.W. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 11549–11552.

BIOLOGICAL CHEMISTRY AS A FOUNDATION OF DNA GENEALOGY, AND EMERGING OF “MOLECULAR HISTORY”

A. A. Klyosov

*MIR International, Inc., 36 Walsh Road, Newton,
Massachusetts 02459, USA; E-mail: aklyosov@comcast.net*

Received November 17, 2010
Revision received December 17, 2010

This paper presents the basis of DNA genealogy, a new field of science that is currently emerging as an unusual blend of biochemistry, history, linguistics, and chemical kinetics. The methodology of the new approach comprises chemical (biological) kinetics applied to a pattern of mutations in non-recombinant fragments of DNA (Y-chromosome and mtDNA; the latter is not considered in this overview). This analysis aims at the translation of the DNA mutation patterns into time spans to the most recent common ancestors of the given population, tribe, to dating of ancient migration routes. To illustrate this approach, time spans to the common ancestors are calculated for the ethnic Russians, that is Eastern Slavs (the R1a1 tribe), Western Slavs (I1 and I2 tribes), and Northern, or Uralic Slavs (the N1c tribe), which were found to live around 4600 years before present (R1a1), 3650 ybp (I1), 3000 and 10500 ybp (I2, two principal DNA-lineages), and 3525 ybp (N1c) (the confidence intervals are given in the main text). The data are compared with the respective dates for the nearest common ancestor of the R1a1 “Indo-European” population in India, who lived 4050 years before present, and descendants of whom present the majority of the upper castes in India today (up to 72%). Furthermore, it was shown that haplotypes of the ethnic Russians of the R1a1 haplogroup (up to 62% of the population in the Russian Federation) and those of the R1a1 Indians (more than 100 million today) are practically identical to each other, up to 67 marker haplotypes. This essentially solves a 200-year old mystery of who were the Aryans who arrived in India around 3500 years before present. Haplotypes and time spans to the ancient common ancestors were also compared for the ethnic Russians of haplogroups I1 and I2, on one hand, and the respective I1 and I2 populations in Eastern and Western Europe and Scandinavia, on the other. It is suggested that the approach described in this overview lays the foundation for “molecular history”, in which the principal tool is the high-tech analysis of the DNA molecules of both our contemporaries and excavated ancient DNA samples, along with their biological kinetics.

Key words: DNA genealogy, haplotypes, haplogroups, mutations, Y chromosome