



УДК 577.113.5:578.828

## СТРУКТУРНЫЕ ХАРАКТЕРИСТИКИ ЧЕТЫРЕХ ДЛИННЫХ КОНЦЕВЫХ ПОВТОРОВ (LTR) ЭНДОГЕННЫХ ЧЕЛОВЕЧЕСКИХ РЕТРОВИРУСОВ И ОСОБЕННОСТИ УЧАСТКОВ ИХ ИНТЕГРАЦИИ

© 1997 г. П. П. Хиль, М. Б. Костина, Т. Л. Ажикина, Т. Б. Колесник,  
Ю. Б. Лебедев, Е. Д. Сverdlov<sup>#</sup>

Институт биоорганической химии им. М.М. Шемякина и Ю.А. Овчинникова РАН,  
117871, Москва, ГСП-7, ул. Миклухо-Маклая, 16/10

Поступила в редакцию 15.01.97 г.

Определена первичная структура четырех областей 19-й хромосомы человека, содержащих длинные концевые повторы (LTR), с использованием разработанного метода шагового секвенирования при помощи комбинированных праймеров с повышенной прочностью связывания с матрицей. Проведен сравнительно-эволюционный анализ последовательностей, гомологичных эндогенным ретровирусам человека (HERV), и определены прототипы изучаемых LTR. Анализ последовательностей 19-й хромосомы, прилегающих к LTR, обнаружил совместную локализацию длинных концевых повторов HERV-K с другими ретропозонами.

*Ключевые слова:* LTR, HERV-K, шаговое секвенирование, первичная структура ДНК.

Среди повторяющихся элементов генома человека последовательности ДНК, гомологичные геномам ретровирусов млекопитающих, – эндогенные ретровирусы человека (HERV) – выделяются своим высоким регуляторным потенциалом. Полагают, что первичное ретровирусное инфицирование экзогенными предшественниками HERV и встраивание их ДНК в ДНК зародышевых клеток произошло от 10 до 40 млн. лет назад [1]. Существует более десятка различных семейств HERV, представленных в геноме человека от одной до десятков тысяч копий [2]. Увеличение числа элементов HERV в геноме человека происходило как в результате интеграции ретровирусов в ДНК зародышевых клеток, так и при ретротранспозиции всего вирусного генома или его отдельных частей в новые локусы. В целом до 1% генома человека занимают ретровирусы и ретровирусподобные последовательности [2].

Функциональные эффекты эндогенных ретровирусподобных последовательностей пока мало изучены. Однако обнаружение транскриптов [3, 4] и некоторых белков [5, 6] HERV в эмбриональных и опухолевых клетках, а также в некоторых нормальных тканях позволяет предполагать участие HERV в процессах клеточной дифферен-

циации, иммунного ответа и злокачественного перерождения.

Длинный концевой повтор ретровирусного элемента – LTR HERV – включает в себя множество регуляторных элементов: промотор, энхансер, сигнальный участок гормонального ответа, сигналы полиаденилирования и др. Благодаря им LTR может существенно модулировать активность соседних генов.

Ранее в нашей лаборатории была определена локализация LTR семейства HERV-K на физической карте 19-й хромосомы человека [7]. Результаты этого картирования позволили нам идентифицировать известные гены, лежащие вблизи этих LTR. Дальнейший анализ структуры самих LTR и окружающих их последовательностей может создать уникальную возможность сопоставления структур элементов сравнительно недавнего эволюционного происхождения, позиции которых в геноме известны. В свою очередь, такое сопоставление важно как для выяснения особенностей регуляции экспрессии генов, так и для лучшего понимания механизмов эволюционных процессов, имевших место после появления приматов.

Данная работа посвящена определению и анализу структуры некоторых из этих потенциально регуляторных последовательностей в четырех случайно выбранных HERV-K LTR-содержащих космидах – 28373, 30181, 28096 и 32408. Они содержат геномную ДНК из участков 19-й хромосомы человека, локализованных в следующих положениях метрической карты: 28373 – район

Сокращения: LTR – длинный концевой повтор. HERV-K – эндогенный ретровирус человека (семейство K), м.п.о. – миллион пар оснований.

<sup>#</sup> Автор для переписки (тел.: (095) 330-65-29, факс: (095) 330-65-38, e-mail: eds@glas.apc.org, sverd@humgen.siobc.ras.ru).

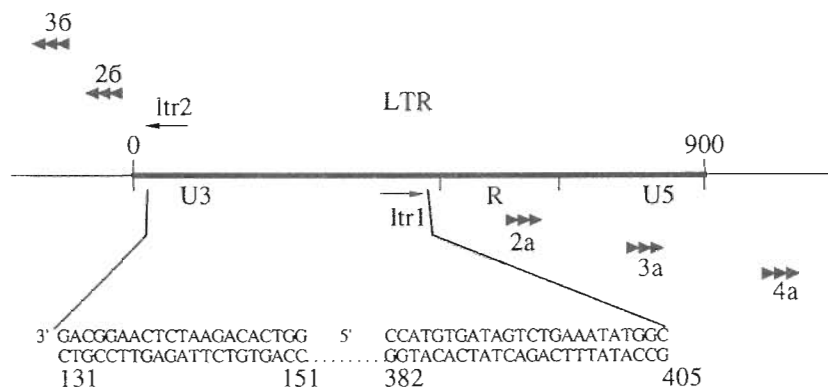


Рис. 1. Схема шагового секвенирования LTR-содержащих фрагментов 19-й хромосомы человека. Жирной линией обозначена последовательность LTR. Стрелки с обозначениями ltr1 и ltr2 соответствуют праймерам, использованным на первой стадии секвенирования, структура которых приведена в нижней части рисунка. Тройными стрелками 2a, 3a и т.д. обозначены комбинированные праймеры, использованные для дальнейшего секвенирования.

локуса q13.4 (55.75 м.п.о. от p-tel); 30181 – q13.2 (46.0 м.п.о.); 28096 – p13.3 (5.0 м.п.о.); 32408 – положение не установлено.

Для определения первичной структуры была использована стратегия шагового секвенирования, основанная на использовании комбинированных праймеров [8] (см. рис. 1).

Согласно представленной схеме, каждый из LTR-содержащих фрагментов секвенировали сначала с использованием двух длинных олигонуклеотидных праймеров, специфичных к консервативным для этого семейства LTR последовательностям, расположенным по краям U3-области. В результате получали две новые последовательности, на основании которых из заранее синтезированной библиотеки гексануклеотидов выбирали подходящую комбинацию трех гексамеров, сумма которых соответствует некоторой 18-звенной непрерывной последовательности в 3'-концевом районе вновь полученной нуклеотидной последовательности. Эти комбинации, получившие название олигонуклеотидных стрингов, использовали как праймеры для следующего шага секвенирования. Далее аналогичная операция проводится еще несколько раз, в результате чего удается выяснить достаточно длинную последовательность без промежуточного субклонирования.

Этот метод, предложенный Студиером [8], позволяет избежать стадии синтеза длинных праймеров по окончании каждого цикла секвенирования и, в принципе, автоматизировать большинство стадий процесса секвенирования. Если использовать полноразмерную библиотеку гексануклеотидных компонентов комбинированных праймеров, то она должна содержать  $4^6 = 4096$  олигонуклеотидов. Однако число гексамеров, необходимых для успешного использования библиотеки, можно значительно снизить, если принять во внимание частоту встречаемости гексануклеотидных последовательностей в геноме человека: расчеты показывают, что для решения структур-

ных задач достаточно располагать библиотекой из 1500 гексануклеотидов, создание которой вполне реально.

В данной работе мы использовали библиотеку модифицированных гексануклеотидов. Ранее мы показали, что олигонуклеотиды, содержащие 5-метилцитозин и 2-аминоаденин вместо цитозина и аденина соответственно и образующие дуплексы, более прочные, чем природные, служат эффективными и специфичными затравками в матричном синтезе ДНК [9]. Было продемонстрировано также, что модифицированные таким образом гекса- и пентамеры служат более эффективными компонентами комбинированных праймеров, чем их немодифицированные аналоги [10]. В дальнейшем мы ввели еще одну модификацию, позволяющую увеличить прочность А • Т-пары. Для этого был использован 5-(3-метоксипропил-1)уранил вместо тимина [11]. В настоящее время нами синтезирована библиотека из 1500 гексануклеотидов, содержащих вышеуказанные модификации.

Для целей секвенирования космидную ДНК расщепляли рестриктазой *EcoRI* в случае космиды 28373 и *PstI* в случае космид 30181, 28096 и 32408. Рестриктные фрагменты размером от 3 до 4 т.п.о., гибридизующиеся с радиоактивно меченным зондом из U3-области LTR, субклонировали в вектор pGEM7zf(+). ДНК плазмид рекомбинантных клонов использовали в качестве матриц для секвенирования.

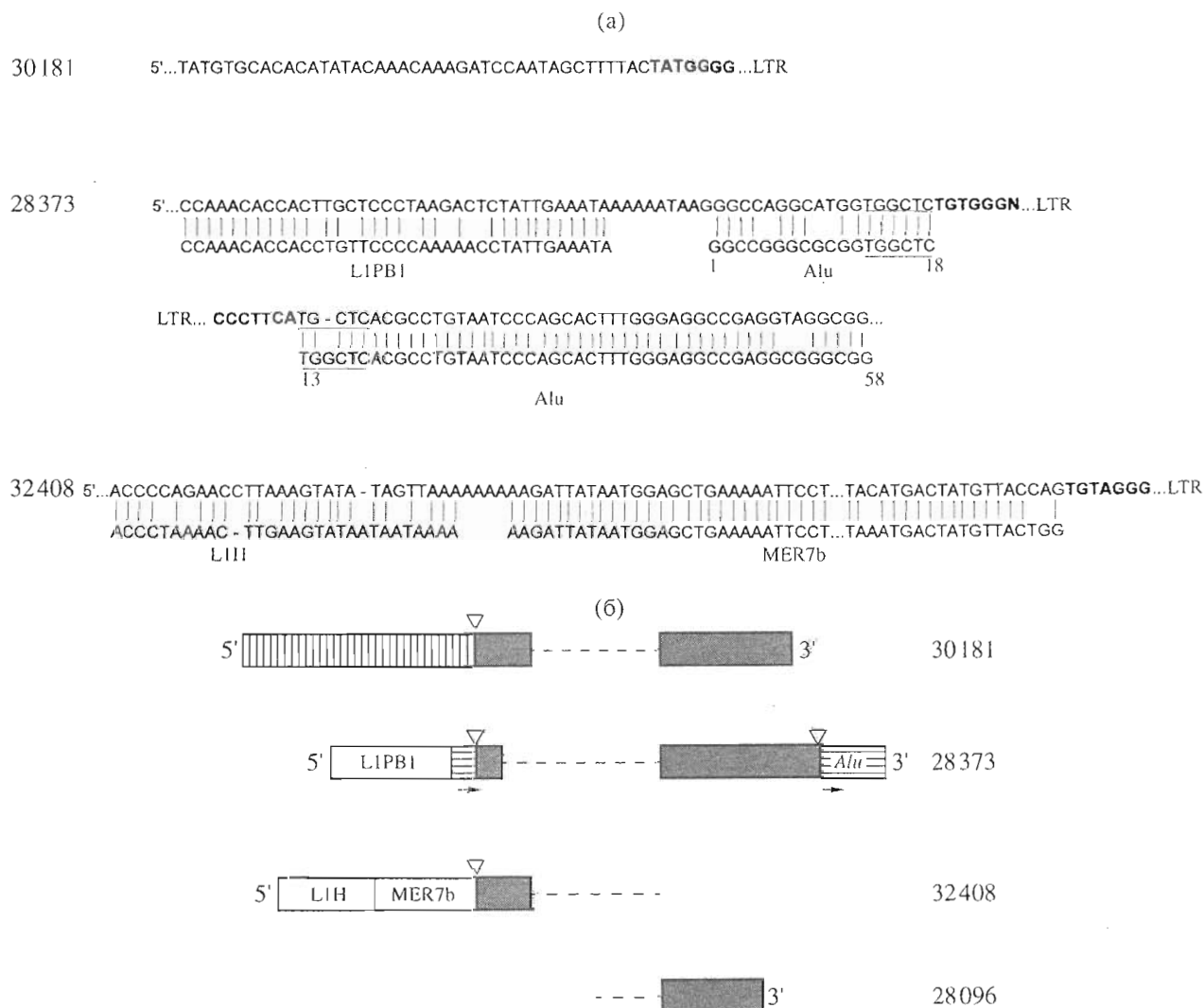
Первый шаг секвенирования проводили с использованием пары праймеров, соответствующих последовательности LTR HERV-K10, -ltr1 и ltr2 (см. рис. 1). Выбор этой пары праймеров обусловлен тем, что их структура высококонсервативна для всех последовательностей LTR-семейства HERV-K. Последующие шаги секвенирования осуществляли с помощью комбинированных праймеров (таблица) согласно вышеописанной стратегии.

Структуры олигонуклеотидных комбинированных праймеров, использованных для секвенирования LTR-содержащих фрагментов космид 28373, 30181, 28096 и 32408

Праймеры*	28373	30181	28096	32408
2a	AGATAG/GGGAAA/ACCGCC	AGGAAG/CATCTG/CTCCT	AGATAG/GGGAAA/ACCGCC	
2б	CCSACA/GAGCCA/CCATGC	TTTAGA/TATGTT/TATGTG		GCAGGA/ATTTTT/CAGCTC
3a	GCATTG/AGATGT/TTAGTG	AGGAAGCATCTGTCTCCT**	CACATC/CCCTCT/CTGAGA	
3б	CACCST/ACTCCA/CTTTTC	AAGGTG/TCACAG/TTTGCC		
4a	ATCCTC/CGTATG/CTGAAC			
4б		TTTGTG/LATTT/TTTTGA		
5a	CACAGGTGTGGAGGGGCA**			
5б		TATTT/TTTTGA/AAAAATA		
6б		TAGTTT/AATGAA/AACAGC		
7б		TGAGATAAATGCTTTATA**		

\* Буквой "а" обозначены праймеры, использованные для секвенирования в направлении R-области LTR, а буквой "б" – в обратном направлении.

\*\* В данном случае вместо комбинированного праймера из трех гексануклеотидов использовали 18-звенный праймер аналогичной структуры.



**Рис. 2.** Структурная организация LTR-содержащих областей 19-й хромосомы человека. (а) – гомология LTR-прилегающих участков из космид 30181, 28373 и 32408 с известными повторяющимися элементами человека. В верхних строках приведены установленные последовательности LTR-прилегающих участков, в нижних строках – фрагменты гомологичных последовательностей: LIPB1 и LIIH – фрагменты подсемейств LINE-повтора, Alu – Alu-повтор, MER7b – фрагмент подсемейства MER-повтора. В приведенном фрагменте Alu-повтора подчеркнут дублированный участок и обозначены координаты, соответствующие консенсусному Alu-повтору. Жирным шрифтом обозначены концевые инвертированные повторы LTR. Вертикальные линии соединяют идентичные основания гомологичных последовательностей; делеции обозначены дефисами. (б) – блок-схема структурной организации LTR-содержащих областей 19-й хромосомы человека. Горизонтальной штриховкой обозначены Alu-гомологичные последовательности, вертикальной – последовательность, соответствующая предположительному *gag*-гену представителя нового семейства эндогенных ретровирусов человека; затемнены LTR-гомологичные участки; штриховые линии соответствуют внутренним участкам U3 области LTR, структура которых в данной работе не определялась. Наименования прочих повторяющихся элементов указаны в соответствующих областях блок-схемы. Стрелками обозначены дублированные последовательности Alu-повтора. Треугольники указывают положение коротких концевых повторов.

В результате для каждой из космид была получена структурная информация, состоящая из двух частей:

1) первичная структура самих LTR, начиная от участков, комплементарных праймерам, использованным на первом шаге секвенирования;

2) структура прилегающих к LTR участков геномной ДНК.

Структуры участков, прилегающих к последовательностям LTR в космидах 30181, 28373 и 32408, приведены на рис. 2. Установленные последовательности LTR заканчиваются гептануклеотидами, гомологичными концевому короткому инвертированному повтору в составе LTR HERV-K (выделен жирным шрифтом на рис. 2а). Этот факт подтверждает присутствие полноразмерного LTR в анализируемых космидах. Для

всех полученных последовательностей ДНК был произведен поиск гомологий в международных базах данных. Результаты анализа суммированы в виде блок-схемы на рис. 2б. Последовательность LTR из космиды 28373 фланкирована с обеих сторон последовательностью Alu-повтора, что свидетельствует от интеграции LTR в 5'-область Alu-повтора (точка интеграции соответствует положениям 13–18 консенсусной Alu-последовательности). В точке интеграции присутствует короткий прямой повтор длиной 6 п.о. В 5'-области Alu-повтор непосредственно примыкает к 3'-области L1-повтора (подсемейство L1PB1 [12]).

Сравнение последовательности, расположенной в 5'-области по отношению к LTR HERV-K на космиде 30181, с международными базами данных выявило наличие нескольких гомологичных ей участков генома человека. Кроме анализируемой последовательности некоторые из обнаруженных локусов содержат длинные прямые повторы, сайт связывания праймера (PBS-элемент) и другие характерные структурные части эндогенных ретровирусов. Полный структурный анализ найденных повторяющихся последовательностей генома человека выходит за рамки данной работы и будет опубликован отдельно. Примыкающая к LTR последовательность из космиды 30181 относится к внутренней области – предположительно гену *gag* – обнаруженного ретровирусподобного элемента.

Последовательность из космиды 32408 фланкирована в 5'-области MER7b-повтором (предположительно являющимся ДНК-транспозоном [13]), а через 180 п.о. находится последовательность, соответствующая 3'-области L1-повтора (подсемейство L1H [12]). В данном случае, по-видимому, LTR и L1-повторы интегрировались в состав ранее присутствующего MER7b-элемента. Таким образом, с высокой, выходящей за рамки случайности частотой на границах с LTR оказываются представители других семейств повторов человека, перемещающихся в геноме путем обратной транскрипции.

Последовательности, по которым происходит интеграция ретровирусов в геном, не вполне случайны [14, 15]. В геноме имеется  $10^3$  сайтов, интеграция в которые происходит преимущественно [14], причем существуют данные об ассоциации этих сайтов с гиперчувствительными к ДНКазе I сайтами [15]. Известно также, что ориентация интегрированных провирусов зависит от сайта встраивания [14–17]. Данные о кластеризации эндогенных ретроэлементов в геноме человека подтверждают эти сведения [18–20]. Рядом авторов отмечается повышенная частота встречаемости кластеров Alu-повторов, а также неслучайное распределение Alu- и L1-повторов в геноме человека и частое совпадение позиций этих двух эле-

ментов [21, 22]. Также описаны случаи тесного сцепления одного из LTR (Spm-LTR) с Alu-повторами и THE1-ретропозоном [23]. Наши данные демонстрируют предрасположенность длинных концевых повторов эндогенных ретровирусов к формированию кластеров с другими ретроэлементами.

Совпадение позиций различных ретроэлементов может вызываться разными причинами, такими, например, как существование “горячих точек” внедрения обратных транскриптов; существование селекции, приводящей к невозможности сохранения внедрившихся ретроэлементов в большинстве мест генома; облегчение внедрения второго ретропозона после внедрения первого; функциональный “симбиоз” различных ретроэлементов в формировании ДНК-ДНК- и ДНК-белковых комплексов, необходимых для внедрения в геномную ДНК. Для понимания причин и следствий этого явления нужны дополнительные эксперименты.

С целью выяснения возможных эволюционных взаимосвязей отдельных LTR эндогенных ретровирусов мы обобщили доступную структурную информацию и использовали ее для построения филогенетического дерева LTR этого семейства эндогенных ретровирусов. Для этого был произведен поиск последовательностей, гомологичных LTR HERV-K, в международных базах данных при помощи BLAST WWW-сервера (доступен через Интернет по адресу: <http://www.ncbi.nlm.nih.gov/BLAST>). Поиск основывался на повторном использовании различных обнаруженных представителей этого семейства, а также различных вариантов промежуточных консенсусных последовательностей. Всего в избыточном подразделе GenBank было выявлено 22 полноразмерных элемента (по состоянию на июль 1996). Поиск в других подразделениях (dbEST и dbSTS) не обнаружил ни одного полноразмерного элемента. Выравнивание последовательностей было произведено при помощи программы CLUSTALW v. 1.5 [24] и отредактировано вручную при помощи GDE 2.0 [25]. Филогенетическое дерево было построено и визуализировано с использованием программ CLUSTALW v. 1.5 и пакета программ Phylip [26]. Результаты представлены на рис. 3. Анализ полученного дерева позволяет выделить два класса длинных концевых повторов HERV-K-элементов. Класс I состоит из более гомологичных между собой последовательностей и, по всей видимости, более молодой. К нему принадлежат LTR прототипного провируса этого семейства [27]. Класс II более гетерогенен и, скорее всего, более древний и многочисленный. Эти два класса различаются характерными “диагностическими” заменами. Во втором из них присутствуют вставки длиной 5, 7, 8 и 21 п.с., отсутствующие в первом. Эти вставки

наряду с точечными заменами позволяют уверенно относить вновь получаемые последовательности к одному из этих типов LTR.

Полученное эволюционное дерево было использовано для отнесения определенных нами последовательностей ДНК LTR к определенному классу и для определения их относительного эволюционного положения по отношению к другим LTR (рис. 3). LTR HERV-K из космиды 30181 (46.0 м.п.о. метрической карты), вероятно, принадлежит ко второму классу. Последовательности LTR из космид 28096 и 28373 (55.75 и 5.0 м.п.о. метрической карты соответственно) относятся, вероятно, к первому классу этих элементов и наиболее близки к LTR HERV-K, отмеченных на рис. 3 фигурной скобкой.

Определение последовательностей большего числа ретровирусных длинных концевых повторов позволит уточнить это дерево и определить последовательность внедрений различных LTR в геном человека.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

В работе использовали: акриламид, бисакриламид (Bio-Rad, США), трис-гидроксиметиламинометан (трис) (Merck, США), этилендиаминтетрауксусную кислоту (EDTA, динатриевая соль) (Sigma, США), [ $\alpha$ - $^{32}$ P]dATP (Обнинск), [ $\alpha$ - $^{33}$ P]dATP (ИБХ РАН), нитроцеллюлозную мембрану Hybond (Amersham, Англия), эндонуклеазы рестрикции *EcoRI* и *PstI* (Boehringer, Германия), T4-ДНК-лигазу, *Taq*-ДНК-полимеразу (ИБХ РАН), SSB protein (белок, специфически связывающий одноцепочечные ДНК) (USB, США).

**Клонирование и секвенирование LTR-содержащих фрагментов.** ДНК космиды 28373 расщепляли эндонуклеазой рестрикции *EcoRI*, а космид 30181, 28096 и 32408 – *PstI*. Смесь рестриционных фрагментов разделяли в 1% агарозном геле и переносили на нитроцеллюлозную мембрану. LTR-содержащие фрагменты, положение которых определяли согласно результатам гибридизации с радиоактивно меченым ПЦР-зондом из области U3 LTR HERV-K, выделяли из легкоплавкой агарозы и клонировали в вектор pGEM7zf(+), расщепленный по сайтам *EcoRI* или *PstI* соответственно. Двухцепочечную форму рекомбинантных плазмид выделяли на колонках QiaGen (США) по рекомендациям производителя, денатурировали щелочным способом и секвенировали по методу Сэнгера с помощью набора “Sequenase version 2.0 sequencing kit” (USB, США). Посадку на матрицу комбинированного праймера проводили в смеси, содержащей 3 мкг матрицы, 3 мкг SSB и по 50 пкмоль каждого гексануклеотида в стандартном буфере.

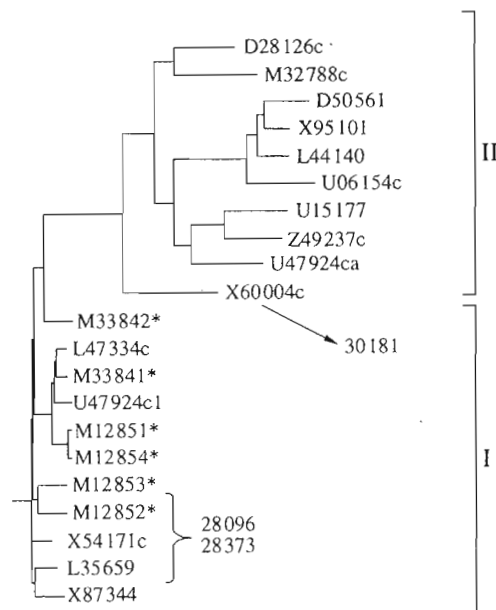


Рис. 3. Эволюционное дерево длинных концевых повторов HERV-K-элементов. Первые шесть символов наименования ветвей соответствуют идентификатору последовательностей в GenBank, в составе которых находятся LTR. Звездочкой обозначены зависимые элементы, использованные для построения дерева (например, 5'- и 3'-LTR одного и того же провируса). Фигурной скобкой указана группа последовательностей, наиболее гомологичных LTR из состава космид 28373 и 28096; стрелкой – последовательность, наиболее гомологичная LTR из космиды 30181.

**Компьютерный анализ.** Поиск последовательностей, гомологичных LTR HERV-K, в международных базах данных производился при помощи BLAST WWW-сервера (доступен через Интернет по адресу: <http://www.ncbi.nlm.nih.gov/BLAST>). Стратегия поиска состояла в повторном использовании для поиска гомологии различных обнаруженных представителей этого семейства, а также различных вариантов промежуточных консенсусных последовательностей. Последовательности, фланкирующие LTR, были использованы для поиска гомологов в международных базах данных при помощи BLAST WWW-сервера, после чего, в случае обнаружения гомологий с известными повторяющимися элементами человека, было произведено выравнивание с консенсусными последовательностями соответствующих повторяющихся элементов, представленными в RIEPBASE, rel. 5.0 (доступна через Интернет по адресу: <ftp://ftp.ncbi.nlm.nih.gov/repository/repbase>). Выравнивание последовательностей было произведено при помощи программы CLUSTALW v.1.5 [24] с использованием параметров, предлагаемых по умолчанию, и отредактировано вручную при помощи GDE 2.0 [25]. Филогенетическое дерево было построено и визуализировано с

использованием программ CLUSTALW v. 1.5 и пакета программ Phylip [26].

Авторы признательны В.К. Потапову, Н.В. Скапцовой и М.Ф. Турчинскому (ИБХ РАН) за синтез и очистку олигонуклеотидных праймеров, а также Р.А. Голубицкому (ИБХ РАН) за участие в экспериментах.

Работа поддержана грантами ННМИ 75195-544201 и LLNL W-7405-ENG-48 при частичной поддержке ГНТП "Геном человека" (договор № 58к/96).

### СПИСОК ЛИТЕРАТУРЫ

1. Steinhuber S., Brack M., Hunsmann G., Schwelberger H., Dierich M.P., Vogetseder W. // *Hum. Genet.* 1995. V. 96. P. 188–192.
2. Lower R., Lower J., Kurth R. // *Proc. Natl. Acad. Sci. USA.* 1996. V. 93. P. 5177–5184.
3. Sauter M., Schommer S., Kremmer E., Remberger K., Dolken G., Lemm I., Buck M., Best B., Neumann-Haefelin D., Mueller-Lantsch K. // *J. Virol.* 1995. V. 69. P. 414–421.
4. Widegren B., Kjellman C., Aminoff S., Sahford L., Sjogren H.-O. // *J. Gen. Virol.* 1996. V. 77. P. 1631–1641.
5. Venables P.J.W., Brookes S.M., Griffiths D., Weiss R.A., Boyd M.T. // *Virology.* 1995. V. 211. P. 589–592.
6. Lower R., Tonjes R.R., Korbmacher C., Kurth R., Lower J. // *J. Virol.* 1995. V. 69. P. 141–149.
7. Lebedev Y.B., Volik S.V., Obradovich D., Ermolaeva O.D., Ashworth L., Lennon G.G., Sverdlov E.D. // *Mol. Gen. Genet.* 1995. V. 247. P. 742–748.
8. Kieleczawa J., Dun J.J., Studier F.W. // *Science.* 1992. V. 258. P. 1787–1791.
9. Ажикина Т.Л., Шевченко Ю.О., Лебедев Ю.Б., Веселовская С.В., Мясников В.А., Потапов В.К., Сverdlov E.D. // *Докл. РАН.* 1993. Т. 330. С. 642–645.
10. Azhikina T., Veselovskaya S., Myasnikov V., Potapov V., Ermolaeva O., Sverdlov E. // *Proc. Natl. Acad. Sci. USA.* 1993. V. 90. P. 11460–11462.
11. Froehler B.C., Wadwani S., Terhorst T.J., Gerard S.R. // *Tetrahedron Lett.* 1992. V. 33. P. 5307–5310.
12. Smit A.F.A., Toth G., Riggs A.D., Jurka J. // *J. Mol. Biol.* 1995. V. 246. P. 401–417.
13. Smit A.F.A., Riggs A.D. // *Proc. Natl. Acad. Sci. USA.* 1996. V. 93. P. 1443–1448.
14. Shih C., Stoye J.P., Coffin J.M. // *Cell.* 1988. V. 53. P. 531–537.
15. Vijaya S., Steffen D.L., Robinson H.L. // *J. Virol.* 1986. V. 60. P. 683–692.
16. Kelley M.R., Kidd S., Berg R.A., Young M.W. // *Mol. Cell Biol.* 1987. V. 7. P. 1545–1548.
17. Swift R.A., Shaller E., Witter R.L., Kung H.-J. // *J. Virol.* 1985. V. 54. P. 869–872.
18. Liu Q.-R., Chan P.K. // *J. Mol. Biol.* 1990. V. 212. P. 453–459.
19. Leib-Mosch C., Haltmeier M., Werner T., Geigl E.-M., Brack-Werner R., Francke U., Erfle V., Hehlmann R. // *Genomics.* 1993. V. 18. P. 261–269.
20. Kambhu S., Falldorf P., Lee J.S. // *Proc. Natl. Acad. Sci. USA.* 1990. V. 87. P. 4927–4931.
21. Sainz J., Pevny L., Wu Y., Cantor C., Smith C. // *Proc. Natl. Acad. Sci. USA.* 1992. V. 89. P. 1080–1084.
22. Stoppa-Lyonett D., Carter P.E., Meo T., Tosi M. // *Proc. Natl. Acad. Sci. USA.* 1990. V. 87. P. 1551–1555.
23. Ericson L.M., Maeda N. // *Genomics.* 1995. V. 27. P. 531–534.
24. Thompson J.D., Higgins D.G., Gibson T.J. // *Nucl. Acids. Res.* 1994. V. 22. P. 4673–4680.
15. Smith S. // *Genetic Data Environment (GDE) version 2.0* 1995.
26. Felsenstein J. // *Cladistics.* 1989. V. 5. P. 164–166.
27. Ono M., Yasunaga T., Miyata T., Ushikubo H. // *J. Virol.* 1986. V. 60. P. 589–598.

## Structural Features of Four Long Terminal Repeats of Human Endogenous Retroviruses and Their Integration Sites

P. P. Khil<sup>1</sup>, M. B. Kostina, T. L. Azhikina, T. B. Kolesnik, Yu. B. Lebedev, and E. D. Sverdlov  
*Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya 16/10, GSP-7 Moscow, 117871 Russia*

Received January 15, 1997

**Abstract**—Four LTR-containing regions of human chromosome 19 were sequenced by the primer walking technique using strings of short oligonucleotides tightly bound to the template. A comparative and evolutionary analysis of sequences homologous to human endogenous retroviruses (HERV) was performed, and the prototypes of the LTRs were determined. Analysis of the chromosome 19 sequences adjacent to LTR revealed that LTRs of HERV-K share a common location with other retroposons.

*Key words:* LTR, HERV-K, primer walking, DNA primary structure.