

## РОЛЬ РЕТРОЭЛЕМЕНТОВ В ЭВОЛЮЦИИ ГЕНОМОВ ЖИВОТНЫХ

© 2021 г. Л. Н. Нефедова<sup>1</sup> \*, А. И. Ким<sup>1</sup>

<sup>1</sup>Московский государственный университет им. М.В. Ломоносова,  
биологический факультет, кафедра генетики  
Ленинские горы, 1/12, Москва, 119234 Россия

\*E-mail: lidia\_nefedova@mail.ru

Поступила в редакцию 12.06.2020 г.

После доработки 26.08.2020 г.

Принята к публикации 28.09.2020 г.

Ретроэлементы (ретротранспозоны и эндогенные ретровирусы) представляют собой класс мобильных генетических элементов, широко представленный у различных групп эукариот. Считается, что большинство ретроэлементов, присутствующих в геноме, не оказывает значительного влияния на организм хозяина, который в процессе коэволюции с ретроэлементами вырабатывает различные механизмы контроля их транспозиции. Но даже лишённые транспозиционной активности ретроэлементы могут играть существенную роль в эволюции, способствуя формированию новых адаптивных механизмов, обусловленных, в свою очередь, формированием новых генетических сетей. Процесс, в котором организм адаптирует отдельные последовательности ретроэлементов, называется молекулярной доместиацией. Экзаптация подразумевает формирование новых генетических структур и функций, как правило, появившихся в результате продолжительной эволюции доместичированных последовательностей мобильных элементов. Для трех ключевых генов ретротранспозонов/эндогенных ретровирусов (*gag*, *pol* и *env*) известны случаи доместики. Феномен доместики/экзаптации свидетельствует о фундаментальной роли, которую играют ретроэлементы в процессе эволюции генома, и значительной эволюционной выгоде присутствия в геноме ретроэлементов, несмотря на то, что их транспозиционная активность может сопровождаться негативными эффектами на уровне индивидуального организма.

DOI: 10.31857/S0044459621010036

Ретроэлементы (ретротранспозоны и эндогенные ретровирусы) представляют собой класс мобильных элементов, широко представленный в геномах различных групп эукариот. Они занимают около 10% генома нематоды *Caenorhabditis elegans*, 15% генома дрозофилы, 35% генома мыши и 40% генома человека (Chénais et al., 2012). Очевидно, что ретроэлементы – весомый компонент, который не может не влиять на функционирование и эволюцию генома в целом.

Считается, что подавляющее большинство копий ретроэлементов, присутствующих в геноме, не оказывают значительного влияния на организм хозяина, поскольку они утратили функциональную активность. Кроме того, в процессе совместной коэволюции генома с его мобильным компонентом сформировались защитные генетические механизмы, эффективно подавляющие транспозицию (Goodier, 2016). Эти механизмы связаны главным образом с РНК-интерференцией и эпигенетической регуляцией. У позвоночных активируются различные типы чувствительных к вирусной нуклеиновой кислоте рецепторов, включая Toll-подобные рецепторы, RIG-I/MDA5,

IFI16, cGAS и другие (Hornung, 2014). У млекопитающих в инактивации ретроэлементов значительную роль играют РНК- и ДНК-редактирующие ферменты семейств ADAR (adenosine deaminase acting on RNA) (Samuel, 2011) и APOBEC (apolipoprotein B mRNA-editing enzyme catalytic polypeptide) (Cullen, 2006). Дезаминазы ADAR1 и ADAR2 участвуют в редактировании РНК большого количества вирусов, в том числе ретровирусов, при этом ADAR может модифицировать вирусную РНК сайт-специфически, как это происходит, например, у вируса гепатита дельта и ВИЧ, а может вызывать летальное гиперредактирование, редактируя множество аденозинов (Samuel, 2011). APOBEC3G проявляет антиретровирусную иммунную активность против ретровирусов: он способен проникать внутрь вирусной частицы и дезаминировать одноцепочечную ДНК, обратную транскрибированную с РНК вируса (Cullen, 2006). Однако лентивирусы, в том числе ВИЧ, в ходе эволюции приобрели ген фактора вирусной инфекционности Vif, который способен взаимодействовать с APOBEC3G и запускать убиквити-

нирование и деградацию APOBEC3G через протеасомный путь (Cullen, 2006).

В ряде исследований показана связь между активностью ретроэлементов и патологиями, снижающими жизнеспособность организма. Выявлены заболевания, обусловленные интеграцией ретроэлементов в гены или их регуляторные последовательности. У человека это различные типы гемофилии и анемии, рак молочной железы, мышечная дистрофия, нейрофиброматоз и аутоиммунные заболевания (Каер, Speek, 2013).

Активация транспозиции чаще всего происходит в том случае, когда вирус или мобильный элемент попадает в организм впервые, и механизмы контроля его транспозиции еще не выработаны, или в том случае, когда имеющиеся механизмы нарушаются (Goodier, 2016). Факторами, нарушающими механизмы контроля мобильных элементов, могут служить мутации в контролирующих генах, а также различные стрессовые факторы (Li, Schmid, 2001). Нарушение контроля транспозиции приводит к геномной нестабильности, выражающейся в значительном повышении темпов мутационного процесса (Vela et al., 2014). Увеличение числа копий ретроэлементов в геноме приводит, в свою очередь, к увеличению частоты рекомбинационных событий и к геномным перестройкам, в том числе к дупликациям последовательностей (Kazazian, 2014). Дупликации могут происходить не только за счет рекомбинации, но и за счет активности обратной транскриптазы ретроэлементов, с которой связывают процесс образования процессированных псевдогенов. Так, в геноме человека насчитывается около восьми тысяч процессированных псевдогенов, 10% из которых транскрипционно активны (Kazazian, 2014). Дупликации, связанные с L1-трандукциями, составляют более 1% генома человека (Pickeral et al., 2000). Таким образом, ретроэлементы способствуют дупликации генов — одному из основных факторов, способствующих эволюционному процессу и приобретению организмом новых функций (Kaessmann et al., 2009).

Но даже лишенные транспозиционной активности, ретротранспозоны могут играть значимую роль в эволюции, способствуя формированию новых адаптивных механизмов. Процесс, в котором организм адаптирует последовательности мобильных элементов для собственной пользы, называют молекулярной доместиацией. Термин был предложен Миллером в 1997 г., чтобы описать эти явления (Miller et al., 1997).

Существует множество примеров, в которых доместикация мобильных элементов привела к формированию новых фундаментальных биологических функций. Наиболее известные из них — адаптация к участию в удлинении теломер ретротранспозонов *HeT-A*, *TAHRE* и *TART* у *Drosophila*

(Silva-Sousa et al., 2012), а также *SART* и *TRAS* у *Bombyx mori* (Fujiwara et al., 2005); доместикация транспозазы транспозона *piggyBac*, участвующей в геномных перестройках у *Paramecium* при образовании соматического макронуклеуса (Baudry et al., 2009); доместикация транспозазы мобильного элемента *Transib* и появление гена *RAG1*, который необходим для рекомбинации V(D)J-сегментов генов иммуноглобулинов (Huang et al., 2016).

Важнейшее понятие, применимое к эволюционной роли последовательностей мобильных элементов, — “экзаптация”. Экзаптация подразумевает формирование новых генетических структур и функций, как правило, появившихся в результате продолжительной эволюции доместичированных последовательностей мобильных элементов (Schrader, Schmitz, 2019).

В настоящем обзоре будут рассмотрены случаи доместикации последовательностей ретроэлементов, которые сопровождались экзаптацией и привели к появлению новых регуляторных последовательностей и генов с адаптивными функциями у животных.

#### СТРУКТУРНОЕ СХОДСТВО ДКП-РЕТРОТРАНСПОЗОНОВ И РЕТРОВИРУСОВ СВИДЕТЕЛЬСТВУЕТ ОБ ИХ ОБЩЕМ ПРОИСХОЖДЕНИИ

В настоящее время ретроэлементы делят на четыре группы: ретроэлементы с длинными концевыми повторами (ДКП) — ДКП-ретротранспозоны, ретроэлементы без ДКП, ретроэлементы с домени тирозиновой рекомбиназы (YR-элементы) и *Penelope* (Eickbush, Jamburuthugoda, 2008). Некоторые ДКП-ретротранспозоны имеют три открытые рамки считывания (ОРС), гомологичные ретровирусным генам *gag*, *pol* и *env*, и проявляют удивительное сходство по строению с интегрированными в геном копиями ретровирусов — про-вирусами.

И ДКП-ретротранспозоны, и интегрированные копии ретровирусов содержат на концах полностью идентичные друг другу ДКП, образующиеся в процессе обратной транскрипции, и имеют довольно консервативную структуру (что свидетельствует об их однократном появлении в процессе эволюции ретроэлементов) и варьирующую обычно в пределах нескольких сотен нуклеотидов длину (Benachene et al., 2013). Почти все ДКП-ретротранспозоны имеют сложно устроенные промоторные области, включающие в себя базальные промоторные элементы, отстоящие на несколько пар нуклеотидов от старта транскрипции, и дистальные промоторные элементы (энхансеры и сайленсеры), влияющие на базальный уровень транскрипции. При этом действие последних может быть тканеспецифично и может

изменяться в зависимости от стадии развития организма (Trizzino et al., 2018). Сравнение нуклеотидных последовательностей двух ДКП одного ретроэлемента может дать информацию о “возрасте” интегрированной копии (Bowen, McDonald, 2001): абсолютная идентичность двух ДКП одной копии ретроэлемента свидетельствует о ее недавней интеграции; чем больше нуклеотидных замен (мутаций) накапливают ДКП, тем дольше они находятся в данном сайте генома.

Кодирующую часть ДКП-ретротранспозонов обычно составляют две открытые рамки считывания (ОРС), гомологичные ретровирусным генам *gag* и *pol*, а у некоторых ДКП-ретротранспозонов имеется еще и ОРС3, гомологичная гену *env* ретровирусов (Kim et al., 1994). Помимо генов *gag*, *pol* и *env*, в геноме ретровирусов (лентивирусов) могут содержаться дополнительные гены — *vif*, *vpr*, *vpu*, *tat*, *rev* и *nef*, продукты которых необходимы для репликации и участвуют в регуляции синтеза и созревания вирусной РНК (Robertson, 2016).

Основной функцией ОРС1 (*gag*) является формирование нуклеопротеиновых комплексов — вирусоподобных частиц (ВПЧ), сходных с вирусными частицами ретровирусов. ВПЧ осуществляют транспорт обратных транскриптов к ядру. Кроме того, образование ВПЧ способствует изоляции обратной транскриптазы и других кодируемых ретроэлементами ферментов, что уменьшает вероятность образования процессированных псевдогенов.

Продукт ОРС2 (*pol*) обычно синтезируется в результате сдвига рибосомы на –1 нуклеотид при трансляции транскрипта *gag-pol*, но у некоторых ретроэлементов встречается сдвиг на +1 нуклеотид (Farabaugh, 1996) и механизм трансляционной супрессии (Yoshinaka et al., 1985). ОРС2 кодирует протеазу, обратную транскриптазу, РНКазу Н и интегразу, которые синтезируются как единый полипептид, процессируемый протеазой. Обратная транскриптаза образует кДНК-копию РНК ретроэлемента, РНКазы Н удаляет скопированную РНК из образующегося гетеродуплекса РНК-ДНК, а интеграна осуществляет встраивание обратного транскрипта в геном хозяина.

Продукт третьей ОРС (*env*) образуется с помощью сплайсинга основного транскрипта. Ген *env* ретровирусов кодирует компоненты, которые ответственны за взаимодействие вирусной частицы с мембранными рецепторами клетки, в результате чего вирус проникает в нее. Раньше считали, что ретровирусы есть только у позвоночных, и основным различием ретровирусов и ретроэлемента является именно наличие у первых гена *env*. Однако были найдены несколько семейств ДКП-ретротранспозонов с функциональной ОРС3 и получены экспериментальные доказательства того,

что некоторые ДКП-ретротранспозоны беспозвоночных (*Drosophila melanogaster*) являются истинными ретровирусами (Kim et al., 1994). В связи с этим классификация ретровирусов была пересмотрена.

Современная классификация ретровирусов включает ретроэлементы и базируется на сходстве последовательностей обратных транскриптаз и интеграз. Ретровирусы и родственные им ДКП-ретротранспозоны объединены в порядок вирусов *Ortervirales*, включающий пять семейств: *Pseudoviridae*, *Belpaoviridae*, *Caulimoviridae*, *Metaviridae* и *Retroviridae* (Krupovic et al., 2018). К сем. *Pseudoviridae* относят ДКП-ретроэлементы группы *Ty1/copia*. Оно включает три рода: *Pseudovirus* (элемент *Ty1* у *Saccharomyces cerevisiae*), *Hemivirus* (элемент *copia* у *D. melanogaster*) и *Sirevirus* (*SIRE*-элементы у хвойных и цветковых растений). Сем. *Belpaoviridae* включает ДКП-ретроэлементы группы *Bel/Pao* (р. *Semotiviridae*). К сем. *Caulimoviridae* относят вирусы растений с необычным жизненным циклом (дцДНК-вирусы с обратной транскрипцией, в отличие от ретровирусов, являющихся оц(+)РНК-вирусами с обратной транскрипцией). Вирусы этого семейства утратили интегразный домен и длинные концевые повторы, но по остальным свойствам близки к ретровирусам. Примечательно, что каулимовирусы формируют огромные семейства эндогенных вирусных элементов — эндогенные параретровирусы (Chabannes, Iskra-Caruana, 2013). Сем. *Metaviridae* включает ДКП-ретроэлементы группы *Ty3/gypsy*. В него входит два рода: *Metavirus* (элементы *Ty3* у *S. cerevisiae*) и *Errantivirus* (элемент *gypsy* у *D. melanogaster* и близких видов). К сем. *Retroviridae* относятся собственно ретровирусы, как инфекционные (экзогенные), так и неинфекционные (эндогенные).

Таким образом, отличия ретровируса и ДКП-ретротранспозона носят формальный характер и заключаются главным образом в наличии/отсутствии инфекционных свойств, которые во многом определяются функциональным геном *env*. Гена *env* нет у многих ДКП-ретротранспозонов, например у ДКП-ретротранспозонов с двумя генами *gag* и *pol*. Но это не означает, что они не могут воспользоваться функцией гена *env* других ДКП-ретротранспозонов или ретровирусов. В то же время существуют “неполноценные” ретровирусы (например, вирусы сарком — Ha-MSV, MoMSV, SSV), не имеющие собственных генов *env* и *pol*. Такие ретровирусы проявляют инфекционные свойства только при коинфекции с другими ретровирусами (дефектный ретровирус использует капсид и ферменты репликации полноценного ретровируса), что мало отличает их от многих ДКП-ретротранспозонов (Horie, Tomonaga, 2011). Процесс эндогенизации является ключевым в переходе от ретровируса к ДКП-ретротранспозону

и заключается в интеграции вируса в геном клеток на ранней эмбриональной стадии и деградации последовательности, отвечающей за инфекционные свойства вируса. Рекомбинация может приводить к обратному переходу ДКП-ретротранспозона в ретровирус.

### НЕКОДИРУЮЩИЕ ПОСЛЕДОВАТЕЛЬНОСТИ РЕТРОТРОЭЛЕМЕНТОВ УЧАСТВУЮТ В РЕГУЛЯЦИИ ЭКСПРЕССИИ ХОЗЯЙСКИХ ГЕНОВ

Ретроэлементы составляют значительную часть генома эукариот и неизбежно оказывают влияние на структуру и функционирование генома организма-хозяина. Очевидно, что интегрированные копии ретроэлементов или их фрагменты, находящиеся рядом с какими-либо регуляторными последовательностями генов, иногда внутри них или даже сами по себе, могут изменять экспрессию генов.

Последовательности ретротранспозонов, интегрированные в интроны генов, могут влиять на сплайсинг, предоставляя новые донорные или акцепторные сайты сплайсинга, а также могут служить донорами новых экзонов (Sela et al., 2010). Так, установлено, что около 4% кодирующих последовательностей белок-кодирующих генов человека включают экзонизированные последовательности Alu и L1 (Nekrutenko, Li, 2001). Кроме того, описан процесс “интронизации” с участием ретроэлементов (Sela et al., 2007). С-концевой экзон рецептора лептина 219.1 происходит от составного неавтономного элемента SVA (из его последовательности, гомологичной эндогенному ретровирусу HERV-K), в результате на С-конце белка добавляется последовательность из 67 аминокислот (Damert et al., 2004).

В геноме человека содержание неДКП-ретроэлементов составляет около 35%. Из-за наличия в их последовательностях полиА-трактов неДКП-ретротранспозоны служат источником микросателлитных повторов во многих локусах генома: около 20% всех микросателлитных повторов, включая 50% мононуклеотидных микросателлитов, находятся внутри Alu-элементов (Kelkar et al., 2008). Ретроэлементы L1 и Alu могут также служить источником внутригенных сигналов полиаденилирования (Roy-Engel et al., 2005). Анализ различных типов альтернативных сайтов полиаденилирования в последовательностях генов человека, мыши, крысы и курицы показал, что все они происходят из последовательностей ретроэлементов (Lee et al., 2008).

Помимо полиА-трактов, последовательности неДКП-ретротранспозонов обогащены CpG-повторами, служащими островками метилирования,

одна треть всех сайтов CpG человека содержится в последовательностях Alu (Rubin et al., 1994). Поскольку элементы L1, Alu и SVA часто обнаруживаются в генах или рядом с ними, образование и распространение гетерохроматина, опосредованное ретротранспозоном, может репрессировать транскрипцию близлежащих генов (Xing et al., 2004).

Ретроэлементы несут сайты связывания факторов транскрипции, которые могут служить для модуляции экспрессии близлежащих генов (Polak, Domany, 2006). Почти 25% промоторных областей человека содержат последовательности, производные от мобильных элементов, в основном SINE (Santangelo et al., 2007). Существует множество примеров участия ДКП эндогенных ретровирусов, функционирующих в качестве новых промоторов или сайтов связывания факторов транскрипции для генов, и теперь стали известны примеры формирования регуляторных длинных некодирующих РНК из последовательностей эндогенных ретровирусов (Kapusta et al., 2013; Deng et al., 2019).

Ретроэлементы для генома человека стали донорами около 28% локализаций хроматиновых меток активных энхансеров — H3K4me1 (Nikitin et al., 2019) — и 20–30% локализаций хроматиновых меток промоторов, активного хроматина и гетерохроматина — H3K4me3, H3K9ac, H3K27ac, H3K27me3, H3K9me3 (Igolkina et al., 2019).

Ретроэлементы могут также служить источником энхансерных последовательностей. Таким примером является регуляция экспрессии гена *Pparg*, одного из компонентов сигнального пути адипогенеза (развития жировой ткани) у человека и животных. Транскрипция гена *Pparg* управляется энхансером, находящимся вблизи *Pparg* и обладающим особой структурой. Этот энхансер состоит из нескольких видоизмененных последовательностей ретроэлемента *LINE1* и служит местом посадки белков ZFP30 и KAP1. ZFP30 является транскрипционным фактором семейства KRAB-ZF, обычно подавляющим транскрипцию генов-мишеней. Однако в случае *Pparg* ZFP30 действует как активатор транскрипции, связываясь с *LINE*-подобной последовательностью энхансера и привлекая коактиватор KAP1 (что прямо противоположно “основной” функции ZFP30 и KAP1) (Chen et al., 2019).

У человека описана длинная некодирующая РНК PRLH1, транскрибируемая с копий последовательностей эндогенного ретровируса ERV-9 (не содержащих в себе функциональных OРС) и участвующая в репарации двунитевых разрывов. Транскрипция этой РНК стимулируется белком p53 (одним из важнейших эффекторов консервативного сигнального пути, реагирующего на двунитевые разрывы). Получившаяся молекула РНК связывается с RNF169, стабилизируя его, а также

привлекая дополнительные белки. Благодаря этому активируется репарация двунитевых разрывов по пути гомологичной рекомбинации (но не по пути NHEJ) (Deng et al., 2019). У ряда других транскриптов ERV-9 выявлена регуляторная функция: они стимулируют транскрипцию ниже лежащих генов, связываясь с ДНК ERV-9 и способствуя сборке энхансерных белковых комплексов. По такому механизму, в частности, регулируется транскрипция генов бета-глобинового кластера у человека (Hu et al., 2017).

В геноме домашней лошади идентифицирована новая микроРНК (обозначенная как eca-miR-1804), образовавшаяся в ходе эволюции из ДКП ретроэлемента E<sub>q</sub>ERV-E1. Авторы предполагают, что ДКП E<sub>q</sub>ERV-E1 стал источником микроРНК, которая экспрессируется тканеспецифично (Jo et al., 2019).

Недавние работы выявили активацию ретроэлементов в клетках мозга человека (Frank et al., 2005), фундаментальную роль экспрессии определенных ретроэлементов в предимплантационный период развития эмбриона (Grow et al., 2015), роль эндогенных ретровирусов в поддержании плюрипотентности эмбриональных стволовых клеток (Tanaka et al., 2012). Эти работы ярко демонстрируют участие регуляторных последовательностей ретроэлементов в регуляции функционирования генома, и значение такой регуляции еще предстоит выяснить.

В последнее время появилось несколько свидетельств того, что эндогенные ретровирусы могут способствовать согласованной эволюции наборов генов, экспрессия которых регулируется согласованно в составе генной регуляторной сети (Chuong et al., 2016). Скоординированная регуляция экспрессии генов обусловлена наличием большого количества последовательностей цис-действующих регуляторных последовательностей, происходящих из ретроэлементов, а полиморфизм таких регуляторных последовательностей может быть источником фенотипических вариантов транскриптомов (Carroll, 2008).

Реструктуризация генома, опосредованная активностью ретроэлементов, может играть важную роль в ответе хозяина на стресс, облегчая адаптацию популяций и видов к изменяющимся условиям среды. Инсерции мобильных элементов в определенные сайты генома происходят с высокой частотой, что может указывать на селективное преимущество инсерционных мутаций (Chung et al., 2007). Примером может служить ген *Cyp6g1*, кодирующий цитохром, участвующий в детоксикации инсектицидов. Некоторые лабораторные линии *D. melanogaster* несут инсерцию ДКП-ретротранспозона *Accord* в 5'-некодирующую область гена *Cyp6g1*, что вызывает его повышенную экспрессию и приводит к инсектицидной устойчивости у

этих линий (Chung et al., 2007). Эти исследования, проведенные на лабораторных линиях, полностью коррелируют с исследованиями на природных популяциях, собранных по всему миру, где также была показана прямая корреляция между резистентностью к инсектициду и наличием инсерции *Accord* (Catania et al., 2004).

Устойчивость к фосфорорганическим инсектицидам в первую очередь обусловлена сверхэкспрессией генов неспецифических эстераз, что наблюдается у многих устойчивых видов, например *Culex quinquefasciatus*, что обусловлено присутствием элемента *LINE* вблизи локуса *Est-2* (Mouches et al., 1990). Наконец, для естественной популяции *C. pipiens* показано, что встраивание мобильного элемента в кодирующую последовательность рецептора токсина кишечника вызывает событие альтернативного сплайсинга новой мРНК, которое вызывает экспрессию измененного мембранного белка, не способного взаимодействовать с бинарным токсином *Bacillus sphaericus* (Bt) (Darboux et al., 2007).

Корреляция между условиями окружающей среды и эволюцией последовательности мобильных элементов у позвоночных все еще остается неизученной областью. К немногочисленным свидетельствам участия ретроэлементов в адаптации к условиям окружающей среды можно отнести работу Кардуччи с соавт. (Carducci et al., 2019). Ретроэлемент *Rex3*, относящийся к неДКП-ретротранспозонам, широко распространен у костистых рыб. Сравнительный анализ последовательностей *Rex3* у 39 различных видов рыб, адаптированных к разным температурным условиям, показал, что определенные структурные особенности последовательности *Rex3* строго коррелируют с условиями обитания видов (Carducci et al., 2019).

## МОЛЕКУЛЯРНАЯ ДОМЕСТИКАЦИЯ ГЕНОВ РЕТРОЭЛЕМЕНТОВ

Важным путем приобретения новых полезных признаков является “заимствование” последовательностей генов ретроэлементов и их модификация для собственных нужд организма-хозяина — процесс, названный молекулярной доместикацией (Miller et al., 1997). Этот процесс начинается с интеграции новой копии ретроэлемента в геном (рис. 1). В некоторых случаях возможна интеграция не целой последовательности ДНК-копии ретроэлемента, а ее отдельных ОРС. Это происходит тогда, когда обратно транскрибируется и затем интегрируется в геном не полноразмерная копия ретроэлемента, а ее слайсированные формы, например ген *env*. Если интеграция последовательности ретроэлемента произошла в герминальных клетках, то она будет передана следующим поколениям, т.е. закрепится в геноме.

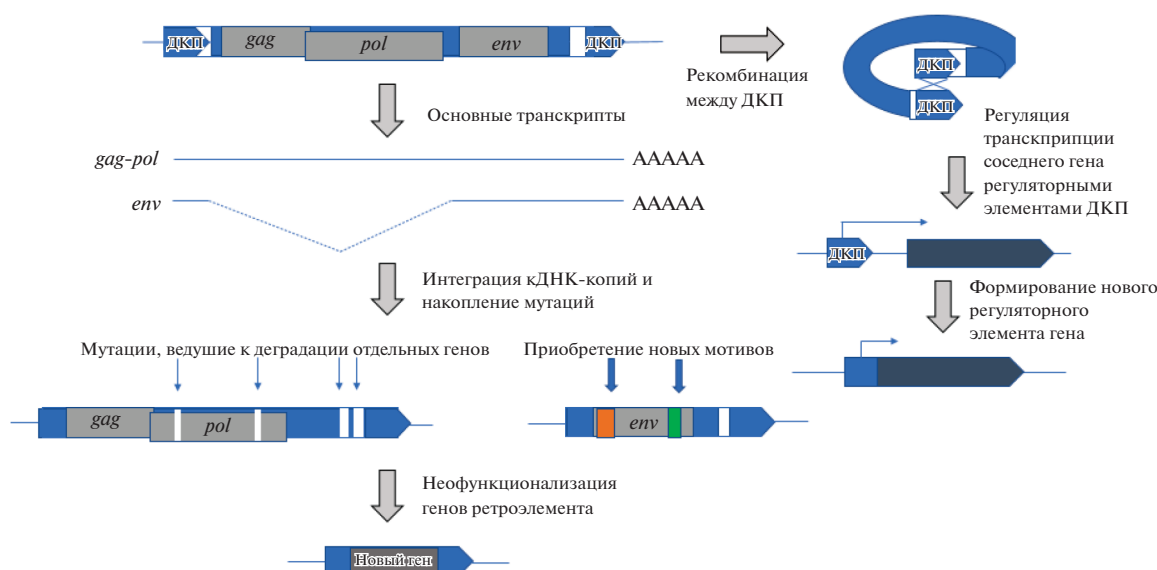


Рис. 1. Процесс доместикиции последовательностей ретроэлементов.

В ходе эволюции интегрированная последовательность ретроэлемента накапливает мутации, как и другие последовательности генома, и вскоре утрачивает функциональность, так как отбор не поддерживает ее сохранение в геноме. Однако в некоторых случаях отдельные последовательности ретроэлемента могут сохраняться отбором. Сохраняться могут как отдельные гены, так и отдельные последовательности генов. Более того, в процессе эволюции отдельные последовательности внутри доместичированного гена могут находиться под давлением движущего отбора, что приводит к формированию мотивов, не характерных для родительских генов. Появление таких мотивов и ведет к неофункционализации гена ретроэлемента — его экзаптации (рис. 1).

Процесс доместикиции может быть ускорен рекомбинацией между ДКП, в результате чего в геноме останется единственный ДКП, включающий регуляторные последовательности. Если расположение ДКП окажется “удачным”, его регуляторные мотивы могут быть “использованы” соседним геном, что впоследствии приведет к адаптивному изменению регуляции гена — образованию нового регуляторного элемента.

Чаще всего доместикиция происходит в тех случаях, когда последовательности ретроэлемента потенциально имеют адаптивное значение и могут оказаться полезными геному. Адаптация может обеспечить приспособляемость организма в условиях конкретного местообитания, устойчивость к воздействию факторов абиотического и биологического характера и др. В частности, доместичированный ген ретровируса может быть использован для защиты от инфекции его “прародителем” (Nair, Rein, 2014; Yap et al., 2014). Это очевидное

применение свойств ретроэлемента. Ниже мы рассмотрим известные случаи доместикиции генов *gag*, *pol* и *env*, которые подразумевают более широкое применение потенциала ретроэлементов в геноме хозяина.

#### Молекулярная доместикиция гена *gag*

Самые многочисленные семейства генов, появившихся в результате доместикиции гена *gag*, — это семейства *MART* (Mammalian Retrotransposon), оно же *SIRH* (Sushi-Ichi Retrotransposon Homologue), *PNMA* (Paraneoplastic Ma antigens) и *SCAN* (Schradler, Schmitz, 2019) (рис. 2). Семейства генов *MART* (*SIRH*) и *PNMA* обнаружены только у млекопитающих и возникли в результате независимых доместикиций ДКП-ретротранспозонов *sushi-ichi* и ДКП-ретротранспозонов группы *Ty3/gypsy* соответственно (Ono et al., 2011; Kokosar, Kordis, 2013). Семейство генов *SCAN* обнаружено у наземных позвоночных и возникло в результате доместикиции ДКП-ретротранспозона *Gmr1* группы *Ty3/gypsy* (Emerson, Thomas, 2011).

Семейство *MART* включает не менее 11 гомологичных генов у плацентарных млекопитающих и еще несколько генов у сумчатых, локализованных чаще всего на X-хромосоме (Ono et al., 2011). Многие из них произошли, по-видимому, в результате дупликации исходного доместичированного варианта. Тем не менее считается, что доместикиция ретротранспозонов *sushi-ichi* могла происходить независимо и многократно.

Гены семейства *MART* являются примером молекулярной экзаптации — ДКП-ретротранспозоны утратили ДКП, необходимые для транспозиции, но сохранили ОРС, гомологичные *gag*,

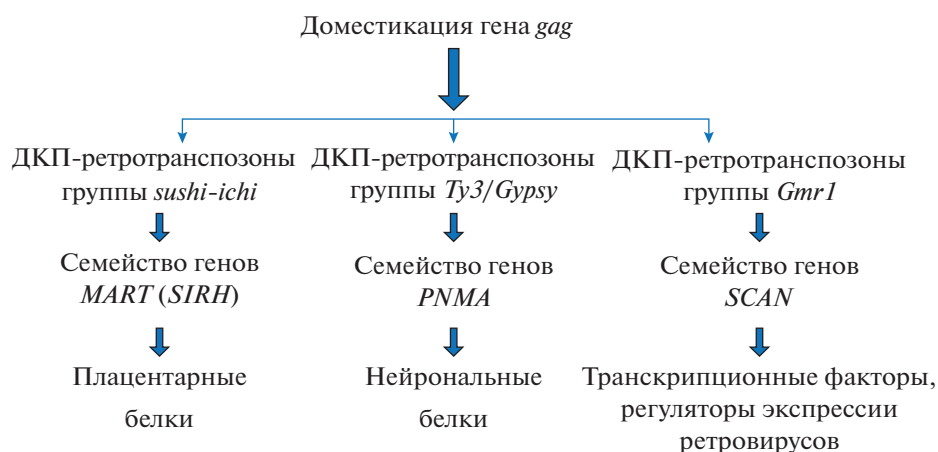


Рис. 2. Молекулярная доместицинация гена *gag* у млекопитающих.

которые приобрели в процессе эволюции новые функции у млекопитающих (Henke et al., 2015) (рис. 1). Некоторые из генов семейства *MART* сохранили последовательность, гомологичную *pol*, и используют сдвиг рамки считывания при трансляции полипротеина-предшественника *gag-pol*, что характерно для ретроэлементов и ретровирусов (Clark et al., 2007). Многие из генов семейства *MART* экспрессируются в плаценте, и по крайней мере три из них необходимы на ранних стадиях формирования плаценты и ее развития (Henke et al., 2015). Наиболее изученные гены этого семейства у человека — *LDOC1* и *PEG10*.

Возникновение генов семейства *PNMA* также происходило в ходе эволюции неоднократно в результате доместицинации разных ДКП-ретротранспозонов группы *Ty3/gypsy* (Kokosar, Kordis, 2013). Известно 12 генов этого семейства у мыши и 15 у человека, большинство из них расположены на X-хромосоме, многие гены семейства образовались в результате дупликаций (Kokosar, Kordis, 2013). Функции этих генов *PNMA* связаны с работой мозга. Аналогично некоторым генам семейства *MART* часть генов семейства *PNMA* сохранили последовательность, гомологичную *pol*, и используют сдвиг рамки считывания, характерный для ДКП-ретротранспозонов и ретровирусов.

Примером одной из наиболее успешных доместицинаций последовательности ДКП-ретротранспозонов являются последовательности, кодирующие домен *SCAN*. Этот домен локализован в N-концевой части ряда транскрипционных факторов млекопитающих. Помимо домена *SCAN*, такие транскрипционные факторы всегда содержат *KRAB*-домен (тандемы цинковых пальцев). Домен *SCAN* гомологичен C-концевой части белка *gag* ДКП-ретротранспозона *Gmr1* группы *Ty3/gypsy* (Emerson, Thomas, 2011). Предполагается, что эта область была приобретена ранними амниотами в результате инсерции последовательности *gag* в

ген семейства *KRAB-ZF*, с последующей потерей (делецией) области, кодирующей N-конец белка *gag*. Доказательством тому служит тот факт, что у ящерицы *Anolis* в геноме присутствуют как гены *KRAB-ZF*, содержащие домен *SCAN*, так и гены *KRAB-ZF*, содержащие полноразмерный *gag*-предшественник (Emerson, Thomas, 2011).

Несмотря на дальнейшую потерю ретротранспозонов *Gmr1* некоторыми группами амниот, в том числе млекопитающими, *SCAN*-домен сохраняется у продуктов многих *ZF*-генов. Геном мыши содержит около 40 таких генов, геном человека — около 70 (Emerson, Thomas, 2013). Многие из них являются транскрипционными факторами, контролирующими разнообразные биологические процессы — эмбриональное развитие, гемопоэз, метаболизм и др. (Schrader, Schmitz, 2019). При этом конкретная функция *SCAN*-домена неизвестна. Существует множество экспериментальных свидетельств его роли в качестве посредника белок-белковых взаимодействий, однако доказательства специфического связывания ограничиваются несколькими примерами.

В недавних исследованиях показано участие *SCAN*-содержащих белков в подавлении транскрипции эндогенных (Matsui et al., 2010; Rowe et al., 2010) и экзогенных ретровирусов (Barde et al., 2009). В то же время важная роль в подавлении транскрипции ретровирусов принадлежит домену *KRAB* (Wolf, Goff, 2009). Так, показано, что белки, содержащие *KRAB*-домен, способны подавлять интеграцию ВИЧ-1 (Allouch et al., 2011).

Наиболее известным примером, демонстрирующим участие доместицированных генов ретроэлементов в защите от экзогенной ретровирусной инфекции, является ген *Fv1* мышей (Best et al., 1996), который обеспечивает устойчивость к вирусу MLV120. *Fv1* произошел от гена *gag* эндогенного ретровируса MERV (Yap et al., 2014). Белок

Fv1 блокирует жизненный цикл ретровируса после стадии обратной транскрипции, взаимодействуя с белком капсида ретровируса и препятствуя его интеграции в геном хозяина. У мыши Fv1 способен предотвращать или задерживать спонтанные или экспериментально индуцированные вирусные опухоли (Nair, Rein, 2014).

Ортологи *Fv1* были идентифицированы у многих видов грызунов, и предполагаемое время внедрения составляет 45–50 млн лет назад (Boso et al., 2018). Некоторые ортологи *Fv1* ограничивают не только инфекцию MLV124, но и других ретровирусов, что, как предполагается, связано с тем, что *Fv1* не распознает консервативные аминокислотные мотивы, но вместо этого может обнаруживать структурно консервативные пространственные паттерны в гексамерной решетке, типичные для ядер ретровирусного капсида (Boso et al., 2018).

Локус EnJS56A1 домашних овец (*Ovis aries*) также кодирует белок Gag, который может действовать как ингибитор родственного экзогенного ретровируса овцы Jaagsiekte (JSRV) (Arnaud et al., 2007). В отличие от Fv1, который блокирует репликацию MLV вскоре после проникновения, EnJS56A1 действует на поздней стадии цикла репликации JSRV, блокируя транспортировку и сборку вирионов (Arnaud et al., 2007).

Gag-опосредованные противовирусные функции эндогенных ретровирусов человека не описаны, хотя было показано, что Gag HERV-K (HML2) ингибирует ВИЧ-1 в культуре клеток (Monde et al., 2012). Пока не известно, проявляется ли этот эффект *in vivo*.

Примером доместикиации *gag* у беспозвоночных служит ген дрозофилы *Gagr* (Gag-related protein), который представляет собой геномный гомолог гена *gag* ДКП-ретротранспозонов группы *gypsy*. Гомологи этого гена также обнаруживаются во всех секвенированных геномах видов р. *Drosophila* и обладают высокой консервативностью последовательности (Nefedova et al., 2014). Для гена *Gagr* проведен ряд исследований, указывающих на его участие в ответе на стресс различного происхождения (Graveley et al., 2011; Kemp et al., 2013; Makhnovskii et al., 2020).

#### Молекулярная доместикиация гена *pol*

Самым известным примером доместикиации гена *pol* является теломераза эукариот, которая произошла от обратной транскриптазы ретротранспозонов и служит для удлинения концов линейных хромосом (Belfort et al., 2011).

Частично происходит от последовательности *pol*, кодирующей интегразу, ген млекопитающих *CGIN1*, гомологи которого имеются и у других групп позвоночных. Считается, что ген *CGIN1*

произошел от слияния фрагмента *pol*, кодирующего интегразу, с одним из генов хозяина около 150 млн лет назад. Предполагается, что функция этого гена у млекопитающих связана с защитой от ретровирусов (Chalopin et al., 2012).

Консервативным геном всех млекопитающих является ген кожной аспарагиновой протеазы, которая произошла от ретровирусной аспарагиновой протеазы. Она играет ключевую роль в процессах регенерации, а также в определении текстуры кожи путем модуляции степени гидратации (Bernard et al., 2005).

#### Молекулярная доместикиация гена *env*

Ярким примером доместикиации гена *env* является семейство синцитинов (*Syncytins*) (рис. 3). Синцитины — это группа генов, которые появились у плацентарных млекопитающих в результате нескольких независимых актов доместикиации гена *env* эндогенных ретровирусов. У приматов известно два таких гена — *Syncytin-1* и *Syncytin-2*, которые происходят от гена *env* разных эндогенных ретровирусов — HERV-W и HERV-FRD, интегрированных 25 и более 40 млн лет назад соответственно (Mallet et al., 2004). Геном мыши также содержит два гена этой группы — *Syncytin-A* и *Syncytin-B*, произошедшие от ретровирусов семейства HERV-F/H примерно 20 млн лет назад. Оба они проявляют специфическую для плаценты экспрессию и фузогенные свойства. Нокаут гена *Syncytin-A* приводит к эмбриональной смерти (Dupressoir et al., 2005).

Синцитины экспрессируются в слое синцитиотрофобласта, структуры, образующей самый внешний компонент плаценты. Этот слой важен для взаимодействия между матерью и плодом. Синцитины способствуют слиянию трофобластов, что приводит к образованию синцитиотрофобласта. Некоторые синцитины обладают иммуносупрессорными свойствами, которые способствуют защите тканей плода от материнской иммунной системы. У человека, помимо образования плаценты, *Syncytin-1* может участвовать в слиянии миобластов (Frese et al., 2015), остеокластов (Soe et al., 2011) и может также регулировать нейровоспаление при рассеянном склерозе (Antony et al., 2007). Кроме грызунов и приматов, синцитины были также идентифицированы у других млекопитающих — зайцеобразных (Heidmann et al., 2009), хищников, парнокопытных и афросорицид (Cornelis et al., 2013, 2014). Все эти гены произошли в результате независимых инсерций эндогенных ретровирусов в зародышевую линию клеток, что указывает на рецидивизирующую конвергентную доместикацию *env*-последовательностей в гены группы синцитинов. Другой человеческий плацентарный белок, полученный из HERV — Suppressyn, ингибирует слияние трофоб-

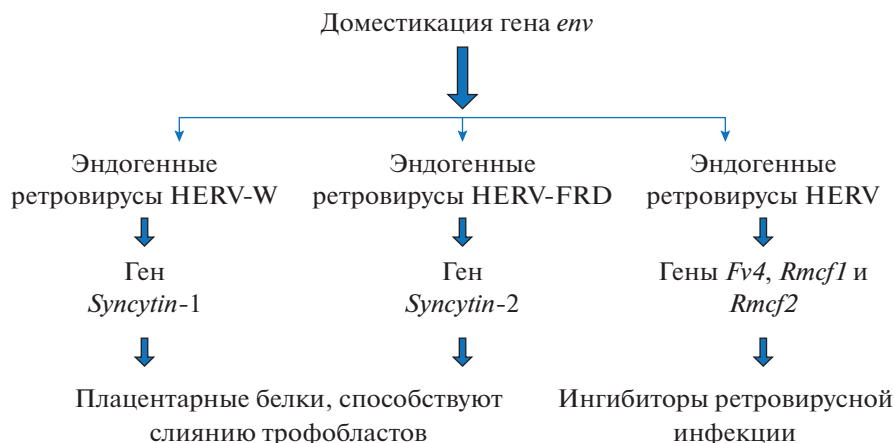


Рис. 3. Молекулярная доместикация гена *env* у млекопитающих.

ластов, конкурируя с Syncytin-1, тем самым контролируя его функцию в плаценте (Sugimoto et al., 2013).

Для доместицированных генов *env*, как и для доместицированных генов *gag*, показана и противовирусная активность.

В 1981 г. было показано, что три эндогенных локуса кур (EV3, EV6 и EV9) блокируют начальную стадию инфицирования вирусом лейкоза птиц, наиболее вероятно за счет интерференции рецепторов (Meyerson, Sawyer, 2011). Авторы исследования предсказали, что аналогичные функции будут обнаружены у других видов, которые содержат эндогенные ретровирусы.

Геном мыши содержит противовирусные гены — *Fv4*, *Rmcf1* и *Rmcf2*, которые представляют собой полноразмерные интактные гены *env* разных ретровирусов (Malfavon-Borja, Feschotte, 2015). Эти гены контролируют восприимчивость к инфекции экотропного вируса лейкоза мыши (MuLV) (Takeda, Matano, 2007).

В результате доместикации генов *env* эндогенных бета-ретровирусов овец JSRV сформировались факторы, ограничивающие инфекцию экзогенными JSRVs (Varela et al., 2009). Примечательно, что некоторые эндогенные JSRV участвуют в развитии плаценты (Arnaud et al., 2007).

Интересно, что инфицирование CD4+ Т-клеток человека ВИЧ-1 индуцирует экспрессию локусов HERV-K (HML2) (Young et al., 2018). Некоторые локусы HERV-K (HML2) кодируют интактные гены *env*, и экспрессия по крайней мере одного из них в лабораторных клеточных линиях привела к ингибированию инфекции ВИЧ-1 (Terry et al., 2017). Показано, что такое ингибирование происходит не из-за интерференции рецепторов, что указывает на существование нового, еще не изученного механизма.

Самыми древними интактными генами *env* эндогенных ретровирусов считаются ген *percomorf* лучеперых рыб (Henzy et al., 2017) и ген *HEMO* приматов (Heidmann et al., 2017). Показано, что эти гены находятся под действием стабилизирующего отбора более 100 млн лет. И тот, и другой сохранили фузогенную функцию и могут действовать в качестве рецепторов. Гомолог *HEMO* обнаружен в геноме человека и экспрессируется как полноразмерный белок Env, который расщепляется неизвестной клеточной протеазой с высвобождением усеченной внеклеточной формы (Heidmann et al., 2017). Секретируемая форма обнаруживается в крови беременных женщин, а также в крови и тканях плаценты, но ее функции остаются неизвестными.

У беспозвоночных явление молекулярной доместикации генов *env* ретротранспозонов изучено слабо. Ген *Iris* представляет собой геномный гомолог гена *env* ретровирусов у *D. melanogaster*. Гомологи гена *Iris* обнаружены в геномах всех видов подрода *Sophophora* (в том числе у группы видов *melanogaster* и *obscura*), для которых проводился их поиск (Malik et al., 2000). У этих видов наблюдается высокая консервативность последовательности гена *Iris*. Анализ гомологии последовательности *Iris* с геном *env* различных ретровирусов и ретротранспозонов насекомых показал, что наиболее близкими последовательностями являются *env* эррантивировусов Kanga. Кроме этого, *Iris* обладает достаточной гомологией с геном *env* ДКП-ретротранспозонов группы *gypsy* и геном *f* бакуловирусов (Malik et al., 2000). У взрослых особей экспрессия *Iris* значительно подавляется при холодовой акклиматизации (6°C) (MacMillan et al., 2016). В популяционных исследованиях показано, что экспрессия гена *Iris* зависит от географической широты, на которой обитает дикая популяция (Juneja et al., 2016).

## ЗАКЛЮЧЕНИЕ

Доместикация ретроэлементов и их белок-кодирующих последовательностей имеет большое значение в эволюции эукариот. На сегодняшний день у млекопитающих молекулярная доместикация привела к появлению нескольких семейств генов, которые играют важную роль в эмбриогенезе и формировании плаценты. Другим примером успешной доместикации ретроэлементов являются последовательности, кодирующие домен SCAN, многие из которых являются транскрипционными факторами, контролирующими разнообразные биологические процессы — эмбриональное развитие, гемопоэз и метаболизм. Молекулярная доместикация ретроэлементов у млекопитающих привела к появлению ряда генов со специфической противовирусной функцией.

К неразрешенным вопросам эволюции ретроэлементов можно отнести вопросы об их происхождении и распространении в геномах высших эукариот. Каким образом ретроэлементы появились в геномах высших эукариот? Были ли это процессы инфекции или горизонтального переноса? Почему именно ретроэлементы, а не другие типы мобильных элементов, например транспозоны, оккупировали геномы высших эукариот?

Процесс молекулярной доместикации ретроэлементов также поднимает много вопросов. Один из основных — почему процессы доместикации некоторых генов ретроэлементов происходили независимо и многократно? Очевидно одно — ретроэлементы обладают значительным потенциалом в процессе генообразования и регуляции функций генов и служат своеобразным резервуаром для эволюционных преобразований.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-14-50415 “Экспансия”.

## СПИСОК ЛИТЕРАТУРЫ

- Allouch A., Di Primio C., Alpi E., Lusic M., Arosio D. et al., 2011. The TRIM family protein KAP1 inhibits HIV-1 integration // *Cell Host Microbe*. V. 9. 16. P. 484–495.
- Antony J.M., Ellestad K.K., Hammond R., Imaizumi K., Mallet F. et al., 2007. The human endogenous retrovirus envelope glycoprotein, syncytin-1, regulates neuroinflammation and its receptor expression in multiple sclerosis: A role for endoplasmic reticulum chaperones in astrocytes // *J. Immunol*. V. 179. № 2. P. 1210–1224.
- Arnaud F., Caporale M., Varela M., Biek R., Chessa B. et al., 2007. Mechanisms of late restriction induced by an endogenous retrovirus // *J. Virol*. V. 81. P. 11441–11451.
- Barde I., Laurenti E., Verp S., Groner A.C., Towne C. et al., 2009. Regulation of episomal gene expression by KRAB/KAP1-mediated histone modifications // *J. Virol*. V. 83. № 11. P. 5574–5580.
- Baudry C., Malinsky S., Restituito M., Kapusta A., Rosa S. et al., 2009. PiggyMac, a domesticated piggyBac transposase involved in programmed genome rearrangements in the ciliate *Paramecium tetraurelia* // *Genes Dev*. V. 23. № 21. P. 2478–2483.
- Belfort M., Curcio M.J., Lue N.F., 2011. Telomerase and retrotransposons: Reverse transcriptases that shaped genomes // *PNAS*. V. 108. № 51. P. 20304–20310.
- Benachenhou F., Sperber G.O., Bongcam-Rudloff E., Anderson G., Boeke J.D., Blomberg J., 2013. Conserved structure and inferred evolutionary history of long terminal repeats (LTRs) // *Mob. DNA*. V. 4. № 1. P. 5.
- Bernard D., Mehul B., Thomas-Collignon A., Delattre C., Donovan M., Schmidt R., 2005. Identification and characterization of a novel retroviral-like aspartic protease specifically expressed in human epidermis // *J. Invest. Dermatol*. V. 25. № 2. P. 278–287.
- Best S., Le Tissier P., Towers G., Stoye J.P., 1996. Positional cloning of the mouse retrovirus restriction gene *Fv1* // *Nature*. V. 382. P. 826–829.
- Boso G., Buckler-White A., Kozak C.A., 2018. Ancient evolutionary origin and positive selection of the retroviral restriction factor *Fv1* in muroid rodents // *J. Virol*. V. 92. P. e00850-18.
- Bowen N.J., McDonald J.F., 2001. Drosophila euchromatic LTR retrotransposons are much younger than the host species in which they reside // *Genome Res*. V. 11. № 9. P. 1527–1540.
- Carducci F., Biscotti M.A., Forconi M., Barucca M., Canapa A., 2019. An intriguing relationship between teleost Rex3 retroelement and environmental temperature // *Biol. Lett*. V. 15. № 9. P. 20190279.
- Carroll S.B., 2008. Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution // *Cell*. V. 134. P. 25–36.
- Catania F., Kauer M.O., Daborn P.J., Yen J.L., Ffrench-Constant R.H., Schlotterer C., 2004. World-wide survey of an *Accord* insertion and its association with DDT resistance in *Drosophila melanogaster* // *Mol. Ecol*. V. 13. P. 2491–2504.
- Chabannes M., Iskra-Caruana M.-L., 2013. Endogenous pararetroviruses — a reservoir of virus infection in plants // *Curr. Opin. Virol*. V. 3. № 6. P. 615–620.
- Chalopin D., Galiana D., Volff J.-N., 2012. Genetic innovation in vertebrates: gypsy integrase genes and other genes derived from transposable elements // *Int. J. Evol. Biol*. V. 2012. P. 724519.
- Chen W., Schwalie P.C., Pankevich E.V., Gubelmann C., Raghav S.K. et al., 2019. ZFP30 promotes adipogenesis through the KAP1-mediated activation of a retrotransposon-derived *Pparg2* enhancer // *Nat. Commun*. V. 10. № 1. P. 1809.
- Chénais B., Caruso A., Hiard S., Casse N., 2012. The impact of transposable elements on eukaryotic genomes: From genome size increase to genetic adaptation to stressful environments // *Gene*. V. 509. № 1. P. 7–15.
- Chung H., Bogwitz M.R., McCart C., Andrianopoulos A., Ffrench-Constant R.H. et al., 2007. Cis-regulatory elements in the *Accord* retrotransposon result in tissue-specific expression of the *Drosophila melanogaster* insecticide resistance gene *Cyp6g1* // *Genetics*. V. 175. P. 1071–1077.

- Chuong E.B., Elde N.C., Feschotte C., 2016. Regulatory evolution of innate immunity through co-option of endogenous retroviruses // *Science*. V. 351. P. 1083–1087.
- Clark M.B., Janicke M., Gottesbuhren U., Kleffmann T., Legge M. et al., 2007. Mammalian gene *PEG10* expresses two reading frames by high efficiency –1 frameshifting in embryonic-associated tissues // *J. Biol. Chem.* V. 282. № 52. P. 37359–37369.
- Cornelis G., Heidmann O., Degrelle S.A., Vernochet C., Lavialle C. et al., 2013. Captured retroviral envelope syncytin gene associated with the unique placental structure of higher ruminants // *PNAS*. V. 110. № 9. P. E828–E837.
- Cornelis G., Vernochet C., Malicorne S., Souquere S., Tzika A.C. et al., 2014. Retroviral envelope syncytin capture in an ancestrally diverged mammalian clade for placentation in the primitive Afrotherian tenrecs // *PNAS*. V. 111. № 41. P. E4332–E4341.
- Cullen B.R., 2006. Role and mechanism of action of the APOBEC3 family of antiretroviral resistance factors // *J. Virol.* V. 80. P. 1067–1076.
- Damert A., Lower J., Lower R., 2004. Leptin receptor isoform 219.1: An example of protein evolution by LINE-1-mediated human-specific retrotransposition of a coding SVA element // *Mol. Biol. Evol.* V. 21. P. 647–651.
- Darboux I., Charles J.F., Pauchet Y., Warot S., Pauron D., 2007. Transposon-mediated resistance to *Bacillus sphaericus* in a field-evolved population of *Culex pipiens* (Diptera: Culicidae) // *Cell. Microbiol.* V. 9. P. 2022–2029.
- Deng B., Xu W., Wang Z., Liu C., Lin P. et al., 2019. An LTR retrotransposon-derived lncRNA interacts with RNF169 to promote homologous recombination // *EMBO Rep.* V. 20. № 11. P. e47650.
- Dupressoir A., Marceau G., Vernochet C., Bénit L., Kanellopoulos C. et al., 2005. Syncytin-A and syncytin-B, two fusogenic placenta-specific murine envelope genes of retroviral origin conserved in *Muridae* // *PNAS*. V. 102. № 3. P. 725–730.
- Eickbush T.H., Jamburuthugoda V.K., 2008. The diversity of retrotransposons and the properties of their reverse transcriptases // *Virus Res.* V. 134. № 1–2. P. 221–234.
- Emerson R.O., Thomas J.H., 2011. Gypsy and the birth of the SCAN domain // *J. Virol.* V. 85. № 22. P. 12043–12052.
- Farabaugh P.J., 1996. Programmed translational frameshifting // *Annu. Rev. Genet.* V. 30. P. 507–528.
- Frank O., Giehl M., Zheng C., Hehlmann R., Leib-Mösch C., Seifarth W., 2005. Human endogenous retrovirus expression profiles in samples from brains of patients with schizophrenia and bipolar disorders // *J. Virol.* V. 79. № 17. P. 10890–10901.
- Frese S., Ruebner M., Suhr F., Konou T.M., Tappe K.A. et al., 2015. Long-term endurance exercise in humans stimulates cell fusion of myoblasts along with fusogenic endogenous retroviral genes *in vivo* // *PLoS One*. V. 10. № 7. P. e0132099.
- Fujiwara H., Osanai M., Matsumoto T., Kojima K.K., 2005. Telomere-specific non-LTR retrotransposons and telomere maintenance in the silkworm, *Bombyx mori* // *Chromosome Res.* V. 13. № 5. P. 455–467.
- Goodier J.L., 2016. Restricting retrotransposons: A review // *Mob. DNA*. V. 7. P. 16.
- Graveley B.R., May G., Brooks A.N., Carlson J.W., Chervas L. et al., 2011. The *D. melanogaster* transcriptome: modENCODE RNA-Seq data for differing treatment conditions. <http://www.modencode.org/celniker/>
- Grow E.J., Flynn R.A., Chavez S.L., Bayless N.L., Wossidlo M. et al., 2015. Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells // *Nature*. V. 522. P. 221–225.
- Heidmann O., Vernochet C., Dupressoir A., Heidmann T., 2009. Identification of an endogenous retroviral envelope gene with fusogenic activity and placenta-specific expression in the rabbit: A new “syncytin” in a third order of mammals // *Retrovirology*. V. 6. P. 107.
- Heidmann O., Béguin A., Paternina J., Berthier R., Deloger M. et al., 2017. HEMO, an ancestral endogenous retroviral envelope protein shed in the blood of pregnant women and expressed in pluripotent stem cells and tumors // *Proc. Natl Acad. Sci. USA*. V. 114. P. E6642–E6651.
- Henke C., Strissel P.L., Schubert M.-T., Mitchell M., Stolt C.C. et al., 2015. Selective expression of sense and antisense transcripts of the sushi-ichi-related retrotransposon-derived family during mouse placentogenesis // *Retrovirology*. V. 12. P. 9.
- Henzy J.E., Gifford R.J., Kenaley C.P., Johnson W.E., 2017. An intact retroviral gene conserved in Spiny-rayed fishes for over 100 My // *Mol. Biol. Evol.* V. 34. P. 634–639.
- Horie M., Tomonaga K., 2011. Non-retroviral fossils in vertebrate genomes // *Viruses*. V. 3. № 10. P. 1836–1848.
- Hornung V., 2014. SnapShot: Nucleic acid immune sensors // *Immunity*. V. 41. № 5. P. 868.e1.
- Hu T., Pi W., Zhu X., Yu M., Ha H. et al., 2017. Long non-coding RNAs transcribed by ERV-9 LTR retrotransposon act in cis to modulate long-range LTR enhancer function // *Nucleic Acids Res.* V. 45. № 8. P. 4479–4492.
- Huang S., Tao X., Yuan S., Zhang Y., Li P. et al., 2016. Discovery of an active RAG transposon illuminates the origins of V(D)J recombination // *Cell*. V. 166. № 1. P. 102–114.
- Igolkina A.A., Zinkevich A., Karandasheva K.O., Popov A.A., Selifanova M.V. et al., 2019. H3K4me3, H3K9ac, H3K27ac, H3K27me3 and H3K9me3 histone tags suggest distinct regulatory evolution of open and condensed chromatin landmarks // *Cells*. V. 8. № 9. P. 1034.
- Jo A., Lee H.E., Kim H.S., 2019. Identification and expression analysis of a novel miRNA derived from ERV-E1 LTR in *Equus caballus* // *Gene*. V. 687. P. 238–245.
- Juneja P., Quinn A., Jiggins F.M., 2016. Latitudinal clines in gene expression and cis-regulatory element variation in *Drosophila melanogaster* // *BMC Genomics*. V. 17. № 1. P. 981.
- Kaer K., Speek M., 2013. Retroelements in human disease // *Gene*. V. 518. № 2. P. 231–241.
- Kaessmann H., Vinckenbosch N., Long M., 2009. RNA-based gene duplication: Mechanistic and evolutionary insights // *Nat. Rev. Genet.* V. 10. P. 19–31.
- Kapusta A., Kronenberg Z., Lynch V.J., Zhuo X., Ramsay L.A. et al., 2013. Transposable elements are major contribu-

- tors to the origin, diversification, and regulation of vertebrate long noncoding RNAs // *PLoS Genet.* V. 9. P. e1003470.
- Kazazian H.H., 2014. Processed pseudogene insertions in somatic cells // *Mob. DNA.* V. 5. P. 20.
- Kelkar Y.D., Tyekucheva S., Chiaromonte F., Makova K.D., 2008. The genome-wide determinants of human and chimpanzee microsatellite evolution // *Genome Res.* V. 18. P. 30–38.
- Kemp C., Mueller S., Goto A., Barbier V., Paro S. et al., 2013. Affiliations expand Broad RNA interference-mediated antiviral immunity and virus-specific inducible responses in *Drosophila* // *J. Immunol.* V. 190. № 2. P. 650–658.
- Kim A., Terzian C., Santamaria P., Péliesson A., Purd'homme N., Bucheton A., 1994. Retroviruses in invertebrates: The gypsy retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster* // *PNAS.* V. 91. P. 1285–1289.
- Kokosar J., Kordis D., 2013. Genesis and regulatory wiring of retroelement-derived domesticated genes: A phylogenomic perspective // *Mol. Biol. Evol.* V. 30. № 5. P. 1015–1031.
- Krupovic M., Blomberg J., Coffin J.M., Dasgupta I., Fan H. et al., 2018. Ortervirales: New virus order unifying five families of reverse-transcribing viruses // *J. Virol.* V. 92. № 12. P. e00515–18.
- Lee J.Y., Ji Z., Tian B., 2008. Phylogenetic analysis of mRNA polyadenylation sites reveals. A role of transposable elements in evolution of the 3'-end of genes // *Nucleic Acids Res.* V. 6. P. 5581–5590.
- Li T.H., Schmid C.W., 2001. Differential stress induction of individual Alu loci: Implications for transcription and retrotransposition // *Gene.* V. 276. P. 135–141.
- MacMillan H.A., Knee J.M., Dennis A.B., Udaka H., Marshall K.E. et al., 2016. Cold acclimation wholly reorganizes the *Drosophila melanogaster* transcriptome and metabolome // *Sci. Rep.* V. 6. P. 28999.
- Makhnovskii P., Balakireva Y., Nefedova L., Lavrenov A., Kuzmin I., Kim A., 2020. Domesticated gag gene of *Drosophila* LTR retrotransposons is involved in response to oxidative stress // *Genes.* V. 11. № 4. P. 396.
- Malfavon-Borja R., Feschotte C., 2015. Fighting fire with fire: Endogenous retrovirus envelopes as restriction factors // *J. Virol.* V. 89. № 8. P. 4047–4050.
- Malik H.S., Henikoff S., Eickbush T.H., 2000. Poised for contagion: Evolutionary origins of the infectious abilities of invertebrate retroviruses // *Genome Res.* V. 10. № 9. P. 1307–1318.
- Mallet F., Bouton O., Prudhomme S., Cheynet V., Oriol G. et al., 2004. The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology // *PNAS.* V. 101. № 6. P. 1731–1736.
- Matsui T., Leung D., Miyashita H., Maksakova I.A., Miyachi H. et al., 2010. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET // *Nature.* V. 464. № 7290. P. 927–931.
- Meyerson N.R., Sawyer S.L., 2011. Two-stepping through time: Mammals and viruses // *Trends Microbiol.* V. 19. P. 286–294.
- Miller W.J., McDonald J.F., Pinsker W., 1997. Molecular domestication of mobile elements // *Genetica.* V. 100. № 1–3. P. 261–270.
- Monde K., Contreras-Galindo R., Kaplan M.H., Markovitz D.M., Ono A., 2012. Human endogenous retrovirus K Gag co-assembles with HIV-1 Gag and reduces the release efficiency and infectivity of HIV-1 // *J. Virol.* V. 86. P. 11194–11208.
- Mouches C., Pauplin Y., Agarwal M., Lemieux L., Herzog M. et al., 1990. Characterization of amplification core and esterase B1 gene responsible for insecticide resistance in *Culex* // *Proc. Natl. Acad. Sci. USA.* V. 87. P. 2574–2578.
- Nair S., Rein A., 2014. Antiretroviral restriction factors in mice // *Virus Res.* V. 193. P. 130–134.
- Nefedova L.N., Kuzmin I.V., Makhnovskii P.A., Kim A.I., 2014. Domesticated retroviral GAG gene in *Drosophila*: New functions for an old gene // *Virology.* V. 450–451. P. 196–204.
- Nekrutenko A., Li W.H., 2001. Transposable elements are found in a large number of human protein-coding genes // *Trends Genet.* V. 17. P. 619–621.
- Nikitin D., Kolosov N., Murzina A., Pats K., Zamyatin A. et al., 2019. Retroelement-linked H3K4me1 histone tags uncover regulatory evolution trends of gene enhancers and feature quickly evolving molecular processes in human physiology // *Cells.* V. 8. № 10. P. 1219.
- Ono R., Kuroki Y., Naruse M., Ishii M., Iwasaki S. et al., 2011. Identification of tammar wallaby SIRH12, derived from a marsupial-specific retrotransposition event // *DNA Res.* V. 18. № 4. P. 211–219.
- Pickeral O.K., Makałowski W., Boguski M.S., Boeke J.D., 2000. Frequent human genomic DNA transduction driven by LINE-1 retrotransposition // *Genome Res.* V. 10. № 4. P. 411–415.
- Polak P., Domany E., 2006. Alu elements contain many binding sites for transcription factors and may play a role in regulation of developmental processes // *BMC Genomics.* V. 7. P. 133.
- Robertson D.S., 2016. Origin and activities of human lentivirus particles // *Biomed. Pharmacother.* V. 83. P. 1311–1314.
- Rowe H.M., Jakobsson J., Mesnard D., Rougemont J., Reynard S. et al., 2010. KAP1 controls endogenous retroviruses in embryonic stem cells // *Nature.* V. 463. № 7278. P. 237–240.
- Roy-Engel A.M., El-Sawy M., Farooq L., Odom G.L., Perepelitsa-Belancio V. et al., 2005. Human retroelements may introduce intragenic polyadenylation signals // *Cytogenet. Genome Res.* V. 110. P. 365–371.
- Rubin C.M., VandeVoort C.A., Teplitz R.L., Schmid C.W., 1994. Alu repeated DNAs are differentially methylated in primate germ cells // *Nucleic Acids Res.* V. 22. P. 5121–5127.
- Samuel C.E., 2011. Adenosine deaminases acting on RNA (ADARs) are both antiviral and proviral // *Virology.* V. 411. P. 180–193.
- Santangelo A.M., Souza F.S., de Franchini L.F., Bumashny V.F., Low M.J., Rubinstein M., 2007. Ancient exaptation of a CORE-SINE retroposon into a highly conserved mammalian neuronal enhancer of the proopiomelanocortin gene // *PLoS Genet.* V. 3. P. 1813–1826.

- Schrader L., Schmitz J.*, 2019. The impact of transposable elements in adaptive evolution // *Mol. Ecol.* V. 28. № 6. P. 1537–1549.
- Sela N., Mersch B., Hotz-Wagenblatt A., Ast G.*, 2010. Characteristics of transposable element exonization within human and mouse // *PLoS One*. V. 5. P. e10907.
- Sela N., Mersch B., Gal-Mark N., Lev-Maor G., Hotz-Wagenblatt A., Ast G.*, 2007. Comparative analysis of transposed element insertion within human and mouse genomes reveals Alu's unique role in shaping the human transcriptome // *Genome Biol.* V. 8. № 6. P. R127.
- Silva-Sousa R., López-Panadés E., Casacuberta E.*, 2012. *Drosophila* telomeres: An example of co-evolution with transposable elements // *Genome Dyn.* V. 7. P. 46–67.
- Soe K., Andersen T.L., Hobolt-Pedersen A.S., Bjerregaard B., Larsson L.I., Delaissé J.M.*, 2011. Involvement of human endogenous retroviral syncytin-1 in human osteoclast fusion // *Bone*. V. 48. № 4. P. 837–846.
- Sugimoto J., Sugimoto M., Bernstein H., Jinno Y., Schust D.*, 2013. A novel human endogenous retroviral protein inhibits cell-cell fusion // *Sci. Rep.* V. 3. P. 1462.
- Takeda A., Matano T.*, 2007. Inhibition of infectious murine leukemia virus production by Fv-4 *env* gene products exerting dominant negative effect on viral envelope glycoprotein // *Microbes Infect.* V. 9. № 14–15. P. 1590–1596.
- Tanaka Y., Chung L., Park I.H.*, 2012. Impact of retrotransposons in pluripotent stem cells // *Mol. Cells*. V. 34. № 6. P. 509–516.
- Terry S.N., Manganaro L., Cuesta-Dominguez A., Brinzevich D., Simon V., Mulder L.C.*, 2017. Expression of HERV-K108 envelope interferes with HIV-1 production // *Virology*. V. 509. P. 52–59.
- Trizzino M., Kapusta A., Brown C.D.*, 2018. Transposable elements generate regulatory novelty in a tissue-specific fashion // *BMC Genomics*. V. 19. P. 468.
- Varela M., Spencer T.E., Palmarini M., Arnaud F.*, 2009. Friendly viruses: The special relationship between endogenous retroviruses and their host // *Ann. N. Y. Acad. Sci.* V. 1178. P. 157–172.
- Vela D., Fontdevila A., Vieira C., García Guerreiro M.P.*, 2014. A genome-wide survey of genetic instability by transposition in *Drosophila* hybrids // *PLoS One*. V. 9. № 2. P. e88992.
- Wolf D., Goff S.P.*, 2009. Embryonic stem cells use ZFP809 to silence retroviral DNAs // *Nature*. V. 458. № 7242. P. 1201–1204.
- Xing J., Hedges D.J., Han K., Wang H., Cordaux R., Batzer M.A.*, 2004. Alu element mutation spectra: Molecular clocks and the effect of DNA methylation // *J. Mol. Biol.* V. 344. P. 675–682.
- Yap M.W., Colbeck E., Ellis S.A., Stoye J.P.*, 2014. Evolution of the retroviral restriction gene *FvI*: Inhibition of non-MLV retroviruses // *PLoS Pathog.* V. 10. № 3. P. e1003968.
- Yoshinaka Y., Katoh I., Copeland T.D., Oroszlan S.*, 1985. Murine leukemia virus protease is encoded by the gene and is synthesized through suppression of an amber termination codon // *Proc. Natl. Acad. Sci. USA*. V. 82. P. 1618–1622.
- Young G.R., Terry S.N., Manganaro L., Cuesta-Dominguez A., Deikus G. et al.*, 2018. Mulder HIV-1 infection of primary CD4+ T cells regulates the expression of specific human endogenous retrovirus HERV-K (HML-2) elements // *J. Virol.* V. 92. P. e01507-17.

## The role of retroelements in the evolution of animal genomes

L. N. Nefedova<sup>a</sup>, \*, A. I. Kim<sup>a</sup>

<sup>a</sup>*Lomonosov Moscow State University, Faculty of Biology, Department of Genetics  
Leninskiye Gory, 1/12, Moscow, 119234 Russia*

*\*e-mail: lidia\_nefedova@mail.ru*

Retroelements (retrotransposons and endogenous retroviruses) are a class of mobile genetic elements that is widely represented in various groups of eukaryotes. It is believed that most of the retroelements present in the genome do not significantly affect the host organism, which, in the process of co-evolution with retroelements, produces various mechanisms for controlling their transposition. But even retroelements deprived of transpositional activity can play a significant role in evolution, contributing to the formation of new adaptive mechanisms, which, in turn, are caused by the formation of new genetic networks. The process in which the body adapts individual sequences of retroelements is called molecular domestication. Exaptation implies the formation of new genetic structures and functions, which, as a rule, appear as a result of the long evolution of domesticated sequences of mobile elements. For three key genes of retrotransposons/endogenous retroviruses (*gag*, *pol*, and *env*), cases of domestication are known. The phenomenon of domestication/exaptation testifies to the fundamental role that retroelements play in the evolution of the genome and the significant evolutionary benefit of the presence of retroelements in the genome despite the fact that their transposition activity can be accompanied by negative effects at the level of an individual organism.