2-2014

Validation and Further Characterization of a Major Quantitative Trait Locus Associated with Host Response to Experimental Infection with Porcine Reproductive and Respiratory Syndrome Virus

Nicholas James Boddicker
Iowa State University

Dorian J. Garrick
Iowa State University, dorian@iastate.edu

Raymond Rowland
Kansas State University

Joan K. Lunney
United States Department of Agriculture

Jack C. M. Reecy
Iowa State University, jreecy@iastate.edu

See next page for additional authors

Validation and Further Characterization of a Major Quantitative Trait Locus Associated with Host Response to Experimental Infection with Porcine Reproductive and Respiratory Syndrome Virus

Abstract
Infectious diseases are costly to the swine industry; porcine reproductive and respiratory syndrome (PRRS) is the most devastating. In earlier work, a quantitative trait locus associated with resistance/susceptibility to PRRS virus was identified on *Sus scrofa* chromosome 4 using approximately 560 experimentally infected animals from a commercial cross. The favorable genotype was associated with decreased virus load and increased weight gain (WG). The objective here was to validate and further characterize the association of the chromosome 4 region with PRRS resistance using data from two unrelated commercial crossbred populations. The validation populations consisted of two trials each of approximately 200 pigs sourced from different breeding companies that were infected with PRRS virus and followed for 42 days post-infection. Across all five trials, heritability estimates were 0.39 and 0.34 for viral load (VL; area under the curve of log-transformed viremia from 0 to 21 days post-infection) and WG to 42 days post-infection respectively. Effect estimates of SNP WUR10000125 in the chromosome 4 region were in the same directions and of similar magnitudes in the two new trials as had been observed in the first three trials. Across all five trials, the 1-Mb region on chromosome 4 explained 15 percent of genetic variance for VL and 11 percent for WG. The effect of the favorable minor allele at SNP WUR10000125 was dominant. Ordered genotypes for SNP WUR10000125 showed that the effect was present irrespective of whether the favorable allele was paternally or maternally inherited. These results demonstrate that selection for host response to PRRS virus infection could reduce the economic impact of PRRS.

Keywords
genome wide association, porcine reproductive and respiratory syndrome, swine chromosome 4, validation

Comments
This article is from *Animal Genetics* 4 5 (2014): 48, doi:10.1111/age.12079.

Authors
Nicholas James Boddicker, Dorian J. Garrick, Raymond Rowland, Joan K. Lunney, James M. Reecy, and Jack C. Dekkers

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi: 10.1111/age.12079

Validation and further characterization of a major quantitative trait locus associated with host response to experimental infection with porcine reproductive and respiratory syndrome virus

N. J. Boddicker*, D. J. Garrick*†, R. R. R. Rowland‡, J. K. Lunney§, J. M. Reecy* and J. C. M. Dekkers*

*Department of Animal Science, Iowa State University, Ames, IA, 50011, USA. †Institute of Veterinary, Animal and Biomedical Sciences, Massey University, Palmerston North 4474, New Zealand. ‡College of Veterinary Medicine, Kansas State University, Manhattan, KS, 66506, USA. §United State Department of Agriculture, Agricultural Research Services, Beltsville Agricultural Research Center, Beltsville, MD, 20705, USA.

Summary

Infectious diseases are costly to the swine industry; porcine reproductive and respiratory syndrome (PRRS) is the most devastating. In earlier work, a quantitative trait locus associated with resistance/susceptibility to PRRS virus was identified on *Sus scrofa* chromosome 4 using approximately 560 experimentally infected animals from a commercial cross. The favorable genotype was associated with decreased virus load and increased weight gain (WG). The objective here was to validate and further characterize the association of the chromosome 4 region with PRRS resistance using data from two unrelated commercial crossbred populations. The validation populations consisted of two trials each of approximately 200 pigs sourced from different breeding companies that were infected with PRRS virus and followed for 42 days post-infection. Across all five trials, heritability estimates were 0.39 and 0.34 for viral load (VL; area under the curve of log-transformed viremia from 0 to 21 days post-infection) and WG to 42 days post-infection respectively. Effect estimates of SNP WUR10000125 in the chromosome 4 region were in the same directions and of similar magnitudes in the two new trials as had been observed in the first three trials. Across all five trials, the 1-Mb region on chromosome 4 explained 15 percent of genetic variance for VL and 11 percent for WG. The effect of the favorable minor allele at SNP WUR10000125 was dominant. Ordered genotypes for SNP WUR10000125 showed that the effect was present irrespective of whether the favorable allele was paternally or maternally inherited. These results demonstrate that selection for host response to PRRS virus infection could reduce the economic impact of PRRS.

Keywords genome wide association, porcine reproductive and respiratory syndrome, swine chromosome 4, validation

Introduction

Porcine reproductive and respiratory syndrome (PRRS) causes reproductive failure in breeding females and respiratory problems in growing animals and collectively costs the US swine industry approximately $664 million annually, of which $365 million is attributed to losses among growing pigs (Holtkamp *et al.* 2013). The PRRS virus (PRRSV) is complex, and many different strains circulate around the world (Fang *et al.* 2007). The PRRS Host Genetics Consortium was established to address this important industry issue by investigating the genetic basis of host response to PRRSV infection in experimentally infected commercial crossbred pigs using PRRSV strain NVSL 97-7985 (Lunney *et al.* 2011; Rowland *et al.* 2012). Previous work, using data from the first three infection trials of approximately 560 pigs, all from the same commercial F1 cross, identified a major quantitative trait locus (QTL) for PRRS resistance/susceptibility on *Sus scrofa* chromosome (SSC) 4 (Boddicker *et al.* 2012). The identified 1-Mb region was associated with viral load (VL), a measure of viremia over 21 days post-infection (dpi), and WG to 42 dpi (WG42). Pigs with the favorable SNP WUR10000125

Address for correspondence

J. C. M. Dekkers, Department of Animal Science, Iowa State University, Ames, IA 50011, USA.
E-mail: jdekkers@iastate.edu

Accepted for publication 24 June 2013

genotypes in the region had reduced VL and increased WG compared with pigs with the unfavorable genotypes. This region explained 15 percent of the variance in genomic estimated breeding values (GEBV) for VL and 11 percent for WG42. The correlation between the GEBV for VL and WG42 for the 1-Mb region was perfect at −1, likely reflecting a single QTL in the region that controls both traits.

The objectives of the current study were to (i) validate the association of the SSC4 region with PRRS resistance in two unrelated commercial crossbred populations and (ii) further characterize the effect of the region in terms of gene action, parent- and breed-of-origin effects and genetic variance explained.

Materials and methods

The Kansas State University Institutional Animal Care and Use Committee approved all experimental protocols for this study.

Study design

A detailed description of the design, data collection and molecular techniques used in PRRS host genetic consortium trials has been previously published (Lunney *et al.* 2011). A total of five trials were analyzed here. Trials 1 through 3 were described in Boddicker *et al.* (2012) and included pigs of the same cross from a single breeding company. Trials 4 and 5 are unique to this study and were sourced from two different breeding companies, and each represents a different cross. Pigs were provided from commercial breeding programs in the United States and Canada. Briefly, each trial of approximately 200 commercial pigs involved transporting animals at weaning (18–28 days of age) to Kansas State University, where they were subjected to a PRRS challenge. Within a trial, pigs were from the same high-health farm, except for trial 5, which included pigs from two farms. All farms were free of PRRSV, *Mycoplasma hyopneumoniae* and swine influenza virus. Upon arrival, pigs were randomly placed into pens of 10–15 pigs. After a 7-day acclimation period, pigs now between 25 and 35 days of age (day 0) were experimentally infected intramuscularly and intranasally with 10^5 (TCID50) of NVSL 97-7985, a highly virulent PRRSV isolate (Fang *et al.* 2007). Blood samples were collected at −6, 0, 4, 7, 11, 14, 21, 28, 35 and 42 dpi. Body weight (BW) was measured at 0, 7, 14, 21, 28, 35 and 42 dpi. Pigs were euthanized at 42 dpi.

Viremia was measured using a semi-quantitative TaqMan PCR assay for PRRSV RNA, as described in the study by Boddicker *et al.* (2012). The assay results were reported as the $\log_{10}$ of PRRSV RNA copies per millilitre of serum. Ear tissue was collected from all pigs for DNA isolation. Tissue or genomic DNA from the sires of pigs in trials 1 through 3 and from sires and dams for trials 4 and 5 was supplied by the breeding companies. Tissues or DNA samples were sent to GeneSeek, Inc. for genotyping with Illumina's Porcine SNP60 BeadChip.

Data from approximately 960 pigs infected with PRRSV across the five trials were evaluated. Table 1 provides an overview of the population structure by trial. Pigs in trials 1 through 3 were F1 individuals from a Large White (LW) and Landrace (LR) cross (see Boddicker *et al.* 2012 for further details). Pigs in trial 4 were from Duroc sires crossed with a LW/LR/Yorkshire composite line, and trial 5 pigs were a cross between Duroc sires and F1 LR/Yorkshire

Table 1 Population structure of the five trials.

Trial	Breed	*n*	WUR10000125 allele frequency[1]	Number offspring per family			Total number of offspring			
				Minimum	Mean	Maximum	Barrows	Gilts	Total	Dead[2]
1–3[3,4]										
Sires	LR	33	0.22	1	17.1	114	565	0	565	48
Dams	LW	204	0.08	1	2.8	6				
4										
Sires	Duroc	6	0.08	8	32.5	50	109	86	195	2
Dams	LW-LR-Y[5]	33	0.10	1	5.9	13				
5										
Sires	Duroc	10	0.12	2	19.9	42	109	90	199	14
Dams	LR/Y[6]	38	0.22	2	5.2	10				

LR, Landrace; LW, Large White; Y, Yorkshire.
[1]Frequency of the favorable allele (B) at SNP WUR10000125. For trials 1–3, dam genotypes were not provided, so frequencies were determined using ordered genotypes of the offspring.
[2]Number of piglets that died prior to 42 days post-infection.
[3]Trials 1–3 consisted of the same cross from one breeding company.
[4]Trials 1 and 2 were LR by LW crosses. For trial 3, 121 piglets were from LR sires by LW dams crosses and 63 piglets were from the reciprocal cross of LW sires by LR dams.
[5]LW-LR-Y composite line.
[6]LR/Y F1 cross.

dams. A total of 64 pigs died before 42 dpi, including two from trial 4 and 14 from trial 5. Dead pigs were necropsied; subsequent gross and microscopic pathology by a board-certified pathologist identified PRRS-associated disease as the major source of mortality.

Pedigree

Pedigree information was provided by the breeding companies and spanned three generations for the first three trials and one generation for trials 4 and 5. Genotypes on parents and offspring were used to validate and correct pedigrees, where possible. Monomorphic SNPs were removed, leaving approximately 58 000 informative markers for each trial. Parent–offspring mismatch frequencies were calculated as the number of SNPs for which the parent and offspring had opposing homozygous genotypes divided by the total number of polymorphic SNPs for which the parent and offspring were both homozygous. If a parent–offspring pair had a mismatch frequency <2 percent, then the named parent was accepted. Otherwise, offspring genotypes were compared with all possible parents and the most likely parent was chosen, provided the mismatch frequency was <10 percent. Otherwise, the parent was set to missing. Dam genotypes were not provided for the first three trials, and therefore, the dams provided by the breeding company were assumed correct. Only one individual from the first three trials was assigned missing parental information. For trial 4, not all parent genotypes were provided, and the pedigree provided by the company was not reliable. Therefore, full- and half-sib families were identified based on genomic relationships computed following VanRaden (2008). Thresholds for assigning full- and half-sib relationships were determined based on frequency distributions of genomic relationships, which clearly showed full- and half-sib groupings. Based on this, pairs of piglets were identified to be full-sibs if their genomic relationship was >0.35 and as half-sibs if their relationship was between 0.10 and 0.35.

Phenotypic traits

Details on the phenotypic traits analyzed are in Boddicker *et al.* (2012). Briefly, VL was quantified as area under the curve for log-transformed viremia for 0, 4, 7, 11, 14 and 21 dpi. WG to 21 or 42 dpi (WG21, WG42) was calculated as BW at day 21 or 42 minus BW at day 0. Edits for trials 1 through 3 are in Boddicker *et al.* (2012). Edits removed 15 individuals from trials 4 and 5 for the analysis of VL, with 13 due to death prior to 21 dpi, one with missing parentage information and one with missing viremia information. For WG42, 18 individuals were removed from trials 4 and 5, with 17 due to death prior to 42 dpi and one with missing parentage information. The number of individuals available after edits is listed in Table 2.

Statistical analyses

Heritabilities and variances due to litter were estimated based on pedigree relationships with a single-trait animal model using the software ASREML (Gilmour *et al.* 2006). Sex and the interaction of trial and parity of the sow were included as fixed factors and pen within trial, animal and litter as random effects. Piglets were born from parities ranging from 1 to 7. Parities 3 through 7 were combined into one parity class. The effect of farm of origin, which was relevant only for trial 5, because piglets came from a single

Table 2 Trait means and estimates (±SE) of heritability and litter effects (proportions of phenotypic variance) for viremia and weight gain after infection obtained from single-trait pedigree-based ASREML analyses.

	Trials 1–5					Trials 1–3		Trial 4		Trial 5	
Trait	*n*	Mean	SD[1]	Heritability	Litter	Heritability	Litter	Heritability	Litter	Heritability	Litter
Viremia 4 dpi[2]	927	5.9	0.54	0.22 (0.11)	0.09 (0.05)	0.39 (0.18)	0.10 (0.07)	0.09 (0.16)	0.04 (0.07)	0	0.16 (0.06)
Viremia 7 dpi[2]	928	6.3	0.31	0	0.30 (0.04)	0	0.33 (0.05)	0.11 (0.17)	0.08 (0.08)	0	0.24 (0.08)
Viremia 11 dpi[2]	906	6.1	0.41	0.33 (0.14)	0.16 (0.06)	0.39 (0.18)	0.17 (0.08)	0.26 (0.24)	0.04 (0.09)	0.27 (0.34)	0.12 (0.15)
Viremia 14 dpi[2]	897	5.6	0.64	0.20 (0.11)	0.20 (0.06)	0.21 (0.14)	0.23 (0.07)	0.04 (0.17)	0.15 (0.10)	0.38 (0.18)	0
Viremia 21 dpi[2]	859	3.8	1.0	0.22 (0.08)	0	0.24 (0.12)	0	0.08 (0.16)	0.08 (0.09)	0.14 (0.13)	0
VL	909	107	7.6	0.39 (0.15)	0.11 (0.06)	0.49 (0.19)	0.06 (0.08)	0.03 (0.19)	0.38 (0.12)	0.34 (0.17)	0
WG21, kg	909	5.6	2.0	0.41 (0.09)	0	0.42 (0.12)	0	0.48 (0.18)	0	0.22 (0.13)	0
WG42, kg	889	15.1	4.0	0.34 (0.14)	0.09 (0.06)	0.37 (0.18)	0.12 (0.08)	0.40 (0.17)	0	0.20 (0.13)	0

dpi, day post-infection; WG21, weight gain from 0 to 21 dpi; WG42, weight gain from 0 to 42 dpi; VL, viral load calculated as area under the curve of log-transformed viremia between 0 and 21 dpi.

[1]Standard deviation calculated as the square root of the sum of animal, litter and residual variances from the joint ASREML analysis of all five trials.

[2]Log_{10} templates/ml (qPCR).

farm for each of the other trials, was not significant ($P > 0.65$) and therefore not included in the analyses. Genetic correlations between traits were estimated using bivariate animal models with the same fixed and random factors as used in the single-trait models.

Genome-wide association analyses

Associations of SNP genotypes with phenotypes were analyzed by fitting all SNPs simultaneously using Bayesian genomic selection methods (Habier *et al.* 2011), as implemented in the software GENSEL (Fernando & Garrick 2009), using the following mixed model:

$$y = Xb + \sum_{i=1}^{k} z_i \alpha_i \delta_i + \varepsilon,$$

where y = vector of phenotypic observations; X = incidence matrix relating fixed factors to phenotypes; b = vector of fixed factors of sex, pen within trial and the interaction of trial and parity class; z_i = vector of the genotype covariate for SNP i based on the number of B alleles using Illumina's genotype calling (coded 0, 1, 2 or equal to the trial average for missing genotypes); α_i = allele substitution effect for SNP i; and δ_i = indicator for whether SNP i was included ($\delta_i = 1$) or excluded ($\delta_i = 0$) in the model for a given iteration of the Monte Carlo Markov chain (MCMC). Pen within trial was included as a fixed factor, as opposed to a random effect in ASREML, because the current version of GENSEL does not allow additional random effects. A total of 50 000 iterations were run for each phenotypic analysis, with the first 5000 iterations discarded as burn-in. The probability of $\delta_i = 0$ was set equal to $\pi = 0.99$. The Bayesian model was implemented using method Bayes-B (Habier *et al.* 2011). Genomic regions associated with traits were identified using 1-Mb non-overlapping windows using build 10.2 of the swine genome (http://www.ncbi.nlm.nih.gov/nuccore?term=199Sus%20scrofa%2C%20whole%20genome%20shotgun%20sequence, accessed 1 November 2011). The approach described by Wolc *et al.* (2012) and implemented in version 4.0 of the GENSEL software (Fernando & Garrick 2009) was used to compute the posterior distribution of the proportion of genetic variance captured by each 1-Mb window.

Validation of the effect of the SSC4 region

Validation of the effect of the 1-Mb region on SSC4, as identified by Boddicker *et al.* (2012), was carried out using SNP effect estimates obtained from the first three trials to predict GEBV of animals in trials 4 and 5:

$$\mathrm{GEBV}_i = \sum_{j=1}^{n} (z_{ij}\, \hat{\alpha}_j),$$

where GEBV_i = genomic estimated breeding value for the ith animal of trial 4 or 5, n = number of SNPs included in the prediction, z_{ij} = genotype covariate of SNP j for animal i (coded 0, 1, 2 or trial average for missing genotypes) and $\hat{\alpha}_j$ = allele substitution effect estimate for SNP j based on the analysis of trials 1 through 3. GEBV from the whole-genome analysis were computed either based on SNPs across the entire genome (n = 58 277 SNPs after removing SNPs that were monomorphic across all five trials) or based on only the 38 SNPs in the 1-Mb region on SSC4. Accuracy of the resulting GEBV was evaluated by correlating the GEBV with phenotypes in trials 4 and 5 adjusted for fixed effects estimated from the ASREML analyses of phenotypes.

Further characterization of the QTL on SSC4

Estimation of SNP effects

Details on single-marker analysis of SNP WUR10000125 identified in the genome-wide analyses for trials 1 through 3 are in Boddicker *et al.* (2012) and were repeated here for trials 1 through 5. Briefly, SNP WUR10000125 was included as an additional class variable in the ASREML analysis of trials 4 and 5 and all five trials jointly, because that SNP contributed over 99 percent of the variance explained by the region on SSC4. Non-additive gene action was investigated by fitting contrasts between genotypic effects. Dominance was investigated by testing whether the AB genotype was significantly different from the average of the AA and BB genotypes (allele names, A vs. B, were based on Illumina's A/B genotype calling). The B allele was favorable (Boddicker *et al.* 2012), so the contrast between genotypes AB and BB was used to test for complete dominance. Dominance was investigated using all five trials jointly. Cause-and-effect relationships of the effects of SNP WUR10000125 on VL and WG42 were investigated by including VL as a fixed linear covariate in the analysis of WG42 and WG42 as a fixed linear covariate in the analysis of VL.

Breed and parental origin effects

The effect of SNP WUR10000125 was evaluated for parent- and breed-of-origin effects. Available genotypes of offspring and their parents for the 38 SNPs (34 after removal of monomorphic SNPs) in the 1-Mb region on SSC4 were phased using PHASE software (Stephens *et al.* 2001) separately for trials 1–3, 4 and 5. Of the 38 SNPs, four were monomorphic across all populations, resulting in 34 informative markers. Based on the identified haplotypes, parental origins were determined for SNP WUR10000125 alleles of heterozygous animals. Parental origin could not be determined for 12 animals in the first three trials because their genotype for SNP WUR10000125 was missing. Ordered genotypes (AA, AB, BA, BB) were fitted as a class variable in the previously described ASREML univariate animal model, and significance of parent- or breed-of-origin

effects was determined based on the contrast between the AB and BA genotypes.

Results

Phenotypic parameters

Sex was not significant for VL but was significant for WG ($P < 0.01$), with barrows having greater WG than gilts. Raw phenotypic means for viremia and WG are shown in Table 2. Viremia peaked, on average, at 7 dpi and thereafter decreased through 21 dpi. The variation in viremia was lowest at 7 dpi and increased thereafter through 21 dpi. The mean WG to 21 dpi was 5.6 kg and 15.1 kg to 42 dpi, with an increase in variation from 21 to 42 dpi.

Phenotypic correlations among traits obtained from the ASREML analyses are shown in Table 3. Both measures of WG were phenotypically negatively correlated with all measures of viremia. The phenotypic correlation between WG42 and VL was moderate. Phenotypic correlations between VL and daily measures of viremia were all positive, with the highest correlation at 14 dpi.

Estimates of genetic parameters

Estimates of heritability and litter components from the single-trait animal model analyses for all five trials jointly, and by trial, are shown in Table 2. Trials 1, 2 and 3 were analyzed together. Heritability estimates for trials 1 through 3 were greater than reported for the same data by Boddicker *et al.* (2012), and litter variances were smaller, likely due to the pedigree corrections that were made. Heritability estimates for trials 4 and 5 were low to moderate but had large SE due to small sample sizes. The estimate of litter variance of VL for trial 5 was 0. These results must be interpreted with caution because estimates have large SE and litter and genetic effects are confounded. For trials 1 through 5 combined, heritability estimates were all moderate, with the exception of viremia at 7 dpi, which had an estimate of 0. Heritability estimates for VL and WG42 across the five trials were 0.39 and 0.34 respectively. The proportion of phenotypic variance explained by litter was 0.11 for VL and 0.09 for WG42.

Estimates of genetic correlations among traits based on joint analysis of all five trials are shown in Table 3. Estimates of the genetic correlations of WG with each measure of viremia were negative but had large SE. The genetic correlation between WG42 and VL was -0.31 ± 0.26. Viremia at 11 and 14 dpi was highly positively correlated (0.94 ± 0.08). VL was also highly positively correlated (>0.9) with viremia at 7, 11 and 14 dpi. The estimate of the genetic correlation between WG21 and WG42 was perfect at 1.

Validation of the effects of the SSC4 region

Correlations between GEBV and phenotypes adjusted for fixed effects for individuals from trials 4 and 5, where GEBV were predicted based on estimates of SNP effects obtained from trials 1 through 3, are shown in Table 4. When dividing by the square root of heritability, these correlations can be interpreted as estimates of the accuracy of the GEBV. All correlations were positive and, on average, the GEBV predicted WG marginally better than it did VL, and predictions based solely on the SSC4 region were marginally better than were the whole-genome predictions.

Results of single-marker analyses for the most significant SNP in the SSC4 region (WUR10000125) that was identified by Boddicker *et al.* (2012) for VL and WG42 are shown in Fig. 1. This SNP was significant for WG42 in trial 4 ($P < 0.0004$); AA animals gained 3.2 kg less BW over the 42-day test period compared with AB animals. Individuals with genotype BB were not significantly different from the other two genotypes, but the number of BB animals was small (Fig. 1). The SNP also was significant ($P < 0.03$) for VL in trial 4; AA individuals had higher VL compared with AB individuals. For trial 5, the SNP was significant for both WG42 ($P < 0.03$) and VL ($P < 0.001$). Individuals with the AA genotype had, on average, lower WG and higher VL compared with AB individuals. Individuals with genotype

Table 3 Estimates of phenotypic (above diagonal) and genetic (below diagonal) correlations based on bivariate analyses of viremia and weight gain following infection for trials 1 through 5.

Trait	4 dpi	7 dpi	11 dpi	14 dpi	21 dpi	VL	WG21, kg	WG42, kg
Viremia 4 dpi	–	0.38 ± 0.03	0.09 ± 0.04	0.06 ± 0.04	0.03 ± 0.04	0.34 ± 0.04	−0.16 ± 0.04	−0.14 ± 0.04
Viremia 7 dpi	−0.01 ± 3.7	–	0.13 ± 0.04	0.09 ± 0.04	0.09 ± 0.04	0.34 ± 0.04	−0.16 ± 0.04	−0.13 ± 0.04
Viremia 11 dpi	0.28 ± 0.34	0.79 ± 1.0	–	0.68 ± 0.02	0.30 ± 0.04	0.71 ± 0.02	−0.05 ± 0.04	−0.14 ± 0.04
Viremia 14 dpi	0.42 ± 0.38	0.96 ± 1.6	0.94 ± 0.08	–	0.39 ± 0.03	0.78 ± 0.02	−0.20 ± 0.04	−0.24 ± 0.04
Viremia 21 dpi	0.11 ± 0.44	0.99 ± 0.45	0.78 ± 0.20	0.81 ± 0.18	–	0.75 ± 0.02	−0.17 ± 0.04	−0.20 ± 0.04
VL	0.40 ± 0.29	0.98 ± 0.71	0.92 ± 0.07	0.96 ± 0.04	0.86 ± 0.10	–	−0.23 ± 0.04	−0.28 ± 0.04
WG21, kg	−0.63 ± 0.62	−0.73 ± 0.56	−0.16 ± 0.28	−0.15 ± 0.30	−0.04 ± 0.34	−0.27 ± 0.24	–	0.78 ± 0.02
WG42, kg	−0.47 ± 0.32	−0.61 ± 0.45	−0.16 ± 0.30	−0.32 ± 0.31	−0.21 ± 0.34	−0.31 ± 0.26	1.00 ± 0.02	–

dpi, day post-infection; WG21, weight gain from 0 to 21 dpi; WG42, weight gain from 0 to 42 dpi; VL, viral load calculated as area under the curve of log-transformed viremia between 0 and 21 dpi.

Table 4 Percentage of genetic variance explained by the 1-Mb region on *Sus scrofa* chromosome 4 for the joint analysis of trials 1 through 5 using Bayes-B and its significance based on the proportion of posterior samples that had variance greater than expected.

Trait	% of total genetic variance	Proportion > expected[1]	Marker-based heritability[2]
Viremia 4 dpi[3]	2.2	0.860	0.43
Viremia 7 dpi[3]	1.3	0.699	0.43
Viremia 11 dpi[3]	10.9	0.998	0.58
Viremia 14 dpi[3]	6.5	0.984	0.49
Viremia 21 dpi[3]	0.77	0.396	0.19
VL	15.1	1.00	0.49
WG21, kg	5.0	1.00	0.70
WG42, kg	11.3	1.00	0.44

dpi, day post-infection; WG21, weight gain from 0 to 21 dpi; WG42, weight gain from 0 to 42 dpi; VL, viral load calculated as area under the curve of log-transformed viremia between 0 and 21 dpi.
[1]Proportion of posterior samples from the Monte Carlo Markov chain for which the 1-Mb region on chromosome 4 explained a greater percentage of the genetic variance than expected if all 1-Mb regions across the genome explained an equal percentage of the genetic variance, that is, 0.039 percent.
[2]Heritability estimates obtained from GENSEL software using method Bayes-B with $\pi = 0.99$.
[3]Log_{10} templates/ml (qPCR).

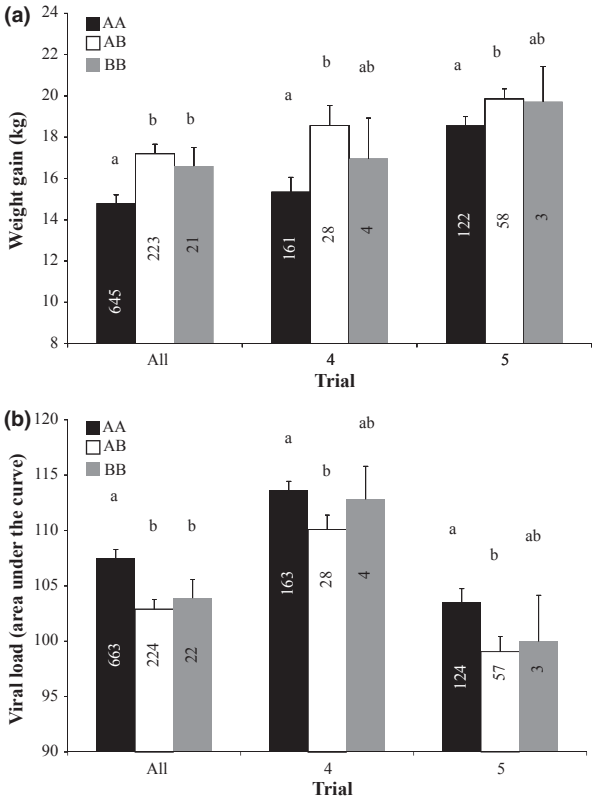

Figure 1 Least squares means by genotype for SNP WUR10000125 across all trials 1 through 5 and within trials 4 and 5 for weight gain from 0 to 42 days post-infection (a) and viral load (b), calculated as area under the curve of log-transformed viremia from 0 to 21 days post-infection. Within a trial, columns with a different letter are significantly different at $P < 0.05$. Numbers of individuals within genotype are listed in the bars.

BB were not significantly different from the other two genotypes; however, there were few animals with the BB genotype (Fig. 1).

Joint analysis of all five trials

The posterior mean of the percentage of genetic variance explained by the 1-Mb region on SSC4 for viremia on individual days ranged from 0.77 to 10.9 percent (Table 4), with 21 dpi having the lowest and 11 dpi having the highest percentage. If all 2592 of the 1-Mb windows across the genome contributed equally to the genetic variance, each window would contribute 0.039 percent. Thus, the SSC4 region was estimated to contribute a substantially greater percentage of genetic variance than expected under an infinitesimal model for each dpi (=0.039% based on 2592 1-Mb windows evaluated). Table 4 also shows a significance test for the hypothesis that the SSC4 region contributes more variance than expected (Wolc *et al.* 2012) based on the percentage of samples from the posterior distribution for which the 1-Mb window on SSC4 contributed more than 0.039 percent of the genetic variance. This percentage was as high as 99.8 for viremia at 11 dpi and not lower than 39.6 for 21 dpi. Table 4 also shows the marker-based heritability estimates obtained from GENSEL; estimates were moderate to high and generally higher than pedigree-based estimates (Table 2).

The 1-Mb region on SSC4 was estimated to explain 15, 5 and 11.3 percent of the genetic variance for VL, WG21 and WG42 respectively, and every sample showed greater variance than expected for the 1-Mb region (Table 4). Figure S1 shows the posterior distribution of the percentage of genetic variance explained by the 1-Mb region on SSC4 for every 100th iteration of the MCMC for WG42 (Fig. S2A) and VL (Fig. S2B). For VL, genetic variance explained by the region was at least 5 percent for all MCMC samples. For WG42, 94 percent of samples had variance >5 percent.

Using $2pq\alpha^2$ to estimate the additive genetic variance captured by SNP WUR10000125 (Falconer & Mackay 1996), this SNP accounted for 31.8, 5.7 and 28.6 percent of the additive genetic variance (using the estimate obtained from joint analysis of all trials) for VL for trials 1–3, 4 and 5 respectively. Similarly for WG42, this SNP accounted for 22.7, 18.5 and 8.0 percent of the additive genetic variance in trials 1–3, 4, and 5 respectively. The frequency of the favorable B allele used to calculate the additive genetic variance captured by SNP WUR10000125 was 0.16, 0.09 and 0.17 for trials 1–3, 4 and 5 respectively.

Single-marker analysis results for SNP WUR10000125 across all trials are presented in Fig. 1. The SNP was highly

significant for WG42 ($P < 1.5 \times 10^{-12}$, Fig. 1a) and VL ($P < 1.6 \times 10^{-14}$, Fig. 1b). The difference between the least squares means for the AB genotype and the average of the AA and BB genotypes was -2.8 ± 1.0 units ($P < 0.007$) for VL and 1.5 ± 0.6 kg ($P < 0.015$) for WG42; therefore, dominance effects were significant for both traits. The BB genotype was not well represented but was not significantly different from the AB genotype ($P > 0.48$) for either trait.

The contrast between AA and AB genotypes for SNP WUR10000125 were still significant ($P < 0.001$) when VL was included as a covariate in the analysis of WG42 and of a similar magnitude (−1.8 vs. −2.3 kg with or without VL included as a covariate). Assuming a linear effect of VL on WG, this indicates that the effect of the SSC4 region on WG42 was not completely caused by the effect of the SSC4 region on VL. Similarly, the contrast between the AA and AB genotypes for VL was still significant (3.7 vs. 4.7 units with or without WG42 as a covariate, $P < 0.001$) when including WG as a covariate in the analysis of VL, suggesting that the effect of the SSC4 region on VL is not through its effect on WG. So the effects of the SSC4 region on WG and VL appear to be a combination of effects that are common to both traits and effects that are specific to each trait.

Parent- or breed-of-origin effects

The effect of SNP WUR10000125 was tested by trial to determine parent- and breed-of-origin effects for the SSC4 region. Haplotype probabilities were high, with an average of 0.997 and SD of 0.03 across all five trials, which indicates that sufficient information was available to phase genotypes for most individuals. Across all five trials, 60 unique haplotypes were present. Estimates of the frequency of the B allele for each parental line by trial are shown in Table 1.

Figure 2 shows the least squares means for the ordered genotypes for WG42 (Fig. 2a) and VL (Fig. 2b). Effects of the heterozygotes ordered by parental origin (AB vs. BA) were not significantly different ($P > 0.30$) from each other for VL or WG42, and at least one of the ordered heterozygotes was significantly different from the AA genotype ($P < 0.02$) within each trial. Analyzed jointly, the AB and BA genotypes were not significantly different from each other ($P > 0.99$), but the AA genotype was significantly different from both AB and BA genotypes ($P < 0.01$) for both VL and WG42 (Fig. 2). Therefore, there is no statistical evidence for parent-of-origin or breed-of-origin effects.

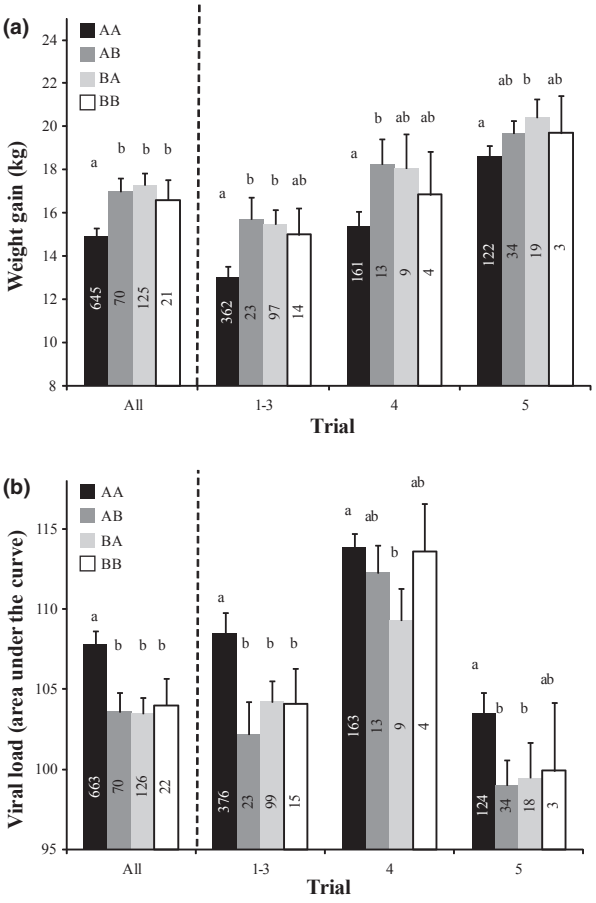

Figure 2 Least squares means by ordered genotype for SNP WUR10000125 across all trials 1 through 5 and within trial for weight gain from 0 to 42 days post-infection (a) and viral load (b), calculated as area under the curve of log-transformed viremia from 0 to 21 days post-infection. Trials 1 through 3 were grouped as animals were of the same cross from the same genetics company. Within a trial, columns with a different letter are significantly different at $P < 0.05$. In the genotype notation, the first allele is of paternal origin and the second allele is of maternal origin.

Discussion

Boddicker *et al.* (2012) identified a region on SSC4 that was associated with VL and WG using pigs from a single commercial cross between a LW line and a LR line that were experimentally infected with PRRSV. In this study, the effects of that region were validated in two additional, unrelated populations, and the favorable allele was shown to have a dominant mode of inheritance. Furthermore, the effect of the region on SSC4 was shown to be present in, and expressed by, both paternal and maternal chromosomes from the LR, LW, Yorkshire and Duroc lines that were involved in the crosses. In the following, these results will be discussed in further detail and compared with the literature. It must be emphasized that comparisons to Boddicker *et al.* (2012) are confounded because the current analyses included the data used in that study (trials 1 through 3).

Genetic parameters

Decreased WG is reportedly associated with the onset of PRRSV infection (Greiner *et al.* 2000; Doeschl-Wilson *et al.*

2009; Boddicker *et al.* 2012). Greiner *et al.* (2000) reported decreased WG at 8 dpi with PRRSV, relative to WG at 8 days pre-infection. Boddicker *et al.* (2012) reported negative phenotypic correlations between VL and WG to 21 dpi and WG to 42 dpi. In the current study, WG21 and WG42 also were found to be phenotypically negatively correlated with all measures of viremia and VL, consistent with previously published results of decreased WG when pigs are infected with PRRSV.

Compared with results of Boddicker *et al.* (2012), correction of pedigree errors resulted in substantial increases in heritabilities for trials 1 through 3 and in a reduction in the proportion of variance due to litter (Table 2). Both measures of WG were moderately heritable in trials 4 and 5, which is consistent with results from the first three trials. When all five trials were analyzed together, heritability of VL increased from 0.31 for trials 1–3 in Boddicker *et al.* (2012) to 0.39 for joint analysis of all five trials in the current study. Heritability estimates for viremia over time for trials 4 and 5 ranged from 0 to 0.38 but had large SE due to small sample sizes within trial. These estimates may still be underestimated as a result of additional pedigree errors that could not be corrected because of missing parental SNP genotypes.

Marker-based heritability estimates were generally substantially higher than were the estimates obtained from pedigree information (Table 4). This is likely because litter was not included in the GENSEL analysis, as the current version of GENSEL does not allow additional random effects to be included. However, for VL, the two estimates of heritability were in agreement with each other (pedigree = 0.39, marker = 0.49). VL is a summary statistic for viremia, which reduces some of the measurement and sampling errors that are present in measures of viremia on individual days.

Within trial, estimates of variance due to litter were not reliable due to small sample sizes and confounding between genetic and litter effects. Joint analysis of all trials revealed that the dam provided a maternal component in addition to her direct genetic contribution that aided her offspring to respond to the virus, despite the dams being from PRRS-free herds. Litter explained a substantial proportion of the phenotypic variance for viremia at 7, 11 and 14 dpi across trials (0.31, 0.16 and 0.18 respectively; i.e. greater for dpi closer to the day of infection). This maternal component could be due to maternal genetics or due to maternal environmental effects. Sows that provide a better maternal environment up to weaning result in stronger piglets that may be able to cope better with challenges, such as PRRS. Boddicker *et al.* (2012) reported a larger litter effect for VL and WG42 for the first three trials likely due to pedigree errors that were corrected in the current analyses.

Estimates of genetic correlations were generally positive among measures of viremia and among measures of WG and negative between measures of viremia and WG (Table 3). However, most of these estimates were not significantly different from zero ($P > 0.10$). Similar to what was observed by Boddicker *et al.* (2012), viremia at 11 and 14 dpi was highly and positively genetically correlated with each other and with VL. Viremia on these 2 days heavily influenced the results observed for VL.

Genome-wide association analysis

A multibreed data set was used for genome-wide association analysis, which raises concerns about the potential impact of population structure in detecting false positives (Pritchard *et al.* 2000). Breed differences between trials were accounted for by including trial as a fixed effect in the analyses. Within a trial, breed composition was homogeneous, with all pigs being generated by the same cross (or its reciprocal in the case of trial 3). The dams in trial 4 were from a three-breed composite, which was assumed to be homogeneous.

A genome-wide association analysis was conducted using the Bayesian genomic prediction statistical models, which analyze all SNPs simultaneously but, other than trial, did not explicitly include additional factors to account for population structure. However, our recent research (ShaarbafToosi 2012) has demonstrated that these models implicitly account for population structure and that not fitting population structure explicitly does not lead to an excess of false positives in these models. One reason for this is that any SNPs that differ in allele frequencies between subpopulations can capture the effects of population structure, and thus, any effects of population structure are spread out across many SNPs across the genome. With over 2500 1-Mb windows, this is expected to increase the variance associated with each window only by a fraction of 1/2500 of the variance contributed by population structure, which is expected to be minimal. Furthermore, if differences due to population structure contain genetic information, allowing the SNPs to capture these effects actually increases the power to detect QTL (ShaarbafToosi 2012).

Validation

Boddicker *et al.* (2012) used Build 10 of the swine genome. The 1-Mb region on SSC4 that was identified by Boddicker *et al.* (2012) included 37 SNPs in Build 10, of which four were monomorphic. Two of these SNPs were located upstream (approximately 469 kb) from the region in Build 10.2, and three SNPs that were unmapped in Build 10 were placed into the region on SSC4 in Build 10.2. Thus, in Build 10.2, the SSC4 region included 38 SNPs, including four monomorphic SNPs.

At the whole-genome level, validation correlations between GEBV and adjusted phenotypes were low to moderate (Table 5). The correlation estimates can be

Table 5 Correlations (±SE) between phenotypes for trials 4 and 5 and genomic estimated breeding values at the whole-genome level and for the 1-Mb region on *Sus scrofa* chromosome 4 (SSC4) using allele substitution effect estimates obtained from trials 1 through 3.

	Trial 4		Trial 5	
Trait[1]	Genome	SSC4 region	Genome	SSC4 region
Viral load	0.07 ± 0.07	0.14 ± 0.07	0.20 ± 0.07	0.18 ± 0.07
Weight gain, kg	0.19 ± 0.07	0.12 ± 0.07	0.15 ± 0.07	0.24 ± 0.07

[1]Viral load was calculated as area under the curve of log-transformed viremia between 0 and 21 days post-infection (dpi). Weight gain is from 0 to 42 dpi.

converted to estimates of the accuracy of GEBV (i.e. the correlation between true and estimated breeding values) by dividing by the square root of heritability of the trait. Using heritabilities of 0.4, this results in estimates of accuracy ranging from 0.11 to 0.32 at the whole-genome level and 0.19–0.38 for the 1-Mb region on SSC4. These accuracies are high, especially considering the limited size of the training data and the fact that prediction was across breeds. Across-breed predictions are typically not accurate due to dominance, epistasis and differences in linkage disequilibrium (LD) between breeds (Garrick 2010). Accuracies of GEBV computed based on the 1-Mb region on SSC4 were on average of similar magnitude as accuracies of whole-genome GEBV. This suggests that the rest of the genome added as much noise as true effects to the GEBV predictions. The SSC4 region had very similar LD patterns in all five trials (Fig. S2), which may explain the predictive power of effects estimated for this region across breeds.

The effects of SNP WUR10000125 in trials 4 and 5 were similar to those previously found for the first three trials (Boddicker *et al.* 2012), in which AA animals had increased VL and decreased WG42 compared with the AB animals for both traits. Therefore, the effect of this region has now been identified in three unrelated populations, suggesting the genetic variant likely existed before the divergence of the breeds represented in this study. Furthermore, the low frequency of the B allele may indicate that, historically, the B allele has not had an advantage for natural or artificial selection.

When the SNP was fitted as a fixed factor in the joint analysis of all five trials, the effect of the favorable allele was shown to be dominant. Expressed in terms of phenotypic SD, the effects were 0.54 and 0.53 for VL and WG42. The minor allele (frequency = 0.15) was associated with the desired phenotype of reduced VL and increased WG. The SSC4 region explained a large percentage of the genetic variance for VL and WG42 (15.1 and 11.3 percent respectively). These percentages are similar to those reported for the first three trials by Boddicker *et al.* (2012) but were calculated in a different manner; percentages in Boddicker *et al.* (2012) were calculated based on variances of GEBV; here, estimates were based on the posterior distribution of the variances of true breeding values. Nevertheless, all these results provide convincing evidence that the previously identified QTL on SSC4 is present in two additional, unrelated commercial populations and has large effects on host response to experimental PRRSV infection.

Potential candidate genes for the QTL identified in the region were discussed in Boddicker *et al.* (2012) and include a family of guanylate-binding proteins (GBP1, GBP2, GBP4, GBP5 and GBP6) and general transcription factor 2b. As discussed in Boddicker *et al.* (2012), these proteins have been reported to have antiviral activity in humans, and the GBP family of genes is associated with the innate immune system. Fine-mapping this QTL and future identification of the causative mutation are clearly hampered by the extensive LD that exists in the region across breeds. Therefore, functional analyses will be required to provide any convincing evidence of causality among candidate mutations.

The frequency of the favorable B allele is low, which begs the question, why? Interestingly, the frequency is low in both maternal and terminal lines, which are selected primarily for reproductive and production traits respectively. Thus, it is not very likely that the QTL on SSC4 associated with PRRS is in LD with a major QTL for reproduction or production traits. Based on the results available from PigQTLdb (http://animalgenome.org/cgi-bin/QTLdb/SS/index), no QTL for production traits, such as growth or feed efficiency, have been reported within 2 Mb of the PRRS QTL on SSC4. However, QTL have been reported for meat quality traits, such as intramuscular fat content, shear force and carcass weight within 9 Mb of the SSC4 region. Non-functional nipple is the only reproduction-related QTL that has been reported within 2 Mb of the PRRS QTL. Selection for performance in the presence of PRRS infection is expected to favor the B allele. The low frequency of that favorable allele is likely due to a combination of the following: (i) PRRSV is a relatively new disease and (ii) selection typically occurs in high-health nucleus farms, which tend to be free of PRRS. Furthermore, random drift could cause a low frequency of the favorable allele and in differences in allele frequencies between breeds.

Origin of SSC4 effects

Parent-of-origin effects are commonly the result of imprinting, and, as reviewed by Guilmatre & Sharp (2012), there is evidence that imprinting contributes to a number of human disorders such as Alzheimer's disease, autism, breast cancer and both type 1 and 2 diabetes. However, to our knowledge, there is no previous evidence of imprinting effects for response to infectious diseases. The parent-of-origin analysis performed here was twofold, in the sense that it tested whether the effect was specific to paternally or maternally

inherited alleles and also whether the effect was breed specific. These effects were, however, partially confounded because some breeds were used only as paternal or maternal breeds in the crosses analyzed. The first three trials were based on a cross between two white breeds (Yorkshire and LR cross), and trials 4 and 5 encompassed the Yorkshire, LR, LW and Duroc breeds. Duroc pigs typically are used as terminal sires in the commercial sector, and Yorkshire and LR are typically used as maternal breeds. The Yorkshire and LR lines used in trials 4 and 5 are, however, expected to be distinct from each other and from the Yorkshire and LR lines used in trials 1 through 3 because of the separate populations and selection strategies used by each breeding company. When SNP WUR10000125 was fitted as an ordered genotype, the effect was present irrespective of whether the allele was paternally or maternally inherited. Additionally, the effect of the favorable allele was present for each of the parental breeds and lines represented in the crosses.

Conclusions and implications

A large QTL was previously identified on SSC4 using approximately 560 pigs from a single commercial cross. The effects of this region have now been validated in two unrelated crosses from two additional breeding companies. The effects were of similar magnitude and in the same direction as previously identified. The QTL acts in a dominant matter, whereby pigs with the favorable allele have reduced viremia and increased WG. The effect is present in lines from the major breeds involved in commercial pork production, and there is no indication of parent-of-origin or breed-of-origin effects, as this study included LR, Yorkshire, LW and Duroc breeds. Frequencies of the favorable allele were <17 percent in each parental line, suggesting ample room for marker-assisted selection on this region. These results are promising for the swine industry and for swine health with respect to PRRS. However, they are based on infection with a specific strain of the PRRSV under experimental conditions. Additional research, such as challenges with other PRRSV isolates and field trials, is required to better understand the effects of this genomic region on pig production in the field. Nevertheless, if the effect of this region is common across breeds and populations, the favorable allele has a low frequency and the effects observed here replicate in the field, selection for the favorable allele could lead to substantial reductions in economic losses due to PRRS.

Acknowledgements

This project was supported by the USDA NIFA PRRS CAP Award 2008-55620-19132, the National Pork Board, the NRSP-8 Swine Genome and Bioinformatics Coordination projects, and the PRRS Host Genetics Consortium consisting of USDA ARS, Kansas State Univ., Iowa State Univ., Michigan State Univ., Washington State Univ., Purdue Univ., University of Nebraska-Lincoln, PIC/Genus, Newsham Choice Genetics, Fast Genetics, Genetiporc, Inc., Genesus, Inc., PigGen Canada, Inc., IDEXX Laboratories, Tetracore, Inc. The authors acknowledge technical assistance to this project from Juan Pedro Steibel for the algorithm to compute area under the curve, Max Rothschild for assistance with genotyping, Nader Deeb for assistance with data and analysis, Eric Fritz for database management, Ania Wolc and Dinesh Thekkoot for statistical assistance, the laboratory of Bob Rowland, specifically Becky Eaves, Maureen Kerrigan, Ben Trible, Jessica Otradovec, Brooke Bloomberg, Aubree Gottlob, Laura O'Brien and Ranjini Chand for animal care and sample collection, and the laboratory of Joan Lunney, specifically Samuel Abrams and Amber Tietgens for preparation of all genomic DNA samples for SNP genotyping.

References

Boddicker N.J., Waide E.H., Rowland R.R.R., Lunney J.K., Garrick D.J., Reecy J.M. & Dekkers J.C.M. (2012) Evidence for a major QTL associated with host response to porcine reproductive and respiratory syndrome virus challenge. *Journal of Animal Science* **90**, 1733–46.

Doeschl-Wilson A.B., Kyriazakis I., Vincent A., Rothschild M.F., Thacker E. & Galina-Pantoja L. (2009) Clinical and pathological responses of pigs from two genetically diverse commercial lines to porcine reproductive and respiratory syndrome virus infection. *Journal of Animal Science* **87**, 1638–47.

Falconer D.S. & Mackay T.F. (1996) *Introduction to Quantitative Genetics*, 4th edn. Pearson, Harlow.

Fang Y., Schneider P., Zhang W.P., Faaberg K.S., Nelson E.A. & Rowland R.R.R. (2007) Diversity and evolution of a newly emerged North American Type 1 porcine arterivirus: analysis of isolates collected between 1999 and 2004. *Archives of Virology* **152**, 1009–17.

Fernando R. & Garrick D. (2009) *User Manual for a Portfolio of Genomic Selection Related Analyses*, 2nd edn. Iowa State University. Available at: http://bigs.ansci.iastate.edu/bigsgui/help.html, accessed 19 April 2012.

Garrick D.J. (2010) Consequences of genomic prediction in cattle. *Interbull Bulletin* **41**, 51–8.

Gilmour R.A., Gogel B.J., Cullis B.R. & Thompson R. (2006) *ASREML User Guide Release 2.0*. VSN International Ltd, Hemel Hempstead, UK.

Greiner L.L., Stahly T.S. & Stabel T.J. (2000) Quantitative relationship of systemic virus concentration on growth and immune response in pigs. *Journal of Animal Science* **78**, 2690–5.

Guilmatre A. & Sharp A.J. (2012) Parent of origin effects. *Clinical Genetics* **81**, 201–9.

Habier D., Fernando R.L., Kizilkaya K. & Garrick D.J. (2011) Extension of the Bayesian alphabet for genomic selection. *BMC Bioinformatics* **12**, 186.

Holtkamp D.J., Kliebenstein J.B., Neumann E.J., Zimmerman J.J., Rotto H., Yoder T.K., Wang C., Yeske P., Mowrer C. & Haley C. (2013) Assessment of the economic impact of porcine reproduc-

tive and respiratory syndrome virus on United States pork producers. *Journal of Swine Health and Production* **21**(2), 72–84.

Lunney J.K., Steibel J.P., Reecy J., Rothschild M., Kerrigan M., Trible B. & Rowland R.R.R. (2011) Probing genetic control of swine responses to PRRSV infection: Current progress of the PRRS Host Genetics Consortium. Intl. Symposium on Animal Genomics for Animal Health, Paris. *BMC Proceedings* **5**(Suppl 4), S30.

Pritchard J.K., Stephens M., Rosenberg N.A. & Donnelly P. (2000) Association mapping in structured populations. *American Journal of Human Genetics* **67**, 170–81.

Rowland R.R.R., Lunney J.K. & Dekkers J.C.M. (2012) Control of porcine reproductive and respiratory syndrome (PRRS) through genetic improvements in disease resistance and tolerance. *Frontiers in Genetics* **3**, 260.

ShaarbafToosi A. (2012) *Genome-Wide Prediction of Breeding Values and Mapping of Quantitative Trait Loci in Stratified and Admixed Populations*. PhD Thesis, Iowa State University. Available at: http://lib.dr.iastate.edu/etd/12756/.

Stephens M., Smith N.J. & Donnelly P. (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics* **68**, 978–89.

VanRaden P.M. (2008) Efficient methods to compute genomic predictions. *Journal of Dairy Science* **91**, 4414–23.

Wolc A., Arango J., Settar P. *et al.* (2012) Genome-wide association analysis and genetic architecture of egg weight and egg uniformity in layer chickens. *Animal Genetics* **43**(Suppl 1), 87–96.

Supporting information

Additional supporting information may be found in the online version of this article.

Figure S1 Posterior distribution for the percentage of genetic variance explained by the 1-Mb region on *Sus scrofa* chromosome 4 for weight gain and viral load.

Figure S2 Linkage disequilibrium plots from HAPLOVIEW for trials 4, 5 and 1 through 5 for the 1-Mb region on *Sus scrofa* chromosome 4.