

Genomics-based selection for reproduction and adaptation in pigs

Egbert F Knol and Pramod K Mathur

TOPIGS Research Center IPC, Schoenaker 6, 6641 SZ Beuningen, The Netherlands

Recent publication of the sequence of the pig genome allows for faster and more accurate selection for traits expressed late in life, sex limited or difficult to measure. Especially reproduction and adaptation traits fall in this category. Existing genetic variation in reproduction traits is substantial and polygenic selection has resulted in genetic trend of around half a piglet per sow per year. A very significant increase in genetic trend is anticipated through genomic selection. Genomic selection capitalizes mainly on Mendelian sampling, variation between full sibs is half of the total genetic variation; estimation of genetic merit based on markers yields, because of quantification of this Mendelian sampling, far higher accuracies than on the basis of parental average. For both polygenic and genomic selection, application appears to follow theory. Current research focuses on application of full sequence per animal, on non-additive genetic effects, and on genomic testing of embryos in order to increase selection intensity. Gradually traits relating to genotype-environment interactions will be added to the selection index, using genomic based methods.

Introduction

Genetic variation helps a population to adapt to changing environments. A population adapts its reproductive performance to a level which can be sustained given available resources. Pork production systems lower cost price by increasing reproduction and minimizing resource constraints through more and better feed, improved climate control and reduction of disease challenges. Pig breeding organizations have exploited genetic variation in reproductive traits by applying selection programs and extensive data recording systems for several decades now (Dekkers et al., 2011). Pigs have been selected based on quantitative polygenic assumptions for the traits of interest. This selection has been quite successful, but has led to noticeable changes in the physiology of the sow (Foxcroft, 2012). The publication of the porcine genome (Groenen et al., 2012) creates the potential for a much higher genetic trend in existing and in new reproduction traits.

Goals of this paper are (1) to quantify the basics of genetic selection for reproductive traits; (2) explore potential changes in reproduction traits through the application of genomic techniques; and (3) explore the role of genetic variation under challenged environmental conditions.

Basics of genetic change

Genetic variation

Litter size, total number born piglets per litter, can vary between zero and thirty. The litter size at a farm depends upon environment, parity, season, nutritional status etc. After correction for

these effects the residual standard deviation is around 3.2. Phenotypic variance is then 10 and heritability, the ratio between genetic and phenotypic variance, is close to 0.10. Genetic variance is therefore (around) 1.0 and genetic standard deviation (σ_H) is, consequently, around 1.0 too.

Polygenic selection

The basic formula to estimate genetic change per year reads:

$$\Delta G_{year} = \frac{\sigma_H * r_{IH} * i}{L}$$

where i is selection intensity and L is generation interval (the average age at which the offspring is born). Let us assume, for the sake of simplicity that $i = 1$ and $L = 1$, a slightly optimistic view. Since, $\sigma_H = 1$, genetic trend equals accuracy. The accuracy of the estimate of the genetic quality of a candidate at the moment he or she is up for potential selection is denoted as r_{IH} . Since the selection candidate does not have its own performance for reproduction at age of selection (around 6 months), this accuracy is half of the average accuracies of the sire and the dam at the moment of selection. The sire has a very low accuracy (around 0.1) at that time; no own performance for female reproduction and no daughters with a performance (yet). If the dam has multiple parities and daughters with own records, her accuracy is around 0.5 at that moment. Consequently genetic trend for litter size would then be $(0.1 + 0.5)/2 = 0.30$. Single trait genetic selection for litter size can therefore yield around 0.3 piglets total born per litter per year in a well-organized selection program. By chance this genetic change of 0.3 pigs per year using the quantitative or polygenic approach also equals 0.3 genetic standard deviations in litter size.

Genomic selection

The international swine genome sequencing consortium effort has led to the availability of large number of genetic markers (Single Nucleotide Polymorphisms). At this moment the 60k SNP chip is the most widely used marker set, with on an average, one genetic marker per 50,000 base pairs. Comparison of the marker results with reproductive performance of a large number of well phenotyped sows yields information on regions of interest and on preferred SNP scores in those regions. Consequently, selection candidates can get a genetic evaluation without having an own performance and even without having known parents. The accuracy of this method should ideally go to 1 in the situation that the genetic markers explain all of the genetic variation. This accuracy of 1.0 is the holy grail for geneticists but currently lower since it is a function of the number of markers and the size and quality of the phenotypic training set, with a large number of assumptions. In the setting of pig improvement with a 60k SNP chip and 1,500 fifth parity sows in the training set, this accuracy tends to be around 0.5.

If parental information is added to this 0.5, the total accuracy could be almost double the original accuracy of 0.3, and could therefore double the genetic trend. If training populations increase and/or chip densities increase or other improvements take place then there is even more space for further improvement (an accuracy of 0.9 could triple the current trend).

Applying selection to reproduction traits

Reproductive Traits

Pork production is about lowering the cost price of good quality meat. Cost price of a piglet is a relevant part of total cost price (around 25%). Increasing numbers of weaned piglets per sow

per year decreases the average cost price of a piglet. Selection for reproductive performance therefore starts with focus on litters per sow per year and piglets weaned per litter. Next in line are traits describing quality of piglets in terms of birth weight and uniformity in birth weight. Finally there is a group of traits describing longevity and adaptation to cope with challenges.

In the first two columns of Table 1 heritabilities (Bidanel, 2011) and genetic standard deviations (TOPIGS database) are given for a number of reproductive traits. According to the litter size example in the previous paragraph around 30% of a genetic standard deviation is feasible under single trait selection. Given the choice, litter size could increase by 0.3 pigs or litter mortality could decrease by 1.2% , or number of teats could increase by 0.20 teats or average birth weight by 30 g per year; or a combination of these traits, weighed by economic interest.

Polygenic selection

Column 3 of Table 1 presents estimates of realized genetic trends over the years 2007-2011 in reproduction traits based on breeding values from a multi trait analysis of TOPIGS dam lines. Genetic parameters were estimated according to analyses same or similar as in Hanenberg et al., 2001. Results presented are the average trends of a Yorkshire and of a Landrace line, resulting in the best estimate for trends in the crossbred sow. Breeding values on litter size were estimated on 7.4 million farrowing records from 1.65 million sows. Birth weight and uniformity in birth weight were available on almost 597 thousand litters with individually weighed piglets, where uniformity was based on the within litter variation in birth weight, including stillborn piglets. Nielsen B, Su G, Lund M & Madsen P 2013, unpublished, reported similar estimates for litter size, based on gilts only, over the past 7 years and based on piglets alive at day 5.

The last column in Table 1 gives the ratio of realized annual genetic trend to genetic standard deviation. In total, Table 1 shows the struggle for balanced breeding. The genetic trend in total number born with 0.07 standard deviations is quite far from the theoretical potential of 0.30.

Table 1: Heritability estimates of reproduction traits with genetic standard deviations and genetic trends according to TOPIGS.

<i>Description</i>	<i>Heritability</i>	<i>Genetic std σ_H</i>	<i>Genetic trend</i>	
			<i>Cumulative 2007-2011</i>	<i>Standardized/ year</i>
Total number born, #	0.11	1.02	0.38	0.07
Stillborn piglets, #	0.07	0.34	-0.19	-0.11
Age first insemination, d	0.37	11	-11.86	-0.22
% prolonged interval ¹	0.22	16	-1.06	-0.01
Farrowing index, %	0.07	7.5	0.98	0.03
Litter mortality, %	0.05	2.83	-1.34	-0.09
Number of weaned piglets, #	0.08	1.93	0.67	0.07
Gestation length, d	0.31	0.75	-0.09	-0.02
Litter variation, g	0.08	30.3	-13.85	-0.09
Number of teats, #	0.45	0.68	0.92	0.27
Longevity, parities	0.09	0.48	0.15	0.06
Birth weight, kg	0.24	0.112	-0.06	-0.11

¹ % prolonged interval from weaning to insemination

However, the genetic trend in weaned piglets per litter (0.67 extra piglets weaned in 5 years time) is higher than the genetic trend in total born (0.38), indicating higher survival during the farrowing as well as lactation. Genetic trend in number of teats is higher than trend in piglets weaned per litter, giving more opportunity in terms of mothering ability of the sows. Number of teats appeals to society as society sees an extra teat for each piglet as very relevant. Genetic trends in farrowing rate and longevity were positive too. If genetic trends for all traits are multiplied by their respective economic values the trend is a 2% reduction in cost price of a piglet, assuming the production and cost levels in 5 years time when the genes of these nucleus populations will be fully expressed at commercial level.

Selection pressure on birth weight was not high enough to prevent birth weight from decreasing during this time period. Litter variation (within litter standard deviation in birth weight), however, decreased almost 14 g in 5 years, creating more uniform piglets with better chances of survival.

Genomic selection

Two approaches for the use of genomic information are currently used. The first estimates the relationship between animals on the basis of SNP's instead of using the average relationship between family members (that is 0.50 for full sibs, 0.25 for half sibs etc.), some half sibs are almost fullsibs, while others are more like first cousins. The second approach is to train a population, that is to genotype a well phenotyped population of animals and establish the relation between the genetic markers and the traits of interest. Young selection candidates can then be genotyped and their phenotype or their breeding value can be estimated from the genetic markers genotyped.

This training phase is referred to as Genome Wide Association Study. The goal is to find variation in traits linked to variation in genetic markers. An example of a GWAS plot is given in Figure 1, where the Y-axis has an indicator of significance (triangles indicating genome wide significance beyond any doubt) and the X-axis indicates the individual markers ordered by chromosome. Linkage equilibrium (LD), causes long stretches of base pairs to be inherited together without recombination; therefore, in single marker analyses more than one marker tends to give a significance signal in a specific interval. Schneider et al, 2012b used groups of 5 consecutive SNP markers in a genome wide association study and identified a number of

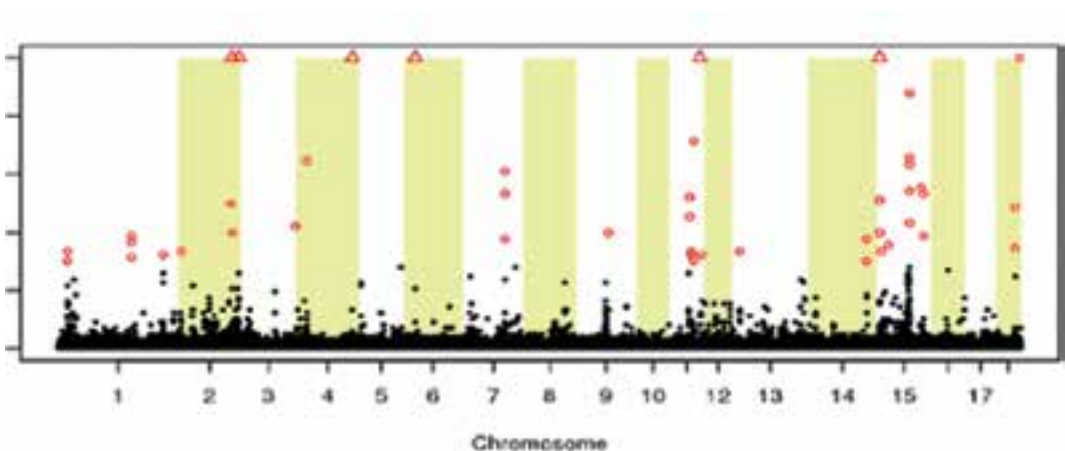


Figure 1. Genome Wide Association plot for mothering ability, the estimate of foster effect in a direct-foster model for piglet survival.

chromosomal regions associated with reproduction traits. Fine mapping of these regions can even lead to identification of functional mutations to dramatically enhance the rate of genetic progress.

Genetic variance explained by markers

Estimates of the marker effects can be used with their MAF (minor allele frequency) to estimate the variance explained by the specific marker. The sum of the variances of all relevant markers is indicative for the expected genetic trend using this set of markers.

In Table 2 variance components are given for some of the traits from Table 1. This table serves only as an example. The ratio between marker variance and phenotypic variance is typically lower than the polygenic heritability. Heritability estimated through polygenic variance is an unbiased estimate of the ratio between genetic variance present in the base population and phenotypic variance. Marker variance is estimated in the present population.

Loosely estimated the variance explained by the markers is around 55% of the variation found under polygenic analysis. Ideally this would be 100%, where all genetic variation is explained by individual markers. There are, at least, three reasons why this is not the case: (1) training populations are still limited in pig genomic work, 1,000-2,000 animals with relatively inaccurate phenotypes, (2) still low marker density, 60k relates to 1 marker per 50,000 base pairs, and (3) non additive effects such as dominance and epigenetic effects.

(1) *limited populations*. Genotyping is becoming cheaper at a faster rate than computer memory. (Pig) Breeding companies are finding efficient methods (imputation) to derive higher density genotyping for large numbers of animals, e.g. derive 60k from only 400 SNP's (Huang et al., 2012). The sex of choice tends to be the female in pigs, with repeated records in terms of reproductive performance instead of offspring tested sires. In dairy cattle a number of countries have formed a European consortium that aims to have 16,000 genotyped bulls with phenotyped daughters.

(2) *low marker density*. With the pig genome being sequenced, millions of SNP's are available for testing. 500-700k chips are becoming available. There is some reluctance in using them since full sequence of individual animals appears to be within reach at reasonable costs. First results in a number of countries did not yield promising results in increasing accuracy going from 50k to 700k in dairy cattle (general consensus at EAAP 2012).

Table 2: Estimates of explained variation by markers in relation to estimated polygenic variance.

<i>Trait</i>	<i>Source</i>	<i>Phenotypic Variance</i>	<i>Proportion explained by markers</i>	<i>Heritability</i>
Total number born, #	Schneider ¹	9.05	0.04	0.09
Litter birth weight, kg		15.93	0.11	0.20
Total number born, #	TOPIGS ²	10.49	0.12	0.17
Average birth weight, kg		0.035	0.32	0.47
Litter variation, g		4670	0.05	0.19
Number of teats, #		0.8	0.19	0.39
Arbitrary Average			0.14	0.25

¹Schneider et al., 2012a, estimated based on 1152 genotyped animals

²TOPIGS, estimated based on 1012 genotyped animals

(3) *non additive genetic framework*. Nature is more complex than we would like. The number of proteins coded for is in the range of 20,000-25,000. Functional mutations in the regions coding for these proteins are rare, they are highly conserved. Genetic variation mainly operates on the regulation of these genes, creating opportunities to adaptation to changing environments, epigenetic effects. Dominance is mainly known from crosses between lines. Obviously, the same phenomenon applies within lines. Until now we did not have the tools to quantify the effects.

In house testing of the use of 60k markers show validation results fully in line with the explained variance of Table 2, the original example stated an accuracy of 50% using genomic selection. The negative side is that family relations are still very important. Predictive ability depends on closeness in terms of pedigree. Application of markers in a different dam or sire line will yield very limited results. Within a line with close phenotyping/genotyping the accuracy of selection should go to 0.7-0.8 with increasing size of trainings population and increasing marker densities.

Adaptation to challenges

The pork industry is an environment, or possibly a set of environments, different than the original environment were pigs were raised for many centuries. The animals have the genetic flexibility to adapt to new environments. The easiest and most straightforward method is to test animals in the environment they have to produce in. The theoretical framework to get around genotype-environment interactions is still under development and can benefit from genomics work.

Theoretical framework

Genotype-environment interactions can be investigated either using multiple trait models (Meyer, 2009) or using random regression reaction norm models (Kolmodin et al., 2002, Calus and Veerkamp, 2003; Cardoso and Tempelman, 2012, Silva FF, Mulder H, Knol EF, Lopes MS, Guimarães SEF, Lopes PS, Bastiaansen JWM 2013, unpublished). In the study of genomic reaction norms Silva FF, Mulder H, Knol EF, Lopes MS, Guimarães SEF, Lopes PS, Bastiaansen JWM 2013, unpublished, investigated litter size records from 295 farms distributed over 22 countries. At first the herd production levels were calculated as a measure of the environmental stress. Then heritabilities were estimated using random regression reaction norm approach using phenotypic as well as genomic information. The results given in Figure 2 suggest higher heritabilities in more favorable environments as well as under extremely unfavorable environmental conditions. One option is to raise the management levels to get higher production and higher response to selection. At the same time the study also suggest that pigs selected under most favorable conditions will have poor correlated response under more stressful environmental conditions and vice versa, indicating the need for selection of specific lines for specific environmental conditions. The selection of robust animals for production in a variety of environmental conditions should then be conducted under moderate or average production levels.

Genomics and new traits

The conclusion of the previous discussion is that traits can be environment dependent and therefore under different genetic control. This is a challenge and an opportunity at the same time. Our own work (Bloemhof et al, 2012) shows clear reactions of both farrowing rate and

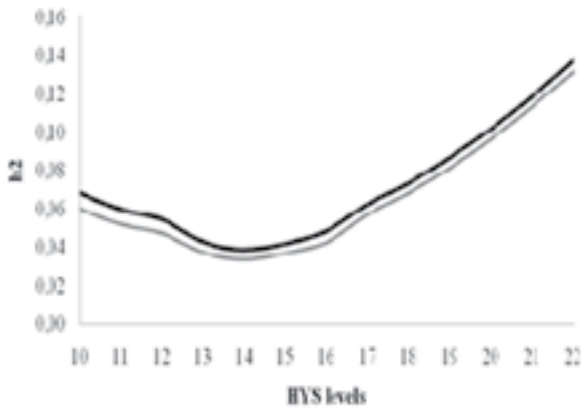


Figure 2. Heritability estimates for total number born in pigs at different herd-year-season (HYS) levels considering conventional and genomic relationship matrices. Source: Silva FF, Mulder H, Knol EF, Lopes MS, Guimarães SEF, Lopes PS, Bastiaansen JWM 2013 Sire evaluation for total number born in pigs using genomic reaction norms approach, unpublished.

litter size above the upper critical temperature (UCT) of 20C and shows genetic variation in the decline above this UCT. They suggest that farrowing rate might be an indicator trait for quality of environment and that farrowing rate shows more genetic variation under adverse environments.

Late in life traits, like longevity/stayability are prime candidates for genomic selection (e.g. Onteru et al., 2011), as are resistance/tolerance to diseases, e.g. PRRS resistance (Boddicker et al., 2012).

Where will we go

Implementation in the breeding program

Breeding companies would like to have high health, high quality, moderate climate nucleus farms (say: on the isle of Man), the right side of Figure 2 shows that it is relatively easy to find the better families because of the higher heritabilities. Use of this type of genetics in a dense pig populated area of the Philippines might not yield the expected result, classic genotype-environment interaction. However, if a closely linked subset of the population would perform in the challenged environment and would be genotyped then selection of the populations could be optimized based on the relative sizes of the markets for high performance and more challenged environments. Theory dictates that nucleus populations could be substantially smaller than the current numbers, because genomic selection allows for the exploitation of Mendelian sampling in contrast to the current/past situation. Investments could then be moved to 'phenotyping stations', where relatives of the nucleus animals could be subjected to disease, heat stress, no cross fostering, super ovulation or other stressors in order to find relevant, but normally hidden genetic variation for specific production environments.

Back to genetic selection; other factors than accuracy

Considering the formula for genetic change from the previous section,

$$\Delta G_{\text{year}} = \frac{\sigma_H * r_{IH} * i}{L}$$

so far, accuracy of selection was under discussion, this r_{IH} could increase from around 0.30 to around 0.75 depending on the trait and the sex of the animal. i is selection intensity. A normal

litter of an elite boar and an elite sow will have, say, 8 males and 8 females for selection. Half of the genetic quality of these offspring is attributed to the quality of the sire and the dam and the other half to 'chance', to the Mendelian sampling. If number of offspring can be increased through super ovulation and or through repeated matings and/or through other reproductive technologies, then selection intensity can increase quite dramatically, assuming that genotyping of embryos will become available at a cost effective level. Generation interval is the average age of the parents at which the offspring is born. Under polygenic selection a balance has to be found between phenotyping (waiting for more performance records) and number of selection cycles (short generation interval). Under genomic selection, accuracy depends mostly on the training population and generation interval can therefore be minimized, offering new possibilities. σ_H is considered fixed, it is existing genetic variation, albeit with new traits and with that new genetic variation can come under selection pressure. σ_H be changed with introgression of new genetic variation, coming from other lines or breeds within species or, theoretically, even from introgression across species. The latter form is not an option for TOPIGS, because of internal ethical considerations.

Biological limitations

All relevant reproductive traits show genetic variation and are therefore accessible to genetic variation. This allows the population to be flexible and allows animal breeders to adapt animals to the current and future pork production systems. There is no limit in sight to genetic increases in litter size or to genetic decreases in mortality. Efficiency, lowering the cost price of pork, is the driving factor. Main and possibly not small risks are in the unexpected, in the genetic correlations between traits.

Foxcroft (2012): "Overall, it is probably fair to conclude that the last two decades of genetic selection in swine have produced changes on the physiology of the sow and her offspring that might have been predicted, but were not expected to be as profound as they now appear to be. However, by adopting a closer interaction among biologists, geneticists and commercial production systems as we move into the 'omic' era of biology,...". The economic value of an extra piglet decreases with increasing level of production. Genetic parameters indicate that birth weight can be kept at least constant even with increasing litter sizes, uterine capacity has to increase in that case. Adult weight might go up and feeding has to be adapted. So we fully agree that closer interaction is needed where genetic trends will go faster and new traits will pop up.

Concluding remarks

Substantial genetic variation exists in reproduction traits. Polygenic theory is correct and quite accurately predicts genetic trend. First signs of genomic selection theory are promising and seem to be equally accurate in their predictions. The past two decades have witnessed increasing genetic trends based on polygenic selection that substantially changed reproductive ability of sows. The next decade will show an almost double trend in 'total genetic standard deviation'. The content of that trend will be different than in the past two decades. Emphasis will be more on 'ease of production', 'robustness', 'flexibility', and 'adaptability to different environments'. And proper care should be put on maintaining quality of finisher pigs, a further increase in reproduction should not lead to reduced lean growth performance, pork quality and overall efficiency of production under given environmental conditions.

References

- Bidanel J-P** 2011 Biology and genetics of reproduction. In *The Genetics of the Pig*, edn 2, pp 218-241. Eds MF Rothschild and A Ruvinsky. Cambridge: CABI.
- Bloemhof S, Kause A, Knol EF, Van Arendonk JA & Misztal I** 2012 Heat stress effects on farrowing rate in sows: genetic parameter estimation using within-line and crossbred models. *Journal of Animal Science* **90** 2109-19.
- Boddicker N, Waide EH, Rowland RR, Lunney JK, Garrick DJ, Reecy JM & Dekkers JC** 2012 Evidence for a major QTL associated with host response to porcine reproductive and respiratory syndrome virus challenge. *Journal of Animal Science* **90** 1733-1746.
- Dekkers JCM, Mathur PK & Knol EF** 2011 Genetic improvement of the pig. In *The Genetics of the Pig*, edn 2, pp 390-425. Eds MF Rothschild and A Ruvinsky. Cambridge: CABI.
- Calus M & Veerkamp R** 2003 Estimation of environmental sensitivity of genetic merit for milk production traits using a random regression model. *Journal of Dairy Science* **86** 3756-3764.
- Cardoso FF & Tempelman RJ** 2012 Linear reaction norm models for genetic merit prediction of Angus cattle under genotype by environment interaction. *Journal of Animal Science* **90** 2130-2141.
- Foxcroft GR** 2012 Reproduction in Farm Animals in an Era of Rapid Genetic Change: Will Genetic Change Outpace Our Knowledge of Physiology? *Reproduction in Domestic Animals* **47**(Suppl. 4) 1-7.
- Groenen MAM, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, Rogel-Gaillard C, Park C, Milan D, Megens H-J, et al.** 2012 Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* **491** 393-398.
- Hanenberg EHAT, Knol EF & Merks JWM** 2001 Estimates of genetic parameters for reproduction traits at different parities in Dutch Landrace pigs. *Livestock Production Science* **69** 179-186.
- Huang Y, Hickey JM, Cleveland MA & Maltecca C** 2012 Assessment of alternative genotyping strategies to maximize imputation accuracy at minimal cost. *Genetics Selection Evolution* **44** 25
- Kolmodin R, Strandberg E, Madsen P, Jensen J & Jorjani H** 2002 Genotype by environment interaction in Nordic dairy cattle studied using reaction norms. *Acta Agriculturae Scandinavica, Section A-Animal Science* **52**(1) 11-24.
- Meyer K** 2009 Factor-analytic models for genotype \times environment type problems and structured covariance matrices. *Genetics Selection Evolution* **41** 21-32.
- Onteru SK, Fan B, Nikkilä MT, Garrick DJ, Stalder KJ & Rothschild MF** 2011 Whole-genome association analyses for lifetime reproductive traits in the pig. *Journal of animal science* **89**(4) 988-995.
- Schneider JF, Rempel LA & Rohrer GA** 2012a GWAS of swine farrowing traits Part I: Genetic and genomic parameter estimates. *Journal of Animal Science* **90** 3353-3359.
- Schneider JF, Rempel LA, Snelling WM, Wiedmann RT, Nonneman DJ & Rohrer GA** 2012b Genome-wide association study of swine farrowing traits. Part II: Bayesian analysis of marker data. *Journal of Animal Science* **90** 3360-3367.

