Simulation Program to Study Genomic Selection to Improve Resistance to PRRS in Swine

Larry Schaeffer

Centre for Genetic Improvement of Livestock
Department of Animal & Poultry Science
University of Guelph, Guelph, Ontario, Canada

1 Introduction

The nature of the PRRS virus and the structure of the swine industry were determining factors in the design of a simulation study to determine the efficacy of genomic selection in swine to improve the resistance of pigs to the PRRS virus. Firstly, the PRRS virus is primarily a problem for the commercial herd owners, affecting both sows and piglets. The strict biosecurity measures at the multiplier and nucleus levels of the pyramid structure mean that disease issues occur infrequently. At the commercial level, however, 15% of litters are affected resulting in 25 to 70% mortality of the fetuses. Growing pigs are also affected, such that around half of the infected pigs eventually die.

Secondly, all of the selection and breeding decisions are made at the nucleus level. In order to evaluate any nucleus animals, the phenotypes observed at the commercial level need to be linked to animals in the nucleus herds. By the time the commercial pigs are observed with or without the virus, their ancestors have been culled. Genomics may be useful to bridge this problem.

Thus, a simulation program was needed which included

1. nucleus, multiplier, and commercial levels of production.
2. many of the important production traits.
3. QTLs with varying numbers of alleles and many QTLs per chro-

mosome.

4. flexibility to study many strategies on the use of genomics.
5. daily activities per pig over ten years or more.

2 Objective

This report presents the details of the simulation program as a background for other papers on strategies of applying genomics to the problem of solving the PRRS virus problem.

The simulation consists of two programs. The first part utilizes a software package called QMSim to generate the QTLs of the genome. The second part simulates the day to day activities of the pyramid structure for each animal in each herd.

3 Part I. Genomic Structure

A swine breeding company typically has two breeds that are selected for maternal traits, like number born and number weaned. A third breed is a terminal cross breed bringing in the growth traits to the crossbred piglets that go to market. Genetically, each breed has 19 pairs of chromosomes (18 autosomal), with many QTLs. Breeds differ in the frequencies of the alleles at each QTL, and some alleles may not exist in one breed.

Swine have undergone many generations of selection and they have built up strong linkage disequilibrium. So the starting populations in a simulation need to reflect this type of background. The software program called QMSim (Sargolzaei and Schenkel, 2009) does the necessary historical generations of matings with recombinations and mutations. The program, QMSim, was used to simulate 3000 historical generations of matings of 50 males and 500 females. In the last 50 generations the population was increased to 5100 individuals. There were 4000 QTL assumed for each of 18 autosomal chromosomes of length 100 cM. QTL

were allowed to have 2 to 7 alleles. The mutation rate was 1×10^{-4}. The input parameter file for QMSim is given in Appendix I.

From the last generation of historical matings, founders for 3 breeds were chosen and 50 generations of random matings were conducted with 200 males and 2000 females for each breed. This was followed by 10 generations that included phenotypic selection for a single trait with heritability of 0.20. If the breeds are called A, B, and C, then the last males and females in the B and C breeds were used to create (B×C) and (C×B) crossbred females for the commercial herds.

Although there were 72,000 QTL, many of those drifted to fixation during the historical generations. Thus, 10,000 QTL with minor allele freqencies greater than 0.01 were randomly chosen. QMSim produced separate files of over 20,000 individuals for each breed, and the two crossbred groups with their genotypes for the 10,000 QTL. The level of linkage disequilibrium that was reached was 0.29 for QTL that were 0 to 0.05 cM apart.

4 Part II. Allelic Effects

4.1 The Traits

Twenty two traits were simulated. The assumption was made that every QTL has allelic effects for each of the 22 traits. Table 1 contains a list of the traits considered in this simulation. Included are the means, phenotypic standard deviation, and heritability of each trait. Means vary by breed and gender. Table 1 has figures for gilts.

Table 1.
Genetic traits in the simulation.

	Trait	h^2	Mean	SD
1	Growth, A parameter	0.30	150	23
2	Growth, B parameter	0.30	1.81	0.75
3	Growth, C parameter	0.30	0.676	0.111
4	Feed intake	0.25	147	50.2
5	Lipid deposition	0.28	0.116	0.015
6	Protein deposition	0.27	0.156	0.008
7	Legs and feet	0.11	100	10
8	Ham	0.13	100	10
9	Loin thickness	0.13	100	10
10	Shoulder	0.13	100	10
11	Length	0.11	100	10
12	Stature	0.20	100	10
13	Teat number	0.10	100	10
14	Age at first mating, d	0.11	245	2
15	Conception rate %	0.04	0.90	0.04
16	Litter size	0.10	14	0.3
17	Survivability	0.02	100	10
18	Boar fertility	0.04	0.80	0.04
19	Gestation length	0.10	114	2
20	Stillbirths	0.03	0.02	0.14
21	Heat tolerance	0.12	100	10
22	PRRS resistance	0.02	100	10

A description of the traits and how they were utilized in the simulation appears in section 5.

4.2 Covariance Matrix

A genetic covariance matrix was constructed in piecemeal fashion with estimates derived from the literature for different breeds, time periods, and countries. Covariances for many combinations of traits were not readily available in the literature. Thus, many genetic covariances ended up being zero. A residual covariance matrix was also constructed in the same manner. Both matrices were required to be positive definite as per

Schaeffer (2013). The genetic covariance matrix is given in Table 5a, 5b, and 5c.

4.3 Generation of Allelic Effects

The problem was to simulate allelic effects for each of L separate loci (quantitative trait loci, QTL) for m traits, such that the genetic variance-covariance matrix among traits was equal to a pre-defined matrix, $\mathbf{G}$, of order m by m.

Usual multiple trait simulation involves the Choleski decomposition of $\mathbf{G} = \mathbf{CC}'$, followed by the generation of a vector, $\mathbf{u}$ of length m of random normal deviates, $N(0, 1)$, followed by the premultiplication by $\mathbf{C}$ to give $\mathbf{a}$, a vector of true genetic values for the m traits for one animal. Below is the derivation to show that this procedure is valid.

$$\begin{aligned} \mathbf{G} &= \mathbf{CC}' \\ \mathbf{a} &= \mathbf{Cu} \\ Var(\mathbf{a}) &= Var(\mathbf{Cu}) \\ &= \mathbf{C} \cdot Var(\mathbf{u}) \cdot \mathbf{C}' \\ &= \mathbf{C} \cdot \mathbf{I} \cdot \mathbf{C}' \\ &= \mathbf{CC}' = \mathbf{G} \end{aligned}$$

The above technique considers all gene loci collectively. The individual genotypic values, or the frequencies of the alleles are not involved. This makes the study of selection on single genes or groups of genes impossible.

Allelic effects need to be assigned for each locus for each of the m traits, say a_{kj} for the k^{th} locus and j^{th} allele, such that the desired $\mathbf{G}$ is obtained in the base population animals.

During the historical generations, mutations occur giving rise to multiple alleles per loci. At the same time drift may cause some alleles to

become extinct. The number of alleles is continually changing. The maximum number of alleles in this study was set to 7 for a given locus, and the minimum number was two. A sample of the results for one population $\mathbf{S}_o$ are given in the following table.

Table 2.
Distribution of Number of Alleles Per QTL

Alleles	Loci
2	2659
3	5411
4	5408
5	3591
6	1850
7	770

After QMSim had been run, population $\mathbf{S}_o$ is ready with alleles and loci. Effects for each allele at each loci were generated as follows:

To illustrate for $m = 3$ traits, and $L = 10000$, where the assumed genetic covariance matrix among traits is given in $\mathbf{G}$, where

$$\mathbf{G} = \begin{pmatrix} 900 & -90 & 60 \\ -90 & 153 & 6 \\ 60 & 6 & 6 \end{pmatrix}.$$

Then the Choleski decomposition of $\mathbf{G}$ is

$$\mathbf{C} = \begin{pmatrix} 30 & 0 & 0 \\ -3 & 12 & 0 \\ 2 & 1 & 1 \end{pmatrix}.$$

For locus k, generate a vector of m random normal deviates for the first allele,

$$\mathbf{u}_1 = \begin{pmatrix} 0.4740 \\ -1.5185 \\ -0.8203 \end{pmatrix},$$

then the allelic effects are

$$\mathbf{C}\mathbf{u}_1/L^{.5} = \mathbf{a}_{k1} = \begin{pmatrix} .142210 \\ -.196437 \\ -.013907 \end{pmatrix}.$$

If there is another allele, then generate another vector of m random normal deviates,

$$\mathbf{u}_2 = \begin{pmatrix} -1.5167 \\ 0.5911 \\ 0.2498 \end{pmatrix},$$

then

$$\mathbf{C}\mathbf{u}_2/L^{.5} = \mathbf{a}_{k2} = \begin{pmatrix} -.455005 \\ .116433 \\ -.021925 \end{pmatrix}.$$

Thus, if an animal has genotype 11, then the genotypic value for that animal at that locus would be,

$$\mathbf{g}_{k11} = \mathbf{a}_{k1} + \mathbf{a}_{k1} = \begin{pmatrix} .284420 \\ -.392874 \\ -.027814 \end{pmatrix}.$$

Similarly, for genotypes 12 and 22,

$$\mathbf{g}_{k12} = \mathbf{a}_{k1} + \mathbf{a}_{k2},$$

and

$$\mathbf{g}_{k22} = \mathbf{a}_{k2} + \mathbf{a}_{k2}.$$

With more than 2 alleles there are more possible genotypes. The above is repeated for every allele at every locus. Every allele has an effect on every trait. Sometimes the effects are large and sometimes small. The QTL with the largest allelic effects, and the top 50 QTLs were determined in each run.

Instead of using random normal deviates, one could use random numbers from a gamma distribution. These numbers are always positive, so that the sign of the number would need to be generated randomly as well. The gamma distribution would tend to have more small allele

effects, but the few large effects could be much larger than from a normal distribution.

Table 3 contains the allelic effects for the first few loci, as an example. Locus 1 has 3 alleles, locus 2 has 2, locus 3 has 5 alleles, locus 4 has 4, and locus 5 has 6 alleles. Within an allele at a locus, the effects were generated to have covariance matrix $\mathbf{G}$ across all alleles. The covariance between effects for the same trait on different alleles is expected to be zero.

Table 3

Example set of genotype effects for L loci.

Locus	Trait	Alleles						
		1	2	3	4	5	6	$\cdots$
1	1	-0.4670	-0.3228	-0.2251	0	0	0	
	2	0.0396	0.2114	0.0941	0	0	0	
	3	-0.0323	0.0100	-0.0122	0	0	0	
2	1	0.1109	-0.0016	0	0	0	0	
	2	0.0501	-0.0110	0	0	0	0	
	3	0.0193	-0.0005	0	0	0	0	
3	1	0.1449	-0.0741	0.2475	-0.0936	-0.3093	0	
	2	0.0699	-0.0134	-0.1259	0.1148	0.0199	0	
	3	0.0028	-0.0197	0.0020	0.0004	-0.0220	0	
4	1	0.1159	-0.0986	0.2561	0.6325	0	0	
	2	0.0330	-0.0392	-0.0365	-0.0050	0	0	
	3	0.0057	-0.0155	0.0128	0.0396	0	0	
5	1	0.0508	0.0829	0.2330	0.1957	0.1636	0.9425	
	2	-0.0302	-0.0228	0.1601	-0.0374	-0.0408	-0.0722	
	3	0.0069	0.0071	0.0304	0.0128	0.0039	0.0586	

4.4 Total Genotypic Effects

To generate the total additive genetic values of an animal for m traits, the genotypes at each locus must be known. Let the genotypes for an animal at the first five loci, be

$$\begin{pmatrix} 23 & 22 & 34 & 11 & 26 \end{pmatrix},$$

then the genotypic values at each locus (from Table 1) would be

$$\begin{pmatrix} -0.5479 & -0.0032 & +0.1539 & +0.2318 & +1.7715 \\ 0.3055 & -0.0220 & -0.0111 & +0.0660 & -0.0950 \\ -0.0022 & -0.0010 & +0.0024 & +0.0114 & +0.0657 \end{pmatrix},$$

or

$$\mathbf{BV}_i = \begin{pmatrix} 1.6061 \\ 0.2434 \\ 0.0763 \end{pmatrix},$$

where $\mathbf{BV}_i$ contains the breeding values of animal i obtained by summing the genotypic effects across all loci within the genome. In this case summation was just over the first five loci, not all 10,000. The genotypes for all 10,000 loci need to be known, as well as the allelic effects.

4.5 The Need for Scaling

The effects of linkage and drift after many historical generations and further generations of the base population is to change the $\mathbf{G}$ matrix values for population $\mathbf{S}_o$. After computing BV_i for all animals in $\mathbf{S}_o$ then $\mathbf{G}$ can be approximated by obtaining the variances and covariances among all of the BV_i across animals. Let the empirical matrix be denoted as $\mathbf{H}$.

Generally, the elements in $\mathbf{H}$ are smaller in magnitude than those in $\mathbf{G}$. This is due to a high degree of homozygosity of the genotypes of the animals in $\mathbf{S}_o$ due to linkage and drift. If $\mathbf{H}$ is desired to be close to $\mathbf{G}$, then the allelic effects need to be scaled upwards. The adjustment factor is obtained by averaging the ratio of diagonals of $\mathbf{G}$ to diagonals of $\mathbf{H}$, and then taking the square root of the average. The adjustment factor

may be in the range from 1 to 2. However, the adjustment factor will be the same for each replicate as long as the same protocol is followed with QMSim to give $\mathbf{S}_o$. Thus, the scaling factor only needs calculation once.

4.6 Matings of Two Individuals

To create a sperm or an egg, the first thing is to determine the number of recombination events per chromosome per meiosis. This was assumed to be either 1, 2, or 3. The next task is to determine where on the chromosome the recombinations occur. This needs to be computed for the sire and the dam separately. Then the haplotypes from each parent can be created and joined together in the new progeny. The process needs to be repeated for each progeny in the litter.

For recombinations, one could identify hot spots on specific chromosomes where more recombinations occur, if desired. For now, recombinations occur anywhere randomly.

5 Description of Traits

5.1 Growth

Growth is defined by a non-linear equation,

$$BW \;=\; A/(1 \;+\; B \cdot e^{-Cx}), \tag{1}$$

where BW is body weight at age t, and x is related to age, namely,

$$x \;=\; 2 \cdot (t - 103.5)/51.5.$$

The A, B, and C parameters of the equation are genetic traits which define the shape of an animal's growth from birth to death. See Schaeffer and De Lange (2013) for details about the nutritional aspects related to growth, feed intake, and protein and lipid deposition.

5.2 Feed Intake

Total feed intake is a linear function of growth, the equation is

$$FI \;=\; b_0 \;+\; b_1 \cdot (BW/160). \tag{2}$$

Phenotypically, what is needed is daily feed intake, or the amount eaten from day k to day $k+1$. Thus, the b_0 parameter drops out when differences are taken between two successive days. BW is divided by 160 to keep the numbers small.

The feed intake function could also have a quadratic parameter on BW squared, but a linear function is a very good approximation for the majority of pigs.

5.3 Lipid and Protein Deposition, Backfat

The amount of total lipid deposition, kg, is given by

$$L \;=\; c_1 \cdot FI, \tag{3}$$

a linear function of total feed intake.

Total protein deposition, kg, is a linear function of BW,

$$P \;=\; d_1 \cdot BW. \tag{4}$$

Both L and P can be converted to daily deposition rates, if desired. From total L and total P up to a certain body weight or age, can be used to derive backfat thickness, mm.

$$BF \;=\; -1.0 + 12.3 \cdot (L/P) + 0.13 \cdot P. \tag{5}$$

5.4 Survivability

Piglet survival in the first 5 days after birth is of great concern. Table 2 contains the cumulative probabilities of mortality in the first five days after birth that were used in the simulation program.

Table 4.
Mortality cumulative probabilities for piglets
in the first five days after birth.

Day	Probability
1	0.050
2	0.080
3	0.105
4	0.135
5	0.150

Let g be the genetic breeding value of an animal for survivability (from a mean of 0 and genetic variance of 10.0), and let r be a random residual value (from a mean of 0 and variance of 90.0) for that animal. To illustrate, let $g = -1.2$ and $r = 0.3$, then a risk ratio would be

$$Risk = (100 - (g + r) \cdot 0.1)/100.0, \tag{6}$$

where 0.1 is one divided by the phenotypic standard deviation. Then $Risk = 0.9991$. So if the average risk of death on day 1 for a piglet is 0.075, then for this pig it would be

$$Risk \cdot 0.075 = 0.0749325.$$

Choose a random uniform variate, and if the value is less than 0.0749325 the animal dies, otherwise it lives.

Beyond 5 days, survival follows a usual S-shaped survival curve, and you can determine the day the animal should die, if it were permitted to live a complete life. Voluntary culling can move that day forward, but not extend it. The day an animal should die is calculated as

$$((g + r) \cdot 0.1) \cdot 230 + 600,$$

which is 579.3 days of age using the same g and r as in equation (6). One should check that this age is greater than 1 and less than 1200. Thus, age at death is a normally distributed trait with mean of 600 d and a standard deviation of 230 d.

5.5 Other Traits

All of the other traits are represented by traits with means of 100 and phenotypic variance of 100. Thus, if the heritability is 0.04, then the genetic variance is 4, and the residual variance is 96. For a trait like age at first breeding, if an animal's genetic value comes out to be 1.3, for example, then this is converted to a risk value by adding the mean (100) plus 1.3 plus a residual component with variance 96, say -3.7. The result is

$$(\mu \ + \ ((1.3 \ - \ 3.7) \cdot 0.1) \cdot sd),$$

where sd is the true phenotypic standard deviation of the trait, and μ is the true phenotypic mean. Thus, if the mean, $\mu = 245$ days, and $sd = 2$, then for this animal its phenotypic age at first breeding would be

$$245 + ((1.3 \ - \ 3.7) \cdot 0.1) \cdot 2 \ = \ 244.52,$$

or 244.5 days.

Traits 7 to 22 work in the same way, as described above. The important thing is to get the heritabilities and genetic correlations as correct as possible, and the true phenotypic standard deviations and means, so that the traits are related correctly, and correspond to the usual range of values, in practice.

Simulation of all genetic traits as continuous, underlying normal distributions, allows for either continuous or discrete expression of phenotypes. Discrete traits require threshold values.

Table 5a.
Non-zero elements of genetic and residual correlation matrices (upper 6 × 6).

Trait 1	Trait 2	Genetic	Residual
Growth, A	Growth, B	0.644	0.708
Growth, A	Growth, C	-0.700	-0.540
Growth, A	Feed Intake	0.709	0.709
Growth, A	Lipid	0.185	0.178
Growth, A	Protein	0.099	0.103
Growth, A	Legs/Feet	0.130	0.130
Growth, A	Ham	0.410	0.410
Growth, A	Loin	0.050	0.050
Growth, A	Shoulder	0.240	0.240
Growth, A	Length	0.050	0.050
Growth, A	PRRS	0.200	0.200
Growth, B	Growth, C	-0.824	-0.686
Growth, B	Feed Intake	0.444	0.485
Growth, B	Lipid	-0.0073	0.0351
Growth, B	Protein	-0.014	-0.042
Growth, C	Feed Intake	-0.592	-0.581
Growth, C	Lipid	0.101	-0.180
Growth, C	Protein	-0.018	0.060
Feed Intake	Lipid	0.296	0.289
Feed Intake	Protein	0.139	0.157
Feed Intake	Loin	0.110	0.110
Feed Intake	Shoulder	0.269	0.270
Feed Intake	Length	0.110	0.110
Feed Intake	PRRS	0.100	0.100
Lipid	Protein	-0.041	-0.161

Table 5b.
Non-zero elements of genetic and residual correlation matrices.

Trait 1	Trait 2	Genetic	Residual
Legs/Feet	Ham	0.368	0.300
Legs/Feet	Loin	0.259	0.250
Legs/Feet	Shoulder	0.284	0.290
Legs/Feet	Length	-0.127	-0.129
Legs/Feet	Stature	-0.229	-0.230
Legs/Feet	Survivability	0.100	0.100
Ham	Loin	0.350	0.344
Ham	Shoulder	0.070	0.069
Ham	Length	0.343	0.300
Ham	Stature	-0.140	-0.132
Loin	Shoulder	0.080	0.079
Loin	Length	0.100	0.099
Loin	Stature	-0.366	-0.300
Loin	Survivability	-0.269	-0.270
Shoulder	Length	0.079	0.080
Shoulder	Stature	-0.100	-0.100

Table 5c.
Non-zero elements of genetic and residual correlation matrices.

Trait 1	Trait 2	Genetic	Residual
Length	Length	1.000	1.000
Length	Stature	-0.140	-0.140
Length	Survivability	-0.250	-0.246
Conception	PRRS	0.130	0.150
Litter size	PRRS	0.100	0.100
Survivability	PRRS	0.100	0.100
Stillbirths	PRRS	0.100	0.100

6 Part III. Simulation of Herd Activities

The simulation procedures follow similarly to those of Thomas et al. (2013) who studied economic weights for traits in dairy cattle. However, their simulation did not involve genomics.

6.1 Industry Structure

The pyramid structure (Figure 1) exists in the Canadian swine industry. There are nucleus herds which are pure-breed operations in which animals are selected for maternal or paternal traits. Nucleus herds generally have very high health standards and as such are almost disease free. Genetic selection and breeding occurs in the nucleus herds.

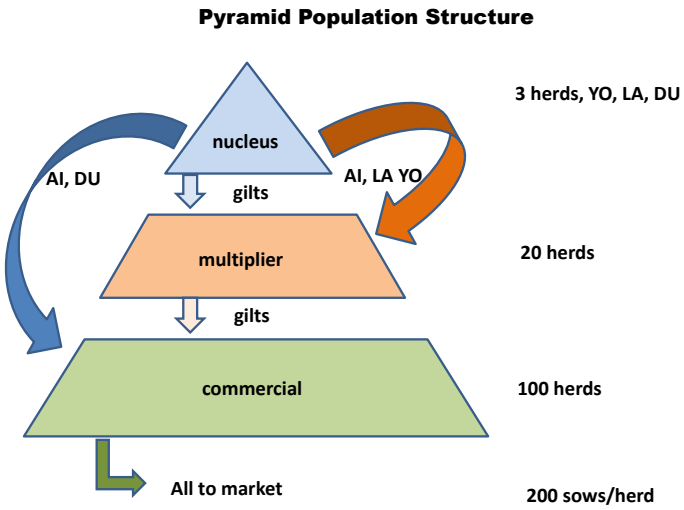
Multiplier herds are produce crossbred gilts for commercial herds. Multiplier herds also have high health standards and were considered disease free in this simulation.

Commercial herds buy crossbred gilts that are mated to boars of a terminal sire line to produce all three-way crossbred progeny for market. Because pigs are constantly coming and going from these herds, most diseases appear in commercial herd settings. This is where the PRRS virus phenotypes are observed. The PRRS phenotypes are used to evaluate the animals in the multiplier and nucleus levels through pedigree

relationships.

The concept is to have three types of herds, namely nucleus, multiplier, or commercial. Herd size was set at 200 sows for all herds. There were 3 nucleus herds, 20 multiplier herds, and 100 commercial herds, which gives 24,600 sows. Nucleus herds supply animals to multiplier herds, and multiplier herds supply gilts to the commercial herds. All breedings were assumed to be made by artificial insemination.

Figure 1



Pools of replacement gilts and boars were maintained for the nucleus and multiplier levels. A separate pool of replacement crossbred gilts were kept for the commercial herds. All animals in the pools were indexed and ranked. If an animal was not selected as a replacement it was slaughtered at 240 days of age. All piglets from the commercial level were sold to market at the appropriate slaughter weights.

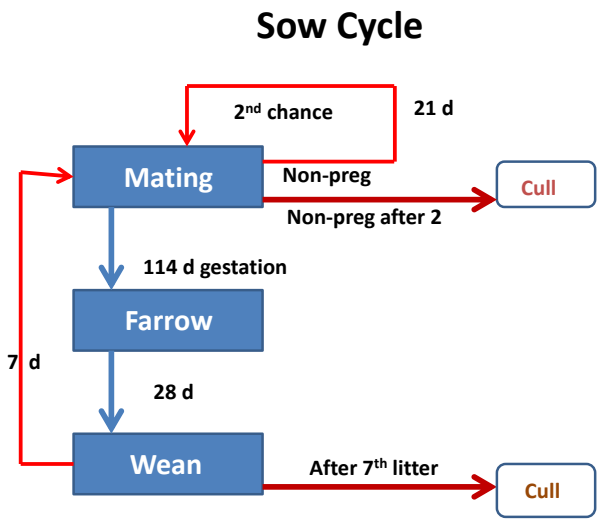
Facilities were assumed to be state of the art and sufficient for each level of production. Management was assumed to make the correct decisions regarding animal movement, housing, and feeding and care of the animals.

6.2 Time Flow Issues

The simulation goes through the daily functions of each herd. Every live sow in each herd is checked daily to determine if it is time to farrow, to wean the litter, or to be bred, and if it should be culled (at weaning of a litter).

Sows go through a 142 d farrowing cycle (Figure 2). After they wean their 7^{th} litter they are culled. Otherwise they die of natural causes or are culled for not conceiving. They were also culled at the commercial level if they ever suffered from PRRS.

Figure 2



After a sow weans its 7^{th} litter it is culled and replaced. If a sow is bred and does not conceive, then it is allowed to be bred a second time at the next estrous cycle. If the animal fails to conceive on the second breeding, then it is culled. Gestation lengths are genetic and randomly determined for each sow. The estrous cycle is fixed at 21 days, and all sows are mated 7 days after weaning their litter, which occurs 28 days after farrowing.

Pigs that are grown for market purposes (Figure 3) are weaned 28 days after birth, at a weight of around 7 kg. At 50 days they go into a grower barn where they are kept until they reach market weight around 120 to 140 kg with about 15 mm of backfat. Pigs that are selected as potential replacements go into the Gilt or Boar pools.

7 Selection Index

Replacement gilts and boars were ranked on the basis of a selection index. The base index equation is

$$\begin{aligned} IND &= 0.32(LS) + 0.05(Surv) + 0.12(Teats) + 0.13(Loin) \\ &\quad + 0.11(A) + 0.03(CR) - 0.22(FI) + 0.02(P) \\ &\quad + 0(PRRS) \end{aligned}$$

where the trait EBVs are standardized by dividing by their genetic standard deviations, and LS is litter size, $Surv$ is survivability, $Teats$ is number of teats, $Loin$ is loin thickness, A is the A parameter of the growth function, CR is conception rate, FI is feed intake, P is protein yield, and $PRRS$ is an evaluation for PRRS resistance. The weight on $PRRS$ and the other traits varies depending on the scenario being considered. When EBVs for PRRS are calculated, then the index weights are changed as shown in Table 6. The amount of weight on the PRRS trait ranges from 0 to 5%, to 10%, and to 20%.

Table 6.
Index weights for different emphasis on the PRRS EBVs.

Trait	IND	.05 $PRRS$	.10 $PRRS$	.20 $PRRS$
Litter Size	0.32	0.3040	0.2880	0.2560
Survival	0.05	0.0475	0.0450	0.0400
No. Teats	0.12	0.1140	0.1080	0.0960
Loin	0.13	0.1235	0.1170	0.1040
Growth	0.11	0.1045	0.0990	0.0880
Conception Rate	0.03	0.0285	0.0270	0.0240
Feed Intake	-0.22	-0.2090	-0.1980	-0.1760
Protein	0.02	0.0190	0.0180	0.0160
PRRS	0.00	0.0500	0.1000	0.2000

When genomics are used, EBVs are replaced by GEBVs representing a blending of EBVs and GEBVs to obtain more accuracy than regular

EBVs. In the case of using a single QTL, in a QTL assisted selection strategy, then only the EBV for PRRS is enhanced in accuracy. Similarly for the case of MQAS when the top 50 QTL for PRRS are used, only the accuracy of the EBV for PRRS is enhanced. When all 10,000 QTLs are used, then accuracies of all trait EBVs are enhanced.

The comparison statistics are the genetic trends in the nucleus breeds for the PRRS trait, and the change in phenotypic results at the commercial level, in terms of number of affected animals and deaths.

8 Validation Results

Simulation programs need to be validated and sensitivity studies should be run to determine the effects of errors in parameters or assumptions. One of the selection strategies for the study was to select replacement animals on the basis of the selection index, ignoring everything about the PRRS virus. The validation comes in terms of plots of genetic trends (in genetic standard deviation units) of all traits in the selection index, including PRRS, depicted in Figures 3 to 11 for sows in the nucleus herds. Given the relative weights in the index and the heritabilities of the traits the observed genetic trends appear to be as expected over 25 sow cycles. A sow cycle was 142 days from farrowing to farrowing. Twenty five cycles is 3550 days or about 9.7 years.

Figure 3

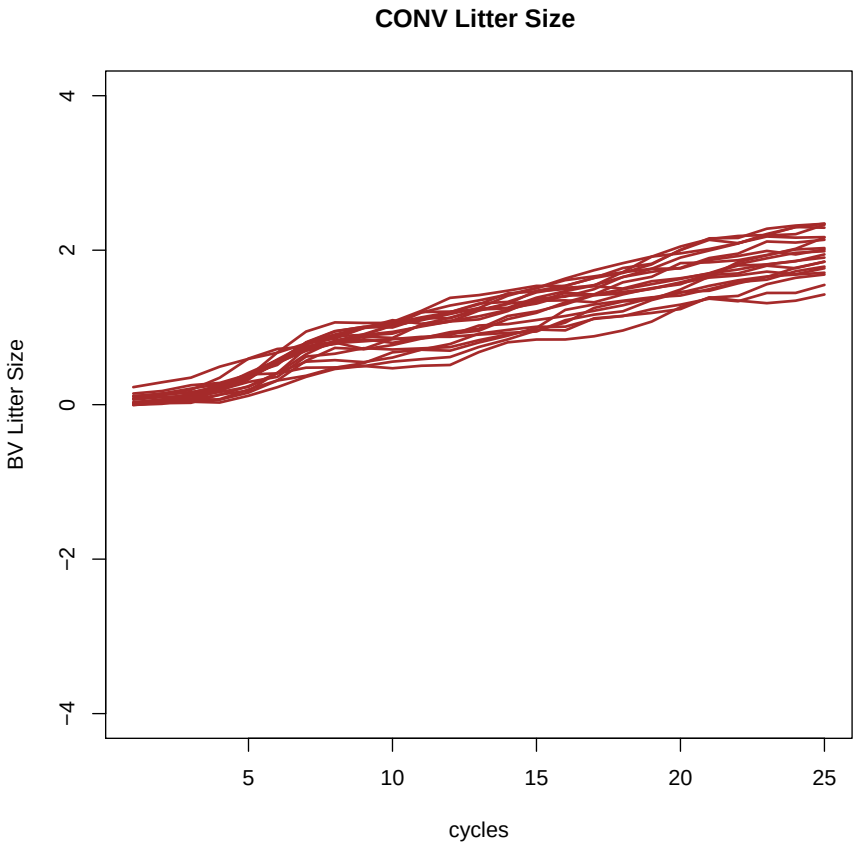


Figure 4

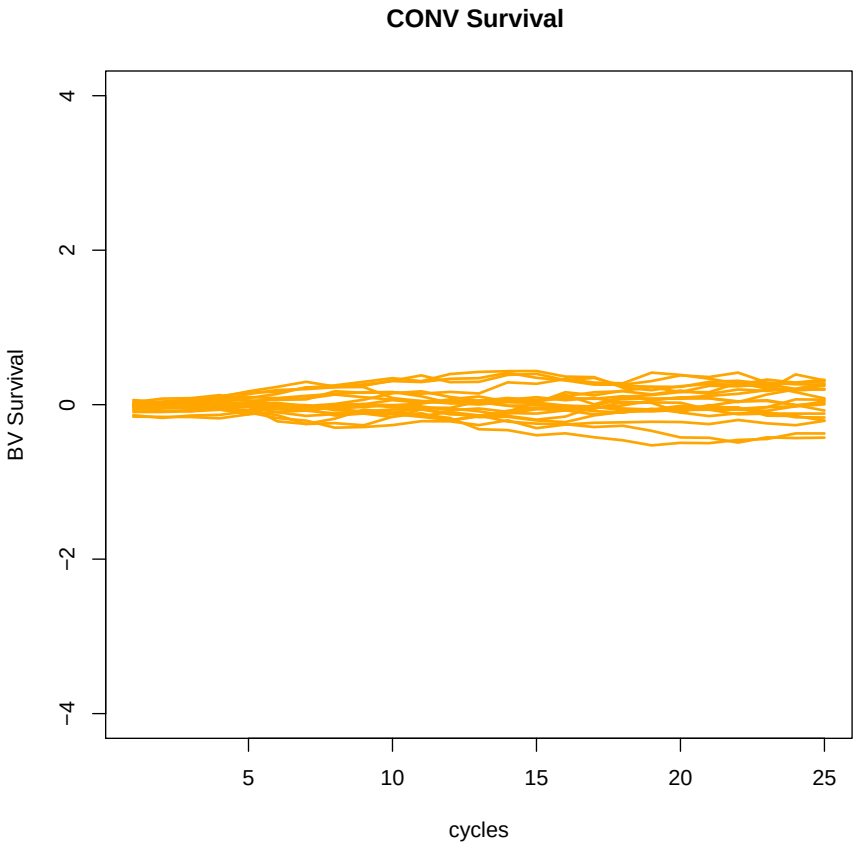


Figure 5

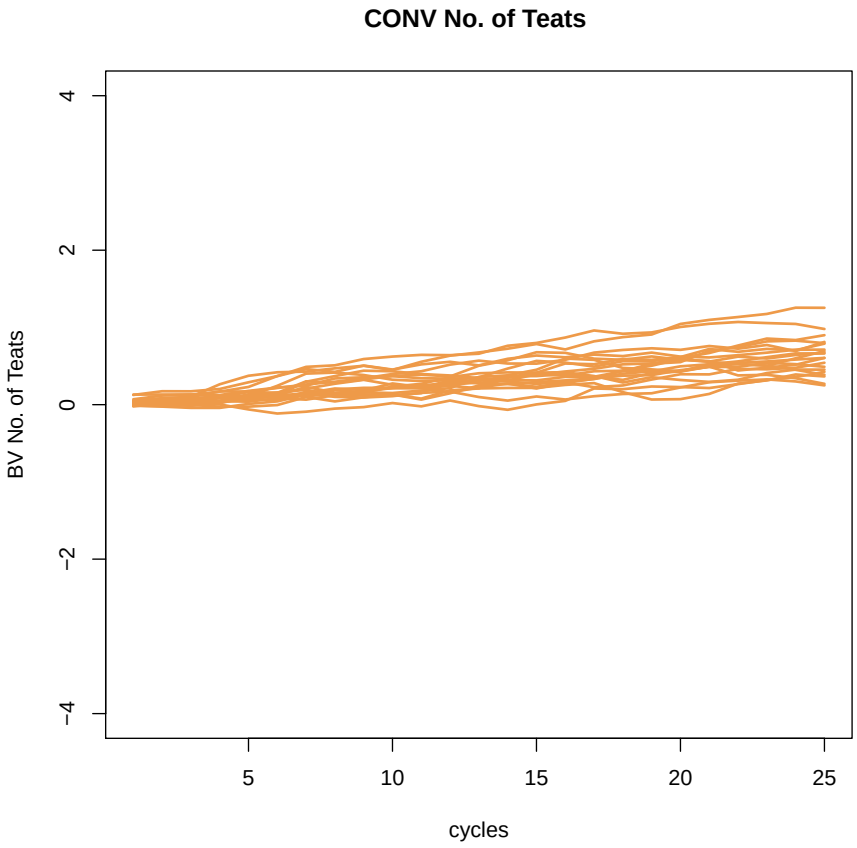


Figure 6

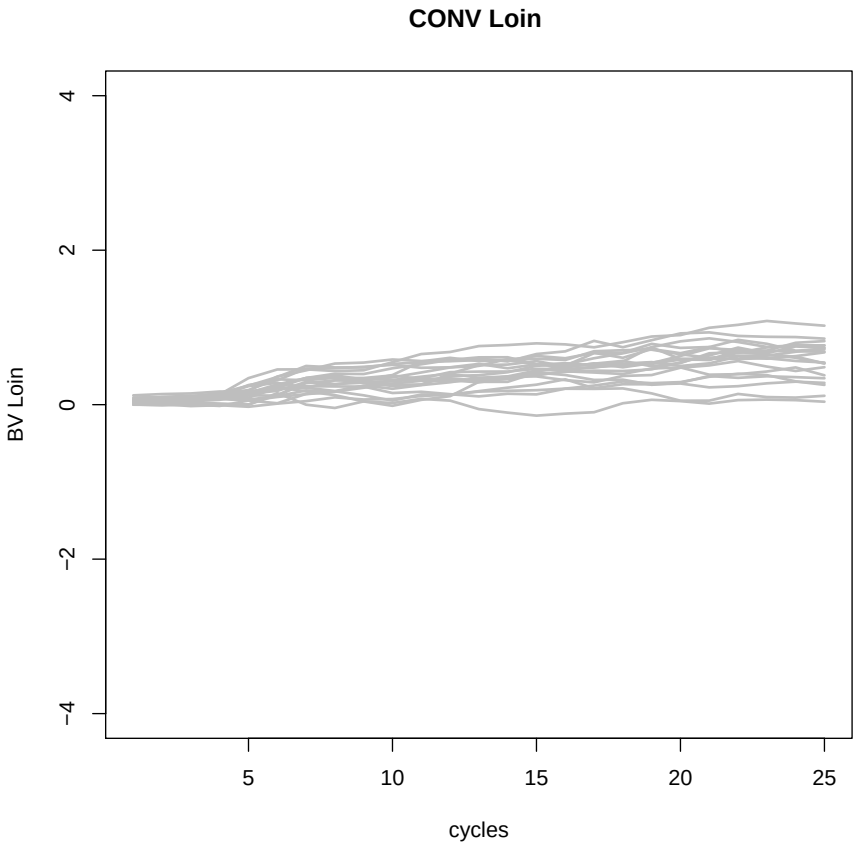


Figure 7

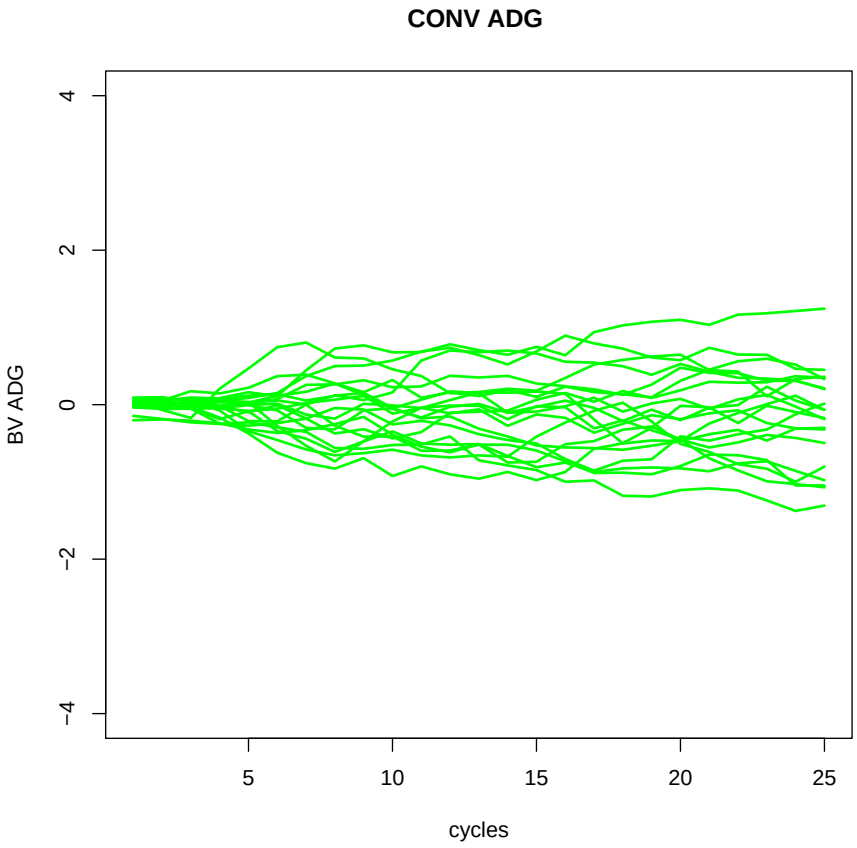


Figure 8

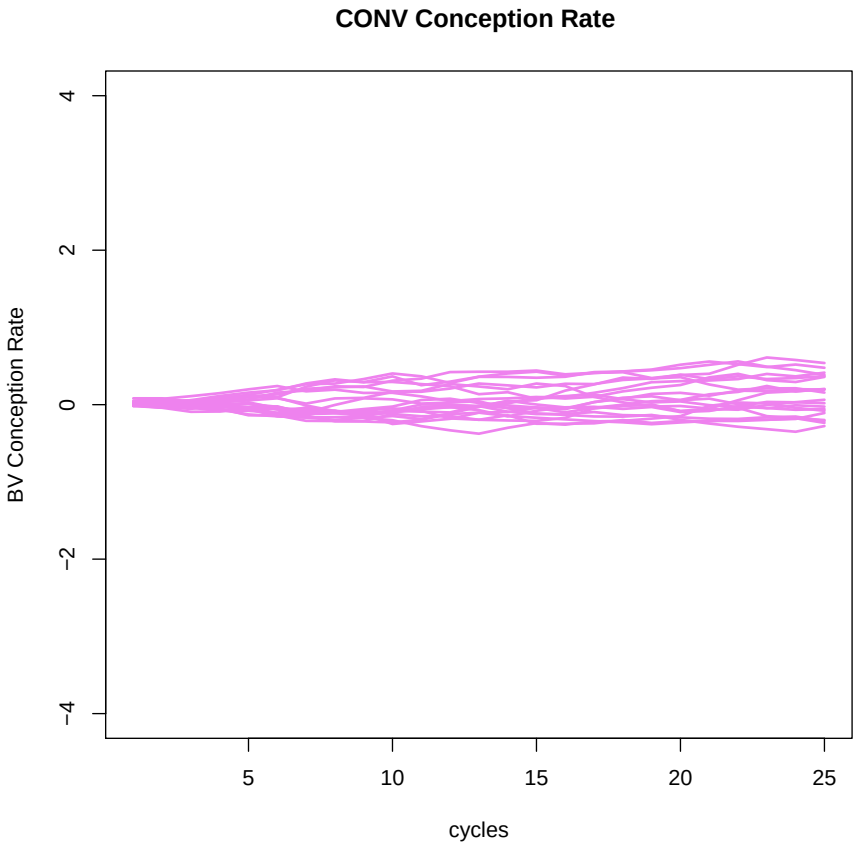

CONV Conception Rate
BV Conception Rate
4
2
0
−2
−4
5
10
15
20
25
cycles

Figure 9

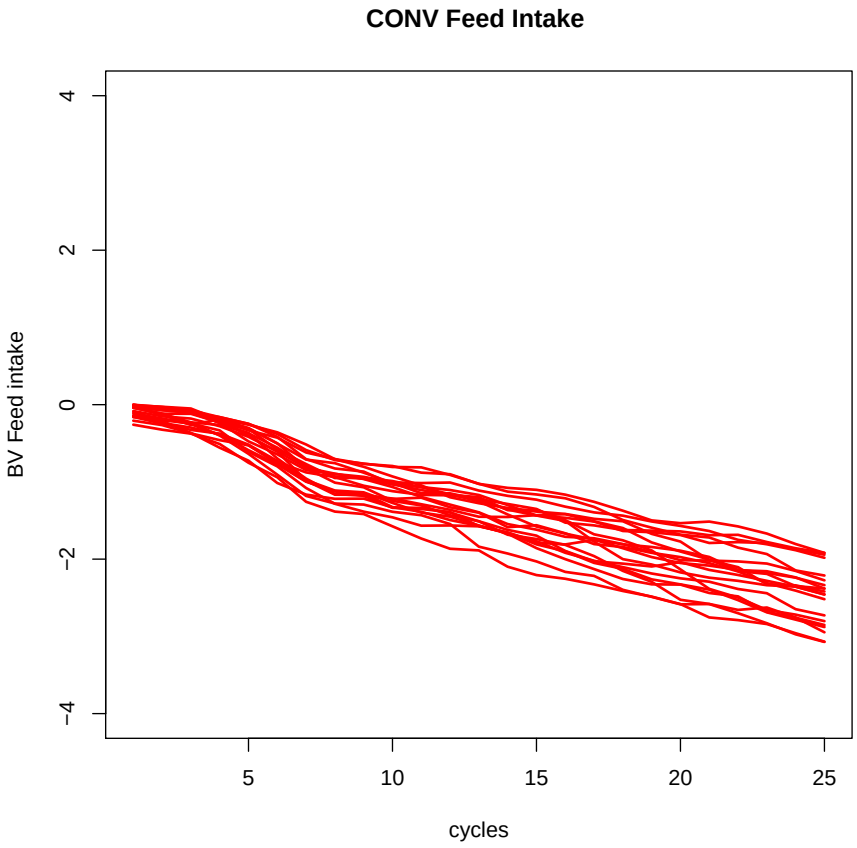


Figure 10

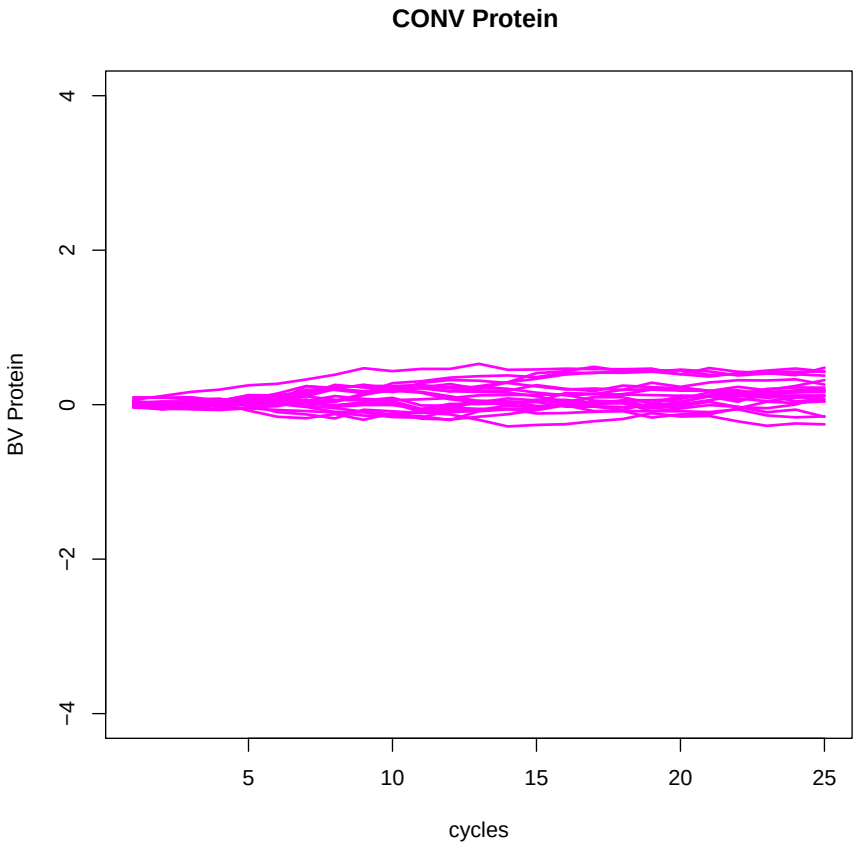
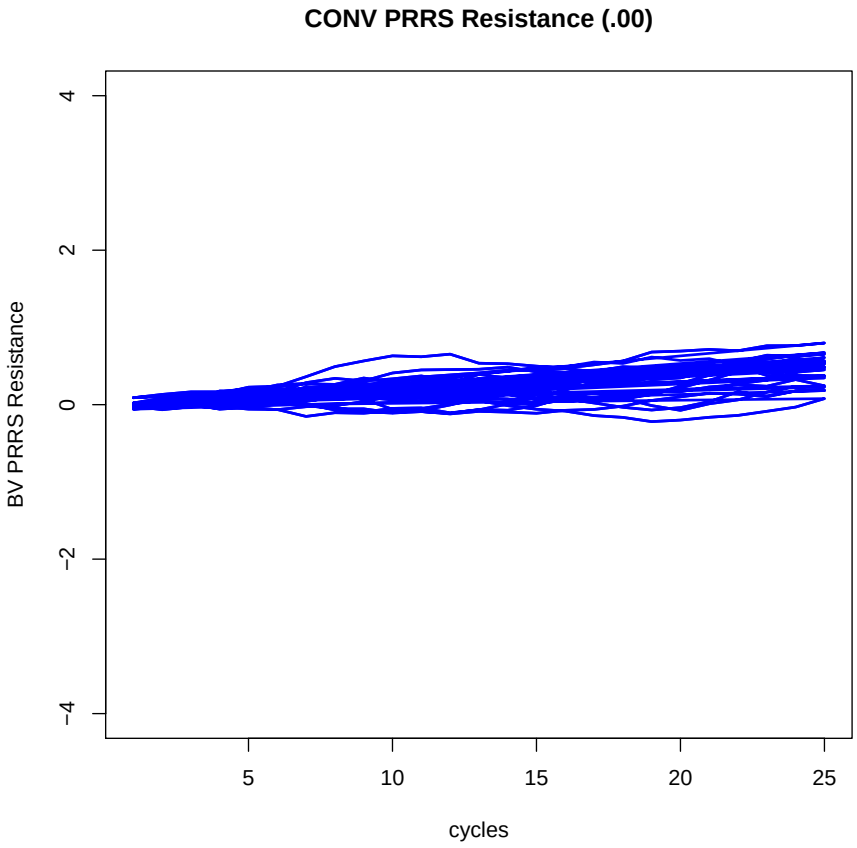


Figure 11



As far as the PRRS virus is concerned, the frequency of pigs infected, pigs that die, sows that are infected, and numbers of mummified fetuses at the commercial level, on a phenotypic basis would be a good measure of a genetic selection strategy against the virus. Plots of those trends are in Figures 12 to 15. These are also over 25 sow cycles. As expected there were little trends over time, except due to correlated responses to selection on the other traits in the index.

Figure 12

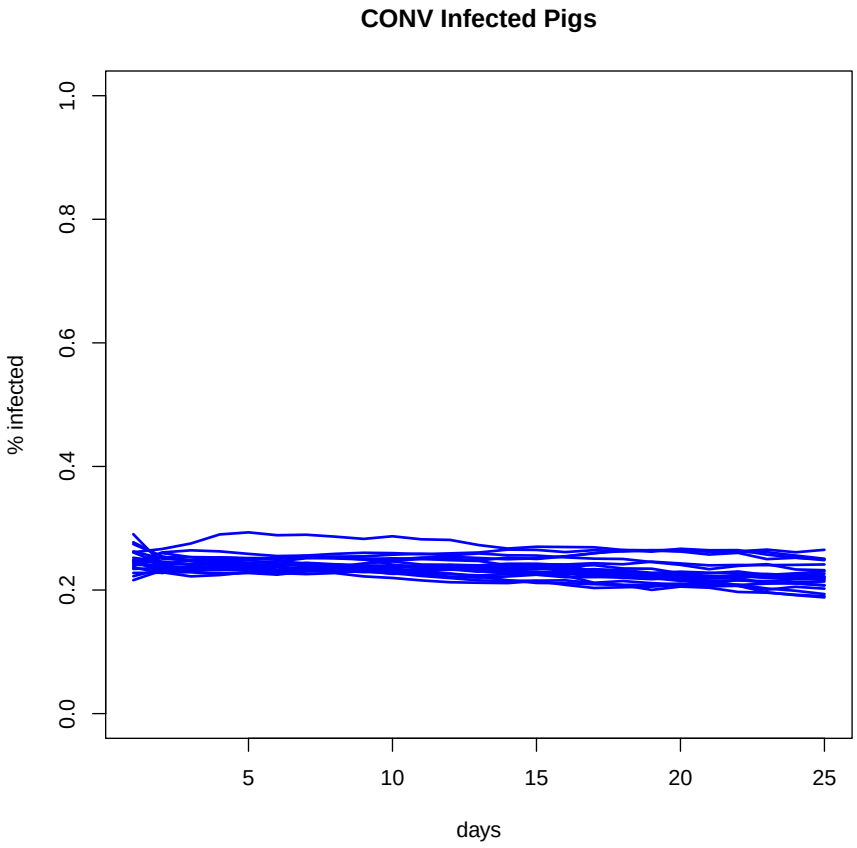


Figure 13

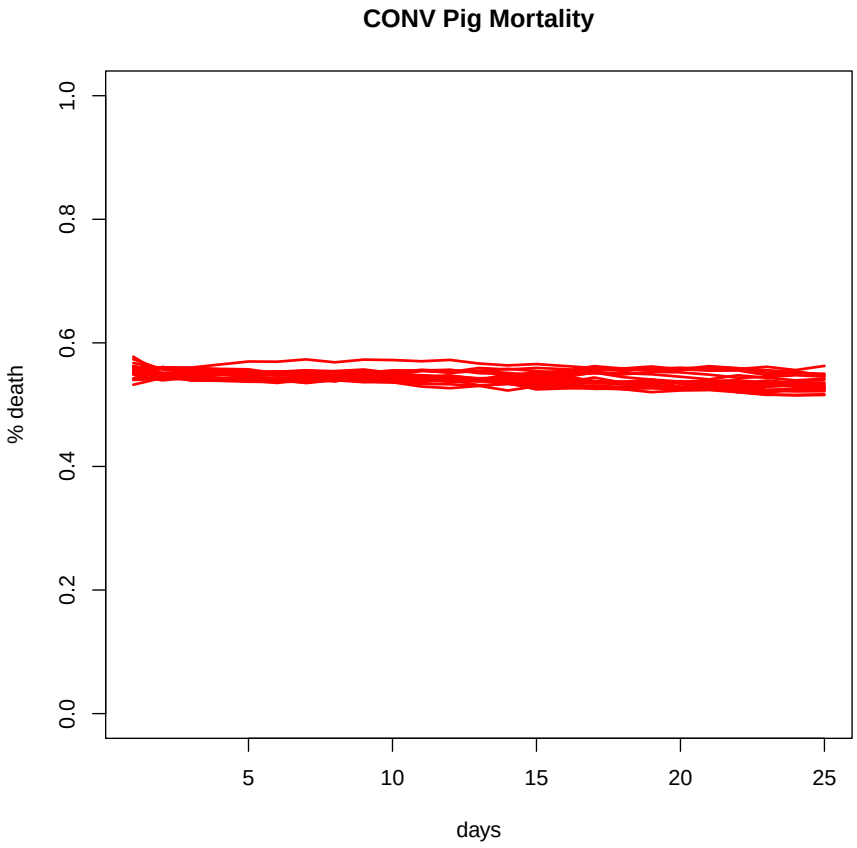


Figure 14

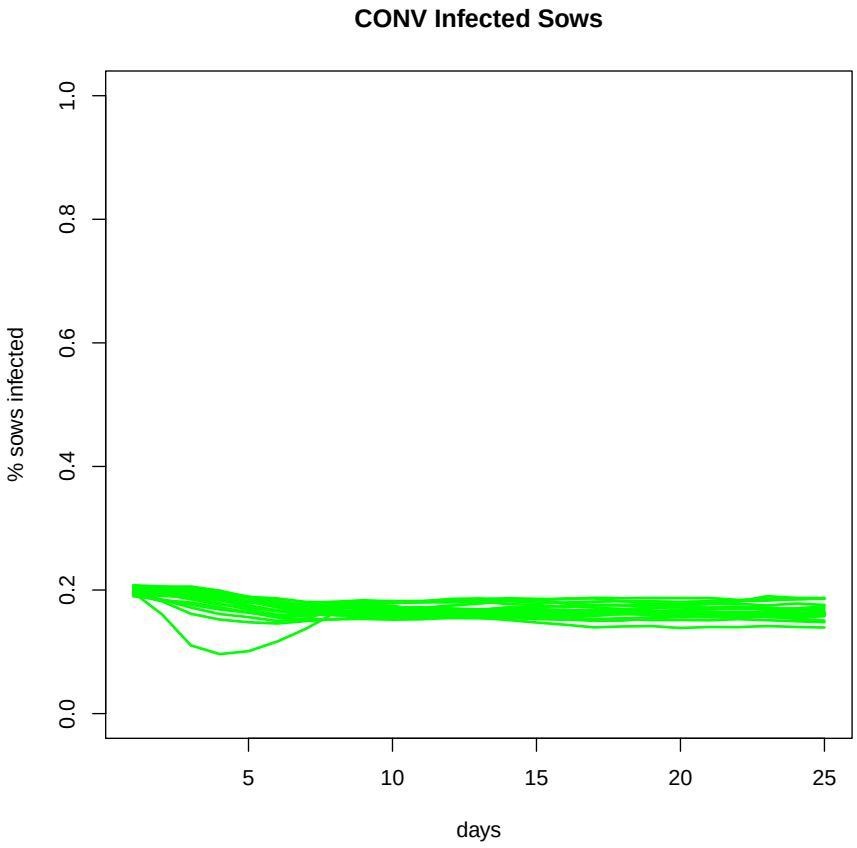
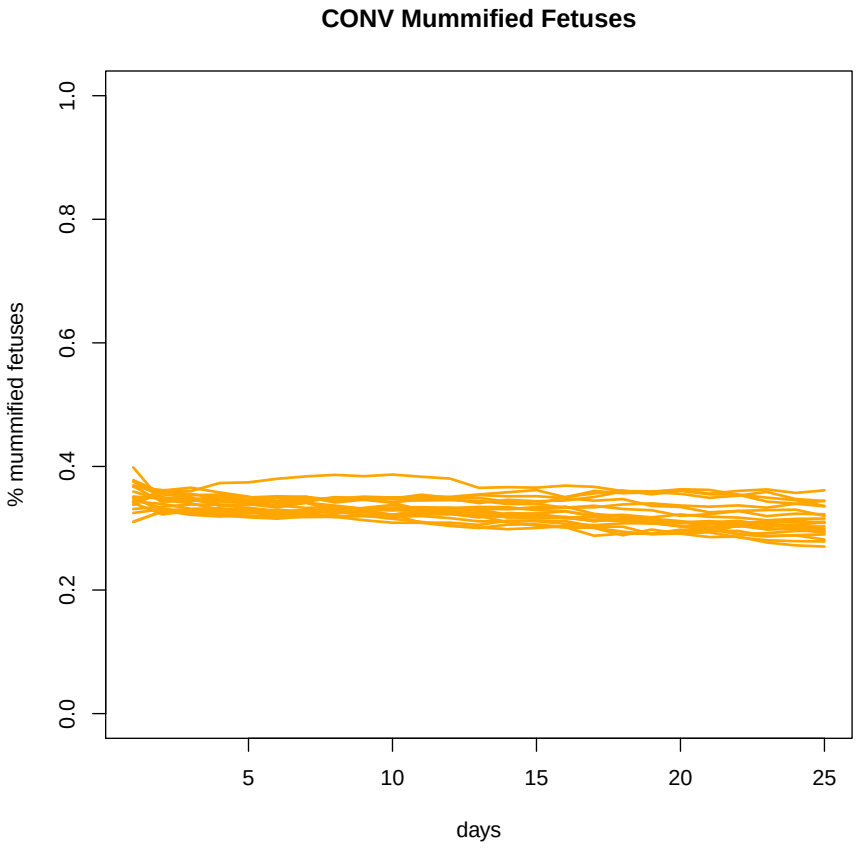

CONV Infected Sows
% sows infected
0.0
0.2
0.4
0.6
0.8
1.0
5
10
15
20
25
days

Figure 15



9 References

Alsop, J. E. , J. B. Rafuse, I. R. Keith, W. Koopman, A. W. Jalvingh, I. R. Dohoo. 1994. Implementation of a computer simulation swine tactical decision, support model in Prince Edward Island. Can. Vet. J. 35:153-157.

Lillehammer, M. , T. H. E. Meuwissen, A. K. Sonesson. 2013. Genomic selection for two traits in a maternal pig breeding scheme. J. Anim. Sci. 91:3079-3087.

Pla-Aragones, L. M. , V. F. Marias, S. V. Rodriguez-Sanchez. 2008. A simulation model for intensive piglet production systems. Proc. of Winter Simulation Conference. IEEE.

Pomar, C. , D. L. Harris, P. Savoie, F. Minvielle. 1991. Computer simulation model of swine production systems: III. A dynamic herd simulation model including reproduction. J. Anim. Sci. 69:2822-2836.

Sargolzaei, M. , F. S. Schenkel. 2009. QMSim: a large-scale genome simulator for livestock. Bioinformatics, 25: 680-681. First published January 28, 2009, doi:10.1093 /bioinformatics/btp045.

Schaeffer, L. R. 2013. Making covariance matrices positive definite. *Elares:44*.

Schaeffer, L. R. , C. F. M. de Lange. 2013. Method to partition growth of pigs on test into protein and lipid contributions. *Elares:52*.

Thomas, A. , V. Osborne, L. R. Schaeffer. 2013. Economic weights for disease traits in dairy cattle. *Elares:43*.

Tribout, T. , C. Larzul, F. Phocas. 2013. Economic aspects of implementing genomic evaluations in a pig sire line breeding scheme. Genet. Sel. Evol. 45:40.

van Steenbergen, E. J. 1990. Relevance of exterior appraisal in pig breeding. Doctoral thesis. Department of Animal Breeding, Wageningen Agricultural University.

10 Appendix I

QMSim parameter file

```
/*******************************
 **     Global parameters     **
 *******************************/
```

```
title = "10k QTL - Swine";
nrep  = 1;                    //Number of replicates
h2    = 0.2;                  //Heritability
qtlh2 = 0.2;                  //QTL heritability
phvar = 1.0;                  //Phenotypic variance

/*******************************
 **   Historical population   **
 *******************************/
begin_hp;
   hg_size = 500  [0]
             500  [2950]
             5100  [3000];   //Size of the historical generations
   nmlhg   = 2000;           //Number of males in the last generation
end_hp;

/*******************************
 **        Populations        **
 *******************************/
begin_pop = "A_base";
   begin_founder;
      male   [n =  200, pop = "hp"];
      female [n = 2000, pop = "hp"];
   end_founder;
   ls  = 12 11[0.35] 10[0.25] 9[0.15] 8[0.05];  //Litter size
   pmp = 0.50;                //Proportion of male progeny
   sr  = 0.50;                //sire replacement rate
   dr  = 0.35;                //dam replacement rate
   ng  = 50;                  //Number of generations
   sd=rnd;                   //the type of selection design
end_pop;

begin_pop = "B_base";
   begin_founder;
      male   [n = 200, pop = "hp"];
      female [n = 2000, pop = "hp"];
   end_founder;
   ls  = 12 11[0.35] 10[0.25] 9[0.1] 8[0.05];  //Litter size
```

```
   pmp = 0.50;             //Proportion of male progeny
   sr  = 0.50;             //sire replacement rate
   dr  = 0.35;             //dam replacement rate
   ng  = 50;               //Number of generations
   sd=rnd;                 //the type of selection design
end_pop;

begin_pop = "C_base";
   begin_founder;
      male   [n = 200, pop = "hp"];
      female [n = 2000, pop = "hp"];
   end_founder;
   ls  = 12 11[0.35] 10[0.25] 9[0.1] 8[0.05];  //Litter size
   pmp = 0.50;           //Proportion of male progeny
   sr  = 0.50;           //sire replacement rate
   dr  = 0.35;           //dam replacement rate
   ng  = 50;             //Number of generations
   sd=rnd;               //the type of selection design
end_pop;

begin_pop = "A_pop";
   begin_founder;
      male   [n =  200, pop = "A_base", gen=50, select=rnd];
      female [n = 2000, pop = "A_base", gen=50, select=rnd];
   end_founder;
   ls  = 12 11[0.35] 10[0.25] 9[0.1] 8[0.05];  //Litter size
   pmp = 0.50;          //Proportion of male progeny
   sr  = 0.50;          //sire replacement rate
   dr  = 0.35;          //dam replacement rate
   ng  = 10;            //Number of generations
   sd=phen;             //the type of selection design
   begin_popoutput;
        data;
        stat;
        ld /maft 0.01 /chr 1 /gen 0 10;
        allele_freq;
        genotype /snp_code /gen 10;
   end_popoutput;
```

```
end_pop;

begin_pop = "B_pop";
   begin_founder;
      male   [n =  200, pop = "B_base", gen=50, select=rnd];
      female [n = 2000, pop = "B_base", gen=50, select=rnd];
   end_founder;
   ls  = 12 11[0.35] 10[0.25] 9[0.1] 8[0.05];  //Litter size
   pmp = 0.50;          //Proportion of male progeny
   sr  = 0.50;          //sire replacement rate
   dr  = 0.35;          //dam replacement rate
   ng  = 10;            //Number of generations
   sd=phen;             //the type of selection design
   begin_popoutput;
        data;
        stat;
        ld /maft 0.01 /chr 1 /gen 0 10;
        allele_freq;
        genotype /snp_code /gen 10;
   end_popoutput;
end_pop;

begin_pop = "C_pop";
   begin_founder;
      male   [n =  200, pop = "C_base", gen=50, select=rnd];
      female [n = 2000, pop = "C_base", gen=50, select=rnd];
   end_founder;
   ls  = 12 11[0.35] 10[0.25] 9[0.1] 8[0.05];  //Litter size
   pmp = 0.50;          //Proportion of male progeny
   sr  = 0.50;          //sire replacement rate
   dr  = 0.35;          //dam replacement rate
   ng  = 10;            //Number of generations
   sd=phen;             //the type of selection design
   begin_popoutput;
        data;
        stat;
        ld /maft 0.01 /chr 1 /gen 0 10;
        allele_freq;
```

```
        genotype /snp_code /gen 10;
   end_popoutput;
end_pop;

begin_pop = "cross_CB";
   begin_founder;
      male   [n =  200, pop = "C_pop", gen=10, select=rnd];
      female [n = 2000, pop = "B_pop", gen=10, select=rnd];
   end_founder;
   ls  = 12 11[0.35] 10[0.25] 9[0.1] 8[0.05];  //Litter size
   pmp = 0.50;          //Proportion of male progeny
   sr  = 0.50;          //sire replacement rate
   dr  = 0.35;          //dam replacement rate
   ng  = 1;             //Number of generations
   sd=rnd;              //the type of selection design
   begin_popoutput;
        data;
        stat;
        ld /maft 0.01 /chr 1 /gen 0 1;
        allele_freq;
        genotype /snp_code /gen 1;
   end_popoutput;
end_pop;

begin_pop = "cross_BC";
   begin_founder;
      male   [n =  200, pop = "B_pop", gen=10, select=rnd];
      female [n = 2000, pop = "C_pop", gen=10, select=rnd];
   end_founder;
   ls  = 12 11[0.35] 10[0.25] 9[0.1] 8[0.05];  //Litter size
   pmp = 0.50;         //Proportion of male progeny
   sr  = 0.50;         //sire replacement rate
   dr  = 0.35;         //dam replacement rate
   ng  = 1;            //Number of generations
   sd=rnd;             //the type of selection design
   begin_popoutput;
        data;
        stat;
```

```
        ld /maft 0.01 /chr 1 /gen 0 1;
        allele_freq;
        genotype /snp_code /gen 1;
   end_popoutput;
end_pop;

/*******************************
 **          Genome           **
 *******************************/
begin_genome;
   begin_chr = 18;
      chrlen = 100;             //Chromosome length
      nmloci = 1000;            //Number of markers
      mpos   = even;            //Marker positions
      nma    = all 2;           //Number of marker alleles
      maf    = eql;             //Marker allele frequencies
      nqloci = 4000;            //Number of QTL
      qpos   = rnd;             //QTL positions
      nqa    = rnd 2 3 4 5 6 7; //Number of QTL alleles
      qaf    = rnd;             //QTL allele frequencies
      qae    = rndn;            //QTL allele effects
   end_chr;
   mmutr = 1e-4 /recurrent;    //marker mutation rate
   qmutr = 1e-4 /recurrent;    //QTL mutation rate
   select_seg_loci /maft 0.01 /nmrk 5000 /nqtl 10000;
   r_mpos_g;
   r_qpos_g;
   interference = 25;
end_genome;

/*******************************
 **       Output options      **
 *******************************/
begin_output;
   linkage_map;
end_output;
```