

УДК 575.852

РОЛЬ ОБРАТНОЙ ТРАНСКРИПТАЗЫ В ВОЗНИКНОВЕНИИ ЖИЗНИ

Обзор

© 2019 Р.Н. Мустафин^{1*}, Э.К. Хуснутдинова²

¹ Башкирский государственный медицинский университет,
450008 Уфа, Россия; электронная почта: ruji79@mail.ru

² Институт биохимии и генетики Уфимского федерального
исследовательского центра РАН, 450054 Уфа, Россия

Поступила в редакцию 08.03.2019

После доработки 14.05.2019

Принята к публикации 14.05.2019

Предполагается, что рибозим РНК-полимераза, обладающая активностью обратной транскриптазы и интегразы, сыграла ключевую роль в закономерностях возникновения жизни на Земле. Представлена гипотеза, согласно которой, благодаря обратной транскриптазе в эволюции были сформированы универсальные предковые единицы всего живого – ретроэлементы (РЭ). Их склонность к мутациям и способность к взаимointegrации стала основой для образования комплексных структур ДНК, первичных геномов, от которых произошли все археи, эукариоты, бактерии и вирусы. Консервативные свойства ретроэлементов сохранились на протяжении всей эволюции – при модификациях их использования возникли новые способы взаимодействий белков и нуклеиновых кислот. Жизнь эволюционировала благодаря инсерционному мутагенезу и противоборству автономно реплицирующихся полинуклеотидов для сохранения структур с адаптивными свойствами. Сделано предположение, что при отборе механизмов защиты от инсерций на основе рибонуклеазной способности рибозима обратной транскриптазы возникли все универсальные ферментативные системы процессинга молекул РНК. Они стали ключевыми источниками дальнейших эволюционных преобразований геномов и их регуляторных особенностей. Приведены данные, позволяющие предположить, что система трансляции, объединившая мир РНК и ДНК с белками, возникла как модификация механизмов защиты от инсерций. Образующиеся с ее помощью полипептиды стали потенцировать работу рибозимов в составе рибонуклеопротеинов (РНП) и даже функционально замещать их при большей успешности катализа биологических реакций. Проведен анализ механизмов использования ретроэлементов в структурных и регуляторных преобразованиях геномов эукариот, которые могут отражать адаптивные принципы, сформированные при зарождении жизни. Параллельно с эволюцией существующих белков, из РЭ возникают и изменяются рибозимы, такие как длинные некодирующие РНК. Они могут функционировать в комплексе с белками в составе РНП, способны к самостоятельной каталитической активности и трансляции. Их гены обладают потенциалом к преобразованию в белок-кодирующие. Т.е. консервативные принципы взаиморегуляции РНК, ДНК и белков, сформированные при возникновении жизни, используются на протяжении всей эволюции.

КЛЮЧЕВЫЕ СЛОВА: обратная транскриптаза, полимеразы, процессинг, рибозимы, ретроэлементы, транспозоны, эволюция.

DOI: 10.1134/S0320972519080037

Гипотеза мира РНК при возникновении жизни на Земле является доминирующей в современной науке. Накоплены многочисленные данные о существовании рибозимов, обладаю-

щих самостоятельной каталитической активностью. Хотя все сохранившиеся в наши дни формы жизни зависят от белковых ферментов, в мире РНК репликацию генетической информа-

Принятые сокращения: РНП – рибонуклеопротеины, нкРНК – некодирующие РНК, мяРНК – малые ядерные РНК, мяоРНК – малые ядрышковые РНК, ТЕ – транспозоны (transposable elements), НП – нуклеотидные последовательности, РНКи – РНК-интерференция, РЭ – ретроэлементы, CRISPR – короткие палиндромные повторы, регулярно расположенные группами (clustered regularly interspaced short palindromic repeats), LTR – длинные концевые повторы (long terminal repeats), LINE – длинные диспергированные элементы ядра (long interspersed nuclear element), nonLTR-РЭ – ретроэлементы, не содержащие длинные концевые повторы, RISC – РНК-индуцированный комплекс выключения гена (RNA-induced silencing complex), RT – обратная транскриптаза (reverse transcriptase), SINE – короткие диспергированные элементы ядра (short interspersed nuclear elements).

* Адресат для корреспонденции.

ции и ее преобразование в функциональные последовательности обеспечивали только молекулы полирибонуклеотидов. В лабораторных условиях подобные механизмы могут быть реконструированы и воспроизведены при полном отсутствии белков. И в экспериментах рибозимы оказались способными к саморепликации с экспоненциальным ростом [1]. Организмам мира РНК требовался рибозим РНК-полимераза для наследования на основе РНК и для экспрессии «генов РНК». Данный предковый фермент был утрачен в ходе эволюции. Однако ключевые функциональные аспекты репликации РНК-катализируемой РНК можно изучить «*by proxy*» с использованием генерируемых селекцией *in vitro* современных рибозимов, таких как РНК-полимераза R18 [2].

В 2016 году Horning и Joyce опубликовали данные об успешном использовании *in vitro* рибозима РНК-полимеразы. Она была способна реплицировать короткие последовательности РНК без белка, синтезируя разнообразные функциональные молекулы, в том числе аптамеры и рибозимы. Продемонстрирована полная самодостаточность репликации и экспрессии молекул РНК. Модулирование активности рибозимов *in vitro* позволило воспроизвести самоподдерживающуюся эволюцию РНК. Полученные результаты могут отражать процессы, сходные с зарождением жизни на Земле [3] при подходящих неравновесных пограничных условиях преодоления эффектов термодинамического равновесия, таких как разбавление и деградация олигонуклеотидов [4].

Предполагается, что поток тепловой энергии через открытые поры в подводных горных породах способствовал репликации олигонуклеотидов с селекцией более длинных цепочек. В эксперименте с использованием взаимодействия молекулярного термофореза и ламинарной конвекции для обеспечения экспоненциальной репликации, цепочки из 75 нуклеотидов выживали, тогда как более короткие разрушались. Взаимодействие данных двух механизмов основано на уравнивании энтропии смешивания. При разнице температур термофорез направляет молекулы горизонтально от теплой стороны к холодной. В то же время жидкость движется вертикально путем ламинарной конвекции и уносит с собой молекулы. Конвекция отклоняет горизонтальное термофоретическое истощение и усиливает вертикальное накопление молекул. Это взаимодействие молекулярного движения и потока жидкости приводит к эффективному сетевому транспорту олигонуклеотидов в нижнюю часть компартмента. В результате накопление уравнивает диффузион-

ное разбавление и разрешает проблему концентрации, связанную с возникновением жизни [4].

Мир РНК без белков функционировал еще до образования клеточных форм жизни — последний универсальный общий предок существовал ~3,9 миллиардов лет назад [5]. Несмотря на это, рибозимы остаются успешными и весьма распространенными структурно-функциональными элементами. К ним относятся консервативные (тРНК, рРНК, малые ядерные РНК (мяРНК), малые ядрышковые РНК (мяоРНК), рибозимы сплайсосомы) и изменчивые (интроны группы I и группы II, некодирующие РНК (нкРНК)) полирибонуклеотиды. Первый рибозим, интрон группы I, был открыт у инфузории *Tetrahymena* еще в 1982 году [6]. В 1983 году обнаружен рибозим рибонуклеаза Р, участвующая в процессинге молекул тРНК [7]. В 1986 году описаны рибозимы типа hammerhead (молоточковые) вирусов-сателлитов растений [8]. Их многочисленные вариации оказались распространенными у представителей всех царств живого [9]. В последние годы опубликованы работы о роли рибозимов длинных нкРНК в возникновении новых белок-кодирующих генов [10–13], регуляции работы генома в онтогенезе и взаимосвязи с мобильными генетическими элементами, транспозонами (TE — transposable elements) [14, 15]. Эти данные могут отражать древние эволюционные принципы, сформированные при возникновении жизни во время перехода мира РНК—ДНК в мир с участием белков.

Важным условием для экспоненциального роста количества и разнообразия полинуклеотидов на Земле была возможность самовоспроизведения более стойких, по сравнению с РНК, биополимеров ДНК, сохраняющих необходимые для выживания адаптивные последовательности. Ключевым событием для этого стало образование обратной транскриптазы (RT — reverse transcriptase), которая обеспечивает считывание генетической информации с полимера РНК на ДНК. Сохранение RT в геномах всех организмов говорит о ее глобальном значении на протяжении эволюции. При возникновении жизни RT могла быть рибозимом, обеспечивающим ДНК новым генетическим материалом даже от одного источника. Это связано с аккумулярованием молекулами РНК мутаций, вероятность которых значительно выше (10^{-3} – 10^{-4}), чем у ДНК (10^{-5} – 10^{-8}) [16]. Действительно, был получен рибозим РНК-полимераза, обладающая свойствами RT в качестве вторичной функции [17]. Кроме того, выявлена взаимосвязь между структурами белковых доменов RT и рибозимами [18], что говорит об их преемствен-

ности в эволюции и альтернативном использовании в идентичных биологических процессах, а также о древнем происхождении белка RT. В связи с этим можно предположить, что свойства белковых аналогов рибозимов могут отражать сходные характеристики их предковых РНК-предшественников.

Белок RT обладает активностью рибонуклеазы Н [19]. В то же время, к аналогам данного фермента, относятся Argonaute, Dicer, Cas9, транспозазы, интегразы и системы сплайсинга [20]. Поэтому, вероятно, рибозимы, обладающие свойствами данных ферментов, могли произойти от RT при ее модификациях под влиянием мутаций. Т.е. рибозим RT обладал потенциалом обеспечивать не только обратную транскрипцию, но и интеграцию образованных молекул полинуклеотидов, а также процессинг молекул РНК. Это способствовало эволюционному успеху в связи удлинением комплексных структур ДНК.

Высокая мутабельность РНК являлась причиной изменчивости первичных геномов, но представляла угрозу для сохранения адаптивных свойств полинуклеотидов, необходимых для их стабильного самовоспроизведения. Выживали такие комплексные молекулы ДНК и их тран-

скрипты, которые создавали механизмы противоборства инсерциям в наиболее важные адаптивные области ДНК. Данные защитные механизмы стали основой для возникновения коротких, регулярно чередующихся, палиндромных повторов (систем CRISPR, clustered regularly interspaced short palindromic repeats) и РНК-интерференции (РНКи), систем сплайсосомного сплайсинга и трансляции. Условием для их образования стала рибонуклеазная активность рибозима RT. Наиболее вероятно, что рРНК и тРНК, а также участвующие в их созревании рибозимы процессинга первично были направлены на защиту от инсерций (рис. 1). Подтверждением служит сохранение в эволюции взаимосвязей рРНК, тРНК, мяРНК, мяоРНК с ретроэлементами [21–25].

В настоящее время доказано, что RT играет ключевую роль в эволюции всех форм жизни. В геномах эукариот РЭ наиболее распространены и занимают от 3 до 85% всех нуклеотидных последовательностей (НП) ДНК [26]. Согласно современной классификации (<http://www.girinst.org/repbase/>), РЭ относятся к I классу ТЕ [27]. Они могут быть автономными и неавтономными. Среди последних, основную долю у эукариот занимают короткие диспергированные элементы

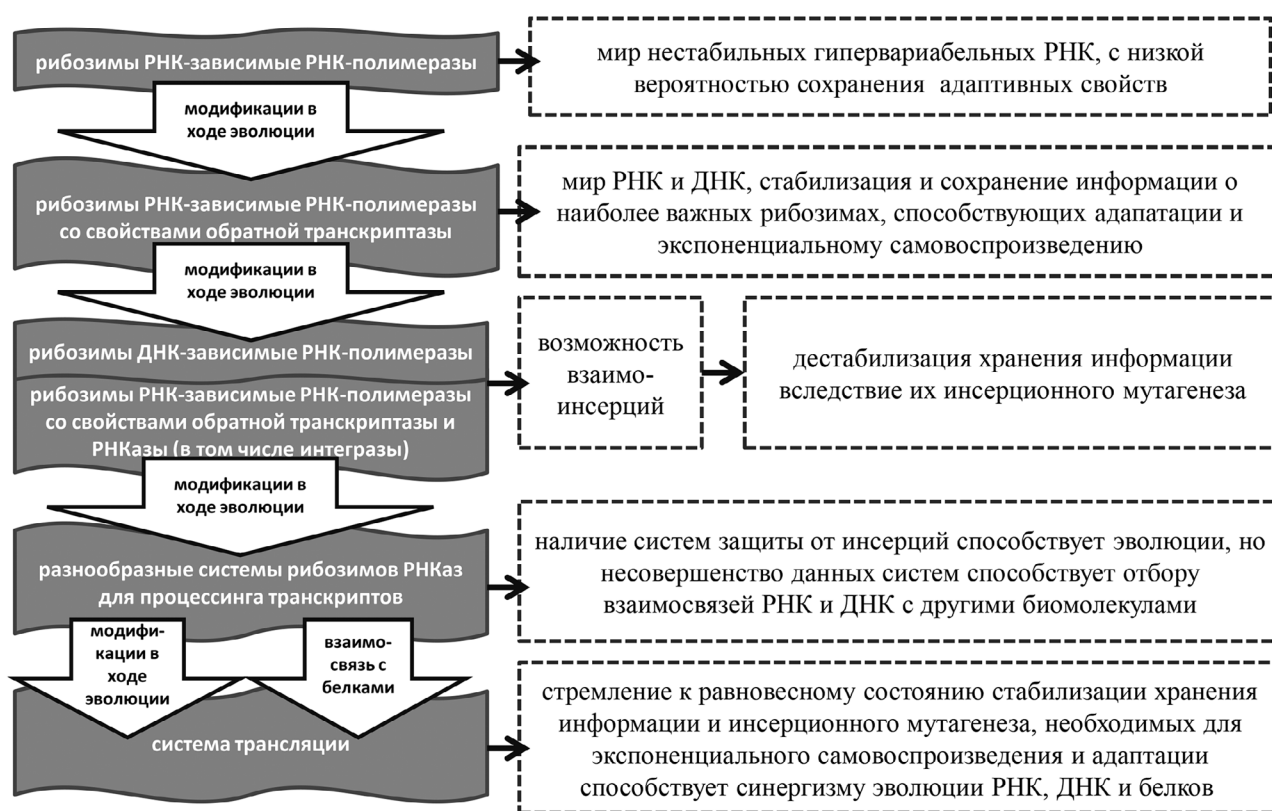


Рис. 1. Эволюция рибозимов при возникновении жизни на Земле

ядра (SINE, short interspersed nuclear elements) [28]. Они перемещаются из одного локуса генома в другие путем «копирования и вставки» при помощи ферментов длинных диспергированных элементов ядра (LINE, long interspersed nuclear elements). В то же время SINE имеют общие НП и свойства с ретроэлементами, не содержащими длинных концевых повторов (non-LTR-РЭ) [29]. SINE являются производными тРНК, мРНК, 5S рРНК [21, 24, 29–31], продукты процессинга которых участвуют в РНК-транспозонах [23, 25]. Это свидетельствует о взаимосвязи РЭ с возникновением системы трансляции на ранних этапах эволюции жизни. Подтверждением данного предположения и эволюционного родства ретроэлементов, содержащих длинные концевые повторы (LTR-РЭ) с non-LTR-РЭ, могут служить данные об использовании молекул тРНК в качестве праймеров для обратной транскрипции LTR-РЭ [28].

Ко II классу ТЕ относятся ДНК-транспозоны (ДНК-ТЕ), которые могут перемещаться по механизму «катящегося кольца» (*Helitron*, *Maverick*) и «вырезания–вставки» (*TIR*, *Crypton*) [32]. Они кодируют фермент транспозазу, катализирующую вырезание ДНК-ТЕ и ее интеграцию в новую область генома [28]. Филогенетический анализ показал, что белки ДНК-ТЕ имеют древнейшее происхождение и содержат складки (специфические структурные мотивы белка), которые существовали еще в период образования самых первых клеточных организмов [33]. При возникновении жизни основное значение имели ферменты, необходимые для амплификации биополимеров и увеличения их длины. Механизм изменчивости путем «вырезания–вставки» стал актуален на более поздних этапах эволюции, поэтому можно предположить, что ДНК-ТЕ произошли от РЭ. Подтверждением является появление у РТ в качестве вторичной активности доменов рибонуклеазы Н [19], которые характерны для сходных по молекулярной структуре транспозазы и интегразы [34, 35]. Рибонуклеаза Н относится к суперсемейству нуклеотидил-трансфераз, в состав которого входят также транспозаза, ретровирусная интеграза, резольваза и нуклеаза РНК-индуцированного комплекса выключения гена (RISC, RNA-induced silencing complex) [36]. Ферменты, подобные рибонуклеазе Н, являются участниками универсальных консервативных для всего живого защитных систем против вирусов и ТЕ, включая RISC (*Argonaute*, *Dicer*) и CRISPR (*Cas9*) [20]. В связи с возможной преемственностью доменов РНК и белков в эволюции [18], рибозимы РТ также могли обладать свойствами одновременно репликаз и рибонуклеаз,

при модификации которых возникали новые системы процессинга.

Предполагается, что решающим событием в происхождении жизни было появление молекулы РНК, способной реплицировать первобытный «геном» РНК. При этом самые ранние и простые биологические процессы зависели от РНК как для наследственности, так и для метаболизма. Доказательствами мира РНК на ранних этапах эволюции могут служить современные данные о центральной каталитической роли РНК в сплайсинге, экспрессии генов и трансляции, а также о многосторонних функциях РНК в формировании специфических рецепторов и катализаторов [2]. При зарождении жизни в качестве универсальных единиц могли служить ТЕ-подобные элементы. Их взаимодействие и противоборство при формировании комплексных первичных геномов послужило основой для всех последующих эволюционных процессов. Однако из-за высокого уровня мутабельности НП транспозонов [16] их глобальную роль в формировании геномов современных организмов трудно идентифицировать. В работе de Koning et al. при анализе ДНК человека с помощью специфических олигонуклеотидов, было обнаружено, что ТЕ занимают >67% всех НП, а не 45%, как считалось ранее [37]. Современные методы анализа позволяют выявлять «скрытые» ТЕ в геномах разных видов животных для определения их роли в формировании регуляторных элементов и белок-кодирующих генов в ходе эволюции. В результате обнаруживается значительно большее количество НП, произошедших от ТЕ [38].

Вероятно, вся имеющаяся в мире биомасса была образована благодаря ТЕ в качестве универсальных единиц, начиная с зарождения жизни на Земле. В качестве подтверждения можно привести работу Horning и Joyce [3], в которой была спроектирована искусственная эволюция РНК *in vitro* путем уменьшения точности работы рибозима в безбелковой среде с формированием 10% ошибок при элонгации новой цепи РНК. Случайные мутации на каждый десятый нуклеотид по всем молекулам РНК привели к образованию 10^{14} различных вариантов РНК. Это стало доказательством эффективности эволюции при участии полимераз, способных образовывать большое количество ошибок [3]. В результате поддерживалась эффективная изменчивость при переходе из мира РНК в мир с участием ДНК, РНК и белков при возникновении жизни на Земле. Это согласуется с характерными свойствами большинства РТ образовывать ошибки из-за отсутствия корректирующего домена (3'–5' экзонуклеазы). Эволюцион-

ными предшественниками РТ считаются РНК-зависимые РНК-полимеразы, которые, как правило, тоже не содержат 3'-5' экзонуклеазного домена, ответственного за корректирующую активность [39]. Вероятно, сохранение способности формировать ошибки говорит о сохранении древних свойств РТ, играющих глобальную роль в эволюции всего живого. Сходство доменов рибозимов и их белковых аналогов [18] позволяет предположить, что в мире РНК обратная транскриптаза также обладала данным свойством.

Согласно альтернативной гипотезе, предложенной в 1999 году Freeland et al. [40], ДНК в триаде РНК—ДНК—белок появилась последней. Преимуществом предлагаемой авторами гипотезы является их утверждение, что биохимическое восстановление рибонуклеотидов до дезоксирибонуклеотидов выходит за рамки каталитических способностей РНК [40]. Однако авторы не учли возможности абиогенеза мономеров ДНК без промежуточной стадии рибонуклеотидов. Кроме того, было доказано, что для механизма трансляции требуется большее количество наследственной информации, чем может поддерживаться РНК-геномами [17]. Это обосновывает предлагаемую нами гипотезу, согласно которой белок в эволюции появляется последним. Наиболее вероятно, система трансляции возникла, как один из множества механизмов защиты первичных геномов ДНК от инсерций с использованием каталитических доменов полипептидов в комплексе с рибозимами. В пользу этого предположения говорят данные о роли продуктов процессинга тРНК, мяоРНК, рРНК [23, 25] в РНКи транспозонов.

ОСОБЕННОСТИ ОБРАТНОЙ ТРАНСКРИПТАЗЫ

Семейство РТ имеет единое древнее эволюционное происхождение, что было доказано на основании гомологии аминокислот и наличия РТ во всех доменах жизни [39]. Фермент РТ ответственна за «превращение» РНК в кДНК и необходима для инвазии РЭ и их распространения в геномах. РТ впервые описана в 1970 году двумя исследовательскими группами у РНК-содержащего вируса лейкемии мышей и вируса саркомы Рауса [41, 42]. Бактериальные РТ были открыты у РЭ, известных как ретроны, в 1989 году на *Mycobacterium xanthus* [43] и *Escherichia coli* B [44]. Более 50% идентифицированных на сегодняшний день бактериальных РТ кодируются интронами группы II, которые являются мо-бильными РЭ со свойствами рибозимов [45]. РТ

послужила не только ключевым звеном для перехода мира РНК в мир РНК—ДНК—белки, но и сохранила свои универсальные способности ускорять эволюцию на всех ее последующих этапах. Предшественниками РТ считаются РНК-зависимые РНК-полимеразы [39], которые при возникновении жизни были представлены рибозимами [2, 3]. Поэтому древняя РТ также могла быть рибозимом, способным синтезировать ДНК на основе матрицы РНК, что и было доказано [17].

Использование нуклеозидтрифосфатов для удлинения молекул РНК или ДНК является фундаментальным свойством для всех живых организмов. Реакция при помощи РНК-рибозима сходна с функционированием теломеразы, специализированной ДНК-полимеразы, в состав которой входит короткая РНК, служащая матрицей для синтеза ДНК теломерных повторов. Предполагается, что рибозимы, сходные по активности с нуклеотидил-трансферазой, использовались как теломеразы еще до появления белкового катализа. Хотя тРНК нуклеотидил-трансферазы прежде всего ответственны за добавление и поддержание ССА-3' хвоста акцепторного стебля тРНК, данный клеточный фермент может функционировать в качестве «теломеразы» для РНК вирусных геномов, которые принимают тРНК-подобные структуры на их 3'-концах [46]. В этом отношении имеются данные о тесной связи тРНК с РЭ у эукариот [21, 25, 28], что может быть отражением их эволюционного родства. В качестве универсальных молекул-посредников тРНК первоначально могли использоваться, как необходимое звено для полимеризации РНК и ДНК (возможно, еще до появления рибозима интегразы и удлинения ДНК при помощи инсерций). Дальнейший отбор определенных тРНК, обладающих свойством взаимодействия с аминокислотами при участии различных рибозимов, позволил объединить мир полинуклеотидов с миром белков за счет универсального для всего живого триплетного кода. Подтверждением данного предположения служит тесная взаимосвязь тРНК с ретроэлементами, которые содержат комплементарные им НП и используют тРНК в качестве праймеров [19]. В живой природе у инфузорий *Tetrahymena thermophila* был выделен естественный рибозим, катализирующий полимеризацию РНК за счет удлинения РНК-праймера. Это является одним из доказательств теории саморепликации пребиотической РНК [47]. В лабораторных условиях было показано, что нестабильные нефункциональные комплексы, собранные из более коротких 3'-усеченных олигонуклеотидов, могут стабилизироваться и обре-

тать функциональность путем неферментативного удлинения праймера [48].

Предполагается, что, в отличие от RT, у других семейств ДНК-полимераз в ходе эволюции развились корректирующие механизмы для повышения точности синтеза ДНК при репликации генома. Чтобы определить, является ли отсутствие корректирующей активности историческим совпадением или функциональным ограничением обратной транскрипции, была разработана термостабильная ДНК-полимераза высокой надежности для эффективного использования матричных РНК. В результате оказалось, что корректура совместима с обратной транскрипцией [39]. Вероятно, полученные данные свидетельствуют об эволюционном сохранении древнего типа RT, так как амплификация с ее помощью сопровождается выраженной мутабельностью, необходимой для изменчивости и обретения новых адаптивных способностей. Это универсальное свойство, обеспечивающее амплификацию ДНК с ее изменчивостью, могло быть ключевым моментом для возникновения жизни.

Сходство структур белковых доменов RT и рибозимов [18] говорит о возможных эволюционных путях преобразований каталитических РНК-молекул в белковые, об их тесной взаимосвязи и взаимозаменяемости (рис. 2). Предшественниками RT считаются РНК-зависимые РНК-полимеразы [39], предковыми формами которых были РНК-рибозимы. Поэтому логично предположить наличие рибозимов RT в живой природе. Действительно, получены доказательства того, что высокоорганизованный рибо-

зим РНК-полимераза способна функционировать в качестве RT. Данная RT-активность имела решающее значение для перехода от РНК к ДНК геномам при возникновении жизни на Земле. Предполагается, что она могла возникнуть как вторичная функция рибозима РНК-зависимой РНК полимеразы [17]. Хотя РЭ кодируют и другие белки, повышающие эффективность обратной транскрипции, их RT обладает двойной ферментативной активностью. В качестве ДНК-полимеразы она может копировать как РНК, так и ДНК матрицу, а в качестве рибонуклеазы Н расщепляет РНК в составе гибрида РНК/ДНК [19].

Наличие у RT активности рибонуклеазы Н позволяет предположить ее глобальную роль в возникновении ключевых систем процессинга транскриптов живых систем. Это связано с принадлежностью рибонуклеазы Н к суперсемейству нуклеотидил-трансфераз, к которому относятся транспозаза, ретровирусная интеграза, резольваза и RISC-нуклеаза Argonaute [36]. В эволюции имеется преемственность строения пространственных доменов РНК-рибозимов и белковых ферментов, выполняющих аналогичную функцию [18]. Поэтому весьма вероятно, что при возникновении жизни модификации рибозима RT могли обладать рибонуклеазной активностью и стать основой для систем процессинга.

При сравнении структур ретровирусных интеграз, транспозаз и сайтов связывания с ними в ДНК, оказалось, что оба фермента могут обладать общими биохимическими и генетическими особенностями. Это говорит о возможном

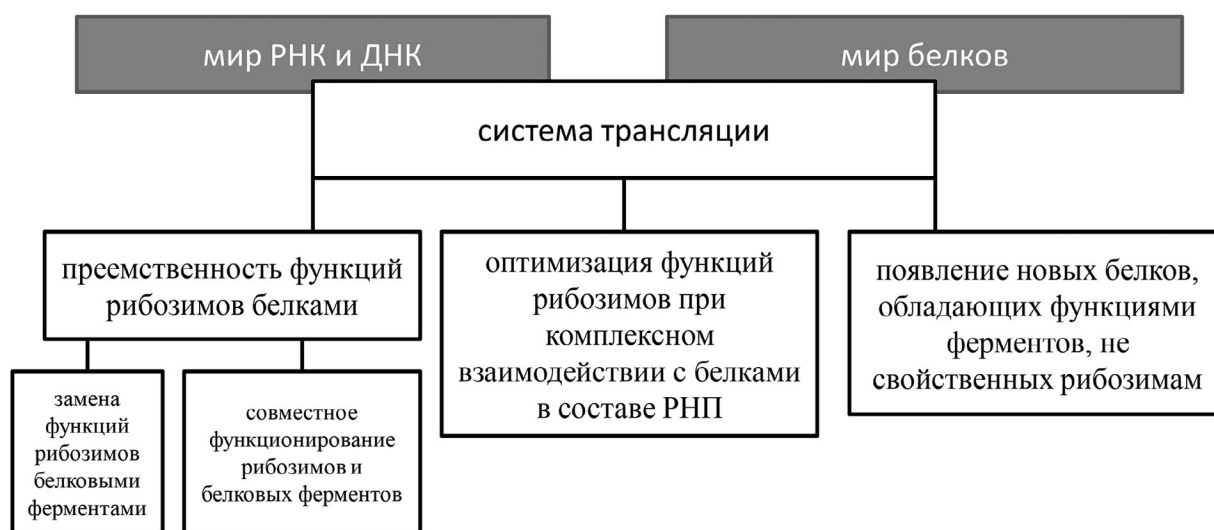


Рис. 2. Взаимосвязь рибозимов и белковых ферментов в эволюции

их эволюционном происхождении от единых предковых последовательностей [34]. Транспозаза и ретровирусная интеграза имеют общий домен, подобный рибонуклеазе H, с каталитической триадой DDE/D, которая координирует двухвалентные катионы, необходимые для расщепления ДНК и интеграции [35]. Описаны химические рибозимы, обладающие свойствами разных ферментов, что подтверждает возможность многофункциональности древней RT. Так, у бактерии *Clostridium difficile* был обнаружен рибозим, имеющий черты интронов группы I и инсерционных элементов [49]. РЭ кодируют интегразу и RT как единый полипептид, который посттранскрипционно процессируется при помощи кодируемой этим же РЭ протеазой [50]. Приведенные данные позволяют предположить, что модификации рибозима RT могли обладать дополнительными свойствами рибонуклеаз и интеграз. Механистическое сходство ферментов ДНК-транспозазы и интегразы РЭ подчеркивает их тесную эволюционную взаимосвязь. Считается, что ретровирусы произошли от LTR-РЭ Ty3/Gypsy. Копии их ДНК можно найти в геномах большинства живых существ [51]. Эволюционное родство и взаимосвязь различных ферментов, ответственных за транспозиции, свидетельствуют о возможном сценарии возникновения жизни от РЭ в качестве структурно-функциональных единиц, от которых произошли другие ТЕ. Их противоборство и взаимодействие, наряду со способностью к амплификации за счет RT, могло стать основой для эволюции.

Взаимопаразитизм ТЕ привел к распространению первично не кодирующих необходимых для транспозиции ферментов неавтономных РЭ. Эти элементы содержат НП, играющие важную роль в регуляторных процессах. Например, Alu РЭ ассоциированы со стимуляцией экспрессии генов и обладают высоким потенциалом регуляции в зависимости от состояния хроматина [52]. Alu относятся к SINE и произошли в эволюции от обратнотранскрибированных 7SL РНК. Последние входят в состав РНП в качестве рибозима и составной части сигнал-распознающей частицы (SRP), необходимой для транспорта рибосом через трансмембранные поры. Интересно, что вторичные структуры SRP животных, растений, грибов, бактерий и архей очень схожи [22], что говорит о древнем эволюционном происхождении 7SL РНК. Это дает основание предположить возникновение SRP от ТЕ, самых первых универсальных источников всех генов в эволюции. Alu и другие SINE, а также 7SL РНК процессируются ферментом Dicer с образованием малых нкРНК [53].

ВОЗНИКНОВЕНИЕ СИСТЕМЫ ТРАНСЛЯЦИИ И ПРЕЕМСТВЕННОСТЬ ФУНКЦИЙ РИБОЗИМОВ БЕЛКАМИ

Отбор и возникновение рибозимов RT из РНК-полимераз был необходимым этапом для сохранения и преумножения компонентов систем автономно реплицирующихся РНК. Это обусловлено большей стабильностью молекул ДНК с сохранением адаптивного генетического материала. При его считывании обратно в РНК при помощи рибозима ДНК-зависимой РНК-полимеразы образовывались успешные для амплификации молекулы полинуклеотидов. Подобно миру РНК, в котором при модификации рибозима RT возникла интегразная функция, множественные инсерции способствовали эволюции и преумножению систем автономно реплицирующихся полинуклеотидов. Однако это стало угрозой для стабильных ДНК-геномов вследствие накопления мутаций, когда вновь возникшие адаптивные свойства нарушались новыми инсерциями. В результате, в ходе отбора стали возникать механизмы противоборства геномов в виде систем процессинга РНК-молекул. Основой для данных способов защиты являлись модифицированные доменные структуры RT-рибонуклеазы. Наиболее вероятно, рибозимы РНК-полимеразы и RT обладали способностью при их модификации вследствие мутабельности становиться источниками различных систем процессинга, направленных на защиту от инсерций. Отражением данной способности являются сходные процессы в мире белков, так как белок RT обладает свойствами рибонуклеазы H [19]. Члены ее суперсемейства являются основными участниками систем процессинга транскриптов всего живого (РНК-интерференции, CRISPR, сплайсинга [20, 36]).

Одним из механизмов процессинга и защиты в мире РНК-ДНК могла стать система трансляции. Этим можно объяснить преемственность доменов рибозимов и белков в эволюции, их взаимосвязь в регуляции единых биологических процессов. Возможность управления при помощи белков реакциями, в которых не участвовали рибозимы, также подтверждает данное предположение, так как отбор и консервативность полипептидов с целью усовершенствования защиты подразумевает преимущество их использования перед рибозимами. Первоначально рРНК, тРНК и процессирующие их рибозимы могли произойти от ТЕ и выполнять иные функции, необходимые для противоборства первичных геномов, в том числе в составе других систем защиты. Так, у РЭ и ретровирусов тРНК используется в качестве праймеров [19],

что может отражать древнее эволюционное свойство тРНК. Кроме того, рРНК [23, 54], мяРНК [23], тРНК [55] и мяоРНК [56, 57] подвергаются неслучайному процессингу с использованием полученных продуктов для РНКи транспозонов [23, 25, 55, 58]. Рибозимы могут участвовать в разных процессах и выполнять самые разнообразные функции. Например, тРНК используется не только в качестве праймера, но также для транспортировки аминокислот и РНКи. Белковые ферменты также могут участвовать в различных процессах и выполнять различные функции – RT теломеразы (TERT), помимо удлинения теломер, используется для регуляции транскрипции генов [59]. Эволюционная взаимосвязь рибозимов системы трансляции с транспозонами отражается также в сайт-специфических особенностях интеграций ТЕ. Напри-

мер, R-element, MITE, *Pokey* характеризуются инсерциями в мультигенные рДНК [60], а ретроэлемент R2 – в области генов 28S рДНК [61].

При рассмотрении других систем процессинга РНК, обнаруживается та же закономерность тесной связи и эволюционного родства всех их компонентов с ТЕ. К примеру, сплайсосомные интроны эволюционировали от интронов группы II. Последние являются ТЕ со свойствами рибозимов и RT активностью [62]. В то же время у эукариот ТЕ служат источниками сплайсосомных интронов [63, 64], сигналов сплайсинга [65, 66], энхансеров и сайленсеров сплайсинга [67, 68], структурно-функциональных компонентов самой сплайсосомы [69]. Кроме того, подобно ТЕ, сплайсосомные интроны подвергаются процессингу специфическими рибонуклеазами с образованием нкРНК [70].

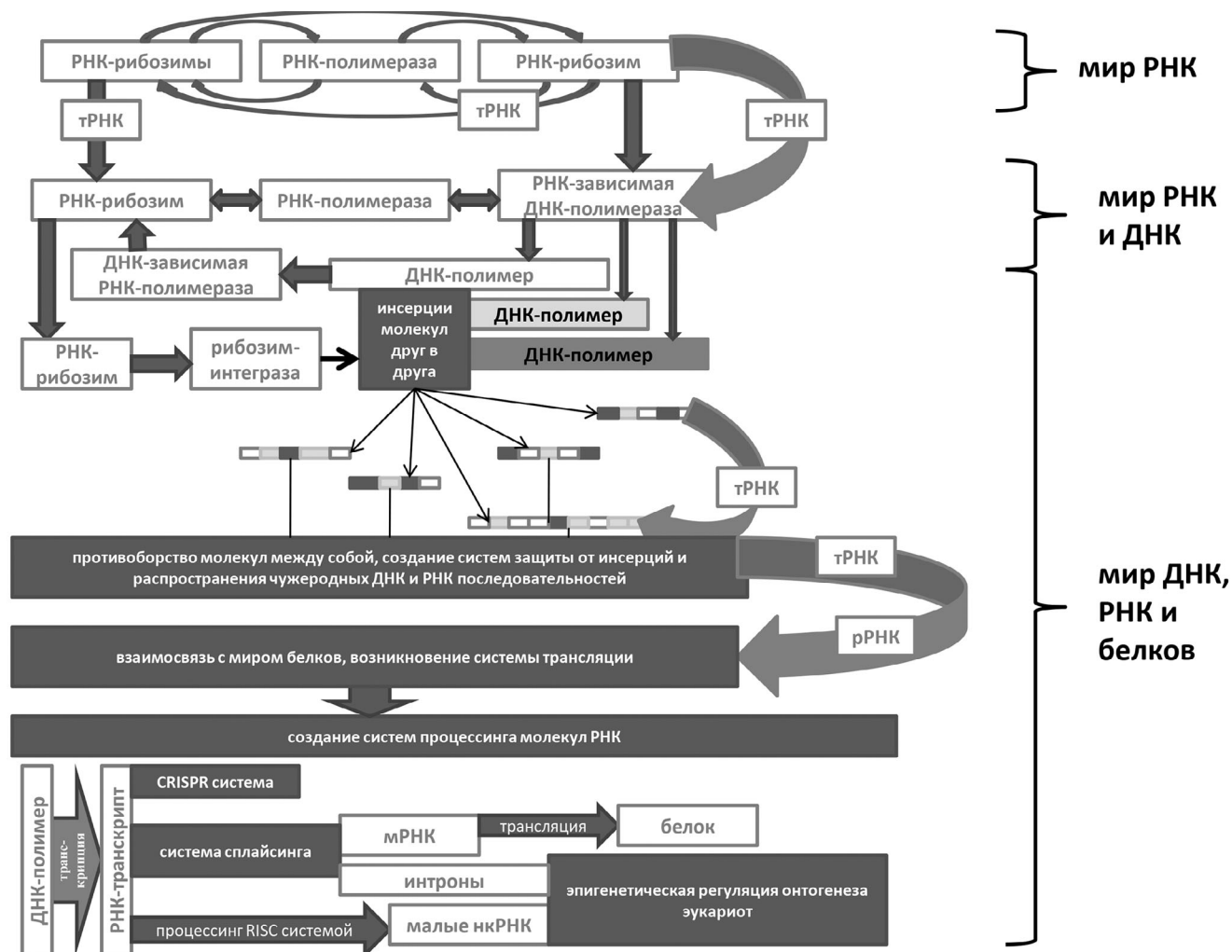


Рис. 3. Схема эволюции систем процессинга как защитных механизмов

Благодаря возникновению системы трансляции, живые организмы перешли на новый уровень структурно-функциональной организации. В результате, в противоборствующих геномных системах белковые ферменты могли замещать функции рибозимов, действовать в комплексе с ними в виде РНП и образовывать домены с новыми не присущими рибозимам функциями. Несмотря на это, в эволюции сохранились консервативные, специфичные исключительно для рибозимов способности, которые служат доказательствами их роли в возникновении системы трансляции. Так, в составе рибосомы синтезом белка управляют непосредственно молекулы рРНК. Было показано, что каталитическая активность рибосом обусловлена активными сайтами, состоящими только из РНК. В отличие от рРНК, рибосомальные белки находятся на значительно большем расстоянии от активных центров рибосом [71]. Не исключено, что рибосомальные рибозимы произошли от модифицированных ТЕ в ходе формирования систем противоборства с другими эгоистическими элементами. Подтверждением является участие нкРНК, образованных при процессинге рРНК, в РНКи транспозонов [23]. Кроме того, наиболее вероятно, что эволюционными предшественниками тРНК также являлись ТЕ. Доказательством служит как использование процессированных транскриптов тРНК в сайленсинге ТЕ (что говорит о наличии между ними комплементарности НП) [25, 55], так и использование тРНК [21] и рРНК [21, 29–31] в геномах в качестве основы для неавтономных ТЕ.

Таким образом, взаимосвязь мира РНК–ДНК и белков возникла благодаря RT и содержащим ее универсальным консервативным структурам – РЭ и их производным (ДНК–ТЕ). Успешное распространение и сохранение в эволюции обратнотранскрибированных рРНК и тРНК в качестве неавтономных ТЕ в геномах различных эукариот подтверждает их эволюционное родство с другими ТЕ. Это обусловлено общими свойствами с автономными ТЕ, а также наличием гомологичных НП в их составе, благодаря чему они активно распространяются и сохраняются в ряду поколений [23]. Интересно, что 3'-конец SINE3, произошедший от 5S рРНК, имеет значительное сходство с CR1-подобными non-LTR РЭ. Наряду с CR1-подобным РЭ, копии SINE3 не фланкированы дубликациями целевых сайтов, а их 3'-конец состоит из микросателлитов (ACATT)_n и (ATT)_n [29]. В геномах млекопитающих выявлены также транскрипционно активные SINE28, происходящие от 3'-конца большой рибосомальной субъеди-

ницы (28S) [31]. У насекомых описаны химерные SINE, HaSE3, состоящие из тРНК и 5S рРНК [30]. Кроме того, в геномах крокодилов, аллигаторов и гавиалов идентифицированы SINEU, происходящие от мяРНК U1 и U2 [24], что может свидетельствовать об эволюционном родстве системы трансляции и сплайсинга. Это согласуется с тем, что мяРНК также подвергаются неслучайному процессингу [23] для РНКи последовательностей ТЕ в геноме. Таким образом, при возникновении жизни компоненты, необходимые для взаимосвязи мира РНК и ДНК с белками, были образованы благодаря RT и ТЕ. Основные принципы данных закономерностей сохраняются в эволюции (рис. 3).

ЗНАЧЕНИЕ ОБРАТНОЙ ТРАНСКРИПТАЗЫ В ЭВОЛЮЦИИ ПРОКАРИОТ

Хотя RT считаются ферментами эукариот, они весьма распространены у бактерий [72]. RT обеспечивают обмен генами между плазмидами и нуклеоидом, что необходимо для передачи адаптивных признаков между отдельными клетками и их потомками [73]. РЭ (ретроны) в геномах прокариот впервые описаны еще в 1989 году у *E. coli* [44] и у *M. xanthus* [43]. В 2002 году выявлен РЭ прокариот, названный DGR (diversity-generating retroelement) [74]. Филогенетический анализ показал, что бактериальные RT могут быть классифицированы на 17 групп: RT интронов группы II, ретронов и ретрон-подобных элементов, DGR RT, Abi-like RT, RT систем CRISPR-Cas, G2L (group II-like RT) и 11 других групп RT с неизвестными функциями [75]. В геномах бактерий идентифицировано множество не охарактеризованных RT и связанных с ними последовательностей, которые демонстрируют большое разнообразие доменных структур. За исключением интронов группы II, содержащие RT элементы не проявляют признаков активной ретромобильности, но вовлечены в защиту геномов от фагов [72]. Эти геномные структуры сохранили древнее эволюционное свойство, направленное на противоборство эгоистических геномов, в ходе которого образовывались новые защитные системы. У бактерий >50% RT, которые могут быть использованы для генерации комплементарной ДНК из матрицы РНК, кодируются интронами группы II. Данные интроны действуют как рибозимы и мобильные РЭ и классифицируются на A, B, C, D, E, F, G, CL1/2 (chloroplast-like), ML (mitochondrion-like) группы. Все они эволюционировали при помощи ассоциации древней RT со структурной каталитической РНК [75].

Интроны группы II кодируют собственную RT, которая стабилизирует каталитически активную РНК для прямого и обратного сплайсинга и конвертирует интегрированную интронную РНК обратно в ДНК [76]. Впервые интроны группы II были выделены еще в 1986 году из транскрипта митохондрии *Saccharomyces cerevisiae* [77]. Предполагается, что данные элементы являлись эволюционными предшественниками не только сплайсосомных интронов, но также самих сплайсосом, РЭ и теломер эукариот [76]. Структура каталитического домена их RT сильно напоминает теломеразу, а активный центр сплайсинга — белок Prp8 сплайсосомы. Это говорит о предковых взаимосвязях между сплайсингом и ретромобильностью [18]. Кроме того, показано, что белок Prp8 в эволюции произошел от RT транспозонов [69].

Была обнаружена взаимосвязь RT системы CRISPR-Cas типа III с кодируемыми интронами группы II обратными транскриптазами. Ассоциация RT с системой CRISPR-Cas происходила множество раз в ходе эволюции. При филогенетическом анализе RT системы CRISPR были распределены на 12 групп. При этом выявлена коэволюция RT с белками Cas1, а также их горизонтальный перенос в геномы архей [45]. RT содержит последовательности, обладающие свойствами рибонуклеазы H, к суперсемейству нуклеотидил-трансфераз которых относятся Cas9 [20]. Поэтому можно предположить, что в эволюции формирование CRISPR происходило из защитных механизмов одних РЭ против других ТЕ. Прокариоты противоборствуют с вирусами при помощи врожденных иммунных систем модификации рецепторов, рестрикции-модификации и abortивной инфекции. РНК-опосредованная адаптивная иммунная система CRISPR-Cas содержится у большинства архей и ~40% бактерий [45]. RT в данной системе играет важную роль для обратной транскрипции чужеродных РНК [78].

У прокариот интеграция новых НП (спейсеров) при помощи CRISPR сопровождается удалением других частей генома, состоящих из повторов, что позволяет сохранять относительно равноценные размеры ДНК их нуклеоидов. В результате максимальная длина генома у *M. xanthus* больше минимальной у *Micoplasma genitalium* всего в 16 раз. Эволюционный успех, огромное разнообразие фенотипов и грандиозные различия в размерах геномов эукариот могут объясняться глобальной ролью инсерций в их эволюции и функционировании, что может отражать механизмы, происходившие при возникновении жизни на Земле. Размеры геномов эукариот могут отличаться в 70 000 раз (от 2,3

мегабаз у *S. cerevisiae* до 148 852 мегабаз у *Paris japonica*) [79]. Однако была обнаружена выраженная гомология между системами CRISPR и РНКи с участием малых интерферирующих РНК [20]. Можно предположить, что у общего предка всех доменов живого существовал защитный механизм, модификации которого дали основу для развития CRISPR у прокариот и RISC эукариот. Специфика их функционирования повлияла на структурные особенности геномов разных доменов. Несмотря на длительный период эволюции с момента возникновения жизни, в современных организмах сохраняются процессы, отражающие преемственность функций рибозимов белками на начальных этапах. Наиболее ярким примером являются длинные нкРНК.

ЗНАЧЕНИЕ ОБРАТНОЙ ТРАНСКРИПТАЗЫ В ЭВОЛЮЦИИ ЭУКАРИОТ

РЭ эукариот классифицируются на 2 основных класса по механизму их транспозиций и организации ДНК. К первому классу относятся LTR-РЭ, которые характеризуются прямыми повторами в несколько сотен п.н. на их концах и содержат гены *pol* и *gag*. Продукт гена *pol* содержит несколько ферментных доменов (RT, рибонуклеаза H, интегразы и протеиназы). Ген *gag* кодирует белки GAG (Group-specific AntiGen). Представители второго класса, nonLTR-РЭ, могут содержать одну открытую рамку считывания (ORF, Open Reading Frame), в которой закодирован белок с доменами RT и EN (эндонуклеазы) или две ORF (первая проявляет сходство с геном *gag* ретровирусов, вторая кодирует домены RT и EN). Предполагается, что LTR-РЭ эволюционировали от non-LTR РЭ за счет возникновения интегразы, а ретровирусы произошли от LTR-РЭ с помощью включения генов оболочки (*env* доменов) от других вирусов [45]. При секвенировании геномов было обнаружено, что эукариоты могут содержать больше генов RT по сравнению с кодирующими любой другой белок. Около 40% НП геномов млекопитающих представлены РЭ, в основном LINE и SINE. У растений содержание ТЕ может достигать 91% их ДНК (*Asparagus officinalis*), главным образом за счет LTR-РЭ [80].

RT эукариот вовлечены в такие процессы, как добавление теломер, репликация митохондриальных плазмид, транскрипция и пролиферация ретровирусных геномов [39], дупликация генов [80]. RT составляет основу РЭ, которые у эукариот служат ключевыми источниками возникновения как регуляторных НП [65], так и бе-

лок-кодирующих генов путем доместикации самих РЭ [81, 82] и экзонизации их НП [66, 83, 84]. РЭ оказались источниками важнейших консервативных свойств геномов эукариот, отличающих их от прокариотических. Так, от РЭ произошли центромерные сателлиты [85] и центромер-связывающие белки CENP/CENH3 [86], с которыми взаимодействуют содержащиеся в центромерах РЭ [87], теломеры [88], и теломераза [89], транскрипционные факторы [65] и сайты связывания с ними [90]. От РЭ и других ТЕ происходят многие микроРНК [91], а также их целевые НП в составе белок-кодирующих генов, что связано с ролью ТЕ в их возникновении [65, 66, 81–84]. Не исключено, что, подобно сплайсосоме, белки процессинга нкРНК, а также основные компоненты системы РНКи в эволюции также произошли от ТЕ. Таким образом, на примере механизмов эволюции геномов эукариот можно предположить, что РЭ и входящая в их состав РТ могли быть источниками происхождения всей жизни. Это связано с использованием РЭ в качестве универсальных единиц, ставших консервативными на всех этапах эволюции. Для эукариот характерно функционирование длинных нкРНК. Они являются наилучшим примером, отражающим принципы участия ТЕ в возникновении саморегуляторных систем и преемственности рибозимов белковыми ферментами.

Длинные нкРНК представляют собой рибозимы, способные функционировать как самостоятельно, так и в составе РНП. Они участвуют в регуляции транскрипции, воздействуя на модификацию гистонов и ДНК-связывающие комплексы, в том числе транскрипционные факторы (ТФ) [92]. При этом НП многих процессированных транскриптов длинных нкРНК полностью совпадают с ТЕ [15], а >80% активных доменов всех длинных нкРНК происходят от транспозонов [14]. РЭ способны функционировать непосредственно в качестве генов длинных нкРНК [93, 94]. Т.е. процессированные транскрипты РЭ проявляют свойства рибозимов, что может отражать их древнюю консервативную характеристику, ставшую ключевым моментом для возникновения жизни. Исследование роли длинных нкРНК в образовании новых белок-кодирующих генов может отображать сходные процессы на ранних этапах эволюции, когда переход из мира РНК и ДНК в мир с участием белков характеризовался преемственностью функций рибозимов полипептидными ферментами. В независимых исследованиях выявлено происхождение эволюционно новых белков эукариот от продуктов трансляции транскриптов длинных нкРНК [10–13]. Это обус-

ловлено способностью длинных нкРНК связываться с рибосомами и транслироваться в функциональные пептиды [95].

Благодаря РТ в геномах эукариот активно распространяются РЭ, которые становятся источниками биохимически активных строго регулируемых некодирующих элементов, таких как сайты связывания с ТФ и нкРНК. В ходе эволюции это способствует перестройке генных регуляторных сетей, за счет чего формируется основа для изменчивости и образования новых видов [96]. Кроме того, РЭ являются важными источниками гибридного дисгенеза при видообразовании. Были выявлены вспышки транспозиций, наблюдаемые после гибридизации и способствующие генетической нестабильности. Связь гибридных транспозиций и деметилирования наблюдается как у млекопитающих, так и у растений [97].

Подобно длинным нкРНК, молекулы пре-мРНК белок-кодирующих генов также могут формировать различные пространственные конфигурации. При этом их вторичные и третичные структуры играют важную роль в процессинге и стабильности пре-мРНК [98], что позволяет предположить функциональную значимость эволюционных предков этих РНК и их древнее происхождение от ТЕ. Например, насыщенность доменов длинных нкРНК последовательностями ТЕ может быть связана с их способностью образовывать конформации, участвующие в важных биологических реакциях [14]. Для предсказания вторичных структур РНК разрабатываются различные алгоритмы, в которых используются минимальная свободная энергия и максимально ожидаемая точность со сравнительными эволюционными методами. Однако эти инструменты не идеальны. Вторичная структура пре-мРНК может оказывать усиливающее или подавляющее воздействие на сплайсинг собственной молекулы в зависимости от особенностей НП в интронах. После сплайсинга вторичная структура РНК также может влиять на стабильность и регуляцию РНКи. На сплайсинг молекул пре-мРНК оказывает воздействие также их третичная структура. Примером является G-квадруплекс, который усиливает или подавляет сплайсинг путем создания или затенения участков связывания РНК с белком [98]. Сходными способностями обладают РЭ. Они могут управлять распознаванием сайтов сплайсинга за счет образования специфических пространственных структур в виде шпилек [99].

Характерен также процессинг интронов белок-кодирующих генов в функциональные нкРНК [70], что объединяет их по свойствам с ТЕ [91]. Сохранение в эволюции РНКи при по-

мощи процессированных транскриптов ТЕ говорит об универсальности этого механизма. Активное использование и консервативность кодирования аминокислот нуклеотидами могли быть связаны с возможностью взаимодействий пептидов с первичной и вторичной структурами ДНК [100] для обеспечения защитных механизмов. В различных биологических реакции более успешны или рибозимы или белки. В тех процессах, где рибозимы обладают уникальными свойствами, они сохранили свои функции. Однако вновь возникающие белковые ферменты способны заместить эти молекулы РНК для лучшей адаптации. Данное состояние неравновесия является источником эволюционных преобразований и создания новых белок-кодирующих генов из областей ДНК, содержащих гены нкРНК [10–13].

Считается, что большинство RT-подобных ферментов принадлежат к ретроэлементам или вирусам и не обладают определенной функцией в клетке-хозяине, за исключением теломеразы. Однако был обнаружен уникальный класс RT-связанных клеточных генов, названных *rvf*. Они являются генами «хозяев» и уже не служат компонентами РЭ. Гены *rvf* представлены единственной копией и могут содержать интроны в эволюционно консервативных позициях, изменяясь под действием отбора. Эти гены обнаруживаются во всех основных таксономических группах, включая протисты, грибы, животные, растения и бактерии с неоднородным филогенетическим распределением [101]. RT используется в создании генов с выраженной амплификацией, таких как *MADS-box* и цитохромы P450 [102]. Гены гистонов также характеризуются кластерной организацией и не содержат интронов, что позволяет предположить их происхождение от ТЕ у общего предка всех эукариот. Подтверждением служит возникновение от ТЕ гена *CENP/CENH*, продукт которого является вариантом гистона H3 [86]. Малые нкРНК способны оказывать свое эпигенетическое воздействие путем модификаций гистонов [103], в то время как сами ТЕ служат источниками нкРНК [91]. Для разных типов ТЕ характерны специфические модификации гистонов в областях их расположения (LTR и LINE – метки H3K9me3, другие ТЕ – H3K27me3) [104], а распределение длины консервативных и неконсервативных областей интронов имеют максимумы, близкие по длине динуклеосомной и нуклеосомной ДНК [105]. Приведенные данные свидетельствуют об универсальных свойствах ТЕ в качестве консервативных единиц всего живого, способных создавать саморегуляторные комплексные структуры, которые эффективно эволюционируют.

Участие гистонов в регуляции работы генома отражает принцип использования белков при возникновении жизни, как необходимых компонентов не только для преемственности функций рибозимов, но и для создания новых, оптимально адаптивных свойств. Кроме того, от ТЕ произошли гены, принимающие непосредственное участие в ремоделировании хроматина [106]. ТЕ оказались источниками возникновения ацетилтрансферазного комплекса HDP1/2 [107], факторов модификации хроматина BEAF-32 и HIM-17 [108], центромерных белков Abp1 [65] и инсуляторов [109].

В литературе накоплены многочисленные данные об универсальности RT в качестве ключевого звена для взаимосвязи мира РНК с ДНК при возникновении жизни. В эволюции бактерий, архей и эукариот проявляется консервативность использования RT для формирования новых принципов саморегуляции и самовоспроизведения, что может отражать сходные механизмы предковых доклеточных форм жизни. Модификации рибозима РНК-полимеразы могли обладать свойствами RT, рибонуклеазы и интегразы для поддержания самовоспроизведения биополимеров РНК и ДНК. Однако сохранение информации об адаптивных свойствах, необходимых для выживания биомолекул на базе ДНК, нуждалось в балансе между инсерциями (для изменчивости и увеличения длины первичных геномов) и защитой важного генетического материала. В связи с этим происходил отбор механизмов защиты от инсерций с формированием систем процессинга РНК, таких как RISC, CRISPR, сплайсинг и система трансляции. Основой для их возникновения должны были быть модификации рибозимов RT, обладающие различными рибонуклеазными активностями. Аминокислоты и пептиды могли использоваться для оптимизации защитных механизмов в противоборствующих эгоистических молекулах полинуклеотидов. В дальнейшем усовершенствование системы трансляции стало основой для тесной взаимосвязи мира ДНК и РНК с полипептидами с преемственностью функций рибозимов белковыми ферментами. Данный принцип сохраняется у эукариот, так как рибозимы длинных нкРНК способны транслироваться с образованием функциональных молекул. В результате отбора гены длинных нкРНК могут преобразовываться в белок-кодирующие, при этом их процессированные транскрипты сохраняют свою ферментативную активность. В результате у эукариот РЭ стали источниками ТФ и сайтов связывания с ними, центромер и центромер-связывающегося белка, теломер и

теломеразы, белок-кодирующих генов и взаимодействующих с ними нкРНК, интронов и компонентов сплайсосомы. Это отражает универсальный принцип саморегуляции РЭ и их производных, необходимый для возникновения жизни и ее эволюции.

Конфликт интересов. Авторы заявляют, что у них нет конфликтов интересов.

Соблюдение этических норм. Настоящая статья не содержит каких-либо исследований с участием людей или использованием животных в качестве объектов исследований.

СПИСОК ЛИТЕРАТУРЫ

- Lincoln, T.A., and Joyce, G.F. (2009) Self-sustained replication of an RNA enzyme, *Science*, **323**, 1229–1232, doi: 10.1126/science.1167856.
- Wöchner, A., Attwater, J., Coulson, A., and Holliger, P. (2011) Ribozyme-catalyzed transcription of an active ribozyme, *Science*, **332**, 209–212, doi: 10.1126/science.1200752.
- Horning, D.P., and Joyce, G.F. (2016) Amplification of RNA by an RNA polymerase ribozyme, *Proc. Natl. Acad. Sci. USA*, **113**, 9786–9791, doi: 10.1073/pnas.1610103113.
- Kreysing, M., Keil, L., Lanzmich, S., and Braun, D. (2015) Heat flux across an open pore enables the continuous replication and selection of oligonucleotides towards increasing length, *Nat. Chem.*, **7**, 203–208, doi: 10.1038/nchem.2155.
- Betts, H.C., Puttick, M.N., Clark, J.W., Williams T.A., Donoghue P.C.J., and Pisani, D. (2018) Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin, *Nat. Ecol. Evol.*, **2**, 1556–1562, doi: 10.1038/s41559-018-0644-x.
- Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., and Cech, T.R. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*, *Cell*, **31**, 147–157.
- Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme, *Cell*, **35**, 849–857.
- Prody, G.A., Bakos, J.T., Buzayan, J.M., Schneider, I.R., and Bruening, G. (1986) Autolytic processing of dimeric plant virus satellite RNA, *Science*, **231**, 1577–1580.
- de la Pena, M., Garcia-Robles, I., and Cervera, A. (2017) The Hammerhead Ribozyme: a long history for a short RNA, *Molecules*, **22**, pii: E78, doi: 10.3390/molecules22010078.
- Levine, M.T., Jones, C.D., Kern, A.D., Lindfors, H.A., and Begun, D.J. (2006) Novel genes derived from noncoding DNA in *Drosophila melanogaster* are frequently X-linked and exhibit testis-biased expression, *Proc. Natl. Acad. Sci. USA*, **103**, 9935–9939.
- Cai, J., Zhao, R., Jiang, H. and Wang, W. (2008) De novo origination of a new protein-coding gene in *Saccharomyces cerevisiae*, *Genetics*, **179**, 487–496, doi: 10.1534/genetics.107.084491.
- Xie, C., Zhang, Y.E., Chen, J.Y., Liu, C.J., Zhou, W.Z., Li, Y., Zhang, M., Zhang, R., Wei, L., and Li, C.Y. (2012) Hominoid-specific de novo protein-coding genes originating from long non-coding RNAs, *PLoS Genet.*, **8**, e1002942, doi: 10.1371/journal.pgen.1002942.
- Ruiz-Orera, J., Messegue, X., Subirana, J.A., and Alba, M.M. (2014) Long non-coding RNAs as a source of new peptide, *Elife*, **3**, e03523, doi: 10.7554/eLife.03523.
- Johnson, R., and Guigo, R. (2014) The RIDL hypothesis: transposable elements as functional domains of long non-coding RNAs, *RNA*, **20**, 959–976, doi: 10.1261/rna.044560.114.
- Kapusta, A., and Feschotte, C. (2014) Volatile evolution of long noncoding RNA repertoires: mechanisms and biological implications, *Trends Genet.*, **30**, 439–452, doi: 10.1016/j.tig.2014.08.004.
- Lukash, L.L. (2007) Mutagenesis induced by integration processes and evolution of nuclear genome, *Biopolym. Cell*, **23**, 172–187.
- Samanta, B., and Joyce G.F. (2017) A reverse transcriptase ribozyme, *Elife*, **6**, e31153, doi: 10.7554/eLife.31153.
- Qu, G., Kaushal, P.S., Wang, J., Shigematsu, H., Piazza, C.L., Agrawal, R.K., Belfort, M., and Wang H.W. (2016) Structure of a group II intron in complex with its reverse transcriptase, *Nat. Struct. Mol. Biol.*, **23**, 549–557, doi: 10.1038/nsm.3220.
- Hughes, S.H. (2015) Reverse transcription of retroviruses and LTR retrotransposons, *Microbiol. Spectr.*, **3**, MDNA3-0027-2014, doi: 10.1128/microbiolspec.MDNA3-0027-2014.
- Moelling, K., and Broecker, F. (2015) The reverse transcriptase-RNase H: from viruses to antiviral defense, *Ann. N. Y. Acad. Sci.*, **1341**, 126–135, doi: 10.1111/nyas.12668.
- Gogolevsky, K.P., Vassetzky, N.S., and Kramerov, D.A. (2009) 5S rRNA-derived and tRNA-derived SINEs in fruit bats, *Genomics*, **93**, 494–500, doi: 10.1016/j.ygeno.2009.02.001.
- Rosenbland, M.A., Larsen, N., Samuelsson, T., and Zwieb, C. (2009) Kinship in the SRP RNA family, *RNA Biol.*, **6**, 508–516.
- Li, Z., Ender, C., Meister, G., Moore, P.S., Chang, Y., and John, B. (2012) Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs, *Nucleic Acids Res.*, **40**, 6787–6799, doi: 10.1093/nar/gks307.
- Kojima, K.K. (2015) A new class of SINEs with snRNA gene-derived heads, *Genome Biol. Evol.*, **7**, 1702–1712, doi: 10.1093/gbe/evv100.
- Martinez, G., Choudury, S.G., and Slotkin, R.K. (2017) tRNA-derived small RNAs target transposable element transcripts, *Nucleic Acids Res.*, **45**, 5142–5152, doi: 10.1093/nar/gkx103.
- Startek, M.P., Nogly, J., Gromadka, A., Grzebelus, D., and Gambin, A. (2017) Inferring transposons activity chronology by TRANAcendence-TEs database and de novo mining tool, *BMC Bioinformatics*, **18**, 422, doi: 10.1186/s12859-017-1824-4.
- Alzohairy, A.M., Gyulai, G., Jansen, R.K., and Bahieldin, A. (2013) Transposable elements domesticated and neofunctionalized by eukaryotic genomes, *Plasmid*, **69**, 1–15, doi: 10.1016/j.plasmid.2012.08.001.
- Kramerov, D.A., and Vassetzky, N.S. (2011) Origin and evolution of SINEs in eukaryotic genomes, *Heredity (Edinb.)*, **107**, 487–495, doi: 10.1038/hdy.2011.43.
- Kapitonov, V.V., and Jurka, J. (2003) A novel class of SINE elements derived from 5S rRNA, *Mol. Biol. Evol.*, **20**, 694–702.
- Wang, J., Wang, A., Han, Z., Zhang, Z., Li, F., and Li, X. (2012) Characterization of three novel SINE families with

- unusual features in *Helicoverpa armigera*, *PLoS One*, **7**, e31355, doi: 10.1371/journal.pone.0031355.
31. Longo, M.S., Brown, J.D., Zhang, C., O'Neill, M.J. and O'Neill, R.J. (2015) Identification of a recently active mammalian SINE derived from ribosomal RNA, *Genome Biol. Evol.*, **7**, 775–788, doi: 10.1093/gbe/evv015.
 32. Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P. and Schulman, A.H. (2007) A unified classification system for eukaryotic transposable elements, *Nat. Rev. Genet.*, **8**, 973–982.
 33. Abrusan, G., Zhang, Y. and Szilagy, A. (2013) Structure prediction and analysis of DNA transposon and LINE retrotransposons proteins, *J. Biol. Chem.*, **288**, 16127–16138, doi: 10.1074/jbc.M113.451500.
 34. Rice, P.A., and Baker, T.A. (2001) Comparative architecture of transposase and integrase complexes, *Nat. Struct. Biol.*, **8**, 302–307.
 35. Wolkowicz, U.M., Morris, E.R., Robson, M., Trubitsyna, M., and Richardson, J.M. (2014) Structural basis of Mos1 transposase inhibition by the anti-retroviral drug Raltegravir, *ACS Chem. Biol.*, **9**, 743–751, doi: 10.1021/cb400791u.
 36. Nowotny, M., Gaidamakov, S.A., Crouch, R.J., and Yang, W. (2005) Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis, *Cell*, **121**, 1005–1016.
 37. De Koning, A.P., Gu, W., Castoe, T.A., Batzer, M.A., and Pollock, D.D. (2011) Repetitive elements may comprise over two-thirds of the human genome, *PLoS Genet.*, **7**, e1002384, doi: 10.1371/journal.pgen.1002384.
 38. Goerner-Potvin, P., and Bourque, G. (2018) Computational tools to unmask transposable elements, *Nat. Rev. Genet.*, **19**, 688–704, doi: 10.1038/s41576-018-0050-x.
 39. Ellefson, J.W., Gollihar, J., Shoroff, R., Shivram, H., Lyer, V.R., and Ellington, A.D. (2016) Synthetic evolutionary origin of a proofreading reverse transcriptase, *Science*, **352**, 1590–1593, doi: 10.1126/science.aaf5409.
 40. Freeland S.J., Knight R.D., Landweber L.F. (1999) Do proteins predate DNA, *Science*, **286**, 690–692.
 41. Baltimore, D. (1970) RNA-dependent DNA polymerase in virions of RNA tumour, *Nature*, **226**, 1209–1211.
 42. Temin, H.M., and Mizutani, S. (1970) RNA-dependent DNA polymerase in virions of Rous sarcoma, *Nature*, **226**, 1211–1213.
 43. Lampson, B.C., Inouye, M. and Inouye, S. (1989) Reverse transcriptase with concomitant ribonuclease H activity in the cell-free synthesis of branched RNA-linked msDNA of *Mycobacterium xanthus*, *Cell*, **56**, 701–707.
 44. Lim, D., and Maas, W.K. (1989) Reverse transcriptase-dependent synthesis of a covalently linked, branched DNA-RNA compound in *E. coli* B, *Cell*, **56**, 891–904.
 45. Toro, N., Martinez-Abarca, F., and Gonzalez-Delgado, A. (2017) The reverse transcriptases associated with CRISPR-Cas systems, *Sci. Rep.*, **7**, 7089, doi: 10.1038/s41598-017-07828-y.
 46. Eklund, E.H., and Bartel, D.P. (1996) RNA-catalysed RNA polymerization using nucleoside triphosphates, *Nature*, **382**, 373–376.
 47. Been, M.D., and Cech, T.R. (1988) RNA as an RNA polymerase: net elongation of an RNA primer catalyzed by the *Tetrahymena* ribozyme, *Science*, **239**, 1412–1416.
 48. Adamala, K., Engelhart, A.E. and Szostak, J.W. (2015) Generation of functional RNAs from inactive oligonucleotide complexes by non-enzymatic primer extension, *J. Am. Chem. Soc.*, **137**, 483–489.
 49. Braun, V., Mehlig, M., Moos, M., Rupnik, M., Kalt, B., Mahony, D.E., and von Eichel-Streiber, C. (2000) A chimeric ribozyme in *clostridium difficile* combines features of group I introns and insertion elements, *Mol. Microbiol.*, **36**, 1447–1459.
 50. Gao, X., and Voytas, D.F. (2005) A eukaryotic gene family related to retroelements integrases, *Trends Genet.*, **21**, 133–137.
 51. Skala, A.M. (2014) Retroviral DNA transposition: themes and variations, *Microbiol. Spectr.*, **2**, doi: 10.1128/microbiolspec.MDNA3-0005-2014.
 52. Zeng, L., Pederson, S.M., Cao, D., Qu, Z., Hu, Z., Adelson, D.L., and Wei, C. (2018) Genome-wide analysis of the association of transposable elements with gene regulation suggests that alu elements have the largest overall regulatory impact, *J. Comput. Biol.*, **25**, 551–562, doi: 10.1089/cmb.2017.0228.
 53. Ren, Y.F., Li, G., Wu, J., Xue, Y.F., Song, Y.J., Lv, L., Zhang, X.J., and Tang, K.F. (2012) Dicer-dependent biogenesis of small RNAs derived from 7SL RNA, *PLoS One*, **7**, e40705, doi: 10.1371/journal.pone.0040705.
 54. Jacob, M.D., Audas, T.E., Mullineux, S.T., and Lee, S. (2012) Where no RNA polymerase has gone before: novel functional transcripts derived from the ribosomal Intergenic spacer, *Nucleus*, **3**, 315–319.
 55. Kumar, P., Anaya, J., Mudunuri, S.B., and Dutta, A. (2014) Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets, *BMC Biol.*, **12**, 78, doi: 10.1186/s12915-014-0078-0.
 56. Ender, C., Krek, A., Friedlander, M.R., Beitzinger, M., Weinmann, L., Chen, W., Pfeffer, S., Rajewsky, N., and Meister, G. (2008) A human snoRNA with microRNA-like functions, *Mol. Cell.*, **32**, 519–528, doi: 10.1016/j.molcel.2008.10.017.
 57. Taft, R.J., Glazov, E.A., Lassmann, T., Hayashizaki, Y., Carninci, P., and Mattick, J.S. (2009) Small RNAs derived from snoRNAs, *RNA*, **15**, 1233–1240, doi: 10.1261/rna.1528909.
 58. Venkatesh, T., Suresh, P.S., and Tsutsumi, R. (2016) tRFs: miRNAs in disguise, *Gene*, **579**, 133–138, doi: 10.1016/j.gene.2015.12.058.
 59. Zhou, J., Ding, D., Wang, M., and Cong, Y.S. (2014) Telomerase reverse transcriptase in the regulation of gene expression, *BMB Rep.*, **47**, 8–14.
 60. Elliott, T.A., Stage, D.E., Crease, T.J., and Eickbush, T.H. (2013) In and out of the rRNA genes: characterization of Pokey elements in the sequenced *Daphnia* genome, *Mob. DNA*, **4**, 20, doi: 10.1186/1759-9753-4-20.
 61. Jamburuthugoda, V.K., and Eickbush, T.H. (2014) Identification of RNA binding motifs in the R2 retrotransposon-encoded reverse transcriptase, *Nucleic Acids Res.*, **42**, 8405–8415, doi: 10.1093/nar/gku514.
 62. Novikova, O., and Belfort, M. (2017) Mobile group II introns as ancestral eukaryotic elements, *Trends Genet.*, **33**, 773–783, doi: 10.1016/j.tig.2017.07.009.
 63. Wang, D., Su, Y., Wang, X., Lei, H., and Yu, J. (2012) Transposon-derived and satellite-derived repetitive sequences play distinct functional roles in mammalian intron size expansion, *Evol. Bioinform. Online*, **8**, 301–319, doi: 10.4137/EBO.S9758.
 64. Yenerall, P., and Zhou, L. (2012) Identifying the mechanisms of intron gain: progress and trends, *Biol. Direct.*, **7**, 29, doi: 10.1186/1745-6150-7-29.
 65. Feschotte, C. (2008) The contribution of transposable elements to the evolution of regulatory networks, *Nat. Rev. Genet.*, **9**, 397–405.
 66. Tajnik, M., Vigilante, A., Braun, S., Hanel, H., Luscombe, N.M., Ule, J., Zarnack, K., and Koning, J. (2015) Ingenic Alu exonisation facilitates the evolution of tissue-specific transcript ends, *Nucleic Acids Res.*, **43**, 10492–10505, doi: 10.1093/nar/gkv956.

67. Lei, H., and Vorechovsky, I. (2005) Identification of splicing silencers and enhancers in sense Alus: a role for pseudoacceptors in splice site repression, *Mol. Cell. Biol.*, **25**, 6912–6920.
68. Pastor, T., Talotti, G., Lewandowska, M.A., and Pagani, F. (2009) An Alu-derived intronic splicing enhancer facilitates intronic processing and modulates aberrant splicing in ATM, *Nucleic Acids Res.*, **37**, 7258–67, doi: 10.1093/nar/gkp778.
69. Galej, W.P., Oubridge, C., Newman, A.J., and Nagai, K. (2013) Crystal structure of Prp8 reveals active site cavity of the spliceosome, *Nature*, **493**, 638–643, doi: 10.1038/nature11843.
70. Rearick, D., Prakash, A., McSweeney, A., Shepard S.S., Fedorova, L., and Fedorov, A. (2011) Critical association of ncRNA with introns, *Nucleic Acids Res.*, **39**, 2357–2366, doi: 10.1093/nar/gkq1080.
71. Nissen, P., Hansen, J., Ban, N., Moore P.B., and Steitz, T.A. (2000) The structural basis of ribosome activity in peptide bond synthesis, *Science*, **289**, 920–930.
72. Zimmerly, S., and Wu, L. (2015) An unexplored diversity of reverse transcriptases in bacteria, *Microbiol. Spectr.*, **3**, MDNA3-0058-2014, doi: 10.1128/microbiolspec.MDNA3-0058-2014.
73. Равин Н.В., Шестаков С.В. (2013) Геном прокариот, *Вавиловский журнал генетики и селекции*, **17**, 972–984.
74. Liu, M., Deora, R., Doulatov, S.R., Gingery, M., Eiserling, F.A., Preston, A., Maskell, D.J., Simons, R.W., Cotter, P.A., Parkhill, J., and Miller, J.F. (2002) Reverse transcriptase-mediated tropism switching in Bordetella bacteriophage, *Science*, **295**, 2091–2094.
75. Toro, N., Martinez-Abarca, F., Gonzalez-Delgado, A., and Mestre, M.R. (2018) On the origin and evolutionary relationships of the reverse transcriptases associated with type III CRISPR–Cas Systems, *Front. Microbiol.*, **9**, 1317, doi: 10.3389/fmicb.2018.01317.
76. Lambowitz, A.M., and Zimmerly, S. (2011) Group II introns: mobile ribozymes that invade DNA, *Cold Spring Harb. Perspect. Biol.*, **3**, a003616, doi: 10.1101/cshperspect.a003616.
77. Peebles, C.L., Perlman, P.C., Mecklenburg, K.L., Petrillo, M.L., Tabor, J.H., Jarrell, K.A., and Cheng, H.L. (1986) A self-splicing RNA excises an intron lariat, *Cell*, **44**, 213–223.
78. Silas, S., Mohr, G., Sidote, D.J., Markham, L.M., Sanchez-Amat, A., Bhaya, D., Lambowitz, A.M., and Fire, A.Z. (2016) Direct CRISPR spacer acquisition from RNA by natural reverse transcriptase-Cas1 fusion protein, *Science*, **351**, aad4234, doi: 10.1126/science.aad4234.
79. Elliott, T.A., and Gregory, T.R. (2015) Do larger genomes contain more diverse transposable elements, *BMC Evol. Biol.*, **15**, 69–81, doi: 10.1186/s12862-015-0339-8.
80. Kubiak, M.R., and Makalowska, I. (2017) Protein-coding genes' retrocopies and their functions, *Viruses*, **9**, pii: E80, doi: 10.3390/v9040080.
81. Zdobnov, E.M., Campillos, M., Harrington, E.D., Torrents, D., and Bork, P. (2005) Protein coding potential of retroviruses and other transposable elements in vertebrate genomes, *Nucleic Acids Res.*, **33**, 946–954.
82. Campillos, M., Doerks, T., Shah, P.K., and Bork, P. (2006) Computational characterization of multiple Gag-like human protein, *Trends Genet.*, **22**, 585–589.
83. Sela, N., Kim, E., and Ast, G. (2010) The role of transposable elements in the evolution of non-mammalian vertebrates and invertebrates, *Genome Biol.*, **11**, R59, doi: 10.1186/gb-2010-11-6-r59.
84. Schmitz, J., and Brosius, J. (2011) Exonization of transposed elements: a challenge and opportunity for evolution, *Biochimie*, **93**, 1928–1934, doi: 10.1016/j.biochi.2011.07.014.
85. Cheng, Z.J., and Murata, M. (2003) A centromeric tandem repeat family originating from a part of Ty3/gypsy-retroelement in wheat and its relatives, *Genetics*, **164**, 665–672.
86. Kipling, D., and Warburton, P.E. (1997) Centromeres, CENP-B and Tigger too, *Trends Genet.*, **13**, 141–145.
87. Mestrovic, N., Mravinac, B., Pavlek, M., Vojvoda-Zeljko, T., Satovic, E., and Plohl, M. (2015) Structural and functional liaisons between transposable elements and satellite DNAs, *Chromosome Res.*, **23**, 583–596, doi: 10.1007/s10577-015-9483-7.
88. Arkhipova, I.R. (2018) Neutral theory, transposable elements, and eukaryotic genome evolution, *Mol. Biol. Evol.*, **35**, 1332–1337, doi: 10.1093/molbev/msy083.
89. Garavis, M., Gonzalez, C., and Villasante, A. (2013) On the origin of the eukaryotic chromosome: the role of non-canonical DNA structures in telomere evolution, *Genome Biol. Evol.*, **5**, 1142–1150, doi: 10.1093/gbe/evt079.
90. De Souza, F.S., Franchini, L.F., and Rubinstein, M. (2013) Exaptation of transposable elements into novel cis-regulatory elements: is the evidence always strong, *Mol. Biol. Evol.*, **30**, 1239–1251, doi: 10.1093/molbev/mst045.
91. Gim, J., Ha, H., Ahn, K., Kim, D.S., and Kim, H.S. (2014) Genome-Wide Identification and Classification of microRNAs derived from repetitive elements, *Genomic Inform.*, **12**, 261–267, doi: 10.5808/GI.2014.12.4.261.
92. Long, Y., Wang, X., Youmans, D.T., and Cech, T.R. (2017) How do lncRNAs regulate transcription, *Sci. Adv.*, **3**, eaao2110, doi: 10.1126/sciadv.aao2110.
93. Lu, X., Sachs, F., Ramsay, L., Jacques, P.E., Goke, J., Bourque, G., and Ng, H.H. (2014) The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity, *Nat. Struct. Mol. Biol.*, **21**, 423–425, doi: 10.1038/nsmb.2799.
94. Honson, D.D., and Macfarlan, T.S. (2018) A lncRNA-like role for LINE1s in development, *Dev. Cell*, **46**, 132–134, doi: 10.1016/j.devcel.2018.06.022.
95. Anderson, D.M., Anderson, K.M., Cang, C.L., Makarewich, C.A., Nelson, B.R., McAnally, J.R., Kasaragod, P., Shelton, J.M., Liou, J., Bassel-Duby, R., and Olson, E.N. (2015) A micropeptide encoded by a putative long noncoding RNA regulates muscle performance, *Cell*, **160**, 595–606, doi: 10.1016/j.cell.2015.01.009.
96. Chuong, E.B., Elde, N.C., and Feschotte, C. (2017) Regulatory activities of transposable elements: from conflicts to benefits, *Nat. Rev. Genet.*, **18**, 71–86, doi: 10.1038/nrg.2016.139.
97. Fontdevila, A. (2005) Hybrid genome evolution by transposition, *Cytogenet. Genome Res.*, **110**, 49–55.
98. Soemedi, R., Cygan, K.J., Rhine, C.L., Glidden, D.T., Taggart, A.J., Lin, C.L., Fredericks, A.M., and Fairbrother, W.G. (2017) The effects of structure on pre-mRNA processing and stability, *Methods*, **125**, 36–44, doi: 10.1016/j.ymeth.2017.06.001.
99. Kralovicova, J., Patel, A., Searle, M. and Vorechovsky, I. (2015) The role of short RNA loops in recognition of a single-hairpin exon derived from a mammalian-wide interspersed repeat, *RNA Biol.*, **12**, 54–69, doi: 10.1080/15476286.2015.1017207.
100. Хавинсон В.Х., Соловьев А.Ю., Шатаева Л.К. (2006) Молекулярный механизм взаимодействия олигопептидов и двойной спирали ДНК, *Бюл. экстр. биол.*, **141**, 443–447.
101. Gladyshev, E.A., and Arkhipova, I.R. (2011) A widespread class of reverse transcriptase-related cellular genes, *Proc. Natl. Acad. Sci. USA*, **108**, 20311–20316, doi: 10.1073/pnas.1100266108.
102. Kim, S., and Choi, D. (2018) New role of LTR-retrotransposons for emergence and expansion of disease-resistance

- genes and high-copy gene families in plants, *BMB Rep.*, **51**, 55–56.
103. Zhang, H., Tao, Z., Hong, H., Chen, Z., Wu, C., Li, X., Xiao, J., and Wang, S. (2016) Transposon-derived small RNA is responsible for modified function of WRKY45 locus, *Nat. Plants*, **2**, 16016–16023, doi: 10.1038/nplants.2016.16.
 104. Trizzino, M., Kapusta, A., and Brown, C.D. (2018) Transposable elements generate regulatory novelty in a tissue-specific fashion, *BMC Genomics*, **19**, 468, doi: 10.1186/s12864-018-4850-3.
 105. Виноградов А.Е. (2011) *Функциональное значение базовых свойств структуры генома эукариот*. Дис. докт. биол. наук, Санкт-Петербург.
 106. Joly-Lopez, Z., and Bureau, T.E. (2018) Exaptation of transposable element coding sequences, *Curr. Opin. Genet. Dev.*, **49**, 34–42, doi: 10.1016/j.gde.2018.02.011.
 107. Duan, C.G., Wang, X., Pan, L., Miki, D., Tang, K., Hsu, C.C., Lei, M., Zhong, Y., Hou, Y.J., Wang, Z., Zhang, Z., Mangrauthia, S.K., Xu, H., Zhang, H., Dilkes, B., Tao, W.A., and Zhu, J.K. (2017) A pair of transposon-derived proteins function in a histone acetyltransferase complex for active DNA demethylation, *Cell Res.*, **27**, 226–240, doi: 10.1038/cr.2016.147.
 108. Sinzelle, L., Izsvak, Z., and Ivics, Z. (2009) Molecular domestication of transposable elements: From detrimental parasites to useful host genes, *Cell. Mol. Life Sci.*, **66**, 1073–1093, doi: 10.1007/s00018-009-8376-3.
 109. Wang, J., Vicente-Garcia, C., Seruggia, D., Molto, E., Fernandez-Minan, A., Neto, A., Lee, E., Gomez-Skarmeta, J.L., Montoliu, L., Lunyak, V.V., and Jordan, I.K. (2015) MIR retrotransposons sequences provide insulators to the human genome, *Proc. Natl. Acad. Sci. USA*, **112**, 4428–4437, doi: 10.1073/pnas.1507253112.

THE ROLE OF REVERSE TRANSCRIPTASES IN THE ORIGIN OF LIFE

R. N. Mustafin^{1*} and E. K. Khusnutdinova²

¹ Bashkir State Medical University, 450008 Ufa, Russia; E-mail: ruji79@mail.ru

² Institute of Biochemistry and Genetics, Ufa Federal Research Centre of the Russian Academy of Sciences, 450054 Ufa, Russia

Received March 8, 2019

Revised May 14, 2019

Accepted May 14, 2019

We proposed that the ribozyme RNA polymerase, possessing reverse transcriptase and integration activity, played a key role in the patterns of the origin of life on Earth. We presented the hypothesis that the universal ancestral units of all living organisms were retroelements with reverse transcriptase and integrase activity. Their insertion capacity was the basis for the formation of complex DNA structures, the primary genomes that gave the rise to all archaea, eukaryotes, bacteria and viruses. Conservative properties of retroelements are preserved throughout the evolution. Modifications of their use have led to the emergence of new ways of interaction of proteins and nucleic acids. Life evolved through insertional mutagenesis and competition of autonomously replicating polynucleotides thus preserving the structures with adaptive properties. We assume that the natural selection of protection mechanisms against insertions based on the ribonuclease ability of the reverse transcriptase ribozyme has led to the emergence of all universal enzymatic systems for processing RNA molecules. These systems were and remain key sources of structural and functional transformations of genomes in evolution. The data are presented enabling us to speculate that the translation system, which united the world of RNA and DNA with proteins, arose as a modification of the protection mechanisms against insertions. Polypeptides formed using this translation system began to potentiate the work of ribozymes in the RNPs and even functionally replace them due to more efficient catalysis of biological reactions. We analyzed the mechanisms of the use of retroelements in structural and regulatory transformations of eukaryotic genomes, which may reflect the adaptive principles formed upon the origin of life. Simultaneously with the evolution of existing proteins, retroelements give rise to ribozymes, such as long non-coding RNA. They can function in combination with proteins in the RNPs and are also capable of independent catalytic activity and translation. Their genes have the potential to transform into protein-coding genes. That is, the conservative principles of the interactions of RNA, DNA, and proteins, formed upon the origin of life on Earth, are used throughout the evolution.

Keywords: reverse transcriptase, polymerase, processing, ribozymes, retroelements, transposable elements, evolution