

Оценка роли однонуклеотидного полиморфизма в гене лимфотоксина бета при доместикации свиньи на основе биоинформационного и экспериментального подходов

Р.Б. Айтназаров^{1,3}, Е.В. Игнатьева^{1,3}, Н.Э. Базарова⁴, В.Г. Левицкий^{1,3}, С.П. Князев⁴, Я. Гон^{3,5}, Н.С. Юдин^{1,2,3}

1 Федеральное государственное бюджетное научное учреждение

«Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук», Новосибирск, Россия

2 Федеральное государственное бюджетное научное учреждение

«Научно-исследовательский институт терапии и профилактической медицины», Новосибирск, Россия

3 Федеральное государственное автономное образовательное учреждение высшего образования «Новосибирский национальный исследовательский государственный университет», Новосибирск, Россия

4 Федеральное государственное бюджетное образовательное учреждение высшего образования «Новосибирский государственный аграрный университет», Новосибирск, Россия

5 Хэйлунцзянский университет, биологический факультет, Харбин, Хэйлунцзянская провинция, Китай

В работах, выполненных на диких и лабораторных животных, показано существование компромисса между репродуктивным успехом и иммунитетом. Поэтому в процессе доместикации могли отбираться особи с повышенными репродуктивными способностями, но со сниженным иммунитетом. Пониженная реактивность иммунной системы могла в дальнейшем стать наследуемой путем фиксации в популяции генов с «неблагоприятными» мутациями. Цель исследования – изучить: 1) частоты генотипов и аллелей однонуклеотидного полиморфизма (SnP – Single nucleotide Polymorphism) SnP rs340283541 в гене цитокина лимфотоксина бета (*LTB*) у домашних свиней и диких кабанов; 2) экспрессию мРНК этого гена у миниатюрных свиней с разными генотипами; 3) провести биоинформатический анализ потенциальной функциональной роли этого SnP. Частота генотипа GG в выборке кабанов была достоверно ниже частоты данного генотипа в объединенной выборке из разных пород и популяций домашних свиней. уровень экспрессии мРНК гена *LTB* в лимфатическом узле у миниатюрных свиней с генотипом GG имел тенденцию к повышению ($p < 0,06$), по сравнению с носителями аллеля A. SNP rs340283541 входит в состав мотива ДНК с высокой степенью консервативности у 12 видов млекопитающих, что косвенно свидетельствует о его важной функциональной роли. С помощью контекстного анализа выявлено, что аллель A содержит потенциальные сайты связывания транскрипционных факторов BRN-2 и AP-1, а аллель G – факторов RFX1, ISGF3 (сайт ISRE) и USF, которые экспрессируются в клетках иммунной системы. Таким образом, в процессе доместикации свиней произошло повышение частоты генотипа GG SNP rs340283541 в 3'- области гена *LTB*. Генотип GG, вероятно, ассоциирован с повышенным уровнем экспрессии мРНК гена *LTB* в ткани лимфатического узла. Повышение уровня экспрессии у свиней с генотипом GG может быть связано с образованием сайтов связывания транскрипционных факторов RFX1, ISRE, USF и/или разрушением сайтов связывания BRN-2 и AP-1. Также не исключено, что полиморфизм rs340283541 находится в неравновесии по сцеплению с другой функционально значимой мутацией.

Ключевые слова: доместикация; свинья; дикий кабан; ген; лимфотоксин бета; однонуклеотидный полиморфизм; SNP; транскрипция; сайт связывания транскрипционного фактора.

Received 21.09.2015

Accepted for publication 15.10.2015

© АВТОРЫ, 2015

e-mail: yudin@bionet.nsc.ru

Dissecting the role of single nucleotide polymorphism of lymphotoxin beta gene during pig domestication using bioinformatic and experimental approaches

R.B. Aitnazarov^{1,3}, E.V. Ignatieva^{1,3}, N.E. Bazarova⁴, V.G. Levitsky^{1,3}, S.P. Knyazev⁴, Y. Gon^{3,5}, N.S. Yudin^{1,2,3}

1 Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

2 Institute of Internal and Preventive Medicine, Siberian Branch of Russian Academy of Medical Sciences, Novosibirsk, Russia

3 Novosibirsk State University, Novosibirsk, Russia

4 Novosibirsk State Agrarian University, Novosibirsk, Russia

5 Heilongjiang University College of Life Sciences, Harbin, China

Studies of wild and laboratory animals have revealed a trade-off between reproductive success and immunity. Therefore, it is likely that domestication favored selection of individuals with high reproductive performance but low immunity. The low responsiveness of the immune system could become hereditary through fixation of genes with “unfavorable” mutations in populations. The objectives of this work are: 1) determination of frequencies of genotypes and alleles of the rs340283541 SnP in the gene for the lymphotoxin beta (*LTB*) cytokine in pigs of domestic breeds and wild boars; 2) investigation of the expression of *LTB* mRNA in minipigs with different genotypes, and 3) bioinformatic analysis of the putative functional role of the SnP. The frequency of the GG genotype in the wild boar sample was significantly lower than in the pooled sample of domestic pigs. The *LTB* mRNA expression rate in the lymph node of minipigs with genotype GG tended to increase ($p < 0.06$) in comparison with carriers of allele A. The rs340283541 SnP occurs in a DNA motif highly conservative among 11 mammalian species; thus, it may be of functional significance. context analysis shows that allele A has putative binding sites for transcription factors BRN-2 and AP-1, whereas allele G has binding sites for transcription factors RFX1, ISGF3 (site ISRE), and USF expressed in cells of the immune system. Thus, pig domestication was accompanied by an increase in the frequency of the GG genotype for the rs340283541 SnP, occurring in the 3' region of the *LTB* gene. it is likely that the GG genotype is associated with elevated *LTB* mRNA expression in the lymph node tissue. This increase may be related to the formation of binding sites for RFX1, ISRE, and USF and/or disruption of binding sites for BRN-2 and AP-1. A linkage disequilibrium between rs340283541 and another functionally significant mutation in *LTB* is also conceivable.

Key words: domestication; pig; wild boar; gene; lymphotoxin beta; single nucleotide polymorphism; SnP; transcription; transcription factor binding site.

HOW TO CITE THIS ARTICLE?

Aitnazarov r .B., Ignatieva E.V., Bazarova n.E., Levitsky V.G., Knyazev S.P., Gon Y., Yudin n.S. Dissecting the role of single nucleotide polymorphism of lymphotoxin beta gene during pig domestication using bioinformatic and experimental approaches. *Vavilovskii Zhurnal Genetiki i Selekcii* = *Vavilov Journal of Genetics and Breeding*. 2015;19(6):699-706. Doi 10.18699/VJ15.088

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ?

Айтназаров Р.Б., Игнатъева е.В., Базарова Н.Э., левицкий В.Г., Князев С.П., Гон Я., Юдин Н.С. Оценка роли однонуклеотидного полиморфизма в гене лимфотоксина бета при доместикации свиньи на основе биоинформационного и экспериментального подходов. *Вавиловский журнал генетики и селекции*. 2015;19(6):699-706. Doi 10.18699/VJ15.088

В работах, выполненных на диких и лабораторных животных, показано существование компромисса между репродуктивным успехом и иммунитетом (Sheldon, Verhulst, 1996; Ardia et al., 2011; van der Most et al., 2011; Balenger, Zuk, 2014). Цитокинам принадлежит центральная роль в регуляции иммунного ответа, а также его интеграции с физиологическими функциями других систем организма – эндокринной и гематопоетической. Одним из наиболее значительных семейств цитокинов является семейство белков фактора некроза опухолей, которое включает около 48 белков (Kim et al., 2005). Типичным представителем этого семейства является лимфотоксин бета (LTB). Он синтезируется активированными Т- и В-лимфоцитами, естественными киллерами и образует гетеротример с лимфотоксином-альфа LTA/LTB/LTB (реже LTA/LTA/LTB) и, таким образом, «заякоривает» лимфотоксин-альфа на клеточной мембране лимфоцита (Nakamura et al., 1995). Такой гетеротример функционирует как лиганд для рецептора TNFRSF3 /LTBR и участвует в развитии иммунного ответа, обеспечивая межклеточную коммуникацию (Crowe et al., 1994). Считается, что основная функция лимфотоксина бета заключается в стимулировании развития лимфоидной ткани, в первую очередь лимфатических узлов (Onder et al., 2013).

Доместикация (одомашнивание) животных представляет собой гигантский биологический эксперимент, главный результат которого состоит в огромном повышении темпа и размаха изменчивости организмов (Беляев, 1981). Известно, что, хотя виды одомашненных животных относятся к далеко отстоящим систематическим группам, их изменчивость по многим признакам носит характер гомологической изменчивости (Трапезов, 2009). Одним из основных результатов доместикации является сильное повышение воспроизводительных способностей животных (Беляев, 1979). Поэтому доместичированных животных и их диких предков можно рассматривать в качестве модели для изучения механизмов компромисса между репродук-

тивными параметрами и иммунитетом. Можно ожидать, что в процессе доместикации и селекции на повышение хозяйственно важных признаков, преимущества при отборе имели более плодовитые особи, а стало быть, особи со сниженным иммунитетом. Пониженная реактивность иммунной системы могла в дальнейшем стать наследуемой путем фиксации в популяции генов с «неблагоприятными» для иммунореактивности мутациями.

Хорошей моделью для изучения мутаций в генах иммунной системы при доместикации являются домашняя свинья и ее дикий предок – кабан. В результате секвенирования полного генома свиньи в 3'-области гена *LTB* на расстоянии 283 п. н. от последнего экзона был обнаружен однонуклеотидный полиморфизм (SNP) – замена А на G в позиции 27547441 7-й хромосомы (rs340283541).

Цель исследования – оценить распространение SNP rs340283541 гена *LTB* в популяциях домашних свиней и диких кабанов, сравнить относительную экспрессию мРНК этого гена у свиней с разными генотипами и провести анализ потенциальной функциональной роли этого SNP методами биоинформатики.

Материалы и методы

Образцы крови и шкур диких кабанов, представляющих разные подвиды *Sus scrofa* L., получены из России (Воронежский биосферный заповедник, Воронежская область) и Украины (Николаевская и Закарпатская области). Образцы крови свиней пород ландрас, скороспелой мясной СМ-1 и сибирских миниатюрных свиней получены из хозяйств Новосибирской области. ДНК выделяли методом протеолитической обработки с последующей экстракцией фенолом. Фрагмент гена *LTB* амплифицировали с использованием праймеров 5'-TCCCCTCAGACTCAACTGCACAC-3' и 5'-TTCAGGCAGCTGGCAGGGAGAA-3'. Ампликон обрабатывали рестриктазой HpySE526 I (СибЭнзим, Россия). Генотип SNP rs340283541 определяли путем электрофореза продуктов рестрикции в 4%-м полиакриламидном геле

(генотип AA – 172 п. н., генотип AG – 172, 148 и 24 п. н., генотип GG – 148 и 24 п. н.).

Для изучения экспрессии мРНК гена *LTB* в паховом лимфоузле миниатюрных свиней с различными генотипами по этому гену использовали животных в возрасте 1 мес., массой 8–12 кг. В эксперимент брали по четыре животных каждого генотипа. Амплификацию фрагмента гена *LTB* длиной 139 п. н. проводили с праймерами *LTB/F* 5'-AACTGGTAACAGGGACCGCT-3' и *LTB/R* 5'-ATCCAAGCGCCAATGAGGT-3'. В качестве гена сравнения использовали ген *GADPH*. С помощью праймеров *GAPDH/F* 5'-CGTCAAGCTCATTTCTGGTACG-3' и *GAPDH/R* 5'-GGGGTCTGGGATGGAACTGGAAG-3' амплифицировали фрагмент размером 223 п. н.

Суммарную РНК выделяли с помощью реактива TRIzol (Invitrogen, США) согласно рекомендациям производителя. Реакцию обратной транскрипции проводили с использованием олиго-(dT)-праймера и обратной транскриптазы M-MuLV (Сибэнзим, Россия). Реакцию ПЦР в реальном времени (ПЦР-РВ) проводили с использованием набора реагентов для проведения ПЦР-РВ в присутствии красителя SYBR Green I (Синтол, Россия) по стандартной схеме. Полученные данные обрабатывали методом относительного количественного анализа $\Delta\text{-}\Delta\text{Ct}$ с помощью программы Rotor-Gene 6000 Series Software.

Частоты аллелей и генотипов SNP rs340283541 сравнивали с применением критерия χ^2 с поправкой Йетса. Относительный уровень экспрессии мРНК гена *LTB* у сибирских миниатюрных свиней с различным генотипом по SNP rs340283541 сравнивали с помощью однофакторного дисперсионного анализа и критерия Стьюдента.

Построение выравнивания между участком 7-й хромосомы свиньи, включающим SNP, и геномными последовательностями других млекопитающих осуществляли с использованием опций геномного браузера UCSC (<https://genome.ucsc.edu/>). Отображение уровня консервативности нуклеотидов в выравнивании выполняли с помощью программы WebLogo (<http://weblogo.berkeley.edu/logo.cgi>).

Потенциальные сайты связывания транскрипционных факторов в окрестностях SNP rs340283541 выявляли с помощью интернет-доступной программы Match-1.0 Public (<http://www.gene-regulation.com/cgi-bin/pub/programs/match/bin/match.cgi>), осуществляющей поиск потенциальных сайтов в нуклеотидных последовательностях на основе весовых матриц (ВМ) методом PWM. Оценку количества ложноположительных сайтов, выявляемых методом PWM, производили на геномной последовательности 1-й хромосомы свиньи. Данные о QTLs экстрагировали их базы Pig Quantitative Trait Locus Database (Pig QTLdb) (<http://www.animalgenome.org/cgi-bin/QTLdb/SS/index>).

Подробное описание всех процедур см. в Дополнительных материалах 1¹.

результаты

Было проведено генотипирование замены А на G в 3'-фланкирующем районе гена *LTB* свиньи (rs340283541). Частота редкого аллеля А у диких кабанов достоверно не отлича-

лась от его частоты как у отдельных пород свиней, так и у домашних свиней в целом (Доп. материалы 2). Однако частота генотипа GG в выборке кабанов (20,0 %) была достоверно ниже частоты этого же генотипа в объединенной выборке домашних свиней (44,5 %) ($p \leq 0,005$).

Однофакторный дисперсионный анализ не выявил достоверного влияния генотипа SNP rs340283541 на уровень относительной экспрессии мРНК гена *LTB* в ткани пахового лимфатического узла у миниатюрных свиней (Доп. материалы 3). При этом уровень экспрессии мРНК гена *LTB* у мини-свиней с генотипом GG имел тенденцию к повышению по сравнению с носителями аллеля А (объединенная группа GA+AA) ($p < 0,07$).

С использованием данных, экстрагированных из геномного браузера UCSC, произведена оценка уровня консервативности участка, включающего исследуемый SNP. У свиньи SNP расположен в 3'-фланкирующем районе гена *LTB* на расстоянии 284 н. п. от точки терминации транскрипции. У 11 видов млекопитающих в этом районе ДНК были выявлены участки гомологии протяженностью от 38 до 74 нуклеотидов (рисунок, а). Наиболее протяженный участок гомологии (74 нуклеотида) выявлен у коровы. Нуклеотид, соответствующей полиморфной позиции у свиньи, располагался у этих видов организмов на расстоянии от 280 (у овцы) до 445 (у гиббона и шимпанзе) нуклеотидов ниже 3'-конца гена *LTB*. Наиболее часто (в 10 случаях из 12) в данной позиции располагается тимин (рисунок, а). Лишь у двух видов в этой позиции обнаружены другие нуклеотиды: полиморфная позиция аденин/гуанин (свинья) и цитозин (мышь). Десять нуклеотидов, непосредственно прилегающих к SNP на 3'-фланге, являются высоко консервативными. На графике WebLogo данный консервативный район отображен в виде мотива TTTCTCTTGG (рисунок, б).

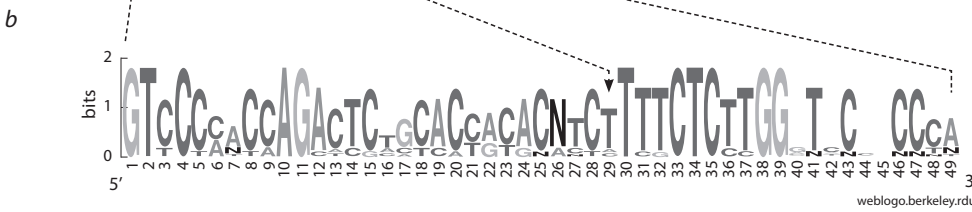
С помощью программы распознавания сайтов связывания транскрипционных факторов Match-1.0 Public были предсказаны потенциальные сайты в двух нуклеотидных последовательностях участка 3'-фланкирующей области гена *LTB* в окрестностях SNP rs340283541, соответствующих аллельным вариантам А и G. Было выявлено, что максимальные значения ВМ имели потенциальные сайты связывания транскрипционных факторов 5 типов (табл. 1). У этих потенциальных сайтов были выявлены: 1) либо значения параметра matrix similarity (сходство последовательности с матрицей) $> 0,85$; 2) либо значения параметра core similarity (сходство последовательности с матрицей по 5 наиболее консервативным позициям матрицы) $> 0,95$; 3) либо значения обоих параметров превышали вышеуказанные пороговые. Два сайта (Brn-2 и AP-1) были обнаружены в последовательности аллеля А. Три сайта (RFX1, USF, ISRE) выявлены в последовательности аллеля G. У двух из пяти сайтов связывания (Brn-2 и RFX1) оба значения ВМ превышали вышеуказанные пороговые. Сравнение участков ДНК, соответствующих предсказанным сайтам, с консенсусными последовательностями показало:

1) в четырех случаях (Brn-2, AP-1, RFX1, ISRE) из пяти полиморфный нуклеотид (А либо G) расположен в коровой части потенциального сайта, соответствующей пяти наиболее консервативным нуклеотидам матрицы;

¹ Дополнительные материалы 1–3 см. в Приложении 4 по адресу: <http://www.bionet.nsc.ru/vogis/download/pict-2015-12/appx4.pdf>

a

Species	The nucleotide sequence at the 3'-flank of the pig <i>LTB</i>	Genomic positions	Distance between SnP (rs340283541) and the 3'-end of the <i>LTB</i> transcript
Pig	gtcccc-tcagactcaacactgcac-tcTttctcttggatccccaccagcagtgatagggcattacagggg-cag	chr7:27,547,415-27,547,485	284
Cow	gtcccc-ccagactcagcactgcac-tcTttctcttggggtccatcccagtagtggtgggtcatgcagggg-aag	chr23:27,532,139-27,532,212	306
Sheep	gtcccc-ccagactcagcactgcac-tcTttctcttggggtccatcccagtagtggtgggtcattgca-----	chr20:26,852,623-26,852,689	280
Mouse	gtcccccaagaccctcacatacaccTtgctcctggctctgcct-----ctgtgaggag	chr17:35,196,687-35,196,744	411
Mouse lemur	gtccccaccagcctgtgcactgcac-tcTttctcttggggtccttcccac---ggaggggccattacagggg-aag	scaffold_8597:54,872-54,801	281
Rat	gttccctcaagaaccctcccatacactTttctcctggctctgcctt-----gctgtga-gaa	chr20:6,941,356-6,941,298	426
Gorilla	gtcccaaccagactctgcaccacac-tcTttctcttgggtaccoccca-----acagtga-gag	chr6:32,473,642-32,473,585	294
Human	gtcccaaccagactctgcaccatac-tcTttctcttgggtaccoccca-----acagtga-gag	chr6:31,548,069-31,548,012	295
Orangutan	gtcccaaccagactctgcaccacac-tcTttctctcgggtcccccca-----gcagtga-gag	chr6:32,078,216-32,078,159	352
Gibbon	gtcctaaccagactctgcaccacgc-tcTttctcttgggtcccccca-----acagtga-gag	chr1a:76,380,055-76,379,998	445
Tarsier	gttcccaaccagattctgcaccacg---cTttctcttggatccctccc-----cagtga-gag	KE944812v1:95,150-95,096	289
Chimpanzee	gtcccaaccagactctgcaccacac-tcTttctcttgg-----	chr6:31,859,616-31,859,579	445



Conservative nucleotides in the 3'-flanking region of the *LTB* gene in twelve mammalian species.

(a) Alignment of regions homologous to the region of pig chromosome 7 containing the rs340283541 SnP. The SnP position in the pig sequence and the corresponding nucleotides in sequences of other species are shown in bold capitals. (b) Presentation of conservative positions with weblogo (Crooks et al., 2004). The overall height of a stack indicates the sequence conservation at the corresponding position (bits on the Y axis), while the height of symbols within the stack reflects their relative frequencies at that position. Numbers on the X axis indicate numbers of positions in the alignment. The rs340283541 SnP (dashed arrow) occurs at position 29 of the pig sequence.

Table 1. Putative transcription factor binding sites predicted in the DnA region comprising SnP rs340283541

Transcription factor binding site	wM scores		Strand	nucleotide sequence of binding site* / c consensus sequence**	wM scores for the DnA region with nucleotide substitution	
	Matrix similarity	core similarity			Matrix similarity	core similarity
Allele A						
Brn-2 [†]	0.857	1.000	-	5'-actc ATTT Ctcttggga-3' 5'-nykn ATTW Ysnatggn-3'	0.713	0.759
AP-1	0.683	0.989	+	5'-tgc ACTCA t-3' 5'-ntg ASTCA g-3'	< 0.550	< 0.600
Allele G						
rFX1	0.888	0.982	-	5'-tc GTTT Ctcttggatccc-3' 5'-nn GTTT RCyatngynacnn-3'	0.693	0.720
iSrE	0.786	1.000	+	5'-tc GTTT Ctcttggat-3' 5'-ca GTTT Cwcttlycc-3'	0.672	0.800
USF	0.846	0.905	+	5'-t GCAC Tcgtt-3' 5'-nn RYCACG trynn-3'	0.732	0.905

The capital letters denote nucleotides that correspond to five most conservative positions in the binding sites predicted by the Match 1.0 tool. The two wM scores that exceed predefined thresholds (matrix similarity>0.85 and core similarity>0.95) are given in bold type and underlined. (†) The false-positive rate in TFBSs prediction using PwM is less than 3E-4 at the recognition threshold corresponding to the score of the putative binding site comprising the SnP (rs340283541). * For the binding sites found in the minus (reverse) strand, we present the DnA sequence from the plus (forward) strand. In such cases, we present the inverted versions of the consensus sequences. ** Positions in the consensus sequences are indicated in the 15-letter IUPAC nucleotide code, where K = G or T; r = A or G; S = c or G; Y = c or T; and w = A or T.

Table 2. Quantitative trait loci (QTLs) co-localized with SnP rs340283541 (position in the porcine genome chr7:27547441) that affect the immune or reproduction traits in pigs.

QTL	QTL iD in Pig QTL db	QTL Span (position on chromosome 7)
Traits related to the immune capacity		
Lymphocyte number	5469	11625414–36993248
Hemolytic complement activity (alternative pathway)	9596	11625414–38992356
Prr SV susceptibility	31803	27019705–27999311
Reproduction-associated traits		
Age at puberty	588	16365408–42509154
corpus luteum number	24284	8014191–92220281
	31864	11136187–116028974
Teat number	5257	10763543–117929721
	584	18392887–31164815

- во всех случаях полиморфный нуклеотид (А либо G) во фрагменте регуляторной последовательности гена *LTB* и соответствующий ему на основании выравнивания нуклеотид консенсусной последовательности совпадают;
- замена нуклеотида в полиморфной позиции приводит к существенному снижению обоих значений ВМ у четырех (Brn-2, AP-1, RFX1, ISRE) из пяти рассматриваемых сайтов (табл. 1, последние две колонки). У сайта связывания USF, выявленного в аллельном варианте G, замена гуанина на аденин приводит к снижению значения matrix similarity и не влияет на значение core similarity.

Для того чтобы выявить сайт, распознанный с наименьшей ошибкой перепредсказания (False Positive Rate), для всех пяти выявленных типов сайтов в качестве оценки было рассчитано количество сайтов данного типа в последовательности хромосомы 1 свиньи. В каждом случае пороговое значение ВМ, использованное функцией распознавания, равнялось значению ВМ соответствующего потенциального сайта, выявленного в окрестностях SNP. Было обнаружено, что минимальная плотность сайтов в геномной последовательности ДНК выявляется при распознавании сайта BRN2 (в табл. 1 помечено знаком «†»).

Для того чтобы охарактеризовать потенциальную роль участка хромосомы 7 свиньи, включающего полиморфизм rs340283541, в формировании фенотипических признаков, мы проанализировали данные из базы Pig QTLdb. Было выявлено, что полиморфизм rs340283541 содержится в 161 локусе количественных признаков (QTLs). Из них три локуса влияют на три различных показателя иммунной системы (табл. 2), а другие пять QTLs ассоциированы с тремя репродуктивными параметрами животных.

Обсуждение

В результате проведенного исследования нами обнаружено, что генотип GG SNP rs340283541 у диких кабанов встречается гораздо реже, чем у домашних свиней (Доп. материалы 2). Это позволяет выдвинуть предположение, что замена А на G в этой позиции может нарушать иммун-

ный ответ, что, согласно гипотезе компромисса, должно приводить к улучшению репродуктивных показателей у животных, гомозиготных по этой мутации. Однако против такого предположения говорит повышенный уровень экспрессии мРНК гена *LTB* у животных с генотипом GG, по сравнению с носителями аллеля А (Доп. материалы 3). Альтернативное объяснение этих фактов заключается в вероятном существовании положительной корреляции между иммунным ответом и репродуктивными признаками при доместикации свиней. Поскольку процесс доместикации, несомненно, сопровождался усилением давления со стороны паразитов, селекция на улучшение репродуктивных показателей могла быть эффективной только у генетически устойчивых животных с сильным иммунитетом.

В литературе имеется ряд работ, косвенно свидетельствующих о вероятном отборе «патологических» аллелей в генах иммунной системы при селекции сельскохозяйственных животных на повышение хозяйственно важных признаков. Например, селекция коров красной норвежской породы на увеличение удоев молока сопровождалась повышением частоты мастита, кератоза и задержания последа, что свидетельствует о снижении у них функционального состояния иммунной системы (Heringstad et al., 2007). При селекции на пониженную частоту мастита у коров той же породы происходило снижение содержания белка в молоке (Heringstad et al., 2005). Интересно, что мутация в гене *LTB* у мышей приводит к нарушению формирования волосяных фолликулов (Cui et al., 2006). Поэтому будет интересно сравнить окраску волосяного покрова у миниатюрных свиней с различными генотипами по гену *LTB*.

Консервативность района ДнК в окрестностях SnP rs340283541

Анализ нуклеотидных последовательностей 3'-районов генов *LTB* млекопитающих показал, что SNP rs340283541 располагается на консервативном участке (рисунок). Это означает, что данный участок с высокой долей вероятности несет определенную функциональную нагрузку.

Известно, что 3'-районы генов могут участвовать в регуляции транскрипции генов. Так же, как и промоторы генов и регуляторные районы, расположенные в 5'-районах, и интронах, 3'-области генов могут включать функционально-активные сайты связывания транскрипционных факторов, оказывающих как активирующие, так и подавляющие эффекты на транскрипцию (Kolchanov et al., 2002; Меркулова и др., 2013; Игнатъева и др., 2015). Основываясь на данном представлении, мы предположили, что замена нуклеотида в полиморфной позиции SNP rs340283541 может нарушать сайт связывания транскрипционного фактора, участвующего в регуляции активности гена *LTB*. Для того чтобы проверить данное предположение, нами был произведен поиск потенциальных сайтов связывания транскрипционных факторов в последовательности ДНК, включающей SNP rs340283541. Было выявлено, что аллельный вариант А содержит потенциальные сайты связывания *Brg-2* и AP-1, которые нарушаются при замене аденина на гуанин в полиморфной позиции. С другой стороны, аллельный вариант G 3'-регуляторной последовательности гена *LTB* содержит потенциальные сайты RFX1, ISRE, USF, которые не удается выявить в случае, если в полиморфной позиции присутствует аденин (табл. 1).

Характеристика предсказанных сайтов

Brg-2. Фактор *Brg-2* (*Brain-2*) кодируется геном *POU3F2* (*POU class 3 homeobox 2*), который, согласно данным базы EntrezGene, в геноме свиньи еще имеет статус «MODEL» (предсказан на основе электронной аннотации и слабо изучен). Относительно белка *Brg-2/Oct7* у человека и у мыши в научной литературе сложилось представление как о транскрипционном факторе, участвующем в дифференцировке нейрональных клеток (Zhao et al., 2013). Гиперпродукция фактора *Brg-2* ассоциирована с усиленной пролиферацией клеток меланомы (Goodall et al., 2004). Согласно данным базы BioGPS, у человека ген *POU3F2*, кодирующий фактор *Brg-2*, экспрессируется не только в нейральных клетках, но и в других тканях и органах, включая лимфатические узлы. Известно, что *Brg-2* может выступать в качестве репрессора транскрипционной активности гена *CDH13* (*T-cadherin*) в клетках меланомы (Ellmann et al., 2012). Это позволяет предположить, что в случае, если данный сайт является функциональным у животных с аллелем А, *Brg-2* также может выполнять роль репрессора и замена аденина на гуанин (аллель G) может приводить к снятию репрессирующего эффекта фактора *Brg-2* и активации транскрипции гена *LTB*.

AP-1. Фактор AP-1 представляет собой гетеродимер, в состав которого могут входить белки таких семейств, как c-Fos, c-Jun, ATF и JDP. Фактор AP-1 участвует в регуляции генов, экспрессирующихся в широком круге тканей, в ответ на множество факторов, включая цитокины, факторы роста, стрессовые сигналы, бактериальные и вирусные инфекции (Hess et al., 2004). У потенциального сайта связывания фактора AP-1 наблюдается наиболее существенное (по сравнению с потенциальными сайтами связывания других факторов) падение значения *core similarity* при замене аденина на гуанин в полиморфной позиции. Как правило, фактор AP-1 является активато-

ром транскрипции, однако имеются свидетельства того, что AP-1 может функционировать в роли репрессора (Mittelstadt et al., 2012). Таким образом, присутствие сайта связывания AP-1 в последовательности аллеля А может быть причиной более низкого уровня транскрипции гена у животных с генотипами AA и AG, по сравнению с животными с генотипом GG.

RFX1. Ген фактора RFX1 (*Regulatory factor X1*) имеет статус «MODEL» в геноме свиньи согласно базе данных EntrezGene. У человека RFX1 регулирует активность генов, экспрессирующихся в различных тканях (включая лимфоидные), среди которых идентифицированы гены 2-го класса главного комплекса гистосовместимости (Fontes et al., 1997), а также гены вируса гепатита С и вируса Эпштейна–Бара (Tierney et al., 2000). Согласно данным базы GeneCards, у человека ген *RFX1* экспрессируется в лимфатических узлах, моноцитах, Т- и В-лимфоцитах.

ISRE. Согласно базе TRANSFAC MATRIX (Release 7), с сайтом ISRE (*interferon-stimulated response element*) (идентификатор матрицы V\$ISRE_01) связывается фактор (мультимерный комплекс) ISGF-3 (*Interferon-Stimulated Gene Factor-3*), включающий три субъединицы из семейства STAT и одну субъединицу из семейства IRF (Kessler et al., 1990). Комплекс ISGF-3 опосредует активацию генов в ответ на интерферон- α и - β , и сайт ISRE присутствует в регуляторных районах целого ряда генов, активно экспрессирующихся в клетках лимфоидного ряда (Kolchanov et al., 2002; Ananko et al., 2007). Таким образом, появление потенциального сайта ISRE в 3'-регуляторной последовательности гена *LTB* вследствие замены аденина на гуанин (аллель G) с высокой вероятностью может повлечь за собой активацию экспрессии гена *LTB* в лимфатических узлах у животных с генотипом GG. Потенциальный сайт ISRE, выявленный в последовательности ДНК аллельного варианта G, является перспективным кандидатом для дальнейшего экспериментального исследования в связи с тем, что, как показано нами ранее, частота носителей полноразмерных геномов ретровирусов PERV типов А, В и С у сибирских мини-свиней очень велика (Айтназаров и др., 2014).

USF. Факторы семейства USF (*upstream transcription factor*) еще слабо изучены у свиньи. В базе EntrezGene статусы для генов *USF1* и *USF2* обозначены как «PROVISIONAL» и «MODEL». У человека и мыши белки USF1 и USF2 известны как повсеместно экспрессирующиеся факторы. USF1 и USF2 участвуют в регуляции генов, функционирующих в регуляторных сетях стрессового и иммунного ответа, клеточного цикла и пролиферации (Corre, Galibert, 2006).

Предположение о том, что выявленные нами потенциальные сайты в районе, включающем SNP rs340283541, могут быть функциональными и их повреждение в результате замены нуклеотида может влиять на экспрессию гена, должно рассматриваться с определенной долей осторожности. Это связано с тем, что компьютерные методы распознавания сайтов (включая метод PWM, реализованный ресурсом Match-1.0 Public) характеризуются существенными погрешностями распознавания (ошибками недо- и перепредсказания) (Levitsky et al., 2007). По нашим оценкам, выполненным на геномной

последовательности хромосомы 1 свиньи, с наименьшим уровнем перепредсказания (т. е. с наибольшей точностью) в нашем исследовании был выявлен сайт связывания *Bm2* (табл. 1).

С другой стороны, известно, что сайты, входящие в состав композиционных элементов (пар близкорасположенных сайтов посадки транскрипционных факторов), могут обладать невысоким сродством к соответствующим факторам (распознаваться методом PWM только при низких значениях *BM*) и, тем не менее, быть функциональными (Kel-Margoulis et al., 2000). В этом случае стабилизация взаимодействия участка ДНК с соответствующим транскрипционным фактором может осуществляться за счет дополнительных белок-белковых контактов с другим фактором, также контактирующим с соседним участком ДНК. Учитывая это обстоятельство, а также биологический контекст, в котором может функционировать сайт связывания, мы предполагаем, что, наряду с сайтом *Bm-2*, для дальнейшего исследования могут быть также интересны сайты связывания факторов, специфичных для клеток иммунной системы: *RFX1* и *ISGF-3* (сайт *ISRE*).

Биоинформатический анализ локусов количественных признаков

В ряде исследований показано, что многие однонуклеотидные полиморфизмы, ассоциированные с фенотипическими признаками, не являются причинными (causative), т. е. не оказывают эффекта ни на функцию белка, ни на интенсивность экспрессии гена. Однако выясняется, что рассматриваемые полиморфные аллели являются маркерами определенных гаплотипов, в составе которых имеются причинные SNP (Seddon et al., 2010; Loos et al., 2014). Учитывая такую возможность, мы проанализировали базу данных Pig QTLdb и выявили, что участок 7-й хромосомы, на котором расположен SNP rs340283541, входит в состав трех QTLs, имеющих отношение к иммунному ответу, и пяти QTLs, имеющих отношение к репродуктивным признакам (табл. 2). Эти данные позволяют рассматривать SNP rs340283541 как маркер определенного гаплотипа, подвергающегося селекции в ходе доместикации. Однако отбору могла подвергаться мутация в другом гене, поскольку в непосредственной близости к гену *LTB* находятся гены фактора некроза опухоли альфа (*TNFA*) и лимфотоксина альфа (*LTA*), которые также кодируют белки иммунной системы. О возможном вкладе SNP rs340283541 в формирование компромисса между репродуктивным успехом и иммунитетом при доместикации говорит и то, что он также входит в состав QTL для возраста полового созревания, числа сосков и числа желтых тел в яичниках, т. е. основных признаков плодovitости у свиней.

Локус генов *LTB-TNFA-LTA* человека насыщен регуляторными элементами, участвующими в транскрипционном и пост-транскрипционном контроле (Недоспасов, Купраш, 2008). У свиньи этот локус характеризуется необычно высокой плотностью генов. Так, промежуток между генами *LTB* и *TNF*, в котором находится SNP rs340283541, составляет всего 2461 п. н., в то время как в геноме человека (имеющего такую же длину, что и геном свиньи), межгенное расстояние в 50 % случаев превышает

величину 3949 п. н. (Djebali et al., 2012), а среднее межгенное расстояние в геномах млекопитающих составляет 91 тыс. п. н. (Zhang et al., 2014). Не исключено, что участок между генами *LTB* и *TNF* содержит регуляторные элементы и аллельные варианты полиморфизма SNP rs340283541 маркируют варианты протяженного гаплотипа, определяющие активность одного из регуляторных участков.

Таким образом, обнаружено, что при доместикации свиньи произошло изменение частоты генотипов SNP в 3'-области гена *LTB*. По-видимому, данный SNP может быть ассоциирован с различным уровнем экспрессии мРНК гена *LTB*. Биоинформатический анализ показал, что участок 3'-района гена *LTB*, окружающий SNP, консервативен и замена нуклеотида в полиморфной позиции может влиять на потенциальные сайты связывания транскрипционных факторов. Выявлено 5 типов сайтов связывания факторов (*Bm-2*, *AP-1*, *RFX1*, *ISRE*, *USF*). Согласно нашим оценкам, сайт фактора *Bm-2* распознан с наименьшей ошибкой перепредсказания, а с биологической точки зрения наиболее интересными в плане дальнейшего исследования являются сайты связывания *RFX1* и *ISRE*. Также не исключено, что полиморфизм rs340283541 находится в неравновесии по сцеплению с другой, пока не известной, функционально значимой мутацией в гене *LTB*.

Acknowledgments

The experiments were financially supported by the Russian Foundation for Basic Research, project 13-04-00968a. The bioinformatical analysis was supported by Budgeted Project VI.61.1.2.

The authors are grateful to A.S. Klimov (Voronezh State University) for assistance in the collection of boar fur samples.

Conflict of interest

The authors declare no conflict of interest.

References

- Aitnazarov R.B., Yudin N.S., Nikitin S.V., Ermolayev V.I., Voevoda M.I. Identification of whole genomes of endogenous retroviruses in Siberian miniature pigs. *Vavilovskii Zhurnal Genetiki i Selektii*=Vavilov Journal of Genetics and Breeding. 2014;18(2): 294-297.
- Ananko E.A., Kondrakhin Y.V., Merkulova T.I., Kolchanov N.A. Recognition of interferon-inducible sites, promoters, and enhancers. *BMC Bioinformatics*. 2007;8:56.
- Ardia D.R., Parmentier H.K., Vogel L.A. The role of constraints and limitation in driving individual variation in immune response. *Functional Ecology*. 2011;25(1):61-73. DOI 10.1111/j.1365-2435.2010.01759.x
- Belyaev D.K. Destabilizing selection as a factor of variability in domestication. *Priroda=Nature (Moscow)*. 1979;2:36-45.
- Belyaev D.K. Destabiliziruyushchiy otkor kak faktor domestikatsii [Destabilizing selection as a factor in domestication]. *Genetika i blagosostoyanie chelovechestva [Genetics and the wellbeing of mankind]*. Moscow, 1981:53-66.
- Balenger S.L., Zuk M. Testing the Hamilton-Zuk hypothesis: past, present, and future. *Integr Comp. Biol*. 2014;54(4):601-613. DOI 10.1093/icb/ucu059
- Corre S., Galibert M.D. *USF* as a key regulatory element of gene expression. *Med. Sci. (Paris)*. 2006;22(1):62-67.

- Crooks G.E., Hon G., Chandonia J.M., Brenner S.E. WebLogo: a sequence logo generator. *Genome Res.* 2004;14(6):1188-1190.
- Crowe P.D., VanArsdale T.L., Walter B.N., Ware C.F., Hession C., Ehrenfels B., Browning J.L., Din W.S., Goodwin R.G., Smith C.A. A lymphotoxin-beta-specific receptor. *Science.* 1994;264(5159):707-710.
- Cui C.Y., Hashimoto T., Grivennikov S.I., Piao Y., Nedospasov S.A., Schlessinger D. Ectodysplasin regulates the lymphotoxin-beta pathway for hair differentiation. *Proc. Natl Acad. Sci. USA.* 2006; 103(24):9142-9147.
- Djebali S., Davis C.A., Merkel A., Dobin A., Lassmann T., Mortazavi A., Tanzer A., Lagarde J., Lin W., Schlesinger F., Xue C., Marinov G.K., Khatun J., Williams B.A., Zaleski C., Rozowsky J., Röder M., Kococinski F., Abdelhamid R.F., Alioto T., Antoshechkin I., Baer M.T., Bar N.S., Batut P., Bell K., Bell I., Chakraborty S., Chen X., Chrast J., Curado J., Derrien T., Drenkow J., Dumais E., Dumais J., Duttagupta R., Falconnet E., Fastuca M., Fejes-Toth K., Ferreira P., Foissac S., Fullwood M.J., Gao H., Gonzalez D., Gordon A., Gunawardena H., Howald C., Jha S., Johnson R., Kapranov P., King B., Kingswood C., Luo O.J., Park E., Persaud K., Preall J.B., Ribeca P., Risk B., Robyr D., Sarmeth M., Schaffer L., See L.H., Shahab A., Skancke J., Suzuki A.M., Takahashi H., Tilgner H., Trout D., Walters N., Wang H., Wrobel J., Yu Y., Ruan X., Hayashizaki Y., Harrow J., Gerstein M., Hubbard T., Reymond A., Antonarakis S.E., Hannon G., Giddings M.C., Ruan Y., Wold B., Carninci P., Guigó R., Gingeras T.R. Landscape of transcription in human cells. *Nature.* 2012;489(7414):101-108. DOI 10.1038/nature11233
- Ellmann L., Joshi M.B., Resink T.J., Bosserhoff A.K., Kuphal S. BRN2 is a transcriptional repressor of CDH13 (T-cadherin) in melanoma cells. *Lab Invest.* 2012;92(12):1788-1800. DOI 10.1038/labinvest.2012.140
- Fontes J.D., Jabrane-Ferrat N., Peterlin B.M. Assembly of functional regulatory complexes on MHC class II promoters *in vivo*. *J. Mol. Biol.* 1997;270(3):336-345.
- Goodall J., Martinozzi S., Dexter T.J., Champeval D., Carreira S., Larue L., Goding C.R. Brn-2 expression controls melanoma proliferation and is directly regulated by beta-catenin. *Mol. Cell Biol.* 2004;24(7):2915-2922.
- Heringstad B., Chang Y.M., Gianola D., Klemetsdal G. Genetic association between susceptibility to clinical mastitis and protein yield in norwegian dairy cattle. *J. Dairy Sci.* 2005;88(4):1509-1514.
- Heringstad B., Klemetsdal G., Steine T. Selection responses for disease resistance in two selection experiments with Norwegian red cows. *J. Dairy Sci.* 2007;90(5):2419-2426.
- Hess J., Angel P., Schorpp-Kistner M. AP-1 subunits: quarrel and harmony among siblings. *J. Cell Sci.* 2004;117(25):5965-5973.
- Ignatieva E.V., Podkolodnaya O.A., Orlov Y.L., Vasiliev G.V., Kolchanov N.A. Regulatory genomics: integrated experimental and computer approaches. *Genetika=Genetics (Moscow).* 2015;51(4): 409-429.
- Kel-Margoulis O.V., Romashchenko A.G., Kolchanov N.A., Winger E., Kel A.E. COMPEL: a database on composite regulatory elements providing combinatorial transcriptional regulation. *Nucl. Acids Res.* 2000;28(1):311-315.
- Kessler D.S., Veals S.A., Fu X.Y., Levy D.E. Interferon-alpha regulates nuclear translocation and DNA-binding affinity of ISGF3, a multimeric transcriptional activator. *Genes Dev.* 1990;4(10):1753-1765.
- Kim J.Y., Moon S.M., Ryu H.J., Kim J.J., Kim H.T., Park C., Kim K., Oh B., Lee J.K. Identification of regulatory polymorphisms in the TNF-TNF receptor superfamily. *Immunogenetics.* 2005;57(5): 297-303.
- Kolchanov N.A., Ignatieva E.V., Ananko E.A., Podkolodnaya O.A., Stepanenko I.L., Merkulova T.I., Pozdnyakov M.A., Podkolodny N.L., Naumochkin A.N., Romashchenko A.G. Transcription regulatory regions database (TRRD): its status in 2002. *Nucl. Acids Res.* 2002;30(1):312-317.
- Levitsky V.G., Ignatieva E.V., Ananko E.A., Turnaev I.I., Merkulova T.I., Kolchanov N.A., Hodgman T.C. Effective transcription factor binding site prediction using a combination of optimization, a genetic algorithm and discriminant analysis to capture distant interactions. *BMC Bioinformatics.* 2007;8:481.
- Loos R.J., Yeo G.S. The bigger picture of FTO: the first GWAS-identified obesity gene. *Nat. Rev. Endocrinol.* 2014;10(1):51-61. DOI 10.1038/nrendo.2013.227
- Mittelstadt M.L., Patel R.C. AP-1 mediated transcriptional repression of matrix metalloproteinase-9 by recruitment of histone deacetylase 1 in response to interferon β . *PLoS One.* 2012;7(8):e42152. DOI 10.1371/journal.pone.0042152
- Merkulova T.I., Ananko E.A., Ignatieva E.V., Kolchanov N.A. Regulatory transcription codes in eukaryotic genomes. *Genetika=Genetics (Moscow).* 2013;49(1):37-54.
- Nakamura T., Tashiro K., Nazarea M., Nakano T., Sasayama S., Honjo T. The murine lymphotoxin-beta receptor cDNA: isolation by the signal sequence trap and chromosomal mapping. *Genomics.* 1995; 30(2):312-319.
- Nedospasov S.A., Kuprash D.V. Tumor necrosis factor and lymphotoxin: physiological function and role in cytokine and anti-cytokine therapy. *Russkiy zhurnal "SPID, rak i obshchestvennoe zdorove" = Russian Journal of AIDS, Cancer, and Public Health.* 2008; 12(1):69-76.
- Onder L., Danuser R., Scandella E., Firner S., Chai Q., Hehlhans T., Stein J.V., Ludewig B. Endothelial cell-specific lymphotoxin- β receptor signaling is critical for lymph node and high endothelial venule formation. *J. Exp. Med.* 2013;210(3):465-473. DOI 10.1084/jem.20121462
- Seddon J.M., Berggren K.T., Fleeman L.M. Evolutionary history of DLA class II haplotypes in canine diabetes mellitus through single nucleotide polymorphism genotyping. *Tissue Antigens.* 2010;75(3):218-226. DOI 10.1111/j.1399-0039.2009.01426.x
- Sheldon B.C., Verhulst S. Ecological immunology: costly parasite defences and trade-offs in evolutionary ecology. *Trends Ecol. Evol.* 1996;11(8):317-321.
- Tierney R., Kirby H., Nagra J., Rickinson A., Bell A. The Epstein-Barr virus promoter initiating B-cell transformation is activated by RFX proteins and the B-cell-specific activator protein BSAP/Pax5. *J. Virol.* 2000;74(22):10458-10467.
- Trapezov O.V. Darwinism and the lessons of practical selection in Russia. *Vavilovskii Zhurnal Genetiki i Selektii = Vavilov Journal of Genetics and Breeding.* 2009;13(2):249-297.
- van der Most P.J., de Jong B., Parmentier H.K., Verhulst S. Trade-off between growth and immune function: a meta-analysis of selection experiments. *Funct. Ecol.* 2011;25(1):74-80. DOI 10.1111/j.1365-2435.2010.01800.x
- Zhao F.Q. Octamer-binding transcription factors: genomics and functions. *Front Biosci.* 2013;18:1051-1071.
- Zhang G., Li C., Li Q., Li B., Larkin D.M., Lee C., Storz J.F., Antunes A., Greenwold M.J., Meredith R.W., Ödeen A., Cui J., Zhou Q., Xu L., Pan H., Wang Z., Jin L., Zhang P., Hu H., Yang W., Hu J., Xiao J., Yang Z., Liu Y., Xie Q., Yu H., Lian J., Wen P., Zhang F., Li H., Zeng Y., Xiong Z., Liu S., Zhou L., Huang Z., An N., Wang J., Zheng Q., Xiong Y., Wang G., Wang B., Wang J., Fan Y., da Fonseca R.R., Alfaro-Núñez A., Schubert M., Orlando L., Mourier T., Howard J.T., Ganapathy G., Pfenning A., Whitney O., Rivas M.V., Hara E., Smith J., Farré M., Narayan J., Slavov G., Romanov M.N., Borges R., Machado J.P., Khan I., Springer M.S., Gatesy J., Hoffmann F.G., Opazo J.C., Håstad O., Sawyer R.H., Kim H., Kim K.W., Kim H.J., Cho S., Li N., Huang Y., Bruford M.W., Zhan X., Dixon A., Bertelsen M.F., Derryberry E., Warren W., Wilson R.K., Li S., Ray D.A., Green R.E., O'Brien S.J., Griffin D., Johnson W.E., Haussler D., Ryder O.A., Willerslev E., Graves G.R., Alström P., Fjeldså J., Mindell D.P., Edwards S.V., Braun E.L., Rahbek C., Burt D.W., Houde P., Zhang Y., Yang H., Wang J., Avian Genome Consortium; Jarvis E.D., Gilbert M.T., Wang J. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science.* 2014;346(6215):1311-1320. DOI 10.1126/science.1251385