



УДК 577.21

ТКАНЕСПЕЦИФИЧНОЕ МЕТИЛИРОВАНИЕ СПЕЦИФИЧНЫХ ДЛЯ ЧЕЛОВЕКА ДЛИННЫХ КОНЦЕВЫХ ПОВТОРОВ ЭНДОГЕННЫХ РЕТРОВИРУСОВ

© 2004 г. К. В. Ходосевич[#], Ю. Б. Лебедев, Е. Д. Свердлов

Институт биоорганической химии им. акад. М.М. Шемякина и Ю.А. Овчинникова РАН,
117997 ГСП, Москва, В-437, ул. Миклухо-Макляя, 16/10

Поступила в редакцию 23.07.2003 г. Принята к печати 27.10.2003 г.

Возможное участие ретроэлементов в эпигенетической регуляции экспрессии генов человека рассматривается на примере метилирования длинных концевых повторов (LTR) эндогенных ретровирусов семейства К (HERV-K). Определен статус метилирования шести LTR эндогенных ретровирусов человека в различных участках генома человека, обогащенных генами. Показано, что метилирование четырех LTR имеет явно выраженный тканеспецифичный характер. Полученные результаты коррелируют с опубликованными данными о тканеспецифичных изменениях уровня экспрессии генов человека, находящихся в пределах возможных регуляторных взаимодействий с исследованными LTR.

Ключевые слова: метилирование, эндогенные ретровирусы, экспрессия генов, эволюция генома человека.

ВВЕДЕНИЕ

Геном человека содержит десятки тысяч эндогенных ретровирусов и их длинных концевых повторов (LTR) [1]. Эти элементы представляют собой потенциальные регуляторные элементы, включающие в себя многие регуляторы транскрипции – промоторы, энхансеры, элементы отклика на гормоны и сигналы полиаденилирования. Показано, что LTR являются функциональными промоторами некоторых генов человека [2, 3]. Результаты экспериментов по экспрессии репортерных генов подтверждают наличие выраженной тканеспецифичной промоторной и энхансерной активности LTR-элементов различных семейств [4–8]. Активность LTR в сочетании с их способностью внедряться в процессе эволюции вблизи генов и таким образом менять регуляцию этих генов делает LTR перспективными кандидатами на роль факторов, влияющих на процесс видообразования. Активность LTR в геноме зависит от многих факторов, одним из которых может являться статус метилирования.

Метилирование цитозинового основания ДНК и модификации гистонов, определяющие компактизацию хроматина, рассматриваются как два основных механизма эпигенетической регуляции

экспрессии генов. Считается, что в ДНК неэмбриональных клеток млекопитающих приблизительно 80% динуклеотидов CpG метилировано [9]. Вместе с тем, распределение CpG по геному неслучайно, так что при сравнительной обедненности последовательности ДНК динуклеотидами CpG, большинство из них сосредоточено в небольших участках, называемых CpG-островками. Эти участки обычно деметилированы в нормальных тканях, а их метилирование в опухолевых клетках может приводить к подавлению экспрессии генов-супрессоров опухолеобразования. Эффекты изменения экспрессии тканеспецифичных генов вследствие метилирования инициируются, по-видимому, связыванием специфических белков с метилированными динуклеотидами CpG, которое приводит, в конечном счете, к подавлению экспрессии гена, содержащего данный CpG-островок. Считается, что изменение статуса метилирования ДНК в различных тканях происходит не по отдельным CpG-сайтам, а затрагивает протяженные участки генома – так называемые блоки метилирования [9]. Предполагается, что выключение LTR за счет метилирования служит средством защиты клеток от нерегулируемой экспрессии множества генов, которая могла бы происходить за счет промоторно-энхансерных функций LTR. Деметилированность LTR, с другой стороны, может означать его функциональную активность. Поэтому исследование статуса метилирования LTR имеет существенное значение для выво-

Сокращения: LTR (Long Terminal Repeat) – длинный концевой повтор; HERV-K (HML-2) – эндогенные ретровирусы человека семейства К, подсемейства HML-2.

[#] Автор для переписки (тел.: (095) 330-63-29; факс: (095) 330-65-38; эл. почта: kosthob@humgen.siobc.ras.ru).

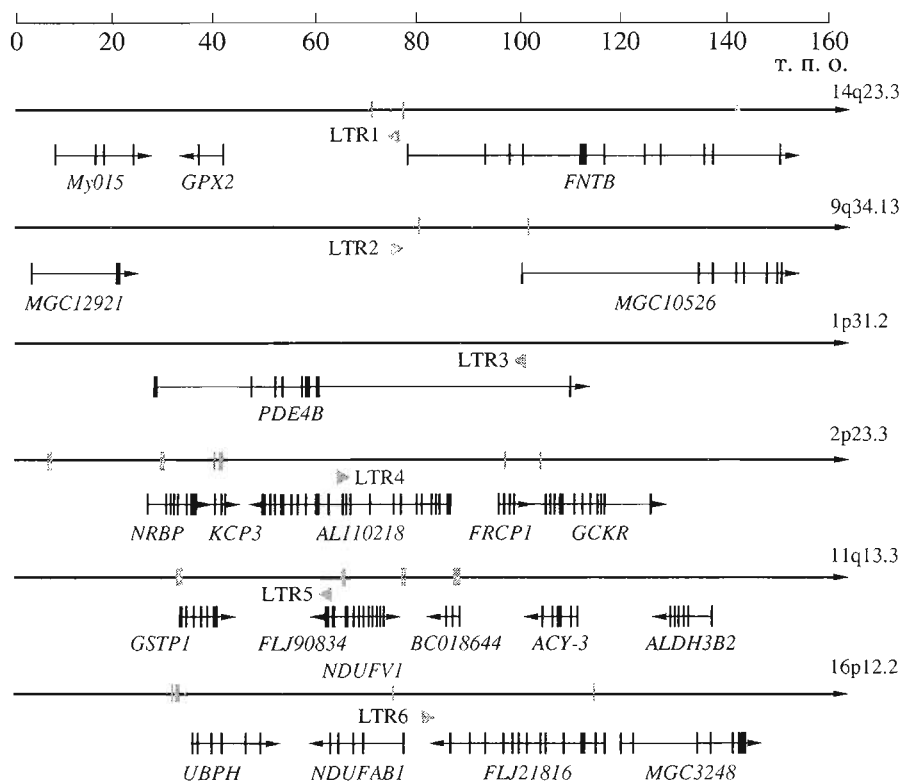


Рис. 1. Схема строения локусов, содержащих исследуемые LTR. Стрелками с соответствующими названиями локусов справа обозначены участки хромосом. Направление стрелок соответствует ориентации локуса в направлении к теломере. Светлые прямоугольники на стрелках соответствуют положению CpG-островков. Тонкие стрелки под схемами локусов обозначают направление транскрипции генов, названия которых приведены под ними. Положения экзонов отмечено черными вертикальными прямоугольниками на стрелках. Светлыми стрелками над схемами генов обозначены положения и U3–U5-направления LTR. Нумерация LTR соответствует использованной в таблицах и в тексте. Над схемами локусов и генов приведена шкала в тысячах пар оснований (т.п.о.).

дов об их функциональном статусе и о возможной роли в регуляции экспрессии близлежащих генов.

В настоящей работе нами впервые предпринята попытка выявить изменения статуса метилирования участков внедрения LTR семейства эндогенных ретровирусов HERV-K в различных тканях человека.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Для анализа статуса метилирования LTR в различных тканях была выбрана группа одиночных LTR семейства HERV-K (HML-2), специфичных для генома человека и, таким образом, потенциально могущих быть вовлеченными в процесс разделения эволюционных линий человека и шимпанзе. Другим критерием для отбора служила локализация последовательностей LTR максимально близко к нескольким известным или предсказанным генам. Ранее с использованием различных методов мы идентифицировали и картировали приблизительно 150 LTR HERV-K (HML-2), отличающих геном человека от генома шимпанзе [10, 11]. Анализ геномного окружения показал, что

около 30% участков человекеспецифичных интеграций LTR расположены не далее 50 т.п.о. от различных генов человека, включая 12 LTR, локализованных в интронных областях [12, 13]. При дополнительном анализе из данной группы LTR с помощью программ Human Genome Draft Browser (см. “Эксперимент. часть”) были отобраны 6 представителей, расположенные в различных обогащенных генами локусах генома человека.

Схематически локусы, содержащие исследованные LTR, приведены на рис. 1. Три LTR, условно обозначенные как LTR4, LTR5 и LTR6, расположены в локусах 2p23.3, 11q13.3 и 16p12.2, содержащих от четырех до шести генов с короткими межгенными областями на участках ДНК менее 100 т.п.о. LTR3 локализован в последнем интроне гена сAMP-специфичной фосфоэстеразы типа 4B. LTR1 и LTR2 находятся в межгенных областях в непосредственной близости от CpG-островков. Повышенная плотность фрагментов ДНК, обогащенных CpG-повторами, отмечена для пяти из шести рассматриваемых локусов (см. рис. 1).

Определение статуса метилирования изучаемых LTR проводили с использованием метода

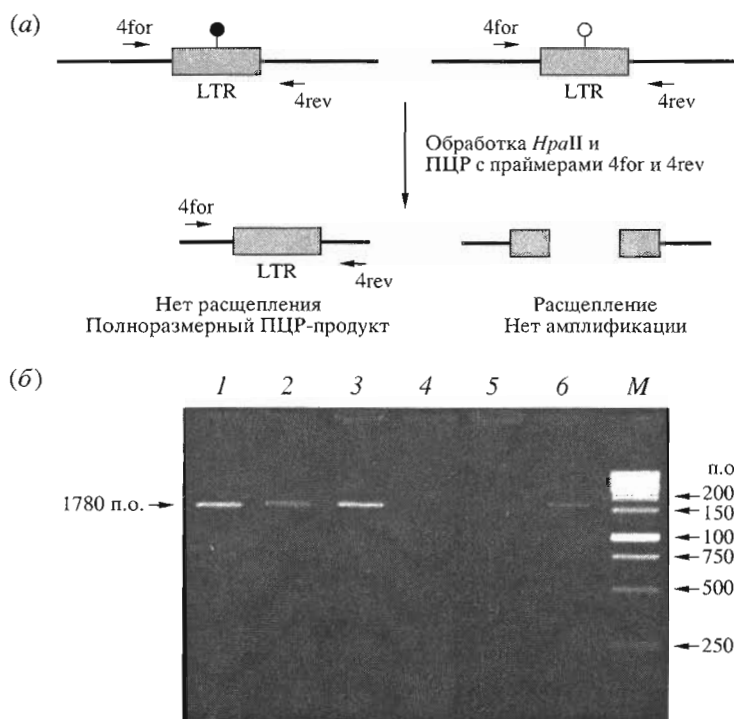


Рис. 2. Определение статуса метилирования индивидуальных LTR. (а) Схема эксперимента. Черным и светлым кружками показаны соответственно метилированный и неметилированный сайты расщепления *HpaII* в LTR. (б). Пример ПЦР-амплификации фрагмента гена *AL110218* из локуса 2p23.3, содержащего LTR4. Электрофоретическое разделение в 1% агарозном геле продуктов ПЦР с праймерами 4for и 4rev (LTR4, табл. 1) после амплификации ДНК из печени, мозжечка, лимфоузла, почки, желудка, легких (дорожки 1, 2, 3, 4, 5, 6 соответственно). М – маркер длин ДНК “1 kb DNA Ladder” (Сибэнзим, Россия). Стрелкой указаны продукты ПЦР-амплификации. В соответствующих контрольных экспериментах с конститутивно экспрессирующимися генами *WT1*, *RBI*, *BRCA1* и *CDKN1A* после расщепления *HpaII* амплификации не происходило (данные не показаны).

ПЦР-амплификации фрагментов геномной ДНК, гидролизованной чувствительной к метилированию эндонуклеазой рестрикции *HpaII* (рис. 2а). Метод основан на использовании пар праймеров (табл. 1), между участками связывания которых в ДНК находится сайт *HpaII*. В том случае, если этот сайт неметилирован, ДНК будет расщепляться и, вследствие этого, амплификация соответствующего фрагмента происходить не будет. Если же сайт метилирован, то расщепления происходить не будет, и ДНК будет амплифицироваться. Для анализа метилированности были использованы препараты ДНК из различных тканей, включая мозжечок, печень, почку, лимфатический узел, желудок и легкое.

Полноту гидролиза образцов ДНК проверяли с помощью контрольных локуспецифичных ПЦР. Для проведения таких ПЦР были использованы 4 пары праймеров, соответствующих уникальным геномным последовательностям, которые фланкируют участки узнавания *HpaII* в конститутивно неметилированных фрагментах генов “домашнего хозяйства” *WT1*, *RBI*, *BRCA1* и *CDKN1A*. Отсутствие продукта ПЦР в результате амплификации ДНК, обработанной *HpaII*, при одновре-

менном образовании продукта ПЦР на матрице исходной ДНК свидетельствовало о том, что рестрикция прошла полностью.

Все использованные в исследовании LTR содержат единственный участок узнавания *HpaII*. Для анализа метилированности этого сайта в различных тканях человека мы проводили ПЦР с праймерами, соответствующими участкам генома, фланкирующим изучаемые LTR (см., например, рис. 2б). В данном случае, ПЦР-фрагмент ожидаемой длины (1780 п.о.) был получен при амплификации гидролизованной *HpaII* геномной ДНК из клеток печени, мозжечка, лимфатического узла и легких, что доказывает метилированное состояние данного LTR в этих тканях. Напротив, LTR4 неметилирован в ДНК из почек и желудка, поскольку локуспецифичный продукт амплификации был получен только на необработанной *HpaII* геномной ДНК (данные не приводятся).

Аналогичный анализ был проведен для всех шести исследуемых LTR. Результаты анализа представлены в табл. 2. Как видно из этой таблицы, большая часть исследованных LTR метилирована в различных тканях человека, за исключением мозжечка. Это согласуется с распространенным

Таблица 1. Праймеры для ПЦР, соответствующие участкам генома, фланкирующим исследованные LTR HERV-K

LTR	Регистрационный номер*	Название праймера	Последовательность праймера 5' → 3' (длина, нт)
LTR1	AL139022	1for	G TGGTAGAAACATGAGCTGTCCA (23)
		1rev	C ACTTAAGACCCCTCTAAGACTG (23)
LTR2	AL354855	2for	C ACGTGTGGACATAAGCAACC(21)
		2rev	G CTGAGTCGTCCTGATTCTTG (21)
LTR3	AL391820	3for	G CAGCATTTGTCATCTTGGG (19)
		3rev	T GCACCCTTGCCTTTCC (17)
LTR4	AC074117	4for	C CCATACCTGGCTTCGATG (19)
		4rev	C CAGTGGCTGATGGACACAC (20)
LTR5	AP003385	5for	C ATACAAGCTGGATCTGCACC (21)
		5rev	A TCGTGCCTTACTCACACTGG(21)
LTR6	AC008870	6for	C ATCTGCTTGACACAGTAGGTATG (24)
		6rev	G TCCCAATATTAAGATCACATCCT (24)

* Приведены регистрационные номера (Accession numbers) в международной базе данных GeneBank нуклеотидных последовательностей фрагментов генома человека, содержащих соответствующие LTR.

мнением, что клетка уменьшает вред, наносимый новыми интеграциями мобильных элементов, используя метилирование как один из способов их репрессии [14]. Вместе с тем, 4 из 6-ти исследованных LTR имеют разный статус метилирования в различных тканях, что может указывать на их участие в функционировании локусов, где они расположены.

LTR5 расположен на расстоянии 3 т.п.о. от первого экзона гена *NDUFV1* (рис. 1) важнейшей субъединицы NADH:убихинон-оксидоредуктазы, фермента, иначе называемого комплексом I дыхательной цепи митохондрий. Эта субъединица содержит в себе сайт связывания NADH, и мутации в гене этого белка приводят к значительным нарушениям работы NADH:убихинон-оксидоредуктазы [15]. LTR в этом локусе неметилирован в клетках всех исследованных тканей, кроме клеток печени. LTR6 в локусе 16p12.2 неметилирован в клетках печени, мозжечка, лимфоузла и легких и метилирован в почке и желудке. Данный

Таблица 2. Статус метилирования исследованных специфичных для человека LTR в различных тканях

LTR	Печень	Мозжечок	Лимф. узел	Почка	Желудок	Легкие
LTR1	+	+	+	+	+	+
LTR2	+	+	+	+	+	+
LTR3	+	-	+	+	+	+
LTR4	+	+	+	-	-	+
LTR5	+	-	-	-	-	-
LTR6	-	-	-	+	+	-

Примечание: * (+) – метилированы, (-) – неметилированы.

LTR находится в обогащенном генами участке генома, в окружении генов *NDUFAB1*, *UBPH*, *MGC3248*, а также нескольких других генов с неизвестными функциями (рис. 1). Ген *NDUFAB1* кодирует еще одну субъединицу NADH:убихинон-оксидоредуктазы [16]. Митохондриальное дыхание – один из основных внутриклеточных процессов, обеспечивающих клетку энергией. Дифференциальное метилирование LTR5 и LTR6, интегрировавшихся вблизи генов, продукты которых обеспечивают митохондриальное дыхание, а также неметилированный статус обоих LTR в клетках печени, где усилена экспрессия этих генов [17], может свидетельствовать об использовании данных LTR в регуляции экспрессии генов *NDUFAB1* и *NDUFV1*.

LTR3 в локусе 1p31.2 метилирован во всех исследованных тканях, кроме мозжечка. Он находится в последнем интроне гена сАМР-специфичной фосфодиэстеразы типа 4В (*PDE4B*). Изозимы PDE4 (A, B, C, D) участвуют в процессах контроля настроения, рвотного рефлекса и передачи сигнала от обонятельных рецепторов [18]. В каждом из изученных отделов головного мозга преобладает один, специфичный для этого отдела, изозим PDE4. Ранее было показано, что изозим PDE4B преобладает в коре мозга, гиппокампе и мозжечке [19]. Экспрессия гена *PDE4B* увеличена в мозжечке человека по сравнению с мозжечком макаки и крысы [18]. Кроме того, ген *PDE4B* в Т-клетках периферической крови транскрибируется в двух вариантах – полноразмерном и без последнего экзона [20, Human Genome Draft Browser]. Возможно данный LTR каким-либо образом принимает участие в этой регуляции.

Встречаются и такие LTR, которые метилированы во всех исследованных тканях, например LTR1, находящийся на участке хромосомы 14q23.3 в промоторной области гена *FNTB* (бета-субъединицы фарнезилтрансферазы). Скорее всего, такие LTR не имеют функционального значения для исследованных тканей. Однако при определенных условиях (например, при тепловом шоке) нельзя исключить деметилирования подобных LTR (как это, например, происходит в случае других ретроэлементов – *Alu* [21]) и активации их регуляторных элементов.

Полученные данные являются первым шагом на пути анализа функционального значения статуса метилирования LTR. В основном метилированность/неметилированность LTR коррелирует с отсутствием/наличием экспрессии близлежащих генов. Возможно, LTR вовлечены в регуляцию их экспрессии. В последнее время широко обсуждается вопрос участия LTR в молекулярной эволюции приматов [22–24]. Влияние деметилированных человекспецифичных LTR на экспрессию соседних генов, участвующих в митохондриальном дыхании, контроле настроения, передачи сигнала от обонятельных рецепторов и др., вместе с другими факторами могло играть определенную роль в эволюции человека.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Выделение геномных ДНК человека из различных тканей проводили по методике выделения высокомолекулярной ДНК с помощью протеиназы K и фенола [25].

Нуклеотидные последовательности и их анализ. Анализ геномного окружения LTR осуществляли при помощи пакетов программ Draft Human Genome Browser (<http://genome.ucsc.edu/goldenPath/hgTracks.html>), RepeatMasker (<http://ftp.genome.washington.edu/>) и BLAST2 (<http://www.ncbi.nlm.nih.gov/BLAST/>). Подбор праймеров проводили с помощью программы Gene Runner (Version 3.00).

Гидролиз геномных ДНК рестриктазой *HpaII* проводили в 100 мкл инкубационной смеси, содержащей 1 мкг геномной ДНК, 30 ед. акт. *HpaII* (NEB, США), 10 мМ Трис-НСl, 10 мМ MgCl₂ и 1 мМ дитиотреит (рН 7.0 при 25°C). Смесь инкубировали при 37°C в течение 16 ч и инактивировали фермент путем инкубации при 65°C в течение 15 мин.

Геномная ПЦР. В качестве матрицы использовали геномную ДНК человека, обработанную и не обработанную эндонуклеазой рестрикции *HpaII* (ДНК, не обработанную *HpaII*, добавляли для контроля ПЦР). В инкубационную смесь для ПЦР, содержащую буфер (50 мМ KCl, 10 мМ Трис-НСl, рН 9.0, 0.1% Тритон X-100), 1.5 мМ MgCl₂ 0.125 мМ dNTPs, 0.4 мкМ праймеры и 10 нг геномной ДНК

добавляли 0.5 ед. акт. *Taq*-ДНК-полимеразы (ИБХ, Россия). Суммарный объем реакционной смеси – 25 мкл. Проводили 28 циклов ПЦР в режиме: 95°C – 20 с, 60°C – 30 с, 72°C – 110 с. Из реакционной смеси отбирали аликвоты по 8 мкл для анализа продуктов ПЦР электрофорезом в 1% агарозном геле.

Авторы выражают благодарность В.К. Потопову и Н.В. Скапцовой за синтез олигонуклеотидов. Исследования проведены при поддержке грантами РФФИ 02-04-48614, INTAS-01-0759 и контрактом 43.073.1.1.2508 Министерства промышленности, науки и технологий РФ.

СПИСОК ЛИТЕРАТУРЫ

1. International Human Genome Sequencing Consortium // *Nature*. 2001. V. 409. P. 860–921.
2. Di Cristofano A., Strazullo M., Longo L., La Mantia G. // *Nucl. Acids Res.* 1995. V. 23. P. 2823–2830.
3. Landry J.R., Rouhi A., Medstrand P., Mager D.L. // *Mol. Biol. Evol.* 2002. V. 19. P. 1934–1942.
4. Domansky A.N., Kopantzev E.P., Snezhkov E.V., Lebedev Y.B., Leib-Mosch C., Sverdlov E.D. // *FEBS Lett.* 2000. V. 472. P. 191–195.
5. Доманский А.Н., Акопов С.Б., Лебедев Ю.Б., Нуколаев Л.Г., Сverdlov E.D. // *Биоорганическая химия*. 2002. Т. 28. С. 341–345.
6. Schon U., Seifarth W., Baust C., Hohenadl C., Erfle V., Leib-Mosch C. // *Virology*. 2001. V. 279. P. 280–291.
7. Patzke S., Lindskog M., Munthe E., Aasheim H.C. // *Virology*. 2002. V. 303. P. 164–173.
8. Ling J., Pi W., Bollag R., Zeng S., Keskinetepe M., Saliman H., Krantz S., Whitney B., Tuan D. // *J. Virol.* 2002. V. 76. P. 2410–2423.
9. Jaenisch R., Bird A. // *Nat. Genet.* 2003. V. 33. P. 245–254.
10. Buzdin A., Khodosevich K., Mamedov I., Vinogradova T., Lebedev Y., Hunsmann G., Sverdlov E. // *Genomics*. 2002. V. 79. P. 413–422.
11. Mamedov I., Batrak A., Arzumanyan E., Lebedev Y., Sverdlov E.D. // *Nucleic Acids Res.* 2002. V. 30. P. e71.
12. Буздин А.А., Лебедев Ю.Б., Сverdlov E.D. // *Биоорганическая химия*. 2003. Т. 29. С. 103–106.
13. Buzdin A., Ustyugova S., Khodosevich K., Mamedov I., Lebedev Y., Hunsmann G., Sverdlov E. // *Genomics*. 2003. V. 81. P. 149–156.
14. Kidwell M.G., Lisch D. // *Proc. Natl. Acad. Sci. USA*. 1997. V. 94. P. 7704–7711.
15. Schuelke M., Smeitink J., Mariman E., Loeffen J., Plecko B., Trijbels F., Stockler-Ipsiroglu S., van den Heuvel L. // *Nat. Genet.* 1999. V. 21. P. 260–261.
16. Triepels R., Smeitink J., Loeffen J., Smeets R., Buskens C., Trijbels F., van den Heuvel L. // *J. Inher. Metab. Dis.* 1999. V. 22. P. 163–173.
17. de Coo R.F., Buddiger P.A., Smeets H.J., van Oost B.A. // *Mamm. Genome*. 1999. V. 10. P. 49–53.
18. Perez-Torres S., Miro X., Palacios J.M., Cortes R., Puigdomenech P., Mengod G. // *J. Chem. Neuroanat.* 2000. V. 20. P. 349–374.

19. D'Sa C., Tolbert L.M., Conti M., Duman R.S. // *J. Neurochem.* 2002. V. 81. P. 745–757.
20. Baroja M.L., Cieslinski L.B., Torphy T.J., Wange R.L., Madrenas J. // *J. Immunol.* 1999. V. 162. P. 2016–2023.
21. Chu W.-M., Ballard R., Carpick B.W., Williams B.G., Schmid C.W. // *Mol. Cell. Biol.* 1998. V. 18. P. 58–68.
22. Sverdlov E.D. // *Bioessays.* 2000. V. 22. P. 161–171.
23. Sverdlov E.D. // *FEBS Lett.* 1998. V. 428. P. 1–6.
24. Khodosevich K., Lebedev Y., Sverdlov E. // *Comp. Funct. Genom.* 2002. V. 3. P. 494–498.
25. Sambrook J., Russell D. *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor. New York; Cold Spring Harbor Laboratory Press, 2001.

The Tissue-Specific Methylation of Human-Specific Endogenous Retroviral LTRs

K. V. Khodosevich[#], Y. B. Lebedev, and E. D. Sverdlov

[#] Phone: +7 (095) 330-6329; fax +7 (095) 330-6538; e-mail: kosthob@humgen.siobc.ras.ru
Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences,
ul. Miklukho-Maklaya 16/10, Moscow, 117997 Russia

A possible involvement of retroelements in the epigenetic regulation of human gene expression was considered by the example of methylation of long terminal repeats (LTRs) of the human endogenous retrovirus family K (HERV-K). The methylation status of six HERV-K LTRs was determined in various gene-enriched regions of the human genome. The methylation of four LTRs was shown to be tissue-specific. Our results correlated with published data on the tissue-specific changes in the expression level of human genes adjacent to the LTRs under study. The English version of the paper: *Russian Journal of Bioorganic Chemistry*, 2004, vol. 30, no. 5; see also <http://www.maik.ru>.

Key words: endogenous retroviruses, gene expression, human genome evolution, methylation