



УДК 577.215

СМЕНА МАТРИЦ ПРИ LINE-ОПОСРЕДОВАННОЙ ОБРАТНОЙ ТРАНСКРИПЦИИ – НАИБОЛЕЕ ВЕРОЯТНЫЙ МЕХАНИЗМ ФОРМИРОВАНИЯ ДВОЙНЫХ И ТРОЙНЫХ ХИМЕРНЫХ РЕТРОЭЛЕМЕНТОВ МЛЕКОПИТАЮЩИХ

© 2005 г. Е. В. Гогвадзе, А. А. Буздин[#], Е. Д. СвердловИнститут биоорганической химии им. акад. М.М. Шемякина и Ю.А. Овчинникова РАН,
117997 ГСП, Москва, В-437, ул. Миклухо-Маклая, 16/10

Поступила в редакцию 06.02.2004 г. Принята к печати 10.06.2004 г.

Показано, что механизм “перетасовывания” транскрибируемых компонентов генома, включающий в себя смену матрицы при обратной транскрипции РНК с помощью энзиматического аппарата ретроэлементов L1, является широко распространенным у млекопитающих. Частота появления образующихся в результате химерных ретроэлементов примерно вдвое выше в геномах грызунов, чем в ДНК приматов. Кроме того, приведены доказательства прохождения *in vivo* не только единичных, но и двойных случаев смены матрицы, что приводит к объединению копий трех различных транскриптов. Многие идентифицированные химеры транскрибируются в тканях млекопитающих.

Ключевые слова: ретроэлементы, транспозоны, обратная транскрипция, LINE, L1.

ВВЕДЕНИЕ

Относящиеся к LINE ретротранспозоны L1 занимают около 20% последовательности геномной ДНК млекопитающих [1, 2]. Большинство элементов L1 представлено неспособными к ретропозиции копиями, укороченными с 5'-конца [3–6], и лишь небольшое количество (80–100) L1-элементов человека транспозиционно активно [7]. Известно, что транспозиция L1 проходит в несколько стадий, включающих транскрипцию ретроэлемента РНК-полимеразой II, обратную транскрипцию образовавшейся РНК с помощью кодируемой L1-элементом РНК-зависимой ДНК-полимеразы (обратной транскриптазы), а также интеграцию получившейся кДНК в новый участок генома [6]. Элемент L1 содержит две открытые рамки считывания: ORF1, которая кодирует РНК-связывающий белок с *M* 40 кДа [8, 9], и ORF2 для обратной транскриптазы – эндонуклеазы [4]. Благодаря эффекту так называемого “цис-предпочтения”, энзиматический аппарат ретропозиционно-компетентных элементов L1 предпочтительно перемещает копии своих собственных

мРНК [10]. Тем не менее L1 также способны перемещать другие последовательности, в основном ретропозоны Alu [11], а также кДНК других типов клеточных РНК, формируя при этом псевдогены [12]. Кроме того, недавно мы показали возможность формирования транспозиционным аппаратом L1 химерных ретроэлементов в процессе смены матрицы при обратной транскрипции [13].

В настоящей статье описаны результаты скрининга баз данных геномов млекопитающих *Homo sapiens*, *Mus musculus* и *Rattus norvegicus*, земноводного *Xenopus laevis*, рыб *Danio rerio* и *Takifugu rubripes*, беспозвоночных *Drosophila melanogaster*, *Anopheles gambiae* и *Caenorhabditis elegans* на присутствие в них химерных псевдогенов, созданных малыми ядерными РНК, – U6, U5 и U3; рибосомной 5S РНК, а также 7SL РНК, которые формируют 5'-концевые части химер, найденных нами ранее в ДНК человека [13].

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Идентифицировано 1800 последовательностей псевдогенов вышеописанных РНК, мало- или среднедивергировавших (<10%) от соответственных консенсусных последовательностей. Эти последовательности были проанализированы на предмет наличия среди них химерных ретрогенов по следующим критериям: 1) 5'- и 3'-концевые части химер должны быть соединены непосредственно, 2) располагаться в одинаковой ориентации;

Сокращения: LINE – длинные диспергированные геномные повторы, от англ. long interspersed nuclear elements; SINE – короткие диспергированные геномные повторы, от англ. short interspersed nuclear elements; EST – маркер экспрессирующейся последовательности, от англ. expressed sequence tag; мЯРНК – малая ядерная РНК; ORF – открытая рамка считывания, от англ. open reading frame.

[#] Автор для переписки (тел.: (095) 330 65 74; факс: (095) 330 65 38; эл. почта: anton@humgen.siobc.ras.ru).

3) химеры должны быть фланкированы прямыми повторами длиной 8–25 п.о., а также 4) нести poly(A)-последовательность на 3'-конце (см. рис. 1). Фланкирование прямыми повторами отражает особенности транспозиционного аппарата LINE отделять созданные им вставки короткими tandemными повторами [14].

Обладающие этими структурными особенностями химеры не были обнаружены в геномах беспозвоночных, рыб и амфибий, в то же время все исследованные геномы млекопитающих (т.е. геномы человека, мыши и крысы) содержали подобные химерные ретрогены. Все найденные химеры содержали вблизи 5'-конца предпочтительно узнаваемую эндонуклеазой L1 последовательность гексонуклеотида T₂A₄ либо ее производные с одноили двунуклеотидными заменами. В соответствии с ранее полученными нами данными [13], 5'-части химер являлись ДНК-копиями РНК, имеющих ядерную локализацию, тогда как 3'-части были сформированы цитоплазматическими РНК: мРНК клеточных генов или ретроэлементов. Обе части химер близки по эволюционному возрасту: чем более молодой (т.е. менее дивергировавшей от консенсусной последовательности) является 5'-часть, тем моложе соответствующая ей 3'-часть химеры, как показано нами ранее для химер человека U6-L1 [15]. Эти данные говорят о том, что 3'- и 5'-концевые части химер внедрялись в геномную ДНК од-



Рис. 1. Схематическое изображение двойных химерных ретрогенов, идентифицированных в геномных базах данных. Химеры фланкированы прямыми повторами, содержат 5'- и 3'-концевые части, мотив ТТАААА, последовательность poly(A). Типы 5'- и 3'-концевых частей перечислены в табл. 1. Коды доступа химер в GenBank приведены в табл. 2.

новременно, в составе единой последовательности. Интересно, что 5'-концевые участки представляли собой полноразмерные последовательности псевдогенов (за исключением фрагментов U3 и U5, всегда укороченных на 145 и 41 нт соответственно), тогда как 3'-концевые части были, как правило, укорочены с 5'-конца по сравнению с консенсусной последовательностью. Такие точки “обрыва” 3'-концевых последовательностей располагались случайным образом и не содержали гомологичных участков между 3'- и 5'-частями химер.

Всего было идентифицировано 82, 116 и 66 химер в геномах человека, мыши и крысы соответственно (см. табл. 1 и 2). 3'-Концевые части химер являлись, как правило, ретроэлементами L1, но в 20, 12 и 18% случаев для ДНК человека, мыши и

Таблица 1. Сопоставление 5'- и 3'-концевых частей химер с известными последовательностями РНК (из GenBank, коды доступа см. табл. 2)

Название транскриптов*	Число представителей (и их доля, %)**			Функция РНК в клетке	Локализация РНК в клетке
	Человек	Мышь	Крыса		
5'-Концевые последовательности химер					
U6	66 (82)	111 (95)	61 (92)	мяРНК, сплайсинг	Ядро
U5	1 (1)	0	0	мяРНК, сплайсинг	Ядро
U3	8 (10)	2 (2)	1 (1)	мяРНК, сплайсинг	Ядро
5S рРНК	3 (4)	3 (3)	4 (6)	Рибосомная РНК	Цитоплазма, ядро
7SL РНК	1 (1)	0	0	Сортинг белков	Цитоплазма, ядро
Alu	2 (2)	0	0	Эгоистическая РНК	Цитоплазма, ядро
3'-Концевые последовательности химер					
L1	65 (80)	103 (88)	54 (82)	Эгоистическая мРНК	Цитоплазма, ядро
мРНК	9 (11)	9 (8)	5 (8)	мРНК гена	Цитоплазма
Alu	7 (9)	0	0	Эгоистическая РНК	Цитоплазма, ядро
B1_Mm	0	2 (2)	0	То же	То же
B2_Mm1	0	1 (1)	0	»	»
B2_Mm2	0	1 (1)	1 (1)	»	»
ID_Rn	0	0	5 (8)	»	»
4,5S РНК	0	0	1 (1)	»	»

* Названия клеточных транскриптов, соответствующие данным частям химер.

** Число представителей и их доля (%) среди идентифицированных химерных ретротранскриптов.

Таблица 2. Обнаруженные химерные ретрогены геномов мыши и крысы

Код доступа в GenBank	5'-часть химеры*	3'-часть химеры**	Код доступа в GenBank	5'-часть химеры*	3'-часть химеры**
Геномная ДНК мыши (<i>Mus musculus</i>)			CAAA01197720	U6	L1_MM
AL672082	U6	L1_MM	CAAA01138310	U6	L1_RN
CAAA01167315	U6	L1_MM	CAAA01047060	U6	L1_MM
CAAA01186746	U6	L1_MM	CAAA01189107	U6	L1_MM
AL772324	U6	L1_MM	CAAA01178342	U6	L1Md-F_3_3
AC121969	U6	L1_MM	CAAA01129371	U6	L1_MM
CAAA01017881	U6	L1_MM	CAAA01187494	U6	L1Md-F_3_3
CAAA01210825	U6	L1_MM	CAAA01031739	U6	Копия мРНК
CAAA01154841	U6	L1_MM	CAAA01005305	U6	L1Md-T_3
CAAA01211991	U6	L1Md-T_3	AC098875	U6	L1_MM
CAAA01087342	U6	L1Md-T_3	CAAA01105106	U6	L1_MM
AL603843	U6	Копия мРНК	CAAA01019438	U6	Lx_3
CAAA01140325	U6	L1_MM	CAAA01208853	U6	L1_MM
CAAA01041190	U6	L1Md-T_3	CAAA01187801	U6	L1Md-F_3_3
AC087891	U6	L1_MM	AC116657	U6	L1Md-T_3
AC083817	U6	L1_MM	AC125092	U6	L1_MM
CAAA01058471	U6	L1_MM	AC121989	U6	L1_MM
CAAA01019220	U6	Копия мРНК	CAAA01051515	U6	L1Md-F_3_3
CAAA01082103	U6	L1_MM	CAAA01103542	U6	L1Md-F_2_3
CAAA01163621	U6	L1_MM	CAAA01008640	U6	L1_MM
AL671671	U6	L1_MM	CAAA01079704	U6	L1_MM
CAAA01139182	U6	L1Md-T_3	CAAA01020902	U6	Копия мРНК
CAAA01090187	U6	L1Md-F2_3	AL646042	U6	L1Md-F_2_3
AL845292	U6	L1_MM	CAAA01154323	U6	L1Md-F_2_3
CAAA01186766	U6	L1Md-F2_3	AL731673	U6	L1Md-F_3_3
CAAA01131310	U6	L1Md-T_3	CAAA01011300	U6	L1_MM
CAAA01180152	U6	L1Md-T_3	CAAA01041034	U6	L1_MM
AL731658	U6	B2_Mm1	CAAA01129898	U6	L1_MM
CAAA01111751	U6	L1_MM	CAAA01013622	U6	L1Md-F_2_3
CAAA01192525	U6	Копия мРНК	AL683829	U6	L1Md-F_2_3
AC124175	U6	L1_MM	CAAA01022952	U6	L1Md-F_2_3
AC124175	U6	L1_MM	CAAA01169542	U6	L1Md-F_2_3
CAAA01180605	U6	L1_MM	CAAA01201318	U6	Копия мРНК
CAAA01003842	U6	L1_MM	CAAA01180476	U6	L1
AL606928	U6	L1_MM	AC124686	U6	L1Md-F_3_3
CAAA01179727	U6	L1_MM	CAAA01134532	U6	L1
AL672051	U6	L1_MM	CAAA01137399	U6	L1_MM
CAAA01078962	U6	L1_MM	CAAA01190370	U6	L1Md-F_2_3
AC090652	U6	L1_MM	AC104327	U6	L1Md-F_2_3
AC024950	U6	L1_MM	CAAA01029067	U6	L1Md-F_2_3
AL844856	U6	L1_MM	AL844896	U6	Копия мРНК
AL929572	U6	L1_MM	CAAA01085936	U6	L1Md-F_3_3
AC138605	U6	L1_MM	CAAA01147472	U6	L1_MM
CAAA01027369	U6	L1Md-F2_3	CAAA01167491	U6	Lx8
CAAA01099480	U6	L1_MM	AC125329	U6	L1_MM
AL590992	U6	L1_MM	CAAA01179761	U6	L1Md-F_2_3
CAAA01005091	U6	L1Md-T_3	AL662895	U6	L1Md-A_3_3

Таблица 2. Окончание

Код доступа в GenBank	5'-часть химеры*	3'-часть химеры**	Код доступа в GenBank	5'-часть химеры*	3'-часть химеры**
CAAA01173279	U6	L1Md-F_2_3	RNOR01147125	U6	L1_RN
CAAA01060255	U6	L1Md-F_2_3	RNOR01128344	U6	RNHAI
CAAA01048872	U6	L1Md-F_2_3	RNOR01093990	U6	L1_RN
CAAA01160526	U6	L1Md-F_2_3	RNOR01089362	U6	L1_RN
CAAA01149884	U6	L1Md-F_2_3	RNOR01058614	U6	Копия мРНК
CAAA01125758	U6	L1Md-F_2_3	RNOR01013966	U6	L1_RN
AC121780	U6	L1Md-F_2_3	RNOR01011948	U6	Копия мРНК
CAAA01007460	U6	L1Md-F_2_3	RNOR01003594	U6	L1_RN
CAAA01149346	U6	L1_MM	RNOR01112589	U6	RNHAI
AL731728	U6	L1_MM	RNOR01071913	U6	Копия мРНК
CAAA01055201	U6	Lx_3	RNOR01009126	U6	L1_RN
AL669884	U6	L1Md-F_2_3	RNOR01148985	U6	L1_RN
AL772234	U6	Lx_3	RNOR01070463	U6	L1_RN
CAAA01064203	U6	L1Md-F_2_3	RNOR01063720	U6	L1_RN
AL691484	U6	Копия мРНК	RNOR01059849	U6	L1_RN
AC121949	U6	L1Md-F_2_3	RNOR01036515	U6	L1_RN
CAAA01194319	U6	Копия мРНК	RNOR01019844	U6	L1_RN
CAAA01110941	U6	L1_MM	RNOR01119709	U6	L1_RN
AL604065	5S рРНК	B2_Mm2	RNOR01078137	U6	L1_RN
AC117240	5S рРНК	B1_Mm	RNOR01023020	U6	L1_RN
AC122879	5S рРНК	L1Md_F_3_3	RNOR01137623	U6	L1_RN
CAAA01006782	U3	L1_MM	RNOR01116297	U6	L1_RN
AL928616	U3	Lx_3	RNOR01109371	U6	L1_RN
Геномная ДНК крысы (<i>Rattus norvegicus</i>)			RNOR01108886	U6	ID_RN (SINE)
RNOR01146973	U6	L1_RN	RNOR01104844	U6	L1_RN
RNOR01065336	U6	L1_RN	RNOR01081522	U6	L1_RN
RNOR01120537	U6	L1_RN	RNOR01078726	U6	ID_RN (SINE)
RNOR01034123	U6	L1_RN	RNOR01041436	U6	L1_RN
RNOR01052629	U6	L1_RN	RNOR01035294	U6	L1_RN
RNOR01067149	U6	L1_RN	RNOR01021032	U6	L1_RN
RNOR01106394	U6	L1_RN	RNOR01002690	U6	L1_RN
RNOR01101953	U6	Копия мРНК	RNOR01073427	U6	L1_RN
RNOR01081579	U6	L1_RN	RNOR01052638	U6	L1Md_F
RNOR01067118	U6	L1_RN	RNOR01008426	U6	L1_RN
RNOR01057080	U6	L1_RN	RNOR01121805	U6	L1_RN
RNOR01043599	U6	L1_RN	RNOR01094232	U6	L1_RN
RNOR01023717	U6	L1_RN	RNOR01007047	U6	RNHAI
RNOR01090287	U6	L1_RN	RNOR01069375	U6	L1_RN
RNOR01089352	U6	L1_RN	RNOR01062956	U6	L1_RN
RNOR01058624	U6	L1_RN	RNOR01036261	U6	L1_RN
RNOR01052367	U6	Копия мРНК	RNOR01017237	U3	ID_Rn (SINE)
RNOR01046897	U6	L1_RN	RNOR01089548	5S рРНК	ID_Rn (SINE)
RNOR01018921	U6	L1_RN	RNOR01110600	5S рРНК	B2_Mm2 (SINE)
RNOR01012737	U6	L1_RN	RNOR01027748	5S рРНК	ID_Rn (SINE)
RNOR01002634	U6	4,5S рРНК	RNOR01013061	5S рРНК	B2_Mm2 (SINE)

* Классификация 5'- и 3'-концевой части химеры.

** Химерные ретротранспозоны генома человека были опубликованы нами ранее [15].

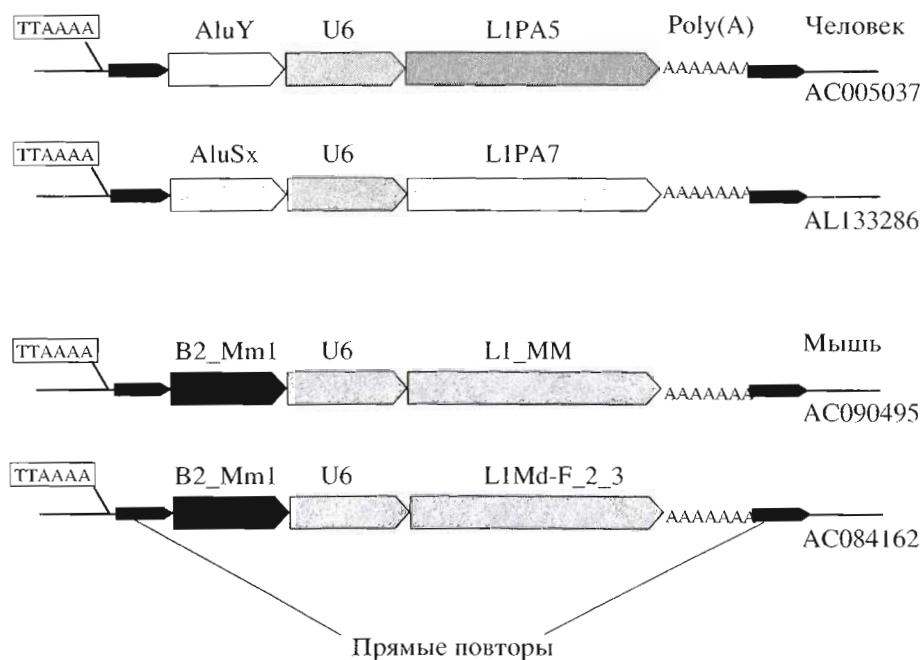


Рис. 2. Схематическое изображение обнаруженных тройных химер.

крысы соответственно, это были либо псевдогены различных клеточных мРНК (11, 8 и 8% соответственно), либо относящиеся к SINE ретропозоны (9, 4 и 10% соответственно). В то же время 5'-частями химер в геномах человека, мыши и крысы являются псевдогены мРНК – U6 (82, 95 и 92% соответственно), U3 (10, 2 и 1% соответственно) и U5 (1, 0 и 0% соответственно); либо также последовательности, относящиеся к SINE (2, 0 и 0% соответственно), 7SL РНК (1, 0 и 0% соответственно) или 5S РНК (4, 3 и 6% соответственно).

Некоторое количество химер в геномах мыши и крысы (107 и 66 соответственно), по-видимому, недооценено, поскольку 3'-концевые части предполагаемых химер были часто прерваны брешами в сборке геномной последовательности, что делало невозможным обнаружение прямых повторов в этих случаях (23 и 33 случая для геномов *M. musculus* и *R. norvegicus* соответственно; нет случаев для *H. sapiens*). Учитывая вышесказанное, количество химерных псевдогенов в геномах мыши, крысы и человека может составлять, таким образом, 139, 99 и 82 соответственно.

Интересно, что доля включенных в химеры псевдогенов U6 значительно выше у грызунов, нежели у *H. sapiens*. 33% всех псевдогенов U6 человека являются частями химер в сравнении с 69% у мыши. Для генома крысы это значение может быть найдено лишь приблизительно, вместе с предполагаемыми химерами оно составляет 64%. Учитывая, что практически все псевдогены U6 (как химеры, так и не химеры), по-видимому, были созданы ретропозиционным аппаратом L1,

можно предположить, что комплексы L1 грызунов сменяют матрицу в ходе обратной транскрипции вдвое чаще, чем L1 человека.

Химеры не были найдены в ДНК амфибий и рыб. Таким образом, механизм образования химер появился, по крайней мере, 75 млн лет назад, до эволюционного расхождения предковых линий приматов и грызунов [16], и остается активным и по сей день [13, 15]. Для более точного датирования возникновения механизма химеризации необходимо определение последовательности нуклеотидов ДНК геномов других позвоночных.

Необходимо заметить, что смена матриц не является уникальным свойством обратной транскриптазы L1. Подобные эффекты были описаны и для другого относящегося к LINE ретропозона R2 [17], а также для рекомбинации ретровирусных геномов [18, 19].

Важным результатом, полученным в ходе выполнения настоящей работы, является обнаружение химер, состоящих из трех компонентов (тройные химеры). Два таких химерных элемента были найдены в ДНК мыши, еще два – в геноме человека (см. рис. 2). Структурная организация всех четырех тройных химерных элементов одинакова: они фланкированы прямыми повторами, содержат на 5'-конце последовательность T₂A₄ и несут poly(A)-хвост на 3'-конце. 5'-Концевые части тройных химер сформированы представителями SINE: AluY или AluSx – для ДНК человека и B2_Mm1 – для ДНК мыши. В середине тройные химеры всегда содержали копии U6 мРНК, а 3'-концевые части представляли собой последователь-

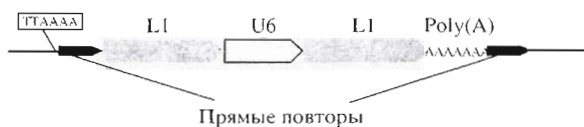


Рис. 3. Схематическое изображение геномных последовательностей типа L1-U6-L1. Многие копии U6 мРНК внедрены в последовательности L1 в прямой ориентации и не фланкированы прямыми повторами.

ности LINE: L1PA5 или L1PA7 у человека и L1_MM или L1_Md-F_2_3 у мыши. Во всех проанализированных геномах млекопитающих неко-

торые копии U6 были внедрены в последовательность L1 в прямой ориентации необычным способом, в отличие от процессированных псевдогенов они не были фланкированы прямыми повторами и не содержали poly(A)-хвостов (рис. 3). Дивергенции U6- и L1-частей таких элементов от соответствующих консенсусных последовательностей линейно коррелируют. Обнаружение таких ретро-транскриптов говорит о том, что *in vivo* при обратной транскрипции происходят не только однократные, но и двукратные смены матрицы L1 (рис. 4).

Полученные в работе результаты показывают, что LINE-опосредованные РНК-РНК-реком-

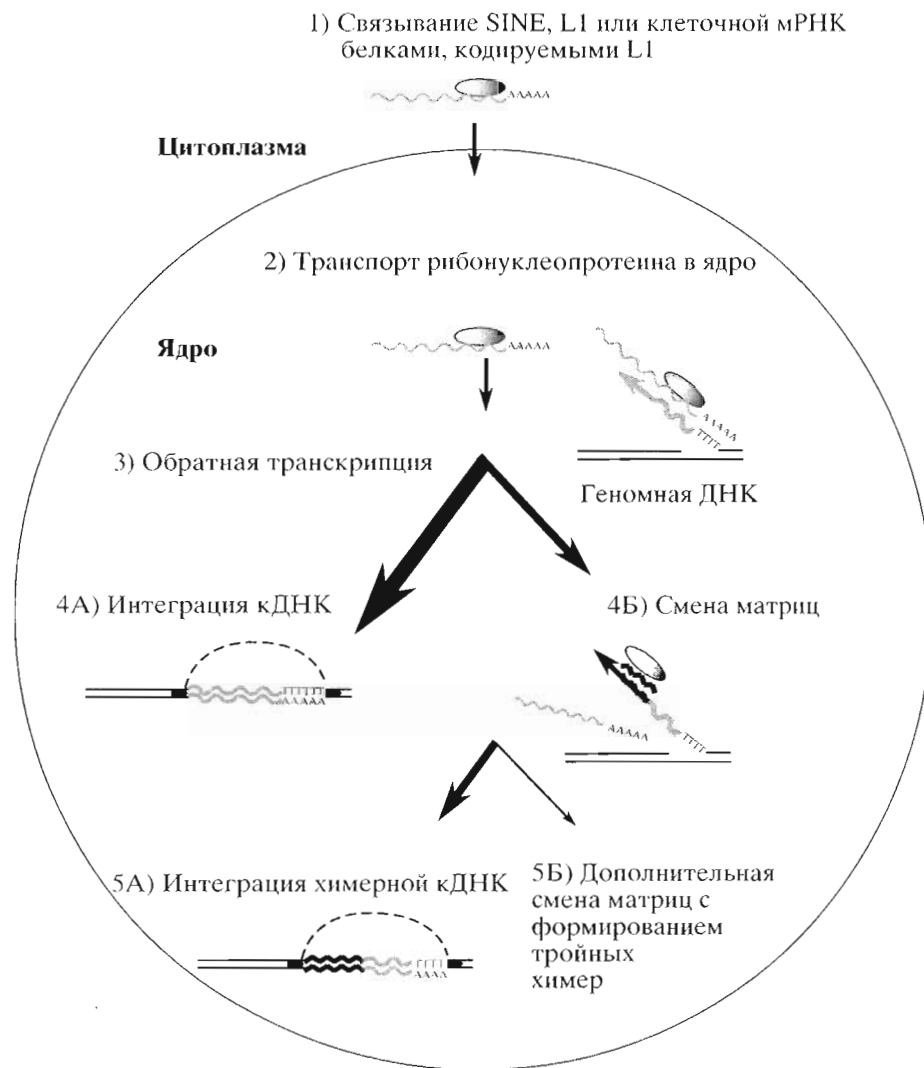


Рис. 4. Предполагаемый механизм формирования двойных и тройных химер. Стадия 1: преинтеграционный комплекс L1 связывается в цитоплазме с мРНК клеточного гена, L1 или Alu. Стадия 2: сформированный рибонуклеопротеин транспортируется в ядро. Стадия 3: обратная транскрипция связанной с комплексом мРНК, при которой в качестве праймера выступает образованный в результате одноцепочечного разрыва фрагмент ДНК внутри последовательности TTTTAA. Стадия 4А: успешная интеграция синтезированной копии кДНК в состав геномной ДНК. Стадия 4Б: смена матрицы на другую РНК при обратной транскрипции. Стадия 5А: интеграция образованной двойной химеры в геномную ДНК. Стадия 5Б: вторая смена матрицы при обратной транскрипции с последующей интеграцией в геном и репарацией ДНК обуславливает формирование тройных химер, фланкированных прямыми повторами и несущих poly(A) на 3'-конце.

Таблица 3. Химерные ретрогены, транскрибирующиеся в тканях млекопитающих

Тип химеры	Код доступа в GenBank	EST/мРНК, код доступа в GenBank	Ткань млекопитающих*	Уровень экспрессии (число транскриптов на клетку)**
Химеры генома человека				
U6-L1PA7	AL121883	BQ720168	Симпатический ствол	–
		AI095257	Стареющие фибробласты	–
U6-L1PA3	AC010894	BQ447264	Остеоартритный хрящ	–
U6-AluY	AC004128	AA581502	Раковая опухоль яичников	–
U6-мРНК	AL591050	AK056682	Моноядерные клетки периферической крови	–
U6-мРНК	AL354668	CB333830	Мозг зародыша: таламус	~10
			Мозг зародыша: гиппокамп	~10
			Паренхима яичка	~10
			Семинома	~10
U6-мРНК	AC021037	CB333829	Мозг зародыша: таламус	~10
			Мозг зародыша: гиппокамп	~200
			Паренхима яичка	~100
			Семинома	0
Химеры генома мыши				
U6-мРНК	CAAA01031739	AV086247	Язык	–
		AV118564	10-суточный зародыш	–
		AV298364	8-суточный зародыш	–
		BB106464	Мочевой пузырь	–
		BM199787	7.5-суточный зародыш	–
		BM199812	То же	–
		BM201119	»	–
		BM201128	»	–
U6-мРНК	CAAA01201318	BI732400	Сетчатка	–
		BY165293	Костный мозг	–
		BY171466	То же	–
U6-L1	CAAA01134532	BB240924	Тимус	–
		AK042346	То же	–
Химеры генома крысы				
5S – ID_RN	RNOR01027748	AW920107	Гомогенизированные ткани	–

* Ткани млекопитающих, в которых обнаружена экспрессия химер.

** Указано число химерных транскриптов на клетку, оцененное с помощью метода ОТ-ПЦР [13]; прочерк – отсутствие данных.

бинации эволюционно консервативны у млекопитающих. Они в конечном итоге приводят к объединению двух или трех транскрибируемых компонентов генома и к интеграции образованных химер в новые участки ДНК. Интересно, что многие такие химеры экспрессируются у различных видов млекопитающих и некоторые из них – тканеспецифично (табл. 3). Это позволяет предположить, что химерные ретротранскрипты могут быть вовлечены в функционирование клетки.

Процесс формирования некоторых семейств L1 также, возможно, включал РНК-РНК-рекомбинацию со сменой матрицы после обратной транскрипции большей части мРНК L1, что приводило к объединению 3'-концевой части L1 с совершенно новыми последовательностями. Так, в частности, 5'-нетранслируемые области и первая треть ORF1 у

семейств L1 человека, мыши, крысы и кролика не гомологичны друг другу [6]. Эти семейства L1-элементов, возможно, были сформированы с привлечением описанного выше механизма. Наконец, механизм формирования химер показывает, что при функционировании клетки между ее компонентами происходит множество сложных взаимодействий, которые могут относиться либо к широко обсуждаемому клеточному “шуму”, либо выполнять важные функции.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Анализ последовательностей ДНК. Использовали консенсусные последовательности псевдогенов из баз данных RepBase Update (<http://www.gir-inst.org/server/RepBase/>) и GenBank. Поиск полнораз-

мерных генов и псевдогенов 5S рРНК, 7SL РНК, U3, U5 и U6 мРНК в геномах млекопитающих и их картирование проводили с помощью программы BLAT (<http://genome.ucsc.edu/cgi-bin/hgBLAT>) и сервера UCSC genome browser. Для анализа ДНК человека, мыши и крысы использовали соответственно следующие даты сборки геномов: апрель, февраль и январь 2003 г. Фланкирующие области псевдогенов исследовали с помощью программы RepeatMasker (<http://ftp.genome.washington.edu/cgi-bin/RepeatMasker>). Прямые повторы и мотивы ТТАААА детектировали при визуальном анализе последовательностей, фланкирующих псевдогены. Поиск гомологий в GenBank для видов, не относящихся к млекопитающим, проводили с использованием сервера BLAST at NCBI (<http://www.ncbi.nlm.nih.gov/BLAST>) [20]. Последовательности L1 и Alu, располагающиеся по соседству с псевдогенами, были отнесены к тем или иным семействам в соответствии с номенклатурой RepBase. Для множественного выравнивания последовательностей использовались программы Clustal X и Clustal W [21].

Авторы выражают искреннюю благодарность Ю.Б. Лебедеву, Т.В. Виноградовой и Л.Г. Николаеву за плодотворную дискуссию и Б.О. Глотову (ИМГ РАН) за ценные комментарии при написании статьи.

Работа была поддержана грантами РФФИ № 02-04-48614-а, 04-04-48564-а, а также Программой РАН по физико-химической биологии.

СПИСОК ЛИТЕРАТУРЫ

1. Venter J.C., Adams M.D., Myers E.W., Li P.W., Mural R.J., et al. // *Science*. 2001. V. 291. P. 1304–1351.
2. International Human Genome Sequencing Consortium // *Nature*. 2001. V. 409. P. 860–921.
3. Buzdin A., Ustyugova S., Gogvadze E., Lebedev Y., Hunsman G., Sverdlov E. // *Hum. Genet.* 2003. V. 112. P. 527–533.
4. Kazazian H.H. // *Science*. 2000. V. 289. P. 1152–1153.
5. Ovchinnikov I., Troxel A.B., Swergold G.D. // *Genome Res.* 2001. V. 11. P. 2050–2058.
6. Furano A.V. // *Prog. Nucleic Acid Res. Mol. Biol.* 2000. V. 64. P. 255–294.
7. Brouha B., Schustak J., Badge R.M., Lutz-Prigge S., Farley A.H., Moran J.V., Kazazian H.H. // *Proc. Natl. Acad. Sci. USA*. 2003. V. 100. P. 5280–5285.
8. Hohjoh H., Singer M.F. // *EMBO J.* 1996. V. 15. P. 630–639.
9. Kolosha V.O., Martin S.L. // *J. Biol. Chem.* 2003. V. 278. P. 8112–8117.
10. Wei W., Gilbert N., Ooi S.L., Lawler J.F., Ostertag E.M., Kazazian H.H., Boeke J.D., Moran J.V. // *Mol. Cell. Biol.* 2001. V. 21. P. 1429–1439.
11. Dewannieux M., Esnault C., Heidmann T. // *Nat. Genet.* 2003. V. 35. P. 41–48.
12. Esnault C., Maestre J., Heidmann T. // *Nat. Genet.* 2000. V. 24. P. 363–367.
13. Buzdin A., Gogvadze E., Kovalskaya E., Volchkov P., Ustyugova S., Illarionova A., Fushan A., Vinogradova T., Sverdlov E. // *Nucl. Acids Res.* 2003. V. 31. P. 4385–4390.
14. Jurka J. // *Proc. Natl. Acad. Sci. USA*. 1997. V. 94. P. 1872–1877.
15. Buzdin A., Ustyugova S., Gogvadze E., Vinogradova T., Lebedev Y., Sverdlov E. // *Genomics*. 2002. V. 80. P. 402–406.
16. Mouse Genome Sequencing Consortium // *Nature*. 2002. V. 420. P. 520–562.
17. Bibillo A., Eickbush T.H. // *J. Mol. Biol.* 2002. V. 316. P. 459–473.
18. Jamain S., Girondot M., Leroy P., Clergue M., Quach H., Fellous M., Bourgeron T. // *Genomics*. 2001. V. 78. P. 38–45.
19. Negroni M., Buc H. // *Annu. Rev. Genet.* 2001. V. 35. P. 275–302.
20. Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. // *J. Mol. Biol.* 1990. V. 215. P. 403–410.
21. Thompson J.D., Higgins D.G., Gibson T.J. // *Nucl. Acids Res.* 1994. V. 22. P. 4673–4680.

Multiple Template Switches on LINE-Directed Reverse Transcription: The Most Probable Formation Mechanism for the Double and Triple Chimeric Retroelements in Mammals

E. V. Gogvadze, A. A. Buzdin[#], and E. D. Sverdlov

[#] Phone: +7 (095) 330-6574; fax: +7 (095) 330-6538; e-mail: anton@humgen.siohc.ras.ru
Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences,
ul. Miklukho-Maklaya 16/10, Moscow, 117997 Russia

It was shown that the shuffling mechanism for transcribed genome components, which involves a template switch on the RNA reverse transcription using the L1 retroelement enzymatic machinery, is common in mammals. The occurrence frequency of the resulting chimeric retroelements in the genomes of rodents is twice as high as in the DNA of primates. Moreover, we proved that not only single but also double switches may occur *in vivo*, which result in the fusion of copies of three different transcripts. Many of the identified chimeras are transcribed in mammals. The English version of the paper: *Russian Journal of Bioorganic Chemistry*, 2005, vol. 31, no. 1; see also <http://www.maik.ru>.

Key words: LINE, L1, retroelements, reverse transcription, transposons