ANIMAL GENETICS AND GENOMICS

Impact of inclusion rates of crossbred phenotypes and genotypes in nucleus selection programs

Garrett M. See,¹ Benny E. Mote, and Matthew L. Spangler

Department of Animal Science, University of Nebraska-Lincoln, Lincoln, NE 68588

¹Corresponding author: garrett.see@huskers.unl.edu

Abstract

Numerous methods have been suggested to incorporate crossbred (CB) phenotypes and genotypes into swine selection programs, yet little research has focused on the implicit trade-off decisions between generating data at the nucleus or commercial level. The aim of this study was to investigate the impact of altering the proportion of purebred (PB) and CB phenotypes and genotypes in genetic evaluations on the response to selection of CB performance. Assuming CB and PB performance with moderate heritabilities ($h^2 = 0.4$), a three-breed swine crossbreeding scheme was simulated and selection was practiced for six generations, where the goal was to increase CB performance. Phenotypes, genotypes, and pedigrees for three PB breeds (25 and 175 mating males and females for each breed, respectively), F_1 crosses (400 mating females), and terminal cross progeny (2,500) were simulated. The genome consisted of 18 chromosomes with 1,800 quantitative trait loci and 72k single nucleotide polymorphism (SNP) markers. Selection was performed in PB breeds using estimated breeding value for each phenotyping/genotyping strategy. Strategies investigated were: 1) increasing the proportion of CB with genotypes, phenotypes, and sire pedigree relationships, 2) decreasing the proportion of PB phenotypes and genotypes, and 3) altering the genetic correlation between PB and CB performance (r_{pc}). Each unique r_{pc} scenario and data collection strategy was replicated 10 times. Results showed that including CB data improved the CB performance regardless of r_{pc} or data collection strategy compared with when no CB data were included. Compared with using only PB information, including 10% of CB progeny per generation with sire pedigrees and phenotypes increased the response in CB phenotype by 134%, 55%, 33%, 23%, and 21% when r_{pc} was 0.1, 0.3, 0.5, 0.7, and 0.9, respectively. When the same 10% of CB progeny were also genotyped, CB performance increased by 243%, 54%, 38%, 23%, and 20% when the r_{pc} was 0.1, 0.3, 0.5, 0.7, and 0.9, respectively, compared with when no CB data were utilized. Minimal change was observed in the average CB phenotype when PB phenotypes were included or proportionally removed when CB were genotyped. Removal of both PB phenotypes and genotypes when CB were genotyped greatly reduced the response in CB performance. In practice, the optimal inclusion rate of CB and PB data depends upon the genetic correlation between CB and PB animals and the expense of additional CB data collection compared with the economic benefit associated with increased CB performance.

Key words: commercial data, genotyping strategy, phenotyping strategy, swine

Introduction

Swine production systems that practice crossbreeding are designed to take advantage of increased performance in crossbred (CB) progeny compared with purebred (PB) offspring

through heterosis. Crossbreeding is flexible in the production of different end products for different markets by changing CB composition (Dickerson, 1973) and benefits from breed complementarity by crossing lines specialized for different

Abbreviations

BIAS	regression coefficients produced from the regression of true breeding value on estimated breeding value
CB	crossbred
CB: Gen	when crossbred contributed phenotypes, genotypes and sire pedigree information to genetic evaluations
CB: Ped	when crossbreds contributed phenotypes and sire pedigree information to genetic evaluations
CCPS	combined crossbred and purebred selection
EBV	estimated breeding value
ΔG	change in average true breeding value of crossbred progeny
LD	linkage disequilibrium
N_e	effective population size
PB	purebred
r_{pc}	the genetic correlation between purebred and crossbred performance
SNP	single nucleotide polymorphism
ssGBLUP	single-step genomic best linear unbiased prediction
TBV	true breeding value

categories of traits (Wientjes and Calus, 2017). Traditionally, selection takes place in PB parental lines, while breeding goal traits are observed in CB progeny. If the genetic correlation between PB and CB performance (r_{pc}) is below 1, PB performance could be a poor indicator of CB performance. In many traits relevant to swine production, the r_{pc} has been found to be below 1 and often less than 0.7 (Wientjes and Calus, 2017). The r_{pc} can be less than 1 due to trait definitions, genotype by environment or genotype by genotype interactions.

Wei and van der Werf (1994) proposed a combined CB and PB selection (CCPS) method to maximize CB performance by including both CB and PB information into genetic evaluations. Using only pedigree relationships, CCPS was shown to be the optimal method in increasing the genetic response in CB performance. When $r_{pc} < 0.7$, CCPS is more effective at increasing CB performance. Despite these positive findings, few breeding schemes incorporated CB performance into genetic evaluations largely due to the increased cost in phenotype collection and pedigree tracking. Dekkers (2007) proposed a modification to CCPS using genomic selection, which fits breed-specific allele substitution effects to avoid the need for tracking pedigrees. To date, many of the studies have focused on the accuracy of genomic prediction of CB performance in PB populations using simulated (Ibáñez-Escriche et al., 2009; Zeng et al., 2013) and real data (Hidalgo et al., 2015; Sewel et al., 2018; Duenk et al., 2019). When the composition of the reference populations was half of the CB records and the r_{pc} was moderate to low, prediction accuracy was equal to or greater than using only PB data (van Grevenhof and van der Werf, 2015). Further studies have extended the CCPS method (Christensen et al., 2014, 2019; Tusell et al., 2016) and evaluated the structure of reference populations using genomic selection in CCPS (van Grevenhof and van der Werf, 2015). Yet, few studies have investigated the response in selection when altering the rate of CB progeny inclusion in a

selection program where PB candidates are selected for CB performance (Esfandyari et al., 2015a; van Grevenhof and van der Werf, 2015). Esfandyari et al. (2015a) reported that including CB genotypes in the reference population used to predict CB performance increased the response to selection compared with a PB-only reference population. Although additional genotyping and phenotypic data collection of CB animals could increase CB performance, these additional data increase costs of the genetic program. Currently, no studies have investigated removal of PB genotypes and/or phenotypes when proposing to increase CB data in CCPS.

The objectives of this study were to investigate the response in CB performance when, 1) increasing the proportion of CB with genotypes, phenotypes, and sire pedigree information; 2) decreasing the proportion of PB phenotypes and genotypes, and 3) altering the r_{pc} .

Methods

All data used in this study were simulated, thus animal care and use approval was not required.

Simulation overview

The simultaneous simulation of genotypic and phenotypic information of historic, PB, and CB populations utilized in this study was performed using AlphaSimR (Gaynor et al., 2019). To reduce computational time required, all developed populations and breeds had a litter size of 4, resulting in two males and two females. The simulation of genotypes and phenotypes follows these general steps: 1) simulation of founder haplotypes and assignment of founder haplotypes to founder animals, 2) assignment of segregating sites to quantitative trait loci (QTL) or single nucleotide polymorphism (SNP) represented on the SNP chip, 3) assignment of additive QTL effects and calculation of subsequent true breeding values (TBVs), and 4) breed formation and development through generations of random and selective mating. After initial founder breeds were simulated, a three-way crossbreeding system was mimicked using one terminal and two maternal PB lines. First, maternal lines were crossed to produce F_1 females to be used as dams of CB animals in subsequent generations. One generation after F_1 females were simulated, they were crossed with the terminal sire breed to produce CB animals. The goal of the breeding scheme was to increase CB performance through selection on estimated breeding values (EBVs). Using the previously mentioned simulation framework, a number of r_{pc} scenarios were simulated, which altered the additive genetic relationship between PB and CB performance. Further, data inclusion strategies were implemented in each r_{pc} scenario, which differed in the proportions of PB and CB data (pedigree, phenotypes, and genotypes) included in the reference population used for genetic evaluations. Data inclusion strategies were then compared by response in CB performance, EBV accuracy, and EBV bias after six generations of selection. Results are the average over 10 replicates. Founder genomes were consistent across strategies within a replicate.

Populations

Different breeds were simulated and utilized in the three-way crossbreeding scheme to emulate a swine breeding program. Historical haplotypes were simulated using the Markovian Coalescence Simulator (Chen et al., 2009). Simulation of the historical haplotypes began 1,000,000 generations prior to the final historical generation with an effective population size (N_e)

of 100,000. The N_e was then gradually reduced to 100 in the final historical generation. Following creation of the historical haplotypes, founder individuals were simulated using AlphaSimR (Gaynor et al., 2019) by randomly sampling founder haplotypes with replacement to create 4,800 founder animals. Founder animals were randomly split into three parental PB breeds (hereafter breeds A, B, and C) each with 1,600 total founders with equal numbers of males and females. Subsequently, within each breed, random mating was performed for 15 discrete generations (generation -66 to -51) as the population size was reduced until each PB population was comprised of 25 males and 175 females. Following the genetic bottleneck, breeds were randomly mated within breed for 50 generations (generation -51

to -1) maintaining a constant population size. The PB mating population size remained constant for the remainder of the simulation (generation 0 to 5). The population sizes of breeds A, B, and C were chosen to ensure that when all PB and CB animals were genotyped and included in the genetic evaluations, the reference population was approximately three-quarters of PB and one-quarter of CB. Further, population sizes were chosen to be large enough to prevent the loss of genetic diversity within breed, yet small enough to ensure that selection could continue for several generations with minimal computational requirements. The genetic contributions of each breed and cross are outlined in Figure 1 from development of the founder population until the end of the simulation.

Figure 1. Crossbreeding scheme. CB progeny began being produced in generation 2 (generation in bold numbers). The number of mating male (M) and female (F) parents per breed and cross (A, B, and C) are in parenthesis. Black arrows denote breeds that produce replacements within a breed for the next generation, while red and blue denote only dams or only sires utilized in producing the next generation, respectively.

In the next step, prior to the initiation of the three-way crossbreeding program, breeds A, B, and C were selectively mated within breed, using EBV for PB performance, for one generation (generation 0) to better differentiate breeds. Founders of the pedigree (i.e., the first generation in the pedigree) were PB animals produced in the last two generations of random mating (generations –2 and –1). The first animals to have genotypic and phenotypic information used for genetic evaluation were sires and dams of generation 0. Beginning from generation 0 with selective mating until the end of the simulation, generations were not discrete. Dam replacement rates for breeds A, B, and C were 0.6, 0.6, and 0.5, respectively. Sire replacement rates for breeds A, B, and C were 0.75, 0.75, and 0.7, respectively. Selection intensities for the females and males from breeds A, B, and C were 0.93 and 1.89, 0.93 and 1.89, and 1.06 and 1.92, respectively. Breed C was used as the terminal sire breed, which in practice would have a greater selection intensity in commercial swine breeding schemes (Bichard, 1971). To accomplish the greater selection intensity in breed C, given all populations were of fixed and equal size in this simulation, sire and dam replacement rates were lower in breed C compared with breeds A and B. Selection of parents for the next generation was determined using EBV produced from within breed genetic evaluations. Data utilized to produce EBV in generation 0 consisted of all available phenotypic, genotypic, and pedigree relationships for a given breed. PB selection candidates were simulated from 175 matings (one litter per female), which produced 700 offspring per generation in each PB breed from generation 0 until the end of the simulation.

For mimicking a three-way crossbreeding program, F₁ progeny were simulated by crossing 25 sires from A and 250 dams from B beginning from generation 1. Dams from B used to produce AB animals were females comprised from two sources/generations: 1) dams that were used to make PB matings in the previous generation and, due to having lower EBV, were not retained to make PB matings in the current generation and 2) selection candidates from the current generation not chosen to make PB matings due to having a lower EBV than their contemporaries. Sires from A and dams from B were mated 250 times, resulting in 1 litter per dam per generation producing 1,000 AB offspring. Of the resulting 1,000 AB progeny, 400 females were selected based on parental average EBV to make up the first generation of AB dams. The replacement rate of AB females was 0.65. Beginning from generation 2 and continuing until the end of the simulation, terminal CB progeny were simulated by crossing dams from AB ($n = 400$) with sires from C ($n = 50$). Sires from C included the same sires used to create PB offspring in a given generation, while, dams of C(AB) were females that were simulated in the previous generation (i.e., dams of C[AB] in generation 2 were simulated in generation 1). Terminal CB progeny dams from AB and sires from C were crossed 400 times (one litter per dam) to produce 1,600 terminal CB progeny.

Genotypes and phenotypes

The genome was simulated to mimic the *Sus Scrofa* genome with a SNP density that is comparable to that of a 60k SNP chip. The genome included 18 chromosomes each with genetic length of 1 Morgan and physical length of 1×10^8 bp. Each simulated chromosome had 4,100 biallelic segregating sites, and 100 were randomly chosen to become nonoverlapping QTL, while the remaining became SNP. A SNP chip was simulated by using the 4,000 segregating loci per chromosome chosen to be SNP on the marker panel for a total of 72,000 markers. The selected

marker set was common among all breeds. Mutation rate was set to 2.5×10^{-8} . SNPs with minor allele frequency less than 0.01 were not used for single-step genomic best linear unbiased prediction (ssGBLUP).

All PB and CB observed traits were simulated under additive gene action free of dominance and epistasis. Simulation resulted in two different traits, the first relating to PB performance (observed in breeds A, B, and C) and the second relating to CB performance (observed in breed C[AB]). The phenotypic variance for both PB and CB performance was set to 10, with heritabilities of 0.4. The r_{pc} varied depending upon the genetic correlation scenario described later in the Scenarios Evaluated section. Residual covariance between PB and CB performance was set to zero as each animal would only have a phenotype for one of the traits. True QTL effects for both traits were simulated as described by Faux et al. (2016). For each animal, TBVs were calculated by summing the breeding values at each QTL locus.

Datasets

For genetic evaluations, the number of records included in an evaluation prior to crossbreeding remained constant for all phenotyping/genotyping strategies and genetic correlation scenarios. The number of PB selection candidates or CB animals with phenotypic or genotypic information included in a genetic evaluation in a given generation after generation 1 is shown in Figure 2. Prior to the production of CB animals in generation 2, all PB selection candidates were included in genetic evaluations. In brief, the pedigree included all PB animals (breeds A, B, and C) produced from generation –2 to 5. Animals belonging to breed AB were not included in the pedigree and did not have recorded phenotypes or genotypes. Individuals from C(AB) were included in the pedigree, yet only sire relationships were considered to be known. Phenotypes observed in breeds A, B, C, and C(AB) were used in the genetic evaluation. The number of CB and PB genotypes included in an analysis depended upon the data inclusion strategy. In general, data available to estimate EBV of selection candidates consisted of the performance records of the selection candidates themselves, performance records from collateral PB (i.e., half- and full-sib) and CB (i.e., half-sib) relatives in the same generation, and performance records from both PB and CB relatives that were produced in prior generations. Due to the simulated breeding scheme, the relationship between PB and CB animals differed across the three PB breeds. The lag between selection candidates of breeds A and B and direct CB descendants was two generations, whereas this lag was only one generation between selection candidates from breed C and CB offspring. The breeding scheme resulted in coefficients of relationship between the selection candidates and CB animals with performance information in an evaluation ranging from 0.0625 to 0.25. By generation 5, when all PB selection candidates and CB animals were included in the genetic evaluation, there were 13,200 phenotypic and genotypic records from breeds A, B, and C and 6,400 phenotypic and genotypic records from C(AB).

Scenarios evaluated

Scenarios that were evaluated can be differentiated by three factors: 1) inclusion of CB genotypes, phenotypes, and sire pedigree relationships in genetic evaluations, 2) removal of PB phenotypes and genotypes in genetic evaluations, and 3) r_{pc} . Values of r_{pc} investigated were 0.1, 0.3, 0.5, 0.7, and 0.9. The proportion and source of genotypic and phenotypic data were evaluated in two categories: first, with CB sire pedigree relationships and phenotypes (CB: Ped) or second, with CB

Figure 2. The number of PB and CB phenotypes (Pheno), genotypes (Geno), and pedigree relationships (Pedig) from a single generation included in a genetic evaluation for a given CB progeny inclusion strategy (expressed as proportion of CB progeny included in an analysis), when a proportion of CB were included in an analysis with CB: Ped or with CB: Gen.

sire pedigree relationships, genotypes, and phenotypes (CB: Gen). Data density of PB either remained constant with all PB candidates included in the evaluation or was reduced as the density of CB data increased (Figure 2). The choice of which PB selection candidates had genotypic and/or phenotypic records removed was random. In each replicate for a given genetic correlation scenario, phenotypic, genotypic, and pedigree information from all simulated animals, prior to generation 1, remained constant for each phenotyping/genotyping strategy. Data inclusion strategies were evaluated by the response to selection in CB phenotype, accuracy, and prediction bias of PB EBV for CB performance.

Genetic evaluation and selection

Selection was performed using EBV produced from either single- or two-trait models, while mate allocation was performed at random excluding matings, which resulted in an accumulation of inbreeding greater than 12.5%. In generations 0 and 1, PB candidates were selected by EBV for PB performance produced from single-trait models. In generation 0, single-trait models for PB performance utilized data within a breed to generate the respective breed EBV. Progressing to generation 1, EBVs were produced using a multi-breed single-trait model that included all three breeds PB performance. From generation 2 until the end of the simulation, PB candidates were selected based on EBV for CB performance produced from two-trait models including varying proportions of PB and CB data depending on the scenario.

Univariate and bivariate models were fitted using ssGBLUP via BLUPf90 application programs (Misztal et al., 2002). The model for the i th trait in both the univariate (PB performance) and bivariate (PB and CB performance) case was:

$$y_i = 1\mu_i + X_i b_i + Z_i u_i + \epsilon_i$$

where for the i th trait, y is the vector of phenotypic records, b is the vector of fixed effects (single breed model: overall mean and generation; multiple breed model: overall mean, generation, proportion of breed A, B, and C fit as covariates), u is a vector of additive genetic effects, ϵ is a vector of residuals, and X and Z are known incidence matrices relating fixed and random effects to each record. In the univariate model, additive genetic effects were assumed to be distributed normal with mean 0 and variance $H\sigma_{a1}^2$, where σ_{a1}^2 is the additive genetic variance associated with PB performance. In the bivariate model, the vector of genetic effects, $u = [u'_1, u'_2]$, was assumed to be distributed multivariate normal with mean 0 and variance $\Phi \otimes H$, where $\otimes$ is the Kronecker product and Φ is the additive genetic (co)variance matrix of PB and CB performance. The vector of residuals, $\epsilon = [\epsilon'_1, \epsilon'_2]$, was assumed to be distributed multivariate normal with mean 0 and variance $R \otimes I$, where I is an identity matrix and R is the residual (co)variance matrix. True variance components were used in ssGBLUP evaluations and were constant throughout the simulation.

In order to calculate H , the genomic relationship matrix (G) was first calculated using the BLUPf90 application programs

(Misztal et al., 2002). Lourenco et al. (2016) suggested that G should not be centered or scaled based upon breed-specific allele frequencies in multibreed ssGBLUP; thus, G was calculated in accordance with VanRaden (2008) as follows:

$$G = \frac{WW'}{2 \sum p_i(1 - p_i)}$$

where W is calculated by the subtraction of P from M , P is a matrix of allele frequencies, and M is a matrix of alleles with dimension n individuals by m markers, and p is the allele frequency of the i th allele. Then, 95% of G was then blended with 5% of the numerator relationship matrix for genotyped animals before inverse pedigree and genomic relationships matrices were used to create H (Aguilar et al., 2010).

Bias of EBV was calculated at the end of the simulation using the `lm` function in R. Bias (BIAS) was calculated for each scenario and PB breed using the following linear model:

$$TBV2_{ijk} = \mu + \beta(EBV2_i) + Gen_j + Rep_k + e_{ijk}$$

where $TBV2_{ijk}$ is the TBV for CB performance of the i th PB animal born in the j th generation from the k th replicate, μ is the overall mean, $EBV2_i$ is the EBV for CB performance, β is the regression coefficient associated with $EBV2$ where $E[\beta] = 1$, Gen_j is the j th generation birth effect, Rep_k is the k th replicate effect, and e_{ijk} is the random residual.

Results and Discussion

Characteristics of simulated data

Breed differentiation between breeds A, B, and C were assessed in generations 1 and 5 using projections of genomic relationships in a principal component analysis in the first replicate of each genetic correlation scenario. The first two principal components were plotted and showed the three parental breeds to be separated (Supplementary Figure S1). Average absolute difference in allele frequencies between the three PB breeds across all r_{pc} scenarios and replicates was 0.24. The estimated

Wright's F_{ST} statistic was estimated from all genotypes of PB animals from generation 0 using the Weir and Hill method (Weir and Hill, 2002) in R using the BEDASSLE package (Bradburd and Bradburd, 2013). The average F_{ST} between the three PB breeds across all r_{pc} scenarios and replicates was 0.25. The linkage disequilibrium (LD) decay of the simulated swine population was similar to previous reports in real swine populations (Badke et al., 2012; Supplementary Figure S2).

Response to selection

Including CB information

Primary comparisons among data inclusion strategies and r_{pc} scenarios were made using the response in CB performance (Figure 3). Including a proportion of CB phenotypes and sire pedigrees, with genotypes or without, in genetic evaluations increased the response in CB performance by generation 5 regardless of r_{pc} (Figure 3). With complete PB data, including 10% of CB progeny phenotypes without genotypes increased CB performance by 134%, 55%, 33%, 23%, and 21% when the r_{pc} was 0.1, 0.3, 0.5, 0.7, and 0.9, respectively, compared with when no CB data entered evaluations (Figure 3 CB: Ped). Similarly, when the same 10% of CB progeny were also genotyped, CB performance increased by 243%, 54%, 38%, 23%, and 20% when the r_{pc} was 0.1, 0.3, 0.5, 0.7, and 0.9, respectively, compared with when no CB data entered evaluations (Figure 3 CB: Gen). Including CB records resulted in an expected increase in CB performance through the selection of PB animals for CB performance. These findings are in agreement with previous simulation studies, which concluded that the response to selection for CB performance is greater when PB animals are selected on CB performance (Dekkers, 2007; Kinghorn et al., 2010; Zeng et al., 2013). Marginal increases in CB performance were seen as the rate of CB inclusion increased and PB data remained constant. The average increase in CB performance across all r_{pc} scenarios from 10% to 100% CB inclusion without genotypes was 15%, which was significantly lower than the increase from 0% to 10% (51%). An increase of CB inclusion with genotypes from 10% to 30% and 70% to 90% across all r_{pc} scenarios resulted in an increase in CB performance of 13% and 1.6%, respectively. These results suggest that the largest improvements in CB

Figure 3. Average CB progeny phenotype in generation 5 when a proportion of CB were included in an analysis with CB: Ped or CB: Gen. Different color bars denote the proportion of CB records included in analyses. Results were averaged over 10 replicates.

performance are from the first 10% to 30% of CB phenotypes/genotypes and improvements thereafter are less dramatic. Given increased CB data collection would incur additional costs, such marginal increases in CB performance at higher inclusion rates of CB data suggest that the most cost-effective CB data inclusion strategies involve a small proportion of CB progeny. With limited increases in CB performance associated with higher proportion of CB data inclusion, care should be taken to make sure that genotyping costs do not outweigh benefits in the increased response to selection.

The benefit of including CB genotypes to the genetic evaluations with complete PB data increased as r_{pc} decreased. Averaged over CB inclusions rates, including CB genotypes in the evaluation increased CB performance in generation 5 by 46.0%, 21.4%, and 11.7% when r_{pc} was 0.1, 0.3, and 0.5, respectively. As r_{pc} increased to 0.7 or greater, the increase in CB performance was less than 8%. Consequently, when r_{pc} was high and information content from PB candidates remains constant, the genotyping cost and economic impact of the CB trait must be considered when determining if collecting additional CB genotypes provides a return on investment for the marginal increase in CB performance. In the current study, genetic evaluations assumed that sire pedigree information is available on CB animals. It is commonplace in commercial swine production to produce CB animals using pooled semen comprised of multiple sires. Such practices restrict tracking pedigree relationships of CB animals without the use of genotyping progeny to determine pedigrees. Traditionally, for a farm to gather sire pedigree information on commercial CB animals, farms must commit to utilizing single sire matings knowing the farm will have a decrease in

farrowing rate and litter size compared with utilizing pooled semen. Additionally, pigs born to these single sire litters must be individually identified at birth and the identification maintained through phenotype collection. In practice, utilizing pooled semen in combination with low-cost genotyping schemes may be more practical and cost effective than using single sire matings to produce the CB animals in commercial test herds.

Reducing PB data

Given collection of CB data would incur additional costs, the reduction in PB data to partially offset these costs was also investigated. The response in CB performance, when CB records were included and when PB genotypes and/or phenotypes were removed, is shown in Figure 4. Overall, when r_{pc} was low, the removal of PB data had a negligible impact on CB performance. The reduction in CB performance, when PB data were removed and CB were not genotyped, was greatest when r_{pc} was high. In agreement with the current study, [van Grevenhof and van der Werf \(2015\)](#) concluded that replacing PB phenotypes with CB phenotypes had a substantial impact on prediction accuracy when PB performance was not the selection goal and the r_{pc} was moderate to low. In the current study, when CB were not genotyped, limited relationships were captured between CB phenotypes and selection candidates belonging to breeds A and B. Relationships connecting CB phenotypes to selection candidates from breeds A and B were inferred genomic relationships with breed C. Consequently, reducing PB genotypes when CB were not genotyped had more of an impact on the response to selection due to limited relationships connecting CB phenotypes to breed A and B selection candidates.

Figure 4. Average CB progeny phenotype in generation 5 when a proportion of CB are included in an analysis with CB: Ped or CB: Gen. A proportion of PB phenotypes (Removing PB: Phen) or PB phenotypes and genotypes (Removing PB: Gen, Phen) were removed according to the CB inclusion rate. Different color bars denote the proportion of CB records included and the simultaneous proportion of PB records removed from an analysis. Results were averaged over 10 replicates.

When CB were genotyped, the removal of PB phenotypes did not reduce CB performance regardless of r_{pc} (Figure 4 CB: Gen). However, when CB were genotyped and both PB phenotypes and genotypes were removed, CB performance was negatively impacted when the r_{pc} was moderate to high. Differences in CB performance between strategies which removed PB phenotypes or removed PB phenotypes and genotypes were less pronounced when CB were not genotyped. This suggests that including PB genotypes is more informative than PB phenotypes when CB animals are genotyped.

Genetic gain

The average rate of genetic gain (ΔG) over four generations of selection, defined as the regression of mean CB progeny TBV on generation of birth, is shown in Figure 5. Comparisons of ΔG were made between individual phenotyping and genotyping strategies when r_{pc} was 0.7, which was nearest to the average r_{pc} in swine reported by Wientjes and Calus (2017). Trends in the ΔG when r_{pc} was 0.1, 0.3, 0.5, or 0.9 were similar to those when r_{pc} was 0.7 (results not shown). The strategy that provided the greatest rate of genetic change was when all PB and CB individuals were included in the genetic evaluation with genotypic, pedigree, and phenotypic data. As the proportion of PB phenotypic data was reduced, a plateau was reached where additional CB information did not improve ($P > 0.01$) ΔG beyond a CB inclusion rate of 10% (Figure 5A). One way to contemplate the trade-off between genetic gain and level of investment is to consider genetic gain per additional genotyped animals required to achieve said gain. Figure 5B shows the percentage increase in average ΔG per 1% additional CB genotypes and reduced PB genotypes required compared with a CB inclusion rate of 0. There was an asymptotic decrease in average ΔG per 1% additional genotypes included in an evaluation. In general, a CB inclusion rate of 10% corresponded with the greatest increase in ΔG per 1% increase in genotypes. Across PB data removal strategies, the greatest increase in ΔG per 1% increase in genotypes occurred when CB inclusion was 10% and 10% of PB phenotypes and genotypes were removed. This suggests that a minimal fraction of PB animals can intentionally

be omitted in phenotyping and genotyping schemes if CB data are included in the evaluation.

Accuracy and bias of CB prediction

In general, accuracy was highest for breed C across all genetic correlations and phenotyping/genotyping strategies as sires from C were the only pedigree link between PB and CB populations. As expected, as r_{pc} increased so did EBV accuracy. The greatest change in accuracy within a r_{pc} scenario was observed in breed A (Figure 6). When CB contributed genotypes, the average EBV accuracy of selection candidates from breed A ranged from 0.07 to 0.77. If CB did not contribute genotypes, the average EBV accuracy of selection candidates from breed A ranged from 0.05 to 0.71. As expected, the average EBV accuracy was lowest when r_{pc} was low and when PB genotypes were removed from analyses. When PB genotypes were removed, there was limited kinship information to link CB performance to PB selection candidates. Trends in EBV accuracy for breed B were similar to those shown in breed A; thus, EBV accuracy for breed B was not reported.

In breed C, the accuracy ranged from 0.23 to 0.81 (Figure 7). The highest accuracy occurred when r_{pc} was 0.9 and all PB candidates as well as CB progeny contributed genotypes, phenotypes, and pedigree information. Conversely, the lowest accuracy achieved in breed C was when r_{pc} was 0.1, 10% of CB progeny contributed phenotypic and sire pedigree data, and 10% of PB phenotypes were removed. When all selection candidates from C had genotypes and phenotypes, EBV accuracy increased as the number of CB records with or without genotypes increased. As the proportion of CB genotypes increased and all PB selection candidates were genotyped, the EBV accuracy increased but at a diminishing rate. As the proportion of PB genotypes and phenotypes decreased and the proportion of CB data increased, the EBV accuracy decreased. Such decreases in accuracy are to be expected as the probability of CB sires not being genotyped increases as the proportion of PB data included in genetic evaluations decreases. Previous studies have estimated EBV accuracy in simulation studies (Esfandiyari et al., 2018) and studies using real data (Hidalgo et al., 2015; Sewel et al.,

Figure 5. Rate of ΔG calculated as the regression coefficient from average CB TBV regressed on generation of birth when CB progeny records included genotypes, phenotypes, and sire pedigrees when the r_{pc} was 0.7. (A) Average ΔG when a proportion of CB records were included in genetic analyses with genotypes, sire pedigrees, and phenotypes, while PB data were included (None), phenotypes were removed (Phen), or phenotypes and genotypes were (Phen, Gen) removed at the CB inclusion rate. (B) Percentage increase in average ΔG per 1% increase in genotyped animals compared with when CB inclusion was 0.

Figure 6. Average correlation (r) between EBV and TBV of selection candidates from breed A in generation 5 CB: Ped or CB: Gen. PB data reduction consisted removing no PB information (Including all PB: Ped, Gen, Phen), a proportion of PB phenotypes (Removing PB: Phen), or PB phenotypes and genotypes (Removing PB: Gen, Phen). Different color bars denote the proportion of CB records included and the proportion of PB records removed from an analysis. Results were averaged over 10 replicates.

2018). Esfandiyari et al. (2018) simulated a CCPS breeding scheme with a CB phenotype, which was impacted by dominance. The authors concluded that accuracy in the reference population was greatest for the CB phenotype when the CB phenotype was the breeding goal and CB data were included in the evaluation. Hidalgo et al. (2015) used a swine population that was focusing on maternal traits with a reference population containing either 10% or 30% of CB progeny. Using a multi-breed genomic relationship matrix, EBVs were predicted using PB and CB data for CB performance. The authors reported marginal increases in prediction accuracy when using CB records in the reference population compared with when only PB data were used. The authors concluded that the marginal increase in accuracy could be a result of CB progeny being distantly related to the PB animals. In terminal swine breeding lines, a greater degree of relationship between PB and CB records is possible due to the swine breeding structure and the use of artificial insemination.

The range of BIAS regression coefficients for selection candidates from breed C was 0.76 to 1.07 (Supplementary Table S1). Generally, EBVs were over-dispersed. When r_{pc} was low and the CB inclusion rate was high, EBVs showed the greatest degree of over-dispersion. This over-dispersion of EBVs is likely caused by the reduction of information provided by PB performance in the covariance between PB and CB performance. Large

deviations from expected values of BIAS for selection candidates from breed C were likely not seen as the degree of relationship between CB phenotypes and PB candidates was high. Although some BIAS estimates were significantly different from 1 when the proportion of PB included in analyses with genotypic data was greater than 30%, numerical differences were negligible. As PB genotypes and phenotypes were removed at a frequency of 30% or greater, BIAS regression coefficients deviated from 1 ($P < 0.01$). These deviations are to be expected as the average EBV accuracy in breeds A and B selection candidates was near zero.

In general, results show that average EBV accuracy for selection candidates was high when CB contributed phenotypic and genotypic data, the r_{pc} was high, and the inclusion of PB data was high. Further, EBV accuracy and resulting BIAS regression coefficients were most favorable for selection candidates from breed C, which is to be expected as CB progeny only share pedigree links to sires from breed C. In practice, strong deviations in BIAS from expectations and decreases in accuracy could result in inflated expectations relative to selection response and reduced realized response to selection in CB performance. Reduced response to selection for CB performance could result from choosing the wrong selection candidates due to inaccurate or biased EBV. Xiang et al. (2016) previously noted the impact of CB genotypes in ssGBLUP evaluations on

Figure 7. Average correlation (r) between EBV and TBV of selection candidates from breed C in generation 5 CB: Ped or CB: Gen. PB data reduction consisted removing no PB information (Including all PB: Ped, Gen, Phen), a proportion of PB phenotypes (Removing PB: Phen), or PB phenotypes and genotypes (Removing PB: Gen, Phen). Different color bars denote the proportion of CB records included and the proportion of PB records removed from an analysis. Results were averaged over 10 replicates.

predictive ability of CB performance of PB boars using real data. Such an increase in predictive ability of CB performance would also result in an expected reduction in EBV bias, in agreement with the findings of the current study. The optimal strategy that maximizes accuracy, while reducing bias and cost involves a minimal proportion of CB progeny contributing both genotypic and phenotypic data to the evaluation.

General discussion

The purpose of this study was to evaluate including CB and removing PB sources of data for genetic evaluations to facilitate selection decisions using EBV for CB performance. Effectiveness of selection for CB performance in PB parents has been shown to be dictated by the r_{pc} (Tusell et al., 2016; Wientjes and Calus, 2017). A growing body of literature has suggested that the genetic divergence in PB and CB performance can be due to dominance effects (Wei and van der Werf, 1993), genotype by environment interactions (Zumbach et al., 2007), or trait definitions (Newcom et al., 2005). Regardless of the causes that manifest in $r_{pc} < 1$, commercial progeny performance can be improved by including performance and relationship information of CB progeny into genetic evaluations. It has been suggested that selecting PB for CB performance would increase genetic progress in CB (Dekkers, 2007; van Grevenhof and van der Werf, 2015) and

improve prediction accuracy (Zeng et al., 2013). Despite potential benefits of including CB performance in PB selection decisions, implementation has been limited due to the requirement of tracking CB pedigrees and collecting CB phenotypes and possibly CB genotypes adding to the cost of the genetics program.

The effectiveness of a multibreed genomic evaluation is dependent upon many factors with several factors not being the principle subject of investigation in the current study. Segregating SNP affecting quantitative traits of interest in PB populations could fall out of LD with QTL in CB populations which would change EBV used to select PB candidates through marker substitution effects (Wientjes et al., 2013). Changes in LD patterns, QTL, and SNP frequency between PB breeds and CB animals and the genetic correlation for presumably the same trait among different PB breeds could reduce the prediction accuracy; however, these issues would be population dependent. To date, there are limited reports that investigate the genetic correlation between the same trait among different PB breeds of swine. If the genetic correlation between PB traits is less than unity, the effectiveness of CCPs could be different amongst PB breeds. Such differences in predictive ability could ultimately reduce the efficiency of CCPs. Esfandiyari et al. (2015b) evaluated the effectiveness of selecting PB animals for CB performance compared with selecting PB animals for PB performance on the rate of genetic gain in CB animals

under different levels of LD. The authors reported that when the LD was high, a combined breed reference population provided greater prediction accuracy compared with separate reference populations. When the correlation of LD phase is high, multi-breed reference populations produce greater rates of genetic gain. However, when the correlation of LD phase is low, multi-breed reference populations might not be the most efficient. In the current study, the LD decay was similar, yet slightly higher than real swine populations (Badke et al., 2012), which could have resulted in increased predictive ability from the use of multi-breed evaluations. However, such increases in predictive ability would be similar across genotyping strategies investigated given combined PB reference populations were used in the current study in the absence of CB data. In practice, SNPs are not a random sample of the genome resulting in the possibility of ascertainment bias in the current simulation, which would inflate the predictive ability of genetic evaluations in the current study. Bijma et al. (2001) reported that implementation of CCPS resulted in a greater reduction in genetic diversity and increased inbreeding rates in PB populations when compared with pure line selection with highly heritable traits and low r_{pc} . Consideration should be given when using CCPS to reduce the increase in inbreeding rates in PB populations.

Crossbreeding schemes are designed to leverage increases in CB performance due to heterosis, which is driven by nonadditive genetic effects. It has been shown that heterosis can increase across generations when alternate heterozygotes can be distinguished with the use of CB phenotypes and genotypes in genomic evaluations (Esfandiyari et al., 2015a). Admittedly, further benefits from CCPS could be realized by accounting for dominance variation if the CB performance is partially influenced by dominance effects. It has been shown that genetic models that included dominance are more robust in the prediction of EBV when dominance is present (Zeng et al., 2013; Duenk et al., 2017). Therefore, in situations where CB performance is included in genetic evaluations, more comprehensive modeling approaches should be utilized to incorporate dominance and inbreeding effects into the ssGBLUP framework (Christensen et al., 2014, 2019). In the current study, the CB performance simulated assumed only additive effects, and the inclusion of dominance effects could reduce the accuracy of EBV used to select PB candidates if only an additive model was used. Such reductions in accuracy would result in a correlated decrease in CB performance, which would be assumed to be constant across all data inclusion strategies. Esfandiyari et al. (2015b) showed that selection for CB performance under the influence of dominance increased heterosis was likely due to divergent changes in allele frequencies. The impact of selection-aided drift due to the presence of partial dominance has been previously discussed by Kinghorn et al. (2010) and was shown to be most impactful when the number of QTL was high. Implications of increased heterosis in successive generations of selection and CB genotyping and phenotyping strategies presented in the current paper need to be further investigated.

Our results suggest that the greatest benefit from incorporating CB data is when the traits under selection and evaluation in the PB populations are genetically distinct from those at the commercial production level ($r_{pc} < 0.7$), which is in agreement with previous literature (Dekkers, 2007; Esfandiyari et al. 2015a; Duenk et al. 2019). The reduction in EBV accuracy in maternal PB breeds used to produce commercial progeny is too great and causes the response in CB performance to decrease dramatically when $r_{pc} < 0.7$ and CB progeny are not genotyped. The current study assumed that a proportion of sires used to make PB matings in breed C and CB matings would be the same within a generation. This assumption

leads to increased connectedness between PB and CB populations ultimately increasing the prediction accuracy for selection candidates. In reality, the connectedness between the sire breed and CB animals could be less than simulated in the current study resulting in a decrease in the performance of CCPS. Results from the current study suggest that a CB reference population with genotypes provides an increase in CB performance compared with a PB reference population when $r_{pc} < 0.7$. Duenk et al. (2019) reported that CB and PB reference populations provided the same response in CB performance when $r_{pc} = 0.8$ when breed-of-origin alleles were ignored. Improvements in CB performance could be realized by accounting for alternative heterozygotes in CB animals when using a CB-only reference population for higher levels of r_{pc} (Esfandiyari et al., 2015b; Duenk et al., 2019). Furthermore, the results from the current study suggest that costs from increased CB data inclusion can be partially offset by a reduction in PB phenotyping and/or genotyping. A reduction in PB genotyping should be limited given the importance of PB genotypes in relating CB performance to PB selection candidates when pedigrees of CB progeny are largely unknown. Given that PB exclusion was performed at random, improvement could likely be made by ensuring that genotyped PB animals are the most likely to contribute substantially to gene flow, thus ensuring greater connectivity overtime. Considerable infrastructure has already been established by swine breeding systems to collect a growing number of phenotypes accurately and efficiently at the PB nucleus level predicated on tracking and maintaining pedigrees. We propose that the collection of CB data should not follow the same system as that observed in the nucleus. Data collection at the PB nucleus level focuses on data collection on a large quantity of traits, which can be used as indicator traits in the prediction of CB performance. We suggest that CB data for economically relevant traits on a subset of animals should be collected in commercial settings and included in the genetic evaluation. Pooled semen could be used when accompanied with genotypes. In the present study, CB animals were chosen to be phenotyped and genotyped at random and considerable improvement of CB performance was seen when only 10% of possible CB were included in the analyses. In swine populations, half-sib sire families relating commercial and terminal nucleus populations can be large, suggesting that strategic selection of animals to genotype and phenotype at the commercial level may not be critical to improving CB performance. Random selection of CB animals to phenotype or genotype could provide reasonable improvements in selection for CB performance without the need for designed data collection schemes focused on maximizing relationships or specific phenotypes captured. As such, results contained herein suggest that the optimal reference population in CCPS, where the breeding goal is CB performance, be comprised of 10% to 30% of CB animals. Further research is, however, warranted to investigate the potential benefits of strategic PB and CB data collection in such a breeding system.

Conclusions

Overall, results showed that including CB data in genetic evaluation improved the response to selection. The largest gains in response and EBV accuracy occurred when all possible candidates and CB progeny were included in the evaluation and when $r_{pc} < 0.5$. Yet, as a larger proportion of CB individuals contributed data, the genetic gain showed a diminishing rate of return for the investment of additional genotypes. A proportion of PB phenotypes can be removed while maintaining high EBV accuracy and an increasing trend in CB phenotypic response. As the number of PB genotypes decreased so did the rate of genetic gain in CB progeny, which was attributed to a reduction in PB EBV accuracy. The optimal strategy

to maximize CB performance while minimizing the amount of data required likely involves only including a small proportion of CB data while simultaneously removing a small proportion of PB phenotypes and genotypes. In practice, the optimal level of CB and PB data utilized to produce EBV is system dependent and is impacted by 1) the degree of connectedness between CB phenotypes and PB selection candidates, 2) the strength of the r_{pc} , and 3) the relationship between the cost of additional CB data collection and the economic benefit associated with increased CB performance.

Supplementary Data

Supplementary data are available at *Journal of Animal Science* online.

Conflict of interest statement

The authors declare no conflicts of interest.

Literature Cited

- Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. Hot Topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* 93:743–752. doi:[10.3168/jds.2009-2730](https://doi.org/10.3168/jds.2009-2730)
- Badke, Y. M., R. O. Bates, C. W. Ernst, C. Schwab, and J. P. Steibel. 2012. Estimation of linkage disequilibrium in four US pig breeds. *BMC Genomics* 13:24. doi:[10.1186/1471-2164-13-24](https://doi.org/10.1186/1471-2164-13-24)
- Bichard, M. 1971. Dissemination of genetic improvement through a livestock industry. *Anim. Prod.* 13:401–411. doi:[10.1017/S000335610010606](https://doi.org/10.1017/S000335610010606)
- Bijma, P., J. A. Woolliams, and J. A. M. Van Arendonk. 2001. Genetic gain of pure line selection and combined crossbred purebred selection with constrained inbreeding. *Anim. Sci.* 72:225–232. doi:[10.1017/S1357729800055715](https://doi.org/10.1017/S1357729800055715)
- Bradburd, G., and M. G. Brdburd. 2013. BEDASSLE: Quantifies effects of geo/eco distances on genetic differentiation. R package version 1.5. <https://CRAN.R-project.org/package=BEDASSLE>. Accessed May 2020.
- Chen, G. K., P. Marjoram, and J. D. Wall. 2009. Fast and flexible simulation of DNA sequence data. *Genome Res.* 19:136–142. doi:[10.1101/gr.083634.108](https://doi.org/10.1101/gr.083634.108)
- Christensen, O. F., P. Madsen, B. Nielsen, and G. Su. 2014. Genomic evaluation of both purebred and crossbred performances. *Genet. Sel. Evol.* 46:23. doi:[10.1186/1297-9686-46-23](https://doi.org/10.1186/1297-9686-46-23)
- Christensen, O. F., B. Nielsen, G. Su, T. Xiang, P. Madsen, T. Ostensen, I. Velander, and A. B. Strathe. 2019. A bivariate genomic model with additive, dominance and inbreeding depression effects for sire line and three-way crossbred pigs. *Genet. Sel. Evol.* 51:45. doi:[10.1186/s12711-019-0486-2](https://doi.org/10.1186/s12711-019-0486-2)
- Dekkers, J. C. M. 2007. Marker-assisted selection for commercial crossbred performance. *J. Anim. Sci.* 85:2104–2114. doi:[10.2527/jas.2006-683](https://doi.org/10.2527/jas.2006-683)
- Dickerson, G. E. 1973. Inbreeding and heterosis in animals. *J. Anim. Sci.* 1973:54–77. doi:[10.2527/1973.1973Symposium54x](https://doi.org/10.2527/1973.1973Symposium54x)
- Duenk, P., M. P. L. Calus, Y. C. J. Wientjes, and P. Bijma. 2017. Benefits of dominance over additive models for the estimation of average effects in the presence of dominance. *G3 (Bethesda)*. 7:3405–3414. doi:[10.1534/g3.117.300113](https://doi.org/10.1534/g3.117.300113)
- Duenk, P., M. P. L. Calus, Y. C. J. Wientjes, V. P. Breen, J. M. Henshall, R. Hawken, and P. Bijma. 2019. Validation of genomic predictions for body weight in broilers using crossbred information and considering breed-of-origin of alleles. *Genet. Sel. Evol.* 51:38. doi:[10.1186/s12711-019-0481-7](https://doi.org/10.1186/s12711-019-0481-7)
- Esfandiyari, H., A. C. Sørensen, and P. Bijma. 2015a. A crossbred reference population can improve the response to genomic selection for crossbred performance. *Genet. Sel. Evol.* 47:76. doi:[10.1186/s12711-015-0155-z](https://doi.org/10.1186/s12711-015-0155-z)
- Esfandiyari, H., A. C. Sørensen, and P. Bijma. 2015b. Maximizing crossbred performance through purebred genomic selection. *Genet. Sel. Evol.* 47:16. doi:[10.1186/s12711-015-0099-3](https://doi.org/10.1186/s12711-015-0099-3)
- Esfandiyari, H., P. Berg, and A. C. Sørensen. 2018. Balanced selection on purebred and crossbred performance increases gain in crossbreds. *Genet. Select. Evol.* 50:8. doi:[10.1186/s12711-018-0379-9](https://doi.org/10.1186/s12711-018-0379-9)
- Faux, A. M., G. Gorjanc, R. C. Gaynor, M. Battagin, S. M. Edwards, D. L. Wilson, S. J. Hearne, S. Gone, and J. M. Hickey. 2016. AlphaSim: software for breeding program simulation. *Plant Genome*. 9:1–14. doi:[10.3835/plantgenome2016.02.0013](https://doi.org/10.3835/plantgenome2016.02.0013)
- Gaynor, R. C. G., G. Gorjanc, E. Wilson, D. Money, and J. M. Hickey. 2019. AlphaSimR: breeding program simulations. R package version 0.9.0. <https://CRAN.R-project.org/package=AlphaSimR>. Accessed September, 2018.
- van Grevenhof, I. E., and J. H. van der Werf. 2015. Design of reference populations for genomic selection in crossbreeding programs. *Genet. Sel. Evol.* 47:14. doi:[10.1186/s12711-015-0104-x](https://doi.org/10.1186/s12711-015-0104-x)
- Hidalgo, A. M., J. W. Bastiaansen, M. S. Lopes, B. Harlizius, M. A. Groenen, and D. J. de Koning. 2015. Accuracy of predicted genomic breeding values in purebred and crossbred pigs. *G3 (Bethesda)*. 5:1575–1583. doi:[10.1534/g3.115.018119](https://doi.org/10.1534/g3.115.018119)
- Ibáñez-Escriche, N., R. L. Fernando, A. Toosi, and J. C. Dekkers. 2009. Genomic selection of purebreds for crossbred performance. *Genet. Sel. Evol.* 41:12. doi:[10.1186/1297-9686-41-12](https://doi.org/10.1186/1297-9686-41-12)
- Kinghorn, B. P., J. M. Hickey, and J. H. J. Van Der Werf. 2010. Reciprocal recurrent genomic selection for total genetic merit in crossbred individuals. In: Proceedings of the 9th World Congress on Genetics Applied to Livestock Production; August 2010; Leipzig, Germany: Proceedings of the World Congress on Genetics Applied to Livestock Production; p. 1–6.
- Lourenco, D. A., S. Tsuruta, B. O. Fragomeni, C. Y. Chen, W. O. Herring, and I. Misztal. 2016. Crossbreed evaluations in single-step genomic best linear unbiased predictor using adjusted realized relationship matrices. *J. Anim. Sci.* 94:909–919. doi:[10.2527/jas.2015-9748](https://doi.org/10.2527/jas.2015-9748)
- Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. H. Lee. 2002. BLUPF90 and related programs (BGF90). In: Proceedings of the 7th World Congress on Genetics Applied to Livestock Production; August, 2002; Montpellier, France: Proceedings of the World Congress on Genetics Applied to Livestock Production; p. 743–744.
- Newcom, D. W., T. J. Baas, C. R. Schwab, and K. J. Stalder. 2005. Genetic and phenotypic relationships between individual subcutaneous backfat layers and percentage of longissimus intramuscular fat in Duroc swine. *J. Anim. Sci.* 83:316–323. doi:[10.2527/2005.832316x](https://doi.org/10.2527/2005.832316x)
- Sewel, A., H. Li, C. Schwab, C. Maltecca, and F. Tiezzi. 2018. On the value of genotyping terminal crossbred pigs for nucleus genomic selection for carcass traits. In: Proceedings of the 9th World Congress on Genetics Applied to Livestock Production; February 2018; Auckland, New Zealand; p. 11.
- Tusell, L., H. Gilbert, J. Riquet, M. J. Mercat, A. Legarra, and C. Larzul. 2016. Pedigree and genomic evaluation of pigs using a terminal-cross model. *Genet. Sel. Evol.* 48:32. doi:[10.1186/s12711-016-0211-3](https://doi.org/10.1186/s12711-016-0211-3)
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423. doi:[10.3168/jds.2007-0980](https://doi.org/10.3168/jds.2007-0980)
- Wei, M., and J. H. van der Werf. 1993. Animal model estimation of additive and dominance variances in egg production traits of poultry. *J. Anim. Sci.* 71:57–65. doi:[10.2527/1993.71157x](https://doi.org/10.2527/1993.71157x)
- Wei, M., and J. H. van der Werf. 1994. Maximizing genetic response in crossbreds using both purebred and crossbred information. *Anim. Prod.* 59:401–413. doi:[10.1017/S0003356100007923](https://doi.org/10.1017/S0003356100007923)
- Weir, B. S., and W. G. Hill. 2002. Estimating F-statistics. *Annu. Rev. Genet.* 36:721–750. doi:[10.1146/annurev.genet.36.050802.093940](https://doi.org/10.1146/annurev.genet.36.050802.093940)

- Wientjes, Y. C. J., and M. P. L. Calus. 2017. Board Invited Review: The purebred-crossbred correlation in pigs: a review of theory, estimates, and implications. *J. Anim. Sci.* **95**:3467–3478. doi:[10.2527/jas.2017.1669](https://doi.org/10.2527/jas.2017.1669)
- Wientjes, Y. C., R. F. Veerkamp, and M. P. Calus. 2013. The effect of linkage disequilibrium and family relationships on the reliability of genomic prediction. *Genetics* **193**:621–631. doi:[10.1534/genetics.112.146290](https://doi.org/10.1534/genetics.112.146290)
- Xiang, T., B. Nielsen, G. Su, A. Legarra, and O. F. Christensen. 2016. Application of single-step genomic evaluation for crossbred performance in pig. *J. Anim. Sci.* **94**:936–948. doi:[10.2527/jas.2015-9930](https://doi.org/10.2527/jas.2015-9930)
- Zeng, J., A. Toosi, R. L. Fernando, J. C. Dekkers, and D. J. Garrick. 2013. Genomic selection of purebred animals for crossbred performance in the presence of dominant gene action. *Genet. Sel. Evol.* **45**:11. doi:[10.1186/1297-9686-45-11](https://doi.org/10.1186/1297-9686-45-11)
- Zumbach, B., I. Misztal, S. Tsuruta, J. Holl, W. Herring, and T. Long. 2007. Genetic correlations between two strains of Durocs and crossbreds from differing production environments for slaughter traits. *J. Anim. Sci.* **85**:901–908. doi:[10.2527/jas.2006-499](https://doi.org/10.2527/jas.2006-499)