

РОЛЬ ТРАНСПОЗОНОВ В ЭПИГЕНЕТИЧЕСКОЙ РЕГУЛЯЦИИ ОНТОГЕНЕЗА

© 2018 г. Р.Н. Мустафин^{1,*}, Э.К. Хуснутдинова^{1,2}

¹Башкирский Государственный Университет
450076, г. Уфа, ул. Заки Валиди, 32

²Институт биохимии и генетики Уфимского научного центра РАН
450054, Уфа, проспект Октября, 71

*E-mail: ruji79@mail.ru

Поступила в редакцию 22.12.2016 г.

Окончательный вариант получен 01.06.2017 г.

В обзорной статье представлен новый взгляд на механизмы реализации геномной информации в индивидуальном развитии эукариот путем динамично меняющихся при каждом клеточном делении взаимодействий транспозонов с эпигенетическими факторами. Данные механизмы поэтапной реализации генетической информации со стадии- и тканеспецифическими особенностями изменений активности определенных семейств мобильных генетических элементов, закреплены эволюционно на уровне вида. При этом индивидуальные различия в результате «незапланированных» перемещений могут привести к существенным изменениям в регуляторной сети генома, приводя к изменениям фенотипа. Данные изменения онтогенетического развития могут вызвать появление новых признаков, приводя либо к появлению болезни, либо к лучшей адаптации, являясь важной составляющей изменчивости для естественного отбора в эволюции. Большая часть транспозонов эукариот изменена мутациями и используется для формирования регуляторной геномной сети, изменения протектирующих и образования новых не кодирующих белки генов. Мобильные генетические элементы, при инсерции в новые локусы, являются основой для образования микроРНК, а также формируют доменные структуры длинных некодирующих РНК, являясь чувствительными к различным видам стресса, что отражается на особенностях индивидуального развития и способствует изменчивости. Эпигенетические факторы, включая некодирующие РНК, метилирование ДНК и модификации гистонов, тесно связаны с мобильными генетическими элементами. Особенности расположения транспозонов отдельных особей, возникшие в результате мутации, либо спонтанной, либо под действием стрессовых факторов, могут вызвать существенные изменения во взаимосвязях геномных сетей, что влияет на вероятность выживания в меняющихся условиях окружающей среды, отражая четкую взаимосвязь механизмов индивидуального развития и эволюции. Имеется параллелизм между механизмами эволюционных преобразований геномов под действием транспозонов с преобразованием генома под их воздействием в онтогенезе. В частности, чувствительность транспозонов к воздействию внешних и внутренних (микроокружение) факторов, дают основу для эволюционного конструирования транспозон-опосредованного тканеспецифического паттерна активации определенных транспозонов при каждом клеточном делении, ведущих к созреванию половозрелого организма. Основой для данного механизма является тесная стадий- и тканеспецифическая взаимосвязь между транспозонами, эпигенетическими факторами и белок-кодирующими генами.

Ключевые слова: длинные некодирующие РНК, микроРНК, метилирование, мобильные генетические элементы, модификация гистонов, нуклеотидные последовательности, окружающая среда, повторяющиеся элементы, транспозоны, repeat-derived miR (RdmiR)

DOI: 10.7868/S0475145018020015

ПЕРЕМЕЩЕНИЯ ТРАНСПОЗОНОВ В ОНТОГЕНЕЗЕ

Повторяющиеся элементы широко распространены в геномах растений и животных. Например, транспозоны (ТЕ – transposable elements) составляют в геномах человека 45%, мыши – 40%, плодовых мушек – 15–22%, нематод – 12%, кур – 8,6%

(Yuan et al., 2011). Хотя истинный вклад ТЕ значительно выше (Burns, Boeke, 2012) – значительная часть геномов эукариот также имеет транспозонное происхождение, но мутирована до такой степени, что не идентифицируется современными методами (Edlefsen, Liu, 2010). Новый вычислительный подход, основанный на «олиго-облаке» – использовании большого количества олигонуклеотидов,

узнающих многие малые фрагменты ретроТЕ, накопленные за сотни миллионов лет эволюции позвоночных, дает основание полагать, что повторы составляют не менее двух третей (69%) всего генома человека (De Koning et al., 2011). В геномах некоторых растений ТЕ составляют до 90% (Yuan et al., 2011).

ТЕ привлекают все большее внимание исследователей в связи получением доказательств их важной роли в формировании структуры и функционировании генома в онтогенезе. При инсерции и интеграции вверх по течению гена, ТЕ могут изменять паттерн экспрессии гена; в экзоне они могут сформировать новый белковый домен; при инсерции в интрон могут привести к продукции белка *de novo*. Также получены доказательства того, что ТЕ могут способствовать рождению новых микроРНК. Впервые происхождение микроРНК из повторов в смысловых и антисмысловых направлениях было задокументировано на растении *Arabidopsis thaliana*. На животных микроРНК (miR), произошедшие из ТЕ (RdmiR – repeat-derived miRNA) впервые были открыты в геномах человека, мыши и крысы. В настоящее время выявлено значительное количество RdmiR в геномах животных и растений (Yuan et al., 2011). ТЕ используются для нужд генома, выполняя ряд полезных функций, например, они незаменимы при выполнении V(D)J-рекомбинации в клетках иммунной системы млекопитающих, в поддержании теломера у дрозофилы и в процессе репарации двунитевых разрывов ДНК у дрожжей, в реакции на стрессовые воздействия. Клетка хозяина контролирует мобильность ТЕ механизмами косупрессии, обычно опосредуемой метилированием, малыми интерферирующими РНК (siРНК) (Юрченко и др., 2011), микроРНК и piwiРНК (Castellano et al., 2015). В то же время, сами ТЕ являются важными источниками для формирования всех малых некодирующих РНК (нкРНК), за счет чего формируется система самоконтроля во взаимосвязи с другими эпигенетическими факторами. Имеются и другие пути самоконтроля ТЕ во взаимосвязи с регуляцией функционирования генома. Например, в исследовании Дуан (Duan, 2017) с соавторами показано, что пара белков транспозонного происхождения (HDP1 и HDP2) функционирует в качестве ацетилтрансферазного комплекса гистонов для активного деметилирования ДНК (данные белки функционируют в качестве факторов антисайленсинга у *Arabidopsis*). Филогенетический анализ показал, что HDP1 и HDP2 были совместно одомашнены из транспозазы и ДНК-связывающего белка, кодируемых транспозоном Harbinger. HDP2 и метил-ДНК-связывающий белок MBD7 имеют большой набор общих сайтов связывания в геноме, что указывает на их совместное определение

целевой специфичности комплекса гистон-ацетилтрансферазы (Duan et al., 2017). Согласно исследованиям Мак-Клинтон, наряду с ответом «теплого шока» у эукариот и SOS-ответом у бактерий (жестко запрограммированных последовательностей клеточных событий), геному свойственны менее запрограммированные реакции на непредвиденные стрессовые воздействия, в число которых входит перемещение транспозонов (McClintock, 1984). Одной из «незапрограммированных» реакций генома, позволяющих ему модифицироваться в ответ на неожиданную смену условий, являются события, происходящие с геномом при преобразовании клеток из нормального для них микроокружения в клеточную культуру (Чересиз и др., 2008). Указанные данные говорят о высокой чувствительности ТЕ к изменениям микроокружения, что позволяет предположить роль ТЕ и их последовательных, эволюционно закрепленных на уровне вида, тканеспецифических изменений активности в регуляции онтогенетической дифференцировки организма. В исследованиях Нали (Nali, 2017) с соавторами показана относительно высокая частота экспрессии эндогенных ретровирусов HERV-K и HERV-W как у ВИЧ-инфицированных женщин и их детей в возрасте до 1 года, так и у неинфицированной контрольной группы. При этом экспрессия HERV-K была сходной в обеих группах, тогда как уровень экспрессии HERV-W у ВИЧ-инфицированных женщин превышал таковой у неинфицированных в 4 раза (Nali et al., 2017). Полученные данные указывают на сохранение активности эндогенных ретровирусов у детей и взрослых, а также специфическое влияние экзогенной ретровирусной инфекции на активность определенных HERV. Таким образом, показана чувствительность мобильных генетических элементов (МГЭ) к инфицированию вирусами. Учитывая сохранение активности LTR в постнатальном периоде как у детей, так и у взрослых, можно предположить роль данных МГЭ в регуляции онтогенетического развития.

Сравнительные цитогенетические исследования показали, что существенную роль в видообразовании у растений играет полиплоидизация. Но у животных полиплоидизация приводит к нарушению хромосомного механизма определения пола, поэтому данный процесс не имел столь важного значения в их эволюции; почти все немногочисленные виды размножаются партеногенетически (Lukash, 2007). В то же время, ТЕ в составе геномов растений составляют значительно большую долю (до 90% и более), чем у животных (Yuan et al., 2011). То есть для различных крупных эволюционных ветвей формировались специфические особенности «использования» ТЕ геномами «хозяев». У растений, размножение которых возможно вегетативно, ТЕ занимали важнейшее значение, вызывая

существенное увеличение размеров геномов при полиплоидизации за счет ослабления контроля транспозиций. У животных, основная часть которых размножается исключительно половым путем, за счет чего формируется значительный резерв для изменчивости при комбинации генетической информации обоих родителей, сформировался более выраженный механизм защиты от инсерций ТЕ.

Хотя значительная часть копий ТЕ остается в эукариотическом геноме неактивной, многие транспозоны транскрибируются и, очевидно, участвуют в регуляции функционирования генома. Например, 75–85% генома человека транскрибируется в первичные и процессирующие транскрипты, однако, только 1,2% генома кодирует белки (Hadjilargyrou, Delihis, 2013). Ретроэлементы часто функционируют в качестве регуляторов транскрипции РНК хозяина. Транскрипты повторов могут служить в качестве мультифункциональных РНК, участвовать в антисмысловой регуляции генов и конкурировать с транскриптами хозяина за связывание с клеточными белками (Баскаев, Буздин, 2011). Спонтанные мутации у эукариот развиваются с низкой частотой (10^{-5} – 10^{-8}) благодаря эффективной работе ферментов с корректорскими и репаративными функциями, что является наследственно закрепленным признаком. Однако гены и функциональные сайты ретротранспозонов подвержены особо сильному мутагенезу, так как их собственный генетический материал при репликации проходит РНК-стадию, имеющую вероятность мутирования 10^{-3} – 10^{-4} , что на несколько порядков выше, чем вероятность мутирования генов (Lukash, 2007). При исследовании мутационного процесса в эволюции дрозофилы, Васильевой с соавторами выявлено, что большинство олигогенных мутаций у *Drosophila melanogaster* – результат инсерций ТЕ. При этом инсерции мобильных генетических элементов (МГЭ) могут изменять активность мажорных и минорных генов, так как содержат в своем составе модули системы управления при связывании с регуляторными белками. При кроссинговере между ТЕ могут возникать хромосомные мутации – делеции, дупликации, инверсии. Кроме того, ТЕ могут дотраивать теломерные концы хромосом, участвовать в горизонтальном переносе генов и откликаться всплесками транспозиций под действием стресса (Васильева и др., 2008). При этом активация ТЕ под действием стресса может происходить независимо от РНК-направленного ДНК-метилирования (RdDM) и независимо от транскрипционных уровней стресс-чувствительных ТЕ (Masuta et al., 2017). В последние годы регистрируется все больше случаев горизонтального переноса генов у многоклеточных эукариот, в том числе у человека. Например, в геноме быков широко распространены

нуклеотидные последовательности (НП) BovB, имевших большее сходство с НП питона и рептилий, чем с НП лошади. Сравнительный полногеномный и филогенетический анализ на 26 видах животных (10 приматов, 12 мух и 4 нематоды) показал, что при горизонтальном переносе генов у животных обычно образуются десятки и сотни активных «чужеродных» генов, в основном связанных с метаболизмом. При этом наиболее активно приобретают чужеродные гены мухи и нематоды в процессе эволюции, тогда как приматы приобрели данным способом относительно мало генов (Huang et al., 2017). В экспериментальных работах на *Drosophila melanogaster*, Васильевой с соавторами показано, что стресс способен увеличивать скорость транспозиций МГЭ на 1–2 порядка величин по сравнению со спонтанным уровнем индукции. Если спонтанная скорость индукции транспозиций составляла $3,8 \times 10^{-4}$ – $1,8 \times 10^{-3}$, то шоковый температурный стресс увеличивал индукцию транспозиций до $1,1 \times 10^{-1}$ на сайт, на геном, за поколение. Пары этанола увеличивали скорость транспозиций до $6,9 \times 10^{-2}$, гамма-облучение – до $1,9 \times 10^{-2}$, изогенизация – до $3,1 \times 10^{-1}$ на сайт, на геном, за поколение. Увеличение скоростей индукции транспозиций МГЭ при стрессовых воздействиях способствует созданию в линиях дополнительного генетического разнообразия, что имеет большое значение для выживания популяций, попадающих в новые экологические условия, а также для селекции и эволюции (Васильева и др., 2007). В ответ на стресс ТЕ служат резервной системой в онтогенетическом развитии. Показано, например, что стрессовое воздействие высоких концентраций железа в почве вызывает мобилизацию многих LTR-ретротранспозонов у риса, помимо активации генов, участвующих в гомеостазе железа. В работе Finatto с соавторами выявлены *cis*-регуляторные элементы, расположенные в промоторных зонах регулируемых генов. При этом *cis*-регуляторные элементы оказались связанными с регуляцией генов и LTR-ретротранспозонов при избытке железа, что подтвердило идею об участии LTR-ретротранспозонов в транскрипционном ответе на стресс как адаптивное преимущество у растений (Finatto et al., 2015). Исследования последних десятилетий определили позитивную роль ТЕ в генерировании геномной новизны и разнообразия растений. В частности, недавние работы показали, что ТЕ-индуцированные эпигенетические изменения и модификация генной экспрессии вносят вклад в фенотипическую изменчивость с адаптацией к стрессу и геосфере. На примере риса показано, что ТЕ играют важную роль в модуляции экспрессии генов и создании новых генов с возможностью фенотипического разнообразия и адаптацией к условиям окружающей среды (Song, Cao, 2017).

L1-элементы регулируют экспрессию генов *in cis* в соответствии с транскрипционной активностью промотора L1, 5'UTR которого контролируется метилированием и транскрибируется как в прямом, так и в обратном направлениях. При транскрипции в прямой ориентации регуляция осуществляется путем продукции РНК, подавляющих экспрессию генов «хозяина» через РНК-индуцированный комплекс сайленсинга. Если сигнал полиаденилирования не функционирует, L1-транскрипты продолжают за пределы L1, приводя к 3'-транскрипции — образуются уникальные РНК. При обратной транскрипции 5'-транскрипция также вызывает образование уникальных РНК (Kitkumthorn, Mutirangura, 2011). Описана также прямая взаимосвязь гипометилирования L1 и гиперметилирования промоторов некоторых локусов. Внутригенные L1 продуцируют антисмысловые РНК в интронах и редуцируют уровни мРНК. Некоторые антисмысловые РНК вызывают метилирование ассоциированных CpG-островков. Гены, содержащие L1 зачастую репрессируются гиперметилированием промоторов (Khowutthitham et al., 2012). LTR-элементы содержат сильные промоторы и энхансеры транскрипции, а также регуляторные нуклеотидные последовательности (НП) для связывания с транскрипционными факторами, такими как рецепторы стероидных гормонов, факторы полиаденилирования и процессинга. При инсерции LTR могут регулировать сплайсинг РНК и выполнять роль альтернативных промоторов (Киселев, 2013).

Транспозиции МГЭ вносят значительный вклад в формирование онтогенетической изменчивости, но их высокая частота, являясь причиной геномной нестабильности, может быть опасна для жизни особи, поэтому в клетке работают механизмы, ограничивающие перемещение ТЕ. Изучение данных механизмов проводится на модельных системах — линиях, в которых повышена частота транспозиции МГЭ. Примером является модельная система нестабильной мутаторной линии дрозофилы. В работе Урусова с соавторами исследована стадийспецифичность транскрипции одиннадцати МГЭ группы *gypsy* в линии *Drosophila melanogaster* SS (Stable Strain), мутантой по гену *flamenco*, контролирующему перемещения ретроТЕ *gypsy*. В линии SS выявлен повышенный уровень транскрипции большинства изученных элементов на эмбриональной стадии и стадии имаго, что говорит об общих механизмах контроля экспрессии/транспозиции для всех элементов группы *gypsy* (Урусов и др., 2011), а также о том, что видоспецифические транспозиции определенных ТЕ участвуют в регуляции высвобождения геномной информации в зависимости от стадии развития клеточного деления и пространственного расположения

(тканеспецифично). У дрозофилы известно 2 типа регуляции, ограничивающей активность Р-элементов. Первый тип регуляции связан со стационарным состоянием Р-элементов в соматических клетках — из первичного транскрипта Р-элемента не вырезается третий интрон и вместо активной транспозазы образуется нефункциональный пептид. При этом активность Р-элемента имеет тканевую специфичность. Второй тип ограничения заключается в отсутствии транспозиционной активности Р-элемента в зародышевых клетках линии Р-типа и называется Р-цитотипом, в котором подавление Р-промотора приводит к ингибированию сплайсинга мРНК и синтезу репрессорного белка (Юшкова, Зайнуллин, 2009).

Инсуляторы являются регуляторными элементами, способствующими организации эукариотического хроматина путем блокирования энхансеров и активации хроматинового барьера. Показано, что некоторые инсуляторы могут иметь транспозонное происхождение. При исследовании роли MIR в качестве инсуляторов, Ванг (Wang, 2015) с соавторами выявили, что MIR-инсуляторы рекрутируют транскрипционные комплексы и энзимы модификации хроматина *in situ*, способствуя установлению хроматина и регуляторных доменов в геноме человека. Представление инсуляторов MIR-транспозонами в геноме человека говорит о специфическом механизме, при котором ТЕ используются для модуляции генных регуляторных сетей (Wang et al., 2015). Сравнительный анализ последовательностей геномов млекопитающих показал, что большое количество белок-некодирующих областей находятся под строгим селективным давлением. Некоторые из данных областей имеют происхождение от SINE. Древние SINE, названные AmnSINE1, имеющие химерную структуру из 5S рРНК и SINE тРНК-происхождения и содержащие общий центральный домен, общий также с SINE3 рыб Данио-рерио, объединены в суперсемейство DeuSINE. Данные AmnSINE1 могут функционировать в качестве промоторов или микроРНК, регулировать репликацию ДНК и организацию генома (Nishihara, 2006).

Таким образом, ТЕ в геноме служат важнейшим резервуаром изменчивости в ответ на стрессовые воздействия окружающей среды, создавая базу для эволюционных преобразований геномов в естественном отборе. При этом ТЕ служат удобной динамической системой, способной к саморегуляции с возможностью одновременного контроля экспрессии генов, а также резервуаром новых НП для построения регуляторных структур и конструирования новых белок-кодирующих генов. За последние годы появились работы, показывающие, что помимо случайных, не обязательных транспозиций, в онтогенезе регистрируются как

активация, так и перемещения ТЕ закономерные, тканеспецифические и зависящие от стадии развития. В 2012 году Ли (Lee, 2012) с соавторами представили работу, подтверждающую гипотезу о том, что структурные изменения в геноме параллельны возраст- и органоспецифическим фенотипам в сочетании с дифференциальной транскрипционной активностью ретроэлементов. В линии мышей C57BL/6J размеры геномов клеток печени были крупнее, чем клеток других органов, что совпадало с увеличением количества геномных копий некоторых ретроэлементов. Кроме того, наблюдалось дифференциальное увеличение размера генома клеток печени с возрастом, который достиг пика в 5 недель. Таким образом, исследователи сделали вывод о том, что структура генома отдельных особей варьирует в зависимости от возраста и типа органа в ассоциации с транспозицией ретроэлементов (Lee et al., 2012). При перемещении ТЕ по геному в онтогенезе может усиливаться или ослабляться функция определенных генов, причем тканеспецифично и в строгой закономерности от стадии развития. В 2015 году Ли (Lee, 2015) с соавторами представили новую работу на линии мышей C57BL/6J для исследования перестройки генома ткане- и стадии-специфически с отбором tandemных повторов мозаичных МГЭ. Были отмечены изменения конфигурации частей массива генома в коже и головном мозге всех мышей в возрасте 6 недель и старше, в то время как в тканях сердца и печени изменения были отмечены в 29 недель. Вариации во времени были подтверждены путем выявления переходов предполагаемых перестроек. Исследователи показали, что преобразования во времени и пространстве некоторых массивов повторяющихся нуклеотидных последовательностей способствуют приобретению новых характеристик информационной системы генома в процессе онтогенеза. При этом было обнаружено, что: 1) tandemный кластер мозаичных повторов в локусе ТЕ изменяется в геноме во времени в головном мозге и коже, начиная с 6-и недельного возраста мышей; 2) в отличие от кожи и мозга, tandemные кластеры повторов в геномах сердца и почек статистически выявлялись не раньше 29 недель; 3) существуют органотканевые специфические полиморфизмы структуры tandemных повторов кластеров среди особой линии C57BL/6J мышей; 4) структурная конфигурация кластера tandemного повтора внутри ТЕ матрицы CHR7.32 локуса может сильно меняться в зависимости от возраста мышей, а также типа органа и ткани. Таким образом, на экспериментальной линии мышей, ученые доказали возраст- и тканеспецифические перемещения ТЕ, меняющие размеры генома в зависимости от возраста и типа органа. Данные перемещения, как предполагают авторы, могут влиять на дифференцировку

клеток, играя решающее значение в онтогенетических преобразованиях многоклеточных эукариот (Lee et al., 2015). В работе Павлицева (Pavlicev, 2015) с соавторами, исследована тканеспецифическая регуляция LTR-элементами смежных генов. Авторы использовали корреляцию паттернов экспрессии в 18-и типах тканей и выявили тканеспецифическое различие в экспрессии генов по шести десяти двум различным LTR-классам. Данные паттерны оказались специфичны для ретровирусных инсерций, так как у видов, у которых те же гены не содержали LTR, подобные эффекты не наблюдались. На основании этого исследователями предполагается, что изменение активности LTR эндогенных ретровирусов в онтогенезе способствуют тканеспецифической регуляции прилегающих генов в процессе дифференцировки клеток (Pavlicev et al., 2015).

В пользу предположения о том, что ТЕ могут играть роль в тканеспецифической регуляции экспрессии генов (путем запрограммированной на уровне вида активаций ТЕ или их транспозиций) говорят некоторые исследования. Были получены данные о закономерном тканеспецифическом изменении активности ТЕ в тканях человека и мыши. В работе Го (Guo, 2016) с соавторами при сравнении СН (где Н может быть А, С, Т) метилирования, характерного для ТЕ в разных типах клеток и разных видах обнаружено изменение СН-метилирования активных L1 элементов (особенно характерное для более молодых L1 мыши и человека). Также было выявлено тканеспецифическое СН-метилирование L1 (Guo et al., 2016). При исследовании сахарной свеклы обнаружено, что вдоль генов и МГЭ метилирование CG и CHG выше, чем в соседних геномных регионах. В отличие от относительно низкого метилирования СНН областей в ретроТЕ и генах, уровень метилирования СНН в ДНК-ТЕ был значительно увеличен, что указывает на функциональную роль асимметричного метилирования в сайленсинге ДНК-ТЕ. При этом сравнение полногеномного метилирования ДНК между листьями и каллусом сахарной свеклы позволило выявить дифференцированное метилирование в культуре ткани. Потенциальные эпигаллы были гипометилированы в участках CG и CHG генов и ретроТЕ и гиперметилированы в участках СНН ДНК-транспозонов в каллусе по сравнению с листьями (Zakrzewski et al., 2017). Это указывает на кардинальное отличие в активности различных ТЕ в разных тканях растения, что позволяет предположить важное значение различия в активности разных ТЕ при дифференцировке тканей.

Экспериментальные данные показали широкое транскрибирование ТЕ и лишь частичную их репрессию теми же эпигенетическими маркерами, которые контролируют их транспозиционные

копии (Faulkner, 2011). Феномен всепроникающей транскрипции в настоящее время подтвержден. Выявлены сотни геномных областей, в которых транскрибируются длинные межгенные некодирующие РНК с тканеспецифическим характером активации (Патрушев, Коваленко, 2014). На разных этапах онтогенеза эукариот происходят закономерные стадии- и тканеспецифические транспозиции МГЭ, формируя в одном организме генетически отличающиеся соматические клетки в связи с геномным мозаицизмом в разных органах. Закономерный соматический мозаицизм выявлен при изучении нейрогенеза — показано, что для перемещений ТЕ не требуется клеточных делений — инсерции в новые области обнаруживаются *in vivo* при помощи обнаружения вариаций числа L1 в геноме. Ретротранспозиции создают генетически различные клеточные субпопуляции в мозге млекопитающих, что влияет на фенотипические особенности данных клеток и может иметь значение для формирования отдельных структур (Faulkner, 2011). Экспериментальные работы Гарсиа Перес (García-Perez, 2007), проведенные еще в 2007 году, показали, что эмбриональные стволовые клетки человека экспрессируют эндогенные L1-элементы и могут аккумулировать ретротранспозиции L1 *in vitro*. В результате ретротранспозиционных событий может происходить подавление целевых сайтов генов, что способствует дифференцировке клеток для выполнения определенной функции (García-Perez et al., 2007). Закономерные ретротранспозиции L1 продемонстрированы во многих экспериментальных исследованиях *in vitro* на различных трансформированных клеточных линиях человека и грызунов (Moran et al., 1996; De Berardinis et al., 1998; Wei et al., 2000; Morrish et al., 2002; Han, Boeke, 2004), в клетках предшественниках нейронов крыс (Muotri et al., 2005) и в первичных фибробластах человека (Kubo et al., 2006; Shi et al., 2007). Исследования *in vivo* на моделях мышей было показано, что экспрессия L1 и их ретротранспозиции закономерно и не случайно происходят на ранних этапах развития при дифференцировке определенных соматических тканей (Ostertag et al., 2002; Prak et al., 2003; Muotri et al., 2005; Marchetto et al., 2013). В недавних экспериментальных исследованиях с использованием клеточных культур, тканей человека и моделей мышей обнаружена широкомасштабная мобилизация L1 в нейронах грызунов и человека, а также активность ТЕ в головном мозге *Drosophila* (Richardson et al., 2014). В исследованиях генома человека и мыши показано, что мобилизация L1, Alu и SVA транспозонов происходит закономерно в раннем эмбриогенезе, вызывая широкомасштабные структурные вариации генома (Prak et al., 2003; Van den Hurk et al., 2007; Coufal et al., 2009; Macia et al., 2011). При этом обнаружено

сохранение потенциала транспозиций ТЕ в полипотентных стволовых клетках, что косвенно доказывает важное значение преобразований генома при дифференцировке клеток под действием активаций и перемещений ТЕ (Wissing et al., 2012; Klawitter et al., 2016), так как ТЕ обладают способностью контролировать экспрессию генов при их активации или инсерции в новые геномные регионы (Feschotte, 2008; Belancio et al., 2010; Kitkumthorn, Mutirangura, 2011).

О неслучайной, обязательной роли активации определенных ТЕ в процессах клеточной дифференцировки говорят исследования значения эндогенных ретровирусов в образовании плаценты. Выявлено, что важнейшими функциями эндогенных ретровирусов человека (HERV) в плаценте являются слияние трофобластов и их дифференцировка в условиях индуцированного HERV синцитиообразования. Эти функции эволюционно закрепились за ERVWE1 (HERV-W) и HERV-FRD (Киселев, 2013). Для выполнения ключевой функции в развитии плаценты необходима экспрессия генов (*env*), кодирующих гликопротеины оболочки ERV. Данные гены названы синцитинами, так как они способствуют образованию плацентарного объединенного клеточного слоя, называемого синцитиотрофобластом, на поверхности матки. Интересно, что использование генов различных эндогенных ретровирусов в эволюции происходило независимо в разнообразных линиях млекопитающих, таких как приматы (в том числе люди), грызуны, зайцевые, хищники и другие, приблизительно между 10 и 85 миллионами лет назад (Dupressoir et al., 2012). Об активации эндогенных ретровирусов как необходимым, а не случайном событии говорит обнаружение ретровирусных частиц в маточных железах опоссумов и родственных видов, у которых эндогенезированные ретровирусы демонстрируют выраженную экспрессию в маточных железах — у некоторых сумчатых также имеется ген синцитин, выполняющий важную функцию и проявляя специфическую экспрессию в короткоживущей плаценте (Cornelis et al., 2015).

На закономерность путей активации ТЕ в онтогенезе и их влияние на дифференцировку тканей с реализацией геномной информации во времени и пространстве указывают также работы, в которых установлена роль ТЕ в качестве источника микроРНК (Smalheiser, Torvik, 2005; Piriyaopongsa et al., 2007a; Piriyaopongsa et al., 2007b; Devor et al., 2009; Yuan et al., 2010; Borchert et al., 2011; Filshtein et al., 2011; Yuan et al., 2011; Nozawa et al., 2012; Tempel et al., 2012; Roberts et al., 2013; Gim et al., 2014; Platt et al., 2014; Spengler et al., 2014; Qin et al., 2015), так как роль микроРНК в эпигенетической регуляции тканеспецифических перестроек тканей доказана (Liu et al., 2012; Lu et al., 2012; Toledano et

al., 2012; Dimmeler, Nicotera, 2013; Du et al., 2013; Samantarrai et al., 2013; Ong et al., 2015; Shen et al., 2015). Также микроРНК влияют на специфичность метилирования генома (Morita et al., 2013; Samantarrai et al., 2013; Zhang et al., 2015), тогда как метилирование признано важным эпигенетическим фактором, влияющим на дифференцировку клеток (Ванюшин, 2013).

РОЛЬ ТРАНСПОЗОНОВ В ОНТОГЕНЕЗЕ ЧЕЛОВЕКА

Диспергированные повторы млекопитающих (MIR) — самое древнее семейство ТЕ в геноме человека. MIR, расположенные вокруг или внутри генов человека, позитивно коррелируют с тканеспецифической генной экспрессией. Анализ взаимосвязей между MIR и тканеспецифичными энхансерами человека с точки зрения локализации, хроматинового окружения, регуляторных функций показал, что MIR концентрированы в энхансерах K562 и HeLa клеточных типах человека. MIR-производные энхансеры показали строгую ассоциацию с уровнями генной экспрессии, тканеспецифической генной экспрессией и тканеспецифической клеточной функцией. MIR-производные энхансеры оказались богатым источником сайтов связывания с транскрипционными факторами (Jjingo et al., 2014). При анализе 10% всех картированных сайтов связывания с транскрипционными факторами (TFBS), обнаружено множество TFBS транспозонного происхождения. Обилие экспериментально охарактеризованных ТЕ-производных TFBS человека говорит о существенных регуляторных эффектах данного класса НП на геноме человека. ТЕ-производные TFBS, несмотря на функциональную ограниченность, эволюционируют чрезвычайно быстро по сравнению с уникальными сайтами. Такие быстро эволюционирующие TFBS использовались для формирования видоспецифических регуляторных фенотипов, то есть дивергентных паттернов экспрессии в эволюционной родословной человека (Polavarapu et al., 2008).

Хотя ДНК-ТЕ не активны в геноме человека, кооперирующая система вырезания и вставки ДНК вовлечена в рекомбинационные события, генерирующие антигенсвязывающее разнообразие лимфоцитов. В настоящее время активные ретроТЕ человека включают подмножество L1 с приблизительно 0,8% дивергенцией от консенсуса и элементов их мобилизации. Видоспецифические инсерции ответственные за минорные определяемые различия между нашим геномом и геномами шимпанзе или горилл. ТЕ ответственны за приблизительно 16 мегабаз различий в последовательности при попарном сравнении геномов, что составляет 0,5% от генома. Для контекста, различия между

видами семейства *Hominidae* составляют около 35 мегабаз замен (1,2% генома). Геном человека содержит около 2000 видоспецифических LINE (L1), 7000 Alu (предоминирует *AluYa5*, *AluYb8*), 1000 SVA, 73 LTR (главным образом HERV-K). Наиболее активными являются AluY, герминативные инсерции *de novo* которых составляют 1 на 20 новорожденных; L1 и SVA — 1 на 100–200 новорожденных (Burns, Boeke, 2012). Многие гены и псевдогены человека идентифицированы либо как произошедшие от ТЕ, либо созданные активностью ТЕ. Четверть экспериментально выявленных промоторов генов млекопитающих содержат НП ТЕ-происхождения. ТЕ используются для манипулирования регуляции генома в экспериментах и потенциально могут применяться в генной терапии человека. Локализация и корреляционная ориентация транспозонов/генов предполагает позитивную селективную роль для транспозонов и может быть использована для разработки алгоритмов поиска генов (Edlefsen, Liu, 2010). Более 1000 генов человека регулируются Alu-ассоциированными CpG-областями. L1-элементы содержат акцепторы/доноры сайтов сплайсинга и сигналы полиаденилирования, используемые в транскрипции и способствующие экспрессии генов и образованию гибридных L1-транскриптов. Alu-элементы обретают *cis*-регуляторные сигналы после инсерции при аккумуляции мутаций (Belancio et al., 2010). Маркеры метилирования в геноме человека, в частности, метилирования *AluYb8*, являются восприимчивыми к эпигенетическим изменениям от воздействия внутриутробной среды. Данные изменения играют решающую роль в посреднической функции плаценты, и в конечном итоге могут быть информативными для определения основных данных о здоровье или болезнях во время развития (Wilhelm-Benartzi et al., 2012). Эволюционно молодые ТЕ играют важную роль в развитии мозга человека. В частности, мутации определенных уникальных некодирующих РНК, содержащих ТЕ-последовательности, связаны с летальной энцефалопатией. Секвенирование ДНК линкерного локуса длиной 400 килобаз выявило точечную мутацию в примат-специфическом ретроТЕ, который транскрибировался как часть уникальной некодирующей РНК, экспрессирующейся в головном мозге. *In vitro* нокдаун данной РНК усиливал апоптоз нейронов (Cartault et al., 2012).

За последние несколько лет получены данные о выраженной ассоциации ТЕ с сайтами связывания транскрипционных факторов в геноме человека. При помощи программы FANTOM4 обнаружено более 20 000 кандидатов регуляторных регионов транспозонного происхождения. В некоторых случаях экспансия ТЕ в геноме человека ведет к существенной модернизации каналов

регуляторной сети. При исследовании отдельного транскрипционного фактора — α рецептора эстрогена ($ER\alpha$) — выявлено, что $ER\alpha$ является преимущественной мишенью строго определенных наборов ТЕ и данные ТЕ в комбинации с регуляторами транскрипции человека вовлечены в работу нескольких известных ко-регуляторов $ER\alpha$. Более того, значительное количество данных ТЕ оказались консервативными между человеком и мышью и локализованы в непосредственной близости (таким образом являясь кандидатами в регуляторы) с важными генами регуляции эстрогенов (Testori et al., 2012). Примерно одна треть всех связанных с альфа-рецепторами эстрогенов ($ER\alpha$) эстроген-чувствительных элементов (ERE) последовательностей находятся внутри ТЕ, чаще в семействах *Alu S*. Значение рецептор-связанных локусов продемонстрировано с использованием анализа репортера люциферазы, который доказал, что повторяющиеся элементы ERE последовательностей способствуют энхансерной функции (Mason et al., 2010). Несмотря на то, что с возрастом наблюдается глобальное гипометилирование генома, некоторые локусы, в число которых входят и рецепторы эстрогена, с возрастом подвергаются гиперметилированию (Cacabelos et al., 2015). Учитывая важное значение ТЕ в регуляции рецепторов эстрогенов, можно сказать, что ТЕ играют важную роль в гормон-зависимом дифференцированном управлении функционирования генома с возрастом.

Особая форма экзаптации ТЕ, также известная как доместикация, наблюдается когда ТЕ-кодируемые белки или домены становятся кооптированы в функциональные белки хозяина. Наиболее склонной к доместикации среди подобных белков является транспозаза. Способность транспозазы к доместикации впервые выявлена при первоначальном анализе генома человека, когда было обнаружено 47 генов, полностью или главным образом произошедших из ТЕ-кодируемых НП, причем все они, за исключением четырех, были связаны с транспозазой, несмотря на тот факт, что ДНК-ТЕ представляют скромную фракцию ТЕ человека (7%) и его генома (3%). Белки ТЕ-происхождения впоследствии были выявлены также в десятках исследований у животных, грибов, растений с применением строгих критериев функциональности ТЕ-производных генов. Данные исследования показали, что некоторые категории ТЕ-кодирующих НП доместизируются во множестве независимых событий, при этом транспозазы продолжают оставаться в качестве рекуррентного поставщика новых белков различных организмов. ТЕ могут быть источниками ДНК-связывающихся доменов. Например, транспозаза может распознавать специфические сайты ДНК в хромосомах, и, подобно ДНК-связывающимся белкам, может открывать доступ к хроматину путем взаимодействия

с комплексом хроматинового ремоделирования хозяина (Feschotte, 2008).

РОЛЬ НЕКОДИРУЮЩИХ РНК В РЕГУЛЯЦИИ ОНТОГЕНЕЗА

Видоспецифическая экспансия микроРНК имеет важное значение в регуляции онтогенеза, экспандирующиеся кластеры микроРНК имеют фундаментальное значение для развития эмбриональных стволовых клеток. Эффективность транскрипции и стабильность микроРНК регулируется тканеспецифично (Du et al., 2013). МикроРНК (miR) регулируют разнообразные биологические процессы, включая рост и дифференциацию стволовых клеток, а также регуляцию эндогенной репарации тканей, что имеет важное значение для развития регенеративной медицины в терапии стволовыми клетками (Ong et al., 2015). МикроРНК играют важную роль в репрограммировании соматических клеток, индуцируя плюрипотентные стволовые клетки (iPSCs). Фибробласты человека и мыши могут репрограммироваться в iPSCs при экспрессии четырех факторов Яманаки: Oct4, Sox2, Klf4, c-Мус. Другие соматические клетки также могут репрограммироваться в iPSCs. В частности, в исследованиях Lu с соавторами обнаружено, что гиперэкспрессия miR-25 или введение имитаторов miR-25 усиливают образование плюрипотентных стволовых клеток (Lu et al., 2012). МикроРНК играют важную роль в регуляции процесса развития, а также в зрелом возрасте и связанных со старостью процессах. Показано, что микроРНК miR-34 регулирует связанные с возрастом события и долгосрочные интеграции у дрозофилы — данная молекула является связующим звеном между старением и нейродегенерацией (Liu et al., 2012).

Предопределенный относительный размер органов — одна из наиболее характерных черт, отличающих многоклеточные виды. Аномальная активация главных эффекторов и транскрипционных ко-активаторов YAP в путях Hippo у дрозофилы вызывает резкие увеличения размеров органов в процессе развития. Передача сигналов YAP поддерживается за счет микроРНК-зависимой позитивной обратной связи. МикроРНК miR-130a, которая напрямую индуцируется при помощи YAP, может эффективно репрессировать VGLL4, ингибитор активности YAP, тем самым усиливая сигналы YAP (Shen et al., 2015). Экспрессия микроРНК у человека физиологически меняется с возрастом, а также при малигнизации, и многие исследования выявили различия в экспрессии микроРНК в неоплазмах и нормальных тканях. МикроРНК могут также эпигенетически модулировать многие гены путем контролирования уровней первичных эпигенетических регуляторов — ДНК-метилтрансфераз

и гистоновых деацетилаз (Samantarrai et al., 2013). В работе Толедано (Toledano, 2012) с соавторами при исследовании стволовых тестикулярных клеток дрозофилы установлено, что при старении клетки ниши вырабатывают микроРНК *let-7*, которая существует у целого ряда биологических видов, включая человека, и помогает определять очередность событий в процессе развития организма (Toledano et al., 2012). Получены многочисленные доказательства регуляции онтогенеза при помощи микроРНК. Например, потеря функции гена *miR-lin-4* у нематод, влияя на сигнальные пути *insulin/IGF-1*, значительно сокращает продолжительность их жизни. При старении многоклеточных эукариот уменьшается экспрессия супрессора стероидных рецепторов *let-7*. В экспериментах на мышах показано усиление экспрессии *miR-29* по мере старения, что вызывает снижение выработки коллагена и ухудшению свойств базальных мембран. У человека с возрастом выявлено подавление экспрессии *miR-24* и *miR-221* в моноцитах периферической крови. В культуре клеток человека с индуцируемым старением снижалась экспрессия *miR-17*, *miR-19b*, *miR-20a*, *miR-106a*. В коре головного мозга человека с возрастом усиливается активность *miR-34*, *miR-33b*, *miR-181* и *miR-1271*. В выборках головного мозга макак наблюдались сходные паттерны, включая дополнительно *miR-29b*. У мышей ассоциированное с возрастом усиление экспрессии *miR-30d*, *miR-34a*, *miR-468*, *miR-669b* и *miR-709* наблюдалось в головном мозге и ткани печени, тогда как уровень экспрессии *miR-22*, *miR-101a*, *miR-720*, *miR-721* был усилен только в головном мозге (Dimmeler, Nicotera, 2013). В то же время доказано, что каскад ретротранспозиций L1 участвует в нейрогенезе человека, что согласуется с происхождением многих микроРНК от ТЕ, определяя важную роль МГЭ в индивидуальном развитии. Перемещения ТЕ происходят при онтогенетическом развитии человека не только в эмбриогенезе человека, но и на протяжении всего онтогенетического развития. Причем вариации в одном организме формируют генетически различающиеся соматические клетки в связи с геномным мозаицизмом по ТЕ в различных тканях. Перемещения ТЕ не нуждаются в клеточных делениях и обнаруживаются в живых тканях при выявлении вариаций числа L1 в тканях человека (Faulkner, 2011). Таким образом, доказанная роль микроРНК в регуляции онтогенеза, а также регистрация закономерных перемещений (Ostertag et al., 2002; Prak et al., 2003; Muotri et al., 2005; Faulkner, 2011; Lee et al., 2012; Marchetto et al., 2013; Lee et al., 2015) или активаций (Prak et al., 2003; Van den Hurk et al., 2007; Coufal et al., 2009; Macia et al., 2011; Richardson et al., 2014; Pavlicev et al., 2015; Guo et al., 2016; Zakrzewski et al., 2017) ТЕ в онтогенезе, наряду с ролью ТЕ в качестве источника микроРНК (Smalheiser, Torvik,

2005; Piriyaopongsa et al., 2007a; Piriyaopongsa et al., 2007b; Devor et al., 2009; Yuan et al., 2010; Borchert et al., 2011; Filshtein et al., 2011; Yuan et al., 2011; Nozawa et al., 2012; Tempel et al., 2012; Roberts et al., 2013; Gim et al., 2014; Platt et al., 2014; Spengler et al., 2014; Qin et al., 2015), позволяет предположить, что ТЕ играют важную роль в процессах онтогенетической дифференцировки клеток и тканей, не случайного, а закономерного, видоспецифического характера, выработанную в процессе эволюции путем отбора наиболее оптимальных вариантов взаимосвязи ТЕ с эпигенетическими факторами.

РОЛЬ ТРАНСПОЗОНОВ В ГЕНЕЗЕ НЕКОДИРУЮЩИХ РНК

Некоторые копии ретроТЕ, инсертированные в некодирующие последовательности генома, эволюционируют либо как новые регуляторные последовательности, либо в качестве источника некодирующих РНК (нкРНК). Сами же ретроТЕ возникли как специфичные и временные мишени для регуляции малыми РНК в половых и в соматических клетках (Cartault et al., 2012). Доказано, что ТЕ могут использоваться как динамический резервуар для новых клеточных функций, так как могут эволюционировать в новые гены, полезные для «хозяина». Аналогичным образом ТЕ, кодирующие малые РНК, могут привести к возникновению новых РНК-кодирующих генов. Приобретая способность к ретротранспозиции, ТЕ-производные нкРНК могут значительно амплифицироваться и распространиться по всему геному. Некоторые из новых копий могут избежать предыдущих эволюционных ограничений, аккумулировать мутации, в результате потерять их первоначальную функцию и приобрести новую активность. В связи с этим транскрипты, происходящие от высокоповторяющихся последовательностей, могут быть богатым резервуаром для эволюции новых функциональных РНК (Matylla-Kulinska et al., 2014). Обширные литературные данные показали, что ТЕ способствуют глубоким эволюционным преобразованиям структуры и функции генома путем инсерции в ранее сформированные НП, как на уровне генов, например, в сайты связывания транскрипционных факторов, сайты сплайсинга, энхансеров, промоторов; так и на уровне функциональных РНК, например, микроРНК, элементов распознавания и белок-кодирующих доменов (Johnson, Guigo, 2014). Имеется тесная взаимосвязь между эпигенетическими факторами и ТЕ в геномах эукариот. На посттранскрипционном уровне в защите от ТЕ играет роль РНК-интерференция. РНК и эндогенные siРНК загружаются в белки семейства Аргонавтов и управляют комплексом сайленсинга против комплементарных НП транспозонов. Подобно рРНК

и siРНК, некоторые микроРНК (miR) также процессируются из геномных локусов, произошедших из ТЕ. Хотя miR не всегда задействованы в защите от ТЕ, канонические мишени для miR-регуляции, мРНК 3'-нетранслируемые области (3'UTR), часто содержат транспозонные НП. Вычислительные и широкомасштабные экспериментальные исследования подтверждают, что miR транспозонного происхождения (RdmiR) могут регулировать мишени мРНК путем комплементарности сайтов мишеней, локализованных внутри 3'UTR-содержащихся гомологов ТЕ. Например, анализ НП деградомных тэгов в клетках человека показал, что произошедшие от L2-элементов *miR-28-5p* и *miR-151* регулируют подмножество генов, включая *LYPD3* и *ATPAF1* через неканонические спаривания с 3'UTR-содержащими L2-элементами. Семейства miR, произошедшие не от ТЕ, также воздействуют на целевые мишени miR ТЕ-происхождения. В работе Шпенглер (Spengler, 2014) с соавторами выявлено, что miR *let-7* регулирует несколько генов человека через консервативные сайты, найденные внутри MIR, а сайты связывания для *miR-24* и *miR-122*, произошедшие от Alu, являются функциональными (Spengler et al., 2014). Основываясь на том, что микроРНК и сайты их мишеней мРНК могли сформироваться в эволюции параллельно в результате непрерывной колонизации древних ТЕ, Фильштейн (Filshtein, 2011) с соавторами разработали методологию выявления мРНК-мишеней для микроРНК, сформированных в более позднем эволюционном периоде, Orbid (Origin-based identification of microRNA targets), наиболее эффективную для древних микроРНК в поиске консервативных между видами сайтов мишеней. При помощи данной методики успешно найдены сайты мишеней для 191 уникальных микроРНК (92% из исследованных) после применения Orbid для серии из 208 отдельных микроРНК человека ТЕ-происхождения (Filshtein et al., 2011).

ТЕ играют важную роль в ингибировании экспрессии генов на посттранскрипционном уровне путем продукции микроРНК в геномах всех многоклеточных эукариот. Некоторые ТЕ имеют палиндромную структуру (MER), которая имеет наибольший потенциал для образования прекурсоров микроРНК (Gim et al., 2014). Инсерции ТЕ в новые сайты генома — одна из движущих сил эволюции, которая создает новые микроРНК. Молекулярное происхождение многих miR подтверждает гипотезу о том, что образование шпилек miR основано на инсерции двух родственных ТЕ, фланкирующих один геномный локус. Транскрипция, проходящая через эти локусы ведет к биогенезу функциональных miR. В одном из самых ранних исследований для выявления происхождения микроРНК от ТЕ с использованием биоинформационного

подхода, проведенном Смолхэйзером и Торвиком (Smalheiser N.R., Torvik V.I., 2005) проанализирована база данных miRNAi с использованием программного обеспечения по специфическому распознаванию известных повторов генома человека. В частности, 11 различных прекурсоров микроРНК содержали НП повторов (4 происходящие от LINE-2, другие — от SINE, LTR и простых повторов). Большинство данных miR оказались высококонсервативными для человека, мыши и крысы. Показано, что miR млекопитающих вырезаются из LINE-2 транспозонов и других геномных повторов (Smalheiser, Torvik, 2005). Последующее более глубокое изучение проведено Пирияпонгза (Piriyaongsa, 2007a) с соавторами, исследовавшими взаимосвязь между человеческими miR и ТЕ путем сравнения геномной локализации экспериментально охарактеризованных генов микроРНК человека с локализацией аннотированных геномных ТЕ. Была обнаружена корреляция, высокая идентичность между 7 членами семейства микроРНК *has-mir-548* и *Made1*. *Made1* и другие МТЕ имеют палиндромную структуру и при их транскрипции обнаруживается сегмент, который имеет несовершенную петлевую структуру. Так как связанные с РНК-интерференцией ферменты могут распознавать данный тип несовершенной петли и процессировать ее в зрелые микроРНК из 22 баз, авторы предположили, что *Made1* транскрипты процессируются в *has-mir-548miR*, что было подтверждено. В соответствии с происхождением от МТЕ, *hsa-mir-548* гены являются примат-специфичными и имеют множество потенциальных паралогов в геноме человека. Предполагается более 3500 целевых генов для *hsa-mir-548*, анализ профилей экспрессии и функциональных свойств которых показал канцер-специфичные регуляторные роли *has-mir-548* (Piriyaongsa et al., 2007a). При дальнейших исследованиях Пирияпонгза (Piriyaongsa, 2007b) с коллегами использовали сравнительные данные НП из Геномного Браузера UCSC и оценивали эволюцию miR ТЕ-происхождения (RdmiR — repeat-derived miR) у человека. При этом было обнаружено 55 экспериментально охарактеризованных RdmiR человека, которые потенциально регулируют тысячи генов. Сравнение НП RdmiR человека показало, что они в среднем менее консервативны, чем miR нетранспозонного происхождения. ТЕ обеспечивают естественный механизм возникновения miR, который способствует регуляторной дивергенции между видами (Piriyaongsa et al., 2007b). В работе Девор (Devor, 2009) с соавторами исследование 14-и ранее выявленных специфичных miR сумчатого домашнего опоссума и их фланкирующих НП, показало, что половина данных микроРНК эволюционировали от специфичных для сумчатых ТЕ. Все данные

ТЕ (за исключением одного Mariner) относились к LINE (Devor et al., 2009). Юань (Yuan, 2010) с соавторами выявили плаценто-специфичное семейство *miR-1302*, все члены которого происходили от одного ТЕ семейства MER53, одиннадцать членов которого распределены в геноме человека (Yuan et al., 2010). В дальнейшей работе Юань (Yuan, 2011) с соавторами было выявлено большое количество генов *miR*, которые пересекались с ТЕ: 226 генов у человека, 115 — у резуса, 141 — у мыши. Определены значительные различия между *miR*, образуемых из ТЕ от остальных классов по локализации в белок-кодируемых и межгенных областях геномов, по минимальной свободной энергии их шпилечных структур и их сохранности в геномах позвоночных. Обнаружены некоторые lineage-специфичные Rdmir и lineage-специфические экспансии семейств ТЕ (Yuan et al., 2011).

Транспозонное происхождение микроРНК предполагает еще одну интересную и предпочтительную роль одомашнивания. При процессинге одной ТЕ-последовательности машина РНК-интерференции может быть загружена с небольшой направляющей РНК, способной определить мишени всех РНК, содержащих ТЕ противоположной цепи. То есть сеть мишеней уже сформирована еще до установления первичных *miR*, которые «возникают» из серии случайных ТЕ-инсерций, когда это выгодно регуляторной нише. Данные транскрипции через интерфейс ТЕ устанавливают многие функциональные *miR*. В работе Борхерт (Borchert, 2011) с соавторами при анализе геномных событий, ответственных за формирование 15 000 аннотированных *miR* против источников базы данных ТЕ и нкРНК, обнаружено, что взаимосвязь между *miR* и ТЕ более значительна, чем оценивалось ранее, и в более широком плане, подтверждая важную роль ТЕ в происхождении *miR*, их экспрессии и формировании регуляторной сети. Выявлено, что 2392 отдельных микроРНК (около 15%) происходят непосредственно от ТЕ. Далее авторы исследовали точное происхождение микроРНК от ТЕ и показали, что чаще всего за генерацию *miR* отвечают ДНК-ТЕ — 891, другие распределены следующим образом: LTR — 414, LINE — 312, SINE — 353, сателлиты — 137 и другие — 136. Из них 136 *miR* произошли от некодирующих РНК-последовательностей (другие), которые также наиболее вероятно произошли из нехарактерных SINE-элементов. При этом определены таксоноспецифические и даже видоспецифические экспансии индивидуальных семейств *miR*. Например, выявлены 72 отдельных геномных локуса *miR-430* (57 у рыб Данио-рерио, 15 у рыбы *Japanese killifish*), каждый из которых формировался из сателлитов; *miR-1302* у лошади и приматов формировался из MER53 (у человека 12 локусов,

у шимпанзе — 8, у орангутанга — 5, у лошади — 6; что составляет 30 отдельных *miR-1302* шпилек — их геномный контекст говорит против общего предка, вместо этого предлагая конвергентную эволюцию). Анализ последовательностей *miR-466*, *miR-467*, *miR-669* семейств раскрыл их таксон-специфическую экспансию (например, 33 геномных локуса *miR-466* у разных животных: 1 — у человека, 1 — у курицы, 4 — у крысы, 27 — у мыши). Для *miR-467* (17 высоко-родственных локусов у мыши) — 50 локусов, каждый из которых возник из CR1 non-LTR РТЗ. Сходным образом, 31 локус *miR-669* (30 у мыши и 1 у коровы) был сформирован из EnSpm ДНК-ТЕ. Геномы растений также показали доказанную таксон-специфичную ТЕ-медирированную пролиферацию семейств *miR* (Borchert et al., 2011).

В работе Нодзава (Nozawa, 2010) с соавторами при исследовании геномов 12-и видов *Drosophila* и при анализе более 100 генов *miR* каждого вида, обнаружено, что, хотя многие гены *miR* образуются из случайных шпилечных структур интронных и межгенных регионов, генерации новых генов *miR* происходят также при их дублировании. Обнаружено, что как только гены *miR* приобретали определенную функцию, они начинали эволюционировать значительно медленнее в течение длительного времени (Nozawa et al., 2010). При исследовании геномов растений, Нодзава (Nozawa, 2012) с соавторами обнаружено, что ТЕ способствовали генерации видоспецифических генов микроРНК двудольных и однодольных растений. Важность указанных механизмов образования генов *miR* значительно отличается от таковых у *Drosophila*, где основными источниками являются шпилечные структуры. При этом молодые гены *miR* менее консервативны, чем древние гены у растений, также как и у *Drosophila* (Nozawa et al., 2012). Работы Шао (Shao, 2012) с соавторами продемонстрировали резкий переход от piРНК-подобных малых РНК ТЕ-происхождения (piРНК) в микроРНК в пре- и постгастрюляционный период куриных эмбрионов. В то же время, онтологический анализ генов показал, что обогащающиеся в ранних стадиях куриных эмбрионов микроРНК нацелены на множество путей сигнальной трансдукции, связанных с репродуктивным процессом и эмбриогенезом, включая Wnt и TGF-бета пути. Интересно, что значительная часть piРНК имеет первичное происхождение от активных и наиболее распространенных ТЕ, которые обогащаются в стадии X бластодермы кур (Shao et al., 2012). В геномах эукариот короткие неавтономные ТЕ могут иметь сходные размеры и шпилечную структуру с прекурсорами РНК. В исследованиях Темпель (Tempel, 2012) с соавторами при картировании всех прекурсоров *miR* в miRBase с использованием автоматизированного метода NcRNAClassifier выявлена корреляция

между количеством и местоположением кандидатов прекурсоров некодирующих РНК с наличием ТЕ. Между 1426 пре-miR у человека и 721 пре-miR у мыши из базы данных miRBase выявлено 235 и 68 mis-аннотированных пре-miR, полностью соответствующих транспозонам (Tempel et al., 2012). В исследованиях Цинь (Qin, 2015) с соавторами обнаружено 409 различных RdmiR у человека, пропорция которых оказалась больше, чем у других позвоночных, особенно не являющихся млекопитающими. Выявлено, что «голова» и «хвост» ТЕ-последовательностей вместе с соседними геномными НП, способствуют генерации микроРНК (Qin et al., 2015).

Наиболее распространенные семейства ТЕ в геноме человека, Alu, MIR, L2, обеспечивают платформу для miR-медирированной регуляции с управлением в 3'UTR мРНК. Возможность регуляции генов обеспечивают ТЕ, расположенные преимущественно внутри или вокруг генов (Spengler et al., 2014). МикроРНК ТЕ-происхождения имеют тенденцию к образованию филогено-специфических микроРНК. В данном отношении Alu-производные микроРНК примат-специфичны, а MITE-производные *miR-548* главным образом обнаруживаются у приматов. В геномах растений инсерции ТЕ могут образовывать как siРНК, так и микроРНК (Gim et al., 2014). В исследовании Роберт (Robert, 2013) с соавторами, проанализировавших происхождение 1213-и локусов miR, обнаружено, что большинство функциональных микроРНК

созданы транспозициями МГЭ. Также доказательно идентифицирован посредством описания формирования 273-х локусов miR новый механизм генерации локусов микроРНК посредством мутирования других форм некодирующих РНК (Roberts et al., 2013). В работе Ду (Du, 2013) с соавторами идентифицированы 9 семейств miR, экспансирующихся в геноме человека, а также ответственные за это новые геномные регионы, расположенные вблизи кодирования белков с консервативными функциями. Выявлено участие Alu и L1 в распространении кластеров микроРНК и сегментарных дупликациях, а также генерация новых miR при процессинге транскриптов псевдогенов. Обнаружено 9 новых семейств микроРНК, распространяющихся короткими фрагментами дупликации (средняя длина 430 баз): *has-mir622*, *has-mir-1282*, *has-mir-220*, *has-mir-199b*, *has-mir-3185*, *has-mir-492*, *has-mir-886*, *has-mir-1826*, *has-mir-1827* (Du et al., 2013). В исследованиях Платт (Platt, 2014) с соавторами подробный анализ секвенирования фракций малых РНК выявил большое количество предполагаемых микроРНК длиной в 17–24 баз у летучей мыши (613), собаки (535) и лошади (729). Из них miR ТЕ-происхождения у летучей мыши оказалось 61%, у собаки – 24%, у лошади – 17%. У собаки и лошади за формирование данных микроРНК, главным образом, были ответственны ретроТЕ, тогда как у летучей мыши ДНК-ТЕ. При этом сроки экспансии ДНК-ТЕ и результаты образования новых микроРНК совпадали с диверсификацией

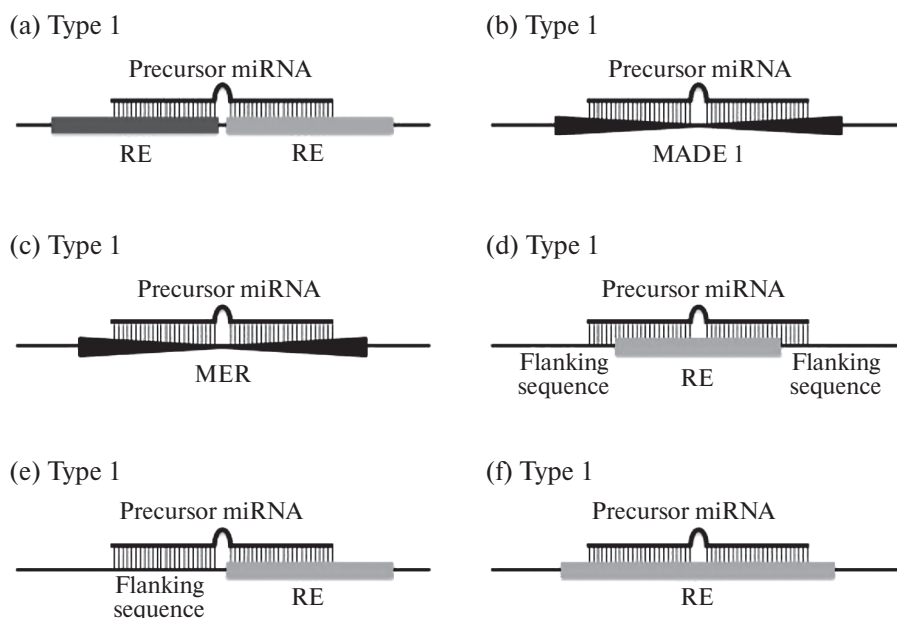


Рис. 1. Схема прекурсоров микроРНК, выделенных из повторяющихся элементов (ПЭ), локализованных в геноме. Тип 1 – прекурсор формируется из 2-х смежных ПЭ. Тип 2 – из MAD1. Тип 3 – из умеренных повторов (MER). Тип 4 – содержит ПН ПЭ. Тип 5 – из ПН, фланкированных между ПЭ и фланкированных ПН. Тип 6 – сохраняется внутри ПН ПЭ (Gim, 2014).

семейства *Vespertilionidae* (гладконосых летучих мышей) (Platt et al., 2014). В базе данных miRBase содержатся 55 экспериментально подтвержденных генов микроРНК, образованных из ТЕ, а также 85 новых микроРНК прогнозируются из потенциально законсервированных вторичных структур 587 ТЕ человека. При этом ТЕ способствуют продукции человеческих микроРНК-генов множеством механизмов. Gim с соавторами выявили 1900 различных микроРНК (406 ранее выявленных и 1494 новых) с подтвержденным ТЕ-происхождением и разделили их по способу образования на 6 типов (рис. 1) (Gim et al., 2014).

Обнаружено, что Alu-элементы встраиваются в гены длинных некодирующих РНК (lncРНК), формируя основу целевого распада мРНК через короткие несовершенные спаривания оснований. Несовершенные спаривания занимают важное место во взаимодействиях нкРНК с целевыми РНК и обнаруживаются во всех биологических царствах. Тысячи длинных межгенных некодирующих РНК (lncРНК) связаны с эндогенными LTR в клетках человека, которые могут обеспечивать lncРНК регуляторными сигналами (Hadjilargyrou, Delias, 2013). Значительная часть экзонов lncРНК организована из ТЕ (41% их НП имеют ТЕ происхождение, а 83% содержат, как минимум, один фрагмент ТЕ). Как следствие, многие зрелые lncРНК транскрипты содержат комбинацию множества повторяющихся фрагментов, напоминающих структуры домена белка. Отдельные семейства ТЕ строго и неслучайно обогащают или покидают lncРНК последовательности: найдены отдельные сверхпредставленные эндогенные ретровирусы человека в экзонах lncРНК в сравнении с общим геномным фоном, хотя другие классы субтипов LTR также обогащают lncРНК. В противоположность этому, многие семейства Alu, L1 и L2 классов в значительном количестве удаляются из lncРНК. Данные паттерны говорят об эволюционной селекции ТЕ в зрелых lncРНК последовательностях. ТЕ оказывают значительное воздействие на генную структуру lncРНК, особенно с точки зрения регуляторных областей и сайтов сплайсинга. Найдено множество примеров, по которым промоторы, сплайсинговые доноры и акцепторы, сайты полиаденилирования lncРНК состояли из НП ТЕ-происхождения. Содержание ТЕ в генах lncРНК намного превосходит в белок-кодирующих генах. Показано, что у 127-и lncРНК в плюрипотентных стволовых клетках специфически усиливается регуляция элементами HERVH. Предполагается, что HERVH фактически отвечают за рождение новых lncРНК путем инсерции активных промоторов в ранее неактивные области генома. Предполагается, что lncРНК обладают модульной организацией, подобно белкам, состоящих из дискретных доменов, комбинация

которых определяет функцию lncРНК. Увеличивается количество экспериментальных данных, подтверждающих идею о том, что фрагменты ТЕ внутри lncРНК способствуют их функционированию. Доказана собственная биологическая активность естественных НП транспозонных РНК – принимающие lncРНК приобретают аспекты первичной активности их ТЕ-повторов. Гипотеза о том, что ТЕ-последовательности могут использоваться в качестве функциональных доменов lncРНК недавно получила поддержку рядом экспериментальных исследований – фрагменты lncРНК ТЕ-происхождения функционируют в качестве РНК-, ДНК- и белок-связывающих доменов. Предполагается, что это отражает более общий феномен экзаптации во время эволюции lncРНК, когда инсертированные ТЕ-последовательности переориентируются в качестве сайтов распознавания как для белков, так и для нуклеиновых кислот. Гипотеза RIDL может объяснить каким образом функциональная эволюция может идти в ногу с быстрой эволюцией генов, наблюдаемым в lncРНК. В практическом отношении картирование ТЕ в будущем может использоваться для прогнозирования функции lncРНК, так как ТЕ обладают потенциалом для объяснения эволюции регуляторных сетей lncРНК (Johnsorn, Guigo, 2014).

ТРАНСПОЗОНЫ И АЛЬТЕРНАТИВНЫЙ СПЛАЙСИНГ В ОНТОГЕНЕЗЕ

Альтернативный сплайсинг имеет центральное значение в развитии и дифференцировке многоклеточных и является главным источником белкового разнообразия высших эукариот. Альтернативному сплайсингу подвергаются до 90% генов эукариот (Kalsotra, Cooper, 2012). При этом альтернативный сплайсинг для генов с жизненно-важными функциями значительно меняет активность их белковых продуктов. Например, изоформы белков натриевых каналов клеточной оболочки характеризуются выраженными различиями по функции (Farmer et al., 2012). Изоформы белка фактора роста эндотелия сосудов (VEGFA) отличаются тем, что одна изоформа проявляет про-ангиогенную активность, а другая – антиангиогенную (Pal et al., 2012). На образование альтернативных сплайсинговых вариантов белков влияют внеклеточные стимулы, такие как медиаторы активированных иммунных клеток и деполяризация нейронов. Например, повышение концентрации КС1 в клеточной культуре вызывает образование сплайсинговых изоформ белков плазматической мембраны. Образование сплайсинговых вариантов белков используется при ответе на стресс и альтерацию ДНК. Например, при вызванном аминокислотным голоданием стрессе у дрожжей ингибируется

сплайсинг генов, кодирующих рибосомальные белки (Cote et al., 2012).

L1 могут вводить дополнительные сайты сплайсинга при внедрении в экзоны — «экзонизация», создавая новые изоформы. При внедрении SINE в интроны генов возможно образование альтернативного сплайсингового продукта, что может способствовать появлению новых функций за счет приобретения дополнительных доменов белка. У растения *Arabidopsis thaliana* в интронных областях генов расположено около 3% ТЕ, сохранение которых указывает на их адаптационное значение (Le et al., 2015). При inserции LTR также могут регулировать сплайсинг РНК (Киселев, 2013). Находящиеся в интронах генов копии L1- и Alu-элементов снижают транскрипцию аллелей, содержащих эти ретроТЕ, по сравнению с аллелями, их не содержащими (Баскаев, Буздин, 2011). Inserции ретроТЕ могут приводить как к разрушению кодирующих НП ДНК, так и к модуляции активности генов за счет изменения свойств регуляторных НП, появлению новых сайтов связывания транскрипционных факторов, сигналов полиаденилирования и сайтов сплайсинга (Курносов, 2015). Транспозиции МГЭ часто служат в качестве движущей силы для превращения интронов в экзоны или для получения новых интронов, а также в качестве альтернативных сплайсинговых транскриптов. Таким образом, интеграции новых НП, а также баланс количества экзонов и интронов, длины и порядкового положения гена обеспечивают основным материалом для эволюции видов. При исследовании экспансии интронов ТЕ-происхождения в различных линиях позвоночных, Wang с соавторами выявили, что ТЕ и сателлиты распределены среди эукариотических геномов не случайным образом. С одной стороны, существует сильный негативный отбор для защиты важных последовательностей (белок-кодирующие последовательности или экзоны) в геномах для передачи основной генетической информации в относительно короткий в эволюционном масштабе промежуток времени. С другой стороны, ТЕ незаменимы в качестве основной силы и сырья в эволюции геномов для лучшей пригодности, усложнения и разнообразия, а также стимулирования видообразования и динамики популяций. В связи с этим ТЕ оказывают сильное воздействие на экспрессию и регуляцию генов косвенно через вариации длины интронов и их содержания. Один из механизмов, общих для всех изученных позвоночных — увеличение размеров интронов путем inserции ТЕ и сателлитов (в меньшей степени) (Wang et al., 2012).

Интроны вырезаются из эукариотических генов крупным динамичным РНК-белковым комплексом — сплайсосомой. Сплайсосома выявляет сайты сплайсинга в пре-мРНК с однонуклеотидной

точностью и часто клеточно-, стадийно- и гендер-специфически, способствуя транскриптомному и белковому разнообразию организмов за счет альтернативного сплайсинга. ТЕ способствуют белковому разнообразию и могут быть «экзонизированы», SINE в данном отношении чрезвычайно активны в эволюции. Сверхпредставление SINE среди экзонизированных ТЕ найдено также в отношении мутационно-индуцированных скрытых экзонов, которые ведут к наследственным болезням. Каждое подсемейство SINE способствует как традиционному, так и добавочному сплайсинговым мотивам экзонизированных последовательностей, повышая тем самым белковое разнообразие и контроль экспрессии генов через альтернативный сплайсинг. SINE сверхпредставлены в интронах, где они часто расположены в обратной ориентации в пределах короткого расстояния друг от друга, что способствует формированию очень стабильных структур створчатых петель и влияет на РНК процессинг. Створчатые петли (шпильки) — наиболее распространенные мотивы вторичных структур РНК, которые играют важную роль в транскрипции, РНК-процессинге, экспорте мРНК, стабильности мРНК, внутриклеточной локализации, трансляции и вирусной репликации. Хорошо известные шпильки, такие как чувствительный к железу элемент, участвуют в важных клеточных функциях, включая поглощение и хранение ионов, энергетический метаболизм, гипоксическая регуляция, нейрональная организация и формирование цитоскелета. Высокочувствительные шпилечные структуры, образованные из инвертированных SINE, могут быть основой для моделей изучения их роли в распознавании сплайсинговых сайтов. Выявлено, что короткие и открытые петли РНК являются ключевыми компонентами структурного кода сплайсинга для большинства репрессоров или промоторов экзонной активности сплайсинговыми сайленсерами или энхансерами. Данные структуры функционируют как ведущие модули экзонной селекции, которые контролируют активацию соседних экзонов, с открытыми петлями и выпуклостями, обеспечивающими ключевые контакты распознавания экзона. Размер шпильки MIR соответствует наиболее часто используемым РНК-складкам человека, предсказанным при помощи полногеномных подходов, со множеством тысяч представленных в интронах. Шпильки MIR также сходны с IRE — регуляторными консервативными створчатыми петлями, часто обнаруживаемыми в мРНК, которые кодируют различные белки с важными функциями (Kralovicova et al., 2015).

Таким образом, имеется сложная система взаимосвязи между альтернативным сплайсингом, транспозонами, ткане- и стадийно-специфической дифференцировкой клеток, а также ответом

на стресс (Cote et al., 2012). ТЕ чувствительны к изменениям микроокружения, воздействию различных белков и нкРНК за счет особенностей структуры, формирующей своеобразные ДНК-домены, отражающиеся в формировании доменов lncРНК (Johnson, Guigo, 2014). Это сказывается на возможности контролирования образования сплайсинговых изомеров стадийно- и ткане-специфично, что может закрепляться эволюционно за счет отбора особей с оптимальным соотношением инсерций ТЕ в гены, отвечающего наиболее подходящим ткане- и стадийно-специфичным дифференцировкам, а также жизнеспособному ответу на стресс с сохранением репродуктивной функции.

ТРАНСПОЗОНЫ И РЕГУЛЯТОРНАЯ СЕТЬ ГЕНОМА

Огромное количество исследований проиллюстрировало множество способов, при помощи которых ТЕ непосредственно влияют на регулирование экспрессии близлежащих генов на транскрипционном и посттранскрипционном уровнях. Было выявлено, что около 25% экспериментально охарактеризованных промоторов генома человека содержат НП ТЕ-происхождения, включая эмпирически определенные *cis*-регуляторные элементы. Многие промоторы и сигналы полиаденилирования генов человека и мыши произошли от специфичных для данных видов ТЕ. Инсерции данных ТЕ способствовали созданию линии дифференцировки специфических паттернов экспрессии генов. Доказано, что ТЕ, обычно приобретающие регуляторную функцию, происходят из фрагментов ТЕ, которые глубоко консервативны среди млекопитающих. Данные элементы имеют тенденцию образовывать кластеры вокруг генов, вовлеченных в развитие и транскрипционную регуляцию. Они представлены в избыточных количествах в пределах прогнозируемых *cis*-регуляторных модулей, содержащих плотные массивы сайтов связывания транскрипционных факторов в геномных сегментах. Одна пятая часть плацентарно-специфических повторяющихся элементов, тысячи из которых произошли от ТЕ, перекрываются с сайтами гиперчувствительности к ДНКазе I, картированных экспериментально в лимфоцитах человека, то есть они обеспечивают промоторные последовательности или сайты связывания для регуляторных белков. Также выявлено множество высококонсервативных ТЕ в качестве транскрипционных энхансеров или как альтернативные сплайсинговые экзоны, вводящие преждевременные стоп-кодоны и иницирующие нонсенс-медиированный распад, тем самым способствуя гомеостазу мРНК в клетке. Повсеместное накопление ТЕ создает сырье для последовательностей, из которых *cis*-регуляторные

элементы эволюционируют *de novo* при точечных мутациях. Другой способ связан с тем, что *cis*-регуляторные элементы предрасположены внутри ТЕ во время их инсерции и кооптируются сразу после их вставки или после модификации окружающей средой (Feschotte, 2008).

Диспергированная экспансия ТЕ семейств в геноме потенциально позволяет одним и тем же регуляторным мотивам, которые будут использоваться во многих хромосомных местах, управлять множеством генов в той же регуляторной сети. Данная модель экспериментально доказана. К примеру, сеть тесно взаимосвязанных LTR-элементов распределена в более 1500 почти идеальных сайтах связывания для регуляторного фактора «хозяина» р53. Интересно, что все содержащиеся в р53 семейства LTR примат-специфичны, и р53 сайты присутствуют во время их инсерции. Это пример того, что распределение LTR-элементов вызывает образование примат-специфичной сети р53-регуляторных генов (Feschotte, 2008). В соответствии с ТЕ, геном рассматривается как динамичная система, активно реагирующая на внешнесредовые изменения. Ген *TP53* — онкосупрессор, мутации в котором широко распространены во всех типах злокачественных опухолей. Его основная функция — поддержание геномной стабильности, участие в апоптозе, репарации и ангиогенезе. Его полифункциональность нуждается в комплексной регуляции, в которую вовлечены МГЭ. В исследованиях Пидпала (Pidpala, 2006) с соавторами было обнаружено, что ТЕ в пределах гена *TP53* и его 5'-фланкирующей области преимущественно образуют кластеры, содержащие различные семейства и подсемейства ТЕ. В гене наиболее часто встречающимся оказались Alu-элементы, среди LINE-элементов L1 присутствовали исключительно внутри гена, тогда как L2 — преимущественно в 5'-фланкирующем участке. Совсем не обнаружено LTR-элементов, ДНК-транспозоны в незначительном количестве были представлены лишь внутри гена. МГЭ распределены случайным образом, главным образом, в межгенных или интронных областях, реже в экзонах. Во многих генах МГЭ или НП их происхождения присутствуют в 5'- или 3'-фланкированных областях промоторов, где они выполняют важную роль для функционирования генома и регуляции генной активности (Pidpala et al., 2006).

Особый интерес в поиске геномных элементов, полиморфизм которых наиболее высок и в большей степени по сравнению с другими вносит вклад в фенотипическую изменчивость представляют полиморфизмы по ТЕ. Выраженный полиморфизм и видоспецифичность ретроТЕ, высокая скорость их дивергенции даже за короткое время расхождения групп организмов от общего предка, наглядно описанная на примере геномов лабораторных

линий мышей, позволяют полагать, что основным источником геномной изменчивости являются ТЕ. Одной из выраженных особенностей геномного размножения ретроТЕ является волнообразная динамика их возникновения, распространения и деградации. Для целого ряда ретроТЕ описаны эволюционные циклы, включающие вертикальную или горизонтальную передачу, вспышку транспозиций с последующим разрушением большинства первоначальных копий. Следы таких циклов является присутствие в хромосомах разных видов их множественных остатков (Глазко, 2013). МГЭ выполняют роль своего рода рецепторов стрессорных сигналов, инициирующих вспышки транспозиционной изменчивости в критические периоды эволюции популяций. Автономные макромолекулярные комплексы (пекеласомы) могут захватить центромеры и превратиться в минихромосомы, эволюционирующие в дальнейшем как самостоятельные генетические структуры (Lukash, 2007). Таким образом, наблюдается взаимосвязь онтогенеза с эволюцией на примере использования ТЕ «хозяином». ТЕ представляют собой с одной стороны резервуар для изменчивости под действием стресса, что изменяет фенотип клеток, позволяя в отдельных случаях лучшей адаптации и выживаемости. С другой стороны, ТЕ являются эволюционно запрограммированными регуляторами онтогенетической программы развития, когда в каждом клеточном делении экспрессируется строго определенная часть генов в зависимости от дифференцировки ткани, микроокружения и стадии деления. Фенотипическая изменчивость в зависимости от стрессового воздействия говорит о роли ТЕ — характер реагирования в зависимости от особенностей распределения ТЕ у конкретной особи способствует такой реакции на изменение окружающей среды, которая вызывает лучшее приспособление и выживаемость. Конкретная особь выживает и получает потомство с большей долей вероятности и передает потомству программу распределения ТЕ в той же последовательности, что способствует закреплению данной особенности и фенотипической адаптации в зависимости от транспозонной структуры генома.

Среди нкРНК длинные нкРНК (lncРНК) и малые нкРНК могут изменять модификации гистонов, целевое метилирование ДНК и сайленсинг генов. Эндогенная ДНК-метилтрансфераза 1 (DNMT1) кооперирует с ингибированием нкРНК. МикроРНК напрямую связываются с DNMT1 с высокой аффинностью. Основываясь на полученных данных, Чжан (Zhang, 2015) с соавторами показали, что гиперэкспрессия *miR-155-5p* приводит к изменению метилирования ДНК путем ингибирования DNMT1 активности, вызывая изменение экспрессии генов (Zhang et al., 2015). Члены

семейства *miR-29* непосредственно нацелены на ДНК-метилтрансферазы — DNMT3A, DNMT3B. Нарушение уровней экспрессии *miR-29* были связаны с онкогенезом и агрессивностью опухолей. Члены семейства *miR-29* подавляют метилирование ДНК и супрессируют туморогенез путем защиты от метилирования *de novo*, они напрямую репрессируют активность DNMT и ДНК-деметилазы (метилцитозин диоксигеназы (TET1) и тимин ДНК гликозилазы (TDG)), которые оказывают противоположное действие в контроле статуса метилирования ДНК (Morita et al., 2013). Учитывая важное значение ТЕ в генезе микроРНК, их тесной взаимосвязи в регуляции функционирования генома в онтогенезе, можно предположить наличие сложной системы взаиморегуляции с участием метилирования, нкРНК, модификации гистонов и ТЕ. При этом любое изменение в одном из звеньев данной системы может вызвать изменение в функционировании генома, отражающееся фенотипически.

Согласно Барбаре Мак-Клинтон, ТЕ являются классическими моделями эпигенетического наследования: «молчащие» ТЕ могут быть активированы и затем наследованы в активном состоянии (McClintock, 1984). Сайленсинг транспозонов связан с метилированием ДНК, деацетилированием гистонов, метилированием лизина-9 в гистоне H3 (H3mK9) и РНК-интерференцией, которые взаимосвязаны между собой. Например, у *Neurospora* и *Arabidopsis* метилирование ДНК может быть вызвано H3mK9 и наоборот. У млекопитающих метилCpG-связанный белок привлекает к участию гистоновую деацетилазу (HDAC) и гистоновую H3 лизин-9 метилтрансферазу. Также поддерживающая ДНК метилтрансфераза млекопитающих, DNMT1, прямо взаимодействует с HDAC. И, наконец, у делящихся дрожжей *Schizosaccharomyces pombe*, РНК-интерферирующая машина направляет взаимосвязь H3mK9 с центромерными повторами. У *Caenorhabditis elegans* и *Drosophila* с потерей ДНК метилирования, сайленсинг транспозонов сохраняется в зародышевой линии при помощи РНК-интерференции. При исследовании *Arabidopsis* Липпман (Lippman, 2003) с соавторами показали, что у мутантов изменения ДНК-метилирования, H3mK9 и РНК-интерференции были вовлечены в эпигенетическую регуляцию транспозонов. Например, хроматиновое ремоделирование АТФазы *DDM1*, *Dnmt1* гомолога *MET1* и *HDAC HDA6* изменяли сайленсинг и метилирование последовательностей ТЕ. В то же время для сайленсинга ретроТЕ *TA3/ATCOPIA44* требуется ДНК метилтрансфераза *CMT3* и *KYP/SUVH4* (Lippman et al., 2003).

В исследовании Го (Guo, 2016) с соавторами собран 51 метилом ДНК человека и мыши из

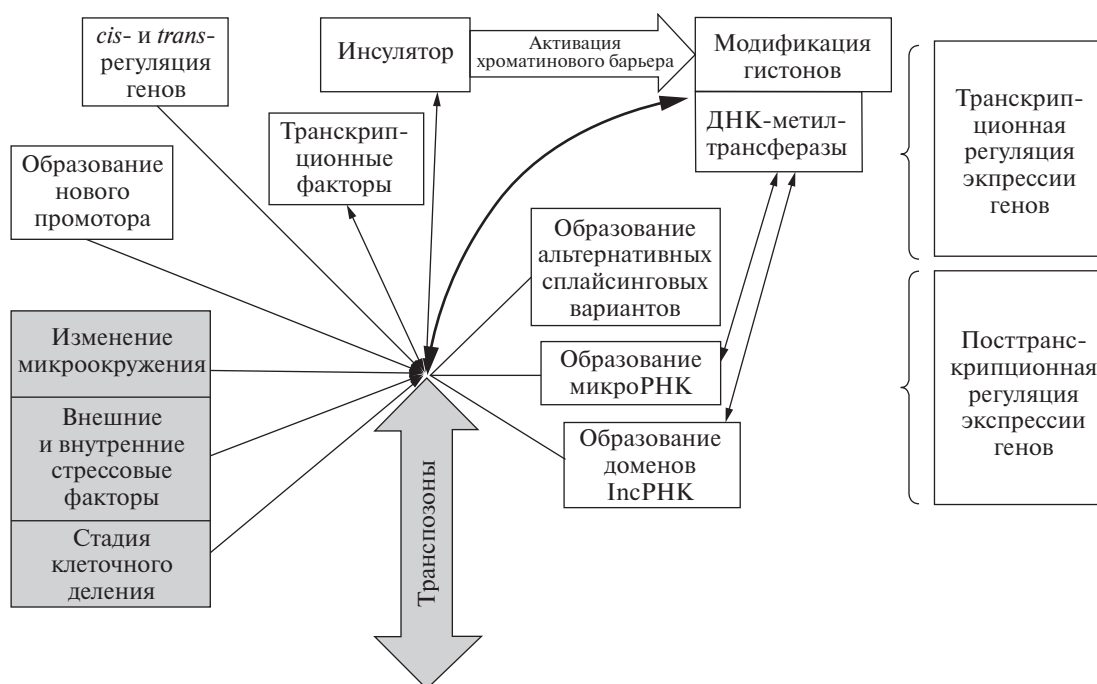


Рис. 2. Схема участия транспозонов в эпигенетической регуляции онтогенеза.

нейронов головного мозга, эмбриональных стволовых клеток и индуцированы полипотентные стволовые клетки, зародышевые клетки и яйцеклетки. Проведено двухмерное сравнение СН (где Н может быть А, С, Т) метилирования, характерного для ТЕ, для разных типов клеток и разных видов. В нейронах головного мозга обнаружен недостаток СН-метилирования активных L1 элементов, тогда как их CG-метилирование не изменялось, что говорит о важном значении СН-метилирования в регуляции развития нейронов. Более молодые L1 элементы в головном мозге показывали меньшую степень метилирования как у человека, так и у мыши, что говорит о консервативной роли СН-метилирования в эволюции L1. Как Alu человека, так и B1 мыши, предпочли высокую степень СН-метилирования в специфических локусах в процессе эволюции. Специфическое СН-метилирование в интронах и LINE также оказалось тканеспецифическим и консервативным. СН-метилирование, таким образом, высоко-консервативно между видами и динамически регулируется в каждом типе клеток, а также потенциально может быть вовлечено в эволюцию ТЕ (Guo et al., 2016). PIWI-взаимодействующие РНК (piРНК) связаны с сайленсингом ТЕ, а также влияют на специфичное метилирование ДНК. Выявлено, что piРНК экзонного происхождения почти для всех видов состоят из 10 или меньшего количества геномных копий, тогда как существующие в большом количестве копий piРНК происходят преимущественно из интронных или

межгенных областей, где располагается подавляющее большинство ТЕ (Fu et al., 2014). На примере работы Цуй (Cui, 2011) с соавторами можно показать взаимосвязь ТЕ с белок-кодирующим геном p53. Так, показано, что примат-специфичные Alu-повторы играют важную роль в формировании p53 регуляторной сети в контексте с хроматином. Один из селективных факторов, ответственных за частую встречаемость Alu-повторов в интронах может быть связан с p53-медиированной регуляцией Alu-транскрипции, которая, в свою очередь, влияет на экспрессию генов. Онкосупрессорный ген p53 вовлечен в сложную регуляторную сеть, медирующую экспрессию около 1000 генов человека. Недавние исследования показали, что многие сайты связывания p53 *in vivo* расположены в повторах транспозонов. В то же время функциональные p53 чувствительные элементы также расположены в ТЕ, особенно обильно в Alu-элементах. Найдено около 400 000 потенциальных p53 сайтов связывания в Alu-элементах во всем геноме. Данные предполагаемые сайты связывания расположены в тех же областях Alu повторов, что и функциональные p53 чувствительные элементы. Аффинитет связывания p53 с находящимися в Alu сайтами, как правило, коррелирует с возрастом подсемейств Alu — наиболее сильные сайты встраиваются в относительно молодые Alu-повторы (Cui et al., 2011).

Вновь возникшие гены miR имеют скорость замен сходную с таковыми в синонимичных

нуклеотидных сайтах белок-кодирующих генов с нейтральным эволюционированием, тогда как древние гены miR проявляют значительно более низкую скорость по сравнению с синонимическими заменами. То есть, как только гены miR приобретали определенную функцию, они начинали эволюционировать значительно медленнее в течение длительного времени (Nozawa et al., 2010), на основании чего можно предположить существование механизма, регулирующего скорость мутирования ТЕ и их производных. Существование подобной сложной системы саморегулирования обеспечивает оптимальный отбор наиболее приспособленных организмов в эволюции. Схема участия ТЕ в регуляции онтогенеза во взаимосвязи с механизмами эволюции приведена на рисунке 2.

ЗАКЛЮЧЕНИЕ

Таким образом, несмотря на сложившееся определение эпигенетики как науки о механизмах регуляции генома, не затрагивающей последовательности ДНК, представлены данные о материальных основах эпигенетической наследственности, закодированной в нуклеотидных последовательностях, главным образом транспозонного происхождения. ТЕ играют важнейшую роль в регуляции онтогенетического развития многоклеточных эукариот в кооперации с эпигенетическими факторами. При инсерции в новую область последовательности транспозонов могут выполняться роль альтернативных промоторов, регуляторов полиаденилирования, метилирования и сплайсинга; их транскрипция может приводить к 3' и 5'-транскрипции с образованием уникальных РНК. Кроме того, в результате мутаций возможно приобретение ТЕ новых функций, образование новых некодирующих РНК, влияние на функционирование геномной сети. Доказаны тканеспецифические и стадийспецифические активации строго определенных ТЕ, характерные для определенных видов, что указывает на эволюционный отбор онтогенетических функций ТЕ. В то же время, высокая мутабельность ТЕ в онтогенезе и чувствительность к стрессу указывает на закономерности изменений генома в индивидуальном развитии, способствующие эволюционным преобразованиям и естественному отбору особей, особенности активации ТЕ которых приводят к наилучшему приспособлению. Важная роль ТЕ в генезе некодирующих РНК и их обратная взаимосвязь дает возможность в дальнейших исследованиях использовать полученную информацию для возможного регулирования онтогенетических преобразований генома, в частности, замедлить процессы старения, находить наиболее оптимальные противораковые препараты. Многие функции ТЕ являются эволюционно закрепленным на уровне вида паттерном активаций при развитии организма, что

выражается в пространственно-временной дифференцировке клеток. Специфическая дифференцировка клеток происходит в связи с чувствительностью ТЕ к микроокружению клеток и тесной взаимосвязью с эпигенетическими факторами, многие из которых сами имеют транспозонное происхождение. Наличие большинства ТЕ в интронах белок-кодирующих генов, наряду с чувствительностью ТЕ к внутренним и внешним стрессовым факторам дает возможность стресс-индуцированному отбору экспрессии сплайсинговых вариантов белков в онтогенезе. Данный феномен дает также возможность регуляции экспрессии сплайсинговых изомеров в зависимости от ткани и стадии дифференцировки.

СПИСОК ЛИТЕРАТУРЫ

- Баскаев К.К., Буздин А.А. Эволюционно недавние группы генетических мобильных элементов в геноме человека // Вавиловский журнал генетики и селекции. 2011. Т. 15. № 2. С. 313–322.
- Ванюшин Б.В. Эпигенетика сегодня и завтра // Вавиловский журнал генетики и селекции. 2013. Т. 17. № 4. С. 805–832.
- Васильева Л.А., Антоненко О.В., Выхристюк О.В. и др. Селекция изменяет паттерн мобильных генетических элементов в геноме *Drosophila melanogaster* // Вестник ВОГиС. 2008. Т. 12. № 3. С. 412–425.
- Васильева Л.А., Выхристюк О.В., Антоненко О.В. и др. Индукция транспозиций мобильных генетических элементов в геноме *Drosophila melanogaster* различными стрессовыми факторами // Вестник ВОГиС. 2007. Т. 11. № 3/4. С. 662–671.
- Глазко В.И. Проблемы «селекции с помощью маркеров» // Генетика. 2013. № 2. С. 16–22.
- Киселев О.И. Эндогенные ретровирусы: структура и функции в геноме человека // Вопросы вирусологии. 2013. № 1. С. 102–115.
- Курносоев А.А. Активность ретроэлементов в нейрональных тканях взрослого организма. Диссертация на соискание ученой степени кандидата биологических наук. М. 2015. 103 с.
- Патрушев Л.И., Коваленко Т.Ф. Функции некодирующих последовательностей генома млекопитающих // Успехи биологической химии. 2014. Т. 54. С. 39–102.
- Урусов Ф.А., Нефедова Л.Н., Ким А.И. Анализ тканевой стадийспецифичности транскрипции ретротранспозонов группы Gypsy у *Drosophila melanogaster* // Вавиловский журнал генетики и селекции. 2011. Т. 15. № 2. С. 283–288.
- Чересиз С.В., Юрченко Н.Н., Иванников А.В., Захаров И.К. Мобильные элементы и стресс // Вестник ВОГиС. 2008. Т. 12. № 1/2. С. 216–241.
- Юрченко Н.Н., Коваленко Л.В., Захаров И.К. Мобильные генетические элементы: нестабильность генов

- и геномов // Вавиловский журнал генетики и селекции. 2011. Т. 15. № 2. С. 261–270.
- Юшкова Е., Зайнуллин В. Транспозиционная активность Р элементов в природных и хронически облученных популяциях дрозофилы // Вестник ИБ (института биологии Коми НЦ УрО РАН). 2009. № 9. С. 21–26.
- Belancio V.P., Roy-Engel A.M., Deininger P.L. All y'all need to know 'bout retroelements in cancer // Semin Cancer Biol. 2010. V. 20. № 4. P. 200–210.
- Borchert G.M., Holton N.W., Williams J.D. et al. Comprehensive analysis of microRNA genomic loci identifies pervasive repetitive-element origins // Mobile Genetic Elements. 2011. V. 1. P. 8–17.
- Burns K.H., Boeke J.D. Human Transposon Tics // Cell. 2012. V.149. № 4. P. 740–52.
- Cacabelos R., Torrellas C. Epigenetics of Aging and Alzheimer's Disease: Implications for Pharmacogenomics and Drug Response // International Journal of Molecular Sciences. 2015. V. 16. P. 30482–30543.
- Cartault F., Munier P., Benko E. et al. Mutation in a primate-conserved retrotransposons reveals a noncoding RNA as a mediator of infantile encephalopathy // Proc Natl Acad Sci USA. 2012. V. 109. № 13. P. 4980–4985.
- Castellano L., Rizzi E., Krell J. et al. The germline of the malaria mosquito produces abundant miRNAs, endo-siRNAs and 29-nt small RNAs // BMC Genomics. 2015. V. 16. № 1. P. 100–106.
- Cornelis G., Vernochet C., Carradec Q. et al. Retroviral envelope gene captures and syncytin exaptation for placentalization in marsupials // Proc. Natl. Acad. Sci. USA. 2015. V. 112. № 5. P. 487–496.
- Cote G.J., Zhu W., Thomas A. et al. Hydrogen Peroxide Alters Splicing of Soluble Guanylyl Cyclase and Selectively Modulates Expression of Splicing Regulators in Human Cancer Cells // PLoS ONE. 2012. V. 7. Is. 7. e41099.
- Coufal N.G., Garcia-Perez J.L., Peng G.E. et al. L1 retrotransposition in human neural progenitor cells // Nature. 2009. V. 460. № 7259. P. 1127–1131.
- Cui F., Sirotin M.V., Zhurkin V.B. Impact of Alu repeats on the evolution of human p53 binding sites // Biology Direct. 2011. V.6. P. 2–22.
- De Berardinis R.J., Goodier J.L., Ostertag E.M., Kazanietz H.H. Rapid amplification of a retrotransposons subfamily is evolving the mouse genome // Nat. Genet. 1998. V. 20. № 3. P. 288–290.
- De Koning A.P., Gu W., Castoe T.A. et al. Repetitive Elements May Comprise Over Two-Thirds of the Human Genome // Plos Genetics. 2011. V.7. Iss.12. e1002384.
- Devor E.J., Peek A.S., Lanier W. et al. Marsupial-specific microRNAs evolved from marsupial-specific transposable elements // Gene. 2009. V. 448. № 2. P. 187–91.
- Dimmeler S., Nicotera P. MicroRNAs in age-related diseases // EMBO Mol. Med. 2013. V. 5. № 2. P. 180–190.
- Du Z., Yang C., Rothschild M.F., Ross J. Novel microRNA families expanded in the human genome // BMC Genomics. 2013. V.14. P. 98–105.
- Duan C.G., Wang X., Pan L. et al. A pair of transposon-derived proteins function in a histone acetyltransferase complex for active DNA demethylation // Cell. Res. 2017. V. 27. № 2. P. 226–240.
- Dupressoir A., Lavialle C., Heidmann T. From ancestral infectious retroviruses to bona fide cellular genes: role of the captured syncytins in placentalization // Placenta. 2012. V. 33. № 9. P. 663–671.
- Edlefsen P.T., Liu J.S. Transposon identification using profile HMMs // BMC Genomics. 2010. V.11. S10.
- Farmer C., Cox J., Fletcher E. Splice Variants of Nav1.7 Sodium Channels Have Distinct beta Subunit-Dependent Biophysical Properties // PLoS ONE. 2012. V. 7. Is. 7. e41750.
- Faulkner G.J. Retrotransposons: mobile and mutagenic from conception to death // FEBS Lett. 2011. V. 585. № 11. P. 1589–1594.
- Feschotte C. The contribution of transposable elements to the evolution of regulatory networks // Nat Rev Genet. 2008. V.9. № 5. P. 397–405.
- Filshstein T.J., Mackenzie C.O., Dale M.D. et al. Origin-based identification of microRNA targets // Mobile Genetic Elements. 2011. V.2. № 3. P.184–192.
- Finatto T., de Oliveira A., Chaparro C. et al. Abiotic stress and genome dynamics: specific genes and transposable elements response to iron excess in rice // Rice. 2015. V. 8. № 13. doi. 10.1186/s12284-015-0045-6.
- Fu A., Jacobs D.I., Zhu Y. Epigenome-wide analysis of piRNAs in gene-specific DNA methylation // RNA Biol. 2014. V. 11. № 10. P. 1301–1312.
- Garcia-Perez J.L., Marchetto M.C., Muotri A.R. et al. LINE-1 retrotransposition in human embryonic stem cells // Hum. Mol. Genet. 2007. V. 16. № 13. P. 1569–1577.
- Gim J., Ha H., Ahn K. et al. Genome-Wide Identification and Classification of microRNAs derived from repetitive elements // Genomic Inform. 2014. V.12. № 4. P. 261–267.
- Guo W., Zhang M.Q., Wu H. Mammalian non-CG methylations are conserved and cell-type specific and may have been involved in the evolution of transposon elements // Sci Rep. 2016. V. 6. P. 32207–32219.
- Hadijargyrou M., Delias N. The Intertwining of Transposable Elements and Non-Coding RNAs // Int J Mol Sci. 2013. V. 14. № 7. P. 13307–1328.
- Han J.S., Boeke J.D. A highly active synthetic mammalian retrotransposons // Nature. 2004. V. 429. № 6989. P. 314–318.
- Huang W., Tsai L., Li Y. et al. Widespread of horizontal gene transfer in the human genome // BMC Genomics. 2017. V. 18. P. 274–285.
- Jjingo D., Conley A., Wang J. et al. Mammalian-wide interspersed repeat (MIR)-derived enhancers and the regulation of human gene expression // Mobile DNA. 2014. V.5. P. 14–25.

- Johnson R., Guigo R.* The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs // *RNA*. 2014. V.20. № 7. P. 959–976.
- Kalsotra A., Cooper T.A.* Functional consequences of developmentally regulated alternative splicing // *Nat. Rev. Genet.* 2012. V. 10. P. 715–729.
- Khowutthitham S., Ngamphiw C., Wanichnopparat W. et al.* Intragenic long interspersed element-1 sequences promote promoter hypermethylation in lung adenocarcinoma, multiple myeloma and prostate cancer // *Genes and Genomics*. 2012. V. 34. Is. 5. P. 517–528.
- Kitkumthorn N., Mutirangura A.* Long interspersed nuclear element-1 hypomethylation in cancer: biology and clinical applications // *Clin Epigenet.* 2011. V. 2. P. 315–330.
- Klawitter S., Fuchs N.V., Upton K.R. et al.* Reprogramming triggers endogenous L1 and Alu retrotransposition in human induced pluripotent stem cells // *Nat. Commun.* 2016. V. 7. P. 10286–10301.
- Kralovicova J., Patel A., Searle M., Vorechovsky I.* The role of short RNA loops in recognition of a single-hairpin exon derived from a mammalian-wide interspersed repeat // *RNA Biology*. 2015. V.12. № 1. P. 54–69.
- Kubo S., Selem M.C., Soifer H.S. et al.* L1 retrotransposition in non-dividing and primary human somatic cells // *Proc. Natl. Acad. Sci. USA*. 2006. V. 103. P. 8036–8041.
- Le T.N., Miyazaki Y., Takuno S. et al.* Epigenetic regulation of intragenic transposable elements impacts gene transcription in *Arabidopsis thaliana* // *Nucleic Acids Res.* 2015. V. 43. № 8. P. 3911–3921.
- Lee K.H., Chiu S., Lee Y.K. et al.* Age-dependent and tissue-specific structural changes in the C57BL/6J mouse genome // *Exp. Mol. Pathol.* 2012. V. 93. № 1. P. 167–172.
- Lee K.H., Yee L., Lim D. et al.* Temporal and spatial rearrangements of a repetitive element array on C57BL/6J mouse genome // *Exp. Mol. Pathol.* 2015. V. 98. № 3. P. 439–445.
- Liu N., Landreh M., Cao K. et al.* The microRNA miR-34 modulates ageing and neurodegeneration in *Drosophila* // *Nature*. 2012. V. 482. № 7386. P. 519–523.
- Lippman Z., May B., Yordan C. et al.* Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification // *PLoS Biology*. 2003. V.1. № 3. P.420–428.
- Lu D., Davis M.P., Abreu-Goodger C. et al.* MiR-25 Regulates Wwp2 and Fbxw7 and Promotes Reprogramming of Mouse Fibroblast Cells to iPSCs // *PLoS One*. 2012. V. 7. № 8. e40938.
- Lukash L.L.* Mutagenesis induced by integration processes and evolution of nuclear genome // *Biopolymers and cell*. 2007. V.23. № 3. P. 172–187.
- Macia A., Munoz-Lopez M., Cortes J.L. et al.* Epigenetic control of retrotransposons expression in human embryonic stem cells // *Mol. Cell. Biol.* 2011. V. 31. № 2. P. 300–316.
- Marchetto M.C., Narvaiza I., Denli A.M. et al.* Differential L1 regulation in pluripotent stem cells of humans and apes // *Nature*. 2013. V. 503. № 7477. P. 525–529.
- Mason C.E., Shu F.J., Wang C. et al.* Location analysis for the estrogen receptor- α reveals binding to diverse ERE sequences and widespread binding within repetitive DNA elements // *Nucleic Acids Res.* 2010. V. 38. № 7. P. 2355–2368.
- Masuta Y., Nozawa K., Takagi H. et al.* Inducible Transposition of a Heat-Activated Retrotransposon in Tissue Culture // *Plant. Cell. Physiol.* 2017. V. 58. № 2. P. 375–384.
- Matylla-Kulinska K., Tafer H., Weiss A., Schroeder R.* Functional repeat-derived RNAs often originate from retrotransposon-propagated ncRNAs // *Wiley Interdiscip Rev RNA*. 2014. V.5. № 5. P. 591–600.
- McClintock B.* The significance of responses of the genome to challenge // *Science*. 1984. V. 226. № 4676. P. 792–801.
- Moran J.V., Holmes S.E., Naas T.P. et al.* High frequency retrotransposition in cultured mammalian cells // *Cell*. 1996. V. 87. P. 917–927.
- Morita S., Horii T., Kimura M. et al.* MiR-29 Represses the Activities of DNA Methyltransferases and DNA Demethylases // *Int J Mol Sci*. 2013. V. 14. P. 14647–14658.
- Morrish T.A., Gilbert N., Myers J.S. et al.* DNA repair mediated by endonuclease-independent LINE-1 retrotransposition // *Nat. Genet.* 2002. V. 31. P. 159–165.
- Muotri A.R., Chu V.T., Marchetto M.C. et al.* Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition // *Nature*. 2005. V. 435. № 7044. P. 903–910.
- Nali L.H., Oliveira A.C., Alves D.O. et al.* Expression of human endogenous retrovirus K and W in babies // *Arch. Virol.* 2017. V. 162. № 3. P. 857–861.
- Nishihara H., Smit A., Okada N.* Functional noncoding sequences derived from SINES in the mammalian genome // *Genome Res*. 2006; 16(7):864–874.
- Nozawa M., Miura S., Nei M.* Origins and evolution of microRNA genes in *Drosophila* species // *Genome Biol Evol*. 2010. V. 2. P. 180–189.
- Nozawa M., Miura S., Nei M.* Origins and evolution of microRNA genes in plant species // *Genome Biol Evol*. 2012. V. 4. № 3. P. 230–239.
- Ong S., Lee W.H., Kodo K. et al.* MicroRNA-mediated Regulation of Differentiation and Transdifferentiation in Stem Cells // *Adv Drug Deliv Rev*. 2015. V.88. P. 3–15.
- Ostertag E.M., De Berardinis R.J., Goodier J.L. et al.* A mouse model of human L1 retrotransposition // *Nat. Genet.* 2002. V. 32. № 4. P. 655–660.
- Pal S., Gupta R., Davuluri R.* Alternative transcription and alternative splicing in cancer // *Pharmacol Ther*. 2012. V. 136. № 3. P. 283–294.
- Pidpala OV, Iatsyshyna AP, Lukash LL.* Analysis of distribution of mobile genetic elements within the human TP53 gene and its 5-flanking region // *Biopolymers and cell*. 2006. V. 22. Is.1.

- Pavlicev M., Hiratsuka K., Swagqart K.A. et al. Detecting endogenous retrovirus-driven tissue-specific gene transcription // *Genome Biol. Evol.* 2015. V. 7 № 4. P. 1082–1097.
- Piriyapongsa J., Jordan I.K. A Family of Human MicroRNA Genes from Miniature Inverted-Repeat Transposable Elements // *PLoS One*. 2007a. V.14. № 2. e203.
- Piriyapongsa J., Marino-Ramirez L., Jordan I.K. Origin and evolution of human microRNAs from transposable elements // *Genetics*. 2007b. V.176. № 2. P.1323–1337.
- Platt R.N., Vandeweyer M.W., Kern C. et al. Large number of novel miRNAs originate from DNA transposons and are coincident with a large species radiation in bats // *Mol Biol Evol.* 2014. V. 31. № 6. P. 1536–1545.
- Polavarapu N., Marino-Ramirez L., Landsman D. et al. Evolutionary rates and patterns for human transcription factor binding sites derived from repetitive DNA // *BMC Genomics*. 2008. V. 9. P. 226–235.
- Prak E.T., Dodson A.W., Farkash E.A., Kazazian H.H. Tracking an embryonic L1 retrotransposition event // *Proc. Natl. Acad. Sci. USA*. 2003. V. 100. № 4. P. 1832–1837.
- Qin S., Jin P., Zhou X. et al. The Role of Transposable Elements in the Origin and Evolution of MicroRNAs in Human // *PLoS One*. 2015. V. 10. № 6. e0131365.
- Richardson S.R., Morell S., Faulkner G.J. L1 retrotransposons and somatic mosaicism in the brain. *Annu Rev Genet* // 2014. V. 48. P. 1–27.
- Roberts J.T., Cooper E.A., Favreau C.J. Formation from transposable element insertions and noncoding RNA mutations // *Mobile Genetic Elements*. 2013. V. 1. № 6. e27755.
- Samantarrai D., Dash S., Chhetri B. et al. Genomic and epigenomic cross-talks in the regulatory landscape of miRNAs in breast cancer // *Mol. Cancer Res*. 2013. V.11. № 4. P. 315–328.
- Shao P., Liao J., Guan D. et al. Drastic expression change of transposon-derived piRNA-like RNAs and microRNAs in early stages of chicken embryos implies a role in gastrulation // *RNA Biology*. 2012. V. 9. № 2. P. 212–227.
- Shen S., Guo X., Yan H. et al. A miR-130a-YAP positive feedback loop promotes organ size and tumorigenesis // *Cell Research*. 2015. V. 25. P. 997–1012.
- Shi X., Seluanov A., Gorbunova V. Cell divisions are required for L1 retrotransposition // *Mol. Cell. Biol.* 2007. V. 27. P. 1264–1270.
- Smallheiser N.R., Torvik V.I. Mammalian microRNAs derived from genomic repeats // *Trends Genet.* 2005. V. 21. № 6. P. 322–326.
- Song X., Cao X. Transposon-mediated epigenetic regulation contributes to phenotypic diversity and environmental adaptation in rice // *Curr Opin Plan. Biol.* 2017. V.36. P. 111–118.
- Spengler R.M., Oakley C.K., Davidson B.L. Functional microRNAs and target sites are created by lineage-specific transposition // *Human Molecular Genetics*. 2014. V. 23. № 7. P. 1783–1793.
- Tempel S., Pollet N., Tahi F. NeRNAclassifier: a tool for detection and classification of transposable element sequences in RNA hairpins // *BMC Bioinformatics*. 2012. V. 13. P. 246–258.
- Testori A., Caizzi L., Cutrupi S. et al. The role of Transposable Elements in shaping the combinatorial interaction of Transcription Factors // *BMC Genomics*. 2012. V.12. P. 400–416.
- Toledano H., D'Alterio C., Czech B. et al. The let-7-lmp axis regulates ageing of the Drosophila testis stem-cell niche // *Nature*. 2012. V. 485. № 7400. P. 605–610.
- Van den Hurk J.A. Meij I.C., Selem M.C. L1 retrotransposition can occur early in human embryonic development // *Hum. Mol. Genet.* 2007. V. 16. № 13. P. 1587–1592.
- Wang D., Su Y., Wang X. et al. Transposon-Derived and Satellite-Derived Repetitive Sequences Play Distinct Functional Roles in Mammalian Intron Size Expansion // *Evol Bioinform Online*. 2012. V.8. P. 301–319.
- Wang J., Vicente-Garcia C., Seruqia D. et al. MIR retrotransposons sequences provide insulators to the human genome // *Proc Natl Acad Sci USA*. 2015. V. 112. № 32. P. 4428–4437.
- Wei W., Morrish T.A., Alisch R.S. A transient assay reveals that cultured human cells can accommodate multiple LINE-1 retrotransposition events // *Anal. Biochem*. 2000. V. 284. P. 435–438.
- Wilhelm-Benartzi C.S., Houseman E.A., Maccani M.A. et al. In utero exposures, infant growth, and DNA methylation of repetitive elements and developmentally related genes in human placenta // *Environmental Health Perspectives*. 2012. V. 120. № 2. P. 296–302.
- Wissing S., Munoz-Lopez M., Macia A. et al. Reprogramming somatic cells into iPS cell activates LINE-1 retroelement mobility // *Hum. Mol. Genet.* 2012. V. 21. № 1. P. 208–218.
- Yuan Z., Sun X., Jiang D. et al. Origin and evolution of a placental-specific microRNA family in the human genome // *BMC Evol Biol*. 2010. V.10. P. 346–358.
- Yuan Z., Sun X., Liu H. et al. MicroRNA genes derived from repetitive elements and expanded by segmental duplication events in mammalian genomes // *PLoS ONE*. 2011. V.6. Iss.3. e17666.
- Zakrzewski F., Schmidt M., Van Lijsebettens M., Schmidt T. DNA methylation of retrotransposons, DNA transposons and genes in sugar beet (*Beta vulgaris* L.) // *Plant. J.* 2017. doi: 10.1111/tpj.13526.
- Zhang G., Esteve P., Chin H.G. et al. Small RNA-mediated DNA (cytosine-5) methyltransferase 1 inhibition leads to aberrant DNA methylation // *Nucleic Acids Research*. 2015. V. 43. № 12. P. 6112–6124.

The Role of Transposons in Epigenetic Regulation of Ontogenesis

R. N. Mustafin^{1,*} and E. K. Khusnutdinova^{1,2}

¹*Bashkir State University, Ufa, 450076 Russia*

²*Institute of Biochemistry and Genetics, Ufa Scientific Center, Russian Academy of Sciences, Ufa, 450054 Russia*

*e-mail: ruji79@mail.ru

Received December 22, 2016; in final form, June 1, 2017

A new insight into the mechanisms underlying implementation of genomic information in the individual development of eukaryotes through interactions of transposons with epigenetic factors dynamically changing during each cell division is described. These mechanisms of stepwise implementation of individual genetic information with characteristic stage- and tissue-specific features in the activities of certain mobile genetic element families are evolutionarily fixed at the species level. In addition, the individual differences caused by their “unscheduled” transpositions can significantly change the regulatory network of the genome altering the phenotype. These changes in individual development can bring about new traits leading to either a disease or better fitness and represent an important component of the variation for natural selection in evolution. A large part of the eukaryotic transposons is altered by mutations and used for formation of the regulatory gene network, changes in the protein-coding genes, and emergence of new nonprotein-coding genes. When inserted into new loci, mobile genetic elements form the basis for microRNA and the domain structures of long noncoding RNA, responding to various types of stress; this is reflected in the specific features of individual development and contributes to variation. The epigenetic factors, including noncoding RNA, DNA methylation, and histone modifications, are tightly associated with mobile genetic elements. The specific features in transposon location in individuals that have emerged owing to spontaneous mutations or those caused by stress impacts can considerably change the interactions in gene networks. This influences the likelihood of survival under changing environmental conditions and reflects a distinct interrelation between the mechanisms of individual development and evolution. There is a parallelism between the mechanisms underlying the rearrangements of genomes caused by transposons in evolution and in individual development. In particular, the responsiveness of transposons to external and internal (microenvironment) factors forms the background for evolutionary construction of transposon-mediated tissue-specific activation patterns of certain transposons during each cell division, which leads to maturation of a reproductive organism. This mechanism is based on tight stage- and tissuespecific interrelation between transposons, epigenetic factors, and protein-coding genes.

Keywords: long noncoding RNA, microRNA (miR), methylation, mobile genetic elements, histone modification, nucleotide sequences, environment, repetitive elements, transposons, repeat-derived miR (RdmiR)