The Addition of Epistatic Genetic Effects Increases Genomic Prediction Accuracy for Reproduction and Production Traits in Duroc Pigs Using Genomic Models

Jian Cheng

North Carolina State University at Raleigh: NC State University

Francesco Tiezzi

North Carolina State University at Raleigh: NC State University

Jeremy Howard

Smithfield Premium Genetics, Rose Hill, NC

Christian Maltecca

North Carolina State University at Raleigh: NC State University

Jicai Jiang (✉ jicai_jiang@ncsu.edu)

North Carolina State University <https://orcid.org/0000-0001-6890-7539>

Research Article

Keywords: non-additive effects, pig, genomic prediction

Posted Date: January 6th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1182452/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Genomic selection has been implemented in livestock genetic evaluations for years. However, currently most genomic selection models only consider the additive effects associated with SNP markers and nonadditive genetic effects have been for the most part ignored.

Methods: Production traits for 26,735 to 27,647 Duroc pigs and reproductive traits for 5,338 sows were used, including off-test body weight (WT), off-test back fat (BF), off-test loin muscle depth (MS), number born alive (NBA), number born dead (NBD), and number weaned (NW). All animals were genotyped with the PorcineSNP60K Bead Chip. Variance components were estimated using a linear mixed model that includes inbreeding coefficient, additive, dominance, additive-by-additive, additive-by-dominance, dominance-by-dominance effect, and common litter environmental effect. Genomic prediction performance, including all nonadditive genetic effects, was compared with a reduced model that included only additive genetic effect.

Results: Significant estimates of additive-by-additive effect variance were observed for NBA, BF, and WT (31%, 9%, and 10%, respectively). Production traits showed significant large estimates of additive-by-dominance variance (9%-23%). MS also showed large estimate of dominance-by-dominance variance (10%). Dominance effect variance estimates were low for all traits (0%-2%). Compared to the reduced model, prediction accuracies using the full model, including nonadditive effects, increased significantly by 12%, 12%, and 1% for NBA, WT, and MS, respectively. A strong dominance association signal with BF was identified near AK5.

Conclusions: Sizable estimates of epistatic effects were found for the reproduction and production traits, while the dominance effect was relatively small for all traits yet significant for all production traits. Including nonadditive effects, especially epistatic effects in the genomic prediction model, significantly improved prediction accuracy for NBA, WT, and MS.

Background

Genomic selection refers to the use of genome-wide dense single nucleotide polymorphism (SNP) markers to predict breeding values, which can then be used to select individuals for breeding [1]. Genomic selection has been implemented in the genetic evaluation of livestock populations for years. However, most genomic selection procedures only consider the additive effects associated with SNP markers. Genomic best linear unbiased prediction is a popular approach that uses genomic information in the form of a genomic relationship matrix that defines the additive genetic covariance between individuals [2–3].

Nonadditive genetic effects have been overlooked in the genetic evaluation of livestock. Nonadditive genetic effects include dominance effect and epistatic effect, which can be further classified as additive-by-additive, additive-by-dominance, and dominance-by-dominance effects. Dominance is the interaction effect of the two alleles at the same locus. The influence of dominance effects on traits related to

production and reproduction has been reported in pigs and other livestock animals [4–5]. Epistatic interactions involve the effect of alleles in different loci and contribute significantly to the variation of complex traits [6–7]. Most previous studies on epistatic effects only consider the additive-by-additive effect [8–9]. However, the additive-by-dominance and dominance-by-dominance interactions may also play a significant role in heterosis as well as inbreeding and outbreeding depression [10]. Recently, Vitezica et al. [11] proposed a flexible and general approach to construct epistatic genomic relationship matrices for two or higher order interactions using Hadamard products of additive and dominance genomic orthogonal relationships and was implemented by Vitezica et al. [12] for litter size in pigs. The total genetic value of an individual is a function of both additive, and nonadditive effects, which combined could result in better predictors of future phenotypes and better mate allocation strategies between selection candidates. Many studies have shown that rates of genetic gain were increased when nonadditive genetic effects were included in the matting program for dairy cattle [13–15].

Many genome-wide association studies (GWAS) have been performed in livestock [16] (Goddard and Hayes, 2009) to identify quantitative trait loci (QTL) associated with economically relevant traits. However, most GWAS studies only reported QTL for additive genetic effect. Nonadditive QTL are challenging to detect in animals, primarily due to the low statistical power in the testing for nonadditive effects of individual loci. Despite this, a few studies have identified QTL for nonadditive effects using GWAS. Specifically, Jiang et al. [5] identified a dominant QTL associated with milk yield using a large dairy cattle dataset. Bolormaa et al. [17] also found significant SNPs with dominance and epistatic effects on beef cattle growth, carcass, and fertility traits. In pigs, Yang et al. [18] detected two SNPs with dominance effects associated with growth rate.

To address these questions related to the nonadditive genetic effects, a large Duroc pig dataset was used, including three production traits and three reproduction traits. The objectives of this study were to 1) estimate the additive, dominance, and epistatic effects for pig production and reproduction traits; 2) identify QTL with dominance effects, and 3) evaluate whether adding these nonadditive genetic effects can improve the prediction accuracy of genomic selection.

Methods

The approval of the animal welfare and ethics committee was not needed for this study since a commercial breeding company provided all data.

Phenotypic and genotypic data

Phenotypes collected in Duroc pigs from 2015 to 2021 were provided by the Smithfield Premium Genetics company. About 27K animals' phenotypes were available for growth traits, including off-test body weight (WT), off-test back fat (BF), and off-test loin muscle depth (MS), and 5,338 sows' parity records were available for number born alive (NBA), number born dead (NBD), and number weaned (NW). The phenotypes for all the traits were pre-adjusted and standardized. First, a mixed model was fitted in order

to obtain effects' solutions. The model for the traits WT, BF and MS included the fixed effects of gender and farm-date of finishing as well as the random batch (pen within farm of finishing), additive genetic and residual effects. The model for the traits NBA, NBD and NW included the fixed effects of parity and farrow date as well as the random additive genetic, sow permanent environmental and residual effects. Variance components were estimated on the same dataset used. Models' parameter and solutions were obtained using the BLUPf90 suite of programs [19]. The pre-adjusted phenotypes were obtained by subtracting the contribution of the fixed effects from the phenotypic values. The phenotypes for the maternal traits were also averaged in order to have one record per individual. All animals were genotyped with the PorcineSNP60K Bead Chip (Illumina Inc., San Diego, CA, USA). Quality control of the genotype data was based on the following criteria: SNP call rate > 0.95, individual call rate > 0.95, Hardy–Weinberg equilibrium (HWE) P -value > 1e-6, and minor allele frequency > 0.01. After quality control, 45,962 SNPs and a total of 26,818 pigs and 3,246 sows with both genotypes and phenotypes were retained for the subsequent analysis. SNP positions are based on the Sus scrofa 11.1 reference genome. Missing genotypes were imputed using FindHap version 4.0 [20].

Variance decomposition with additive, dominance, and epistasis components

Variance components were firstly estimated by MMAP [21] using the following full model:

$$\mathbf{y} = \mu + \mathbf{f}b + \mathbf{Z}\mathbf{g}_A + \mathbf{Z}\mathbf{g}_D + \mathbf{Z}\mathbf{g}_{AA} + \mathbf{Z}\mathbf{g}_{AD} + \mathbf{Z}\mathbf{g}_{DD} + \mathbf{Q}\mathbf{p} + \mathbf{e}$$

1

where $\mathbf{y}$ is the pre-adjusted trait phenotype; μ is the mean; $\mathbf{f}$ is the genomic inbreeding coefficient; b is the regression coefficient measuring inbreeding depression; $\mathbf{g}_A$, $\mathbf{g}_D$, $\mathbf{g}_{AA}$, $\mathbf{g}_{AD}$, and $\mathbf{g}_{DD}$ are genetic values corresponding to additive, dominance, additive-by-additive, additive-by-dominance, and dominance-by-dominance effects, respectively; $\mathbf{p}$ is common litter environmental effect (each sow as one litter); $\mathbf{Z}$ and $\mathbf{Q}$ are the corresponding incidence matrices; and $\mathbf{e}$ is residual effect. The (co)variance matrix of phenotypes is

$$\begin{aligned} \text{Var}(\mathbf{y}) &= \text{Var}(\mathbf{g}_A) + \text{Var}(\mathbf{g}_D) + \text{Var}(\mathbf{g}_{AA}) + \text{Var}(\mathbf{g}_{AD}) + \text{Var}(\mathbf{g}_{DD}) + \text{Var}(\mathbf{p}) + \text{Var}(\mathbf{e}) \\ &= \mathbf{G}_A\sigma_A^2 + \mathbf{G}_D\sigma_D^2 + \mathbf{G}_{AA}\sigma_{AA}^2 + \mathbf{G}_{AD}\sigma_{AD}^2 + \mathbf{G}_{DD}\sigma_{DD}^2 + \mathbf{Q}\mathbf{Q}^T\sigma_p^2 + \mathbf{I}\sigma_e^2 \end{aligned} \quad (2)$$

where $\mathbf{G}_A$, $\mathbf{G}_D$, $\mathbf{G}_{AA}$, $\mathbf{G}_{AD}$, and $\mathbf{G}_{DD}$ are additive, dominance, additive-by-additive, additive-by-dominance, and dominance-by-dominance genomic relatedness matrices (GRM), respectively, and $\mathbf{I}$ is an identity matrix. σ_A^2 , σ_D^2 , σ_{AA}^2 , σ_{AD}^2 , and σ_{DD}^2 are additive, dominance, additive-by-additive, additive-by-dominance, and dominance-by-dominance genetic variance components explained by SNPs, respectively. σ_p^2 is the

variance component of common litter effects, and σ_e^2 is the residual variance. The ij th entry of $\mathbf{G}_A$ is calculated using the equation by VanRaden [3],

$$G_A(i, j) = \sum_k (Z_{ik} - 2p_k)(Z_{jk} - 2p_k) / \sum_k 2p_k(1 - p_k)$$

3

,

where Z_{ik} and Z_{jk} are the additive genotype code (0, 1, and 2 for genotypes AA, AB, and BB, respectively) for marker k of individual i and j , respectively, and p_k is the population frequency of the allele B. $\mathbf{G}_D$ is computed using the formula by Vitezica, et al. [22],

$$G_D(i, j) = \sum_k H_{ik}H_{jk} / \sum_k [2p_k(1 - p_k)]^2$$

4

,

where H_{ik} and H_{jk} are the dominance genotype code $[-2p_k^2, 2p_k(1 - p_k), \text{ and } -2(1 - p_k)^2$ for genotypes AA, AB, and BB, respectively] for marker k of individual i and j , respectively. Epistatic genomic relationship matrices can be constructed directly from $\mathbf{G}_A$ and $\mathbf{G}_D$ [12] assuming HWE,

$$\mathbf{G}_{ij} = (\mathbf{G}_i \# \mathbf{G}_j) / [\text{trace}(\mathbf{G}_i \# \mathbf{G}_j) / n]$$

5

,

where i and j represent A or D, n is the number of animals included in the genomic relationship matrix, and symbol $\#$ represents the Hadamard matrix product. In addition to the full model (1), we used a reduced model that excluded the nonadditive genetic effects. The significance of the variance component estimates was investigated using a Wald test.

Genomic prediction

Additive, dominance, and epistatic effects were estimated using model (1) in MMAP. The prediction of total genetic values (TGVs) can be computed as

$$\hat{\mathbf{g}}_T = \hat{\mathbf{g}}_A + \hat{\mathbf{g}}_D + \hat{\mathbf{g}}_{AA} + \hat{\mathbf{g}}_{AD} + \hat{\mathbf{g}}_{DD}$$

A reduced model that only includes additive genetic effects was also fitted in comparison to the full model with respect to predictive performance, based on a ten-fold cross-validation strategy. We evaluated the prediction accuracy using the Pearson's correlations in the validation population between estimated genetic values (either breeding values from the reduced model or TGVs from the full model) and phenotypes adjusted by common litter environmental effect and inbreeding depression. The phenotype adjustment was done once using all individuals. Unbiasedness was assessed as the regression coefficient of adjusted phenotypes on estimated genetic values in the validation population. We also investigated the predictive accuracy of the estimated additive genetic values and the sum of additive and dominance genetic values in the full model.

Genome-wide association study of dominance effects

Both a two-step and a one-step strategy were used for genome-wide association studies (GWAS). In the two-step strategy, we first fitted model (1) with MMAP and obtained the residuals. We then fitted the residual as a response variable in a linear model. The full linear model was implemented by the software package PLINK [23] using a 2-df joint test of both additive and dominance. Association p -values were calculated from t -tests for the dominance effects. To validate the results from the two-step strategy, a one-step linear mixed model was further used,

$$\mathbf{y} = \mu + \mathbf{fb} + \mathbf{H}_k d_k + \mathbf{a} + \mathbf{e}$$

$$\mathbf{a} \sim N(0, \mathbf{G}_A \sigma_A^2)$$

where d_k and $\mathbf{H}_k$ are the dominance effect and the dominance genotype code for SNP k , respectively. The linear mixed model was implemented in MMAP, which allowed only one term of polygenic effects to be fitted. Bonferroni correction was used to define the genome-wide significant threshold ($P < 1e-6$). Genes located in each quantitative trait locus (QTL) region were identified in BioMart [24], using the ENSEMBL pig gene database (Sscrofa 11.1). Overlap with previously identified QTLs was based on queries of the Animal QTL database (QTLdb) developed by Hu et al. [25], which houses all publicly available QTL and single nucleotide polymorphism (SNP) gene association data on livestock animal species. The pig QTLdb (accessed August 1st, 2021) was used in this study and included 30,869 QTL from 706 publications for 692 different traits.

Results

Estimates of variance components for additive, dominance, and epistatic effects

Estimates of inbreeding depression, common litter environmental effect variance, variance components of additive, dominance, and epistatic effects in Duroc pigs are shown in Table 1. Corresponding maternal genetic effects were initially analyzed but resulted in small and non-significant estimates, and were not included in the full model. In total, 3,246 sows were considered for reproductive traits, whereas 26,189 to 26,813 animals were included in the production traits analysis. Common litter environmental effect variance for all production traits were generally large and significantly greater than zero, in the case of WT even greater than the additive genetic variance (10% versus 8%, based on the full model). Common litter environmental effect variance for all reproductive traits were small, and were not included in the model. Inbreeding depression was significant for NBA ($P = 0.01$) and all production traits ($P < 0.001$) and was the largest for WT ($b = -23.71$). In general, large nonadditive genetic variances were observed for all reproduction and production traits. For the reproductive traits, the largest proportion of additive-by-additive variance to total phenotypic variance was found for NBA, and it was six times greater than the additive genetic variance (31% versus 5%). However, the full model with additive-by-dominance and dominance-by-dominance effects did not converge; thus, these effects were not included. Dominance-by-dominance explained a large proportion of variation (22%) in NBD, which was greater than the additive genetic variance (3%). However, the dominance-by-dominance genetic variance was not significant with P -value = 0.55. The additive-by-additive genetic variance for NBD also explained twice the additive genetic variance (6% versus 3%), but was still not significant. Additive-by-additive genetic variance accounted for about 12% of total phenotypic variance for NW, nearly four times the additive genetic variance (3%). BF had the largest additive effect and its variance explained 33% of phenotypic variation for production traits, while WT and MS contributed much less (8% and 19%, respectively) based on the full model. Most nonadditive genetic variances, i.e., dominance effect, additive-by-additive, and additive-by-dominance effect, were significantly greater than zero for BF. For this trait, the additive-by-additive and additive-by-dominance genetic effects explained the largest proportion of phenotypic variance (9% for both), followed by the dominance effect (1%). For WT, the additive-by-dominance effect accounted for the largest proportion of phenotypic variance (23%), followed by the additive-by-additive effect (10%), then the dominance effect (2%). MS had large and significant additive-by-dominance and dominance-by-dominance genetic variances (17% and 10% of variance explained, respectively). Finally, dominance variance was also small (1%) but significantly greater than zero.

Table 1
Estimates of inbreeding depression and proportions of variance component in Duroc pigs

Trait	Sample size	b (P-value)	A (P-value)	D (P-value)	AA (P-value)	AD (P-value)	DD (P-value)	p (P-value)
NBA (Full)	3246	-21.26 (0.01)	0.05 (0.01)	0 (1)	0.31 (0.001)	-	-	-
NBA (Reduced)	3246	-20.16 (0.01)	0.09 (< 0.001)	-	-	-	-	-
NBD (Full)	3246	-3.48 (0.75)	0.03 (0.14)	0 (1)	0.06 (0.50)	0 (1)	0.22 (0.55)	-
NBD (Reduced)	3246	-3.26 (0.74)	0.04 (0.01)	-	-	-	-	-
NW (Full)	3246	-13.95 (0.11)	0.03 (0.10)	0 (1)	0.12 (0.2)	0 (1)	-	-
NW (Reduced)	3246	-14.12 (0.14)	0.06 (< 0.001)	-	-	-	-	-
BF (Full)	26189	-15.35 (0.001)	0.33 (< 0.001)	0.01 (0.001)	0.09 (0.001)	0.09 (0.003)	0 (1)	0.05 (0.001)
BF (Reduced)	26189	-18.55 (0.001)	0.36 (< 0.001)	-	-	-	-	0.07 (0.001)
WT (Full)	26813	-23.71 (0.001)	0.08 (< 0.001)	0.02 (< 0.001)	0.10 (0.001)	0.23 (< 0.001)	0.02 (0.64)	0.10 (0.001)
WT (Reduced)	26813	-33.00 (0.001)	0.09 (< 0.001)	-	-	-	-	0.13 (0.001)
MS (Full)	26189	-12.27 (0.001)	0.19 (< 0.001)	0.01 (0.01)	0.05 (0.001)	0.17 (< 0.001)	0.10 (0.02)	0.04 (0.001)
MS (Reduced)	26189	-15.35 (0.001)	0.20 (<0.001)	-	-	-	-	0.06 (0.001)

b: inbreeding depression; A: additive; D: dominance; AA: additive-by-additive; AD: additive-by-dominance; DD: dominance-by-dominance; p: common litter environmental effect; Full: statistical model includes all genetic effects and common litter environmental effect; Reduced: statistical model includes only additive genetic effect (also common litter environmental effect for production traits). NBA: number born alive; NBD: number born dead; NW: number weaned; BF: off-test back fat; WT: off-test weight; MS: off-test loin muscle depth

Genomic prediction

Prediction performance using the full model that included the additive, dominance, and epistatic effects and the reduced model that includes only the additive effect is shown in Figure 1. Figure 1A displays how the prediction accuracies for NBA, WT, and MS were significantly greater using the full model than the reduced model. Compared to the reduced model, the full model's prediction accuracies increased by 15%, 8%, and 1% for NBA, WT, and MS, respectively. BF had the highest prediction accuracy (> 0.5), followed by MS, WT, NBA, NW, then NBD. Prediction accuracies using only the additive effect or the sum of additive and dominance effects estimated from the full model were not improved compared to the reduced model. In general, the regression coefficients of phenotypes on estimated genetic values were not significantly different between the full and reduced models. However, using the full model resulted in substantially less biasedness for BF, as shown in Figure 1B.

Genome-wide association study of the dominance effects

Manhattan plots of the dominance effects for all reproduction and production traits are shown in Figures 2 and 3, respectively. In general, both one-step and two-step strategies detected largely the same dominance signals, but the one-step strategy GWAS resulted in relatively smaller p-values. Herein, only signals that were detected by both approaches were reported. Six SNP markers were detected with dominance effect for NBD on SSC3, SSC4, SSC12, and SSC14, with P-values ranging from 1×10^{-17} to 1×10^{-12} . The marker on SSC12 had the strongest dominance association. However, these dominance signals are likely false positive since a cluster of significant markers was not observed. A strong dominance signal was identified for BF on SSC6 (135-136 Mb) with the smallest P-value = 3×10^{-13} using the two-step GWAS. The estimate of dominance effect for the most significant marker (SSC6: 135,370,427 bp) identified for BF was greater than the additive effect estimate (-1.01 versus 0.70) using the two-step GWAS. All SNPs with significant dominance effect were also found with a significant additive genetic effect based on the two-step approach using 2-df genotypic test in the PLINK software.

Discussion

Although many studies attempted to obtain estimates of nonadditive genetic effects from the total genetic variance in cattle and pigs [4, 13–14, 17, 26–27], the data size used in these studies were relatively small, possibly hampering their identification. Additionally, the number of traits studied was limited. In this study, a larger number of pigs were used (26,735 to 27,647 pigs and 5,338 sows with both genotypes and phenotypes for the production and reproduction traits). In addition, in the current study additive-by-dominance and dominance-by-dominance effects were also investigated. A similar approach was taken by Vitezica et al. [12]. However, the data size used in their study was relatively small (3,619 sows), and only one trait was analyzed (litter size).

Variance decomposition with additive, dominance, and epistatic effects

In general, in the current study dominance genetic variance explained a small proportion of total variation (0 for reproductive traits and 1-2% for production traits), smaller than estimates in pigs reported by Guo et al. [4], Xiang et al. [27], and Norris et al. [28] (4%-10% for NBA and 5% for daily gain and BF). Duroc pigs were used by Guo et al. [4] and Norris et al. [28], while Landrace and Yorkshire pigs were used by Xiang et al. [27]. However, these studies did not include the epistatic effects. Same as our study, Vitezica et al. [12] included the additive-by-additive, additive-by-dominance, and dominance-by-dominance effects in the statistical model for litter size and observed about 1.9% of dominance genetic variance, which is much lower than the estimates reported by the above studies but still larger than the estimates by our research. However, in Vitezica et al. [12] the breed used for the study was not specified, making it challenging to compare results from the current research. Regardless, the above results suggest that excluding the epistatic effects from the statistical model could result in biased estimates of dominance genetic variance. Vitezica et al. [12] also compared a full model that included the epistatic effect versus the reduced model that did not include the epistatic effect using the same data. Indeed, they found that excluding the epistatic effects generated a larger estimate of dominance genetic variance. The same was true for the additive genetic effect, estimates of additive genetic variance were biased upward when excluding the dominance and epistatic effects from the model. This is in agreement with our study (Table 1). In the same paper, Vitezica et al. [12] discussed how variance component estimates could differ if the effects in the model are not orthogonal. The assumption of orthogonality depends on HWE and does not hold in our study. This might have affected our estimates to some degree. This could also explain why the additive effect variance increased when nonadditive effects were excluded from the model, especially for NBA. Joshi et al. [29] used both the method that assumed HWE and the natural and orthogonal interaction approach that did not require HWE for constructing genomic relationship matrices. They found that estimates of variance components were only marginally different whether HWE was attained or not. Still, this had little to no impact on the qualitative outcomes related to the genetic architecture of the trait. The same was true for the traits investigated in our study.

In our study, the proportion of variation explained by epistatic effects was large for reproductive traits, in contrast to what was found by Vitezica et al. [12]. A large percentage of additive-by-additive effect variances were found for NBA and NW (31% and 12%, respectively), while the dominance-by-dominance variance was the largest for NBD (22%). Estimates of the epistatic effects for reproductive traits in pigs are scarce in literature. Vitezica et al. [12] reported a very low percentage of epistatic effect (4%) for litter size. However, no specific information on the pig breed was provided in their study. To further validate our results, two additive-by-additive genomic relationship matrices were compared using two approaches proposed by Jiang and Reif [9] and Vitezica et al. [11], respectively. The method proposed by Vitezica et al. [11] included interactions of loci with themselves (intra-locus interaction), which was excluded by Jiang and Reif [9]. Based on our data, the models using two different additive-by-additive genomic relationship matrices resulted in very similar estimates of variance components (Table 2). This suggests

that when the number of markers is large enough, the effect of intra-locus interactions can be ignored. It is normally accepted that fitness-related traits display relatively high nonadditive variance [30]. Litter size is an important aspect of fitness in mammals and is lowly heritable in pigs [31–33]. Nonadditive genetic effects could be partially responsible for low heritability estimates observed in combination with high total genetic variance. Our studies found a large proportion of nonadditive genetic variance, particularly the epistatic genetic variances for NBA and NW. Literature in swine is scarce, however, Peripato et al. [34] observed that QTL and their interactions explain nearly 49% of the phenotypic variation for litter size in mice.

Table 2

Estimates of inbreeding depression and proportions of variance components in Duroc pigs using alternative genomic additive-by-additive relationship matrices

Trait	Sample size	b (P-value)	A (P-value)	D (P-value)	AA (P-value)	p (P-value)
NBA ¹	3246	-21.26 (< 0.001)	0.05 (0.01)	0 (1)	0.31 (< 0.001)	-
NBA ²	3246	-21.26 (< 0.001)	0.05 (0.01)	0 (1)	0.31 (< 0.001)	-
NBD ¹	3246	-3.25 (0.67)	0.03 (0.05)	0 (1)	0.07 (0.36)	-
NBD ²	3246	-3.25 (0.67)	0.03 (0.05)	0 (1)	0.07 (0.36)	-
NW ¹	3246	-14.06 (0.07)	0.04 (0.02)	0 (1)	0.10 (0.18)	-
NW ²	3246	-14.06 (0.07)	0.04 (0.02)	0 (1)	0.10 (0.18)	-
BF ¹	26189	-15.03 (< 0.001)	0.34 (< 0.001)	0.01 (< 0.001)	0.11 (< 0.001)	0.06 (< 0.001)
BF ²	26189	-15.03 (< 0.001)	0.34 (< 0.001)	0.01 (< 0.001)	0.11 (< 0.001)	0.06 (< 0.001)
WT ¹	26813	-22.81 (< 0.001)	0.08 (< 0.001)	0.02 (< 0.001)	0.16 (< 0.001)	0.11 (< 0.001)
WT ²	26813	-22.81 (< 0.001)	0.08 (< 0.001)	0.02 (< 0.001)	0.16 (< 0.001)	0.11 (< 0.001)
MS ¹	26189	-11.26 (< 0.001)	0.18 (< 0.001)	0.01 (< 0.001)	0.11 (< 0.001)	0.05 (< 0.001)
MS ²	26189	-11.26 (< 0.001)	0.18 (< 0.001)	0.01 (< 0.001)	0.11 (< 0.001)	0.05 (< 0.001)

¹ used the additive-by-additive matrix proposed by Vitezica et al. (2017), ² used the additive-by-additive matrix proposed by Jiang and Reif (2015). b: inbreeding depression; A: additive; D: dominance; AA: additive-by-additive; p: common litter environmental effect; NBA: number born alive; NBD: number born dead; NW: number weaned; BF: off-test back fat; WT: off-test weight; MS: off-test loin muscle depth

Interestingly, large epistatic genetic variances were also observed for the production traits despite large additive effects being found, especially for BF and MS. Su et al. [8] reported that the additive-by-additive genetic effect accounted for 9.5% of total phenotypic variance for growth rate in Duroc pigs, mostly without genotype data. This estimate is very similar to WT (off-test weight, 10%) in our study, a trait related to growth rate. However, the additive-by-dominance effect, which was not included in Su et al. [8], having genetic variance much larger for WT (23%) than the additive-by-additive effect. Joshi et al. [29] observed a sizeable additive-by-additive genetic variance (17% of total phenotypic variance) for body weight at harvest in tilapia. Carcass trait BF showed large estimates of additive-by-additive (9%) and additive-by-dominance variance (9%), and MS showed large estimates of additive-by-dominance variance (17%) and dominance-by-dominance variance (10%). To our knowledge, no previous studies have reported estimates of epistatic effects on carcass traits in pigs. However, in beef cattle, Bolormaa et al. [17] found very low estimates of additive-by-additive variance for several carcass traits, which disagrees with our studies, in which 6% and 4% of additive-by-additive effects were observed for BF and MS in pigs, respectively. This may be because different species and slightly different carcass traits were used. Our findings also provide empirical evidence that nonadditive effects can be considerable for traits with relatively higher heritability, e.g., carcass traits.

Inbreeding depression is widely assumed to be deleterious for economically important traits, thus, can have a significant impact on animal production [35–37]. In the present study, NBA and more prevalently production traits showed a substantial inbreeding depression (Table 1), consistent with results reported by Christensen et al. [35] in pigs. Additionally, Xiang et al. [27] proved analytically that inclusion of genomic inbreeding as a covariate is necessary to obtain correct estimates of dominance variance in the presence of directional dominance. Vitezica et al. [12] also confirmed that genomic models that include nonadditive effects must also consider inbreeding depression to avoid upward bias of estimates of dominance variance for litter size in pigs. This agrees with our results. As shown in Table 3, a consistent reduction of estimates of dominance variance was observed for NBA and production traits, especially for WT, when inbreeding coefficient was included as a fixed covariate in the model. However, the effect of fitting genomic inbreeding on estimates of epistatic variance components for production traits in pigs was previously unknown. For the first time, our results show that the epistatic effects for production traits are largely unaffected by the inclusion of genomic inbreeding. While slight changes of estimates are observed these are likely due to random errors.

Table 3

Estimates of proportions of variance components in Duroc pigs with and without inbreeding coefficient

Trait	Sample size	b (P-value)	A (P-value)	D (P-value)	AA (P-value)	AD (P-value)	DD (P-value)	p (P-value)
NBA (W)	3246	-21.26 (0.01)	0.05 (0.01)	0 (1)	0.31 (< 0.001)	-	-	-
NBA (WO)	3246	-	0.05 (0.01)	0.004 (0.83)	0.30 (< 0.001)	-	-	-
NBD (W)	3246	-3.48 (0.75)	0.03 (0.14)	0 (1)	0.06 (0.50)	0 (1)	0.22 (0.55)	-
NBD (WO)	3246	-	0.03 (0.14)	0 (1)	0.06 (0.48)	0 (1)	0.20 (0.57)	-
NW (W)	3246	-13.95 (0.11)	0.03 (0.10)	0 (1)	0.12 (0.2)	0 (1)	-	-
NW (WO)	3246	-	0.03 (0.10)	0 (1)	0.12 (0.19)	0 (1)	-	-
BF (W)	26189	-15.35 (< 0.001)	0.33 (< 0.001)	0.008 (0.001)	0.09 (< 0.001)	0.09 (0.003)	0 (1)	0.05 (< 0.001)
BF (WO)	26189	-	0.33 (< 0.001)	0.011 (< 0.001)	0.09 (< 0.001)	0.08 (< 0.001)	0 (1)	0.05 (< 0.001)
WT (W)	26813	-23.71 (< 0.001)	0.08 (< 0.001)	0.02 (< 0.001)	0.10 (< 0.001)	0.23 (< 0.001)	0.02 (0.64)	0.10 (< 0.001)
WT (WO)	26813	-	0.08 (< 0.001)	0.03 (< 0.001)	0.10 (< 0.001)	0.22 (< 0.001)	0.02 (0.64)	0.10 (< 0.001)
MS (W)	26189	-12.27 (< 0.001)	0.19 (< 0.001)	0.007 (0.01)	0.05 (< 0.001)	0.17 (< 0.001)	0.10 (0.02)	0.04 (< 0.001)
MS (WO)	26189	-	0.19 (< 0.001)	0.01 (< 0.001)	0.06 (< 0.001)	0.17 (< 0.001)	0.10 (0.02)	0.04 (< 0.001)

b: inbreeding depression; A: additive; D: dominance; AA: additive-by-additive; AD: additive-by-dominance; DD: dominance-by-dominance; p: common litter environmental effect; W: statistical model includes inbreeding coefficient; WO: statistical model does not include inbreeding coefficient; NBA: number born alive; NBD: number born dead; NW: number weaned; BF: off-test back fat; WT: off-test weight; MS: off-test loin muscle depth

Prediction performance

Epistasis did not improve the accuracy of prediction of litter size phenotype in pigs in Vitezica et al. [12]. In contrast, our results show that the addition of dominance and epistatic effects significantly improve the genomic prediction accuracy for NBA (Figure 1). Notably, a large percentage of additive-by-additive effect variance was observed in our study, compared to the 1.6% observed by Vitezica et al. [12]. This

may occur due to two reasons. First, different pig breeds were used (reproductive traits haven't been under selection in Duroc pigs); second, the trait studied were different with total number of pigs born per litter used by Vitezica et al. [12], while our study analyzed total number of pigs born alive per litter. It is noteworthy that total number of pigs born may also include total number of pigs born dead (NBD), which had a much lower estimate of additive-by-additive effect variance than NBA (Table 1), as shown in our study. There were very few studies reporting the genomic prediction of production traits by including epistatic effects in pigs. Su et al. [8] reported that the inclusion of dominance and additive-by-additive effect in a genomic prediction model increased prediction accuracy for daily gain in Duroc pigs. This is consistent with our findings. Genomic prediction accuracy was significantly improved when dominance and epistatic effects were included in the prediction model for off-test weight (WT, Figure 1), a trait similar to daily weight gain. However, in our study, the additive-by-dominance effect also contributed to increased accuracy. To the best of our knowledge, no studies have reported the genomic prediction accuracy, including dominance and epistasis, for carcass traits in pigs. Our results suggest that genomic prediction accuracy for loin muscle depth (MS) can be significantly improved when dominance and epistasis are included in the prediction model (Figure 1). The improvement can be mostly attributed to large additive-by-dominance and dominance-by-dominance effects. Guo et al. [4] reported improved genomic prediction accuracy for backfat and daily gain in Danish Duroc pigs when dominance effect was included in the prediction model. However, epistatic effects were not included, which may have resulted in upward biases for dominance variance estimates. It is noteworthy that in our study, predictions using only additive effect or the sum of additive and dominance effect estimated from the full model did not improve the accuracy for any of the traits. Our results suggest that the significant improvement of prediction accuracy for NBA, WT, and MS is to be attributed mainly to the epistatic effects, however further research is needed to confirm our findings.

Application of genomic selection with epistatic effects

Inclusion of nonadditive effects in the genomic prediction model can increase the prediction accuracy of phenotypes. It can also be used to predict the performance of future mating between two individuals. Studies have shown that inclusion of dominance effects in the matting program for dairy cattle can increase rates of genetic gain [13–14]. The inclusion of dominance effect requires probabilities of the genotypes, e.g. A_1A_1 , A_1A_2 , and A_2A_2 , for the combination of the two individuals and the estimates of the additive and dominance effects for the same marker [38]. However, the use of models that include more complex interactions such as epistasis in the matting program is more challenging because the predicted performance of a mate should integrate the predictive performance over all possible future genotypic configurations of the expected progeny. One solution to this is to select the top markers with significant epistatic interaction and incorporate these markers in the matting program. Carlborg et al. [39] and Carter et al. [40] showed that epistatic interactions between four loci mediated a considerably higher response to selection of growth in chicken than predicted by a single-locus model. Furthermore, Martini et al. [41] also showed that selecting just a subset of the largest epistatic interaction effects has the potential to improve predictive ability. However, further studies are needed to evaluate the epistatic effects between markers across the genome using our data.

Genome-wide association study of dominance effect

Several GWAS studies have been performed to identify the dominance QTL for reproduction and production traits in pigs [18, 42–43]. Specifically, Coster et al. [42] detected significant markers with dominance effect for NBA and litter size in pigs of unspecified breed. Lopes et al. [43] identified one significant marker with dominance effect for teat number in Landrace pigs. Yang et al. [18] further detected two significant markers associated with dominance effect on growth rate in large white pigs. However, these studies were based on relatively small data size, which may result in lower statistical power to detect the dominance signals. A larger phenotypic and genotypic dataset was used in our study. Additionally, both one-step and two-step GWAS strategies were applied in this study to validate the results (Figures 2 and 3). Our study suggested that the two-step GWAS approach was comparable to the one-step full mixed model. The two-step method used in this study was proved to be an efficient alternative to detect dominance effects when fast implementations of full mixed models are not available. Interestingly, a strong dominance signal was identified on SSC6 (135-136 Mb) for BF with high LD between SNPs in this QTL region ($r^2 = 0.42-0.75$). The same trait was investigated by Yang et al. [18], who did not find any dominance signals associated with BF. This may be because, first, a much smaller dataset was used by Yang et al. [18] than by our study; second, different pig breeds were used, Yang et al. [18] studied the large white pigs, whereas Duroc pigs were used in our research. It is known that the large white pigs have been selected mainly for reproductive traits. In contrast, the Duroc breed has been selected for production traits for a long time, resulting in different genetic architecture between the two populations.

A candidate gene *AK5* was harbored in the QTL region (135-136 Mb on SSC6) identified for BF. *AK5* was found to be strongly associated with increased body fat and may be helpful in the treatment of obesity [44]. From our results, *AK5* may also play a role in determining the back fat of Duroc pigs. Larger estimate of dominance effect than additive effect for the most significant maker for BF suggested overdominance for this gene. Furthermore, the QTL region identified for BF overlapped with previously identified QTL in pigs associated with back fat at the 10th and last rib, average backfat thickness, marbling score, and intramuscular fat content, indicating a significant association between this QTL and backfat.

Conclusions

A large dataset was used to investigate the contribution of dominance and three epistatic effects to reproduction and production traits in Duroc pigs. Large estimates of additive-by-additive, additive-by-dominance, and dominance-by-dominance genetic variance were found for the reproduction and production traits. In contrast, the dominance variance was relatively small for all traits but still significant for all production traits. Including nonadditive effects in the genomic prediction model significantly improved prediction accuracy for number born alive, off-test loin muscle depth, and body weight. A strong dominance association signal with off-test back fat was identified near *AK5*.

Abbreviations

BF: off-test back fat

GWAS: genome-wide association study

GRM: genomic relatedness matrices

HWE: Hardy–Weinberg equilibrium

MS: off-test loin muscle depth

NBA: number born alive

NBD: number born dead

NW: number weaned

QTL: quantitative trait loci

SNP: single nucleotide polymorphism

WT: off-test body weight

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Data Availability

All the data supporting the results of this study are included in the article. Scripts for all the analyses are available at https://github.com/jiang18/projects/tree/main/swine_varcomp. The raw phenotypic and genotypic data used in this study are the sole property of SPG (Rose Hill, NC). Restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are, however, available from the authors upon reasonable request and with permission of SPG. A request to SPG for accessing data may be sent to Kent Gray, General Manager (kgray@smithfield.com).

Competing interests

The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

This work was supported by JJ's startup grant from North Carolina State University.

Authors' contributions

JJ and CM conceived and designed the study. JC and JJ performed the data analysis. JC and JJ wrote the manuscript. FT and JH contributed materials. All authors read and approved the final manuscript.

Acknowledgements

We thank the Smithfield Premium Genetics (Rose Hill, NC) breeding company for providing access to their phenotypic, pedigree, and genomic data.

References

1. Muewissen THE, Hayes BJ, Goddard ME. Prediction of total genetic effect using genome-wide dense marker maps. *Genetics*. 2001;157:1819–29.
2. Hayes BJ, Bowman PJ, Chamberlain AL, Goddard ME. Invited review: Genomic selection in daily cattle: progress and challenges. *J Dairy Sci*. 2009;92:433–43.
3. VanRaden PM. Efficient methods to compute genomic predictions. *J Dairy Sci*. 2008;91(11):4414–23.
4. Guo X, Christensen OF, Ostersen T, Wang Y, Lund MS, Su G. Genomic prediction using models with dominance and imprinting effects for backfat thickness and average daily gain in Danish Duroc pigs. *Genet Sel Evol*. 2016;48:67.
5. Jiang J, Shen B, O'Connell JR, VanRaden PM, Cole JB, Ma L. Dissection of additive, dominance, and imprinting effects for production and reproduction traits in Holstein cattle. *BMC Genom*. 2017;18:425.
6. Carlborg Ö, Haley CS. Epistasis: too often neglected in complex trait studies? *Nat Rev Genet*. 2004;5:618–25. <https://doi.org/10.1038/nrg1407>.
7. Mackay TF. Epistasis and quantitative traits: using model organisms to study gene–gene interactions. *Nat Rev Genet*. 2014;15:22–33. <https://doi.org/10.1038/nrg3627>.

8. Su G, Christensen OF, Ostersen T, Henryon M, Lund MS. Estimating additive and nonadditive nonadditive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers. *PLoS One*. 2012;7:e45293.
9. Jiang Y, Reif JC. Modeling epistasis in genomic selection. *Genetics*. 2015;201:759–68.
10. Lynch M, Walsh B. *Genetics and analysis of quantitative traits*. New York: Sinauer associates Inc; 1998. p. 233.
11. Vitezica ZG, Legarra A, Toro MA, Varona L. Orthogonal estimates of variances for additive, dominance, and epistatic effects in populations. *Genetics*. 2017;206:1297–307.
12. Vitezica ZG, Reverter A, Herring W, Legarra A. Dominance and epistatic genetic variances for litter size in pigs using genomic models. *Genet Sel Evol*. 2018;50:71. <https://doi.org/10.1186/s12711-018-0437-3>.
13. Sun C, VanRaden P, O’Connell J, Weigel K, Gianola D. Mating programs including genomic relationships and dominance effects. *J Dairy Sci*. 2013;96(12):8014–23.
14. Aliloo H, Pryce J, González-Recio O, Cocks B, Goddard M, Hayes B. Including nonadditive genetic effects in mating programs to maximize dairy farm profitability. *J Dairy Sci*. 2016;100:1203–22.
15. VanRaden P. Practical implications for genetic modeling in the genomics era. *J Dairy Sci*. 2016;99(3):2405–12.
16. Goddard ME, Hayes BJ. Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nat Rev Genet*. 2009;10:381–91.
17. Bolormaa S, Pryce JE, Zhang Y, Reverter A, Barendse W, Hayes BJ, et al. Nonadditive genetic variation in growth, carcass and fertility traits of beef cattle. *Genet Sel Evol*. 2015;47:26.
18. Yang W, Wu J, Yu J, Zheng X, Kang H, Wang Z, et al. A genome-wide association study reveals additive and dominance effects on growth and fatness traits in large white pigs. *Anim Genet*. 2021;52:749–53. <https://doi.org/10.1111/age.13131>.
19. Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T, Lee DH, et al. BLUPF90 and related programs (BGF90). In: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*: 19-23 August 2002. Montpellier; 2002. p. 743–4.
20. VanRaden PM, Null DJ, Sargolzaei M, Wiggans GR, Tooker ME, Cole JB, et al. Genomic imputation and evaluation using high-density Holstein genotypes. *J Dairy Sci*. 2013;96(1):668–78.
21. O’Connell JR. *MMA User Guide*. Available: <http://edn.som.umaryland.edu/mmap/index.php>. Accessed 8 August 2021. 2015.
22. Vitezica ZG, Varona L, Legarra A. On the additive and dominant variance and covariance of individuals within the genomic selection scope. *Genetics*. 2013;195:1223–30.
23. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am. J. Hum*. 2007;81.
24. Smedley D, Haider S, Ballester B, Holland R, London D, Thorisson G, et al. BioMart – biological queries made easy. *BMC Genom*. 2009;10:22. <https://doi.org/10.1186/1471-2164-10-22>.

25. Hu ZL, Fritz ER, Reecy JM. AnimalQTLdb: a livestock QTL database tool set for positional QTL information mining and beyond. *Nucleic Acids Res.* 2007;35:D604-9.
26. Wittenburg D, Melzer N, Reinsch N. Genomic additive and dominance variance of milk performance traits. *J Anim Breed Genet.* 2015;132(1):3–8.
27. Xiang T, Christensen OF, Vitezica ZG, Legarra A. Genomic evaluation by including dominance effects and inbreeding depression for purebred and crossbred performance with an application in pigs. *Genet Sel Evol* 2016. 48(1):92.
28. Norris D, Varona L, Ngambi JW, Visser DP, Mbajiorgu CA, Voordewind SF. Estimation of the additive and dominance variances in SA Duroc pigs. *Livest Sci.* 2010;131:144–7.
29. Joshi R, Meuwissen THE, Woolliams JA, Gjøen HM. Genomic dissection of maternal, additive and nonadditive genetic effects for growth and carcass traits in Nile tilapia. *Genet Sel Evol.* 2020;52:1.
30. Hill WG, Goddard ME, Visscher PM. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet.* 2008;4(2):e1000008.
31. Bidanel JP. Biology and genetics of reproduction. In: Rothschild MF, Ruvinsky A, editors. *The genetics of the pig.* 2nd ed. London: AB International; 1998. pp. 313–43.
32. Nielsen B, Su G, Lund MS, Madsen P. Selection for increased number of piglets at d 5 after farrowing has increased litter size and reduced piglet mortality. *J Anim Sci.* 2013;91:2575–82.
33. Guo X, Christensen OF, Ostersen T, Wang Y, Lund MS, Su G. Improving genetic evaluation of litter size and piglet mortality for both genotyped and nongenotyped individuals using a single-step method. *J Anim Sci.* 2015;93:503–12.
34. Peripato AC, De Brito RA, Matioli SR, Pletscher LS, Vaughn TT, Cheverud JM. Epistasis affecting litter size in mice. *J Evol Biol.* 2004;17:593–602.
35. Christensen K, Jensen P, Jørgensen JN. A note on effect of inbreeding on production traits in pigs. *Anim Sci.* 1994;58:298–300.
36. Smith LA, Cassell BG, Pearson RE. The effects of inbreeding on the lifetime performance of dairy cattle. *J Dairy Sci.* 1998;81:2729–37.
37. Howard JT, Pryce JE, Baes C, Maltecca C. Invited review: Inbreeding in the genomics era: Inbreeding, inbreeding depression, and management of genomic variability. *J Dairy Sci.* 2017;100:6009–24. [10.3168/jds.2017-12787](https://doi.org/10.3168/jds.2017-12787).
38. Varona L, Legarra A, Toro MA, Vitezica ZG. Non-additive effects in genomic selection. *Front. Genet.* 2018. 9:78. doi: [10.3389/fgene.2018.00078](https://doi.org/10.3389/fgene.2018.00078).
39. Carlborg O, Jacobsson L, Ahgren P, Siegel P, Andersson L. Epistasis and the release of genetic variation during long term selection. *Nat Genet.* 2006;38:418420.
40. Carter AJR, Hermisson J, Hansen TF. The role of epistatic gene interactions in the response to selection and the evolution of evolvability. *Theor Pop Biol.* 2005;68:179–96.
41. Martini JWR, Wimmer V, Erbe M, Simianer H. Epistasis and covariance: how gene interaction translates into genomic relationship. *Theor Appl Genet.* 2016;129(5):963–76.

<https://doi.org/10.1007/s00122-016-2675-5>.

42. Coster A, Madsen O, Heuven HC, Dibbits B, Groenen MA, et al. The imprinted gene DIO3 is a candidate gene for litter size in pigs. PLoS ONE. 2012;7:e31825.

43. Lopes MS, Bastiaansen JW, Harlizius B, Knol EF, Bovenhuis H. A genome-wide association study reveals dominance effects on number of teats in pigs. PLoS One. 2014;9:e105867.

44. Powell DR, Revelli JP, Doree DD, DaCosta CM, Desai U, Shadoan MK, et al. High-throughput screening of mouse gene knockouts identifies established and novel high body fat phenotypes. Diabetes, metabolic syndrome and obesity: targets and therapy. 2021. 14, 3753–3785.
<https://doi.org/10.2147/DMSO.S322083>.

Figures

Figure 1

Prediction performance of two models for reproduction and production traits in Duroc pigs

* indicates P-value is less than 0.05; A: additive effect; D: dominance effect; AA: additive-by-additive effect; AD: additive-by-dominance effect; DD: dominance-by-dominance effect; NBA: number born alive; NBD: number born dead; NW: number weaned; BF: off-test back fat; WT: off-test weight; MS: off-test loin muscle depth

Figure 2

Manhattan plots for associations of marker dominance effects with reproductive traits in Duroc pigs

Figure 3

Manhattan plots for associations of marker dominance effects with production traits in Duroc pigs