

Структура генома и генетическое разнообразие

УДК 636.2:577.212.3

doi: 10.15389/agrobiology.2017.4.658rus

ДОМЕННАЯ ОРГАНИЗАЦИЯ МОБИЛЬНЫХ ГЕНЕТИЧЕСКИХ ЭЛЕМЕНТОВ В 1-й ХРОМОСОМЕ КРУПНОГО РОГАТОГО СКОТАВ.И. ГЛАЗКО^{1, 2}, О.И. СКОБЕЛЬ¹, Г.Ю. КОСОВСКИЙ¹, Т.Т. ГЛАЗКО^{1, 2}

Организация генома крупного рогатого скота (КРС), его «геномный ландшафт», в последние годы привлекает особое внимание в связи со сложностью решения задач геномной селекции — использованием полилокусных генотипов для ускорения и упрощения селекционной работы. Накоплены данные, свидетельствующие о высокой скорости эволюции различных геномных элементов и выраженности их структурно-функционального полиморфизма (L. Chen с соавт., 2017). Около 50 % всех нуклеотидных последовательностей в геноме крупного рогатого скота представлено диспергированными повторами (R.L. Tellam с соавт., 2009), некоторые из них образуют консервативные внутригеномные домены при совместной локализации (D.L. Adelson с соавт., 2009). Особенности распределения консервативных и варибельных доменов в геномах крупного рогатого скота до сих пор недостаточно исследованы, несмотря на их важность для решения задач контроля и управления генетическими ресурсами. В настоящей работе с использованием базы данных мобильных генетических элементов программы RepeatMasker (A.F.A. Smit с соавт., <http://repeatmasker.org>) и аналитической программы Integrated Genome Browser (J.W. Nikol с соавт., 2009) выполнен анализ доменной организации мобильных генетических элементов и продуктов их рекомбинаций в нуклеотидных последовательностях (13436028 п.н.) 1-й хромосомы КРС. Обнаружено, что в исследованном участке наиболее часто встречались элементы SINE/tRNA-Core-RTE, LINE/RTE-BovB, LINE/L1 и LTR/ERV. Их взаимная локализация представляет собой сложную структуру. Чаще всего наблюдались двучленные ассоциации SINE и LINE, SINE/tRNA-Core-RTE и LTR/ERV, (LTR/ERVK)/(LINE/RTE-BovB) и (LTR/ERVK)/LINE/L1. Последние два варианта служат основой для образования трехчленных кластеров — (LINE/RTE-BovB)/(BTLTR1)/(LINE/RTE-BovB) и (LINE/L1)/(BTLTR1J)/(LINE/L1), причем другие ретротранспозоны такие трехчленные кластеры фактически не формируют. Было выявлено некоторое смещение (относительно повышенной плотности локализации) этих трехчленных кластеров к дистальному концу исследованного участка 1-й хромосомы. С помощью программы Integrated Genome Browser мы определили расположение трехчленных продуктов рекомбинаций между LINE и LTR ERV по отношению к структурным генам. Оказалось, что 34 такие конструкции локализируются в 12 структурных генах (остальные — в межгенных пространствах), причем в основном по 10 и 12 копий в двух генах — *grik1* и *app*, тесно связанных у млекопитающих с функцией центральной нервной системы. Тот факт, что в каждом из этих двух генов трехчленная конструкция (LINE/RTE-BovB)/(BTLTR1)/(LINE/RTE-BovB) была представлена девятью копиями, а конструкцию (LINE/L1)/(BTLTR1J)/(LINE/L1) обнаружили в одной копии в *grik1* и в трех — в *app*, дает основание считать указанные гены древними мишенями для встройки и сохранения мобильных генетических элементов. Следует отметить, что (LINE/L1)/(BTLTR1J)/(LINE/L1) выявили только в этих двух генах, но не обнаружили в остальных 10, где присутствовал продукт рекомбинации (LINE/RTE-BovB)/(BTLTR1)/(LINE/RTE-BovB). Особенности распределения продуктов рекомбинаций между LINE и LTR ERV по исследованному участку 1-й хромосомы и их локализация в структурных генах позволяют предполагать возможное присутствие в них специфических структурно-функциональных элементов, выявление которых составляет предмет дальнейших исследований.

Ключевые слова: мобильные генетические элементы, ретротранспозоны, ДНК транспозоны, продукты рекомбинации, доменная организация, геномный ландшафт, крупный рогатый скот.

Наблюдения за мобильными генетическими элементами (МГЭ) в геномах млекопитающих имеют достаточно длинную историю. Наибольшую часть (от 40 до 50 % общей длины) в геномах млекопитающих занимают диспергированные повторы. У большинства видов млекопитающих, в том числе у крупного рогатого скота (КРС), в геномах среди мобильных генетических элементов доминируют ретротранспозоны (РТ), которые для своего перемещения используют механизмы размножения экзогенных ретровирусов (1). Один из отрядов пресмыкающихся — чешуйчатые (*Squamata*) мог быть источником ряда мобильных генетических элементов для жвачных (2-4). Предполагается, что горизонтальному переносу РТ между таксономически удаленными формами способствует общность среды обитания (5).

Многие выявленные РТ оказались общими для всех млекопитающих, что, вероятно, свидетельствует о древности их происхождения. В большинстве своем РТ утратили активность, но повышенный полиморфизм некоторых РТ может отражать их относительно недавнее происхождение и вовлеченность в процессы геномных реорганизаций, как правило, имеющих функциональное и эволюционное значение (6). Так, инсерции некоторых РТ в промоторные области структурных генов существенно меняют экспрессию последних, в кодирующие аминокислоты последовательности — приводят к появлению новых белков (7). У крупного рогатого скота описан ряд мутаций по структурным генам, связанных с инсерциями РТ, приводящих в гомозиготном состоянии к летальным эффектам (8).

У млекопитающих основной РТ — длинный диспергированный ядерный элемент 1 (Long Interspersed Nuclear Element, LINE1). Они несут в геноме активные и другие LINE, такие как семейства LINE RTE. Неавтономные короткие диспергированные ядерные элементы (Short Interspersed Nuclear Element, SINE) для транспозиций нуждаются в LINE. У приматов для транспозиций SINE Alu требуется LINE L1. Древний РТ LINE2 (L2) кодирует белки, необходимые для распространения широко представленного в геномах млекопитающих SINE MIR. У жвачных и сумчатых LINE RTE кодирует белки, необходимые для транспозиций SINE BovA — соответственно (BOV-A2, Bov-tA1, 2, 3)/SINE ART2A или SINE RTE. RTE LINE содержат BovB повторы, что позволяет предполагать их горизонтальный перенос от рептилий к жвачным и сумчатым (4, 9-11). Древние кластеры повторов L2/MIR образуют домены, консервативные в геномах человека и крупного рогатого скота, причем в таких доменах отсутствуют более молодые варианты повторов, например RTE/ART2A. Поскольку древние повторы кластеризуются в эволюционно консервативные домены, это позволяет предполагать наличие специальных механизмов, обеспечивающих подобную консервативность, которые могут быть связаны с блоками экспрессирующихся генов, со спецификой локализации в пространстве интерфазного ядра, различиями в рисунке метилирования или особенностями геномных участков, разных по защищенности от повторной интеграции ретротранспозонов (2). Накоплено большое количество данных, свидетельствующих о том, что именно ретротранспозиции у млекопитающих участвуют в появлении новых генов и функциональной эволюции (12), а также в генных дупликациях (13). Эндогенные ретровирусы (ERV), содержащие длинные концевые повторы (LTR), — еще один вариант РТ, широко распространенный в геномах млекопитающих (в частности, у КРС), выделяется высоким полиморфизмом; описаны межпородные различия по присутствию некоторых из них (14).

Обнаружение консервативных и варьируемых доменов локализации РТ имеет существенное практическое значение для выбора наиболее полиморфных геномных элементов, удобных для использования в качестве якорей в геномном сканировании (полилокусном генотипировании) при контроле генетической структуры и ее динамики на уровне вида и внутривидовой дифференциации (15). Особую значимость такие исследования приобретают в связи с выявленной в геномах человека зональностью распределения разных типов РТ и изменчивостью по копияности коротких участков ДНК (Copy Number Variability, CNV) в клетках герминативного ряда и соматических клетках (16).

Несмотря на важность изучения консервативных и варьируемых доменов в геномах животных сельскохозяйственных видов для решения традиционных задач контроля и управления генетическими ресурсами, до

сих пор такие работы остаются достаточно редкими. Более того, как правило, в них рассматривается колокализация полноразмерных РТ, в то время как в геномах присутствует огромное количество фрагментов РТ, маркирующих участки транспозиций и рекомбинаций между ними (17, 18).

В настоящем исследовании выявлены закономерности колокализации и кластеризации участков гомологии к наиболее часто встречающимся ретротранспозонам различных семейств и продуктам их рекомбинации (двучленные и трехчленные ассоциации) в самой длинной аутосоме крупного рогатого скота — 1-й хромосоме.

Цель работы заключалась в анализе распределения и позиционирования мобильных генетических элементов для выяснения возможных закономерностей их структурной организации в геноме.

Методика. В качестве исходных данных использовали информацию о геномной локализации мобильных генетических элементов у КРС с указанием нуклеотидных координат, представленную в программе RepeatMasker (19), — архив bosTau7.fa.out.gz (20), созданный версией программы RepeatMasker open-4.0.5 в октябре 2011 года. Сведения о распределении мобильных генетических элементов в пределах первичной последовательности 1-й хромосомы (161428367 п.н.), полученные из архива, проанализировали в отношении количества и частоты встречаемости различных мобильных элементов, используя возможности программы Microsoft Office Word в биоинформатических целях. Для последующего изучения наиболее часто встречающихся мобильных генетических элементов выбрали нуклеотидную последовательность 1-й хромосомы длиной 13436028 п.н. В этой последовательности, выявляя ближайшие соседства для участков гомологии между различным семействам мобильных генетических элементов, определили число доменов и частоту их встречаемости. На основании полученных данных о количестве мобильных генетических элементов, наиболее часто встречающихся в выделенном участке, и образуемых ими доменов была построена таблица для оценки закономерностей распределения мобильных генетических элементов в альтернативных цепях геномной ДНК. Двучленные домены с частотой встречаемости более 60 %, содержащие один и тот же мобильный генетический элемент в разных цепях, изучали с целью поиска более интересных закономерностей распределения и структурной организации РТ в геноме. О функциональных характеристиках выявленных кластеров судили на основании оценки их позиционирования внутри структурных генов с помощью программы Integrated Genome Browser (21). Информация об отобранных структурных генах получена из международной базы данных GenBank (22).

Результаты. Использованный нами подход позволил выявить не только попарную колокализацию разных элементов, но и кластеры, состоящие из трех мобильных элементов. В проанализированной по длине 1-й хромосоме КРС (161428367 п.н.) среди всех диспергированных повторов наиболее часто встречались SINE (38,277 %) и LINE (34,002 %) (табл. 1). Следует отметить, что ретротранспозон LINE1 широко распространен во всех геномах эукариот, в том числе у млекопитающих. Известно, что у КРС присутствуют полноразмерные варианты РТ этого семейства, сохраняющие активность в отношении ретротранспозиций и участвующие в геномных преобразованиях (23). Число микросателлитов (simple repeat) и DNA-транспозонов составило соответственно 9,549 % и 6,105 %, а число микросателлитов с низкой комплексностью (low complexity) не превышало 1,452 %. Основными семействами SINE были tRNA-Core-RTE (20,989 %), Core-RTE (6,979 %) и MIR (6,822 %). В семействе LINE наиболее полно

1. Распределение мобильных генетических элементов по длине 1-й хромосомы крупного рогатого скота (161428367 п.н.)

Семейство	Число	Частота, %
SINE	122589	38,277
/tRNA-Core-RTE	67222	20,989
/Core-RTE	22351	6,979
/MIR	21850	6,822
/tRNA	10781	3,366
/tRNA-RTE	221	0,069
/S-Deu-L2	126	0,039
/tRNA-Deu	36	0,011
SINE?/tRNA	2	0,001
LINE	108898	34,002
/L1	49258	15,380
/RTE-BovB	37251	11,631
/L2	18665	5,828
/CR1	2954	0,922
/RTE-X	669	0,209
/Penelope	61	0,019
/Dong-R4	27	0,008
/Jockey	9	0,003
/L1-Tx1	4	0,001
LTR ERV	33222	10,373
/ERVl-MaLR	10035	3,133
/ERV1	7105	2,218
/ERVl	6698	2,091
/ERVk	7526	2,350
/Gypsy	680	0,212
LTR?	336	0,105
LTR	341	0,106
/Gypsy?	302	0,094
/ERVl?	154	0,048
/ERV1?	45	0,014
Simple_repeat	30581	9,549
DNA	19553	6,105
/hAT-Charlie	10763	3,361
/TcMar-Tigger	3214	1,004
/hAT-Tip100	2114	0,660
/TcMar-Mariner	634	0,198
/hAT-Blackjack	881	0,275
/hAT	377	0,118
/hAT-Ac	256	0,080
/hAT-Tip100?	79	0,025
DNA	283	0,088
/TcMar-Tc2	315	0,098
DNA?	184	0,057
/hAT?	101	0,032
/TcMar	80	0,025
/hAT-Tag1	136	0,042
/PIF-Harbinger	16	0,005
DNA?/hAT-Tip100?	34	0,011
/PiggyBac	31	0,010
/TcMar-Tc1	25	0,008
DNA?/PiggyBac?	11	0,003
/TcMar?	6	0,002
/Kolobok	10	0,003
/TcMar-Pogo	3	0,001
Low_complexity	4650	1,452
Other	773	0,241
Unknown	315	0,098
snRNA	100	0,031
tRNA	116	0,036
RC/Helitron	95	0,030
rRNA	66	0,021
Satellite/centr	33	0,010
RNA	29	0,009
srpRNA	2	0,001
scRNA	1	0,000
RC?/Helitron?	16	0,005
Bcero	320266	

были представлены L1 (15,380 %), RTE-BovB (11,631 %) и L2 (5,828 %), в LTR ERV чаще встречались ERVL-MaLR (3,133 %), ERV1 (2,218 %), ERVL (2,091 %), ERVK (2,350 %) и Gypsy (0,212 %), а среди DNA-транспозонов больше половины приходилось на hAT-Charlie (3,361 %).

В выделенной первичной последовательности 13436028 п.н. для указанных семейств изучили образование доменов с попарной локализацией (табл. 2, процентное соотношение полученных результатов см. в приложениях 2.1 и 2.2 к электронной версии статьи на <http://www.agrobiology.ru>). Сравнение показало, что разница в частоте встречаемости доменов в обеих цепях не превышает 2,35 % для указанных семейств SINE (377 MIR/tRNA-Core-RTE и 346 tRNA-Core-RTE/MIR) на 1320 элементов MIR), 1,35 % — для LINE (280 L2/tRNA-Core-RTE и 296 tRNA-Core-RTE/L2 на 1183 L2), 4,76 % — для LTR ERV (например, 3 Gypsy/Core-RTE и 0 Core-RTE/Gypsy на 63 элемента Gypsy, хотя в случае Gypsy/L1 и L1/Gypsy разница составила 7,94 % — 14,29 против 6,35 %), для DNA-транспозонов — 2,58 %. Таким образом, в дальнейшем рассматривали пары в одной цепи.

Анализ колокализации SINE с другими мобильными генетическими элементами показал, что с наивысшей частотой SINE образует домены с элементами своего же семейства. Особенно активно в этом отношении семейство tRNA-Core-RTE: частота доменов Core-RTE/tRNA-Core-RTE для Core-RTE — 27,33 % (568 пар на 2078 элементов), доменов tRNA/tRNA-Core-RTE для tRNA — 21,48 % (290 пар на 1350 элементов), а MIR соседствует с tRNA-Core-RTE в 28,56 % случаев (377 пар на 1320 элементов). При этом само семейство tRNA-Core-RTE активнее всего взаимодействует с RTE-BovB (21,06 %, или 1611 пар на 7651 элемент).

Core-RTE кластеризуется с RTE-BovB и L1 с частотой соответственно 16,46 (342 домена на 2078 Core-RTE) и 15,21 % (316 доменов на 2078 Core-RTE). Семейство tRNA в 16,74 % случаев соседствует с L1 (226 доменов на 1350 tRNA). Домены MIR/L1 встречаются в 2 раза чаще, чем MIR/L2 — соответственно 154 (11,67 %) и 75 (5,68 %) на 1320 MIR. Полученные дан-

2. Число пар мобильных генетических элементов (МГЭ), наиболее часто встречающихся по длине участка 13436028 п.н. 1-й хромосомы крупного рогатого скота

Семейство	SINE				LINE			LTR					DNA	Всего МГЭ
	/tRNA-Core-RTE	/Core-RTE	/tRNA	/MIR	/L1	/RTE-BovB	/L2	/ERV1-MaLR	/ERV1	/ERV1	/ERVK	/Gypsy	/hAT-Charlie	
SINE														
/tRNA-Core-RTE	1425	568	290	377	1038	1574	280	322	197	201	11	10	211	7651
/Core-RTE	584	107	73	77	307	373	81	64	42	40	6	3	56	2078
/tRNA	299	64	76	89	217	58	82	68	20	46	3	2	63	1350
/MIR	346	90	87	101	149	98	75	31	23	28	3	4	56	1320
LINE														
/L1	1055	316	226	154	925	318	118	129	67	72	127	9	88	4162
/RTE-BovB	1611	342	70	71	350	324	80	82	70	49	422	3	53	3756
/L2	296	78	83	75	125	65	142	46	19	30	5	1	45	1183
LTR														
/ERV1-MaLR	312	72	84	54	132	81	39	106	26	22	1	4	19	1071
/ERV1	189	43	41	26	70	71	20	34	134	26	3	2	15	743
/ERV1	191	45	34	24	78	59	25	32	32	69	2	3	16	709
/ERVK	10	2	1	4	135	416	2	2	4	3	7	0	3	599
/Gypsy	13	0	1	2	4	5	3	2	2	4	0	9	5	63
DNA														
/hAT-Charlie	190	55	67	46	97	41	57	28	15	15	2	2	79	813

ные противоречат выводам D.L. Adelson с соавт. (2), которые именно этот вариант рассматривали как образующий консервативные и наиболее древние двучленные домены. По-видимому, такие расхождения могут быть обусловлены тем, что D.L. Adelson с коллегами рассматривали совместную локализацию полноразмерных генов ретротранспозонов MIR/L2, в нашем же исследовании оценивалась колокализация участков гомологии к РТ, размеры которых могли быть меньше полной длины идентификационных генов мобильных элементов. Полученные нами данные свидетельствуют о том, что при анализе формирования доменов ретротранспозонов в геномном ландшафте в случае видоспецифичных «молодых», активно участвующих в транспозициях мобильных элементов необходимо учитывать не только полноразмерные последовательности, но и продукты рекомбинаций МГЭ.

Семейства LTR ERV и SINE формировали домены, в частности 84 домена tRNA/ERVL-MaLR, не более чем в 6,22 % случаев. Тем не менее, всем семействам LTR ERV близко сопутствовали семейства SINE, чаще всего tRNA-Core-RTE. Так, домен ERVL-MaLR/tRNA-Core-RTE встречался 322 раза (с частотой 30,07 %), ERV1/tRNA-Core-RTE — 197 (26,51 %), ERVL/tRNA-Core-RTE — 201 (28,35 %), Gypsy/tRNA-Core-RTE — 10 раз (15,87 %). Аналогично в случае DNA-транспозонов tRNA соседствовал с hAT-Charlie 67 раз с частотой 4,96 %, остальные формировали двучленные домены менее чем в 4 % случаев. При этом hAT-Charlie колокализовался с tRNA 63 раза с частотой 7,75 %, а с tRNA-Core-RTE — 211 раз (25,95 %). Семейства LINE тоже чаще всего соседствовали с семействами SINE, а именно с tRNA-Core-RTE. В частности, на RTE-BovB/tRNA-Core-RTE (1574 домена) приходилось 41,91 % последовательностей, на L1/tRNA-Core-RTE (1038 доменов) — 24,94 %, а на L2/tRNA-Core-RTE (280 доменов) — 23,67 %. Так же часто встречалась пара L1/L1 (925 раз, что составило 22,22 %). В ассоциациях с семейством LTR ERV чаще всего присутствовал домен RTE-BovB/ERVK — 416 раз, или в 11,08 % случаев. Частота соседства остальных семейств LINE с семействами LTR ERV не превышала 3,30 %. Колокализация с DNA-транспозонами происходила реже, чем в 4,82 % случаев, что отмечали для L2/hAT-Charlie (57 доменов на 1183 L2), а hAT-Charlie формировал домены с семейством LINE с частотой не более 10,82 %, которую выявили для hAT-Charlie/L1 (88 доменов). DNA-транспозоны тоже редко соседствовали с семейством LTR ERV: частоту не более 2,34 % отмечали для hAT-Charlie/ERVL-MaLR (19 доменов), не более 3,17 % — для Gypsy/hAT-Charlie (2 домена).

Наибольший интерес представляет колокализация LTR/ERVK с семействами LINE. Так, домены вида ERVK/RTE-BovB встречались 422 раза с частотой 70,45 %, ERVK/L1 — 127 (21,20 %), ERVK/L2 — 5 раз (0,83 %). В альтернативной цепи частота для RTE-BovB/ERVK составила 69,45 % (416 доменов), для L1/ERVK — 22,54 % (135 доменов) и для L2/ERVK — 0,33 % (5 доменов).

Совместная распространенность семейства ERVK с LINE составила 92,49 % и 92,32 % в прямой и обратной цепях, в то время как с остальными МГЭ этот показатель не превышал 4 %. Подобный факт предполагает наличие трехчленного домена вида LINE/ERVK/LINE.

Более тщательный анализ показал, что семейство ERVK действительно встречалось в составе домена LINE/ERVK/LINE в 85,31 % случаев (511 из 599). Среди доменов LINE/ERVK/LINE выделялись RTE-BovB/BTLTR1/RTE-BovB, которые имели частоту 74,74 % (382 из 511), и L1/BTLTR1/L1 частотой 21,51 % (110 из 511) (табл. 3). Кроме того, наблю-

дали 12 перекрывающихся трехчленных кластеров, что свидетельствует о высокой изменчивости их локализации.

3. Виды доменов мобильных генетических элементов LINE/ERVК/LINE, расположенных на участке 13436028 п.н. 1-й хромосомы крупного рогатого скота

Вид домена	Доменов в участке хромосомы	
	число	частота, %
RTE-BovB/BTLTR1/RTE-BovB	382	74,74
RTE-BovB/BTLTR1B/RTE-BovB	1	0,20
RTE-BovB/BTLTR1E2/RTE-BovB	1	0,20
RTE-BovB/BTLTR1J4/RTE-BovB	1	0,20
RTE-BovB/ERV2-1-LTR BT/RTE-BovB	1	0,20
RTE-BovB/LTR2_BT/RTE-BovB	1	0,20
L1/BTLTR1J/L1	110	21,51
L1/BTLTR1/L1	3	0,59
L1/BTLTR1F/L1	1	0,20
L2/BTLTR1/L2	1	0,20
L1/BTLTR1/RTE-BovB	4	0,78
RTE-BovB/BTLTR1/L2	2	0,39
RTE-BovB/BTLTR1/L1	2	0,39
RTE-BovB/BTLTR1J/L1	1	0,20
Всего	511	

4. Распределение трехчленных доменов мобильных генетических элементов на участке 13436028 п.н. 1-й хромосомы крупного рогатого скота

Число и доля (%) доменов на 12 равных отрезках 1119669 п.н.												
1	2	3	4	5	6	7	8	9	10	11	12	Всего
LINE/ERVК/LINE												
33	21	29	40	38	33	41	39	34	56	62	85	511
6,46 %	4,11 %	5,68 %	7,83 %	7,44 %	6,46 %	8,02 %	7,63 %	6,65 %	10,96 %	12,13 %	16,63 %	100 %
В том числе												
RTE-BovB/BTLTR1/RTE-BovB												
26	15	21	29	31	25	28	30	29	40	44	64	382
6,81 %	3,93 %	5,50 %	7,59 %	8,12 %	6,54 %	7,33 %	7,85 %	7,59 %	10,47 %	11,52 %	16,75 %	100 %
L1/BTLTR1J/L1												
7	6	8	9	6	5	11	6	5	16	13	18	110
6,36 %	5,45 %	7,27 %	8,18 %	5,45 %	4,55 %	10,00 %	5,45 %	4,55 %	14,55 %	11,82 %	16,36 %	100 %

Анализ локализации трехчленных кластеров показал, что они покрывают 6,86 % первичной последовательности 1-й хромосомы длиной 13436028 п.н. и расположены неравномерно (табл. 4). Причем повышение плотности размещения таких кластеров наблюдалось ближе к дистальному концу этого участка. Неравномерность распределения семейств мобильных генетических элементов внутри и между хромосомами крупного рогатого скота описана в литературе (24). В то же время анализ таких данных затруднен, поскольку сложно различать последствия новых инсерций и транспозиций и результаты их делетирования или естественного отбора против неблагоприятных вариантов (очищающая селекция).

В программе Integrated Genome Browser мы оценили позиционирование трехчленных кластеров МГЭ вида RTE-BovB/BTLTR1/RTE-BovB и L1/BTLTR1J/L1 по отношению к структурным генам. Было обнаружено, что с высокой частотой эти кластеры выявляются внутри структурных генов *Bos taurus*, кодирующих glutamate ionotropic receptor kainate type subunit 1 (*grik1*) и amyloid beta precursor protein (*app*). В последовательностях гена *grik1* локализовались 9 трехчленных кластеров RTE-BovB/BTLTR1/RTE-BovB из 30 выявленных в структурных генах в исследованном участке хромосомы и один из 4 кластеров L1/BTLTR1J/L1. В гене *app* из 30 кластеров RTE-BovB/BTLTR1/RTE-BovB обнаружили 9 и из 4 кластеров вида L1/BTLTR1J/L1 — 3 (табл. 5). Следует отметить, что в последовательностях этих структурных генов именно RTE-BovB/BTLTR1/RTE-BovB встречались с наибольшей частотой (см. табл. 5), несмотря на повышенное количество

L1 по сравнению с RTE-BovV в рассмотренном фрагменте 1-й хромосомы (см. табл. 1). Это соответствует заключению о том, что RTE-BovV — более древний и накопивший большее количество мутаций элемент генома крупного рогатого скота по сравнению с L1 (2). Важно подчеркнуть, что L1 и RTE-BovV — исторически достаточно удаленные друг от друга РТ, хотя и принадлежащие к LINE; тем не менее, именно в обоих генах присутствуют RTE-BovV/BTLTR1/RTE-BovV и L1/BTLTR1J/L1 — сходные трехчленные продукты (см. табл. 5). Можно предположить, что такая колокализация ассоциирована с имеющимися в этих продуктах рекомбинациями РТ элементов структурно-функционального сходства, что будет предметом наших дальнейших исследований.

5. Позиционирование трехчленных кластеров мобильных элементов по отношению к структурным генам по длине участка (13436028 п.н.) 1-й хромосомы крупного рогатого скота

Наименование структурного гена	Число кластеров	
	1	2
Bos taurus potassium voltage-gated channel subfamily E regulatory subunit 2 (<i>kcne2</i>), mRNA (-)	1	0
Bos taurus phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase (<i>gart</i>), mRNA (+)	1	0
Bos taurus transmembrane protein 50B (<i>tmem50b</i>), mRNA (+)	1	0
Bos taurus interleukin 10 receptor subunit beta (<i>il10rb</i>), mRNA (-)	2	0
Bos taurus interferon alpha and beta receptor subunit 2 (<i>ifnar2</i>), mRNA (-)	1	0
Bos taurus URB1 ribosome biogenesis 1 homolog (<i>S. cerevisiae</i>) (<i>urb1</i>), mRNA (+)	1	0
Bos taurus glutamate ionotropic receptor kainate type subunit 1 (<i>grik1</i>), mRNA (+)	9	1
Bos taurus ubiquitin specific peptidase 16 (<i>usp16</i>), mRNA (-)	1	0
Bos taurus listerin E3 ubiquitin protein ligase 1 (<i>ltn1</i>), mRNA (+)	1	0
Bos taurus cysteine and tyrosine rich 1 (<i>cyyr1</i>), mRNA (+)	1	0
Bos taurus amyloid beta precursor protein (<i>app</i>), mRNA (+)	9	3
Bos taurus junctional adhesion molecule 2 (<i>jam2</i>), mRNA (-)	2	0

Примечание. 1 — RTE-BovV/BTLTR1/RTE-BovV, 2 — L1/BTLTR1J/L1; «←» и «→» — ген расположен в обратной или в прямой цепи

В литературе имеются данные об ассоциации мутаций субъединицы 1 ионотропного каинатного рецептора глутамата *grik1* с поведенческими патологиями человека — шизофренией, эпилепсией, депрессией, биполярным расстройством (25-27). Согласно имеющимся данным, белок предшественника β-амилоида APP задействован в процессах нейропластичности и необходим для выживания нервных клеток (28). Фрагмент указанного белка, так называемый β-амилоидный пептид (Aβ), представляет собой основной компонент сенильных бляшек, образование которых считается основным патоморфологическим признаком болезни Альцгеймера, причем Aβ пептид, обнаруженный в мозге крупного рогатого скота, демонстрирует определенное сходство с аналогичными пептидами человеческого мозга на ранних стадиях старения (29). Высокая плотность локализации трехчленных продуктов рекомбинации видоспецифичных для *Bos taurus* ретро-транспозонов BTLN1 и BTLTRERV с постоянной архитектурой (прямые повторы BTLN1 на флангах трехчленной конструкции в одной цепи и участок гомологии к BTLTRERV в центре в альтернативной цепи) в двух генах, тесно связанных с функциями центральной нервной системы, позволяет предполагать определенную связь с теми признаками доместикизации (сниженная агрессивность по отношению к человеку), которые Д.К. Беляев выделял в качестве ведущих в процессе одомашнивания животных (30). Интересно отметить, что ранее были выявлены инсерции Bov-V в структурный ген (ассоциирован у крупного рогатого скота с особенностями развития краниофасциальных соотношений), которые отсутствуют в этом гене у человека и мыши (31).

В целом выявленное распределение ретро-транспозонов и продуктов

их рекомбинаций в нуклеотидных последовательностях 1-й хромосомы КРС длиной 13436028 п.н. позволило сделать следующие заключения. В исследованном участке хромосомы часто встречаются tRNA-Core-RTE, RTE-BovB, L1 и LTR ERV. Их взаимная локализация в геноме сложно организована, наиболее часты двучленные ассоциации — SINE и LINE, tRNA-Core-RTE и LTR ERV, ERVK/RTE-BovB, ERVK/L1. Последние два варианта служат основой трехчленных кластеров RTE-BovB/BTLTR1/RTE-BovB и L1/BTLTR1J/L1, при этом другие РТ таких трехчленных кластеров фактически не формируют. Обнаружено некоторое смещение относительно повышенной плотности локализации этих трехчленных кластеров к дистальному концу исследованного участка 1-й хромосомы. Анализ локализации выявленных трехчленных продуктов рекомбинаций между LINE и LTR ERV по отношению к структурным генам показал, что 34 такие конструкции выявляются в 12 структурных генах (остальные — в межгенных пространствах), причем 10 и 12 копий в двух генах (*grik1* и *app*), тесно связанных у млекопитающих с функцией центральной нервной системы. То обстоятельство, что в каждом из этих двух генов встречалось по 9 копий трехчленной конструкции RTE-BovB/BTLTR1/RTE-BovB, а конструкция L1/BTLTR1J/L1 обнаружена только в одной копии в *grik1* и в трех копиях — в *app*, дает основание считать указанные гены древними мишенями для встройки и сохранения. Отметим, что конструкцию L1/BTLTR1J/L1 обнаружили только в двух этих генах, но не в остальных 10, в которых присутствует продукт рекомбинации RTE-BovB/BTLTR1/RTE-BovB.

Итак, нами получены данные о наличии закономерностей в распределении фрагментов ретротранспозонов и продуктов их рекомбинации в геноме крупного рогатого скота. Специфические особенности распределения продуктов рекомбинаций между LINE и LTR ERV по изученному участку 1-й хромосомы, их локализация в структурных генах позволяет предполагать возможное присутствие консервативных структурно-функциональных элементов, выполняющих регуляторную роль, выявление которых составляет предмет наших дальнейших исследований.

ЛИТЕРАТУРА

1. Elsik C.G., Tellam R.L., Worley K.C. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*, 2009, 324(5926): 522-528 (doi: 10.1126/science.1169588).
2. Adelson D.L., Raison J.M., Edgar R.C. Characterization and distribution of retrotransposons and simple sequence repeats in the bovine genome. *PNAS USA*, 2009, 106(31): 12855-12860 (doi: 10.1073/pnas.0901282106).
3. Walsh A.M., Kortschak R.D., Gardner M.G., Bertozzi T., Adelson D.L. Widespread horizontal transfer of retrotransposons. *PNAS USA*, 2013, 110(3): 1012-1016 (doi: 10.1073/pnas.1205856110).
4. Годакова С.А., Севастьянова Г.А., Семенова С.К. Особенности структуры и распространения ретротранспозона Bov-B LINE. *Молекулярная генетика, микробиология и вирусология*, 2016, 34(1): 9-12 (doi: 10.18821/0208-0613-2016-34-1-9-12).
5. Wang X., Liu X. Close ecological relationship among species facilitated horizontal transfer of retrotransposons. *BMC Evolutionary Biology*, 2016, 1: 201 (doi: 10.1186/s12862-016-0767-0).
6. Chen L., Chamberlain A.J., Reich C.M., Daetwyler H.D., Hayes B.J. Detection and validation of structural variations in bovine whole-genome sequence data. *Genet. Sel. Evol.*, 2017, 49: 13 (doi: 10.1186/s12711-017-0286-5).
7. Krull M., Petrusma M., Makalowski W., Brosius J., Schmitz J. Functional persistence of exonized mammalian-wide interspersed repeat elements (MIRs). *Genome Res.*, 2007, 17(8): 1139-1145 (doi: 10.1101/gr.6320607).
8. Schütz E., Wehrhahn C., Wanjek M., Bortfeld R., Wemheuer W.E., Beck J., Brenig B. The Holstein Friesian Lethal Haplotype 5 (HH5) results from a complete deletion of TBF1M and cholesterol deficiency (CDH) from an ERV-(LTR) insertion into the coding region of APOB. *PLoS ONE*, 2016, 11(4): e0154602 (doi: 10.1371/journal.pone.0154602).

9. Kordis D., Gubensek F. Horizontal transfer of non-LTR retrotransposons in vertebrates. *Genetica*, 1999, 107(1-3): 121-128.
10. Kordis D., Gubensek F. Unusual horizontal transfer of a long interspersed nuclear element between distant vertebrate classes. *PNAS USA*, 1998, 95(18): 10704-10709.
11. Gentles A.J., Wakefield M.J., Kohany O., Gu W., Batzer M.A., Pollock D.D., Jurka J. Evolutionary dynamics of transposable elements in the short-tailed opossum *Monodelphis domestica*. *Genome Res.*, 2007, 17(7): 992-1004 (doi: 10.1101/gr.6070707).
12. Carelli F.N., Hayakawa T., Go Y., Imai H., Warnefors M., Kaessmann H. The life history of retrocopies illuminates the evolution of new mammalian genes. *Genome Res.*, 2016, 26(3): 301-314 (doi: 10.1101/gr.198473.115).
13. Tan S., Cardoso-Moreira M., Shi W., Zhang D., Huang J., Mao Y., Jia H., Zhang Y., Chen C., Shao Y., Leng L., Liu Z., Huang X., Long M., Zhang Y.E.. LTR-mediated retroposition as a mechanism of RNA-based duplication in metazoans. *Genome Res.*, 2016, 26(12): 1663-1675 (doi: 10.1101/gr.204925.116).
14. Garcia-Etxebarria K., Sistiaga-Poveda M., Jugo B.M. Endogenous retroviruses in domestic animals. *Curr. Genomics*, 2014, 15(4): 256-265 (doi: 10.2174/1389202915666140520003503).
15. Mei L., Ding X., Tsang S.-Y., Pun F.W., Ng S.-K., Yang J., Zhao C., Li D., Wan W., Yu C.-H., Tan T.-C., Poon W.-S., Leung G.K.-K., Ng H.-K., Zhang L., Xue H. AluScan: a method for genome-wide scanning of sequence and structure variations in the human genome. *BMC Genomics*, 2011, 12: 564 (doi: 10.1186/1471-2164-12-564).
16. Ng S.-K., Hu T., Long X., Chan C.-H., Tsang S.-Y., Xue H. Feature co-localization landscape of the human genome. *Sci. Rep.*, 2016, 6: 20650 (doi: 10.1038/srep20650).
17. Liu M., Eiden M.V. Role of human endogenous retroviral long terminal repeats (LTRs) in maintaining the integrity of the human germ line. *Viruses*, 2011, 3(6): 901-905 (doi: 10.3390/v3060901).
18. Глазко В.И., Феофилов А.В., Бардуков Н.В., Глазко Т.Т. Видоспецифичные ISSR-PCR-маркеры и пути их формирования. *Изв. ТСХА*, 2012, 1: 118-125.
19. Smit A.F.A., Hubley R., Green P. RepeatMasker. Режим доступа: <http://repeatmasker.org>. Без даты.
20. bosTau7.fa.out.gz, Oct 2011. RepeatMasker open-4.0.5, Repeat Library 20140131. Режим доступа: <http://www.repeatmasker.org/species/bosTau.html>. Без даты.
21. Nikol J.W., Helt G.A., Blanchard S.G. Jr., Raja A., Loraine A.E. The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics*, 2009, 25(20): 2730-2731 (doi: 10.1093/bioinformatics/btp472).
22. GenBank. Режим доступа: <https://www.ncbi.nlm.nih.gov/genbank/>. Без даты.
23. Ivancevic A.M., Kortschak R.D., Bertozzi T., Adelson D.L. LINEs between species: evolutionary dynamics of LINE-1 retrotransposons across the eukaryotic tree of life. *Genome Biol. Evol.*, 2016, 8(11): 3301-3322 (doi: 10.1093/gbe/evw243).
24. Saylor B., Elliott T.A., Linquist S., Kremer S.C., Gregory T.R., Cottenie K. A novel application of ecological analyses to assess transposable element distributions in the genome of the domestic cow, *Bos taurus*. *Genome*, 2013, 56(9): 521-533 (doi: 10.1139/gen-2012-0162).
25. Hirata Y., Zai C.C., Souza R.P., Lieberman J.A., Meltzer H.Y., Kennedy J.L. Association study of GRIK1 gene polymorphisms in schizophrenia: case-control and family-based studies. *Hum. Psychopharmacol.*, 2012, 27(4): 345-351 (doi: 10.1002/hup.2233).
26. Le-Niculescu H., Patel S.D., Bhat M., Kuczenski R., Faraone S.V., Tsuang M.T., McMahon F.J., Schork N.J., Nurnberger J.I. Jr., Niculescu III A.B. Convergent functional genomics of genome-wide association data for bipolar disorder: comprehensive identification of candidate genes, pathways and mechanisms. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, 2009, 150B(2): 155-181 (doi: 10.1002/ajmg.b.30887).
27. Башкатов С.А., Нургалиева А.Х., Еникеева Р.Ф., Казанцева А.В., Хуснутдинова Э.К. Перспективы разработки объективных индикаторов субъективного благополучия на основе данных психолого-генетического анализа. *Вестник ЮУрГУ. Серия Психология*, 2016, 9(4): 25-39 (doi: 10.14529/psy160403).
28. Татарникова О.Г., Орлов М.А., Бобкова Н.В. Бета-амилоид и Тау-белок: структура, взаимодействие и прионоподобные свойства. *Успехи биологической химии*, 2015, 55: 351-390.
29. Costassa E.V., Fiorini M., Zanusso G., Peletto S., Acutis P., Baioni E., Maurella C., Tagliavini F., Catania M., Gallo M., Lo Faro M., Chiappa M.N., Meloni D., D'Angelo A., Paciello O., Ghidoni R., Tonoli E., Casalone C., Corona C. Characterization of amyloid- β deposits in bovine brains. *Journal of Alzheimer's Disease*, 2016, 51: 875-887 (doi: 10.3233/JAD-151007).
30. Трут Л.Н. Доместикация в историческом процессе и в эксперименте. *Вестник ВОГиС*, 2007, 11(2): 273-289.
31. Iwashita S., Itoh T., Takeda H., Sugimoto Y., Takahashi I., Nobukuni T., Sezaki M., Masui T., Hashimoto K. Gene organization of bovine BCNT that contains a portion corresponding to an endonuclease domain derived from an RTE-1 (Bov-B LINE), non-LTR retrotransposable element: duplication of an intramolecular repeat unit downstream of the

¹ФГБНУ Центр экспериментальной эмбриологии
и репродуктивных биотехнологий,
127422 Россия, г. Москва, ул. Костякова, 12, стр. 4,
e-mail: vigvalery@gmail.com, skobelolga@gmail.com, gkosovsky@mail.ru,
tglazko@rambler.ru;

Поступила в редакцию
5 декабря 2016 года

²ФГБОУ ВПО Российский государственный
аграрный университет—МСХА им. К.А. Тимирязева,
127550 Россия, г. Москва, ул. Тимирязевская, 49

Sel'skokhozyaistvennaya biologiya [Agricultural Biology], 2017, V. 52, № 4, pp. 658-668

DOMAIN DISTRIBUTION OF MOBILE GENETIC ELEMENTS IN THE BOVINE GENOME

V.I. Glazko^{1, 2}, O.I. Skobel¹, G.Yu. Kosovsky¹, T.T. Glazko^{1, 2}

¹Center for Experimental Embryology and Reproductive Biotechnology, Federal Agency of Scientific Organizations, 12/4, ul. Kostyakova, Moscow, 127422 Russia, e-mail vigvalery@gmail.com (corresponding author), skobelolga@gmail.com, gkosovsky@mail.ru, tglazko@rambler.ru;

²K.A. Timiryazev Russian State Agrarian University—Moscow Agrarian Academy, 49, ul. Timiryazevskaya, Moscow, 127550 Russia

ORCID:

Glazko V.I. orcid.org/0000-0002-8566-8717

Kosovsky G.Yu. orcid.org/0000-0003-3808-3086

Skobel O.I. orcid.org/0000-0002-0599-9787

Glazko T.T. orcid.org/0000-0002-3879-6935

The authors declare no conflict of interests

Received December 5, 2016

doi: 10.15389/agrobiol.2017.4.658eng

Abstract

Genetic landscape of bovine genome attracts a lot of attention in recent years. This is due to the complexity of genomic selection task solution, i.e. the use of multilocus genotypes in order to simplify and hasten breeding. Accumulated data show that there is high evolutionary speed of different genetic elements and also they have structure functional polymorphism intensity (L. Chen et al., 2017). It was shown that interspersed repeats account for about 50 % of nucleotide sequence of the bovine genome (R.L. Tellam et al., 2009). Also it was found that some of the interspersed repeats cluster into conservative domains along the bovine genome due to joint localization (D.L. Adelson et al., 2009). The characteristics of domain distribution are still not fully studied despite the fact that it is very important to identify conservative and variable domains throughout the bovine genome to solve traditional tasks of their genetic resources management and controlling. In this work domain distribution of mobile genetic elements and their products of recombination in nucleotide sequences of 13,436,028 nucleotides of bovine chromosome 1 were analyzed by means of Repeat Masker mobile genetic elements database and Integrated Genome Browser software. It was revealed that the most prevalent types throughout analyzed region are SINE/tRNA-Core-RTE, LINE/RTE-BovB, LINE/L1 and LTR/ERV. Their joint localization in bovine genome has complicated structure. The most common pairwise clusters are SINE and LINE, SINE/tRNA-Core-RTE and LTR ERV, (LTR/ERVK)/(LINE/RTE-BovB), (LTR/ERVK)/(LINE/L1). Two last pairs are the bases for such triple clusters as (LINE/RTE-BovB)/(BTLTR1)/(LINE/RTE-BovB) and (LINE/L1)/(BTLTR1J)/(LINE/L1). It should be mentioned that there is no such clustering with other retrotransposons. It was revealed that there is some certain bias of these triple clusters high density to the distal end of studied region of chromosome 1. By the means of Integrated Genome Browser software the localization of obtained triple products of recombination between LINE and LTR ERV to structural genes was analyzed. It was found that only 34 clusters are localized in 12 structural genes (other are located in intergenic space). Besides, 10 and 12 copies are located in two genes that are closely connected with the function of central nervous system in mammals, *grik1* and *app*. The fact that 9 copies of triple gene construct (LINE/RTE-BovB)/(BTLTR1)/(LINE/RTE-BovB) are found in each of two genes and (LINE/L1)/(BTLTR1J)/(LINE/L1) had only 1 copy in *grik1* and 3 copies in *app*, suggests that these genes are ancestral targets for such insertions and their conservations. It also should be mentioned that (LINE/L1)/(BTLTR1J)/(LINE/L1) construct was found only in these two genes but not in other 10 genes where (LINE/RTE-BovB)/(BTLTR1)/(LINE/RTE-BovB) was also located. Specific features of distribution of products of recombination between LINE and LTR ERV throughout the studied chromosome 1 area and their localization in structural genes suggest the possible presence of structure functional elements there. Revealing of such elements is the subject of our further study.

Keywords: mobile genetic elements, retrotransposons, DNA transposons, products of recombination, domain distribution, genomic landscape, cattle.