



DOI: 10.18413/2658-6533-2025-11-2-0-2

УДК 575.852

Роль транспозонов в обеспечении эволюционного разнообразия протеомов эукариот (обзор)

Р.Н. Мустафин¹ , Э.К. Хуснутдинова^{2,3}

¹ Федеральное государственное бюджетное образовательное учреждение высшего образования «Башкирский государственный медицинский университет», ул. Ленина, д.3, г. Уфа, 450008, г. Уфа, Российская Федерация

² Федеральное государственное бюджетное научное учреждение Уфимский федеральный исследовательский центр Российской академии наук, пр-т. Октября, д. 71, г. Уфа, 450054, Российская Федерация

³ Федеральное государственное бюджетное образовательное учреждение высшего образования «Уфимский университет науки и технологий», ул. Заки Валиди, д. 32, г. Уфа, 450076, Российская Федерация

Автор для переписки: Р.Н. Мустафин (ruji79@mail.ru)

Резюме

Актуальность: Транспозоны составляют значительную долю геномов эукариот и являются объектами для современных генетических исследований, в том числе в проектировании таргетной терапии опухолей. Для оптимального проектирования таких методов важно определить возможные взаимосвязи мобильных генетических элементов с регуляторными элементами генома в связи с происхождением от транспозонов многих белок-кодирующих генов и способных к трансляции некодирующих РНК. **Цель исследования:** Описать роль мобильных генетических элементов в обеспечении эволюционного разнообразия протеомов эукариот как за счет непосредственного возникновения белок-кодирующих генов от транспозонов, так и путем возникновения от них некодирующих РНК, способных к трансляции. Определить практическую ценность полученных результатов путем анализа вовлечения образуемых из некодирующих РНК пептидов в канцерогенезе. **Материалы и методы:** Использованы базы данных Scopus, WoS, PubMed для анализа роли транспозонов в возникновении белок-кодирующих генов, микроРНК, длинных некодирующих РНК и кольцевых РНК, об участии образуемых при трансляции данных молекул РНК пептидов в канцерогенезе. **Результаты:** Согласно проанализированной литературе, транспозоны являются важнейшими источниками возникновения и эволюции белок-кодирующих генов за счет одомашнивания генов мобильных генетических элементов, экзонизации инсертированных ретроэлементов и образования псевдогенов. Большинство длинных некодирующих РНК, кольцевых РНК, многих генов микроРНК, а также их регуляторных элементов в эволюции произошли от транспозонов. У человека достоверно выявлена трансляция 15 длинных некодирующих РНК, 4 при-микроРНК и 6 кольцевых РНК с образованием функциональных пептидов, вовлеченных в механизмы канцерогенеза. **Заключение:** Возникновение в эволюции большинства генов некодирующих РНК и многих белок-кодирующих генов от транспозонов свидетельствует о возможности использования данных геномных элементов в качестве мишеней для перспективных генетических

исследований, в том числе в лечении заболеваний у человека. Свидетельством служат полученные данные о вовлечении в канцерогенез 25 пептидов, образуемых при трансляции некодирующих РНК.

Ключевые слова: белок-кодирующие гены; канцерогенез; длинные некодирующие РНК; микроРНК; транспозоны; эволюция

Для цитирования: Мустафин РН, Хуснутдинова ЭК. Роль транспозонов в обеспечении эволюционного разнообразия протеомов эукариот (обзор). Научные результаты биомедицинских исследований. 2025;11(2):215-242. DOI:10.18413/2658-6533-2025-11-2-0-2

The role of transposable elements in formation of evolutionary diversity of eukaryotic proteomes (review)

Rustam N. Mustafin¹ , Elza K. Khusnutdinova^{2,3} 

¹ Bashkir State Medical University,
3 Lenin St., Ufa, 450008, Russia

² Ufa Scientific Center, RAS,
71 Oktyabrya Ave., Ufa, 450054, Russia

³ Ufa University of Science and Technology,
32 Zaki Validi St., Ufa, 450076, Russia

Corresponding author: Rustam N. Mustafin (ruji79@mail.ru)

Abstract

Background: Transposable elements make up a significant proportion of eukaryotic genomes and are objects of modern genetic research, including in the design of targeted tumor therapies. For the optimal design of such methods, it is important to determine the possible relationships of transposons with genome regulatory elements, as many protein-coding genes and translatable non-coding RNAs originate from transposons. **The aim of the study:** To describe the role of transposons in ensuring the evolutionary diversity of eukaryotic proteomes, both through the direct origin of protein-coding genes and translation-capable non-coding RNAs from transposons. To determine the practical value of the results obtained by analysing the involvement of peptides formed from non-coding RNAs in carcinogenesis. **Materials and methods:** Scopus, WoS, PubMed databases were used to analyse the role of transposons in the origin of protein-coding genes, microRNAs, long non-coding RNAs and circular RNAs, and the involvement of peptides formed during the translation of these RNA molecules in carcinogenesis. **Results:** According to the literature reviewed, transposons are the most important sources of origin and evolution of protein-coding genes due to their domestication, exonisation of inserted retroelements and formation of pseudogenes. Most long non-coding RNAs, circular RNAs, many microRNA genes and their regulatory elements have evolved from transposons. In humans, the translation of 15 long non-coding RNAs, 4 pri-microRNAs and 6 circular RNAs with the formation of functional peptides involved in carcinogenesis mechanisms was reliably detected. **Conclusion:** The evolutionary origin of most non-coding RNA genes and many protein-coding genes from transposable elements suggests the prospects of using these genomic elements as targets for advanced genetic research, including the treatment of human diseases. This is evidenced by the data obtained on the involvement of 25 peptides formed during the translation of non-coding RNA in carcinogenesis.

Keywords: protein-coding genes; carcinogenesis; long non-coding RNA; microRNA; transposons; evolution

For citation: Mustafin RN, Khusnutdinova EK. The role of transposable elements in formation of evolutionary diversity of eukaryotic proteomes (review). Research Results in Biomedicine. 2025;11(2):215-242. Russian. DOI:10.18413/2658-6533-2025-11-2-0-2

Введение. Транспозоны или мобильные генетические элементы (МГЭ) представляют собой специфические локусы генома, способные к перемещениям в новый локус того же генома с помощью ферментов, кодируемых их собственными генами (автономные МГЭ) или генами других МГЭ (неавтономные). У эукариот различают класс I МГЭ – ретроэлементы (РЭ) и класс II – ДНК-транспозоны. МГЭ класса I перемещаются за счет обратной транскрипции собственных РНК со встраиванием кДНК в новый локус генома по механизму «копирования и вставки». Их классифицируют на порядки LTR, DIRS, PLE, LINE, SINE. ДНК-транспозоны классифицируют на подкласс 1, включающий TIR и Crypton и подкласс 2, содержащий порядки Helitron и Maverick [1]. Геномы большинства бактерий и археев практически не содержат МГЭ, за некоторыми исключениями. Например, у бактерии *Clostridium difficile* МГЭ составляют 11% последовательностей генома, у *Enterococcus faecalis* – 25%, у *Orientia tsutsugamushi* – 46,7% и отличаются способностью взаимодействовать между собой путем рекомбинации, формируя новые химерные элементы [2]. Однако наибольшую регуляторную и эволюционную роль МГЭ приобрели у эукариот, о чем свидетельствует их глобальная распространенность и использование последовательностей МГЭ для образования и эволюции новых белок-кодирующих генов (БКГ) и некодирующих РНК (нкРНК). Благодаря этому эукариоты характеризуются выраженными различиями в размерах геномов. Например, у растений наиболее мелкий геном микроспоридии *Encephalitozoon intestinalis* размером в 2,3 миллиард п.н. меньше

крупного генома *Paris japonica* (148 852 миллиард п.н.) в 70 000 раз, а количество нуклеотидов ДНК амебы почти в 200 раз больше по сравнению с геномом человека [3], который представляет наибольший интерес в современных исследованиях в связи с открывающимися возможностями геномного редактирования.

В ядерной ДНК человека LINE занимают 0,63 млрд. п.н., SINE – 0,39 млрд. п.н., LTR – 0,27 млрд. п.н., ДНК-транспозоны – 0,108 млрд. п.н. Всего МГЭ составляют около 1,4 млрд. п.н., что составляет 46,7% всех последовательностей генома [4]. При помощи специфических олигонуклеотидов, комплементарных транспозонам, проведенный анализ генома человека показал, что последовательности МГЭ занимают не менее 2/3 всего генома человека [5]. Такое несоответствие обусловлено ключевой ролью МГЭ в возникновении и эволюции генов нкРНК и белок-кодирующих генов, что предполагает многофункциональность таких генов как специфическое свойство МГЭ [6], поскольку при возникновении мира ДНК-РНК-белки из мира РНК-ДНК, транспозоны служили универсальными источниками происхождения бесчисленного разнообразия пептидных и белковых молекул благодаря процессингу их транскриптов [7]. Соответственно, транскрипты МГЭ процессируются различными ферментативными системами с образованием множества вариантов молекул РНК, которые могут связываться с рибосомами и формировать огромное разнообразие белков и пептидов. Действительно, в эволюции МГЭ становились источниками не только множества БКГ [8-19], но также генов

микроРНК [20] и длинных некодирующих РНК (днРНК) [21, 22].

Согласно опубликованной в 2022 году статье о расшифровке полной последовательности генома человека, в гаплоидном наборе содержится 3,055 миллиардов пар нуклеотидов. Всего выявлено 63494 генов, из которых лишь 19969 БКГ [4]. Значительную долю составляют гены нкРНК, количество которых в базе данных GENECODE все более возрастает благодаря современным молекулярно-генетическим методам исследования. В 2024 году статистические данные этой базы (<https://www.gencodegenes.org/human/stats.html>) свидетельствуют о наличии 20424 генов днРНК, ответственных за образование 59719 транскриптов. Аннотировано 7565 генов малых нкРНК, а также 14719 псевдогенов [4], которые также образованы в эволюции благодаря МГЭ [23, 24]. Происхождение генов нкРНК [20, 21, 22] и БКГ [8-19] от МГЭ предполагает наличие специфических свойств таких генов и продуктов транскрипции, отражающих универсальные свойства МГЭ. Анализ научной литературы показывает появление сведений, подтверждающих это предположение. И в данной статье будут представлены доказательства того, что нкРНК являются важнейшими источниками возникновения новых белок-кодирующих генов в эволюции, что связано с их функциональной взаимосвязью с МГЭ. Исследование роли МГЭ в обеспечении разнообразия транскриптомов и протеомов имеет практическое значение, поскольку транспозоны являются мишенями для проведения таргетной терапии злокачественных новообразований [25]. Соответственно, полученные результаты о значении МГЭ в происхождении белков и нкРНК могут свидетельствовать о возможном вовлечении этих молекул в канцерогенез.

Цель исследования. Описать роль МГЭ в обеспечении эволюционного разнообразия протеомов эукариот как за счет непосредственного возникновения

белок-кодирующих генов от МГЭ, так и путем возникновения от транспозонов некодирующих РНК, способных к трансляции. Определить практическую ценность полученных результатов путем анализа вовлечения образуемых из нкРНК пептидов в канцерогенезе.

Материалы и методы исследования. Использованы базы данных Scopus, WoS, PubMed для анализа роли транспозонов в возникновении белок-кодирующих генов, микроРНК, длинных некодирующих РНК и кольцевых РНК, об участии образуемых при трансляции данных молекул РНК пептидов в канцерогенезе.

Результаты. Анализ научной литературы позволил выявить важную роль МГЭ в возникновении БКГ и различных нкРНК. Возможность трансляции таких нкРНК обеспечивает образование функциональных пептидов, многие из которых участвуют в регуляции канцерогенеза и могут быть использованы в качестве инструментов для таргетной противоопухолевой терапии.

Роль транспозонов в возникновении белок-кодирующих генов

Формирование в эволюции новых БКГ возможно тремя отдельными механизмами при помощи МГЭ: 1) путем непосредственного одомашнивания генов МГЭ, 2) за счет экзонизации инсертированных в интроны или некодирующие области существующих БКГ последовательностей МГЭ (с образованием новых вариантов сплайсинга) и 3) путем образования ретрогенов (псевдогенов) в результате использования ферментов РЭ (рис. 1). Даже формирование псевдогенов способствует изменению генных регуляторных сетей, так как ретрокопии фланкированы последовательностями МГЭ [26], находящимися под управлением сложных эпигенетических систем, с участием происходящих от МГЭ некодирующих РНК [6]. Сплайсосомные интроны также произошли от МГЭ в эволюции [24], с последующими новыми инсерциями МГЭ в

эти интроны, что способствовало формированию альтернативных сплайсинговых транскриптов. Если в ходе эволюции образованные новые белковые изоформы способствовали лучшей адаптации, эти варианты закреплялись на уровне вида. Возникла экзонизация. В геноме человека выявлены различные

экзоны, произошедшие от МГЭ [27]. С помощью модели глубокого обучения eXAlu, было показано, что количество элементов Alu в геноме человека, подверженных экзонизации составляет около 110 000, что в 21 раз больше, чем представлено в базе данных GENOCODE [28].

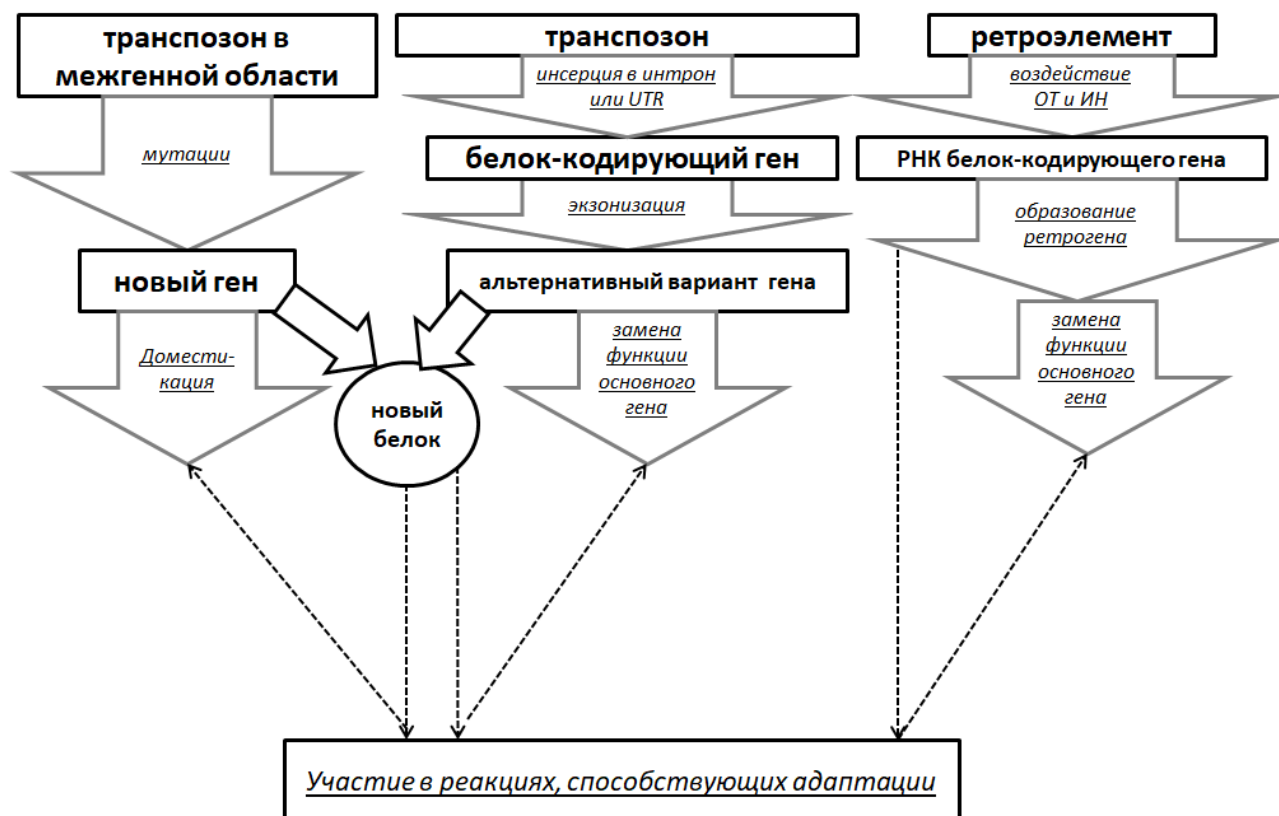


Рис. 1. Схема участия транспозонов в возникновении новых белок-кодирующих генов
(ОТ – обратная транскриптаза, ИН – интеграна,
UTR (untranslated region) – нетранслируемая область гена)

Fig. 1. Scheme of the participation of transposable elements in the emergence of new protein-coding genes (OT – reverse transcriptase, IN – integrase, UTR – untranslated region of the gene)

При образовании ретрогенов используются кодируемые ретроэлементами обратная транскриптаза для образования кДНК из транскриптов БКГ и эндонуклеаза для встраивания копии в геном. В результате образованные ретрогены содержат новые регуляторные последовательности, содержащиеся в локусе встраивания. Кроме того, такие новые БКГ во фланкирующих областях содержат специфические для LINE последовательности, такие как удвоенные

целевые сайты и последовательности расщепления эндонуклеазы TTTT/AA. У животных описано также участие ферментов ERV (эндогенных ретровирусов) в образовании ретрогенов. В таких случаях новые БКГ будут фланкированы длинными концевыми повторами (LTR), характерными для ERV [23]. Однако использование LTR-содержащих РЭ при образовании псевдогенов наиболее характерно для растений, поскольку в их геномах, по

сравнению с животными, преобладают LTR-содержащие РЭ. Например, у арабидопсиса описаны фланкированные длинными концевыми повторами функциональные ретрогены *CYP98A8* и *CYP98A9*, а у томатов – псевдоген *Sun* [29]. У растений отмечена также способность ретрогенов транскрибироваться на более низком уровне (подобно генам днРНК), что отражает свойства происходящих от МГЭ генов днРНК в эволюции. Например, у *Oryza L.* более 2/3 всех псевдогенов экспрессируются со специфическими для тканей особенностями, что также отражает сходство с генами днРНК [30].

Подобно генам днРНК, эволюционно молодые псевдогены, состоящие только из экзонов, постепенно приобретают мозаичную структуру за счет многократных встраиваний новых МГЭ, которые содержат донорные и акцепторные сайты сплайсинга, обеспечивая взаимодействие мРНК таких генов со сплайсосомой. Это дает возможность участвовать белковым продуктам таких генов в регуляторных сетях, в которых задействованы МГЭ. Соответственно, такие псевдогены в определенных условиях обеспечивают лучшую адаптацию особей и могут быть отобраны для экспрессии на более высоком уровне и использоваться вместо БКГ, копией которого они являются [24]. Кроме того, образованные ретрокопии служат источниками транскрипции некодирующих РНК, что было доказано при анализе распространения кластеров микроРНК с участием *Alu* и *L1* в геноме человека. За счет этого образуются новые возможности взаиморегуляции с другими генами и локусами генома, что потенциально служит источником новых адаптивных функций [31].

Для определения происхождения БКГ от транспозонов используют различные подходы, одним из которых является идентификация консервативных доменов, специфичных для МГЭ. В результате можно выявить множество генов, возникших в эволюции из последовательностей МГЭ и используемых

для нужд хозяев. Некоторые из таких БКГ образуют тандемные кластеры семейств генов [32]. Формирование новых генов путем вербовки доменов МГЭ является причиной их распространения в связи с возможностью участия в адаптивных реакциях. Доказательства одомашнивания МГЭ в эволюции были получены в связи с наличием функциональных различий между автономными и доместифицированными транспозонами [15]. МГЭ оказались источниками возникновения консервативных генов, играющих ключевую роль в эволюционных преобразованиях живых организмов. Например, обратная транскриптаза РЭ стала основой для формирования теломеразы [27]. У всех эукариот было идентифицировано множество БКГ, произошедших от МГЭ [10]. Так, у позвоночных описано более 1000 генов, источниками которых оказались РЭ [33]. Древняя эволюционная способность МГЭ образовывать белковые структуры, взаимодействующие с последовательностями собственных ДНК в геноме стала основой для возникновения от МГЭ различных транскрипционных факторов и сайтов связывания с ними, формируя сложные регуляторные генные сети [10, 34], находящиеся также под регуляторным влиянием МГЭ и произошедших от них некодирующих РНК (рис. 2).

Несмотря на то, что в геномах эукариот преобладают РЭ, важным источником БКГ различных животных, грибов и растений стал ген транспозазы ДНК-транспозонов (табл. 1). От транспозазы произошли такие консервативные белки как центромер-связывающий протеин CENP-B у животных, белок *harb1* у млекопитающих, рыб и лягушек. Важнейшие компоненты иммунной системы RAG у позвоночных также произошли от транспозазы, которая стала также источником возникновения белков Metnase и Pgbd у человека и мыши, протеинов BUSTER1-3, ZBED1, ZBED4, ZBED5, P52rlPK у млекопитающих. У арабидопсиса от транспозазы произошел белок Daysleeper [9, 15].

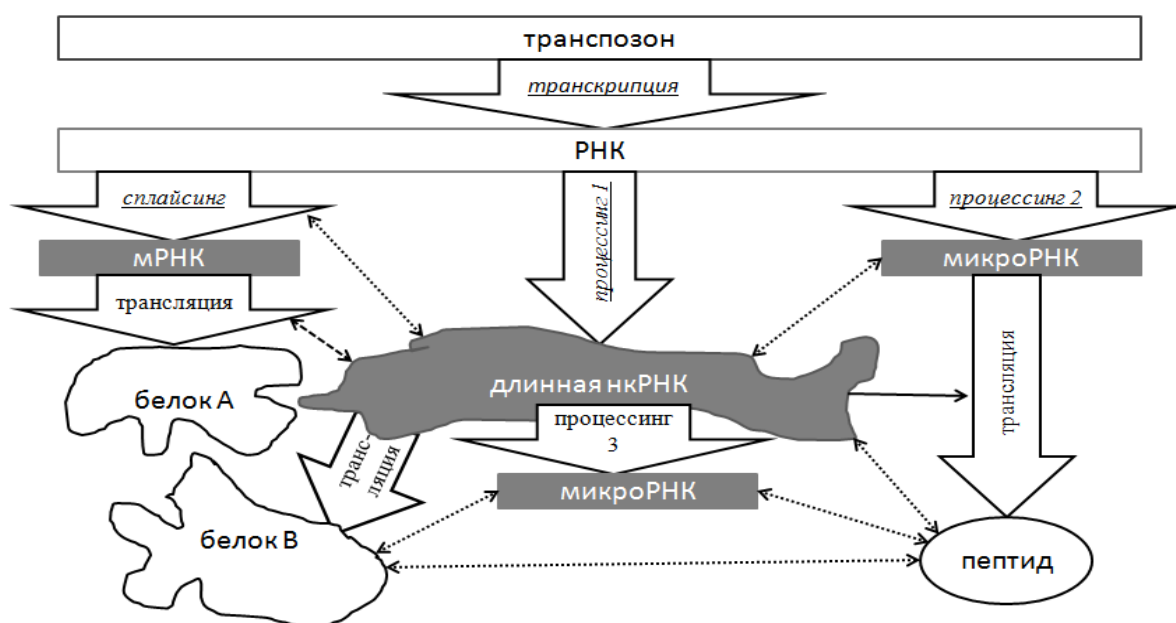


Рис. 2. Роль транспозонов в формировании регуляторных сетей управления функционированием генов

Fig. 2. Role of transposable elements in the formation of regulatory networks that control genes functioning

Специфические нуклеазы, используемые для перегруппировки ДНК у животных, произошли от транспозазы ДНК-транспозонов *Harbinger* [8]. Транспозазы стали основой для формирования таких ДНК-связывающих доменов эукариот как спираль-поворот-спираль (HTH) и цинковые пальцы (ZF). Кроме того, от транспозазы произошли HD (гомеодомен), KRAB (Kruppel-связанный бокс), BTB (Broad-Complex, Tramtrack, and Bric-a-brac), SET (Su(var), E(z) and Trithorax), SWIM (SWI2/SNF2 and MuDR), hATC (hAT C-terminal dimerization), LZ (лейциновая молния) [10]. Следует отметить роль ДНК-транспозонов в качестве эволюционных источников генов белков, участвующих в эпигенетической регуляции. Так, от ДНК-транспозона *Harbinger* у Арабидопсиса произошли гены *HDP1* (от транспозазы) и *HDP2* (от ДНК-связывающего белка), которые взаимодействуют с компонентами IDM1, IDM2, IDM3 и MBD7 ацетилтрансферазного комплекса, участвующего в деметилировании ДНК [18]. От транспозазы *hAT* произошли факторы модификации хроматина BEAF-32 и HIM-17 [11], центромерный белок Abp1 (у грибов) [10],

инсуляторы, участвующие в модификации хроматина. Инсуляторы – это регуляторные элементы, которые способствуют организации хроматина эукариот за счет блокирования энхансеров и активации хроматинового барьера [17]. Гены, произошедшие от ДНК-транспозонов, экспрессируются в белки THAP0, THAP1, E93 (ген *Eip93F*), участвующие в апоптозе, а также THAP, LIN-36, LIN-15B, контролирующие клеточный цикл [11]. У дрожжей интегразы ДНК-транспозонов стала источником возникновения гена *Fob1p*, продукт которого управляет рекомбинацией рРНК [9, 15].

От гена интегразы LTR-ретроэлементов произошел ген *Gin-1* у млекопитающих, белковый продукт которого участвует в регуляции эмбриогенеза. От генов GAG ретроэлементов произошли гены, вовлеченные в иммунные ответ (*Ma*), в управление транскрипцией миелина (*MyEF-3*), в регуляцию апоптоза и пролиферации клеток (*Mart*), противовирусный ответ (*Fv1*), регуляцию экспрессии отцовских генов (*Rtl1*), партеногенетическое развитие (*PEG10*) [9, 15].

Таблица 1

Белок-кодирующие гены эукариот, произошедшие от транспозазы [9-18]

Table 1

Protein-coding genes of eukaryotes derived from transposase [9-18]

ДНК-транспозон	Организм	Ген хозяина (функция белка)
<i>CACTA</i>	львиный зев (антирринум)	<i>ROSINA</i> (модуляция развития лепестков и тычинок)
<i>hAT</i>	арабидопсис	<i>Daysleeper</i> (регуляция развития)
	травянистые растения	<i>Gary</i> (неизвестна)
	сахарный тростник	<i>SchAT</i> (транспозаза с мотивом DDE)
	дрозофила	<i>BEAF-32</i> (регуляция хроматина, инсуляторная активность)
		<i>DREF</i> (репликация ДНК, дифференцировка клеток)
	круглые черви	<i>GON-14</i> (регуляция развития)
		<i>LIN-15B</i> (ингибитор клеточного цикла G1/S)
	млекопитающие	<i>GTF2IRD2</i> (фактор транскрипции)
		<i>BUSTER1-3</i> <i>P52rIPK</i> (ингибирование протеинкиназы PKR) <i>ZBED1, ZBED4, ZBED5</i> (регуляция транскрипции)
<i>Harbinger</i>	млекопитающие, птицы, рыбы, лягушки	<i>HARBI1</i> (перегруппировка ДНК)
	позвоночные	<i>NAIF1</i> (ядерный перенос HARBI1)
	растение арабидопсис	<i>HDP1, HDP2</i> (деметилирование ДНК)
<i>Mariner</i>	млекопитающие, растения и грибы	<i>haT-like</i> (участие в репродукции)
	человек, мышь	<i>SETMAR</i> (негомологичная репарация двойных разрывов ДНК)
		<i>METNASE</i> (интеграция и репарация ДНК)
<i>Mutator-like</i>	арабидопсис	<i>FHY3, FAR1</i> (пути фитохрома A)
<i>Mutator</i>	грибы	<i>Aft1</i> (утилизация ионов и гомеостаз)
		<i>Rcs1</i> (утилизация ионов и гомеостаз)
	арабидопсис, рис	<i>MUG1</i> (регуляция транскрипции)
<i>P-element</i>	дрозофила	<i>phsa/pgga</i> (репрессор-подобный белок с ДНК-связывающим мотивом)
	млекопитающие	<i>THAP0, THAP1</i> (индуцированный интерфероном-γ апоптоз)
		<i>THAP2, 3 – 6, 10, 11</i> (неизвестна)
		<i>THAP7</i> (рекрутирует деацетилазу HDAC3 и NcoR в специфические сайты ДНК)
		<i>THAP-E2F6</i> (репрессор E2F-зависимой транскрипции в S-фазу)
	круглые черви	<i>CDC14B</i> (ингибитор клеточного цикла G1/S)
		<i>CTB-1</i> (регуляция транскрипции)
		<i>HIM-17</i> (модификация хроматина)
<i>Pogo</i>	дрозофила, млекопитающие, растения	<i>LIN-36</i> (регуляция клеточного цикла G1/S)
		<i>CENP-B</i> (центромерный белок)
	дрозофила	<i>Bab1, Bab2</i> (морфогенез яичников и тела)
		<i>Eip93F</i> (регуляция опосредованной стероидами программированной клеточной гибели при метаморфозе)
	грибы	<i>Cbh1, Cbh2</i> (центромерные белки)
		<i>Abp1</i> (сегрегация хромосом, формирование гетерохроматина центромер)
	млекопитающие	JRK, JRKL (ДНК- и РНК-связывающая активность в нейронах)
		<i>PGBD1-5</i> (неизвестна)
<i>PiggyBac-like</i>	позвоночные	<i>KOBUTA</i> (неизвестна)
	шпорцевые лягушки	
<i>Tc1</i>	позвоночные	<i>RAG1, RAG2</i> (V(D)J рекомбинация)
	человек	<i>PAX6</i> (транскрипционный фактор регуляции развития органов)
<i>Transib</i>	челюстные позвоночные	<i>RAG1, RAG2</i> (V(D)J рекомбинация)

Ранее предполагалось, что обратная транскриптаза и подобные ей ферменты относятся исключительно к РЭ или вирусам, а к происходящим от обратной транскриптазы генам относились лишь теломеразы. Однако было выявлено, что от гена обратной транскриптазы РЭ произошли гены *rvt*, которые содержатся в геномах бактерий, протистов, грибов, животных и растений с неоднородным филогенетическим распределением [13]. Обратная транскриптаза ERV эволюционировала в белок Prp8, являющийся компонентом сплайсосомы эукариот [12]. У млекопитающих в эволюции были одомашнены гены *Env*, кодирующие оболочку ERV, с образованием генов синцитинов *Syncytin-1*, *-2*, *-A*, *-B*, важной функцией которых является регуляция развития плаценты [9, 15].

Белки, произошедшие от *Env* LTR-ретроэлементов, функционируют в противовирусной защите хозяев у позвоночных. Они образуют поверхностную (детерминирует клеточную

специфичность, тропизм к хозяину и типу клеток) и трансмембранную (необходима для слияния оболочек вируса и клетки-мишени) субъединицы. Во всех случаях эндогенные *env* гены, произошедшие от генов различных ERVs, действуют как факторы ограничения для родственных экзогенных ретровирусов [16]. От ORF1 LINE-ретроэлементов у человека и других млекопитающих в эволюции произошел ген *L1TD1*, который кодирует РНК-связывающий белок, функционирующий в недифференцированных клетках [19]. Доместикация генов МГЭ для нужд хозяев приводит к тому, что новые БКГ содержат последовательности транспозонов не только в своих интронах и регуляторных областях, но и в консервативных экзонах. Это дает основу для эпигенетического контроля работы БКГ на транскрипционном и посттранскрипционном уровнях с участием происходящих от транспозонов нкРНК в качестве универсальных инструментов РНК интерференции (рис. 3).

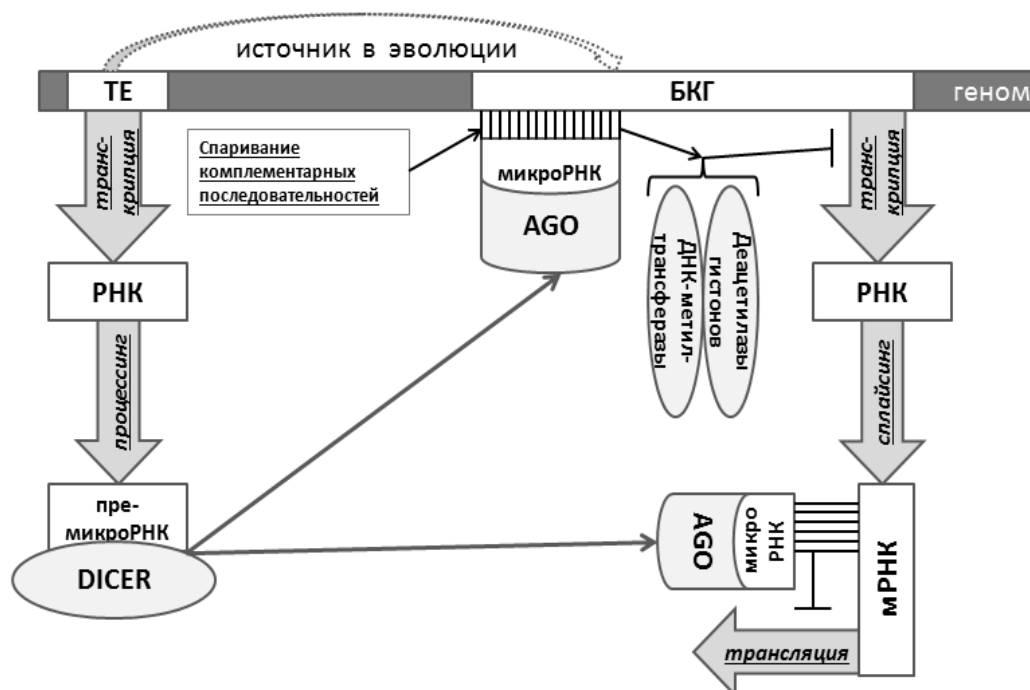


Рис. 3. Роль транспозонов в эпигенетическом контроле генов
Fig. 3. Role of transposable elements in epigenetic gene control

Происхождение некодирующих РНК от транспозонов

К некодирующим РНК относятся днРНК длиной более 200 нуклеотидов и малые нкРНК (короче 200 нуклеотидов, включая микроРНК длиной 20-24 нуклеотида). МикроРНК негативно регулируют их целевые мРНК путем связывания с их 3'-нетранслируемой областью и подавляя трансляцию или нарушая стабильность молекул. Биогенез микроРНК состоит из процессинга более крупных молекул при-микроРНК в более короткие пре-микроРНК с дальнейшим их процессингом до зрелых микроРНК. Одна микроРНК может регулировать множество мишеней, участвуя в разнообразных функциях организма. Например, онкосупрессорная miR-34a ингибирует экспрессию 700 отдельных БКГ [35]. Гены днРНК могут располагаться как между БКГ, так и внутри них. Названия многих днРНК происходят в соответствии с БКГ, вблизи или в которых они расположены. Если они комплементарны генам, то называются антисмысловыми днРНК или антисмысловыми транскриптами (NAT – natural antisense transcript). Различают cis-NAT (перекрываются с комплементарными генам) и trans-NAT (транскрибируются от псевдогенов) [36]. Экспрессия генов днРНК имеет много общего с БКГ, поскольку днРНК транскрибируются РНК-полимеразой II, подвергаются сплайсингу и экзпириванию [37].

Молекулы днРНК способны самостоятельно выполнять каталитические функции и называются рибозимами, однако в большинстве случаев они являются структурными компонентами рибонуклеопротеинов, играющих важные роли в клетках. Примерами являются рибосомы, представляющие собой сложные рибонуклеопротеидные комплексы. ДнРНК участвуют в регуляции транскрипции путем взаимодействия с ДНК-связывающими белками, с гистон-модифицирующими комплексами и РНК-полимеразой, за счет чего служат своеобразными драйверами

эпигенетической регуляции [38]. К таким днРНК относятся *Airn*, *ecCEBP*, *H19*, *Kcnq1ot1*, *PAPAS*, *pRNA*, *PTENpa1-AS*, *TARID*, *Xist*, которые вызывают образование метилцитозина в специфических локусах; *ANRASSF1*, *ANRIL*, *BORDERLINE*, *Kcnq1ot1*, *NeST*, *PINT* – являются гидами для факторов модификации гистонов. ДнРНК *AK141205*, *AK028326*, *ES1*, *ES2*, *ES3*, *linc-RoR*, *Evx1as*, *Hoxb5/6as* регулируют транскрипцию белок-кодирующих генов путем взаимодействий с факторами транскрипции. Кроме того, днРНК обладают сходными с микроРНК свойствами, влияя на экспрессию генов за счет ингибирования трансляции мРНК. К таким днРНК относятся *Uchl1-as1*, *lincMD1*, *lincRNA-p21*, $\frac{1}{2}$ -sbsRNA [37]. Некоторые днРНК (*lincRNA-Cox2*, *lincRNA-p21*) формируют рибонуклеопротеины, которые регулируют транскрипцию специфических наборов генов. Например, *lincRNA-p21* вместе с *hnRNP-K* образуют комплексы, связывающиеся со специфическими областями генома и подавляющие транскрипцию генов по пути p53 [36].

ДнРНК могут выполнять функции приманки для транскрипционных факторов путем мимикрии последовательности или структуры ДНК-мишеней. К таким днРНК относятся *gas5*, *PANDA*, *DHFR minor*, *Lethe*. Общее происхождение днРНК и микроРНК от МГЭ предполагает наличие у них идентичных и комплементарных последовательностей, благодаря чему реализуется важная функция днРНК в качестве губок для микроРНК. Эта способность позволяет ингибировать специфические микроРНК [36]. Действительно, согласно литературным данным, последовательности многих зрелых микроРНК [20] и днРНК [21] идентичны или комплементарны последовательностям МГЭ. Большинство (более 80%) генов днРНК содержат один и более фрагмент МГЭ, а почти половина всех экзонов днРНК имеют транспозонное происхождение. Доместикация

последовательностей МГЭ в эволюции обусловлена способностью РНК-структур МГЭ образовывать функциональные домены [22].

Одной из причин одомашнивания геномами хозяев генов МГЭ в качестве источников днРНК является функциональность транскриптов транспозонов [6], что связано с их древнейшими универсальными свойствами в качестве источников жизни на Земле [7]. Например, *Alu* встраиваются в гены днРНК и формируют структуры, необходимые для взаимодействия с мРНК за счет коротких несовершенных спариваний нуклеотидов. Регуляторными сигналами для днРНК обладают также LTR-содержащие РЭ. Например, в геноме человека выявлены тысячи генов днРНК, происходящие от LTR-РЭ [39], которые контролируют сеть плюрипотентности путем изменения структуры хроматина. Они участвуют в формировании бластоцисты и в дальнейшем эмбриогенезе [40]. Кроме того, LTR-РЭ могут непосредственно служить генами днРНК [41]. Транскрипты LINE1 также сами функционируют в качестве днРНК, взаимодействуя со специфическими участками хроматина и регулируя экспрессию генов в раннем эмбриогенезе. Например, при связывании с Nucleolin и KAP1, они вызывают как активацию генов рДНК, так и подавление многих генов двухклеточного эмбриона путем сайленсинга Dux [42]. То есть, помимо транспозиций с помощью белковых продуктов, подвергнутые альтернативному процессингу транскрипты РЭ способны функционировать самостоятельно в качестве молекул некодирующих РНК, что обусловлено их древним свойством, сформированном при возникновении жизни в мире РНК-ДНК [7]. Соответственно, взаимосвязи транспозонов с днРНК и их роль в возникновении новых БКГ могут отображать сходные процессы на ранних этапах эволюции живого, когда появилась преемственность функций рибозимов белковыми молекулами.

Тканевая специфичность днРНК превышает таковую для белков. При этом в регуляции дифференцировки стволовых клеток они взаимодействуют с МГЭ [43], что характерно для днРНК, образующихся из межгенных областей геномов эукариот, а также из перекрывающихся и антисмысловых паттернов относительно примыкающих БКГ, которые они и регулируют [44]. Это позволяет им в значительной степени детерминировать разнообразие клеточных фенотипов, особенно в нейронах центральной нервной системы [45]. Пространственные 3D-структуры молекулы днРНК характеризуются модульной организацией с образованием специфических доменов, которые состоят главным образом из последовательностей МГЭ [22]. Способность транскриптов МГЭ подвергаться процессингу с образованием мРНК или функциональных нкРНК с помощью рибонуклеаз, представляет собой эволюционный процесс приспособления генных сетей к меняющимся условиям при участии новых некодирующих РНК и их транскриптов [46, 47, 48]. В результате из наиболее оптимальных в естественном отборе генов нкРНК формируются новые БКГ [49-52]. Многофункциональность некодирующих РНК, происходящих от МГЭ, проявляется в том, что некоторые молекулы микроРНК могут образовываться из транскриптов днРНК [53]. Соответственно, такие днРНК являются также генами микроРНК, открытая рамка считывания прекурсоров которых может связываться с рибосомами с образованием функциональных белковых или пептидных молекул [54, 55, 56]. МГЭ также сами являются важными источниками генов микроРНК. Еще в 2016 году была опубликована база данных MDTE DB, в которую были включены 661 микроРНК человека, произошедшие от МГЭ [20].

Происхождение пептидов и белков от некодирующих РНК в эволюции

Возникновение днРНК и микроРНК от МГЭ в эволюции могло стать причиной многофункциональности данных нкРНК.

Это означает, что, несмотря на название «некодирующие», они обладают грандиозным потенциалом трансляции в функциональные белки и пептиды. Причиной является универсальное свойство МГЭ содержать последовательности нуклеотидов, связывающихся с рибосомами даже при наличии неканонической короткой открытой рамки считывания (ORF – open reading frame), что обусловлено ролью МГЭ в качестве универсальных эволюционных источников жизни на Земле [7]. В нескольких филогенетических исследованиях было выявлено происхождение эволюционно новых БКГ различных эукариот от генов днРНК [49-52]. Так, у гриба *Saccharomyces cerevisiae* выявлен новый ген *BSC4*, содержащий ORF (которая отсутствует у других близкородственных видов грибов), кодирующую белок длиной 132 аминокислот. Продукт гена *BSC4* участвует в путях восстановления ДНК в стационарную фазу гриба и способствует его устойчивости при перемещении в среду с низким содержанием питательных веществ. У грибов видов *S. paradoxus*, *S. mikatae*, *S. bayanus* идентичные некодирующие последовательности также транскрибируются, поэтому их можно отнести к категории гена днРНК, который у *Saccharomyces cerevisiae* в эволюции стал источником нового БКГ, доместифицированного в связи с адаптивными функциями [50]. Сравнительный анализ генома мухи *Drosophila melanogaster* с другими видами рода *Drosophila* позволил выявить 5 генов, 4 из которых расположены на X хромосоме и транслируются в пептиды длиной 58, 79, 97 и 158 аминокислот, 1 ген – на 2 хромосоме с кодированием пептида из 127 аминокислот. Данные гены произошли от некодирующих последовательностей ДНК, транскрибируемых у других видов (гены днРНК) [49].

В 2012 году были выявлены 24 эволюционно молодых БКГ, которые произошли в эволюции от генов днРНК. Из

них 11 генов кодируют специфичные только для человека белки, а 13 других – консервативные для человека и шимпанзе. Эти гены содержат от 1 до 7 экзонов, а длина продуктов трансляции генов варьирует от 72 до 423 аминокислот [51]. Такие гены называют «орфанными», то есть характерными для специфического вида животного и не обнаруживаемые у других видов, что связано с их недавним возникновением в эволюции из генов днРНК и транспозонов и участием в видоспецифических адаптивных реакциях. Около 53% «орфанных» генов приматов содержат последовательности МГЭ, что свидетельствует об использовании их в качестве источников происхождения данных генов [57]. Подтверждение роли днРНК в качестве источников «орфанных» генов было получено в исследованиях 2014 года, результаты которых показали, что связанные с рибосомами днРНК отличаются низкой эволюционной консервативностью и содержат гомологи у других видов от 0 до 15,6%, тогда как для БКГ этот показатель более 95% для позвоночных и 70-73% для растений и грибов. Анализ экспрессии днРНК в клетках 6 различных эукариот (*A. thaliana*, *S. cerevisiae*, *Danio rerio*, *D. melanogaster*, *Homo sapiens*, *Mus musculus*) видов показал аналогичный кодирующий потенциал и ограничения последовательностей с эволюционно молодыми белками. Полученные данные свидетельствуют о взаимосвязи с рибосомами значительной доли днРНК (от 28,6% у *S. cerevisiae* до 81,9% у мыши), что свидетельствовало об их трансляции. Было также обнаружено, что эволюционно молодые БКГ, кодирующие экспериментально проверенные белки, характеризовались общими с генами днРНК свойствами. Такие БКГ содержали короткую открытую рамку считывания, занимающую небольшую часть транскрипта, а также имели низкий показатель кодирования, аналогичный таковому для днРНК. Это говорит о происхождении таких эволюционно молодых БКГ от генов днРНК благодаря

отбору пептидов, участвующих в адаптивных для выживания организмов процессах [52].

Трансляция некодирующих РНК в функциональные пептиды и белки

Возникновение новых БКГ из генов днРНК у различных животных и растений обусловлено их трансляцией в пептиды, которые принимают участие в разнообразных биологических процессах. Данные о роли днРНК в качестве источников БКГ говорят о возможности образования из транскриптов их генов более крупных молекул полипептидов [49-52]. Это происходит в ходе эволюции образуемых генов с увеличением их размеров благодаря инсерциям МГЭ с их последующей экзонизацией [27, 28]. Помимо эволюционной роли в возникновении БКГ по данным филогенетических исследований, появляется все больше свидетельств непосредственной трансляции днРНК у эукариот. При этом механизмы возникновения более крупных БКГ из всей или большей части последовательности гена днРНК может отличаться от механизмов образования транслируемой молекулы РНК из днРНК, длина которой значительно меньше зрелой днРНК. Так, у мыши и человека обнаружен специфический для скелетных мышц пептид миорегулин (MLN), образуемый из ORF эволюционно консервативной днРНК, которая обозначена для человека как LINC00948, а для мыши как AK009351. У обоих организмов гены днРНК состоят из 3 экзонов и имеют длину более 15000 п.н. В то же время OFR, из которой транслируется пептид MLN, имеет длину лишь 138 п.н. и расположен в 3 экзоне генов. образуемый пептид регулирует прохождение через мембрану саркоплазматического ретикулула ионов кальция, тем самым управляя расслаблением мышц [46]. Сходным механизмом действия обладает пептид, DWORF, также транслируемый из днРНК и экспрессируемый в тканях миокарда [47]. У растений при трансляции ORF днРНК образуются регулирующие временем

цветения пептиды COLDAIR и COOLAIR, важный для симбиоза с бактериями пептид ENOD40 и управляющий поглощением фосфатов белок IPS1 [48].

Наибольший интерес представляют исследования трансляции днРНК человека, поскольку образуемые при этом пептиды специфически экспрессируются или подавляются в злокачественных новообразованиях и могут служить объектами для таргетной терапии опухолей (табл. 2). Поскольку МГЭ играют важную роль в возникновении, эволюции и регуляции днРНК, участие образуемых от нкРНК пептидов и белков подтверждает роль МГЭ в канцерогенезе [25]. Можно предположить, что одними из мишеней таких пептидов в опухолевых клетках являются МГЭ, что является предметом дальнейших исследований. Опубликованные в 2019 году результаты полноразмерного секвенирования транслируемых РНК и профилирования рибосом свидетельствуют о том, что 3330 днРНК человека связываются с рибосомами с активной элонгацией трансляции. Протеомный анализ позволил авторам обнаружить 308 новых белков, образуемых в результате трансляции днРНК [58]. Подобные исследования были проведены и другими исследовательскими группами, в которых были определены 128 пептидов, закодированных в последовательностях днРНК [59]. Следует отметить, что некоторые пептиды образуются при трансляции днРНК не в нормальных клетках организма, а только в опухолевых, что свидетельствует о механизме эволюции опухолей за счет образования новых БКГ. Это отражает общие эволюционные механизмы, когда отбираются возможные варианты трансляции нкРНК для формирования новых адаптивных свойств живых организмов. Например, днРНК LINC00675 транслируется с образованием малого консервативного белка FORCP из 79 аминокислот, который экспрессируется клетками колоректального рака, тогда как данный протеин не характерен для всех нормальных клеток человека [60].

Таблица 2

Пептиды человека, транслируемые из открытых рамок считывания днРНК

Table 2

Human peptides translated from lncRNA open reading frames

Название днРНК	Название/ длина пептида	Функция	Автор
CRNDE	CRNDEP/ 84	регулирует пролиферацию клеток парабазального слоя плоского эпителия, кишечных крипт и сперматоцитов	[63]
HOXB-AS3	HOXB-AS3/ 53	подавляет опосредованный hnRNP A1 сплайсинг пируваткиназы М, блокируя остатки аргинина в мотиве RGG hnRNP A1, ингибирует пролиферацию клеток рака ободочной кишки	[64]
LINC-PINT	PINT87aa/ 87	напрямую взаимодействует с комплексом факторов PAF1c, ассоциированных с полимеразой, ингибируя элонгацию транскрипции множества онкогенов и рост глиобластомы	[65]
LINC01420	NoBody/ 68	взаимодействует с белками, деапитирующими мРНК, которые удаляют 5'-кэп мРНК, способствуя распаду мРНК от 5' до 3', стимулирует пролиферацию клеток рака молочной железы и лейкоза	[66]
LINC-00266-1	RBRP/ 71	взаимодействует с РНК-связывающими белками, такими как m6A считыватель IGF2BP1, повышая стабильность и экспрессию с-Мус, способствуя канцерогенезу колоректального рака	[67]
LINC00278	YY1BM/ 21	ингибирует взаимодействие YY1 с рецептором андрогена, что снижает экспрессию eEF2K, способствуя апоптозу клеток рака пищевода	[62]
LINC00467	ASAP/94	стимулирует активность АТФ-синтазы путем взаимодействия с ATP5A и ATP5C, способствуя пролиферации клеток колоректального рака	[61]
LINC00665	CIP2A-BP/ 52	напрямую связывается с онкогеном CIP2A, активируя PP2A, ингибирующей путь PI3K/AKT/NFκB, что приводит к снижению экспрессии MMP-2, MMP-9 и Snail, подавляет пролиферацию клеток рака молочной железы	[68]
LINC00675	FORCP/ 79	регулирует апоптоз и канцерогенез в клетках колоректального рака	[60]
LINC00908	ASRPS/ 60	напрямую связывается с STAT3 через спиральный домен CCD и подавляет фосфорилирование STAT3, что приводит к снижению экспрессии VEGF; ингибирует развитие тройного негативного рака молочной железы	[69]
LOC90024	SRSP/ 130	взаимодействует с регуляторами сплайсинга SRSF3, стимулируя опухолевый сплайсинг РНК транскрипционного фактора Sp4, индуцирует развитие колоректального рака	[70]
LINC00998	SMIM30/ 59	стимулирует пролиферацию и миграцию клеток, способствует закреплению на мембране нерецепторных тирозинкиназ SRC/YES, активируя сигнальный путь MAPK и индуцируя гепатоцеллюлярную карциному	[71]
NCBP2-AS2	KRASIM/90	взаимодействует с белком KRAS, ингибируя сигнальные пути ERK, подавляя пролиферацию и клональную эволюцию клеток рака печени	[72]
NR029453	CASIMO1/ 80	взаимодействует со скваленэпоксидазой, регулируя кластеризацию липидных капель и пролиферацию клеток рака молочной железы	[73]
UBAP1-AST6	UBAP1-AST6	стимулирует размножение и формирование клонов клеток рака легкого	[58]

Для того же типа злокачественного новообразования выявлено образование из днРНК LINC00467 пептида ASAP, который

не только экспрессируется в нормальных клетках, но также является консервативным для высших млекопитающих, регулируя

функцию митохондрий [61]. Это свидетельствует о развитии различных путей эволюции злокачественных новообразований с использованием как вновь образуемых, так и консервативных пептидов, участвующих в канцерогенезе. Было показано также, что на экспрессию днРНК влияют изменения внутренней среды организма и внешней среды, что свидетельствует о потенциале исследований в данном направлении для выявления факторов риска злокачественных опухолей и возможности их коррекции. Так, на экспрессию пептида YY1BM влияют не только андрогены (в связи с чем он специфично транслируется у мужчин), но также сигаретный дым [62].

Некоторые днРНК являются также прекурсорами для микроРНК, поэтому их кодирующие области в ДНК одновременно являются генами днРНК и генами микроРНК. Более того, продукты транскрипции таких генов также обладают потенциалом транслироваться в функциональные пептиды. Примерами являются днРНК MIR497HG [56], днРНК MIR22HG [54], днРНК MIR155HG [55]. Транскрибируемые из генов микроРНК незрелые при-микроРНК также способны связываться с рибосомами и формировать функциональные пептиды, обозначаемые как miPEPs [74, 75] в связи с наличием в них коротких открытых рамок считывания (smORF), что позволяет им взаимодействовать с рибосомами [74]. Пептиды и белки, кодируемые при-микроРНК, экспрессируются в клетках растений и животных, регулируя рост и развитие нормальных и раковых клеток [76]. Пептиды miPEPs, которые образуются при трансляции при-микроРНК, участвуют в регуляции как БКГ, так и собственных генов микроРНК. Молекулы miPER характеризуются способностью участвовать в ряде важных биологических реакций у растений, в связи с чем могут быть использованы для улучшения свойств растений. Наилучшим примером является, miPER172c, который стимулирует трансляцию собственного гена miR172c. В

результате этого усиливается формирование узелков в корнях сои с индукцией симбиоза с азотфиксирующими бактериями [74].

SmORFs, обладающие потенциалом связываться с рибосомами, выявлены у 2% мРНК генов микроРНК растений [75]. Сходными свойствами обладают также процессированные при-микроРНК животных. Был идентифицирован функциональный пептид miPER31 у человека, образуемый при трансляции pri-miRNA-31, который подавляет экспрессию miR-31, усиливает индукцию регуляторных Т-лимфоцитов, действуя как репрессор транскрипции [77]. Согласно проведенному в 2023 году систематическому обзору научной литературы, в настоящее время известны miPER различных растений и животных, в том числе miPER-156a у *Brassica rapa*, miPER397a у *Brassica oleacera*, miPER-164b, miPER165a, miPER-397a, miPER858a у *Arabidopsis thaliana*, miPER164b у *Barbarea vulgaris*, miPER-164c, miPER-171d1, miPER-172d, miPER3635b у *Vitis vinifera*, miPER171a/b/c/d/e/f у *Medicago truncatula*, miPER171b у *Lotus japonicas*, miPER171i у *Oryza sativa*, miPER171e у *Solanum lycopersicum*, miPER172c у *Glycine max*, miPER-31 у мыши, miPER8 у дрозофилы [78]. Недавно описан также miPER408, участвующий в реакции на мышьяковый стресс и ассимиляцию серы у *Arabidopsis* [79]. Способностью к трансляции с образованием функциональных пептидов характеризуются не только днРНК и при-микроРНК, но также и другие молекулы известных РНК, такие как малые ядрышковые РНК [59]. Это говорит о грандиозном количестве различных пептидных молекул, которые еще не открыты, функционирующих как специфично для отдельных тканей или органов, так и для патологических процессов, таких как развитие злокачественных новообразований. То есть регуляторные системы геномов эукариот значительно сложнее сложившихся в генетике классических представлений. В таблице 3 содержатся данные об обнаруженных у человека miPER и их функции.

Таблица 3

Пептиды, кодируемые генами микроРНК (а также генами днРНК – предшественниками микроРНК) человека

Table 3

Peptides encoded by human microRNA genes (as well as lncRNA genes - microRNA precursors)

Название прекурсора	Название/длина пептида	Функция	Автор
lnc MIR7-3HG	miPEP7-3/127	защищает бета-клетки поджелудочной железы от вызванной DEX дисфункции путем активации пути PI3K/AKT	[80]
pri-miR-147b	MOCCI/83	регуляция противовирусных иммунных реакций	[81]
lnc MIR155HG	miPEP155/17	взаимодействует с аденозин-5'-трифосфат-связывающим доменом родственного белка теплового шока 70 (HSC70), шаперона, необходимого для транспортировки и презентации антигена в дендритных клетках	[55]
pri-miR-200a	miPEP-200a/187	регулируют экспрессию виментина, ингибируют эпителиально-мезенхимальный переход клеток рака простаты, подавляя их миграцию	[76]
pri-miR-200b	miPEP-200b/54		
lnc MIR22HG	miPEP22/57	противовирусные реакции организма	[54]
pri-miR-31	miPEP-31/48	способствует индукции регуляторных Т-лимфоцитов	[77]
pri-miR-34a	miPEP-133/133	в митохондриях взаимодействует с митохондриальным белком теплового шока, предотвращает взаимодействие HSPA9 с белками, усиливает транскрипционную активность p53 и стимулирует экспрессию miR-34a	[35]
lnc MIR497HG	miPEP497/21	способствует пролиферации опухолевых клеток, ингибирует развитие рака шейки матки	[56]

Функциональность транскриптов белок-кодирующих генов

Интроны БКГ также подвергаются процессингу с образованием функциональных нкРНК [82]. Это подтверждает их эволюционную взаимосвязь с МГЭ, которые являются ключевыми источниками нкРНК [20, 21, 22]. О роли МГЭ в возникновении как днРНК, так и БКГ свидетельствует общее свойство их транскриптов образовывать специфические вторичные и третичные структуры, которые содержат функциональные домены и влияют на собственный процессинг [83], что позволяет предположить их древнее происхождение от МГЭ. Обилие доменов днРНК транспозонными последовательностями связано с тем, что МГЭ образуют пространственные структуры, участвующие в биологических процессах [22]. Биоинформационный

анализ позволил установить способность транскриптов БКГ эукариот регулировать собственный сплайсинг, которая зависит от последовательностей нуклеотидов, произошедших в эволюции от МГЭ, в составе их интронов [83]. Данное свойство преемственно от МГЭ, молекулы РНК которых характеризуются формированием шпилек в пространстве, регулирующих распознавание сайтов сплайсинга [84].

О функциональности молекул мРНК белок-кодирующих генов свидетельствует образование из их интронов специфических кольцевых РНК (кРНК), обладающих множеством регуляторных способностей. Кроме того, кРНК также содержат короткие рамки считывания, связываются с рибосомами и транслируются в пептиды и белки. Кольцевые РНК широко экспрессируются в геномах различных эукариот, поскольку образуются путем обратного сплайсинга, что приводит к

образованию ковалентных структур РНК с замкнутой петлей без 3'-5'-полярности. Они не содержат полиадениновых хвостов и экспрессируются на относительно низких уровнях. Подобно днРНК, кРНК могут функционировать в качестве губок микроРНК, участвуют в регуляции развития нервной системы, в механизмах канцерогенеза, экспрессии БКГ [85]. Кроме того, кРНК связываются с промоторами целевых генов, рекрутируя факторы транскрипции, влияя тем самым на транскрипцию БКГ [86]. Наибольший интерес представляют пептиды, кодируемые кРНК у человека в связи с

обнаружением их вовлеченности в развитие злокачественных новообразований, поскольку они могут быть использованы в качестве мишеней для противоопухолевой терапии [85-88] (табл. 4). Однако одной из первых кРНК человека, способных к трансляции, стала *circ-ZNF609*, вовлеченная в регуляцию дифференцировки скелетной мускулатуры. В отличие от белкового продукта основного БКГ (транскрипционный фактор, белок цинковых пальцев), из которого образуется эта кРНК, продукт ее трансляции не содержит доменов цинковых пальцев [89].

Таблица 4

Пептиды, кодируемые кольцевыми РНК человека, участвующие в канцерогенезе

Table 4

Peptides encoded by human circular RNAs involved in carcinogenesis

Название кРНК	Название/длина пептида	Функция	Автор
circ-FBXW7	FBXW7-185aa/ 185	ингибирует пролиферацию и ускорение клеточного цикла, подавляет развитие глиобластомы	[85]
circ-Gprc5a	Circ-Gprc5a/ 11	в комбинации с мембранным белком Gprc5a активирует передачу сигналов GPCR, способствует развитию рака мочевого пузыря	[86]
circ-FNDC3B	circ-FNDC3B- 218aa/ 218	ингибирует экспрессию Snail, способствуя онкосупрессорному эффекту FBP1, подавляет пролиферацию, инвазию и миграцию клеток рака толстой кишки	[87]
circ-PPP1R12A	circ-PPP1R12A - 73aa/ 73	стимулирует пролиферацию, миграцию и инвазию клеток рака толстой кишки путем активации сигнальных путей Hippo-YAP	[88]
circ-β-catenin	β-catenin-370aa/ 370	активирует Wnt пути за счет стабилизации полноразмерного β-катенина, стимулирует развитие рака печени	[90]
circ-SHPRH	SHPRH-146aa/ 146	защищает полноразмерную <i>circ-SHPRH</i> от деградации убиквитинового протеасомой, убиквитинирует ядерный антиген пролиферирующих клеток, ингибирует развитие глиобластомы	[91]

Поскольку МГЭ в эволюции являлись источниками возникновения сплайсосомных интронов, а также их изменчивости [24], логично предположить, что способность кРНК транслироваться в функциональные пептиды также отражают универсальные свойства МГЭ и обусловлены возникновением кРНК от транспозонов. Действительно, кРНК, подобно днРНК, характеризуются тканеспецифической экспрессией, а также низкой консервативностью между видами [85-88], что говорит о динамичности их

эволюции с помощью перемещений МГЭ. Было показано, что среди всех комплементарных последовательностей, SINE (особенно Alu) человека вносят наибольший вклад в образование кРНК и их разнообразие. Соответственно, разное распределение МГЭ между видами приводит к увеличению сложности экспрессии кРНК при эволюции видов [92]. Это обусловлено ролью Alu в качестве источников акцепторов сплайсинга и ингибиторов трансляции мРНК. Alu являются основными мишенями фермента

ADAR, редактирующего РНК, а образование экзонов *Alu* подавляется рибонуклеопротеином HNRNPC. Было обнаружено, что ядерная РНК-хеликаза DHX9 специфически связывается с инвертированными повторами *Alu* в составе мРНК белок-кодирующих генов. Увеличение количества содержащих *Alu* генов, образующих кРНК, происходит при потере DHX9, который действует в качестве резольвазы ядерной РНК, нейтрализующей угрозу нарушений транскрипции и трансляции вследствие вставок *Alu* [93]. Было предложено, что кРНК длиной 100-1000 нуклеотидов, образующиеся при участии малых саморасщепляющихся РНК, которые обозначаются как рибозимы типа «молот», закодированы неавтономными РЭ ретрозимами [94]. Это новое семейство РЭ оказалось широко распространенным в геномах эукариот. Образующиеся в результате экспрессии этих РЭ транскрипты, содержащие мотивы типа «молот», легко обнаружить с помощью РНК-блоттинга [95]. Не исключено, что широкое распространение *Alu* в геномах эукариот (занимают 11% генома человека) обусловлено свойством образовывать кРНК, которые участвуют в регуляции экспрессии генов в ходе адаптивной эволюции организмов. Действительно, для формирования кРНК необходимо спаривание нуклеотидов РНК между последовательностями, фланкирующими сайты обратного сплайсинга, благодаря наличию *Alu* в составе интронов генов [96].

Заключение. В данной статье представлены данные, подтверждающие роль МГЭ в качестве универсальных единиц эволюции, свойства которых к противоборству с образованием систем процессинга стало основой многофункциональности их продуктов транскрипции и трансляции. Описана роль МГЭ в возникновении новых БКГ, генов микроРНК, днРНК и кРНК в эволюции. Представлены сведения об образовании 50 БКГ только от транспозазы ДНК-транспозонов в эволюции у различных

эукариот, в том числе 16 генов у человека, а также возникновение различных БКГ от генов ретроэлементов. Одним из преимуществ одомашнивания генов МГЭ геномами эукариот является вовлеченность образуемых мРНК в регуляторные адаптивные сети, а также функциональность самих молекул мРНК после их процессинга. Способность транскриптов генов днРНК и микроРНК к трансляции отражает свойство МГЭ образовывать одновременно функциональные РНК и их белковые продукты. Исследование данных свойств наиболее перспективно в современной онкологии, поскольку появляется все больше данных о вовлеченности, образуемых из днРНК, при-микроРНК и кРНК пептидов в канцерогенезе. Сами молекулы этих нкРНК могут быть использованы в качестве мишеней для воздействия, а их пептидные продукты трансляции как инструменты. В статье представлены данные об образовании 15 участвующих в канцерогенезе специфических функциональных пептидов в результате трансляции днРНК, 4 пептидов – из при-микроРНК и 6 пептидов – из кРНК.

Информация о финансировании

Финансирование данной работы не проводилось.

Financial support

No financial support has been provided for this work.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Conflict of interests

The authors have no conflict of interest to declare.

Список литературы

1. Makalowski W, Gotea V, Pande A, et al. Transposable Elements: Classification, Identification, and Their Use As a Tool For Comparative Genomics. In: Anisimova M, editor. Evolutionary Genomics. Methods in Molecular Biology, vol 1910. Humana, New York, NY; 2019.

DOI: https://doi.org/10.1007/978-1-4939-9074-0_6

2. Roberts AP, Chandler M, Courvalin P, et al. Revised Nomenclature for Transposable Genetic Elements. *Plasmid*. 2008;60(3):167-173. DOI:

<https://doi.org/10.1016/j.plasmid.2008.08.001>

3. Elliott TA, Gregory TR. Do larger genomes contain more diverse transposable elements. *BMC Evolutionary Biology*. 2015;15:69. DOI: <https://doi.org/10.1186/s12862-015-0339-8>

4. Nurk S, Koren S, Rhie A, et al. The complete sequence of a human genome. *Science*. 2022;376(6588):44-53. DOI: <https://doi.org/10.1126/science.abj6987>

5. De Koning APJ, Gu W, Castoe TA, et al. Repetitive Elements May Comprise Over Two-Thirds of the Human Genome. *PLoS Genetics*. 2011;7(12):e1002384. DOI: <https://doi.org/10.1371/journal.pgen.1002384>

6. Mustafin RN. Functional dualism of transcripts of transposons in the evolution of eukaryotic genomes. *Russian Journal of Developmental Biology*. 2018;49(4S1):339-355. DOI: <https://doi.org/10.1134/S1062360418070019>

7. Mustafin RN, Khusnutdinova EK. The Role of Reverse Transcriptase in the Origin of Life. *Biochemistry*. 2019;84(8):870-883. DOI: <https://doi.org/10.1134/S0006297919080030>

8. Kapitonov VV, Jurka J. Harbinger Transposons and an Ancient HARBI1 Gene Derived From a Transposase. *DNA and Cell Biology*. 2004;23(5):311. DOI: <https://doi.org/10.1089/104454904323090949>

9. Volff JN. Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. *Bioessays*. 2006;28(9):913-922. DOI: <https://doi.org/10.1002/bies.20452>

10. Feschotte C. Transposable elements and the evolution of regulatory networks. *Nature Reviews Genetics*. 2008;9(5):397-405. DOI: <https://doi.org/10.1038/nrg2337>

11. Sinzelle L, Izsvak Z, Ivics Z. Molecular domestication of transposable elements: From detrimental parasites to useful host genes. *Cellular and Molecular Life Sciences*. 2009;66:1073-1093. DOI: <https://doi.org/10.1007/s00018-009-8376-3>

12. Dlakic M, Mushegian A. Prp8, the Pivotal Protein of the Splisosomal Catalytic Center, Evolved From a Retroelement – Encoded Reverse Transcriptase. *RNA*. 2011;17(5):799-808. DOI: <https://doi.org/10.1261/rna.2396011>

13. Gladyshev EA, Arkhipova IR. A widespread class of reverse transcriptase-related cellular genes. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108(51):20311-20316. DOI: <https://doi.org/10.1073/pnas.1100266108>

14. Abrusan G, Zhang Y, Szilagyi A. Structure prediction and analysis of DNA transposon and LINE retrotransposons proteins. *Journal of Biological Chemistry*. 2013;288(22):16127-16138. DOI: <https://doi.org/10.1074/jbc.M113.451500>

15. Alzohairy AM, Gyulai G, Jansen RK, et al. Transposable elements domesticated and neofunctionalized by eukaryotic genomes. *Plasmid*. 2013;69(1):1-15. DOI: <https://doi.org/10.1016/j.plasmid.2012.08.001>

16. Malfavon-Borja R, Feschotte C. Fighting Fire with Fire: Endogenous Retrovirus Envelopes as Restriction Factors. *Journal of Virology*. 2015;89(8):4047-4050. DOI: <https://doi.org/10.1128/JVI.03653-14>

17. Wang J, Vicente-Garcia C, Seruggia D, et al. MIR retrotransposons sequences provide insulators to the human genome. *Proceedings of the National Academy of Sciences of the United States of America*. 2015;112(32):4428-4437. DOI: <https://doi.org/10.1073/pnas.1507253112>

18. Duan CG, Wang X, Xie S, et al. A pair of transposon-derived proteins function in a histone acetyltransferase complex for active DNA demethylation. *Cell Research*. 2017;27(2):226-240. DOI: <https://doi.org/10.1038/cr.2016.147>

19. Joly-Lopez Z, Bureau TE. Exaptation of transposable element coding sequences. *Current Opinion in Genetics and Development*. 2018;49:34-42. DOI: <https://doi.org/10.1016/j.gde.2018.02.011>

20. Wei G, Qin S, Li W, et al. MDTE DB: a database for microRNAs derived from Transposable element. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. 2016;13(6):1155-1160. DOI: <https://doi.org/10.1109/TCBB.2015.2511767>

21. Kapusta A, Kronenberg Z, Lynch VJ, et al. Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. *PLoS Genet*. 2013;9:e1003470. DOI: <https://doi.org/10.1371/journal.pgen.1003470>

22. Johnson R, Guigo R. The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs. *RNA*.

- 2014;20:959-976. DOI: <https://doi.org/10.1261/rna.044560.114>
23. Tan S, Cardoso-Moreira M, Shi W, et al. LTR-mediated retroposition as a mechanism of RNA-based duplication in metazoans. *Genome Research*. 2016;26(12):1663-1675. DOI: <https://doi.org/10.1101/gr.204925.116>
24. Kubiak MR, Makalowska I. Protein-Coding Genes' Retrocopies and Their Functions. *Viruses*. 2017;9(4):80. DOI: <https://doi.org/10.3390/v9040080>
25. Мустафин РН, Хуснутдинова ЭК. Ретроэлементы как мишени таргетной терапии опухолей (обзор). *Научные результаты биомедицинских исследований*. 2024;10(1):5-22. DOI: <https://doi.org/10.18413/2658-6533-2024-10-1-0-1>
26. Schrader L, Schmitz J. The Impact of Transposable Elements in Adaptive Evolution. *Molecular Ecology*. 2019;28(6):1537-1549. DOI: <https://doi.org/10.1111/mec.14794>
27. Abascal F, Tress ML, Valencia A. Alternative splicing and co-option of transposable elements: the case of TMPO/LAP2 α and ZNF451 in mammals. *Bioinformatics*. 2015;31(14):2257-2261. DOI: <https://doi.org/10.1093/bioinformatics/btv132>
28. He Z, Chen O, Phillips N, et al. Predicting Alu exonization in the human genome with a deep learning model. *BioRxiv*. 2024;8:727537. DOI: <https://doi.org/10.1101/2024.01.03.574099>
29. Zhu Z, Tan S, Zhang Y, et al. LINE-1-like retrotransposons contribute to RNA-based gene duplication in dicots. *Scientific Reports*. 2016;6:24755. DOI: <https://doi.org/10.1038/srep24755>
30. Sakai H, Mizuno H, Kawahara Y, et al. Retrogenes in rice (*Oryza sativa* L. ssp. *japonica*) exhibit correlated expression with their source genes. *Genome Biology and Evolution*. 2011;3:1357-1368. DOI: <https://doi.org/10.1093/gbe/evr111>
31. Du Z, Yang C, Rothschild MF, et al. Novel microRNA families expanded in the human genome. *BMC Genomics*. 2013;14:98. DOI: <https://doi.org/10.1186/1471-2164-14-98>
32. Hoen DR, Buraeu TE. Discovery of Novel Genes Derived from Transposable Elements Using Integrative Genomic Analysis. *Molecular Biology and Evolution*. 2015;32(6):1487-1506. DOI: <https://doi.org/10.1093/molbev/msv042>
33. Zdobnov EM, Campillos M, Harrington ED, et al. Protein coding potential of retroviruses and other transposable elements in vertebrate genomes. *Nucleic Acids Research*. 2005;33(1):946-954. DOI: <https://doi.org/10.1093/nar/gki236>
34. De Souza FS, Franchini LF, Rubinstein M. Exaptation of transposable elements into novel cis-regulatory elements: is the evidence always strong. *Molecular Biology and Evolution*. 2013;30(6):1239-1251. DOI: <https://doi.org/10.1093/molbev/mst045>
35. Kang M, Tang B, Li J, et al. Identification of miPEP133 as a novel tumor-suppressor microprotein encoded by miR-34a pri-miRNA. *Molecular Cancer*. 2020;19:143. DOI: <https://doi.org/10.1186/s12943-020-01248-9>
36. Fitzgerald KA, Caffrey DR. Long noncoding RNAs in innate and adaptive immunity. *Current Opinion in Immunology*. 2014;26:140-146. DOI: <https://doi.org/10.1016/j.coi.2013.12.001>
37. Fico A, Fiorenzano A, Pascale E, et al. Long non-coding RNA in stem cell pluripotency and lineage commitment: functions and evolutionary conservation. *Cellular and Molecular Life Sciences*. 2019;76:1459-1471. DOI: <https://doi.org/10.1007/s00018-018-3000-z>
38. Long Y, Wang X, Youmans DT, et al. How do lncRNAs regulate transcription. *Science Advances*. 2017;3(9):2110. DOI: <https://doi.org/10.1126/sciadv.aao2110>
39. Hadjiargyrou M, Delihias N. The Intertwining of Transposable Elements and Non-Coding RNAs. *International Journal of Molecular Sciences*. 2013;14(7):13307-13328. DOI: <https://doi.org/10.3390/ijms140713307>
40. Gerdes P, Richardson SR, Mager DL, et al. Transposable elements in the mammalian embryo: pioneers surviving through stealth and service. *Genome Biology*. 2016;17:100. DOI: <https://doi.org/10.1186/s13059-016-0965-5>
41. Lu X, Sachs F, Ramsay L, et al. The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. *Nature Structural and Molecular Biology*. 2014;21:423-425. DOI: <https://doi.org/10.1038/nsmb.2799>
42. Honson DD, Macfarlan TS. A lncRNA-like Role for LINE1s in Development. *Developmental Cell*. 2018;46(2):132-134. DOI: <https://doi.org/10.1016/j.devcel.2018.06.022>
43. Ramsay LA, Marchetto MC, Caron M, et al. Conserved expression of transposon-derived non-coding transcripts in primate stem cells. *BMC*

- Genomics. 2017;18:214. DOI: <https://doi.org/10.1186/s12864-017-3568-y>
44. Arendt T, Ueberham U, Janitz M. Non-coding transcriptome in brain aging. *Aging*. 2017;9(9):1943-1944. DOI: <https://doi.org/10.18632/aging.101290>
45. Lapp HE, Hunter RG. The dynamic genome: transposons and environmental adaptation in the nervous system. *Epigenomics*. 2016;8(2):237-249. DOI: <https://doi.org/10.2217/epi.15.107>
46. Anderson DM, Anderson KM, Cang CL, et al. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell*. 2015;160(4):595-606. DOI: <https://doi.org/10.1016/j.cell.2015.01.009>
47. Nelson BR, Makarewich CA, Anderson DM, et al. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science*. 2016;351(6270):271-275. DOI: <https://doi.org/10.1126/science.aad4076>
48. Zhang J, Mujahid H, Hou Y, et al. Plant Long ncRNAs: A New Frontier for Gene Regulatory Control. *American Journal of Plant Sciences*. 2013;4(5):32139. DOI: <https://doi.org/10.4236/ajps.2013.45128>
49. Levine MT, Jones CD, Kern AD, et al. Novel genes derived from noncoding DNA in *Drosophila melanogaster* are frequently X-linked and exhibit testis-biased expression. *Proceedings of the National Academy of Sciences of the United States of America*. 2006;103(26):9935-9939. DOI: <https://doi.org/10.1073/pnas.0509809103>
50. Cai J, Zhao R, Jiang H, et al. De novo origination of a new protein-coding gene in *Saccharomyces cerevisiae*. *Genetics*. 2008;179(1):487-496. DOI: <https://doi.org/10.1534/genetics.107.084491>
51. Xie C, Zhang YE, Chen JY, et al. Hominoid-specific de novo protein-coding genes originating from long non-coding RNAs. *PLoS Genetics*. 2012;8:e1002942. DOI: <https://doi.org/10.1371/journal.pgen.1002942>
52. Ruiz-Orera J, Messeguer X, Subirana JA, et al. Long non-coding RNAs as a source of new peptide. *eLife*. 2014;3:e03523. DOI: <https://doi.org/10.7554/eLife.03523>
53. Guo L, Zhao Y, Yang S, et al. An integrated evolutionary analysis of miRNA-lncRNA in mammals. *Molecular Biology Reports*. 2014;41:201-207. DOI: <https://doi.org/10.1007/s11033-013-2852-4>
54. Razooky BS, Obermayer B, O'May JB, et al. Viral Infection Identifies Micropeptides Differentially Regulated in smORF-Containing lncRNAs. *Genes*. 2017;8(8):206. DOI: <https://doi.org/10.3390/genes8080206>
55. Niu L, Lou F, Sun Y, et al. A micropeptide encoded by lncRNA MIR155HG suppresses autoimmune inflammation via modulating antigen presentation. *Science Advances*. 2020;6(21):eaaz2059. DOI: <https://doi.org/10.1126/sciadv.aaz2059>
56. Prel A, Dozier C, Combier JP, et al. Evidence that regulation of Pri-miRNA/miRNA expression is not a general rule of miPEPs function in humans. *International Journal of Molecular Sciences*. 2021;22(7):3432. DOI: <https://doi.org/10.3390/ijms22073432>
57. Toll-Riera M, Bosch N, Bellora N, et al. Origin of primate orphan genes: a comparative genomics approach. *Molecular Biology and Evolution*. 2009;26(3):603-612. DOI: <https://doi.org/10.1093/molbev/msn281>
58. Lu S, Zhang J, Lian X, et al. A hidden human proteome encoded by 'non-coding' genes. *Nucleic Acids Research*. 2019;47(15):8111-8125. DOI: <https://doi.org/10.1093/nar/gkz646>
59. van Heesch S, Witte F, Schneider-Lunitz V, et al. The Translational Landscape of the Human Heart. *Cell*. 2019;178(1):242-260. DOI: <https://doi.org/10.1016/j.cell.2019.05.010>
60. Li XL, Pongor L, Tang W, et al. A Small Protein Encoded by a Putative lncRNA Regulates Apoptosis and Tumorigenicity in Human Colorectal Cancer Cells. *eLife*. 2020;9:e53734. DOI: <https://doi.org/10.7554/eLife.53734>
61. Ge Q, Jia D, Cen D, et al. Micropeptide ASAP encoded by LINC00467 promotes colorectal cancer progression by directly modulating ATP synthase activity. *Journal of Clinical Investigation*. 2021;131(22):e152911. DOI: <https://doi.org/10.1172/JCI152911>
62. Wu S, Zhang L, Deng J, et al. A Novel Micropeptide Encoded by Y-Linked LINC00278 Links Cigarette Smoking and AR Signaling in Male Esophageal Squamous Cell Carcinoma. *Cancer Research*. 2020;80(13):2790-2803. DOI: <https://doi.org/10.1158/0008-5472.can-19-3440>
63. Szafron LM, Balcerak A, Grzybowska EA, et al. The novel gene journal pre-proof 25 CRNDE encodes a nuclear peptide (CRNDEP) which is overexpressed in highly proliferating tissues. *PLoS ONE*. 2015;10(5):e0127475. DOI: <https://doi.org/10.1371/journal.pone.0127475>

64. Huang JZ, Chen M, Chen D, et al. A peptide encoded by a putative lncRNA HOXB-AS3 suppresses colon cancer growth. *Molecular Cell*. 2017;68(1):171-184. DOI: <https://doi.org/10.1016/j.molcel.2017.09.015>
65. Zhang M, Zhao K, Xu X, et al. A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma. *Nature Communications*. 2018;9:4475. DOI: <https://doi.org/10.1038/s41467-018-06862-2>
66. D'Lima NG, Ma J, Winkler L, et al. A human microprotein that interacts with the mRNA decapping complex. *Nature Chemical Biology*. 2017;13(2):174-180. DOI: <https://doi.org/10.1038/nchembio.2249>
67. Zhu S, Wang JZ, Chen D, et al. An oncopeptide regulates m(6)A recognition by the m(6)A reader IGF2BP1 and tumorigenesis. *Nature Communications*. 2020;11:1685. DOI: <https://doi.org/10.1038/s41467-020-15403-9>
68. Guo B, Wu S, Zhu X, et al. Micropeptide CIP2A-BP encoded by LINC00665 inhibits triple-negative breast cancer progression. *EMBO Journal*. 2020;39(1):e102190. DOI: <https://doi.org/10.15252/embj.2019102190>
69. Wang Y, Wu S, Zhu X, et al. LncRNA-encoded polypeptide ASRPS inhibits triple-negative breast cancer angiogenesis. *Journal of Experimental Medicine*. 2020;217(3):jem.20190950. DOI: <https://doi.org/10.1084/jem.20190950>
70. Meng N, Chen M, Chen D, et al. Small protein hidden in lncRNA LOC90024 promotes "cancerous" RNA splicing and tumorigenesis. *Advanced Science*. 2020;7(10):1903233. DOI: <https://doi.org/10.1002/advs.201903233>
71. Pang Y, Liu Z, Han H, et al. Peptide SMIM30 promotes HCC development by inducing SRC/YES1 membrane anchoring and MAPK pathway activation. *Journal of Hepatology*. 2020;73(5):1155-1169. DOI: <https://doi.org/10.1016/j.jhep.2020.05.028>
72. Xu W, Deng B, Lin P, et al. Ribosome profiling analysis identified a KRAS-interacting microprotein that represses oncogenic signaling in hepatocellular carcinoma cells. *Science China Life Sciences*. 2020;63(4):529-542. DOI: <https://doi.org/10.1007/s11427-019-9580-5>
73. Polycarpou-Schwarz M, Gross M, Mestdagh P, et al. The cancer-associated microprotein CASIMO1 controls cell proliferation and interacts with squalene epoxidase modulating lipid droplet formation. *Oncogene*. 2018;37:4750-4768. DOI: <https://doi.org/10.1038/s41388-018-0281-5>
74. Couzigou JM, Andre O, Guillotin B, et al. Use of microRNA-encoded peptide miPEP172c to stimulate nodulation in soybean. *New Phytologist*. 2016;211(2):379-381. DOI: <https://doi.org/10.1111/nph.13991>
75. Laressergues D, Couzigou JM, Clemente HS, et al. Primary transcripts of microRNAs encode regulatory peptides. *Nature*. 2015;520(7545):90-93. DOI: <https://doi.org/10.1038/nature14346>
76. Fang J, Morsalin S, Rao VN, et al. Decoding of non-coding DNA and non-coding RNA: pri-micro RNA-encoded novel peptides regulate migration of cancer cells. *Journal of Pharmaceutical Sciences and Pharmacology*. 2017;3:23-27. DOI: <https://doi.org/10.1166/jpsp.2017.1070>
77. Zhou H, Lou F, Bai J, et al. A peptide encoded by pri-miRNA-31 represses autoimmunity by promoting Treg differentiation. *EMBO Reports*. 2022;23(5):e53475. DOI: <https://doi.org/10.15252/embr.202153475>
78. Ormancey M, Thuleau P, Combier JP, et al. The Essentials on microRNA-Encoded Peptides from Plants to Animals. *Biomolecules*. 2023;13(2):206. DOI: <https://doi.org/10.3390/biom13020206>
79. Kumar RS, Sinha H, Datta T, et al. microRNA408 and its encoded peptide regulate sulfur assimilation and arsenic stress response in Arabidopsis. *Plant Physiology*. 2023;192(2):837-856. DOI: <https://doi.org/10.1093/plphys/kiad033>
80. Mao X, Zhou J, Kong L, et al. A peptide encoded by lncRNA MIR7-3 host gene (MIR7-3HG) alleviates dexamethasone-induced dysfunction in pancreatic β -cells through the PI3K/AKT signaling pathway. *Biochemical and Biophysical Research Communications*. 2023;647:62-71. DOI: <https://doi.org/10.1016/j.bbrc.2023.01.004>
81. Lee CQE, Kerouanton B, Chothani S, et al. Coding and non-coding roles of MOCCI (C15ORF48) coordinate to regulate host inflammation and immunity. *Nature Communications*. 2021;12:2130. DOI: <https://doi.org/10.1038/s41467-021-22397-5>
82. Rearick D, Prakash A, McSweeney A, et al. Critical association of ncRNA with introns. *Nucleic Acids Research*. 2011;39(6):2357-2366. DOI: <https://doi.org/10.1093/nar/gkq1080>
83. Soemedi R, Cygan KJ, Rhine CL, et al. The effects of structure on pre-mRNA processing

and stability. *Methods*. 2017;125:36-44. DOI: <https://doi.org/10.1016/j.ymeth.2017.06.001>

84. Kralovicova J, Patel A, Searle M, et al. The role of short RNA loops in recognition of a single-hairpin exon derived from a mammalian-wide interspersed repeat. *RNA Biology*. 2015;12(1):54-69. DOI: <https://doi.org/10.1080/15476286.2015.1017207>

85. Yang Y, Gao X, Zhang M, et al. Novel role of FBXW7 Circular RNA in repressing glioma tumorigenesis. *Journal of the National Cancer Institute*. 2018;110(3):304-315. DOI: <https://doi.org/10.1093/jnci/djx166>

86. Gu C, Zhou N, Wang Z, et al. circGprc5a promoted bladder oncogenesis and metastasis through Gprc5a-targeting peptide. *Molecular Therapy - Nucleic Acids*. 2018;13:633-641. DOI: <https://doi.org/10.1016/j.omtn.2018.10.008>

87. Pan Z, Cai J, Lin J, et al. A novel protein encoded by circFNDC3B inhibits tumor progression and EMT through regulating Snail in colon cancer. *Molecular Cancer*. 2020;19:71. DOI: <https://doi.org/10.1186/s12943-020-01179-5>

88. Zheng X, Chen L, Zhou Y, et al. A novel protein encoded by a circular RNA circPPP1R12A promotes tumor pathogenesis and metastasis of colon cancer via Hippo-YAP signaling. *Molecular Cancer*. 2019;18:47. DOI: <https://doi.org/10.1186/s12943-019-1010-6>

89. Legnini I, Di Timoteo G, Rossi F, et al. Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis. *Molecular Cell*. 2017;66(1):22-37. DOI: <https://doi.org/10.1016/j.molcel.2017.02.017>

90. Liang WC, Wong CW, Liang PP, et al. Translation of the circular RNA circbeta-catenin promotes liver cancer cell growth through activation of the Wnt pathway. *Genome Biology*. 2019;20:84. DOI: <https://doi.org/10.1186/s13059-019-1685-4>

91. Zhang M, Huang N, Yang X, et al. A novel protein encoded by the circular form of the SHPRH gene suppresses glioma tumorigenesis. *Oncogene*. 2018;37(13):1805-1814. DOI: <https://doi.org/10.1038/s41388-017-0019-9>

92. Dong R, Ma XK, Chen LL, et al. Increased complexity of circRNA expression during species evolution. *RNA Biology*. 2017;14(8):1064-1074. DOI: <https://doi.org/10.1080/15476286.2016.1269999>

93. Aktaş T, Avşar Ilık İ, Maticzka D, et al. DHX9 suppresses RNA processing defects originating from the Alu invasion of the human

genome. *Nature*. 2017;544(7648):115-119. DOI: <https://doi.org/10.1038/nature21715>

94. de la Peña M. Circular RNAs Biogenesis in Eukaryotes Through Self-Cleaving Hammerhead Ribozymes. In: Xiao J, editor. *Circular RNAs. Advances in Experimental Medicine and Biology*, vol 1087. Springer, Singapore; 2018. DOI: https://doi.org/10.1007/978-981-13-1426-1_5

95. Cervera A, de la Peña M. Cloning and Detection of Genomic Retrozymes and Their circRNA Intermediates. In: Scarborough RJ, Gatignol A, editors. *Ribozymes. Methods in Molecular Biology*, vol 2167. Humana, New York, NY; 2021. DOI: https://doi.org/10.1007/978-1-0716-0716-9_3

96. Welden JR, Stamm S. Pre-mRNA structures forming circular RNAs. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*. 2019;1862(11-12):194410. DOI: <https://doi.org/10.1016/j.bbagr.2019.194410>

References

1. Makalowski W, Gotea V, Pande A, et al. Transposable Elements: Classification, Identification, and Their Use As a Tool For Comparative Genomics. In: Anisimova M, editor. *Evolutionary Genomics. Methods in Molecular Biology*, vol 1910. Humana, New York, NY; 2019. DOI: https://doi.org/10.1007/978-1-4939-9074-0_6

2. Roberts AP, Chandler M, Courvalin P, et al. Revised Nomenclature for Transposable Genetic Elements. *Plasmid*. 2008;60(3):167-173. DOI: <https://doi.org/10.1016/j.plasmid.2008.08.001>

3. Elliott TA, Gregory TR. Do larger genomes contain more diverse transposable elements. *BMC Evolutionary Biology*. 2015;15:69. DOI: <https://doi.org/10.1186/s12862-015-0339-8>

4. Nurk S, Koren S, Rhie A, et al. The complete sequence of a human genome. *Science*. 2022;376(6588):44-53. DOI: <https://doi.org/10.1126/science.abj6987>

5. De Koning APJ, Gu W, Castoe TA, et al. Repetitive Elements May Comprise Over Two-Thirds of the Human Genome. *PLoS Genetics*. 2011;7(12):e1002384. DOI: <https://doi.org/10.1371/journal.pgen.1002384>

6. Mustafin RN. Functional dualism of transcripts of transposons in the evolution of eukaryotic genomes. *Russian Journal of Developmental Biology*. 2018;49(4S1):339-355. DOI: <https://doi.org/10.1134/S1062360418070019>

7. Mustafin RN, Khusnutdinova EK. The Role of Reverse Transcriptase in the Origin of Life. *Biochemistry*. 2019;84(8):870-883. DOI: <https://doi.org/10.1134/S0006297919080030>
8. Kapitonov VV, Jurka J. Harbinger Transposons and an Ancient HARBI1 Gene Derived From a Transposase. *DNA and Cell Biology*. 2004;23(5):311. DOI: <https://doi.org/10.1089/104454904323090949>
9. Volff JN. Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. *Bioessays*. 2006;28(9):913-922. DOI: <https://doi.org/10.1002/bies.20452>
10. Feschotte C. Transposable elements and the evolution of regulatory networks. *Nature Reviews Genetics*. 2008;9(5):397-405. DOI: <https://doi.org/10.1038/nrg2337>
11. Sinzelle L, Izsvak Z, Ivics Z. Molecular domestication of transposable elements: From detrimental parasites to useful host genes. *Cellular and Molecular Life Sciences*. 2009;66:1073-1093. DOI: <https://doi.org/10.1007/s00018-009-8376-3>
12. Dlakic M, Mushegian A. Prp8, the Pivotal Protein of the Spliseosomal Catalytic Center, Evolved From a Retroelement – Encoded Reverse Transcriptase. *RNA*. 2011;17(5):799-808. DOI: <https://doi.org/10.1261/rna.2396011>
13. Gladyshev EA, Arkhipova IR. A widespread class of reverse transcriptase-related cellular genes. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108(51):20311-20316. DOI: <https://doi.org/10.1073/pnas.1100266108>
14. Abrusan G, Zhang Y, Szilagyi A. Structure prediction and analysis of DNA transposon and LINE retrotransposons proteins. *Journal of Biological Chemistry*. 2013;288(22):16127-16138. DOI: <https://doi.org/10.1074/jbc.M113.451500>
15. Alzohairy AM, Gyulai G, Jansen RK, et al. Transposable elements domesticated and neofunctionalized by eukaryotic genomes. *Plasmid*. 2013;69(1):1-15. DOI: <https://doi.org/10.1016/j.plasmid.2012.08.001>
16. Malfavon-Borja R, Feschotte C. Fighting Fire with Fire: Endogenous Retrovirus Envelopes as Restriction Factors. *Journal of Virology*. 2015;89(8):4047-4050. DOI: <https://doi.org/10.1128/JVI.03653-14>
17. Wang J, Vicente-Garcia C, Seruggia D, et al. MIR retrotransposons sequences provide insulators to the human genome. *Proceedings of the National Academy of Sciences of the United States of America*. 2015;112(32):4428-4437. DOI: <https://doi.org/10.1073/pnas.1507253112>
18. Duan CG, Wang X, Xie S, et al. A pair of transposon-derived proteins function in a histone acetyltransferase complex for active DNA demethylation. *Cell Research*. 2017;27(2):226-240. DOI: <https://doi.org/10.1038/cr.2016.147>
19. Joly-Lopez Z, Bureau TE. Exaptation of transposable element coding sequences. *Current Opinion in Genetics and Development*. 2018;49:34-42. DOI: <https://doi.org/10.1016/j.gde.2018.02.011>
20. Wei G, Qin S, Li W, et al. MDTE DB: a database for microRNAs derived from Transposable element. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. 2016;13(6):1155-1160. DOI: <https://doi.org/10.1109/TCBB.2015.2511767>
21. Kapusta A, Kronenberg Z, Lynch VJ, et al. Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. *PLoS Genet*. 2013;9:e1003470. DOI: <https://doi.org/10.1371/journal.pgen.1003470>
22. Johnson R, Guigo R. The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs. *RNA*. 2014;20:959-976. DOI: <https://doi.org/10.1261/rna.044560.114>
23. Tan S, Cardoso-Moreira M, Shi W, et al. LTR-mediated retroposition as a mechanism of RNA-based duplication in metazoans. *Genome Research*. 2016;26(12):1663-1675. DOI: <https://doi.org/10.1101/gr.204925.116>
24. Kubiak MR, Makalowska I. Protein-Coding Genes' Retrocopies and Their Functions. *Viruses*. 2017;9(4):80. DOI: <https://doi.org/10.3390/v9040080>
25. Mustafin RN, Khusnutdinova EK. Retroelements in targeted antitumor therapy (review). *Research Results in Biomedicine*. 2024;10(1):5-22. Russian. DOI: <https://doi.org/10.18413/2658-6533-2024-10-1-0-1>
26. Schrader L, Schmitz J. The Impact of Transposable Elements in Adaptive Evolution. *Molecular Ecology*. 2019;28(6):1537-1549. DOI: <https://doi.org/10.1111/mec.14794>
27. Abascal F, Tress ML, Valencia A. Alternative splicing and co-option of transposable elements: the case of TMPO/LAP2 α and ZNF451 in mammals. *Bioinformatics*. 2015;31(14):2257-2261. DOI: <https://doi.org/10.1093/bioinformatics/btv132>

28. He Z, Chen O, Phillips N, et al. Predicting Alu exonization in the human genome with a deep learning model. *BioRxiv*. 2024;8:727537. DOI: <https://doi.org/10.1101/2024.01.03.574099>
29. Zhu Z, Tan S, Zhang Y, et al. LINE-1-like retrotransposons contribute to RNA-based gene duplication in dicots. *Scientific Reports*. 2016;6:24755. DOI: <https://doi.org/10.1038/srep24755>
30. Sakai H, Mizuno H, Kawahara Y, et al. Retrogenes in rice (*Oryza sativa* L. ssp. *japonica*) exhibit correlated expression with their source genes. *Genome Biology and Evolution*. 2011;3:1357-1368. DOI: <https://doi.org/10.1093/gbe/evr111>
31. Du Z, Yang C, Rothschild MF, et al. Novel microRNA families expanded in the human genome. *BMC Genomics*. 2013;14:98. DOI: <https://doi.org/10.1186/1471-2164-14-98>
32. Hoen DR, Bureau TE. Discovery of Novel Genes Derived from Transposable Elements Using Integrative Genomic Analysis. *Molecular Biology and Evolution*. 2015;32(6):1487-1506. DOI: <https://doi.org/10.1093/molbev/msv042>
33. Zdobnov EM, Campillos M, Harrington ED, et al. Protein coding potential of retroviruses and other transposable elements in vertebrate genomes. *Nucleic Acids Research*. 2005;33(1):946-954. DOI: <https://doi.org/10.1093/nar/gki236>
34. De Souza FS, Franchini LF, Rubinstein M. Exaptation of transposable elements into novel cis-regulatory elements: is the evidence always strong. *Molecular Biology and Evolution*. 2013;30(6):1239-1251. DOI: <https://doi.org/10.1093/molbev/mst045>
35. Kang M, Tang B, Li J, et al. Identification of miPEP133 as a novel tumor-suppressor microprotein encoded by miR-34a pri-miRNA. *Molecular Cancer*. 2020;19:143. DOI: <https://doi.org/10.1186/s12943-020-01248-9>
36. Fitzgerald KA, Caffrey DR. Long noncoding RNAs in innate and adaptive immunity. *Current Opinion in Immunology*. 2014;26:140-146. DOI: <https://doi.org/10.1016/j.coi.2013.12.001>
37. Fico A, Fiorenzano A, Pascale E, et al. Long non-coding RNA in stem cell pluripotency and lineage commitment: functions and evolutionary conservation. *Cellular and Molecular Life Sciences*. 2019;76:1459-1471. DOI: <https://doi.org/10.1007/s00018-018-3000-z>
38. Long Y, Wang X, Youmans DT, et al. How do lncRNAs regulate transcription. *Science Advances*. 2017;3(9):2110. DOI: <https://doi.org/10.1126/sciadv.aao2110>
39. Hadjiargyrou M, Delihias N. The Intertwining of Transposable Elements and Non-Coding RNAs. *International Journal of Molecular Sciences*. 2013;14(7):13307-13328. DOI: <https://doi.org/10.3390/ijms140713307>
40. Gerdes P, Richardson SR, Mager DL, et al. Transposable elements in the mammalian embryo: pioneers surviving through stealth and service. *Genome Biology*. 2016;17:100. DOI: <https://doi.org/10.1186/s13059-016-0965-5>
41. Lu X, Sachs F, Ramsay L, et al. The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. *Nature Structural and Molecular Biology*. 2014;21:423-425. DOI: <https://doi.org/10.1038/nsmb.2799>
42. Honson DD, Macfarlan TS. A lncRNA-like Role for LINE1s in Development. *Developmental Cell*. 2018;46(2):132-134. DOI: <https://doi.org/10.1016/j.devcel.2018.06.022>
43. Ramsay LA, Marchetto MC, Caron M, et al. Conserved expression of transposon-derived non-coding transcripts in primate stem cells. *BMC Genomics*. 2017;18:214. DOI: <https://doi.org/10.1186/s12864-017-3568-y>
44. Arendt T, Ueberham U, Janitz M. Non-coding transcriptome in brain aging. *Aging*. 2017;9(9):1943-1944. DOI: <https://doi.org/10.18632/aging.101290>
45. Lapp HE, Hunter RG. The dynamic genome: transposons and environmental adaptation in the nervous system. *Epigenomics*. 2016;8(2):237-249. DOI: <https://doi.org/10.2217/epi.15.107>
46. Anderson DM, Anderson KM, Cang CL, et al. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell*. 2015;160(4):595-606. DOI: <https://doi.org/10.1016/j.cell.2015.01.009>
47. Nelson BR, Makarewich CA, Anderson DM, et al. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science*. 2016;351(6270):271-275. DOI: <https://doi.org/10.1126/science.aad4076>
48. Zhang J, Mujahid H, Hou Y, et al. Plant Long ncRNAs: A New Frontier for Gene Regulatory Control. *American Journal of Plant Sciences*. 2013;4(5):32139. DOI: <https://doi.org/10.4236/ajps.2013.45128>
49. Levine MT, Jones CD, Kern AD, et al. Novel genes derived from noncoding DNA in *Drosophila melanogaster* are frequently X-linked

and exhibit testis-biased expression. *Proceedings of the National Academy of Sciences of the United States of America*. 2006;103(26):9935-9939. DOI: <https://doi.org/10.1073/pnas.0509809103>

50. Cai J, Zhao R, Jiang H, et al. De novo origination of a new protein-coding gene in *Saccharomyces cerevisiae*. *Genetics*. 2008;179(1):487-496. DOI: <https://doi.org/10.1534/genetics.107.084491>

51. Xie C, Zhang YE, Chen JY, et al. Hominoid-specific de novo protein-coding genes originating from long non-coding RNAs. *PLoS Genetics*. 2012;8:e1002942. DOI: <https://doi.org/10.1371/journal.pgen.1002942>

52. Ruiz-Orera J, Messeguer X, Subirana JA, et al. Long non-coding RNAs as a source of new peptide. *eLife*. 2014;3:e03523. DOI: <https://doi.org/10.7554/eLife.03523>

53. Guo L, Zhao Y, Yang S, et al. An integrated evolutionary analysis of miRNA-lncRNA in mammals. *Molecular Biology Reports*. 2014;41:201-207. DOI: <https://doi.org/10.1007/s11033-013-2852-4>

54. Razooky BS, Obermayer B, O'May JB, et al. Viral Infection Identifies Micropeptides Differentially Regulated in smORF-Containing lncRNAs. *Genes*. 2017;8(8):206. DOI: <https://doi.org/10.3390/genes8080206>

55. Niu L, Lou F, Sun Y, et al. A micropeptide encoded by lncRNA MIR155HG suppresses autoimmune inflammation via modulating antigen presentation. *Science Advances*. 2020;6(21):eaaz2059. DOI: <https://doi.org/10.1126/sciadv.aaz2059>

56. Prel A, Dozier C, Combier JP, et al. Evidence that regulation of Pri-miRNA/miRNA expression is not a general rule of miPEPs function in humans. *International Journal of Molecular Sciences*. 2021;22(7):3432. DOI: <https://doi.org/10.3390/ijms22073432>

57. Toll-Riera M, Bosch N, Bellora N, et al. Origin of primate orphan genes: a comparative genomics approach. *Molecular Biology and Evolution*. 2009;26(3):603-612. DOI: <https://doi.org/10.1093/molbev/msn281>

58. Lu S, Zhang J, Lian X, et al. A hidden human proteome encoded by 'non-coding' genes. *Nucleic Acids Research*. 2019;47(15):8111-8125. DOI: <https://doi.org/10.1093/nar/gkz646>

59. van Heesch S, Witte F, Schneider-Lunitz V, et al. The Translational Landscape of the Human Heart. *Cell*. 2019;178(1):242-260. DOI: <https://doi.org/10.1016/j.cell.2019.05.010>

60. Li XL, Pongor L, Tang W, et al. A Small Protein Encoded by a Putative lncRNA

Regulates Apoptosis and Tumorigenicity in Human Colorectal Cancer Cells. *eLife*. 2020;9:e53734. DOI: <https://doi.org/10.7554/eLife.53734>

61. Ge Q, Jia D, Cen D, et al. Micropeptide ASAP encoded by LINC00467 promotes colorectal cancer progression by directly modulating ATP synthase activity. *Journal of Clinical Investigation*. 2021;131(22):e152911. DOI: <https://doi.org/10.1172/JCI152911>

62. Wu S, Zhang L, Deng J, et al. A Novel Micropeptide Encoded by Y-Linked LINC00278 Links Cigarette Smoking and AR Signaling in Male Esophageal Squamous Cell Carcinoma. *Cancer Research*. 2020;80(13):2790-2803. DOI: <https://doi.org/10.1158/0008-5472.can-19-3440>

63. Szafron LM, Balcerak A, Grzybowska EA, et al. The novel gene journal pre-proof 25 CRNDE encodes a nuclear peptide (CRNDEP) which is overexpressed in highly proliferating tissues. *PLoS ONE*. 2015;10(5):e0127475. DOI: <https://doi.org/10.1371/journal.pone.0127475>

64. Huang JZ, Chen M, Chen D, et al. A peptide encoded by a putative lncRNA HOXB-AS3 suppresses colon cancer growth. *Molecular Cell*. 2017;68(1):171-184. DOI: <https://doi.org/10.1016/j.molcel.2017.09.015>

65. Zhang M, Zhao K, Xu X, et al. A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma. *Nature Communications*. 2018;9:4475. DOI: <https://doi.org/10.1038/s41467-018-06862-2>

66. D'Lima NG, Ma J, Winkler L, et al. A human microprotein that interacts with the mRNA decapping complex. *Nature Chemical Biology*. 2017;13(2):174-180. DOI: <https://doi.org/10.1038/nchembio.2249>

67. Zhu S, Wang JZ, Chen D, et al. An oncopeptide regulates m(6)A recognition by the m(6)A reader IGF2BP1 and tumorigenesis. *Nature Communications*. 2020;11:1685. DOI: <https://doi.org/10.1038/s41467-020-15403-9>

68. Guo B, Wu S, Zhu X, et al. Micropeptide CIP2A-BP encoded by LINC00665 inhibits triple-negative breast cancer progression. *EMBO Journal*. 2020;39(1):e102190. DOI: <https://doi.org/10.15252/embj.2019102190>

69. Wang Y, Wu S, Zhu X, et al. lncRNA-encoded polypeptide ASRPS inhibits triple-negative breast cancer angiogenesis. *Journal of Experimental Medicine*. 2020;217(3):jem.20190950. DOI: <https://doi.org/10.1084/jem.20190950>

70. Meng N, Chen M, Chen D, et al. Small protein hidden in lncRNA LOC90024 promotes “cancerous” RNA splicing and tumorigenesis. *Advanced Science*. 2020;7(10):1903233. DOI: <https://doi.org/10.1002/advs.201903233>
71. Pang Y, Liu Z, Han H, et al. Peptide SMIM30 promotes HCC development by inducing SRC/YES1 membrane anchoring and MAPK pathway activation. *Journal of Hepatology*. 2020;73(5):1155-1169. DOI: <https://doi.org/10.1016/j.jhep.2020.05.028>
72. Xu W, Deng B, Lin P, et al. Ribosome profiling analysis identified a KRAS-interacting microprotein that represses oncogenic signaling in hepatocellular carcinoma cells. *Science China Life Sciences*. 2020;63(4):529-542. DOI: <https://doi.org/10.1007/s11427-019-9580-5>
73. Polycarpou-Schwarz M, Gross M, Mestdagh P, et al. The cancer-associated microprotein CASIMO1 controls cell proliferation and interacts with squalene epoxidase modulating lipid droplet formation. *Oncogene*. 2018;37:4750-4768. DOI: <https://doi.org/10.1038/s41388-018-0281-5>
74. Couzigou JM, Andre O, Guillotin B, et al. Use of microRNA-encoded peptide miPEP172c to stimulate nodulation in soybean. *New Phytologist*. 2016;211(2):379-381. DOI: <https://doi.org/10.1111/nph.13991>
75. Lauressergues D, Couzigou JM, Clemente HS, et al. Primary transcripts of microRNAs encode regulatory peptides. *Nature*. 2015;520(7545):90-93. DOI: <https://doi.org/10.1038/nature14346>
76. Fang J, Morsalin S, Rao VN, et al. Decoding of non-coding DNA and non-coding RNA: pri-micro RNA-encoded novel peptides regulate migration of cancer cells. *Journal of Pharmaceutical Sciences and Pharmacology*. 2017;3:23-27. DOI: <https://doi.org/10.1166/jpsp.2017.1070>
77. Zhou H, Lou F, Bai J, et al. A peptide encoded by pri-miRNA-31 represses autoimmunity by promoting Treg differentiation. *EMBO Reports*. 2022;23(5):e53475. DOI: <https://doi.org/10.15252/embr.202153475>
78. Ormancey M, Thuleau P, Combier JP, et al. The Essentials on microRNA-Encoded Peptides from Plants to Animals. *Biomolecules*. 2023;13(2):206. DOI: <https://doi.org/10.3390/biom13020206>
79. Kumar RS, Sinha H, Datta T, et al. microRNA408 and its encoded peptide regulate sulfur assimilation and arsenic stress response in *Arabidopsis*. *Plant Physiology*. 2023;192(2):837-856. DOI: <https://doi.org/10.1093/plphys/kiad033>
80. Mao X, Zhou J, Kong L, et al. A peptide encoded by lncRNA MIR7-3 host gene (MIR7-3HG) alleviates dexamethasone-induced dysfunction in pancreatic β -cells through the PI3K/AKT signaling pathway. *Biochemical and Biophysical Research Communications*. 2023;647:62-71. DOI: <https://doi.org/10.1016/j.bbrc.2023.01.004>
81. Lee CQE, Kerouanton B, Chothani S, et al. Coding and non-coding roles of MOCCI (C15ORF48) coordinate to regulate host inflammation and immunity. *Nature Communications*. 2021;12:2130. DOI: <https://doi.org/10.1038/s41467-021-22397-5>
82. Rearick D, Prakash A, McSweeney A, et al. Critical association of ncRNA with introns. *Nucleic Acids Research*. 2011;39(6):2357-2366. DOI: <https://doi.org/10.1093/nar/gkq1080>
83. Soemedi R, Cygan KJ, Rhine CL, et al. The effects of structure on pre-mRNA processing and stability. *Methods*. 2017;125:36-44. DOI: <https://doi.org/10.1016/j.ymeth.2017.06.001>
84. Kralovicova J, Patel A, Searle M, et al. The role of short RNA loops in recognition of a single-hairpin exon derived from a mammalian-wide interspersed repeat. *RNA Biology*. 2015;12(1):54-69. DOI: <https://doi.org/10.1080/15476286.2015.1017207>
85. Yang Y, Gao X, Zhang M, et al. Novel role of FBXW7 Circular RNA in repressing glioma tumorigenesis. *Journal of the National Cancer Institute*. 2018;110(3):304-315. DOI: <https://doi.org/10.1093/jnci/djx166>
86. Gu C, Zhou N, Wang Z, et al. circGprc5a promoted bladder oncogenesis and metastasis through Gprc5a-targeting peptide. *Molecular Therapy - Nucleic Acids*. 2018;13:633-641. DOI: <https://doi.org/10.1016/j.omtn.2018.10.008>
87. Pan Z, Cai J, Lin J, et al. A novel protein encoded by circFNDC3B inhibits tumor progression and EMT through regulating Snail in colon cancer. *Molecular Cancer*. 2020;19:71. DOI: <https://doi.org/10.1186/s12943-020-01179-5>
88. Zheng X, Chen L, Zhou Y, et al. A novel protein encoded by a circular RNA circPPP1R12A promotes tumor pathogenesis and metastasis of colon cancer via Hippo-YAP signaling. *Molecular Cancer*. 2019;18:47. DOI: <https://doi.org/10.1186/s12943-019-1010-6>
89. Legnini I, Di Timoteo G, Rossi F, et al. Circ-ZNF609 Is a Circular RNA that Can Be

Translated and Functions in Myogenesis. *Molecular Cell*. 2017;66(1):22-37. DOI: <https://doi.org/10.1016/j.molcel.2017.02.017>

90. Liang WC, Wong CW, Liang PP, et al. Translation of the circular RNA circbeta-catenin promotes liver cancer cell growth through activation of the Wnt pathway. *Genome Biology*. 2019;20:84. DOI: <https://doi.org/10.1186/s13059-019-1685-4>

91. Zhang M, Huang N, Yang X, et al. A novel protein encoded by the circular form of the SHPRH gene suppresses glioma tumorigenesis. *Oncogene*. 2018;37(13):1805-1814. DOI: <https://doi.org/10.1038/s41388-017-0019-9>

92. Dong R, Ma XK, Chen LL, et al. Increased complexity of circRNA expression during species evolution. *RNA Biology*. 2017;14(8):1064-1074. DOI: <https://doi.org/10.1080/15476286.2016.1269999>

93. Aktaş T, Avşar Ilık İ, Maticzka D, et al. DHX9 suppresses RNA processing defects originating from the Alu invasion of the human genome. *Nature*. 2017;544(7648):115-119. DOI: <https://doi.org/10.1038/nature21715>

94. de la Peña M. Circular RNAs Biogenesis in Eukaryotes Through Self-Cleaving Hammerhead Ribozymes. In: Xiao J, editor. *Circular RNAs. Advances in Experimental Medicine and Biology*, vol 1087. Springer, Singapore; 2018. DOI: https://doi.org/10.1007/978-981-13-1426-1_5

95. Cervera A, de la Peña M. Cloning and Detection of Genomic Retrozymes and Their circRNA Intermediates. In: Scarborough RJ, Gatignol A, editors. *Ribozymes. Methods in Molecular Biology*, vol 2167. Humana, New York, NY; 2021. DOI: https://doi.org/10.1007/978-1-0716-0716-9_3

96. Welden JR, Stamm S. Pre-mRNA structures forming circular RNAs. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*. 2019;1862(11-12):194410. DOI: <https://doi.org/10.1016/j.bbagr.2019.194410>

Статья поступила в редакцию 5 апреля 2024 г.
Поступила после доработки 1 июня 2024 г.
Принята к печати 12 сентября 2024 г.

Received 5 April 2024

Revised 1 June 2024

Accepted 12 September 2024

Информация об авторах

Рустам Наилевич Мустафин, кандидат биологических наук, доцент кафедры медицинской генетики и фундаментальной медицины ФГБОУ ВО «Башкирский государственный медицинский университет», г. Уфа, Российская Федерация, E-mail: ruji79@mail.ru, ORCID: <https://orcid.org/0000-0002-4091-382X>.

Эльза Камилевна Хуснутдинова, доктор биологических наук, профессор, член-корреспондент Российской академии образования, академик Академии Наук Республики Башкортостан, директор Института биохимии и генетики – обособленного структурного подразделения ФГБНУ Уфимский федеральный исследовательский центр Российской академии наук; заведующий кафедрой генетики и фундаментальной медицины ФГБОУ ВО «Уфимский университет науки и технологий», г. Уфа, Российская Федерация, E-mail: elzakh@mail.ru, ORCID: <https://orcid.org/0000-0003-2987-3334>.

Information about the authors

Rustam N. Mustafin, Cand. Sci. (Biology), Associate Professor of the Department of Medical Genetics and Fundamental Medicine, Bashkir State Medical University, Ufa, Russia, E-mail: ruji79@mail.ru, ORCID: <https://orcid.org/0000-0002-4091-382X>.

Elza K. Khusnutdinova, Doct. Sci. (Biology), Professor, Corresponding Member of the Russian Academy of Education, Academician of the Academy of Sciences of the Republic of Bashkortostan, Director of the Institute of Biochemistry and Genetics, Ufa Scientific Center of RAS; Head of the Department of Genetics and Fundamental Medicine, Ufa University of Science and Technology, Ufa, Russia, E-mail: elzakh@mail.ru, ORCID: <https://orcid.org/0000-0003-2987-3334>.