Linkage disequilibrium and signatures of selection on chromosome 1 in a commercial sire and dam line

A.E. Huisman[*], P. Cherel[†], B. van Haandel[*]

Introduction

Linkage disequilibrium (LD), which refers to nonrandom association of alleles at different loci, has received increasing attention in recent years and studying LD has gained unprecedented momentum as the result of the availability of genome sequences and large numbers of identified single-nucleotide polymorphisms (SNP). An association between genetic variation at a locus and a phenotype indicates that either the genetic variation at that locus directly affects the phenotype or the locus is in LD with the causative mutation (Mueller, 2004). With the sequencing of the pig genome, it has become possible to estimate the extent of LD using SNP information. Here we focus on chromosome 1; several QTL of economic importance have been mapped to this particular chromosome (Malek et al. (2001); Milan et al. (2002); de Koning et al. (2003); Evans et al. (2003); Kim et al. (2005); Buske et al. (2006)). Genes with a known effect on growth performance are also located on this chromosome. In this paper we will test the hypothesis that the selection history is reflected in the genome of two commercial lines.

Material and methods

DNA was collected from a commercial sire line (n = 234) and a commercial dam line (n = 288). The sire line was derived from Pietrain pigs, whereas the dam line was derived from Landrace pigs. All animals were genotyped using the porcine Illumina SNP-chip, only SNP that were located on SSC1 and segregating in both populations, were in Hardy-Weinberg equilibrium and had a minor allele frequency greater than 0.05 were used in this analysis, resulting in 3,762 SNP spanning approximately 300Mb. The locations for the SNP were obtained from the Illumina website. Genotypes were checked for Mendelian inheritance based on the pedigree, and the few erroneously typed SNP were set to missing. The phase of the SNP genotypes was determined in blocks of 50 SNP using PedPhase (Li and Jiang, 2003). Subsequently, maternal haplotypes were used to estimate LD. Linkage disequilibrium was measured as the square of the correlation coefficient (r^2) between marker alleles. The r^2 values were binned according to the physical distance separating the SNP. To estimate signatures of selection, allele frequency based on the 1 allele was estimated in both populations. The allele frequencies were then computed for each line as a rolling average across the chromosome, and average frequencies for the sire line were subtracted from average frequencies estimated in the dam line.

---

[*] Hypor, Hendrix Genetics, Boxmeer, The Netherlands

[†] Research and Technology Centre, Hendrix Genetics, St. Jean de Braye, France

Results and discussion

The average minor allele frequency was 0.274 and 0.287 in the dam line and the sire line, respectively. The distribution of minor allele frequencies for the SNP used in the analysis is shown in Figure 1. This distribution of minor allele frequencies deviated from the uniform distribution; however the distributions are similar for both lines.

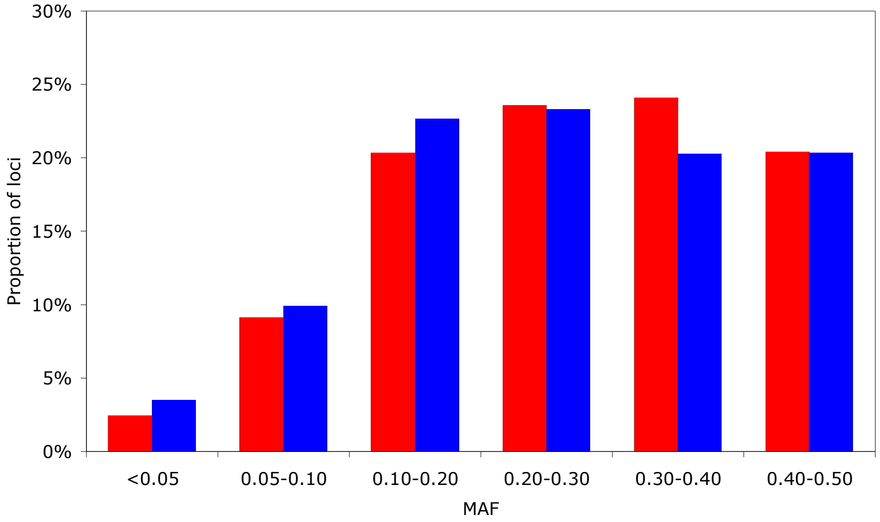

Figure 1: Distribution of minor allele frequencies for the SNP markers (red = dam line; blue = sire line)

Linkage disequilibrium for both lines is in Figure 2, LD is somewhat higher on short ranged distances distance for the sire line compared to the dam line, whereas on the long range LD is quite similar for both lines. This most likely reflects the size and history of the populations. The dam line population size has always been higher compared to the sire line. Previous studies that investigated the extent of LD in pig populations (e.g. Du et al. (2007) estimated LD to be a lot higher.

Differences between allelic frequencies are plotted in Figure 3. The locations which showed most deviation from zero were matched with known QTL locations from the pigQTLdb (http://www.animalgenome.org/cgi-bin/QTLdb/SS/index), only QTL with a significant effect and a confidence interval less than 50 Mb and the interval mid-point not further away than 10 Mb from the SNP location were reported here. Most differences were found for QTL that had an effect on production traits. Also a difference in allelic frequency was found for a SNP in the region of a QTL for ear erectedness (Wei et al. 2007). Although this is probably not due to recent selection, it reflects the differences between the two lines in phenotypic appearance.

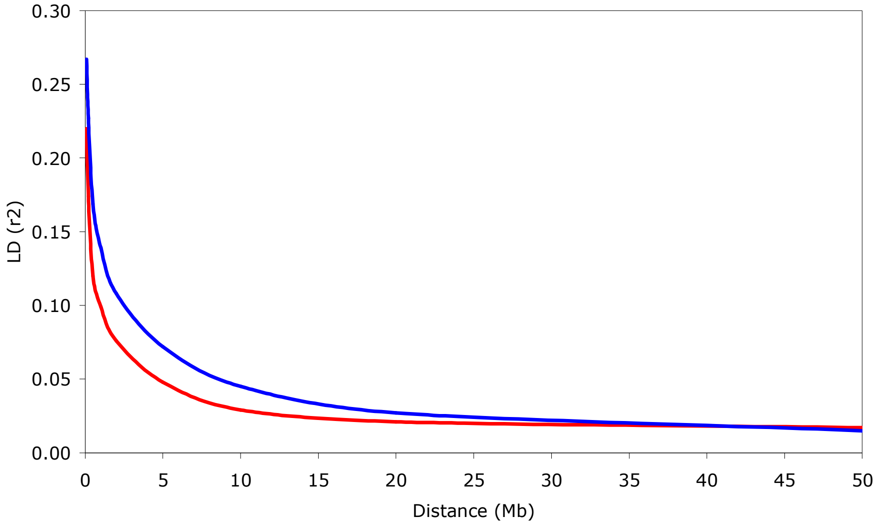

Figure 2: Plot of estimated linkage disequilibrium in sire (blue) and dam line (red)

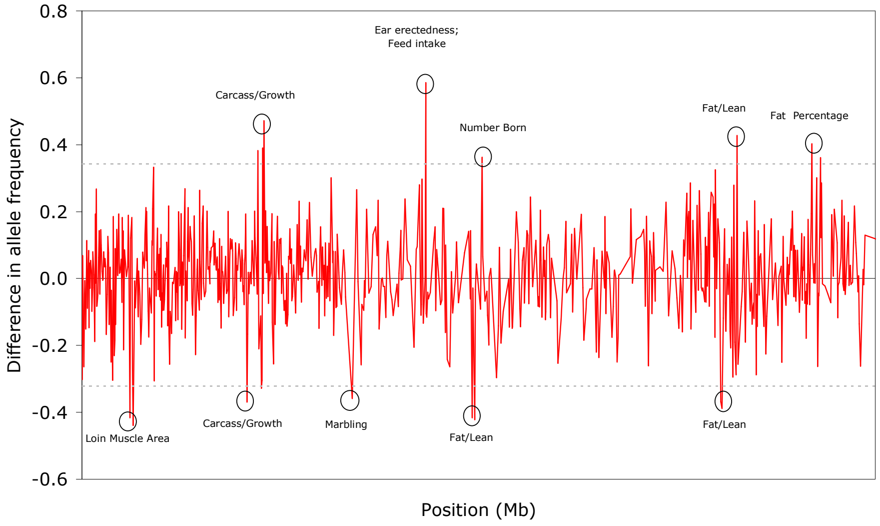

Figure 3: Difference between sire and dam line rolling average allele frequencies (positive when higher in dam line)

Conclusion

Although no major differences in minor allele frequencies were found in these two commercial lines, there was a difference in close range LD. Linkage disequilibrium appears to be somewhat lower in these two lines compared to earlier studies in pigs, it is in between LD found in cattle and chicken (e.g. McKay et al. (2007); Abasht (2009)). Signatures of selection were detected for some production traits (Carcass/Growth related traits; feed intake; loin muscle area and fat related traits, marbling), for number of piglets born and for ear erectedness.

References

Abasht, B., Sandford, E., Arango, J., Settar, P., Fulton, J.E., O'Sullivan, N.P., Hassen, A. Habier, D., Fernando, F.L., Dekkers, J.C.M., and Lamont, S.J. (2009) *BMC Genomics* 10(Suppl2):S2.

Buske, B., Sternstein, I., Reissmann, M., and Brockmann, G. (2006). *J Anim. Breed. Genet.* 123:343-348.

Du, F., Clutter, A.C. and Lohuis, M.M. (2007). *Int. J. Biol. Sci.* 3:166-178.

de Koning D. J., Pong-Wong, R., Varona, L., Evans, G.J., Giuffra, E., Sanchez, A., Plastow G., Noguera, J.L., Andersson, L., and Haley, C. S. (2003). *J. Anim. Sc.* 81:2155-2163.

Evans, G.J., Giuffra, E., Sanchez, A., Kerje, S., Davalos, G., Vidal, O., Illán, S., Noguera, J.L., Varona, L., Velander, I., Southwood, O.I., de Koning, D.J., Haley, C.S., Plastow, G.S., and Andersson, L. (2003). *Genetics* 164:621-627.

Kim, J.J., Zhao, H., Thomsen, H., Rothschild, M.F., and Dekkers, J.C. (2005). *Genet. Res.* 85:235-248.

Li, J., and Jiang, T. (2003). *J. Bioinf. Comp. Biol.* 1:41-69

Malek, M., Dekkers, J.C., Lee, H.K., Baas, T.J., Prusa, K., Huff-Lonergan, E., and Rothschild, M.F. (2001). *Mamm. Genome* 12:637-645.

Milan, D., Bidanel, J.P., Iannuccelli, N., Riquet, J., Amigues, Y., Gruand, J., Le Roy, P., Renard, C., and Chevalet, C. (2002). *Genet. Sel. Evol.*34:705-728.

Mueller, J. C. (2004). *Briefings in Bioinformatics* 5:355-364.

Wei, W.H., de Koning, D.J., Penman, J.C., Finlayson, H.A., Archibald, A.L., and Haley, C.S. (2007). *Anim. Genet.* 38:222-226.

McKay, S.D., Schnabel, R.D., Murdoch, B.M., Matukumalli, L.M., Aerts, J., Coppieters, W., Crews, D., Dias Neto, E., Gill, C.A., Gao, C., Mannen, H., Stothard, P., Wang, Z., Van Tassell, C.P., Williams, J.L., Taylor, J.F., and Moore, S.S. (2007). *BMC Genetics* 8:74.