ОНТОГЕНЕЗ, 2018, том 49, № 4 ПРИЛОЖЕНИЕ, с. S3–S20

ОБЗОРЫ

УДК 575.829

ФУНКЦИОНАЛЬНЫЙ ДУАЛИЗМ ТРАНСКРИПТОВ ТРАНСПОЗОНОВ В ЭВОЛЮЦИИ ЭУКАРИОТИЧЕСКИХ ГЕНОМОВ

© 2018 г. Р. Н. Мустафин[a, b, *]

[a]*Федеральное государственное бюджетное образовательное учреждение высшего образования "Башкирский государственный медицинский университет", Россия 450008, г. Уфа, ул. Ленина, 3*

[b]*Федеральное государственное бюджетное образовательное учреждение высшего образования "Башкирский государственный университет", Россия 450076, г. Уфа, ул. Заки Валиди, 32*

******E-mail: ruji79@mail.ru*

Поступила в редакцию 11.08.2018 г.
Окончательный вариант получен 30.10.2018 г.

Способность транспозонов объединять разобщенные их инсерциями гены, кодирующие единые биологические процессы, в регуляторные сети, способствовало одновременно с усложнением эукариот их эволюционному успеху путем формирования новых универсальных систем. При помощи данных систем, включающих метилирование ДНК, модификации гистонов, взаимосвязь теломер с транспозонами, регуляцию сплайсинга и РНК-интерференцию, глобальное распространение транспозонов в геномах сопровождалось возникновением их структурных инноваций, динамичных регуляторных последовательностей и белок-кодирующих генов. Мобильные элементы способствовали эволюции белок-кодирующих генов путем их дупликации, а также экзонизации и одомашнивания самих транспозонов. Образуемые при этом новые гены содержат транспозонные последовательности, участвующие в их управлении при помощи регуляторных сетей и некодирующих РНК, также происходящих от мобильных элементов. В ходе эволюции сформировалась стратегия, когда трансляция генов некодирующих РНК способствовала отбору полученных полипептидов в качестве функциональных клеточных белков. При этом некодирующие РНК процессируются также в молекулы, участвующие в регуляторных процессах самостоятельно или в составе белковых комплексов. Двойственность функций оказалась присущей всем некодирующим РНК, неслучайный распад/процессинг которых приводит к образованию молекул, оказывающих регуляторное воздействие на транспозоны и белок-кодирующие гены. В эволюции эукариот сформировалась стратегия, когда первичные транскрипты транспозонов взаимодействуют с разными системами их процессинга (возникших для защиты хозяев от транспозонов), образуя функциональные молекулы РНК, транслируемые в пептиды. Универсальными источниками для данных стратегий служат транспозоны, что объясняет их глобальное распространение в геномах эукариот и одомашнивание в системе "двойного поиска" мишеней для функционального взаимодействия некодирующих РНК и процессированных продуктов их трансляции. Первичные транскрипты некоторых белок-кодирующих генов, помимо сплайсинга, также могут процессироваться в функциональные некодирующие РНК, участвующие в единых биологических реакциях с белковым продуктом гена. Это обосновывает ассоциации мультифакториальных заболеваний с SNP генов, так как они могут вызвать инактивацию доменов РНК. Сделано предположение, что функциональный дуализм транскриптов транспозонов мог быть важным условием возникновения жизни, а мобильные элементы являются одним из фундаментальных свойств живого.

Ключевые слова: белок-кодирующие гены, гистоны, некодирующие РНК, метилирование, пептиды, транспозоны

DOI: 10.1134/S0475145018070017

СПИСОК СОКРАЩЕНИЙ:

БКГ – белок-кодирующие гены.

мяРНК – малые ядерные РНК (класс малых нкРНК, которые обнаруживаются в ядрах эукариот и принимают участие в регуляции сплайсинга, факторов транскрипции и в поддержании целостности теломер).

мяоРНК – малые ядрышковые РНК (класс малых нкРНК, которые участвуют в метилировании и псевдоуридилировании транспортных РНК, малых ядерных РНК и рибосомных РНК).

нкРНК – некодирующие РНК.

НП – нуклеотидные последовательности.

рРНК – рибосомные РНК.

ТФ – транскрипционные факторы (белки, которые связываются со специфическими последовательностями ДНК и контролирующие транскрипцию мРНК на матрице ДНК).

тРНК – транспортные РНК.

ERV – эндогенные ретровирусы (endogenous retroviruses).

IGS – межгенные спейсеры (intergenic spacer).

LINE – длинные диспергированные элементы (Long Interspersed Nuclear Elements).

lncPHK – длинные некодирующие РНК (long noncoding RNA).

LTR – длинные концевые повторы (Long Terminal Repeats).

RT – обратная транскриптаза (Reverse Transcriptase).

sdPHK – малые РНК, произошедшие от малых ядрышковых РНК (sno-derived RNA).

SINE – корокие диспергированные ядерные элементы (Short Interspersed Nuclear Elements).

SNP – однонуклеотидные полиморфизмы (Single Nucleotide Polymorphism).

SR – саркоплазматический ретикулум (sarcoplasmic reticulum).

TE – транспозоны (Transposable Elements).

tRF – фрагменты РНК, вырезанные из тРНК (tRNA derived RNA fragments).

tsRNA – малые РНК, образуемые из тРНК (tRNA-derived small RNA).

UTR – нетранслируемая область (Untranslated Region).

ВВЕДЕНИЕ

Транспозоны (TE – transposable elements) – это генетические объекты, способные изменять свою локализацию в геноме (Startek et al., 2017). Они отличаются по структуре и механизму перемещения и подразделяются на классы, подклассы (для ДНК-транспозонов), порядки, надсемейства, семейства, подсемейства. Класс I – ретротранспозоны (ретроТЕ) – перемещаются по механизму “копирование – вставка”. Класс II – ДНК транспозоны (ДНК-ТЕ) – по механизму “вырезание–вставка” (порядки TIR и Crypton) и “катящегося кольца” (порядки Helitron и Maverick). Подразделение на порядки проводится на основании общей организации элемента, особенностей ферментов перемещения и механизмов встраивания в геном (рис. 1) (Wicker et al., 2007). Самая распространенная классификация ТЕ была реализована в Repbase (http://www.girinst.org/repbase/). Транспозиция ТЕ класса I осуществляется через промежуточное звено РНК, когда исходная последовательность остается на прежнем месте (Startek et al., 2017). Все ретроТЕ имеют два общих гена: *gag* и *pol*. Гены *gag* кодируют белки GAG (Group-specific AntiGen), которые формируют вирусоподобные частицы VLPs (Virus-Like Particles). Гены *pol* кодируют ферменты обратную транскриптазу (RT – reverse transcriptase), интегразу (IN) и протеазу (PR) (Alzohairy et al., 2013). Оба класса ТЕ встречаются у прокариот и эукариот, однако у последних ретроэлементы оказались наиболее распространены, занимая у разных представителей от 3 до 85% нуклеотидных последовательностей (НП) генома (Startek et al., 2017). ТЕ содержатся даже в геномах гигантских вирусов, участвуя в их функционировании (Filee, 2018). Было показано, что в составе пандоравирусов ТЕ способствуют их эволюционным преобразованиям (Zhang et al., 2018).

Обмен генами вспомогательного набора при помощи горизонтального переноса является важным условием выживания бактерий и архей. Поэтому у прокариот ТЕ также играют ключевую роль в эволюции, способствуя возникновению новых генов за счет перехода вспомогательных генов в базовые (Равин, Шестаков, 2013) при помощи ретроэлементов. К примеру, ТЕ бактерий обычно переносят гены устойчивости к антибиотикам (Babakhani, Oloomi, 2018). Еще в 1989 году у *E. coli* и *M. xanthus* были обнаружены ретроны, а в 2002 году – у рода *Bordetella* – ретроэлементы DGR. Последние, несмотря на кодирование обратной транскриптазы, не проявляют мобильности, но обеспечивают защиту от бактериофагов. Помимо этого, в геномах бактерий обнаружены CRISPR (clastered regularly interspaced short palindromic repeats) и Abi, также кодирующие RT и защищающие их от чужеродных последовательностей (Toro, Nisa-Martinez, 2014). Более 50% RT бактерий кодируются интронами группы II, которые подразделяют на несколько основных групп: A, B, C, D, E, F, G, CL1/2 (chloroplast-like) и ML (mitochondrion-like) (Toro et al., 2018). Классификация ТЕ, плазмид и фагов прокариот представлена в базе данных ACLAME (http://aclame.ulb.ac.be) (Leplae et al., 2004). Однако несмотря на наличие ТЕ и сложных систем защиты, геномы прокариот значительно проще эукариотических и содержат гораздо меньшее количество ТЕ, что отражается в их размерах. Причиной усложнений в структуре и функционировании геномов эукариот является возникновение ряда консервативных механизмов, способствующих использованию новых инсерций ТЕ для изменчивости и адаптации в эволюции. В изменчивости прокариот большее значение имеет горизонтальный обмен генами, который также может носить глобальный характер, подобно эндогенному распространению ТЕ в геномах эукариот. Например, две энтеробактерии *Escherichia coli* и *Salmonella* разделяют около 70% их генов – аналогичный уровень дивергенции наблюдается между геномами двух штаммов

Класс 1 МГЭ – ретротранспозоны

Порядок	Супер-семейство	Структурная организация
LTR	*Copia*	GAG AP INT RT RH
	Gypsy	GAG AP RT RH INT
	Bel-Pao	GAG AP RT RH INT
	Retrovirus	GAG AP RT RH INT ENV
	ERV	GAG AP RT RH INT ENV
DIRS	*DIRS*	GAG AP RT RH YR
	Ngaro	GAG AP RT RH YR
	VIPER	GAG AP RT RH YR
PLE	*Penelope*	RT EN
LINE	*R2*	RT EN
	RTE	APE RT
	Jockey	ORF1 APE RT
	LI	ORF1 APE RT
	I	ORF1 APE RT RH
SINE	*tRNA*	
	7SL	
	5S	

Класс II МГЭ – ДНК транспозоны

Порядок	Супер-семейство	Структурная организация
TIR	*Tc1/mairner*	ТРАНСПОЗАЗА
	hAT	
	Mutator	
	Merlin	
	Transib	
	P	
	PiggyBac	
	PIF-Harbinger	ТРАНСПОЗАЗА ORF2
	CACTA	ТРАНСПОЗАЗА ORF2
Crypton	*Crypton*	YR
Helitron	*Helitron*	RPA Y2HEL
Maverick	*Maverick*	C-INT ATP CYP POL B

Рис. 1. Классификация транспозонов по Wicker. LTR – long terminal repeats (длинные концевые повторы). GAG – капсидный белок, AP – aspartic proteinase (аспартатная протеиназа), INT – integrase (интеграза), RT – reverse transcriptase (обратная транскриптаза), RH – RNase H (РНКаза-Н), TIR – terminal inverted repeats (инвертированные концевые повторы), ENV – envelope protein (оболочечный белок), YR – тирозиновая рекомбиназа, EN – endonucleases (эндонуклеаза), APE – apurinic endonucleases (апуриновая эндонуклеаза), ORF – open reading frame (открытая рамка считывания), RPA – replication protein A (репликативный белок А), HEL – helicase (хеликаза), ATP – упаковывающая АТФаза, CYP – cysteine protease (цистеиновая протеаза), POLB – DNA polymerase B (ДНК-полимераза В).

E. coli (лабораторным штаммом K12 и патогенном O157-H7, которые отличаются на 30% их геномов). При этом большинство различий объясняются профагами. Аналогичным образом, основные отличия между геномами *Listeria monocytogenes* и *Listeria innocua* соответсвуют профагам, интегрированным в последний (Leplae et al., 2004).

Поскольку активность ТЕ, подобно другим формам мутаций, может быть вредной с точки зрения фенотипических эффектов, понимание широкого распространения транспозонов в геномах эукариот нуждается в анализе смысла эволюционного накопления ТЕ. В данном отношении имеет значение инактивация ТЕ из-за накопления ими мутаций, сайленсинг ТЕ и их доместикация хозяевами, а также горизонтальный перенос ТЕ (Song, Schaack, 2018). Однако ТЕ являются ключевыми компонентами адаптивной эволюции эукариот (Barry, 2018). Несмотря на развитие и совершенствование механизмов борьбы с эгоистическими последовательностями, в геномах эукариот распространение ТЕ характеризуется глобальными масштабами, а системы защиты от ТЕ способствовали возникновению кардинальных отличий от прокариотических геномов. Данные отличия включают нуклеосомную упаковку ДНК с возможностью ее управления при помощи модификации гистонов и метилирования ДНК, линейную структуру хромосом с теломерами на концах, глобальное распространение интронов и ретрогенов, использование некодирующих РНК (нкРНК) в управлении всеми биологическими процессами. Эукариоты произошли от архей, многие из которых обитают в агрессивных условиях (Henikoff, Smith, 2015; Шестаков, 2017), способствующих отбору организмов с формированием максимально

возможного количества биомолекул, необходимых для появления новых адаптационных свойств для выживания. Это могло стать причиной распространения ТЕ и формирования универсальных систем их использования хозяином. Способность ТЕ изменять локализацию внутри генома приводит к возникновению новых структурных и функциональных активностей, которые способствуют эволюции геномов и видообразованию (Alzohairy et al., 2013).

У эукариот эндогенное распространение ТЕ в эволюции привело к использованию их НП хозяевами в глобальном масштабе. ТЕ стали ключевыми источниками нкРНК (Borchert et al., 2011; Li et al., 2011; Yuan et al., 2011; Roberts et al., 2013; Gim et al., 2014; Johnson, Guigo, 2014; Kapusta, Feschotte, 2014; Lorenzetti et al., 2016; Cho, 2018), белок-кодирующих генов (БКГ), в том числе транскрипционных факторов (ТФ) (Zdobnov et al., 2005; Volff, 2006; Feschotte, 2008; Sinzelle et al., 2009; Alzohairy et al., 2013; Garavis et al., 2013; Duan et al., 2017; Schrader, Schmitz, 2018), а также сайтов связывания с ТФ и других регуляторных структур (de Souza et al., 2013; Jaques et al., 2013; Mak et al., 2014; Ito et al., 2017). Кроме того, ТЕ могут служить источниками сателлитных тандемных массивов в геномах (McGurk, Barbash, 2018). При этом ТЕ изменяются, благодаря накоплению мутаций, до неузнаваемости, и становятся неотъемлемыми консервативными структурами и генами. Предполагается, что ТЕ составляют значительно большую часть геномов эукариот, чем идентифицированные стандартными методами исследования (Goerner-Potvin, Bourque, 2018). Например, принято считать, что ТЕ занимают около 45% всей ДНК человека, однако анализ при помощи специфических олигонуклеотидов, узнающих фрагменты ТЕ, показал, что их доля значительно выше, и занимает более 60% генома (de Koning et al., 2011). В настоящее время активно разрабатываются методы идентификации “скрытых” ТЕ в геномах различных видов для более точного определения их роли в формировании регуляторных НП и БКГ (Goerner-Potvin, Bourque, 2018). Даже с учетом несовершенства методов обнаружения, согласно современным данным, у эукариот в эволюции ТЕ стали источниками значительной части НП геномов, расположенных в межгенных областях и в интронах БКГ. Например, в геноме дрозофилы ТЕ составляют до 22, мыши – 40, растений – до 90% (Yuan et al., 2011). Благодаря глобальному распространению ТЕ, размеры геномов эукариот различаются в 70000 раз от 2.3 (*Saccharomyces cerevisiae*) до 150000 млн. пар нуклеотидов (*Paris japonica*). Для сравнения, минимальный размер генома прокариот у *Micoplasma genitalium* меньше максимального у *Myxococcus xanthus* всего в 16 раз (Патрушев, Минкевич, 2007). Несмотря на то, что в геномах эукариот доля НП, кодирующих белки, составляет от доли до нескольких процентов всей ДНК, ТЕ оказались важнейшими источниками их возникновения и эволюции путем дупликации существующих генов (Huang et al., 2009; Sakai et al., 2011; Grandi et al., 2015; Tan et al., 2016; Zhu et al., 2016; Kubiak, Makalowska, 2017; Cerbin, Jiang, 2018), экзонизации (Wang et al., 2012; Elkon et al., 2013; Tajnik et al., 2015) и одомашнивания самих ТЕ (Zdobnov et al., 2005; Volff et al., 2006; Feschotte, 2008; Sinzelle et al., 2009; Alzohairy et al., 2013; Garavis et al., 2013; Duan et al., 2017; Schrader, Schmitz, 2018).

ВОЗНИКНОВЕНИЕ НОВЫХ БЕЛКОВ ПРИ ПОМОЩИ ТРАНСПОЗОНОВ

ТЕ способствуют дупликации генов хозяев путем инсерции процессированных РНК-продуктов генов с помощью обратной транскрипатзы (Cerbin, Jiang, 2018). При этом образованные ретрокопии фланкированы НП транспозонов, что способствует их управлению регуляторными сетями с участием ТЕ (Schrader, Schmitz, 2018). Например, у млекопитающих ретрокопии содержат во фланкирующих областях сайт расщепления эндонуклеазы TTTT/AA и поли(А) хвост, специфичные для LINE-1 (Long Interspersed Nuclear Elements). У дрозофилы ретрогены могут быть фланкированы длинными концевыми повторами (LTR – long terminal repeats) (Tan et al., 2016). У растений, включая арабидопсис, рис и многие другие виды, также идентифицировано множество ретрогенов, часть из которых обладают функциональностью (например, ген *Sun* у томатов, гены *CYP98A8* и *CYP98A9* у арабидопсиса). Ретрокопии растений фланкированы главным образом LTR, так как их образованию способствуют в основном LTR-содержащие ТЕ, хотя L1-подобные ретроэлементы также участвуют в данных процессах (Zhu et al., 2016). Безинтронные ретрогены могут приобретать новые интроны, сохранение которых также говорит о включении ретрокопий в функционирование регуляторных сетей (как в качестве источников продуктов транскрипции, так и трансляции). За счет этого ретрогены могут даже функционально заместить родительский ген или развить новую функцию (Kubiak, Makalowska, 2017). Таким образом, ТЕ способствуют белковому разнообразию в эволюции за счет ретротранспозиции существующих БКГ, сохранение которых способствует обретению генами взаимосвязей в формирующихся новых регуляторных сетях в эволюции. Предполагается, что количество ретрогенов значительно больше определяемых стандартными методами идентификации в связи с 5'-усечением, хотя даже с этим учетом в среднем около 10% всех БКГ содержат одну и более ретрокопий (Grandi et al., 2015). В геноме человека, к примеру, содержится более 1000 ретрогенов, которые транскрибируются и проявля-

ют биологическую активность (Huang et al., 2009). У риса определена экспрессия не менее 66% выявленных ретрокопий, хотя и с более низкими значениями, чем их исходные гены, но с тканеспецифической корреляцией их уровней (Sakai et al., 2011).

Распространенным способом возникновения новых БКГ в эволюции оказалось также одомашнивание генов ТЕ (Joly-Lopez, Bureau, 2018; Schrader, Schmitz, 2018). Самым ярким примером является теломераза, консервативная для большинства эукариот и произошедшая от RT ретроэлементов (Garavis et al., 2013). Теломераза обладает значительными сходствами с RT L1 элементов (Kopera et al., 2011), а в клеточной линии яичников китайского хомячка с нарушенной функцией теломер, L1 преимущественно транспозируют в области теломер, помогая сохранить их функцию (Morrish et al., 2007). Потеря гена теломеразы у дрозофилы в эволюции была успешно замещена ретро-ТЕ HeT-A, TART и TAHRE (Casacuberta, 2017). У животных, растений и дрожжей было обнаружено множество генов, происходящих от ТЕ, особенно от транспозазы (Feschotte, 2008). У позвоночных выявлено 1000 генов, произошедших от ретро-ТЕ (Zdobnov et al., 2005). При этом вновь образованные БКГ могут кодировать ТФ за счет использования ДНК-связывающих доменов ТЕ и осуществляют важнейшие функции в управлении работы геномом. В то же время, многие сайты связывания с ТФ также происходят от НП транспозонов, что говорит о глобальной роли ТЕ в формировании генных регуляторных сетей (Feschotte, 2008; de Souza et al., 2013; Jaques et al., 2013; Mak et al., 2014; Ito et al., 2017). Способность эгоистических элементов к саморегуляции, сохраняющаяся в эволюции геномов их хозяев, говорит о существовании стратегии “ТЕ–нкРНК–пептиды” (рис. 2), в которой транскрипты транспозонов отбирались на функциональную пригодность параллельно с их белковыми продуктами, так как нкРНК способны транслироваться в пептиды (Zhang et al., 2013; Anderson et al., 2015; Lauressergues et al., 2015; Couzigou et al., 2016; Lv et al., 2016; Nelson et al., 2016), а ТЕ служат важнейшими источниками нкРНК (Borchert et al., 2011; Li et al., 2011; Yuan et al., 2011; Roberts et al., 2013; Gim et al., 2014; Johnson, Guigo, 2014; Kapusta, Feschotte, 2014; Lorenzetti et al., 2016). Кроме того, БКГ, произошедшие от ТЕ, выполняют важнейшие клеточные функции (Sinzelle et al., 2009). Помимо одомашнивания для регуляции работы генома, ТЕ стали источниками белков, участвующих в ремоделировании хроматина и клеточных функциях (Joly-Lopez, Bureau, 2018). Это указывает на ключевую роль ТЕ в формировании отличительных особенностей эукариотических хромосом. Например, из транспозазы образованы связывающий хроматин с ядерным матриксом инсулятор BEAF-32 и центромерсвязывающие белки Abp1 у дрожжей (Feschotte, 2008), а также белки ацетилтрансферазного комплекса гистонов HDP1/2 у арабидопсиса (Duan et al., 2017). У человека выявлены инсуляторы, произошедшие от ТЕ семейства MIR (mammalian-wide interspersed repetitive) элементов, которые вербуют ферменты модификации хроматина для изменения регуляторных генных сетей (Wang et al., 2015). У растений гены *far1*, *fhy3* и *frs* произошли путем одомашнивания ДНК-ТЕ (Alzohairy et al., 2013). Множество новых БКГ в эволюции произошли при доместикации различных генов ТЕ. Из транспозазы образовались гены RAG у позвоночных, *Daysleeper* у дрозофилы, *Tram, Buster1-3*, *Zbed4*, *P52rlPK* у млекопитающих, *Cenp-B* у всех эукариот, *Metnase* и *Pgbd* у человека и мыши, гены *harbi1* у рыб, лягушек и млекопитающих, *Daysleeper* у арабидопсиса, *Fob1p* у дрожжей. От интегразы произошли гены *Gin-1*, *Fob1p*. От *Gag*-генов возникли гены *Mart*, *Ma*, *Fv1*, *PEG10, Rtl1, MyEF-3* у млекопитающих. От генов оболочки ретровирусов произошли гены *Syncytin-1*, *-2*, *-A*, *-B* (Volff, 2006; Alzohairy et al., 2013). Белковые продукты одомашненных генов ТЕ управляют структурой хроматина (белки Cenp-B, BEAF-32, HIM-17), участвуют в апоптозе (THAP0, THAP1, E93), контролируют клеточный цикл (семейство белков THAP, LIN-36, LIN-15B) (Sinzelle et al., 2008). Крупномасштабный структурный анализ протеома ТЕ позволил предсказать структуры сотен белков из репрезентативного набора ДНК-ТЕ и LINE для общей структурной характеристики белков, произошедших в эволюции от ТЕ. Белки ORF1 и Gag ретроэлементов содержат большое количество структурных изменений, и, несмотря на их низкую консервативность, сохраняют их в эволюции. Белки ДНК-ТЕ оказались наиболее древними и содержат складки, которые существовали уже тогда, когда появились первые клеточные организмы (Abrusan et al., 2013). Согласно недавним исследованиям, последний универсальный общий предок клеточной жизни появился около 3.9 млрд лет назад. Эубактерии и археи появились около 3.4 млрд лет назад, а современные эукариоты возникли значительно позже – менее чем 1.84 млрд лет назад (Betts et al., 2018). Белки ДНК-ТЕ имеют более низкий порядок контактов, чем случайно выбранные контрольные белки, что позволяет им быстро складываться и избегать агрегации (Abrusan et al., 2013).

Инсерции ТЕ в интроны и межгенные области способствуют экзонизации и альтернативному сплайсингу, что играет важную роль в эволюции белков и обретению ими новых функциональных доменов. В данном отношении значительную активность проявляют SINE (Short Interspersed Nuclear Elements) и LINE, содержащие латентные сигналы сплайсинга, что объясняет их распространенность в интронах. Например, в геноме человека до 5% альтернативных экзонов и 3'-не-

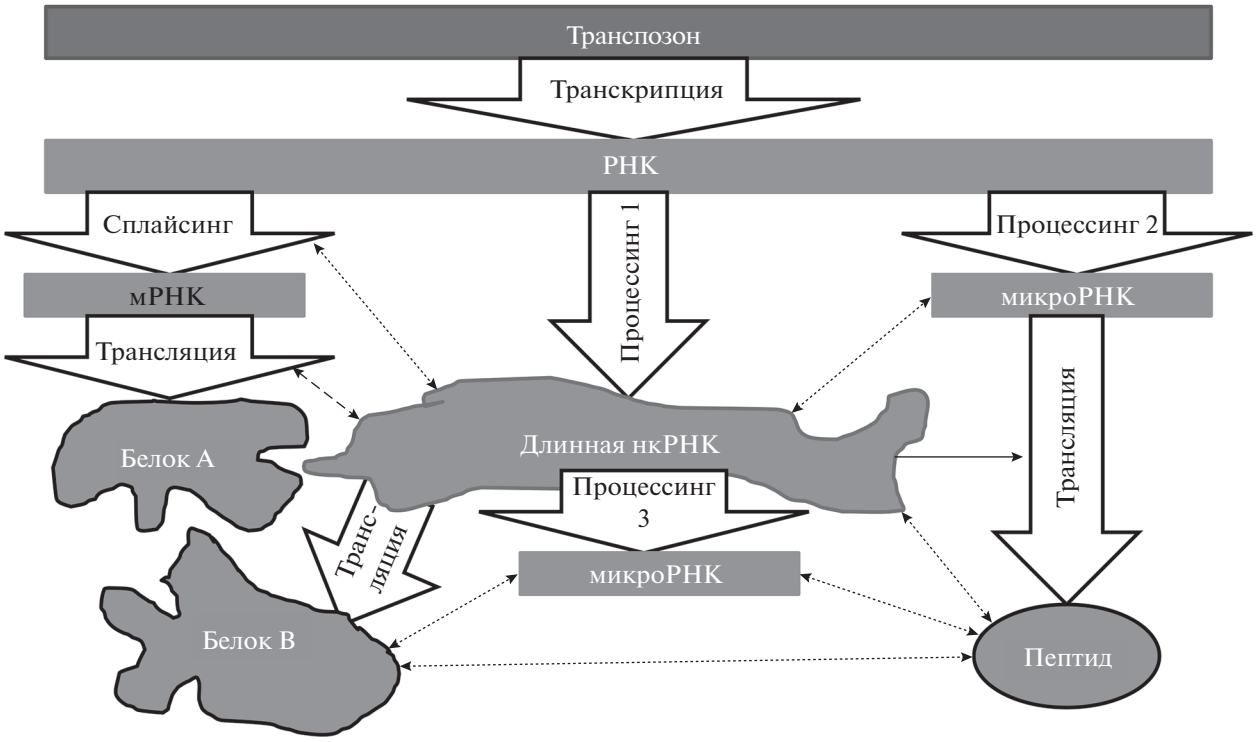

Рис. 2. Схема “двойного поиска” мишеней в стратегии “транспозоны–некодирующие РНК–(поли)пептиды” в возникновении новых белок-кодирующих генов в эволюции.

транслируемых областей (UTR – untranslated region) возникли путем экзонизации межгенных *Alu*-элементов (Tajnik et al., 2015). Эпигенетическая регуляция внутригенных ТЕ, инсертированных в интроны, часто способствует изменению генной регуляции, фенотипической экспрессии и геномной эволюции (Saze, 2018). У человека более 50% генов характеризуются наличием альтернативных 3'UTR, способствующих их ткане- и стадие-специфической регуляции (Elkon et al., 2013). Содержание НП транспозонов в 3'UTR генов имеет эволюционное значение для отбора оптимальных регуляторных сетей управления экспрессией генов, так как ТЕ служат источниками нкРНК, эпигенетическое воздействие которых распространяется на гены, в 3'UTR (и в интронах) которых содержатся ТЕ или их остатки. ТЕ также способствуют возникновению новых интронов с использованием их НП для альтернативного сплайсинга и экзонизации (Wang et al., 2012). Предполагается, что сплайсосомные интроны, характерные для эукариот, произошли от интронов группы II, которые обладают способностью к ретротранспозиции (Novikova, Belfort, 2017) и кодируют собственную RT (Toro et al., 2018). При этом глобальное распространение интронов происходило главным образом путем транспозиций, инсерций ТЕ, репарации двойных разрывов ДНК (в которой также участвуют ТЕ) (Yenerall, Zhou, 2012), что объясняется наличием в НП ТЕ акцепторных и донорных сайтов сплайсинга (Belancio et al., 2010). При этом в эволюции ТЕ оказались источниками основного компонента сплайсосомы, Prp8, который образовался из RT с утраченной каталитической активностью (Arkhipova, 2018). Роль ТЕ в возникновении универсальных систем регуляции геномов объясняет сохранение в эволюции феномена мозаичной структуры генов, консервативной для эукариот. Подтверждением служат данные о наличии взаимосвязей множества нкРНК с интронами с отчетливым синергетическим эффектом. Более того, после сплайсинга часть интронов сами способны процессироваться в функциональные молекулы нкРНК (Rearick et al., 2011), что говорит о важной роли ТЕ в обеспечении функционального дуализма транскриптов генов, необходимого для возникновения максимального количество регуляторных взаимосвязей как источника динамической и пластичной изменчивости и адаптации в естественном отборе, а также служит одним из объяснений распространения и сохранения интронов в эволюции.

ВОЗНИКНОВЕНИЕ РЕГУЛЯТОРНОЙ СЕТИ ПРИ ПОМОЩИ ТРАНСПОЗОНОВ

Сохранение в эволюции новых генов, содержащих последовательности ТЕ, неизбежно приводит к тому, что экспрессия нового белка управляется

новой регуляторной сетью с большим количеством взаимосвязей и контролем эпигенетическими факторами. То есть возникает двойная система контроля (рис. 3). С одной стороны, ТЕ в эволюции служат источниками как ТФ, так и сайтов связывания с данными ТФ (Feschotte, 2008; Ito et al., 2017). С другой стороны, возникающие из ТЕ новые белки содержат участки, связывающиеся с нкРНК транспозонного происхождения, так как ТЕ являются важными источниками siРНК, микроРНК и длинных некодирующих РНК (lncРНК – long noncoding RNA) (Borchert et al., 2011; Li et al., 2011; Yuan et al., 2011; Roberts et al., 2013; Gim et al., 2014; Johnson, Guigo, 2014; Kapusta, Feschotte, 2014; Lorenzetti et al., 2016). За счет этого создаются сложные регуляторные взаимосвязи между эпигенетическими факторами и ТФ, что обеспечивает координированные последовательные этапы управления экспрессией генов при клеточных делениях. Это служит ключом к регуляции дифференцировки клеток в онтогенезе с одновременной динамичностью данного процесса в связи с присущей ТЕ мобильностью и чувствительностью к стрессу (Feng et al., 2013; Wheeler, 2013; Мустафин, Хуснутдинова, 2018), обуславливая возможность эволюции эукариот. В данных регуляторных системах принимают также важное участие сплайсинговая машина, что объясняет глобальную распространенность у эукариот сплайсосомных интронов, которые взаимодействуют со множеством различных микроРНК, snoРНК, siРНК, piРНК и lncРНК при управлении экспрессией генов (рис. 4) (Rearick et al., 2011). Так, у дрозофилы piРНК регулируют сплайсинг пре-мРНК транспозонов как в соматических, так и в половых клетках (Barry, 2018). Несмотря на происхождение от интронов группы II (Novikova, Belfort, 2017), распространение в геномах эукариот получили только сплайсосомные интроны. Это говорит об их важной роли в регуляции экспрессии генов, а взаимосвязь с нкРНК и ТЕ говорит о существовании консервативной для эукариот системы содействия между ТЕ, сплайсингом, РНК-интерференцией (РНКи), гистонами и теломерами. При этом в ходе эволюции интроны могут теряться и распространяться главным образом путем транспозиций и при помощи ТЕ (Yenerall, Zhou, 2012). Участие ТЕ в распространении интронов предполагает наличие их НП в интронах, что способствует управлению экспрессией генов не только на уровне сплайсосомы, но и транскрипционно (за счет воздействия нкРНК транспозонного происхождения). Помимо этого, ТЕ в интронах способствуют образованию альтернативных сплайсинговых вариантов (Feschotte, 2008).

Глобальное распространение ТЕ в геномах эукариот способствует использованию их НП в генной регуляции в качестве сайтов связывания с ТФ, промоторов и энхансеров (de Souza et al., 2013; Sahebi et al., 2018). ТЕ могут изменять экспрессию не только соседних, но также не связанных с инсерциями генов, изменять модификацию хроматина, служить источниками *цис*-действующих регуляторных элементов, а также могут инсертировать в существующие энхансеры, которые регулируют транскрипцию (Sahebi et al., 2018). При этом ТЕ ассоциированы с видоспецифическими изменениями генной экспрессии у высших позвоночных, что говорит об их роли в видообразовании (Zeng et al., 2018) и тканеспецифической регуляции работы геномов (Trizzino et al., 2018). Выявлена выраженная корреляция между активностью ТЕ и видообразованием (Ricci et al., 2018), а также роль ТЕ в возникновении генетических дефектов у гибридов, препятствующих скрещиванию видов (Serrato-Capuchina, Matute, 2018). В отличие от изменений в самих генах, настройка регуляторных сетей экспрессии генов играет более значимую роль в морфологической эволюции. Например, при исследовании сайтов связывания с ТФ у представителей 29 различных видов млекопитающих было выявлено, что фрагменты последовательностей ТЕ используются хозяевами в качестве регуляторных структур в глобальных масштабах (Lowe, Haussler, 2012). При сравнительном анализе 29 видов млекопитающих идентифицировано 280 000 произошедших от ТЕ регуляторных структур, в том числе обладающих свойствами энхансеров (de Souza et al., 2013). В геноме человека из MIR, помимо инсуляторов (Wang et al., 2015), образованы также энхансеры, играющие важную роль в *цис*-регуляции экспрессии БКГ (Jjingo et al., 2014). Для многоклеточных большинство тканеспецифически активируемых регуляторов генов имеют происхождение от ТЕ, что говорит о важнейшей роли транспозонов в эволюции эукариот. Интересно, что в раннем эмбриональном развитии из ретро-ТЕ инициируется до 20% транскриптома, а многочисленные перемещения ТЕ в ходе эволюции вызвали перераспределение сайтов связывания с ТФ, выполняющих тканеспецифическую регуляцию (Mak et al., 2014). Количество активных регуляторных элементов транспозонного происхождения, управляющих экспрессией генов хозяина, имеет колоссальные масштабы. Например, в геноме человека оказалось около 800000 сайтов связывания с ТФ, которые произошли только от LTR-содержащих ТЕ. Многие из этих регуляторных элементов характеризуются тканеспецифическим характером активации (Ito et al., 2017). Около 15 млн лет назад семейство ретровирусов, породившее LTR5HS, распространилось в зародышевой линии гоминид, создавая сходные НП в геноме человека. В результате при помощи всего 12 специфических РНК удалось идентифицировать около 90% данных элементов. Несмотря на то, что LTR5HS являются эволюционно молодыми ТЕ в геноме человека, была отмечена их спо-

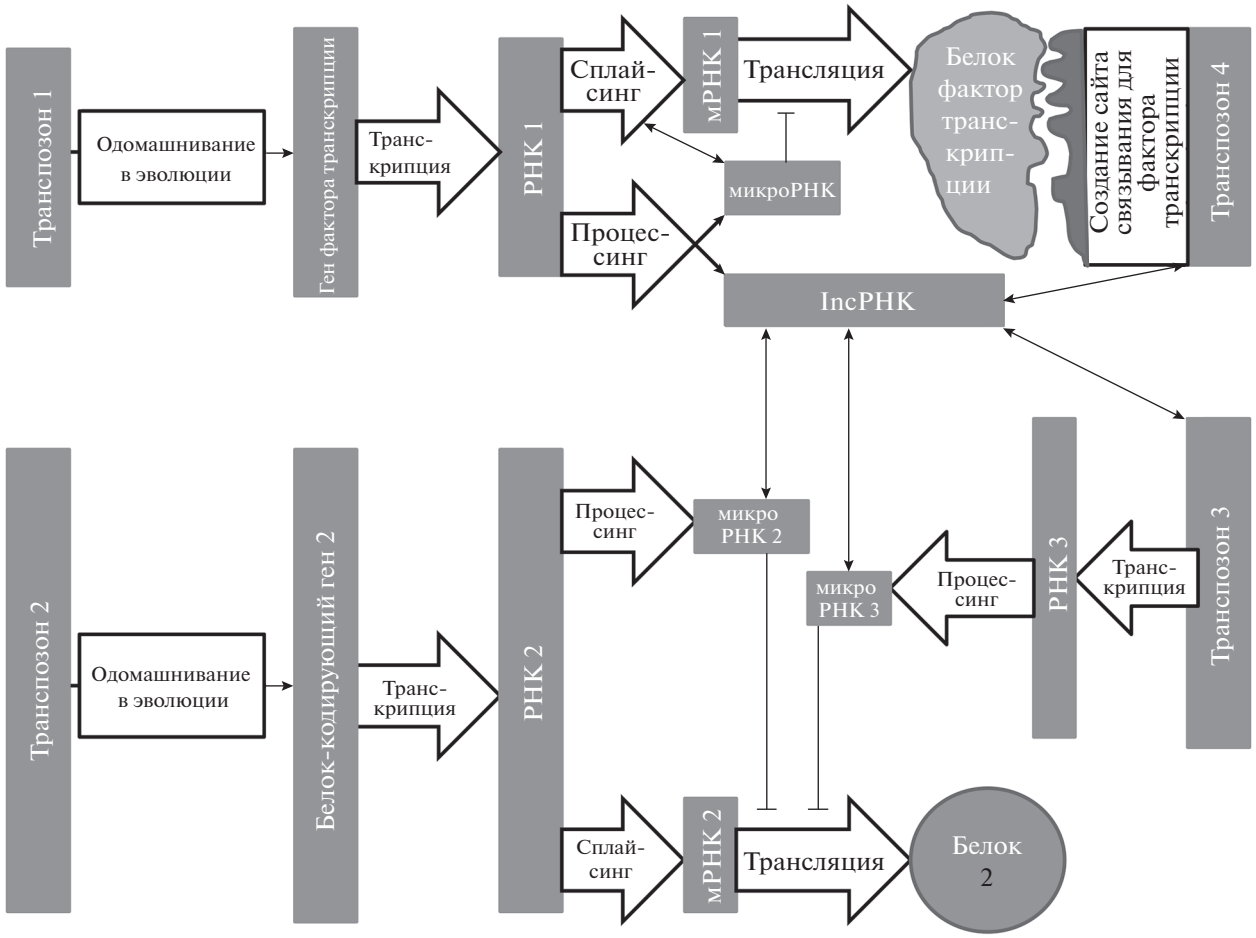

Рис. 3. Роль транспозонов в создании регуляторных сетей геномов с двойной системой контроля экспрессии генов – при помощи некодирующих РНК (нкРНК) и транскрипционных факторов.

собность функционировать в качестве энхансеров для контроля работы генов (Judd, Feschotte, 2018). При этом параллельно с использованием одомашненных генов ТЕ в онтогенезе, происходила кооптация НП транспозонов для регуляторных взаимосвязей данных генов (Мустафин, Хуснутдинова, 2018). Примером может служить доместикация гена *env* ретровирусов (Dupressoir et al., 2012), способствующего формированию плаценты, в то время как экспансия транспозонов RLTR13D5 и MER20 была необходима для экспрессии генов в эндометрии, взаимодействующих с продуктом данного гена (Chuong et al., 2013). Оказалось, что около половины функциональных регуляторных последовательностей в геноме человека имеют транспозонное происхождение (Jacques et al., 2013). Выявлены сотни тысяч энхансеров различных видов животных и растений, произошедших от ТЕ, что говорит об универсальности принципа кооптации транспозонных регуляторных последовательностей геномами эукариот (de Souza et al., 2013). Глобальные масштабы кооптации транспозонных НП как для регуляции экспрессии генов в онтогенезе, так и для возникновения новых БКГ можно оценить на примере взаимосвязи ТЕ с теломерами и гистонами, консервативность которых для всех эукариот говорит о ключевом значении ТЕ как универсальных систем моделирования геномов в эволюции.

ВЗАИМОСВЯЗЬ ТРАНСПОЗОНОВ С ТЕЛОМЕРАМИ И ГИСТОНАМИ

В естественном отборе формирование теломер на концах хромосом произошло благодаря ТЕ, так как теломераза в эволюции произошла от RT ретроэлементов (Garavis et al., 2013). Данное свойство является характерным отличием эукариот от бактерий и архей. Хотя некоторые представители бактерий могут обладать линейной структурой ДНК, теломер на концах хромосом они не содержат, а для большинства из них характерны кольцевые нуклеоиды (Равин, Шестаков, 2013). У дрозофилы в эволюции была потеряна теломераза, и

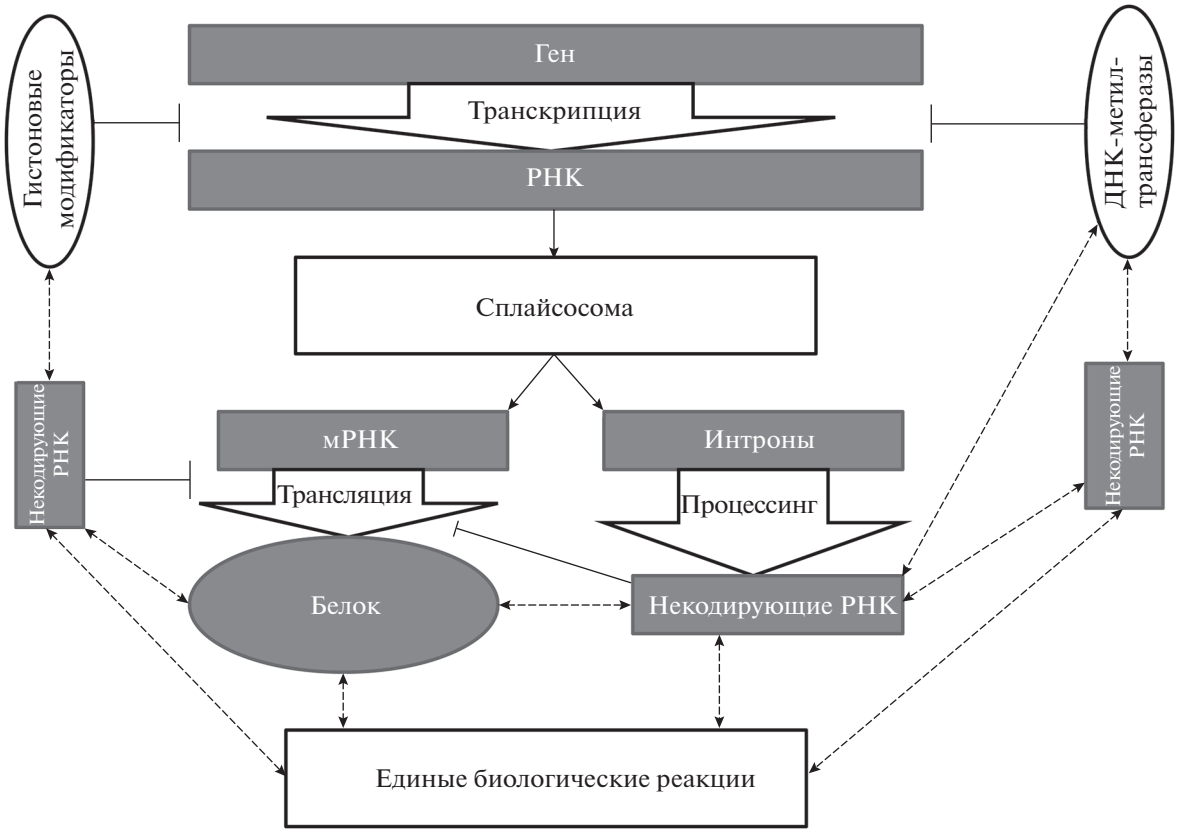

Рис. 4. Взаимосвязь сплайсинга с другими регуляторными системами эукариот.

для формирования теломер используются ретроэлементы HeT-A, TART и TAHRE (Casacuberta, 2017). В неоплазмах с мутациями в гене *ATRX*, продукт которого ремоделирует нуклеосомы за счет взаимодействия с H3 хвостом, обнаружено выраженное увеличение размера теломер без участия теломеразы (Henikoff, Smith, 2015). Кроме того, Atrx связывается с участками ретроэлементов, называемых IAP (intracisternal A particle) и способствует их сайленсингу, способствуя формированию гетерохроматина (Sadic et al., 2015). Роль белка, ремоделирующего нуклеосомы, в альтернативном удлинении теломер и регуляции ТЕ, говорит о наличии взаимосвязей между ТЕ, теломерами и гистонами, позволяющих сохранившимся в эволюции герминативным инсерциям ТЕ изменять структуру и количество хромосом. Действительно, теломерные последовательности обнаруживаются в интерстициальных частях хромосом, возникающих за счет слияния теломер при редукции числа хромосом (Бадаева, Салина, 2013). Роль ремоделирования хроматина в данных участках говорит о сложных взаимосвязях с ТЕ, которые стали источниками теломер в эволюции (Garavis et al., 2013).

У бактерий и архей в упаковке ДНК участвует белок HU, у некоторых архей и всех эукариот – гистоны. Предполагается, что октамерные нуклеосомы эукариот эволюционировали от более простых тетрамерных нуклеосом архей. Но эукариотические гистоны, в отличие от археальных, содержат хвосты, модификация которых способствует эпигенетическому управлению экспрессией генов (Henikoff, Smith, 2015). Хотя, наиболее вероятно, первоначально данный механизм возник как система защиты для замалчивания внедренных в геном ТЕ. Это говорит о важной роли ТЕ в эволюции гистонов. Можно предположить, что возникновение октамерного нуклеосомного комплекса и консервативной системы гистоновых белков у эукариот происходило параллельно с адаптацией геномов к многочисленным инсерциям ТЕ и созданием других универсальных систем, таких как РНКи, сплайсинговая машина и теломеры. Отбор наиболее оптимальных наборов гистонов на пригодность мог происходить при участии ТЕ, на что указывают некоторые особенности. Например, гены гистонов располагаются в виде тандемных повторов, а их мРНК не содержат полиаденилированных участков, заканчиваясь специфическими вторичными структурами “стебель–петля”. Эти структуры взаимодействуют с белком SLBP для дальнейших этапов метаболизма (Marzluff, Koreski, 2017). Пространственная организация

мРНК гистонов проявляет функциональность наподобие lncРНК, домены которых построены главным образом из последовательностей ТЕ (Kapusta, Feschotte, 2014). То есть продукты транскрипции генов гистонов проявляют функциональный дуализм – пространственные домены РНК способны взаимодействовать с белками, одновременно со способностью к трансляции. Данное свойство позволяет предположить роль ТЕ в эволюции генов гистонов, так как НП транспозонов играют ключевую роль в формировании вторичных и третичных структур молекул РНК, способных участвовать в различных биологических реакциях (Johnson, Gujigo 2014), а функциональный дуализм транскриптов ТЕ является их характерной чертой. Это обусловлено воздействием защитных систем хозяина, осуществляющих процессинг РНК транспозонов и выработанными в эволюции свойствами полученных продуктов процессинга к функциональности в интересах хозяина для обретения им адаптивных свойств за счет участия нкРНК в различных биологических реакциях либо самостоятельно (в качестве рибозимов или рибосвитчей), либо в составе РНП или RISC.

Не исключено, что функциональные домены РНК гистонов построены при участии ТЕ, что указывает на их тесную взаимосвязь в эволюции, так как ТЕ способствовали возникновению важнейших клеточных белков путем экзонизации, дупликации и одомашнивания, создавая основу для появления линий с новыми адаптационными свойствами. Примером служит возникновение теломеразы из обратной транскриптазы ретроэлементов (Garavis et al., 2013) и центромерсвязывающих белков Cenp-B от ТЕ *pogo* всех эукариот (Volff, 2006), одомашнивание гена *env* оболочки ретровирусов геномами млекопитающих с использованием его для создания объединенного слоя плаценты на поверхности матки. Данный ген, названный синцитином, играет важнейшую роль в эмбриональном развитии млекопитающих (Sinzelle et al., 2009). При этом помимо кооптации для создания новых белков, ТЕ активно использовались в эволюции в качестве источников регуляторных последовательностей (рис. 5). Интересно, что диапазон размеров консервативных и неконсервативных участков интронов БКГ имеет максимумы, близкие к длине нуклеосомной и динуклеосомной ДНК. Это говорит о том, что длина интронов определяется функциональной нагрузкой (Виноградов, 2011). Кроме того, это служит подтверждением роли ТЕ в формировании консервативного комплекса гистонов в виде нуклеосом путем “поиска” возможных вариантов в эволюции с участием ТЕ, так как интроны обязаны своим происхождением мобильным элементам (Yenerall, Zhou, 2012; Novikova, Belfort, 2017). Данный пример характеризует универсальность принципа “двойного поиска” структурно-функциональных взаимосвязей последовательностей нуклеиновых кислот с пространственной конфигурацией доменов вторичных и третичных структур ДНК, РНК и белков. То есть в ходе естественного отбора эукариот взаимодействие хозяев с ТЕ происходило благодаря универсальной закономерности саморегуляции транспозонов, способствующей эволюции их хозяев. ТЕ служат как источниками регуляторных последовательностей, так и взаимодействующих с ними одомашненных белков. ТЕ распространяются в геномах эукариот, но при этом участвуют в эволюции систем, направленных на борьбу с ТЕ (эволюция нуклеосом). Универсальность саморегуляции продуктами собственной экспрессии способствовала эволюции структуры и функции геномов. Например, центромеры возникли благодаря ТЕ и теломерам (Garavis et al., 2013; Klein, O’Neill, 2018; McGurk, Barbash, 2018), а центромерсвязывающие белки Cenp-B эукариот произошли от ТЕ *pogo* в эволюции (Volff, 2006). Кроме того, ТЕ служат важными источниками микроРНК (Borchert et al., 2011; Li et al., 2011; Yuan et al., 2011; Roberts et al., 2013; Gim et al., 2014; Lorenzetti et al., 2016), которые вызывают сайленсинг собственных ТЕ и других генов, содержащих гомологичные НП, а продукты трансляции при-микроРНК стимулируют экспрессию собственных микроРНК (Couzigou et al., 2015; Lauressergues et al., 2015; Couzigou et al., 2016; Lv et al., 2016).

СТРАТЕГИЯ ТРАНСПОЗОНЫ – НКРНК – ПЕПТИДЫ В ВОЗНИКНОВЕНИИ НОВЫХ ГЕНОВ

В эволюции эукариот ТЕ оказались не только источниками БКГ и регуляторных последовательностей, но и нкРНК, включая siРНК, микроРНК (Borchert et al., 2011; Li et al., 2011; Yuan et al., 2011; Roberts et al., 2013; Gim et al., 2014; Lorenzetti et al., 2016; Cho, 2018) и lncРНК (Johnson, Guigo, 2014; Kapusta, Feschotte, 2014). При этом большая часть геномов эукариот образована из ТЕ последовательностей, значительно превышая идентифицируемую стандартными методами определения долю (de Koning et al., 2011; Goerner-Potvin, Bourque, 2018). Так как большинство нкРНК транскрибируются из межгенных и интронных областей, где расположены ТЕ, можно предположить, что транспозоны являются основными источниками нкРНК. Действительно, в различных работах были выявлены тысячи микроРНК, происходящих от ТЕ у животных (Borchert et al., 2011; Yuan et al., 2011; Roberts et al., 2013; Gim et al., 2014). У растений большинство микроРНК гомологичны ТЕ, что указывает на их происхождение от ТЕ (Li et al., 2011; Lorenzetti et al., 2016; Cho, 2018). ТЕ оказывают важное влияние на эволюцию благодаря их роли в возникновении lncРНК. LncРНК являются регуля-

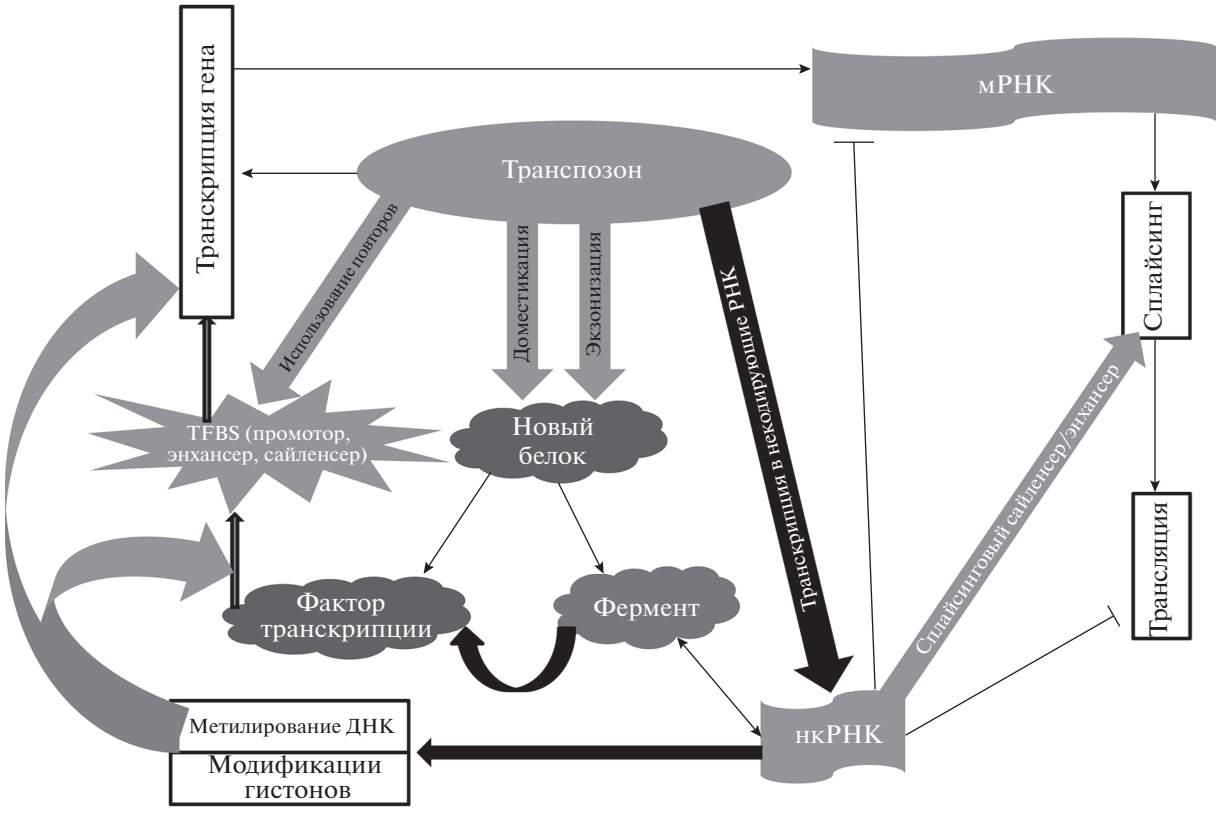

Рис. 5. Взаимосвязь транспозонов, некодирующих РНК, транскрипционных факторов и регуляторных структур геномов эукариот.

торами транскрипции за счет взаимодействия с ТФ и другими ДНК-связывающими белками, РНК-полимеразой II и гистон-модифицирующими комплексами. LncРНК могут выполнять свои функции как самостоятельные активные молекулы в качестве рибосвитчей или рибозимов, а также в составе РНП (Long et al., 2017). Вторичные структуры молекул lncРНК, подобно белкам, обладают модульной организацией и содержат дискретные домены, которые создаются главным образом благодаря ТЕ. Около 83% доменов lncРНК содержат не менее 1 фрагмента ТЕ (Johnson, Guigo, 2014), а многие зрелые lncРНК полностью состоят из ТЕ (Kapusta, Feschotte, 2014). LINE1 ретротранспозоны обладают lncРНК-подобной функцией, участвуя в регуляции генной экспрессии для управления самообновлением эмбриональных стволовых клеток и преимплантационного развития (Honson, Macfarlan, 2018). Более того, ТЕ могут использоваться непосредственно в качестве генов lncРНК (Lu et al., 2012), что говорит о первичном свойстве ТЕ к функциональному дуализму их транскриптов. То есть транскрипты транспозонов способны процессироваться как функциональные молекулы РНК, обладающие доменной структурой, так и во взаимодействующие с рибосомами РНК, продукты трансляции которых участвуют в различных биологических реакциях. Данная стратегия “ТЕ–нкРНК–(поли)пептиды” служит в качестве источника для “поиска” в эволюции оптимальных путей регуляции генных сетей и может служить универсальным свойством построения динамических биологических систем. Более того, удивительным открытием оказалось то, что lncРНК также способны транслироваться в пептиды (Zhang et al., 2013; Anderson et al., 2015; Lv et al., 2016; Nelson et al., 2016). Еще удивительнее то, что, помимо трансляции в пептиды, lncРНК способны к дальнейшему процессингу с образованием микроРНК (Guo et al., 2014; Lv et al., 2018). В то же время обнаружено, что при-микроРНК также способны транслироваться с образованием функциональных пептидов (Couzigou et al., 2015; Lauressergues et al., 2015; Couzigou et al., 2016; Lv et al., 2016). То есть регуляторные системы геномов эукариот оказываются намного сложнее сложившихся в генетике представлений. Изучение новых данных позволяет прийти к инновационным выводам о том, что ТЕ могли быть древнейшей универсальной системой, обеспечивающей самовоспроизведение живой материи при возникновении жизни, так как консервативные нкРНК, включающие тРНК (транспортные РНК), рРНК

(рибосомные РНК), мяРНК (малые ядерные РНК) и мяоРНК (малые ядрышковые РНК) способны к неслучайному процессингу (подобно транскриптам ТЕ) с использованием РНК-продуктов для интерференции транспозонных последовательностей (Ender et al., 2008; Jacob et al., 2012; Li et al., 2012; Kumar et al., 2014; Venkatesh et al., 2016; Martinez et al., 2017).

Сама по себе функциональность одновременно как молекул РНК, так и их белковых продуктов может свидетельствовать о происхождении от ТЕ, что важно для определения эволюционной истории БКГ. Действительно, филогенетические анализы еще в 2006 позволили выявить образование новых БКГ за счет отбора различных вариантов экспрессии lncРНК (Levine et al., 2006). В 2008 году та же закономерность в возникновении новых БКГ обнаружена на *S. cerevisiae* (Cai et al., 2008). В 2012 г. было идентифицировано 24 различных БКГ, возникших из lncРНК путем их отбора при "поиске" вариантов молекулярных и биохимических взаимодействий (Xie et al., 2012). В 2014 г. на основании изучения шести видов животных (*Mus musculus*, *Danio rerio*, *D. melanogaster*, *A. thaliana*, *S. cerevisiae*, *Homo sapiens*) было доказано, что возникновение БКГ *de novo* из нкРНК является распространенным явлением и служит источником инноваций в эволюции. Оказалось, что большинство lncРНК всех шести видов взаимодействуют с рибосомами и проявляют сходный потенциал кодирования с эволюционного молодыми БКГ (Ruiz-Orera et al., 2014). Кроме того, у животных достоверно подтверждена трансляция lncРНК в пептиды, которые, подобно БКГ, участвуют в разнообразных биологических реакциях. Например, у мыши выявлена специфическая для скелетных мышц экспрессия lncРНК, которая транслируется в пептид миорегулин (MLN). MLN проявляет структурное и функциональное сходство с фосфаламбаном и сарколипином, которые контролируют мышечную релаксацию за счет регуляции поглощения ионов кальция в саркоплазматический ретикулум (SR) (Anderson et al., 2015). В SR обнаружено функционирование другого пептида, также образуемого путем трансляции lncРНК. Данный пептид назван DWORF (dwarf open reading frame). Он усиливает активность SERCA (sarco/endoplasmic reticulum Ca^{2+}–ATPase), вытесняя его ингибиторы фосфоламбан, сарколипин и MLN в кардиомиоцитах и скелетных мышцах (Nelson et al., 2016). У растений пептиды, образуемые путем трансляции lncРНК, участвуют в различных биологических процессах: IPS1 регулирует поглощение фосфатов, ENOD40 участвует в симбиотических взаимосвязях с бактериями, COLDAIR и COOLAIR управляют временем цветения (Zhang et al., 2013).

Оказалось, что помимо lncРНК, у эукариот транслироваться в функциональные пептиды способны также при-микроРНК. Образуемые при этом miPEP оказывают регуляторное влияние на БКГ (Lv et al., 2016). При-микроРНК, одновременно со способностью к процессингу в зрелые микроРНК, содержат короткие открытые рамки считывания (smORF) с трансляцией в miPEP, которые способны усиливать транскрипцию связанных с ними микроРНК (Couzigou et al., 2015). Функциональная активность miPEP доказана, а некоторые из них (miPEP172c) уже успешно применяются в сельском хозяйстве с целью оптимизации агрономических свойств растений. Например, miPEP172c, за счет усиления экспрессии miR172c, способствует образованию узелков корневой системы сои, что повышает урожайность путем стимуляции симбиоза с ризобиями (Couzigou et al., 2016). У арабидопсиса выявлен miPEP165a, а не менее 1 из 50 при-микроРНК содержат smORF, необходимую для трансляции при-микроРНК в пептид. У *Medicago truncatula* обнаружен pri-miR171b (Lauressergues et al., 2015). Не исключено наличие подобных свойств у при-микроРНК животных и грибов, что является предметом для дальнейших исследований.

Интересно, что, помимо БКГ, произошедших от генов lncРНК (Cai et al., 2008; Levine et al., 2006; Xie et al., 2012; Ruiz-Orera et al., 2014), функциональным дуализмом обладают гены консервативных нкРНК (тРНК, рРНК, мяРНК, мяоРНК) (Li et al., 2012; Kumar et al., 2014; Venkatesh et al., 2016; Martinez et al., 2017). Это позволяет предположить, что в эволюции, у истоков зарождения жизни мобильные элементы могли быть первичными универсальными структурами, способными увеличивать размеры биополимеров благодаря транспозиции и служить источниками для возникновения всего многообразия других функциональных молекул путем транскрипции и трансляции. ТЕ могли участвовать в возникновении и эволюции генов всех известных нкРНК. Показано, что конститутивно экспрессируемые тРНК, рРНК, мяоРНК и мяРНК продуцируют преимущественно малые 5' и 3' концевые фрагменты. Подобно процессингу микроРНК, эти концевые фрагменты генерируются асимметричным способом с преимущественной поддержкой 5' или 3' концов. Данные фрагменты оказались более распространенными, чем микроРНК, и, подобно микроРНК, они процессируются для участия в разнообразных биологических процессах, являясь консервативными для отдаленно связанных видов. Доказано, что произошедшие от мяоРНК фрагменты функционируют в качестве микроРНК длиной от 18 нуклеотидов и более (Li et al., 2012).

Из тРНК образуются 3 типа фрагментов, называемых tRF (tRNA derived RNA fragments): полученные из 3'-конца зрелых тРНК tRF-3s, из крайнего 5'-конца зрелых тРНК tRF-5s и из 3'-фрагмента прекурсора тРНК tRF-1s. У человека tRF-5s и

tRF-3s связываются белками семейства Аргонавт, функционируя подобно микроРНК и вызывая их сайленсинг целевых РНК. Молекулы tRF образуются не случайно, а являются точно сформированными независимо от DROSHA и DICER фрагментами, присутствующими во всех живых организмах от бактерий до человека. Несмотря на то, что tRF более эволюционно консервативны, чем микроРНК, и присутствуют в том же изобилии в клетках, для них до сих пор не существует общепринятой номенклатуры и уникальных идентификаторов (Kumar et al., 2014). Молекулы tRF являются неслучайными продуктами асимметрического процессинга тРНК длиной 18–26 нуклеотидов. Данные нкРНК образуются в большом количестве в инфицированных вирусом или подвергнутых стрессу клетках. Кроме того, tRF накапливаются в большом количестве в мужских гаметах дикого типа, не подверженных стрессу, у цветковых и нецветковых растений. У арабидопсиса, tRF процессируются при помощи Dicer-like1 и соединяются с AGO1, подобно микроРНК. Образованный комплекс tRF-AGO1 воздействует на специфические мишени и расщепляет мРНК транскрипционно активных ТЕ (Martinez et al., 2017). Интересно, что tRF-3s обладают высокой комплементарностью эндогенным ретровирусным (ERV) последовательностям в геноме человека. Предполагается роль tRF-3s в регуляции экспрессии ERV путем РНК-интерференции. Было показано, что тРНК и рРНК подвергаются стресс-индуцированному расщеплению с образованием стабильных РНК-продуктов – данный механизм консервативен от дрожжей до человека. Стресс-индуцированный распад тРНК вызывает образование продуктов длиной от 30 до 50 нуклеотидов, названных sitРНК (stress-induced tRNA-derived RNAs), tiРНК (tRNA-derived stress-induced RNAs) и тРНК-"половинки" (tRNA halves). Обнаружен также еще один класс tsRNA (tRNA-derived small RNAs) – произошедшая от тРНК малая РНК (tRNA-derived small RNA), функционирующая подобно микроРНК. Перечисленные tsРНК процессируется при помощи Dicer или РНКазы-Z, в зависимости от локализации tRF как в зрелой тРНК, так и в ее прекурсоре (Li et al., 2012). Поразительным открытием стало то, что, помимо перечисленных классов фрагментов тРНК, из ранее существовавших тРНК могут образовываться также микроРНК. Например, было выявлено 20 микроРНК, произошедших от тРНК, которые разделяют последовательности с tRF (Venkatesh et al., 2016). Полученные данные о роли неслучайных продуктов расщепления тРНК в сайленсинге ERV (Li et al., 2012) могут свидетельствовать об общих механизмах возникновения ТЕ и генов тРНК на ранних этапах эволюции живого, так как данные структуры отбирались в качестве универсальных систем самовоспроизведения биополимеров. Подтверждением может служить обнаружение функциональных взаимосвязей ТЕ с тРНК. Так, у дрожжей ТЕ Ty3 распознают гены тРНК путем взаимодействия их интегразы с субъединицей TFIIIB РНК-полимеразы-III, участвующей в транскрипции тРНК. У клеточного слизевка *Dictyostelium discoideum* рибонуклеаза, кодируемая ретроТЕ DGLT-A, взаимодействует с TFIIIC субъединицей РНК-полимеразы-III (Kling et al., 2018).

Помимо тРНК, асимметричный процессинг (специфический распад) с 3' или 5' концов мяРНК, мяоРНК и рРНК приводит к образованию фрагментов РНК длиной около 20 нуклеотидов (Li et al., 2012). Выявлено также, что длинные межгенные спейсеры (IGS – intergenic spacer), расположенные между генами рРНК, не являются инертными, а проявляют полифункциональность, необходимую для правильного функционирования клеток. Благодаря индукции различных нкРНК транскриптов из IGS, клетка способна как регулировать синтез рРНК, так и связывать большое количество белков, тем самым модулируя важные молекулярные сети. IGS, в отличие от кодирующих рРНК, характеризуются высоким уровнем изменчивости как по НП, так и по длине, составляя у дрожжей 2.5 kb, у дрозофилы 5.1 kb, у *Xenopus laevis*, кур, мыши и приматов – около 30 kb. При этом изменения НП обусловлены инсерциями большого количества ретроТЕ (Jacob et al., 2012). Были обнаружены также малые РНК, произошедшие из мяоРНК (играющие важную роль в созревании рРНК). Например, из мяоРНК ACA45 процессируется РНК длиной от 20 до 25 нуклеотидов, которая стабильно связывается с белками Ago. Процессинг не зависит от комплекса Drosha/DGCR8, но нуждается в Dicer. Данные малые нкРНК, произошедшие от мяоРНК, функционируют подобно микроРНК с целевым воздействием на специфические мРНК (Ender et al., 2008). Данные микроРНК-подобные РНК, названные sdРНК (sno-derived RNA), с эволюционно консервативными размерами и расположением, происходят от широкого диапазона мяоРНК локусов у животных (человека, мыши, кур, дрозофил), арабидопсиса и делящихся дрожжей. МяоРНК направляют модификацию РНК, локализуясь у эукариот в ядрышках и телах Кахала. С этими структурами связаны компоненты пути РНК. У животных sdРНК длиной 20–24 нуклеотида образуются из 3' конца H/ACA мяоРНК. Из C/D мяоРНК происходят sdРНК с бимодальным размером приблизительно 17–19 и более 27 нуклеотидов преимущественно из 5' конца. SdРНК связываются с AGO7 у арабидопсиса и с Ago1 у делящихся дрожжей с характерными 5' нуклеотидными смещениями. То есть у эукариот существует взаимосвязь между РНК-сайленсингом и опосредованными мяоРНК системами процессинга РНК, а sdРНК представляют собой

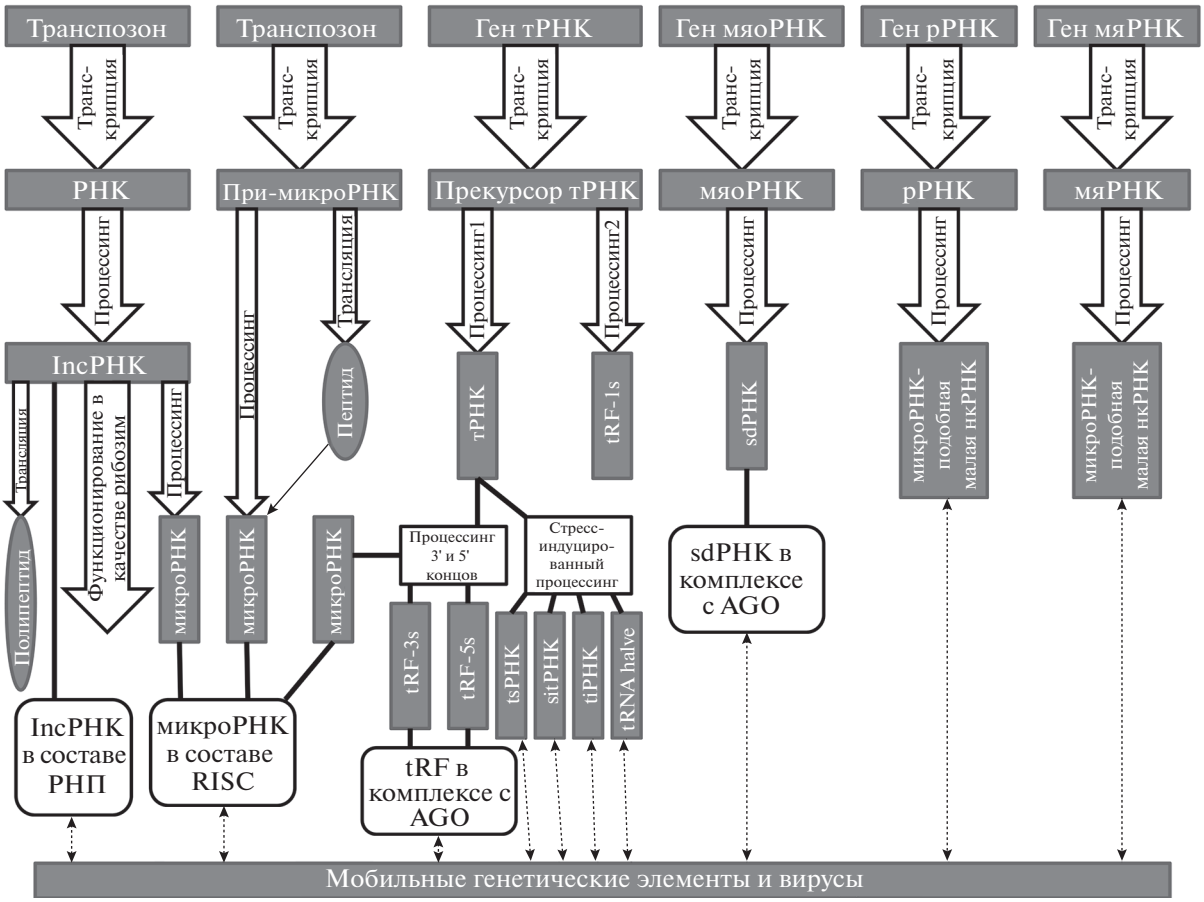

Рис. 6. Функциональный дуализм некодирующих РНК. LncРНК – длинная некодирующая РНК.

древний класс малых нкРНК (Taft et al., 2009). Таким образом, все известные нкРНК характеризуются способностью к процессингу в новые нкРНК, связь которых со специфическими белками позволяет им функционировать для сайленсинга транскриптов ТЕ и БКГ, обеспечивая динамический гомеостаз эукариотических клеток (рис. 6). Функциональный дуализм характерен для транскриптов ТЕ, обусловленный воздействием систем процессинга, возникших для защиты хозяина от распространения ТЕ. Поэтому функциональный дуализм и возможность процессинга всех известных наиболее древних нкРНК с их использованием для РНК-интерференции ТЕ, говорит о роли мобильных элементов в возникновении данных нкРНК в эволюции при возникновении живого. Универсальная способность ТЕ к самовоспроизведению и саморегуляции процессированными продуктами собственной транскрипции позволяет предположить, что ТЕ являются фундаментальным свойством живого, необходимым для самовоспроизведения биополимеров при возникновении жизни на Земле.

ЗАКЛЮЧЕНИЕ

Физическое разобщение инсерциями транспозонов в структуре ДНК генов, участвующих в единых биологических реакциях (составляющих оперон у прокариот), у эукариот было компенсировано формированием сложных регуляторных сетей. Данные сети состоят из ТЕ, происходящих от них регуляторных последовательностей, белков и нкРНК, которые обеспечивают более динамическое управление работой генов в общих биохимических реакциях. Образуемые при этом сложные молекулярные взаимосвязи между разными биологическими процессами, стали более лабильными, обеспечивая возможность лучшей адаптации и усложнения организмов с появлением новых клеточных функций и межклеточных взаимосвязей. Сохранение герминативных инсерций ТЕ в эволюции, в связи с их адаптивным значением (образование новых регуляторных последовательностей и новых доменов БКГ) сопровождалось изменением структуры и количества хромосом. Это обусловлено происхождением теломер от ТЕ, их динамической взаимосвязью с ТЕ в эволюции и онтогенезе, а также наличием слож-

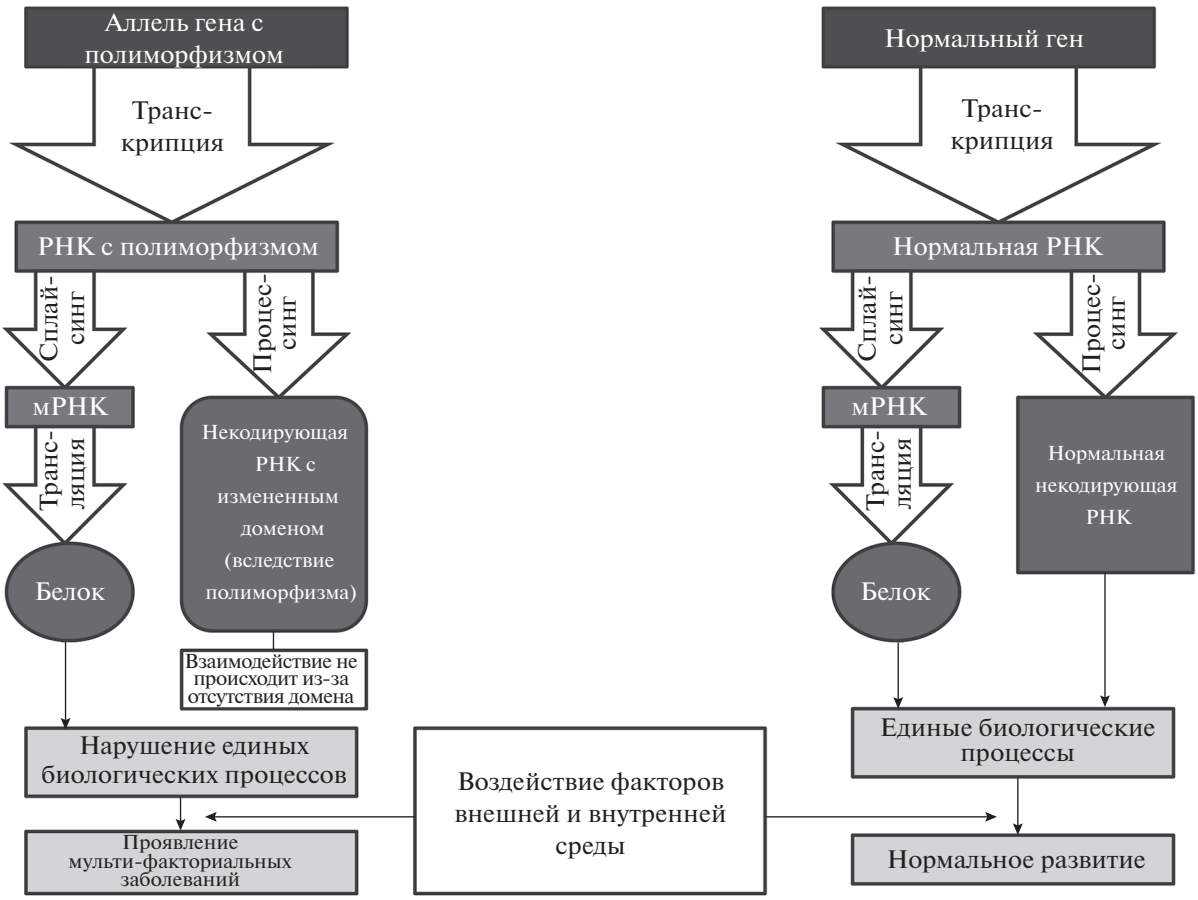

Рис. 7. Схема влияния однонуклеотидного полиморфизма (SNP) на развитие ассоциированного с ним мультифакториального заболевания.

ной системы взаимодействий между ТЕ, гистонами и теломерами. Взаимосвязь ТЕ с гистонами обусловлена их специфической модификацией, наряду с изменением метилирования ДНК, под влиянием происходящих от ТЕ siРНК и микроРНК. Таким образом, для эукариот характерно функционирование сложных систем между ТЕ, теломерами, гистонами, метилированием ДНК, системами сплайсинга и РНКи, которые обеспечивают динамическую регуляцию использования инсертированных ТЕ для нужд хозяина в эволюции и онтогенезе. Взаимосвязь этих систем обеспечила возможность возникновения новых БКГ из нкРНК, происходящих от ТЕ. При этом вновь образованные в эволюции белки характеризуются взаиморегуляцией с молекулами РНК, из которых они транслировались, так как данные РНК проявляют функциональность либо самостоятельно, либо в составе РНП (для lncРНК) или RISC (для микроРНК). Стратегия “ТЕ – нкРНК – пептиды” в эволюции геномов говорит о потенциальной функциональности большинства мРНК, транслируемых в белки. С одной стороны, это может стать ключом к объяснению роли SNP, не вызывающих изменение аминокислотных последовательностей белков, но играющих роль в генезе мультифакториальных заболеваний (рис. 7). С другой стороны, обнаружение функциональности мРНК клеточных белков может свидетельствовать об их возникновении в эволюции при участии ТЕ с возможным использованием стратегии “ТЕ–нкРНК–пептиды”. Можно предположить, что данная стратегия возникла первично при зарождении жизни на Земле как универсальное свойство, за счет которого происходил “поиск” наиболее оптимальных взаимодействий между пространственными структурами РНК и белков с последовательностями нуклеотидов (главным образом повторами) ДНК и РНК.

СПИСОК ЛИТЕРАТУРЫ

Бадаева Е.Д., Салина Е.А. Структура генома и хромосомный анализ растений // Вавиловский журн. генетики и селекции. 2013. Т. 17. С. 1017–1042.

Виноградов А.Е. Функциональное значение базовых свойств структуры генома эукариот. Автореф. ... дис. док. биол. наук. СПб. 2011.

Мустафин Р.Н., Хуснутдинова Э.К. Роль транспозонов в эпигенетической регуляции онтогенеза // Онтогенез. 2018. Т. 49. № 2. С. 69–90.

Патрушев Л.И., Минкевич И.Г. Проблема размера геномов эукариот // Успехи биологической химии. 2007. Т. 47. С. 293–370.

Равин Н.В., Шестаков С.В. Геном прокариот // Вавиловский журн. генетики и селекции. 2013. Т. 17. № 4/2. С. 972–984.

Шестаков С.В. Роль архей в происхождении эукариот // Экологическая генетика. 2017. Т. 15. № 4. С. 52–59.

Abrusan G., Zhang Y., Szilagyi A. Structure prediction and analysis of DNA transposon and LINE retrotransposons proteins // J. Biol. Chem. 2013. V. 288. P. 16127–16138.

Anderson D.M., Anderson K.M., Cang C.L. et al. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance // Cell. 2015. V. 160. P. 595–606.

Alzohairy A.M., Gyulai G., Jansen R.K., Bahieldin A. Transposable elements domesticated and neofunctionalized by eukaryotic genomes // Plasmid. 2013. V. 69. № 1. P. 1–15.

Arkhipova I.R. Neutral Theory, Transposable Elements, and Eukaryotic Genome Evolution // Mol. Biol. Evol. 2018. V. 35. № 6. P. 1332–1337.

Babakhani S., Oloomi M. Transposons: the agents of antibiotic resistance in bacteria // J. Basic Microbiol. 2018. doi 10.1002/jobm.201800204

Barry G. Small RNAs and Transposable Elements Are Key Components in the Control of Adaptive Evolution in Eukaryotes // Bioessays. 2018. V. 40. № 8. e1800070.

Belancio V.P., Roy-Engel A.M., Deininger P.L. All y'all need to know 'bout retroelements in cancer // Semin. Cancer Biol. 2010. V. 20. P. 200–210.

Betts H.C., Puttick M.N., Clark J.W. et al. Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin // Nat. Ecol. Evol. 2018. doi 10.1038/s41559-018-0644-x

Borchert G.M., Holton N.W., Williams J.D. et al. Comprehensive analysis of microRNA genomic loci identifies pervasive repetitive-element origins // Mobile Genetic Elements. 2011. V. 1. № 1. P. 8–17.

Cai J., Zhao R., Jiang H., Wang W. De novo origination of a new protein-coding gene in Saccharomyces cerevisiae // Genetics. 2008. V. 179. P. 487–496.

Casacuberta E. Drosophila: Retrotransposons Making up Telomeres // Viruses. 2017. V. 9. № 7. pii:E192.

Cerbin S., Jiang N. Duplication of host genes by transposable elements // Curr. Opin. Genet. Dev. 2018. V. 49. P. 63–69.

Cho J. Transposon-Derived Non-coding RNAs and Their Function in Plants // Front. Plant. Sci. 2018. V. 9. P. 600.

Chuong E.B., Rumi M.A., Soares M.J., Baker J.C. Endogenous retroviruses function as species-specific enhancer elements in the placenta // Nat. Genet. 2013. V. 45. № 3. P. 325–329.

Couzigou J.M., Lauressergues D., Becard G., Comier J.P. miRNA-encoded peptides (miPEPs): A new tool to analyze the role of miRNAs in plant biology // RNA Biol. 2015. V. 12. P. 1178–1180.

Couzigou J.M., Andre O., Cuillotin B. et al. Use of microRNA-encoded peptide miPEP172c to stimulate nodulation in soybean // New Phytol. 2016. V. 211. P. 379–381.

de Koning A.P., Gu W., Castoe T.A. et al. Repetitive Elements May Comprise Over Two-Thirds of the Human Genome // Plos Genetics. 2011. V. 7. № 12. e1002384.

de Souza F.S., Franchini L.F., Rubinstein M. Exaptation of transposable elements into novel cis-regulatory elements: is the evidence always strong // Mol. Biol. Evol. 2013. V. 30. № 6. P. 1239–1251.

Duan C.G., Wang X., Xie S. et al. A pair of transposon-derived proteins function in a histone acetyltransferase complex for active DNA demethylation // Cell. Res. 2017. V. 27. № 2. P. 226–240.

Dupressoir A., Lavialle C., Heidmann T. From ancestral infectious retroviruses to bona fide cellular genes: role of the captured syncytins in placentation // Placenta. 2012. V. 33. P. 663–671.

Elkon R., Ugalde A.P., Agami R. Alternative cleavage and polyadenylation: extent, regulation and function // Nat. Rev. Genet. 2013. V. 14. № 7. P. 496–506.

Ender C., Krek A., Friedlander M.R. et al. Human snoRNA with MicroRNA-Like Functions // Mol. Cell. 2008. V. 32. № 4. P. 519–528.

Feng G., Leem Y.E., Levin H.L. Transposon integration expression of stress response genes // Nucleic Acids Res. 2013. V. 41. № 2. P. 775–789.

Feschotte C. Transposable elements and the evolution of regulatory networks // Nat. Rev. Genet. 2008. V. 9. № 5. P. 397–405.

Filee J. Giant viruses and their mobile genetic elements: the molecular symbiosis hypothesis // Curr. Opin. Virol. 2018. V. 33. P. 81–88.

Garavis M., Gonzalez C., Villasante A. On the origin of the eukaryotic chromosome: the role of noncanonical DNA structures in telomere evolution // Genome Biol. Evol. 2013. V. 5. P. 1142–1150.

Gim J., Ha H., Ahn K. et al. Genome-Wide Identification and Classification of microRNAs derived from repetitive elements // Genomic Inform. 2014. V. 12. № 4. P. 261–267.

Grandi F.C., Rosser J.M., Newkirk S.J. et al. Retrotransposition creates sloping shores: a graded influence of hypomethylated CpG islands on flanking CpG sites // Genome Res. 2015. V. 25. № 8. P. 1135–1146.

Goerner-Potvin P., Bourque G. Computational tools to unmast transposable elements // Nat. Rev. Genet. 2018. V. 19. № 11. P. 688–704.

Guo L., Zhao Y., Yang S. et al. An integrated evolutionary analysis of miRNA-lncRNA in mammals // Mol. Biol. Rep. 2014. V. 41. P. 201–207.

Henikoff S., Smith M.M. Histone variants and epigenetics // Cold. Spring Harb. Perspect. Biol. 2015. V. 7. № 1. a019364. doi 10.1101/cshperspect.a019364

Honson D.D., Macfarlan T.S. A lncRNA-like Role for LINE1s in Development // Dev. Cell. 2018. V. 46. № 2. P. 132–134.

Huang C.J., Lin W.Y., Chang C.M., Choo K.B. Transcription of the rat testis-specific Rtdpoz-T1 and -T2 retrogenes during embryo development: co-transcription and frequent exonisation of transposable element sequences // BMC Mol. Biol. 2009. V. 10. P. 74–88.

Ito J., Suqimoto R., Nakaoka H. et al. Systematic identification and characterization of regulatory elements derived from human endogenous retroviruses // PLoS Genet. 2017. V. 13. № 7. e1006883.

Jacob M.D., Audas T.E., Mullineux S.T., Lee S. Where no RNA polymerase has gone before: novel functional transcripts derived from the ribosomal Intergenic spacer // Nucleus. 2012. V. 3. № 4. P. 315–319.

Jacques P.E., Jeyakani J., Bourgue G. The majority of primate-specific regulatory sequences are derived from

transposable elements // PLoS Genet. 2013. V. 9. № 5. e1003504.

Jjingo D., Conley A.B., Wang J. et al. Mammalian-wide interspersed repeat (MIR)-derived enhancers and the regulation of human gene expression // Mob. DNA. 2014. V. 5. P. 5–14.

Johnson R., Guigo R. The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs // RNA. 2014. V. 20. № 7. P. 959–976.

Joly-Lopez Z., Bureau T.E. Exaptation of transposable element coding sequences // Curr. Opin. Genet. Dev. 2018. V. 49. P. 34–42.

Judd J., Feschotte C. Transposons take remote control // Elife. 2018. V. 7. pii: e40921.

Kapusta A., Feschotte C. Volatile evolution of long noncoding RNA repertoires: mechanisms and biological implications // Trends Genet. 2014. V. 30. № 10. P. 439–452.

Klein S.J., O'Neill R.J. Transposable elements: genome innovation, chromosome diversity, and centromere conflict // Chromosome Res. 2018. V. 26. № 1–2. P. 5–23.

Kopera H.C., Moldovan J.B., Morrish T.A. et al. Similarities between long interspersed element-1 (LINE-1) reverse transcriptase and telomerase // Proc. Natl. Acad. Sci. USA. 2011. V. 108. P. 20345–20350.

Kubiak M.R., Makalowska I. Protein-Coding Genes' Retrocopies and Their Functions // Viruses. 2017. V. 9. pii: E80.

Kumar P., Anaya J., Mudunuri S.B., Dutta A. Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets // BMC Biol. 2014. V. 12. P. 78.

Lauressergues D., Couzigou J.M., Clemente H.S. et al. Primary transcripts of microRNAs encode regulatory peptides // Nature. 2015. V. 520. № 7545. P. 90–93.

Leplae R., Hebrant A., Wodak S.J., Toussaint A. ACLAME: a CLAssification of Mobile genetic Elements // Nucleic Acids Res. 2004. V. 32. P. 45–49.

Levine M.T., Jones C.D., Kern A.D. et al. Novel genes derived from noncoding DNA in Drosophila melanogaster are frequently X-linked and exhibit testis-biased expression // Proc. Natl. Acad. Sci. USA. 2006. V. 103. P. 9935–9939.

Li Y., Li C., Xia J., Jin Y. Domestication of transposable elements into microRNA genes in plants // PLoS One. 2011. V. 6. e19212.

Li Z., Ender C., Meister G. et al. Extensive terminal and asymmertric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs // Nucleic Acids Res. 2012. V. 40. № 14. P. 6787–6799.

Long Y., Wang X., Youmans D.T., Cech T.R. How do lncRNAs regulate transcription? // Sci. Adv. 2017. V. 3. № 9. eaao2110.

Lorenzetti A.P., de Antonio G.Y., Paschoal A.R., Domingues D.S. Plant TE-MIR DB: a database for transposable element-related microRNAs in plant genomes // Funct. Integr. Genomics. 2016. V. 16. P. 235–242.

Lowe C.B., Haussler D. 29 mammalian genomes reveal novel exaptations of mobile elements for likely regulatory functions in the human genome // PLoS One. 2012. V. 7. № 8. P. e43128.

Lu X., Sachs F., Ramsay L. et al. The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity // Nat. Struct. Mol. Biol. 2014. V. 21. № 4. P. 423–425.

Lv S., Pan L., Wang G. Commentary: Primary Transcripts of microRNAs Encode Regulatory Peptides // Front Plant Sci. 2016. V. 7. P. 1436.

Lv J., Wang L., Zhang J. et al. Long noncoding RNA H19-derived miR-675 aggravates restenosis by targeting PTEN // Biochem. Biophys. Res. Commun. 2018. V. 497. P. 1154–1161.

Martinez G., Choudury S.G., Slotkin R.K. tRNA-derived small RNAs target transposable element transcripts // Nucleic Acids Res. 2017. V. 45. № 9. P. 5142–5152.

Marzluff W.F., Koreski K.P. Birth and Death of Histone mRNAs // Trends Genet. 2017. V. 33. № 10. P. 745–759.

Mak K.S., Burdach J., Norton L.J. et al. Repression of chimeric transcripts emanating from endogenous retrotransposons by a sequence-specific transcription factor // Genome Biol. 2014. V. 15. № 4. R58.

McGurk M.P., Barbash D.A. Double insertion of transposable elements provides a substrate for the evolution of satellite DNA // Genome Res. 2018. V. 28. № 5. P. 714–725.

Morrish T.A., Garcia-Perez J.L., Stamato T.D. et al. Endonuclease-independent LINE-1 retrotransposition at mammalian telomeres // Natrure. 2007. V. 446. № 7132. P. 208–212.

Nelson B.R., Makarewich C.A., Anderson D.M. et al. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle // Science. 2016. V. 351. P. 271–275.

Novikova O., Belfort M. Mobile Group II Introns as Ancestral Eukaryotic Elements // Trends Genet. 2017. V. 33. № 11. P. 773–783. doi 10.1016/j.tig2017.07.009

Rearick D., Prakash A., McSweeny A. et al. Critical association of ncRNA with introns // Nucleic Acids Res. 2011. V. 39. № 6. P. 2357–2366.

Ricci M., Peona V., Guichard E. et al. Transposable Elements Activity is Positively Related to Rate of Speciation in Mammals // J. Mol. Evol. 2018. V. 86. P. 303–310.

Roberts J.T., Cooper E.A., Favreau C.J. Formation from transposable element insertions and noncoding RNA mutations // Mob. Genet. Elements. 2013. V. 1. e27755.

Ruiz-Orera J., Messeguer X., Subirana J.A., Alba M.M. Long non-coding RNAs as a source of new peptides // Elife. 2014. V. 3. e03523. doi 10.7554/eLife.03523

Sadic D., Schmidt K., Groh S. et al. Atrx promotes heterochromatin formation at retrotransposons // EMBO Rep. 2015. V. 16. № 7. P. 836–850.

Sakai H., Mizuno H., Kawahara Y. et al. Retrogenes in rice (*Oryza sativa* L. ssp. japonica) exhibit correlated expression with their source genes // Genome Biol. Evol. 2011. V. 3. P. 1357–1368.

Saze H. Epigenetic regulation of intragenic transposable elements: a two-edged sword // J. Biochem. 2018. V. 164. P. 323–328.

Schrader L., Schmitz J. The impact of transposable elements in adaptive evolution // Mol. Ecol. 2018. doi 10.111/mec.14794

Serrato-Capuchina A., Matute D.R. The Role of Transposable Elements in Speciation // Genes (Basel). 2018. V. 9. № 5. pii: E254.

Sinzelle L., Izsvak Z., Ivics Z. Molecular domestivation of transposable elements: From detrimental parasites to useful host genes // Cell. Mol. Life Sci. 2009. V. 66. P. 1073–1093.

Song M.J., Schaack S. Evolutionary Conflict between Mobile DNA and Host Genomes // Am. Nat. 2018. V. 192. № 2. P. 263–273.

Startek M.P., Nogly J., Gromadka A. et al. Inferring transposons activity chronology by TRANAcendence-TEs database and de-novo mining tool // BMC Bioinformatics. 2017. V. 18. P. 422.

Taft R.J., Glazov E.A., Lassmann T. et al. Small RNAs derived from snoRNAs // RNA. 2009. V. 15. P. 1233–1240.

Tan S., Cardoso-Moreira M., Shi W. et al. LTR-mediated retroposition as a mechanism of RNA-based duplication in metazoans // Genome Res. 2016. V. 26. № 12. P. 1663–1675.

Tajnik M., Vigilante A., Braun S. et al. Inergenic Alu exonisation facilitates the evolution of tissue-specific transcript ends // Nucleic Acids Res. 2015. V. 43. № 21. P. 10492–10505.

Toro N., Nisa-Martinez R. Comprehensive phylogenetic analysis of bacterial reverse transcriptases // PLoS One. 2014. V. 9. № 11. e114083.

Toro N., Martinez-Abarca F., Gonzalez-Delgado A., Mestre M.R. On the origin and evolutionary relationships of the reverse transcriptases associated with type III CRISPR-Cas Systems // Front. Mircrobiol. 2018. V. 9. P. 1317.

Trizzino M., Kapusta A., Brown C.D. Transposable elements generate regulatory novelty in a tissue-specific fashion // BMC Genomics. 2018. V. 19. № 1. P. 468.

Venkatesh T., Suresh P.S., Tsutsumi R. tRFs: miRNAs in disguise // Gene. 2016. V. 579. P. 133–138.

Volff J.N. Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes // BioEssays. 2006. V. 28. P. 913–922.

Wang D., Su Y., Wang X. et al. Transposon-Derived and Satellite-Derived Repetitive Sequences Play Distinct Functional Roles in Mammalian Intron Size Expansion // Evol Bioinform Online. 2012. V. 8. P. 301–319.

Wang J., Vicente-Garcia C., Seruqqia D. et al. MIR retrotransposons sequences provide insulators to the human genome // Proc Natl Acad Sci USA. 2015. V. 112. P. 4428–4437.

Wheeler B.S. Small RNAs, big impact: small RNA pathways in transposon control and their effect on the host stress response // Chromosome Res. 2013. V. 21. P. 587–600.

Wicker T., Sabot F., Hua-Van A. et al. A unified classification system for eukaryotic transposable elements // Nat. Rev. Genet. 2007. V. 8. № 12. P. 973–982.

Xie C., Zhang Y.E., Chen J.Y. et al. Hominoid-specific de novo protein-coding genes originating from long noncoding RNAs // PLoS Genet. 2012. V. 8. e1002942.

Yenerall P., Zhou L. Identifying the mechanisms of intron gain: progress and trends // Biol. Direct. 2012. V. 7. P. 29. doi 10.1186/1745-6150-7-29

Yuan Z., Sun X., Liu H., Xie J. MicroRNA genes derived from repetitive elements and expanded by segmental duplication events in mammalian genomes // PLoS One. 2011. V. 6. e17666.

Zdobnov E.M., Campillos M., Harrington E.D. et al. Protein coding potential of retroviruses and other transposable elements in vertebrate genomes // Nucleic Acids Res. 2005. V. 33. P. 946–954.

Zhang J., Mujahid H., Hou Y. et al. Plant Long ncRNAs: A New Frontier for Gene Regulatory Control // American J. of Plant Sciences. 2013. V. 4. ID:32139.

Zeng L., Pederson S.M., Kortschak R.D., Adelson D.L. Transposable elements and gene expression during the evolution of amniotes // Mob. DNA. 2018. V. 9. P. 17.

Zhang H.H., Zhou Q.Z., Wang P.L. et al. Unexpected invasion of miniature inverted-repeat transposable elements in viral genomes // Mob. DNA. 2018. V. 9. P. 19.

Zhu Z., Tan S., Zhang Y., Zhang Y.E. LINE-1-like retrotransposons contribute to RNA-based gene duplication in dicots // Sci. Rep. 2016. V. 6. P. 24755.

Functional Dualism of Transcripts of Transposons in the Evolution of Eukariotic Genomes

R. N. Mustafin[1, 2, *]

[1]Bashkir State Medical University, Lenin str., 3, Ufa, Russia

[2]Bashkir State University, Zaki Validi str., 32, Ufa, Russia

**e-mail: ruji79@mail.ru*

Received August 11, 2018
Revised October 30, 2018

In the evolution of eukaryotes, the insertion of transposons structurally disconnected genes encoding the same biological processes. But at the same time, transposons contributed to more complex regulation of these genes, which is necessary for adaptation. This became possible through the creation of universal conservative systems that include DNA methylation, histone modifications, the relationship of telomeres to transposons, regulation of splicing and RNA interference. Due to the functioning of these systems, transposons have been used to create new genes through duplication, exonization and domestication. The new genes thus formed contain sequences of transposons that are necessary for their control in regulatory gene networks in which the transposons themselves participate. The sources of new proteins of eukaryotes are also noncoding RNA, which, simultaneously with translation, are processed into functional RNAs. This duality of functions of non-coding RNA is due to their origin from transposons, which are attacked by protective systems of host processing. We suggested that the primary transcripts of genes encoding proteins can be subjected to both splicing and processing into functional RNAs. This justifies the effect of SNP of genes on the development of multifactorial diseases, since polymorphisms can cause inactivation of RNA domains while maintaining the function of the protein. We hypothesize that the functional dualism of transposon transcripts was an important condition for the onset of life on Earth, and mobile genetic elements are a fundamental property of the living.

Keywords: protein-coding genes, histones, long noncoding RNA, methylation, microRNA, peptides, regulation, transposable elements