Genome-wide association study for 13 udder traits from linear type classification in cattle

C. Flury[*], C. Boschung[*], M. Denzler[*], B. Bapst[†], U. Schnyder[†], B. Gredler[†] and H. Signer-Hasler[*]
[*]Bern University of Applied Sciences, School of Agricultural, Forest and Food Sciences, Zollikofen, Switzerland [†]Qualitas AG, Zug, Switzerland

ABSTRACT:
Udder conformation traits are known to correlate with the incidence of clinical mastitis and the length of productive life. The results of a genome-wide association study based on imputed high-density genotypes of 1,637 -Brown Swiss sires and de-regressed breeding values for 13 udder traits are presented here. For seven traits significant signals could be observed in five regions on BTA3, BTA5, BTA6, BTA17, and BTA25. For fore udder length and teats diameter significant SNPs were found in a known region around 90 Mb on BTA6. For the trait rear udder height significant SNPs are positioned in the coding region of the *SNX29* gene. Several significant SNPs around 62 Mb on BTA17 are associated with the traits rear udder width, front teat placement and rear teat placement. The function of potential candidate genes and the influence of substructure will be addressed as next steps.

Keywords:
GWAS
udder conformation traits

Introduction

The availability of dense SNP-data led to an increase in results from genome-wide association studies (GWAS). In cattle several independent studies were published on the detection of quantitative trait loci (QTL) for stature (e.g. Pausch et al. 2011, Pryce et al. 2011). Cole et al. (2011) and Wu et al. (2013) reported various significant SNPs by applying GWAS on conformation traits in the US- and Chinese-Holstein population, respectively. The analysis of data from the international genetic evaluation of the Brown Swiss breed revealed a region of strong association on chromosome 25 for the two conformation traits stature and body depth (Guo et al. 2012). Additionally, this region showed significant signals for milk production traits. In the same study two SNPs on chromosome 6 (90.30 and 90.50 Mb) reached genome-wide significance level for milking speed (Guo et al. 2012).
In the meantime accurate imputation results of 50K-genotypes to HD-genotypes were reported for different cattle breeds (e.g. Pausch et al. 2013, Gredler et al. 2013). The availability of SNP-data at much higher density is expected to allow for the detection of a high fraction of the genetic variation of complex traits (Pausch et al. 2013). Therefore imputed HD-genotypes for 1,637 sires from the Swiss Brown Swiss population were used here for GWAS of de-regressed proofs of 13 udder traits from linear type classification.

Materials and Methods

Animals and phenotypes

Estimated breeding values (EBVs) for 12 udder traits from linear type classification and the corresponding breakdown score were provided by Qualitas AG. EBVs and de-regressed EBVs (dEBVs) were available for totally 1,637 Brown Swiss sires with daughter proofs. The investigated traits are listed in table 1, each with the corresponding number of sires with dEBV, average dEBV, standard deviation of dEBV, mean reliability (based on EBVs), the difference in the mean dEBVs between the two subpopulations (Figure 1) and the corresponding heritability.

Genotypes and quality control

The genotypes of around 9,000 individuals used for imputation in this study are an extended dataset of that described by Kramer et al. (2014). The chromosomal position of the SNPs was defined according UMD3.1 of the bovine genome sequence (Zimin et al. 2009). For separate quality control in both datasets (i.e. Illumina BovineHD Beadchip genotypes and Illumina BovineSNP50k Beadchip genotypes) animals with more than 10% missing genotypes were omitted. Additionally, SNPs with call rates below 90% and SNPs with a minor allele frequency below 5% were not considered for further analysis. A subset of 38,936 common SNPs was genotyped with the different genotyping arrays. The missing genotypes of totally 628,417 SNPs were imputed using FImpute (Sargolzaei et al. 2011). Pausch et al. (2013) identified 5,039 SNPs with poor imputation quality due to probable misplacements. Out of these 3,713 SNPs were common with our data set and excluded from further analysis. After quality control and imputation procedure the final data set contained 1,637 individuals and 624,704 SNPs.

Genome-wide association study

To derive potential substructure in the sample the genome-wide IBD matrix was transformed to a distance matrix and the first two principal components were plotted after multi-dimensional-scaling procedure (Figure 1) as implemented in the R-package GenABEL (Aulchenko et al. 2007).
Genome-wide association study was performed by using the mixed-model approach considering genome-wide IBD as implemented in the function mmscore of the above mentioned R-package. Before correction for stratification the genomic inflation factor for the analyzed traits ranged between 1.08 and 1.20. Stratification control was empirically assessed via QQ-plots. SNPs were considered to be genome-wide significantly associated if their p-values were below the 5% Bonferroni-corrected threshold for 624,704 independent tests ($p_{BONF} < 8.0038 \times 10^{-8}$).

Results and Discussion

Genome-wide association study

For seven of the totally 13 investigated traits significant SNPs could be identified. These traits and their five most significant SNPs - with the corresponding chromosome number, physical position in bp, and p-value - are listed in Table 2.

For fore udder length (FUL) a region with 11 highly significant SNPs (p-values<p_{BONF}) was identified between 88.44 and 88.99 Mb on chromosome 6. On the same chromosome 27 SNPs with p-values below p_{BONF} were found between 89.19 Mb and 91.25 Mb for the trait teats diameter (TED). In the same region (i.e. 90.30 to 90.5 Mb) two significant SNPs were observed by Guo et al. (2013) for milking speed based on dEBVs and 50k-IlluminaBeadchip genotypes of 4,411 Brown Swiss bulls. Sodeland et al. (2011a,b) revealed a QTL for protein yield around 88 Mb and a QTL for clinical mastitis around 90 Mb on chromosome 6 for the breed Norwegian Red.

For the trait fore udder attachment (FUA) 11 SNPs built a significant signal on chromosome 3 in the region of 117.28 Mb and 119.81 Mb. Thereof two SNPs are positioned in the coding region of *LRRFIP1* gene and two in the coding region of the *MLPH* gene. However, with *PRLH*, *RAB17* and *COL6A3* other genes are found in proximity of the significant SNPs in this region.

All of the 13 significant SNPs for the trait rear udder height (RUH) are positioned in the coding region of the *SNX29* gene (chromosome 25; 10,743,976-11,273,969 bp).

In total 19 SNPs on chromosome 17 are significantly associated with rear teat placement (RTP). The most significant SNP BovineHD1700017896 is also significant for the trait RUW. In addition five of the significant SNPs for RTP are also significant for the trait FTP. For these three traits significant SNPs are spanning the region of 62.47 - 62.79 Mb whereof the SNP BovineHD1700017839 is positioned in the coding region of the *TBX5* gene. *TBX5* is located between the *TBX3* gene (62.35 – 62.36 Mb) and the *RBM19* gene (62.89 – 63.01 Mb). Pausch et al. (2012) reported significant SNPs in this region for the trait "udder clearness" in the Fleckvieh cattle breed. Beside the above mentioned significant SNPs on chromosome 17 three SNPs on chromosome 5 are significantly associated with FTP.

In comparison with GWAS for body conformation traits in the Holstein breed (Cole et al. 2011, Wu et al. 2013) none of the significant SNPs reported here were in common with the results for udder traits by these authors. However, based on a granddaughter-design (16 halfsib-families with 872 sons) and 264 microsatellite markers Hiendleder et al. (2003) found significant F-values for udder traits in the region of 88-89 cM on chromosome 6. Beside *TBX5* (e.g. Pausch et al. 2012) no more candidate genes related to udder conformation traits and the five significant regions presented in our study could be found in literature. Therefore further exploration of the relationships between significant SNPs and potential candidate genes is necessary.

Influence of population structure?

As described by Hagger (2005) the Swiss Brown cattle population is characterized by two sub-populations: The majority of the Swiss Brown cattle population (BV) is influenced by genetic material from the US-Brown Swiss population. Where the second subpopulation – the Original Brown cattle breed (OB) - is kept closed and used as dual purpose breed. This sub-structure becomes obvious in Figure 1, where the 208 sires with an OB-gene proportion greater 25% (assessed based on pedigree information) build the separate, small cluster.

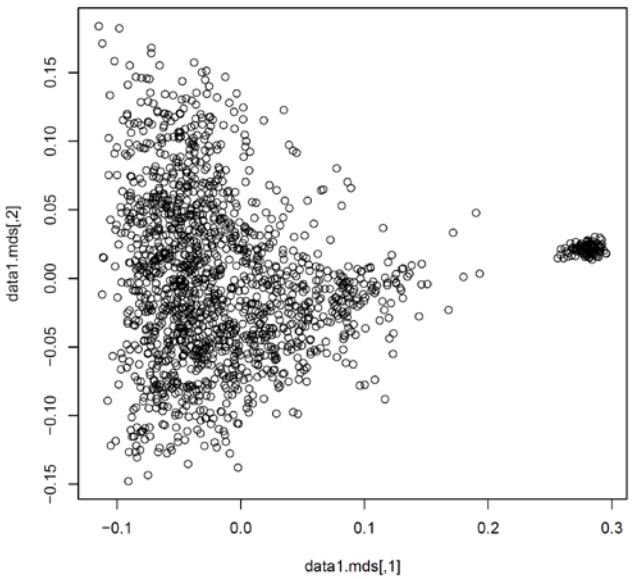

Figure 1: MDS-plot for the first two principal components of the genome-wide IBD estimated from the HD-genotypes of 1,637 sires indicating population sub-structuring.

It is assumed that the 254 bulls identified as a separate cluster by Guo et al. (2012) were also influenced by or originated from the OB-breed. These authors mention that the results for separate GWAS of the two sub-classes were similar to those of a joint analysis of the whole population. However, as also the phenotypes (dEBVs) between the two subgroups vary almost in the range of one standard deviation of the total sample for some traits (FUL, RUW, TED, RTP, OUS) (Table 1) the GWAS-results presented here need further validation.

Conclusion

The results from GWAS of udder traits from linear type classification by using imputed HD-genotypes from 1,637 Brown Swiss bulls are reported here. For seven of the totally 13 investigated traits significant SNPs could be observed. Significant signals are observed on chromosome 3 (117.31- 118.36 Mb); chromosome 5 (31.50 – 31.52 Mb), chromosome 6 (88.74 – 90.80 Mb), chromosome 17 (63.12 – 65.39), and chromosome 25 (10.89 – 10.93 Mb). The region on chromosome 6 is known from other studies to influence clinical mastitis and milking speed. All significant

SNPs for the trait rear udder attachment height are positioned in the *SNX29* gene. The validity of the results will be checked for the two sub-populations separately. Functions and relations of positional candidate genes with significant SNPs need further exploration. Additionally, GWAS for other traits from linear type classification is planned for the near future. The availability of high-density genotypes allows uncovering the genetic architecture of conformation traits and with this improves the understanding of known correlations with economically important traits on the molecular-genetic level.

Acknowledgements

The authors thank Braunvieh Schweiz (Zug, Switzerland), the genotype pool Germany-Austria, Associazione Nazionale Allevatori Bovini della Razza Bruna (Busolengo, Italy), and Beltsville Agricultural Research Center (Beltsville, MD, USA) for the provision of genotypes for imputing procedure.

Literature Cited

Aulchenko, Y.S., Ripke, S., Isaacs, A. et al. (2007). Bioinformatics, 23 (10) :1294-1296.

Cole, J.B., Wiggans, G.R., Li M., et al. (2011). BMC Genomics, 12 : 408.

Gredler, B., Sargolzaei, M., Bapst, B. et al. (2013). EAAP, Book of abstracts No 19: Session 16, p.229.

Guo, J., Jorjani, H. and Ö. Carlborg. (2012) BMC Genetics, 13: 82.

Hagger, C. (2005). J. Anim. Breed. Genet. 122(6): 405-413.

Hiendleder, S., Thomsen, H., Reinsch, N. et al. (2003). J. Hered. 94(6): 496-506.

Kramer, M., Erbe, M., Seefried, F.R. et al. (2014). J. Dairy Sci. 97:1774-1781.

Pausch, H., Flisikowski, K., Jung, S. et al. (2011). Genetics 187(1):289-297.

Pausch, H., Jung, S., Edel, C. et al. (2012). Anim. Genet. 43(6):689-695.

Pausch, H., Aigner, B., Emmerling, R. et al. (2013). Genet. Sel. Evol. 45: 3.

Pryce, J.E., Hayes, B.J., Bolormaa, S. et al. (2011). Genetics :187 : 981-984.

Sargolzaei, M., Chesnais, J.P. and F.S. Schenkel. (2011). J. Dairy Sci. 94(E-Suppl.1): 421 (abstr. 333).

Sodeland, M., Kent, M.P., Olsen, H.G. et al. (2011a). Anim. Genet. 42(5): 457-465.

Sodeland, M., Grove, H., Kent, M. et al. (2011b). BMC Genetics 12:70.

Wu, X., Fang, M., Liu, L. et al. (2013). BMC Genomics 14: 897.

Zimin, A.V., Delcher, A.L., Florea, L. et al. (2009). Genome Biol. 10(4): R42.

Table 1: Descriptive statistics of the phenotypic information and the 13 investigated traits.

Trait	# sires with dEBV	mean dEBV	STD dEBV	mean relia-bility[1] EBV	Diff Mean[2] dEBVs OB / BV	h^2
Fore udder length (FUL)	1605	-0.148	1.161	0.861	-1.713	0.20
Fore udder attachment (FUA)	1610	-0.184	0.996	0.871	-0.487	0.23
Rear udder width (RUW)	1621	-0.256	1.034	0.862	-0.955	0.19
Rear udder height (RUH)	1625	-0.212	1.062	0.869	-0.805	0.23
Udder support (USU)	1621	-0.102	1.073	0.861	-0.405	0.21
Udder depth (UDE)	1628	-0.209	1.017	0.906	-0.311	0.32
Udder balance (UDB)	1602	-0.035	0.869	0.855	-0.632	0.24
Teats length (FTL)	1629	0.434	1.266	0.915	-0.364	0.37
Teats diameter (TEM)	1614	0.167	1.296	0.910	0.699	0.37
Teats direction (TED)	1611	-0.244	1.387	0.886	-1.137	0.28
Front teat placement (FTP)	1628	-0.271	1.250	0.905	-0.942	0.34
Rear teat placement (RTP)	1618	-0.232	1.344	0.900	-1.272	0.32
Overall udder score (OUS)	1621	-0.412	1.214	0.873	-1.174	0.29

[1] Reliability of EBVs; [2] Difference of mean dEBVs between the subgroup of OB-individuals (n=208) and the subgroup of BV-individuals (n=1,429).

Table 2: Udder traits with signifcant SNPs. The top five associated SNPs are listed and significant p-values < p_{BONF} are written bold.

Trait	Chr	SNP-id	Position (bp)	P-values
Fore udder length (FUL)	6	BovineHD0600024355	88'919'352	**9.14×10^{-9}**
	6	BovineHD0600024357	88'922'396	**9.14×10^{-9}**
	6	BovineHD0600024354	88'913'092	**1.66×10^{-8}**
	6	BovineHD0600024289	88'728'581	**2.38×10^{-8}**
	6	BovineHD0600024297	88'744'593	**2.96×10^{-8}**
Fore udder attachment (FUA)	3	BovineHD0300034340	117'768'104	**1.56×10^{-9}**
	3	BovineHD0300034294	117'625'064	**3.59×10^{-9}**
	3	BTA-69789-nors	117'760'734	**8.40×10^{-9}**
	3	BovineHD0300034233	117'537'703	**1.57×10^{-8}**
	3	BovineHD0300034137	117'305'327	**1.62×10^{-8}**
Rear udder width (RUW)	17	BovineHD1700017896	62'697'699	**7.03×10^{-8}**
	17	BovineHD1700017901	62'707'985	9.68×10^{-7}
	17	BovineHD1700017875	62'640'837	1.05×10^{-6}
	17	BovineHD1700017881	62'656'132	1.07×10^{-6}
	17	BovineHD1700017884	62'668'270	1.14×10^{-6}
Rear udder height (RUH)	25	BovineHD2500002973	10'931'374	**3.17×10^{-9}**
	25	BovineHD2500002949	10'891'363	**8.07×10^{-9}**
	25	BovineHD2500002950	10'892'254	**8.07×10^{-9}**
	25	BovineHD2500002951	10'893'792	**8.07×10^{-9}**
	25	BovineHD2500002954	10'898'471	**8.07×10^{-9}**
Teats diameter (TEM)	6	BovineHD0600024745	90'370'978	**2.74×10^{-12}**
	6	BovineHD0600024780	90'491'374	**1.92×10^{-11}**
	6	BovineHD0600024859	90'797'634	**4.81×10^{-11}**
	6	BovineHD0600024639	89'876'621	**7.20×10^{-11}**
	6	BovineHD0600024718	90'277'007	**8.56×10^{-11}**
Front teat placement (FTP)	17	BovineHD1700017839	62'533'120	**3.21×10^{-8}**
	5	BovineHD0500009167	31'504'622	**4.27×10^{-8}**
	5	BovineHD0500009169	31'520'865	**4.27×10^{-8}**
	17	ARS-BFGL-BAC-36617	62'623'054	**4.96×10^{-8}**
	17	BovineHD1700017837	62'522'492	**6.12×10^{-8}**
Rear teat placement (RTP)	17	BovineHD1700017896	62'697'699	**7.38×10^{-11}**
	17	BovineHD1700017839	62'533'120	**1.23×10^{-10}**
	17	ARS-BFGL-BAC-36617	62'623'054	**2.32×10^{-10}**
	17	BovineHD1700017865	62'613'160	**3.91×10^{-10}**
	17	BovineHD1700017837	62'522'492	**7.85×10^{-10}**