

Selective genotyping in commercial populations

Naomi Duijvesteijn

Boar taint is an off-flavour in pig meat caused primarily by high levels of skatol and androstenone in some entire males. In many countries, male piglets are castrated early in life to prevent boar taint. However, keeping entire males in the future would be desirable for animal welfare and because of the higher efficiency of entire males (higher feed conversion rate, less fat deposition). Many studies have been conducted to investigate the underlying genetic components affecting androstenone and skatol. In the SABRE project WP9, one of the interests is to identify new candidate genes for skatol and androstenone by comparative genome-wide scans. Recent developments have made it possible for Illumina to provide a 60k SNP chip for pigs.

IPG (Institute for Pig Genetics) has been collecting androstenone and skatol samples from 2005 till now and so far collected 2880 samples, of which most samples are from a synthetic Duroc line. In total there are 1921 samples from the Duroc line that are also analysed for skatol and androstenone by the same company. From these animals, 1000 need to be selected to genotype with the 60k SNP chip. The aim of this 8 week placement at the Roslin Institute is to simulate phenotypes and genotypes for different selection methods from an existing pedigree from IPG, to be able to make a decision about which animals should be selected for genotyping in order to obtain the highest power for detecting relevant SNPs. To achieve this, WP9 and WP1 are collaborating to exchange knowledge between IPG and the Roslin Institute. Phenotypes and genotypes were simulated using software programme MORGAN, which uses the actual pedigree provided by IPG. Ten markers and 1 QTL were simulated on 1 chromosome and also 10 markers, but no QTL were simulated on another chromosome for determining the false-positive rate. After this simulation, a different selection methods were applied to select only 1000 animals from the 1920. Four selection methods were compared: random selection (**rand**), selecting large families (>35 half sibs; **fam**), 500 high and 500 low based on androstenone phenotypes (**hilo**), and high and low phenotypes within families (**hilo-fam**). ANOVA was used to analyse each marker against the phenotype and F-test values for each simulated SNP were saved from each replicate. The whole analysis was repeated 1000 times for each selection method.

The 95th percentile of the F-test values of the simulated chromosome without a QTL were computed and used as threshold for the analysis of the chromosome with a QTL. The highest average F-test value was for the hilo selection method and the lowest for hilo-fam. Random selection and selecting large families resulted in a significantly lower power compared with hilo, hilo-fam and no selection (i.e. typing all 1920 individuals; $p < 0.01$). Hilo_fam shows a trend to get a higher power of analysis than just hilo across all families, but was not significant ($p > 0.05$). Genotyping all animals would result in similar power as selective genotyping using hilo or hilo-fam.