== ОБЗОРЫ ==

Разнообразие некодирующих РНК в геномах эукариот

Назипова Н.Н.

Институт математических проблем биологии РАН – филиал Института прикладной математики им. М.В. Келдыша РАН

Аннотация. Геномы крупных многоклеточных эукариот в основном состоят из ДНК, кодирующей не белки, а РНК. Неожиданное открытие примерно одинакового количества генов белков у Homo sapiens и Caenorhabditis elegans привело к пониманию, что не количество белков определяет сложность развития и функционирования организма. Феномен всепроникающей транскрипции геномов находит все больше подтверждений. Появляются данные о новых видах РНК, которые работают в разных компартментах клетки, экспрессируются на разных стадиях развития, в разных тканях и выполняют разнообразные функции, главной из которых является тонкая регуляция основного клеточного процесса – трансляции белков. Наличие богатого арсенала регуляторов, которые могут взаимодействовать друг с другом и работать на принципе взаимозаменяемости определяет физиологическую сложность организма и его способность к адаптации к изменяющимся условиям окружающей среды. Здесь представлен обзор известных на настоящий момент функциональных РНК, экспрессирующихся в геномах эукариот. Нет сомнений, что в ближайшее время с применением высокотехнологичных транскриптомных технологий будет выявлено и охарактеризовано много новых РНК, но вполне вероятно, что многие из экспрессирующихся транскриптов не имеют выраженной функции, а являются эволюционным резервом организмов.

Ключевые слова: РНК-интерференция, микроРНК, малые интерферирующие РНК, мяРНК, мякРНК, пиРНК, длинные некодирующине РНК, кольцевые РНК, тРНК-фрагменты, рРНК-фрагменты транспозонов.

введение

До недавнего времени считалось, что главное содержание генома – это гены, кодирующие белки в процессе трансляции матричной РНК (мРНК). Результат секвенирования генома человека стал для многих сюрпризом, тогда было определено, что в нем нашлось только около 20 тысяч генов, кодирующих белок, что составляет меньше 2% от общей геномной последовательности [1, 2]. Поскольку другие менее сложные эукариоты, например, нематода *Caenorhabditis elegans*, имеют почти такое же количество генов, кодирующих белок, стало ясно, что сложность развития и физиологическая сложность человека не может быть объяснена только количеством генов, кодирующих белок. Механизмы альтернативного сплайсинга кодирующих белок транскриптов, а также посттрансляционные модификации синтезируемых белков, безусловно, увеличивают функциональное разнообразие протеома. Это частично объясняет повышенную сложность эукариотических геномов по срвавнению с прокариотическими. Остальная часть генома долгое время при этом считалась мусорной, или эгоистичной.

Бурное развитие высокопроизводительных технологий секвенирования и компьютерного исследования геномов вызвало рост числа работ, касающихся возможной функциональной нагрузки на другую часть генома, которая не кодирует

белки и поэтому названа некодирующей. Эта часть занимает около 98 % генома человека. Оказалось, что транскрипции подвержены не только области, кодирующие белки, но почти весь геном млекопитающих [3]. Это было продемонстрировано крупномасштабными проектами по клонированию комплементарных ДНК [4, 5] и исследованиями с применением массивов геномных плиток [3, 6, 7]. Так был открыт феномен всепроникающей транскрипции (pervasive transcription) [6, 8, 9], в результате которой более 90% генома человека, вероятно, будет транскрибироваться хотя бы в одном типе клеток, хотя бы на одном этапе онтогенеза [6, 10, 11]. В результате получается сложная сеть перекрывающихся транскриптов, которая включает десятки тысяч длинных РНК с небольшой способностью кодировать белок или вовсе без нее [12]. До сих пор ведутся споры о том, представляет ли пул жизнеспособных транскриптов транскрипционный «шум» [13–16] или это некодирующие РНК (нкРНК), которые имеют функции, которые просто еще не идентифицированы [17].

Уже более 60-ти лет, с той поры, когда в 50-е годы прошлого века был открыт механизм трансляции белков, известны первые некодирующие РНК – это транспортные РНК (тРНК) и рибосомальные РНК (рРНК). Они обеспечивают синтез белка (экспрессии гена). В конце 70-х – начале 80-х годов были открыты первые представители основных современных классов некодирующих РНК. Оказалось, что, в отличие от матричной РНК (мРНК), некодирующие РНК не транслируются в белки, но выполняют важные клеточные функции либо сами по себе, либо в комплексе с белками [18–23]. Функции некодирующих РНК варьируются от обеспечения процессинга РНК, модификации, регуляции транскрипции, стабильности и трансляции мРНК до секреции белка [19].

Было обнаружено и биохимически выделено несколько малых РНК, отличных от мРНК, pPHK и тPHK, в том числе богатые уридином U-PHK [24–27]. Названные малыми ядерными PHK (мяPHK) U1, U2, U4, U5 и U6, оказались компонентами сплайсосомы, участвующими в сплайсинге мPHK [28, 29]. Другие U-PHK - U4atac, U6atac, U11 и U12 - оказались компонентами сплайсосом второго типа [30, 31].

Неожиданные открытия расширили набор возможных функций отдельных некодирующих РНК. Например, было обнаружено, что биохимические фракции содержат РНК с каталитическими свойствами, которые принимают участие в метаболизме транспортных РНК, которые назвали РНКазами Р [32]. А обнаружение 7SL РНК [33] – длинной некодирующей РНК, входящей в состав эукариотической частицы узнавания сигнала (signal recognition particle, SRP) – привело к тому, что считавшаяся до этого белком, SRP получила статус цитоплазматического рибопротеинового комплекса, который участвует в котрансляционном транспорте белков [33, 34].

Размеры многих известных некодирующих РНК обычно намного меньше размеров мРНК и колеблются от двух десятков нуклеотидов у микроРНК [35, 36] до примерно 500 нуклеотидов (например, теломеразная РНК [37]. Кроме того, наблюдались также очень большие нкРНК, включая Xist РНК человека длиной 17000 нуклеотидов [38–40] или мышиную РНК Air длиной 108 тысяч нуклеотидов [41].

С тех пор количество новых и предполагаемых функциональных некодирующих РНК значительно увеличилось [20, 42] и процесс идентификации и характеристики нкРНК продолжается с многократно увеличившейся скоростью. Высокопроизводительные транскриптомные технологии [43] и развитие биоинформатических методов и программных средств привели к тому, что мир постоянно узнает о существовании новых видов РНК и их функциях. В начале 90-х годов 20-го века были открыты микроРНК [44–46], а в конце 90-х – малые интерферирующие РНК (миРНК) [47] и РНК, взаимодействующиес белками РІШІ (пиРНК) [48].

В 2013 году было обнаружено существование кольцевых некодирующих РНК [49– 51]. Недавние исследования значительно расширили список известных РНК открытием некодирующих регуляторных фрагментов, генерируемых из РНК с ранее известными и записанными в учебники функциями, такими как транспортные РНК (тРНК) [52, 53], рибосомальные РНК [54-56], а также транпозоны [57]. В таблице 1 приведена классификация основных типов некодирующих РНК, известных на сегодняшний день.

Таблица 1. Классификация некодирующих РНК

Тип	Название	Функцио- нальная длина, нукл.	Ссылки
Структурные некодирующие РНК			
Рибосомные РНК	5S pPHK	120	[58, 59]
	5,8S pPHK	150	
	18S pPHK	1800	
	28S pPHK	4000–5000	
Транспортные РНК	тРНК	70–100	[60]
Малые ядерные РНК	мяРНК	100-200	[61]
Малые ядрышковые РНК	мякРНК	80–1000	[62]
Регуляторные некодирующие РНК			
Малые некодирующие РНК	микроРНК	22	[46]
	миРНК	20–23	[63]
	пиРНК	24–32	[64]
	фрагменты тРНК	14–50	[65]
	фрагменты рРНК	16–25	[66]
	фрагменты транспозонов	21	[57]
Длинные некодирующие РНК	длинные межгенные РНК	200-100000	[67]
	кольцевые РНК	500	[68]

СТРУКТУРНЫЕ НЕКОДИРУЮЩИЕ РНК

Рибосомальные РНК

Рибосомы эукариот могут содержать до 80 рибосомных белков и до четырех структурных рибосомных РНК, названных в соответствии с их размерами: это 5S (~ 120 нуклеотидов), 5,8S (~ 150 нуклеотидов), 18S (~ 1800 нуклеотидов) и 28S (~ 4000-5000 нуклеотидов) [58]. Молекулы 5.8S-, 18S- и 28S-рРНК транскрибируются, как один оперон, 5S-рРНК находится в отдельном локусе. 18S и 5.8S разделены внутренним транскрибируемым спейсером (ITS1), а 5.8S и 28S разделены другим спейсером ITS2. Исходный транскрипт расщепляется на три функциональные РНК путем удаления ITS1 и ITS2 [59].

Рибосома - это древняя молекулярная машина, ответственная за трансляцию генетического кода, которая имеет консервативную центральную вторичную структуру рибосомной РНК, она универсально присуща всем царствам жизни. Однако в ходе

эволюции размер рибосомы значительно увеличился. Например, эукариотическая рибосома более чем на 1 МДа больше, чем бактериальная рибосома, отчасти из-за вставки до 30 специфических сегментов расширения (expansion segments) в рРНК [69]. Интересно, что сегменты расширения демонстрируют высокую изменчивость с точки зрения длины и последовательности у разных видов и, по-видимому, с ними не взаимодействую рибосомальные белки [70].

Вторичная структура сегментов расширения недостаточно исследована, у человека они GC-богаты и на электронно-микроскопических снимках похожи на гибкие структуры, напоминающие щупальца [71]. Можно резюмировать, что длина pPHK у эукариот увеличена из-за вставок сегментов расширения, а тот факт, они выступают из рибосомы, говорит о том, что они, вероятно, доступны для связывания. Хотя человек и дрозофила содержат одинаковый набор сегментов расширения, что и дрожжи, и простейшие, их сегменты расширения, как правило, намного длиннее. Например, сравнивая длины сегментов расширения ES3S, ES7L, ES9L, ES15L, ES27L и ES39L у дрожжей (110, 200, 70, 20, 160 и 140 нуклеотидов) и человека (160, 870, 110, 190, 710 и 240 нуклеотидов соответственно) [71], можно заметить два самых длиных из них, ES7L и ES27L, они являются самыми длинными для обоих организмов.

и точность синтеза белка являются основополагающими Скорость для функционирования живых клеток. Высококонсервативная рРНК и видоспецифические вставки сегментов расширения, очевидно, обеспечивают механизм управления качеством трансляции. Эти сегменты эволюционировали, чтобы соответствовать специфическим потребностям синтеза белка в каждом организме. Как рибосома, которая является окаменелостью мира РНК, эволюционировала, чтобы приспособиться к такой видоспецифической адаптации, включая повышение точности трансляции от прокариот к эукариотам, остается пока недостаточно изученным явлением. В работе [72] показано, что вариабельные сегменты расширения рибосомы играют важную регуляторную роль и служат в качестве платформ для связывания дополнительных белков, которые модулируют работу рибосом. Загадочные межвидовые изменения в ходе эволюции нуклеотидных последовательностей и длины дополнительных 30-ти сегментов расширения, а внутри видов – на уровне нуклеотидных вариаций, могут обеспечить дополнительную регуляцию этой древней молекулярной машины.

Транспортные РНК

Впервые описанные в 1975 году [60], молекулы тРНК долгое время считались пассивными участниками процесса трансляции, доставляющими аминокислотные остатки к рибосомам в соответствии с генетическим кодом без какой-либо активной регуляторной роли. Открытия последних лет, такие, как большое общее количество генов, кодирующих тРНК (около 600 у человека [73]), и неожиданные вариации числа копий тРНК в разных клетках или тканях, настоятельно предполагают дополнительную роль этих древних молекул [74, 75]. Обнаружилось, что клетки могут реагировать на различные сигналы окружающей среды, изменяя паттерны экспрессии тРНК, модифицируя ее нуклеотиды, устанавливая новые правила использования тРНК, и, таким образом, динамично способствовать адаптации трансляции к этим сигналам [76].

Оказалось, что человеческий тРНК-ом (включая различные функциональные фрагменты, происходящие от тРНК), влияет на клеточный цикл, индивидуальное развитие организма и определяет многие патологии, такие как рак, нейродегенеративные и метаболические заболевания [77, 78]. Например, было показано, что активация онкогенных сигнальных путей, включая АКТ-mTOR, RAS-MAPK и MYC, или потеря опухолевого супрессора TP53 может тонко регулировать экспрессию PHK-полимеразы III, что каскадно приводит к изменению экспрессии тPHK [79-81]. В результате повышенние уровня синтеза определенных тPHK может способствовать прогрессированию опухоли, за счет обеспечения преимущественной трансляции

опухолеспецифических мРНК, имеющих использование кодонов, отличающееся от такового у здоровых клеток [82].

Свойство вырожденности кодонов для 20 аминокислот требует до пяти тРНК с различными антикодонами (изоакцепторы тРНК) для считывания кодонов для каждой аминокислоты [83]. Семейство изоакцепторов может состоять из одной тРНК, например, tRNA^{Trp}, или до пяти различных тРНК, например, семейство tRNA^{Leu}. Каждое семейство содержит множество изо-декодеров. Это гены тРНК, имеющие один и тот же антикодон, но разные последовательности в другом месте тРНК [84]. Например, девять генов кодируют изо-декодер tRNA^{Trp} с антикодоном ССА в геноме человека, согласно данным базы данных GtRNAdb [73]. Сначала считали, что количество разных тРНК в клетке коррелирует с числом копий гена тРНК изо-декодера, однако недавние исследования показывают, что регуляция экспрессии тРНК намного сложнее, чем предполагалось [85]. Вариации уровней экспрессии тРНК изодекодера модулируют уровни изоакцепторных тРНК, изменения которых в настоящее время рассматриваются как соответствующие нормальному состоянию клетки или вызванные адаптацией к патологическим изменениям использования кодонов в трансформированных клетках [86–88].

Единственный этап трансляционного цикла, на котором, по определению, не должно быть места регуляторным тРНК, - это терминация трансляции. Действительно, для прекращения синтеза белка наличие одного из трех стоп-кодонов UAA, UAG и UGA в рибосомном A-сайте распознается комплексом факторов высвобождения eRF1 и eRF3. После распознавания стоп-кодона синтезирующийся пептид гидролизуется с помощью eRF1. В этот момент высвобождается зрелый белковый продукт, а мРНК удаляется из рибосомы с последующей утилизацией в клетке (обзор см. в [89, 90]). Однако иногда стоп-кодон интерпретируется как смысловой кодон из-за конкуренции между фактором высвобождения и близкородственной тРНК в A-сайте. Это приводит к образованию удлиненной на C-конце изоформы белка [91, 92]. Хотя такое считывание стоп-кодона встречается довольно редко (не более чем в 0.1 % случаев), его так называемая запрограммированная форма с частотами до 20 % действительно играет важную биологическую роль в клетках [93, 94]. Молекулы тРНК, следовательно, контролируют каждую фазу трансляции. Еще одной известной функцией тРНК является ее роль в вирусной репликации [95].

Однако, многочисленность генов тРНК у эукариот, тем не менее, считается несообразной с известными функциями. Многие гены тРНК были описаны как неактивные, что вызывает вопросы об их реальной роли [96]. Однако в литературе описываются неканонический вариант процессинга тРНК. Например, у красной водоросли *Cyanidioschyzon merolae* нашлись 11 генов тРНК, в которых 3'-половина тРНК расположена в геноме раньше 5'-половины. Был выявлен необычный путь процессинга, в котором концы предшественника тРНК лигируются, что приводит к образованию промежуточной кольцевой РНК, которая затем процессируется с образованием зрелой тРНК с правильным порядком расположения половинок [97].

В то же время, взгляд на архей обнаруживает гены тРНК с множественными интронами, а также сборку необходимых зрелых тРНК у Caldivirga maquilingensis из транскриптов, получаемых с так называемых расщепленных генов. Эти гены индивидуально транскрибируются и затем транс-сплайсируются для получения недостающих тРНК, генов которых нет в геноме этого организма. Межгенные расщепленной тРНК высокую идентичность промежутки имеют с последовательностями интронов тРНК, расположенными в тех же положениях в интронсодержащих тРНК у родственных видов Thermoproteales. Это свидетельствует о существовании эволюционной взаимосвязи между интронсодержащими И расщепленными тРНК [97, 98, 99]. Приведенные примеры неканонического процессинга тРНК доказывают приспособляемость генов тРНК к разного рода ограничениям,

направленную на обеспечение процесса трансляции необходимым количеством тРНК нужного вида.

Малые ядерные РНК

Большинство генов у высших эукариот транскрибируются как предшественники мРНК (пре-мРНК) которые содержат промежуточные последовательности (интроны), а также экспрессируемые последовательности (экзоны). Открытые в конце 1970-х годов, интроны удаляются в процессе сплайсинга пре-мРНК, который объединяет экзоны вместе с образованием зрелых мРНК [100, 101]. Поскольку большинство человеческих генов содержат несколько интронов, сплайсинг является решающим шагом в экспрессии генов. Сплайсинг катализируется в два этапа с помощью динамической рибонуклеопротеиновой машины, называемой сплайсосомой [102], реакция требует гидролиза большого количества АТФ [103]. Считается, что эта повышенная сложность гарантирует точность и регулируемость сплайсинга [104].

Сплайсосома состоит из пяти различных рибонуклеопротеиновых субъединиц, а также множества ассоциированных белковых кофакторов [103, 105]. Чтобы отличить их от других клеточных рибонуклеопротеинов (РНП), таких, как рибосомные субъединицы, сплайсосомные субъединицы были названы малыми ядерными РНП (мяРНП). Как и в случае сборки рибосом, биогенез сплайсосомных мяРНП представляет собой многоступенчатый процесс, который происходит в различных субклеточных компартментах.

Малые ядерные РНК представляют собой группу многочисленных некодирующих неполиаденилированных транскриптов, которые выполняют свои функции в нуклеоплазме. МяРНК можно разделить на два класса на основе общих признаков первичных последовательностей и белковых кофакторов [106].



Рис 1. Схематическое строение мяРНП двух классов Sm и Lsm. Структурные особенности каждого класса выделены красными рамками. а – Все мяРНП Sm-класса (за исключением U7 РНП), имеют общее ядро из семи Sm-белков и уникальный набор дополнительных белков, специфичный для каждого мяРНП; b – мяРНП Lsm-класса имеют полиуридиновый сайт (Lsm-сайт), который направляет сборку специфичного гептамерного белкового кольца. Заимствовано из [106].

Малые ядерные РНК класса Sm имеют 5'-триметилгуанозиновый кэп (TMG), 3'шпильку и сайт связывания для группы из семи белков Sm (сайт Sm), которые образуют гетерогептамерную кольцевую структуру (рис. 1,а). Класс Sm snRNA состоит из мяРНК

261

U1, U2, U4, U4atac, U5, U7, U11 и U12. Все РНК этого класса транскрибируются РНКполимеразой II. Отдельно стоит U7 мяРНК, которая имеет более узкое назначение, она участвует в процессинге 3'-конца пре-мРНК гистонов.

МяРНК класса Lsm содержат метилфосфатный кэп (MPG) и 3'-шпильку, оканчивающуюся отрезком из нескольких уридинов, которые образуют сайт связывания для отдельного гетерогептамерного кольца белков Lsm (рис. 1,b). Класс Lsm состоит из двух мяРНК: U6 и U6atac, которые транскрибируются РНК-полимеразой III. Они никогда не покидают ядро в отличие от мяРНК Sm-класса, у которых вновь синтезированные мяРНК сначала экспортируются в цитоплазму, где они проходят дополнительные этапы созревания, прежде чем они будут импортированы обратно в ядро. Возможные причины такого сложного пути созревания мяРНК анализируются в [104].

Основная сплайсосома, которая состоит из пяти малых ядерных РНП (мяРНП; U1, U2, U4, U5 и U6) и различных дополнительных белков, отвечает за удаление более, чем 99 % всех интронов пре-мРНК (тип U2). Функционально подобная минорная сплайсосома, состоящая из snRNP U11, U12, U4atac и U6atac вместе с мяРНП U5, опосредует удаление определенного подмножества атипичных интронов (тип U12) [107]. Поэтапная сборка основной сплайсосомы на субстрате предшественника матричной РНК (пре-мРНК) подробно описана в [108, 109, 104].

Малые ядерные РНК играют критическую роль в позиционировании сплайсосомы на пре-мРНК и в некоторых случаях также вносят вклад в каталитические реакции. Во время своего созревания мяРНК подвергаются многочисленным стадиям процессинга и связываются со многими белками с образованием сворачивания, а также функциональных мяРНП. Сборка и функция сплайсосом включает динамические взаимодействия спаривания оснований между мяРНК, между мяРНК и пре-мРНК, а также бесчисленное множество взаимодействий РНК-белок. Точная настройка этих взаимодействий, вероятно, важна для поддержания эффективности и точности сплайсинга пре-мРНК. Малые ядерные РНК содержат многочисленные модифицированные нуклеотиды, которые, как предполагается, вносят вклад в такую оптимизацию [110]. Подробный обзор всех данных о сайтах и функциях модификаций мяРНК человека, а также о ферментах, которые их осуществляют, сделан в [107].

Непрерывное расширение знаний о механизмах, запускаемых модификациями РНК в процессах сборки и сплайсинга, способствует развитию технологий модуляции сплайсинга. В обзоре [111] проанализированы текущие знания о механизмах и функциях модификаций сплайсосомных мяРНК и их биологическом значении в процессе сплайсинга. С ростом числа клинических исследований, основанных на терапии путем модуляции сплайсинга (пропуск экзона или включение экзона) с применением одобренных FDA препаратов, основанных на этом механизме действия [112, 113], интерес к этой области биологии, безусловно, будет расти.

Малые ядрышковые РНК

Малые ядрышковые РНК (мякРНК) представляют собой консервативный, многочисленный и повторяющийся тип некодирующих РНК, присутствующих у всех эукариот и у подмножества архей [114–116]. Обнаруженные более четырех десятилетий назад, мякРНК были охарактеризованы, благодаря их роли в биогенезе рибосом. Многие из них служат в качестве природного инструмента сайт-специфической химической модификации рибосомальной РНК (рРНК), а небольшое их число также участвует в процессинге рРНК [117, 118].

Описаны два основных класса мякРНК, различающиеся по содержащимся в них мотивам последовательностей нуклеотидов, структуре, взаимодействующим с ними белкам и, как следствие, химической модификации, которую они катализируют [119].

Малые ядрышковые РНК с С / D-боксами варьируются по длине от 70 до 130 нуклеотидов и характеризуются наличием сайтов С (RUGAUGA) и D (CUGA), обнаруженных около 5'- и 3'-концов молекулы и взаимодействующих посредством неканонического спаривания оснований, образующего кинк-поворот [120–122]. Дополнительные мотивы, С' и D', с такими же консенсусными последовательностями, как блоки С и D, соответственно, но часто менее консервативные, обнаруживаются в середине молекулы [118, 122]. Малые ядрышковые РНК с С / D-боксами идентифицируют свои мишени с помощью антисмыслового элемента, отрезка из 10-20 нуклеотидов непосредственно перед сайтами D или D', строго комплементарного мишени (рис 2A) [118, 122].

Малые ядрышковые РНК с Н / АСА-боксами бывают длиннее (обычно от 110 до 145 нуклеотидов) и состоят из двух шпилек, разделенных шарниром или Н-боксом (ANANNA, где N может быть любым нуклеотидом) и оканчиваются АСА-боксом, который находится на расстоянии в 3 нуклеотида от 3'-конца молекулы (рис. 2,В) [118, 122, 123]. Малые ядрышковые РНК с Н / АСА-боксами взаимодействуют с четырьмя консервативными белками, включая дискерин псевдоуридинтрансферазы, который катализирует модификацию мишени [124]. Антисмысловые области этих РНК с комплементарностью к мишени являются двудольными и расположены в выпуклостях на шпильках, что указывает на вполне определенный уридин, который должен быть модифицирован путем реакции псевдоуридилирования [118, 122].



Рис 2. Два основных класса мякРНК: A - C / D-мякРНК характеризуются наличием сайтов C и D, обнаруживаемых около 5'- и 3'-концов, соответственно, и взаимодействующих посредством неканонического спаривания оснований, образующих кинк-поворот. Дополнительные сайты C 'и D' можно найти в середине молекулы. Антисмысловой элемент, или направляющая область, которая спаривается с мишенью, изображена красной линией; остаток, который должен быть метилирован (Me), находится непосредственно выше боксов D' и/или D. B – H / ACA- мякРНК состоят из двух шпилек, разделенных сайтом H и оканчиваются сайтом ACA, обнаруживаемым за 3 нуклеотида перед 3'-концом молекулы. Направляющие области, определяющие позицию в мишени, которая должна быть псевдоуридилирована (Ψ), находятся в выпуклостях на шпильках. Заимствовано из [119].

Несколько мякРНК из обоих классов не функционируют как направляющие РНК, но необходимы для эндонуклеолитического процессинга рРНК, процесса, также включающего эволюционно консервативную мякРНК, которая не может быть отнесена ни в один из двух вышеуказанных классов: это – рибонуклеазная MRP-PHK [62].

Геномная организация мякРНК варьируется от независимо транскрибируемых генов, наделенных своими собственными промоторными элементами, до интронных кодирующих единиц, лишенных независимого промотора. И у дрожжей, и у животных процессинг интрон-кодируемых мякРНК в значительной степени зависит от сплайсинга. Напротив, экспрессия растительных мякРНК, кодируемых интронами, по-видимому,

263

зависит от процесса, независимого от сплайсинга [125]. Более того, в обоих контекстах (межгенном или интронном) гены этих РНК могут быть, как отдельными транскрипционными единицами, так и частью кластеров. В последнем случае генерация индивидуальных мякРНК включает ферментативный процессинг полицистронных РНК-предшественников. Такой процессинг, по крайней мере, у дрожжей, включает ту же комбинацию эндо- и экзорибонуклеаз, необходимых для созревания моноцистронных предшественников мякРНК [126–128].

Помимо своей канонической роли в биогенезе pPHK, есть подмножество мякPHK, которое направляет модификацию малых ядерных PHK, а остальные мякPHK упоминаются, как орфанные, т.е. не имеют гомологов в других эволюционных ветвях [100, 129]. Однако за последние 15 лет о неканонических функциях мякPHK сообщалось в контексте различных уровней регуляции экспрессии генов [117, 128, 130]. Было подтверждено при этом, что небольшое число мякPHK может выполнять как канонические, так и неканонические функции [131].

РЕГУЛЯТОРНЫЕ НЕКОДИРУЮЩИЕ РНК

РНК-интерференция

Открытие РНК-интерференции [132-134] произвело революцию в понимании регуляции генов, выявив ряд связанных с этим явлением путей в клетке, в которых небольшие (длиной 20-30 нуклеотидов) некодирующие РНК и ассоциированные с ними белки контролируют экспрессию генетической информации [135]. В процессах, широко распространенных у растений и животных, каждая малая РНК связывается с белком семейства Argonaute с образованием специфичного лля последовательности рибонуклеопротеина (PHП), подавляющего экспрессию конкретного гена, специфичность которого обеспечивается спариванием оснований между малой (направляющей) РНК и ее мРНК-мишенью.

РНК-интерференция отвечает за контроль жизненно важных процессов, включая рост клеток, дифференцировку тканей, образование гетерохроматина и пролиферацию клеток. Было показано, что дисфункция РНК-интерференции связана с сердечнососудистыми заболеваниями, неврологическими расстройствами и многими типами рака [136]. Огромные усилия были приложены в надежде разработать методы лечения на основе миРНК для борьбы с генетическими или вирусными заболеваниями [137–139]. Несмотря на то, что эта область исследований является многообещающей, проблемы с адресной доставкой лечебных миРНК, а также вредные побочные эффекты лечения надолго затормозили выход препаратов на рынок. И только в августе 2018 года Управление по санитарному надзору за качеством пищевых продуктов и медикаментов разрешило применение США (FDA) первого препарата, основанного на олигонуклеотидах и РНК-интерференции, одобрив патисиран (Onpattro; Alnylam Pharmaceuticals) для лечения редкого наследственного заболевания амилоидной полинейропатии [140, 141]. Сейчас создание препаратов на основе малых некодирующих РНК вновь развивается очень активно. Разработаны и совершенствуются химические вспомогательные агенты доставки: наночастицы, липидные наночастицы, полимеры, дендримеры, наноструктуры из нуклеиновых кислот, экзосомы и специфические пептиды [141]

Два пути РНК-интерференции (ассоциированные с микроРНК или миРНК) имеют общий способ действия: минимальный эффектор – это комплекс РНП, содержащий белок семейства Argonaute, связанный с одноцепочечной РНК, которая обеспечивает специфичность за счет спаривания нуклеотидных оснований с геном-мишенью. Называется этот эффектор РНК-индуцированным комплексом выключения генов (RISC). Он управляет подавлением экспрессии мРНК-мишени посредством её деградации и / или репрессии её трансляции (рис. 3).



Рис. 3. Схема двух путей РНК-интерференции: для миРНК (слева) и для микроРНК (справа). Заимствовано из [133].

Клеточное происхождение (биогенез) микроРНК и миРНК различается: микроРНК происходят из генома, тогда как миРНК могут быть эндогенными или возникать в результате вирусной инфекции или из других экзогенных источников [142]. Другое ключевое различие возникает в предшественниках двухцепочечной РНК каждой из них: дуплексы миРНК характеризуются идеальным спариванием оснований, в то время как спирали микроРНК содержат несовпадения и выпетливания в стержне, имеют более протяженные концевые петли. Несмотря на различное происхождение, эти пути процессинга сходятся, как только любой из двух типов РНК собирается в RISC.

Обычно биогенез микроРНК происходит в ядре, где в результате транскрипции гена микроРНК получается первичная микроРНК. Такие транскрипты имеют длину не менее 1000 нуклеотидов, содержат одиночные шпильки или кластеры двухцепочечных шпилек, которые несут одноцепочечные 5'- и 3'- выпетливания и дистальные петли [143]. Первичная микроРНК обрезается микропроцессорным комплексом (microprocessor), включающим фермент Drosha из семейства РНКаз III и, белок DGCR8, содержащий два домена, связывающих двухцепочечные РНК. DGCR8 распознает соединение ствола и одноцепочечной РНК (дистальной петли), что помогает позиционировать Drosha для эндонуклеолитического расщепления, которое он выполняет на стержне, отступив 11 пар оснований [144]. Образовавшийся предшественник микроРНК длиной 65–70 нуклеотидов связывается с транспортными белками Exportin-5 и RanGTP и экспортируется в цитоплазму [145].

В цитоплазме пути процессинга сходятся для эндогенных микроРНК и, как правило, экзогенных, миРНК. Оба типа предшественников РНК-интерференции сокращаются до дуплекса подходящего размера для загрузки на белок Argonaute (Ago); это обычно выполняется ферментом Dicer. Ферменты семейства Dicer – это большие эндорибонуклеазы, содержащие домен хеликазы и внутренне димеризованную пару доменов РНКазы III, но этот состав может варьировать между организмами.

Полученная двухцепочечная РНК представляет собой дуплекс из 21–25 нуклеотидов, несущих двухнуклеотидные выступы на каждом 3'-конце и фосфатную группу на каждом 5'-конце [146]. На этом ферментативном этапе и последующей загрузке на Argonaute эндорибонуклеазе может помогать РНК-связывающий белок dsRBP. Эти три белка (Dicer, Argonaute и dsRBP) составляют минимальный RISC-нагружающий комплекс, который отвечает за нарезку двухцепочечной РНК и загрузку ее на Argonaute [147].

Загрузка RISC – это этап выбора преимущественной цепи дуплекса, которая будет связана с Argonaute, другая цепь отбрасывается. Эти цепи называются, соответственно, направляющей (guide) и пассажирской (passenger) цепями, и их выбор является ключевым детерминантом последующего выключения гена. В случае микроРНК-опосредованной РНК-интерференции наиболее часто загружаемая цепь из дуплекса известна как микроРНК, тогда как противоположная цепь называется микроРНК* (star, звездочка).

RISC осуществляет комплементарное связывание одноцепочечной мРНК, с направляющей цепью, связанной с Argonaute. Нуклеотиды в позициях 2–6 направляющей цепи представляют собой затравочную последовательность (seed) и инициализируют связывание с мишенью. Степень спаривания оснований влияет на то, как будет происходить последующее выключение экспрессии. В случаях полной комплементарности может происходить расщепление мишени, если Argonaute обладает каталитической активностью. RISC может также вызывать неэндонуклеолитическую репрессию трансляции до или после инициации; это может сопровождаться деаденилированием и деградацией.

Дополнительный клеточный аппарат отвечает за процессы, происходящие после связывания с мишенью, а Ago-связывающий белок GW182 является ключевым медиатором в привлечении дополнительных компонентов к RISC и в локализации активности по выключению гена в цитоплазматических локусах, известных как процессинговые тельца [148].

Данный процесс имеет значение, как в регуляции активности генов с помощью микроРНК, так и в защите от вирусных инфекций, так как геном вирусов часто представляет собой двухцепочечные РНК. Понимание механизма микроРНКиндуцированной РНК-интерференции дало начало синтетическим терапевтическим средствам на основе микроРНК. Искусственные конструкции, именуемые как anti-mirs и block-mirs, которые ингибируют активность определенной микроРНК или конкретной предотвращают выключение мишени с помощью микроРНК. соответственно. Идет создание также микроРНК-миметиков (синтетических версий микроРНК) для направленного выключения генов [141].

МикроРНК

МикроРНК были обнаружены почти 30 лет назад [44]. Сейчас это самый изученный класс молекул РНК, которые играют важную регуляторную роль у различных видов эукариот [149, 150].

Биогенез микроРНК, приведенный выше при описании механизма РНКинтерференции, касается только канонических микроРНК, их первичные транскрипты первоначально расщепляются ядерным ферментом Drosha с образованием шпилекпредшественников микроРНК. После экспорта в цитоплазму шпильки превращаются в

дуплексы микроРНК под действием фермента Dicer [151]. Одна дуплексная цепь длиной 22 нуклеотида сохраняется в комплексе Argonaute и направляет его к комплементарным мишеням мРНК, тогда как ее комплементарная цепь miRNA* преимущественно разрушается [152, 153].

Помимо канонического пути, были обнаружены неканонические пути биогенеза микроРНК с участием РНКаз, которые функционируют в других процессах [154]. Главным среди них является путь миртронов, в котором расщепление Drosha заменяется действием сплайсосомы [155–157].

Неканонические микроРНК используют специфический путь транскрипции, отличный от пути созревания канонических микроРНК. Они используют машину сплайсинга, которая работает при созревании всех мРНК генов, кодирующих белки у эукариот. Неканонические микроРНК кодируются в коротких интронах белоккодирующих генов, которые имеют большое число альтернативных транскриптов [158].

Сравнительное исследование репертуаров миртронов у человека и мыши [159] показало, что в этих геномах кодируются по меньшей мере 478 и 488 миртронов, соответственно, из них 13 общих. Из числа общих только 3 имеют консервативную последовательность, а 10 регулируют трансляцию одинаковых по функции генов у этих двух организмов. В работах [158, 160] были проанализированы различные библиотеки глубокого секвенирования РНК из клеток разных тканей человека и мыши, в результате были получены достоверные выборки неканонических микроРНК, существование которых экспериментально подтверждено.

На основе геномной локализации неканонические микроРНК были поделены на 4 группы:

- миртроны с 5'-хвостом (5'-tailed), самая многочисленная группа (87 % всего пула выявленнных миртронов у человека и 86 % у мыши), с 3'-конца вырезаются по акцепторному сайту сплайсинга;

- миртроны с 3'-хвостом (3'-tailed), составляют 3.8 % у человека и 3 % у мыши, вырезаются по донорному сайту сплайсинга;

- традиционные (conventional) миртроны, которые занимают весь интрон, т.е. ограничиваются сайтами сплайсинга, их 6.9 % у человека и 8.1 % у мыши;

- миртроны с двумя хвостами (two-tailed), они не примыкают ни к одному сайту сплайсинга, находятся внутри интронов, их 3.6 % у человека и 1.4 % у мыши.



Рис. 4. Схема канонических и опосредованных сплайсингом путей биогенеза микроРНК. Заимствовано из [159].

На рисунке 4 показаны два пути биогенеза микроРНК – каконический и неканонический (сплайсинг-опосредованный). Канонические шпильки расщепляются Drosha для высвобождения предшественника микроРНК, которая экспортируется в цитоплазму и попадают под действие Dicer для высвобождения дуплекса микроРНК / микроРНК^{*}. Одна нить загружается в белок Argonaute (AGO) и направляет

267

его к мишеням. Обычные миртроны представляют собой короткие интронные шпильки, которые сплайсируются И разветвляются, чтобы сформировать шпильку предшественника микроРНК. У 5'- и 3'-хвостовых миртронов сплайсинг формирует только один конец шпильки с неструктурированным хвостом, который простирается до или 5'-донорного сайтов сплайсинга, соответственно. 3'-акцепторного Такие предшественники требуют дополнительного процессинга для образования шпильки премиРНК.

Показано, что при биогенезе З'-хвостового миртрона, образуется шпилька с З'концевым удлинением, которая удаляется ядерной экзосомой с образованием предшественника микроРНК [157]. Кроме того, микроРНК, происходящие из миртронов и канонических микроРНК, очень часто подвергаются З'-укорачиванию. Это, вероятно, происходит внутри зрелых комплексов RISC. Такая обрезка часто проводится на 1 или 2 нуклеотида, но она расширена для определенных миртронов, у которых расщепление Dicer приводит к появлению аномально длинных цепей в диапазоне 24–27 нуклеотидов. Следовательно, в дополнение к хорошо известной роли эндонуклеазы РНКазы III, различные другие ядерные и цитоплазматические экзонуклеазы также играют существенную роль в биогенезе определенных микроРНК [160].

МикроРНК животных способны осуществлять регуляцию мРНК посредством шестисеми комплементарных взаимодействий на 5'-конце предпочтительно в положениях 2–8 зрелой микроРНК [161–165]. Такие взаимодействия в затравочном сайте достаточны для обеспечения регуляции, и результаты сравнительного геномного анализа показывают, что эти сайты длиной в 7 нуклеотидов часто и специфически подвергаются стабилизирующему отбору эволюции [166–168, 164].

Малые итерферирующие РНК

Еще одними участниками РНК-интерференции – специфичного механизма регуляции генов, консервативного у эукариот, кроме эндогенных микроРНК, являются малые интерферирующие РНК (миРНК).

Анализ биохимического механизма РНК-интерференции в лизате эмбриона *D. melanogaster* показал, что длинная двухцепочечная РНК процессируется до 21нуклеотидной малой РНК [169]. После процессинга 21-нуклеотидные малые РНК имеют двухнуклеотидный 3'-выступ, смежный с 5'-фосфатом и свободный 3'-гидроксильный конец [170]. Поскольку такие промежуточные продукты процессинга эффективно нацелены на мРНК для её деградации, они были названы малыми интерферирующими РНК (миРНК) [63].

МиРНК генерируются из длинной двухцепочечной РНК и, в основном, участвуют в защите эукариотической клетки от молекулярного паразитизма. Исключение атак клетки и её генома РНК-вирусами, транспозонами и воздействия путем трансгенеза может эффективно осуществляться посредством миРНК-опосредованной РНК-интерференции [171, 172]. Отличие миРНК от микроРНК состоит в том, что для регуляции экспрессии гена первые действуют посредством полной комплементарности с последовательностью мишени, а микроРНК связываются более свободно. Малые интерферирующие РНК образуются при экзогенной доставке двухцепочечной РНК или трансгенной экспрессии длинной двухцепочечной РНК, но эндогенные источники миРНК немногочисленны [63].

Эндогенные миРНК подразделяются на три подкласса: миРНК, ассоциированные с повторами (rasiRNA); транс-активирующие миРНК (tasiRNA) и миРНК, полученные из природных антисмысловых транскриптов (natsiRNA) [173–175].

Ассоциированные с повторами миРНК репрессируют повторяющиеся последовательности на уровне транскрипции или посттранскрипционной регуляции и поддерживают центромерную гетерохроматическую структуру. Эти РНК были клонированы и секвенированы у *S. pombe, T. brucei, C. elegans, D. melanogaster, D. rerio* и *A. thaliana*, но не у млекопитающих [63]. У *D. melanogaster* и *D. rerio*, однако, длина

миРНК, ассоциированных с повторами, больше, чем у микроРНК [176, 177]. Источником миРНК, ассоциированных с повторами, предположительно, является двухцепочечная РНК, продуцируемая отжигом смысловых и антисмысловых транскриптов, которые содержат повторяющиеся последовательности, часто связанные с мобильными элементами [178]. Молекулы РНК этого вида обычно менее многочисленны, чем микроРНК.

Транс-активирующие миРНК процессируются у наземных растений из двухцепочечных РНК, синтезируемых эндогенной РНК-зависимой РНК-полимеразой, использующей эндогенные мРНК в качестве матрицы. Малые интерферирующие РНК этого вида отличаются от других миРНК тем, что они связывают свои мишени с меньшей специфичностью. В этом их механизм действия похож на активность микроРНК, так как они не нуждаются в полной комплементарности со своей мишенью, чтобы направлять её распад [174].

Малые интерферирующие РНК, полученные из природных антисмысловых транскриптов, процессируются из двухцепочечных РНК, образованных эндогенными смысловыми и антисмысловыми транскриптами, они имеют длину от 21 до 24 нуклеотидов [179] и участвуют в ряде этапов индивидуального развития и адаптогенного ответа у растений, такого как устойчивость к патогенам [180], солеустойчивость и биосинтез клеточной стенки [181]. Малые интерферирующие РНК этого вида также изменяют у растений экспрессию генов, реагирующих на стресс-факторы окружающей среды [182].

РНК, взаимодействующие с белками PIWI

В 2006 году одновременно пять групп сообщили об открытии новых малых РНК, отличных от микроРНК и миРНК [183–187], которые были обнаружены в семенниках мышей. Эти молекулы – РНК, взаимодействующие с белками РІШІ (пиРНК) – представляют собой эндогенные малые некодирующие РНК, которые действуют как хранители генома, защищая его от инвазивных мобильных элементов в зародышевой линии животных. РНК, взаимодействующие с белками PIWI, как правило, длиннее микроРНК и малых интерферирующих РНК. Они имеют длину 26 -32 нуклеотида и, в отличие от микроРНК, не так консервативны. Белки PIWI относятся к большой группе белков Argonaute и экспрессируются почти исключительно в клетках зародышевой линии; они необходимы для поддержания стволовых клеток. сперматогенеза и репрессии мобильных элементов. Комплексы PIWI с пиРНК не только задействованы в сайленсинге ретротранспозонов и других генетических элементов на пост-трансляционном уровне, но имеют и некоторые другие функции [64, 188, 189].

Опубликованные в последние годы результаты показали, что механизм PIWI / пиРНК также способствует деградации транскриптов мРНК через известные механизмы с участием микроРНК или миРНК в генетической регуляции в половых клетках различных животных [190–196].

Совсем недавно была показана двойная роль MIWI/пиPHK-аппарата в регуляции трансляции во время сперматогенеза у мышей. (МIWI – один из трех мышиных гомологов PIWI). Ранее было установлено что механизм MIWI/пиPHK в комплексе с деаденилазой CAF1 отвечает за элиминацию мPHK на поздних стадиях сперматогенеза при подготовке к производству сперматозоидов [190]. Позже обнаружилось [197], что тот же механизм активирует трансляцию мPHK на ранних стадиях сперматогенеза. Такое действие требует специфичной гибридизации пиPHK с 3'-нетранслируемой областью целевых мPHK, а также определенного сочетания цис-действующих AU-богатых элементов в целевой мPHK и транс-действующих факторов HuR и eIF3f. Также было обнаружено, что MIWI связан с белками HuR и eIF3f преимущественно на ранних стадиях развития сперматидов, а с CAF1 – на поздних. Следовательно, функционирование механизма MIWI/пиPHK отрегулировано в соответствии с

конкретной требованиями стадии сперматогенеза И специфики зависит от ассоциированных белковых факторов, контекста последовательности мишени мРНК и стадии индивидуального развития [198].

РНК, взаимодействующие с белками PIWI, кодируются в межгенных повторяющихся элементах генома, называемых кластерами пиРНК [199]. Кластеры пиРНК охватывают широкую область генома, иногда состоящую из сотен тысяч оснований. Эти области генома, в основном, состоят из различных мобильных элементов и их дефектных остатков. У мух предшественники пиРНК происходят из гетерохроматических локусов, тогда как у млекопитающих кластеры пиРНК неотличимы от транскрипционных единиц РНК полимеразы II канонической эухроматической ДНК [200]. Функции пиРНК в последнее время активно изучаются (см., например, обзоры в [201–204]).

Малые РНК, происходящие из тРНК

Первые фрагменты тРНК были обнаружены в экспериментах по секвенированию микроРНК. В этих экспериментах многочисленные и вездесущие тРНК-фрагменты (tRNA-derived frafments, tRFs) были обнаружены во множестве секвенированных образцов и были исключены из анализа, как шум. Потребовалось больше 10 лет, чтобы после детальных исследований [205-207] понять их роль, хотя специфические продукты разрушения тРНК обнаруживались в моче больных раком пациентов и гораздо раньше [208].

РНК, происходящие от тРНК, или тРНК-фрагменты (тРФ), ассоциированы с развитием раковых опухолей, и их сигнатуры сейчас уже предложены в качестве биомаркеров рака [209, 210]. Во многих экспериментальных исследованиях сообщалось, что тРФ ингибируют синтез белка в ответ на индуцированный стресс [211], а также подавляют рост раковой опухоли, клеточную инвазию и метастазирование [212]. Фрагменты тРНК были идентифицированы у представителей всех ветвей древа жизни. Исследования показали, что они обладают такими же регуляторными функциями, какими обладают микроРНК [213-215], а также другими специфичными функциями [216].

Это далеко не случайные продукты деградации тРНК [217, 218], биогенез тРФ фактически контролируется набором высококонсервативных и точных сайтспецифических механизмов разрезания, которые производят фрагменты РНК длиной 14-50 нуклеотидов [219].

Выделяют два класса РНК, происходящих из тРНК [220]. Первый – это 5'- и 3'половинки тРНК, также называемые индуцированными стрессом РНК из тРНК (tRNAderived stress-induced RNAs, tiRNAs), длиной 30-40 оснований. Они образуются путем специфического расщепления в антикодоновых петлях зрелых тРНК в стрессовых условиях, таких как гипоксия, голодание, вирусная инфекция, тепловой шок или клеточный стресс, вызванный отравлением тяжелыми металлами. Этот специфический гидролиз редко происходит в нормальных условиях, и при этом обычно продуцируется очень мало таких фрагментов тРНК. Эти РНК стоят особняком среди РНК, происходящих из тРНК [221, 222].

Другой класс – это фрагменты тРНК. Они имеют размер 18–22 нуклеотидов и образуются из зрелых тРНК и их предшественников нуклеазами Dicer или РНКазой Z (рис. 5). По месту их разрезания и местоположению последовательности на тРНК, в настоящее время существует три типа тРФ, в том числе: 5'- тРФ, 3'- тРФ, 1- тРФ [220].

Большинство тРФ могут играть важную роль в выключении генов, опосредованном РНК-интерференцией. Эксперименты показали, что тРФ участвуют в подавлении генов, прямо воздействуя на мРНК аналогично микроРНК и пиРНК, или даже конкурируя с микроРНК за связывание с ее мишенями. Другой механизм подавления гена с помощью тРФ заключается в конкурентном с белками связывании с мРНК. Например, 1-тРФ

(фрагмент тРНК^{Ser-TGA}) при вирусных инфекциях может напрямую связываться с белком La / SSB и ингибировать La / SSB-зависимую экспрессию вирусного гена [223].



Рис 5. Различные типы фрагментов РНК, происходящих из тРНК: 1-тРФ продуцируется путем расщепления РНКазой Z во время процессинга предшественника тРНК. Зрелая тРНК может быть расщеплена в петле антикодона с помощью ангиогенина с образованием 5'-половинки и 3'-половинки тРНК в стрессовых условиях, 5'-тРФ происходит от 5'-конца зрелой тРНК путем эндонуклеолитического расщепления и расщепления экзонуклеазой в D-петле. Расщепление T-петли приводит к образованию серии 3'-тРФ. Заимствовано из [220].

Было обнаружено, что индуцированная стрессом 5'-половинка РНК ингибирует синтез белка путем вытеснения эукариотических факторов инициации трансляции eIF4g и eIF4a с мРНК и даже вытеснения кэп-связывающего комплекса eIF4f с кэпа m7G [224].

Фрагменты тРНК также могут регулировать трансляцию, воздействуя на сборку рибосом. Было обнаружено, что одна специфическая 3'-тРФ, полученная из тРНК^{Leu-CAG}, связывается с кодирующей и 3'-некодирующей последовательностью в мРНК рибосомного белка S28, что увеличивает ее трансляцию и, в конечном итоге, количество рибосомного белка [220].

Как правило, активность мобильных элементов может ингибироваться различными эпигенетическими факторами на уровне транскрипции, например, метилированием ДНК, модификацией гистонов, ремоделированием хроматина и регуляции с помощью других некодирующих РНК. Процесс передвижения включает промежуточную стадию специфической последовательности синтеза ЛНК ретротранспозона. Новосинтезированный ретротранспозон встраивается в любой участок генома. Так называемый сайт связывания праймера, структурированный элемент РНК, с которого РНК-зависимая ДНК-полимераза начинает обратную транскрипцию, является уникальной мишенью для специфического ингибирования LTR-ретротранспозонов. Регуляторные З'-тРФ нацелены именно на него, и это является потенциально высококонсервативным механизмом контроля транспозонов, опосредованного малой РНК [220].

Малые РНК, происходящие из рибосомальных РНК

Самый распространенный вид клеточной РНК, рибосомная РНК (рРНК), повидимому, является источником огромного количества неслучайно генерируемых фрагментов. До недавнего момента шел процесс накопления разрозненных фактов о существовании и функционировании в разных организмах малых РНК, произошедших из рРНК, или фрагментов рРНК (рРФ). Однако пока эти данные не дают представления о роли этих новых регуляторных РНК.

Первое биоинформатическое исследование было проведено совсем недавно [66]. Авторы исследовали рРФ с помощью вычислительного метаанализа больших экспериментальных наборов малых регуляторных РНК в клетках человека и мыши. У человека проанализировали данные, полученные методом фиксации in vivo PHK-PHKвзаимодействий с помощью перекрестного связывания, лигирования и секвенирования гибридов (CLASH). Эта методика разработана для экспериментальной идентификации связывания микроРНК с мишенями и была использована для идентификации фрагментов тРНК, как потенциальных микроРНК-подобных регуляторов. Авторы использовали эти данные ранее при изучении паттернов взаимодействия фрагментов тРНК с их мишенями, предложив новые мотивы связывания и механизмы взаимодействия [225]. Они предсказали взаимодействующие области для двух дюжин тРФ, подтвержденные совпадениями с их мотивами во многих случаях, когда такие области были определены экспериментально [226, 227]. В новой работе, проведенной с помощью метода, уже давшего хороший результат по поиску новых тРФ [66], авторами было предсказано 680 мотивов pPФ, которые могут управлять взаимодействиями с тысячами мишеней на мРНК, кодирующие белки.

Обнаруженные потенциальные pPФ имели длину от 16 до 25 нуклеотидов, при этом наиболее часто встречались фрагменты длиной 20 нуклеотидов, 99 % которых неслучайным образом вырезаются из предшественника 45S-pPHK и попадают в границы зрелой pPHK. Фрагменты, в антисенс-ориентации, попадающие в нетранслируемые области транскрипта, составляли пренебрежимо малые доли. Области внешних (ETS) и внутренних транскрибируемых спейсеров (ITS) не были задействованы в этом процессе, то же самое касается и двух самых больших сегментов расширения (ES7L и ES27L) [228]. То есть регуляторные фрагменты вырезаются из коровых областей молекулы pPHK. Такое расположение потенциальных pPФ указывает на то, что состав, функция или доступность больших сегментов расширения не позволяет им генерировать pPФ. Другое объяснение состоит в том, что pPФ могли быть сохранены в процессе эволюции от более древних организмов, чьи pPHK были лишены сегментов расширения. Это согласуется с тем, что большинство мишеней pPФ нашлись в генах основных путей метаболизма и обработки генетической информации.

Наибольшая часть рРФ вырезается из 18S рРНК. В человеческом транскриптоме богатые гуанином рРФ предпочтительно вырезаются в одноцепочечных областях зрелых рРНК между пиримидинами и аденозином и неслучайно спариваются с клеточными транскриптами. Многочисленные идентичные рРФ были обнаружены в цитоплазме и ядре мыши. Авторы сообщают также о конкретных мотивах взаимодействия. Расположение таких мотивов на рРФ находится в полном соответствии со структурными особенностями Ago и паттернами взаимодействия Ago-PHK у обоих видов. Многие из этих мотивов могут связываться с двухцепочечными областями на PHK-мишенях, что указывает на потенциальный путь регуляции трансляции посредством раскручивания мPHK.

У всех малых регуляторных РНК участок связывания с мишенью содержит короткий фрагмент (затравку) для комплементарного связывания. У канонических микроРНК этот фрагмент имеет длину 7 нуклеотидов, у рРФ – до 12 нуклеотидов. Мишенями являются преимущественно зрелые мРНК, связанные с эволюционно старыми основными метаболическими путями и путями обработки генетической информации. В частности, в

272

качестве мишеней были выявлены рибосомальные белки, факторы инициации трансляции и факторы элонгации. Исключение составляют короткие интроны, так называемые аготроны (agotrones), которые известны как регуляторы, связанные с AGO (Ago-associated regulators) [229].

Малые РНК, происходящие из транспозонов

Мобильные элементы, или транспозоны, являются основной составляющей геномов многих эукариот, сообенно растений. Более 80 % геномов зерновых культур (например, ячменя, пшеницы и кукурузы) состоят из транспозонов [230]. Транспозоны делятся на два класса по способу их транспозиции (перемещения по геному) [231].

Мобильные элементы I класса, известные как ретротранспозоны, перемещаются посредством молекул РНК, превращающихся в комплементарные ДНК, которые дополнительные копии встраиваются обратно в геном, создавая элемента. Ретротранспозоны с длинными концевыми повторами (LTR) И длинными диспергированными повторами (LINE) являются двумя основными типами автономных ретротранспозонов, поскольку они сами кодируют белки, необходимые ЛЛЯ транспозиции. Есть еще группа неавтономных ретротранспозонов.

Мобильные элементы II класса, или ДНК-транспозоны, вырезаются из одного места и вставляются в другое положение генома с помощью белка транспозазы, который кодируется внутри самого элемента.

Во многих геномах растений ретротранспозоны более многочисленны по сравнению с мобильными элементами II класса. В частности, ретротранспозоны LTR являются преобладающими семействами мобильных элементов у многих растений [232].

Цикл репликации ретротранспозонов LTR начинается с транскрипции геномной копии с помощью РНК-полимеразы II хозяина. Матричные РНК ретротранспозонов LTR преремещаются в цитоплазму, там подвергаются трансляции, а затем обратной транскрипции [233]. Автономные ретротранспозоны LTR продуцируют множество белков, которые необходимы для завершения цикла ретротранспозиции, включая обратную транскриптазу и интегразу. В результате обратной транскрипции образуется линейная двухцепочечная ДНК, которая транспортируется обратно в ядро и интегрируется в геномную хромосомную ДНК с помощью интегразы.

Поскольку мобилизация транспозонов может быть мутагенной, в геномах хозяев выработались сложные механизмы для подавления их активности [234]. Мобильные элементы, в первую очередь, репрессируются эпигенетическими путями сайленсинга. Области генома, отмеченные метилированием, распознаются растительной РНКполимеразой IV, которая транскрибирует относительно короткие участки РНК [235]. Эти РНК дуплицируются с помощью РНК-зависимой РНК-полимеразы II и затем разрезаются до малых интерферирующих РНК (миРНК), имеющих длину 24 нуклеотида, посредством DICER-подобного DCL3. Эти миРНК связываются с белками AGO4 и взаимодействуют с формирующейся РНК, транскрибируемой РНК-полимеразой V. Затем AGO4 связывает ещё несколько белков, которые опосредуют репрессивную модификацию гистонов и метилирование ДНК, тем самым способствуя усилению состояния молчания транспозонов [236]. Те транспозоны, которые избежали сайленсинга, или вновь встроенные в геном транспозоны, инактивируются другим механизмом, в результате которого процессируются миРНК длиной 21-22 нуклеотида [237]. Эти миРНК связываются с AGO1 и нацелены на мРНК транспозона для деградации.

Интересно, что транспозонные миРНК могут также взаимодействовать с транскриптами-мишенями, которые не имеют отношения к мобильным элементам, выполняя определенные регуляторные роли в различных биологических процессах. У млекопитающих РНК, взаимодействующие с белками РІШ, регулируют большое количество мРНК и длинных некодирующих РНК в семенниках, что указывает на

широкое распространение регуляторных ролей малых РНК, полученных из транспозонов, как у растений, так и у животных [238].



Рис 6. Доля длинных некодирующих РНК, происходящих из транспозонов, в общем пуле длинных некодирующих РНК для ряда растений (кукуруза, рис, помидор, два вида арабидопсис) в сравнении с человеком. **Заимствовано из** [57].

Мобильные элементы часто рассматриваются как вредные геномные факторы, они сильно подавляются эпигенетическими механизмами сайленсинга. С другой стороны, мобилизация транспозонов приводит к изменчивости генома и транскриптома, которые необходимы для выживания и эволюции вида-хозяина. Подавляющее большинство контролирующих мобильных элементов влияют на соседние гены, либо стимулируя, либо репрессируя транскрипционную активность. Хотя транспозоны очень часто повторяются в геномах и транскрибируются в определенных стрессовых условиях или на стадиях развития, транс-действующая регуляция происходящих из транспозонов РНК описываются редко. У растений транспозоны являются обильным источником малых РНК, таких как миРНК и микроРНК. Эти малые РНК обладают потенциалом влиять на транскрипты, не являющиеся транспозонами, счет комплементарности за последовательностей, тем самым создавая новые регуляторные сети генов, включая стрессоустойчивость и барьер гибридизации. Помимо существования малых РНК, большинство длинных некодирующих РНК у животных и растений происходит из транспозонов (см. рис. 6) [239, 240]. За последнее десятилетие транскриптомный анализ резко расширил каталог длинных некодирующих РНК в различных тканях и стрессовых условиях многих видов растений. Однако, несмотря на большое количество идентифицированных длинных некодирующих РНК растений, их биологические роли все еще в значительной степени не исследованы [57].

Длинные некодирующие РНК

Длинные некодирующие РНК (днРНК) в зрелом виде имеют длину больше 200 нуклеотидов, расположены в ядре или цитоплазме и редко кодируют белки [241–243, 67]. Геномная локализация днРНК может быть отнесена к одной или нескольким из пяти широких категорий: 1) смысловая или 2) антисмысловая, когда один или несколько экзонов другого транскрипта перекрываются с геном днРНК на той же или противоположной цепи соответственно; 3) двунаправленная, когда экспрессия днРНК и кодирующего транскрипта на противоположной цепи инициируются в непосредственной близости, 4) интронная, когда РНК полностью происходит из интрона другого транскрипта или 5) межгенная, когда длинная некодирующая РНК кодируется в пределах интервала между двумя другими генами [416].

Хорошо известно, что по сравнению с кодирующими белками последовательностями и структурными РНК длинные некодирующие РНК слабо консервативны в эволюции [245]. Поэтому многие ранние исследования заклеймили днРНК «транскрипционной темной материей» и считали их в целом нефункциональными [246, 247]. Однако низкий

уровень или отсутствие детектируемой консервации не обязательно означает, что эти молекулы не имеют функции [248, 249].

Выделяют четыре архетипа механизма регуляции с помощью днРНК экспрессии генов: сигналы, ловушки, направляющие и каркасы [250], при этом отдельная длинная некодирующая РНК может участвовать в нескольких архетипах; очевидно, сложные функции могут быть построены путем комбинаторного использования разных архетипических молекулярных механизмов.

Большинство днРНК транскрибируются с помощью РНК-полимеразы II, как мРНК они сплайсированы, 5'-кэпированы и 3'-полиаденилированы [251]. Длинные некодирующие РНК демонстрируют специфическую экспрессию и реагируют на различные стимулы, они могут служить молекулярными сигналами, потому что транскрипция индивидуальных днРНК происходит в очень определенное время и в очень определенном месте для интеграции сигналов развития, интерпретации клеточного контекста или ответа на различные стимулы [250].

Большинство генов, считающихся неактивными, поскольку они не производят поддающегося обнаружению транскрипта, тем не менее, подвергаются инициации транскрипции [252]. Эти результаты позволяют предположить, что гены, кодирующие белок, делятся на три группы регуляторного поведения. Активно транскрибируемые гены заняты нуклеосомами с модификациями гистонов, которые являются отличительными признаками как инициации, так и элонгации; и они обычно производят детектируемые транскрипты. Вторая группа испытывает инициацию транскрипции без доказательств удлинения или накопления транскрипта. В этой популяции генов, которая включает большинство генов, кодирующих регуляторы развития, регуляция событий, инициацией транскрипции, должна играть важную слелующих за роль В предотвращении продукции или накопления транскриптов. Третья группа состоит из генов, которые исключены из процесса инициации транскрипции, где должны преобладать механизмы, предотвращающие инициацию транскрипции. Таким образом, повсеместная транскрипция [252] указывает на центральную роль днРНК в регуляции транскрипции, как положительно, так и отрицательно. Способы, с помощью которых такие нкРНК регулируют транскрипцию, расширяются и включают в себя множество механизмов, главный из которых – действовать как молекулярные ловушки. Этот архетип днРНК транскрибируется, а затем связывает и титрует белок-мишень, но не выполняет никаких дополнительных функций [250].

Третий архетип – это направляющие РНК: длинная некодирующая РНК связывает белок (белки), а затем направляет рибонуклеопротеиновый комплекс к конкретным мишеням. Длинные некодирующие РНК могут управлять изменениями в экспрессии генов либо цис- (на соседних генах), либо транс- (на удаленно расположенных генах) способом, который нелегко предсказать на основе последовательности днРНК [250].

Длинные некодирующие РНК могут служить платформами, на которых собираются соответствующие молекулярные компоненты; во многих разнообразных биологических процессах передачи сигналов эта характеристика жизненно важна для точного контроля специфичности и динамики межмолекулярных взаимодействий и сигнальных событий [253]. Традиционно считалось, что белки являются основными участниками различных строительных комплексов [254]. Недавно появившиеся данные повышают вероятность того, что днРНК также могут играть аналогичную роль [250].

В литературе описано пять возможных способов возникновения днРНК: 1) при нарушении трансляционной рамки считывания гена, кодирующей белок; 2) в результате хромосомной реорганизации, например, путем объединения двух нетранскрибируемых участков ДНК таким образом, чтобы способствовать транскрипции слитых некодирующих последовательностей; 3) в результате репликации некодирующего гена ретротранспозицией; 4) с помощью генерации некодирующей РНК, содержащей повторы, посредством механизма частичной тандемной дупликации; и 5) в результате

вставки мобильного элемента в ген таким образом, чтобы получить функциональную транскрибируемую некодирующую РНК [244, 255, 256].

Исследования разных групп [257–263] показали, что, транспозоны явно внесли значительный вклад в первичную структуру не только днРНК, но и даже микроРНК. Палиндромные последовательности определенных семейств мобильных элементов играют решающую роль в шпилечной структуре микроРНК, и разные транспозоны связаны с разными семействами микроРНК. Фрагменты последовательности транспозонов также были обнаружены в зрелых микроРНК, не являющихся шпильками. Присутствие транспозонов во всех областях генов днРНК (промоторы, интроны и экзоны) подчеркивает вклад мобильных элементов в образование днРНК [264].

Группой Кунина [265] было проведено сравнение наборов экспериментально подтвержденных днРНК у человека и мыши. Исходя из предположения, что наборы транскриптов днРНК являются случайными выборками, авторы оценили общий размер совокупности примерно в 40–50 тысяч единиц, что, по крайней мере, вдвое превышает количество генов, кодирующих белок. Доля эухроматических генов человека и мыши, кодирующих днРНК, более чем в два раза больше, чем доля последовательностей, кодирующих белок. Хотя первичные последовательности большинства днРНК гораздо менее консервативны, чем гены белков, оказалось, что от 60 % до 70 % генов днРНК являются общими для человека и мыши. Авторы предположили, что ортологичные гены днРНК млекопитающих должны выполнять эквивалентные функции; соответственно, кажется вероятным, что тысячи эволюционно консервативных неизвестных функциональных ролей днРНК будут охарактеризованы [266].

Кольцевые РНК

Поскольку большинство эукариотических генов прерываются интронами, РНКтранскрипты обычно подвергаются сплайсингу для удаления интронов, после чего экзоны коллинеарно сливаются с образованием зрелых линейных РНК-транскриптов. Сплайсинг - это строго регулируемый процесс, который посредством альтернативного сплайсинга может генерировать множество зрелых изоформ РНК из одного гена, и эти изоформы могут проявлять разные функции [267]. Более 95 % генов человека подвержены альтернативному сплайсингу [268], их экспрессия определяется как трансрегуляторными факторами, так и цис-регуляторными элементами, включая факторы сплайсинга и связывающие их мотивы.

Кольцевые РНК образуются с помощью особого механизма сплайсинга, называемого обратным сплайсингом, при котором 5'-конец пре-мРНК вышележащего экзона неколлинеарно сплайсирован с 3'-концом нижележащего экзона. Кольцевые РНК преимущественно обнаруживаются в цитоплазме, а отсутствие 5'-кэпа и 3'-poly(A)хвоста делает кольцевые молекулы более устойчивыми к деградации РНКазами по сравнению с их линейными аналогами [269]. Существование кольцевых РНК у млекопитающих впервые было обнаружено в 1979 году в цитоплазме клеток HeLa и других клеток млекопитающих с помощью электронной микроскопии [270]. Однако изза технических ограничений в течение следующих двух десятилетий было идентифицировано только несколько специфических кольцевых РНК, и потенциальные функции их оставались неясными [271–275]. С развитием секвенирования следующего полных последовательностей геномов поколения, появлением И развитием биоинформатических технологий, исследователи смогли обнаружить, что экспрессия кольцевых РНК у млекопитающих часто сохраняется у разных видов и демонстрирует тканевую и клеточную специфичность. Уровень экспрессии некоторых кольцевых РНК может быть выше, чем у линейных родственников [276-280].

Кольцевые РНК – это большой класс РНК, обладающих регуляторной способностью. Кольцевые РНК, состоящие из одной экзонной последовательности, представляют собой малоизученную форму нкРНК, которая была открыта более 20 лет назад из горстки

транскрибированных генов. В статьях [49, 50, 281, 282] описывается множество примеров кольцевых РНК у людей, мышей и нематод, а также механизм функционирования, по крайней мере, некоторых из них, в качестве «губок микроРНК» (miRNAs sponge) в клетке. Отдельные кольцевые РНК могут нести множество сайтов связывания микроРНК и ингибировать активность одной или нескольких микроРНК.

Хотя большинство кольцевых РНК сплайсируются из предшественников мРНК, кодирующих белки, кольцевые РНК обычно классифицируются как длинные некодирующие РНК. Подобно другим днРНК, они могут служить в качестве РНК- или белков-приманок для регулирования экспрессии генов. Кольцевые РНК также образуют комплексы с белками для регулирования клеточного цикла [283] или трансляции [284], или служат в качестве межклеточных сигнальных молекул в высвобождаемых экзосомах [285, 286]. Интересно, что некоторые кольцевые РНК могут кодировать функциональные пептиды, как было продемонстрировано в недавних работах, показывающих, что эти РНК способны транслироваться *in vitro* и *in vivo* [287–290].

Как правило, большинство кольцевых РНК содержат два или три экзона без сегрегации интронов, в то время как те экземпляры, которые содержат только один экзон, обычно имеют длину экзона больше средней [291]. Тем не менее, анализ последовательностей кольцевых экзонов и изучение эффекта замены кольцевых экзонов позволяют предположить, что не существует специфических экзонных последовательностей, которые контролируют образование кольцевой РНК [292]. С другой стороны, фланкирующие интроны обычно длиннее среднего и обогащены комплементарными повторами [278, 291, 271, 292-296]. Более того, повторяющимися элементами Alu обогащены фланкирующие интроны кольцевых РНК человека [278], а спаривание между повторяющимися элементами Alu с обратной ориентацией регулирует экспрессию линейных и круговых изоформ [291, 297].



Рис. 7. Возможные механизмы образования кольцевых РНК. Предполагаемый механизм образования CDR1as и его функция губки микроРНК (слева). Два других возможных механизма: при первом закольцовывание обеспечивается случайной гибридизацией двух инвертированных повторов; при втором альтернативный сплайсинг вырезает два экзона, окаймленных инвертированными ALU-повторами, они обеспечивают замыкание концов сплайсированного фрагмента (справа). Заимствовано из [51].

Несколько механизмов образования кольцевых РНК показано на рисунке 7, например, за счет закольцовывания одного экзона, чему способствует наличие смежной повторяющейся последовательности. В качестве альтернативы кольцевые транскрипты могут образовываться либо после выбора неправильного акцепторного сайта сплайсинга или удаления нескольких последовательных экзонов в транскрипте при альтернативном сплайсинге. Самая основная форма кольцевой РНК включает в себя обратную петлю к

277

предыдущему экзону внутри транскрипта, например, соединение между 3'-концом экзона 3 и 5'-концом экзона 2 в гене [51].

Описаны случаи участия кольцевых РНК в пролиферации клеток и прогрессировании раковых заболеваний, в онтогенезе и других клеточных процессах [298–300]. В последнее десятилетие количество опубликованных исследований кольцевых РНК росло экспоненциально, что сделало кольцевые РНК одними из самых заметных молекул в биологии РНК. Несколько баз данных (например, circBase[301], circNet [302], Circ2Traits [303], Circ2Disease [304], exoRBase [305] и CSCD [306]) были созданы для хранения кольцевых РНК разных видов и представления дополнительной информации о их связи с заболеваниями, клеточными местоположениями и другими некодирующими РНК.

Несмотря на внимание широкого круга исследователей к вопросам, касающихся функций и регуляции кольцевых РНК, многие вопросы остаются нерешенными. Например, то, как кольцевые РНК деградируют в клетке, и как деградация работает в сочетании с биогенезом, чтобы реагировать на динамические клеточные состояния, является важной областью, которая требует дальнейшего изучения. Хотя экспрессия кольцевых РНК была изучена в контексте многих заболеваний человека, понимание различных ролей в нормальной физиологии и болезненных условиях для подавляющего большинства идентифицированных кольцевых РНК ограничено [51].

ЗАКЛЮЧЕНИЕ

В этом обзоре представлена картина разнообразия клеточных РНК, описаны места их кодирования в геноме, пути биогенеза и представлены их помощники в деле регуляции различных клеточных процессов. Крупными мазками обозначены гибкие механизмы их функционирования и мишени, на которые направлена деятельность этих молекул. Картина вырисовывается довольно сложная, и белых пятен на ней еще много. Есть и небольшие точки, которые являются началом новых линий. Они соответствуют немногочисленным свидетельствам существования иных типов РНК, которые в нарисованной здесь картине пока не отражены, например, теломеразные РНК [307], рибонуклеаза Р [32, 308], энхансерные РНК [309] и другие функциональные РНК.

Исследования последних лет значительно расширили спектр некодирующих РНК, открыв существование фрагментов больших рибонуклеотидных молекул с определенными функциями, генерируемых самыми неожиданными классами хорошо изученных ранее РНК. Так, давно отнесенные к структурным, или РНК «домашнего хозяйства» – рРНК и тРНК – являются генераторами регуляторных РНК-фрагментов. Транспозоны генерируют миРНК, которые их же держат в угнетенном состоянии. Интроны первичной матричной РНК являются местом обитания других малых РНК (например, миртронов и мякРНК), которые воздействуют на экспрессию мРНК.

Важнейшими функциями некодирующих РНК являются регуляторная и каталитическая функции. Реакции, катализируемые рибозимами (РНК, обладающие каталитическим действием), малочисленны и менее разнообразны, чем катализируемые белковыми ферментами, но они имеют важнейшее значение в современных клетках. Первейшим примером является сама рибосома, где ключевая пептидилтрансферазная реакция катализируется рРНК большой субъединицы без непосредственного участия белков [310]. Есть данные о побочных регуляторных функциях, осуществляемых транспортными РНК.

Эпигенетика - это дисциплина, изучающая наследственные изменения в экспрессии генов, не связанные с изменением последовательности ДНК. За последнее десятилетие исследователи показали, что эпигенетическая регуляция играет важную роль в росте, дифференцировке клеток, аутоиммунных заболеваниях и раке. Основные эпигенетические механизмы включают феномен метилирования ДНК, модификации гистонов и регуляцию некодирующими РНК. МикроРНК, пиРНК, эндогенные миРНК и длинные некодирующие РНК являются наиболее распространенными регуляторными

РНК, и, что важно, появляется все больше доказательств того, что именно некодирующие РНК играют важную роль в эпигенетическом контроле.

Хотя подавляющее большинство длинных некодирующих РНК еще предстоит полностью охарактеризовать, многие из этих транскриптов вряд ли представляют транскрипционный «шум», так как было показано, что значительное их количество специфическую экспрессию, локализацию демонстрирует в субклеточных компартментах и связь с болезнями человека. В некоторых случаях оказывается, что простого акта транскрипции некодирующей РНК достаточно, чтобы положительно или отрицательно повлиять на экспрессию близлежащих генов. Однако во многих случаях длинные некодирующие РНК сами выполняют ключевые регуляторные роли, которые, как предполагалось ранее, зарезервированы для белков, такие как регулирование активности или локализации белков, они служат организационными каркасами субклеточных структур. Кроме того, многие длинные некодирующие РНК процессируются с образованием малых РНК или, наоборот, модулируют процессинг других РНК. Таким образом, становится все более очевидным, что длинные некодирующие РНК могут функционировать разными способами и являются ключевыми регуляторными молекулами в клетке.

Исследования транскритомов в настоящее время ведутся очень интенсивно, способствуя появлению новых данных об РНКомах эукариот. Особенно много новых статей в последнее время появляется о функционировании некодирующих РНК в растениях. Эти данные почти не вошли в эту статью, на эту тему есть много хороших обзоров (см., например, [311–313]).

У растений регуляция осуществляется сложнее. Растительная РНК-интерференция требует полной гибридизации малой РНК с мишенью, что значительно ограничивает количество целевых генов, но при этом система кажется еще более сложной (четыре типа DICER, семь dsRBP по сравнению с одним DICER и парой dsRBP у *Homo sapiens*). Более того, механизмы генерации микроРНК тоже отличаются, зрелые микроРНК продуцируются и модифицируются в ядре растений, но эти процессы разделены между ядром и цитоплазмой у животных. Непонятно, как эти разные процессы развивались параллельно с созданием сложных функций AGO (четыре разных AGO у *H. sapiens*, десять – у *A. thaliana*, 26 - y C. elegans), способных использовать эти точно разрезанные микроРНК для регуляции мРНК. Для этого сценария требуется небольшая РНК, действующая вместе с некоторым прототипным белком Ago. В соответствии с этой точкой зрения, последний общий предок эукариот, вероятно, появился после того, как уже были изобретены функциональные принципы РНК-интерференции (а также некоторые его компоненты, кроме, вероятно, микроРНК) [314, 315].

Интерес к изучению РНК постоянно подогревается гипотезой об изначальном мире РНК. Термин «мир РНК» впервые был введен Гилбертом [316]. Но представление о РНК, как о первичной молекуле было сформулировано раньше независимо Криком [317], Оргелем [318] и Вёзе [319]. Не все ученые поддерживают гипотезу о первичности РНК. Например, в работе пионера исследований некодирующих РНК Eddy в 2002 году [320] утверждалось, что многие гены РНК - это недавние филогенетические инновации, которые хорошо адаптированы к их современным ролям в посттранскрипционной регуляции, процессинге РНК и модификации РНК.

Однако идея мира РНК не оставляет умы ученых, которые ищут и находят все новые ее подтверждения. Различные аспекты гипотезы мира РНК и подтверждающие ее данные основательно рассмотрены в классическом сборнике статей «Мир РНК», выдержавшем четыре издания [321]. А в книге Кунина [322] приведены три убедительные свидетельства в поддержку важнейшей роли РНК на самых ранних этапах истории жизни: 1) разнообразие белкового мира сформировалось, когда система трансляции состояла в основном из РНК; 2) в современных клетках действует несколько классов рибозимов, их свойства совместимы с гипотезой о том, что они являются реликтами

первичного мира РНК; 3) рибозимы способны катализировать значительное разнообразие реакций, включая и те, что играют центральную роль в эволюции трансляции. Исследования недавнего времени продолжают находить новые свидетельства первичности РНК. Существует точка зрения, согласно которой современные 23S рРНК, при всём их разнообразии, являются продуктом последовательной эволюции относительно небольшого участка цепи РНК в сложную мультидоменную рибонуклеотидную структуру. В большинстве случаев таким древнейшим «ядром» рибосомы признаётся основная часть V домена 23S рРНК пептидил-трансферазный центр [323]. Именно эта часть рибосомы непосредственно отвечает за реализацию её важнейшей роли в процессе формирования транслируемой пептидной цепи. Еще одним примером такого рода результатов является выявление консервативного пептидного мотива, сохраняющего в одних и тех же участках структурно и функционально различных белков у филогенетически далёких микроорганизмов. Все эти белки принадлежат важнейшим метаболическим путям синтеза тРНК и формирования рибосомы [324].

Открытия последних двух десятков лет в биоинформатике, в том числе, и в области исследования некодирующих РНК, делаются с помощью технологий обработки больших данных [325], получаемых в результате использования высокопроизводительных технологий секвенирования. В дополнение к технологии секвенирования РНК RNA-seq в последнее время было разработано множество высокопроизводительных методов секвенирования для идентификации интерактома (полного графа взаимодействий) некодирующих РНК in vivo, такие, как CHIRP-seq, R-CHIP-seq, DRIP-seq, CLIP-seq, SHAPE-seq, Frag-seq (см. таблицу 2 в [266]). Эти инструменты способны помочь не только в деле обнаружения и каталогизации новых РНК, но и в решении более сложных и интересных проблем. Например, решить задачу составления карт РНК-ДНК-, РНКбелок- и РНК-РНК-взаимодействий в клетке. Или изучать эволюцию разных классов регуляторных РНК методами сравнительной геномики [315]. Эти работы не должны быть нацелены на исследования наличия или отстутствия отдельных локусов РНК в разных органихмах, они могут оценить видо- и кладоспецифичные вариации в идентичности, численности и процессинге конкретного вида РНК, а также сравнить скорости эволюции РНК в зависимости от особенностей их биогенеза и стадии развития организма. Подобные исследования уже проводятся [322] и их становится все больше.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-07-00996. Автор выражает благодарность С. А. Шабалиной за всестороннюю помощь в подборе материала для статьи. Отдельная благодарность А. А. Зимину за внимательное прочтение рукописи и ценные замечания.

СПИСОК ЛИТЕРАТУРЫ

- International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*. 2004. V. 431. P. 931–945. doi: 10.1038/nature03001
- 2. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012. V. 489. P. 57–74 doi: <u>10.1038/nature11247</u>
- Clark M.B., Amaral P.P., Schlesinger F.J., Dinger M.E., Taft R.J., Rinn J.L., Ponting C.P., Stadler P.F., Morris K.V., Morillon A. et al. The reality of pervasive transcription. *PLoS Biol.* 2011. V. 9. Article No. e1000625. doi: <u>10.1371/journal.pbio.1000625</u>
- 4. Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N., Oyama R., Ravasi T., Lenhard B., Wells C. et al. The transcriptional landscape of the mammalian genome. *Science*. 2005. V. 309. P. 1559–1563. doi: <u>10.1126/science.1112014</u>

- Katayama S., Tomaru Y., Kasukawa T., Waki K., Nakanishi M., Nakamura M., Nishida H., Yap C.C., Suzuki M., Kawai J. et al. Antisense transcription in the mammalian transcriptome. *Science*. 2005. V. 309. P. 1564–1566. doi: <u>10.1126/science.1112009</u>
- Birney E., Stamatoyannopoulos J.A., Dutta A., Guigo R., Gingeras T.R., Margulies E.H., Weng Z., Snyder M., Dermitzakis E.T., Thurman R.E. et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*. 2007. V. 447. P. 799–816. doi: 10.1038/nature05874
- Kapranov P., Cheng J., Dike S., Nix D.A., Duttagupta R., Willingham A.T., Stadler P.F., Hertel J., Hackermuller J., Hofacker I.L. et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science*. 2007. V. 316. P. 1484–1488. doi: <u>10.1126/science.1138341</u>
- Bertone P., Stolc V., Royce T.E., Rozowsky J.S., Urban A.E., Zhu X., Rinn J.L., Tongprasit W., Samanta M., Weissman S. et al. Global identification of human transcribed sequences with genome tiling arrays. *Science*. 2004. V. 306. P. 2242–2246. doi: <u>10.1126/science.1103388</u>
- Cheng J., Kapranov P., Drenkow J., Dike S., Brubaker S., Patel S., Long J., Stern D., Tammana H., Helt G. et al. Transcriptional maps of 10 human chromosomes at 5nucleotide resolution. *Science*. 2005. V. 308. P. 1149–1154. doi: <u>10.1126/science.1108625</u>
- Kapranov P., Cawley S.E., Drenkow J., Bekiranov S., Strausberg R.L., Fodor S.P.A., Gingeras T.R. Large-scale transcriptional activity in chromosomes 21 and 22. *Science*. 2002. V. 296. P. 916–919. doi: <u>10.1126/science.1068597</u>
- Rinn J.L., Euskirchen G., Bertone P., Martone R., Luscombe N.M., Hartman S., Harrison P.M., Nelson F.K., Miller P., Gerstein M. et al. The transcriptional activity of human Chromosome 22. *Genes Dev.* 2003. V. 17. P. 529–540. doi: <u>10.1101/gad.1055203</u>
- Kapranov P., Willingham A.T., Gingeras T.R. Genome-wide transcription and the implications for genomic organization. *Nat. Rev. Genet.* 2007. V. 8. P. 413–423. doi: <u>10.1038/nrg2083</u>
- Wang J., Zhang J., Zheng H., Li J., Liu D., Li H., Samudrala R., Yu J., Wong G.K. Mouse transcriptome: Neutral evolution of 'non-coding' complementary DNAs. *Nature*. 2004. V. 431. P. 757. doi: <u>10.1038/nature03016</u>
- 14. Struhl K. Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nat. Struct. Mol. Biol.* 2007. V. 14. P. 103–105. doi: <u>10.1038/nsmb0207-103</u>
- 15. Ebisuya M., Yamamoto T., Nakajima M., Nishida E. Ripples from neighbouring transcription. *Nat. Cell Biol.* 2008. V. 10. P. 1106–1113. doi: <u>10.1038/ncb1771</u>
- 16. Palazzo A.F., Lee E.S. Non-coding RNA: what is functional and what is junk? *Front. Genet.* 2015. V. 6. doi: <u>10.3389/fgene.2015.00002</u>
- Mattick J.S., Taft R.J., Faulkner G.J. A global view of genomic information-moving beyond the gene and the master regulator. *Trends Genet*. 2010. V. 26. P. 21–28 doi: <u>10.1016/j.tig.2009.11.002</u>
- Huttenhofer A., Brosius J., Bachellerie J.P. RNomics: identification and function of small, non-messenger RNAs. *Curr. Opin. Chem. Biol.* 2002. V. 6. P. 835–843. doi: <u>10.1016/s1367-5931(02)00397-6</u>
- 19. Huttenhofer A., Schattner P., Polacek N. Non-coding RNAs: hope or hype? *Trends Genet*. 2005. V. 21. P. 289–297.
- 20. Eddy S.R. Non-coding RNA genes and the modern RNA world. *Nature Reviews Genetics*. 2001. V. 2. P. 919–929. doi: 10.1038/35103511
- 21. Kawano M., Reynolds A.A., Miranda-Rios J., Storz G. Detection of 5' and 3'-UTRderived small RNAs and cis-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Res*. 2005. V. 33. P. 1040–1050.
- 22. Mattick J.S. RNA regulation: a new genetics? Nat. Rev. Genet. 2004. V. 5. P. 316–323.

- 23. Mattick J.S., Makunin I.V. Small regulatory RNAs in mammals. *Hum. Mol. Genet.* 2005. V. 14. P. R121–R132.
- 24. Weinberg R.A., Penman S. Small molecular weight monodisperse nuclear RNA. J. Mol. Biol. 1968. V. 38. P. 289–304.
- 25. Zieve G.W. Two groups of small stable RNAs. Cell. 1981. V. 25. P. 296–297.
- Busch H., Reddy R., Rothblum L., Choi Y.C. SnRNAs, SnRNPs, and RNA processing. *Annu. Rev. Biochem.* 1982. V. 51. P. 617–654. doi: 10.1146/annurev.bi.51.070182.003153
- Yang V.W., Lerner M.R., Steitz J.A., Flint S.J. A small nuclear ribonucleoprotein is required for splicing of adenoviral early RNA sequences. *PNAS*. 1981. V. 78. P. 1371– 1375. doi: <u>10.1073/pnas.78.3.1371</u>
- Yu Y.T., Scharl E.C., Smith C.M., Steitz J.A. In: *The RNA World*, 2nd edn. Eds. Gesteland R.F., Cech T.R., Atkins J.F. New York: Cold Spring Harbor Laboratory Press, 1999. P. 487–524.
- 29. Burge C.B., Tuschl T., Sharp P.A. Splicing of precursors to mRNAs by the spliceosome. In: *The RNA World*, 2nd edn. Eds. Gesteland R.F., Cech T.R., Atkins J.F. New York: Cold Spring Harbor Laboratory Press, 1999. P. 525–560.
- 30. Tarn W.Y., Steitz J.A. Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. *Science*. 1996. V. 273. P. 1824–1832.
- Sharp P.A., Burge C.B. Classification of introns: U2-type or U12-type. *Cell*. 1997. V. 91. P. 875–879.
- 32. Stark B.C., Kole R., Bowman E.J., Altman S. Ribonuclease P: an enzyme with an essential RNA component. *PNAS*. 1978. V. 75. P. 3717–3721. doi: <u>10.1073/pnas.75.8.3717</u>
- Walter P., Blobel G. Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. *Nature*. 1982. V. 299. P. 691–698. doi: <u>10.1038/299691a0</u>
- 34. Lewin R. Surprising discovery with a small RNA. Science. 1982. V. 218. P. 777–778.
- 35. Bartel D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*. 2004. V. 116. P. 281–297.
- 36. Bartel D.P., Chen C.Z. Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nat. Rev. Genet.* 2004. V. 5. P. 396–400.
- 37. Fragnet L., Kut E., Rasschaert D. Comparative functional study of the viral telomerase RNA based on natural mutations. *J. Biol. Chem.* 2005. V. 280. P. 23502–23515.
- 38. Plath K., Mlynarczyk-Evans S., Nusinow D.A., Panning B. Xist RNA and the mechanism of X chromosome inactivation. *Annu. Rev. Genet.* 2002. V. 36. P. 233–278.
- Brockdorff N., Ashworth A., Kay G.F., McCabe V.M., Norris D.P., Cooper P.J., Swift S., Rastan S. The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell*. 1992. V. 71. P. 515–526. doi: <u>10.1016/0092-8674(92)90519-1</u>
- 40. Kelley R.L., Kuroda M.L. Noncoding RNA genes in dosage compensation and imprinting. *Cell*. 2000. V. 103. P. 9–12.
- 41. Sleutels F., Zwart R., Barlow D.P. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature*. 2002. V. 415. P. 810–813.
- 42. Ulitsky I., Bartel D.P. lincRNAs: genomics, evolution, and mechanisms. *Cell*. 2013. V. 154. P. 26–46. doi: <u>10.1016/j.cell.2013.06.020</u>
- 43. Mutz K.O., Heilkenbrinker A., Lönne M., Walter J.G., Stahl F. Transcriptome analysis using next-generation sequencing. *Current Opinion in Biotechnology*. 2013. V. 24. № 1. P. 22–30.
- 44. Lee R.C., Feinbaum R.L., Ambros V. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*. 1993. V. 75. P. 843–854.
- 45. Fabian M.R., Sonenberg N., Filipowicz W. Regulation of mRNA translation and stability by microRNAs. *Annu. Rev. Biochem.* 2010. V. 79. P. 351–379.

- Bissels U., Wild S., Tomiuk S., Holste A., Hafner M., Tuschl T., Bosio A. Absolute quantification of microRNAs by using a universal reference. *RNA*. 2009. V. 15. P. 2375–2384. doi: <u>10.1261/rna.1754109</u>
- 47. Hamilton A.J., Baulcombe D.C. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*. 1999. V. 286. P 950–952.
- Ishizu H., Siomi H., Siomi M.C. Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. *Genes Dev.* 2012. V. 26. P. 2361–2373. doi: <u>10.1101/gad.203786.112</u>
- 49. Hansen T.B., Jensen T.I., Clausen B.H., Bramsen J.B., Finsen B., Damgaard C.K., Kjems J. Natural RNA circles function as efficient microRNA sponges. *Nature*. 2013. V. 495. № 7441. P. 384–388. doi: <u>10.1038/nature11993</u>
- 50. Memczak S., Jens M., Elefsinioti A., Torti F., Krueger J., Rybak A., Maier L., Mackowiak S.D., Gregersen L.H., Munschauer M. et al. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature*. 2013. V. 495. № 7441. P. 333–338.
- Valdmanis P.N., Kay M.A. The expanding repertoire of circular RNAs. *Mol. Ther.* 2013.
 V. 21. № 6. P. 1112–1114. doi: <u>10.1038/mt.2013.101</u>
- 52. Lee S.R., Collins K. Starvation-induced cleavage of the tRNA anticodon loop in *Tetrahymena thermophila*. J. Biol. Chem. 2005. V. 280. P. 42744–42749.
- 53. Haiser H.J., Karginov F.V., Hannon G.J., Elliot M.A. Developmentally regulated cleavage of tRNAs in the bacterium *Streptomyces coelicolor*. *Nucleic Acids Res.* 2008. V. 36. P. 732–741.
- 54. Chak L.L., Mohammed J., Lai E.C., Tucker-Kellogg G., Okamura K. A deeply conserved, noncanonical miRNA hosted by ribosomal DNA. *RNA*. 2015. V. 21. P. 375–384.
- 55. Asha S., Soniya E.V. The sRNAome mining revealed existence of unique signature small RNAs derived from 5.8SrRNA from Piper nigrum and other plant lineages. *Sci. Rep.* 2017. V. 7. Article No. 41052.
- 56. Chen Z., Sun Y., Yang X., Wu Z., Guo K., Niu X., Wang Q., Ruan J., Bu W., Gao S. Two featured series of rRNA-derived RNA fragments (rRFs) constitute a novel class of small RNAs. *PLoS One.* 2017. V. 12. Article No. e0176458.
- 57. Cho J. Transposon-Derived Non-coding RNAs and Their Function in Plants. *Front. Plant Sci.* 2018. V. 9. Article No. 600. doi: <u>10.3389/fpls.2018.00600</u>
- Wolf S.F., Schlessinger D. Nuclear metabolism of ribosomal RNA in growing, methionine-limited, and ethionine-treated HeLa cells. *Biochemistry (Mosc.)* 1977. V. 16. P. 2783–2791. doi: <u>10.1021/bi00631a031</u>
- Natsidis P., Schiffer P.H., Salvador-Martínez I., Telford M.J. Computational discovery of hidden breaks in 28S ribosomal RNAs across eukaryotes and consequences for RNA Integrity Numbers. *Sci. Rep.* 2019. V. 9. Article No. 19477. doi: <u>10.1038/s41598-019-55573-1</u>
- 60. Waldron C., Lacroute F. Effect of growth rate on the amounts of ribosomal and transfer ribonucleic acids in yeast. *J. Bacteriol.* 1975. V. 122. P. 855–865.
- 61. Matera A.G., Terns R.M., Terns M.P. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat. Rev. Mol. Cell Biol.* 2007. V 8. P. 209–220. doi: <u>10.1038/nrm2124</u>
- 62. Dieci G., Preti M., Montanini B. Eukaryotic snoRNAs: a paradigm for gene expression flexibility. *Genomics*. 2009. V. 94. P. 83–88. doi: <u>10.1016/j.ygeno.2009.05.002</u>
- Aravin A., Tuschl T. Identification and characterization of small RNAs involved in RNA silencing. *FEBS Lett.* 2005. V. 579. № 26. P. 5830–5840. doi: 10.1016/j.febslet.2005.08.009
- Siomi M., Sato K., Pezic D., Aravin A.A. PIWI-interacting small RNAs: the vanguard of genome defence. *Nat. Rev. Mol. Cell. Biol.* 2011. V. 12. P. 246–258. doi: <u>10.1038/nrm3089</u>

- Lee Y.S., Shibata Y., Malhotra A., Dutta A. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev.* 2009. V. 23. P. 2639–2649. doi: <u>10.1101/gad.1837609</u>
- 66. Guan L., Grigoriev A. Computational meta-analysis of ribosomal RNA fragments: potential targets and interaction mechanisms. *Nucleic Acids Res.* 2021. V. 49. № 7. P. 4085–4103. doi: <u>10.1093/nar/gkab190</u>
- 67. Wilusz J.E., Sunwoo H., Spector D.L. Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev.* 2009. V. 23. № 13. P. 1494–1504. doi: <u>10.1101/gad.1800909</u>
- Chen L., Huang C., Wang X., Shan G. Circular RNAs in eukaryotic cells. *Curr. Genom.* 2015. V. 16. № 5. P. 312–318. doi: <u>10.2174/1389202916666150707161554</u>
- 69. Gerbi S.A. Expansion segments: regions of variable size that interrupt the universal core secondary structure of ribosomal RNA. In: *Ribosomal RNA: Structure, Evolution, Processing and Function in Protein Synthesis.* Eds: R.A. Zimmermann and A.E. Dahlberg. Boca Raton, FL: Telford CRC Press, 1996. P. 71–87.
- Armache J.-P., Jarasch A., Anger A.M., Villa E., Becker T., Bhushan S., Jossinet F., Habeck M., Dindar G., Franckenberg S., et al. Cryo-EM structure and rRNA model of a translating eukaryotic 80S ribosome at 5.5-A resolution. *PNAS*. 2010. V. 107. P. 19748– 19753.
- Anger A.M., Armache J.-P., Berninghausen O., Habeck M., Subklewe M., Wilson D.N., Beckmann R. Structures of the human and Drosophila 80S ribosome. *Nature*. 2013. V. 497. P. 80–85.
- 72. Fujii K., Susanto T.T., Saurabh S., Barna M. Decoding the function of expansion segments in ribosomes. *Mol. Cell.* 2018. V. 72. P. 1013–1020.
- 73. Chan P.P., Lowe T.M. GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res.* 2009. V. 37. P. D93–D97.
- 74. Parisien M., Wang X., Pan T. Diversity of human tRNA genes from the 1000-genomes project. *RNA Biol.* 2013. V. 10. P. 1853–1867.
- 75. Orioli A. tRNA biology in the omics era: stress signalling dynamics and cancer progression. *Bioessays*. 2017. V. 39. Article No. 1600158. doi: 10.1002/bies.201600158
- 76. Kirchner S., Ignatova Z. Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. *Nat. Rev. Genet.* 2015. V. 16. P. 98–112.
- Huang S.Q., Sun B., Xiong Z.P., Shu Y., Zhou H.H., Zhang W., Xiong J., Li Q. The dysregulation of tRNAs and tRNA derivatives in cancer. *J. Exp. Clin. Cancer Res.* 2018. V. 37. Article No. 101.
- 78. Santos M., Fidalgo A., Varanda A.S., Oliveira C., Santos M.A.S. tRNA deregulation and its consequences in cancer. *Trends Mol. Med.* 2019. V. 25. P. 853–865.
- 79. Gomez-Roman N., Grandori C., Eisenman R.N., White R.J. Direct activation of RNA polymerase III transcription by c-Myc. *Nature*. 2003. V. 421. P. 290–294.
- 80. Felton-Edkins Z.A., Fairley J.A., Graham E.L., Johnston I.M., White R.J., Scott P.H. The mitogen-activated protein (MAP) kinase ERK induces tRNA synthesis by phosphorylating TFIIIB. *EMBO J.* 2003. V. 22. P. 2422–2432.
- 81. Wei Y., Tsang C.K., Zheng X.F. Mechanisms of regulation of RNA polymerase IIIdependent transcription by TORC1. *EMBO J.* 2009. V. 28. P. 2220–2230.
- 82. Truitt M.L., Ruggero D. New frontiers in translational control of the cancer genome. *Nat. Rev. Cancer*. 2016. V. 16. P. 288–304.
- Beznosková P., Bidou L., Namy O., Valášek L.S. Increased expression of tryptophan and tyrosine tRNAs elevates stop codon readthrough of reporter systems in human cell lines. *Nucleic Acids Res.* 2021. V. 49. № 9. P. 5202–5215. doi: <u>10.1093/nar/gkab315</u>
- 84. Goodenbour J.M., Pan T. Diversity of tRNA genes in eukaryotes. *Nucleic Acids Res.* 2006. V. 34. P. 6137–6146.
- 85. Mahlab S., Tuller T., Linial M. Conservation of the relative tRNA composition in healthy and cancerous tissues. *RNA*. 2012. V. 18. P. 640–652.

- Pavon-Eternod M., Gomes S., Geslain R., Dai Q., Rosner M.R., Pan T. tRNA overexpression in breast cancer and functional consequences. *Nucleic Acids Res.* 2009. V. 37. P. 7268–7280.
- Gingold H., Tehler D., Christoffersen N.R., Nielsen M.M., Asmar F., Kooistra S.M., Christophersen N.S., Christensen L.L., Borre M., Sorensen K.D. et al. A dual program for translation regulation in cellular proliferation and differentiation. *Cell*. 2014. V. 158. P. 1281–1292.
- Goodarzi H., Nguyen H.C.B., Zhang S., Dill B.D., Molina H., Tavazoie S.F. Modulated expression of specific tRNAs drives gene expression and cancer progression. *Cell*. 2016. V. 165. P. 1416–1427.
- 89. Jackson R.J., Hellen C.U., Pestova T.V. Termination and post-termination events in eukaryotic translation. *Adv. Protein Chem. Struct. Biol.* 2012. V. 86. P. 45–93.
- Valasek L.S., Zeman J., Wagner S., Beznoskova P., Pavlikova Z., Mohammad M.P., Hronova V., Herrmannova A., Hashem Y., Gunisova S. Embraced by eIF3: structural and functional insights into the roles of eIF3 across the translation cycle. *Nucleic Acids Res.* 2017. V. 45. P. 10948–10968.
- 91. Dabrowski M., Bukowy-Bieryllo Z., Zietkiewicz E. Translational readthrough potential of natural termination codons in eucaryotes–The impact of RNA sequence. *RNA Biol.* 2015. V. 12. P. 950–958.
- Schueren F., Thoms S. Functional translational readthrough: a systems biology perspective. *PLoS Genet*. 2016. V. 12. Article No. e1006196. doi: <u>10.1371/journal.pgen.1006196</u>
- 93. Schueren F., Lingner T., George R., Hofhuis J., Dickel C., Gartner J., Thoms S. Peroxisomal lactate dehydrogenase is generated by translational readthrough in mammals. *Elife*. 2014. V. 3. Article No. e03640.
- 94. Loughran G., Chou M.Y., Ivanov I.P., Jungreis I., Kellis M., Kiran A.M., Baranov P.V., Atkins J.F. Evidence of efficient stop codon readthrough in four mammalian genes. *Nucleic Acids Res.* 2014. V. 42. P. 8928–8938.
- 95. Avcilar-Kucukgoze I., Kashina A. Hijacking tRNAs from Translation: Regulatory Functions of tRNAs in Mammalian Cell Physiology. *Front. Mol. Biosci.* 2020. V. 7. Article No. 610617. doi: <u>10.3389/fmolb.2020.610617</u>
- 96. Torres A.G. Enjoy the Silence: Nearly Half of Human tRNA Genes Are Silent. *Bioinformatics Biol. Insights.* 2019. V. 13. doi: <u>10.1177/1177932219868454</u>
- 97. Soma A., Onodera A., Sugahara J., Kanai A., Yachie N., Tomita M., Kawamura F., Sekine Y. Permuted tRNA Genes Expressed via a Circular RNA Intermediate in Cyanidioschyzon Merolae. Science. 2007. V. 318. № 5849. P. 450–453. doi: 10.1126/science.1145718
- Randau L., Münch R., Hohn M.J., Jahn D., Söll D. Nanoarchaeum Equitans Creates Functional tRNAs from Separate Genes for Their 5'- and 3'-halves. *Nature*. 2005. V. 433. P. 537–541. doi: <u>10.1038/nature03233</u>
- 99. Fujishima K., Sugahara J., Kikuta K., Hirano R., Sato A., Tomita M., Kanai A. Tri-split tRNA Is a Transfer RNA Made from 3 Transcripts that Provides Insight into the Evolution of Fragmented tRNAs in Archaea. *PNAS*. 2009. V. 106. № 8. P. 2683–2687. doi: 10.1073/pnas.0808246106
- 100. Cheng J., Kapranov P., Drenkow J., Dike S., Brubaker S., Patel S., Long J., Stern D., Tammana H., Helt G. et al. Transcriptional maps of 10 human chromosomes at 5nucleotide resolution. *Science*. 2005. V. 308. P. 1149–1154. doi: <u>10.1126/science.1108625</u>
- 101. Mattick J.S., Makunin I.V. Non-coding RNA. Hum. Mol. Genet. 2006. V. 15. P. R17–R29.
- 102. Huttenhofer A., Schattner P., Polacek N. Noncoding RNAs: hope or hype? *Trends Genet*. 2005. V. 21. P. 289–297.

- 103. Storz G., Altuvia S., Wassarman K.M. An abundance of RNA regulators. Annu. Rev. Biochem. 2005. V. 74. P. 199–217.
- 104. Matera A.G., Wang Z. A day in the life of the spliceosome. *Nat. Rev. Mol. Cell Biol.* 2014.
 V. 15. № 2. P. 108–121. doi: <u>10.1038/nrm3742</u>
- 105. Szymanski M., Barciszewska M.Z., Zywicki M., Barciszewski J. Noncoding RNA transcripts. J. Appl. Genet. 2003. V. 44. P. 1–19.
- 106. Matera A.G., Terns R.M., Terns M.P. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat. Rev. Mol. Cell Biol.* 2007. V. 8. № 3. P. 209–220. doi: 10.1038/nrm2124
- 107. Bohnsack M.T., Sloan K.E. Modifications in small nuclear RNAs and their roles in spliceosome assembly and function. *Biological Chemistr*. 2018. V. 399. № 11. P. 1265–1276. doi: <u>https://doi.org/10.1515/hsz-2018-0205</u>
- 108. Will C.L., Lührmann R. Spliceosome structure and function. *Cold Spring Harb. Perspect. Biol.* 2011. V. 3. doi: <u>10.1101/cshperspect.a003707</u>
- 109. Hoskins A.A., Moore M.J. The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem. Sci.* 2012. V. 37. P. 179–188.
- Karijolich J., Yu Y.T. Spliceosomal snRNA modifications and their function. *RNA Biol.* 2010. V. 7. P. 192–204.
- Morais P., Adachi H., Yu Y.T. Spliceosomal snRNA Epitranscriptomics. *Front Genet*. 2021. V. 12. Article No. 652129. doi: <u>10.3389/fgene.2021.652129</u>
- 112. Stein C.A., Castanotto D. FDA-approved oligonucleotide therapies in 2017. *Mol. Ther.* 2017. V. 25. P. 1069–1075. doi: <u>10.1016/j.ymthe.2017.03.023</u>
- 113. Rüger J., Ioannou S., Castanotto D., Stein C.A. Oligonucleotides to the (gene) rescue: FDA approvals 2017-2019. *Trends Pharmacol. Sci.* 2020. V. 41. P. 27–41. doi: <u>10.1016/j.tips.2019.10.009</u>
- 114. Eliceiri G.L. Small nucleolar RNAs. *Cell. Mol. Life Sci.* 1999. V. 56. P. 22–31. doi: 10.1007/s000180050003
- 115. Matera A.G., Terns R.M., Terns M.P. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat. Rev. Mol. Cell Biol.* 2007. V. 8. P. 209–220. doi: <u>10.1038/nrm2124</u>
- 116. Boivin V., Faucher-Giguère L., Scott M., Abou-Elela S. The cellular landscape of midsize noncoding RNA. Wiley Interdiscip Rev RNA. 2019. V. 10. Article No. e1530. doi: <u>10.1002/wrna.1530</u>
- 117. Bergeron D., Fafard-Couture É., Scott M.S. Small nucleolar RNAs: continuing identification of novel members and increasing diversity of their molecular mechanisms of action. *Biochem. Soc. Trans.* 2020. V. 48. P. 645–656. doi: 10.1042/BST20191046
- 118. Kiss T. Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. EMBO J. 2001. V. 20. P. 3617–3622. doi: <u>10.1093/emboj/20.14.3617</u>
- Bergeron D., Laforest C., Carpentier S., Calvé A., Fafard-Couture É., Deschamps-Francoeur G., Scott M.S. SnoRNA copy regulation affects family size, genomic location and family abundance levels. *BMC Genomics*. 2021. V. 22. Article No. 414. doi: <u>10.1186/s12864-021-07757-1</u>
- Deschamps-Francoeur G., Garneau D., Dupuis-Sandoval F., Roy A., Frappier M., Catala M., Couture S., Barbe-Marcoux M., Abou-Elela S., Scott M.S. Identification of discrete classes of small nucleolar RNA featuring different ends and RNA binding protein dependency. *Nucleic Acids Res.* 2014. V. 42. P. 10073–10085. doi: <u>10.1093/nar/gku664</u>
- 121. Marmier-Gourrier N., Cléry A., Senty-Ségault V., Charpentier B., Schlotter F., Leclerc F., Fournier R., Branlant C. A structural, phylogenetic, and functional study of 15.5-kD/Snu13 protein binding on U3 small nucleolar RNA. *RNA*. 2003. V. 9. P. 821– 838. doi: <u>10.1261/rna.2130503</u>

- 122. Lestrade L., Weber M.J. snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res.* 2006. V. 34. P. D158. doi: 10.1093/nar/gkj002
- 123. Ganot P., Caizergues-Ferrer M., Kiss T. The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. *Genes Dev.* 1997. V. 11. P. 941–956. doi: 10.1101/gad.11.7.941
- 124. Ganot P., Bortolin M.L., Kiss T. Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell.* 1997. V. 89. P. 799–809. doi: <u>10.1016/S0092-8674(00)80263-9</u>
- 125. Brown J.W., Marshall D.F., Echeverria M. Intronic noncoding RNAs and splicing. *Trends Plant Sci.* 2008. V. 13. P. 335–342.
- 126. Chanfreau G., Legrain P., Jacquier A. Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism. *J. Mol. Biol.* 1998. V. 284. P. 975–988.
- Petfalski E., Dandekar T., Henry Y., Tollervey D. Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. *Mol. Cell. Biol.* 1998. V. 18. P. 1181–1189.
- 128. Qu L.H., Henras A., Lu Y.J., Zhou H., Zhou W.X., Zhu Y.Q., Zhao J., Henry Y., Caizergues-Ferrer M., Bachellerie J.P. Seven novel methylation guide small nucleolar RNAs are processed from a common polycistronic transcript by Rat1p and RNase III in yeast. *Mol. Cell. Biol.* 1999. V. 19. P. 1144–1158.
- 129. Dupuis-Sandoval F., Poirier M., Scott M.S. The emerging landscape of small nucleolar RNAs in cell biology. Wiley Interdiscip. Rev. RNA. 2015. V. 6. P. 381–397. doi: 10.1002/wrna.1284
- Bratkovič T., Bozič J., Rogelj B. Functional diversity of small nucleolar RNAs. *Nucleic Acids Res.* 2020. V. 48. P. 1627–1651. doi: <u>10.1093/nar/gkz1140</u>
- Falaleeva M., Welden J.R., Duncan M.J., Stamm S. C/D-box snoRNAs form methylating and non-methylating ribonucleoprotein complexes: old dogs show new tricks. *BioEssays*. 2017. V. 39. doi: <u>10.1002/bies.201600264</u>
- 132. Shabalina S.A., Koonin E.V. Origins and evolution of eukaryotic RNA interference. *Trends Ecol. Evol.* (*Amst.*). 2008. V. 23. № 10. P. 578–587.
- 133. Wilson R.C., Doudna J.A. Molecular mechanisms of RNA interference. Annu. Rev. Biophys. 2013. V. 42. P. 217–239. doi: 10.1146/annurev-biophys-083012-130404
- 134. Fabian M.R., Sonenberg N. The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. *Nature Structural & Molecular Biology*. 2012. V. 19. № 6. P. 586–593.
- Carthew R.W., Sontheimer E.J. Origins and Mechanisms of miRNAs and siRNAs. *Cell*. 2009. V. 136. № 4. P. 642–655.
- 136. Lu M., Zhang Q., Deng M., Miao J., Guo Y., Gao W., Cui Q. An analysis of human microRNA and disease associations. *PLoS ONE*. 2008. V. 3. № 10. Article No. e3420.
- Kubowicz P., Zelaszczyk D., Pekala E. RNAi in clinical studies. *Curr. Med. Chem.* 2013. V. 20. P. 1801–1816.
- Martinez T., Wright N., López-Fraga M., Jiménez A.I., Pañeda C. Silencing human genetic diseases with oligonucleotide-based therapies. *Hum. Genet.* 2013. V. 132. P. 481– 493.
- 139. Ramachandran P.V., Ignacimuthu S. RNA interference–a silent but an efficient therapeutic tool. *Appl. Biochem. Biotechnol.* 2013. V. 169. P. 1774–1789.
- 140. Davidson B.L., McCray P.B. Current prospects for RNA interference-based therapies; *Nat. Rev. Genet.* 2011. V. 12. P. 329–340. doi: <u>10.1038/nrg2968</u>
- 141. Setten R.L., Rossi J.J., Han S. The current state and future directions of RNAi-based therapeutics. *Nature Reviews Drug Discovery*. 2019. V. 18. № 6. P. 421–446. doi: 10.1038/s41573-019-0017-4

- 142. Carthew R.W., Sontheimer E.J. Origins and Mechanisms of miRNAs and siRNAs. *Cell*. 2009. V. 136. № 4. P. 642–655.
- 143. Saini H.K., Griffiths-Jones S., Enright A.J. Genomic analysis of human microRNA transcripts. *PNAS*. 2007. V. 104. № 45. P. 17719–17724.
- 144. Han J., Lee Y., Yeom K.-H., Nam J.-W., Heo I., Rhee J.-K., Sohn S.Y., Cho Y., Zhang B.-T., Kim V.N. Molecular Basis for the Recognition of Primary microRNAs by the Drosha-DGCR8 Complex. *Cell*. 2006. V. 125. № 5. P. 887–901.
- 145. Lund E., Dahlberg J.E. Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs. *Cold Spring Harb. Symp. Quant. Biol.* 2006. V. 71. P. 59–66.
- 146. Schwarz D.S., Hutvágner G., Du T., Xu Z., Aronin N., Zamore P.D. Asymmetry in the assembly of the RNAi enzyme complex. *Cell*. 2003. V. 115. № 2. P. 199–208.
- 147. MacRae I.J., Ma E., Zhou M., Robinson C.V., Doudna J.A. In vitro reconstitution of the human RISC-loading complex. *PNAS*. 2008. V. 105. № 2. P. 512–517.
- 148. Eulalio A., Huntzinger E., Izaurralde E. GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. *Nature Structural & Molecular Biology*. 2008. V. 15. № 4. P. 346–353.
- 149. Flynt A.S., Lai E.C. Biological principles of microRNA-mediated regulation: shared themes amid diversity. *Nat. Rev. Genet.* 2008. V. 9. P. 831–842.
- 150. Axtell M.J., Westholm J.O., Lai E.C. Vive la différence: biogenesis and evolution of microRNAs in plants and animals. *Genome Biol.* 2011. V. 12. P. 221.
- Westholm J.O., Lai E.C. Mirtrons: microRNA biogenesis via splicing. *Biochimie*. 2011. V. 93. P. 1897–1904.
- 152. Czech B., Hannon G.J. Small RNA sorting: matchmaking for Argonautes. *Nat. Rev. Genet.* 2010. V. 12. P. 19–31.
- 153. Ambros V., Bartel B., Bartel D.P., Burge C.B., Carrington J.C., Chen X., Dreyfuss G., Eddy S.R., Griffiths-Jones S., Marshall M. et al. A uniform system for microRNA annotation. *RNA*. 2003. V. 9. P. 277–279.
- 154. Yang J.S., Lai E.C. Alternative miRNA biogenesis pathways and the interpretation of core miRNA pathway mutants. *Mol. Cell.* 2011. V. 43. P. 892–903. doi: <u>10.1016/j.molcel.2011.07.024</u>
- 155. Okamura K., Hagen J.W., Duan H., Tyler D.M., Lai E.C. The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. Cell. 2007. V. 130. P. 89–100.
- 156. Ruby J.G., Jan C.H., Bartel D.P. Intronic microRNA precursors that bypass Drosha processing. *Nature*. 2007. V. 448. P. 83–86.
- 157. Flynt A.S., Chung W.J., Greimann J.C., Lima C.D., Lai E.C. microRNA biogenesis via splicing and exosome-mediated trimming in *Drosophila*. *Mol. Cell*. 2010. V. 38. P. 900– 907.
- 158. Ladewig E., Okamura K., Flynt A.S., Westholm J.O., Lai E.C. Discovery of hundreds of mirtrons in mouse and human small RNA data. *Genome Res.* 2012. V. 22. № 9. P. 1634– 1645. doi: 10.1101/gr.133553.111
- 159. Wen J., Ladewig E., Shenker S., Mohammed J., Lai E.C. Analysis of Nearly One Thousand Mammalian Mirtrons Reveals Novel Features of Dicer Substrates. *PLoS Comput. Biol.* 2015. V. 11. № 9. Article No. e1004441. doi: 10.1371/journal.pcbi.1004441
- Westholm J.O., Ladewig E., Okamura K., Robine N., Lai E.C. Common and distinct patterns of terminal modifications to mirtrons and canonical microRNAs. *RNA*. 2012. V. 18. P. 177–192.
- Lai E.C. microRNAs: Runts of the genome assert themselves. Curr. Biol. 2003. V. 13. P. R925–R936.
- 162. Doench J.G., Sharp P.A. Specificity of microRNA target selection in translational repression. *Genes Dev.* 2004. V. 18. № 5. P. 504–511. doi: 10.1101/gad.1184404

- 163. Brennecke J., Stark A., Russell R.B., Cohen S.M. Principles of microRNA-target recognition. *PLoS Biol.* 2005. V. 3. № 3. Article No. e85. doi: 10.1371/journal.pbio.0030085
- 164. Lewis B.P., Burge C.B., Bartel D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*. 2005. V. 120. № 1. P. 15–20. doi: 10.1016/j.cell.2004.12.035
- 165. Bartel D.P. MicroRNAs: target recognition and regulatory functions. *Cell*. 2009. V. 136. № 2. P. 215–233.
- 166. Krek A., Grün D., Poy M.N., Wolf R., Rosenberg L., Epstein E.J., MacMenamin P., da Piedade I., Gunsalus K.C., Stoffel M., Rajewsky N. Combinatorial microRNA target predictions. *Nat. Genet.* 2005. V. 37. № 5. P. 495–500. doi: <u>10.1038/ng1536</u>
- 167. Friedman R.C., Farh K.K., Burge C.B., Bartel D.P. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res. 2009. V. 19. № 1. P. 92–105. doi: 10.1101/gr.082701.108
- 168. Simkin A., Geissler R., McIntyre A.B.R., Grimson A. Evolutionary dynamics of microRNA target sites across vertebrate evolution. *PLoS Genet*. 2020. V. 16. № 2. Article No. e1008285. doi: <u>10.1371/journal.pgen.1008285</u>
- 169. Zamore P.D., Tuschl T., Sharp P.A., Bartel D.P. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*. 2000. V. 101. P. 25–33.
- 170. Elbashir S.M., Lendeckel W., Tuschl T. RNA interference is mediated 1- and 22nucleotide RNAs. *Genes Dev.* 2001. V. 15. P. 188–200.
- 171. Sijen T., Plasterk R.H. Transposon silencing in the *Caenorhabditis elegans* germline by natural RNAi. *Nature*. 2003. V. 426. P. 310–314.
- 172. Shi H., Djikeng A., Tschudi C., Ullu E. Argonaute protein in the early divergent eukaryote *Trypanosoma brucei*: Control of small interfering RNA accumulation and retroposon transcript abundance. *Mol. Cell. Biol.* 2004. V. 24. P. 420–427.
- 173. Lippman Z., Martienssen R. The role of RNA interference in heterochromatic silencing. *Nature*. 2004. V. 431. P. 364–370.
- 174. Peragine A., Yoshikawa M., Wu G., Albrecht H.L., Poethig R.S. SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of *trans*acting siRNAs in *Arabidopsis*. *Genes & Dev.* 2004. V. 18. P. 2368–2379. doi: <u>10.1101/gad.1231804</u>
- 175. Borsani O., Zhu J., Verslues P.E., Sunkar R., Zhu J.K. Endogenous siRNAs derived from a pair of natural *cis*-antisense transcripts regulate salt tolerance in *Arabidopsis*. *Cell*. 2005. V. 123. P. 1279–1291.
- 176. Ambros V., Lee R.C., Lavanway A., Williams P.T., Jewell D. MicroRNAs and other tiny endogenous RNAs in *C. elegans. Curr. Biol.* 2003. V. 13. P. 807–818.
- 177. Chen P.Y., Manninga H., Slanchev K., Chien M., Russo J.J., Ju J., Sheridan R., John B., Marks D.S., Gaidatzis D., et al. The developmental miRNA profiles of zebrafish as determined by small RNA cloning. *Genes & Dev.* 2005. V. 19. P. 1288–1293.
- 178. Aravin A.A. The small RNA profile during *Drosophila melanogaster* development. *Dev. Cell.* 2003. V. 5. P. 337–350. doi: 10.1016/s1534-5807(03)00228-4
- 179. Moldovan D., Spriggs A., Dennis E.S., Wilson I.W. The hunt for hypoxia responsive natural antisense short interfering RNAs. *Plant Signaling & Behavior*. 2010. V. 5. P. 247– 251. doi: <u>10.4161/psb.5.3.10548</u>
- 180. Katiyar-Agarwal S., Morgan R., Dahlbeck D., Borsani O., Villegas A., Zhu J.-K., Staskawicz B.J., Jin H. A pathogen-inducible endogenous siRNA in plant immunity. *PNAS*. 2006. V. 103. № 47. P. 18002–18007. doi: 10.1073/pnas.0608258103
- 181. Held M.A., Penning B., Brandt A.S., Kessans S.A., Yong W., Scofield S.R., Carpita N.C. Small-interfering RNAs from natural antisense transcripts derived from a cellulose

synthase gene modulate cell wall biosynthesis in barley. *PNAS*. 2008. V. 105. № 51. P. 20534–20539. doi: <u>10.1073/pnas.0809408105</u>

- 182. Zhang X., Xia J., Lii Y.E., Barrera-Figueroa B.E., Zhou X., Gao S., Lu L., Niu D., Chen Z., Leung C. et al.Genome-wide analysis of plant nat-siRNAs reveals insights into their distribution, biogenesis and function. *Genome Biology*. 2012. V. 13. № 20. P. R20. doi: <u>10.1186/gb-2012-13-3-r20</u>
- 183. Watanabe T., Takeda A., Tsukiyama T., Mise K., Okuno T., Sasaki H., Minami N., Imai H. Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. *Genes Dev.* 2006. V. 20. № 13. P. 1732–1743. doi: <u>10.1101/gad.1425706</u>
- 184. Aravin A., Gaidatzis D., Pfeffer S., Lagos-Quintana M., Landgraf P., Iovino N., Morris P., Brownstein M.J., Kuramochi-Miyagawa S., Nakano T.et al. A novel class of small RNAs bind to MILI protein in mouse testes. *Nature*. 2006. V. 442. P. 203–207.
- 185. Girard A., Sachidanandam R., Hannon G.J., Carmell M.A. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature*. 2006. V. 442. № 7099. P. 199–202. doi: <u>10.1038/nature04917</u>
- 186. Lau N.C., Seto A.G., Kim J., Kuramochi-Miyagawa S., Nakano T., Bartel D.P., Kingston R.E. Characterization of the piRNA Complex from Rat Testes. *Science*. 2006. V. 313. P. 363–367. doi: <u>10.1126/science.1130164</u>
- 187. Grivna S.T., Beyret E., Wang Z., Lin H. A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev.* 2006. V. 20. P. 1709–1714.
- 188. Malone C.D., Hannon G.J. Small RNAs as guardians of the genome. *Cell*. 2009. V. 136. № 4. P. 656–668.
- 189. Ishizu H., Siomi H., Siomi M.C. Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. *Genes Dev.* 2012. V. 26. № 21. P. 2361–2373. doi: 10.1101/gad.203786.112
- 190. Gou L.T., Dai P., Yang J.H., Xue Y., Hu Y.P., Zhou Y., Kang J.Y., Wang X., Li H., Hua M.M. et al. Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell Res.* 2014. V. 24. P. 680–700.
- 191. Zhang P., Kang J.Y., Gou L.T., Wang J., Xue Y., Skogerboe G., Dai P., Huang D.W., Chen R., Fu X.D. et al. MIWI and piRNAmediated cleavage of messenger RNAs in mouse testes. *Cell Res.* 2015. V. 25. P. 193–207.
- 192. Goh W.S.S., Falciatori I., Tam O.H., Burgess R., Meikar O., Kotaja N., Hammell M., Hannon G.J. piRNA-directed cleavage of meiotic transcripts regulates spermatogenesis. *Genes Dev.* 2015. V. 29. P. 1032–1044.
- 193. Watanabe T., Cheng E., Zhong M., Lin H. Retrotransposons and pseudogenes regulate mRNAs and lncRNAs via the piRNA pathway in the germline. *Genome Res.* 2015. V. 25. P. 368–380.
- 194. Vourekas A., Alexiou P., Vrettos N., Maragkakis M., Mourelatos Z. Sequence-dependent but not sequence-specific piRNA adhesion traps mRNAs to the germ plasm. *Nature*. 2016. V. 531. P. 390–394.
- 195. Shen E.Z., Chen H., Ozturk A.R., Tu S., Shirayama M., Tang W., Ding Y.H., Dai S.Y., Weng Z., Mello C.C. Identification of piRNA binding sites reveals the Argonaute regulatory landscape of the *C. elegans* germline. *Cell*. 2018. V. 172. P. 937–951.e18. doi: <u>10.1016/j.cell.2018.02.002</u>
- 196. Zhang D., Tu S., Stubna M., Wu W.S., Huang W.C., Weng Z., Lee H.C. The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. *Science*. 2018. V. 359. P. 587–592.
- 197. Dai P., Wang X., Gou L.T., Li Z.T., Wen Z., Chen Z.G., Hua M.M., Zhong A., Wang L., Su H., Wan H., Qian K., Liao L., Li J., Tian B., Li D., Fu X.D., Shi H.J., Zhou Y., Liu M.F. A Translation-Activating Function of MIWI/piRNA during Mouse Spermiogenesis. *Cell*. 2019. V. 179. № 7. P. 1566–1581.E16. doi: 10.1016/j.cell.2019.11.022

- 198. Dai P., Wang X., Liu MF. A dual role of the PIWI/piRNA machinery in regulating mRNAs during mouse spermiogenesis. *Sci. China Life Sci.* 2020. V. 63. P. 447–449. doi: <u>10.1007/s11427-020-1632-5</u>
- 199. Brennecke J., Aravin A.A., Stark A., Dus M., Kellis M., Sachidanandam R., Hannon G.J. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell*. 2007. V. 128. P. 1089–1103.
- 200. Ozata D.M., Gainetdinov I., Zoch A., O'Carroll D., Zamore P.D. PIWI-interacting RNAs: small RNAs with big functions. *Nat. Rev. Genet.* 2019. V. 20. P. 89–108. doi: <u>10.1038/s41576-018-0073-3</u>
- 201. Seto A.G., Kingston R.E., Lau N.C. The Coming of Age for Piwi Proteins. *Molecular Cell*. 2007. V. 26. № 5. P. P603–609. doi: <u>10.1016/j.molcel.2007.05.021</u>
- 202. Jochl C., Rederstorff M., Hertel J., Stadler P.F., Hofacker I.L., Schrettl M., Haas H., Huttenhofer A. Small ncRNA transcriptome analysis from *Aspergillus fumigatus* suggests a novel mechanism for regulation of protein synthesis. *Nucleic Acids Res.* 2008. V. 36. P. 2677–2689.
- 203. Li Y., Zhang Y., Liu M. Knockout Gene-Based Evidence for PIWI-Interacting RNA Pathway in Mammals. *Front. Cell Dev. Biol.* 2021. V. 9. P. 681188. doi: 10.3389/fcell.2021.681188
- 204. Su J.F., Concilla A., Zhang D.Z., Zhao F., Shen F.F., Zhang H., Zhou F.Y. PIWIinteracting RNAs: Mitochondria-based biogenesis and functions in cancer. *Genes Dis.* 2020. V. 8. № 5. P. 603–622. doi: <u>10.1016/j.gendis.2020.09.006</u>
- 205. Haussecker D., Huang Y., Lau A., Parameswaran P., Fire A.Z., Kay M.A. Human tRNAderived small RNAs in the global regulation of RNA silencing. *RNA*. 2010. V. 16. P. 673– 695. doi: <u>10.1261/rna.2000810</u>
- 206. Ivanov P., Emara M.M., Villen J., Gygi S.P., Anderson P. Angiogenin-induced tRNA fragments inhibit translation initiation. *Mol. Cell.* 2011. V. 43. P. 613–623. doi: 10.1016/j.molcel.2011.06.022
- 207. Gebetsberger J., Zywicki M., Kunzi A., Polacek N. tRNA-Derived Fragments Target the Ribosome and Function as Regulatory Non-coding RNA in *Haloferax Volcanii*. Archaea. 2012. V. 2012. Article ID 260909. doi: 10.1155/2012/260909
- 208. Speer J., Gehrke C.W., Kuo K.C., Waalkes T.P., Borek E. tRNA Breakdown Products as Markers for Cancer. 1979. V. 44. № 6. P. 2120–2123. doi: <u>10.1002/1097-0142(197912)44:6<2120::aid-cncr2820440623>3.0.co;2-6</u>
- 209. Balatti V., Nigita G., Veneziano D., Drusco A., Stein G.S., Messier T.L., Farina N.H., Lian J.B., Tomasello L., Liu C.G. et al. tsRNA signatures in cancer. *PNAS*. 2017. V. 114. P. 8071–8076.
- 210. Guzman N., Agarwal K., Asthagiri D., Yu L., Saji M., Ringel M.D., Paulaitis M.E. Breast cancer-specific miR signature unique to extracellular vesicles includes "microRNA-like" tRNA fragments. *Mol. Cancer Res.* 2015. V.13. P. 891–901.
- 211. Keam S.P., Hutvagner G. tRNA-Derived Fragments (tRFs): Emerging New Roles for an Ancient RNA in the Regulation of Gene Expression. *Life (Basel)*. 2015. V. 5. № 4. P. 1638–1651. doi: 10.3390/life5041638
- 212. Goodarzi H., Liu X., Nguyen H.C., Zhang S., Fish L., Tavazoie S.F. Endogenous tRNAderived fragments suppress breast cancer progression via YBX1 displacement. *Cell*. 2015. V. 161. P. 790–802.
- 213. Anderson P., Ivanov P. tRNA fragments in human health and disease. *FEBS Lett.* 2014. V. 588. P. 4297–4304.
- 214. Kumar P., Anaya J., Mudunuri S.B., Dutta A. Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets. *BMC Biol.* 2014. V. 12. P. 78.
- 215. Karaiskos S., Grigoriev A. Dynamics of tRNA fragments and their targets in aging mammalian brain. *F1000Res*. 2016. V. 5. P. 2758. doi: <u>10.12688/f1000research.10116.1</u>

- 216. Jin F., Guo Z. Emerging role of a novel small non-coding regulatory RNA: tRNA-derived small RNA. *ExRNA*. 2019. V. 1. Article No. 39. doi: <u>10.1186/s41544-019-0036-7</u>
- 217. Guzzi N., Bellodi C. Novel insights into the emerging roles of tRNA-derived fragments in mammalian development. *RNA Biol.* 2020. V. 17. P. 1214–1222. doi: 10.1080/15476286.2020.1732694
- 218. Megel C., Hummel G., Lalande S., Ubrig E., Cognat V., Morelle G., Salinas-Giegé T., Duchêne A.-M., Maréchal-Drouard L. Plant RNases T2, but not Dicer-like proteins, are major players of tRNA-derived fragments biogenesis. *Nucleic Acids Res.* 2019. V. 47. P. 941–952. doi: <u>10.1093/nar/gky1156</u>
- 219. Zheng L.L., Xu W.L., Liu S., Sun W.J., Li J.H., Wu J., Yang J.-H., Qu L.-H. tRF2Cancer: a web server to detect tRNA-derived small RNA fragments (tRFs) and their expression in multiple cancers. *Nucleic Acids Res.* 2016. V. 44. P. W185–W193. doi: 10.1093/nar/gkw414
- 220. Yu M., Lu B., Zhang J., Ding J., Liu P., Lu Y. tRNA-derived RNA fragments in cancer: current status and future perspectives. J. Hematol Oncol. 2020. V. 13. P. 121. doi: <u>10.1186/s13045-020-00955-6</u>
- 221. Thompson D.M., Lu C., Green P.J., Parker R. tRNA cleavage is a conserved response to oxidative stress in eukaryotes. *RNA*. 2008. V. 14. № 10. P. 2095–2103. doi: 10.1261/rna.1232808
- 222. Shen Y., Yu X., Zhu L., Li T., Yan Z., Guo J. Transfer RNA-derived fragments and tRNA halves: biogenesis, biological functions and their roles in diseases. J. Cell Mol. Med. 2018. V. 96. P. 1167–76.
- 223. Cho H, Lee W, Kim GW, Lee SH, Moon JS, Kim M, Kim HS, Oh JW. Regulation of La/SSB-dependent viral gene expression by pre-tRNA 3' trailer-derived tRNA fragments. *Nucleic Acids Res.* 2019. V 47. P. 9888–9901. doi: <u>10.1093/nar/gkz732</u>
- 224. Ivanov P., Emara M.M, Villen J., Gygi S.P, Anderson P. Angiogenin-induced tRNA fragments inhibit translation initiation. *Mol Cell*. 2011. V. 43. P. 613–623. doi: 10.1016/j.molcel.2011.06.022
- 225. Guan L., Karaiskos S., Grigoriev A. Inferring targeting modes of Argonaute-loaded tRNA fragments. *RNA Biol.* 2020. V. 17. P. 1070–1080.
- 226. Kuscu C., Kumar P., Kiran M., Su Z., Malik A., Dutta A. tRNA fragments (tRFs) guide Ago to regulate gene expression post-transcriptionally in a Dicer-independent manner. *RNA*. 2018. V. 24. P. 1093–1105.
- 227. Hasler D., Lehmann G., Murakawa Y., Klironomos F., Jakob L., Grasser F.A., Rajewsky N., Landthaler M., Meister G. The lupus autoantigen La prevents mis-channeling of tRNA fragments into the human microRNA pathway. *Mol. Cell.* 2016. V. 63. P. 110–124.
- 228. Anger A.M., Armache J.P., Berninghausen O., Habeck M., Subklewe M., Wilson D.N., Beckmann R. Structures of the human and Drosophila 80S ribosome. *Nature*. 2013. V. 497. P. 80–85.
- 229. Hansen T.B., Veno M.T., Jensen T.I., Schaefer A., Damgaard C.K., Kjems J. Argonauteassociated short introns are a novel class of gene regulators. *Nat. Commun.* 2016. V. 7. P. 11538.
- 230. Tenaillon M.I., Hollister J.D., Gaut B.S. A triptych of the evolution of plant transposable elements. *Trends Plant Sci.* 2010. V. 15. P. 471–478. doi: 10.1016/j.tplants.2010.05.003
- 231. Wicker T., Sabot F., Hua-Van A., Bennetzen J.L., Capy P., Chalhoub B., Flavell A., Leroy P., Morgante M., Panaud O. et al. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* 2007. V. 8. P. 973–982. doi: <u>10.1038/nrg2165</u>
- 232. Vitte C., Fustier M.A., Alix K., Tenaillon M.I. The bright side of transposons in crop evolution. *Brief. Funct. Genomics.* 2017. V. 13. P. 276–295. doi: <u>10.1093/bfgp/elu002</u>
- 233. Grandbastien M.A. Activation of plant retrotransposons under stress conditions. *Trends Plant Sci.* 1998. V. 3. P. 181–187. doi: <u>10.1016/S1360-1385(98)01232-1</u>

- 234. Matzke M.A., Mosher R.A. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* 2014. V. 15. P. 394–408. doi: <u>10.1038/nrg3683</u>
- 235. Blevins T., Podicheti R., Mishra V., Marasco M., Wang J., Rusch D., Tang H., Pikaard C.S. Identification of Pol IV and RDR2-dependent precursors of 24 nt siRNAs guiding de novo DNA methylation in Arabidopsis. *Elife*. 2015. V. 4. Article No. e09591. doi: <u>10.7554/eLife.09591</u>
- 236. Zhong X., Du J., Hale C.J., Gallego-Bartolome J., Feng S., Vashisht A.A., Chory J., Wohlschlegel J.A., Patel D.J., Jacobsen S.E. et al. Molecular mechanism of action of plant DRM de novo DNA methyltransferases. *Cell*. 2014. V. 157. P. 1050–1060. doi: <u>10.1016/j.cell.2014.03.056</u>
- 237. Creasey K.M., Zhai J., Borges F., Van Ex F., Regulski M., Meyers B.C., Martienssen R.A. miRNAs trigger widespread epigenetically activated siRNAs from transposons in Arabidopsis. *Nature*. 2014. V. 508. P. 411–415. doi: <u>10.1038/nature13069</u>
- 238. Watanabe T., Cheng E.C., Zhong M., Lin H. Retrotransposons and pseudogenes regulate mRNAs and lncRNAs via the piRNA pathway in the germline. *Genome Res.* 2015. V. 25. P. 368–380. doi: 10.1101/gr.180802.114
- 239. Wang D., Qu Z., Yang L., Zhang Q., Liu Z.H., Do T., Adelson D.L., Wang Z.-Y., Searle I., Zhu J.-K. Transposable elements (TEs) contribute to stress-related long intergenic noncoding RNAs in plants. *Plant J.* 2017. V. 90. P. 133–146. doi: <u>10.1111/tpj.13481</u>
- 240. Wang X., Ai G., Zhang C., Cui L., Wang J., Li H., Zhang J., Ye Z. Expression and diversification analysis reveals transposable elements play important roles in the origin of Lycopersicon-specific lncRNAs in tomato. *New Phytol.* 2015. V. 209. P. 1442–1455. doi: 10.1111/nph.13718
- 241. Liu J., Jung C., Xu J., Wang H., Deng S., Bernad L., Arenas-Huertero C., Chua N.-H. Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in Arabidopsis. *Plant Cell*. 2012. V. 24. P. 4333–4345. doi: <u>10.1105/tpc.112.102855</u>
- 242. Cao X., Yeo G., Muotri A.R., Kuwabara T., Gage F.H. Noncoding RNAs in the mammalian central nervous system. *Annu Rev. Neurosci.* 2006. V. 29. P. 77–103. doi: <u>10.1146/annurev.neuro.29.051605.112839</u>
- 243. Mercer T.R., Dinger M.E., Mattick J.S. Long non-coding RNAs: Insights into functions. *Nat. Rev. Genet.* 2009. V. 10. P. 155–159.
- 244. Ponting C.P., Oliver P.L., Reik W. Evolution and functions of long noncoding RNAs. *Cell.* 2009. V. 136. P. 629–641. doi: <u>10.1016/j.cell.2009.02.006</u>
- 245. Managadze D., Rogozin I.B., Chernikova D., Shabalina S.A., Koonin E.V. Negative correlation between expression level and evolutionary rate of long intergenic noncoding RNAs. *Genome Biol. Evol.* 2011. V. 3. P. 1390–1404. doi: <u>10.1093/gbe/evr116</u>
- 246. Ponting C.P., Belgard T.G. Transcribed dark matter: meaning or myth? *Hum. Mol. Genet.* 2010. V. 19. P. R162–R168. doi: <u>10.1093/hmg/ddq362</u>
- 247. Ponjavic J., Ponting C.P., Lunter G. Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. *Genome Res.* 2007. V. 17. P. 556–565. doi: <u>10.1101/gr.6036807</u>
- 248. Pang K.C., Frith M.C., Mattick J.S. Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends Genet*. 2006. V. 22. P. 1–5. doi: <u>10.1016/j.tig.2005.10.003</u>
- 249. Rinn J.L., Chang H.Y. Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* 2012. V. 81. P. 145–166. doi: <u>10.1146/annurev-biochem-051410-092902</u>
- 250. Wang K.C., Chang H.Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell.* 2011. V. 43. P. 904–914. doi: <u>10.1016/j.molcel.2011.08.018</u>
- 251. Guttman M., Amit I., Garber M., French C., Lin M.F., Feldser D., Huarte M., Zuk O., Carey B.W., Cassady J.P. et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*. 2009. V. 458. P. 223–227.

- 252. Guenther M.G., Levine S.S., Boyer L.A., Jaenisch R., Young R.A. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell*. 2007. V. 130. P. 77–88. doi: <u>10.1016/j.cell.2007.05.042</u>
- 253. Spitale R.C., Tsai M.C., Chang H.Y. RNA templating the epigenome: Long noncoding RNAs as molecular scaffolds. *Epigenetics*. 2011. V. 6. P. 539–543. doi: <u>10.4161/epi.6.5.15221</u>
- 254. Good M.C., Zalatan J.G., Lim W.A. Scaffold proteins: hubs for controlling the flow of cellular information. *Science*. 2011. V. 332. P. 680–686. doi: <u>10.1126/science.1198701</u>
- 255. Duret L., Chureau C., Samain S., Weissenbach J., Avner P. The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. *Science*. 2006. V. 312. P. 1653–1655. doi: <u>10.1126/science.1126316</u>
- 256. Elisaphenko E.A., Kolesnikov N.N., Shevchenko A.I., Rogozin I.B., Nesterova T.B., Brockdorff N., Zakian S.M. A dual origin of the Xist gene from a protein-coding gene and a set of transposable elements. *PLoS ONE*. 2008. V. 3. Article No. e2521. doi: <u>10.1371/journal.pone.0002521</u>
- 257. Hadjiargyrou M., Delihas N. The intertwining of transposable elements and non-coding RNAs. *Int. J. Mol. Sci.* 2013. V. 14. P. 13307–13328. doi: <u>10.3390/ijms140713307</u>
- 258. Piriyapongsa J., Jordan I.K. A family of human microRNA genes from miniature inverted-repeat transposable elements. *PLoS ONE*. 2007. V. 2. Article No. e203. doi: 10.1371/journal.pone.0000203
- 259. Yuan Z., Sun X., Jiang D., Ding Y., Lu Z., Gong L., Liu H., Xie J. Origin and evolution of a placental-specific microRNA family in the human genome. *BMC Evol. Biol.* 2010. V. 10. P. 346. doi: <u>10.1186/1471-2148-10-346</u>
- 260. Ahn K., Gim J.-A., Ha H.-S., Han K., Kim H.-S. The novel MER transposon-derived miRNAs in human genome. *Gene.* 2013. V. 512. P. 422–428. doi: <u>10.1016/j.gene.2012.08.028</u>
- 261. Kannan S., Chernikova D., Rogozin I.B., Poliakov E., Managadze D., Koonin E.V., Milanesi L. Transposable element insertions in long intergenic non-coding RNA genes. *Front. Bioeng. Biotechnol.* 2015. V. 3. P. 71. doi: <u>10.3389/fbioe.2015.00071</u>
- 262. Kelley D., Rinn J. Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome Biol.* 2012. V. 13. № 11. P. R107. doi: <u>10.1186/gb-2012-13-11-r107</u>
- 263. Cabili M.N., Trapnell C., Goff L., Koziol M., Tazon-Vega B., Regev A., Rinn J.L. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 2011. V. 25. P. 1915–1927. doi: <u>10.1101/gad.17446611</u>
- 264. Ali A., Han K., Liang P. Role of Transposable Elements in Gene Regulation in the Human Genome. *Life (Basel)*. 2021. V. 11. № 2. P. 118. doi: <u>10.3390/life11020118</u>
- 265. Managadze D., Lobkovsky A.E., Wolf Y.I., Shabalina S.A., Rogozin I.B., Koonin E.V. The Vast, Conserved Mammalian lincRNome. *PLoS Comput Biol.* 2013. V. 9. № 2. Article No. e1002917. doi: 10.1371/journal.pcbi.1002917
- 266. Song Z., Lin J., Li Z., Huang C. The nuclear functions of long noncoding RNAs come into focus. Noncoding RNA Res. 2021. V. 6. № 2. P. 70–79. doi: 10.1016/j.ncrna.2021.03.002
- 267. Braunschweig U., Gueroussov S., Plocik A.M., Graveley B.R., Blencowe B.J. Dynamic integration of splicing within gene regulatory pathways. *Cell*. 2013. V. 152. № 6. P. 1252–1269. doi: 10.1016/j.cell.2013.02.034
- 268. Wang E.T., Sandberg R., Luo S., Khrebtukova I., Zhang L., Mayr C., Kingsmore S.F., Schroth G.P., Burge C.B. Alternative isoform regulation in human tissue transcriptomes. *Nature*. 2008. V. 456. № 7221. P. 470–476. doi: 10.1038/nature07509

- 269. Enuka Y., Lauriola M., Feldman M.E., Sas-Chen A., Ulitsky I., Yarden Y. Circular RNAs are long-lived and display only minimal early alterations in response to a growth factor. *Nucleic Acids Res.* 2016. V. 44. № 3. P. 1370–1383. doi: <u>10.1093/nar/gkv1367</u>
- 270. Hsu M.T., Coca-Prados M. Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. *Nature*. 1979. V. 280. № 5720. P. 339–340. doi: <u>10.1038/280339a0</u>
- 271. Capel B., Swain A., Nicolis S., Hacker A., Walter M., Koopman P., Goodfellow P., Lovell-Badge R. Circular transcripts of the testis-determining gene Sry in adult mouse testis. *Cell*. 1993. V. 73. № 5. P. 1019–1030. doi: <u>10.1016/0092-8674(93)90279-y</u>
- 272. Cocquerelle C., Daubersies P., Majerus M.A., Kerckaert J.P., Bailleul B. Splicing with inverted order of exons occurs proximal to large introns. *EMBO J.* 1992. V. 11. № 3. P. 1095–1098.
- 273. Nigro J.M., Cho K.R., Fearon E.R., Kern S.E., Ruppert J.M., Oliner J.D., Kinzler K.W., Vogelstein B. Scrambled exons. *Cell*. 1991. V. 64. № 3. P. 607–613. doi: <u>10.1016/0092-</u> <u>8674(91)90244-s</u>
- 274. Zaphiropoulos P.G. Circular RNAs from transcripts of the rat cytochrome P450 2C24 gene: correlation with exon skipping. *PNAS*. 1996. V. 93. № 13. P. 6536–6541. doi: <u>10.1073/pnas.93.13.6536</u>
- 275. Zaphiropoulos P.G. Exon skipping and circular RNA formation in transcripts of the human cytochrome P-450 2C18 gene in epidermis and of the rat androgen binding protein gene in testis. *Mol. Cell Biol.* 1997. V. 17. № 6. P. 2985–2993. doi: 10.1128/mcb.17.6.2985
- 276. Rybak-Wolf A., Stottmeister C., Glazar P., Jens M., Pino N., Giusti S., Hanan M., Behm M., Bartok O., Ashwal-Fluss R. et al. Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Mol. Cell.* 2015. V. 58. № 5. P. 870–885. doi: 10.1016/j.molcel.2015.03.027
- 277. Salzman J., Chen R.E., Olsen M.N., Wang P.L., Brown P.O. Cell-type specific features of circular RNA expression. *PLoS Genet.* 2013. V. 9. P. e1003777. doi: <u>10.1371/journal.pgen.1003777</u>
- 278. Jeck W.R., Sorrentino J.A., Wang K., Slevin M.K., Burd C.E., Liu J., Marzluff W.F., Sharpless N.E. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA*. 2013. V. 19. № 2. P. 141–157. doi: <u>10.1261/rna.035667.112</u>
- 279. Salzman J., Gawad C., Wang P.L., Lacayo N., Brown P.O. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One*. 2012. V. 7. № 2. Article No. e30733. doi: <u>10.1371/journal.pone.0030733</u>
- 280. Guo J.U., Agarwal V., Guo H., Bartel D.P. Expanded identification and characterization of mammalian circular RNAs. *Genome Biol.* 2014. V. 15. № 7. Article No. 409. doi: <u>10.1186/s13059-014-0409-z</u>
- 281. Salzman J., Gawad C., Wang P.L., Lacayo N., Brown P.O. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS ONE*. 2012. V. 7. Article No. e30733. doi: 10.1371/journal.pone.0030733
- 282. Guo J.U., Agarwal V., Guo H., Bartel D.P. Expanded identification and characterization of mammalian circular RNAs. *Genome Biol.* 2014. V. 15. Article No. 409. doi: <u>10.1186/s13059-014-0409-z</u>
- 283. Du W.W., Yang W., Liu E., Yang Z., Dhaliwal P., Yang B.B. Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2. *Nucleic Acids Res.* 2016. V. 44. P. 2846–2858. doi: <u>10.1093/nar/gkw027</u>
- 284. Abdelmohsen K., Panda A.C., Munk R., Grammatikakis I., Dudekula D.B., De S., Kim J., Noh J.H., Kim K.M., Martindale J.L., Gorospe M. Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1. *RNA Biol.* 2017. V. 14. № 3. P. 361–369. doi: 10.1080/15476286.2017.1279788

- 285. Zhao R.T., Zhou J., Dong X.L., Bi C.W., Jiang R.C., Dong J.F., Tian Y., Yuan H.J., Zhang J.N. Circular Ribonucleic Acid Expression Alteration in Exosomes from the Brain Extracellular Space after Traumatic Brain Injury in Mice. *J. Neurotrauma*. 2018. V. 35. P. 2056–2066. doi: 10.1089/neu.2017.5502
- 286. Preusser C., Hung L.H., Schneider T., Schreiner S., Hardt M., Moebus A., Santoso S., Bindereif A. Selective release of circRNAs in platelet-derived extracellular vesicles. J. Extracell Vesicles. 2018. V. 7. № 1. P. 1424473. doi: 10.1080/20013078.2018.1424473
- 287. Legnini I., Di Timoteo G., Rossi F., Morlando M., Briganti F., Sthandier O., Fatica A., Santini T., Andronache A., Wade M., et al. Circ-ZNF609 is a circular RNA that can be translated and functions in Myogenesis. *Mol. Cell.* 2017. V. 66. № 1. P. 22–37.e9. doi: 10.1016/j.molcel.2017.02.017
- 288. Wang Y., Wang Z. Efficient backsplicing produces translatable circular mRNAs. *RNA*. 2015. V. 21. № 2. P. 172–179. doi: <u>10.1261/rna.048272.114</u>
- 289. Pamudurti N.R., Bartok O., Jens M., Ashwal-Fluss R., Stottmeister C., Ruhe L., Hanan M., Wyler E., Perez-Hernandez D., Ramberger E. et al. Translation of CircRNAs. *Mol. Cell.* 2017. V. 66. № 1. P. 9–21.e7. doi: 10.1016/j.molcel.2017.02.021
- 290. Zhang M., Zhao K., Xu X., Yang Y., Yan S., Wei P., Liu H., Xu J., Xiao F., Zhou H. et al. A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma. *Nat. Commun.* 2018. V. 9. № 1. P. 4475. doi: 10.1038/s41467-018-06862-2
- 291. Zhang X.O., Wang H.B., Zhang Y., Lu X., Chen L.L., Yang L. Complementary sequencemediated exon circularization. *Cell*. 2014. V. 159. № 1. P. 134–147. doi: <u>10.1016/j.cell.2014.09.001</u>
- 292. Ashwal-Fluss R., Meyer M., Pamudurti N.R., Ivanov A., Bartok O., Hanan M., Evantal N., Memczak S., Rajewsky N., Kadener S. circRNA biogenesis competes with pre-mRNA splicing. *Mol. Cell.* 2014. V. 56. № 1. P. 55–66. doi: 10.1016/j.molcel.2014.08.019
- 293. Westholm J.O., Miura P., Olson S., Shenker S., Joseph B., Sanfilippo P., Celniker S.E., Graveley B.R., Lai E.C. Genome-wide analysis of drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Rep.* 2014. V. 9. № 5. P. 1966–1980. doi: 10.1016/j.celrep.2014.10.062
- 294. Barrett S.P., Wang P.L., Salzman J. Circular RNA biogenesis can proceed through an exon-containing lariat precursor. *Elife*. 2015. V. 4. Article No. e07540. doi: <u>10.7554/eLife.07540</u>
- 295. Kristensen L.S., Okholm T.L.H., Veno M.T., Kjems J. Circular RNAs are abundantly expressed and upregulated during human epidermal stem cell differentiation. *RNA Biol.* 2018. V. 15. № 2. P. 280–291. doi: 10.1080/15476286.2017.1409931
- 296. Kramer M.C., Liang D., Tatomer D.C., Gold B., March Z.M., Cherry S., Wilusz J.E. Combinatorial control of Drosophila circular RNA expression by intronic repeats, hnRNPs, and SR proteins. *Genes Dev.* 2015. V. 29. № 20. P. 2168–2182. doi: 10.1101/gad.270421.115
- 297. Zhang Y., Xue W., Li X., Zhang J., Chen S., Zhang J.L., Yang L., Chen L.L. The biogenesis of nascent circular RNAs. *Cell Rep.* 2016. V. 15. № 3. P. 611–624. doi: 10.1016/j.celrep.2016.03.058
- 298. Wilusz J.E. A 360 degrees view of circular RNAs: from biogenesis to functions. *Wiley Interdiscip. Rev. RNA*. 2018. V. 9. № 4. Article No. e1478. doi: <u>10.1002/wrna.1478</u>
- 299. Xiao M.S., Ai Y., Wilusz J.E. Biogenesis and functions of circular RNAs come into focus. *Trends Cell Biol.* 2020. V. 30. № 3. P. 226–240. doi: <u>10.1016/j.tcb.2019.12.004</u>
- 300. Yu C.Y., Kuo H.C. The emerging roles and functions of circular RNAs and their generation. J. Biomed. Sci. 2019. V. 26. P. 29. doi: <u>10.1186/s12929-019-0523-z</u>
- 301. Glazar P., Papavasileiou P., Rajewsky N. circBase: a database for circular RNAs. *RNA*. 2014. V. 20. № 11. P. 1666–16670. doi: <u>10.1261/rna.043687.113</u>

- 302. Liu Y.C., Li J.R., Sun C.H., Andrews E., Chao R.F., Lin F.M., Weng S.L., Hsu S.D., Huang C.C., Cheng C. et al. CircNet: a database of circular RNAs derived from transcriptome sequencing data. *Nucleic Acids Res.* 2016. V. 44. № D1. P. D209–D215. doi: 10.1093/nar/gkv940
- 303. Ghosal S., Das S., Sen R., Basak P., Chakrabarti J. Circ2Traits: a comprehensive database for circular RNA potentially associated with disease and traits. *Front. Genet.* 2013. V. 4. Article No. 283. doi: <u>10.3389/fgene.2013.00283</u>
- 304. Yao D., Zhang L., Zheng M., Sun X., Lu Y., Liu P. Circ2Disease: a manually curated database of experimentally validated circRNAs in human disease. *Sci. Rep.* 2018. V. 8. № 1. P. 11018. doi: 10.1038/s41598-018-29360-3
- 305. Li S., Li Y., Chen B., Zhao J., Yu S., Tang Y., Zheng Q., Li Y., Wang P., He X., Huang S. exoRBase: a database of circRNA, lncRNA and mRNA in human blood exosomes. *Nucleic Acids Res.* 2018. V. 46. P. D106–D112. doi: <u>10.1093/nar/gkx891</u>
- 306. Xia S., Feng J., Chen K., Ma Y., Gong J., Cai F., Jin Y., Gao Y., Xia L., Chang H. et al. CSCD: a database for cancer-specific circular RNAs. *Nucleic Acids Res.* 2018. V. 46. № D1. P. D925–D929. doi: 10.1093/nar/gkx86
- 307. Laudadio I., Carissimi C., Fulci V. How RNAi machinery enters the world of telomerase. *Cell Cycle*. 2019. V. 18. P. 1056–1067. doi: <u>10.1080/15384101.2019.1609834</u>
- 308. Altman S. Ribonuclease P. *Philos. Trans. R. Soc. Lond. B. Biol Sci.* 2011. V. 366. P. 2936–2941. doi: <u>10.1098/rstb.2011.0142</u>
- 309. Arnold P.R., Wells A.D., Li X.C. Diversity and Emerging Roles of Enhancer RNA in Regulation of Gene Expression and Cell Fate. *Frontiers in Cell and Developmental Biology*. 2020. V. 7. Article No. 377. doi: <u>10.3389/fcell.2019.00377</u>
- 310. Bhogireddy S., Mangrauthia S.K., Kumar R., Pandey A.K., Singh S., Jain A., Budak H., Varshney R.K., Kudapa H. Regulatory non-coding RNAs: a new frontier in regulation of plant biology. *Funct. Integr. Genomics.* 2021. V. 21. P. 313–330. doi: <u>10.1007/s10142-021-00787-8</u>
- 311. Song L., Fang Y., Chen L., Wang J., Chen X. Role of non-coding RNAs in plant immunity. *Plant Commun.* 2021. V. 2. Article No. 100180. doi: <u>10.1016/j.xplc.2021.100180</u>
- 312. Wu N., Yang B.B. The Biological Functions of Non-coding RNAs: From a Line to a Circle. *Discoveries (Craiova)*. 2015. V. 3. Article No. e48. doi:10.15190/d.2015.40
- 313. Cerutti H., Casas-Mollano J.A. On the origin and functions of RNA-mediated silencing: from protists to man. *Curr. Genet.* 2006. V. 50. P. 81–99. doi: <u>10.1007/s00294-006-0078-</u>
 <u>x</u>
- 314. Grigoriev A. Transfer RNA and Origins of RNA Interference. *Front Mol Biosci*. 2021.
 V. 8. Article No. 708984. doi:<u>10.3389/fmolb.2021.708984</u>
- 315. Mohammed J., Flynt A.S., Panzarino A.M., Mondal M.M.H., DeCruz M., Siepel A., Lai E.C. Deep experimental profiling of microRNA diversity, deployment, and evolution across the *Drosophila* genus. *Genome Res.* 2018. V. 28. P. 52–65. doi: 10.1101/gr.226068.117
- 316. Gilbert W. Origin of life: The RNA world. Nature. 1986. V. 319. P. 618.
- 317. Crick F.H. The origin of the genetic code. J. Mol. Biol. 1968. V. 38. P. 367–379.
- 318. Orgel L.E Evolution of the genetic apparatus. J. Mol. Biol. 1968. V. 38. P. 381-393.
- 319. Woese C.R. *The genetic code: The molecular basis for genetic expression*. Harper & Row, 1967. p. 186.
- 320. Eddy S.R. Computational genomics of noncoding RNA genes. *Cell*. 2002. V. 109. P. 137–140.
- 321. *RNA Worlds: From Life's Origins to Diversity in Gene Regulation*. Eds.: Atkins J.F., Gesteland R.F., Cech T.R. Cold Spring Harbor Laboratory Press, 2011. 366 p.
- 322. Кунин Е.В. Логика случая. О природе и происхождении биологической эволюции. М.: ЗАО Издательство Центрполиграф, 2017. 527 стр.

- 323. Bokov K., Steinberg S.V. A hierarchical model for evolution of 23S ribosomal RNA. *Nature*. 2009. V. 457. P. 977–980. doi: <u>10.1038/nature07749</u>
- 324. Skoblikow N.E., Zimin A.A. A search for relict ribonucleotide and amino acid sequences that played a key role in the development of the ribosome and modern protein diversity. *Math. Biol. Bioinf.* 2015. V. 10. P. 116–130. doi: <u>10.17537/2015.10.116</u>
- 325. Nazipova N.N., Isaev E.A., Kornilov V.V., Pervukhin D.V., Morozova A.A., Gorbunov A.A., Ustinin M.N. Big Data in Bioinformatics. *Math. Biol. Bioinf.* 2018. V. 13. P. t1–t16. doi: 10.17537/2018.13.t1

Рукопись поступила в редакцию 03.06.2021, переработанный вариант поступил 13.08.2021. Дата опубликования 26.08.2021.

Variety of Non-Coding RNAs in Eukaryotic Genomes

Nafisa Nazipova

Institute of Mathematical Problems of Biology RAS, Keldysh Institute of Applied Mathematics of Russian Academy of Sciences, Pushchino, Russia

Abstract. The genomes of large multicellular eukaryotes mainly consist of DNA that encodes not proteins, but RNAs. The unexpected discovery of approximately the same number of protein genes in Homo sapiens and *Caenorhabditis elegans* led to the understanding that it is not the number of proteins that determines the complexity of the development and functioning of an organism. The phenomenon of pervasive transcription of genomes is finding more and more confirmation. Data are emerging on new types of RNA that work in different cell compartments, are expressed at different stages of development, in different tissues and perform various functions. Their main purpose is fine regulation of the main cellular processes. The presence of a rich arsenal of regulators that can interact with each other and work on the principle of interchangeability determines the physiological complexity of the organism and its ability to adapt to changing environmental conditions. An overview of the currently known functional RNAs expressed in eukaryotic genomes is presented here. There is no doubt that in the near future, using high-tech transcriptome technologies, many new RNAs will be identified and characterized. But it is likely that many of the expressed transcripts do not have a function, but are an evolutionary reserve of organisms.

Key words: RNA-interference, pervasive transcription, miRNA, siRNA, snRNA, snoRNA, piRNA, long noncoding RNA, circular RNA, tRNA derived fragments, rRNA-derived fragments, transposon-derived noncoding RNA.