

# ЭКСПЕРИМЕНТАЛЬНАЯ ИДЕНТИФИКАЦИЯ И ФУНКЦИОНАЛЬНЫЙ АНАЛИЗ РЕТРОЭЛЕМЕНТОВ, СПЕЦИФИЧНЫХ ДЛЯ ГЕНОМА ЧЕЛОВЕКА

*А.А. Буздин, Т. В.Виноградова\*, Ю.Б. Лебедев, Е. Д.Свердлов*

*Институт биоорганической химии им. М.М. Шемякина и Ю.А. Овчинникова РАН, 117997, Москва, ГСП, ул. Миклухо-Маклая, 16/10*

## 1. Введение

После расхождения линий предков человека и шимпанзе гоминиды, то есть представители предковой линии человека, прошли через много эволюционных этапов прежде чем наконец стать существующим сегодня видом *Homo sapiens* [1]. Генетическая история этих явлений отражена в геноме современного человека, где некоторые последовательности ДНК можно проследить до очень древних предков, тогда как другие появились в более поздние времена. В геноме приматов на протяжении последних 65 миллионов лет шло непрерывное распространение мобильных элементов, в первую очередь ретроэлементов (РЭ). Многие из них встречаются только в геноме человека и могли каким-то образом участвовать в становлении человека как вида.

Известно, что РЭ активно реформируют геном, вызывая геномные перестройки, создавая новые гены и модифицируя регуляторный механизм существующих генов, таким образом активно участвуя в эволюции [2-4]. В частности, так называемые длинные концевые повторы (long terminal repeat, LTR) эндогенных ретровирусов человека, десятки и тысячи которых рассеяны по геному, содержат множество потенциальных регуляторных элементов. Вновь интегрированные LTR могли играть роль альтернативных промоторов, энхансеров, входить в состав так называемых локусов контролирующей области, а также могли являться источником некоторых других регуляторных сигналов для окружающих генов (см. обзоры [2, 3, 5]). Каждый индивидуальный LTR может играть определённую роль в сети регуляторных взаимодействий клетки, причём эта роль может отличаться от ролей других индивидуальных LTR. Таким образом, не существует некоей универсальной роли для ретроэлементов и для понимания их функций необходимо рассматривать каждый ретроэлемент по отдельности в том геномном контексте, в котором он был зафиксирован эволюцией.

---

\* Автор для переписки (тел.: (095) 330-6547; факс: (095) 330-6538; e-mail: [tv@humgen.siohc.ras.ru](mailto:tv@humgen.siohc.ras.ru))

Многие ретроэлементы обеспечивают транскрипцию своих собственных РНК, некоторые из которых, вероятно, важны для функционирования клетки. Например, транскрипты LINE (long interspersed nuclear element) кодируют обратную транскриптазу и эндонуклеазу, которые, возможно, участвуют в ретротранспозициях LINE и SINE (short interspersed nuclear element), а также в образовании псевдогенов [2]. Кроме того, недавно нами был обнаружен новый механизм образования химерных РЭ, опосредованный ретроэлементом L1, представителем LINE-семейства [6, 7].

Отметим также, что в последние годы открыт новый тип мощных регуляторов генной активности с помощью так называемых некодирующих РНК (ncRNA), в синтезе которых могут принимать участие РЭ.

Поскольку предполагается, что РЭ могут играть роль в видообразовании, то идентификация специфичных для генома человека интеграций РЭ вблизи человеческих генов, особенно тех, которые принимают участие в эмбриональном развитии, могла бы открыть таких представителей РЭ, которые ответственны за возникновение фенотипических различий между человеком и человекообразными обезьянами.

Такие РЭ следует искать среди тех, которые внедрялись в геном непосредственно перед разделением двух видов (примерно 6 миллионов лет назад) и на протяжении времени, в течение которого формировался наш вид. Анализ последовательности генома человека выявил три молодых в эволюционном отношении группы РЭ (предположительно интегрировавших в геном менее чем 6 миллионов лет назад), включающих AluYa5, AluYb8, SVA и SINE-R повторы, субпопуляцию повторов L1 (L1-Ta, L1PA1, L1PA2), а также семейство эндогенных ретровирусов человека HERV-K (HML-2), в основном представленное одиночными элементами LTR [8, 9]. Однако вопрос о том, какие из них специфичны для ДНК человека и какие из этих последних могли быть эволюционно существенными, остаётся открытым.

В рамках настоящего обзора мы рассматриваем в основном наши работы, посвященные разработке подходов к выявлению человек-специфичных интеграций РЭ и их функциональной значимости.

## **2. Идентификация специфичных для человека интеграций РЭ. Новые подходы обнаруживают большое количество интеграций ретроэлементов, специфичных для генома человека**

Современные стратегии поиска интеграций РЭ, специфичных для генома человека, могут быть условно разделены на две категории. В первой из них используется поиск РЭ в геноме человека и компьютерный анализ первичных структур ДНК обнаруженных элементов, который по-

зволяет выделять «эволюционно молодые» группы РЭ. Использование первой стратегии дает результаты, имеющие предварительный характер и требует последующего экспериментального подтверждения. Альтернативная стратегия, которую развиваем мы, основана на прямом экспериментальном сравнении геномной ДНК человека с ДНК приматов. Обе стратегии получили мощный импульс в результате определения нуклеотидной последовательности генома человека, которая позволяет проводить точную локализацию изучаемых РЭ в геноме.

Сравнение нуклеотидных последовательностей Alu-повторов, известных на момент проведения исследований, позволило охарактеризовать семейство AluY как самый эволюционно молодой класс SINE элементов генома человека [10–12]. Подробный анализ структуры всех представителей AluY-семейства, найденных в опубликованном геноме человека, привел к выделению отдельных предположительно специфичных для человека групп AluYa5, AluYb8 и AluYc1 [13]. Для подтверждения выдвинутого предположения авторами был проведен крупномасштабный ПЦР-анализ сотен локусов генома, в результате которого обнаружены многочисленные человек-специфические и полиморфные интеграции AluY-элементов [14, 15].

Методы сравнения нуклеотидных последовательностей использовали также для изучения молекулярной эволюции LINE-элементов генома человека. В настоящее время можно считать доказанной специфичность для человека семейства L1-Ta [16–18]. Уточненную консенсусную последовательность L1-Ta применяли для поиска полиморфных инсерций L1 [19, 20].

Аналогичные попытки идентифицировать недавно интегрированные HERV элементы делались преимущественно для наиболее широко представленных в геноме семейств эндогенных ретровирусов, а именно – HERV-H, HERV-E, ERV-9, HERV-W и некоторых других. В результате была показана различная эволюционная динамика распространения эндогенных ретровирусов в геномах приматов. Однако все семейства, за исключением семейства HERV-K, по-видимому, утратили транспозиционную активность ко времени расхождения линий шимпанзе и человека (см. обзоры [3, 5]).

Хотя биоинформационные методы анализа вариабельности структуры и степени родства всех представителей изучаемой группы элементов традиционно считают мощным инструментом выявления недавно сформировавшихся субпопуляций РЭ, для идентификации всех инсерций РЭ, специфичных для генома человека, необходимы экспериментальные подходы полногеномного сравнения ДНК человека и шимпанзе. Без такого исчерпывающего сравнения геномов невозможно сделать определенные

выводы о влиянии внедрений ретроэлементов на процессы видообразования, проходившие в суперсемействе *Hominoidea*.

Недавно мы предложили такие подходы для полногеномного сравнения близкородственных видов [21, 22]. Они включают несколько стадий (рис. 1): (1) ПЦР-амплификацию всех фрагментов ДНК, примыкающих к РЭ-повторам в обоих сравниваемых геномах; (2) вычитающую гибридизацию полученных ампликонов с последующим созданием клонотеки, обогащенной фрагментами ДНК, специфичными для одного из сравниваемых образцов; (3) дифференциальную гибридизацию созданной клонотеки, секвенирование и картирование дифференциальных клонов, для окончательной идентификации видоспецифичных последовательностей.

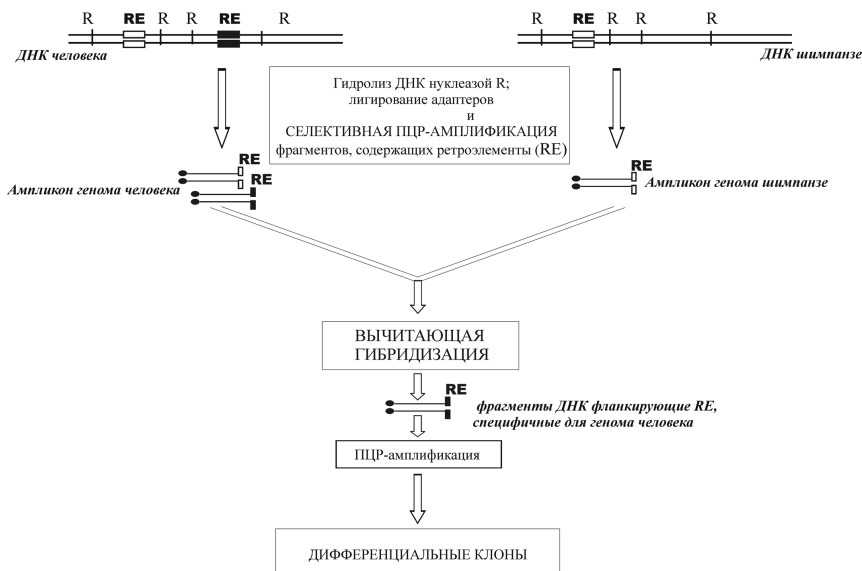


Рис. 1. Общая схема метода полногеномного сравнения близкородственных видов в участках интеграции РЭ.

Тонкими параллельными линиями обозначены фрагменты геномных ДНК; вертикальными отрезками с символом R – участки узнавания используемой эндонуклеазы рестрикции; положение ретроэлементов (RE) обозначено светлыми (общие для сравниваемых геномов) и темными (специфичные для генома человека) прямоугольниками; лигированные супрессионные адаптеры отмечены темными кружками

Этот метод позволил нам осуществить первый, в масштабах целых геномов, экспериментальный поиск различий в интеграции LTR- и L1-

элементов между геномами человека и шимпанзе. Дальнейший анализ участков интеграции в геномах приматов и структурный анализ РЭ человека привел нас к следующим выводам.

- 1) Не менее 4000 инсерций L1 и 150 инсерций HERV-K элементов отличают геном человека от генома шимпанзе [21, 23, 24].
- 2) Значительная часть интеграций LTR, специфичных для генома человека, неравномерно распределяется как между отдельными хромосомами, так и вдоль каждой из них. Интеграции LTR часто локализуются в непосредственной близости от различных генов человека или даже внутри них, что предполагает вероятное участие LTR в регуляции экспрессии окружающих генов [25]. Отдельные примеры расположения вблизи или внутри генов выявленных нами человек-специфичных LTR представлены на рисунке 2. Данные о концентрации LTR рядом с кодирующими областями согласуется с результатами других авторов о функциональном участии различных типов РЭ в регуляции экспрессии генов [26–30].

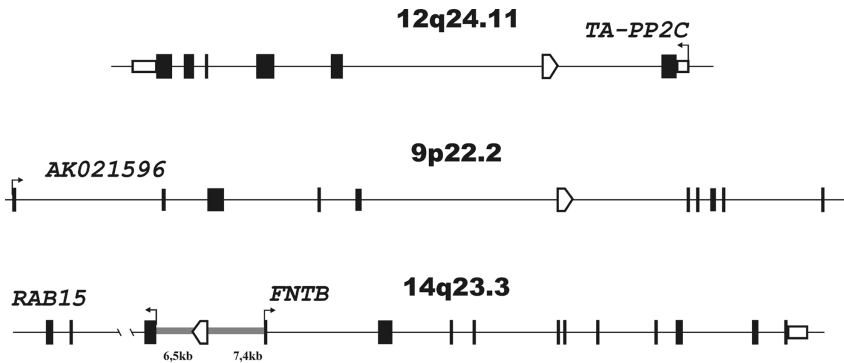


Рис. 2 Схемы трех локусов генома, содержащих человек-специфические интеграции LTR.

Горизонтальными отрезками обозначены фрагменты локусов генома человека, хромосомное расположение локусов приведено над каждым из отрезков; черными (транслируемые области) и светлыми (нетранслируемые) прямоугольниками отмечено положение экзонов генов, названия которых даны курсивом, а направление транскрипции указано стрелками; светлые пятиугольники указывают положение и ориентацию LTR элементов; серый прямоугольник – межгенный участок локуса 14q23.3; расстояния от LTR до старта транскрипции генов *RAB15* и *FNTB* указано в тыс. пар оснований (kb)

### **3. Анализ клеточной РНК демонстрирует большое количество транскрибируемых LTR и их разное содержание в нормальных и опухолевых тканях**

Сейчас уже стало ясно, что по крайней мере часть РЭ сохранила свой функциональный потенциал на протяжении миллионов лет эволюции, но их истинный вклад в функционирование и эволюцию генома хозяина еще очень далек от понимания. Вновь встроившиеся РЭ могут по-разному влиять на соседние гены: подавлять или усиливать их транскрипцию, изменять тканеспецифичность их экспрессии, а также создавать гены с изменёнными функциями при помощи механизма альтернативного сплайсинга [27].

Например, одиночные LTR эндогенных ретровирусов могут функционировать как (см. обзоры [2-4, 27]):

- энхансеры, усиливающие транскрипцию генов,
- промоторы транскрипции,
- терминаторы транскрипции,
- «горячие точки» рекомбинации,
- сигналы, изменяющие структуру хроматина.

Кроме того, LTR потенциально могут участвовать в синтезе антисмысловых РНК, подавляющих транскрипцию соседних генов.

В клетках различных типов один и тот же LTR может реализовывать различные комбинации вышеперечисленных функций. При этом разные LTR могут выполнять разные функции в одной и той же клетке или ткани и функции одного и того же LTR могут быть разными в разных, даже близкородственных, клетках или тканях [31, 32].

Несмотря на принципиальную возможность выполнять все вышеперечисленные функции, на сегодняшний день достоверно показано, что LTR могут выполнять роль промоторов и терминаторов транскрипции, а также участвовать в альтернативном сплайсинге. Во всех этих случаях LTR входят, по крайней мере частично, в состав образующихся транскриптов. В принципе любой современный метод определения профиля экспрессии генов может быть использован для анализа характера экспрессии РЭ в разных тканях. Однако ни один из существующих в настоящее время методов не позволяет анализировать все предполагаемые функции LTR. В частности, методы, основанные на гибридизации, успешно применявшиеся для определения и сравнительного анализа экспрессии генов, такие как вычитающая гибридизация и микрочипы, практически не используются при анализе РЭ. Очевидной слабостью этих методов применительно к исследованию транскриптов РЭ является нежелательная перекрестная гибридизация повторяющихся последовательностей, что может серьезно нарушить интерпретацию данных по экспрессии РЭ. Поэтому для изучения РНК, содержащих РЭ, более предпочтительны

ния РНК, содержащих РЭ, более предпочтительны методы определения и сравнения транскриптов, позволяющие избежать стадии гибридизации.

Нами был разработан метод селективного дифференциального дисплея РНК, содержащих повторяющиеся элементы (Selective Differential Display of RNAs containing Interspersed Repeats, SDDIR), который позволяет анализировать наборы РНК, содержащих определенный тип повторов [32].

SDDIR впервые дал возможность изучить весь спектр LTR-содержащих транскриптов HERV-K или, по крайней мере, его значительную часть. Было показано, что транскрибируется неожиданно большое число LTR, причем относительное содержание многих транскриптов различается в нормальных и опухолевых тканях. SDDIR позволил показать, что количество и относительное содержание транскриптов, содержащих LTR, могут как уменьшаться, так и увеличиваться в опухолях по сравнению с нормальными тканями (рис. 3).

Анализ 60-ти обнаруженных LTR-содержащих транскриптов, в разной степени представленных в нормальных и опухолевых тканях, позволил разделить найденные транскрибируемые LTR на несколько категорий: (1) LTR эндогенных ретровирусов. Этот результат не является неожиданным, так как известно, что экспрессия генов эндогенных ретровирусов изменяется в опухолях; (2) LTR, расположенные в кластерах других повторов. Принято считать, что такие кластеры не проявляют транскрипционную активность. Однако полученные нами данные позволяют предполагать, что они сохраняют определенный транскрипционный потенциал; (3) LTR, находящиеся в межгенных участках генома, вдалеке от известных генов или картированных мРНК; (4) LTR, расположенные вблизи аннотированных генов или картированных мРНК; (5) LTR, расположенные в интронах генов. Группы (4) и (5) могут участвовать в регуляции экспрессии близлежащих генов.

Последняя группа заслуживает особого рассмотрения. Тринадцать LTR были обнаружены в интронах известных генов или генов, предсказанных на основе последовательностей мРНК. В 8 случаях направление транскрипции LTR совпадало с направлением транскрипции «хозяйских» генов. Однако в 5 случаях направления транскрипции LTR и хозяйских генов были противоположны. Эти найденные нами транскрипты могут играть роль в антисмысловой регуляции экспрессии генов.

Мы показали, что SDDIR является высокоспецифичным методом, позволяющим анализировать наборы РНК, содержащие повторяющиеся элементы генома. Он позволяет практически со 100% специфичностью амплифицировать только LTR-содержащие транскрипты и с 90% специфичностью идентифицировать транскрипты, различающиеся по представленности в нормальных и опухолевых тканях. Таким образом, проведённый

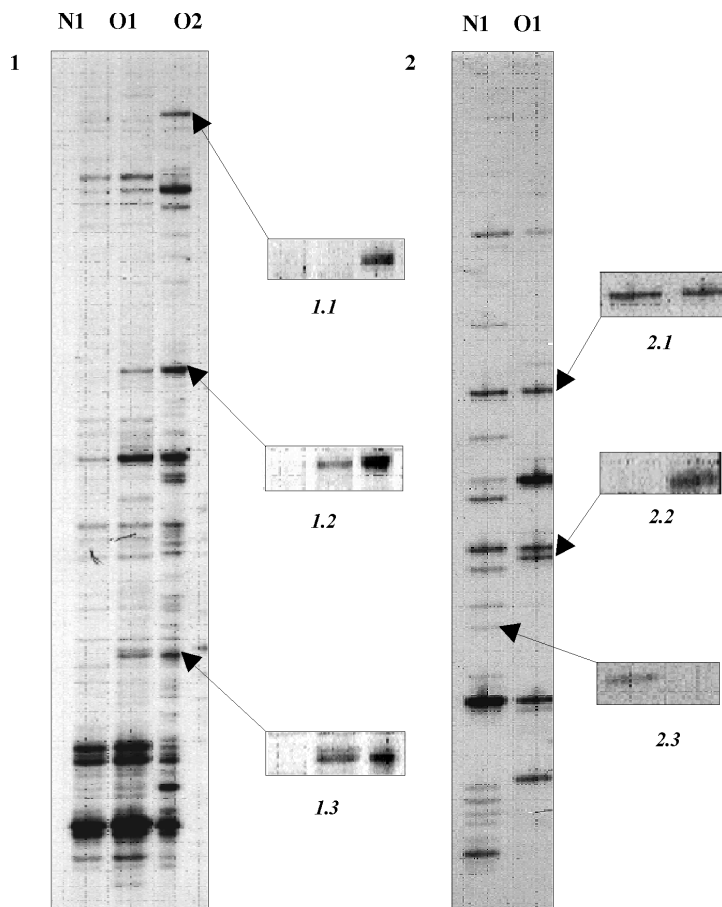


Рис. 3. Результаты анализа LTR-содержащих транскриптов методом SDDIR (Selective Differential Display of RNAs containing Interspersed Repeats).

S1, S2 – два разных образца семиномы; N1 – нормальная паренхима яичка, взятая у пациента с семиномой S1. (1) 1-ая цепь кДНК была синтезирована с LTR-специфического праймера; (2) 1-ая цепь кДНК была синтезирована с oligo(dT) праймера. Стрелки указывают положение выделенных фрагментов на геле: 1.2, 1.3 и 2.2 – транскрипты присутствуют в семиномах и отсутствуют в нормальной паренхиме яичка; 2.3 – транскрипты присутствуют только в нормальной паренхиме; 1.1 – транскрипты присутствуют только в одной семиноме и отсутствуют в нормальной паренхиме яичка и другой семиноме; 2.1 – транскрипты присутствуют и в семиномах, и в нормальной паренхиме



нами первый анализ LTR-содержащих мРНК в полном транскриптоме клетки показал, что в клетках человека транскрибируется множество LTR. Многие такие LTR располагаются вблизи генов или даже внутри них и таким образом вполне могли бы участвовать в процессе видообразования. Однако их истинная функциональная роль пока, к сожалению, далека от понимания.

#### **4. Образование химерных ретроэлементов у млекопитающих может происходить в результате смены матриц в ходе LINE-зависимой обратной транскрипции**

В настоящее время очевидно, что РЭ принимают весьма активное участие в эволюции геномов эукариот [3, 33]. В этом отношении наиболее хорошо изучена роль представителей LINE, относящихся к группе L1. Надо отметить, что внедрения L1 являются горячими точками рекомбинации, таким образом обуславливая множественные перестройки ДНК [34]. Кроме того, L1 способны «зашивать» разрывы двухцепочечной ДНК, вставляя новую копию ретроэлемента в поврежденный локус [35]. Другое интересное свойство L1 – так называемая «L1-трансдукция», то есть способность переносить фрагменты ДНК, примыкающие к их 3'-концевой области, в новые геномные локусы [36, 37]; [38]. Как оказалось, возможности LINE в изменении генома хозяина этим не исчерпываются.

Недавно мы обнаружили в геноме человека химерные ретроэлементы, состоящие из двух компонентов: ДНК-копий малой ядерной РНК U6 и фрагмента L1 [7]. Впоследствии оказалось, что в состав таких химерных ретроэлементов в сочетании с L1 могут входить не только копии U6 мРНК, но и ДНК-копии различных клеточных транскриптов: рибосомных РНК, различных малых ядерных РНК, мРНК клеточных генов, 7SL РНК, а также РНК мобильных элементов. Все идентифицированные химеры имели следующие общие свойства: (1) их 5'- и 3'- части были непосредственно соединены, (2) обе части имели одинаковую ориентацию относительно начала транскрипции, (3) все химеры были фланкированы прямыми повторами длиной 8–25 пар оснований и несли на своих 5' концах гексануклеотид или его производные с одной или двумя нуклеотидными заменами [6, 7]. Два последних свойства обусловлены способностью энзиматического аппарата L1 предпочтительно узнавать сайты TТАААА и образовывать на границах интегрированной ДНК короткие тандемные повторы [39].

Мы предположили [7], что химеры были образованы при помощи ранее неизвестного механизма, включающего рекомбинацию в ходе обратной транскрипции различных клеточных РНК. Было высказано предположение, что такой тип рекомбинации возник в результате смены мат-

рицы, происходящей в процессе обратной транскрипции РНК энзиматическим аппаратом L1 (рис. 4). Эта гипотеза была подтверждена детальным анализом последовательности химер и структурным анализом геномов различных видов приматов. Таким образом, химеры представляют собой объединённые ДНК-копии различных клеточных транскриптов, одновременно интегрировавшие в геном хозяина в составе единого блока.

Далее нами были исследованы доступные в базах данных геномные последовательности ДНК 9 различных видов животных: млекопитающих (*Homo sapiens*, *Mus musculus* и *Rattus norvegicus*), амфибий (*Xenopus laevis*), рыб (*Danio rerio* и *Takifugu rubripes*) и беспозвоночных (*Drosophila melanogaster*, *Anopheles gambiae* и *Caenorhabditis elegans*) на наличие генов и псевдогенов малых ядерных РНК U6, U5, U3, рибосомной 5S РНК, а также 7SL РНК, которые образовывали 5'-части химер у человека. Проведённый анализ показал, что у млекопитающих вышеописанный механизм является эволюционно консервативным (Гогвадзе и др. Биоорг. Хим., принято к печати). Геномы млекопитающих (человека, мыши и крысы) содержали химеры, тогда как в ДНК беспозвоночных, рыб и амфибий они обнаружены не были. Примечательно, что 5'-концевые части химер представляли собой копии РНК, имеющих либо ядерную, либо ядерную и цитоплазматическую локализацию, в то время как 3'-концевые части представляли собой копии цитоплазматических РНК: мРНК клеточных генов или мобильных элементов. Кроме того, обнаружена корреляция между эволюционными возрастами частей химер: чем моложе были 5'-части, тем моложе были и 3'- части химер, и наоборот.

Было обнаружено, что доля псевдогенов U6 мРНК, участвующих в образовании химер, существенно выше у грызунов, чем у *Homo sapiens*: у человека 33% всех псевдогенов U6 входят в состав химер, по сравнению с 69% у мыши и 64% у крысы. Предположив, что почти все псевдогены U6 (как химерные, так и нехимерные) появились в результате побочной активности энзиматического аппарата L1, можно сделать вывод, что L1 комплексы грызунов меняют матрицы в ходе обратной транскрипции приблизительно вдвое чаще, чем L1 человека. Наши данные также показали, что механизм образования химер возник приблизительно 75 миллионов лет назад, то есть прежде, чем разошлись предковые линии приматов и грызунов [8], и является активным и поныне [6]. Кроме того, недавно нами были обнаружены химеры сходного строения, состоящие не из двух, а из трёх компонентов. Два таких «тройных» элемента были обнаружены в мышинной ДНК и два – в геноме человека. Многие из химер экспрессируются в различных тканях млекопитающих, что позволяет рассматривать описанный феномен как новый РЭ-зависимый и эволюционно консервативный механизм образования генов.

## 1. Молекулярная и клеточная регуляция

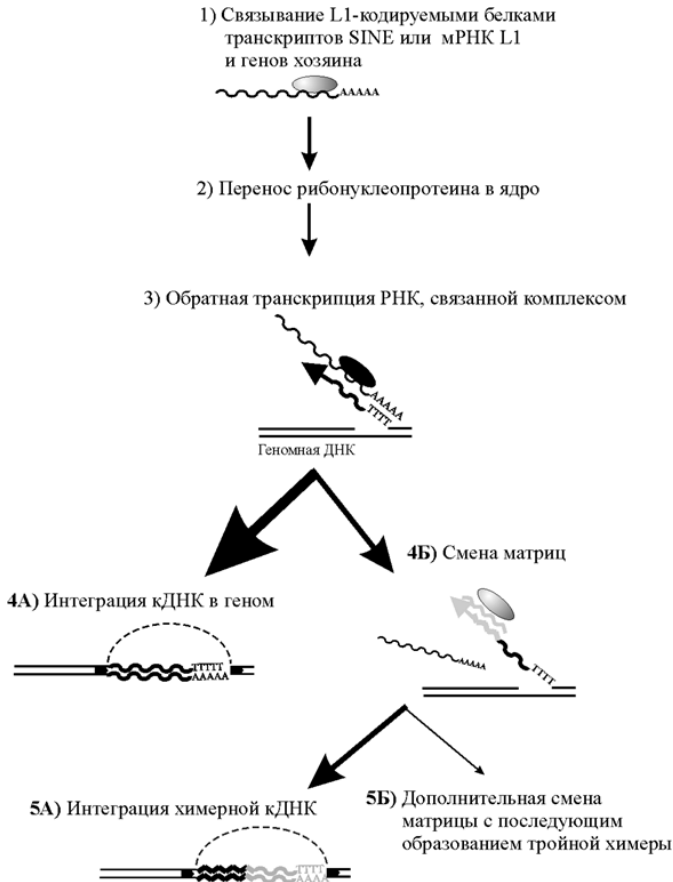


Рис. 4. Предполагаемый механизм формирования двойных и тройных химер. (Стадия 1) Преинтеграционный комплекс L1 связывает в цитоплазме мРНК клеточного гена, L1 или Alu. (Стадия 2) Сформированный рибонуклеопротеин транспортируется в ядро. (Стадия 3) Обратная транскрипция связанной комплексом мРНК, при которой в качестве праймера выступает одноцепочечный разрыв ДНК внутри последовательности ТТТТАА. (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 последовательность поли (А) на 3'-конце.

## **Заключение**

В данном обзоре мы постарались показать, что разработанные нами новые экспериментальные подходы позволили выявить некоторые важные признаки генома человека:

1. Имеется множество интеграций ретроэлементов, отличающих геном человека от геномов его ближайших родственников – человекообразных обезьян. Эти интеграции в ряде случаев находятся в непосредственной близости от генов, что создает возможность изменения их транскрипционных характеристик.
2. Неожиданно большое число РЭ сохранили свою функциональную активность на протяжении миллионов лет эволюции. Характеристики этой активности индивидуальны для каждого РЭ.
3. Способность РЭ участвовать в эволюции геномов не ограничивается хорошо известными механизмами, такими как изменения структуры генов, рекомбинационные перестройки ДНК или образование псевдогенов. L1, основные преобразователи генома млекопитающих, способны создавать новые ретроэлементы с использованием ранее неизвестного механизма: РНК-РНК рекомбинации и последующей интеграции химерных копий кДНК в новые позиции генома.
4. Все эти свойства РЭ делают их серьезными кандидатами на роль эволюционных преобразователей, способных фиксироваться в геномной ДНК при помощи механизмов естественного отбора.

Вместе с тем, взаимодействие генома хозяина и РЭ представляется чрезвычайно сложным и многофункциональным процессом. Современная наука находится в самом начале пути к пониманию этих сложных взаимодействий. Вместе с тем возможно, что в относительно недалеком будущем сопоставление всей совокупности таких взаимодействий у организмов, принадлежащих к разным видам, позволит сузить круг претендентов на роль эволюционных революционеров и сфокусироваться на анализе их возможного вклада в процессы эволюции и видообразования. Тем не менее, уже сейчас можно предполагать, что многие представители РЭ являются вероятными кандидатами на эту роль.

## **Благодарности**

Работа поддержана грантами РФФИ НШ-2006.20034 и президиума РАН «Молекулярная и клеточная биология».

## Словарь терминов и сокращений

**Гоминиды** (*Hominidae*) – семейство отряда приматов, включающее современного человека и вымершие виды рода *Homo*.

**Гоминоиды** (*Hominoidea*) – надсемейство, включающее семейства гоминид (человек), понгид (шимпанзе, горилла, орангутан) и гиббонообразных (гиббон, сиаманг).

**РЭ** – ретроэлементы – мобильные элементы генома, способные к размножению посредством синтеза РНК-копий (транскрипции) их превращению в комплементарную ДНК (кДНК, обратная транскрипция) и интеграций кДНК в другие места генома.

**Интеграза** – кодируемый ретроэлементами белок, осуществляющий внедрение ДНК-копий ретроэлементов в геном.

**LINE** – тип ретроэлементов, кодирующих обратную транскриптазу и интегразу (автономные ретроэлементы). LINE-элементы геномов млекопитающих включают три отдельных группы – **LINE1 (L1)**, **LINE2 (L2)** и **LINE3 (L3)**. **L1-Ta, L1PA1, L1PA2 – L1PA14** – отдельные семейства LINE1-элементов, специфичные для геномов приматов.

**L1\_MM, L1\_Md-F\_2\_3** – отдельные семейства (группы) LINE1-элементов генома мыши.

**SINE** – короткие диспергированные ядерные элементы – самый многочисленный тип ретроэлементов геномов млекопитающих. SINE-элементы не кодируют собственной ревертазы (неавтономные ретроэлементы) и используют для размножения в геноме ретропозиционный аппарат L1 элементов.

**Alu, SVA, SINE-R** – отдельные классы SINE-элементов генома человека.

**B1, B2** – классы SINE элементов мышинного генома.

**Alu**-элементы – специфичный для генома приматов класс SINE повторов, включает три суперсемейства (**AluY, AluS** и **AluJ**), разделяемые на отдельные семейства и группы (**AluYa5, AluYb8, AluSx, AluSq** и др.)

**HERV** – эндогенные ретровирусы человека (от англ. human endogenous retroviruses) – автономные LTR-содержащие ретроэлементы, найденные в геномах человека и других видов приматов; общая структура **HERV** сходна со строением интегрированной ДНК копии (провирусом) генома экзогенных ретровирусов позвоночных; нуклеотидные последовательности **HERV** в геноме человека представлены как дефектными провирусами различных семейств – **HERV-H, HERV-E, ERV-9, HERV-W** и др., так и одиночными LTR (см. далее), образованными в результате рекомбинации между последо-

вательностями идентичных LTR одного интегрировавшего HERV элемента.

**LTR** – длинный концевой повтор (от англ. long terminal repeat) – регуляторные элементы, ограничивающие с обоих концов структуру эндогенного ретровируса; нуклеотидная последовательность LTR включает три функциональные области – **U3**, **R** и **U5**.

**Транскриптом** – полное содержание РНК в данной клетке в данных условиях. Транскриптом динамичен. Его состав меняется при изменении условий в которых находится клетка. Разные типы клеток содержат различные транскриптомы.

**Транскрипция** – процесс копирования ДНК в комплементарную РНК. Осуществляется РНК-полимеразами. Транскрипция инициируется с помощью специфичных последовательностей ДНК промоторов и энхансеров. Терминация транскрипции также осуществляется в специфичных участках ДНК.

**Обратная транскрипция** – процесс копирования РНК с образованием комплементарной ДНК (кДНК). Осуществляется ферментами обратными транскриптазам.

## **Литература**

1. Foley R. The context of human genetic evolution. *Genome Res.* 1998;339–347.
2. Kazazian H.H., Jr. Mobile elements: drivers of genome evolution. *Science* 2004; 303:1626–1632.
3. Sverdlov E.D. Retroviruses and primate evolution. *Bioessays* 2000; 22:161–171.
4. Sverdlov E.D. Perpetually mobile footprints of ancient infections in human genome. *FEBS Lett.* 1998; 428:1–6.
5. Mager D.L., Medstrand P. Retroviral Repeat Sequences. In: *Nature Encyclopedia of the human genome*. Vol. Science (in press); Macmillan Online Publishing, 2003.
6. Buzdin A., Ustyugova S., Gogvadze E., Vinogradova T., Lebedev Y., Sverdlov E. A new family of chimeric retrotranscripts formed by a full copy of U6 small nuclear RNA fused to the 3' terminus of I1. *Genomics* 2002; 80:402–406.
7. Buzdin A., Gogvadze E., Kovalskaya E., Volchkov P., Ustyugova S., Il-larionova A., et al. The human genome contains many types of chimeric retrogenes generated through in vivo RNA recombination. *Nucleic Acids Res.* 2003; 31:4385–4390.

8. Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* 2002; 420:520–562.
9. Venter J.C., Adams M.D., Myers E.W., Li P.W., Mural R.J., Sutton G.G., et al. The sequence of the human genome. *Science* 2001; 291:1304–1351.
10. Britten R.J., Baron W.F., Stout D.B., Davidson E.H. Sources and evolution of human Alu repeated sequences. *Proc. Natl. Acad. Sci. U S A* 1988; 85:4770–4774.
11. Jurka J., Smith T. A fundamental division in the Alu family of repeated sequences. *Proc. Natl. Acad. Sci. U S A* 1988; 85:4775–4778.
12. Quentin Y. The Alu family developed through successive waves of fixation closely connected with primate lineage history. *J. Mol. Evol.* 1988; 27:194–202.
13. Batzer M.A., Deininger P.L. Alu repeats and human genomic diversity. *Nat. Rev. Genet.* 2002; 3:370–379.
14. Carroll M.L., Roy-Engel A.M., Nguyen S.V., Salem A.H., Vogel E., Vincent B., et al. Large-scale analysis of the Alu Ya5 and Yb8 subfamilies and their contribution to human genomic diversity. *J. Mol. Biol.* 2001; 311:17–40.
15. Roy-Engel A.M., Carroll M.L., Vogel E., Garber R.K., Nguyen S.V., Salem A.H., et al. Alu insertion polymorphisms for the study of human genomic diversity. *Genetics* 2001; 159:279–290.
16. Skowronski J., Fanning T.G., Singer M.F. Unit-length line-1 transcripts in human teratocarcinoma cells. *Mol. Cell. Biol.* 1988; 8:1385–1397.
17. Myers J.S., Vincent B.J., Udall H., Watkins W.S., Morrish T.A., Kilroy G.E., et al. A comprehensive analysis of recently integrated human Ta L1 elements. *Am. J. Hum. Genet.* 2002; 71:312–326.
18. Boissinot S., Chevret P., Furano A.V. L1 (LINE-1) retrotransposon evolution and amplification in recent human history. *Mol. Biol. Evol.* 2000; 17:915–928.
19. Sheen F.M., Sherry S.T., Risch G.M., Robichaux M., Nasidze I., Stoneking M., et al. Reading between the LINEs: human genomic variation induced by LINE-1 retrotransposition. *Genome Res.* 2000; 10:1496–1508.
20. Moran J.V., Holmes S.E., Naas T.P., DeBerardinis R.J., Boeke J.D., Kazazian H.H., Jr. High frequency retrotransposition in cultured mammalian cells. *Cell* 1996; 87:917–927.
21. Buzdin A., Khodosevich K., Mamedov I., Vinogradova T., Lebedev Y., Hunsmann G., et al. A technique for genome-wide identification of differences in the interspersed repeats integrations between closely related genomes and its application to detection of human-specific integrations of HERV-K LTRs. *Genomics* 2002; 79:413–422.
22. Mamedov I., Batrak A., Buzdin A., Arzumanyan E., Lebedev Y., Sverdlov E.D. Genome-wide comparison of differences in the integration sites

- of interspersed repeats between closely related genomes. *Nucleic Acids Res.* 2002; 30:e71.
23. Buzdin A., Ustyugova S., Gogvadze E., Lebedev Y., Hunsmann G., Sverdlov E. Genome-wide targeted search for human specific and polymorphic L1 integrations. *Hum. Genet.* 2003; 112:527–533.
  24. Buzdin A., Ustyugova S., Khodosevich K., Mamedov I., Lebedev Y., Hunsmann G., et al. Human-specific subfamilies of HERV-K (HML-2) long terminal repeats: three master genes were active simultaneously during branching of hominoid lineages. *Genomics* 2003; 81:149–156.
  25. Buzdin A.A., Lebedev Y. B., Sverdlov E.D. Human-specific HERV-K intron LTRs have nonaccidental opposite orientation relative to the direction of gene transcription and might be involved in the antisense regulation of gene expression. *Rus. J. Bioorg. Khim.* 2003; 29:103–106.
  26. Kreaehling J., Graveley B.R. The origins and implications of Aluternative splicing. *Trends Genet.* 2004; 20:1–4.
  27. Van de Lagemaat L.N., Landry J.R., Mager D.L., Medstrand P. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet.* 2003; 19:530–536.
  28. Kowalski P.E., Mager D.L. A human endogenous retrovirus suppresses translation of an associated fusion transcript, PLA2L. *J. Virol.* 1998; 72:6164–6168.
  29. Baust C., Seifarth W., Schon U., Hehlmann R., Leib-Mosch C. Functional activity of HERV-K-T47D-related long terminal repeats. *Virology* 2001; 283:262–272.
  30. Mager D.L., Hunter D.G., Schertzer M., Freeman J.D. Endogenous retroviruses provide the primary polyadenylation signal for two new human genes. *Genomics* 1999; 59:255–263.
  31. Vinogradova T.V., Leppik L.P., Nikolaev L.G., Akopov S.B., Kleiman A.M., Senyuta N.B., et al. Solitary human endogenous retroviruses-K LTRs retain transcriptional activity in vivo, the mode of which is different in different cell types. *Virology* 2001; 290:83–90.
  32. Vinogradova T., Leppik L., Kalinina E., Zhulidov P., Grzeschik K.H., Sverdlov E. Selective Differential Display of RNAs containing interspersed repeats: analysis of changes in the transcription of HERV-K LTRs in germ cell tumors. *Mol. Genet. Genomics* 2002; 266:796–805.
  33. Deininger P.L., Moran J.V., Batzer M.A., Kazazian H.H., Jr. Mobile elements and mammalian genome evolution. *Curr. Opin. Genet. Dev.* 2003; 13:651–658.
  34. Kazazian H.H., Jr., Goodier J.L. LINE drive. retrotransposition and genome instability. *Cell* 2002; 110:277–280.
  35. Teng S.C., Kim B., Gabriel A. Retrotransposon reverse-transcriptase-mediated repair of chromosomal breaks. *Nature* 1996; 383:641–644.



36. Pickeral O.K., Makalowski W., Boguski M.S., Boeke J.D. Frequent human genomic DNA transduction driven by LINE-1 retrotransposition. *Genome Res.* 2000; 10:411–415.
37. Moran J.V., DeBerardinis R.J., Kazazian H.H., Jr. Exon shuffling by L1 retrotransposition. *Science* 1999; 283:1530–1534.
38. Wei W., Gilbert N., Ooi S.L., Lawler J.F., Ostertag E.M., Kazazian H.H., et al. Human L1 retrotransposition: cis preference versus trans complementation. *Mol. Cell. Biol.* 2001; 21:1429–1439.
39. Jurka J. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. *Proc. Natl. Acad. Sci. U S A* 1997; 94:1872–1877.