

This author's accepted manuscript may be used for non-commercial purposes in accordance with [Wiley Terms and Conditions for Self-Archiving](#).

The full details of the published version of the article are as follows:

TITLE: Genome-wide association study for body weight in cattle populations from Siberia

AUTHORS: Igoshin, A V; Yudin, N S; Belonogova, N M; Larkin, D M

JOURNAL: Animal Genetics

PUBLISHER: Wiley

PUBLICATION DATE: 7 April 2019 (online)

DOI: <https://doi.org/10.1111/age.12786>

1 **Genome-wide Association Study for Body Weight in Cattle Populations from Siberia**

2 A.V. Igoshin¹, N.S. Yudin^{1,2}, N.M. Belonogova¹, D.M. Larkin^{1,3*}

3 ¹*Institute of Cytology and Genetics SB RAS, Novosibirsk 630090, Russia*

4 ²*Novosibirsk State University, Novosibirsk 630090, Russia*

5 ³*Royal Veterinary College, University of London, London NW1 0TU, UK*

6 *Corresponding author

7 *e-mail: dmlarkin@gmail.com*

8 Tel: +447577057645

9 Fax: +442074681196

10 **Summary**

11 Body weight is a complex trait in cattle associated with commonly used commercial breeding
12 measurements related to growth. Although many quantitative trait loci (QTLs) for body weight have been
13 identified in cattle so far, searching for genetic determinants in different breeds or environments is
14 promising. Therefore, we carried out a genome-wide association study (GWAS) in two cattle populations
15 from the Russian Federation (Siberian region) using the GGP HD150K array containing 139,376 single
16 nucleotide polymorphism markers (SNPs). Association tests for 107,550 SNPs left after filtering revealed
17 five statistically significant SNPs on BTA5, considering a false discovery rate (FDR) < 0.05. The
18 chromosomal region containing these five SNPs contains the *CCND2* gene which was previously
19 associated with average daily weight gain and body mass index in US beef cattle populations and in
20 humans, respectively. Our study is the first GWAS for body weight in the beef cattle populations from the
21 Russian Federation. The results provided here suggest that despite the existence of breed- and species-
22 specific QTLs, the genetic architecture of body weight could be evolutionarily conserved in mammals.

23 **Key words:** body weight, GWAS, cattle

24

25 Beef cattle are one of the world's major meat sources, including the Russian Federation. With a
26 recent development of national breeding programs that utilize local cattle populations, the genetic studies
27 of economically important traits (e.g., body weight and growth) in the context of local environments and
28 breeds are required. For over 120 QTLs reported for the body weight in cattle, only a small fraction is
29 attributed to multiple breeds while most QTLs are breed-specific. Each cattle chromosome harbors one or
30 more QTLs for growth-related traits (Hu *et al.* 2013). Among the QTLs with major effects are two
31 mapped to/or near the *PLAG1* and *LEP* candidate genes. It was proposed that the *PLAG1* transcription
32 factor plays a key role in growth-related traits (review by (Takasuga 2016) because the product of this
33 gene regulates expression of insulin-like growth factors (Juma *et al.* 2016). The *LEP* gene encodes for
34 leptin which influences food intake and metabolism in mammals therefore also affecting the body weight
35 (Friedman & Halaas 1998). Multiple association studies have revealed the effect of sequence variants in
36 *LEP* on the body weight and related traits in multiple cattle breeds (Collis *et al.* 2012; Woronuk *et al.*
37 2012; Hernandez *et al.* 2016; Ardicli *et al.* 2017).

38 The evolutionarily older (ancestral) QTLs segregate across many breeds, while younger QTLs are

39 more likely to be breed-specific (Kemper *et al.* 2015). Consistent with this Saatchi and colleagues (2014)
40 demonstrated that most of the QTLs associated with dry matter intake, metabolic weight, growth, and
41 feed efficiency in four beef cattle populations had no overlap (Saatchi *et al.* 2014). Given that there are
42 more than a thousand cattle breeds existing worldwide (Food and Agriculture Organization of the United
43 Nations, <http://www.fao.org/3/a-mm278e.pdf>), it is important to search for QTLs in previously
44 unexplored breeds or populations dwelling in various environments. The fact that body weight in cattle is
45 highly heritable (e.g. (Ryu & Lee 2014) makes identifying QTLs in different populations promising for
46 breed improvement through gene-assisted or genomic selection. Herein, we performed a genome-wide
47 association study for body weight in two highly related cattle breeds (Supporting Information: Figure S1)
48 bred in the Russian Federation (Siberian region).

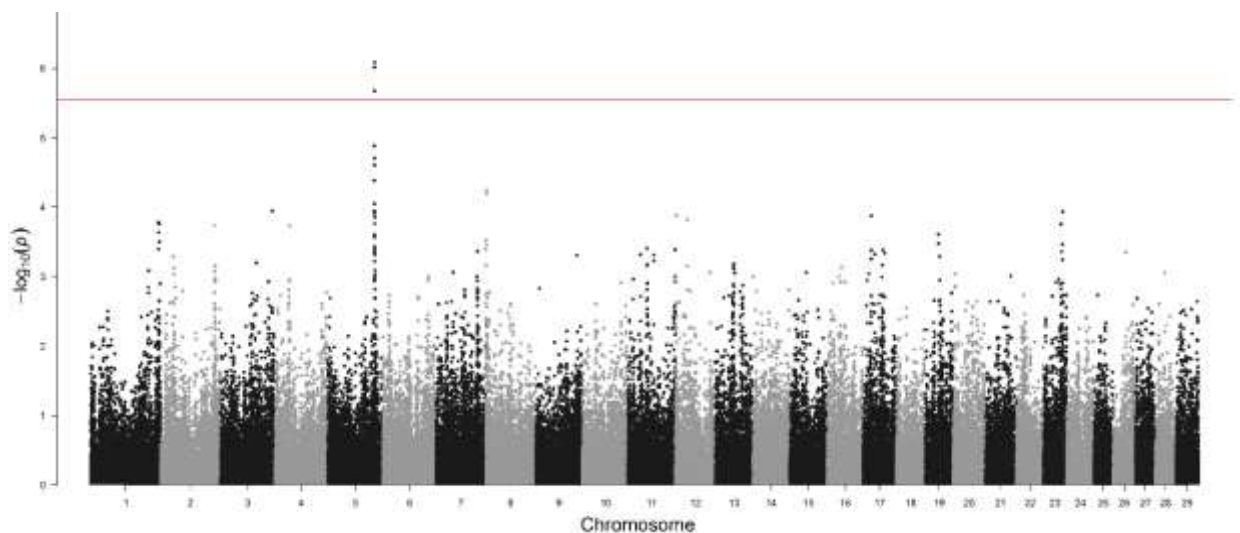
49 The Kazakh Whiteheaded breed was established between 1930 and 1950 in the Kazakh Republic
50 of the USSR by crossing the Turano-Mongolian Kazakh and the Kalmyk breeds with the Hereford
51 (Dmitriev *et al.* 1989). The close relationship between the Kazakh Whiteheaded and Hereford cattle has
52 been recently confirmed in our study of the genetic history of Russian cattle breeds (Yurchenko *et al.*
53 2018; also see Figure S1). A total of 174 animals from the Siberian population of Herefords (92 dams and
54 58 bulls, bred locally since 1960s) and the Kazakh Whiteheaded (4 dams and 20 bulls) breeds aged from
55 6 months to 13 years were used in this study under the ICG SB RAS Ethical Committee approval (№
56 37/28.11.2017). The data on body weight were obtained from the breeding record cards (Supporting
57 Information: Table S1). Blood samples (5 to 10 ml) from the tail vein were collected from each animal
58 into EDTA vacutainer tubes (Weihai Hongyu Medical Devices Co., Ltd, China). DNA were isolated
59 using the cell lysis followed by phenol-chloroform extraction (Sambrook *et al.* 2006). Genotyping was
60 performed on the GGP HD150K array containing 139,376 single nucleotide polymorphism (SNP)
61 markers. Genotypes were called using the GenomeStudio 2 software (Illumina, San Diego, USA). Quality
62 control for the genotypic data obtained was carried out using PLINK software (version 1.9). We removed
63 all SNPs from sex chromosomes (5,171), as well as those not assigned to any chromosomal position
64 (14,364). Additional 11,135 SNPs have been discarded due to a low minor allele frequencies in our
65 dataset (MAF<0.05), 1,065 had low call rate (called in <90% of samples) and 91 SNPs significantly
66 deviated from Hardy-Weinberg equilibrium (p-value <10⁻⁶). As a result, 107,550 SNPs were used in the
67 subsequent analysis.

68 GWAS for body weight was conducted using a variance component model implemented in
69 EMMAX (Kang *et al.* 2010). This method accounts for both population stratification and relatedness
70 between individuals being studied by using Balding-Nichols estimate of the kinship matrix (BN-matrix)
71 to model random effects to the phenotype. Statistical analysis was carried out using the following
72 regression model:

$$73 \quad \mathbf{y} = \boldsymbol{\mu} + \mathbf{X}\mathbf{b} + \mathbf{g} + \mathbf{e},$$

74 where \mathbf{y} represents the $\mathbf{n} \times \mathbf{1}$ vector of phenotype ($\mathbf{n}=174$); $\boldsymbol{\mu}$ is the $\mathbf{n} \times \mathbf{1}$ population mean vector;
75 \mathbf{X} is the $\mathbf{n} \times \mathbf{4}$ design matrix for the fixed effects and \mathbf{b} is the $\mathbf{4} \times \mathbf{1}$ vector of regression coefficients; \mathbf{g} is a
76 $\mathbf{n} \times \mathbf{1}$ random vector of polygenic effects whose covariance matrix is proportional to kinship matrix; \mathbf{e} is
77 the $\mathbf{n} \times \mathbf{1}$ random vector of residuals. Square root of age (at the time of weighing), sex (coded as 1 or 2),
78 breed (coded as 1 or 2) and genotype of the marker were taken as fixed effects. To control for the multiple
79 testing error rate the Benjamini-Hochberg FDR method was applied. Q-value equal to 0.05 was
80 considered as threshold for significance.

81 Genomic inflation factor λ (based on median chi-squared) in our analysis was 1.005 suggesting
82 no inflation of statistics due to population structure, therefore no corrections were applied. We revealed
83 genotype-to-phenotype associations at significant level (q-value <0.05) for five SNPs (Table 1, Figure 1).



84
85 **Figure 1.** Manhattan plot of the genome-wide association study for body weight in Kazakh Whiteheaded
86 and Hereford cattle from Siberia. Red line represents the significance threshold (q-value=0.05).

87 **Table 1.** Significant single nucleotide polymorphisms identified in the genome-wide association study for
88 body weight in the Kazakh Whiteheaded and Hereford cattle from Siberia.

SNP	BTA	Position	MAF	Alleles (minor/major)	Closest genes (± 250 Kbp)	p-value	q-value	Effect size (95% confidence interval)
BovineHD0500030398	5	105804923	0.164	G/A	KCNA6; NDUFA9; AKAP3; RAD51AP1; C5H12orf4	9.65E-07	0.026	62.32 (37.38 - 87.26)
BovineHD0500030494	5	106294449	0.172	G/A	C5H12orf4; RAD51AP1; C5H12orf4; FGF6; TIGAR; CCND2	8.22E-07	0.026	61.3 (36.93 - 85.67)
ARS-BFGL-NGS-106674	5	106296860	0.172	G/A	RAD51AP1; C5H12orf4; FGF6; TIGAR; CCND2	8.22E-07	0.026	61.3 (36.93 - 85.67)
BovineHD0500030501	5	106329896	0.17	A/G	RAD51AP1; C5H12orf4; FGF6; TIGAR; CCND2	2.13E-06	0.046	58.94 (34.57 - 83.31)
BovineHD0500030503	5	106337540	0.164	G/A	C5H12orf4; FGF6; TIGAR; CCND2	9.65E-07	0.026	62.32 (37.39 - 87.26)

89

90 All five statistically significant SNPs were found on bovine chromosome (BTA) 5. Four of these
91 SNPs (HD0500030503, HD0500030501, HD0500030494 and ARS-BFGL-NGS-106674) were located
92 close to each other within a 43 Kbp interval (BTA5: 106,294,449 - 106,337,540), while the
93 HD0500030398 was located ~0.5 Mbp upstream from this cluster (BTA5:105,804,923). The cluster of
94 four SNPs is found 18 Kbp upstream from *CCND2*. The region containing this gene has previously been
95 associated with average daily weight gain in Hereford cattle (Seabury *et al.* 2017). *CCND2* was also
96 associated with both height and body mass index in humans (Steinthorsdottir *et al.* 2014) making this
97 gene the functional candidate for the body weight traits in two Siberian cattle populations. The
98 BovineHD0500030398 SNP was found 40 Kbp upstream from the *KCNA6* gene. The region containing
99 this gene has been associated with metabolic body weight in Holstein dairy cows (Hardie *et al.* 2017).
100 However, further work is required reveal any possible associations of this QTL in Siberian populations
101 with commonly used commercial measurements related to growth e.g., the feed conversion ratio or birth
102 weight.

103 Our study is the first GWAS for body weight in beef cattle populations from Siberia. Candidate
104 QTL regions associated with the body weight in the Hereford and Kazakh Whiteheaded reported herein
105 likely to be a QTL present in the ancestral Hereford population and admixed into the two Siberian
106 populations. The fact that it contains candidate genes that have been previously associated with the body

107 weight and/or growth-related traits in both cattle and in humans, suggests the contribution of this region
108 to weight/growth-related traits in mammals.

109

110 **Acknowledgements**

111 The work was supported by the Russian Science Foundation grant (RSF, 16–14–00090).

112 **References**

- 113 Ardicli S., Dincel D., Samli H. & Balci F. (2017) Effects of polymorphisms at LEP, CAST, CAPN1, GHR,
114 FABP4 and DGAT1 genes on fattening performance and carcass traits in Simmental bulls.
115 *Archives Animal Breeding* **60**, 61-70.
- 116 Collis E., Fortes M.R.S., Zhang Y., Tier B., Schutt K., Barendse W. & Hawken R. (2012) Genetic variants
117 affecting meat and milk production traits appear to have effects on reproduction traits in cattle.
118 *Animal Genetics* **43**, 442-6.
- 119 Dmitriev N.G., Ernst L.K. & Food and Agriculture Organization of the United Nations. (1989) *Animal genetic*
120 *resources of the USSR*. Food and Agriculture Organization of the United Nations, Rome.
- 121 Friedman J.M. & Halaas J.L. (1998) Leptin and the regulation of body weight in mammals. *Nature* **395**, 763-
122 70.
- 123 Hardie L.C., VandeHaar M.J., Tempelman R.J., Weigel K.A., Armentano L.E., Wiggans G.R., Veerkamp
124 R.F., de Haas Y., Coffey M.P., Connor E.E., Hanigan M.D., Staples C., Wang Z., Dekkers J.C.M.
125 & Spurlock D.M. (2017) The genetic and biological basis of feed efficiency in mid-lactation Holstein
126 dairy cows. *Journal of Dairy Science* **100**, 9061-75.
- 127 Hernandez N., Martinez-Gonzalez J.C., Parra-Bracamonte G.M., Sifuentes-Rincon A.M., Lopez-Villalobos
128 N., Morris S.T., Briones-Encinia F., Ortega-Rivas E., Pacheco-Contreras V.I. & Meza-Garcia
129 L.A. (2016) Association of polymorphisms in growth hormone and leptin candidate genes with
130 live weight traits of Brahman cattle. *Genetics and Molecular Research* **15**.
- 131 Hu Z.L., Park C.A., Wu X.L. & Reecy J.M. (2013) Animal QTLdb: an improved database tool for livestock
132 animal QTL/association data dissemination in the post-genome era. *Nucleic Acids Research* **41**, D871-
133 D9.
- 134 Juma A.R., Damdimopoulou P.E., Grommen S.V.H., Van de Ven W.J.M. & De Groef B. (2016) Emerging
135 role of PLAG1 as a regulator of growth and reproduction. *Journal of Endocrinology* **228**, R45-R56.
- 136 Kang H.M., Sul J.H., Service S.K., Zaitlen N.A., Kong S.Y., Freimer N.B., Sabatti C. & Eskin E. (2010)
137 Variance component model to account for sample structure in genome-wide association studies. *Nature*
138 *genetics* **42**, 348-U110.
- 139 Kemper K.E., Hayes B.J., Daetwyler H.D. & Goddard M.E. (2015) How old are quantitative trait loci and how
140 widely do they segregate? *Journal of Animal Breeding and Genetics* **132**, 121-34.

- 141 Ryu J. & Lee C. (2014) Genomic Heritability of Bovine Growth Using a Mixed Model. *Asian-Australasian*
142 *Journal of Animal Sciences* **27**, 1521-5.
- 143 Saatchi M., Beever J.E., Decker J.E., Faulkner D.B., Freetly H.C., Hansen S.L., Yampara-Iquise H., Johnson
144 K.A., Kachman S.D., Kerley M.S., Kim J., Loy D.D., Marques E., Neiberghs H.L., Pollak E.J.,
145 Schnabel R.D., Seabury C.M., Shike D.W., Snelling W.M., Spangler M.L., Weaber R.L., Garrick D.J.
146 & Taylor J.F. (2014) QTLs associated with dry matter intake, metabolic mid-test weight, growth and
147 feed efficiency have little overlap across 4 beef cattle studies. *BMC Genomics* **15**.
- 148 Sambrook J., Russell D.W. & Sambrook J. (2006) *The condensed protocols from Molecular cloning : a*
149 *laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- 150 Seabury C.M., Oldeschulte D.L., Saatchi M., Beever J.E., Decker J.E., Halley Y.A., Bhattarai E.K., Molaei
151 M., Freetly H.C., Hansen S.L., Yampara-Iquise H., Johnson K.A., Kerley M.S., Kim J., Loy D.D.,
152 Marques E., Neiberghs H.L., Schnabel R.D., Shike D.W., Spangler M.L., Weaber R.L., Garrick D.J. &
153 Taylor J.F. (2017) Genome-wide association study for feed efficiency and growth traits in US beef
154 cattle. *BMC Genomics* **18**.
- 155 Steinthorsdottir V., Thorleifsson G., Sulem P., Helgason H., Grarup N., Sigurdsson A., Helgadottir H.T.,
156 Johannsdottir H., Magnusson O.T., Gudjonsson S.A., Justesen J.M., Harder M.N., Jorgensen M.E.,
157 Christensen C., Brandslund I., Sandbaek A., Lauritzen T., Vestergaard H., Linneberg A., Jorgensen T.,
158 Hansen T., Daneshpour M.S., Fallah M.S., Hreidarsson A.B., Sigurdsson G., Azizi F., Benediktsson
159 R., Masson G., Helgason A., Kong A., Gudbjartsson D.F., Pedersen O., Thorsteinsdottir U. &
160 Stefansson K. (2014) Identification of low-frequency and rare sequence variants associated with
161 elevated or reduced risk of type 2 diabetes. *Nat Genet* **46**, 294-8.
- 162 Takasuga A. (2016) PLAG1 and NCAPG-LCORL in livestock. *Animal Science Journal* **87**, 159-67.
- 163 Woronuk G.N., Marquess F.L., James S.T., Palmer J., Berryere T., Deobald H., Howie S. & Kononoff P.J.
164 (2012) Association of leptin genotypes with beef cattle characteristics. *Animal Genetics* **43**, 608-10.
- 165 Yurchenko A., Yudin N., Aitnazarov R., Plyusnina A., Brukhin V., Soloshenko V., Lhasaranov B., Popov
166 R., Paronyan I.A., Plemyashov K.V. & Larkin D.M. (2018) Genome-wide genotyping uncovers
167 genetic profiles and history of the Russian cattle breeds. *Heredity* **120**, 125-37.
- 168
- 169 **Supporting Information**
- 170 Supplemental Table S1. Information on animal phenotypes and breeds.

171 Supplemental Figure S1. Results of Principal Component Analysis depicting the relationship between the
172 Hereford, Kazakh Whiteheaded and other Russian native and worldwide cattle breeds.

173

174 **Access to Data**

175 Data available from the Dryad Digital Repository: [10.5061/dryad.4p3f8c8](https://doi.org/10.5061/dryad.4p3f8c8).