Genome-wide linkage disequilibrium linkage analysis (LDLA) of body fat traits in an F2 porcine model for human obesity

S. D. Pant[*]**, P. K. Mortensen**[*]**, S. Cirera**[*]**, L. J. A. Kogelman**[*]**, M. J. Jacobsen**[*]**, C. S. Bruun**[*]**,
C. B. Jørgensen**[*]**, T. H. E. Meuwissen**[†]**, H. N. Kadarmideen***[‡]** and M. Fredholm***[‡]**.**
([‡] *Equal Contribution & Correspondence*)
[*]University of Copenhagen, Denmark; [†]University of Life Sciences, As, Norway

ABSTRACT: Purebred Duroc and Yorkshire sows were crossed with Göttingen minipig boars to obtain two separate F2 intercross resource populations (n=287 and 279 respectively). Several obesity, metabolic and slaughter measurements were recorded from birth to slaughter (220 ± 45 days). In addition, body composition was determined at about two months of age (64 ± 11 days) via dual-energy x-ray absorptiometry (DXA) scanning. All pigs were genotyped using Illumina Porcine 60k SNP Beadchip and a combined LDLA approach was used to perform genome-wide linkage and association analysis for body fat traits. Subsequently bioinformatic analysis was performed to identify genes in close proximity of chromosomal positions where statistically significant QTLs were identified. Several important genes previously linked to obesity (e.g. *BBS4, CHRNA2, DLK1)*, along with other novel genes were identified, that together provide novel insights that may further the current understanding of the molecular mechanisms underlying human obesity.
Keywords: LDLA; Pig model; Genomics; QTL mapping

Introduction

Obesity, a condition represented by excessive accumulation of body fat, incurs massive economic costs and predisposes individuals to a number of other diseases including diabetes, cardiovascular disorders and osteoarthritis (Whitlock et al., 2009; Cawley and Meyerhoefer, 2012). Obesity is estimated to increase per capita annual medical expenses by as much as 2,741 US dollars (Cawley and Meyerhoefer, 2012), and its prevalence is rapidly increasing worldwide. The etiology of obesity is highly complex and influenced by numerous factors including individual genetics and the environment. Past studies (Maes et al., 1997) have demonstrated genetic factors to determine as much as 60 – 70% of phenotypic variation, though genetic determinants underlying only 10% of the total genetic variance have been identified thus far (Speliotes et al., 2010). Genetic heterogeneity, confounding between genetics and environmental factors, and imprecise, costly and difficult measurement systems associated with obesity and obesity related phenotypes, are some of the factors that are likely to contribute to this discrepancy between overall genetic contribution to obesity and identified genetic determinants.

For a complex trait like obesity, animal models can aid and accelerate the identification of underlying genetic determinants. Mouse models have been widely used primarily due to their evolutionary proximity to humans, their ability to recapitulate human disease pathophysiology, and the relatively low costs involved in housing, handling and breeding them. However, findings from murine models of obesity have often failed to translate to humans largely due to pathophysiological differences (Arner, 2005). Given these differences, alternative animal models for human obesity are needed where research findings have a greater probability of being translatable to humans. Pig models are of interest in this regard as they have a sequenced genome and are genetically closer to humans especially in the context of energy metabolism and obesity (Heinritz et al., 2013; Nielsen et al., 2014). Pigs, unlike mice, also exhibit almost all of the pathophysiological features related to obesity and metabolic syndrome in a relatively short time span (Spurlock and Gabler, 2008).

Given the potential benefits of using pigs to model human obesity, two separate, comprehensively phenotyped and genotyped, porcine F2 intercross resource populations were established (See Kogelman et al., 2012 for further details). QTL fine mapping was performed using combined linkage disequilibrium linkage analysis (LDLA) to identify novel genetic determinants. A brief description of the resource population, statistical methods and preliminary results are presented herein.

Materials and Methods

Experimental Design and Genotyping. Two separate F2 intercross resource populations were created by crossing purebred Duroc and Yorkshire sows (obtained from DanBred breeding herd) with Göttingen minipig boars (obtained from Ellegaard A/S). Both Duroc and Yorkshire are production breeds that have undergone extensive selection for leanness and growth traits, while Göttingen minipigs are mainly used for research purposes and are bred primarily for their small size and ease of handling. Unlike the production pigs, Göttingen minipigs are also susceptible to diet induced obesity and share many metabolic impairments associated with human obesity (Johansen et al., 2001). All pigs were genotyped using Illumina Porcine 60k SNP Beadchip.

Collection of Phenotypes. Extensive phenotypic collection was performed from birth to slaughter (220 ± 45 days) and included obesity and obesity related, metabolic, and slaughter phenotypes. In addition, body composition was also determined after weaning using dual-energy x-ray absorptiometry (DXA) scanning at about two months of age (64 ± 11 days). Further details of pedigree and phenotyping are available in Kogelman et al. (2013). Preliminary genome-wide analysis was performed on two obesity traits i.e. total fat and total fat percentage that were obtained via DXA scanning.

Statistical Genetics and Bioinformatic Analyses. Genotype data was subjected to preliminary quality control separately within the Duroc and Yorkshire crosses by excluding all SNPs that had a minor allele frequency (MAF) < 0.05, Hardy Weinberg equilibrium test p-value < 0.001, and a genotype call rate < 0.95. Subsequently, identity by descent (IBD) probabilities were estimated chromosome-wise for each marker bracket using a linkage disequilibrium (LD) multilocus iterative peeling (LDMIP) algorithm described in Meuwissen and Goddard (2010). Genome-wide association analysis was performed using the following model:

$$Y_{ijkl} = \mu + Gender_i + \beta_j Age + IBD_k + Animal_l + e_{ijkl}$$

where μ is the intercept, $Gender_i$ is the fixed effect of the sex i, Age represents the age of the pig in days at the time of measurement, and β_j is the associated fixed regression coefficient. Random effects included were the *IBD* probability associated with a chromosomal segment k (Fitted QTL), (Meuwissen and Goddard, 2010), the additive genetic effect of the *Animal l*, and *e*. the error term. A log-likelihood ratio test-statistic (LRT) was subsequently computed as

$$LogLikRatio = LogLik_{QTL} - LogLik_{null}$$

Where $LogLik_{QTL}$ represents the restricted maximum log-likelihood (REML) of the model as described above (full model), and $LogLik_{null}$ represents the REML log-likelihood of the same model excluding the IBD probabilities (null model). Under the null-hypothesis of no QTL effect in the tested marker brackers, *2 x LogLikRatio* was assumed to follow a *chi-squared* distribution with one degree of freedom. In order to account for multiple testing, a modified version of the Bonferroni test was used where the genome-wide significance threshold was obtained as: $\frac{p}{\sqrt{nsnp}}$ where p is the nominal significance threshold of 0.05 and *nsnp is* total number of SNPs used in the analysis (Mantel, 1980).

REML Likelihoods of the full and null models were estimated in Asreml (Gilmour et al. 2009), and subsequently LRT and generation of Manhattan plots and bioinformatic analysis were performed in R statistical environment (R Core Team, 2013).

Results and Discussion

Phenotyping and Genotyping. A broad range of obesity, obesity related, slaughter and metabolic phenotypes were recorded at different time points that included, but is not limited to the phenotypes described in Table 1. Genotyping data was available on 287 pigs within the Duroc x Minipig cross; and 279 pigs within the Yorkshire x Minipig Cross. After initial quality control, a total of 31,540 SNPs within the Duroc x Minipig cross; and 31,730 SNPs within the Yorkshire x Minipig cross remained for statistical analysis.

Table 1. Phenotypes collected in porcine F2 resource population.

Type	Phenotypes
Obesity	birth weight; body weight at 13±16 days, 34±38 days, 64±11 days, 115±27 days, and 220±45 days; body length; daily weight gain; abdominal circumference; body mass index; body adiposity index
DXA scan	Whole body total fat; whole body lean tissue; percentage of fat tissue to total tissues in whole body; bone mineral density
Slaughter	Backfat; omental fat; intestinal fat; retroperitoneal fat
Metabolic	Fasting glucose at 7 months (220±45 days); total, free and esterified cholesterol in plasma; triglycerides in plasma; low and high-density lipoproteins in plasma

Statistical Genetics Analyses. Since single marker Genome-wide association studies (GWAS) only leverage population-wide LD and linkage information was available given the 3 generation intercross pedigree, QTL fine mapping using a LDLA approach was implemented based on a previously described LDMIP algorithm. This approach is more powerful than traditional GWAS as it leverages LD that exists across families as well as the LD that exists within families. Within family LD is especially useful to map QTLs in genomic regions with low population-wide LD. A total of 404 and 10 maker brackets were associated with total fat; and another 18 and 7 marker brackets were associated with total fat percentage in the Duroc and Yorkshire intercrosses respectively, after application of genome-wide correction for multiple comparisons (Figures 1-4).
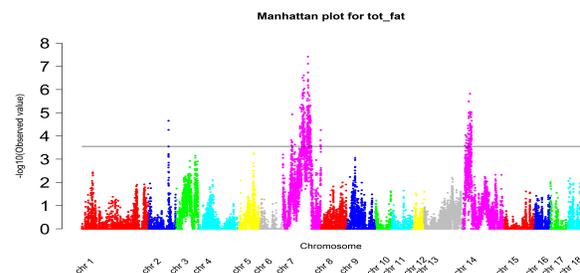

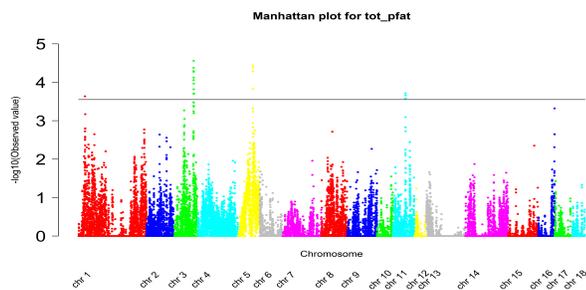**Figure 1. Manhattan Plot for Total Fat in Duroc X Minipig Intercross**

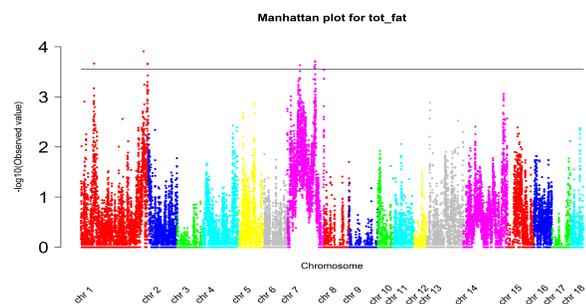**Figure 2: Manhattan Plot for Total Fat Percentage in Duroc x Minipig Intercross**

Figure 3: Manhattan Plot for Total Fat in Yorkshire x Minipig Intercross
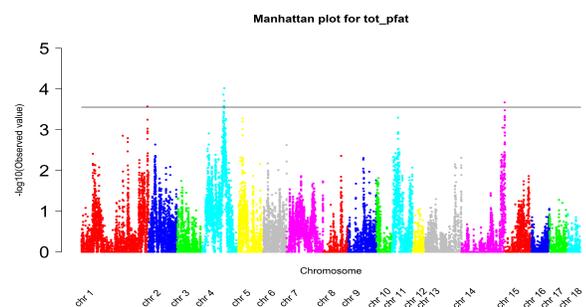

Figure 4: Manhattan Plot for Total Fat Percentage in Yorkshire x Minipig Intercross

Identification of Genes. Genes located within 50 kb of the QTLs in the middle of each marker bracket detected by LDLA were retrieved by querying the NCBI gene database. A non-inclusive list of important genes in vicinity of identified QTLs for each trait is provided in Table 2. More than 150 genes were identified in vicinity of the QTLs identified for total fat in the Duroc intercross that is in line with the high number of QTLs identified for this trait. Several of these genes (Table 2) have previously been linked to obesity which indicates that findings from the porcine model may be of high translational value to humans. Our downstream post-LDLA bioinformatics analyses that includes gene-set enrichment analyses and genetic interaction analyses further provide novel insights into the biology underlying obesity in the porcine resource population that are translatable to humans (results not shown). Furthermore, since LD structure should differ in the purebred founder generations, comparison of QTL regions for the same phenotype between the two crosses is

expected to further fine map QTL regions aiding the identification of genes that are potentially causal.

Table 2. Genes within 50kbs of SNPs associated with total fat in Duroc and Yorkshire F2 intercrosses.

Intercross	Phenotype	Genes
Duroc x Minipig	Total Fat	*BBS4, CHRNA2, DLK1, EPHX2, FUT8, MAX, NFKBIA, RAD51B RETN, SCARB1, SIPA1L1, SLCO3A1, SPTB, TMEM132D, TMEM229B*
	Total Fat Percentage	*MTRF1L, FBXO5, GDDR, GKN3, ARHGAP25, APLF, PNO1, WDR92, C1D, PPP3R1, MEIS1*
Yorkshire x Minipig	Total Fat	*LPAR1, MUSK, KIAA0368, DPF3, FOS, FLVCR2*
	Total Fat Percentage	*PLAG1, CHCHD7*

Conclusion

The three generational F2 structure of the resource population enabled the application of combined LDLA analysis on a genome-wide scale that is much more powerful than traditional GWAS approaches. A number of genes previously linked to obesity were found in close proximity to QTLs associated with total fat and total fat percentage; and several novel genes were also identified. These results indicate that the described porcine model could prove to be a relevant model for human obesity.

Acknowledgment

References

Arner, P. (2005). *Diabetologia* 48: 2203-2205.

Cawley, J., and Meyerhoefer, C. (2012). *J. Health Econ.* 31: 219-230.

Heinritz, S. N., R. Mosenthin, and E. Weiss. (2013). *Nutr. Res. Rev.* 26: 191-209.

Johansen, T., Hansen, H. S., Richelsen, B. et al. (2001). *Comparative Med.* 51: 150-155.

Kogelman, L. J., Kadarmideen, H. N., Mark, T., et al. (2013). *Front. Genet.* 4: 29.

Maes, H. H., Neale, M. C., and Eaves., L. J. (1997). *Behav. Genet.* 27: 325-351.

Mantel, N. (1980). *Biometrics* 36: 381-399.

Meuwissen, T., and Goddard, M. (2010). *Genetics* 185: 1441-1449.

Nielsen, K. L., Hartvigsen, M. L., Hedemann, M. S. et al. (2014). *Am. J. Clin. Nutr.* Ahead of Print: doi:10.3945/ ajcn.113.074724.

R Core Team. (2013). *R Foundation for Statistical Computing*, Vienna, Austria.

Speliotes, E. K., Willer, C. J., Berndt, S. I., et al. (2010). *Nat. Genet.* 42: 937-948.

Spurlock, M. E., and Gabler, N. K. (2008). *J. Nutr.* 138: 397-402.

Whitlock, G., Lewington, S, Sherliker, P. et al. (2009). *Lancet* 373: 1083-1096.