



УДК 575.117.2

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ОБЛАСТЕЙ РЕГУЛЯЦИИ ГЕНА ТРАНСТИРЕТИНА В ГЕНОМАХ МЫШИ, ЧЕЛОВЕКА И ШИМПАНЗЕ

© 2004 г. Е. В. Надеждин[#], Т. В. Виноградова, Е. Д. Свердлов

Институт биоорганической химии им. М.М. Шемякина и Ю.А. Овчинникова РАН,

117997 ГСП, Москва, В-437, ул. Миклухо-Маклая, 16/10

Поступила в редакцию 29.01.2004 г. Принята к печати 09.02.2004 г.

При полногеномном анализе различий в экспрессии генов в мозге человека и шимпанзе найдено, что ген переносчика тиреоидных гормонов транстиретина дифференциально транскрибируется в мозжечках этих организмов. Для выявления возможных регуляторных последовательностей, определяющих разницу в экспрессии, мы определили первичную структуру интересующего нас фрагмента ДНК шимпанзе длиной 7 т.п.о. В последовательности шимпанзе по сравнению с человеком было выявлено более 113 замен, при этом ~40% замен находились внутри повторяющихся элементов генома, положение и размер которых в исследуемой области совпадают у человека и шимпанзе, а нуклеотидные последовательности имеют высокую степень идентичности. Сравнение нуклеотидных последовательностей данной области генов транстиретина человека, шимпанзе и мыши обнаружило существенное различие в распределении G+C-состава вдоль исследуемого фрагмента у человека (шимпанзе) и мыши и позволило локализовать три участка, имеющих более высокую степень идентичности у всех трех видов. Один из этих участков находится в области промотора гена, два других, возможно, определяют специфичность экспрессии гена транстиретина в печени и мозге. В одном из консервативных участков в геноме шимпанзе обнаружены одна единичная нуклеотидная замена и тройная нуклеотидная замена, которая отличает шимпанзе от человека и мыши, имеющих в этом месте идентичные последовательности. Возможно, что эти замены могут быть причиной различий в уровне экспрессии гена транстиретина в мозге человека и шимпанзе.

Ключевые слова: транстиретин, геном шимпанзе, регуляция экспрессии генов.

ВВЕДЕНИЕ

С момента разделения эволюционных ветвей человека (*Homo sapiens*) и шимпанзе (*Pan troglodytes* и *Pan paniscus*), произошедшего около 5 млн. лет назад, между этими видами образовалось множество фенотипических различий. Среди них такие фундаментальные, как способность к усвоению информации, наличие у человека речи и разная восприимчивость к болезням, в частности к психическим расстройствам. Вопрос о генетических основах этих различий представляет собой важнейшую фундаментальную проблему, решение которой поможет понять, какие факторы определяют основные характеристические признаки, отличающие человека от других видов животных.

Данные о высокой степени идентичности геномных последовательностей в среднем (95%) и кодирующих областей генома (98%) человека и шимпанзе позволяют предположить, что в основе фенотипических различий этих двух видов лежит сравнительно небольшое число изменений в регуляторных системах, в частности в регуляции экспрессии генов, и, прежде всего, генов, вовлеченных в развитие организма [1, 2].

Несмотря на фундаментальную важность этой проблемы, различия в экспрессии генов человека и шимпанзе исследованы недостаточно. Сегодня опубликованы данные лишь об относительно небольшом числе генов, по-разному экспрессирующихся у двух видов [3, 4]. Исследуя такие дифференциально экспрессирующиеся гены, мы обнаружили существенное различие в уровне экспрессии гена транстиретина – переносчика тиреоидных гормонов в мозге человека и шимпанзе [5]. Наши данные по транскрипции гена транстиретина в мозге согласуются с результатами, полученными практически одновременно в группе А. Варки [6] при сравнительном ана-

Сокращения: *Ttr* – ген транстиретина; *Enh_h* – энхансер, обеспечивающий экспрессию гена транстиретина в печени; *Enh_{cp}* – гипотетический энхансер, обеспечивающий экспрессию гена транстиретина в сосудистом сплетении мозга.

[#] Автор для переписки (тел.: (095) 330 69 92; факс: (095) 330 65 38; эл. почта: neugene@humgen.siobc.ras.ru).

лизе белков из цереброспинальной жидкости человека и шимпанзе.

Транстиретин отвечает за связывание тиреоидных гормонов и ретинола и их транспорт из плазмы крови в мозг. Он привлекает особое внимание, поскольку тиреоидные гормоны являются важнейшими регуляторами развития, дифференцировки и функционирования мозга. Разница в содержании этого белка может иметь весьма существенное значение для функций мозга. Ввиду этого изменение регуляции транспорта и биосинтеза тиреоидных гормонов с высокой вероятностью может быть одним из факторов, определивших возникновение эволюционно значимых различий человека и шимпанзе.

В данной работе мы предприняли попытку сравнить регуляторные последовательности гена транстиретина человека и шимпанзе, разница в структуре которых может определять наблюдаемые различия в уровне экспрессии транстиретина у этих видов. С этой целью мы установили первичную структуру интересующего нас фрагмента ДНК шимпанзе длиной 7 т.п.о. (последовательность депонирована в GenBank AY505106 и AY530327) и провели сравнение этой последовательности с известными последовательностями соответствующих участков генов транстиретина человека и мыши.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

У млекопитающих ген транстиретина (*Ttr*) экспрессируется преимущественно в хороидном плексусе (сосудистом сплетении мозга) и в печени [7]. У человека для экспрессии *Ttr* необходим участок генома длиной 6 т.п.о. [8], прилегающий к промотору гена *Ttr* с 5'-конца, тогда как у мыши, по имеющимся данным [8, 9], гомологичный участок составляет ~3 т.п.о. В настоящее время из регуляторных элементов, определяющих тканеспецифичность экспрессии гена *Ttr*, экспериментально определено только положение энхансерной последовательности, отвечающей за экспрессию транстиретина в печени мышей [9–11]. Структура и локализация регуляторных элементов гена *Ttr* человека и высших обезьян неизвестны.

Геном шимпанзе полностью не секвенирован (в настоящее время появился несовершенный черновик последовательности on-line <http://genome.ucsc.edu/cgi-bin/hgGateway>). Мы определили первичную структуру участка генома шимпанзе длиной 7 т.п.о. (последовательность депонирована в GenBank AY505106 и AY530327), соответствующего областям гена транстиретина человека и мыши, содержащим потенциальные регуляторные элементы.

Ниже приведены результаты сравнения этой последовательности с соответствующей последовательностью генома человека. Средняя степень идентичности последовательностей составила 98.2%. Эти данные хорошо согласуются с опубликованными ранее результатами сравнения различных участков геномов человека и шимпанзе [12, 13]. В исследованной последовательности шимпанзе по сравнению с соответствующим участком генома человека было выявлено: 104 единичные нуклеотидные замены, четыре динуклеотидные и две тринуклеотидные замены, две динуклеотидные делеции и одна делеция семи нуклеотидов.

Из вышеперечисленных нуклеотидных изменений ~40% замен находятся внутри повторяющихся элементов генома, местоположение и размеры которых в исследуемой области совпадают у человека и шимпанзе, а нуклеотидные последовательности имеют высокую степень идентичности.

Для выявления потенциальных регуляторных элементов гена *Ttr* было проведено сравнение нуклеотидных последовательностей исследуемых 7 т.п.о. – участков генома человека и шимпанзе с нуклеотидной последовательностью аналогичного участка генома мыши, для которого известна локализация промотора и энхансера, обеспечивающего экспрессию гена *Ttr* в печени [10]. Анализ основывался на принятой гипотезе о более высокой консервативности функционально значимых участков генома в эволюции.

При сравнении обнаружено существенное различие в распределении G+C-состава вдоль исследуемого фрагмента у человека (и шимпанзе) и мыши (рис. 1), что дает основания считать, что в процессе эволюции этот участок подвергался значительным преобразованиям. Изменение в распределении G+C-состава привело к существенному различию в распределении CpG-динуклеотидов внутри этих участков у мыши и человека.

Поскольку CpG-динуклеотиды являются мишенью метилирования геномной ДНК, различие в их распределении наводит на мысль о различии эпигенетических факторов, регулирующих экспрессию гена у грызунов и высших приматов. В частности, на расстоянии примерно 800 п.о. от точки начала транскрипции у человека и шимпанзе обнаруживается участок с повышенным, по сравнению с геномом мыши, содержанием G+C и CpG. Возможно, этот участок в геноме человека играет роль CpG-островка, существенного для регуляции экспрессии генов [14]. Другой потенциальный CpG-островок, отсутствующий в последовательности мыши, находится на расстоянии примерно 4800 п.о. от точки начала транскрипции в участ-

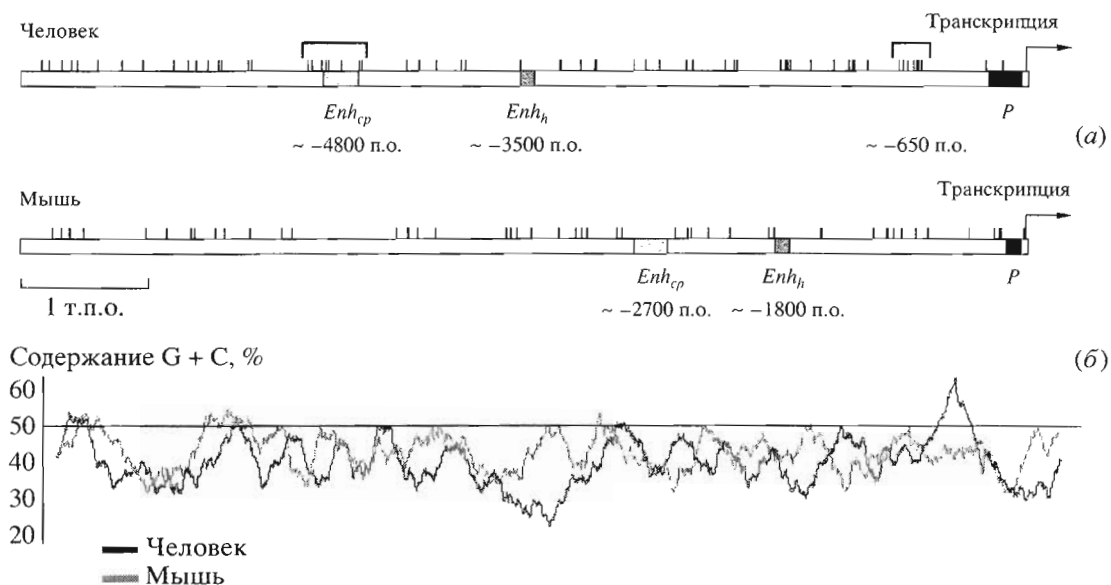


Рис. 1. Схема распределения регуляторных элементов, CpG-островков (а) и G+C-состава (б) в участках генома (7 т.п.о.) человека и мыши (номера локусов в GenBank AC017100 и AC129078), прилегающих к 5'-концам генов транстиретина. (а): Вертикальными линиями обозначены положения динуклеотидов CpG. CpG-островки в последовательности человека обозначены горизонтальными скобками. Стрелкой показано направление транскрипции гена транстиретина. Темными прямоугольниками отмечены экспериментально определенные [10] для мыши промотор (P) и энхансер (Enh_h), обеспечивающий экспрессию транстиретина в печени, а также участки генома человека, имеющие высокую степень идентичности с этими регуляторными элементами генома мыши. Серыми прямоугольниками показаны участки генома человека и мыши, имеющие высокую степень идентичности (85%) и предположительно содержащие энхансер, обеспечивающий экспрессию в хорoidalном плексусе (Enh_{cp}). Цифрами обозначено расстояние от потенциальных регуляторных элементов до точки начала транскрипции гена *Ttr*.

ке, консервативном для геномов мыши и человека (см. ниже).

В результате сравнительного анализа было локализовано три области (рис. 1), имеющих, по сравнению с остальной последовательностью исследуемого участка, более высокую степень идентичности:

1. Область промотора длиной ~200 п.о. В геномах человека и мыши степень идентичности этого участка – 85%. Далее гомология между последовательностями человека и мыши практически теряется. Степень идентичности последовательностей промотора человека и шимпанзе составляет 100%. Здесь находятся базовые элементы проксимального промотора, такие, как ТАТА-бокс, САТ-бокс и гипотетический, частично дублированный элемент связывания гормонов. В силу полной идентичности, эта область не может определять различия в уровне транскрипции генов *Ttr* у человека и шимпанзе.

2. Область длиной ~100 п.о., гомологичная фрагменту гена транстиретина мыши, который содержит энхансер (Enh_h), определяющий экспрессию гена транстиретина в печени (рис. 2). Степень идентичности этого участка у мыши и человека составляет ~60%, что достоверно выше степени идентичности остальных проанализиро-

ванных нами последовательностей человека и мыши на протяжении 7 т.п.о. данной последовательности (~50%).

Было экспериментально показано, что этот энхансер гена транстиретина мыши содержит в себе участки связывания факторов транскрипции С/ЕВР, АР1, HNF-3, которые влияют на уровень экспрессии гена [10, 11]. Анализ соответствующего фрагмента гена транстиретина человека показал, что участки связывания этих факторов транскрипции в ходе эволюции претерпели существенные изменения.

Степень идентичности этой области у человека и шимпанзе составляет 98.2%. Нами были идентифицированы две замены в геноме шимпанзе по сравнению с геномом человека, но обе замены находятся вне участков связывания с транскрипционными факторами. Таким образом, по-видимому, этот участок также не может быть ответствен за различия в экспрессии гена у человека и шимпанзе.

3. Область длиной ~150 п.о., со степенью идентичности для человека и мыши около 85% (рис. 3). Если обсуждаемый участок действительно является регулятором транскрипции, то, поскольку энхансер транстиретина мыши Enh_h локализован в другом месте, можно предположить, что этот



Рис. 2. Выравнивание нуклеотидных последовательностей человека и шимпанзе с фрагментом генома мыши, содержащим энхансер *Enh_h*. Жирным шрифтом выделена последовательность экспериментально определенного мышино-го энхансера *Enh_h*. В рамки заключены экспериментально определенные для мыши участки связывания с транскрипционными факторами C/EBP, HNF-3, AP1. Серым фоном закрашены идентичные последовательности.



Рис. 3. Выравнивание области предполагаемого энхансера *Enh_{cp}*, обеспечивающего экспрессию гена *Tir* в сосудистом сплетении мозга человека, шимпанзе и мыши. Жирным шрифтом выделены последовательности, имеющие для трех геномов высокую степень идентичности (85%) и предположительно содержащие энхансер *Enh_{cp}*. Серым фоном закрашены идентичные последовательности.

фрагмент вовлечен в регуляцию экспрессии в сосудистом сплетении мозга. Анализ методами биоинформатики (использованы on-line-программы, расположенные на интернет-сайте компании Genomatix) показал, что на этом участке имеется много потенциальных сайтов связывания с различными транскрипционными факторами. Среди них такие известные, как факторы транскрипции, со-

державшие гомео- и POU-домены, принимающие участие в развитии и дифференцировке.

Степень идентичности этой области у человека и шимпанзе составляет 98.1%: в геноме шимпанзе находятся одна единичная нуклеотидная замена и тройная нуклеотидная замена, которая отличает шимпанзе от человека и мыши, имеющих в этом месте идентичные последова-

тельности. Возможно, что идентифицированные нами у шимпанзе нуклеотидные замены могут быть причиной различий в уровне экспрессии гена транстиретина в мозге человека и шимпанзе.

В заключение следует заметить, что вопрос о местоположении регуляторов, определяющих различия в экспрессии гена транстиретина у человека и шимпанзе, остается открытым и должен решаться прямым экспериментальным анализом. Полученные структурные данные дают указания на участки, которые могут быть использованы в этом анализе.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Геномная ПЦР. В качестве матрицы использовали геномную ДНК человека и шимпанзе. Геномную ДНК из ткани мозга человека и шимпанзе выделяли, используя систему для выделения геномной ДНК "Wizard SV Genomic DNA Purification System" (Promega, США) в соответствии с рекомендациями производителя. Предварительно смешивали 1.75 ед. акт. *Taq*-ДНК-полимеразы (ИБХ, Россия) и 0.7 мкг *TaqStart*-антител (Clontech, США) в буфере, содержащем 50 мМ КСl, 10 мМ Трис-НСl, рН 7.0, инкубировали при комнатной температуре 5–10 мин и добавляли в инкубационную смесь для ПЦР, содержащую буфер (50 мМ КСl, 10 мМ Трис-НСl, рН 9.0, 0.1% Тритон Х-100), 2 мМ MgCl₂, 0.125 мМ dNTPs, 0.4 мМ олигонуклеотидные праймеры и 10 нг геномной ДНК. Суммарный объем реакционной смеси – 50 мкл. Из реакционной смеси отбирали аликваты по 10 мкл для анализа продуктов ПЦР электрофорезом в 1% агарозном геле.

ПЦР проводили в автоматических ПЦР-амплификаторах "PTC-200 Peltier Thermal Cycler" и "PTC-150 MiniCycler" (MJ Research, США).

Аmplification, клонирование, определение первичной структуры регуляторных участков гена транстиретина шимпанзе. Регуляторные участки гена транстиретина шимпанзе амплифицировали с парами праймеров (5'–3'):

ССТGCCAAGAATGAGTGGAC и
GCCTGTGCAGCTAGATGTCA,
АТАСТСАСТТСТССТGAGCTAGGC и
СТСАСGCATGATATGCCACT,
ГТСАСТСТСАGACATCAAGACTCC и
AGTGGGCATTTCAAGTATCAACA,
GTTGCCTGGCTCTTAGTTTAGG и
GCTTCAATCTCACCTCTTCACA.

Осуществляли от 28 до 34 циклов ПЦР в режиме: 94°C – 30 с, 60°C – 30 с, 72°C – 2 мин.

Фрагменты ДНК, полученные в результате ПЦР, клонировали, используя систему для клонирования продуктов ПЦР "pGEM-T Vector System" (Promega, США) в соответствии с рекомендациями производителя. Первичную структуру клонированных фрагментов ДНК определяли на автоматическом секвенаторе ABI Prism 3100-Avant Genetic Analyzer.

Для выравнивания и редактирования нуклеотидных последовательностей, определения G+C-состава использовали программу GeneRunner v. 3.00 (Hastings Software, Inc.). Поиск и анализ транскрипционных факторов, связывающихся с последовательностями гипотетических энхансеров гена транстиретина, проводили с помощью on-line-программы MatInspector компании Genomatix (<http://www.genomatix.de>).

БЛАГОДАРНОСТИ

Авторы выражают признательность В.К. Потопову и Н.В. Скапцовой за синтез олигонуклеотидов.

Исследования проведены при поддержке грантов РФФИ № 02-04-48712 и НШ-2006.20034.

СПИСОК ЛИТЕРАТУРЫ

1. *Duboule D., Wilkins A.S.* // Trends Genet. 1998. V. 14. P. 54–59.
2. *King M.C., Wilson A.C.* // Science. 1975. V. 188. P. 107–116.
3. *Enard W., Khaitovich P., Klose J., Zollner S., Heissig F., Giavalisco P., Nieselt-Struwe K., Muchmore E., Varki A., Ravid R., Doxiadis G.M., Bontrop R.E., Paabo S.* // Science. 2002. V. 296. P. 340–343.
4. *Gagneux P., Varki A.* // Mol. Phylogenet. Evol. 2001. V. 18. P. 2–13.
5. *Надеждин Е.В., Виноградова Т.В., Свердлов Е.Д.* // ДАН. 2001. Т. 381. С. 697–700.
6. *Gagneux P., Amess B., Diaz S., Moore S., Patel T., Dillmann W., Parekh R., Varki A.* // Am. J. Phys. Anthropol. 2001. V. 115. P. 99–109.
7. *Schreiber G., Richardson S.J.* // Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 1997. V. 116. P. 137–160.
8. *Nagata Y., Tashiro F., Yi S., Murakami T., Maeda S., Takahashi K., Shimada K., Okamura H., Yamamura K.* // J. Biochem. (Tokyo). 1995. V. 117. P. 169–175.
9. *Costa R.H., Lai E., Darnell J.E., Jr.* // Mol. Cell. Biol. 1986. V. 6. P. 4697–4708.
10. *Costa R.H., Lai E., Grayson D.R., Darnell J.E., Jr.* // Mol. Cell. Biol. 1988. V. 8. P. 81–90.
11. *Costa R.H., Grayson D.R., Darnell J.E., Jr.* // Mol. Cell. Biol. 1989. V. 9. P. 1415–1425.
12. *Chen F.C., Li W.H.* // Am. J. Hum. Genet. 2001. V. 68. P. 444–456.
13. *Fujiyama A., Watanabe H., Toyoda A., Taylor T.D., Itoh T., Tsai S.F., Park H.S., Yaspo M.L., Lehrach H., Chen Z.*

Fu G., Saitou N., Osoegawa K., de Jong P.J., Suto Y., Hattori M., Sakaki Y. // *Science*. 2002. V. 295. P. 131–134.

14. Antequera F. // *Cell. Mol. Life Sci.* 2003. V. 60. P. 1647–1658.

A Comparative Analysis of Regulatory Regions of the Transthyretin Gene in the Mouse, Human, and Chimpanzee Genomes

E. V. Nadezhdin[#], T. V. Vinogradova, and E. D. Sverdlov

[#]Phone: +7 (095) 330-6992; fax: +7 (095) 330-6538; e-mail: neugene@humgen.siocb.ras.ru
Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences,
ul. Miklukho-Maklaya 16/10, GSP Moscow, 117997 Russia

A full genome analysis of differences between the gene expression in the human and chimpanzee brains revealed that the gene for transthyretin, the carrier of thyroid hormones, is differently transcribed in the cerebella of these species. A 7-kbp DNA fragment of chimpanzee was sequenced to identify possible regulatory sequences responsible for the differences in expression. One hundred and thirteen substitutions were found in the chimpanzee sequence in comparison with the human sequence. About 40% of the substitutions were revealed within the repeating elements of the genome; their location and sizes did not differ from those in the corresponding fragments of the human genome, and the nucleotide sequences had a high degree of identity. A comparison of nucleotide sequences of the transthyretin region of human, chimpanzee, and mouse genes revealed substantial differences in the distribution of G + C content along the examined fragment in the human (chimpanzee) and mouse genes and allowed us to localize three sequence tracts with a higher degree of identity in the three species. One of these tracts is located in the promoter region of the gene, and the other two probably determine the specificity of transthyretin gene expression in the liver and brain. One of the conserved tracts of the chimpanzee genome was found to have a single and a triple nucleotide substitution. The triple substitution distinguishes chimpanzees from humans and mice, which have identical sequences of this site. It is likely that these substitutions are responsible for the differences in the expression levels of the transthyretin gene in the human and chimpanzee brains. The English version of the paper: *Russian Journal of Bioorganic Chemistry*, 2004, vol. 30, no. 4; see also <http://www.maik.ru>.

Key words: chimpanzee genome, regulation of gene expression, transthyretin