RESEARCH REPORT
ANIMAL SCIENCE

Title: Identification of markers associated with sow lifetime productivity for whole genomic selection - **NPB project #09-032**

Investigator: Daniel Ciobanu

Institution: University of Nebraska, Lincoln.

Date Submitted: 06.30.2010

Industry Summary:
Sows are more productive today than ever before. However, concurrent with increased prolificacy, high sow death losses and replacement rates are serious economic and welfare issues facing producers. Reproductive failure is the most frequent reason for culling sows. Length of productive life is moderately heritable and has high variance; thus, substantial genetic variation is expected to exist in most populations. Genetic improvement in the swine industry occurs from selection in nucleus herds and is then transmitted through the breeding pyramid in the multiplication process. Thus, it is critical to identify selection methods that can be applied in nucleus herds that will improve length of productive life in commercial herds. This trait is quite complex and believed to be affected by many genes with relatively small to moderate effects. As a consequence, the response to traditional selection methods is low; generation intervals will be long and accuracy of identifying genetically superior young animals will be reduced. Therefore, this trait is one for which application of genomic tools will be especially helpful. Using genome wide association studies we identified several major and minor loci associated with developmental and sow reproductive traits as well as with sow life-time reproductive and productive traits. As expected, we identified common loci that influence the variation of different traits. For example, same chromosomal regions potentially influence variation of both age at puberty and the lifetime number of live piglets produced by a sow. We are in process of assembling a panel of DNA markers associated with sow reproductive and productive longevity and identify procedures to apply the information in whole genome selection. The markers effects estimated in our population will be validated in other commercial populations and discussions about potential collaborations were initiated with major breeding organizations and commercial producers. We expect that our research will provide a molecular tool that will reduce culling rates, sow death losses, and enhance the productive life of sows.
Contact information:
Daniel C. Ciobanu, PhD
Assistant Professor, Molecular Genetics.
A218k Animal Science Bldg.
Department of Animal Science
University of Nebraska-Lincoln
Lincoln, NE 68583-0908
Phone: (402) 472 8066
Fax: (402) 472 6362
Email: dciobanu2@unl.edu

Keywords: sow, reproduction, longevity, DNA, markers.

Scientific Abstract:

We performed genome wide association studies in a population resource well characterized to identify loci that influence traits such as growth, fatness, age at puberty, breeding performance, and sow lifetime productivity traits until completion of parity 4. Genotyping of 639 gilts of two maternal crossbreds that reached puberty was performed using the PorcineSNP60 BeadChip (Illumina) that contains assays for a total of 62,183 SNPs. Quality scores were assigned for each genotype obtained. A genotype quality score of 0.15 was used as a cutoff threshold for removing low quality genotypes. There were 57.8% of the samples with a call rate of at least 0.995. There were 58,982 SNPs with a call rate of at least 0.900. In this group of reliable marker assays there were 9,627 SNPs with a Minor Allelic Frequency (MAF)<0.05, 6,450 SNPs with a MAF<0.01 and finally 4,147 SNPs completely fixed for one of the alleles.

The level of DNA degradation is influencing the fraction of SNPs successfully genotyped and the genotyping rate. The samples with no or low level of DNA degradation had a higher number of the SNPs genotyped (>99.2%) compared to samples with advanced level of DNA degradation (87%). The number of discrepancies between called genotypes of the samples with absent and limited DNA degradation was extremely small (<0.008%) compared to the samples with increased level of DNA degradation (0.1%).

A Bayesian analysis was conducted to estimate proportion of variance for each developmental, reproductive and lifetime trait accounted for by the SNPs from the PorcineSNP60 BeadChip. From all the reproductive traits, variation in the age at puberty was associated with the largest SNP effects. Clusters of SNPs associated with large effects on puberty onset were located on SSC 1 (21, 29, 88, 127 and 269 Mb regions), SSC2 (56 and 80 Mb), SSC4 (16 Mb), SSC10 (3Mb) and SSC14 (27 Mb). Lower marker effects were identified for litter size from parity 1 to parity 4 and for lifetime productive traits. Variation of sow lifetime productivity traits is potentially influenced by genes located on most swine chromosomes. Clusters of SNPs clearly associated with the largest effects on lifetime number of live-born piglets are located on SSC2, SSC8, SSC13, SSC16 and SSC17. Clear clusters of SNPs associated with lifetime number of weaned pigs are located on SSC1, SSC2, SSC3, SSC8, SSC9, SSC11, SSC12, SSC16 and SSC17. Common major loci that affect both lifetime number of live-born and weaned piglets are located on SSC1, SSC2, SSC16 and SSC17.

Combined association analysis of age at puberty and sow lifetime production traits revealed common regions/genes located on SSC1 (21, 29, 88, 127 Mb) that influence phenotypic variation of age at puberty and lifetime number of live-born and weaned piglets produced by sows. In addition, one of the clusters with the richest number of SNPs associated with the total number of live-born pigs and located on the distal end of SSC17 is influencing also litter size on parity 3.

Main clusters of SNPs associated with gilt weaning weights are located on proximal SSC8 and distal SSC11. Major loci influencing growth from weaning to the initiation of the treatment are located on SSC1, SSC2, SSC3, SSC4, SSC16 and SSC18. Growth during the developmental phase appeared to be influenced by major loci identified on most swine chromosomes. There are loci that potentially influence variation of growth during both developmental phases.

We are in process of assembling a panel of DNA markers associated with sow reproductive and productive longevity and identify procedures to apply the information in whole genome selection. The effects of the major markers will be validated in other commercial populations. We expect that our research will provide a panel of molecular markers that can be successfully used for reducing culling rates, sow death losses, and enhance the sow productive life.

Introduction:

Sow fertility, death losses, and poor health that lead to early culling and high replacement rates are major economic and welfare issues for the swine industry. Our primary objective is to identify SNP markers and combinations of markers associated with sow longevity by conducting a genome-wide characterization of two maternal lines that differ in rate of lean growth, litter size and lifetime production. Our central hypotheses are: 1) there are many genes with relatively small to moderate effects on sow longevity, and 2) moderate heritability and high phenotypic variance indicate

that considerable genetic variation exists and sow reproductive longevity can be improved by whole genome selection.

Objectives:

Objective 1. Identify SNP markers and combinations of markers associated with sow longevity by conducting a genome-wide characterization of two maternal lines that differ in rate of lean growth, litter size, and lifetime production.

Objective 2. Determine the degree to which markers associated with components of sow longevity (BW, BF, and LMA during gilt development, AP, litter size and weight, sow BW and BF loss during lactation, etc.) explain variation in lifetime productivity.

Objective 3. Identify procedures to apply markers association results in WGS programs to enhance sow reproductive longevity.

Materials & Methods:

Population: Using previous Nebraska Pork Producers Association funding through NPB, we collected developmental and longevity data in 631 gilts of two genetic lines known to differ in lean growth rate and reproductive rates. Dams of the gilts were either Large White-Landrace cross females or Nebraska Line 45 females. The Large White-Landrace dams were from the two-breed rotation breeding system used to produce replacement females for the UNL nutrition research program. The Line 45 dams were from Generations 23 and 24 of the Nebraska line (Line 45) that was selected for increased litter size only (Generations 1-18) or increased litter size and within litter selection for increased growth and decreased backfat (BF)(Generations 19-24). Both Large White–Landrace and L45 females were inseminated with semen from boars of an unrelated industry maternal line to produce litters from which project gilts were selected. The semen was collected and packaged from individual boars with known identities and used to inseminate Large White-Landrace and Line 45 females to produce families of half-sib litters that contained gilts of both genetic lines of dam. Project gilts were produced in four replicates and selected randomly from these litters when pigs were approximately 50 d of age. When possible, at least two gilts were selected per litter so that gilts of each litter could be assigned to each of two gilt developmental regimens. Gilts by Large White-Landrace dams are designated LW/LR and those by L45 dams are designated L45X. A total of 661 project gilts from 211 litters by 32 sires started the experiment at approximately 56 d of age.

Phenotypes: Recording of feed intake and backfat (BW) of gilts began at 53 d of age and recording of BF and longissimus muscle area (LMA) began at 123 d of age. Estrus detection began at 140 d of age to determine age at puberty (AP). Traits were recorded bi-weekly until gilts were approximately 230 d of age when they were moved to the breeding barn. Only gilts that could be mated at 2nd or later post-pubertal estrus were designated as breeders. Thus, some gilts with late age at puberty were culled. The length of the breeding period for gilts (24 to 26 days within replicates) permitted all gilts an opportunity to express estrus and be mated. Normal production and feeding practices were used during gestation and lactation. The breeding period for sows ranged from 24 to 32 days and continued until 10 days after the last sow in the replication was weaned. Estrus, breeding, conception, and reasons for death or culling were carefully recorded during the breeding periods for all parities. Females were given only one opportunity to breed during the breeding period. Females that returned to estrus during the breeding period were not re-mated. All females that were not mated and those diagnosed open or that did not farrow a litter were culled. The number and weight of pigs at birth and weaning and sow BW and BF pre-farrowing and at weaning through parity 4 were recorded. Three replicates are completed; replicate 4 gilts have completed two parities and will complete 4 parities during the period of this project.

DNA isolation: DNA was extracted from all 639 individuals that reached puberty. DNA was isolated from tail and ear tissue using the DNeasy blood and tissue kit (Qiagen). The isolated DNA was scored for quality based on the level of DNA degradation that was evaluated by electrophoresis.

DNA genotyping: Genotyping was performed using the PorcineSNP60 BeadChip (Illumina) that contains assays for a total of 62,183 SNPs. The genotyping was contracted out to Gene Seek as described in the proposal. There were 639 samples submitted for genotyping. Genotyping data was analyzed for quality across samples, SNP markers and physical location of the SNPs. Allelic, genotype frequency and deviation from Hardy-Weinberg equilibrium was estimated for all SNPs. We selected for genotyping additional SNPs that are not present in the Porcine 60KSNP BeadArray.

Haplotype analysis: We evaluated Recursive Long Range Phasing (RLRP) (Kingorn et al., 2009) software for haplotype analysis.

GWAS: GenSel software (Fernando and Garrick, 2009) was used for conducting a Bayesian analysis to estimate the proportion for reproductive and developmental traits accounted by the Porcine 60KSNP BeadArray markers.

Results:

Objective 1

- GenCall data analysis software was used to assign quality scores (GeneTrain and GeneCall) for each genotype. A GeneCall genotype quality score of 0.15 was used as a cutoff threshold for removing low quality genotypes. There were 57.8% of the samples with a call rate of at least 0.995, 95.2% had a call rate of at least 0.950 and 97% of the samples had a call rate of at least 0.900.

- There were 46.1% of the markers with a call frequency of at least 0.995, 93.1% with a call frequency of at least 0.950, and 94.8% of the samples with a call frequency of at least 0.900. There were 3.6% of the markers with a call frequency of 0. There were 58,982 SNPs with a call frequency of at least 0.900. In this group of reliable marker assays there were 9,627 SNPs with a MAF<0.05, 6,450 SNPs with a MAF<0.01 and finally 4,147 SNPs completely fixed for one of the alleles.

- The level of DNA degradation is influencing the fraction of SNPs successfully genotyped. The samples with no or low level of DNA degradation had 99.5% and, respectively 99.2% of the SNPs genotyped. Samples with advanced level of DNA degradation have 87% of the SNPs genotyped.

- Rate of DNA degradation affects the genotyping accuracy. Three of the samples with different level of DNA degradation (absent, limited and increased degradation) were submitted for genotyping in duplicates. The number of discrepancies between called genotypes of the samples with absent and limited DNA degradation was extremely small (0.008% and 0.007% respectively) compared to the samples with increased level of DNA degradation (0.1%).

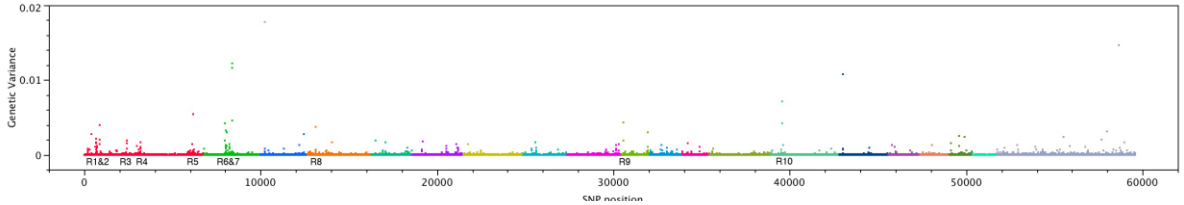
- We used RLRP software for haplotype analysis. A 92.5% phasing call rate was obtained for a selected set of 357 well-characterized SNPs. The software is currently under improvement to allow haplotype phasing of all markers located on the same chromosome.

- GenSel (Fernando and Garrick, 2008) was used to conduct a Bayesian analysis to estimate proportion of variance for each trait accounted for by the 59,608 SNPs from the Porcine SNPK60 BeadChip. The probability of a particular SNP to have an effect on a trait was set to 0.005. Variation in the age at puberty was associated with the largest SNP effects from all reproductive traits (Figure 1). This is not surprising since the heritability of age at puberty in our study is 0.42. Clusters of SNPs associated with large effects on puberty onset were located on Sus scrofa chromosome (SSC) 1 (in the 21, 29, 88, 127 and 269 Mb regions), SSC2 (56 and 80 Mb), SSC4 (16 Mb), SSC10 (3Mb) and SSC14 (27 Mb). For simplicity we coded these regions R1 to R10 (Table 1). Suggestive effects were also detected on SSC3, SSC6, SSC8, SSC9, SSC10, SSC11, SSC15 and SSC18.

Table 2. Major chromosomal regions associated with variation in age at puberty

Region	Chromosome	Position (Mb)
R1	1	21.482
R2	1	29.215
R3	1	88.261
R4	1	126.999
R5	1	269.336
R6	2	55.970
R7	2	80.205
R8	4	16.190
R9	10	2.964
R10	14	26.542

Figure 1. Genome wide association between 59,608 SNPs and age at puberty. Each dot represents the variance explained by a single SNP. X axis represents the location of the SNPs on the 18 autosomes and chromosome X. Y axis represents the contribution of that locus to the genetic variance. Each color represents a chromosome, from SSC1 to 18, followed by chromosome X and by a group of SNPs (represented in grey) without a precise location. R1 to R10 represent regions with the highest effects.



As expected, lower marker effects were identified for litter size from parity 1 to parity 4 (Figure 2) and lifetime productive traits (Figure 3 and 4). Variation of lifetime number of live-born piglets is potentially influenced by genes located on most swine chromosomes (Figure 3). Clusters of SNPs clearly associated with the largest effects are located on SSC2, SSC8, SSC13, SSC16 and SSC17. Similarly, lifetime number of weaned piglets is affected by many chromosomes but clear clusters of SNPs associated with major effects are located on SSC1, SSC2, SSC3, SSC8, SSC9, SSC11, SSC12, SSC16 and SSC17 (Figure 4). Common major loci that affect lifetime number of live-born and number of weaned piglets are located on SSC1, SSC2, SSC16 and SSC17 (Figures 3 and 4).

In our analysis, age at puberty iss negatively correlated with reproductive and lifetime productive performance in pigs (-0.18). Given that both traits are dependent on the function of the hypothalamic-pituitary-gonadal axis, we anticipated that the variation of both traits would be influenced by the same gene variants. We found such evidence in a combined association analysis of age at puberty and sow lifetime production traits. For example, the same regions/genes from SSC1 (21, 29, 88, 127 Mb) seem to influence variation of both age at puberty and the lifetime number of live-born piglets (Figure 1 and 3). Likewise, the first three regions detected on SSC1 (21, 29 and 88 Mb) also influence the lifetime number of weaned pigs (Figures 1 and 4). Interestingly, one of the clusters with the richest number of SNPs associated with the total number of live-born and weaned pigs and located on the distal end of SSC17 is also influencing litter size on parity 3 (Figures 2-4).

Figure 2. Genetic variance of SNPs for litter size. Potential positional candidate genes located in areas associated with important effects are represented in squares.

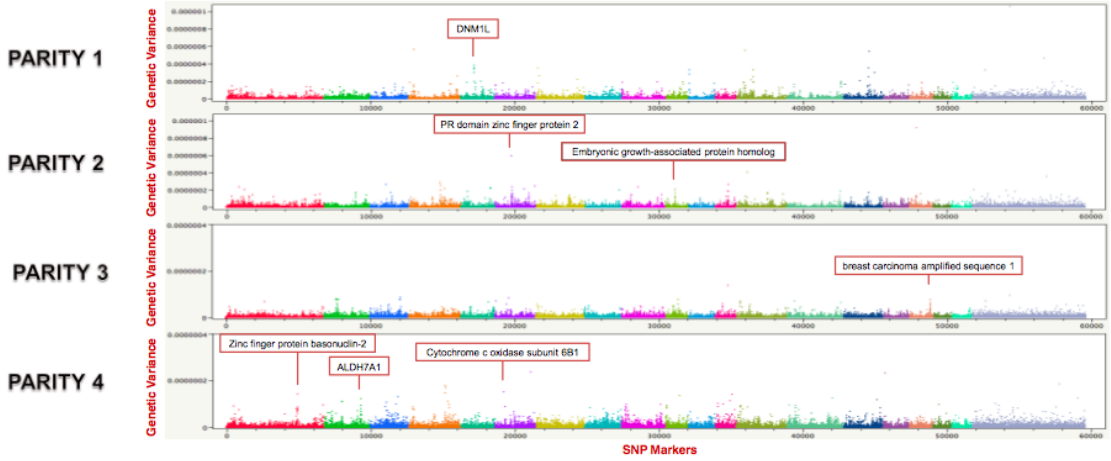
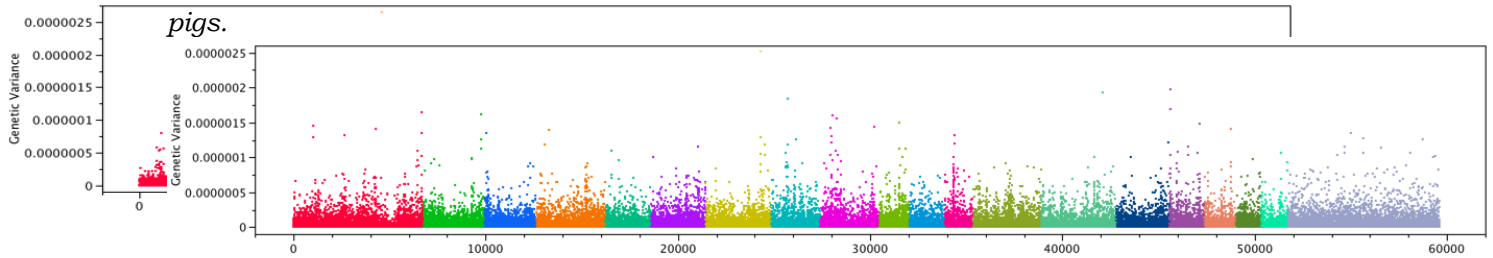


Figure 3. Genome wide association between 59,608 SNPs and variation of lifetime number of live-born piglets.

Figure 4. Genome wide association between 59,608 SNPs and variation of lifetime number of weaned pigs.

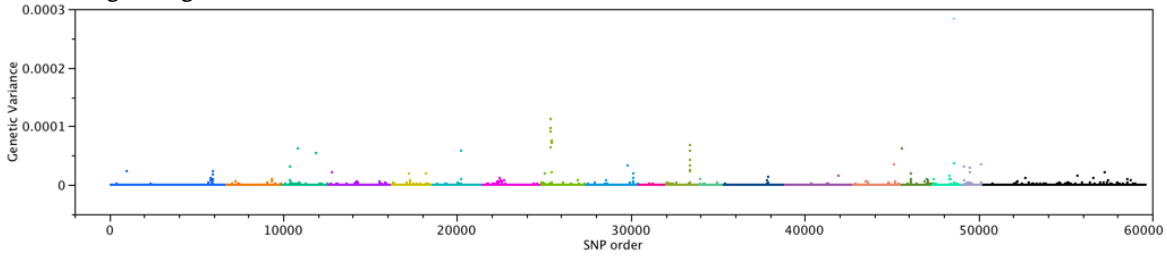


- We selected for genotyping additional SNPs that are not present in the Porcine 60KSNP BeadArray. All the selected polymorphisms are located in genes known to influence either reproductive traits, growth and energy metabolism. We selected two SNP from *MC4R* (Fan et al., 2009), one SNP from *PRKAG3* (V199I, Ciobanu et al., 2001) and one from *EPOR* (Vallet et al., 2005). *PRKAG3* polymorphism was informative in our resource population while *MC4R* and *EPOR* polymorphisms were not informative. Single marker association analysis did not uncover any significant relationships between the genotypes of *V199I PRKAG3* polymorphism and reproductive traits.

- While the majority of the SNPs could be mapped on the latest assemble (build 9) of the porcine reference genome, we hope to have almost all SNPs mapped on the new assemble (build 10) expected to be released in the summer of 2010.

Objective 2. We used the same statistical approaches described above to estimate proportion of variance for sow developmental traits accounted for by the 59,608 SNPs from the PorcineSNP60 BeadChip (GenSel, Fernando and Garrick, 2008). For example, the main clusters of SNPs associated with gilt weaning weights are located on proximal SSC8 and distal SSC11 (Figure 5).

Figure 5. Genome wide association between 59,608 SNPs and variation in gilt weaning weights.



Major loci influencing growth from weaning to the initiation of the treatment (average of 123 d) when the gilts were subjected to different feeding regimen are located mainly on SSC1, SSC2, SSC3, SSC4, SSC16 and 18 (Figure 6). Additional suggestive loci appear to be located on SSC5, SSC6 and SSC10. Growth during the developmental phase (treatment, from approximately 123 d till breeding) when the gilts were subjected to two feeding regimens appeared to be influenced by regions located on most swine chromosomes (Figure 7). As expected there are loci that potentially influence variation in growth during both developmental phases as are those located distal on SSC2 and proximal on SSC3 (Figures 6 and 7).

Figure 6. Genome wide association between 59,608 SNPs and gilts growth from weaning to treatment initiation (123 d).

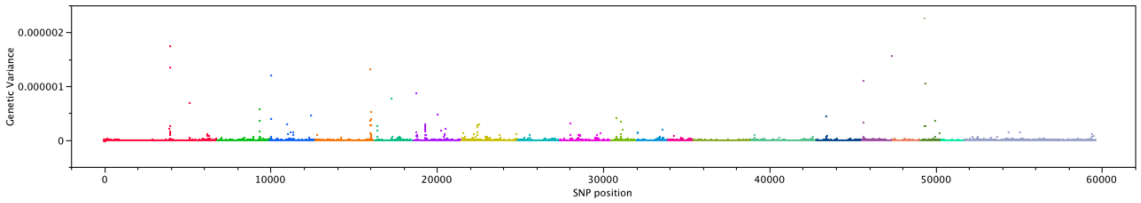
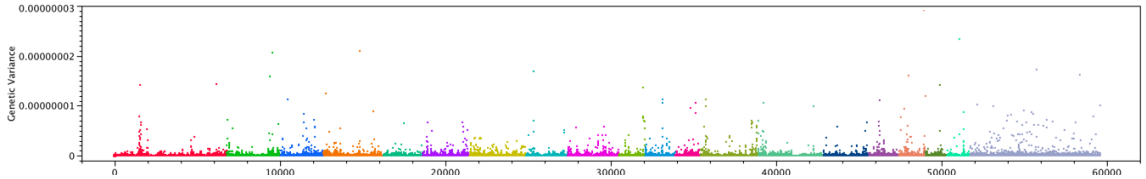


Figure 7. Genome wide association between 59,608 SNPs and gilts growth during developmental phase (123 d till breeding) when gilts were subjected to different feeding regimens.



Additional efforts are currently underway to identify potential candidate genes and markers in high linkage disequilibrium with the most important traits. Panels of markers are constructed for these traits and molecular breeding values estimated.

Objective 3. Discussions were initiated with Newsham (Dr. Archie Clutter), PIC (Genus) and Danbred (Dr. Tom Rathje) to test the value of some of the markers in predicting sow lifetime productivity using genome-wide assisted selection. All these major breeding organizations expressed an interest in evaluating our research results and use our results as training data.

Discussion:

Genome wide association studies for identification of loci that influence developmental, lifetime productive and reproductive traits were performed in 639 gilts of two maternal crossbreds. Variation in age at puberty was associated with the largest SNP effects from all the reproductive traits. Clusters of SNPs associated with large effects on puberty onset were located on SSC 1, SSC2, SSC4, SSC10 and SSC14. As expected we found the variation of sow lifetime productivity traits to be potentially influenced by genes located on most swine chromosomes. Clusters of SNPs clearly associated with major effects on lifetime number of live-born piglets were detected on SSC2, SSC8, SSC13, SSC16 and SSC17. Clusters of SNPs associated with important effects on lifetime number of weaned pigs were detected on SSC1, SSC2, SSC3, SSC8, SSC9, SSC11, SSC12, SSC16 and SSC17. Loci detected on SSC1, SSC2, SSC16 and SSC17 clearly influence both lifetime number of live-born and weaned pig.

Combined association analysis of age at puberty and sow life production traits revealed common regions/genes located on SSC1 (21, 29, 88, 127 Mb) that influence phenotypic variation of age at puberty and lifetime number of live-born and weaned piglets per sow. In addition, one of the clusters with the richest number of SNPs associated with the total number of live-born pigs and located on the distal end of SSC17 is influencing variation in litter size on parity 3 as well.
Main clusters of SNPs associated with gilts weaning weights are located on proximal SSC8 and distal SSC11. Major loci influencing growth from weaning to the initiation of the treatment are located on SSC1, SSC2, SSC3, SSC4, SSC16 and SSC18. Growth during the developmental phase appeared to be influenced by major loci identified on most swine chromosomes. There are loci that potentially influence variation in growth during both developmental phases.

We are in process of assembling a panel of DNA markers associated with sow reproductive and productive longevity and identify procedures to apply the information in whole genome selection. The effects of the significant markers will be validated in other commercial populations. We expect that our research will provide a panel of molecular markers that can be successfully used for reducing culling rates, sow death losses, and enhance the productive life of sows.