Linkage disequilibrium pattern and genome-wide association mapping for meat traits in multiple porcine F2-crosses

P. Stratz¹, M. Schmid¹, R. Wellmann¹, S. Preuß¹, I. Blaj², J. Tetens³, Georg Thaller² & J. Bennewitz¹

¹ University of Hohenheim, Institute of Animal Science, Garbenstraße 17, 70599 Stuttgart, Germany
² Christian-Albrechts-University Kiel, Institute of Animal Breeding and Husbandry, Hermann-Rodewald-Straße 6, 24118 Kiel, Germany
³ Georg-August-University Göttingen, Department of Animal Science, Functional Breeding Group, Burckhardtweg 2, 37077 Göttingen, Germany

Patrick.Stratz@uni-hohenheim.de (Corresponding Author)

Summary

In the current study, data from four F2-crosses were pooled. This data set was used to study the linkage disequilibrium (LD) within and across the crosses. Genome-wide association studies for conductivity and dressing out were conducted using BayesD. Three porcine F2-crosses generated from distantly related founder breeds Wild Boar, Piétrain and Meishan and one porcine F2-cross from closely related founder breeds Piétrain and the F1 cross Large White × Landrace were pooled. Around 50k SNPs localized on the current genome assembly from 2,570 F2-individuals were usable for LD studies. The breakdown of the LD was faster in crosses generated from closely related founders, compared to crosses from distantly related founders. By pooling the data the fastest breakdown of LD was observed. Dominance which is of special interest for the traits was included in BayesD. Based on the LD structure, 1Mb windows were built. Window posterior probability of associations ≥ 0.8 were found for conductivity on chromosomes 2, 6, 8, 11, 14, and 15.

Keywords: meat traits, multi-marker association analysis, multiple porcine crosses, windows

Introduction

In pig breeding numerous F2-crosses were established in the past. The individuals were usually genotyped for a relatively small number of microsatellite markers and QTL mapping was done using linkage analysis. With the advent of the porcine 62K SNP chip it is possible to conduct genome wide association analysis (GWAS) in pig breeding.

Schmid et al. (2016) used stochastic simulation studies to analyse the prospects of GWAS in porcine F2-crosses genotyped with dense SNP markers. The main results of Schmid et al. (2016) were that power as well as precision in GWAS was high, when the founder breeds to establish the cross were closely related and sample size is large.

In the current study, data from four F2-crosses were pooled. All individuals were genotyped with the porcine 62K SNP chips to study the linkage disequilibrium (LD), within and across the crosses. Subsequently GWAS for meat traits were conducted using BayesD.

Material and methods
Genotypes and phenotypes

Three porcine F2-crosses were generated from three distantly related founder breeds Meishan (M), Piétrain (P) and Wild Boar (W) (Rückert & Bennewitz, 2010). The experimental design of closely related founder breeds was a cross of the breeds Piétrain and a F1 Large White × Landrace (PxLWL) (Boysen et al., 2010). Approximately 2570 F2-individuals were available.

Porcine 62K SNP chip genotypes were available for all individuals. The location of the markers was retrieved from the current genome assembly. Genotype filtering was done using Illumina GenomeStudio software. Haplotypes were reconstructed and sporadic missing genotypes were imputed using Beagle. Approximately 50K autosomal SNPs retained for further analyses.

Phenotypes were available for conductivity, (conductivity in mS/cm 45 min post-mortem) and dressing out (as ratio of carcass weight and live weight at slaughter).

Population analysis

The population differentiation index ($F_{ST}$) was used to quantify levels of differentiation between the founder breeds. $F_{ST}$-values were estimated for each SNP between the founder breeds Piétrain, Meishan and the F1 Large White x Landrace cross.

Reconstructed haplotypes were used to estimate $r^2$ within and across the crosses. This was done for marker pairs being $\leq$ 5MB apart.

Statistical analyses

Variance components were estimated in GCTA and used as prior information for BayesD (Wellmann & Bennewitz, 2012) which was applied to account for imperfect LD between the marker and the causal mutation. In BayesD it was assumed that the distribution of the effect of SNP $m$ is a mixture of two $t$-distributions that differ by a scaling factor which was set to 0.01. The marker effect is either allocated in the $t$-distribution with the larger variance with prior probability $pLD$=0.2, or in the $t$-distribution with the smaller variance with prior probability 1-$pLD$. $T$-distributions were set up with 3 degrees of freedom for the additive effects for BayesD. The Markov chain was generated by Gibbs sampling. 100,000 Gibbs sampling iterations were performed, 50,000 were discarded as burn-in.

Every 25th sample of the additive and dominance effects was stored to calculate the window posterior probability of association ($WPPA$; Fernando & Garrick, 2013). $WPPAs$ were calculated for 1Mb sliding windows. A detailed description of the $WPPA$ calculation for BayesD can be found in Bennewitz et al. (2017).
Results and Discussion

Population analysis

The mean $F_{ST}$ index over the SNPs was $F_{ST} \approx 0.17$ between Piétrain and Meishan. Between Large White × Landrace and Piétrain (Meishan) the index was $F_{ST} \approx 0.07$ ($F_{ST} \approx 0.25$).

In Figure 1 the extent of LD is exemplarily shown for chromosome 1. For low distances the level of LD is highest and decreases especially for distances >0.5Mb and >1.5Mb. In the WxM cross the LD is highest and decreases slowest. In the MxP cross the decrease of LD is slow. The decrease of LD was fastest in the PxLW and WxP crosses. Compared to single crosses the LD in the pooled crosses is lowest and the decrease is highest.

![Figure 1. Extent of LD on chromosome 1 in single - (left) and in pooled crosses (right) as a function of distance between pairs of SNPs up to 5 Mb.](image)

The lowest differentiation of the breeds which was found between the European-type breeds positively impacts the mapping resolution within that cross. In contrast, the greater differentiation between the European-type - and Asian-type breeds leads to a lower mapping resolution. The highest mapping resolution in GWAS was observed when pooling the data. The LD results of this study are in agreement with Schmid et al. (2016) and support the benefit from pooling data from several F2-crosses.

Statistical analyses

For conductivity and dressing out the proportion of dominance on the genetic variance is 47.4% and 13.8%. In BayesD a $WPPA \geq 0.6$ was found for conductivity and dressing out on chromosomes 1, 2, 4, 6, 8, 11, 14, 15 and 17 and on chromosomes 2, 7 and 10 (Figure 2).
Figure 2. BayesD WPPAs calculated for 1Mb windows for conductivity (left) and dressing out (right). The x-axis denotes the 18 chromosomes, the y-axis the WPPAs.

Dominance seems to be important for meat traits. Although a lot of BayesD WPPA signals were found, further investigations in the current Piétrain reference population are needed.

Conclusion

The LD results of this study support the benefit from pooling data from several F2-crosses. Dominance seems to be important for the traits and was included in BayesD to increase the mapping signals. The results of this study serve to preselect target regions in the genome for subsequent fine-mapping.

Acknowledgements

The study was funded by the Deutsche Forschungsgemeinschaft, DFG.

List of References


