

## СКОРОСТЬ ГЕРМИНАЛЬНЫХ МУТАЦИЙ ЧЕЛОВЕКА – ВАРИАБЕЛЬНЫЙ ФАКТОР ЭВОЛЮЦИИ И БОЛЕЗНЕЙ

© 2019 г. Н. Я. Успенская<sup>1</sup>, \*, С. Б. Акопов<sup>1</sup>, Е. В. Снежков<sup>1</sup>, Е. Д. Свердлов<sup>1</sup>

<sup>1</sup>Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова  
Российской академии наук, Москва, 117997 Россия

\*e-mail: natusp@gmail.com

Поступила в редакцию 17.05.2018 г.

После доработки 22.05.2018 г.

Принята к публикации 04.07.2018 г.

Возникновение генетических болезней и эволюционные процессы связаны с потоком генетической информации от одного поколения к другому, при котором генетическая информация, переносимая гаметой, может меняться путем возникновения *de novo* герминальных мутаций. Скорость, с которой происходит возникновение герминальных мутаций, определяет темпы эволюции и частоту возникновения наследуемых патологий. Несмотря на большую теоретическую и практическую значимость, проблема установления скоростей мутаций и их зависимости от различных факторов остается недостаточно изученной, и значения величин скорости мутаций, получаемые разными методами, заметно различаются. В обзоре обсуждаются разные способы оценки скорости этих мутаций и дается попытка объяснения возможных причин расхождения получаемых данных. При этом рассмотрены три уровня формирования мутаций: 1) мутации, которые формируются в процессе развития данного индивидуума при гаметогенезе (базовые мутации); 2) мутации, передаваемые потомству и определяющие различия в геномах последовательных поколений (родители и их дети), включают базовые мутации и возможные изменения в результате сложных процессов переноса спермы к ооциту, оплодотворения и последующих событий, которые приводят только к одному жизнеспособному потомку из сотен миллионов сперматозоидов и ооцитов; 3) мутации, которые образуются на уровне 2, фиксируются в эволюции и определяют эволюционные процессы и различия между геномами, в частности, гоминоидов, гоминоидов и гоминоидов.

**Ключевые слова:** мутации *de novo*, секвенирование геномов, дивергенция, датирование, замещения.

**DOI:** 10.1134/S001667581905014X

### ГЕРМИНАЛЬНЫЕ *de novo* МУТАЦИИ

Возникновение генетических болезней и эволюционные процессы связаны с потоком генетической информации от одного поколения к другому, осуществляемой половыми клетками – сперматозоидами и ооцитами. Образование гамет, начиная от оплодотворения яйцеклетки с образованием зиготы до формирования зрелых гамет, происходит путем развития зародышевой линии клеток, включающего многочисленные акты репликации и рекомбинации. При этом генетическая информация может меняться путем возникновения *de novo* разнообразных мутаций: от точечных, связанных с ошибками в репликации ДНК, до крупных делеций и инсерций, которые могут возникать в процессе рекомбинации. Мутации могут происходить также под влиянием экзогенных факторов. Скорость, с которой происходит возникновение мутаций, называется скоростью мутаций зародышевой линии и имеет центральное значение для понимания закономерностей появ-

ления геномных различий между людьми и популяциями и для эволюционной генетики. Эта скорость определяет темпы эволюции и частоту возникновения наследуемых патологий.

Первые попытки определения скорости мутаций у людей делались еще до того, как была установлена структура ДНК, на основе частоты патологических, фенотипически идентифицируемых мутаций (см. [1]). Они вынужденно были ограничены мутациями, вызывающими фенотипические различия, например у доминантных локусов болезни. К таким оценкам относятся классические работы Джона Холдейна (<http://vikent.ru/author/1052/Haldane>). Последующие оценки были основаны на филогенетических сопоставлениях видов с временами дивергенции, полученными на основании палеонтологического и археологического датирования ископаемых остатков. В последнее время разработки в технологии секвенирования генома позволили оценить скорость мутаций, основываясь на подсчете мутаций *de novo*, путем сравне-

ния последовательностей геномов близких родственников — трио родителей и их детей или более крупных родословных.

Несмотря на большую теоретическую и практическую значимость, проблема установления скоростей мутаций и их зависимости от различных факторов остается недостаточно изученной и значения величин скорости мутаций, получаемые разными методами, заметно различаются. Так, оценки скоростей с использованием прямого секвенирования геномов родственных трио человека дают значение  $0.5 \times 10^{-9}$  замещений на пн в год, что в 2 раза ниже оценок, получаемых филогенетическими методами.

Известный исследователь нестабильности генома М. Линч [2] писал: “Несмотря на большое значение для здоровья и фенотипической эволюции, факторы, влияющие на скорости и спектр мутаций, спонтанно возникающих в геноме человека, остаются малоизученными, и цифры, приводимые разными авторами, существенно различаются. Как часто возникают зародышевые и соматические мутации и в какой степени это различается между полами? Какова относительная частота различных форм мутаций, например вставки, дубликации и делеции, особенно среди мутаций, имеющих фенотипические эффекты? Насколько мутационный спектр у людей сопоставим с другими видами? И, что самое важное, каковы последствия мутаций для долгосрочного генетического благополучия наших видов?”

В настоящем обзоре речь пойдет о скоростях герминальных мутаций, т.е. мутациях, возникающих в зародышевой линии клеток, начиная от зиготы и завершая зрелыми гаметами [3–5]. Мы обсудим разные способы оценки скорости этих мутаций и попытаемся объяснить причины расхождения получаемых данных. Проблему соматических мутаций мы детально обсуждали в наших недавних обзорах [6, 7].

### ТРИ РАЗНЫХ УРОВНЯ ОПРЕДЕЛЕНИЯ СКОРОСТИ ГЕРМИНАЛЬНЫХ МУТАЦИЙ

В данном обзоре мы будем пользоваться следующими определениями используемых понятий. Скорость мутаций — это число мутаций, приобретаемых геномом за единицу времени. Эта единица может быть абсолютной или условной. Например, год, миллион лет или клеточное деление, или генерация.

Важно отметить, что эти единицы времени обычно различаются для соматических и герминальных клеток. Для соматических клеток это, как правило, число мутаций на пару оснований за клеточное деление, а для герминальных клеток чаще используют число мутаций на пару основа-

ний или на целый геном *за генерацию*. В популяционной биологии и эволюционной биологии время генерации — это среднее время между двумя последовательными поколениями в популяции. В человеческих популяциях время генерации обычно принимают равным 22–32 годам. Время генерации может также определяться как время, необходимое для рождения человека, достижения половой зрелости и воспроизведения. В эволюции используют и такую универсальную единицу как число мутаций на пару оснований в миллион лет.

Под спектром мутаций понимают соотношение мутаций различного типа: замещения оснований, делеции, вставки, рекомбинации, изменения числа копий повторяющегося элемента и т.д.

В соответствии с методом, используемым при определении скорости, в литературе оперируют и сравнивают друг с другом три разных понятия зародышевых мутаций:

1. Мутации, которые, как и в соматических клетках, формируются в процессе развития данного индивидуума при гаметогенезе и определяют различия (мозаицизм) между его отдельными гаметами (сперматозоидами или ооцитами). Мы определим этот тип мутаций как базовый.

2. Мутации, передаваемые потомству и определяющие различия в геномах последовательных поколений (родители и их дети). Они являются результатом не только базовых мутаций, определенных в п. 1, но также и дополнительных сложных процессов переноса спермы к ооциту, оплодотворения и последующих событий, которые приводят только к одному жизнеспособному потомку из сотен миллионов сперматозоидов и ооцитов. При этом успешная гамета может содержать селектированные по каким-либо признакам или просто стохастические выборки мутаций. Некоторые из них могут вызывать генетические заболевания.

3. Мутации, которые передаются потомкам, как в п. 2, но еще и *фиксируются в эволюции* и определяют эволюционные процессы и различия между геномами, в частности, гоминоидов, гоминидов и гоминин.

В первых двух случаях скорости мутаций определяются в данном поколении или в двух-трех последующих поколениях, тогда как мутации типа 3 зависят от процессов, происходящих на протяжении множества поколений. Вполне возможно, что скорости и спектры мутаций менялись от поколения к поколению и те значения, которые мы получаем на основании филогенетического подхода, представляя собой некую, усредненную по скоростям разных поколений, величину, которая *обязана* отличаться от первых двух. Интуитивно, скорости этих мутаций не могут быть больше, а скорее всего должны быть меньше, чем те, кото-

рые определяются в первых двух случаях. Однако, как мы увидим дальше, реально экспериментальные оценки скоростей, определяемых в первых двух случаях, ниже, что требует объяснения и вызывает множество дебатов.

### МЕТОДИЧЕСКИЕ ПОДХОДЫ К ОПРЕДЕЛЕНИЮ СКОРОСТИ ГЕРМИНАЛЬНЫХ МУТАЦИЙ

Герминальные мутации являются предметом количественных исследований уже длительное время, начиная с 20-х гг. XX в. (см. обзор [8]). Тем не менее усилия прошлого века породили лишь фрагментарную картину. Изучение скорости спонтанных мутаций не достигало такого уровня точности, чтобы рассматриваться как количественная наука. Авторы вышеуказанного обзора задают вопрос: “Почему через 100 лет после определения массы электрона и скорости света мы все еще в процессе попыток сделать точные измерения этого фундаментального параметра?”. Они перечисляют шесть причин этого “отставания”.

Прежде всего это связано с тем, что частота мутаций в расчете на одну пару оснований очень низка. Для измерения таких низких скоростей необходимы чрезвычайно тщательные и кропотливые методы. Во-вторых, мутация является *случайным* процессом, и нет способа узнать, где и когда произойдет мутация. В-третьих, существует несколько видов мутаций, от точечных генных мутаций до хромосомных и геномных, и их скорости необходимо измерять отдельно. В-четвертых, скорость мутации может зависеть от *генетического фона*, особенно от контекста вокруг рассматриваемого сайта или его метилирования. Этот фон различен у разных индивидуумов. В-пятых, на скорость мутации могут оказывать значительное влияние многие мутагенные экологические и физиологические факторы, в том числе температура, химические вещества (список опасных веществ, <http://www.evol.nw.ru/~spirov/hazard/hazard.html>, содержит около 700 известных мутагенов), радиация, возраст, пол и т.д. Наконец, сама скорость мутации подвержена селекции и эволюции, причем некоторые авторы предполагают, что организмы могут увеличивать скорость мутации во время стресса. Поэтому существует очень большая естественная вариабельность, которая приводит к существенным разбросам в величинах скорости мутаций, получаемых разными авторами и разными методами.

В случае мутаций типа 1, которые, как и в соматических клетках, формируются в процессе развития данного индивидуума при репликации ДНК в процессе гаметогенеза и определяют различия (мозаицизм) между его отдельными гаметами, скорость герминальных мутаций может определяться секвенированием индивидуальных

сперматозоидов (ооцитов). В случае мутаций типа 2, передаваемых потомству и определяющих различия в геномах последовательных поколений (в основном родители и их дети), скорость мутаций может определяться анализом частоты возникновения генетических болезней в популяции, а также путем сравнительного анализа последовательностей геномов родителей и потомков. Наконец, в случае мутаций типа 3, которые передаются потомкам, фиксируются в эволюции и определяют эволюционные процессы, в частности различия между геномами гоминоидов, гоминидов и гоминин, скорость определяется путем сравнения последовательностей геномов современного человека и его эволюционных родственников, произошедших от общего предка, выявления числа различий между ними (дивергенции) и отнесения числа различий ко времени или числу генераций, отделяющих современного человека (или его эволюционных родственников) от общего предка. Последнее может определяться независимо, путем палеонтологического и археологического датирования, включая радиоуглеродный и другие виды радиоактивных изотопов. О возникающих при этом проблемах см. ниже.

Неудивительно, что разная история происхождения мутаций, выявляемых разными методами, обуславливает различные значения скорости мутаций, получаемые этими методами.

Рассмотрим эти методы несколько более подробно. При этом речь будет идти в основном о мутациях, вызываемых ошибками репликации, приводящими к замене оснований. Это наиболее частый источник ошибок [9]. Другие типы мутационных изменений изучены гораздо меньше и, соответственно, им будет уделено меньшее внимание.

### ВНУТРИОРГАНИЗМЕННАЯ (ТИП 1) СКОРОСТЬ МУТАЦИЙ В ЗАРОДЫШЕВОЙ ЛИНИИ

Single cell ДНК-секвенирование [10–12] обеспечивает новый подход к изучению механизмов, которые приводят к изменению зародышевой линии. В одном из первых исследований по этой теме были выделены отдельные клетки спермы, которые выявили в среднем 22.8 события рекомбинации и 25–36 *de novo* мутаций в каждой клетке спермы [13].

Можно приблизительно оценить скорость мутаций –  $10^{-8}$  на генерацию. Это не вполне согласуется со скоростью мутаций, определенной путем оценок *de novo* мутаций в сперме при сравнении геномов родственников, о чем речь пойдет ниже. Так, [4] дает общую скорость герминальных мутаций в среднем 64 новых мутаций на одного ребенка, 78% из которых, в среднем 50, име-

ют отцовское происхождение, т.е. представляют число мутаций на сперматозоид. Это выше 25–36 *de novo* мутаций, которое дает [13]. К сожалению, такие сравнения вряд ли правомерны, если учитывать низкую точность определений.

Авторы [13] также рассчитали, что 7% одиночных сперматозоидов имеют анеуплоидные геномы. В соответствии с этим исследованием другая группа использовала секвенирование целого генома для определения гаплотипов в отдельных клетках спермы у одного человека, что выявило в среднем 25.3 случая рекомбинации на клетку [14]. Это также несколько ниже, чем количество рекомбинаций, определяемых другими методами, 36 событий рекомбинации на одно поколение [15]. В другом исследовании определяли последовательности 99 сперматозоидов у азиатского индивидуума и сообщали об анеуплоидии в 4% клеток и 26 событиях рекомбинации на одну сперматозоидную клетку [16].

В то время как в большинстве исследований зародышевой линии основное внимание уделяется сперматозоидам, недавно был выполнен анализ оплодотворенных ооцитов [17]. Были проанализированы ооциты от восьми отдельных женщин и выявлены 43 события рекомбинации на ооцит, причем скорость рекомбинации была в 1.6 раза выше, чем в сперме. В этом исследовании также сообщалось о гораздо более высокой скорости анеуплоидии в ооцитах (17.6%) по сравнению со спермой (4–7%). В совокупности эти исследования показали сильное геномное разнообразие, которое возникает в зародышевых клетках перед передачей генетического материала потомству [18].

### СКОРОСТЬ ГЕРМИНАЛЬНЫХ МУТАЦИЙ ТИПА 2, ОПРЕДЕЛЯЕМЫХ ПО ЧАСТОТЕ НАСЛЕДСТВЕННЫХ БОЛЕЗНЕЙ В ПОПУЛЯЦИИ

Можно получить определенную (ограниченную) картину оценки скорости точечной мутации из баз данных о мутациях в *локусах*, которые, как известно, лежат в основе моногенных нарушений с явными фенотипическими эффектами. В случае аутосомно-доминантных и X-связанных расстройств больные индивидуумы могут быть легко идентифицированы как мутанты *de novo*. Оценка скорости мутаций человека, исходя из частоты моногенных (менделевских) заболеваний, как уже говорилось выше, имеет долгую историю, восходящую к Холдейну (John Haldane).

Исторически скорость мутаций в зародышевой линии у людей была рассчитана путем анализа данных по частоте возникновения новых случаев моногенных наследственных заболеваний [8, 19, 20].

Вычисления основывались на гипотезе о том, что тяжелое менделевское заболевание возникает из-за присутствия в популяции мутантных аллелей данного локуса. Аутосомно-доминантные менделевские болезни, вызванные одиночными высокопроникающими (*highly penetrant*) и вредными аллелями, возможно, легче других поддаются исследованию. Частота встречаемости данной болезни (обозначим ее  $f$ ) в популяции определяется балансом между скоростью  $\mu$  появления больных индивидуумов в популяции с каждой новой генерацией и скоростью их удаления из популяции (негативной селекцией), которое происходит, поскольку больные хуже приспособлены к выживанию в окружающей среде. В сбалансированном состоянии скорость появления больных равна скорости их удаления, так что частота  $f$  сохраняется примерно постоянной (конечно, если условия, в которых существует популяция, не меняются существенно). В первом приближении  $f = \mu/s$  (где  $\mu$  – скорость, при которой мутации приводят к возникновению данного заболевания, а  $s$  – коэффициент селекции против фенотипа, соответствующего данной болезни, который в первом приближении представляет собой долю мутантных аллелей, которые не передаются и поэтому теряются в результате отбора). Если доминирующее заболевание не смертельно, больные могут воспроизводиться, но давать меньше, чем “нормальное”, число потомков, т.е. их приспособленность уменьшена. Такой мутантный аллель путем селекции теряется из генофонда. Частота мутантных аллелей, ответственных за заболевание в популяции, представляет собой баланс между потерей мутантных аллелей за счет селекции и появления мутантных аллелей за счет новых мутаций. Когда генетическое заболевание ограничивает размножение настолько сильно, что  $s = 1$ , его называют генетическим летальным. Скорость мутации на пару оснований на генерацию можно затем оценить по частоте заболевания  $f$ , учитывая оценки значений  $s$  и мутационного размера мишени (т.е. количество сайтов, при мутации любого из которых возникает данное заболевание). Одна болезнь может вызываться несколькими мутациями в разных участках гена, ответственного за нее. Это повышает скорость, с которой эта болезнь возникает. Разные мутации образуют разные аллели. Каждый аллель вызывает болезнь. И скорость появления болезни равна сумме скоростей появления мутантных аллелей, каждая из которых примерно равна скорости мутаций на пару оснований, если речь идет о точечных мутациях. Эти подходы, основанные на заболеваемости, делают несколько предположений не только о коэффициенте селекции и размере мишени, но и требуют полной уверенности, что наблюдаемый фенотип – это именно тот, который

вызывается данным генотипом, а не генокопия или фенокопия [21].

Наш соотечественник Алексей Кондрашев [22] впервые использовал информацию о последовательностях мутантных аллелей, вызывающих моногенные генетические дефекты. Он сопоставлял данные о частотах мутаций локусов с последовательностями нуклеотидных замещений, делеций, инсерций и более сложных событий в девяти локусах, вызывающих аутосомно-доминантные заболевания, и 12 локусах, вызывающих X-связанные заболевания. Средняя оценка скорости всех мутаций составляет  $1.8 \times 10^{-8}$  на каждый нуклеотид на одно поколение, а коэффициент изменения этой скорости по 20 локусам составляет 0.53. Из общих соображений эта скорость не может быть более высокой, чем скорость мутаций зародышевой линии (если исключить дополнительные мутации, происходящие в процессе фертилизации), и то, что она все-таки больше, наводит в очередной раз на мысли о высокой степени неточности всех определений подобного типа.

Одиночные нуклеотидные замены происходят примерно в 25 раз чаще, чем все другие мутации, делеции — примерно в 3 раза чаще, чем вставки, сложные мутации очень редки, а в CpG богатых областях скорость замещений на порядок выше.

Со времени исследований Кондрашова наблюдался существенный приток новых данных о большем количестве аллелей и исследовалось больше расстройств, что позволило уточнить предварительные оценки скоростей и их вариаций между разными локусами (для более новой информации см. [23, 24]). Однако цифры по скорости мутаций остаются примерно такими же.

#### “ПРЯМЫЕ” МЕТОДЫ ОЦЕНКИ СКОРОСТИ МУТАЦИЙ ТИПА 2 ПРИ ПЕРЕНОСЕ ОТ РОДИТЕЛЕЙ К ДЕТАМ

Общий подход, который стал возможен только в течение последних нескольких лет, состоит в том, чтобы подсчитывать вновь возникающие мутации по данным глубокого секвенирования членов семей, особенно трио родителей-детей. Этот подход дает прямую оценку, но может быть технически осложнен, поскольку может давать ошибки, свойственные современным высокопроизводительным методам. В частности, случайные ошибки при определении последовательности и ошибки выравнивания могут быть трудно отличимы от истинных мутаций *de novo*. Кроме того, соматические мутации в геномах потомства секвенированных семей не всегда могут быть отделены от новых зародышевых вариантов [25].

Неожиданно, что эти оценки оказались заметно ниже (примерно в 2 раза), чем те, которые основаны на сравнении дивергенции видов (см. ниже).

Сравнение семейных трио давало оценки  $1-1.8 \times 10^{-8}$  замен на пару оснований на генерацию, в отличие от  $2-2.5 \times 10^{-8}$ , получаемых при межвидовых сравнениях [1, 4, 20, 26, 27]. При этом скорость существенно варьировала от семьи к семье. Это означает, что в геноме среднего человека содержится 40–100 однонуклеотидных мутаций *de novo*, причем одна–две из них попадают на кодирующую последовательность [28, 29].

Современные геномные подходы позволяют определить дополнительные характеристики мутаций *de novo*, например от кого из родителей произошла данная мутация, произошла она в зародышевой линии или постзиготически. Теперь мы знаем, что большинство мутаций *de novo* в зародышевой линии имеют отцовское происхождение и что более высокий возраст отцовства при зачатии приводит к увеличению числа *de novo* мутаций у потомства [20] (см. также ниже).

Исследование скорости герминальных мутаций генома путем полногеномного секвенирования 78 исландских трио родителей-потомков показывает, что при среднем возрасте отца при зачатии ребенка равном 29.7 лет средняя частота мутации *de novo* составляет  $1.20 \times 10^{-8}$  на пн на генерацию. Эти наблюдения указывают на важность влияния возраста отца на риск таких заболеваний как, в частности, шизофрения и аутизм [29]. Каждый новорожденный получает 30–100 (в среднем 60) новых мутаций. Это исследование было расширено дальше [30]. Цифры будут не безынтересны. Они дают представление о масштабе мутационного процесса. Были секвенированы геномы 1548 исландцев, их родителей, и в 225 случаях, по крайней мере, одного ребенка. Обнаружили 108778 *de novo* мутаций, как мононуклеотидных полиморфизмов, так и *indels*. Число мутантов *de novo* от матерей увеличивается с возрастом значительно медленнее, чем у отцов. Интересно, что эти возрастные изменения распределены неравномерно по всему геному. Ярким примером является область 20 Mb на хромосоме 8p с частотой мутации, в 50 раз превышающей частоту в остальных участках генома. Эта особенность наблюдалась у шимпанзе, в меньшей степени — у горилл и почти отсутствует у орангутанов. Это демонстрирует, что разнообразие последовательностей у людей обусловлено эволюционирующим взаимодействием между возрастом, полом, типом мутаций и их геномным положением [30].

Исследование, оценивающее частоту новых вариантов числа копий повторяющихся элементов (copy number variants, CNV) зародышевой линии, показало, что можно ожидать, что около 0.0065 (с длиной >500 kb) до 0.0123 (~30 kb) новых CNV будут возникать на геном на генерацию [31, 32]. Другое исследование *de novo* структурных из-

менений дает скорость 2.94 инделей (длиной 1–20 пн) и 0.16 структурных вариантов (>20 пн) на геном на генерацию [33]. Несмотря на то что число описанных структурных вариантов *de novo* меньше, чем число новых точечных мутаций, число оснований на один геном на одно поколение, которые вовлечены в эти структурные изменения, на самом деле в 50–100 раз больше [32].

### ОПРЕДЕЛЕНИЕ СКОРОСТЕЙ ГЕРМИНАЛЬНЫХ МУТАЦИЙ ТИПА 3 НА ОСНОВАНИИ СРАВНЕНИЯ ГЕНОМОВ СОВРЕМЕННОГО ЧЕЛОВЕКА И ЕГО ЭВОЛЮЦИОННЫХ РОДСТВЕННИКОВ – ОБЕЗЬЯН. МОЛЕКУЛЯРНОЕ ДАТИРОВАНИЕ В ЭВОЛЮЦИИ ПРИМАТОВ

Приняв существование глобальных эволюционных часов (т.е. часов, предполагающих одинаковую скорость мутаций для всех видов или хотя бы для больших таксонов, например приматов или гоминоидов, что в дальнейшем оказалось неверным) и имея независимую оценку времени, когда сравниваемые виды разделились, можно перевести число замен между двумя современными геномами в годовую скорость мутации. Например, летопись окаменелостей указывает на время разделения между *Homo* и макаками резус (Old World Monkey, OWM) 30 млн лет. Принимая нуклеотидную дивергенцию между двумя видами ~6.2%, а длину диплоидного генома  $6 \times 10^9$  пн, можно вычислить скорость мутации  $10^{-9}$  на пн в год. Эта единица фигурировала вплоть до 2010 г. как скорость замещений в эволюции приматов (см. [27]).

Сейчас есть данные, которые указывают на то, что скорость замещений меняется от вида к виду и, более того, на разных этапах эволюции видов [27, 34]. Естественно, что изменения, определяющие траекторию эволюции вида и зависящие от различий в признаках жизненного цикла (например, возраст и размер индивидуумов, принадлежащих к данному виду, в зрелости; возраст при первом воспроизведении; количество, размер и соотношение полов потомков; репродуктивный срок жизни и старение; продолжительность жизни), должны также быть найдены среди приматов. Группа авторов [34] проанализировала полные геномы 10 видов приматов, в том числе обезьян Старого Света (OWM), обезьян Нового Света (NWM) и высших обезьян. Эти авторы рассматривали предположительно нейтральные аутомомные участки и пришли к выводу, что скорости замещения в эволюционных линиях, ведущих к NWM от общего предка гоминоидов, и NWM примерно на 64% выше, чем в линиях, ведущих от общего предка к обезьянам. А у обезьян, в свою очередь, скорости замещений у шимпанзе на ~2% выше, а у горилл – на 7% выше, чем у лю-

дей. Здесь работают так называемые локальные часы.

Однако не все типы мутаций ведут себя одинаково. В частности, транзиции в сайтах CpG проявляют более регулярное (clock-like) поведение, чем другие типы, по-видимому из-за их нерепликативного происхождения. (Напомним, что CpG является последовательностью, в которой C подвергается метилированию и, легко дезаминируясь, превращается в T. В результате по этим сайтам более часто происходят транзиции C > T). Таким образом, среди приматов варьирует не только общая скорость мутаций, но и спектр мутаций.

Авторы полагают, что события в эволюции приматов наиболее надежно датируются с использованием CpG-транзиций. Принимая такой подход, они оценивают время дивергенции человека и шимпанзе 12.1 млн лет (по сравнению с традиционно принимаемым около 6 млн), а разделение человека и гориллы – 15.1 млн лет (по сравнению с традиционными 7–8 млн) [34]. Время покажет, кто ближе к истине.

Как говорилось в предыдущем параграфе, недавние прямые сравнения последовательностей родственных геномов дали оценки скорости примерно в 2 раза меньшие – приблизительно  $0.5 \times 10^{-9}$  на пн в год. Разница оказывается существенной – эта цифра отодвигает очень сильно назад время разделения между видами приматов и оно становится несовместимым с палеонтологическими данными. Например, с этой скоростью генетическое время дивергенции между обезьянами и обезьянами старого мира оценивается как 47 млн лет и, следовательно, приблизительное время видообразования – 40 млн лет, в то время как данные по датированию ископаемых дают оценку расщепления видов не более чем 25–30 млн лет назад [1]. Здесь мы сталкиваемся с расхождением времени генетической дивергенции и фактического расщепления видов [35].

Существует громадное число статей, пытающихся объяснить расхождения между методами оценки скорости. Сводка литературы и таблица скоростей мутаций, полученных разными авторами, приведены в [27]. Одно из возможных объяснений заключается в том, что в эволюции гоминоидов происходило снижение скорости мутаций [1]. Другое – что прямой анализ соматических мутаций и вывод из них скорости герминальных мутаций может приводить к многочисленным ошибкам и к занижению числа мутаций [36]. Указывалось также, что необходим учет особенностей признаков жизненного цикла в оценке скорости мутаций гоминоидов.

Попытка такого учета приводит к непротиворечивым результатам и оценивает время разделения человека и шимпанзе как 6.6 млн лет назад [37]. Эта цифра вполне согласуется с традицион-

ными оценками. Детальный анализ ситуации был дан недавно в хорошем обзоре [1]. Автор, как и многие другие исследователи, предполагает, что за последние 15 млн лет скорость мутаций могла замедлиться, что объясняет вышеуказанные противоречия. Анализ ископаемых останков свидетельствует о том, что предковые обезьяны были меньше современных, а мелкие животные склонны к более быстрому воспроизведению и, следовательно, более высокой скорости мутации. В частности, автор этого обзора отмечает, что дополнительные доказательства, подтверждающие более низкий уровень герминальных мутаций у современных людей по сравнению с их предками, можно получить из сравнения ДНК древних предков человека и ДНК современного человека (см. ниже).

Дискуссии на тему о скорости герминальных мутаций будут, по-видимому, продолжаться долго. Еще очень коротко упомянем подходы, основанные на калибровке скорости мутаций путем сравнения с другим процессом, данные по скорости которого более точны. Таким сравнительным процессом может быть скорость мутаций микросателлитов или скорость рекомбинации. В первом случае была получена скорость замещений на пн на генерацию  $1.4\text{--}2.3 \times 10^{-8}$  ( $0.56\text{--}0.9 \times 10^{-9}$  в год), во втором —  $1.6 \times 10^{-8}$  замещений на пн на генерацию ( $0.64 \times 10^{-9}$  в год) [26].

### ОПРЕДЕЛЕНИЕ СКОРОСТЕЙ ГЕРМИНАЛЬНЫХ МУТАЦИЙ ТИПА 3 НА ОСНОВАНИИ СРАВНЕНИЯ ГЕНОМОВ СОВРЕМЕННОГО ЧЕЛОВЕКА С ДАТИРОВАННЫМИ ДРЕВНИМИ ГЕНОМАМИ ТОГО ЖЕ ВИДА

ДНК из многих древних образцов (древней ДНК, дДНК. Ancient DNA, aDNA), которые раньше считались непригодными для анализа из-за сильной деградации, теперь могут успешно секвенироваться. Высокоскоростные методы новой генерации позволяют резко увеличить число считываемых последовательностей и надеяться, что это поможет выявить артефакты, связанные с загрязнением древней ДНК, особенно в исследованиях человека (см. ниже). В настоящее время секвенированы полные геномы древних анатомически современных людей, архаичных гомининов, древних патогенов и растений. Они выявили важную функциональную и генетическую информацию [38].

Недавнее сравнение последовательности генома современного человека и верхнепалеолитического 45000-летнего человека привело к оценке скорости  $0.44\text{--}0.63 \times 10^{-9}$  на пн в год или  $1.3\text{--}1.8 \times 10^{-8}$  на пн на генерацию (при условии, что она составляла 29 лет) [26].

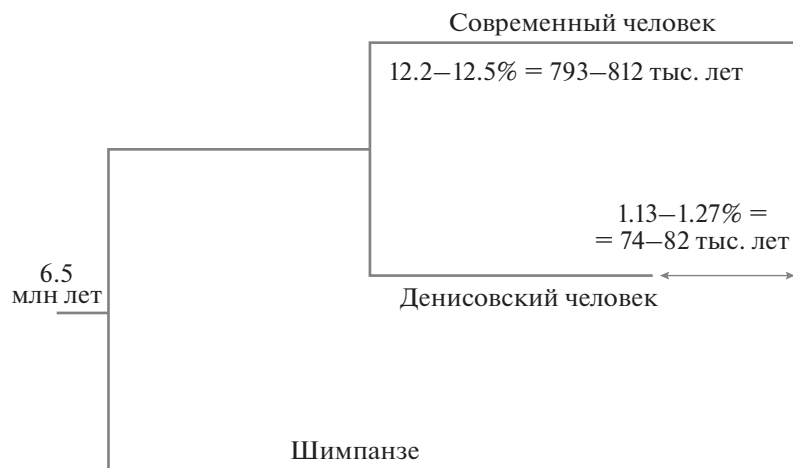
Изысканный метод “сокращения ветвей” (рис. 1) использовался при изучении истории популяций, определении скоростей мутаций и установлении временных масштабов для филогенетических деревьев у людей и других видов. Впервые в полномасштабном масштабе он был использован Мейером и др. [39] для оценки возраста кости Денисовского человека. Ее возраст, вероятно, больше 50000 лет и поэтому она не могла быть датирована радиоуглеродным методом. Авторы сравнили расхождение между Денисовской последовательностью и геномами современных людей. Результат демонстрируется на рис. 1.

Сравнение числа замещений, которые предположительно произошли между геномом предка человека, шимпанзе и геномом Денисовского человека, и числа замещений между геномами предка и современного человека показывает, что число замещений в геноме “Денисовца” на 1.16% ниже ( $1.13\text{--}1.27\%$  [25, 39, 40], соответствующая ветвь короче на рис. 1). Это, по-видимому, отражает меньший эволюционный возраст кости “Денисовца”, у которой было меньше времени для накопления изменений, чем у современных людей. Предполагая, что время расхождения в последовательности между людьми и шимпанзе 6.5 млн лет, по сокращению ветви Денисовца можно условно определить возраст кости — примерно от 74000 до 82000 лет, в согласии с археологическими данными. Из этих данных затруднительно вывести скорость мутаций, поскольку на эту оценку может влиять множество источников ошибок. Использование времени дивергенции шимпанзе и человека для калибровки основывается на оценках скорости мутации, которые, как мы видели, неточны. Сравнение миллионолетних интервалов с десятилетиями может приводить к большим ошибкам в оценке существования Денисовского человека. А десятки тысяч лет — это как раз период времени, охватывающий возраст подавляющего большинства древних человеческих образцов, секвенированных к настоящему времени [41].

Поэтому недавно был предложен идеологически подобный подход, но использующий “рекомбинационные часы”.

### ОЦЕНКА ВОЗРАСТА И ВРЕМЕНИ ГЕНЕРАЦИИ ГЕНЕТИЧЕСКИМ МЕТОДОМ С ИСПОЛЬЗОВАНИЕМ РЕКОМБИНАЦИЙ В ИСКОПАЕМОЙ И СОВРЕМЕННОЙ ДНК

Priya Moogjani и др. [41] разработали генетический подход для датирования древних геномов путем сопоставления количества событий генетической рекомбинации, накопившихся в геноме современных неафриканцев за время, прошедшее с момента внедрения неандертальской ДНК в ДНК предка древних людей (рис. 2). Это более



**Рис. 1.** Средняя дивергенция последовательностей и разница в длине ветвей между Денисовским человеком и 11 современными людьми. Дивергенция здесь представлена как доля (%) ветви, ведущей от человека к общему предшественнику человека и шимпанзе, и пересчитана в десятки тысяч лет, принимая время дивергенции человека и шимпанзе 6.5 млн лет назад (цит. по [39]).

или менее точно датированное событие сравнительно недавнего прошлого. Включение в геном предшественников современных людей ДНК неандертальцев происходило вследствие скрещиваний между двумя популяциями 37000–86000 лет назад. Эти вкрапления составляют от 1 до 4% генома современного человека в Евразии (но не в Африке). Поскольку подавляющее большинство древних образцов, секвенированных до настоящего времени, были обнаружены в Евразии (с оцененными возрастными значениями ~2000–45000 лет), авторы использовали их, чтобы определить число рекомбинаций с момента “вброса” неандертальской ДНК в геном современных людей и их вымерших родственников, Денисовских людей [41]. Предполагая приблизительно постоянную скорость рекомбинации за поколение, число рекомбинаций, которое произошло с момента первоначального введения неандертальской ДНК, можно использовать в качестве молекулярных часов для определения: 1) числа поколений, прошедших с момента внедрения неандертальской ДНК в геном древнего вымершего человека до момента смерти древнего обладателя этого генома; 2) числа поколений, прошедших между временем смерти древнего человека и нашим временем, временем существования современных геномов.

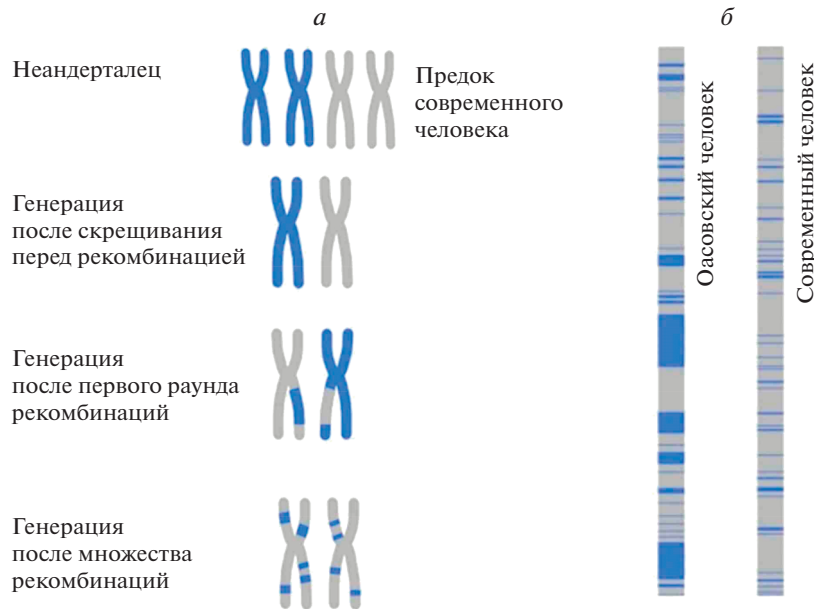
Идея метода та же, что была предложена [39] в описанном выше методе “укорочения ветвей”: древний геном вымершей популяции прошел меньше поколений в своей эволюции после отделения от общего предшественника по сравнению с геномами существующих родственников. Это значит, что он претерпел меньше событий рекомбинации. Чем более недавно ушел из жизни древний предок, тем меньше разница в числе рекомбинаций по сравнению с современным челове-

ком. Поскольку скорости рекомбинаций более постоянны в эволюции, чем скорости замещений, накопленное число событий рекомбинации обеспечивает молекулярные часы, позволяющие оценить число поколений, разделяющих время существования древнего родственника и время существования современника. Разницы в числе рекомбинаций могут быть переведены в абсолютное время в годах, используя оценку времени генерации.

Авторы опробовали свой метод, сопоставив даты, полученные на основании скорости рекомбинаций пяти секвенированных геномов древних человеческих образцов из Северной Америки, Европы и Сибири, с радиоуглеродными датами, и получили согласующиеся оценки возраста — между 12000 и 45000 лет назад. Используя корреляцию между радиоуглеродным датированием, измеренным в годах, и датами внедрения неандертальской ДНК в геном современного человека, измеренными в поколениях, авторы подсчитали, что в этот исторический интервал время одной генерации составляло приблизительно 26–30 лет, предполагая, что время генерации существенно не изменилось за последние 45000 лет. По мнению авторов, эта оценка согласуется с оценкой для сегодняшних западных евразийцев [41].

Идея хороша, но трудности заключаются в палеонтологическом датировании ископаемых остатков, которое не всегда можно сделать уверенно. Даже если возраст ископаемых останков можно определить точно, определение их правильных филогенетических позиций на эволюционном древе часто затруднено. Более того, ряд данных [42] указывает на то, что время, отделяющее нас от последнего общего предка, определяемого по дивер-





**Рис. 2.** Оценка возраста останков, основанная на распределении длин или количества рекомбинантных фрагментов (см. текст) (а). Сравнение хромосомы 6 из 40000-летних останков из Пештера-ку-Оасе (Pestera cu Oase) в Румынии с хромосомой современного человека (б). Синие полосы представляют собой сегменты неандертальской ДНК от прошлого скрещивания. Сегменты Оаса длиннее, потому что у него был неандертальский предок всего за 4–6 поколений перед его рождением. Оценки основаны на числе актов рекомбинации (цит. по [15]).

генции генов, всегда больше, а иногда намного больше, чем время дивергенции собственно видов, т.е. время, когда они прекратили контакты [26]. Это означает, что палеонтологические и археологические калибровки, основанные на определении времени раздельного существования видов (популяций), не всегда могут быть непосредственно применены к генетическим расхождениям [26, 35]. Поэтому и по ряду других причин, которые мы не имеем возможности обсудить в рамках этого обзора, остается неопределенность в отношении времени расхождения между людьми, шимпанзе и другими гоминидами.

Завершая, еще раз подчеркнем, что все измерения крайне приблизительны и нужно уважительно, но скептически, относиться ко всем выводам, которые делаются на сравнительном анализе получаемых цифр.

Основные результаты оценок скоростей герминальных мутаций сведены в табл. 1.

### ВОЗРАСТ РОДИТЕЛЕЙ, ОСОБЕННО ОТЦА, ВЛИЯЕТ НА ВЕРОЯТНОСТЬ ГЕРМИНАЛЬНЫХ МУТАЦИЙ

Сперма более пожилых отцов претерпевает больше раундов клеточного деления, поэтому накапливает больше мутаций. С другой стороны, матери передают меньше мутаций (около 0.25 в год), так как ооциты женщины формируются в основном одновременно, до ее рождения. Ско-

рость мутаций также зависит от таких факторов, как начало полового созревания, возраст воспроизводства и скорость производства спермы. Эти признаки жизненного цикла варьируют у живых приматов и, вероятно, также различаются между вымершими видами предков человека [15].

Было подсчитано, что мужская зародышевая линия испытывает 160 репликаций генома у 20-летнего мужчины, увеличиваясь до 610 повторов генома у 40-летнего мужчины [4].

Это согласуется с тем, что описано в исследовании [29]: общее число мутаций *de novo* у ребенка сильно коррелировало с возрастом отца при зачатии, увеличиваясь на ~1 или 2 *de novo* мутаций за каждый дополнительный год отцовского возраста, в то время как мамы вносили небольшую долю (~10 *de novo* мутаций) в существенно меньшей зависимости от их возраста. Аналогичные результаты были получены в ряде других исследований [4, 43]. На сегодняшний день консенсус относительно скорости герминальных мутаций и, следовательно, риска возникновения генетических болезней связан с тем, что основным виновником является подверженный ошибкам процесс репликации ДНК, число циклов которого увеличивается с возрастом отца [44].

В недавнем широкомасштабном исследовании [45] были использованы данные по 816 семейным трио и идентифицированы 36441 *de novo* мутаций (в среднем ~45 мутаций на ребенка). Это более низкая цифра, чем было описано ранее. Ав-

**Таблица 1.** Основные результаты анализа скоростей герминальных мутаций в геноме человека

Тип мутации (см. текст)	Метод определения	Значение замещений (рекомбинаций) на пн (рекомбинаций на геном) на генерацию	Ссылка
1	Секвенирование сперматозоидов	$10^{-8}$	[13, 18]
1	Секвенирование ооцитов	43 рекомбинации	[17]
2	Прямое секвенирование родственных геномов*	$1 - 1.8 \times 10^{-8}$	[1, 4, 20, 25–29, 32]
2	Исследование частоты наследственных болезней в популяции и секвенирование	$1.8 \times 10^{-8}$	[22]
3	Секвенирование геномов человека—обезьян	$2-2.5 \times 10^{-8}$	[27]
3	Сравнение частоты замещений со скоростью рекомбинаций или мутаций микросателлитов	$1.4-2.3 \times 10^{-8}$	[26]
3	Сравнение геномов современного человека с датированными древними геномами того же вида	$1.3-1.8 \times 10^{-8}$	[26]

\* Детальная сводка и таблица скоростей мутаций, полученных разными авторами этим методом, приведены в [27].

торы подтвердили некоторые предыдущие результаты: в том числе 80% мутаций, возникающих во время сперматогенеза, и эффект отцовского возраста — появление ~ одной новой мутации с каждым годом увеличения возраста отца. В более ранних исследованиях отсутствовала способность обнаруживать небольшой, но значительный, возрастной эффект для материнских мутаций *de novo*; в данной работе показано, что матери вносят ~ одну мутацию на каждые четыре дополнительных года материнского возраста. Выяснилось также, что отдельные районы генома обогащены материнскими мутациями.

Механизмы, приводящие к увеличению генетических рисков для потомков отцов более пожилого возраста, становятся яснее с улучшением технологий геномики. Помимо мутаций, связанных с репликацией ДНК, есть множество других вариантов мутаций. При этом замены оснований, вставки-делеции, экспансии повторов и хромосомные изменения следуют различным правилам (см. [46]).

Например, в настоящее время есть убедительные генетические доказательства того, что помимо приблизительно линейного увеличения с возрастом числа точечных мутаций, происходящих от отца [4, 29], с возрастом происходит более резкий рост мутаций, приводящих к приобретению определенных функций, например в генах *FGFR2*, *FGFR3*, *HRAS*, *PTPN11* и *RET*. Все эти гены участвуют в передаче сигнала рецептора фактора роста — RAS, который активен в сперматогониях и обычно дисрегулируется при раке. Предполагают, что эти мутации приводят к селективным преимуществам для пролиферации и/или выживания мутантных сперматогониев [47].

#### ЗАКЛЮЧЕНИЕ. СУЩЕСТВУЕТ ДИСБАЛАНС МЕЖДУ СЛОЖНЫМИ ПРОГРАММНЫМИ АЛГОРИТМАМИ АНАЛИЗА МОЛЕКУЛЯРНЫХ ЧАСОВ И СКУДНЫМИ ДАННЫМИ ГЕНЕТИЧЕСКИХ, ПАЛЕОНТОЛОГИЧЕСКИХ И АРХЕОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ

Многие исследователи отмечали громадный дисбаланс между сложными программными алгоритмами, используемыми для анализа молекулярных часов, с одной стороны, и плохими экспериментальными данными, с другой стороны.

Этот дисбаланс не позволяет рационально объяснить расхождения между данными по скорости герминальных мутаций, получаемых различными методами.

В недавних статьях приводится много объяснений, почему эти оценки отличаются друг от друга (см. например, [21, 48, 49]). Возможные недостатки включают: а) небольшие размеры выборки, с точки зрения числа индивидуумов, использованных для получения данных; б) неточные оценки числа ложнопозитивных и ложнонегативных результатов, весьма вероятные, поскольку сравниваются последовательности, полученные от разных людей и имеющие разные уровни этих ошибок; в) рассмотрение только мутаций, происходящих в одном поколении, что приводит к неполному выявлению постзиготических мутаций [4]; г) недостаточный учет корреляции с отцовским возрастом и ряд других.

Интересно, что в перечисляемом списке не фигурируют колоссальные ошибки, свойственные методам, используемым для получения данных для анализа. Например, секвенирование древней ДНК несет в себе массу источников оши-

бок, которые были недавно проанализированы [50–52]. Это – недостаточно высокое качество ископаемых ДНК, которые, естественно, деградировали с разной для разных образцов скоростью в зависимости от температуры, влажности, рН среды и т.д. Очень важную роль играют загрязнения этой ДНК современными примесями микробиологической природы. Неизбежно сравниваются образцы, полученные из разных мест и, следовательно, хранившиеся в разных условиях. Обнаруживаемые различия могут быть связаны не с эволюционными процессами, а с процессами, происходившими во время, когда образец находился в месте захоронения. Caldararo [50] в весьма критическом анализе статьи [53] пишет: “Хотя авторы сообщения по двум Денисовским последовательностям интерпретируют большое число различий между ними как доказательство генетического разнообразия и, следовательно, длительного нахождения Денисовцев в данном месте пребывания, эти различия могут быть результатом контаминации и деградации”.

Статья [50] заслуживает гораздо более детального анализа, чем тот, который мы можем дать в коротком заключении к короткому обзору, но вместе с другими подобными статьями она сеет серьезные сомнения не только в надежности данных, получаемых в результате секвенирования древних ДНК, но, более того, в разумности их интерпретации. К сожалению, здесь, несмотря на многие предостережения (см. из совсем недавнего [52]), складывается ситуация, которую Уильям Фолкнер в своем романе “Шум и ярость” (“The Sound and the Fury”) выразил таким образом: “Все сразу заговорили – горячо, наперебой, запальчиво, нереальное обращая в возможное, затем в вероятное, затем в неоспоримый факт, как это у людей всегда выходит, когда они желают облакают в слова”.

Конечно, экспериментальные данные по секвенированию, получаемые с современными образцами, значительно менее подвержены ошибкам, но и их интерпретация, по-видимому, нуждается в серьезном совершенствовании. Следует более тщательно учитывать зависимость скорости мутаций от признаков жизненного цикла, нельзя исключать что у одного и того же вида скорость мутаций могла быть различной в зависимости от места обитания, где меняются условия окружающей среды и источники питания; по-видимому в эволюции гоминоидов происходило снижение скорости мутаций по мере приближения к сегодняшним популяциям. Наконец, прямой анализ соматических мутаций и вывод из него скорости герминальных могут приводить к многократным ошибкам и к занижению числа мутаций [36]. Эти ошибки могут быть систематическими, поскольку в процессе развития сомы могли происходить различные селекционные

процессы, отбрасывающие значительную часть мутаций, которые являются нейтральными для условий развития герминальной линии, и наоборот. Возможно ли корректно учесть все эти источники вариабельности?

Если мы хотим знать, кто мы, откуда мы взяли и когда, и куда идем, следует вложить в эту проблему интеллект и опыт специалистов разного профиля, чтобы избежать проблемы GIGO (англ. Garbage In, Garbage Out, “Мусор на входе – мусор на выходе”) – принцип в информатике, подразумевающий, что при неверных входящих данных будут получены недостоверные результаты, даже если используется верный алгоритм обработки.

### СПИСОК ЛИТЕРАТУРЫ

1. Scally A. Mutation rates and the evolution of germline structure // *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 2016. V. 371. № 1699. doi 10.1098/rstb.2015.0137
2. Lynch M. Rate, molecular spectrum, and consequences of human mutation // *Proc. Natl Acad. Sci. USA.* 2010. V. 107. № 3. P. 961–968. doi 10.1073/pnas.0912629107
3. Campbell I.M., Yuan B., Robberecht C. et al. Parental somatic mosaicism is underrecognized and influences recurrence risk of genomic disorders // *Am. J. Hum. Genet.* 2014. V. 95. № 2. P. 173–182. doi 10.1016/j.ajhg.2014.07.003
4. Rahbari R., Wuster A., Lindsay S.J. et al. Timing, rates and spectra of human germline mutation // *Nat. Genet.* 2016. V. 48. № 2. P. 126–133. doi 10.1038/ng.3469
5. Tang W.W., Kobayashi T., Irie N. et al. Specification and epigenetic programming of the human germ line // *Nat. Rev. Genet.* 2016. V. 17. № 10. P. 585–600. doi 10.1038/nrg.2016.88
6. Alekseenko I.V., Kuzmich A.I., Pleshkan V.V. et al. The cause of cancer mutations: Improvable bad life or inevitable stochastic replication errors? // *Mol. Biol. (Mosk.)*. 2016. V. 50. № 6. P. 906–921. doi 10.7868/S0026898416060033
7. Sverdlov E.D., Mineev K. Mutation rate in stem cells: an underestimated barrier on the way to therapy // *Trends Mol. Med.* 2013. V. 19. № 5. P. 273–280. doi 10.1016/j.molmed.2013.01.004
8. Kondrashov F.A., Kondrashov A.S. Measurements of spontaneous rates of mutations in the recent past and the near future // *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 2010. V. 365. № 1544. P. 1169–1176. doi 10.1098/rstb.2009.0286
9. Lynch M. Evolution of the mutation rate // *Trends Genet.* 2010. V. 26. № 8. P. 345–352. doi 10.1016/j.tig.2010.05.003
10. Wang J., Song Y. Single cell sequencing: a distinct new field // *Clin. Transl. Med.* 2017. V. 6. № 1. P. 10. doi 10.1186/s40169-017-0139-4
11. Golov A.K., Razin S.V., Gavrilov A.A. Single-cell genome-wide studies give new insight into nongenetic cell-to-cell variability in animals // *Histochem. Cell Biol.* 2016. V. 146. № 3. P. 239–254. doi 10.1007/s00418-016-1466-z

12. *Junker J.P., van Oudenaarden A.* Every cell is special: genome-wide studies add a new dimension to single-cell biology // *Cell*. 2014. V. 157. № 1. P. 8–11. doi 10.1016/j.cell.2014.02.010
13. *Wang J., Fan H.C., Behr B. et al.* Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm // *Cell*. 2012. V. 150. № 2. P. 402–412. doi 10.1016/j.cell.2012.06.030
14. *Kirkness E.F., Grindberg R.V., Yee-Greenbaum J. et al.* Sequencing of isolated sperm cells for direct haplotyping of a human genome // *Genome Res*. 2013. V. 23. № 5. P. 826–832. doi 10.1101/gr.144600.112
15. *Alex B., Moorjani P.* DNA dating: How molecular clocks are refining human evolution's timeline // *The Conversation*. 2017. 7 April.
16. *Lu S., Zong C., Fan W. et al.* Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing // *Science*. 2012. V. 338. № 6114. P. 1627–1630. doi 10.1126/science.1229112
17. *Hou Y., Fan W., Yan L. et al.* Genome analyses of single human oocytes // *Cell*. 2013. V. 155. № 7. P. 1492–1506. doi 10.1016/j.cell.2013.11.040
18. *Wang Y., Navin N.E.* Advances and applications of single-cell sequencing technologies // *Mol. Cell*. 2015. V. 58. № 4. P. 598–609. doi 10.1016/j.molcel.2015.05.005
19. *Keightley P.D.* Rates and fitness consequences of new mutations in humans // *Genetics*. 2012. V. 190. № 2. P. 295–304. doi 10.1534/genetics.111.134668
20. *Acuna-Hidalgo R., Veltman J.A., Hoischen A.* New insights into the generation and role of *de novo* mutations in health and disease // *Genome Biol*. 2016. V. 17. № 1. P. 241. doi 10.1186/s13059-016-1110-1
21. *Segurel L., Wyman M.J., Przeworski M.* Determinants of mutation rate variation in the human germline // *Annu. Rev. Genomics Hum. Genet*. 2014. V. 15. P. 47–70. doi 10.1146/annurev-genom-031714-125740
22. *Kondrashov A.S.* Direct estimates of human per nucleotide mutation rates at 20 loci causing Mendelian diseases // *Hum. Mutat*. 2003. V. 21. № 1. P. 12–27. doi 10.1002/humu.10147
23. *de Ligt J., Veltman J.A., Vissers L.E.* Point mutations as a source of *de novo* genetic disease // *Curr. Opin. Genet. Dev*. 2013. V. 23. № 3. P. 257–263. doi 10.1016/j.gde.2013.01.007
24. *Smith T., Ho G., Christodoulou J. et al.* Extensive variation in the mutation rate between and within human genes associated with mendelian disease // *Hum. Mutat*. 2016. V. 37. № 5. P. 488–494. doi 10.1002/humu.22967
25. *Campbell C.D., Eichler E.E.* Properties and rates of germline mutations in humans // *Trends Genet*. 2013. V. 29. № 10. P. 575–584. doi 10.1016/j.tig.2013.04.005
26. *Lipson M., Loh P.R., Sankararaman S. et al.* Calibrating the human mutation rate via ancestral recombination density in diploid genomes // *PLoS Genet*. 2015. V. 11. № 11. P. e1005550. doi 10.1371/journal.pgen.1005550
27. *Moorjani P., Gao Z., Przeworski M.* Human germline mutation and the erratic evolutionary clock // *PLoS Biol*. 2016. V. 14. № 10. P. e2000744. doi 10.1371/journal.pbio.2000744
28. *Conrad D.F., Keebler J.E., DePristo M.A. et al.* Variation in genome-wide mutation rates within and between human families // *Nat. Genet*. 2011. V. 43. № 7. P. 712–714. doi 10.1038/ng.862
29. *Kong A., Frigge M.L., Masson G. et al.* Rate of *de novo* mutations and the importance of father's age to disease risk // *Nature*. 2012. V. 488. № 7412. P. 471–475. doi 10.1038/nature11396
30. *Jonsson H., Sulem P., Kehr B. et al.* Parental influence on human germline *de novo* mutations in 1.548 trios from Iceland // *Nature*. 2017. V. 549. № 7673. P. 519–522. doi 10.1038/nature24018
31. *Itsara A., Wu H., Smith J.D. et al.* De novo rates and selection of large copy number variation // *Genome Res*. 2010. V. 20. № 11. P. 1469–1481. doi 10.1101/gr.107680.110
32. *Forsberg L.A., Gisselsson D., Dumanski J.P.* Mosaicism in health and disease – clones picking up speed // *Nat. Rev. Genet*. 2017. V. 18. № 2. P. 128–142. doi 10.1038/nrg.2016.145
33. *Kloosterman W.P., Francioli L.C., Hormozdiari F. et al.* Characteristics of *de novo* structural changes in the human genome // *Genome Res*. 2015. V. 25. № 6. P. 792–801. doi 10.1101/gr.185041.114
34. *Moorjani P., Amorim C.E., Arndt P.F. et al.* Variation in the molecular clock of primates // *Proc. Natl Acad. Sci. USA*. 2016. V. 113. № 38. P. 10607–10612. doi 10.1073/pnas.1600374113
35. *Wilke T., Schultheiß R., Albrecht C.* As time goes by: A simple fool's guide to molecular clock approaches in invertebrates // *Am. Malac. Bull*. 2009. V. 27. P. 25–45. doi 10.4003/006.027.0203
36. *Chen C., Qi H., Shen Y. et al.* Contrasting determinants of mutation rates in germline and soma // *Genetics*. 2017. V. 207. № 1. P. 255–267. doi 10.1534/genetics.117.1114
37. *Amster G., Sella G.* Life history effects on the molecular clock of autosomes and sex chromosomes // *Proc. Natl Acad. Sci. USA*. 2016. V. 113. № 6. P. 1588–1593. doi 10.1073/pnas.1515798113
38. *Der Sarkissian C., Allentoft M.E., Avila-Arcos M.C. et al.* Ancient genomics // *Philos. Trans. R. Soc. Lond. B. Biol. Sci*. 2015. V. 370. № 1660. P. 20130387. doi 10.1098/rstb.2013.0387
39. *Meyer M., Kircher M., Gansauge M.T. et al.* A high-coverage genome sequence from an archaic Denisovan individual // *Science*. 2012. V. 338. № 6104. P. 222–226. doi 10.1126/science.1224344
40. *Shendure J., Akey J.M.* The origins, determinants, and consequences of human mutations // *Science*. 2015. V. 349. № 6255. P. 1478–1483. doi 10.1126/science.aaa9119
41. *Moorjani P., Sankararaman S., Fu Q. et al.* A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45000 years // *Proc. Natl Acad. Sci. USA*. 2016. V. 113. № 20. P. 5652–5657. doi 10.1073/pnas.1514696113
42. *Segurel L., Thompson E.E., Flutre T. et al.* The ABO blood group is a trans-species polymorphism in primates // *Proc. Natl Acad. Sci. USA*. 2012. V. 109. № 45. P. 18493–18498. doi 10.1073/pnas.1210603109
43. *Francioli L.C., Polak P.P., Koren A. et al.* Genome-wide patterns and properties of *de novo* mutations in humans //

- Nat. Genet. 2015. V. 47. № 7. P. 822–826. doi 10.1038/ng.3292
44. *Goriely A.* Decoding germline *de novo* point mutations // Nat. Genet. 2016. V. 48. № 8. P. 823–824. doi 10.1038/ng.3629
45. *Goldmann J.M., Wong W.S., Pinelli M. et al.* Parent-of-origin-specific signatures of *de novo* mutations // Nat. Genet. 2016. V. 48. № 8. P. 935–939. doi 10.1038/ng.3597
46. *Crow J.F.* The origins, patterns and implications of human spontaneous mutation // Nat. Rev. Genet. 2000. V. 1. № 1. P. 40–47. doi 10.1038/35049558
47. *Maher G.J., Rajpert-De Meyts E., Goriely A. et al.* Cellular correlates of selfish spermatogonial selection // Andrology. 2016. V. 4. № 3. P. 550–553. doi 10.1111/andr.12185
48. *Narasimhan V.M., Rahbari R., Scally A. et al.* Estimating the human mutation rate from autozygous segments reveals population differences in human mutational processes // Nat. Commun. 2017. V. 8. № 1. P. 303. doi 10.1038/s41467-017-00323-y
49. *Scally A., Durbin R.* Revising the human mutation rate: implications for understanding human evolution // Nat. Rev. Genet. 2012. V. 13. № 10. P. 745–753. doi 10.1038/nrg3295
50. *Caldararo N.* Denisovans, Melanesians, Europeans, and Neandertals: The confusion of DNA assumptions and the biological species concept // J. Mol. Evol. 2016. V. 83. № 1–2. P. 78–87. doi 10.1007/s00239-016-9755-7
51. *Allentoft M.E., Collins M., Harker D. et al.* The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils // Proc. Biol. Sci. 2012. V. 279. № 1748. P. 4724–4733. doi 10.1098/rspb.2012.1745
52. *Llamas B., Valverde G., Fehren-Schmitz L. et al.* From the field to the laboratory: Controlling DNA contamination in human ancient DNA research in the high-throughput sequencing era // STAR: Sci. Technol. Archaeol. Res. 2017. V. 3. № 1. P. 1–14. doi 10.1080/20548923.2016.1258824
53. *Vernot B., Tucci S., Kelso J. et al.* Excavating Neandertal and Denisovan DNA from the genomes of Melanesian individuals // Science. 2016. V. 352. № 6282. P. 235–239. doi 10.1126/science.aad9416

## The Rate of Human Germinal Mutations – Variable Factor of Evolution and Disease

N. Y. Uspenskaya<sup>a, \*</sup>, S. B. Akopov<sup>a</sup>, E. V. Snezhkov<sup>a</sup>, and E. D. Sverdlov<sup>a</sup>

<sup>a</sup>*Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, 117997 Russia*

*\*e-mail: natusp@gmail.com*

The emergence of genetic diseases and evolutionary processes are associated with the flow of genetic information from one generation to another. During this transfer genetic information carried by gametes changes by *de novo* germinal mutations. The mutation rate determines the rate of evolution and the incidence of inherited pathologies. Despite the great theoretical and practical significance, the problem of establishing the rates of mutations and their dependence on various factors is far from being understood and the values of the rates obtained by different methods vary significantly. In this review, we discuss different ways of estimating the mutation rates and try to explain possible causes of discrepancies in the data obtained. Three levels of mutation formation will be considered: (1) mutations formed in the process of development of a given individual during gametogenesis (basic mutations); (2) mutations transmitted to the offspring and determining differences in genomes of successive generations (parents and their children), which include basic mutations and possible changes emerging during complex sperm transfer processes to the oocyte, fertilization and subsequent events; (3) mutations formed at level 2 and, in addition, are fixed in evolution and determine evolutionary processes and differences between genomes, in particular, hominoids, hominids and hominins.

**Keywords:** *de novo* mutations genome sequencing, divergence, dating, recombination, substitution rate.