

МНОЖЕСТВЕННОСТЬ ВАРИАНТОВ ОРГАНИЗАЦИИ МИТОХОНДРИАЛЬНОГО ГЕНОМА

©2012 г. А. А. КОЛЕСНИКОВ, Е. С. ГЕРАСИМОВ

*Биологический факультет Московского государственного
университета им. М.В.Ломоносова, Москва*

I. Введение. II. Варианты организации мт-генома. III. Редактирование первичных транскриптов мт-ДНК. IV. Заключение.

I. ВВЕДЕНИЕ

Кажущаяся простота организации митохондриального генома очень обманчива. С момента установления первых примеров организации мтДНК молекулы [1–3] круг вопросов, как и почему растет в геометрической прогрессии. Долгое время казалось, что основной формой организации является кольцевая молекула ДНК, практически лишенная ассоциированных с ней белков, а линейные варианты, обнаруженные у представителей Ciliata (*Tetrahymena* и *Paramecium*), – это исключение из общего правила. Однако в настоящее время исследователи столкнулись с большим набором вариантов, и число их продолжает увеличиваться, особенно в последнее десятилетие. Различия касаются как формы организации молекулы, суммарного размера молекулы, так и форм организации «нуклеоида», различающегося набора белков, кодируемых митохондриальным геномом, разнообразия механизмов, привлекаемых к реализации генетической информации. Широкая доступность методов установления первичной структуры полных геномов открыла возможности для анализа огромного количества геномов организмов от простейших до млекопитающих. На начало 2012 г. в GenBank депонировано около 3000 последовательностей полноразмерных митохондриальных геномов, в том числе, 2683 Metazoa,

Принятые сокращения: мт-геном – митохондриальный геном; Кbp –тысяча пар оснований; мРНК – матричная РНК; РД – редактируемый домен, участок первичного транскрипта, требующий редактирования; пан-редактирование – редактирование первичного транскрипта по всей длине молекулы.

Адрес для корреспонденции: aak330@yandex.ru

109 представителей Fungi, 69 – Viridiplantae и 90 – Protista (www.ncbi.nlm.nih.gov/genomes). В данном обзоре мы остановимся на вариантах структурной организации (архитектуре) митохондриального генома, в которую входят следующие характерные особенности: размер и форма молекул ДНК, набор кодируемых генов, наличие криптогенов и редактирование первичных транскриптов.

II. ВАРИАНТЫ ОРГАНИЗАЦИИ МИТОХОНДРИАЛЬНОГО ГЕНОМА. РАЗМЕР И ФОРМА МОЛЕКУЛ ДНК

За более, чем 50-ти летнюю историю изучения мт-генома стало очевидным, что структурная организация может быть представлена большим разнообразием форм. Можно выделить 6 основных вариантов:

- 1) Кольцевая молекула с размером от 11–12 Кbp до 28 Кbp.
- 2) Кольцевая молекула с размером от 22 Кbp до 1000 Кbp.
- 3) Кольцевая молекула более 22 Кbp с одновременным присутствием плазмидоподобных молекул.
- 4) Популяция гетерогенных кольцевых молекул.
- 5) Популяция гомогенных линейных молекул.
- 6) Популяция гетерогенных линейных молекул.

Остановимся более подробно на характеристике вариантов.

Описание вариантов подразумевает все формы молекул, обнаруживаемых в митохондриях.

1) Классическое представление о митохондриальном геноме выглядит так: геном митохондрий клеток большинства животных выглядит однотипно. Это, как правило, кольцевая молекула размером 14–20 Кbp. Обычно такой геном кодирует 2 рибосомных РНК, 13 специфических полипептидов и до 25 тРНК (количество может меняться от 2 до 24 у разных видов). Геномы животных содержат очень незначительный процент некодирующей ДНК в виде единственной контрольной области (однако, встречаются исключения, когда контрольная область приобретает большую протяженность, [4–11]). Интроны в генах белков отсутствуют. Гены тРНК расположены, как правило, между генами белков и, кроме основной функции – транспорта аминокислот при биосинтезе белка, выполняют функцию сигналов («запятых»), по которым происходит процессинг полицистронной пре-мРНК. Иногда встречаются кольцевые мультимерные (чаще всего димерные) молекулы, устроенные по типу «голова–хвост». Примером такого варианта является мт-геном клеток животных (рис 1а).

2) Второй вариант – это также кольцевая молекула, размерные границы которой находятся в пределах 19 – 1000 Кbp. Мт-геном

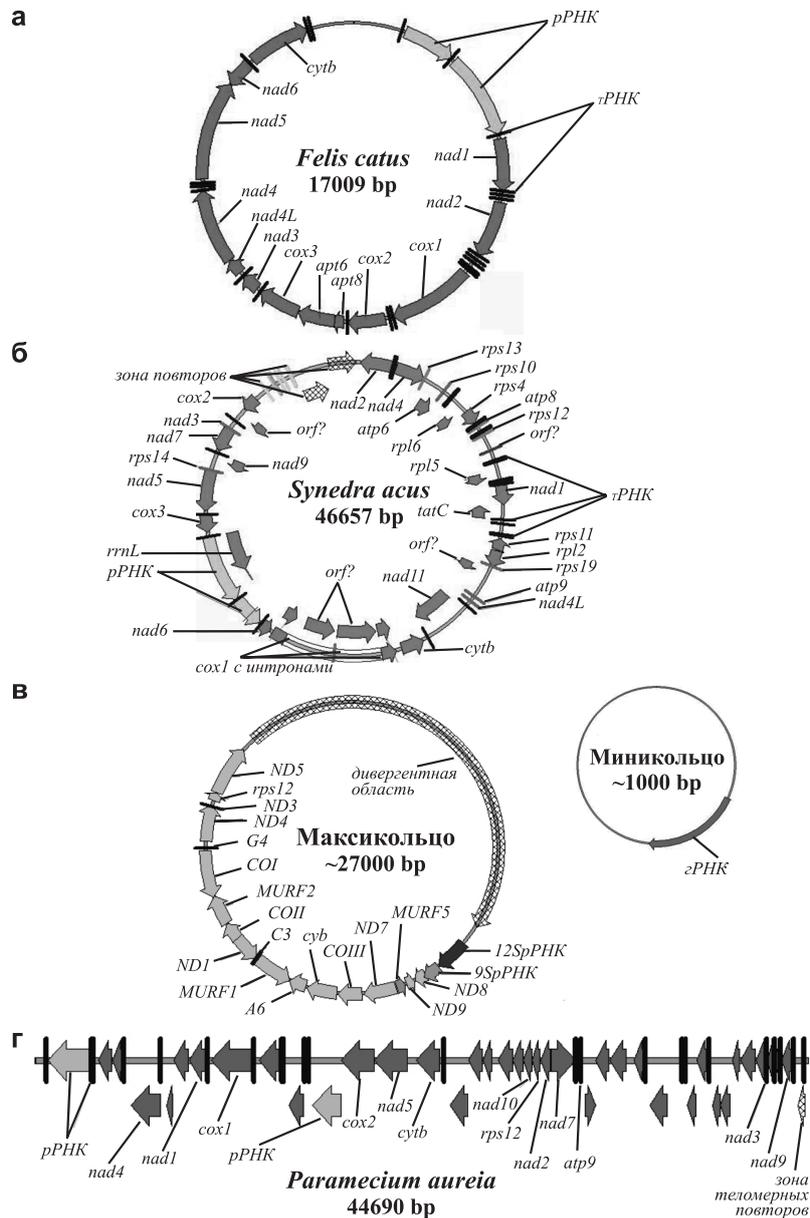


Рис. 1. Примеры структуры митохондриальных геномов 1, 2, 4 и 5 типов.
 1а – *Felis catus*; 1б – *Synedra acus*; 1в – *Trypanosoma brucei*; 1г – *Paramesicium aurelia*. Схемы построены в программе Vector NTI с использованием аннотации из GenBank (NC_001700, GU002153, M94286, NC_001324).

высших растений представлен молекулами от 100 до 1000 Кбр. Отличия от первого варианта касаются в первую очередь наличия межгенных спейсерных последовательностей разной протяженности и присутствием дополнительных генов. Количество идентифицированных генов колеблется от 40 до 156. Структурные гены у некоторых видов содержат интроны I или II типа. Примером этого варианта организации может быть геном диатомовой водоросли *Synedra acus* [12] (рис 1б), геномы зеленых и красных водорослей [13] или представителя воротничковых жгутиковых – *Monosiga brevicollis* [14].

3) Для третьего варианта характерно присутствие помимо кольцевой молекулы (размеры молекулы как во втором варианте от 20 до 1000 Кбр) плазмидоподобных кольцевых или линейных молекул размеры и количество которых может варьировать в широких пределах. Такой вариант выявлен у многих грибов и у высших растений.

4) В четвертом варианте геном представлен несколькими различающимися кольцевыми молекулами в одной митохондрии. В качестве примера сошлемся на мт-геном паразитического примитивного многоклеточного (содержит всего 20–30 клеток) *Dicyemida misakiense* (или *Rhombozoa*). Геном представлен набором миникольцевых ДНК с размером 1–2 Кбр. Одна молекула, как правило, содержит одну рамку считывания. Количество молекул колеблется от 100 до 1000 на клетку. [15]. У представителя диплонемид *Rhynchopus euleeides* выявлены два варианта кольцевых молекул 7,0 и 7,7 Кбр [16].

Подобный вариант выявлен у 6 из 10 видов пухоедов и вшей (Insecta: Phthiraptera). Геном этих видов представлен несколькими вариантами: тип 1 – полноразмерная хромосома с 37 генами и набор гетерогенных миникольцевых молекул; тип 2 – набор миникольцевых молекул, меньших по размеру, чем полная хромосома с одной короткой контрольной областью; тип 3 – набор миникольцевых молекул, содержащих 1–3 гена с протяженной сложной контрольной областью [17].

Уникальная организация мт-генома обнаружена у представителей группы кинетопластид. Представителей Kinetoplastida разделяют на группы Prokinetoplastina (бодониды *Ichthyobodo* и *Perkinsiella*) и Metakinetoplastina (оставшиеся бодониды и трипаносоматиды). В пределах Metakinetoplastina выделяются четыре клада: Neobodonida, Parabodonida, Eubodonida (*Bodo saltans*) и Трипаносоматиды [18]. У Кинетопластид в единственной митохондрии клетки имеется сложный ассоциат, представленный двумя популяциями кольцевых молекул ДНК: миникольцевыми (размером 0,5 – 12 Кбр в количестве до 10^5) и максикольцевыми (20–50 молекул с размером 20–40 Кбр) (рис 1в). Молекулы сцеплены друг с другом по типу катенанов и

формируют сложную структуру, которая встречается в литературе под разными названиями: сеть (network), корзина (basket) или ассоциат. При этом, истинным аналогом мт-генома (кодирующим стандартные митохондриальные белки) являются максикольцевые молекулы. Установленная на сегодняшний момент функция миникольцевых молекул заключается в кодировании малых «гидовых» РНК, которые необходимы для процесса уридилового редактирования множества первичных транскриптов максиколец. Не исключено, что миникольца могут выполнять и какие-то другие функции. У бодонид ассоциат мт-ДНК менее компактный и миникольцевой компонент представлен более протяженными молекулами.

Сюрпризом оказалось присутствие в клетках некоторых двусторчатых моллюсков (в частности, представители рода *Mytilis*) двух различающихся вариантов кольцевого мт-генома: «мужского, М» и «женского, F», наследование которых происходит двумя путями: через яйцеклетку (F) и через сперму (M). В связи с этим появился термин – двуродительское наследование мт-генома (DUI), в отличие от привычного однородительского (материнского) наследования (SMI) [19]. М и F геномы у *Mytilus galloprovincialis* различаются по последовательности нуклеотидов примерно на 20%, хотя набор генов и их расположение в геноме идентичны [20].

5) Для пятого варианта характерны линейные однокомпонентные митохондриальные геномы, обнаруженные в ряде независимых филогенетических линий: у некоторых представителей Ciliata (*Tetrahymena pyriformis* [21] и *Paramecium aurelia* [22] (рис 1г)), Apicomplexa (*Plasmodium* и родственные виды) [23], грибов, зеленых водорослей (*Chlamydomonas* и родственные виды), животных группы Cnidaria [24–26]. Один из наиболее крупных линейных мт-геномов у Ciliata охарактеризован для *Oxitricha trifallax* [29]. Геном представлен линейной хромосомой около 70 Kbp и линейной плазмидой около 5 Kbp.

Линейные молекулы мт-ДНК содержат разнообразные кластеры тандемных повторов [27], выполняющих роль теломеров. Нозек и Томашка [28] систематизировали теломерные последовательности линейных мт-ДНК в виде 6 вариантов. 5 из них формируются собственно последовательностью ДНК: 1) ковалентно замкнутые шпильки на обоих концах (например, у инфузории *Paramecium aurelia*, дрожжей *Pichia pijperii* и *Williopsis* (раньше *Hansenula*) *mrakii*); 2) ковалентно замкнутая шпилька на одном конце и кластер коротких повторов на другом конце молекулы (плазмиды в митохондрии аскомицета *Fusarium*); 3) кластеры тандемных повторов (например, у инфузории *Tetrahymena thermophila*), в некоторых случаях уложен-

ные в теломерные петли (t-петли) (у зеленой водоросли *Chlamydomonas parapsilosis*); 4) инвертированные повторы и длинные одноцепочечные участки на 3'-концах (например, у *Chlamydomonas reinhardtii*); 5) сложные кластеры повторов различных типов, например, у *Theileria parva* (Аpicомплекса), *Amoebidium parasiticum* (Ichthyosporaea), в линейных плаزمидах миксомицета *Physarum polycephalum* (Мухогастриды). Еще один вариант обозначения конца хромосомы определяется белком, который специфически ковалентно прикрепляется к 5'-концам каждой цепи ДНК (например, у линейных митохондриальных плазмид грибов).

6) В шестом варианте геном представлен гетерогенной популяцией линейных молекул. В качестве примера приведем *Amoebidium parasiticum*, геном которого представлен несколькими сотнями линейных молекул (суммарная сложность генома более 200 Kbp) [30]. Этот же вариант обнаружен также у представителя группы Euglenozoa (*E. gracilis*). Общая генетическая сложность генома оценивается в 60 Kbp, но выделяется как набор коротких линейных фрагментов [31, 32]. Возможно, подобная структура генома существует и у других представителей Euglenozoa – Diplonemid и Euglenids [16]. Складывается впечатление, что сложный набор линейных молекул составляет мт-геном динофлагеллят [33, 34]. Уникальная структура мт-генома обнаружена у представителя Cubozoa *Alatina moseri*. 18 генов расположены на 8 линейных хромосомах размером от 2,9 до 4,6 Kbp, причем все хромосомы имеют идентичные теломерные последовательности [26].

Значительное количество работ по оценке структуры мт-генома проводилось с помощью электронной микроскопии. При этом авторы наблюдали структуры типа кольца с длинным линейным хвостом или ветвящиеся. Как правило, такие структуры являются промежуточными продуктами при репликации ДНК по типу катящегося кольца [23]. В этой связи следует заметить, что выделение мт-ДНК целесообразно вести из клеток, находящихся в стадии покоя. В активно делящихся клетках препараты мт-ДНК могут содержать большое количество промежуточных продуктов репликации. В частности, это обстоятельство вызывает оживленную дискуссию о структуре мт-генома у *Saccharomyces cerevisiae* [35].

ЯВЛЕНИЕ ГЕТЕРОПЛАЗМИИ

Гетероплазмия – присутствие в клетке различающихся молекул митохондриальной ДНК. Как правило, речь в данном случае идет об однотипных молекулах, имеющих небольшие отличия. Различают

два варианта гетероплазмии: размерную и сайтовую. Обычно первая является результатом присутствия в контрольной области молекулы разного количества сравнительно коротких повторяющихся последовательностей (10–100 нуклеотидов). Считается, что увеличение количества повторов связано с интенсификацией процесса транскрипции мт-ДНК. Размерную гетероплазмию фиксировали у многих организмов, в частности у рыб, насекомых, летучих мышей и др. [36, 37]. Сайтовая гетероплазмия чаще всего связана с накоплением мутаций [37]. В рамках данного обзора мы не будем подробно останавливаться на этом явлении.

СУЩЕСТВУЕТ ЛИ СПЕЦИФИЧНОСТЬ ФОРМЫ ОРГАНИЗАЦИИ МТ-ГЕНОМА У ОПРЕДЕЛЕННЫХ ГРУПП ОРГАНИЗМОВ?

Существует ли специфичность той или иной формы организации мт-генома у родственных групп организмов? Для хордовых животных характерна только 1 форма митохондриального генома, т.е. компактной кольцевой молекулы. С другой стороны, у беспозвоночных животных мы видим примеры форм 1, 2 и 5. Правда, в этой группе животных преимущественно представлена форма 1, остальные формы встречаются как исключение из правила. Меньшая консервативность наблюдается у растений и водорослей. Скорее здесь можно говорить о тенденциях. Как правило, это – геном формы 2, достаточно компактный (100–500 Кbp у высших растений, 40–90 Кbp у зеленых водорослей, за исключением представителей хлорофитовых водорослей, размер от 15 до 40 Кbp) и имеющий специфические признаки. Например, мт-геном высших растений имеет преимущественно интроны II группы, тогда как у других организмов этот тип интронов встречается в виде исключения. В среднем мт-геном растений кодирует 50–70 белков (без учета неидентифицированных рамок считывания и рамок считывания в интронах). Мт-геном высших растений кодирует 5S рРНК, ген которой не обнаруживается у других организмов. Что касается остальных групп организмов, то сходство наблюдается на уровне семейств. Особенно это относится к представителям простейших. Следует отметить, что чаще всего, значительные отличия в структуре генома наблюдаются у паразитических организмов.

В последнее десятилетие благодаря накоплению огромных биоинформационных массивов данных (включающих морфологию, биохимические характеристики и данные о структуре геномов) была предложена система эукариот, предполагающая группирование всех организмов в 5–8 больших «супергрупп», описывающих дивергенцию эукариот [38–40]. Интересно проанализировать распре-

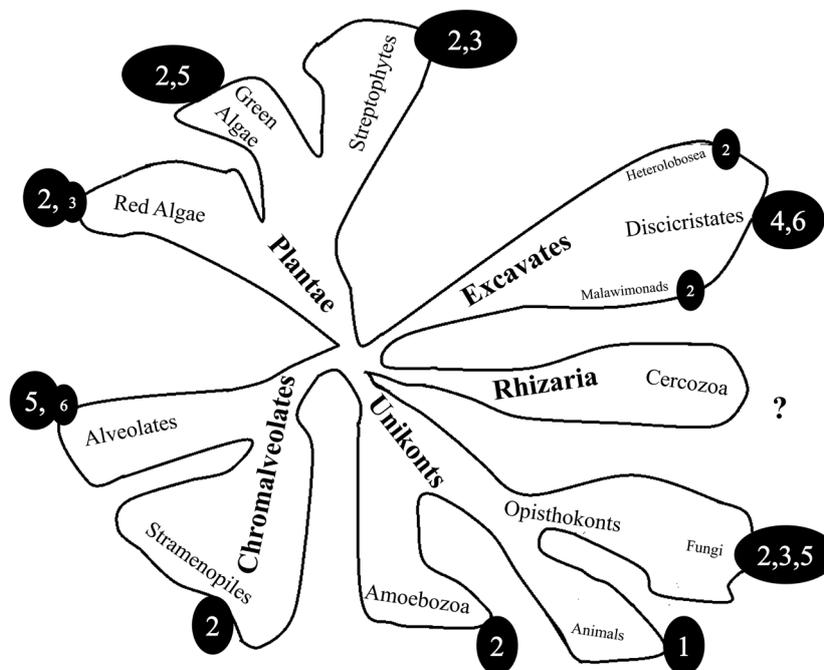


Рис. 2. Схематическое изображение древа эукариот [37].

Выделенным шрифтом обозначены супергруппы. Обычным шрифтом – подгруппы. Цифры в кружках – варианты организации мт-генома.

деление различных вариантов организации мт-генома на древе эукариот. На рис. 2 схематически изображено древо эукариот [37] на котором мы обозначили цифрами варианты организации генома, которые обсуждались выше. Из рассмотрения мы исключили единичные примеры, которые отличаются от мажорных вариантов для «супергрупп». Сразу следует оговориться, что база данных по структуре мт-генома чрезвычайно неравномерна. Если животные представлены тысячами примеров, то представители супергруппы Plantae – сотнями, а остальные группы десятками или даже единицами. Для представителей супергруппы Rhizaria вообще нет информации о полном размере мт-генома ни для одного вида. Поэтому понятно, что рассмотрение распределения форм носит предварительный характер. Для более достоверной оценки следует существенно расширить базу данных. Из рисунка видно, что наиболее распространенным является вариант 2; следующими по частоте встречаемости являются варианты 3 и 5. Варианты 1 и 4 являются высокоспециализированными.

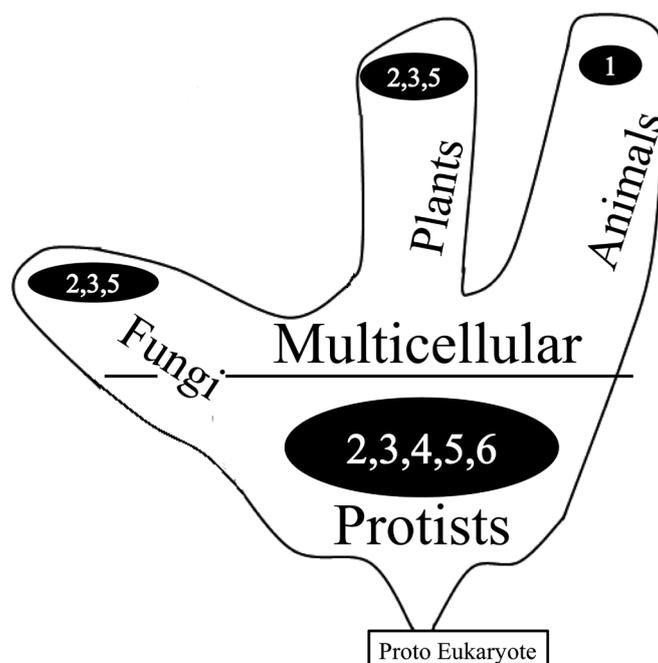


Рис. 3. Встречаемость различных вариантов организации мт-генома у одноклеточных и многоклеточных организмов.

На рис. 3 мы представили распределение вариантов организации мт-генома у одноклеточных и многоклеточных организмов. Из этого рисунка видно, что наибольшее разнообразие характерно для одноклеточных эукариот (Protists).

Аналогичное заключение можно сделать и из анализа набора генов, кодируемых мт-геномом.

МИТОХОНДРИАЛЬНЫЙ НУКЛЕОИД

Структурированность мт-ДНК в митохондрии (нуклеоид) также может быть источником разнообразия мт-генома. К сожалению, прямых экспериментальных исследований структуры нуклеоида очень мало. В последнее десятилетие активно исследовали белки, специфически связывающиеся с мт-ДНК. На сегодняшний день идентифицировано около 30–35 белков, которые могут специфически участвовать в организации нуклеоида [42]. Исследования проводились на ограниченном количестве объектов, включая дрожжи, ооциты лягушки и культуру клеток человека [43–46]. Предложена модель организации

нуклеоида [45], согласно которой существует центральная часть, в которой концентрируются процессы репликации и транскрипции, и периферическая, в которой происходит трансляция. Однако, на настоящий момент ощущается явный недостаток данных для проведения сравнительного анализа организации нуклеоидов у различных групп организмов

НАБОР ГЕНОВ, КОДИРУЕМЫХ МИТОХОНДРИАЛЬНЫМ ГЕНОМОМ

Большие вариации в размере мт-генома, позволили предположить, что это определяется разным количеством генов, кодируемых данным геномом. То, что принято считать типовым набором митохондриальных генов включает набор генов, кодирующих субъединицы комплекса I (NADH: убихинон оксидоредуктазы) *nad* 1–9; комплекса III (убихинон:цитохром с оксидоредуктазы) *cob*; комплекса IV (цитохром с оксидазы) *cox* 1–3; комплекса V (АТФ синтазы) *atp* 1,6,8. Кроме того, в геноме имеются гены рРНК большой и малой субчастиц рибосом и варьирующее количество генов тРНК. Этот набор генов за редким исключением присутствует практически во всех мт-геномах и в частности в геномах первого варианта. Имеющаяся база данных свидетельствует о том, что у многих организмов с другими вариантами организации мт-генома набор генов может значительно превышать этот список. Среди дополнительных генов следует отметить от 3 до 27 генов рибосомных белков (чаще малой субчастицы), гены субъединицы комплекса II (сукцинат:убихинон оксидоредуктазы) *sdh* 2–4 и другие, включая, как правило, уникальные неидентифицированные рамки считывания [47]. Количество дополнительных генов может различаться даже у близкородственных видов. Для сравнения часто используют набор мт-геном *Reclinomonas americana* (97 генов) [48]. Считается, что этот геном ближе всего может напоминать геном про-митохондриального предшественника. Однако, число генов у высших растений может достигать до 156 у *Beta macrocarpa* и *Nicotiana tabacum* [49]. Присутствие дополнительных генов, скорее всего, отражает индивидуальные особенности развития организмов. Что касается генов рибосомных белков, то их сохранение в геноме в процессе эволюции можно объяснить тем, что они могут быть полифункциональными белками и выполнять в митохондриии какие-то дополнительные функции, помимо участия в формировании рибосом [50]. Действительно, по меньшей мере, для 25 рибосомных белков у разных организмов (*E.coli*, *Saccharomyces cerevisiae*, *Drosophila melanogaster*, *Xenopus laevis*, *Ascaris lumbricoides*, *Arabidopsis thaliana*, *Strongilocentrotus purpuratus*, мышь, крыса, человек) показаны дополнительные функции в клетке.

В рамках данного обзора мы посчитали нецелесообразным рассматривать вопросы репликации и транскрипции мт-генома, и остановимся только на процессах редактирования первичных транскриптов.

III. РЕДАКТИРОВАНИЕ ПЕРВИЧНЫХ ТРАНСКРИПТОВ МТ-ДНК

Существенный вклад в оценку разнообразия вариантов организации мт-генома вносит оценка наличия/отсутствия криптогенов. Криптоген – это открытая рамка считывания, первичный транскрипт которой для превращения в полноценную матрицу для синтеза белка должен быть дополнительно изменен (отредактирован). Иногда при определении набора генов, кодируемых мт-геномом сложно идентифицировать ген, пока он находится в виде криптогена. Например, криптоген СОII у *Trypanosoma brucei*, состоящий из 438 нуклеотидов реализуется в результате редактирования в мРНК размером 866 нуклеотидов [41].

Редактирование РНК – термин, охватывающий широкий круг разнообразных процессов модификации РНК почти всех типов: матричных, рибосомных, транспортных, а также малых некодирующих РНК [51, 52]. Исторически первым был открыт процесс редактирования гена СОII мт-генома у двух видов трипаносоматид: *Trypanosoma brucei* и *Crithidia fasciculata* [53]. Данный ген обладает высокой степенью консервативности среди трипаносоматид, что указывает на его функциональность, однако рамка считывания его нарушена. Venne с соавторами [53] показали, что вставка 4 уридиловых оснований в пре-мРНК происходит посттранскрипционно и это восстанавливает нормальную рамку считывания. В последующие годы изучение организмов других систематических типов привело к открытию новых вариантов редактирования.

Скоро стало ясно, что редактирование РНК широко распространено в природе от вирусов [54] до млекопитающих и высших растений. По сути можно выделить три основных типа редактирования РНК: модификация основания, посттранскрипционная вставка/удаление основания и котранскрипционная вставка/удаление основания. [55].

В дальнейшем тексте мы будем пользоваться двумя терминами, которые следует расшифровать. РД – редактируемый домен, участок последовательности пре-мРНК, который подвергается редактированию. Пан-редактирование – процесс редактирования происходит по всей длине пре-мРНК, в отличие от частично редактируемого криптогена.

РЕДАКТИРОВАНИЕ В ЦАРСТВЕ ЖИВОТНЫХ

Первый и пока единственный случай редактирования у многоклеточных животных был продемонстрирован у трематоды *Teratocephalus lirellus*. Обнаружена вставка 6 U в мРНК цитохрома b [84]. Авторы анализировали также мРНК у 12 других представителей нематод и не обнаружили такого феномена (оценка проводилась путем сравнения гена *cyt b* и его мРНК).

РЕДАКТИРОВАНИЕ В МИТОХОНДРИЯХ РАСТЕНИЙ

РНК редактирование в митохондриях цветковых растений заключается в дезаминировании C до U [56–58]. В изученных митохондриальных геномах насчитывается около 400–500 сайтов дезаминирования внутри открытых рамок считывания [59–61]. Дезаминирование осуществляется специализированным ферментом, скорее всего, класса дезаминаз, эволюционно приспособившимся к использованию полинуклеотидов в качестве субстратов [62, 63]. Реакция дезаминирования C→U происходит строго сайт-специфично и в настоящее время выявлены цис-факторы специфичности. Прежде всего, ими оказались последовательности в 20–40 нуклеотидов непосредственно прилегающие с 3'-конца к модифицируемому цитозину. Изменение этих последовательностей влияло на число отредактированных по данному сайту транскриптов. В некоторых случаях значение имели также около 10 нуклеотидов с 5'-стороны мРНК [64, 65]. Биоинформатический анализ этих цис-элементов у *A. thaliana* не выявил какого-либо заметного сходства между ними и при этом не удалось даже сгруппировать эти элементы. Оказалось, что последовательности, окружающие разные редактируемые сайты, совершенно не похожи между собой [59]. Цис-проксимальные элементы, описанные выше, не образовывали консенсус и редко имели хотя бы какие-то общие черты. Попытки выявить транс-факторы специфичности долгое время не приводили к успехам. Поэтому первоначально существовала идея поиска молекул РНК, подобных гРНК трипаносоматид, являющихся транс-факторами, определяющими сайты дезаминирования. Однако, до сих пор подобные поиски не увенчались успехом

Исследования, проведенные *in vitro*, показали, что сайты редактирования узнаются независимо. Обнаруживается целый ряд частично отредактированных интермедиатов, которые сами равновероятно могут служить субстратами для редактирования по оставшимся сайтам. До сих пор не выявлено процесса полярности (т.е. направленности от 3'- к 5'-концу молекулы или наоборот).

Тип редактирования $U \rightarrow C$ не характерен для митохондрий цветковых растений. Во всяком случае не выявлено ни одного случая такого редактирования у *Arabidopsis thaliana*, *Brassica napus* и *Oryza sativa* – наиболее изученных видах [59–61]. Зато этот тип редактирования распространен в митохондриях роголистника, где достигает частоты, сопоставимой с $C \rightarrow U$. У других мохообразных, как и у представителей всех порядков печёночников, кроме Marchantiales, распространены оба типа редактирования ($C \rightarrow U$, $U \rightarrow C$).

Единственные наземные растения, для которых не показано редактирования – печёночники порядка Marchantiales, например, *Marchantia polymorpha* [66]. На сегодняшний день редактирования не найдено среди ближайших родственником наземных растений – зеленых водорослей [67]

РЕДАКТИРОВАНИЕ У МИКСОМИЦЕТОВ

О редактировании у миксомицетов известно достаточно давно. Масштабное редактирование COI транскрипта первой субъединицы цитохром оксидазы показано Готт и соавт. у *Physarum polycephalum* [68]. В транскрипте этого гена происходит 66 вставок оснований и не менее 4 конверсий. В процессе редактирования используются 4 разных по механизму типа редактирования: динуклеотидная вставка UU (и, возможно, CU), конверсия $C \rightarrow U$, а также моонуклеотидные вставки C и U. Вставочные типы редактирования происходят котранскрипционно [69]. Изучение возможных путей эволюции разных типов редактирования COI привело к пониманию, что эти четыре типа редактирования имеют свою уникальную историю [70, 71].

Редактирование путем вставки U присутствует у всех миксомицетов, что указывает на раннее возникновение этого типа редактирования у их общего предка. Вставки C отсутствуют у филогенетически рано ответвившегося *Clastoderma debaryanum* и единственная вставка C наблюдается у *Arcyria cinerea*. У более поздно ответвившихся на филогенетическом дереве групп обнаружено около 30–40 вставок C. Эти данные указывают на то, что механизм C вставки возникает внутри таксона миксомицетов где-то в районе ответвившегося *A. cinerea* и, по-видимому, развивается в более поздних группах [72]. Редактирование по типу динуклеотидной вставки возникает только в пределах одной филогенетической ветви миксомицетов и отсутствуют в других ветвях. Надо отметить, что механизм динуклеотидной вставки не сводится к последовательной вставке двух моно-U [73, 74].

Конверсия $C \rightarrow U$, судя по филогенетическим данным, либо несколько раз возникала, либо несколько раз утрачивалась разными

представителями миксомицетов [75]. Такое разнообразие эволюционных путей разных систем редактирования, обеспечивающих функционирование одного гена, указывает на необычайно высокую степень пластичности систем редактирования [72].

Еще около 500 сайтов редактирования найдено и около 500 сайтов предсказано в митохондриальном геноме *Physarum polycephalum* [76]. В случае редактирования по типу динуклеотидной вставки обнаружены варианты AA, UU, GU, UA, GC, CU вставок. Уникальный для миксомицетов тип редактирования обнаружен для транскрипта гена *nad2*: удаление 4 остатков аденина [77]. Гены митохондриальных рибосомных белков у *Didymium iridis* (*rpS12*, *rpS7*, *rpL2*, *rpS19*, *rpS3* и *rpL16*) подвергаются интенсивному редактированию с использованием тех же механизмов, что и транскрипт гена COI [78].

Редактированию двух типов подвергаются также митохондриальные тРНК *Physarum polycephalum*. Две метиониновые тРНК ($tRNA^{met1}$, $tRNA^{met2}$) имеют неспаренные основания в акцепторном стволе, устраняемые посттранскрипционной вставкой единственного G. Это первый известный и пока единственный случай посттранскрипционной вставки G [79]. Таким образом, система редактирования митохондриальных транскриптов миксомицетов чрезвычайно усложнена. Она включает в себя несколько механизмов, сочетает как котранскрипционное, так и посттранскрипционное редактирование. Редактированию подвергаются разные типы молекул РНК: мРНК белков дыхательной цепи, мРНК рибосомных белков и тРНК.

Описав таксоны живой природы, для которых характерно C→U редактирование, мы бы хотели подчеркнуть, что эта форма редактирования наиболее лабильна. Независимо и в разных вариантах она возникает у трипаносоматид (редактирование тРНК Trp [80]); в митохондриях и хлоропластах высших растений [81]; в случае аполипротеина В (ядерного генома млекопитающих); и у клеточного слизевика *Dictyostelium discoideum*, [82].

Редактирование у *Dinoflagellates*

У трех представителей динофлагеллят было обнаружено обширное редактирование в генах *cox1* и *cob*. Редактирование по типу замен нуклеотидов (А на G, U на C, C на U) затрагивает около 2% кодирующей последовательности генов [83]. Редактирование затрагивает нуклеотиды в первой и втором положениях кодона, изменяя кодируемую аминокислоту.

Механизм уридилowego редактирования у трипаносоматид

Наиболее детально на сегодняшний день изучен исторически открытый первый процесс редактирования у трипаносоматид. Ниже мы опишем более подробно, как реализуется генетическая информация кинетопластного генома (митохондриального) генома трипаносоматид, группы одноклеточных, для большинства представителей которой характерен паразитический образ жизни. Мы покажем тесную взаимосвязь редактирования со структурной организацией молекул ДНК и проследим некоторые пути эволюции криптогенов.

Отличительной чертой редактирования в митохондриях трипаносоматид является участие в этом процессе специализированных молекул РНК – гидовых РНК (гРНК). Кратко процесс редактирования можно описать так. Молекула гРНК комплементарно связывается 5'- и 3'-концевыми участками с молекулой пре-мРНК, затем мультиферментный комплекс катализирует последовательно разрезание пре-мРНК, добавление/удаление уридиловых оснований до установления полной комплементарности с центральным участком гРНК и лигирование пре-мРНК [85, 86].

Первыми были идентифицированы гРНК, кодирующиеся в межгенных и прилегающих к дивергентной области участках максиколяца [87,88], однако затем стало ясно, что основная масса гРНК кодируется миниколяцами [89, 90]. На сегодняшний день это единственная доказанная генетическая функция миниколец [91]. Миниколяца могут содержать 1–4 гена гРНК, каждый из которых фланкирован короткими инвертированными повторами. Гены гРНК транскрибируются в составе полицистрона, который затем процессируется на отдельные гРНК 19S РНК процессирующим комплексом [92]. Уникальной является гРНК максиколяцевого кодирования gMURF2-II, ген которой расположен внутри кодирующей области гена ND4 [93]. Это единственный на сегодня известный случай внутригенного кодирования функциональной гРНК. Несмотря на то, что транскрипция максиколяцевых генов идет полицистронно, гены гРНК максиколяцевого кодирования являются индивидуальными транскрипционными единицами [93]. У *L. tarentolae* альтернативное редактирование было выявлено для нескольких генов, например, COIII, RPS12 и ND3 [94]. Пример альтернативного редактирования известен и у *T. brucei*: одна из гРНК приводит к образованию альтернативной формы мРНК COIII, которая транслируется в альтернативную форму белка [95]. Таким образом, альтернативное редактирование может быть распространенным, хотя пока и малоизученным, механизмом генерации разнообразия белков в кинетопласте.

Рассмотрим более подробно структуру гРНК. гРНК – короткие (около 60 п.н.) молекулы РНК. Функционально можно выделить три участка молекулы гРНК [88]:

– 5'-концевой участок, называемый также «якорным» в ряде работ (anchor), участвует в закреплении гРНК на пре-мРНК перед сайтом редактирования. Эта часть обычно длиной около 10 п.н. полностью комплементарна пре-мРНК редактируемого криптогена.

– Информационный участок гРНК, начинающийся с первого некомплемментарного нуклеотида. Данный участок несет информацию о количестве вставляемых или удаляемых уридилловых оснований. В процессе редактирования за счет вставки или удаления уридиллов в пре-мРНК восстанавливается комплементарность в районе информационного участка с гРНК. Такой участок называется блоком редактирования. В образовании дуплекса мРНК-гРНК в районе редактируемого блока принимают участие как канонические Ватсон-Криковские пары, так и неканонические G:U пары [96].

– 3'-концевой участок, состоящий из уридилловых оснований, добавляемых к гРНК посттрансляционно и имеющих нефиксированную длину. Добавление уридиллов осуществляет специальный фермент KRET1, терминальная трансфераза уридилловых нуклеотидов, отличная от аналогичного фермента, участвующего в добавлении уридиллов к пре-мРНК [97, 98]. Он участвует в стабилизации дуплекса пре-мРНК и гРНК в процессе редактирования, взаимодействуя с пурин-богатой (обычно, G-богатой) последовательностью, предшествующей редактируемому блоку. В клетках, где экспрессия KRET1 подавлялась РНК-интерференцией, редактирование не происходило [97].

гРНК обеспечивает только передачу информации, но не выполняют никаких каталитических функций в процессе редактирования. Катализ реакций осуществляется белковым комплексом, обладающим множеством активностей [85]. Больших успехов в изучении процесса редактирования удалось достичь, когда появились первые *in vitro* системы, в которых осуществлялось редактирование. Остановимся на основных этапах механизма редактирования.

На первом этапе гРНК «узнает» за счет комплементарного взаимодействия 5'-конца редактируемую пре-мРНК и образует с ней дуплекс. Последние исследования свидетельствуют, что ключевую роль играет формирование на данном этапе так называемой трехспиральной структуры гРНК:пре-мРНК (имеется ввиду формирование трех двухспиральных участков) [99]. Первая спираль формируется якорным участком гРНК и пре-мРНК, вторую спираль образуют основания информационного участка гРНК, а третья возникает в результате

взаимодействия полиU участка гРНК и пре-мРНК. Такая структура дуплекса гРНК-пре-мРНК, стабилизированная белками, является основой для сборки редактирующего комплекса. Происходит разрезание пре-мРНК в месте, определяемом первым неспаренным нуклеотидом информационной части гРНК. В результате образуются 5'- и 3'-фрагменты пре-мРНК. гРНК удерживает их вместе, взаимодействуя якорным участком с 3'-фрагментом, а полиU участком с 5'-фрагментом. Далее в зависимости от структуры информационного участка гРНК происходит вставка или удаление уридиллов с 3'-конца 5'-фрагмента пре-мРНК [100]. Следует отметить, что не найдено гРНК, осуществляющей исключительно удаление уридиллов. Дело в том, что такая гРНК должна была бы быть короче пре-мРНК, и в дуплексе было бы невозможно образование трехспиральной структуры, одну из спиралей которой образует информационный участок гРНК, некомплементарный пре-мРНК до вставки уридиллов. Как правило, удаление уридиллов обеспечивается за счет части информационного участка гРНК, в то время как оставшаяся его часть образует трехспиральную структуру и отвечает за вставку уридиллов. Благодаря этому любая гРНК всегда образует трехспиральную структуру, необходимую для сборки редактирующего комплекса.

Источником уридиллов для вставки U служит свободный UTP [101]. Количество уридиллов определяется количеством А или G в информационном участке гРНК, так как с этими основаниями будут образованы комплементарные пары. Вставка заканчивается, когда гРНК становится полностью комплементарной редактируемому ей участку пре-мРНК [102]. Удаление уридиллов происходит, когда 3'-концевые уридиллы 5'-фрагмента выпетливаются из образуемого дуплекса. Удаление также заканчивается восстановлением полной комплементарности гРНК и пре-мРНК [96]. Таким образом, независимо от того, происходит ли вставка или удаление уридиллов, гРНК привносит генетическую информацию в мРНК. Затем происходит лигирование 5'-и 3'-фрагментов пре-мРНК [100].

БЕЛКИ RECC, 20S КОМПЛЕКС, МОДЕЛЬ ФЕРМЕНТАТИВНОГО КАСКАДА

Общепринятое понимание последовательности реакций в процессе редактирования отражает модель ферментного каскада [103]. Разрезание пре-мРНК осуществляется специальной эндонуклеазой. При делеционном типе редактирования выпетливающиеся уридиллы удаляются с 3'-конца 5'-фрагмента специальной экзонуклеазой, а при инсерционном типе редактирования добавляются терминальной

трансферазой уридилowych нуклеотидов – KRET2. Затем фрагменты пре-мРНК сшиваются лигазой – KREL1-2. Все эти активности обнаруживаются в составе общего белкового комплекса – эдитосомы [104].

Первоначально в ряде работ был выделен и охарактеризован так называемый 20S комплекс. У ряда авторов подобный набор белков получил название L-комплекса [83, 84, 90]. Далее мы будем чаще использовать для упрощения изложения название «20S комплекс». Этот комплекс включает перечисленные выше ферментативные активности, которые соосаждаются в градиенте глицерола на уровне пика 20S у *Trypanosoma brucei* [105] и 25S для *Leishmania tarentolae* [106]. Данный комплекс способен осуществлять только один раунд редактирования *in vitro*, то есть осуществлять вставку или удаление уридилых только в первом редактируемом блоке одной гРНК. Поэтому, по мнению ряда авторов, 20S комплекс, скорее всего, представляет собой только часть реальной действующей эдитосомы или наоборот, в пик попадают сопутствующие клеточные активности, не являющиеся *in vivo* частью редактирующего аппарата [107]. До сих пор обсуждается вопрос о возможных белках-кандидатах на роль экзонуклеазы, так как подобной активностью обладает целый ряд белков 20S комплекса.

Показано существование 3 типов редактирующих основных (core) комплексов [108, 109]: инсерционный субкомплекс (комплекс получил название RECC2 – RNA Editing Core Complex 2), делеционный субкомплекс (комплекс RECC1) и субкомплекс, осуществляющий редактирование с помощью цис-гРНК (RECC3). Получена трехмерная реконструкция данных комплексов и установлен порядок взаимодействия входящих в данные комплексы белковых молекул. Предположительно, все типы субкомплексов входят в состав общей 20S эдитосомы. Согласно альтернативной гипотезе в клетке существует три типа эдитосом, соответствующих трем типам субкомплексов

Функциональную роль ряда белков, входящих в состав эдитосомы, недавно удалось установить благодаря технике нокаута генов и подавления экспрессии с помощью РНК-интерференции.

ЭВОЛЮЦИЯ РЕДАКТИРУЕМЫХ ДОМЕНОВ КРИПТОГЕНОВ ТРИПАНОСОМАТИД

Следует отметить, что присутствие того или иного гена в форме криптогена придется исследовать для каждого индивидуального организма, даже в рамках одного таксона. Два нижеследующих раздела иллюстрируют эту ситуацию.

Структура криптогенов представителей бодонид (ближайших родственников трипаносоматид) на сегодняшний день изучена

очень фрагментарно и слабо. У *Trypanoplasma borreli* редактированию подвергаются гены *cyb* и *COI*. Оба криптогена имеют два редактируемых домена: один на 5'- и другой на 3'-конце, и каждый редактируется независимо. Это напоминает ситуацию с криптогеном ND7 трипаносоматид. Ген *COII*, редактируемый у трипаносоматид вставкой 4 U, у бодонид *T. borreli* и *Criptobia helicis* [110] не редактируется, и цис-гРНК в участке после *COII* не обнаруживается [111]. У бодониды *Bodo saltans*, более близкого к трипаносоматидам вида, *COII* и *MURF2* редактируются в позициях, сходных с трипаносоматидами. В гене *COII* два из 4 вставляемых у трипаносоматид U закодированы на уровне ДНК. Цис-гРНК гена *COII* обнаруживается непосредственно за геном в том же положении, как у трипаносоматид. Неожиданно у *B. saltans* транскрипт ND5 редактируется по всей длине. Напомним, что у трипаносоматид этот ген никогда не является криптогеном [110]. Эти данные показывают, что редактирование явно предшествовало переходу трипаносом к паразитическому образу жизни, а эволюция криптогенов, по-видимому, шла одновременно по путям утраты редактирования одними генами (ND5) и перехода к пан-редактированию других.

РЕТРОПОЗИЦИОННАЯ МОДЕЛЬ УМЕНЬШЕНИЯ РАЗМЕРА РЕДАКТИРУЕМОГО ДОМЕНА

Как обсуждалось выше, редактирование в пределах РД носит полярный характер и продвигается по молекуле пре-мРНК от 3'-конца к 5'-концу. Приведенное выше сравнение структур РД гомологичных криптогенов разных видов трипаносоматид показало, что РД криптогена может прогрессивно уменьшаться в длине, сокращаясь к 5'-концу, вплоть до РД, состоящего из единственного редактируемого блока (редактируемого одной гРНК, например, оба домена ND7 *L. tarentolae*). Как было сказано выше, криптоген A6 у *T. brucei* пан-редактируется с участием 8–10 гРНК [112], у *L. tarentolae* и других лейшманий он 5'-редактируется с участием около 4–6 гРНК [113]. Этот же ген редактируется четырьмя гРНК у *Crithidia fasciculata* и одной гРНК у *Angomonas culicis* [114]. Этот факт привел к появлению модели, объясняющей сокращение длины РД путем ретропозиции частично отредактированной матрицы в максикольцо [114–116]. Таким образом, наиболее древние пан-редактируемые криптогены могли замещаться на частично отредактированные [114, 116]. Наличие в геноме трипаносомид гена, а в клеточных экстрактах активности обратной транскриптазы [118, 119] определяют теоретическую возможность этого процесса. Сходство 5'-редактируемой формы гена с

частично отредактированной мРНК пан-редактируемого гомолога из других видов согласуется с такой моделью. В то же время в кинето-пластном ассоциате находится около 50 копий максиколецевого генома, тогда как ретропозиция, по-видимому, является крайне редким явлением и происходит лишь в одном максиколеце одной клетки. Предполагается, что в основе закрепления произошедшей в одном максиколеце ретропозиции может быть явление потери части типов миниколец в ходе стохастической сегрегации их при делении клетки [114]. Таким образом, в культурах, утрачивающих часть типов миниколец, а, следовательно, гРНК, утрачивается и возможность редактирования соответствующих транскриптов. В таких случаях преимущество получают клетки, имеющие максиколецо с произошедшей ретропозицией на частично отредактированный криптоген. Таким образом, ретропозицию можно рассматривать как мутацию, компенсирующую утрату определенных типов миниколец.

Диксенные трипаносоматиды на одной из стадий жизненного цикла, являются облигатными анаэробами с неактивной дыхательной цепью митохондрий. В этих условиях отсутствие нужных классов миниколец не будет проявляться, пока не наступят аэробные условия. Можно предположить, что среди моноксенных трипаносоматид может наблюдаться тенденция к сокращению длины РД у ряда генов.

Суммируя сказанное в этом разделе можно отметить, что варианты и механизмы редактирования первичных транскриптов кардинально различаются. Можно предположить, что типы редактирования возникали и эволюционировали независимо для каждого варианта. Возможно, феномен редактирования не ограничивается описанными процессами. Это связано с тем, что для установления факта редактирования матрицы необходима детальная информация о первичной структуре гена и отредактированной матрицы, а накопление информации о структурах генов явно опережает формирование базы данных о структуре их транскриптов.

IV. ЗАКЛЮЧЕНИЕ

Накопленный к настоящему времени экспериментальный материал показывает, что митохондриальный геном может быть организован в клетке чрезвычайно разнообразно. В первую очередь это относится к форме, собственно, митохондриальной ДНК (т.е. ее архитектуре). В отношении формы и размера молекулы можно выделить как минимум 6 базовых вариантов. Наибольшее разнообразие форм существования мт-генома характерно для представителей царства

Protista, что отражает многообразие условий существования этих организмов. С усложнением организма, которое происходит параллельно с функциональным упрощением каждой отдельной его клетки (специализацией), разнообразие форм организации мт-генома уменьшается. Складывается впечатление, что кольцевая форма, в том или ином варианте, оказывается более предпочтительной, чем линейная. Абсолютное большинство многоклеточных организмов имеет кольцевые митохондриальные геномы.

Вторым важным фактором является размер молекулы. Логично, что увеличение размера позволяет клетке кодировать больше белков. Однако, не обнаружено пропорциональности между увеличением размера ДНК и количеством генов, при этом значительная часть ДНК оказывается занятой межгенными спейсерами. Возможно, что это определяется какими-то другими задачами, которые возлагаются на эти последовательности. Возможно, эти участки выполняют какие-то регуляторные функции и взаимодействуют с белками, формирующими нуклеоид митохондрии или выполняют иные функции. Количество молекул мт-ДНК в клетке исчисляется сотнями и, понятно, что они должны быть структурированы в митохондрии. К сожалению, вопрос организации мт-нуклеоида остается недостаточным выясненным. Предложена модель «слоистой» структуры [45], в которой предлагаются разные варианты распределения функциональных зон внутри нуклеоида. Выявляется значительное количество белков, которые способны взаимодействовать с мт-ДНК, но их участие в формировании нуклеоида, в обеспечении функционирования генома в системе *in vivo* у разных организмов пока далеко не ясно. Множественность молекул ДНК в митохондрии ставит вопрос об идентичности этих ДНК. Известно явление гетероплазмии. Кроме того, известно, что в процессе онтогенеза могут происходить сложные структурные перестройки, в результате чего в одном и том же организме могут выявляться разные варианты мт-генома.

Для мт-генома характерно большое разнообразие механизмов реализации генетической информации и процессов экспрессии генома, описанный нами во второй части обзора. Редактирование транскриптов оказывается чрезвычайно распространенным способом регуляции и генерации разнообразия в геноме митохондрий. Для разных групп организмов характерны различные формы реализации механизма редактирования, а также набор генов (криптогенов), транскрипты которых вовлечены в этот процесс. Подчас лишь половина, а то и меньшая часть генетической информации остается закодированной в криптогене, а недостающая часть привносится в матрицу из других

локусов молекулы или других молекул, что позволяет организовать геном, в котором информация распределена между абсолютно разными компартментами (например, максиколецями и миниколецями кинетопласта трипаносоматид). Такая дополнительная структурная сложность организации, вероятно, обуславливается тем, что помимо канонической функции выработки энергии, митохондрии способны вовлекаться в специфические для вида метаболические и регуляторные пути и соответствовать потребностям сложного жизненного цикла, процессам деления и реорганизации клетки.

В заключении следует отметить, что многие экспериментальные данные получены на модельных объектах. Поэтому, учитывая большое разнообразие уже известных вариантов мт-генома, следует с большой осторожностью экстраполировать их на все организмы. Так, например, в последнее время за счет внедрения технологий секвенирования транскриптомов становится все более очевидным, что редактирование транскриптов является куда более распространенным феноменом, и скорее правилом, чем исключением в геномах органелл.

ЛИТЕРАТУРА

1. Steinert, G., Firket, H., Steinert, M. (1958) *Exp. Cell Res.*, **15**, 632–635.
2. Ris, H., Plaut, W. (1962) *J. Cell Biol.*, **13**, 383–391.
3. Nass, S., Nass, M.M.K. (1963) *J. Cell Biol.*, **19**, 593–611.
4. Boyce, T.M., Zwick, M.E., Aquadro, C.F. (1989) *Genetics*, **123**, 825–836.
5. Okimoto, R., Macfarlane, J.L., Clary, D.O., Wolstenholme, D.R. (1992) *Genetics*, **130**, 471–498.
6. Azevedo, J., Hyman, B. (1993) *Genetics*, **133**, 933–942.
7. Fuller, K.M., Zouros, E. (1993) *Curr. Genet.*, **23**, 365–369.
8. Boore, J.L. (1999) *Nucleic Acids Res.*, **27**, 1767–1780.
9. Clayton, D.A. (2000) *Exp. Cell Res.*, **255**, 4–9.
10. Delarbre, C., Rasmussen, A.-S., Amason, U., Gachelin, G. (2001) *J. Mol. Evol.*, **53**, 634–641.
11. Hu, M., Gasser, R.B. (2006) *Trends Parasitol.*, **22**, 78–84.
12. Ravin, N.V., Galachyants, Y.P., Mardanov, A.V., Beletsky, A.V., Petrova, D.P., Sherbakova, T.A., Zakharova, Y.R., Likhoshway, Y.V., Skryabin, K.G., Grachev, M.A., (2010) *Curr. Genet.*, **56**, 215–223.
13. Одинцова М.С., Юрина Н.П. (2002) *Генетика*, **38**, 773–788.
14. Lang, B.F., O'Kelly, C., Nerad, T., Gray, M.W., Burger, G. (2002) *Curr. Biol.*, **12**, 1773–1778.
15. Watanabe, K.I., Bessho, Y., Kawasaki, M., Hori, H. (1999) *J. Mol. Biol.*, **286**, 645–650.
16. Roy, J., Faktorova, D., Lukes, J., Burger, G. (2007) *Protist*, **158**, 385–396.
17. Cameron, S.L., Yoshikawa, K., Mizukoshi, A., Whiting, M.F., Lohson, K.P. (2011) *BMC Genomics*, **12**, 394.
18. Moreira, D., Lopez-Garcia, P., Vickers, K. (2004) *Int. J. Syst. Evol. Microbiol.*, **54**, 1861–1875.

19. *Theologidis, I., Fodelianakis, S., Gaspar, M.B., Zouros, E.* (2008) *Evolution*, **62**, 959–970.
20. *Mizi, A., Zouros, E., Moschonas, N., Rodakis, G.C.* (2005) *Mol. Biol. Evol.*, **22**, 952–967.
21. *Suyama, Y., Miura, K.* (1968) *Proc. Nat. Acad. Sci. USA*, **60**, 235–242.
22. *Goddard, J.M., Cummings, D.J.* (1977) *J. Mol. Biol.*, **109**, 327–344.
23. *Wilson, R.J.M., and Williamson, D.H.* (1997) *Microbiology and Molecular biology Reviews*, **61**, 1–16.
24. *Bridge, D., Cunningham, C.W., Schierwater, B., DeSalle, R., Buss, L.W.* (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 8750–8753.
25. *Kayal, E., Bentlage, B., Collins, A.G., Kayal, M., Pirro, S., Lavrov, D.V.* (2012) *Genome Biol. Evol.*, **4**, 1–12.
26. *Smith, D.R., Kayal, E., Yanaqihara, A.A., Collins, A.G., Pirro, S., Keeling, P.J.* (2012) *Genome Biol. Evol.*, **4**, 52–58.
27. *Morin, G.B., Cech, T.R.* (1988) *Cell*, **52**, 367–374.
28. *Nosek, J., and Tomaska, L.* (2003) *Curr. Genet.*, **44**, 73–84.
29. *Swart, E.C., Nowacki, M., Shum, J., Stiles, H., Higgins, B.P., Doak, T.G., Schotanus, K., Magrini, V.J., Minx, P., Mardis, E.R., Landweber, L.F.* (2012) *Genome Biol. Evol.*, **4**, 136–154.
30. *Burger, G., Forget, L., Zhu Yun, Gray, M.W., Lang B.F.* (2003) *Proc. Natl. Acad. Sci. USA*, **100**, 892–897.
31. *Yasuhira, S., Simpson, L.* (1997) *J. Mol. Evol.*, **44**, 341–347.
32. *Spencer, D.F., Gray, M.W.* (2010) *Mol. Gen. Genomics*, **285**, 19–31.
33. *Nash, E.A., Nisbet, R.E.R., Barbrook, A.C., Howe, C.J.* (2008) *Trends in Genetics*, **24**, 328–335.
34. *Jackson, C.J., Norman, J.T., Schnare, M.N., Gray, M.W., Keeling, P.J., Waller, R.F.* (2007) *BMC Biology*, **5**:41 doi: 10.1186/1741-7007-5-41.
35. *Bendich, A.J.* (2010) *Molecular Cell*, **39**, 831–832.
36. *Arnason, E., Rand, D.M.* (1992) *Genetics* **132**, 211–220.
37. *Sato, A., Endo, H., Umetsu, K., Sone, H., Yanagisawa, Y., Saigusa, A., Aita, S., Kagawa, Y.* (2003) *Bioscience reports*, **23**, 313–337.
38. *Cavalier-Smith, T.* (2003) *Eur. J. Protistol.*, **39**, 338–348.
39. *Baldauf, S.L.* (2003) *Science*, **300**, 1703–1706.
40. *Keeling, P.J., Burger, G., Durnford, D.G., Lang, B.F., et al.* (2005) *Trends in Ecology and Evolution*, **20**, 670–676.
41. *Feagin, J.E., Abraham, J.M., and Stuart, K.* (1988) *Cell*, **53**, 413–422.
42. *Wang, Y., Bogenhagen, D.F.* (2006) *J. Biol. Chem.*, **281**, 25791–25802.
43. *Bogenhagen, D.F., Wang, Y., Shen, E.L., Kobayashi, R.* (2003) *Mol. Cell Proteomics*, **2**, 1205–1216.
44. *Holt, I.J., He, J., Mao, C.-C., Boyd-Kirkup, J.D., et al.* (2007) *Mitochondrion*, **7**, 311–321.
45. *Bogenhagen, D.F., Rousseau, D., Burke, S.* (2008) *J. Biol. Chem.*, **283**, 3665–3675.
46. *Miyakawa, I., Okamuro, A., Kinski, S., Visacka, K., Tomashka, L., Nosek, J.* (2009) *Microbiology*, **155**, 1558–1568.
47. *Gray, M.W.* (1999) *Current Opinion in Genetics and Development*, **9**, 678–687.
48. *Lang, B.F., Burger, G., O'Kelly, C.J., Cedergren, R., Golding, G.B., Lemieux, C., Sankoff, D., Turmel, M., and Gray, M.W.* (1997) *Nature*, **387**, 493–497.
49. *Sugiyama, Y., Watase, Y., Nagase, M., Makita, N., Yagura, S., Hirai, A., and Sugiura, M.* (2005) *Mol. Genet. Genomics*, **272**, 603–615.
50. *Warner, J.R., McIntosh K.B.* (2009) *Molecular Cell*, **34**, 3–11.
51. *Gott, J.M., Emeson, R.B.* (2000) *Annu Rev. Genet.*, **34**, 499–531.

52. Ohman, M. (2007) *Biochimie*, **89**, 171–1176.
53. Benne, R., Van den Burg, J., Brakenhoff, J., Sloof, P., Van Boom, J., Tromp, M. (1986) *Cell*, **46**, 819–826.
54. Vidal, S., Curran, J., Kolakofsky, D. (1990) *EMBO J.*, **9**, 2017–2022.
55. Chateigner-Boutin, A.-L., Small, I. (2011) *WIREs RNA*, **2**, 493–506.
56. Covello, P.S., Gray, M.W. (1989) *Nature*, **341**, 662–666.
57. Gualberto, J.M., Lamattina, L., Bonnard, G., Weil, J.-H., Grienenberger, J.M. (1989) *Nature*, **341**, 660–662.
58. Hiesel, R., Wissinger, B., Schuster, W., Brennicke, A. (1989) *Science*, **246**, 1632–1634.
59. Giege, P., Brennicke, A. (1999) *Proc. Natl. Acad. Sci. USA*, **96**, 15324–15329.
60. Handa, H. (2003). *Nucleic Acids Res.*, **31**, 5907–5916.
61. Notsu, Y., Masood, S., Nishikawa, T., Kubo, N., Akiduki, G., Nakazono, M., Hirai, A., Kadowaki, K. (2002) *Mol. Genet. Genomics*, **268**, 434–445.
62. Faivre-Nitschke, S.E., Grienenberger, J.M., Gualberto, J.M. (1999) *Eur. J. Biochem.*, **263**, 896–903.
63. Takenaka, M., Brennicke, A. (2003) *J. Biol. Chem.*, **278**, 47526–47533.
64. Lippok, B., Wissinger, B., Brennicke, A. (1994) *Mol. Gen. Genet.*, **243**, 39–46.
65. Kubo, N., Kadowaki, K. (1997) *FEBS Lett.*, **413**, 40–44.
66. Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N., Akashi, K., Kanegae, T., Ogura, Y., Kohchi, T., Ohyama, K. (1992) *J. Mol. Biol.*, **223**, 1–7.
67. Turmel, M., Otis, C., Lemieux, C. (2003) *Plant Cell*, **15**, 1888–1903.
68. Gott, J.M., Visomirski, L.M., Hunter, J.L. (1993) *J. Biol. Chem.*, **268**, 25483–25486.
69. Gott, J.M., Rhee, A.C. (2008) In: *RNA editing* / Gott, J.M., Rhee, A.C. (eds) / Springer, Berlin, 85–104.
70. Visomirski-Robic, L.M., Gott, J.M. (1995) *RNA*, **1**, 681–691.
71. Wang, S.S., Mahendran, R., Miller, D.L. (1999) *J. Biol. Chem.*, **274**, 2725–2731.
72. Horton, T., Landweber, L. (2000) *RNA*, **6**, 1339–1346.
73. Visomirski-Robic, L.M., Gott, J.M. (1997) *RNA*, **3**, 821–837.
74. Visomirski-Robic, L.M., Gott, J.M. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 4324–4329.
75. Traphagen, S.J., Dimarco, M.J., Silliker, M.E. (2010) *RNA*, **16**, 828–838.
76. Beargie, C., Liu, T., Corriveau, M., Lee, H.Y., Gott, J., Bundschuh, R. (2008) *Bioinformatics*, **24**, 2571–2578.
77. Gott, J.M., Parimi, N., Bundschuh, R. (2005) *Nucleic Acids Res.*, **33**, 5063–5072.
78. Hendrickson, P.G., Silliker, M.E. (2010) *Curr. Genet.*, **56**, 203–213.
79. Gott, J.M., Somerlot, B.H., Gray, M.W. (2010) *RNA*, **16**, 482–488.
80. Alfonzo, J.D., Thiemann, O.H., Simpson, L. (1998) *J. Biol. Chem.*, **273**, 30003–30011.
81. Pring, D., Brennicke, A., Schuster, W. (1993) *Plant Mol. Biol.*, **21**, 1163–1170.
82. Barth, C., Greferath, U., Kotsifas, M., Fisher, P.R. (1999) *Curr. Genet.*, **36**, 55–61.
83. Lin, S., Zhang, H., Spencer, D. F., Norman, J.E., Gray, M.W. (2002) *J. Mol. Biol.*, **320**, 727–739.
84. Vanfleteren, J.R., Vierstraete, A.R. (1999) *RNA*, **5**, 622–624.
85. Lukes, J., Hashimi, H., Zikova, A. (2005) *Curr. Genet.*, **48**, 277–299.
86. Simpson, L., Sbicego, S., Aphasizhev, R. (2003) *RNA*, **9**, 265–276.
87. Blum, B., Bakalara, N., Simpson, L. (1990) *Cell*, **60**, 189–198.

88. Blum, B., Simpson, L. (1990) *Cell*, **62**, 391–397.
89. Bhat, G.J., Koslowsky, D.J., Feagin, J.E., Smiley, B.L., Stuart, K. (1990) *Cell*, **61**, 885–894.
90. Koslowsky, D.J., Bhat, G.J., Perrolaz, A.L., Feagin, J.E., Stuart, K. (1990) *Cell*, **62**, 901–911.
91. Sturm, N.R., Simpson, L. (1990) *Cell*, **61**, 879–884.
92. Grams, J., McManus, M.T., Hajduk, S.L. (2000) *EMBO J.*, **19**, 5525–5532.
93. Clement, S.L., Mingler, M.K., Koslowsky, D.J. (2004) *Eukaryot Cell*, **3**, 862–869.
94. Maslov, D.A., Simpson, L. (1992) *Cell*, **70**, 459–467.
95. Ochsenreiter, T., Hajduk, S.L. (2006) *EMBO Rep.*, **7**, 1128–1133.
96. Igo, R.P. Jr., Lawson, S.D., Stuart, K. (2002) *Mol. Cell Biol.*, **22**, 1567–1576.
97. Aphasizhev, R., Aphasizheva, I., Nelson, R.E., Simpson, L. (2003). *RNA*, **9**, 62–76.
98. Aphasizhev, R., Aphasizheva, I., Simpson, L. (2003). *Proc. Natl. Acad. Sci. USA*, **100**, 10617–10622.
99. Reifur, L., Koslowsky, D.J. (2008) *RNA*, **14**, 2195–2211.
100. Seiwert, S.D., Heidmann, S., Stuart, K. (1996) *Cell*, **84**, 831–841.
101. Kable, M.L., Seiwert, S.D., Heidmann, S., Stuart, K. (1996) *Science*, **273**, 1189–1195.
102. Seiwert, S.D., Stuart, K. (1994) *Science*, **266**, 114–117.
103. Stuart, K.D., Feagin, J.E., Abraham, J.M. (1989) *Gene*, **15**, 155–160.
104. Aphasizhev, R., Aphasizheva I., Nelson, R.E., Gao, G.H., Simpson, A.M., Kang, X.D. Falick, A.M., Sbicego, S., Simpson, L. (2003) *EMBO J.*, **22**, 913–924.
105. Falick, A.M., Sbicego, S., Simpson, L. (2003) *EMBO J.*, **22**, 913–924.
106. Peris, M., Frech, G.C., Simpson, A.M., Bringaud, F., Byrne, E., Baker, A., Simpson, L. (1994) *EMBO J.*, **13**, 1664–1672.
107. Corell, R.A., Read, L.K., Riley, G.R., Nellissery, J.K., Allen, T.E., Kable, M.L., Wachal M.D., Seiwert S.D., Myler P.J., Stuart K.D., (1996) *Mol Cell Biol.* **16**:1410–1418.
108. Seiwert, S.D., Myler, P.J., Stuart, K.D. (1996) *Mol. Cell Biol.*, **16**, 1410–1418.
109. Simpson, L., Aphasizhev, R., Gao, G., Kang, X. (2004) *RNA*, **10**, 159–170.
110. Schnauffer, A., Wu, M., Park, Y.J., Nakai, T., Deng, J., Proff, R., Hol, W.G., Stuart, K.D. (2010) *J. Biol. Chem.*, **285**, 5282–5295.
111. Panigrahi, A.K., Schnauffer, A., Carmean, N., Igo, R.P., Gygi, S.P., Ernst, N.L., Palazzo, S.S., Weston, D.S., Abersold, R., Salavati, R., Stuart, K.D. (2001) *Mol. Cell Biol.*, **21**, 6833–6840.
112. Weston, D.S., Abersold, R., Salavati, R., Stuart, K.D. (2001). *Mol. Cell Biol.*, **21**, 6833–6840.
113. Blom, D., de Haan A., van den Berg, M., Sloof, P., Jirku, M., Lukes, J., Benne, R. (1998) *Nucleic Acids Res.*, **26**, 1205–1213.
114. Lukes, J., Arts, G.J., van den Burg, J., de Haan, A., Opperdoes, F., Sloof, P., Benne, R. (1994) *EMBO J.*, **13**, 5086–5098.
115. Ochs, D.E., Ots, K., Teixeira, S.M.R., Moser, D.R., Kirchhoff, L.V. (1996) *Mol. Biochem. Parasitol.*, **76**, 267–278.
116. Колесников А.А., Шониан Г., Пресбер В. (2000) *Молекулярная биология*, **34**, 183–185.
117. Simpson, L., Maslov, D.A. (1999) *Ann. NY Acad. Sci.*, **870**, 190–205.
118. Колесников А.А., Мерзляк Е.М., Бессолицина Е.А., Федяков А.В.,

- Шониан Г.* (2003) Молекулярная биология **37**, 539–543.
119. *Landweber, L.F., Gilbert, W.* (1994) Proc. Natl. Acad. Sci. USA, **91**, 918–921.