БИОМИКА/BIOMICS

ISSN 2221-6197 http://biomicsj.ru

ULTRA-LONG SEQUENCING TECHNOLOGY AND ITS APPLICATION TO TEST THE DIVERSITY OF GENES OF INNATE IMMUNITY OF CATTLE

Karel Novák[1], Alexander E. Kalashnikov[2, 3, *]

[1]Institute of Animal Science, Přátelství 815 104 00 Praha Uhříněves, Czech Republic
[2]All-Russian Research Institute of Animal Breeding, 141212 Lesnye Polyany, Russia
[3]Ernst Russian Research Institute for Animal Husbandry, 142132 Dubrovitsy, Russia,
*E-mail - aekalashnikov@yandex.ru

Resume

The application of the developed approach for the screening polymorphism of TLR genes in a representative group of historical and modern cattle breeds from Russia is relevant today. The pipeline included the stages of obtaining overlapping amplification products from the coding regions of all ten cattle *TLR* genes, their subsequent purification and normalization. While the antibacterial group included *TLR1*, *TLR2*, *TLR4*, *TLR5*, and *TLR6*, the antiviral group includes *TLR3*, *TLR7*, *TLR8*, *TLR9*, and *TLR10* (despite unclear specificity). For the analysis, 575 animals were used, both bulls and cows. Combined samples containing equimolar concentrations of amplicons obtained from pooled genomic DNA were sequenced on the PacBio platform. Identified structural 36 variants of TLR were annotated according to their biological significance. Both new and already identified patches of variability, already annotated and documented in dbSNP, were found. The data are necessary for the further breeding of local breeds in Russia with regard to their natural resistance to various diseases.

Keywords: cattle, immunity, Toll-like receptors, SMRT sequencing

Citation: Novák K., Kalashnikov A.E. Ultra-long sequencing technology and its application to test the diversity of genes of innate immunity of cattle. *Biomics.* 2019. V.11(4). P. 418-421. DOI: 10.31301/2221-6197.bmcs.2019-34

© The Author(s)

Introduction

The application of the developed approach for the screening polymorphism of *TLR* genes in a representative group of historical and modern cattle breeds from Russia is relevant today. *TLR* genes code for Toll-like receptors, which represent a key part of the system recognizing microorganisms and viral nucleic acids associated with bovine diseases and disorders. Among others, they are known to activate. For this, the signalling pathways that involve the interferon regulatory factors (IRFs), as a family of transcription factors known to play a critical role in antiviral defense, cell growth and immune regulation. Three IRFs (IRF3, IRF5 and IRF7) functioned as direct transducers of virus-mediated TLR signalling, and their activation depends on the TLR stimulated: thus, while TLR3 activates IRF3 and IRF7, while TLR7 and TLR8 activate IRF5 and IRF7. In addition to these IRFs, several kinase-pathways are involved in fine-tuning the intracellular response.

Toll-like family (TLR) are essential in vital when the formingation of and the innate immune response from of animals to various pathogens, for example fungal, bacterial and viral origin [10-11]. The variability of immunity genes with high-throughput technology has been studied and shown to be a modern and real, and cattle immunity genes [16], and TLR-1-10 genes that recognize viral nucleic acids and microorganisms of bovine diseases and disorders was investigated.

The diversity of the Toll-like receptors is of direct relevance to the selection of highly productive animals demonstrating high disease resistance. The obtained data are also expected to be used in the implementation of conservation programs for native breeds in the diverse regions of Russia.

Materials and methods

The tissue samples comprised cows and bulls of Kholmogor breed (Pechora type), Yakut breed, Yaroslavl breed, Simmental dairy and meat breed variant, Black-white breed (black Holstein), and the cattle-forest buffalo hybrid. Altogether, 575 samples of 14 populations were used for the study. DNA from the sperm and cartilage specimen was hydrolyzed with proteinase K and after 8-hour incubation, DNA was isolated by separation by MagSep magnetic particles using the robotic station epMotion 5070 (Eppendorf, Germany). The DNA concentration was determined by UV-spectrophotometry and its integrity were checked electrophoretically. DNA samples were normalized to 50 ng/μL.

The coding sequences of the *TLR* genes were amplified in a series of PCR reactions performed on an eqimolar DNA pool. Subsequently, an equimolar combined sample of all amplicons (31 of anti-bacterial *TLR*s, 52 of anti-viral *TLR*s) [9, 17] was prepared according to the electrophoretically determined yield of PCR reactions and the product molecular weight. The pooled amplicon sample was purified using a NucleoSpin column (Macherey-Nagel, Germany). After additional purity checks with gel and Agilent chip electrophoresis, the amplicons were sequenced using SMRT sequencing technology (Pacific Biosciences, USA) on the RS-II instrument in the GATC Biotech (Constanz, Germany) sequencing center. The average coverage for sequencing were up to 76 reads of genes from 400 to 1200 bp in length, with a depth of coverage of 3-12 per individual. During sequencing, reading lengths of up to 35-37 kb were obtained, which were then cut according to the repeat cycle and sent to the duplicate removal process.

Results and Discussions

The primary processing of our data was carried out according to the flowchart implemented in the UGENE software package (version 1.23.1, www.ugene.net) [12]. The quality assessment was performed with FastQC [9]. The reads of appropriate quality were aligned to the reference genome Bostau6 [1] using BWA-MEM [2]. Removing duplicates was progressed in Cigar MDWMC [3], SNV was detected using FreeBayes [7] and SAMTOOLS [14] (Tables 1-2).

Table 1

Parameters of the *TLR* gene resequencing

Gene	Chromosome	Start	End	Coverage	Reads/ coverage	Reads/ coverage RemDb	Combi of coverage RemDb
TLR1	chr6	59,678,173	59,689,488	22746-38236	455901-553591	2265-4380	386753-464166
TLR2	chr17	3,949,870	3,963,092	14784-23262	448428-543892	2140-3327	394409-478635
TLR4	chr8	108,828,899	108,839,911	19641-32480	455348-553106	2076-3686	381340-459607
TLR5	chr16	27,303,742	27,306,323	13335-22907	453639-550379	1539-2650	387898-467064
TLR6	chr6	59,686,794	59,720,509	22746-38236	455901-553591	2265-4380	386753-464166
TLR7	X	141,044,355	141,063,596	22750-38350	455950-553600	2270-4390	386755-464167
TLR8	X	141,063,596	141,063,596	14800-23300	448430-543750	2145-3330	394550-478700
TLR9	chr22	49,229,610	49,233,939	19550-32600	455200-553250	2080-3690	381450-459750
TLR10	chr6	59,670,230	59,677,223	13252-22950	453700-550400	1540-2655	387920-467124

Table 2

Variants indentified in the *TLR* genes

Gene	Coverage/sample	Mutations (from Byers calc)	Mutations (from SAMTOOLS	QC, score	GC, %
TLR1	5.2-15.9 (for 2 genes)	36 (for 2 genes)	30 (for 2 genes)	9-19	44
TLR2	7.7-12.1	8	30	6-14	43
TLR4	7.5-13.4	13	28	5-13	43
TLR5	5.6-9.6	28	70	5-13	43
TLR6	5.2-15.9 (for 2 genes)	36 (for 2 genes)	30 (for 2 genes)	9-19	44
TLR7	3.7-6.2	4	6	6-14	44
TLR8	7.5-14.2	8	12	5-13	43
TLR9	5.5-9.5	12	22	5-13	44
TLR10	5.2-8.0	5	7	3-12	44

The application of the FreeBayes algorithm in combination with SAMTOOLS allowed in an automatic mode to identify meaningful variants of mutations "in the flow".

Abstract and validation of the variant alleles obtained were carried out with the help of VeIP [18].

In the future work, the found genotypes will require validation in order to exclude errors that occur during poor-quality sequencing. Currently, the genotyping assays exploiting the primer extension technology as enabled by the SNaPshot Multiplex Kit (Applied Biosystems, USA) are applied to this task. The validated novel SNPs will be added to the dbSNP database [4].

Only after the completion of these stages, the found single nucleotide variation (SNP/SNV) will be used for the association studies with the disease resistance/susceptibility traits. Application of the algorithms for factor statistics implemented in the SAS package [15] and the definition of haplotypes is presumed [13].

We plan to refine animal diagnostics by the monitoring of the expression of immunity genes, as elicited in experimental groups of animals with external factors [19]. The available techniques of parallel sequencing or real-time PCR will be applied according to the number of targets and the expediency of scaling, as exemplified in infected animals with different viral load [5, 8].

In conclusion, sequences of *TLR* genes of cattle from seven different breeds were obtained. Method of next generation sequencing made it possible to identify from 5 to 36 suspected mutations in the coding sequences of the screened genes. This information allows for the determination of their occurrence in the population using direct genotyping assays. The SAM-TOOLS-method as a whole revealed more mutations in genes than the method of Bayers [7]. Identified mutations require further assignment at the individual level to identify associations with the susceptibility to diseases.

Acknowledgments

The work was supported by the Program in the Research Institution Development of the Ministry of Agriculture of the CR no MZERO0714 and by National Program of Conservation and Use of Animal Genomic Resources.

References

1. Bostau6. https://genome.ucsc.edu/
2. BWA-MEM algorithm. www.bio-bwa.sourceforge.net/
3. Cigar MDWMC. http://broadinstitute.github.io/picard/
4. dbSNP database. https://www.ncbi.nlm.nih.gov/SNP/
5. Farias, M.V., Lendez, P.A., Marin, M., Quintana, S., Martínez-Cuesta, L., Ceriani, M.C., Dolcini, G.L. Toll-like receptors, IFN-γ and IL-12 expression in bovine leukemia virus-infected animals with low or high proviral load. *Res. Vet. Sci*. 2016. V.107. P. 190-195. doi: 10.1016/j.rvsc.2016.06.016
6. FastQC.(http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
7. FreeBayes. https://github.com/ekg/freebayes
8. Harrison, E., Koufopanou, V., Burt, A., MacLean, R.C. The cost of copy number in a selfish genetic element: the 2-μm plasmid of *Saccharomyces cerevisiae*. *J. Evol. Biol*. 2012. V.25. P. 2348-2356. doi: 10.1111/j.1420-9101.2012.02610.x
9. IAD85455 197 Termo Ion Torrent Design *Bos Taurus* TLR1-10. DOI: 10.13140/RG.2.2.28465.02405
10. Mansilla, F.C., Quintana, M.E., Cardoso, N.P., Capozzo, A.V. Fusion of foreign T-cell epitopes and addition of TLR agonists enhance immunity against *Neospora caninum* profilin in cattle. *Parasite*

Immunol. 2016. V.38. P. 663-669. doi: 10.1111/pim.12354
11. Novák, K. Functional polymorphisms in Toll-like receptor genes for innate immunity in farm animals. *Vet. Immunol. Immunopathol*. 2014. V.157. P. 1-11.
12. Okonechnikov, K., Golosova, O., Fursov, M. UGENE team Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*. 2012. V.28. P.1166–1167. doi: 10.1093/bioinformatics/bts091
13. Patel, S.M., Koringa, P.G., Nathani, N.M., Patel, N.V., Shah, T.M., Joshi, C.G. Exploring genetic polymorphism in innate immune genes in Indian cattle (*Bos indicus*) and buffalo (*Bubalus bubalis*) using next generation sequencing technology. *Meta Gene*. 2015. V.3. P. 50-58. DOI: 10.1016/j.mgene.2015.01.002
14. SAMTOOLS. www.htslib.org/
15. SAS package. www.sas.com
16. Schaut, R.G., Ridpath, J.F., Sacco, R.E. Bovine viral diarrhea virus type 2 impairs macrophage responsiveness to Toll-like receptor ligation with the exception of Toll-like receptor 7. *PLoS One*. 2016. 11: e0159491. doi: 10.1371/journal.pone.0159491
17. TLR 1-10 *Bos taurus* genome design to Illumina MiSEQ Sequencing (114014) Illumina Design Studio. DOI:10.13140/RG.2.2.26787.30247
18. VeIP. www.ensembl.org
19. Werling, D., Piercy, J., Coffey, T.J. Expression of TOLL-like receptors (TLR) by bovine antigen-presenting cells-potential role in pathogen discrimination. *Vet. Immunol. Immunopathol*. 2006. V.112. P. 2–11. DOI: 10.1016/j.vetimm.2006.03.007