

Imputation-based whole-genome analysis identifies candidate variants affecting body temperature under the cold stress in Siberian cattle

Igoshin A.V.^{1*}, Yudin N.S.^{1,2}, Larkin D.M.³

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

² *Novosibirsk State University, Novosibirsk, Russia*

³ *Royal Veterinary College, University of London, London, UK*

* igoshin@bionet.nsc.ru

Key words: genome-wide association study, imputation, cold adaptation

Motivation and Aim: Adaptability of cattle to environmental stresses has gained much interest over the last decades. Cold stress resilience, among others, represents a promising phenotype amenable to genetic improvement by means of selection or gene editing techniques. Previously, using SNP array genotyping, we identified *GRIA4* as a candidate gene associated with body temperature maintenance under cold stress in Siberian populations of Hereford and Kazakh Whiteheaded breeds [1]. Now, we performed imputation of the SNP array data to whole genome sequence level to identify candidate variants underlying this phenotype.

Methods and Algorithms: For our target SNP array set comprising 183 individuals, genomic coordinates of 119,841 SNPs were lifted from Btau8 to Btau9 assembly using the UCSC LiftOver tool. Then, we used autosomal variants of 141 Hereford animals from the 1000 Bull Genomes Project as an imputation reference panel. Low quality (GQ < 15) genotypes were removed and only biallelic SNPs with at least 80 % non-missing genotypes were retained. The target and reference sets were checked for an allele mismatch using *check_strands.py* script and then pre-phased with Eagle v.2.4.1. The imputation was performed using Minimac4 software. After removing variants with low frequency and quality (maf < 0.025, rsq < 0.3) from the imputed dataset, we conducted genome-wide association analysis using EMMAX software. The area under the curve of body temperature over the period of five coldest days in February 2017 was taken as a phenotype. Sex and breed were taken as covariates.

Results: As a result, 769 SNPs, of which 597 found in *GRIA4* gene, reached a significance level (q-value < 0.05). The highest level of association ($p = 1.03\text{E-}07$, $q = 0.0045$) was observed for the variant BTA15:1,862,815 in the *MSANTD4-GRIA4* intergenic region. As the candidate functional variation, we consider three consecutive SNPs (BTA15:1,905,535-1,905,537; $p = 1.10\text{E-}06$, $q = 0.017$) located in the poly-T tract near the splice acceptor site of the *GRIA4* exon 8. In our imputed set, these SNPs are in a strong ($r^2 = 1$) linkage disequilibrium and thus represent a trinucleotide substitution. Another promising (though reaching a suggestive level of association, $q < 0.1$) candidate is BTA15:1,866,804 ($p = 1.57\text{E-}05$, $q = 0.074$) whose alternative allele frequency in the Russian Herefords is 4-fold higher than that of this breed's worldwide population (0.265 vs. 0.066) according to the 1000 Bull Genomes Project. Also, this variant can be found at moderate (0.136–0.238) frequencies in cold-tolerant Turano-Mongolian breeds (Yakut, Buryat, Yanbian) and at moderate-to-high frequencies (0.142–0.5) in several

heat-adapted indicine breeds (e.g. Red Sindhi, NDama, Ankole, Senepol, Nelore) implying its possible role in thermoregulation efficiency.

Conclusion: After an experimental genotype confirmation in carrier animals and validation of association in an independent sample set, the *GRIA4* variants highlighted in this study may be used in breeding practice.

Acknowledgements: This study was supported by the grant from the Russian Science Foundation (project No. 19-76-20026).

References

1. Igoshin A.V., Yurchenko A.A., Belonogova N.M., Petrovsky D.V., Aitnazarov R.B., Soloshenko V.A., Yudin N.S., Larkin D.M. Genome-wide association study and scan for signatures of selection point to candidate genes for body temperature maintenance under the cold stress in Siberian cattle populations. *BMC Genetics*. 2019;20(Suppl 1):26.