БИОФИЗИКА, 2024, том 69, № 6, с. 1167–1174

МОЛЕКУЛЯРНАЯ БИОФИЗИКА

УДК 577.323.7

АНАЛИЗ РЕДАКТИРОВАНИЯ РНК В КОНСЕРВАТИВНЫХ УЧАСТКАХ ПОСЛЕДОВАТЕЛЬНОСТИ ГЕНА RPS12 ТРИПАНОСОМАТИД

© 2024 г. Е.С. Герасимов*,#, Ю.А. Руденская*, Е.А. Брюшкова*, О.А. Коржавина*, А.А. Колесников*

**Биологический факультет Московского государственного университета имени М.В. Ломоносова, Ленинские горы, 1/12, Москва, 119234, Россия*

#E-mail: jalgard@gmail.com

Поступила в редакцию 02.05.2024 г.
После доработки 02.05.2024 г.
Принята к публикации 08.05.2024 г.

Ген RPS12 митохондриального генома кодирует важный белковый фактор малой субъединицы рибосомы. Транскрипт данного гена у ряда организмов подвергается РНК редактированию, например, он активно редактируется у всех трипаносоматид. В данной работе мы сравнили последовательности отредактированных мРНК и третичные структуры транслируемых с них белков ряда представителей семейства Trypanosomatidae и показали, что третичная структура RPS12 остается консервативной у всех видов, тогда как на уровне первичной структуры белка выделяется лишь два коротких консервативных участка. С помощью компьютерного моделирования продемонстрировано, что обнаруженные участки кодируют аминокислоты, обращенные в сторону декодирующего центра рибосомы. Характер их редактирования свидетельствует о том, что их эволюционная консервативность не связана с наличием системы редактирования, более того, прослеживается эволюционная тенденция к утрате статуса редактирования у ряда сайтов в пределах консервативных участков гена RPS12.

Ключевые слова: RPS12, рибосома, редактирование РНК, молекулярная эволюция, структура белка, консервативные области.

DOI: 10.31857/S0006302924060035, **EDN:** NLXCDL

Около половины генов митохондриального генома трипаносоматид подвергается процессу РНК-редактирования, в результате которого последовательность зрелой матричной РНК (мРНК) формируется из первичного транскрипта митохондриальной мРНК путем множественных вставок или удалений уридиловых оснований [1, 2]. Позиции и количество вставок и удалений уридилов определяется в результате взаимодействия с первичным транскриптом митохондриальной мРНК специальных молекул РНК, называемых гидовыми (гРНК), таким образом, генетическая информация зрелой мРНК оказывается разнесенной по разным локусам генома [3, 4]. Митохондриальные гены трипаносоматид расположены на относительно крупных кольцевых молекулах, называемых макси-кольцами. Гены гРНК в большинстве случаев располагаются на более мелких, однако более многочисленных мини-кольцах [3]. Разнообразие мини-колец в клетках создает основу для так называемого альтернативного редактирования, когда один локус может быть отредактирован разными гРНК [5].

Современные представления о транскриптоме трипаносоматид, полученные на основании данных высокопроизводительного секвенирования РНК, выявляют его чрезвычайную сложность, выраженную в присутствии множества изоформ мРНК [6, 7]. В формировании сложности транскриптома задействованы как минимум два механизма: привлечение нескольких альтернативных гРНК к редактированию одного локуса [7–9] и формирование альтернативных вариантов локуса в результате взаимодействия с отдельно взятой гРНК [10].

В редактировании гена RPS12, кодирующего белок S12 малой субъединицы митохондриальной рибосомы, участвует несколько десятков гидовых РНК, производя в общей сложности вставку и удаление 124–172 уридиловых оснований в зависимости от вида организма [6, 7, 11–13]. В данной работе мы показываем, что ген RPS12 стабильно

Сокращения: мРНК – матричная РНК, гРНК – гидовая РНК, RPS12 – ген митохондриального рибосомального белка S12 (Ribosomal Protein S12), ЭТЦ – электрон-транспортная цепь.

экспрессируется у всех исследованных нами видов трипаносоматид, а также стабильно редактируется, в отличие от генов электрон-транспортной цепи (ЭТЦ). В белковой последовательности RPS12 можно выделить два чрезвычайно консервативных участка (блока). Сравнительный анализ редактирования этих участков мРНК RPS12 разных видов показывает, что консервативность паттерна редактирования не зависит от последовательности гидовой РНК. Данные результаты вносят большой вклад в изучение эволюции системы РНК редактирования.

МАТЕРИАЛЫ И МЕТОДЫ

Митохондриальные геномы и транскриптомы. В данной работе анализировали митохондриальные геномы и транскриптомы различных видов трипаносоматид из разных филогенетических клад семейства, их данные приведены в табл. 1.

В случае, если необходимые анализы для вида были проведены ранее, приводится ссылка на публикацию и используются последовательности матричной РНК, геномной ДНК и результаты секвенирования РНК из данной работы. В других случаях последовательности митохондриального генома и чтения РНК скачивался из соответствующих баз данных, а анализ производился, как описано ниже.

Анализ митохондриального транскриптома. Последовательность гена и отредактированной (зрелой) мРНК RPS12 для всех исследованных видов были получены с использованием программы T-Aligner [7, 10] либо в рамках данной работы из публично доступных данных РНК-секвенирования, либо в ходе предыдущих исследований (см. табл. 1) по методу, описанному ранее [6, 10, 13]. Кратко суть метода заключается в следующем: результаты транскриптомного секвенирования картируются с помощью программы alignlib пакета T-Aligner (алгоритм программы создан для работы с сильно измененными уридиловым редактированием чтениями РНК) на последовательность митохондриального макси-кольца, выявленные события вставки и удаления уридилов используются для аннотации границ редактируемых генов. Затем чтения, картированные на ген RPS12, отбираются и используются программой findorfs пакета T-Aligner для транскриптомной сборки и анализа изоформ. Последовательность зрелой мРНК гена RPS12 выделяется, как наиболее похожая на известные последовательности RPS12 изоформа с помощью blastp из ncbi-blast-suite [14].

Анализ транскриптома каждого из видов проводили следующим образом: количество картированных чтений и долю отредактированных чтений оценивали по результату работы инструмента alignlib, события редактирования устанавливали с использованием скрипта coverage.py пакета T-Aligner. В работе использованы полученные ранее карты взаимодействия гРНК с мРНК RPS12 для *Trypanoplasma borreli* [6], *Blastocrithidia nonstop* [12], *Leishmania tarentolae* [11] и *Leptomonas pyrrhocoris* [10].

Множественное выравнивание последовательности RPS12 и выделение консервативных позиций. Множественное выравнивание отредактированных мРНК RPS12 осуществляли путем группирования попарных потриплетных выравниваний мРНК и геномной ДНК для каждого из исследуемых видов, полученной с помощью ин-

Таблица 1. Данные анализированных видов трипаносоматид

Вид	РНК-секвенирование (номер в SRA или ссылка на источник)	Макси-кольцо (номер в GenBank, TriTrypDB или ссылка на источник)
Trypanosoma cruzi	SRR17349272; SRR17349274	[13]
Leishmania donovani	ERR2191874	MN904518.1
Leishmania major	SRR2136265; SRR2136270	LR697138
Leishmania mexicana	ERR789787; ERR789788	TriTryp: MHOM/GT/2001/U1103
Leptomonas pyrrhocoris	[7]	[7]
Wallacemonas sp. WSD	SRR24223636	GCA_030849645.1 (скаффолд JASANQ010001068.1)
Blastocrithidia nonstop	[12]	[12]
Trypanoplasma borreli	[6]	[6]

струмента mrna2ref из пакета T-Aligner. Консервативность позиции оценивали путем вычисления энтропии колонки выравнивания по формуле – $\sum_x p_x \log p_x$, где p_x – наблюдаемая вероятность соответствующей аминокислоты в столбце. Консервативные блоки были выбраны в зонах выравнивания, где энтропия была меньше 0.5 для окна длиной 10 аминокислотных позиций вокруг выбранной позиции.

Моделирование третичной структуры белка RPS12. Для моделирования третичной структуры белка RPS12 использовалb последовательность, полученне. путем *in silico* трансляции отредактированной мРНК. Моделирование проводилb с помощью web-сервиса SWISS-MODEL [15] Швейцарского института биоинформатики (swissmodel.expasy.org/).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Характеристика экспрессии и редактирования гена RPS12 у разных представителей семейства Trypanosomatidae. Большинство имеющихся транскрипционных данных для трипаносоматид получено в результате секвенирования лабораторных культур, однако имеются основания полагать, что в условиях культивирования на искусственных средах ЭТЦ митохондрии либо не активна, либо работает не в полную силу, что может сказаться на экспрессии генома митохондрии. Так, например, известно, что у *Leishmania tarentolae* ряд генов белков ЭТЦ транскрибируется на низком уровне, а транскрипты этих генов не редактируются до зрелых мРНК [11]. Ген RPS12 кодирует важный для трансляции белковый фактор малой субъединицы митохондриальной рибосомы, можно ожидать, что данный ген будет экспрессироваться и редактируется в любых условиях. Наши предыдущие работы подтверждают данное предположение: ген RPS12 активно транскрибируется и редактируется у *Leptomonas pyrrhocoris* [7], *Blastocrithidia nonstop* [12], трех штаммов *Trypanosoma cruzi* [13], а также у криптобии *Trypanoplasma borreli* [6]. При этом в ряде исследований были использованы дифференциальные по содержанию глюкозы и других питательных веществ условия культивирования, но уровень транскрипции RPS12 оставался стабильным. В данной работы мы расширили круг видов, проведя анализ экспрессии и характера редактирования мРНК гена RPS12 у трех представителей рода *Leishmania* и одного вида из рода *Wallacemonas*. При этом для *Leishmania mexicana* и *Leishmania major* нами проанализированы данные по экспрессии на промастиготной (характерной для нахождения паразита в насекомом) и амастиготной (проходящей в теплокровном хозяине) стадиях жизненного цикла. Во всех исследованных образцах ген RPS12 экспрессируется на относительно высоком уровне (от ~200 до ~45 000 чтений, в зависимости от глубины секвенирования), при этом для большинства видов это самый или один из самых высокоэкспрессируемых генов митохондриального генома. Во всех исследованных образцах удается реконструировать последовательность зрелой мРНК RPS12 с помощью T-Aligner, что свидетельствует о том, что ген активно и полностью редактируется. Интересно отметить, что при этом почти в половине образцов исследованных лейшманий, в том числе на амастиготной стадии, не удается реконструировать зрелые мРНК ряда генов ЭТЦ, например, ND8 и ND9.

Таким образом, можно охарактеризовать паттерн экспрессии и редактирования RPS12 как стабильный, что делает данный ген удобным объектом для исследования эволюции механизма редактирования.

Консервативные участки RPS12 и их расположение на модели 3D-структуры белка. Используя полученные ранее и реконструированные в данной работе последовательности отредактированных (зрелых) мРНК RPS12 разных видов трипаносоматид, мы построили множественное выравнивание белковых продуктов. С помощью метода оценки энтропии позиций в данном выравнивании четко выделяются два консервативных блока (длиной 9 и 11 аминокислот), которые мы обозначили по характерным аминокислотам, как PSSG и 3GFID.

Модель 3D-структуры белка RPS12 *Trypanosoma cruzi* (модели, полученные для последовательностей других видов, имеют схожие параметры) по шаблонам (SWISS-MODEL) показана на рис. 1. Очевидно, что оба консервативных участка белка ориентированы на одну сторону и, вероятно, участвуют в важных взаимодействиях, что объясняет высокую степень эволюционной консервации данных участков. Известно, что белок RPS12 является важным фактором сборки рибосомы [16, 17] и находится в непосредственной близости от декодирующего центра малой субъединицы [16], при этом участок белка, отмеченный стрелкой, направлен в канал рибосомы, где проходит матричная РНК.

Интересно, что белок RPS12 кодируется митохондриальным геномом у разных организмов, что, вероятно, связано с крайней гидрофобностью данного белка и необходимостью в котрансляционном включении его в рибосомы [17]. Также известно, что RPS12 редактируется не только у трипаносоматид. У растений транскрипт гена *rps12* также подвергается редактированию по механизму конверсии цитозина (C→U) [18, 19], при этом показано, что в состав рибосомы могут включаться частично отредактированные формы [18], хотя в случае с другими белками ЭТЦ расте-

ний присутствует отбор в пользу трансляции только полностью отредактированных матриц [19].

Выравнивание геномных ДНК, мРНК и соответствующих аминокислотных последовательностей в участках, соответствующих консервативным участкам RPS12, приведено на рис. 2. Из выравнивания видно, что последовательности геномной ДНК обоих блоков очень консервативны, отличия же в основном касаются кодируемых в геноме тимидинов, наличие или отсутствие которых в геномной ДНК разных видов компенсируется затем действием системы редактирования. Логично предположить, что соответствующие мутации в геномной ДНК эволюционно компенсировались мутациями в генах гРНК, редактирующих данный регион. Ниже мы рассмотрим эту гипотезу детально. Отметим также, что за пределами выделенных нами консервативных участков белковая последовательность RPS12 трипаносоматид достаточно дивергентна, при этом согласно результатам моделирования 3D-структура белка RPS12 разных видов оказывается схожей со структурой белка *Trypanosoma brucei*, используемой в качестве шаблона SWISS-MODEL.

Изменения в геномной ДНК не связаны с компенсаторными мутациями генов гидовых РНК. Мы сравнили доступные карты гРНК RPS12 четырех видов *Trypanoplasma borreli* [6], *Blastocrothidia nonstop* [12], *Leptomonas pyrrhocoris* [10] и *Leishmania tarentolae* [11]. Выравнивания гРНК на фрагмент мРНК в зонах консервативных участков для данных видов представлены на рис. 3. Из приведенного выравнивания хорошо видно, что для редактирования консервативных участков разных видов используются негомологичные гРНК. Во-первых, последовательности самих гРНК обладают низким уровнем сходства. Во-вторых, что более важно, гомологичные сайты на мРНК разных видов могут редактироваться в составе разных гРНК (соседних по каскаду, см. ниже). Например, область участка PSSG мРНК *Trypanoplasma borreli* редактируется одной гРНК, тогда как у *Leptomonas pyrrhocoris* данная область редактируется двумя гРНК, граница между которыми приходится между парой серинов в участке PSSG, а у относительно близкородственной *Leishmania tarentolae* граница между гРНК находится после второго серина. Аналогичную картину можно наблюдать в блоке выравнивания 3GFID, где у трех видов сайты редактирования попадают в одну гРНК, а у *T. borreli* – в две с границей перед тремя глицинами. Таким образом, в образование кодонов гомологичных и крайне консервативных аминокислотных остатков белка RPS12 вовлекаются негомологичные и структурно различные молекулы гРНК. Из этого можно сделать вывод, что консервативность белковой структуры RPS12 сохраняется вопреки высокой вариабельности последовательности гРНК, характер взаимодействия гРНК с мРНК в зоне консервативных участков мРНК сильно варьирует у разных видов, при этом структура зрелой мРНК остается практически неизменной.

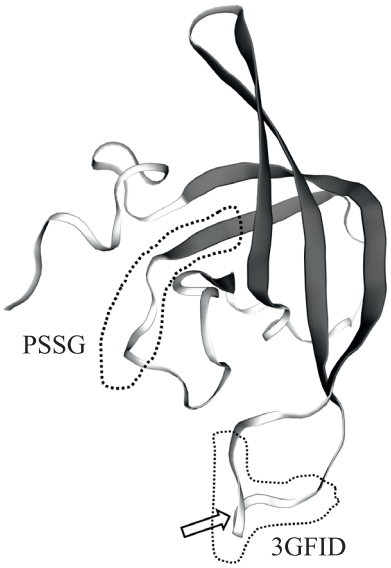

Рис. 1. Структурная модель белка RPS12 *Trypanosoma cruzi* штамма Sylvio X10. Модель 3D-структуры белка RPS12, полученная с помощью web-сервиса SWISS-MODEL [15] с использованием установленной нами последовательности зрелой мРНК *Trypanosoma cruzi* [13] в качестве запроса. Использован шаблон 7pub.8, оценка качества полученной модели (Global Model Quality Estimate, GMQE) равна 0.50. Лентой условно изображены участки со стабильной вторичной структурой согласно алгоритму DSSP. Пунктиром обведены зоны консервативных участков мРНК, обозначенные нами, как PSSG и 3GFID в данной работе. Стрелка указывает на место в структуре, наиболее близко расположенное к декодирующему центру на малой субъединице рибосомы [16].

Эволюция паттернов редактирования в консервативных участках последовательности гена RPS12. Выше мы показали, что гРНК, привлекаемые для редактирования RPS12 у разных видов, не являются эволюционно консервативными и не являются гомологами. Однако характерные паттерны редактирования остаются достаточно консервативными. На рис. 4 приведены формулы редактирования консервативных блоков PSSG и 3GFID у исследуемых видов, которые отражают последовательность редакционных действий, имеющую место при создании системной редактирования зрелой мРНК. Интересно, что паттерны редактирования оказываются намного более консервативными, чем сами последовательности гРНК, но менее консервативными, чем геномные

```
Lpyr AG-CCTCGA    --ACCT     AGTTCC        GG- AA-AGACG-  --A---ATAG--GAATTTTTTGG-  GG-GG----A--   GATCTACCC     GGA
Lpyr AGuCCUCGA    uuACCU     AGUUCC        GGu AAuAGACGu  uuAuuuAUAGuuGAA------GGu  GGuGGuuuuAuu   GAUCUACCC     GGA
Lpyr S  P  R      L  P       S  S          G   N  R  R   L  F  I  V  E        G    G  G  F  I     D  L  P       G
Lmaj AG-CCACGA    --ACCA     AGCTCC        GG- AACCGACG-  TTG---A-AG--GAATTTT  GG-  GG-GG----A--   GA---GCCA     GGA
Lmaj AGuCCACGA    uuACCA     AGCUCC        GGu AACCGACGu  UUGuuuAuAGuuGAA----  GGu  GGuGGuuuuAuu   GAuuuGCCA     GGA
Lmaj S  P  R      L  P       S  S          G   N  R  R   L  F  I  V  E        G    G  G  F  I     D  L  P       G
Lmex AG-CCACGA    --ACCA     AGCTCC        GG- AATCGACG-  TTG---A-AG--GAATT    GG-  GG-GG----A--   GA---GCCA     GGA
Lmex AGuCCACGA    uuACCA     AGCUCC        GGu AAUCGACGu  UUGuuuAuAGuuGAA--    GGu  GGuGGuuuuAuu   GAuuuGCCA     GGA
Lmex S  P  R      L  P       S  S          G   N  R  R   L  F  I  V  E        G    G  G  F  I     D  L  P       G
Ldon AG-CCACGA    --ACCA     AGTTCC        GG- AACCGACG-  TTG---A-AG--GAATTTTTTGG-  GG-GG----A--   GA---GCCA     GGA
Ldon AGuCCACGA    uuACCA     AGUUCC        GGu AACCGACGu  UUGuuuAuAGuuGAA------GGu  GGuGGuuuuAuu   GAuuuGCCA     GGA
Ldon S  P  R      L  P       S  S          G   N  R  R   L  F  I  V  E        G    G  G  F  I     D  L  P       G
Bnon AG-CCACG-    --GCCT     AGTTCA        GG- AA-CG---A  G-AT-GGGA---GAAT     GG-  GG-GG----A--   GA-G--CCA     GG-
Bnon AGuCCACGt    uuGCCU     AGUUCA        GGu AAuCGuuuA  GuAUuGGGAuuuGAA-     GGu  GGuGGuuuuAuu   GAuGuuCCA     GGu
Bnon S  P  R      L  P       S  S          G   N  R  L   V  L  G  F  E        G    G  G  F  I     D  V  P       G
Wsd  AGTCCACTTGG  T-GCCA     AG--CC        GG- AATCG-CG-  --A---A-AA--GATTTA   GG-  GG-GG----A--   GA--G-CCA     GG-
Wsd  AGUCCAC--GG  UuGCCA     AGuuCC        GGu AAUCGuCGu  uuAuuuAuAAuuGA---A   GGu  GGuGGuuuuAuu   GAuuGuCCA     GGu
Wsd  S  P  R      L  P       S  S          G   N  R  R   L  F  I  I  E        G    G  G  F  I     D  C  P       G
Tcru AGCCCACTTTTGA--ACCC     AGT-CC        GG- AACCGTCG-  --A---A--A-AGATATT   GTTG-GG-GG----A--   GA---ACCCTTTTTGTTT
Tcru AGCCCAC----GAuuACCC     AGUuCC        GGu AACCGUCGu  uuAuuuAuuAuAGA-A--   G--GuGGuGGuuuuAuu   GAuuuACCC-----G---
Tcru S  P  R      L  P       S  S          G   N  R  R   L  F  I  I  E        G    G  G  F  I     D  L  P       G
Tbor -GTCCACG-    --ACTTTTTCCAG--CCTTTTTTTTGTGTAA-CG-A-G  A---A-GTTG-GGAA     GGC  GG-GG----ATTTTTGA---ACTC-    AG-
Tbor uGUCCACGu    uuAC-----CCAGuuCC--------G-GUAAuCGuAuG  AuuuAuGUUGuGGAA     GGC  GGuGGuuuuAUU---GAuuuAC-Cu   AGu
Tbor C  P  R      L  P       S  S          G   N  R  M   I  Y  V  V  E        G    G  G  F  I     D  L  P       S
```

PSSG 3GFID

Рис. 2. Множественное выравнивание консервативных участков PSSG и 3GFID гена RPS12. Приведены два консервативных блока (участки PSSG и 3GFID) глобального потриплетного выравнивания последовательностей мРНК, геномной ДНК и соответствующего пептида для RPS12 исследуемых видов трипаносоматид. Для каждого вида выравнивание имеет три строки в следующем порядке сверху вниз: геномная ДНК, мРНК и последовательность белка. Вставляемые на этапе редактирования нуклеотиды показаны строчными буквами (u), удаляемые – знаком минуса (-). Видовые сокращения: Lpyr – *Leptomonas pyrrhocoris*, Lmaj – *Leishmania major*, Lmex – *Leishmania Mexicana*, Ldon – *Leishmania donovani*, Bnon – *Blastocrithidia nonstop*, Wsd – *Wallacemonas sp.* WSD, Tcru – *Trypanosoma cruzi*, Tbor – *Trypanoplasma borelli*.

последовательности. Также заметно, что паттерн редактирования хорошо согласуется с филогенетическим положением видов, причем наиболее поздно ветвящимся родам *Leptomonas-Leishmania* присущи относительно более простые формулы редактирования.

Особенностью системы уридилового редактирования является направленность процесса по мРНК, называемая полярностью [20, 21]. Она обусловлена тем, что сайты посадки следующих гРНК обычно создаются в результате редактирования предыдущими гРНК, при этом редактиро-

```
Tbor ......    CACGUG-----AUUUAGGG--------U-CAUUAGCAUAC...  AAAUAAAUACAACACCUU]      [CACCAUAAUAA---UUAAAUG...
Tbor ......    ||#:||     ##|:|:||        : ||||||||||      ||||||||||||||||||       |||||#|||||   :||||||
Tbor ......CCACGuuuAC-----CCAGuuCC--------G-GUAAuCGuAuG...  uuuAuuuAuGUUGuGGAA    GGCGGuGGuuuuAUU---GAuuuAC...
Tbor ......P  R  L  P       S  S          G   N  R  M      F  I  Y  V  V  E     G  G  G  F  I     D  L  P...
Bnon ...UCAGGUGCAAGCG     GAUCUUGC       CC]UUAGCAAGTG...       GAUUGUAAGCAA    UCAUCGCUGGAGUGA   CUACAAGAGUCCCUC
Bnon ...||||||||||:||     ||||##|#       ||[||||||||||:       :|::#|||:|##    :||:|:|:::|:|:|   |||||||##:|#|:|
Bnon ...AGuCCACGuuuGC     CUAGUUCA       GGuAAuCGuuuAu...       uUGGGAuuuGAA    GGuGGuGGuuuuAuu   GAuGuuCCAGGuGGG
Bnon ...S  P  R  L  P       S  S          G  N  R  L            L  G  F  E     G  G  G  F  I     D  V  P  G  G
Lpyr ...UUAAGAGUUGAUG     GA] [AGU       CCAUUAUCCAUAU      AUAAGUAUCAACUU    AUAUCGUUAGAAUGA   UUAGAUGGGCUUACAUU
Lpyr ...|:|#|||:|:|||     ||   ||#       ||||||||##:|:       ||||:|||||||||    #:|:|:::|:|||:|   :||||||||||:||||||
Lpyr ...AGuCCuCGAuuAC     CuAGuuCC       GGuAAuAGACGuG      uAuuuAuAGuuGAA    GGuGGuGGuuuuAuu   GAuCuACCCGGAuGuAA
Lpyr ...S  P  R  L  P       S  S          G   N  R  R            F  I  V  E     G  G  G  F  I     D  L  P  G  C
Ltar    UCAGGUGUUAAUG     GAUUAAGG ]    [UUAUUAGUUGCAA..         AGAUGUUGGCUC----UCACUAUCAAGAUAA   UUAGACGGUCCU
Ltar    |||||||:|||||     |||:||||        ::|||||:                |:||:||::|||     ||||:|:|||:||||   :||||||||||
Ltar    AGuCCACGAuuAC     CUAGUUCC       GGuAAUCGCACGu..         uuuAuAGuuGAA    GGuGGuGGuuuuAuu   GAuuuGCCAGGA
Ltar    S  P  R  L  P       S  S          G  N  R  R            F  I  V  E     G  G  G  F  I     D  L  P  G
```

PSSG 3GFID

Рис. 3. Карты взаимодействия гидовых РНК с мРНК в областях консервативных участков гена RPS12. Для каждого вида сверху вниз даны последовательности гРНК, строка выравнивания, последовательность мРНК и соответствующая им трансляция открытой рамки считывания. В строке выравнивания канонические пары оснований обозначены символом вертикальной черты (|); пары G:U, характерные для дуплексов гРНК:мРНК [2, 4] – символом двоеточия (:); несовпадающие нуклеотиды – символом (#). Малыми буквами u обозначены вставляемые урацилы. Границы гРНК обозначены квадратными скобками. Видовые сокращения: Lpyr – *Leptomonas pyrrhocoris*, Ltar – *Leishmania tarentolae*, Bnon – *Blastocrithidia nonstop*, Tbor – *Trypanoplasma borelli*.

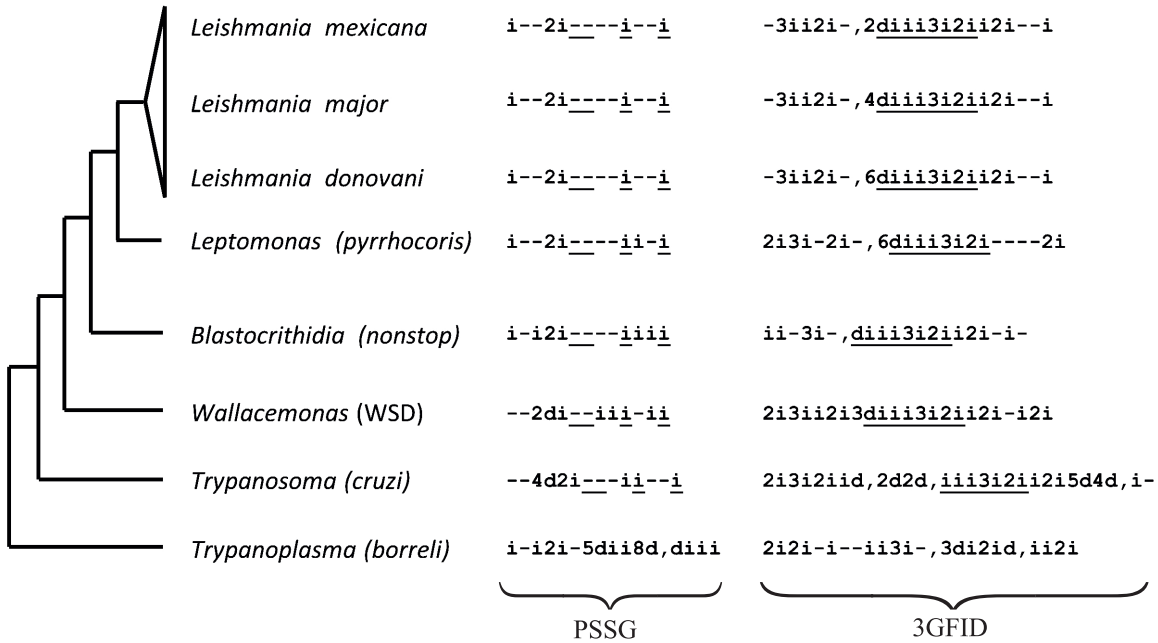

Рис. 4. Формулы редактирования консервативных участков мРНК гена RPS12 в разных кладах трипаносоматид. Кладограмма слева отражает лишь порядок ветвления исследуемых таксонов трипаносоматид, справа от нее приведены формулы редактирования участков консервативных участков мРНК гена RPS12. В случаях, когда таксон является родом, в скобках указывается вид, мРНК которого была использована для определения формулы (например, род *Wallacemonas* представлен изолятом WSD). В таксономической группе, представляющей род лейшманий, для демонстрации крайней консервативности структуры редактирования приведены 3 разных вида. Формулы редактирования являются компактной формой записи событий редактирования каждого из кодонов мРНК в пределах консервативного блока. Обозначены вставки (i) и удаления (d) уридиловых остатков системой редактирования (в случае, когда происходит вставка/удаление нескольких уридилов, соответствующей букве предшествует число вставляемых/удаляемых остатков, например, вставка двух уридилов обозначается 2i); нередактируемые кодоны обозначены знаком минуса (-). В случае, когда триплет возникает путем двух событий редактирования, такие события перечисляются через запятую (например, (i,2d) – вставка одного уридила сочетается с удалением 2 уридилов при формировании данного кодона). В формулах подчеркнуты символы, отвечающие высококонсервативным типам редактирования в соответствующих кодонах (см. раздел «Результаты и обсуждение»).

вание начинается с 3’-сайтов транскрипта и заканчивается на 5’-конце. В пределах одной гРНК редактирование также протекает полярно, сайты на 3’ (по мРНК) редактируются первыми [21].

Полученные нами данные свидетельствуют в пользу того, что структура каскадов гРНК, возникающая при редактировании мРНК, не является эволюционно консервативной. Гомологичные локусы мРНК разных видов могут редактироваться в составе разных гРНК, при этом последовательность аминокислот кодируемого белка может быть идентичной у очень дальних филогенетически видов (например, роды *Trypanosoma* и *Leishmania*), а структура самих каскадов сильно различаться у близких родственников (например, роды *Leptomonas* и *Leishmania*).

Интересно также отметить общую тенденцию к сокращению числа событий редактирования, характерную для семейства трипаносоматид. Так, например, у криптобии *T. borreli* наблюдаются наиболее сложные паттерны редактирования, а у ветвящихся последними на кладограмме *Leptomonas-Leishmania* большинство сайтов не требует изменения (обозначены знаком «минус» на схеме, представленной на рис. 4). Подобная тенденция ранее была прослежена на уровне редактирования генов в целом [22, 23], когда ген или его часть в макси-кольце замещается на отредактированную версию, таким образом, снижая общее число редактируемых генов в геноме или длину редактируемого участка определенного гена. Предполагалось, что механизмом здесь служит обратная транскрипция частично или полностью отредактированной мРНК в геном макси-кольца. Данная теория хорошо согласуется с полярным характером редактирования, что объясняет редукцию редактирования на 3'-концах транскриптов. Но в случае с RPS12 ген редактируется по всей длине у всех представителей семейства, и в данной работе мы впервые указываем на тенден-

цию в эволюции к утрате редактирования на уровне отдельных сайтов. Это наблюдение может указывать на существование более точечных механизмов эволюции, направленных на отдельные сайты внутри гена.

ЗАКЛЮЧЕНИЕ

Уридиловое редактирование РНК у трипаносоматид – уникальный молекулярный феномен, по-видимому, присущий только данной группе паразитических простейших. Роль редактирования в эволюции представителей семейства на сегодняшний день остается объектом интенсивного изучения, предположительно редактирование может быть связано с паразитизмом. Изучение эволюции редактирования в наиболее консервативных участках мРНК гена RPS12, стабильная экспрессия которого говорит о его важной роли для клеток всех видов, указывает на эволюционные тенденции, направленные на уменьшение числа редактируемых сайтов в генах макси-кольца и на отсутствие гомологии между аналогичными по локализации на мРНК генами гРНК.

БЛАГОДАРНОСТИ

Авторы благодарят своих коллег, профессора Остравского университета Вячеслава Юрченко и доцента Университета Миннесоты Сару Лауру Циммер за помощь в обсуждении.

ФИНАНСИРОВАНИЕ РАБОТЫ

Исследование выполнено при финансовой поддержке Российского научного фонда (проект № 19-74-10008).

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Настоящая статья не содержит описания собственных исследований с участием людей или животных в качестве объектов.

СПИСОК ЛИТЕРАТУРЫ

1. Lukeš J., Kaur B., and Speijer D. RNA editing in mitochondria and plastids: weird and widespread. *Trends Genet.*, **37**, 99–102 (2021). DOI: 10.1016/j.tig.2020.10.004
2. Read L. K., Lukeš J., and Hashimi H. Trypanosome RNA editing: the complexity of getting U in and taking U out. *Wiley Interdisciplinary Rev.*: *RNA*, **7**, 33–51 (2016). DOI: 10.1002/wrna.1313
3. Shlomai J. The structure and replication of kinetoplast DNA. *Curr. Mol. Med.*, **4**, 623–647 (2004). DOI: 10.2174/1566524043360096
4. Aphasizheva I., Alfonzo J., Carnes J., Cestari I., Cruz-Reyes J., Göringer H. U., Hajduk S., Lukeš J., Madison-Antenucci S., Maslov D. A., McDermott S. M., Ochsenreiter T., Read L. K., Salavati R., Schnaufer A., Schneider A., Simpson L., Stuart K., Yurchenko V., Zhou Z. H., and Aphasizhev R. Lexis and grammar of mitochondrial RNA processing in trypanosomes. *Trends Parasitol.*, **36**, 337–355 (2020). DOI: 10.1016/j.pt.2020.01.006
5. Zimmer S. L., Simpson R. M., and Read L. K. High throughput sequencing revolution reveals conserved fundamentals of U-indel editing. *Wiley Interdisciplinary Rev.: RNA*, **9**, e1487 (2018). DOI: 10.1002/wrna.1487
6. Gerasimov E. S., Afonin D. A., Korzhavina O. A., Lukeš J., Low R., Hall N., Tyler K., Yurchenko V., and Zimmer S. L. Mitochondrial RNA editing in *Trypanoplasma borreli*: New tools, new revelations. *Comput. Struct. Biotechnol. J.*, **20**, 6388–6402 (2022). DOI: 10.1016/j.csbj.2022.11.023
7. Gerasimov E. S., Gasparyan A. A., Kaurov I., Tichý B., Logacheva M. D., Kolesnikov A. A., Lukeš J., Yurchenko V., Zimmer S. L., and Flegontov P. Trypanosomatid mitochondrial RNA editing: dramatically complex transcript repertoires revealed with a dedicated mapping tool. *Nucl. Acids Res.*, **46**, 765–781 (2018). DOI: 10.1093/nar/gkx1202
8. Ochsenreiter T., Anderson S., Wood Z. A., and Hajduk S. L. Alternative RNA editing produces a novel protein involved in mitochondrial DNA maintenance in trypanosomes. *Mol. Cell. Biol.*, **28**, 5595–5604 (2008). DOI: 10.1128/MCB.00637-08
9. Kirby L. E. and Koslowsky D. Mitochondrial dual-coding genes in *Trypanosoma brucei*. *PLoS Neglected Tropic. Dis.*, **11**, e0005989 (2017). DOI: 10.1371/journal.pntd.0005989
10. Gerasimov E. S., Gasparyan A. A., Afonin D. A., Zimmer S. L., Kraeva N., Lukeš J., Yurchenko V., and Kolesnikov A. A. Complete minicircle genome of *Leptomonas pyrrhocoris* reveals sources of its non-canonical mitochondrial RNA editing events. *Nucl. Acids Res.*, **49**, 3354–3370 (2021). DOI: 10.1093/nar/gkab114
11. Simpson L., Douglass S. M., Lake J. A., Pellegrini M., and Li F. Comparison of the mitochondrial genomes and steady state transcriptomes of two strains of the trypanosomatid parasite, *Leishmania tarentolae*. *PLoS Neglected Tropic. Dis.*, **9**, e0003841 (2015). DOI: 10.1371/journal.pntd.0003841
12. Afonin D. A., Gerasimov E. S., Sveráková I. Š., Záhonová K., Gahura O., Albanaz A. T. S., Myšková E., Bykova A., Paris Z., Lukeš J., Opperdoes F. R., Horváth A., Zimmer S. L., and Yurchenko V. *Blastocrithidia nonstop* mitochondrial genome and its expression are remarkably insulated from nuclear codon reassignment. *Nucl. Acids Res.*, 52 (7), 3870–3885 (2024). DOI: 10.1093/nar/gkae168

13. Gerasimov E. S., Ramirez-Barrios R., Yurchenko V., and Zimmer S. L. *Trypanosoma cruzi* strain and starvation-driven mitochondrial RNA editing and transcriptome variability. *RNA*, **28**, 993–1012 (2022). DOI: 10.1261/rna.079088.121

14. Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., and Madden T. L. BLAST+: architecture and applications. *BMC Bioinformatics*, **10**, 421 (2009). DOI: 10.1186/1471-2105-10-421

15. Waterhouse A., Bertoni M., Bienert S., Studer G., Tauriello G., Gumienny R., Heer F. T., de Beer T. A. P., Rempfer C., Bordoli L., Lepore R., and Schwede T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucl. Acids Res.*, **46**, W296–W303 (2018). DOI: 10.1093/nar/gky427

16. Lenarčič T., Niemann M., Ramrath D. J. F., Caldera ro S., Flügel T., Saurer M., Leibundgut M., Boehringer D., Prange C., Horn E. K., Schneider A., and Ban N. Mitoribosomal small subunit maturation involves formation of initiation-like complexes. *Proc. Natl. Acad. Sci. USA*, **119**, e2114710118 (2022). DOI: 10.1073/pnas.2114710118

17. Ramrath D. J. F., Niemann M., Leibundgut M., Bieri P., Prange C., Horn E. K., Leitner A., Boehringer D., Schneider A., and Ban N. Evolutionary shift toward protein-based architecture in trypanosomal mitochondrial ribosomes. *Science*, **362**, eaau7735 (2018). DOI: 10.1126/science.aau7735

18. Small I. D., Schallenberg-Rüdinger M., Takenaka M., Mireau H., and Ostersetzer-Biran O. Plant organellar RNA editing: what 30 years of research has revealed. *Plant J. Cell Mol. Biol.*, **101**, 1040–1056 (2020). DOI: 10.1111/tpj.14578

19. Lu B., Wilson R. K., Phreaner C. G., Mulligan R. M., and Hanson M. R. Protein polymorphism generated by differential RNA editing of a plant mitochondrial rps12 gene. *Mol. Cell. Biol.*, **16**, 1543–1549 (1996). DOI: 10.1128/MCB.16.4.1543

20. Koslowsky D. J., Bhat G. J., Read L. K., and Stuart K. Cycles of progressive realignment of gRNA with mRNA in RNA editing. *Cell*, **67**, 537–546 (1991). DOI: 10.1016/0092-8674(91)90528-7

21. Maslov D. A. and Simpson L. The polarity of editing within a multiple gRNA-mediated domain is due to formation of anchors for upstream gRNAs by downstream editing. *Cell*, **70**, 459–467 (1992). DOI: 10.1016/0092-8674(92)90170-h

22. Aravin A. A., Yurchenko V., Merzlyak E. M., and Kolesnikov A. A. The mitochondrial ND8 gene from *Crithidia oncopelti* is not pan-edited. *FEBS Lett.*, **431**, 457–460 (1998). DOI: 10.1016/s0014-5793(98)00813-8

23. Gerasimov E. S., Kostygov A. Y., Yan S., and Kolesnikov A. A. From cryptogene to gene? ND8 editing domain reduction in insect trypanosomatids, *Eur. J. Protistology*, **48**, 185–193 (2012). DOI: 10.1016/j.ejop.2011.09.002

Analysis of RNA Editing in Conserved Sequence Blocks of the Trypanosomatid RPS12 Gene

E.S. Gerasimov*, Y.A. Rudenskaya*, E.A. Bryushkova*, O.A. Korzhavina*, and A.A. Kolesnikov*

**Department of Biology, M.V. Lomonosov Moscow State University, Leninskie Gory 1/12, Moscow, 119234 Russia*

The RPS12 gene in the mitochondrial genome encodes the important protein from the small ribosomal subunit. The transcript of this gene in a number of organisms can undergo RNA editing, for instance it is edited actively in all tripanosomatida. In this study, a comparison between the sequences of edited mRNA and tertiary structures of the proteins when mRNA were translated from some of the members of Tripanosomatidae family showed that tertiary structures of the RPS12 gene remained unchanged in all species while only two short conserved regions were found in the primary structure of the gene. By means of computer modeling it was demonstrated that the found regions encode amino acids turned to the direction of the decoding center of ribosome. The nature of editing patterns demonstrates that evolutionary conservation of these regions is independent of the editing process and, moreover, the evolutionary tendencies on the reduction of editing process in a number of sites was seen within the conserved sites of the RPS12 gene.

Keywords: RPS12, ribosome, RNA editing, molecular evolution, protein structure, conserved domains