

Некодирующие части генома как основа эпигенетической наследственности

Р.Н. Мустафин¹✉, Э.К. Хуснутдинова^{1, 2}

¹ Башкирский государственный университет, Уфа, Россия

² Институт биохимии и генетики Уфимского научного центра РАН, Уфа, Россия

Анализ литературных данных дает возможность предположить, что основой эпигенетических преобразований геномов в онтогенезе являются особенности распределения, количества и состава мобильных генетических элементов. Транспозоны составляют большую часть геномов многоклеточных эукариот, эволюционное сохранение данных структур сопряжено с двумя универсальными механизмами управления дифференцировки клеток – процессингом некодирующих РНК и регуляцией сплайсинга. Данные универсальные механизмы первоначально были направлены на защиту от вирусов и мобильных генетических элементов, однако в дальнейшем кооперация защитных систем с механизмами управления взаимосвязи клеток и их дифференцировкой стала причиной возникновения и эволюции многоклеточных. В пользу этого говорят эволюционное сохранение комплекса взаимосвязанных ферментов Droscha, Dicer, Argonaut, RdRP и их гомологов практически у всех многоклеточных, а также отсутствие данных ферментов у одноклеточных. Интроны происходят от мобильных генетических элементов, в распространении и регуляции интронов важную роль играют транспозоны с их продуктами экспрессии. Транспозоны регулируют экспрессию генов *in cis* и *in trans*, а также опосредованно путем продукции малых РНК, влияющих на собственную активность мобильных генетических элементов, как путем изменения метилирования ДНК и модификацией гистонов, так и посттранскрипционно. Кроме того, транспозоны рассматриваются в качестве важных источников длинных некодирующих РНК, участвующих в регуляции дифференцировки клеток. Закономерное изменение активности транспозонов в онтогенезе тканеспецифично и стадийно-специфично и сопряжено с экспрессией специфических некодирующих РНК транспозонного происхождения, изменяющих активность генов при дифференцировке клеток. Предполагается, что видоспецифические особенности активации транспозонов при каждом последующем делении клеток проходят эволюционный отбор и используются в качестве ключевых регуляторов роста и развития организма. Начиная с первого деления зиготы, расположение и состав транспозонов в геноме влияют на их наследуемую активацию в каждом последующем клеточном делении. Это вызывает изменение экспрессии определенных генов и дифференцировку клеток, в результате чего развивается целостный многоклеточный организм.

Ключевые слова: альтернативный сплайсинг; интроны; мобильные генетические элементы; некодирующие РНК; система РНК-интерференции; транспозоны; процессинг.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Мустафин Р.Н., Хуснутдинова Э.К. Некодирующие части генома как основа эпигенетической наследственности. Вавиловский журнал генетики и селекции. 2017;2017;21(6):742-749. DOI 10.18699/VJ17.30-0

HOW TO CITE THIS ARTICLE:

Mustafin R.N., Khusnutdinova E.K. Non-coding parts of genomes as the basis of epigenetic heredity. Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding. 2017;2017;21(6):742-749. DOI 10.18699/VJ17.30-0 (in Russian)

Received 30.05.2017

Accepted for publication 07.07.2017

Online 02.10.2017

© AUTHORS, 2017

✉ e-mail: ruji79@mail.ru

Non-coding parts of genomes as the basis of epigenetic heredity

R.N. Mustafin¹✉, E.K. Khusnutdinova^{1, 2}

¹ Bashkir State University, Ufa, Russia

² Institute of Biochemistry and Genetics, Ufa Research Center, Ufa, Russia

We hypothesized that the basis of epigenetic regulation of genomes in ontogenesis is the specificity of the distribution, number and composition of transposons. Transposons constitute the major part of the genomes of multicellular eukaryotes. The evolutionary preservation of transposons is associated with universal mechanisms for controlling cell differentiation: processing of non-coding RNAs and splicing regulation. These universal mechanisms were originally aimed at protecting against viruses and transposons. The cooperation of these protective systems with mechanisms for controlling the interrelation of cells and their differentiation became the basis for the emergence and evolution of multicellular eukaryotes. The evolutionary conservation of a complex enzymes Droscha, Dicer, Argonaut, RdRP and their homologues in all multicellular eukaryotes, and their absence in unicellular organisms supports this assumption. Introns originated from mobile genetic elements. Transposons played an important role in the propagation of introns in evolution and their regulation in ontogenesis. Transposons regulate the expression of genes *in cis* and *in trans*, and also indirectly by the production of small RNAs that affect their own activity, both by altering the DNA methylation and modifying histones, and at the posttranscriptional level. Tissue-specific and stage-specific changes in the activity of transposons in ontogenesis are associated with the expression of transposon-derived noncoding RNAs and altering the activity of genes, which leads to cell differentiation. We proposed that the species-specific features of activation of transposons for each subsequent cell division undergo evolutionary selection and are key regulators of the growth and development of the organism. We proposed that transposons in the genome affect their inherited activation in each subsequent cell division, which causes a change in cell differentiation.

Key words: alternative splicing; introns; mobile genetic elements; noncoding RNA; RNA interference; transposons; processing.

В настоящее время эпигенетику принято считать наукой о наследуемых свойствах организма, не связанных с изменением собственно нуклеотидной последовательности ДНК. К числу известных эпигенетических механизмов относятся энзиматическое метилирование ДНК, гистоновый код и замалчивание генов малыми РНК (Ванюшин, 2013). Однако за последние годы накапливается все больше данных, позволяющих предположить, что основой эпигенетической наследственности являются конкретные структуры генома, связанные с некодирующей ДНК в межгенных и интронных областях. Причиной любых эпигенетических процессов должны быть изменения материальной структуры, для которой наиболее подходят состав, распределение и количество транспозонов (TE – transposable elements), специфичные для особей одного вида. В пользу данного предположения свидетельствует выявление у многоклеточных животных после оплодотворения яйцеклетки во вновь образованной зиготе глобального эпигенетического перепрограммирования генома, при котором вплоть до стадии бластоцисты обнаруживается глобальное деметилирование генома эмбриона, при котором стираются импринты мужских и женских геномов. Дифференцировка клеток внутренней клеточной массы с образованием трех зародышевых листков сопровождается установлением характерного для каждого из них паттерна метилирования (Баранов, Кузнецова, 2007). То есть метки метилирования вначале удаляются, после чего тканеспецифично и стадийспецифично распределяются. За процесс перераспределения меток метилирования должны отвечать конкретные структуры. В то же время установлены тканеспецифические и стадийспецифические особенности экспрессии некодирующих РНК (Dimmeler, Nicotera, 2013; Du et al., 2013; Samantarrai et al., 2013; Ong et al., 2015; Shen et al., 2015), регуляции альтернативных сплайсинговых вариантов генов (Feschotte, 2008; Belancio et al., 2010; Luco et al., 2011; Dumesic, Madhani, 2013), а также характер активации транспозонов (Ostertag et al., 2002; Prak et al., 2003; Muotri et al., 2005; Van den Hurk et al., 2007; Coufal et al., 2009; Macia et al., 2011; Marchetto et al., 2013). Транспозоны играют важную роль в генезе микроРНК (Borchert et al., 2011; Yuan et al., 2011; Gim et al., 2014; Platt et al., 2014; Qin et al., 2015), siРНК (Shabalina, Koonin, 2008; Xu et al., 2013; Zhang et al., 2016), piРНК (Biryukova, Ye, 2015), lncРНК (Singh, Rath, 2012; Hadjiargyrou, Delihis, 2013; Johnson, Guigo, 2014) и регуляции сплайсинга (Luco et al., 2011). Так как регуляция сплайсинга, микроРНК, siРНК, piРНК, lncРНК имеет важное значение в онтогенетической регуляции развития органов и тканей, можно предположить, что транспозоны – материальная основа эпигенетической наследственности.

Влияние некодирующих РНК на дифференцировку клеток, метилирование ДНК и модификацию гистонов

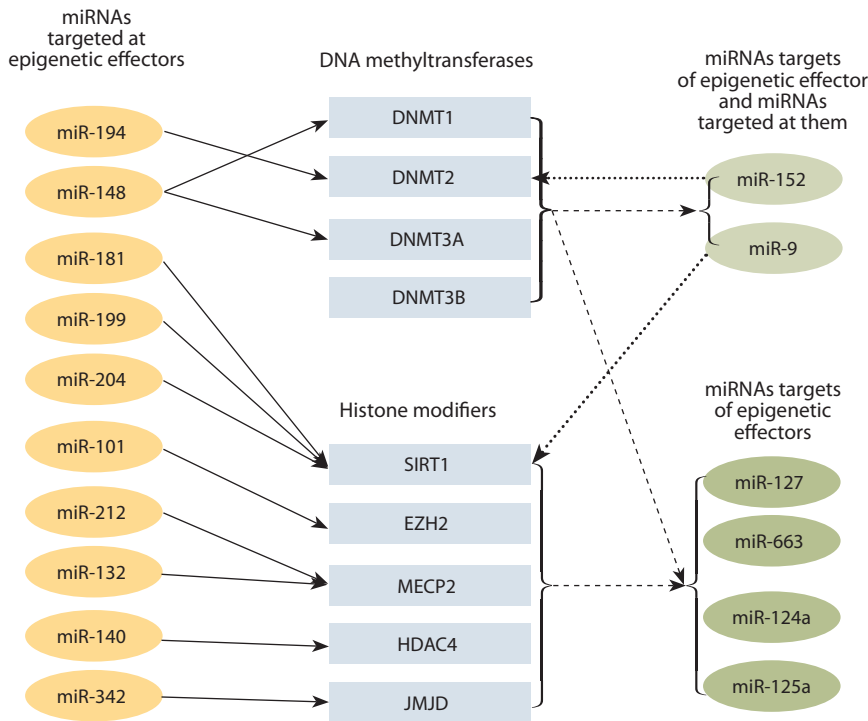
Длинные некодирующие (lncРНК) и малые некодирующие (нкРНК) РНК регулируют активность генов не только посттранскрипционно, но и путем воздействия на модификацию гистонов и целевое метилирование ДНК (Zhang et al., 2015). Показано, что siРНК, происходящие от ТЕ, могут способствовать метилированию ДНК и де-

метилованию лизина 9 в гистоне H3, изменяя активность ТЕ и экспрессию генов (Xu et al., 2013; Zhang et al., 2016). Помимо siРНК, микроРНК также могут эпигенетически модулировать множество генов путем контролирования уровней первичных эпигенетических регуляторов – ДНК-метилтрансферазы (DNMT) и гистоновой деацетилазы (HDAC). Например, мишенью *miR-320* служит метил-СрG-связывающий белок-2 (MECP2), усиливающий клеточную пролиферацию. Взаимосвязь некоторых микроРНК с DNMT и HDAC представлена на рисунке (Samantarrai et al., 2013). Члены семейства *miR-29* воздействуют на ДНК-метилтрансферазы DNMT3A и DNMT3B, подавляя метилирование и туморогенез путем защиты от *de novo* метилирования. *miR-29* вовлечены также в деметилирование ДНК, влияя на метилцитозин диоксидазу 1 (TET1) и тимин-ДНК-гликозилазу (TDG) (Morita et al., 2013). Связывание некоторых микроРНК, таких как *miR-155-5p*, с DNMT1 ведет к ингибированию ее ферментативной активности (Zhang et al., 2015).

Доказано, что нкРНК регулируют развитие и дифференциацию стволовых клеток, воздействуя на образование органов и тканей многоклеточного организма путем регуляции экспрессии белок-кодирующих генов (Ong et al., 2015). Видоспецифическая экспансия специфических малых нкРНК оказывает выраженное влияние на онтогенез; кластеры малых нкРНК играют фундаментальную роль в развитии эмбриональных клеток, при этом регуляция эффективности транскрипции и стабильности микроРНК происходит тканеспецифично (Du et al., 2013). У дрозофилы выявлена микроРНК *miR-130a*, оказывающая воздействие на предопределенный размер органов (Shen et al., 2015). Обнаружена также микроРНК *miR-25*, способная обратно репрограммировать зрелые фибробласты мышей и человека в полипотентные стволовые клетки (Lu et al., 2012). Показано изменение экспрессии определенных малых нкРНК видоспецифически и тканеспецифически в различные стадии онтогенетического развития. Например, в коре головного мозга человека с возрастом усиливается активность микроРНК *miR-33b*, *miR-34*, *miR-181*, *miR-1271*. У мышей при старении в головном мозге возрастает экспрессия *miR-22*, *miR-101a*, *miR-720*, *miR-721*, в ткани печени – *miR-29*, *miR-30d*, *miR-34a* – и другие тканеспецифические особенности экспрессии микроРНК (Dimmeler, Nicotera, 2013). Имеется ряд работ, свидетельствующих как о закономерных перемещениях ТЕ в раннем эмбриональном развитии многоклеточных животных *in vivo*, так и об их активации в разные периоды онтогенеза (Ostertag et al., 2002; Prak et al., 2003; Muotri et al., 2005; Van den Hurk et al., 2007; Coufal et al., 2009; Macia et al., 2011; Marchetto et al., 2013). Учитывая важную роль ТЕ в качестве источников малых нкРНК, можно предположить, что сложившиеся в эволюции видоспецифические особенности активации транспозонов являются основой для эпигенетического регулирования геномов делящихся клеток, способствуя их дифференцировке с развитием целостного организма.

Транспозоны как источники некодирующих РНК

Транспозоны у многоклеточных эукариот составляют значительную долю геномов и служат причиной их



Crosslinking network of miRNAs and epigenetic effectors according to Samantarrai et al. (2013).

крупных размеров, значительно превышающих таковые у прокариот. Наименьший размер эукариотического генома определен у *Encephalitozoon intestinalis* (2.3 мегабазы), который в 70000 раз меньше размера генома другого эукариотического организма – *Paris japonica* (150000 мегабаз) (Elliott, Gregory, 2015). У прокариот размер наиболее крупного генома у *Mycococcus xanthus* (9.5 мегабазы) превышает размер наименьшего генома *Micoplasma genitalium* (0.58 мегабазы) всего в 16 раз (Патрушев, Минкевич, 2007). ТЕ занимают большую часть геномов растений – у некоторых представителей до 90 % всей ДНК. У животных ТЕ также распространены, например у мыши ТЕ составляют 40 %, у дрозофилы – 15–22, у нематоды – 12, у кур – 8.6 % генома (Yuan et al., 2011). У растений наиболее значимая фракция ТЕ геномов представлена LTR-содержащими ТЕ и ассоциируется с увеличением размеров геномов (например, у *Zea mays* за счет LTR размер генома увеличен в два раза за три миллиона лет). LTR-содержащие ТЕ занимают более 58 % всего генома *Allium cepa*, более 76 – *Hordeum vulgare* и более 91 % – *Asparagus officinalis* (Kubiak, Makalowska, 2017). Предполагается, что ТЕ занимают значительную долю в генезе последовательностей геномов, однако вследствие их высокой мутабельности распознать многие участки ДНК как результат инсерции мобильных элементов часто не удается. Например, в работе (de Koning et al., 2011) при анализе генома человека с помощью олигонуклеотидов, узнающих фрагменты ТЕ, накопленные за сотни миллионов лет эволюции, обнаружено, что ТЕ составляют более 60 % генома, тогда как, согласно современным методам исследования, ТЕ обнаруживают лишь в 45 % генома человека.

У эукариот ТЕ – важная движущая сила эволюции, участвующая в построении межгенных участков и интронов. Последовательности ТЕ-происхождения выявляются в экзонах белок-кодирующих генов (Yuan et al., 2011). Кроме того, имеется ряд примеров использования геномами «хозяев» последовательностей ТЕ для участия в важных этапах онтогенеза. Например, гены RAG, полученные от древнего ТЕ, применяются для V(D)J рекомбинации (Lescale, Deriano, 2016). Белки HDP1 и HDP2, участвующие в активном деметилировании ДНК путем ацетилтрансферазного воздействия на гисто-

ны, одомашнены из транспозазы и ДНК-связывающего белка, кодируемых транспозоном суперсемейства *Harbinger* (Duan et al., 2017). Белки Env оболочки эндогенных ретровирусов одомашнены для формирования объединенного клеточного слоя плаценты на поверхности матки (Dupressoir et al., 2012).

ТЕ служат важнейшим источником малых интерферирующих РНК (siРНК) у животных и растений. Более того, siРНК ТЕ-происхождения воздействуют на РНК-направленное метилирование ДНК (RdDM), вызывая сайленсинг не только ТЕ, но и белок-кодирующих генов, в том числе транскрипционных факторов, оказывая важное влияние на онтогенетическое развитие (Shabalina, Koonin, 2008; Zhang et al., 2016). Основными источниками piРНК, например, у насекомых, также оказались транспозоны (Biruykova, Ye, 2015). Происхождение некоторых микроРНК от геномных повторов в смысловых и антисмысловых направлениях впервые выявлено на *Arabidopsis thaliana* (Llave et al., 2002). У животных первые RdmIRs (repeat-derived miRNAs) выявлены N.R. Smalheiser и V.I. Torvik (2005). Позднее J. Piriyaongsa с коллегами (2007) обнаружили 55 микроРНК ТЕ-происхождения у человека. В настоящее время получено множество доказательств, что микроРНК могут происходить от ТЕ у животных и растений. Например, в работе (Yuan et al., 2011) выявлено 226 RdmIRs в геноме человека, 115 – в геноме резуса, 141 – в геноме мыши. G.M. Borchert с коллегами (2011) сообщили об обнаружении 2392 микроРНК ТЕ-происхождения. 374 микроРНК ТЕ-происхождения описаны у летучей мыши, 128 – у собаки и 124 – у лошади (Platt et al., 2014). J. Gim с коллегами (2014) описали 1900 микроРНК ТЕ-происхождения, а S. Qin с коллегами (2015) – 409 ТЕ-микроРНК.

Таким образом, ТЕ служат важнейшими источниками нкРНК, участвующих в специфической регуляции активности генов, управляя дифференцировкой клеток в зависимости от ткани и стадии онтогенеза. Учитывая регуляторное влияние самих ТЕ *in cis* и *in trans* (Belancio et al., 2010; Kitkumthoron, Mutirangura, 2011; Fi-

natto et al., 2015) на гены и возможность саморегуляции продуктами собственной экспрессии (малые нкРНК), можно предположить, что ТЕ являются материальной основой эпигенетической наследственности – их распределение, состав и количество, специфичное для геномов одного вида, регулируют поэтапную активацию и сайленсинг специфических генов, способствующих дифференцировке клеток в зависимости от ткани, стадии развития и пространственного расположения, отражаясь на фенотипических особенностях. В пользу данного предположения говорит наследуемая активация ТЕ в онтогенезе и даже их перемещение на ранних этапах эмбрионального развития. Вероятно, что кооперация ТЕ с универсальными системами процессинга малых нкРНК и регуляции сплайсинга стала причиной возникновения многоклеточных эукариот в эволюции, о чем свидетельствует, например, отсутствие гомологов Droscha и Pasha у одноклеточных эукариот, тогда как они обнаруживаются даже у наиболее примитивных многоклеточных, например у типа *Cnidaria* (Moran et al., 2013), у представителей которых выявляются также консервативные для многих многоклеточных животных микроРНК *miR-100*, *miR-2022*, *miR-2023*, *miR-2030*, *miR-2036* (Liew et al., 2014). У прокариот нет системы, гомологичной эукариотической системе РНК-интерференции (RNAi), но они обладают независимо развитыми аналогичными механизмами защиты от экспрессии ТЕ (Shabalina, Koonin, 2008).

Изменение активности транспозонов в онтогенезе

Доказано, что ТЕ влияют на экспрессию близлежащих генов на транскрипционном и посттранскрипционном уровнях. Многие промоторы и сигналы полиадезилации генов человека и мыши произошли от ТЕ – их инсерции способствовали созданию линии дифференцировки специфических паттернов экспрессии генов. ТЕ, приобретающие регуляторную функцию, образуют кластеры вокруг генов, вовлеченных в развитие и транскрипционную регуляцию. Около 25 % промоторов в геноме человека содержат последовательности транспозонов. ТЕ создают сырье, из которого *cis*-регуляторные элементы эволюционируют *de novo* при точечных мутациях. Кроме того, *cis*-элементы предрасположены в последовательностях ТЕ (Feschotte, 2008). L1-элементы регулируют экспрессию генов *in cis* в соответствии с транскрипционной активностью промотора L1, 5'UTR которого контролируется метилированием и транскрибируется в прямом и обратном направлениях. Гипометилированные L1 могут контролировать экспрессию гена *in trans*, например в некоторых раковых клетках ингибирование обратной транскриптазы L1 может изменять экспрессию множества генов (Kitkumthoron, Mutirangura, 2011). В *Alu*-элементах появляются *cis*-регуляторные сигналы после их инсерции при аккумуляции мутаций, обретая способность воздействовать на многие гены. Например, у человека под регуляторным влиянием *Alu*-ассоциированных CpG-областей находится около 1000 генов (Belancio et al., 2010). LTR-элементы также регулируют активацию белок-кодирующих генов *in cis* (Finatto et al., 2015). LTR содержат сильные промоторы и энхансеры транскрипции,

а также регуляторные последовательности для связывания с транскрипционными факторами, включая рецепторы стероидов и факторы процессинга. При инсерции LTR могут выполнять роль альтернативных промоторов и участвовать в регуляции сплайсинга РНК (Киселев, 2013).

Многочисленные экспериментальные исследования доказали закономерные наследуемые транспозиции различных ТЕ (LINE, *Alu*, SVA) в раннем эмбриогенезе мышей и человека *in vivo*, в зависимости от характера дифференцировки и вида ткани (Ostertag et al., 2002; Prak et al., 2003; Muotri et al., 2005; Van den Hurk et al., 2007; Coufal et al., 2009; Macia et al., 2011; Marchetto et al., 2013), что говорит о возможной роли данных явлений в качестве регулятора экспрессии генов. Сам характер транспозиций ТЕ, начиная с первого деления зиготы, с дальнейшей дифференцировкой вновь образованных клеток может быть унаследован видоспецифически и зависеть от количества, состава и распределения ТЕ в геноме. Аккумуляция инсерций в эмбриональных стволовых клетках человека доказано в экспериментах *in vitro*. Данные перемещения способствовали сайленсингу одних генов и активации других, что оказывало регуляторное воздействие для специфической дифференцировки клеток для выполнения определенных функций (Garcia-Perez et al., 2007). Помимо эмбрионального развития, ТЕ сохраняют высокий потенциал транспозиций в полипотентных стволовых клетках, что необходимо для их трансформации с дифференцировкой за счет избирательной регуляции специфических генов (Wissing et al., 2012; Klawitter et al., 2016). Более того, перемещения ТЕ происходят в различных клетках организма на протяжении всего онтогенеза, при этом вариации в одном организме формируют генетически различающиеся соматические клетки с геномным мозаицизмом в разных тканях. Доказано, что закономерные перемещения ТЕ имеют важное значение в нейрогенезе человека (Faulkner, 2011) и грызунов (Richardson et al., 2014). Выявлено, что у исследуемых линий мышей геномы клеток печени крупнее, чем в других органах. Данные изменения связаны с перемещениями ТЕ со стадийспецифическими особенностями с пиком активности в 5 нед (Lee et al., 2012). Тканеспецифическое изменение размеров геномов под влиянием поэтапных наследуемых перемещений ТЕ выявлено, помимо печени, в клетках сердца исследуемых мышей в возрасте 29 нед, а также в коже и головном мозге мышей в возрасте 6 нед. При этом показано, что данные закономерные перемещения ТЕ способствовали специфическим преобразованиям в функционировании генома, необходимым для дифференцировки клеток при формировании органов и тканей (Lee et al., 2015). Многочисленные исследования доказали также характерную для вида наследуемую активацию определенных ТЕ в онтогенезе, что говорит об их возможном значении в регуляции экспрессии генов, необходимой для дифференцировки клеток при формировании органов и тканей (Prak et al., 2003; Van den Hurk et al., 2007; Coufal et al., 2009; Macia et al., 2011; Guo et al., 2016; Klawitter et al., 2016; Zakrzewski et al., 2017). Учитывая важнейшую роль ТЕ в генезе некодирующих РНК, регулирующих экспрессию генов, можно предположить, что ТЕ, являясь материальной основой эпигенетической наследственно-

сти, эволюционно запрограммированы на уровне вида на последовательный каскад транспозиций и/или их активаций в каждом клеточном делении. Данный закономерный цикл активации мобильных генетических элементов ведет к формированию целостного организма. Тканеспецифическая динамическая экспрессия РНК LINE-элементов с их мозаичным распределением представляет собой стационарный геномный поток при развитии млекопитающих. Показана выраженная экспрессия РНК P1-LINE в различных тканях и органах крыс с тканеспецифическим и динамическим паттерном в виде длинных некодирующих РНК (lncРНК) и малых нкРНК (Singh, Rath, 2012). Транспозоны – важный источник lncРНК, они формируют основу целевого распада мРНК через короткие несовершенные спаривания нуклеотидов. У человека тысячи lncРНК связаны с LTR-элементами, обеспечивающими их регуляторными сигналами (Hadjiarugou, Delihis, 2013). Около 41 % экзонов генов lncРНК непосредственно образованы из ТЕ, а 83 % экзонов lncРНК имеют предположительное ТЕ-происхождение в связи с содержанием фрагментов их последовательностей (Johnson, Guigo, 2014). В пользу регуляторной роли закономерных наследуемых транспозиций и активации ТЕ в онтогенезе свидетельствуют высокие тканеспецифические особенности экспрессии lncРНК, влияющие на характер дифференцировки тканей (Ramsay et al., 2017).

Взаимосвязь сплайсинга с транспозонами и некодирующими РНК

В возникновении интронов важную роль играли ТЕ, способствуя их распространению в геномах (Lee, Stevens, 2016). Показано, например, что короткие неавтономные ТЕ независимо генерировали тысячи интронов у зеленых и бурых водорослей. Каждый ТЕ содержит один сайт сплайсинга, другой сайт кооптируется из последовательности гена, дублированной при инсерции ТЕ, способствуя идеальному сплайсингу (Huff et al., 2016). Взаимосвязь интронов с ТЕ существует даже у прокариот. У бактерий обнаружены химерные элементы IStrons, состоящие из интронов группы I и IS-элементов. В данных структурах интроны группы I, сами являющиеся мобильными генетическими элементами, обеспечивают IStron самосплайсирующей способностью. При этом сплайсинг возможен по альтернативным путям (Tougas et al., 2014). У животных гены микроРНК расположены в основном внутри интронов белок-кодирующих генов, зачастую формируя кластеры с транскрипцией в единый первичный транскрипт. В образовании микроРНК у многоклеточных животных, помимо Drosha-опосредованного механизма, применяется альтернативный путь с участием сплайсинга подходящих РНК, транскрибированных из интронов (миРтронов), которые имитируют структурные особенности пре-микроРНК (Shabalina, Koonin, 2008). Учитывая важную роль ТЕ как источников микроРНК и их участие в распространении интронов, можно предположить, что в регуляции сплайсинга большое значение могут иметь специфические особенности активации ТЕ в онтогенезе, выработанные эволюционно на уровне вида. В пользу данного предположения говорит выявление множества высококонсервативных ТЕ в качестве альтернативных

сплайсинговых экзонов или транскрипционных энхансеров, вводящих преждевременные стоп-кодона и иницирующих нонсенс-опосредованный распад. Введение преждевременных стоп-кодонов и инициация нонсенс-опосредованного распада способствуют гомеостазу мРНК в клетке (Feschotte, 2008). Имеется ряд свидетельств о способности ТЕ регулировать сплайсинг, вызывая тканеспецифическую и стадийспецифическую экспрессию определенных генов при дифференцировке клеток (Feschotte, 2008; Belancio et al., 2010; Dumesic, Madhani, 2013). Последовательности транспозонов содержат акцепторные и донорные сайты сплайсинга, данные сайты могут быть использованы во время транскрипции, что может способствовать изменению экспрессии генов (Belancio et al., 2010). Транскрипты ТЕ могут использоваться в качестве субстратов для продукции малых нкРНК из-за их тенденции фиксироваться в сплайсосомах. Имеется общая значимость объединенного с сплайсосомой биогенеза малых нкРНК для других систем регуляции работы генома. В качестве субстратов для биогенеза siРНК используются не полностью сплайсированные прекурсоры мРНК, и транскрипты ТЕ-происхождения считаются триггерами для РНК-сайленсинга в процессе сплайсинга в ядре (Dumesic, Madhani, 2013). Обнаружено также, что сплайсинг интронов способен подавлять систему RNAi, в частности, у *Arabidopsis* библиотеки малых нкРНК значительно обогащены для безинтронных генов. Экспериментально доказано, что внедрение интрона в трансген понижает РНК-сайленсинг в четыре раза (Christie et al., 2011).

Основная часть первичных транскриптов многоклеточных содержит более одного интрона и способна к альтернативному сплайсингу, в результате которого образуются различные мРНК из одного гена, что используется в дифференцировке клеток (Ulrich, Wahl, 2017). Поэтому количество генов, способных к альтернативному сплайсингу, возрастает при эволюционной диверсификации типов клеток в связи с усложнением программ развития (Bush et al., 2017). Помимо специфической продукции малых нкРНК, в тканеспецифической дифференцировке огромное значение имеет регуляция альтернативного сплайсинга, в которой важную роль также играют ТЕ. Традиционно считается, что альтернативный сплайсинг регулируется сплайсинговыми энхансерами и сайленсерами. Данные короткие консервативные последовательности РНК, обычно длиной 10 нуклеотидов, локализуются в экзонах или интронах и действуют либо изолированно, либо в виде кластеров. Функция указанных РНК заключается в стимуляции или ингибировании использования сайтов путем стимуляции или ингибирования использования сайтов сплайсинга посредством специфического связывания с такими регуляторными белками, как SR или малые ядерные рибонуклеопротеины (snRNP) (Pastor et al., 2009; Lucio et al., 2011). Интересно, что источниками сплайсинговых энхансеров и сайленсеров также могут служить ТЕ (Lei, Vorechovsky, 2005). Например, выявлено, что из последовательности *Alu* экспрессируются сплайсинговые энхансеры, действующие за счет привлечения U1 snRNP (Pastor et al., 2009). Показано также, что подмножество SR-белков активирует криптический 3'-сайт сплайсинга в смысловом повторе *Alu*, локализованном в интроне 4 гена

LST1 человека. Использование данного криптического сайта сплайсинга контролируется сопоставляющимися сплайсинговыми сайленсерами и энхансерами, происходящими от *Alu* (Lei, Vorechovsky, 2005). На основании этого можно предположить, что малые нкРНК также могут выступать в качестве энхансеров и сайленсеров сплайсинга. В связи с наличием большого количества доказательств важного значения ТЕ в качестве источника малых нкРНК, выработанная в эволюции последовательность онтогенетической активации ТЕ может быть основой важнейших путей геномных преобразований при дифференцировке клеток (Borchert et al., 2011; Yuan et al., 2011; Xu et al., 2013; Gim et al., 2014; Platt et al., 2014; Biryukova, Ye, 2015; Qin et al., 2015; Zhang et al., 2016).

Заключение

Получены многочисленные экспериментальные данные, позволяющие предположить, что в основе возникновения и эволюции многоклеточных эукариот ключевую роль играли универсальные системы процессинга малых нкРНК и регуляции сплайсинга. Данные универсальные механизмы имеют непосредственную связь с ТЕ (Lei, Vorechovsky, 2005; Feschotte, 2008; Shabalina, Koonin, 2008; Pastor et al., 2009; Belancio et al., 2010; Borchert et al., 2011; Yuan et al., 2011; Singh, Rath, 2012; Dumesic, Madhani, 2013; Hadjiargyrou, Delihis, 2013; Xu et al., 2013; Gim et al., 2014; Johnson, Guigo, 2014; Platt et al., 2014; Biryukova, Ye, 2015; Qin et al., 2015; Huff et al., 2016; Lee, Stevens, 2016; Zhang et al., 2016) и первоначально были направлены на защиту генома хозяев от чужеродных генетических последовательностей и мобильных генетических элементов (Shabalina, Koonin, 2008). Однако кооперация универсальных систем процессинга малых нкРНК и регуляции сплайсинга, помимо защиты от ТЕ с возможностью дифференцированного регулирования функциями генов, способствовала координации работы клеток в многоклеточных организмах. Большое количество инсерций различных ТЕ в геномах эукариот в эволюции способствовало усовершенствованию взаимосвязи универсальных систем регуляции онтогенеза с работой различных генов, что содействовало последовательной регуляции экспрессии определенных генов с фенотипическими проявлениями в виде формирования органов, тканей и целостного организма. ТЕ обладают способностью саморегуляции с чувствительностью к стрессу (Zhou, Kishima, 2017) и хозяин-опосредованным механизмам контроля экспрессии (Pizarro, Cristofari, 2016). ТЕ-опосредованный контроль экспрессии осуществляется благодаря образованию усеченных супрессорных копий для транспозаз-опосредованной ауторегуляции, использованию собственных антисмысловых промоторов, регуляторных факторов хозяина и образованию нкРНК из транспозонных последовательностей (Shabalina, Koonin, 2008; Borchert et al., 2011; Yuan et al., 2011; Singh, Rath, 2012; Hadjiargyrou, Delihis, 2013; Xu et al., 2013; Gim et al., 2014; Johnson, Guigo, 2014; Platt et al., 2014; Biryukova, Ye, 2015; Qin et al., 2015; Zhang et al., 2016). Хозяин-опосредованный контроль связан с воздействием эпигенетических факторов, таких как модификации гистонов, регуляторами которых являются сами ТЕ (Miousse et al., 2015), а также нкРНК

ТЕ-происхождения (Shabalina, Koonin, 2008; Borchert et al., 2011; Yuan et al., 2011; Singh, Rath, 2012; Hadjiargyrou, Delihis, 2013; Xu et al., 2013; Gim et al., 2014; Johnson, Guigo, 2014; Platt et al., 2014; Biryukova, Ye, 2015; Qin et al., 2015; Zhang et al., 2016). На возможность формирования саморегуляторной системы ТЕ, способствующей росту и развитию организма, указывает чувствительность ТЕ к воздействию стероидных гормонов (Киселев, 2013). Учитывая важную роль ТЕ в формировании геномов, генезе некодирующих РНК, влияющих на метилирование ДНК и модификации гистонов (Upadhyay et al., 2017), можно предположить, что видоспецифические особенности расположения, количества и состава ТЕ служат основой эпигенетической наследственности.

Conflict of interest

The authors declare no conflict of interest.

References

- Baranov V.S., Kuznetsova T.V. Tsitogenetika embrionalnogo razvitiya cheloveka [Cytogenetics of Human Embryonic Development]. St.-Petersburg, 2007. (in Russian)
- Belancio V.P., Roy-Engel A.M., Deininger P.L. All y'all need to know 'bout retroelements in cancer. *Semin. Cancer Biol.* 2010;20(4): 200-210.
- Biryukova I., Ye T. Endogenous siRNA and piRNAs derived from transposable elements and genes in the malaria vector mosquito *Anopheles gambiae*. *BMC Genomics.* 2015;16:278.
- Borchert G.M., Holton N.W., Williams J.D., Hernan W.L., Bishop I.P., Dembosky J.A., Elste J.E., Gregoire N.S., Kim J.A., Koehler W.W., Lengerich J.C., Medema A.A., Nguyen M.A., Ower G.D., Rarick M.A., Strong B.N., Tardi N.J., Tasker N.M., Wozniak D.J., Gatto C., Larson E.D. Comprehensive analysis of microRNA genomic loci identifies pervasive repetitive-element origins. *Mob. Genet. Elements.* 2011;1(1):8-17.
- Bush S.J., Chen L., Tovar-Corona J.M., Urrutia A.O. Alternative splicing and the evolution of phenotypic novelty. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 2017;372(1713):pii:20150474.
- Christie M., Croft L.J., Carroll B.J. Intron splicing suppresses RNA silencing in *Arabidopsis*. *Plant J.* 2011;68(1):159-167.
- Coufal N.G., Garcia-Perez J.L., Peng G.E., Yeo G.W., Mu Y., Lovci M.T., Morell M., O'Shea K.S., Moran J.V., Gage F.H. L1 retrotransposition in human neural progenitor cells. *Nature.* 2009; 460(7259):1127-1231.
- de Koning A.P., Gu W., Castoe T.A., Batzer M.A., Pollock D.D. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet.* 2011;7(12):e1002384.
- Dimmeler S., Nicotera P. MicroRNAs in age-related diseases. *EMBO Mol. Med.* 2013;5(2):180-190.
- Du Z., Yang C., Rothschild M.F., Ross J. Novel microRNA families expanded in the human genome. *BMC Genomics.* 2013;14:98-105.
- Duan C.G., Wang X., Xie S., Pan L., Miki D., Tang K., Hsu C.C., Lei M., Zhong Y., Hou Y.J., Wang Z., Zhang Z., Mangrauthia S.K., Xu H., Zhang H., Dilkes B., Tao W.A., Zhu J.K. A pair of transposon-derived proteins function in a histone acetyltransferase complex for active DNA demethylation. *Cell Res.* 2017;27(2):226-240.
- Dumesic P.A., Madhani H.D. The spliceosome as a transposon sensor. *RNA Biol.* 2013;10(11):1653-1660.
- Dupressoir A., Lavialle C., Heidmann T. From ancestral infectious retroviruses to bona fide cellular genes: role of the captured syncytins in placentation. *Placenta.* 2012;33(9):663-671.
- Elliott T.A., Gregory T.R. Do larger genomes contain more diverse transposable elements? *BMC Evol. Biol.* 2015;15(1):69-81.
- Faulkner G.J. Retrotransposons: mobile and mutagenic from conception to death. *FEBS Lett.* 2011;585(11):1589-1594.

- Feschotte C. The contribution of transposable elements to the evolution of regulatory networks. *Nat. Rev. Genet.* 2008;9(5):397-405.
- Finatto T., de Oliveira A., Chaparro C., da Maia L.C., Farias D.R., Woyann L.G., Mistura C.C., Soares-Bresolin A.P., Llauro C., Panaud O., Picault N. Abiotic stress and genome dynamics: specific genes and transposable elements response to iron excess in rice. *Rice.* 2015; 8(13). DOI 10.1186/s12284-015-0045-6.
- Garcia-Perez J.L., Marchetto M.C., Muotri A.R., Coufal N.G., Gage F.H., O'Shea K.S., Moran J.V. LINE-1 retrotransposition in human embryonic stem cells. *Hum. Mol. Genet.* 2007;16(13):1569-1577.
- Gim J., Ha H., Ahn K., Kim D.S., Kim H.S. Genome-wide identification and classification of microRNAs derived from repetitive elements. *Genomics Inform.* 2014;12(4):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;6:32207-32219.
- Hadjjargyrou M., Delihans N. The Intertwining of transposable elements and non-coding RNAs. *Int. J. Mol. Sci.* 2013;14(7):13307-13328.
- Huff J.T., Zilberman D., Roy S.W. Mechanism for DNA transposons to generate introns on genomic scales. *Nature.* 2016;538(7626):533-536.
- Johnson R., Guigo R. The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs. *RNA.* 2014;20(7):959-976.
- Kiselev O.I. Endogenous retroviruses: structure and functions in the human genome. *Voprosy virusologii = Problems of Virology.* 2013; 1:102-115. (in Russian)
- Kitkumthorn N., Mutirangura A. Long interspersed nuclear element-1 hypomethylation in cancer: biology and clinical applications. *Clin. Epigenet.* 2011;2:315-330.
- Klawitter S., Fuchs N.V., Upton K.R., Munoz-Lopez M., Shukla R., Wang J., Garcia-Canadas M., Lopez-Ruiz C., Gerhardt D.J., Sebe A., Grabundajia I., Merkert S., Gerdes P., Pulgarin J.A., Bock A., Held U., Witthuhn A., Haase A., Sarkadi B., Lower J., Wolvetang E.J., Martin U., Ivics Z., Izsvak Z., Garcia-Perez J.L., Faulkner G.J., Schumann G.G. Reprogramming triggers endogenous L1 and Alu retrotransposition in human induced pluripotent stem cells. *Nat. Commun.* 2016;7:10286-10301.
- Kubiak M.R., Makalowska I. Protein-coding genes' retrocopies and their functions. *Viruses.* 2017;9(4):pii:E80.
- Lee K.H., Chiu S., Lee Y.K., Greenhalgh D.G., Cho K. Age-dependent and tissue-specific structural changes in the C57BL/6J mouse genome. *Exp. Mol. Pathol.* 2012;93(1):167-172.
- Lee K.H., Yee L., Lim D., Greenhalgh D., Cho K. Temporal and spatial rearrangements of a repetitive element array on C57BL/6J mouse genome. *Exp. Mol. Pathol.* 2015;98(3):439-445.
- Lee S., Stevens S.W. Spliceosomal intronogenesis. *Proc. Natl. Acad. Sci. USA.* 2016;113(23):6514-6519.
- Lei H., Vorechovsky I. Identification of splicing silencers and enhancers in sense Alus: a role for pseudoacceptors in splice site repression. *Mol. Cell. Biol.* 2005;25(16):6912-6920.
- Lescale C., Deriano L. The RAG recombinase: Beyond breaking. *Mech. Ageing Dev.* 2016;16:30263-30269. DOI 10.1016/j.mad.2016.11.003.
- Liew Y.J., Aranda M., Carr A., Baumgarten S., Zoccola D., Tambutte S., Allemant D., Micklem G., Voolstra C.R. Identification of microRNA in the coral *Styphora pistillata*. *PLoS One.* 2014;9(3):e91101.
- Llave C., Kasschau K.D., Rector M.A., Carrington J.C. Endogenous and silencing-associated small RNAs in plants. 2002;14(7):1605-1619.
- Lu D., Davis M.P., Abreu-Goodger C., Wang W., Campos L.S., Siede J., Vigorito E., Skarnes W.C., Dunham I., Enright A.J., Liu P. MiR-25 regulates Wwp2 and Fbxw7 and promotes reprogramming of mouse fibroblast cells to iPSCs. *PLoS One.* 2012;7(8):e40938.
- Luco R.F., Allo M., Schor I.E., Kornblitt A.R., Misteli T. Epigenetics in alternative pre-mRNA splicing. *Cell.* 2011;144(1):16-26.
- Macia A., Munoz-Lopez M., Cortes J.L., Hastings R.K., Morell S., Lucena-Aguilar G., Marchal J.A., Badge R.M., Garcia-Perez J.L. Epigenetic control of retrotransposons expression in human embryonic stem cells. *Mol. Cell. Biol.* 2011;31(2):300-316.
- Marchetto M.C., Narvaiza I., Denli A.M., Benner C., Lazzarini T.A., Nathanson J.L., Paguola A.C., Desai K.N., Herai R.H., Weitzman M.D., Yeo G.W., Muotri A.R., Gage F.H. Differential L1 regulation in pluripotent stem cells of humans and apes. *Nature.* 2013; 503(7477):525-529.
- Miousse I.R., Chalbot M.G., Lumen A., Ferguson A., Kavouras I.G., Koturbash I. Response of transposable elements to environmental stressors. *Mutat. Res. Rev. Mutat. Res.* 2015;765:19-39.
- Moran Y., Praher D., Fredman D., Technau U. The evolution of microRNA pathway protein components in Cnidaria. *Mol. Biol. Evol.* 2013;30(12):2541-2552.
- Morita S., Horii T., Kimura M., Ochiya T., Tajima S., Hatada I. miR-29 represses the activities of DNA methyltransferases and DNA demethylases. *Int. J. Mol. Sci.* 2013;14:14647-14658.
- Muotri A.R., Chu V.T., Marchetto M.C., Deng W., Moran J.V., Gage F.H. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature.* 2005;435(7044):903-910.
- Ong S., Lee W.H., Kodo K., Wu J.C. MicroRNA-mediated regulation of differentiation and trans-differentiation in stem cells. *Adv. Drug Deliv. Rev.* 2015;88:3-15.
- Ostertag E.M., De Berardinis R.J., Goodier J.L., Zhang Y., Yang N., Gerton G.L., Kazazian H.H., Jr. A mouse model of human L1 retrotransposition. *Nat. Genet.* 2002;32(4):655-660.
- Pastor T., Talotti G., Lewandowska M.A., Pagani F. An Alu-derived intronic splicing enhancer facilitates intronic processing and modulates aberrant splicing in ATM. *Nucleic Acids Res.* 2009;37(21):7258-7267.
- Patrushev L.I., Minkevich I.G. The problem of the eukaryotic genome size. *Uspekhi biologicheskoi khimii = Advances of Biological Chemistry.* 2007;47:293-370. (in Russian)
- Piriyaopongsa J., Marino-Ramirez L., Jordan I.K. Origin and evolution of human microRNAs from transposable elements. *Genetics.* 2007; 176(2):1323-1337.
- Pizarro J.G., Cristofari G. Post-transcriptional control of LINE-1 retrotransposition by cellular host factors in somatic cells. *Front. Cell. Dev. Biol.* 2016;4:14-23.
- Platt R.N., Vandeweque M.W., Kern C., Schmidt C.J., Hoffmann F.G., Ray D.A. Large number of novel miRNAs originate from DNA transposons and are coincident with a large species radiation in bats. *Mol. Biol. Evol.* 2014;31(6):1536-1545.
- Prak E.T., Dodson A.W., Farkash E.A., Kazazian H.H. Jr. Tracking an embryonic L1 retrotransposition event. *Proc. Natl. Acad. Sci. USA.* 2003;100(4):1832-1837.
- Qin S., Jin P., Zhou X., Chen L., Ma F. The role of transposable elements in the origin and evolution of microRNAs in human. *PLoS One.* 2015;10(6):e0131365.
- Ramsay M., Marchetto M.C., Caron M., Chen S.H., Busche S., Kwan T., Pastinen T., Gage F.H., Bourque G. Conserved expression of transposon-derived non-coding transcripts in primate stem cells. *BMC Genomics.* 2017;18(1):214-226.
- Richardson S.R., Morell S., Faulkner G.J. L1 retrotransposons and somatic mosaicism in the brain. *Annu. Rev. Genet.* 2014;48:1-27.
- Samantarrai D., Dash S., Chhetri B., Mallick B. Genomic and epigenomic cross-talks in the regulatory landscape of miRNAs in breast cancer. *Mol. Cancer Res.* 2013;11(4):315-328.
- Shabalina S.A., Koonin E.V. Origins and evolution of eukaryotic RNA interference. *Trends Ecol. Evol.* 2008;23(10):578-587.
- Shen S., Guo X., Yan H., Lu Y., Ji X., Li L., Liang T., Zhou D., Feng X.H., Zhao J.C., Yu J., Gong X.G., Zhang L., Zhao B. A miR-130a-YAP positive feedback loop promotes organ size and tumorigenesis. *Cell Res.* 2015;25:997-1012.
- Singh D.K., Rath P.C. Long interspersed nuclear elements (LINEs) show tissue-specific, mosaic genome and methylation-unrestricted, widespread expression of noncoding RNAs in somatic tissues of the rat. *RNA Biol.* 2012;9(11):1380-1396.
- Smalheiser N.R., Torvik V.I. Mammalian microRNAs derived from genomic repeats. *Trends Genet.* 2005;21(6):322-326.
- Tourasse N.J., Stabell F.B., Kolsto A.B. Survey of chimeric IStron elements in bacterial genomes: multiple molecular symbioses between

- group I intron ribozymes and DNA transposons. *Nucleic Acids Res.* 2014;42(20):12333-12351.
- Ulrich A.K.C., Wahl M.C. Human MFAP1 is a cryptic ortholog of the *Saccharomyces cerevisiae* Spp381 splicing factor. *BMC Evol. Biol.* 2017;17:91-107.
- Upadhyay U., Srivastava S., Khatri I., Nanda J.S., Subramanian S., Aro-
ra A., Singh J. Ablation of RNA interference and retrotransposons
accompany acquisition and evolution of transposases to hetero-
chromatin protein CENPB. *Mol. Biol. Cell.* 2017;28(8):1132-1146.
- Van den Hurk J.A., Meij I.C., Seleme M.C. L1 retrotransposition can
occur early in human embryonic development. *Hum. Mol. Genet.*
2007;16(13):1587-1592.
- Vanyushin B.F. Epigenetics today and tomorrow. *Vavilovskii Zhurnal
Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding.*
2013;17(4/2):805-832. (in Russian)
- Wissing S., Munoz-Lopez M., Macia A., Yang Z., Montano M., Col-
lins W., Garcia-Perez J.L., Moran J.V., Greene W.C. Reprogram-
ming somatic cells into iPS cell activates LINE-1 retroelement mo-
bility. *Hum. Mol. Genet.* 2012;21(1):208-218.
- Xu C., Tian J., Mo B. siRNA-mediated DNA methylation and H3K9
dimethylation in plants. *Protein Cell.* 2013;4(9):656-663.
- Yuan Z., Sun X., Liu H., Xie J. MicroRNA genes derived from re-
petitive elements and expanded by segmental duplication events in
mammalian genomes. *PLoS One.* 2011;6(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., Terragni J., Dai N., Correa Jr. I.R.,
Pradhan S. Small RNA-mediated DNA (cytosine-5) methyltransfer-
ase 1 inhibition leads to aberrant DNA methylation. *Nucleic Acids
Res.* 2015;43(12):6112-6124.
- Zhang H., Tao Z., Hong H., Chen Z., Wu C., Li X., Xiao J., Wang S.
Transposon-derived small RNA is responsible for modified function
of WRKY45 locus. *Nat. Plants.* 2016;2:16016-16023.
- Zhou H., Kishima Y. Alternative plant host defense against transposon
activities occurs at the post-translational stage. *Plant Signal. Behav.*
2017;e1318238. DOI 10.1080/15592324.2017.