

УДК 577.21

НОВОЕ СЕМЕЙСТВО ДИСПЕРГИРОВАННЫХ ПОВТОРОВ ИЗ ГЕНОМА ЧЕШУЙЧАТЫХ РЕПТИЛИЙ

© 2006 г. С. А. Косушкин, О. Р. Бородулина, В. В. Гречко*, Д. А. Крамеров

Институт молекулярной биологии Российской академии наук, Москва, 119991

Поступила в редакцию 23.11.2005 г.

Ключевые слова: SINE, ретропозон, рептилии.

A NEW FAMILY OF INTERSPERSED REPEATS FROM SQUAMATE REPTILES, by S. A. Kosushkin, O. R. Borodulina, V. V. Grechko*, D. A. Kramerov (Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia, *e-mail: grechko@eimb.ru).

Диспергированные повторы типа SINE (Short Interspersed Elements) составляют значительную долю генома многих позвоночных животных. Эти последовательности наиболее изучены у млекопитающих и рыб, на примере которых были установлены основные характеристики структуры и эволюции SINEs. На 5'-конце этой последовательности имеется участок, как правило, родственный какой-либо тРНК, за которым следует специфический участок, нередко обладающий сходством с 3'-концевым регионом повтора типа LINE из генома того же вида [1]. На 3'-конце SINE находится либо участок коротких прямых тандемных повторов, либо олиго(А)-последовательность, которые необходимы, как полагают, для осуществления обратной транскрипции и встраивания ДНК-копии в геном [2, 3].

До последнего времени SINEs чешуйчатых рептилий (ящериц и змей, отр. Squamata) оставались неизученными. Описан лишь повтор такого типа из генома ящерицы *Podarcis sicula* (сем. Lacertidae), 5'-область которого сходна с лизиновой тРНК, а 3'-фрагмент – с LINE-подобным элементом [4]. Однако авторами представлена единственная копия повторов этого семейства и не приведено данных по его распространению в ДНК других рептилий.

*Эл. почта: grechko@eimb.ru

В нашей работе описано новое семейство диспергированных повторов типа SINE из генома ящериц рода *Darevskia* (Squamata: Lacertidae) – *Darevskia praticola* и *D. raddei*, различия между копиями которых не превышают внутривидового уровня дивергенции (рисунок). Повтор обнаружен при секвенировании ПЦР-фрагментов с геномной ДНК *D. raddei* (методика описана ранее [5, 6]). В качестве праймера использовали последовательность, комплементарную участку, содержащему А-бокс промотора для РНК-полимеразы III (А-ПЦР). Среди продуктов амплификации выявлено девять фрагментов ДНК, содержащих 3'-участок SINE, начинающийся с бокса А (отмечен знаком > на рисунке а, 5'-TRGCTCAGTGG-3'). На основании последовательностей этих фрагментов сконструирован обратный праймер (обозначен знаком < на рисунке а, 5'-GCTCTCYAG-GRTTTCA-3'), с использованием которого в ПЦР получили еще три последовательности, содержащие 5'-участок SINE. Кроме того, восемь полноразмерных копий получено при скрининге геномной клонотеки из *D. praticola* с помощью меченого зонда. Этот зонд получен с использованием праймеров А1 и Т2 к консервативным участкам последовательности внутри SINE (рисунок б).

Анализ всех имеющихся последовательностей (рисунок а) показал, что в них имеются элементы структуры, характерной для канонического типа

а – Последовательности SINE Squam-2 из генома ящериц сем. Lacertidae (Dra – *Darevskia raddei*, Dpr – *Darevskia praticola*). Верхние девять последовательностей получены с помощью ПЦР с праймером к боксу А (обозначен символом ">"), следующие три – с обратным праймером ("<"). Нижние восемь полноразмерных последовательностей получены путем секвенирования клонированных фрагментов из геномной клонотеки. tRNA-Ala – последовательность тРНК^{Ala(CGC)} из генома мыши и человека (идентичная последовательность обнаружена нами в геномах различных позвоночных путем компьютерного поиска). б – Выравнивание консенсусных последовательностей SINE, сходных с тРНК^{Ala(CGC)}: Squam-2 из генома ящериц, Das-1 из генома броненосца [6, 14], Vic-1 из генома викунии [13] и ID-Rno грызунов [12]. Внизу рисунка указаны праймеры А1, А2, Т1 и Т2. Схематическая структура Squam-2 над последовательностями отражает уменьшение консервативности последовательности в направлении 3'-конца, на ней указаны также боксы промотора РНК-полимеразы III (бокс А и бокс В).

SINE. 5'-участок обладает значительным сходством с последовательностью тРНК^{Ala(CGC)} млекопитающих (рисунок а), в которой содержатся элементы внутреннего промотора для РНК-полимеразы III – так называемые боксы А и В. Затем следует неродственный тРНК участок длиной около 100 п.н., специфичный для рептилий изучаемого отряда. тРНК-родственный регион (85 п.н.) во всех исследованных копиях наиболее консервативен, причем консервативными оказываются и 10–15 н. далее к 3'-концу, за пределами собственно тРНК. За ним следует короткий, более вариабельный участок. Степень дивергенции следующего участка неоднородна. Таким образом, после консервативного участка расположено около 15 н. с малой степенью сходства, затем степень вырожденности вначале несколько уменьшается и далее возрастает. В целом, длина тРНК-неродственного участка лишь немного превышает длину соответствующей тРНК.

В то же время, обнаруженный нами повтор обладает некоторыми свойствами, отличающими его от канонического типа SINE млекопитающих и рыб [1, 7, 8]. Так, не обнаружено ни поли(А)-“хвоста”, ни прямых коротких повторов на 3'-конце элемента. SINE, не содержащие поли(А)-“хвоста” (так называемые “tailless retroposons”), обнаружены в геномах большинства млекопитающих и, по-видимому, должны использовать другие механизмы взаимодействия с обратной транскриптазой при ретропозиции [9]. Кроме того, в окружающих последовательностях обнаруженного нами элемента не найдены короткие прямые фланкирующие повторы, представляющие собой дубликации сайта-мишени, по которому, как полагают, происходит встраивание новых копий SINE. Это не позволяет точно определить границы самого элемента и его длину, которая примерно составляет 160–180 п.н. Можно предположить, что, как и в случае других древних семейств SINE (например, MIR [10]), обнаруженное нами семейство потеряло участки, связанные с ретропозицией, а именно, короткие тандемные повторы на 3'-конце элемента и фланкирующие прямые повторы. Альтернативно, нельзя исключить, что механизм ретропозиции SINE рептилий отличается от других SINE, что объясняет отсутствие упомянутых структурных элементов.

Чтобы определить, имеется ли изучаемый SINE в ДНК других таксонов, использовали метод дот-блот-гибридизации специфичного зонда с геномной ДНК представителей 12 семейств ящериц отр. Squamata и двух семейств змей, а также других таксонов: черепах, крокодила, амфибий (травяная лягушка), млекопитающих (мышь и человек) и трех отрядов птиц (таблица). Гибридизацию ставили с двумя мечеными зондами по отдельности, один из которых, полученный с праймерами А1 и А2, комплементарен наиболее

консервативной (тРНК-родственной) части повтора, а другой, амплифицированный с помощью праймеров Т1 и Т2, – тРНК-неродственной части SINE, специфичной для изучаемых рептилий.

Сигнал гибридизации ДНК и обоих зондов наблюдали при анализе всех представителей чешуйчатых, однако в случае ДНК хамелеона, игуаны и агамы (входящих в кладу Iguania), а также змей (*Boa constrictor* и *Elaphe diene*) гибридизация была слабой (таблица). ДНК всех остальных животных, не входящих в отр. Squamata, или не давали гибридизационного сигнала ни с одним из праймеров, или слабо гибридизовались с праймером, специфичным для тРНК. Основываясь на этих данных, мы предложили для обнаруженного нами диспергированного повтора наименование Squam-2.

В случае ДНК человека, мыши и лягушки слабый сигнал наблюдали при гибридизации с зондом, соответствующим участку сходства с тРНК^{Ala(CGC)}. Степень этого сходства с тРНК-частью повтора представлена на рисунке б. Это может быть обусловлено присутствием в геноме грызунов семейства SINE ID, также ведущего свое происхождение от аланиновой тРНК. Кроме того, сигнал с “тРНК-овым” зондом у травяной лягушки и человека может быть обусловлен присутствием в их геномах генов или псевдогенов тРНК, однако отсутствие у этих видов гибридизации с тРНК-неродственным зондом свидетельствует о том, что в их ДНК повтора Squam-2 нет.

Результаты гибридизации подтверждены также в опытах, в которых ДНК всех упомянутых животных амплифицировали с помощью ПЦР и праймеров А1 и Т2. Продукт реакции, соответствующий по размеру повтору SINE, обнаружен в ДНК практически всех представителей отряда чешуйчатых, за исключением хамелеона. То, что у этой ящерицы отсутствует специфический ПЦР-продукт, соответствует тому, что дот-гибридизация ее ДНК также была очень слабой (таблица). Этот вопрос требует дальнейшего исследования. Обращает на себя внимание также тот факт, что две змеи из разных семейств, как было показано обоими методами, отличаются по наличию повтора типа Squam-2 ящериц: *Boa constrictor* содержит этот повтор в большем количестве, а *Elaphe diene*, если и содержит, то в более вырожденном состоянии или в меньшем числе копий. Вместе с другими данными это дает основание полагать, что изучаемый SINE появился в геноме предка чешуйчатых рептилий, судя по данным палеонтологии о возникновении этого отряда, около 250 млн. лет назад [11]. После появления это семейство SINEs, по-видимому, распространилось и в разной степени дивергировало в геномах представителей разных семейств чешуйчатых рептилий.

Распространение Squam-2 в геномах животных различных таксонов по данным гибридизации и ПЦР

| Таксономическое положение таксонов | | Гибридизация | | | ПЦР |
|------------------------------------|----------------|--------------|-------|---|-----|
| | | A1-A2 | T1-T2 | | |
| Mammalia: | Hominidae | 1 | ± | – | ± |
| Eutheria | Rodentia | 2 | ± | – | – |
| Mammalia: | Macropodidae | 3 | – | – | – |
| Marsupialia | | | | | |
| Aves | Strigidae | 4 | – | – | – |
| | Falconidae | 5 | – | – | – |
| | Passeridae | 6 | – | – | – |
| Amphibia: | Ranidae | 7 | + | – | – |
| Anura | | | | | |
| Reptilia: | Chamaeleonidae | 8 | + | ± | – |
| Squamata | Agamidae | 9 | + | + | + |
| | Iguanidae | 10 | + | + | + |
| | Lacertidae | 11 | + | + | + |
| | Lacertidae | 12 | + | + | + |
| | Cordylidae | 13 | + | + | + |
| | Scincidae | 14 | + | + | + |
| | Varanidae | 15 | + | + | + |
| | Anguidae | 16 | + | + | + |
| | Teiidae | 17 | + | + | ± |
| | Helodermatidae | 18 | + | + | + |
| | Eublepharidae | 19 | + | + | + |
| | Gekkonidae | 20 | + | + | + |
| | Colubridae | 21 | ± | ± | ± |
| | Boidae | 22 | ± | + | + |
| Reptilia: | Testudinidae | 23 | – | – | – |
| Testudines | | | | | |
| Reptilia: | Crocodylidae | 24 | – | – | – |
| Crocodylia | | | | | |

Примечание. ДНК гибридизовали с зондами, полученными с помощью праймеров A1 + A2 и T1 + T2 (см. рисунок). Знак “+” – интенсивный сигнал гибридизации и мажорная полоса продукта ПЦР; знак “–” – отсутствие сигнала гибридизации и продукта ПЦР; “±” – слабый сигнал гибридизации и слабая полоса продукта ПЦР. Исследованы следующие виды, обозначенные в таблице цифрами: 1 – человек, *Homo sapiens*; 2 – мышь, *Mus musculus*; 3 – древесный кенгуру, *Dendrolagus bennettianus*; 4 – ушастая сова, *Asio otus*; 5 – пустельга, *Falco tinnunculus*; 6 – домовый воробей, *Passer domesticus*; 7 – травяная лягушка, *Rana temporaria*; 8 – йеменский хамелеон, *Chamaeleo calyptratus*; 9 – плащеносная ящерица, *Chlamidosaurus kingii*; 10 – обыкновенная игуана, *Iguana iguana*; 11 – скальные ящерицы, *Darevskia raddeii/praticola*; 12 – канарская ящерица, *Gallotia* sp.; 13 – гигантский полоховост, *Cordylus giganteus*; 14 – синезыкий сцинк, *Tiliqua* sp.; 15 – смарагдовый варан, *Varanus prasinus*; 16 – ломкая веретеница, *Anguis fragilis*; 17 – тегу, *Tupinambis teguixin*; 18 – жилая ящерица, *Heloderma suspectum*; 19 – пятнистый эubleфар, *Eublepharis macularius*; 20 – токи, *Gekko gecko*; 21 – узорчатый полоз, *Elaphe dione*; 22 – обыкновенный удав, *Boa constrictor*; 23 – средиземноморская черепаха, *Testudo graeca*; 24 – нильский крокодил, *Crocodylus niloticus*.

При поиске последовательностей, сходных с изучаемым SINE в его специфической (тРНК-неродственной) части, в Genbank'e с помощью программы BLAST таких последовательностей не обнаружено. Сходство в области аланиновой тРНК имеется лишь в случае SINE млекопитающих, также, как уже отмечалось, происходящих от этой тРНК, а именно – ID грызунов [12], Vic-1 мозолоногих [13] и Das1 [6, 14] броненосцев (рисунок б). Вероятно, данный тип тРНК неоднократно в эволюции давал начало таксон-специфическим семействам SINEs, имеющим сравнительно простую структуру последовательности.

Таким образом, в ДНК изученных рептилий отряда Squamata нами найден и охарактеризован новый повтор типа SINE – Squam-2. Он относится к ряду повторов, ведущих свое происхождение от тРНК^{Ala(CGC)}, известных из геномов млекопитающих, но обладает свойствами, отличающими его от повторов с канонической структурой, выведенной на основании изучения диспергированных повторов у представителей других классов. Индивидуальные копии этого повтора рептилий сильно дивергированы в своей РНК-неродственной области и, по существу, не имеют сколько-нибудь определенного 3'-конца, характерного для всех них.

Одновременно нами обнаружено у чешуйчатых рептилий другое семейство SINE (Squam-1), данные о котором подготовлены к печати.

Благодарим Д.Б. Васильева и И.С. Даревского за консультации и предоставленные образцы крови некоторых животных.

Работа получила финансовую поддержку Российского фонда фундаментальных исследований (03-04-49157; 05-04-49553).

СПИСОК ЛИТЕРАТУРЫ

- Ohshima K., Hamada M., Terai Y. et al. 1996. The 3' ends of tRNA-derived short interspersed repetitive elements are derived from the 3' ends of long interspersed repetitive elements. *Mol. Cell. Biol.* **16**, 3756–3764.
- Okada N. 1991. SINEs. *Curr. Opin. Genet. Dev.* **1**, 498–504.
- Schmid C.W. 1996. Alu: structure, origin, evolution, significance and function of one-tenth of human DNA. *Prog. Nucleic Acid Res. Mol. Biol.* **53**, 283–319.
- Fantaccione S., Russo C., Palomba P. et al. 2004. A new pair of CR1-like LINE and tRNA-derived SINE elements in *Podarcis sicula* genome. *Gene*. **339**, 189–198.
- Borodulina O.R., Kramerov D.A. 1999. Wide distribution of short interspersed elements among eukaryotic genomes. *FEBS Lett.* **457**, 409–413.
- Borodulina O.R., Kramerov D.A. 2005. PCR-based approach to SINE isolation: simple and complex SINEs. *Gene*. **349**, 197–205.
- Kajikawa M., Okada N. 2002. LINEs mobilize SINEs in the eel through a shared 3' sequence. *Cell*. **111**, 433–444.

8. Ohshima K., Okada N. 2005. SINEs and LINEs: symbionts of eukaryotic genomes with a common tail. *Cytogenet. Genome Res.* **110**, 475–490.
9. Schmitz J., Churakov G., Zischler H. et al. 2004. A novel class of mammalian-specific tailless retropseudogenes. *Genome Res.* **14**, 1911–1915.
10. Jurka J., Zietkiewicz E., Labuda D. 1995. Ubiquitous mammalian-wide interspersed repeats (MIRs) are molecular fossils from the mesozoic era. *Nucleic Acids Res.* **23**, 170–175.
11. Carroll R.L. 1987. *Vertebrate Paleontology and Evolution*. N.Y.: W.H. Freeman and Company.
12. Kim J., Martignetti J.A., Shen M.R. et al. 1994. Rodent BC1 RNA gene as a master gene for ID element amplification. *Proc. Natl. Acad. Sci. USA.* **91**, 3607–3611.
13. Lin Z., Nomura O., Hayashi T. et al. 2001. Characterization of a SINE species from vicuna and its distribution in animal species including the family Camelidae. *Mamm. Genome.* **12**, 305–308.
14. Churakov G., Smit A.F., Brosius J. et al. 2005. A novel abundant family of retroposed elements (DAS-SINEs) in the nine-banded armadillo (*Dasypus novemcinctus*). *Mol. Biol. Evol.* **22**, 886–893.