

МОБИЛЬНЫЕ И НЕМОБИЛЬНЫЕ: МНОГООБРАЗИЕ ОБРАТНЫХ ТРАНСКРИПТАЗ И ИХ РЕКРУТИРОВАНИЕ ГЕНОМ О ХОЗЯИНА

Обзор

© 2023 И.Р. Архипова*, И.А. Юшенова

*Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory,
Woods Hole, MA 02543 USA; e-mail: iarkhipova@mbl.edu, iyushenova@mbl.edu*

Поступила в редакцию 13.09.2023

После доработки 18.09.2023

Принята к публикации 20.09.2023

Обратные транскриптазы (reverse transcriptase, RT), или РНК-зависимые ДНК-полимеразы — это необычные ферменты, которые впервые дали возможность пересмотреть общепринятое представление об одностороннем потоке генетической информации в клетке от ДНК к РНК и белку. RT были впервые обнаружены в ретровирусах позвоночных, а впоследствии — повторно обнаружены у большинства эукариот, бактерий и архей (что, по сути, охватывает все надцарства живых организмов). В ретровирусах RT обеспечивают возможность копировать РНК-геном в ДНК для последующего включения в геном хозяина, что важно для репликации и выживания. В клеточных организмах большинство последовательностей RT происходит от ретротранспозонов — типа самореплицирующихся генетических элементов, которые полагаются на обратную транскрипцию для копирования и вставки своих последовательностей в новые места генома. Однако некоторые ретроэлементы могут быть «одомашнены» и в конечном итоге стать ценным дополнением к общему репертуару клеточных ферментов. Они могут быть полезными и при этом либо вспомогательными — например, как элементы, генерирующие разнообразие (diversity-generating elements) — либо даже незаменимыми, как теломеразные RT. В настоящее время обнаруживают всё большее количество «одомашненных» генетических элементов, несущих гены RT. Можно утверждать, что «одомашненные» RT и обратная транскрипция в целом более широко распространены в клеточных организмах, чем считалось ранее, и что многие важные клеточные функции, такие как поддержание стабильности концов хромосом, могли возникнуть из изначально «эгоистичного» процесса преобразования РНК в ДНК.

КЛЮЧЕВЫЕ СЛОВА: обратная транскрипция, РНК-зависимая ДНК-полимераза, теломеразная обратная транскриптаза.

DOI: 10.31857/S0320972523110088, **EDN:** MLIDVY

ВВЕДЕНИЕ

На заре молекулярной биологии, когда ещё мало что было известно о молекулярной природе биологических явлений, авторы многочисленных теоретических работ пытались предвидеть будущие открытия и обоснованно предсказать молекулярные механизмы, объясняющие фундаментальные генетические принципы. Примечательно, что лишь относительно малая часть таких работ выдержала испытание временем и экспериментальную проверку, последовавшую в предстоящие годы. Среди

таких дальновидных работ заслуженное место занимает теоретическое предсказание Алексея Оловникова о недорепликации концов ДНК в линейных хромосомах и существовании специализированного фермента, способного решить эту проблему [1, 2]. Хотя одновременно открытие проблемы репликации концов ДНК было также сделано в статье Джеймса Уотсона [3], внимание в ней было уделено в основном фаговой ДНК, не указывалось на необходимость специализированной полимеразы, а вместо этого акцент был смещён на нуклеазы, процессирующие концы ДНК.

Принятые сокращения: RT — reverse transcriptase, обратная транскриптаза.

* Адресат для корреспонденции.

За открытие теломеразы — специализированной полимеразы, которая может добавлять простые повторяющиеся последовательности к концам линейных хромосом, чтобы компенсировать потерю концевой ДНК после каждого цикла репликации — была присуждена Нобелевская премия, но путь к нему был долгим и непростым. В первоначальном сообщении Грейдер и Блэкберн найденный в *Tetrahymena* фермент обозначался как концевая трансфераза [4], поскольку его обнаруженная активность заключалась в добавлении tandemных повторов к теломерным праймерам без очевидной матрицы. Однако впоследствии соответствующая матричная РНК была выявлена как неотъемлемый компонент рибонуклеопротеинового холофермента, что послужило экспериментальным доказательством РНК-зависимого синтеза ДНК [5], хотя по-прежнему считалось, что классифицировать теломеразный фермент как настоящую обратную транскриптазу преждевременно.

Процесс синтеза ДНК с использованием РНК в качестве матрицы в целом обозначается термином «обратная транскрипция», а соответствующий фермент, способный осуществлять эту реакцию, носит название «обратная транскриптаза» (reverse transcriptase, RT), также известная как «ревертаза» в русскоязычной литературе. Его экспериментальному открытию Теминым и Балтимором более 50 лет назад [6, 7] (которое также было отмечено Нобелевской премией) аналогичным образом предшествовала концепция синтеза ДНК на матрице вирусной ДНК Говарда Темина, известная как «гипотеза провируса» [8]. Мало кто догадывался, что, помимо открытия обратного потока генетической информации от вирусной РНК к ДНК, этой гипотезой также были созданы предпосылки для открытия самореплицирующихся подвижных генетических элементов и для последующего осознания того, что некоторые дополнительные или даже незаменимые функции клетки-хозяина

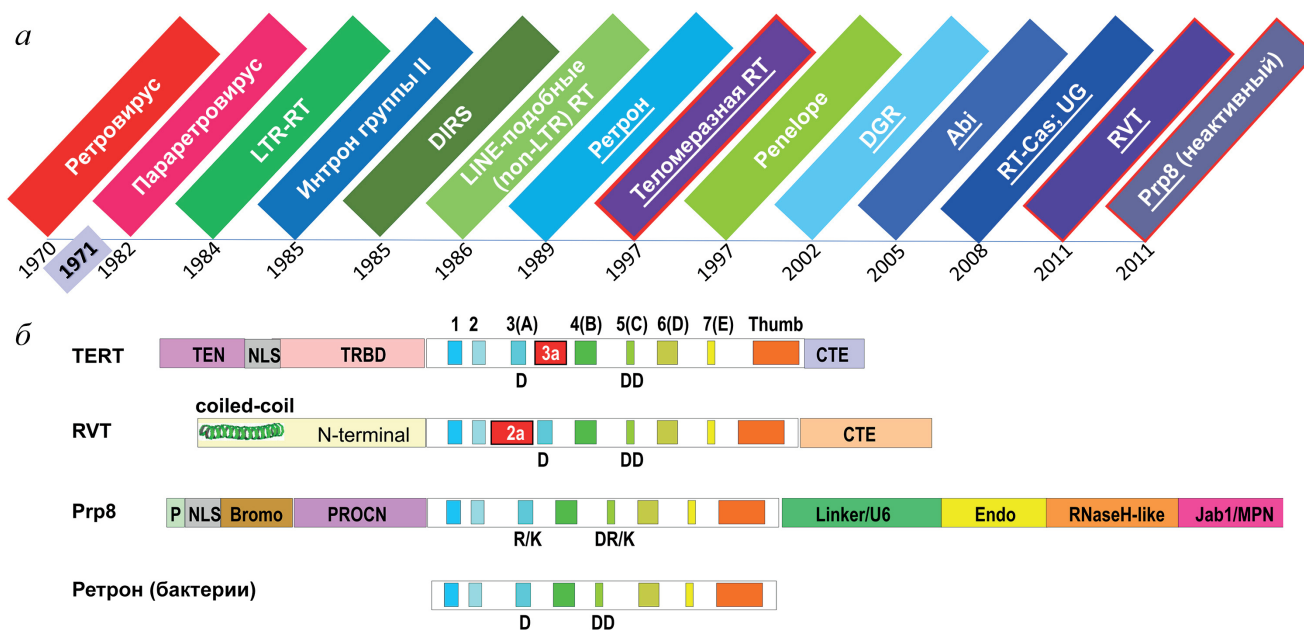


Рис. 1. Основные типы обратных транскриптаз (RT) из трёх надцарств живых организмов. *a* — Хронология открытия RT. Основные типы RT, описанные в тексте, показаны следующими цветами: вирусные RT — оттенки красного; RT эукариотических мобильных элементов — оттенки зелёного; прокариотические RT — оттенки синего; «одомашненные» эукариотические RT — оттенки фиолетового. «Одомашненные» RT подчеркнуты. Годы соответствуют первым сообщениям о выявлении гомологии с каталитическим ядром RT. В 1971 г. впервые была обозначена проблема неполной репликации концов хромосом [1]. *б* — Примеры структурной организации «одомашненных» эукариотических RT. Бактериальные ретроны включены для сравнения. Расположенное в центре каталитическое ядро RT представлено семью консервативными мотивами, разделёнными спейсерами переменной длины с характерными длинными петлями 2a и 3a (также называемая IFD), которые отмечены красным. Указаны аминокислотные остатки каталитической триады D..DD и их некаталитические замены. Дополнительные домены по обе стороны от ядра RT и «большого пальца»: TEN — незаменимый N-концевой домен теломеразы (telomerase essential N-terminal domain); TRBD — теломеразный РНК-связывающий домен (telomerase RNA binding domain); CTE — C-концевое удлинение (C-terminal extension); P — полипролиновый участок (polyproline stretch); NLS — сигнал ядерной локализации (nuclear localization signal); Bromo — бромодомен; PROCN — центральный домен PRO8 (PRO8 central domain); Endo — эндонуклеазоподобный домен (endonuclease-like); Jab1/MPN — предполагаемый деубиквитиназоподобный домен. Масштаб приблизительный. Доменная структура представлена по различным литературным источникам [55, 57, 59].

могут взять на себя потомки таких мобильных элементов. Примечательно, что RT были обнаружены примерно в то же время, когда была впервые выявлена проблема неполной репликации концов хромосом (рис. 1).

РАЗВИТИЕ ПОДХОДОВ К ВЫЯВЛЕНИЮ РЕТРОЭЛЕМЕНТОВ

С момента открытия RT в ретровирусах представление об их разнообразии необычайно расширилось: от представления о них как о чисто вирусных компонентах до открытия их удивительно разнообразных структурных и функциональных ролей у эукариотических и прокариотических хозяев (рис. 1, а). Ранние достижения в области вирусологии привели к дальнейшему открытию обратной транскрипции в репликативных циклах гепаднавирусов и каулимовирусов (совместно названных параретровирусами [9]) — этому способствовала доступность методов выделения вирусов и биохимического анализа RT. Вскоре после этого перспективы новых открытий сместились к обнаружению гомологии последовательностей. Этот процесс ускорило появление технологий секвенирования и знаковое открытие общих мотивов аминокислотных последовательностей в каталитическом ядре ДНК-полимераз вирусов, осуществляющих обратную транскрипцию [10]. Вскоре после этого неотъемлемой частью идентификации новых RT стал поиск остатков аспартата, образующих каталитическую триаду D..DD в активном центре RT. На представленной временной шкале изучения RT (рис. 1, а) основополагающим работам, в которых были впервые выявлены характерные остатки RT, был присвоен приоритет по сравнению с теми, в которых сообщалось о первоначальном биохимическом обнаружении РНК-зависимой полимеризации ДНК. Это связано с тем, что полноценное экспериментальное подтверждение активности RT непременно должно включать в себя сайт-направленный мутагенез остатков активного центра, присутствующих в двух из семи консервативных мотивов, формирующих каталитическое ядро RT (рис. 1, б).

Первая половина временной шкалы, до 1990-х гг., представлена в основном открытиями RT в различных типах вирусов и мобильных генетических элементах. Действительно, мультикопийные мобильные элементы были одними из первых компонентов эукариотических геномов, клонированных на молекулярном уровне [11, 12], наряду с другими ак-

тивно транскрибируемыми мультикопийными генами, такими как повторяющиеся единицы рибосомальной ДНК или кластеры генов гистонов [13, 14]. Общее структурное сходство LTR-ретротранспозонов и ретровирусов сразу же стало очевидным при их клонировании из дрожжи и дрожжей [15]. Однако окончательное доказательство их близкого родства с ретровирусами было получено в результате анализа их полных нуклеотидных последовательностей, выявившего их способность кодировать фермент RT [16, 17]. Более того, характерные участки гомологии с консервативными мотивами RT вскоре были идентифицированы не только в ретровирусоподобных мобильных элементах, но также и в мобильных интронах митохондрий грибов II группы и других типах мультикопийных эукариотических транспозонов, таких как DIRS и LINE-подобные ретро-транспозоны [18–21]. В конце первых двух десятилетий изучения RT появились первые сообщения об их существовании у бактерий в форме ретронов, мультикопийных внехромосомных химерных молекул ДНК–РНК, связанных через точку ветвления 2'-5' [22, 23].

На следующем этапе в истории открытий новых RT исследователи по-прежнему опирались на обнаружение гомологии последовательностей, но преимущественно обнаруживались RT, присутствующие в меньшем количестве копий, и большинство из них не относилось к мобильным элементам, а представляло собой однокопийные гены клетки-хозяина (рис. 1, а, подчеркнуто). Фактически известный на данный момент спектр эукариотических ретро-транспозонов не расширился с момента открытия *Penelope*-подобных ретро-элементов (*Penelope-like elements*, PLE) [24]. Первый и наиболее известный случай «одомашнивания» RT у эукариот был выявлен после доказательства того, что подлинную RT представляет собой теломераза. Проведение связи между активностью RT и соответствующим ферментом потребовало много сил и времени, в течение которого происходили и ошибочные определения [25]. Окончательного успеха в определении каталитической субъединицы теломеразы как RT удалось достичь благодаря выявлению консервативных мотивов в домене «пальцев» и «ладони» RT, что подтвердила потеря активности ферментом при направленном мутагенезе трёх инвариантных каталитических остатков аспартата [26]. Таким образом, было обнаружено, что однокопийный ген RT, присутствующий почти у всех видов эукариот, отвечает за важную функцию клетки-хозяина по удлинению концов линейных хромосом

в результате захвата RT ДНК-вирусом [29]. Ty1/*copi*a-подобные LTR-ретротранспозоны (Pseudoviridae) соответствуют общей структуре LTR, но имеют другой порядок доменов. Все ретровирусоподобные элементы, входящие в таксономический порядок Ortervirales (Retroviridae, Metaviridae, Pseudoviridae и Belraoviridae) [29], мобилизуются с помощью интегразы (integrase, IN), которая отвечает за встраивание копии кДНК в новые участки хромосом. Отдельная группа элементов под названием **DIRS** мобилизуется с помощью тирозинрекомбиназы (YR) вместо IN.

Ретротранспозоны подкласса LINE (также известные как non-LTR) мобилизуются без образования цитоплазматической промежуточной кДНК: их RT использует механизм обратной транскрипции, праймируемой в сайте интеграции (target-primed reverse transcription, TPRT) для синтеза кДНК непосредственно на сайте интеграции в хромосому, который разрезается одним из двух различных типов ассоциированных эндонуклеаз (EN) — AP-подобной или REL-подобной. Наконец, RT **Penelope-подобных элементов (PLE)** используют для мобилизации ещё один тип EN (GIY-YIG), в результате чего количество типов эндонуклеаз, связанных с ретротранспозонами эукариот, достигает пяти. Более подробное свежее описание механизмов ретромобильности можно найти в обзоре Paul et al. [30].

ПРОКАРИОТИЧЕСКИЕ МОБИЛЬНЫЕ ЭЛЕМЕНТЫ: ИНТРОНЫ ГРУППЫ II, РЕТРОПЛАЗМИДЫ

Интроны группы II (G2I) представляют собой самосплайсинговые ретроэлементы, обнаруженные у бактерий, некоторых архей и органелл эукариот [31]. Они были впервые найдены в митохондриях грибов и, как было показано, обладают такой же структурной организацией у бактерий и архей и часто рассматриваются в качестве эволюционных предшественников эукариотических сплайсосомных интронов. Их ретромобильность обеспечивается совместным действием каталитически активной РНК, выполняющей функцию рибозима в реакциях самосплайсинга и обратного сплайсинга, а также кодируемой интроном RT, которая синтезирует кДНК-копию интронной РНК в сайте-мишени с использованием механизма TPRT.

Ретроплазмиды были обнаружены в митохондриях грибов [32] и долгое время служили

модельной системой для изучения нестандартных способов праймирования обратными транскриптазами (белковое праймирование, в процессе которого RT использует в качестве праймера гидроксильную группу остатка тирозина или серина, или инициация *de novo*, при которой вообще не используется праймер). Их распространение всё же весьма ограничено, поскольку они присутствуют лишь в нескольких десятках видов грибов среди сотен секвенированных грибных геномов. Ожидается, что, будучи внехромосомными образованиями, они не подвергаются интеграции, но формально составляют часть мобилома благодаря способности реплицироваться автономно.

НЕМОБИЛЬНЫЕ РЕТРОЭЛЕМЕНТЫ В БАКТЕРИЯХ И АРХЕЯХ: РЕТРОНЫ, DGR, Abi/UG, Cas-АССОЦИИРОВАННЫЕ, G2I-ПОДОБНЫЕ

Ретроны представляют собой своеобразные «одомашненные» бактериальные элементы, состоящие из ковалентно связанной РНК и мультикопийной одноцепочечной ДНК (оцДНК) в одной разветвлённой молекуле, соединённой 2'-5'-фосфодиэфирными связями [22, 23]. Каждый отдельный участок ретрона кодирует последовательность белка RT, некодирующую РНК, которая подвергается обратной транскрипции с помощью RT с образованием химерных одноцепочечных молекул ДНК/РНК, а также эффекторный ген, необходимый для антифаговой активности. Несмотря на то, что ретроны были первыми прокариотическими немобильными ретроэлементами, открытыми более 30 лет назад, их клеточная функция была выяснена только в 2020 г. [33–35]. Ретроны обеспечивают защиту хозяина от широкого спектра фагов посредством abortивной инфекции и последующей гибели клеток. Они широко распространены у бактерий, являясь одним из основных компонентов бактериальной иммунной системы. Однако точные механизмы, посредством которых они дают бактериям устойчивость к фагам за счёт обратной транскрипции, до сих пор неизвестны. Появление RT в трёхчастных модулях вместе с матрицей РНК и множеством предполагаемых эффекторных генов предполагает их прямое взаимодействие в индукции антифагового ответа [36]. Действительно, такое взаимодействие наблюдалось в комплексе RT, родственной ей оцДНК и связанной эффекторной нуклеозиддезоксирибозилтрансферазы [37].

Ретроэлементы, генерирующие разнообразие (diversity-generating retroelements, DGR), представляют собой немобильные RT, которые модифицируют соседние с ними последовательности ДНК у бактерий, архей и вирусов [38, 39]. Несмотря на то что DGR не являются жизненно необходимыми ретроэлементами, они тем не менее полезны для своих хозяев. В наиболее детально описанной модельной системе DGR создают разнообразие в С-концевой вариабельной области генамишени (*mtd*), кодирующего белок бактериофага BPP-1 *Bordetella pertussis*. Возникающая в результате этого гипервариабельность белка хвоста фага — области, которая контактирует с бактериальной клеткой во время инфекции — позволяет фагу инфицировать бактериальные клетки с изменёнными рецепторами на поверхности. Используя подверженную ошибкам обратную транскрипцию, DGR помогают увеличить разнообразие продуктов генов, особенно тех, которые участвуют в связывании лигандов и прикреплении к клетке-хозяину. До сих пор остаётся загадкой то, как достигается адениновая специфичность целенаправленного гипермутагенеза. Более того, проверка соседних генов в модулях DGR позволяет предположить, что гипервариабельность может не ограничиваться переключением тропизма и поверхностным дисплеем [40, 41].

Системы abortивной инфекции (abortive infection systems, Abi), представленные AbiA, AbiK и Abi-P2, представляют собой бактериальные ретроэлементы, служащие для защиты определённых бактерий от фаговых инфекций. Эти гены обнаружены только в геномах некоторых бацилл (в основном у *Lactococcus lactis*), где в основном кодируются плазмидами (AbiA и AbiK), а также в P2-подобных профагах в *Escherichia coli* (Abi-P2). Хотя подробный механизм их действия до сих пор неизвестен, ясно, что белки Abi необходимы для блокирования репликации фагов с последующей запрограммированной гибелью клеток или удалением фага [42, 43]. Интересно, что белок AbiK, как было показано, осуществляет нематричную полимеризацию ДНК *in vitro* и ковалентно присоединяется к ДНК, что указывает на белковое праймирование [44]. Таким образом, Abi представляют собой ещё один (помимо ретронов) тип активных RT, который даёт преимущество субпопуляции бактерий при атаке фагами. Следует отметить, что RT AbiP2 и AbiK отличаются исключительной возможностью образовывать компактные тримеры или гексамеры в растворе, а также отсутствием RT-домена «большого пальца», который

заменён α -спиральным доменом, состоящим из повторов HEAT [45, 46]. Для значительной части так называемых неизвестных групп RT (unknown groups, UG) [47], некоторые из которых были независимо названы DRT (защитные RT) [33], в более ранних исследованиях указывалось, что их невозможно отнести к конкретному типу RT, но позже было обнаружено, что они родственны RT Abi и играют роль в антифаговой защите, находясь в так называемых защитных островках, которые содержат множество других генов, обеспечивающих защиту от вторжения чужеродной ДНК [33, 45].

RT-Cas: домены RT были обнаружены рядом с CRISPR-ассоциированными генами или даже слиты с белками Cas [48–50]. Потенциально эти RT могут обеспечивать бактериальный иммунитет, выполняя синтез кДНК на РНК из бактериофагов, и действительно, было показано, что они опосредуют наследуемое приобретение коротких последовательностей (спейсеров) из чужеродных РНК-элементов [51]. Слияние с белками Cas не является необходимым, хотя и способствует более эффективной кооперации взаимодействующих доменов [52]. Эти RT не являются монофилетическими, поскольку включены в системы CRISPR-Cas из нескольких бактериальных линий RT [50].

Интроноподобные RT группы II — гетерогенная группа немобильных RT, последовательности которых имеют сходство с таковыми G2I, но лишены рибозимного компонента — были впервые описаны в работе Simon et al. [48]. Недавно было обнаружено, что RT G2L из *Pseudomonas aeruginosa* (RT G2L4) участвует в транслезионном синтезе ДНК и репарации двухцепочечных разрывов посредством микрогомологичного соединения концов (microhomology-mediated end-joining, MMEJ) [53]. Интересно, что замена YADD на YIDD в активном центре RT G2L4 ответственна за сдвиг в сторону выполнения MMEJ вместо удлинения праймера, что характерно для каноничных RT G2I с YADD в каталитическом центре. Тем не менее каноничная RT G2I также была способна выполнять репарацию ДНК.

НЕМОБИЛЬНЫЕ ЭУКАРИОТИЧЕСКИЕ RT И ИХ ПРОИЗВОДНЫЕ: ТЕЛОМЕРАЗА, RVT, Prp8

Теломеразная обратная транскриптаза (TERT, telomerase reverse transcriptase), описанная выше, несомненно, является наиболее

известной RT с ключевой клеточной функцией. За счёт своей главной функции поддержания длины линейных хромосом она играет хорошо описанную роль в старении, развитии рака и других заболеваний человека (апластическая анемия, синдром кошачьего крика, врождённый дискератоз и т.д.). Разрабатываются многочисленные методы фармацевтического воздействия на активную теломеразу и связанную с ней матрицу РНК TERT в контексте противораковой терапии и лечения возрастных заболеваний (недавний обзор приведён в работе Fragkiadaki et al. [54]).

Гены, родственные обратной транскриптазе (reverse transcriptase-related genes, *rvt*), представляют собой последний обнаруженный тип «одомашненных» эукариотических RT, широко распространённый у грибов и иногда встречающийся у отдельных растений, простейших и беспозвоночных [55]. Поразительно, что эти гены присутствуют как у прокариот, так и у эукариот, в отличие от всех других типов RT. Примечательно, что RVT изо всех типов бактерий образуют монофилетическую группу, что позволяет предположить, что они не были переданы горизонтально от эукариот, как изначально предполагалось, но, возможно, присутствовали в бактериях до эукариогенеза [56]. Гены *rvt* кодируют активные RT-подобные белки, которые у грибов могут полимеризовать как dNTP, так и NTP. Белки RVT также способны к белковому праймированию. Хотя биологическая функция генов *rvt* ещё не полностью изучена, они явно сохраняются в результате естественного отбора, что указывает на их важность для клеток-хозяев. Эти гены сильно активируются при аминокислотном голодании и применении некоторых антибиотиков у грибов, что позволяет предположить их участие в ответе на такие агенты [55].

Фактор процессинга пре-мРНК 8 (pre-mRNA-processing factor 8, Prp8) представляет собой необычное «одомашненное» производное RT, которое потеряло два из трёх каталитических остатков аспартата, в результате чего утратило способность полимеризовать нуклеотиды [57]. Тем не менее Prp8 является важной частью эукариотической сплайсосомы, регулирующей её сборку и конформацию во время сплайсинга пре-мРНК [58]. Было высказано предположение, что RT-фрагмент Prp8 происходит от мобильных интронов группы II [59]. Это даёт нам ещё один пример того, как в ходе эволюции «эгоистичные» ретротранспозоны могут давать начало важным компонентам эукариотических клеток, в данном случае в качестве структурного элемента,

который включает в себе центральную часть крупного мультидоменного белка, связывающую U5-мРНК (рис. 1, б). Отсутствие каталитических остатков и очень высокая консервативность последовательностей, обусловленная эволюционными ограничениями, налагаемыми функцией сплайсосомы, препятствуют однозначному филогенетическому размещению этого RT-производного домена, но его происхождение, несомненно, восходит к последнему общему предку всех эукариот.

ЗАКЛЮЧЕНИЕ

Из приведённых выше описаний RT легко заключить, что те из них, которые принадлежат к типам, открытым в более ранние годы, как правило, обнаруживались у многочисленных организмов с большим количеством копий RT. Вначале это были вирусы, а затем клеточные многокопийные мобильные генетические элементы: у эукариот — LTR, DIRS и LINE-подобные ретротранспозоны, у прокариот — мобильные интроны группы II и ретроплазмиды, а также ретроны, продуцирующие многочисленные разветвлённые молекулы ДНК–РНК в бактериальных клетках. Ретромобильность обычно обеспечивается определённым типом эндонуклеазы, связанной с каждым мобильным элементом и дающей возможность осуществлять внутривнутрихромосомную вставку копии кДНК. На начальных этапах многие эукариотические мобильные элементы были выявлены по своей способности вызывать инсерционные мутации с видимыми фенотипами в штаммах, где происходит транспозиция мультикопийных элементов [60]. Сейчас очевидно, что RT могут выполнять широкий спектр функций, помимо роли в распространении «эгоистичных» генетических элементов. Мы утверждаем, что многообразие «одомашненных» RT было сильно занижено, а их роль существенно недооценена, и существует множество возможностей рекрутирования RT клетками-хозяевами, несмотря на то, что в целом они не являются незаменимыми и распределены неравномерно. Неудивительно, что иногда от первоначальной идентификации того или иного элемента до правильного определения его функции в клетке-хозяине может пройти много времени, вплоть до десятилетий, если он даёт хозяину селективное преимущество лишь в определённых условиях. Теломеразная RT, кодируемая геном с одной копией, представляет собой существенное исключение, поскольку она практически повсеместно при-

сутствует в эукариотах, и открытие того факта, что она кодирует специализированную RT, т.е. фермент, который ранее считался характерным лишь для вирусов и мобильных элементов, произвело настоящую революцию в данной области исследований [26]. Тем не менее даже критически важная функция поддержания стабильности теломер может поддерживаться за счёт независимых запасных путей [61].

Стоит подчеркнуть, что «приручение» RT у эукариот неразрывно связано с появлением дополнительных функциональных доменов, которые предотвращают произвольный синтез ею кДНК с использованием случайных комбинаций праймер/матрица. В целом, не ожидается, что синтез копий кДНК на случайных матрицах РНК клетки-хозяина принесёт ей пользу, и этот процесс следует предотвращать. Самый простой способ это сделать — устранить каталитическую активность путём замены остатков активного центра, как в случае Prp8. Другой вариант — изменить конфигурацию активного центра за счёт вставки дополнительных структурных петель, как в генах *rvf*. Наконец, TERT достигли строгой субстратной специфичности благодаря высокой степени специализации в отношении несвязанной высокоструктурированной РНК (называемой TER или TR), которая содержит короткий обратный комплемент теломерной повторяющейся единицы, служащей матрицей, и специфически взаимодействует с доменом TRBD, выполняя высокопроцессивный синтез ДНК по механизму обратной транскрипции, праймируемой в сайте интеграции (target-primed

reverse transcription, TPRT) с 3'-концов экспонированных коротких G-богатых tandemных повторов на концах линейных хромосом [62]. Трудно не удивляться тому, что специализированный фермент, для которого Оловниковым была предсказана способность решать проблему утраты концевой ДНК и сохранять целостность хромосом, происходит от мобильных элементов, изначально предназначенных для нарушения стабильности хромосом.

Вклад авторов. И.Р. Архипова, И.А. Юшенкова — концепция, написание и редактирование текста статьи; рисунки были адаптированы И.А. из своей презентации на конференции 2022 г. «Пятьдесят лет обратной транскриптазы» (Cold Spring Harbor, “Fifty Years of Reverse Transcriptase”).

Благодарности. Работа посвящена памяти Алексея Оловникова и заменяет собой запланированную ранее личную беседу, которая должна была произойти в Москве, но так и не состоялась.

Финансирование. Лабораторные исследования поддержаны грантами Национального института здравоохранения США (И.А., R01GM111917) и Национального научного фонда США (И.А. и И.Ю. — MCB-2139001, MCB-2326038).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Соблюдение этических норм. Данная статья не содержит описаний каких-либо исследований с участием людей или использованием животных, выполненных авторами.

СПИСОК ЛИТЕРАТУРЫ

1. Olovnikov, A. M. (1971) Principle of marginotomy in template synthesis of polynucleotides [in Russian], *Dokl. Akad. Nauk SSSR*, **201**, 1496-1499.
2. Olovnikov, A. M. (1973) A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon, *J. Theor. Biol.*, **41**, 181-190, doi: 10.1016/0022-5193(73)90198-7.
3. Watson, J. D. (1972) Origin of concatemeric T7 DNA, *Nat. New Biol.*, **239**, 197-201, doi: 10.1038/newbio239197a0.
4. Greider, C. W., and Blackburn, E. H. (1985) Identification of a specific telomere terminal transferase activity in Tetrahymena extracts, *Cell*, **43**, 405-413, doi: 10.1016/0092-8674(85)90170-9.
5. Greider, C. W., and Blackburn, E. H. (1989) A telomeric sequence in the RNA of *Tetrahymena* telomerase required for telomere repeat synthesis, *Nature*, **337**, 331-337, doi: 10.1038/337331a0.
6. Temin, H. M., and Mizutani, S. (1970) RNA-dependent DNA polymerase in virions of Rous sarcoma virus, *Nature*, **226**, 1211-1213, doi: 10.1038/2261211a0.
7. Baltimore, D. (1970) RNA-dependent DNA polymerase in virions of RNA tumour viruses, *Nature*, **226**, 1209-1211, doi: 10.1038/2261209a0.
8. Temin, H. M. (1964) Nature of the provirus of Rous sarcoma, *Nat. Cancer Inst. Monogr.*, **17**, 557-570.
9. Temin, H. M. (1985) Reverse transcription in the eukaryotic genome: retroviruses, pararetroviruses, retrotransposons, and retrotranscripts, *Mol. Biol. Evol.*, **2**, 455-468, doi: 10.1093/oxfordjournals.molbev.a040365.
10. Toh, H., Hayashida, H., and Miyata, T. (1983) Sequence homology between retroviral reverse tran-

- scriptase and putative polymerases of hepatitis B virus and cauliflower mosaic virus, *Nature*, **305**, 827-829, doi: 10.1038/305827a0.
11. Georgiev, G. P., Ilyin, Y. V., Ryskov, A. P., Tchurikov, N. A., Yenikolopov, G. N., Gvozdev, V. A., and Ananiev, E. V. (1977) Isolation of eukaryotic DNA fragments containing structural genes and the adjacent sequences, *Science*, **195**, 394-397, doi: 10.1126/science.401545.
 12. Finnegan, D. J., Rubin, G. M., Young, M. W., and Hogness, D. S. (1978) Repeated gene families in *Drosophila melanogaster*, *Cold Spring Harb. Symp. Quant. Biol.*, **42**, 1053-1063, doi: 10.1101/sqb.1978.042.01.106.
 13. Glover, D. M., White, R. L., Finnegan, D. J., and Hogness, D. S. (1975) Characterization of six cloned DNAs from *Drosophila melanogaster*, including one that contains the genes for rRNA, *Cell*, **5**, 149-157, doi: 10.1016/0092-8674(75)90023-9.
 14. Schaffner, W., Gross, K., Telford, J., and Birnstiel, M. (1976) Molecular analysis of the histone gene cluster of *Psammecinus miliaris*: II. The arrangement of the five histone-coding and spacer sequences, *Cell*, **8**, 471-478, doi: 10.1016/0092-8674(76)90214-2.
 15. Georgiev, G. P. (1984) Mobile genetic elements in animal cells and their biological significance, *Eur. J. Biochem.*, **145**, 203-220, doi: 10.1111/j.1432-1033.1984.tb08541.x.
 16. Saigo, K., Kugimiya, W., Matsuo, Y., Inouye, S., Yoshioka, K., and Yuki, S. (1984) Identification of the coding sequence for a reverse transcriptase-like enzyme in a transposable genetic element in *Drosophila melanogaster*, *Nature*, **312**, 659-661, doi: 10.1038/312659a0.
 17. Emori, Y., Shiba, T., Kanaya, S., Inouye, S., Yuki, S., and Saigo, K. (1985) The nucleotide sequences of *copia* and *copia*-related RNA in *Drosophila* virus-like particles, *Nature*, **315**, 773-776, doi: 10.1038/315773a0.
 18. Michel, F., and Lang, B. F. (1985) Mitochondrial class II introns encode proteins related to the reverse transcriptases of retroviruses, *Nature*, **316**, 641-643, doi: 10.1038/316641a0.
 19. Cappello, J., Handelsman, K., and Lodish, H. F. (1985) Sequence of Dictyostelium DIRS-1: an apparent retrotransposon with inverted terminal repeats and an internal circle junction sequence, *Cell*, **43**, 105-115, doi: 10.1016/0092-8674(85)90016-9.
 20. Hattori, M., Kuhara, S., Takenaka, O., and Sakaki, Y. (1986) L1 family of repetitive DNA sequences in primates may be derived from a sequence encoding a reverse transcriptase-related protein, *Nature*, **321**, 625-628, doi: 10.1038/321625a0.
 21. Fawcett, D. H., Lister, C. K., Kellett, E., and Finnegan, D. J. (1986) Transposable elements controlling I-R hybrid dysgenesis in *D. melanogaster* are similar to mammalian LINEs, *Cell*, **47**, 1007-1015, doi: 10.1016/0092-8674(86)90815-9.
 22. Lampson, B. C., Sun, J., Hsu, M. Y., Vallejo-Ramirez, J., Inouye, S., and Inouye, M. (1989) Reverse transcriptase in a clinical strain of *Escherichia coli*: production of branched RNA-linked msDNA, *Science*, **243**, 1033-1038, doi: 10.1126/science.2466332.
 23. 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, doi: 10.1016/0092-8674(89)90693-4.
 24. Evgen'ev, M. B., Zelentsova, H., Shostak, N., Kozitsina, M., Barskyi, V., Lankenau, D. H., and Corces, V. G. (1997) *Penelope*, a new family of transposable elements and its possible role in hybrid dysgenesis in *Drosophila virilis*, *Proc. Natl. Acad. Sci. USA*, **94**, 196-201, doi: 10.1073/pnas.94.1.196.
 25. Lundblad, V., and Blackburn, E. H. (1990) RNA-dependent polymerase motifs in ESTI: Tentative identification of a protein component of an essential yeast telomerase, *Cell*, **60**, 529-530, doi: 10.1016/0092-8674(90)90653-v.
 26. Lingner, J., Hughes, T. R., Shevchenko, A., Mann, M., Lundblad, V., and Cech, T. R. (1997) Reverse transcriptase motifs in the catalytic subunit of telomerase, *Science*, **276**, 561-567, doi: 10.1126/science.276.5312.561.
 27. Olovnikov, A. M. (1996) Telomeres, telomerase, and aging: origin of the theory, *Exp. Gerontol.*, **31**, 443-448, doi: 10.1016/0531-5565(96)00005-8.
 28. Arkhipova, I. R., Mazo, A. M., Cherkasova, V. A., Gorelova, T. V., Schuppe, N. G., and Ilyin, Y. V. (1986) The steps of reverse transcription of *Drosophila* mobile genetic elements and U3-R-U5 structure of their LTRs, *Cell*, **44**, 555-563, doi: 10.1016/0092-8674(86)90265-5.
 29. Krupovic, M., Blomberg, J., Coffin, J. M., Dasgupta, I., Fan, H., Geering, A. D., Gifford, R., Harrach, B., Hull, R., Johnson, W., Kreuze, J. F., Lindemann, D., Llorens, C., Lockhart, B., Mayer, J., Muller, E., Olszewski, N., Pappu, H. R., Pooggin, M., Richert-Poggeler, K. R., et al. (2018) Ortervirales: A new viral order unifying five families of reverse-transcribing viruses, *J. Virol.*, **92**, e00515-18, doi: 10.1128/jvi.00515-18.
 30. Paul, B. G., Yushenova, I. A., and Arkhipova, I. R. (2022) *The Diversity of Reverse Transcriptases in Retrotransposons and Human Disease* (Gabriel, A., ed.) World Scientific, Singapore, pp. 1-28, doi: 10.1142/9789811249228_0001.
 31. Lambowitz, A. M., and Belfort, M. (2015) Mobile bacterial group II introns at the crux of eukaryotic evolution, *Microbiol. Spectr.*, **3**, doi: 10.1128/microbiolspec.MDNA3-0050-2014.
 32. Arkhipova, I. R., and Yushenova, I. A. (2019) Giant transposons in eukaryotes: Is bigger better? *Genome Biol. Evol.*, **11**, 906-918, doi: 10.1093/gbe/evz041.
 33. Gao, L., Altae-Tran, H., Böhning, F., Makarova, K. S., Segel, M., Schmid-Burgk, J. L., Koob, J.,

- Wolf, Y. I., Koonin, E. V., and Zhang, F. (2020) Diverse enzymatic activities mediate antiviral immunity in prokaryotes, *Science*, **369**, 1077-1084, doi: 10.1126/science.aba0372.
34. Millman, A., Bernheim, A., Stokar-Avihail, A., Fedorenko, T., Voichek, M., Leavitt, A., Oppenheimer-Shaanan, Y., and Sorek, R. (2020) Bacterial retrons function in anti-phage defense, *Cell*, **183**, 1551-1561, doi: 10.1016/j.cell.2020.09.065.
 35. Bobonis, J., Mitosch, K., Mateus, A., Karcher, N., Kritikos, G., Selkig, J., Zietek, M., Monzon, V., Pfalz, B., Garcia-Santamarina, S., Galardini, M., Sueki, A., Kobayashi, C., Stein, F., Bateman, A., Zeller, G., Savitski, M. M., Elfenbein, J. R., Andrews-Polymenis, H. L., and Typas, A. (2022) Bacterial retrons encode phage-defending tripartite toxin-antitoxin systems, *Nature*, **609**, 144-150, doi: 10.1038/s41586-022-05091-4.
 36. Mestre, M. R., González-Delgado, A., Gutiérrez-Rus, L. I., Martínez-Abarca, F., and Toro, N. (2020) Systematic prediction of genes functionally associated with bacterial retrons and classification of the encoded tripartite systems, *Nucleic Acids Res.*, **48**, 12632-12647, doi: 10.1093/nar/gkaa1149.
 37. Wang, Y., Guan, Z., Wang, C., Nie, Y., Chen, Y., Qian, Z., Cui, Y., Xu, H., Wang, Q., Zhao, F., Zhang, D., Tao, P., Sun, M., Yin, P., Jin, S., Wu, S., and Zou, T. (2022) Cryo-EM structures of *Escherichia coli* Ec86 retron complexes reveal architecture and defence mechanism, *Nat. Microbiol.*, **7**, 1480-1489, doi: 10.1038/s41564-022-01197-7.
 38. Guo, H., Arambula, D., Ghosh, P., and Miller, J. F. (2014) Diversity-generating retroelements in phage and bacterial genomes, *Microbiol. Spectr.*, **2**, MDNA3-0029-2014, doi: 10.1128/microbiolspec.MDNA3-0029-2014.
 39. Paul, B. G., Burstein, D., Castelle, C. J., Handa, S., Arambula, D., Czornyj, E., Thomas, B. C., Ghosh, P., Miller, J. F., Banfield, J. F., and Valentine, D. L. (2017) Retroelement-guided protein diversification abounds in vast lineages of Bacteria and Archaea, *Nat. Microbiol.*, **2**, 17045, doi: 10.1038/nmicrobiol.2017.45.
 40. Roux, S., Paul, B. G., Bagby, S. C., Nayfach, S., Allen, M. A., Attwood, G., Cavicchioli, R., Chistoserdova, L., Gruninger, R. J., Hallam, S. J., Hernandez, M. E., Hess, M., Liu, W. T., McAllister, T. A., O'Malley, M. A., Peng, X., Rich, V. I., Saleska, S. R., and Elie-Fadrosh, E. A. (2021) Ecology and molecular targets of hypermutation in the global microbiome, *Nat. Commun.*, **12**, 3076, doi: 10.1038/s41467-021-23402-7.
 41. Paul, B. G., and Eren, A. M. (2022) Eco-evolutionary significance of domesticated retroelements in microbial genomes, *Mobile DNA*, **13**, 6, doi: 10.1186/s13100-022-00262-6.
 42. Fortier, L. C., Bouchard, J. D., and Moineau, S. (2005) Expression and site-directed mutagenesis of the lactococcal abortive phage infection protein AbiK, *J. Bacteriol.*, **187**, 3721-3730, doi: 10.1128/jb.187.11.3721-3730.2005.
 43. Lopatina, A., Tal, N., and Sorek, R. (2020) Abortive infection: bacterial suicide as an antiviral immune strategy, *Annu. Rev. Virol.*, **7**, 371-384, doi: 10.1146/annurev-virology-011620-040628.
 44. Wang, C., Villion, M., Semper, C., Coros, C., Moineau, S., and Zimmerly, S. (2011) A reverse transcriptase-related protein mediates phage resistance and polymerizes untemplated DNA *in vitro*, *Nucleic Acids Res.*, **39**, 7620-7629, doi: 10.1093/nar/gkr397.
 45. Mestre, M. R., Gao, L. A., Shah, S. A., López-Beltrán, A., González-Delgado, A., Martínez-Abarca, F., Iranzo, J., Redrejo-Rodríguez, M., Zhang, F., and Toro, N. (2022) UG/Abi: a highly diverse family of prokaryotic reverse transcriptases associated with defense functions, *Nucleic Acids Res.*, **50**, 6084-6101, doi: 10.1093/nar/gkac467.
 46. Figiel, M., Gapińska, M., Czarnocki-Cieciura, M., Zajko, W., Sroka, M., Skowronek, K., and Nowotny, M. (2022) Mechanism of protein-primed template-independent DNA synthesis by Abi polymerases, *Nucleic Acids Res.*, **50**, 10026-10040, doi: 10.1093/nar/gkac772.
 47. Zimmerly, S., and Wu, L. (2015) An unexplored diversity of reverse transcriptases in bacteria, *Microbiol. Spectrum*, **3**, MDNA3-0058-2014, doi: 10.1128/microbiolspec.MDNA3-0058-2014.
 48. Simon, D. M., and Zimmerly, S. (2008) A diversity of uncharacterized reverse transcriptases in bacteria, *Nucleic Acids Res.*, **36**, 7219-7229, doi: 10.1093/nar/gkn867.
 49. Kojima, K. K., and Kanehisa, M. (2008) Systematic survey for novel types of prokaryotic retroelements based on gene neighborhood and protein architecture, *Mol. Biol. Evol.*, **25**, 1395-1404, doi: 10.1093/molbev/msn081.
 50. Toro, N., Martinez-Abarca, F., Mestre, M. R., and Gonzalez-Delgado, A. (2019) Multiple origins of reverse transcriptases linked to CRISPR-Cas systems, *RNA Biol.*, **16**, 1486-1493, doi: 10.1080/15476286.2019.1639310.
 51. 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 a natural reverse transcriptase-Cas1 fusion protein, *Science*, **351**, aad4234, doi: 10.1126/science.aad4234.
 52. Mohr, G., Silas, S., Stamos, J. L., Makarova, K. S., Markham, L. M., Yao, J., Lucas-Elio, P., Sanchez-Amat, A., Fire, A. Z., Koonin, E. V., and Lambowitz, A. M. (2018) A reverse transcriptase-Cas1 fusion protein contains a Cas6 domain required for both CRISPR RNA biogenesis and RNA spacer acquisition, *Mol. Cell*, **72**, 700-714, doi: 10.1016/j.molcel.2018.09.013.

53. Park, S. K., Mohr, G., Yao, J., Russell, R., and Lambowitz, A. M. (2022) Group II intron-like reverse transcriptases function in double-strand break repair, *Cell*, **185**, 3671-3688.e3623, doi: 10.1016/j.cell.2022.08.014.
54. Fragkiadaki, P., Renieri, E., Kalliantasi, K., Kouvidi, E., Apalaki, E., Vakonaki, E., Mamoulakis, C., Spandidos, D. A., and Tsatsakis, A. (2022) Telomerase inhibitors and activators in aging and cancer: A systematic review, *Mol. Med. Rep.*, **25**, 158, doi: 10.3892/mmr.2022.12674.
55. 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.
56. Yushenova, I. A., and Arkhipova, I. R. (2018) Biochemical properties of bacterial reverse transcriptase-related (*rvt*) gene products: multimerization, protein priming, and nucleotide preference, *Curr. Genet.*, **64**, 1287-1301, doi: 10.1007/s00294-018-0844-6.
57. Dlakic, M., and Mushegian, A. (2011) Prp8, the pivotal protein of the spliceosomal catalytic center, evolved from a retroelement-encoded reverse transcriptase, *RNA*, **17**, 799-808, doi: 10.1261/rna.2396011.
58. Grainger, R. J., and Beggs, J. D. (2005) Prp8 protein: at the heart of the spliceosome, *RNA*, **11**, 533-557, doi: 10.1261/rna.2220705.
59. 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.
60. Lambert, M. E., McDonald, J. F., and Weinstein, I. B. (1988) *Eukaryotic Transposable Elements as Mutagenic Agents*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
61. Arkhipova, I. R. (2012) *Telomerase, retrotransposons, and evolution*, in *Telomerases: Chemistry, Biology, and Clinical Applications* (Lue, N. F., and Autexier, C., eds.) John Wiley & Sons, Inc., Hoboken, NJ, pp. 265-299, doi: 10.1002/9781118268667.ch11.
62. Lue, N. F., and Autexier, C. (2006) The structure and function of telomerase reverse transcriptase, *Annu. Rev. Biochem.*, **75**, 493-517, doi: 10.1146/annurev.biochem.75.103004.142412.
63. Geer, L. Y., Domrachev, M., Lipman, D. J., and Bryant, S. H. (2002) CDART: protein homology by domain architecture, *Genome Res.*, **12**, 1619-1623, doi: 10.1101/gr.278202.
64. Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., and Geer, L. Y. (2015) CDD: NCBI's conserved domain database, *Nucleic Acids Res.*, **43**, D222-D226, doi: 10.1093/nar/gku1221.

TO BE MOBILE OR NOT: THE VARIETY OF REVERSE TRANSCRIPTASES AND THEIR RECRUITMENT BY HOST GENOMES

Review

I. R. Arkhipova* and I. A. Yushenova

*Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory,
Woods Hole, MA 02543 USA; e-mail: iarkhipova@mbl.edu, iyushenova@mbl.edu*

Reverse transcriptases (RT), or RNA-dependent DNA polymerases, are unorthodox enzymes that originally added a new angle to the conventional view of the unidirectional flow of genetic information in the cell from DNA to RNA to protein. First discovered in vertebrate retroviruses, RTs were since re-discovered in most eukaryotes, bacteria, and archaea, spanning essentially all domains of life. For retroviruses, RTs provide the ability to copy the RNA genome into DNA for subsequent incorporation into the host genome, which is essential for their replication and survival. In cellular organisms, most RT sequences originate from retrotransposons, the type of self-replicating genetic elements that rely on reverse transcription to copy and paste their sequences into new genomic locations. Some retroelements, however, can undergo domestication, eventually becoming a valuable addition to the overall repertoire of cellular enzymes. They can be beneficial yet accessory, like the diversity-generating elements, or even essential, like the telomerase reverse transcriptases. Nowadays, ever-increasing numbers of domesticated RT-carrying genetic elements are being discovered. It may be argued that domesticated RTs and reverse transcription in general is more widespread in cellular organisms than previously thought, and that many important cellular functions, such as chromosome end maintenance, may evolve from an originally selfish process of converting RNA into DNA.

Keywords: reverse transcription, RNA-dependent DNA polymerase, telomerase reverse transcriptase