

РЕГУЛЯЦИЯ ТРАНСКРИПЦИИ РЕТРОТРАНСПОЗОНОВ LINE1 МЛЕКОПИТАЮЩИХ

© А. В. Федоров

Институт цитологии РАН, Санкт-Петербург;
электронный адрес: an_tn@mail.ru

Всестороннее изучение геномов высших эукариот, и в частности результаты успешно реализуемых проектов по их секвенированию, убеждают нас в том, что не только уникальные последовательности ДНК, но и многочисленные семейства повторяющихся последовательностей несут существенную функциональную нагрузку, что не позволяет рассматривать повторы лишь как «паразитическую» ДНК. Ретротранспозоны LINE1, занимающие около 1/5 геномов млекопитающих, способны к перемещению и участвуют в ретротранспозиции клеточной ДНК, организации структуры хроматина, геномных перестройках и регуляции транскрипции генов, а также вовлечены в процессы рекомбинации и репарации. В обзоре обсуждаются структурная организация, особенности функционирования и механизмы тканеспецифичной регуляции транскрипции ретротранспозонов LINE1.

Ключевые слова: ретротранспозоны LINE1, регуляция транскрипции, эпигенетика, РНК-сайленсинг.

Принятые сокращения: п. н. — пары нуклеотидов, DAPI — 4'-6'-диамидино-2-фенилиндо́л, LINE1 — длинные диспергированные ядерные элементы, ORF — открытая рамка считывания, UTR — нетранслируемый район.

Большую часть повторяющихся последовательностей ДНК млекопитающих (примерно 40—45 % генома) составляют мобильные элементы — транспозоны и ретротранспозоны (Lander et al., 2001). Широкое распространение мобильных элементов в геномах обусловлено существованием эффективного механизма размножения этих последовательностей. Автономные семейства мобильных элементов кодируют белки, необходимые для перемещения и встраивания в геном их копий. У млекопитающих наиболее распространенным семейством автономных мобильных элементов, содержащим способные к перемещению копии, является семейство LINE1 (long interspersed nuclear element 1, L1). Элементы LINE1 занимают около 20 % геномов млекопитающих, они способны размножаться в геноме с помощью механизма обратной транскрипции и поэтому относятся к ретротранспозонам (Ostertag, Kazazian, 2001).

Элементы LINE1 оказывают существенное влияние на геном, главным образом посредством ретротранспозиции, которая требует экспрессии РНК LINE1 и кодируемых ими белков. Неконтролируемое перемещение элементов LINE1 может приводить к катастрофическим последствиям, таким как дестабилизация генома (Bourc'his, Bestor, 2004) и малигнизация клеток (Yoder et al., 1997). Несмотря на это, ретротранспозиция элементов LINE1 не является абсолютно запрещенной у млекопитающих и действительно наблюдается в определенных типах клеток (Brouha et al., 2002; Ostertag et al., 2002; Prak et al., 2003). Следовательно, существует четкий механизм контроля как активации, так и репрессирования этих повторов *in vivo*. Очевидно, что одним из основных способов конт-

роля ретротранспозиции элементов LINE1 является регуляция их экспрессии. Экспрессия элементов LINE1 тканеспецифична, однако механизмы ее регуляции до сих пор мало изучены.

В настоящем обзоре рассматриваются структура элементов LINE1, их предполагаемые функции, механизмы ретротранспозиции и факторы, влияющие на экспрессию LINE1, в разных типах клеток. Обсуждаются особенности строения промоторных областей LINE1 и роль известных белков, связывающихся с промотором, а также эпигенетических факторов и механизмов РНК-сайленсинга в регуляции тканеспецифичной экспрессии этих мобильных элементов.

Структура повторов LINE1

К настоящему времени с помощью рестрикционного анализа ДНК в геномах эукариот охарактеризовано несколько семейств диспергированных повторяющихся последовательностей. Так, в частности, были выявлены повторы *KpnI* у приматов (Maio et al., 1981) и *BamHI* или *MIF-1* у грызунов (Brown, Dover, 1981; Meunier-Rotival et al., 1982). Структурное сходство этих семейств повторов позволило отнести их к одному типу повторяющихся последовательностей, названных LINEs (long interspersed nuclear elements) (Singer, 1982). Повторы *KpnI* человека и *BamHI* мыши гомологичны (Singer et al., 1983) и, следовательно, являются представителями одного семейства, названного LINE1, или L1 (Burton et al., 1986). В геномах многих других эукариот также присутствуют повторы с

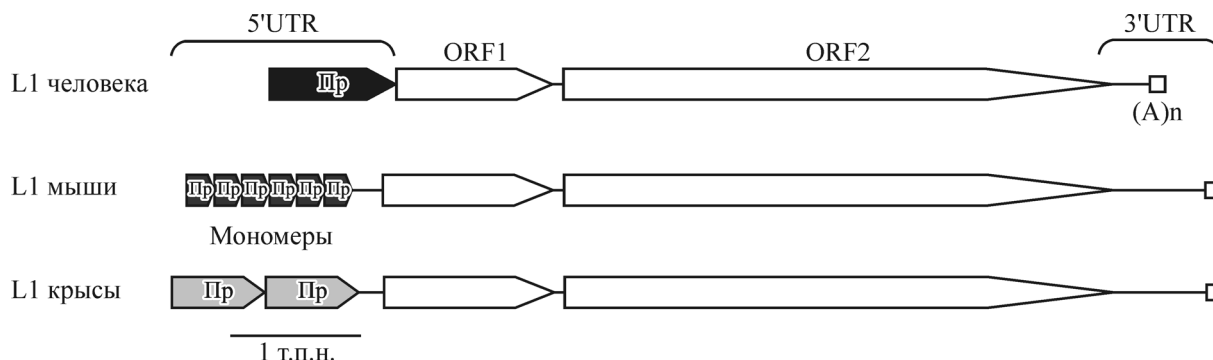


Рис. 1. Структура элементов L1 млекопитающих.

В составе повторов L1 выявлены 5'- и 3'-нетранслируемые районы (5'UTR и 3'UTR, untranslated region), две открытые рамки считывания (ORF1 и ORF2, open reading frames) и участок, состоящий из нескольких аденинов — (A)_n. У грызунов 5'-нетранслируемые районы содержат тандемно повторяющиеся промоторные мономеры. Пр — промоторные последовательности (здесь и далее). По: Furano, 2000, с изменениями.

подобной организацией, первичные последовательности которых, однако, отличаются от первичной последовательности L1 млекопитающих. Такие повторы обнаружены у простейших (Kimmel et al., 1987), насекомых (Fawcett et al., 1986), растений (Noma et al., 2000), амфибий (Pont-Kingdon et al., 1997) и птиц (Vandergon, Reitman, 1994).

Секвенирование повторов L1 человека (Scott et al., 1987) и грызунов (D'Ambrosio et al., 1986; Loeb et al., 1986) позволило определить особенности их строения (рис. 1). Промоторные последовательности, располагающиеся в видоспецифичных 5'-нетранслируемых районах элементов L1, отвечают за транскрипционную активность этих ретротранспозонов (Furano et al., 1988; Swergold, 1990; DeBerardinis, Kazazian, 1999). Интересная черта организации элементов L1 заключается в том, что в отличие от областей, кодирующих белки, 5'-нетранслируемые регуляторные районы ретротранспозонов L1 у разных видов, в частности у человека, мыши и крысы, являются абсолютно негомологичными (Scott et al., 1987; Furano et al., 1988; Padgett et al., 1988).

Современные семейства повторов L1 произошли от одного предшественника, существовавшего у общего предка всех млекопитающих более 100 млн лет назад (Furano, 2000). В секвенированных геномах человека, мыши и крысы обнаружено более 500 000 копий повторов L1 (Lander et al., 2001; Waterston et al., 2002; Gibbs et al., 2004). Однако подавляющее число копий L1 являются неактивными, так как содержат делеции и точечные мутации, лишаящие их способности к дальнейшей ретротранспозиции (Fanning, 1983; Scott et al., 1987). Следовательно, за размножение элементов L1 в геномах млекопитающих отвечает лишь небольшая доля этих повторов.

Предполагаемый механизм ретротранспозиции элементов LINE1

Изучение перемещения ретротранспозонов L1 в экспериментальных клеточных системах позволило получить данные, значительно проясняющие механизм размножения этих повторов в геноме (Dombroski et al., 1994; Feng et al., 1996; Moran et al., 1996). Первым ключевым этапом ретротранспозиции является транскрипция геномных элементов L1 со своих внутренних видоспецифич-

ных промоторов, расположенных в 5'UTR. Синтезированная в результате полноразмерная смысловая РНК L1 является матрицей для построения новой копии ДНК L1 и в то же время кодирует белки ORF1p и ORF2p. Белок ORF1p имеет мол. массу около 40 кДа, способен формировать мультимерные комплексы, связывать одонитивную РНК и, вероятно, участвует в процессе обмена нити во время праймирования обратной транскрипции L1 (Martin et al., 2003). Белок ORF2p имеет мол. массу около 150 кДа, обладает активностями эндонуклеазы (Feng et al., 1996; Moran et al., 1996) и обратной транскриптазы (Mathias et al., 1991). Установлено, что белки ORF1p и ORF2p, формирующие вместе с РНК L1 РНП-частицы, необходимы для ретротранспозиции L1 (Mathias et al., 1991; Feng et al., 1996; Moran et al., 1996; Kulpa, Moran, 2006), а также способствуют ретротранспозиции и другой клеточной ДНК, в том числе ДНК неавтономных ретротранспозонов (Dewannieux et al., 2003). Интеграция элементов L1 в геном происходит с помощью механизма TPRT (target primed reverse transcription), т. е. обратной транскрипции, праймированной в сайте интеграции (Luan et al., 1993). Для синтеза новой копии ДНК L1 по матрице РНК в качестве праймера используется одонитивной участок геномной ДНК с 3'-ОН-концом, образующийся после расщепления эндонуклеазой одной нити ДНК сайта интеграции. Встроенные по такому механизму элементы L1 фланкированы короткими прямыми повторами длиной 7—20 п. н., которые представляют собой участки дупликации сайта интеграции.

Влияние элементов L1 на геном млекопитающих

К настоящему времени выявлено множество примеров участия элементов L1 в различных клеточных процессах. Повторы L1 составляют значительную часть генома и уже благодаря этому могут оказывать на него значительное влияние. Копии элементов L1 представляют собой гомологичные последовательности ДНК, которые могут участвовать в рекомбинации. Так, описаны примеры рекомбинации между элементами L1, которая приводила к появлению новых подсемейств L1 (Hayward et al., 1997; Saxton, Martin, 1998) или дупликации генов (Fitch et al., 1991). Кроме этого, неравная гомологичная рекомбинация между элементами L1 может порождать крупные ге-

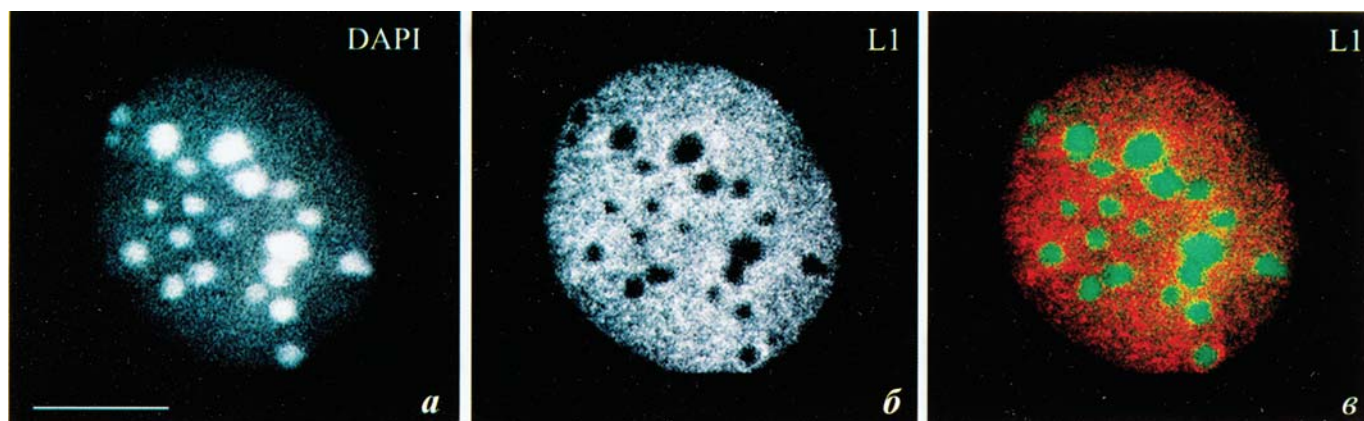


Рис. 2. Расположение элементов L1 в ядрах клеток линии L929.

a — интерфазное ядро окрашено DAPI; *б* — выявление элементов L1 методом флуоресцентной гибридизации in situ с зондом к 3'UTR; *в* — совмещенное изображение. Масштабная линейка — 10 мкм.

номные перестройки, приводя к делециям и дупликациям, которые вызывают серьезные заболевания или оказываются летальными (Burwinkel, Kilimann, 1998). Недавно обнаружено, что элементы L1 принимают активное участие в рекомбинации на ранних стадиях эмбриогенеза мыши (Wurtele et al., 2005); такие перестройки могут давать начало наследуемым генетическим изменениям и участвовать в функциональной реорганизации геномов млекопитающих.

Известно, что элементы L1 неравномерно распределены по хромосомам человека и мыши; на метафазных хромосомах повторы L1 главным образом локализованы в G-бэндах, AT-богатых областях хромосом, обедненных генами (Korenberg, Rykowski, 1988; Boyle et al., 1990); кроме этого, показано, что хромосомы X и Y существенно обогащены элементами L1 (Boissinot et al., 2001). Считается, что характерное распределение элементов L1 обусловлено процессом их специфического исключения из GC-богатых областей хромосом, которые обогащены генами и обладают повышенным уровнем рекомбинации (Ovchinnikov et al., 2001). Существование процесса исключения элементов L1 из генома может объяснить и обогащение половых хромосом этими повторами. По-видимому, элементы L1 присутствуют в двукратном избытке на половых хромосомах по сравнению с аутосомами, потому что не могут быть удалены с хромосом X и Y с помощью мейотической рекомбинации (Boissinot et al., 2001). Распределение элементов L1 на хромосоме X не является равномерным — область, с которой транскрибируется РНК, принимающая участие в инактивации одной из хромосом X в женских соматических клетках, обогащена повторами L1, а районы, не подвергающиеся инактивации, обеднены повторами L1 (Bailey et al., 2000). Предполагают, что элементы L1 участвуют в инактивации хромосомы X, выступая в качестве «переносчиков» сигналов инактивации (Lyon, 1998, 2000). Обогащение гомологичными копиями элементов L1 определенных областей аутосом, по-видимому, также может влиять на состояние хроматина в этих районах (Allen et al., 2003).

Центромерные районы хромосом высших эукариот обычно содержат протяженные, как правило очень однородные, массивы сателлитных повторов, которые крайне обеднены элементами L1 (Korenberg, Rykowski, 1988; Boyle et al., 1990; Wong, Choo, 2004). Предполагают, что непрерывность массивов центромерных сателлитов обеспе-

чивается механизмом специфического исключения из центромерных районов вновь встраивающихся мобильных элементов за счет неравного кроссинговера (Wong, Choo, 2004). Так, детальный анализ распределения ретротранспозонов L1 мыши в метафазных хромосомах и в интерфазном ядре с помощью конфокальной лазерной сканирующей микроскопии показал, что в центромерном гетерохроматине и в хромоцентрах эти элементы не выявляются (рис. 2).

Поскольку ретротранспозоны L1 являются не только высокоповторяющимися диспергированными последовательностями, но и мобильными элементами, они, несомненно, оказывают серьезное воздействие на геном за счет активности своих регуляторных последовательностей и ретротранспозиции. Наиболее критичные последствия ретротранспозиции L1 связаны с нарушениями первичных последовательностей генов, а также некодирующих функциональных участков хромосом, в частности центромерных районов. У человека выявлено по крайней мере 14 случаев генетических заболеваний, которые вызваны мутациями, связанными с инсерциями L1 (Kazazian, 2004). Повторы L1 содержат последовательности, затрудняющие элонгацию транскрипции (Han et al., 2004), внутренние сигналы полиаденилирования (Perelitsa-Belancio, Deininger, 2003), а также функциональные сайты сплайсинга (Belancio et al., 2006). Поэтому влияние на экспрессию соответствующих белков может оказывать встраивание повторов L1 как в регуляторные элементы и экзоны, так и в интроны генов. Входящие в состав 5'UTR элементов L1 регуляторные последовательности могут влиять на транскрипцию близлежащих генов. Это, в частности, показано для 5'-нетранслируемого района L1 человека (Yang et al., 1998), который обладает как смысловой, так и антисмысловой промоторными активностями (Sprek, 2001; Nigumann et al., 2002).

Известно, что белки ORF1p и ORF2p преимущественно ретротранспозируют именно ту РНК L1, с которой они только что считались, т. е. участвуют в *цис*-ретротранспозиции (Wei et al., 2001). *Цис*-ретротранспозиция способствует существенному ограничению размножения дефектных копий L1 и повышает вероятность того, что элементы L1 останутся активными в геномах млекопитающих (Wei et al., 2001). Однако установлено, что активные элементы L1 способны ретротранспозировать и другие клеточные мРНК, т. е. участвовать в *транс*-ретротранспозиции, хотя

и с низкой эффективностью (Wei et al., 2001). За счет *транс*-ретротранспозиции сплайсированных продуктов транскрипции РНК-полимеразы II в геноме появляются процессированные псевдогены. Кроме этого, показано, что элементы L1 способствуют ретроинтеграции Alu-повторов (Dewannieux et al., 2003).

Элементы L1 обладают слабым сигналом полиаденилирования, поэтому часто терминация транскрипции L1 происходит на каких-то следующих, более сильных сайтах полиаденилирования, что приводит к включению в транскрипты L1 ДНК, фланкирующей их с 3'-стороны. В случае ретроинтеграции таких L1 фланкирующая ДНК переносится на новое место (Moran et al., 1999). Анализ генома человека показал, что примерно 20 % элементов L1 участвовали в таком переносе фланкирующих последовательностей длиной 30—1000 п. н. (Pickeral et al., 2000). Иногда элементы L1 участвуют в переносе последовательности, фланкирующей их с 5'-стороны. Это может произойти, если РНК L1 образовалась за счет транскрипции с вышестоящего промотора (Lander et al., 2001). Еще один способ переноса геномных транскриптов в новый локус обнаружили с помощью компьютерного анализа. Выявлено несколько химерных ретроэлементов, 3'-конец которых соответствовал 3'-области повтора L1, а 5'-конец соответствовал клеточному гену какой-либо ядерной РНК (Buzdin et al., 2003). Предполагают, что химерные последовательности образовались в процессе интеграции элемента L1 за счет переключения его обратной транскриптазы с матрицы РНК L1 на матрицу другой ядерной РНК (Buzdin et al., 2003).

Изучение процесса интеграции элементов L1 в геном показывает, что существует связь между ретроинтеграцией L1 и клеточными системами репарации. Предполагают, что репарационная система клетки может как мешать интеграции L1, репарируя большинство только что образованных эндонуклеазой L1 разрывов в клеточной ДНК, так и способность ретроинтеграции, завершая процесс интеграции L1 (Farkash et al., 2006; Gasior et al., 2006). С другой стороны, в клетках с нарушенной системой репарации элементы L1 в отсутствие эндонуклеазной активности белка ORF2p сохраняют низкий уровень ретроинтеграции, встраиваясь, по-видимому, в двухцепочечные разрывы ДНК (Eickbush, 2002; Morrish et al., 2002). Эндогенная эндонуклеазная активность элементов L1 может без последующей интеграции L1 вносить вклад в образование двухцепочечных разрывов ДНК в клетках животных; в частности, предполагают, что повышенная экспрессия L1 может приводить к геномным перестройкам в опухолевых клетках (Gasior et al., 2006).

Описанные выше данные демонстрируют, что главным результатом активности элементов L1 является структурное ремоделирование генома. Следует подчеркнуть, что наиболее существенное воздействие на геном элементы L1 реализуют посредством своей экспрессии и последующей ретроинтеграции, а также предоставляли клетке обратную транскриптазу и эндонуклеазу.

Тканеспецифичный характер экспрессии LINE1

Экспрессию L1, которая является первым необходимым этапом их ретроинтеграции, изучают в разных типах клеток как на уровне транскрипции, выявляя смысловую дискретную РНК L1, так и на уровне трансляции, вы-

являя белок ORF1p. Такие исследования показали, что у человека, мыши и крысы экспрессия L1, регулируемая видоспецифичными промоторами, тканеспецифична. Экспрессия дискретной смысловой полноразмерной РНК L1 и белка ORF1p наблюдается в эмбриональных карциномах NTera2D1 и F9 (Skowronski, Singer, 1985; Leibold et al., 1990; Martin, Branciforte, 1993), соматических опухолях (Asch et al., 1996), а также в клетках, передаваемых в следующее поколение, а именно на определенных стадиях развития мужских и женских половых клеток (Branciforte, Martin, 1994; Ostertag et al., 2002; Ergun et al., 2004) и в раннем эмбриогенезе (Packer et al., 1993; Trelogan, Martin, 1995; Prak et al., 2003). Специфичная экспрессия L1 в клетках, передаваемых в следующее поколение, способна обеспечить наследование вновь встроившихся копий L1 и, следовательно, успешное размножение элементов L1 в геномах млекопитающих.

Подавление транскрипции L1 происходит в большинстве дифференцированных соматических клеток (Ostertag, Kazazian, 2001). Так, при дифференцировке клеток NTera2D1 и F9 экспрессия смысловой полноразмерной РНК L1 уменьшается (Skowronski, Singer, 1985; Martin, Branciforte, 1993). В клетках печени мыши и крысы вообще не обнаружено дискретных смысловых транскриптов L1, зато выявляется дисперсный набор транскриптов, содержащих смысловую или антисмысловую РНК L1 (Nur et al., 1988; Packer et al., 1993). В клетках печени крысы установлено, что подавляющее большинство таких транскриптов L1 локализовано в ядре (Schmitz et al., 1991). Сходный набор главным образом неполиаденилированных, смысловых и антисмысловых транскриптов L1 всех возможных длин в диапазоне от нескольких сотен до нескольких тысяч нуклеотидов присутствует также во фракции ядерной РНК из клеток NTera2D1 и HeLa; в клетках HeLa существенной экспрессии белка ORF1p не наблюдается (Shafit-Zagardo et al., 1983; Skowronski, Singer, 1985; Leibold et al., 1990).

Регуляция экспрессии LINE1

Одинаковый характер тканеспецифичной экспрессии элементов L1 человека и грызунов свидетельствует об общих принципах регуляции L1 у разных видов, несмотря на отсутствие гомологии между их промоторными областями. Сравнительный анализ структуры видоспецифичных промоторных районов L1 и изучение белков, с ними связывающихся, является первым необходимым шагом для понимания общих механизмов тканеспецифичной регуляции экспрессии L1.

Структурная и функциональная организация промоторных последовательностей элементов L1. Большинство копий L1 человека и грызунов усечено на разном расстоянии от своего 3'-конца (Hutchinson et al., 1989; Cabot et al., 1997; Szak et al., 2002; Martin et al., 2005). При этом усеченные копии L1 не содержат промоторов и поэтому не требуют регуляции своей экспрессии. Вероятно, утрата промоторных последовательностей обусловлена особенностями механизма ретроинтеграции элементов L1 (Zingler et al., 2005). Существует предположение о том, что элементы L1 используют эти особенности своей репликации для захвата новых регуляторных последовательностей и таким образом избегают подавления своей транскрипции, преодолевая последствия инактивирующих мутаций, а также клеточных меха-

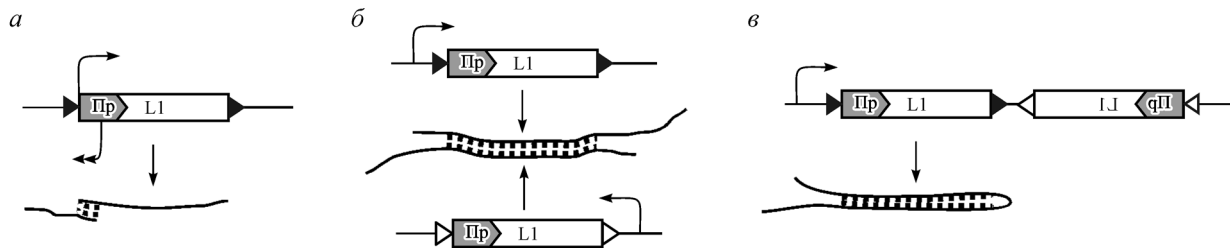


Рис. 3. Возможные способы образования двухцепочечной РНК L1.

a — транскрипция со смыслового и антисмыслового промоторов L1; *б* — транскрипция L1 с соседних промоторов в составе разных транскрипционных единиц; *в* — транскрипция элементов L1, располагающихся в непосредственной близости друг от друга в антипараллельной ориентации. По: Norman et al., 2006, с изменениями.

низмов репрессии (Adey et al., 1994; Furano, 2000; Martin et al., 2005). В рамках этой гипотезы полное отсутствие гомологии между 5'-нетранслируемыми промоторными районами элементов L1 разных видов объясняется повторяющимся приобретением повторами L1 новых регуляторных последовательностей, которое можно рассматривать как постоянный поиск новых активных промоторов (Scott et al., 1987; Furano et al., 1988; Padgett et al., 1988; Schichman et al., 1993).

Установлено, что даже у одного вида подсемейства элементов L1 могут иметь разные промоторные области. Например, у мыши выявлены 4 разных типа мономеров — F, A, Tf и Gf, каждый из которых организует 5'UTRs подсемейств L1Md_F, L1Md_A, L1Md_{Tf} и L1Md_{Gf}, которые появились в геноме мыши около 5.0, 2.5, 0.5 и 0.5 млн лет назад соответственно (Padgett et al., 1988; Schichman et al., 1993; Adey et al., 1994; Naas et al., 1998; Goodier et al., 2001). Эксперименты с репортерными генами показали, что нетандемный 5'-нетранслируемый район L1 человека и тандемные повторы, которые называют мономерами, из 5'UTR L1 грызунов обладают промоторной активностью (Nur et al., 1988; Padgett et al., 1988; Swergold, 1990; Severynse et al., 1992; Naas et al., 1998). Промоторная активность 5'UTR L1 грызунов увеличивается пропорционально числу мономеров в их составе (Segal-Bendirdjian, Heidmann, 1991; DeBerardinis, Kazazian, 1999). Промоторы наиболее «молодых» подсемейств L1 обладают большей транскрипционной активностью и ответственны за синтез большинства специфических смысловых транскриптов L1 (Schichman et al., 1992).

Накопленные к настоящему моменту данные позволяют считать, что элементы L1 транскрибируются РНК-полимеразой II. Промоторы элементов L1 человека, мыши и крысы не содержат консервативных ТАТА-боксов (Furano et al., 1988; Swergold, 1990; DeBerardinis, Kazazian, 1999), следовательно, посадка РНК-полимеразы II на эти промоторы и инициация базальной транскрипции осуществляются с помощью альтернативных механизмов, в частности, возможно, при участии присутствующих в промоторах L1 инициаторных последовательностей (Inr-последовательностей). Обычно РНК-полимераза II иницирует транскрипцию с 3'-стороны от своего промотора, поэтому регуляторные последовательности не попадают в образующийся транскрипт. Для сохранения способности к транскрипции новых копий элементов L1 должен существовать механизм, обеспечивающий включение промоторной области в РНК L1. Элементы LINE1 разных видов решают эту проблему различными способами. В промоторных районах LINEs дрозофилы обнаружены консервативные регуляторные элементы DPEs (downstre-

am promoter elements), которые организуют инициацию транскрипции на Inr-подобных последовательностях, расположенных на строго определенном расстоянии с 5'-стороны от элементов DPE (Minchiotti et al., 1997).

Собственный компьютерный анализ не выявил элементов DPE в необходимых положениях относительно обнаруженных Inr-последовательностей в промоторных областях L1 человека и грызунов. Тандемная организация 5'-нетранслируемых районов L1 грызунов может обеспечивать попадание в транскрипт промоторной последовательности от следующего по ходу транскрипции мономера. Предполагают, что такая тандемная организация 5'-нетранслируемых районов может поддерживаться за счет рекомбинации (Hutchison et al., 1989; Saxton, Martin, 1998). В случае L1 человека сиквенс-специфичное взаимодействие транскрипционного фактора YY1 с промотором играет важную роль в инициации транскрипции ретротранспозона L1 в окрестностях первого нуклеотида промоторной области (Athaniar et al., 2004) и, таким образом, обеспечивает попадание в транскрипт всех необходимых регуляторных последовательностей.

Для некоторых типов клеток характерна экспрессия не только дискретной смысловой полноразмерной РНК L1, но и дисперсного набора смысловой и антисмысловой РНК L1. Вопрос о механизмах формирования таких транскриптов до сих пор остается открытым. Один из них может быть обусловлен функциональными особенностями промоторов некоторых семейств L1, в частности L1 человека, которые обладают антисмысловой активностью (рис. 3, *a*) (Speek, 2001). Однако этот механизм, по-видимому, не может быть универсальным для ретротранспозонов L1, так как промоторы элементов L1 крысы и подсемейства L1Md_F мыши антисмысловой активностью не обладают (Nur et al., 1988; Adey et al., 1994). Чрезвычайно большое количество элементов L1 в геноме позволяет допустить существование других механизмов образования смысловой и антисмысловой РНК L1. В частности, высказано предположение о том, что комплементарные РНК L1 могут появляться в результате попадания некоторых копий повторов L1 в состав разных транскрипционных единиц (рис. 3, *б*) (Nur et al., 1988). В пользу этой гипотезы свидетельствуют данные компьютерного анализа, демонстрирующие, что фрагменты элементов L1 присутствуют в смысловой и антисмысловой ориентации в интронах и нетранслируемых районах 9 % генов человека (Szak et al., 2002). Еще одним способом образования смысловой и антисмысловой РНК L1, вероятно, может быть однонаправленная транскрипция инвертированных повторов, сформированных элементами L1, располагающимися в антипараллельной ориентации на близком расстоянии друг от

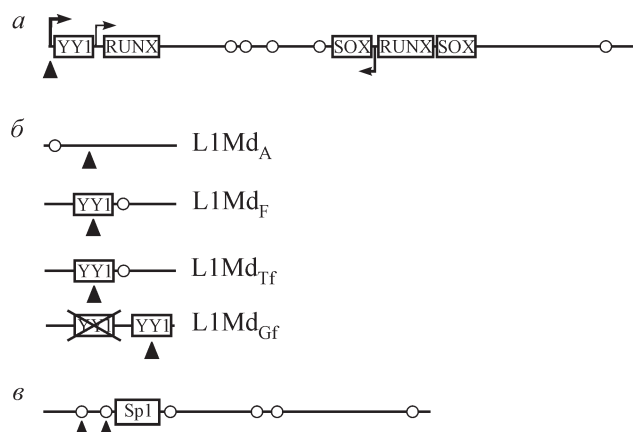


Рис. 4. Регуляторные последовательности в составе видоспецифичных промоторов элементов L1 человека (а), мыши (б) и крысы (в).

Треугольники указывают области, в окрестностях которых обрывается большинство полноразмерных геномных элементов L1, т. е. предполагаемые области инициации транскрипции L1; стрелки указывают положение точек инициации смысловой и антисмысловой транскрипции с промотора L1 человека, кружки обозначают положение инициаторных последовательностей, прямоугольники — положения сайтов связывания транскрипционных факторов.

друга (рис. 3, в). В ходе компьютерного анализа генома человека выявлено несколько десятков таких инвертированных повторов (Norman et al., 2006).

Белки, связывающиеся с промоторами элементов L1 человека и грызунов. Взаимодействие ядерных белков с промоторными районами L1, очевидно, является одним из механизмов, регулирующих транскрипцию L1. Существуют примеры структурно-специфичного связывания ядерных белков с негомологичными регуляторными последовательностями ДНК (Fukuda, 2000; Gadhavi et al., 2001), которые обладают сходной вторичной структурой и поэтому могут участвовать в одинаковых регуляторных процессах. С помощью компьютерного анализа элементов L1 крысы, мыши и человека, используя ресурсы (<http://www.lfd.uci.edu/~gohlke/curve/>) (Shpigelman et al., 1993) и MAR-Wiz 1.0 (<http://www.futuresoft.org/MAR-Wiz/>) (Singh et al., 1997), мы установили, что их промоторные области являются прямыми последовательностями ДНК и не содержат элементов вторичной структуры, которые могли бы связывать одинаковые структурно-специфичные белки. Несмотря на отсутствие гомологии, промоторные последовательности элементов L1 разных видов могут содержать сайты связывания одинаковых транскрипционных факторов и взаимодействовать с ними. Некоторые сиквенса-специфичные белки, связывающие промоторные области L1 разных видов, выявлены и идентифицированы. Результаты сравнительного анализа организации регуляторных последовательностей в составе видоспецифичных промоторов элементов L1 человека, мыши и крысы представлены на рис. 4.

Последовательность, располагающаяся в положениях +13—20 п. н. промотора L1 человека, специфически связывает ядерный белок YY1, влияющий на транскрипционную активность 5'-нетранслируемой области L1 (Minakami et al., 1992; Becker et al., 1993). В промоторе элемента L1 человека сайт связывания для белка YY1 работает как компонент корового промотора, обеспечивающего

инициацию транскрипции около положения +1 области 5'UTR (Athankar et al., 2004). Так как в отсутствие сайта связывания фактора YY1 промотор L1 человека сохраняет транскрипционную активность, в его составе должны существовать другие регуляторные элементы. К таким элементам могут относиться сайты связывания белков RUNX-семейства и факторов SOX-семейства в положениях +83—101, +526—508 и +472—477, +573—578 н. п. соответственно (рис. 4, а). Установлено, что мутация сайтов связывания белков RUNX-семейства уменьшает транскрипционную активность смыслового и антисмыслового промоторов L1 (Yang et al., 2003), а экзогенная экспрессия белков RUNX3 и SOX11 увеличивает транскрипцию элемента L1 человека (Tchenio et al., 2000; Yang et al., 2003). Анализ 5'-нетранслируемых районов элементов L1 грызунов показал, что промоторные мономеры подсемейств L1MdF и L1MdTF мыши также содержат консервативный сайт связывания белка YY1 в окрестностях положения +75 н. п. (рис. 4, б) (DeBerardinis, Kazazian, 1999), а последовательность, располагающаяся в положениях +125—139 п. н. промоторного мономера L1 крысы, ответственна за взаимодействие с неидентифицированными ядерными белками (Furano, 2000). Транскрипционные факторы Sp1/Sp3 способны специфично связываться с этой областью промоторного мономера L1 крысы *in vitro* (рис. 4, в) (Fedotov et al., 2006). Вместе с тем в промоторных мономерах L1 грызунов не обнаружено сайтов связывания для белков семейства RUNX и SOX. Кроме того, промоторы элементов L1 крысы и подсемейства L1MdA мыши не содержат консервативного сайта связывания для фактора YY1 (рис. 4, б, в).

Известно, что большинство геномных копий элементов L1 мыши подсемейств L1MdF и L1MdTF обрывается в окрестностях выявленного консервативного сайта YY1 (DeBerardinis, Kazazian, 1999; Goodier et al., 2001). Наиболее вероятно, что 5'-концы геномных элементов L1 соответствуют 5'-концам РНК L1, т. е. точкам инициации транскрипции. Поэтому можно предположить, что белок YY1 участвует в позиционировании транскрипционного комплекса в окрестности своего сайта связывания на промоторных мономерах подсемейств L1MdF и L1MdTF (DeBerardinis, Kazazian, 1999). Интересно отметить, что у промоторов подсемейства L1MdGF сайт для связывания белка YY1 в окрестностях положения +75 п. н. отличается от консенсуса одной нуклеотидной заменой (Goodier et al., 2001) (рис. 4, б, *перечеркнутый прямоугольник*). Однако компьютерный анализ показал, что в 3'-области промоторов подсемейства L1MdGF содержится еще один консервативный сайт для связывания белка YY1 и большинство полноразмерных элементов подсемейства L1MdGF обрывается в окрестностях этого сайта (рис. 4, б). Следовательно, белок YY1, по-видимому, может участвовать также в инициации транскрипции элементов подсемейства L1MdGF. Большинство полноразмерных геномных элементов L1 крысы обрывается в окрестностях Inr-последовательности в положении +54 п. н. промоторного мономера (Fedorov et al., 2006) (рис. 4, в). Поэтому можно предположить, что эта Inr-последовательность является главным сайтом инициации транскрипции L1 крысы: отсутствие первых 56 п. н. в промоторе L1 существенно уменьшает его транскрипционную активность (Segal-Bendirdjian, Heidmann, 1991). Белок Sp1 может способствовать инициации транскрипции посредством привлечения фактора TFIID к Inr-элементам (Lania et al., 1997; Li et al., 2004); таким образом, можно предположить его участие в

организации транскрипции на Inr-содержащих промоторах L1 крысы.

Мы предполагаем, что транскрипционный фактор YY1 в случае промоторов L1 человека и промоторных мономеров подсемейств L1Md_F и L1Md_{Tf} и L1Md_{Gf} мыши и белки Sp-семейства в случае промотора L1 крысы могут играть роль в привлечении к промоторам элементов L1 РНК-полимеразы II. Последняя, по-видимому, инициирует транскрипцию в окрестности сайтов YY1 промоторов L1 человека и мыши и на выявленных Inr-последовательностях промотора L1 крысы. Таким образом, можно заключить, что сходный характер экспрессии элементов L1 разных видов, по-видимому, обеспечивают разные транскрипционные факторы.

Однако остается не до конца понятным, каким образом выявленные белки могут влиять на тканеспецифичность экспрессии L1, поскольку фактор YY1 и белки Sp1/Sp3 экспрессируются повсеместно (Shrivastava, Calame, 1994; Li et al., 2004), а уровень эндогенной экспрессии белков SOX11 и RUNX3 не коррелирует с уровнем транскрипции дискретных РНК L1 (Wegner, 1999; Yang et al., 2003). Кроме этого, установлено, что промоторы элементов L1 человека в составе экзогенных конструкций с репортерными генами работают одинаково эффективно в клетках, где наблюдается экспрессия эндогенных элементов L1, и в клетках, где такая экспрессия отсутствует (Steinhoff, Schulz, 2003). Все эти данные показывают, что помимо выявленных регуляторных белков существуют дополнительные факторы, обеспечивающие наблюдаемую тканеспецифичность экспрессии L1.

Метилирование промоторных районов L1. Как в опухолевых клетках, так и в нормальных тканях повышенный уровень экспрессии элементов L1 сопровождается общим явлением частичного деметилирования промоторов L1. В частности, установлено, что промоторы геномных элементов L1 человека полностью метилированы в клетках HeLa и частично деметилированы в клетках NTera2D1 (Thayer et al., 1993). Кроме этого, частичное деметилирование промоторной области L1 наблюдалось нами в семенниках крысы, а также выявлено в опухолевых клетках печени человека и крысы (Takai et al., 2000; Pogribny et al., 2006), у мыши в эмбриогенезе на стадии бластоцисты (Lane et al., 2003) и в первичных половых клетках на 15.5-е сут развития (Lees-Murdock et al., 2003), т. е. в тех клетках, где ранее была выявлена экспрессия L1 (Packer et al., 1993; Trelogan, Martin, 1995). Напротив, в нормальных клетках печени промоторы L1 полностью метилированы (Nur et al., 1988; Takai et al., 2000). Следовательно, тканеспецифичное метилирование промоторных последовательностей, по-видимому, связано с регуляцией экспрессии элементов L1. Это предположение находит дополнительные экспериментальные подтверждения. Так, установлено, что метилирование *in vitro* цитозинов в динуклеотидах CpG промоторных последовательностей L1 человека и крысы уменьшает их транскрипционную активность более чем на 70 %, в том числе и в тех клетках, где наблюдается экспрессия эндогенных L1 (Nur et al., 1988; Steinhoff, Schulz, 2003). Обработка культуры клеток мыши линии NIH3T3 5-азациитидином, вызывающая деметилирование ДНК, приводит к появлению смыслового полноразмерного транскрипта L1 и белка ORF1p (Tchenio et al., 1993). Метилирование промоторных областей может оказывать влияние на экспрессию посредством привлечения белков, связывающих метилированную ДНК (Boyes, Bird, 1991). Взаимодействие таких белков с мети-

лированными промоторами элементов L1 разных видов может служить общим механизмом регуляции L1. Эта гипотеза согласуется с данными о том, что представитель семейства метил-CpG-связывающих белков — MeCP2 — способен дополнительно репрессировать транскрипцию и ретротранспозицию экзогенного элемента L1 человека, промотор которого метилирован (Yu et al., 2001).

Делеция гена ДНК-метилазы Dnmt3L предотвращает метилирование промоторов L1 мыши *de novo* и приводит к повышению уровня полноразмерных транскриптов L1 в мужских половых клетках (Bourc'his, Bestor, 2004). Таким образом, эпигенетические механизмы установления метилирования, а также частичного деметилирования промотора L1 в определенных типах клеток, по-видимому, могут обеспечивать тканеспецифичную экспрессию элементов L1.

Роль двухцепочечной РНК в регуляции экспрессии ретротранспозонов L1. Известно, что синтезированная РНК может играть существенную роль в регуляции экспрессии генов; это явление получило название РНК-сайленсинг (подавление экспрессии). Недавно было установлено, что в клетках эукариот длинная двухцепочечная РНК, а также РНК, формирующая шпилечные структуры, может процессироваться с образованием малых РНК. Оказалось, что такие малые РНК в свою очередь могут участвовать в РНК-сайленсинге посредством нескольких механизмов — РНК-интерференции (специфичной деградации гомологичной РНК), подавления трансляции гомологичных транскриптов, а также транскрипционного сайленсинга генов (индукции гетерохроматинизации гомологичной области ДНК за счет модификации гистонов и(или) метилирования ДНК) (Аравин и др., 2002; Bayne, Allshire, 2005; Kim, 2005; Weinberg et al., 2006).

Исследования последних лет убедительно доказали участие РНК-сайленсинга в регуляции мобильных элементов у многих эукариот. В частности, показана роль этих механизмов в репрессии ДНК-транспозонов в половых клетках *C. elegans* (Vastenhouw, Plasterk, 2004), LINE-подобных ретротранспозонов у *N. crassa* (Nolan et al., 2005) и *T. brucei* (Shi et al., 2004), а также ретротранспозонов, содержащих длинные концевые повторы у *D. melanogaster* (Kalmykova et al., 2005). Открытие явления РНК-сайленсинга у млекопитающих позволило изменить существующие представления о роли смысловых и антисмысловых транскриптов L1 в клетках и предположить участие этого механизма в регуляции экспрессии L1 (Soifer et al., 2005; Horman et al., 2006; Soifer, 2006). Вскоре появились первые экспериментальные подтверждения такого предположения. Так, в клетках HEK293 удалось обнаружить эндогенные смысловые и антисмысловые малые РНК, гомологичные ретротранспозону L1 человека, а именно его промоторной области (Yang, Kazazian, 2006). Репрессия в клетках HeLa белка Dicer1, участвующего в процессировании двухцепочечной РНК до малых РНК, приводила к увеличению экспрессии и эффективности ретротранспозиции конструкций под контролем полноразмерного промотора L1 (Yang, Kazazian, 2006). Авторы предполагают, что малые РНК, гомологичные элементам L1, участвуют в репрессировании этих ретротранспозонов на посттранскрипционном уровне посредством деградации мРНК, т. е. механизма РНК-интерференции. Однако нельзя исключать и другие пути регуляции, значимость которых еще предстоит выяснить. Роль РНК в регуляции экспрессии L1 подкрепляется собственными

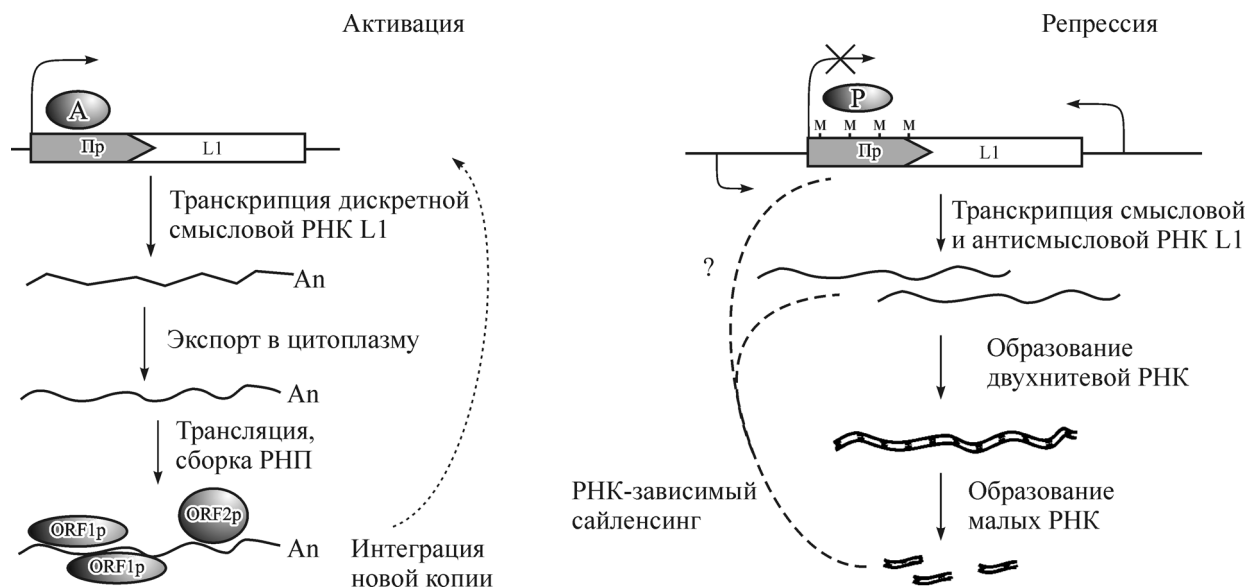


Рис. 5. Предполагаемый механизм регуляции экспрессии элементов L1.
А — белок-активатор, Р — белок-репрессор, М — метилирование промоторной области.

данными о выявлении малых РНК, гомологичных ретротранспозонам L1, в семенниках крысы и результатами других авторов, обнаруживших малые РНК L1 в ооцитах и семенниках мыши (Watanabe et al., 2006; Aravin et al., 2007). Предполагают, что повышенная способность к ретротранспозиции элементов L1 с модифицированной последовательностью РНК при сохраненной аминокислотной последовательности (Perelitsa-Belancio, Deininger, 2003; Nan, Voeke, 2004) может являться следствием участия РНК-сайленсинга в репрессировании L1 (Horman et al., 2006).

Совокупность полученных к настоящему времени данных можно обобщить в следующей схеме регуляции экспрессии элементов L1 (рис. 5). При активации транскрипции элементов L1 происходит синтез полноразмерной смысловой РНК L1 и белков ORF1p, ORF2p; промоторы таких элементов L1 деметилированы и взаимодействуют с активаторами транскрипции, определяемыми конкретной видоспецифичной регуляторной последовательностью. В случае репрессии транскрипции элементов L1 их промоторы метилированы и не взаимодействуют с белками-активаторами. Дополнительное репрессирование метилированных промоторов может быть обусловлено привлечением метил-СрG-связывающих белков, а также взаимодействующих с ними корепрессоров. Синтезируемые смысловые и антисмысловые РНК L1 могут образовывать двухцепочечную РНК и предположительно участвовать в РНК-зависимом сайленсинге повторов L1.

Метилирование ДНК считается универсальным механизмом репрессирования мобильных элементов генома (Yoder et al., 1997; Bestor, 2003). Однако во время гаметогенеза и эмбрионального развития, когда ретротранспозоны частично деметилируются, ключевую роль в их регуляции могут играть механизмы РНК-сайленсинга (White-law, Martin, 2001; Soifer, Rossi, 2006). Существует предположение о том, что именно пути РНК-зависимого транскрипционного сайленсинга используются для распознавания и установления специфического метилирования повторяющихся мобильных элементов (Meunier et al., 2005).

Заключение

Несомненно, что механизмы регуляции ретротранспозиции элементов L1 требуют всестороннего изучения, так как их специфичная экспрессия оказывает большое влияние на remodelирование генома и сопровождается малигнизацией клеток. У разных организмов, несмотря на отсутствие гомологии в регуляторных районах элементов L1, выработаны сходные стратегии контроля их тканеспецифичной транскрипции. Ключевую регуляторную роль, очевидно, оказывают не связывающиеся с промоторами широко распространенные транскрипционные факторы, но эпигенетические механизмы, включающие метилирование ДНК промоторных районов.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект 04-04-48941).

Список литературы

- Аравин А. А., Кленов М. С., Вагин В. В., Розовский Я. М., Гвоздев В. А. 2002. Роль двухцепочечной РНК в подавлении экспрессии генов эукариот. Молекуляр. биол. 36 (2) : 240—251.
- Adey N. B., Schichman S. A., Graham D. K., Peterson S. N., Edgell M. H., Hutchison C. A. 1994. Rodent L1 evolution has been driven by a single dominant lineage that has repeatedly acquired new transcription regulatory sequences. Mol. Biol. Evol. 11 : 778—789.
- Allen E., Horvath S., Tong F., Kraft P., Spiteri E., Riggs A. D., Marahrens Y. 2003. High concentrations of long interspersed nuclear element sequence distinguish monoallelically expressed genes. Proc. Nat. Acad. Sci. USA. 100 : 9940—9945.
- Aravin A. A., Sachidanandam R., Girard A., Fejes-Toth K., Hannon G. J. 2007. Developmentally regulated piRNA clusters implicate MILI in transposon control. Science. 316 : 744—747.
- Asch H. L., Eliacin E., Fanning T. G., Connolly J. L., Bratthauer G., Asch B. B. 1996. Comparative expression of the LINE-1 p40 protein in human breast carcinomas and normal breast tissues. Oncol. Res. 8 : 239—247.

- Athanikar J. N., Badge R. M., Moran J. V. 2004. A YY1-binding site is required for accurate human LINE-1 transcription initiation. *Nucl. Acids Res.* 32 : 3846—3855.
- Bailey J. A., Carrel L., Chakravarti A., Eichler E. E. 2000. Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. *Proc. Nat. Acad. Sci. USA.* 97 : 6634—6639.
- Bayne E. H., Allshire R. C. 2005. RNA-directed transcriptional gene silencing in mammals. *Trends Genet.* 21 : 370—373.
- Becker K. G., Swergold G. D., Ozato K., Thayer R. E. 1993. Binding of the ubiquitous nuclear transcription factor YY1 to a cis regulatory sequence in the human LINE-1 transposable element. *Hum. Mol. Genet.* 2 : 1697—1702.
- Belancio V. P., Hedges D. J., Deininger P. 2006. LINE-1 RNA splicing and influences on mammalian gene expression. *Nucl. Acids Res.* 34 : 1512—1521.
- Bestor T. H. 2003. Cytosine methylation mediates sexual conflict. *Trends Genet.* 19 : 185—190.
- Boissinot S., Entezam A., Furano A. V. 2001. Selection against deleterious LINE-1-containing loci in the human lineage. *Mol. Biol. Evol.* 18 : 926—935.
- Bourc'his D., Bestor T. H. 2004. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature.* 431 : 96—99.
- Boyes J., Bird A. 1991. DNA methylation inhibits transcription indirectly via a methyl-CpG-binding protein. *Cell.* 64 : 1123—1134.
- Boyle A. L., Ballard S. G., Ward D. C. 1990. Differential distribution of long and short interspersed element sequences in the mouse genome: chromosome karyotyping by fluorescence *in situ* hybridization. *Proc. Nat. Acad. Sci. USA.* 87 : 7757—7761.
- Branciforte D., Martin S. L. 1994. Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. *Mol. Cell. Biol.* 14 : 2584—2592.
- Brouha B., Meischl C., Ostertag E., de Boer M., Zhang Y., Neijens H., Roos D., Kazazian H. H., Jr. 2002. Evidence consistent with human L1 retrotransposition in maternal meiosis I. *Amer. J. Hum. Genet.* 71 : 327—336.
- Brown S. D., Dover G. 1981. Organization and evolutionary progress of a dispersed repetitive family of sequences in widely separated rodent genomes. *J. Mol. Biol.* 150 : 441—466.
- Burton F. H., Loeb D. D., Voliva C. F., Martin S. L., Edgell M. H., Hutchison C. A. 1986. Conservation throughout mammalia and extensive protein-encoding capacity of the highly repeated DNA long interspersed sequence one. *J. Mol. Biol.* 187 : 291—304.
- Burwinkel B., Kilimann M. W. 1998. Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. *J. Mol. Biol.* 277 : 513—517.
- Buzdin A., Gogvadze E., Kovalskaya E., Volchkov P., Ustyugova S., Illarionova A., Fushan A., Vinogradova T., Sverdlov E. 2003. The human genome contains many types of chimeric retrogenes generated through *in vivo* RNA recombination. *Nucl. Acids Res.* 31 : 4385—4390.
- Cabot E. L., Angeletti B., Usdin K., Furano A. V. 1997. Rapid evolution of a young L1 (LINE-1) clade in recently speciated *Rattus* taxa. *J. Mol. Evol.* 45 : 412—423.
- D'Ambrosio E., Waitzkin S. D., Witney F. R., Salemme A., Furano A. V. 1986. Structure of the highly repeated, long interspersed DNA family (LINE or L1Rn) of the rat. *Mol. Cell. Biol.* 6 : 411—424.
- DeBerardinis R. J., Kazazian H. H., Jr. 1999. Analysis of the promoter from an expanding mouse retrotransposon subfamily. *Genomics.* 56 : 317—323.
- Dewannieux M., Esnault C., Heidmann T. 2003. LINE-mediated retrotransposition of marked Alu sequences. *Nat. Genet.* 35 : 41—48.
- Dombroski B. A., Feng Q., Mathias S. L., Sassaman D. M., Scott A. F., Kazazian H. H., Jr., Boeke J. D. 1994. An *in vivo* assay for the reverse transcriptase of human retrotransposon L1 in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 14 : 4485—4492.
- Eickbush T. H. 2002. Repair by retrotransposition. *Nat. Genet.* 31 : 126—127.
- Ergun S., Buschmann C., Heukeshoven J., Dammann K., Schnieders F., Lauke H., Chalajour F., Kilic N., Stratling W. H., Schumann G. G. 2004. Cell type-specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues. *J. Biol. Chem.* 279 : 27 753—27 763.
- Fanning T. G. 1983. Size and structure of the highly repetitive BAM HI element in mice. *Nucl. Acids Res.* 11 : 5073—5091.
- Farkash E. A., Kao G. D., Horman S. R., Prak E. T. 2006. Gamma radiation increases endonuclease-dependent L1 retrotransposition in a cultured cell assay. *Nucl. Acids Res.* 34 : 1196—1204.
- Fawcett D. H., Lister C. K., Kellett E., Finnegan D. J. 1986. Transposable elements controlling I-R hybrid dysgenesis in *D. melanogaster* are similar to mammalian LINES. *Cell.* 47 : 1007—1015.
- Fedorov A. V., Lukyanov D. V., Podgornaya O. I. 2006. Identification of the proteins specifically binding to the rat LINE1 promoter. *Biochem. Biophys. Res. Commun.* 340 : 553—559.
- Feng Q., Moran J. V., Kazazian H. H., Jr., Boeke J. D. 1996. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell.* 87 : 905—916.
- Fitch D. H., Bailey W. J., Tagle D. A., Goodman M., Sieu L., Slightom J. L. 1991. Duplication of the gamma-globin gene mediated by L1 long interspersed repetitive elements in an early ancestor of simian primates. *Proc. Nat. Acad. Sci. USA.* 88 : 7396—7400.
- Fukuda Y. 2000. Interaction of nuclear proteins with intrinsically curved DNA in a matrix attachment region of a tobacco gene. *Plant Mol. Biol.* 44 : 91—98.
- Furano A. V. 2000. The biological properties and evolutionary dynamics of mammalian LINE-1 retrotransposons. *Prog. Nucl. Acid Res. Mol. Biol.* 64 : 255—294.
- Furano A. V., Robb S. M., Robb F. T. 1988. The structure of the regulatory region of the rat L1 (L1Rn, long interspersed repeated) DNA family of transposable elements. *Nucl. Acids Res.* 16 : 9215—9231.
- Gadhavi P. L., Greenwood M. D., Strom M., King I. A., Buxton R. S. 2001. The regulatory region of the human desmocollin 3 promoter forms a DNA four-way junction. *Biochem. Biophys. Res. Commun.* 281 : 520—528.
- Gasior S. L., Wakeman T. P., Xu B., Deininger P. L. 2006. The human LINE-1 retrotransposon creates DNA double-strand breaks. *J. Mol. Biol.* 357 : 1383—1393.
- Gibbs R. A., Weinstock G. M., Metzker M. L., Muzny D. M., Sodergren E. J. et al. 2004. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature.* 428 : 493—521.
- Goodier J. L., Ostertag E. M., Du K., Kazazian H. H., Jr. 2001. A novel active L1 retrotransposon subfamily in the mouse. *Genome Res.* 11 : 1677—1685.
- Han J. S., Boeke J. D. 2004. A highly active synthetic mammalian retrotransposon. *Nature.* 429 : 314—318.
- Han J. S., Szak S. T., Boeke J. D. 2004. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. *Nature.* 429 : 268—274.
- Hayward B. E., Zavanelli M., Furano A. V. 1997. Recombination creates novel L1 (LINE-1) elements in *Rattus norvegicus*. *Genetics.* 146 : 641—654.
- Horman S. R., Svoboda P., Prak E. T. 2006. The potential regulation of 11 mobility by RNA interference. *J. Biomed. Biotechnol.* 2006 : 32 713.
- Hutchison C. A., Hardies S. C., Loeb D. D., Shehee W. R., Edgell M. H. 1989. LINES and related retroposons: long interspersed repeated sequences in the eucaryotic genome. In: *Mobile DNA*. Washington: Amer. Soc. Microbiol. 593—617.
- Kalmykova A. I., Klenov M. S., Gvozdev V. A. 2005. Argonaute protein PIWI controls mobilization of retrotransposons in the *Drosophila* male germline. *Nucl. Acids Res.* 33 : 2052—2059.
- Kazazian H. H., Jr. 2004. Mobile elements: drives of genome evolution. *Science.* 303 : 1626—1632.
- Kim V. N. 2005. Small RNAs: classification, biogenesis, and function. *Mol. Cells.* 19 : 1—15.
- Kimmel B. E., ole-MoiYoi O. K., Young J. R. 1987. Ingi, a 5.2-kb dispersed sequence element from *Trypanosoma brucei* that

carries half of a smaller mobile element at either end and has homology with mammalian LINES. *Mol. Cell Biol.* 7 : 1465—1475.

Korenberg J. R., Rykowski M. C. 1988. Human genome organization: Alu, Lines, and the molecular structure of metaphase chromosome bands. *Cell.* 53 : 391—400.

Kulpa D. A., Moran J. V. 2006. Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. *Nat. Struct. Mol. Biol.* 13 : 655—660.

Lander E. S., Linton L. M., Birren B., Nusbaum C., Zody M. C. et al. 2001. Initial sequencing and analysis of the human genome. *Nature.* 409 : 860—921.

Lane N., Dean W., Erhardt S., Hajkova P., Surani A., Walter J., Reik W. 2003. Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis.* 35 : 88—93.

Lania L., Majello B., De Luca P. 1997. Transcriptional regulation by the Sp family proteins. *Int. J. Biochem. Cell. Biol.* 29 : 1313—1323.

Lees-Murdock D. K., De Felici M., Walsh C. P. 2003. Methylation dynamics of repetitive DNA elements in the mouse germ cell lineage. *Genomics.* 82 : 230—237.

Leibold D. M., Swergold G. D., Singer M. F., Thayer R. E., Dombroski B. A., Fanning T. G. 1990. Translation of LINE-1 DNA element *in vitro* and in human cells. *Proc. Nat. Acad. Sci. USA.* 87 : 6990—6994.

Li L., He S., Sun J. M., Davie J. R. 2004. Gene regulation by Sp1 and Sp2. *Biochem. Cell. Biol.* 82 : 460—471.

Loeb D. D., Padgett R. W., Hardies S. C., Shehee W. R., Comer M. B., Edgell M. H., Hutchison C. A. 1986. The sequence of a large L1Md element reveals a tandemly repeated 5' end and several features found in retrotransposons. *Mol. Cell. Biol.* 6 : 168—182.

Luan D. D., Korman M. H., Jakubczak J. L., Eickbush T. H. 1993. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism, for non-LTR retrotransposition. *Cell.* 72 : 595—605.

Lyon M. F. 1998. X-chromosome inactivation: a repeat hypothesis. *Cytogenet. Cell Genet.* 80 : 133—137.

Lyon M. F. 2000. LINE-1 elements and X chromosome inactivation: a function for «junk» DNA? *Proc. Nat. Acad. Sci. USA.* 97 : 6248—6249.

Maio J. J., Brown F. L., McKenna W. G., Musich P. R. 1981. Toward a molecular paleontology of primate genomes. II. The KpnI families of alphoid DNAs. *Chromosoma.* 83 : 127—144.

Martin S. L., Branciforte D. 1993. Synchronous expression of LINE-1 RNA and protein in mouse embryonal carcinoma cells. *Mol. Cell. Biol.* 13 : 5383—5392.

Martin S. L., Branciforte D., Keller D., Bain D. L. 2003. Trimeric structure for an essential protein in L1 retrotransposition. *Proc. Nat. Acad. Sci. USA.* 100 : 13 815—13 820.

Martin S. L., Li W. L., Furano A. V., Boissinot S. 2005. The structures of mouse and human L1 elements reflect their insertion mechanism. *Cytogenet. Genome Res.* 110 : 223—228.

Mathias S. L., Scott A. F., Kazazian H. H., Jr., Boeke J. D., Gabriel A. 1991. Reverse transcriptase encoded by a human transposable element. *Science.* 254 : 1808—1810.

Meunier J., Khelifi A., Navratil V., Duret L. 2005. Homology-dependent methylation in primate repetitive DNA. *Proc. Nat. Acad. Sci. USA.* 102 : 5471—5476.

Meunier-Rotival M., Soriano P., Cuny G., Strauss F., Bernardi G. 1982. Sequence organization and genomic distribution of the major family of interspersed repeats of mouse DNA. *Proc. Nat. Acad. Sci. USA.* 79 : 355—359.

Minakami R., Kurose K., Etoh K., Furuhashi Y., Hattori M., Sakaki Y. 1992. Identification of an internal cis-element essential for the human L1 transcription and a nuclear factor(s) binding to the element. *Nucl. Acids Res.* 20 : 3139—3145.

Minchiotti G., Contursi C., Di Nocera P. P. 1997. Multiple downstream promoter modules regulate the transcription of the *Drosophila melanogaster*. I. Doc and F elements. *J. Mol. Biol.* 267 : 37—46.

Moran J. V., DeBerardinis R. J., Kazazian H. H., Jr. 1999. Exon shuffling by L1 retrotransposition. *Science.* 283 : 1530—1534.

Moran J. V., Holmes S. E., Naas T. P., DeBerardinis R. J., Boeke J. D., Kazazian H. H., Jr. 1996. High frequency retrotransposition in cultured mammalian cells. *Cell.* 87 : 917—927.

Morrish T. A., Gilbert N., Myers J. S., Vincent B. J., Stamatou T. D., Taccioli G. E., Batzer M. A., Moran J. V. 2002. DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. *Nat. Genet.* 31 : 159—165.

Naas T. P., DeBerardinis R. J., Moran J. V., Ostertag E. M., Kingsmore S. F., Seldin M. F., Hayashizaki Y., Martin S. L., Kazazian H. H. 1998. An actively retrotransposing, novel subfamily of mouse L1 elements. *EMBO J.* 17 : 590—597.

Nigumann P., Redik K., Matlik K., Speek M. 2002. Many human genes are transcribed from the antisense promoter of L1 retrotransposon. *Genomics.* 79 : 628—634.

Nolan T., Braccini L., Azzalin G., De Toni A., Macino G., Congoni C. 2005. The post-transcriptional gene silencing machinery functions independently of DNA methylation to repress a LINE1-like retrotransposon in *Neurospora crassa*. *Nucl. Acids Res.* 33 : 1564—1573.

Noma K., Ohtsubo H., Ohtsubo E. 2000. ATLN elements, LINEs from *Arabidopsis thaliana*: identification and characterization. *DNA Res.* 7 : 291—303.

Nur I., Pascale E., Furano A. V. 1988. The left end of rat L1 (L1Rn, long interspersed repeated) DNA which is a CpG island can function as a promoter. *Nucl. Acids Res.* 16 : 9233—9251.

Ostertag E. M., DeBerardinis R. J., Goodier J. L., Zhang Y., Yang N., Gerton G. L., Kazazian H. H., Jr. 2002. A mouse model of human L1 retrotransposition. *Nat. Genet.* 32 : 655—660.

Ostertag E. M., Kazazian H. H., Jr. 2001. Biology of mammalian L1 retrotransposons. *Annu. Rev. Genet.* 35 : 501—538.

Ovchinnikov I., Troxel A. B., Swergold G. D. 2001. Genomic characterization of recent human LINE-1 insertions: evidence supporting random insertion. *Genome Res.* 11 : 2050—2058.

Packer A. I., Manova K., Bachvarova R. F. 1993. A discrete LINE-1 transcript in mouse blastocysts. *Develop. Biol.* 157 : 281—283.

Padgett R. W., Hutchison C. A., Edgell M. H. 1988. The F-type 5' motif of mouse L1 elements: a major class of L1 termini similar to the A-type in organization but unrelated in sequence. *Nucl. Acids Res.* 16 : 739—749.

Perpelitsa-Belancio V., Deininger P. 2003. RNA truncation by premature polyadenylation attenuates human mobile element activity. *Nat. Genet.* 35 : 363—366.

Pickeral O. K., Makalowski W., Boguski M. S., Boeke J. D. 2000. Frequent human genomic DNA transduction driven by LINE-1 retrotransposition. *Genome Res.* 10 : 411—415.

Pogribny I. P., Ross S. A., Tryndyak V. P., Pogribna M., Poirier L. A., Karpinet T. V. 2006. Histone H3 lysine 9 and H4 lysine 20 trimethylation and the expression of Suv4-20h2 and Suv39h1 histone methyltransferases in hepatocarcinogenesis induced by methyl deficiency in rats. *Carcinogenesis.* 27 : 1180—1186.

Pont-Kingdon G., Chi E., Christensen S., Carroll D. 1997. Ribonucleoprotein formation by the ORF1 protein of the non-LTR retrotransposon Tx1L in *Xenopus* oocyte. *Nucl. Acids Res.* 25 : 3088—3094.

Prak E. T., Dodson A. W., Farkash E. A., Kazazian H. H., Jr. 2003. Tracking an embryonic L1 retrotransposition event. *Proc. Nat. Acad. Sci. USA.* 100 : 1832—1837.

Saxton J. A., Martin S. L. 1998. Recombination between subtype creates a mosaic lineage of LINE-1 that is expressed and actively retrotransposing in the mouse genome. *J. Mol. Biol.* 280 : 611—622.

Schichman S. A., Adey N. B., Edgell M. H., Hutchison C. A. 1993. L1 A-monomer tandem arrays have expanded during the course of mouse L1 evolution. *Mol. Biol. Evol.* 10 : 552—570.

Schmitz E., Mohr E., Richter D. 1991. Rat vasopressin and oxytocin genes are linked by a long interspersed repeated DNA element (LINE): sequence and transcriptional analysis of LINE. *DNA Cell. Biol.* 10 : 81—91.

Scott A. F., Schmeckpeper B. J., Abdelrazik M., Comey C. T., O'Hara B., Rossiter J. P., Cooley T., Heath P., Smith K. D., Margole L. 1987. Origin of the human L1 elements: proposed progeni-

- tor genes deduced from a consensus DNA sequence. *Genomics*. 1 : 113—125.
- Segal-Bendirdjian E., Heidmann T. 1991. Evidence for a reverse transcription intermediate for a marker line transposon in tumoral rat cells. *Biochem. Biophys. Res. Commun.* 181 : 863—870.
- Severynse D. M., Hutchison C. A., Edgell M. H. 1992. Identification of transcriptional regulatory activity within the 5' A-type monomer sequence of the mouse LINE-1 retroposon. *Mamm. Genome*. 2 : 41—50.
- Shafit-Zagardo B., Brown F. L., Zavodny P. J., Maio J. J. 1983. Transcription of the KpnI families of long interspersed DNAs in human cells. *Nature*. 304 : 277—280.
- Shi H., Djikeng A., Tschudi C., Ullu E. 2004. Argonaute protein in the early divergent eukaryote *Trypanosoma brucei*: control of small interfering RNA accumulation and retroposon transcript abundance. *Mol. Cell. Biol.* 24 : 420—427.
- Shpigelman E. S., Trifonov E. N., Bolshoy A. 1993. Curvature: software for the analysis of curved DNA. *Comput. Appl. Biosci.* 9 : 435—440.
- Shrivastava A., Calame K. 1994. An analysis of genes regulated by the multi-functional transcriptional regulator Yin Yang-1. *Nucl. Acids Res.* 22 : 5151—5155.
- Singer M. F. 1982. SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes. *Cell*. 28 : 433—434.
- Singer M. F., Thayer R. E., Grimaldi G., Lerman M. I., Fanning T. G. 1983. Homology between the KpnI primate and BamHI (MIF-1) rodent families of long interspersed repeated sequences. *Nucl. Acids Res.* 11 : 5739—5745.
- Singh G., Kramer J., Krawetz S. 1997. Mathematical model to predict regions of chromatin attachment to the nuclear matrix. *Nucl. Acids Res.* 25 : 1419—1425.
- Skowronski J., Singer M. F. 1985. Expression of a cytoplasmic LINE-1 transcript is regulated in a human teratocarcinoma cell line. *Proc. Nat. Acad. Sci. USA*. 82 : 6050—6054.
- Soifer H. S. 2006. Do small RNAs interfere with LINE-1? *J. Biomed. Biotechnol.* 2006 : 29 049.
- Soifer H. S., Rossi J. J. 2006. Small interfering RNAs to the rescue: blocking L1 retrotransposition. *Nat. Struct. Mol. Biol.* 13 : 758—759.
- Soifer H. S., Zaragoza A., Peyvan M., Behlke M. A., Rossi J. J. 2005. A potential role for RNA interference in controlling the activity of the human LINE-1 retrotransposon. *Nucl. Acids Res.* 33 : 846—856.
- Speek M. 2001. Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. *Mol. Cell. Biol.* 21 : 1973—1985.
- Steinhoff C., Schulz W. A. 2003. Transcriptional regulation of the human LINE-1 retrotransposon L1. 2B. *Mol. Genet. Genomics*. 270 : 394—402.
- Swergold G. D. 1990. Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol. Cell. Biol.* 10 : 6718—6729.
- Szak S. T., Pickeral O. K., Makalowski W., Boguski M. S., Landsman D., Boeke J. D. 2002. Molecular archeology of L1 insertions in the human genome. *Genome Biol.* 3 : research0052.
- Takai D., Yagi Y., Habib N., Sugimura T., Ushijima T. 2000. Hypomethylation of LINE1 retrotransposon in human hepatocellular carcinomas, but not in surrounding liver cirrhosis. *Jap. J. Clin. Oncol.* 30 : 306—309.
- Tchenio T., Casella J. F., Heidmann T. 2000. Members of the SRY family regulate the human LINE retrotransposons. *Nucl. Acids Res.* 28 : 411—415.
- Tchenio T., Segal-Bendirdjian E., Heidmann T. 1993. Generation of processed pseudogenes in murine cells. *EMBO J.* 12 : 1487—1497.
- Thayer R. E., Singer M. F., Fanning T. G. 1993. Undermethylation of specific LINE-1 sequence in human cells producing a LINE-1-encoded protein. *Gene*. 133 : 273—277.
- Trelogan S. A., Martin S. L. 1995. Tightly regulated, developmentally specific expression of the first open reading frame from LINE-1 during mouse embryogenesis. *Proc. Nat. Acad. Sci. USA*. 92 : 1520—1524.
- Vandergon T. L., Reitman M. 1994. Evolution of chicken repeat 1 (CR1) elements: evidence for ancient subfamilies and multiple progenitors. *Mol. Biol. Evol.* 11 : 886—898.
- Vastenhouw N. L., Plasterk R. H. 2004. RNAi protects the *Caenorhabditis elegans* germline against transposition. *Trends Genet.* 20 : 314—319.
- Watanabe T., Takeda A., Tsukiyama T., Mise K., Okuno T., Sasaki H., Minami N., Imai H. 2006. Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. *Genes Develop.* 20 : 1732—1743.
- Waterston R. H., Lindblad-Toh K., Birney E., Rogers J., Abril J. E. et al. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature*. 420 : 520—562.
- Wegner M. 1999. From head to toes: the multiple facets of Sox proteins. *Nucl. Acids Res.* 27 : 1409—1420.
- Wei W., Gilbert N., Ooi S. L., Lawler J. F., Ostertag E. M., Kazazian H. H., Boeke J. D., Moran J. V. 2001. Human L1 retrotransposition: cis preference versus trans complementation. *Mol. Cell. Biol.* 21 : 1429—1439.
- Weinberg M. S., Villeneuve L. M., Ehsani A., Amarzguioui M., Aagaard L., Chen Z. X., Riggs A. D., Rossi J. J., Morris K. V. 2006. The antisense strand of small interfering RNAs directs histone methylation and transcriptional gene silencing in human cells. *RNA*. 12 : 256—262.
- Whitelaw E., Martin D. I. 2001. Retrotransposons as epigenetic mediators of phenotypic variation in mammals. *Nat. Genet.* 27 : 361—365.
- Wong L. H., Choo K. H. 2004. Evolutionary dynamics of transposable elements at the centromere. *Trends Genet.* 20 : 611—616.
- Wurtele H., Gusew N., Lussier R., Chartrand P. 2005. Characterization of *in vivo* recombination activities in the mouse embryo. *Mol. Genet. Genomics*. 273 : 252—263.
- Yang N., Kazazian H. H., Jr. 2006. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat. Struct. Mol. Biol.* 13 : 763—771.
- Yang N., Zhang L., Zhang Y., Kazazian H. H., Jr. 2003. An important role for RUNX3 in human L1 transcription and retrotransposition. *Nucl. Acids Res.* 31 : 4929—4940.
- Yang Z., Boffelli D., Boonmark N., Schwartz K., Lawn R. 1998. Apolipoprotein(a) gene enhancer resides within a LINE element. *J. Biol. Chem.* 273 : 891—897.
- Yoder J. A., Walsh C. P., Bestor T. H. 1997. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* 13 : 335—340.
- Yu F., Zingler N., Schumann G., Stratling W. H. 2001. Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Aly transcription. *Nucl. Acids Res.* 29 : 4493—4501.
- Zingler N., Willhoeft U., Brose H. P., Schoder V., Jahns T., Hanschmann K. M., Morrish T. A., Lower J., Schumann G. G. 2005. Analysis of 5' junctions of human LINE-1 and Alu retrotransposons suggests an alternative model for 5'-end attachment requiring microhomology-mediated end-joining. *Genome Res.* 15 : 780—789.

REGULATION OF MAMMALIAN LINE1 RETROTRANSPOSONES TRANSCRIPTION

A. V. Fedorov

Institute of Cytology RAS, St. Petersburg;
e-mail: an_tn@mail.ru

In-depth investigation of genomes of higher eukaryotes and, especially, results of successful genome sequencing projects bring us over that not only unique DNA sequences but also multiple families of repeats possess certain functions and are not just «junk» DNA. LINE1 retrotransposons are mobile elements that account for about 1/5 of mammalian genomes and participate in the retrotransposition of cellular DNA, organization of chromatin structure, genome rearrangements, and regulation of gene transcription. LINE1 elements are also involved in recombination and repair processes. In this review, LINE1 retrotransposons structural organization, characteristics of action and mechanisms of tissue-specific transcription regulation are discussed.

Key words: LINE1 retrotransposons, transcription regulation, epigenetics, RNA silencing.
