Proceedings, 10th World Congress of Genetics Applied to Livestock Production

Targeted Association Mapping in Merinoland Crossbred Lambs

K.F. Schiller[1], P. Stratz[1], S. Preuss[1] and J. Bennewitz[1].
[1]University of Hohenheim, Germany

ABSTRACT: This study reports the results of targeted association analysis in multiple F1 Merinoland crossbred lambs. A number of 384 SNPs in chromosomal regions with reported QTL for growth, carcass and meat quality were genotyped at 1493 crossbred lambs. These lambs were produced from Merinoland ewes and rams from five different meat-type breeds (Charollais, Ile de France, German Blackheaded Mutton, Suffolk, and Texel). Single SNP association analysis was conducted across the crosses or nested within the crosses. The traits daily gain, carcass yield, drip loss, haunch circumference, and fat layer were considered. Modeling SNP effect across the crosses identified weak associations with the same effect sign across the crosses. The nested analysis revealed significant associations with different effects signs in the crosses, which were not detected in the model where SNP effect was fitted across the crosses. Positional and functional candidate genes were identified and discussed.
Keywords: Crossbred sheep; Targeted association analysis; Meat-type traits

Introduction

The Merinoland (ML) is a widespread breed of sheep in Germany. ML ewes are crossed frequently with a meat-type sire breed in order to produce high quality lamb meat. In a previous study we investigated which sire breed is most appropriate to produce F1 crossbred lambs with ML. Five sire breeds and in addition the ML were used to produce F1 lambs, which were fattened and slaughtered and comprehensively phenotyped for various growth, carcass and meat quality traits (Henseler et al., 2014a; b).

Identifying genetic markers that are associated with economically relevant traits will be helpful to select rams within and across sire breeds. Targeted association study by selecting a low number of SNPs in chromosomal regions that have been frequently reported to harbor genes affecting the traits of interest is a cost effective alternative to genome wide association studies. Especially in situations where the empirical power of the study design is limited (e. g. due to limited number of individuals) it has its advantages due to a lower multiple testing problem.

The aim of the present study was to apply a targeted association study for five meat-type traits using 1493 ML x sire breed F1 crossbred lambs and 384 selected SNPs.

Materials and Methods

Data. The dataset included 1511 F1 crossbreed and purebred ML-lambs. For production of crossbreed lambs rams of the meat-type breeds Charollais, Ile de France, German Blackheaded Mutton (Deutsches Schwarzköpfiges Fleischschaf), Suffolk, and Texel were crossed with ML ewes. The crosses are listed in Table 1. All lambs were raised, fattened and slaughtered under standardised conditions. Lambs were raised on seven farms till weaning at 17 kg bodyweight (BW). Fattening took place on a single farm in group housing with 200-300 g hay per animal and concentrate *ad libitum*. The lambs had a mean BW at slaughter of 43.14 ± 3.78 kg at an age of 102 to 161 days. During and after slaughter growth, carcass, and meat quality traits were recorded. Details can be found in Henseler et al. (2014a; b). The following traits were considered in this study: daily bodyweight gain during fattening (BWG [g]), carcass yield (CY [%]), haunch circumference (HC [cm]), fat cover (FAT [cm]), and drip loss (DRIP [%]). Lambs were genotyped for 384 SNPs These SNPs were located on chromosome 1, 2, 3, 18 and 21, in order to focus on chromosomes where QTL for these traits have been reported in the literature (Hu et al., 2013).

Table 1. Crosses, cross abbreviation, number of sires and number of F1 lambs

Cross	Abbrev.	n sires	n lambs
Charollais x ML[1]	CH	5	298
Ile de France x ML	IF	5	329
ML x ML	ML	4	225
Blackheaded Mutton x ML	SK	5	221
Suffolk x ML	SU	5	277
Texel x ML	TX	4	143

[1]ML = Merinoland sheep

Statistical analysis. SNP filtering was done using following criteria. A SNP was excluded if it had a minor allele frequency <3%, and a call rate <95%. A number of 313 SNP passed the data filtering. Single marker association mapping was done using two different models. Model one estimated one effect per SNP *k* across all six crosses. The model was

$$y_{ij} = x_{ij}\beta + sire_{ij} + b_k * x_{ijk} + e_{ij}, \qquad (1)$$

where y_{ij} is the trait record of individual *j* of cross *i*, the term $\mathbf{x}_{ij}$ denotes for the *ij*th row vector of a design matrix

linking the phenotypic observation of the individual to some fixed effects stored in β (i. e. the effect of the cross, the sex, and the weight at slaughter). The effect of the SNP *k* was modelled as a regression on the number of copies of the allele with the higher frequency (x = 0, 1, or 2), with b_k being the regression coefficient. Pedigree data were not available. Therefore, the sire effect was included as an uncorrelated random effect to capture some population structure effects. The term e_{ij} is a random residual with heterogeneous variance, i. e. $e_{ij} \sim N(0, \sigma_i^2)$. The null (alternative) hypothesis was that $b_k = 0$ ($b_k \neq 0$). The test statistic was an *F*-test.

In the second model the SNP effects were nested within the crosses, i. e.

$$y_{ij} = x_{ij}\beta + sire_{ij} + b_{ik} * x_{ijk} + e_{ij} . \qquad (2)$$

The terms are as defined for the previous model. The null (alternative) hypothesis was that $b_{ik} = 0$ for every cross *i* ($b_{ik} \neq 0$ for at least one cross *i*). The test statistic was a pooled *F*-test. This model was applied, because the marker density was low even in the targeted regions, and hence, the Linkage Disequillibrium (LD) between an SNP and a causal mutation might be different across the crosses. If this LD holds across the crosses, then this model will be of reduced power, because six regression coefficients have to be estimated instead of one (as in model (1)). In order to control for multiple testing an FDR *q*-value was calculated for each test using the software QVALUE (Storey and Tibshirani, 2003). The association analysis was undertaken using ASReml 3.0 (Gilmour et al, 2009).

Gene annotation and ontology. Significant SNPs were arranged in clusters based on trait association. Candidate genes were searched in the vicinity of significant SNPs. The super-set of cDNA sequences for Ovis Aries (taxid:9940) was obtained from Ensembl (Flicek et al., 2014) known, novel and pseudo gene predictions. cDNA sequences were used as queries against the non-redundant protein database using Blast2GO version 2.7.0. A relaxed statistical significance threshold for reporting matches against database sequences was chosen. The gene matches were used for the gene ontology (GO) term assignment. After gene ID mapping, GO term assignment and annotation augmentation the final annotation file was produced. Results were categorized with respect to the Blast2GO categories Biological Process, Molecular Function and Cellular Component. GO terms were searched at several levels, in order to establish links to considered traits.

Results and Discussion

The number of significant SNPs from both models is shown in Table 2. A low threshold level was chosen because no extensive multiple testing was done and in addition the empirical power of the study is limited. The FDR *q*-values of the significant associations are relatively high (not shown), suggesting a number of false positives. Nearly the same number of significant SNPs was identified by the two models. However, these were not always the same. Model (1) had more power to detect associations with same effect in the crosses. Model (2) detected additional significant associations that showed opposite effect signs in the crosses.

Table 2. Number of significant SNPs, results from both models

	Model (1)		Model (2)	
Trait	p≤0.01	p≤0.001	p≤0.01	p≤0.001
BWG[1]	4	1	3	2
CY[2]	4	1	6	2
HC[3]	8	2	6	2
DRIP[4]	2	0	5	0
FAT[5]	5	0	4	0

[1] BWG = daily bodyweight gain during fattening
[2] CY = carcass yield
[3] HC = haunch circumference
[4] DRIP = drip loss
[5] FAT= fat layer

Some of highly significant SNPs and their chromosomal position and candidate genes are shown in Table 3. ATF2 showed significant results for the trait CY and BWG. GO terms of the gene's transcripts are connected to terms like muscle organ development, embryo development, regulation of protein metabolic process and therefore were of functional interest. SNP *OAR18_68269251.1* seemed to be of special interest because of possible homolog functions to the human *DLK1* gene, which is known to be involved in cell differentiation of several cell types also in other species (Appelbe et al., 2013).

Conclusion

Targeted association analysis revealed weak significant SNP associations for all traits. Modeling SNP effects nested within crosses revealed additional significant associations that would have been missed if the SNP would have been fitted solely across the crosses. Interesting candidate genes were identified. The study will be continued using additional targeted and untargeted SNPs. This will allow also an SNP-based modeling of the population effects.

Acknowledgements

K. F. Schiller received funding from the H. Wilhelm Schaumann Stiftung, Hamburg, Germany.

Table 3. Significant SNPs per trait, candidate genes, error probability (*p*) effect estimates ($\hat{b}$ and $\hat{b}_i$) (standard error in parenthesis), results from both models

Trait[1] ($\mu \pm \sigma_P$)	SNP bp[2]	candidate gene	Model (1)		Model (2)	
			p	$\hat{b}$	*p*	$\hat{b}_i$
BWG (330.1 ±70.9)	OAR2_142354112.1 133865504	ATF2	<0.001	-5.169(1.520)	<0.001	CH: -5.930 (3.494) IF: -0.583(2.917) ML: 2.912 (4.255) SK: -13.863 (4.687) SU: -8.768 (3.200) TX: -7.692 (5.454)
	OAR18_68269251.1 64385940	DLK1, BEGAIN, oar-mir-136	0.394	-1.680(1.970)	0.001	CH: -12.255(5.873) IF: 4.72137 (5.824) ML: 15.751 (4.879) SK: -5.256 (4.313) SU: -7.370 (3.490) TX: -3.306 (5.987)
CY (53.6 ±4.6)	OAR2_142354112.1 133865504	ATF2	<0.001	-0.331(0.091)	0.010	CH: -0.320 (0.20) IF: -0.464 (0.175) ML: -0.548 (0.272) SK: -0.328 (0.313) SU: -0.092 (0.188) TX: -0.227 (0.320)
	OAR2_205872952.1 194324700	PCGEM1	0.145	0.126(0.086)	<0.001	CH: 0.297 (0.174) IF: 0.341 (0.186) ML: 0.458 (0.244) SK: 0.130 (0.218) SU: 0.057 (0.210) TX: -0.905 (0.255)
	OAR3_197402139.1 183368930	DENND5B, FAM60A, CAPRIN2,	0.010	0.227(0.088)	<0.001	CH: -0.095 (0.192) IF: -0.022 (0.202) ML: 1.043 (0.248) SK: 0.671 (0.239) SU: 0.153 (0.184) TX: 0.087 (0.239)
HC (64.0 ±2.7)	OAR1_140104902.1 129332577	-	<0.001	-0.214(0.065)	0.053	CH: -0.173 (0.143) IF: -0.196 (0.144) ML: -0.288 (0.164) SK: -0.323 (0.168) SU: -0.193 (0.150) TX: -0.164 (0.186)
	OAR1_145988855.1 135244346	-	0.990	-0.008	<0.001	CH: 0.151 (0.135) IF: 0.684 (0.168) ML: -0.271 (0.172) SK: -0.301 (0.151) SU: -0.258 (0.152) TX: 0.014 (0.189)
	OAR2_222903133.1 210644328	ENSOARG 0000001949	<0.001	-0.228(0.064)	0.035	CH: -0.343 (0.164) IF: -0.199 (0.142) ML: -0.322 (0.164) SK: -0.091 (0.158) SU: -0.197 (0.149) TX: -0.187 (0.157)
	OAR3_169440758.1 158312976	-	0.305	-0.067(0.065)	<0.001	CH: -0.502 (0.177) IF: -0.113 (0.145) ML: -0.219 (0.146) SK: 0.062 (0.165) SU: -0.040 (0.139) TX: -0.239 (0.204)

[1] Abbreviations are shown in Table 1 and 2.
[2] Flicek et al., 2014

Literature Cited

Appelbe, O. K., Yevtodiyenko, A., Muniy-Talavera, H. et al. (2013) Mechanisms of Development, 130:143-159.
Flicek, P., Amode, M. R., Barrell, et al. (2014). Ensembl 2014. Nucleic Acids Research, 42:D749-D755.
Gilmour, A., Gogel, B., Cullis, B. et al. (2009) VSN International Ltd., UK., vol. 275.
Henseler, S., Preuss, S., Bennewitz, J. (2014a) Züchtungskunde, 86:95-103.
Henseler, S., Preuss, S., Bennewitz, J. (2014b) Züchtungskunde, 86:104-114.
Hu, Z.-L., Park, C. A., Wu, L. et. al. (2013) Nucleic Acids Research, 41 (D1): D871-D879
Storey, J. D., and Tibshirani, R. (2003) PNAS 100:9440-9445.

Table 3. Significant SNPs per trait, candidate genes, error probability (*p*) effect estimates ($\hat{b}$ and $\hat{b}_i$) (standard error in parenthesis), results from both models

Trait[1] ($\mu \pm \sigma_P$)	SNP bp[2]	candidate gene	Model (1)		Model (2)	
			p	$\hat{b}$	*p*	$\hat{b}_i$
BWG (330.1 ±70.9)	OAR2_142354112.1 133865504	ATF2	<0.001	-5.169(1.520)	<0.001	CH: -5.930 (3.494) IF: -0.583(2.917) ML: 2.912 (4.255) SK: -13.863 (4.687) SU: -8.768 (3.200) TX: -7.692 (5.454)
	OAR18_68269251.1 64385940	DLK1, BEGAIN, oar-mir-136	0.394	-1.680(1.970)	0.001	CH: -12.255(5.873) IF: 4.72137 (5.824) ML: 15.751 (4.879) SK: -5.256 (4.313) SU: -7.370 (3.490) TX: -3.306 (5.987)
CY (53.6 ±4.6)	OAR2_142354112.1 133865504	ATF2	<0.001	-0.331(0.091)	0.010	CH: -0.320 (0.20) IF: -0.464 (0.175) ML: -0.548 (0.272) SK: -0.328 (0.313) SU: -0.092 (0.188) TX: -0.227 (0.320)
	OAR2_205872952.1 194324700	PCGEM1	0.145	0.126(0.086)	<0.001	CH: 0.297 (0.174) IF: 0.341 (0.186) ML: 0.458 (0.244) SK: 0.130 (0.218) SU: 0.057 (0.210) TX: -0.905 (0.255)
	OAR3_197402139.1 183368930	DENND5B, FAM60A, CAPRIN2,	0.010	0.227(0.088)	<0.001	CH: -0.095 (0.192) IF: -0.022 (0.202) ML: 1.043 (0.248) SK: 0.671 (0.239) SU: 0.153 (0.184) TX: 0.087 (0.239)
HC (64.0 ±2.7)	OAR1_140104902.1 129332577	-	<0.001	-0.214(0.065)	0.053	CH: -0.173 (0.143) IF: -0.196 (0.144) ML: -0.288 (0.164) SK: -0.323 (0.168) SU: -0.193 (0.150) TX: -0.164 (0.186)
	OAR1_145988855.1 135244346	-	0.990	-0.008	<0.001	CH: 0.151 (0.135) IF: 0.684 (0.168) ML: -0.271 (0.172) SK: -0.301 (0.151) SU: -0.258 (0.152) TX: 0.014 (0.189)
	OAR2_222903133.1 210644328	ENSOARG 0000001949	<0.001	-0.228(0.064)	0.035	CH: -0.343 (0.164) IF: -0.199 (0.142) ML: -0.322 (0.164) SK: -0.091 (0.158) SU: -0.197 (0.149) TX: -0.187 (0.157)
	OAR3_169440758.1 158312976	-	0.305	-0.067(0.065)	<0.001	CH: -0.502 (0.177) IF: -0.113 (0.145) ML: -0.219 (0.146) SK: 0.062 (0.165) SU: -0.040 (0.139) TX: -0.239 (0.204)

[1] Abbreviations are shown in Table 1 and 2.
[2] Flicek et al., 2014