

## ANIMAL GENETICS AND GENOMICS

# Purebred-crossbred genetic parameters for reproductive traits in swine

Luke M. Kramer,<sup>†,1</sup> Anna Wolc,<sup>†,‡</sup> Hadi Esfandyari,<sup>†,‡,§</sup> Dinesh M. Thekkoot,<sup>¶</sup> Chunyan Zhang,<sup>¶</sup> Robert A. Kemp,<sup>¶</sup> Graham Plastow,<sup>\*\*</sup> and Jack C. M. Dekkers<sup>†</sup>

<sup>†</sup>Department of Animal Science, Iowa State University, Ames IA 50011, USA, <sup>‡</sup>Hy-Line International, 2583 240th Street, Dallas Center, IA 50063, USA, <sup>§</sup>Department of Molecular Biology and Genetics, Center for Quantitative Genetics and Genomics, Aarhus University, Tjele, Denmark, <sup>¶</sup>TYR, Norwegian beef cattle organization, 2315, Hamar, Norway, <sup>¶</sup>Genesus Inc., Oakville, MB, R0H 0Y0, Canada, <sup>\*\*</sup>Department of Agricultural, Food, and Nutritional Science, Livestock Genetec, University of Alberta, Edmonton AB, T6G 2P5, Canada

<sup>1</sup>Corresponding author: [lmkramer@iastate.edu](mailto:lmkramer@iastate.edu)

ORCID numbers: 0000-0002-1570-6271 (H. Esfandyari); 0000-0002-3774-3110 (G. Plastow).

## Abstract

For swine breeding programs, testing and selection programs are usually within purebred (PB) populations located in nucleus units that are generally managed differently and tend to have a higher health level than the commercial herds in which the crossbred (CB) descendants of these nucleus animals are expected to perform. This approach assumes that PB animals selected in the nucleus herd will have CB progeny that have superior performance at the commercial level. There is clear evidence that this may not be the case for all traits of economic importance and, thus, including data collected at the commercial herd level may increase the accuracy of selection for commercial CB performance at the nucleus level. The goal for this study was to estimate genetic parameters for five maternal reproductive traits between two PB maternal nucleus populations (Landrace and Yorkshire) and their CB offspring: Total Number Born (TNB), Number Born Alive (NBA), Number Born Alive > 1 kg (NBA > 1 kg), Total Number Weaned (TNW), and Litter Weight at Weaning (LWW). Estimates were based on single-step GBLUP by analyzing any two combinations of a PB and the CB population, and by analyzing all three populations jointly. The genomic relationship matrix between the three populations was generated by using within-population allele frequencies for relationships within a population, and across-population allele frequencies for relationships of the CB with the PB animals. Utilization of metafounders for the two PB populations had no effect on parameter estimates, so the two PB populations were assumed to be genetically unrelated. Joint analysis of two (one PB plus CB) vs. three (both PB and CB) populations did not impact estimates of heritability, additive genetic variance, and genetic correlations. Heritabilities were generally similar between the PB and CB populations, except for LWW and TNW, for which PB populations had about four times larger estimates than CB. Purebred-crossbred genetic correlations ( $r_{pc}$ ) were larger for Landrace than for Yorkshire, except for NBA > 1 kg. These estimates of  $r_{pc}$  indicate that there is potential to improve selection of PB animals for CB performance by including CB information for all traits in the Yorkshire population, but that noticeable additional gains may only occur for NBA > 1 kg and TNW in the Landrace population.

**Key words:** genetic parameters, purebred-crossbred correlation, reproductive traits, swine

## Abbreviations

CB	crossbred
GBLUP	genomic best linear unbiased prediction
PB	purebred
ssGBLUP	single-step genomic best linear unbiased prediction

## Introduction

In the swine breeding industry, meeting consumer and producer demands drives innovation and decision making at all levels. Typical breeding programs are implemented in well-managed and highly bio-secure purebred (PB) populations that drive the genetic merit of pigs at the commercial producer level, which are typically crossbred (CB). In addition to the unique genetic composition of PB vs. CB animals, pigs in the PB breeding populations are not exposed to the same environmental stressors that their CB descendants will experience in a typical production environment (Wei and van der Werf, 1994; Bijma and van Arendonk, 1998; Dekkers, 2007; Zumbach et al., 2007; Zeng et al., 2013; Tusell et al., 2016). As such, selection for production traits observed in the PB animals at the nucleus level may not be entirely indicative of how their CB descendants will perform at the commercial level. As a result, recording phenotypes at the commercial CB level and integration of these to estimate breeding values of PB nucleus animals may improve the accuracy of selection in the nucleus PB populations for CB performance improvement. Many studies have examined the use of PB and CB data or a combination of these to determine which scenarios benefit most from the addition of CB data (Christensen et al., 2014; Esfandyari et al., 2015; van Grevenhof and van der Werf, 2015; Wientjes and Calus, 2017). A key parameter that determines this benefit is the genetic correlation between the same traits recorded in the PB vs. CB populations ( $r_{pc}$ ). A high  $r_{pc}$  indicates that the additive genetic component of the trait is essentially the same in the two populations and that CB data may be of limited use. However, if  $r_{pc}$  is lower than, e.g., 0.75, the use of CB data for genetic evaluation in the nucleus may improve the accuracy of selection for commercial CB performance substantially and be worth the extra costs (Dekkers, 2007; Iversen et al., 2017; Wientjes et al., 2020).

Design of breeding programs to improve commercial CB performance requires estimates of the variance components and of  $r_{pc}$  to make informed decisions about what data to collect, what estimates of genetic merit to be used for selection, and what models to use for the estimation of breeding values. Previous studies have evaluated various strategies to use CB data to improve selection for commercial CB performance in PB populations (Esfandyari et al., 2015), such as exploring the impact of single-step genomic best linear unbiased prediction (ssGBLUP) (Legarra et al., 2009) for total number born (Xiang et al., 2016), comparing breed-specific effect models for genomic selection for litter size and gestation length (Lopes et al., 2017), and the inclusion of CB data for carcass and production traits (Christensen et al., 2019). Further studies included the use of metafounders, which provide the opportunity to connect the PB parental populations and the CB population through ancestral individuals when estimating covariance(s) between phenotypes collected on CB and PB populations (Xiang et al., 2017; van Grevenhof et al., 2019).

Although multiple studies have reported estimates of  $r_{pc}$  for the same phenotype between PB and CB populations, published estimates of  $r_{pc}$  between different reproductive phenotypes in pigs are limited. Therefore, the objectives of this work were to 1) estimate variance components, heritabilities, and  $r_{pc}$  for reproductive traits for CB and their two parental PB populations both between the same and different phenotypes recorded on PB and CB animals, and 2) evaluate the effect of using metafounders on estimates of variance components and of  $r_{pc}$ . To achieve these goals, ssGBLUP was used to combine genomic and pedigree relationships for two- (one PB plus CB) and three-population (both PB plus CB) analyses. The three-population analyses allow for information from both parental PB populations to be utilized.

## Materials and Methods

The animals used in this study were cared for according to the Canadian Council on Animal Care (Olfert et al., 1993) guidelines.

### Populations

Maternal reproduction phenotypes were collected on two maternal PB lines (Yorkshire and Landrace) and on CB produced from the two PB. The PB animals were housed in a single high-health nucleus herd with 2,540 Landrace and 2,451 Yorkshire females with recorded data. The CB population was housed in a typical commercial production herd and consisted of 2,570 CB females with records. Phenotypes were collected over a 3-yr period for CB individuals (2017–2020) and over an 8-yr period for the two PB populations (2011–2019). Sires and dams of the CB individuals were included in the PB data and pedigree, although not all purebred parents were included in the genotypic data, as some were not genotyped. Additionally, while the matings to produce CB females were primarily in one direction (Landrace sires and Yorkshire dams), some matings were in the opposite direction and, as such, most of those dams were also present in the PB populations.

### Phenotypic data

Data were collected on PB and CB females across multiple parities of each sow's productive lifetime (parities 1 to 7 for CB; parities 1 to 4 for PB) for five economically important reproductive phenotypes (Table 1): Total Number Born (TNB), Number Born Alive (NBA), Number Born Alive with birth weight greater than 1 kg (NBA > 1 kg), Total Number Weaned (TNW), and Litter Weaning Weight (LWW) of piglets weaned by the sow. Not all measurements were collected on all individuals and not all parities were recorded for each individual. Number born alive was determined as the number of piglets alive at the first observation after farrowing (normally within 12 h). Lactation length and net number fostered (number of piglets fostered onto the sow minus the number of piglets fostered off the sow over the length of her lactation) were also recorded for each litter. Early weaning of some piglets within a litter was practiced in the CB population and the number of early weaned piglets and their weights were included in litter records for TNW and LWW. Phenotypic data were edited based on the following criteria: lactation length between 14 and 28 d; total number born > 6; average piglet wean weight  $\geq$  3 kg; total piglet differential  $\pm$  1 per parity [number of piglets weaned – (number of piglets born alive + net fostered – deaths)]. Phenotypes for TNW and LWW

**Table 1.** Summary statistics for phenotypes and covariates recorded in the two purebred populations (Landrace and Yorkshire) and their CB offspring<sup>1</sup>

	Total records			Average			SD		
	Landrace	Yorkshire	CB	Landrace	Yorkshire	CB	Landrace	Yorkshire	CB
TNB	4483	4067	9671	13.37	14.83	14.5	3.2	3.54	3.08
NBA	4483	4067	9671	11.17	12.09	13.32	2.76	3.08	3.05
NBA > 1 kg	4483	4067	9173	10.44	10.92	11.25	2.58	2.9	2.87
LWW	4483	4067	7083	60.82	60.37	62.16	12.55	13.46	11.98
TNW	4483	4067	7083	10.28	10.35	11.66	1.53	1.68	1.65
LL	4483	4067	7083	19.91	20.61	19.35	2.3	2.14	1.89
#Repeated Records	-	-	-	1.76	1.66	3.76	0.83	0.78	2.22
NF	4483	4067	7083	0.11	-0.66	-0.12	2.72	3.02	2.54
NEW	N/A	N/A	7083	N/A	N/A	1.56	N/A	N/A	0.88

<sup>1</sup>Number of records for each population by trait, average, and standard deviation.

TNB, Total Number Born, NBA, Number Born Alive, NBA > 1 kg, Number Born Alive > 1 kg, LWW, Litter Weight at Weaning (kg), TNW, Total Number Weaned, LL, Lactation Length (days), NF, Net fostered, NEW, Number of Early Weaned piglets per litter.

were set to missing if lactation length, average piglet wean weight, or total piglet differential were out of range, whereas all phenotypes for a sow's parity were set to missing if TNB was out of range.

### Genotypes and pedigree

Genotypes were available for 1,965 Landrace, 1,981 Yorkshire, and 2,060 CB individuals from the Illumina porcine SNP60 v2 BeadChip (Illumina Inc., San Diego, CA) and from a custom Affymetrix Axiom Porcine Genotyping Array with approximately 55k SNPs (Affymetrix, Santa Clara, CA). Genotyped females were from the same period of time as the period for which phenotypes were available and the majority of phenotyped animals were also genotyped (77% of Landrace, 81% of Yorkshire, and 80% of CB females). Genotyped animals were reduced to common SNPs and imputed to the custom Affymetrix genotyping array using FIMPUTE (Sargolzaei et al., 2014). A total of 53,079 SNPs remained after excluding SNPs with minor allele frequency less than 0.05 across all three populations, call rate less than 0.90, and SNPs that were not mapped in Sus scrofa build 11.1 ([https://www.ncbi.nlm.nih.gov/genome/gdv/browser/genome/?id=GCF\\_000003025.6](https://www.ncbi.nlm.nih.gov/genome/gdv/browser/genome/?id=GCF_000003025.6)). The genomic relationship matrix across the three populations was generated using VanRaden method 1 (VanRaden, 2008; Wientjes et al., 2017), using within-population allele frequencies for relationships within a population and across-population allele frequencies for relationships of the CB with PB animals. The two PB populations were assumed to be genetically unrelated and their relationships were set to zero.

Pedigree was available only for the PB populations. Individuals with unknown parentage, including all CB sows, were assigned to an unknown parentage group based on year of birth, using the unknown parent group coding of REMLF90 (Misztal et al., 2018). The use of metafounders was also evaluated, with a single metafounder created for each PB population to allow a connection between purebred populations and CB offspring, as described in Xiang et al. (2017). The use of metafounders, however, did not have a noticeable effect on estimates of variance components and genetic correlations. Therefore, only models and results without the use of metafounders will be presented.

### Statistical models and analyses

Briefly, five maternal trait phenotypes were analyzed using either two-population (one PB plus CB) or three-population (both PB plus CB) models. Traits were considered to be population

specific (i.e., CB TNB, Landrace TNB, and Yorkshire TNB). This was done for both the same trait phenotype and for different trait phenotypes between populations (e.g., between TNB in PB and LWW in CB). The two- and three-population models yielded very similar estimates of variance components and of  $r_{pc}$ . Thus, only results from the three-population approach will be presented. Genetic correlations between traits within a population were estimated using bivariate models, considering only phenotypes for that population.

The following trait-level linear mixed model designates the appropriate matrices, fixed effects, and permanent environmental effects associated with each population-specific trait and was used as the basis for all analyses, allowing each population to express a different but correlated trait:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{Sow} + \mathbf{e} \quad (1)$$

where  $\mathbf{y}$  is the vector of phenotypic observations for the trait within a population;  $\mathbf{X}$ ,  $\mathbf{Z}_1$ , and  $\mathbf{Z}_2$  are incidence matrices relating records to fixed, animal, and permanent environmental effects, respectively;  $\mathbf{b}$  is the vector of fixed effects,  $\mathbf{a}$  is the vector of random animal additive genetic effects, with  $\mathbf{a} \sim N(0, \mathbf{H} \otimes \Sigma_g)$ , where  $\mathbf{H}$  is the combined genomic and pedigree relationship matrix (Legarra et al., 2014), and  $\Sigma_g$  is the genetic co-variance matrix of dimension 2 and 3 for the bi-variate and three-variate analyses, respectively;  $\mathbf{Sow}$  is the vector of permanent environmental effects of sow to account for repeated records, with  $\mathbf{Sow} \sim N(0, \mathbf{I} \otimes \Sigma_{pe})$ , where  $\Sigma_{pe}$  is the appropriate permanent environmental co-variance matrix for the bi- or three-variate analyses; and  $\mathbf{e} \sim N(0, \mathbf{I} \otimes \Sigma_e)$ , where  $\Sigma_e$  is the appropriate residual co-variance matrix. For the three-population analyses, variance-covariance matrices were as

$$\text{follows: } \Sigma_g = \begin{bmatrix} \sigma_{g1}^2 & \sigma_{g1g2} & \sigma_{g1g3} \\ \sigma_{g1g2} & \sigma_{g2}^2 & \sigma_{g2g3} \\ \sigma_{g1g3} & \sigma_{g2g3} & \sigma_{g3}^2 \end{bmatrix}, \Sigma_{pe} = \begin{bmatrix} \sigma_{pe1}^2 & 0 & 0 \\ 0 & \sigma_{pe2}^2 & 0 \\ 0 & 0 & \sigma_{pe3}^2 \end{bmatrix},$$

$$\Sigma_e = \begin{bmatrix} \sigma_{e1}^2 & 0 & 0 \\ 0 & \sigma_{e2}^2 & 0 \\ 0 & 0 & \sigma_{e3}^2 \end{bmatrix}, \text{ where subscripts 1, 2, and 3 represent}$$

the phenotype recorded on CB, Landrace, and Yorkshire, respectively. Covariances of permanent environmental and residual effects between populations were set to zero because sows only have observations in one population. Fixed effects

included contemporary group (CG, year-week-herd of farrowing, with 112 to 342 levels depending on population and trait) and parity (1 to 4 for PB; 1 to 7 for CB). For TNW and LWW, linear covariates of lactation length (LL), net number fostered (NF), and number early weaned (NEW) were included as fixed effects (see Table 1 for general statistics on these covariates).

Two-population (one PB plus CB) analyses were run in ASReml 4.1 (Gilmour et al., 2015) to obtain priors for (co)variance components between the CB and each PB population. Final estimates were obtained from the three-population model using ssGBLUP of the BLUPF90 series of programs, including renumf90, remlf90, and airemlf90 (Misztal et al., 2018). EM-REML was used for the first 50 iterations and the convergence criterion was set to  $10^{-9}$ .

The bivariate model used to estimate variance components and genetic correlations between trait phenotypes within a population assumed a similar variance-covariance structure for  $\Sigma_g$  to the three-population model, whereas  $\Sigma_{pe}$  and  $\Sigma_e$  differed by having off-diagonal values as phenotypes were on the same animal.

Estimates of heritability were calculated as the ratio of the estimates of additive genetic to phenotypic variance, which was the sum of estimates of genetic, permanent environment, and residual variance. Estimates of repeatability were calculated as the sum of the estimates of additive genetic plus permanent environmental variances divided by the estimate of phenotypic variance.

## Results

### Descriptive statistics

The two PB populations had approximately equal numbers of records, whereas the CB population had about double the number of records (Table 1). The number of CGs was smaller for the CB than for the PB populations by a factor of 3, but the numbers of records per CG were higher for the CB than for the PB populations (average of 41 records per CG for CB and 13/12 for Landrace/Yorkshire). Parity ranged from 1 to 7 for CB sows, with all PB sows having between 1 and 4 parities. The frequency of CB records from parities 1 to 5 was relatively constant and then decreased sharply for parities 6 on.

All recorded PB litters had complete data for all five of the sow reproduction traits. All recorded litters on CB females had TNB and NBA data but not for NBA > 1 kg, TNW, and LWW (Table 1). A given phenotype had similar averages and standard deviations in the three populations. The lower number of records for LWW and TNW in the CB population was because phenotypes that were outside acceptable ranges were removed, as described in Materials and Methods.

### Genetic parameter estimates

Estimates of variance components, heritabilities, and repeatabilities for each phenotype were obtained using three-trait analyses, considering the phenotype to be a different trait in each population (Table 2). Heritability estimates were similar between the three populations, except for LWW and TNW, for which heritability estimates were approximately four times larger for the PB than for the CB populations (Landrace 0.16, Yorkshire 0.21, and CB 0.04 for LWW, and Landrace 0.07, Yorkshire 0.07, and CB 0.02 for TNW). Standard errors of heritability estimates were consistently low across traits and populations.

Trends in estimates of genetic variance between populations for the five phenotypes were different from the trends in heritability estimates. The two PB populations had similar genetic variance estimates for all phenotypes, except for NBA > 1 kg and LWW, for which the Yorkshire population had approximately 75% higher estimates than the Landrace population. The CB population had the lowest numerical estimates of genetic variance of the three populations for TNB, LWW, and TNW, and the largest numerical estimates of genetic variation for NBA and the second largest for NBA > 1 kg, although high standard errors make these differences not significant. Similar to the estimates of heritability, the difference between the estimates for the PB and CB populations was again largest for LWW and TNW. Repeatability estimates were generally moderate, with low standard errors across all phenotypes and populations. Repeatability estimates were similar between the three populations for TNB, TNW, and LWW, but higher for NBA and NBA > 1 kg in the CB population.

Estimates of genetic correlations between phenotypes within the PB and CB populations ( $r_g$ , Figure 1) ranged from low to high (0.14 to 0.76 for CB, -0.1 to 0.95 for Landrace, and 0.06 to 0.78 for Yorkshire) and many had large standard errors. The genetic

**Table 2.** Estimates of heritabilities, additive genetic variance, repeatability, and genetic correlation (SE) for traits recorded in the two purebred populations (Landrace and Yorkshire) and their CB offspring<sup>1</sup>

	Heritability			Genetic variance			Repeatability		
	Landrace	Yorkshire	CB	Landrace	Yorkshire	CB	Landrace	Yorkshire	CB
TNB	0.12 (0.02)	0.1 (0.02)	0.09 (0.02)	1.23 (0.24)	1.21 (0.27)	0.84 (0.22)	0.25 (0.03)	0.21 (0.03)	0.23 (0.02)
NBA	0.09 (0.02)	0.08 (0.02)	0.08 (0.02)	0.7 (0.16)	0.73 (0.19)	0.79 (0.22)	0.19 (0.03)	0.19 (0.03)	0.26 (0.02)
NBA > 1kg	0.07 (0.02)	0.09 (0.02)	0.08 (0.02)	0.45 (0.13)	0.72 (0.17)	0.61 (0.19)	0.21 (0.03)	0.19 (0.03)	0.27 (0.02)
LWW	0.16 (0.03)	0.21 (0.03)	0.04 (0.02)	16.69 (3.01)	31 (5.01)	4.09 (2.43)	0.32 (0.03)	0.29 (0.03)	0.31 (0.02)
TNW	0.07 (0.02)	0.07 (0.02)	0.02 (0.02)	0.14 (0.04)	0.18 (0.05)	0.06 (0.05)	0.2 (0.03)	0.18 (0.03)	0.19 (0.02)
	CB/Landrace Genetic Correlation			CB/Yorkshire Genetic Correlation					
TNB	0.88 (0.22)			0.68 (0.23)					
NBA	0.91 (0.26)			0.52 (0.27)					
NBA > 1 kg	0.24 (0.34)			0.52 (0.28)					
LWW	0.86 (0.37)			0.65 (0.3)					
TNW	0.67 (0.56)			0.69 (0.55)					

<sup>1</sup>Estimates (and standard errors) from three-trait ssGBLUP REMLF90 analysis per trait. TNB, Total Number Born, NBA, Number Born Alive, NBA > 1 kg, Number Born Alive > 1 kg, LWW, Litter Weight at Weaning (kg), TNW, Total Number Weaned.

correlation between TNW and NBA > 1 kg could not be estimated in CB, possibly due to the low genetic variance of TNW in CB. Estimates of genetic correlations between phenotypes within a population were generally 50% to 100% larger for Landrace than for Yorkshire, although estimates that involved LWW

were larger in Yorkshire. Estimates of phenotypic correlations between traits within a population were in general larger than the corresponding genetic correlation estimates, although not always, such as for LWW with TNB and LWW with TNW in Landrace (Figure 2).

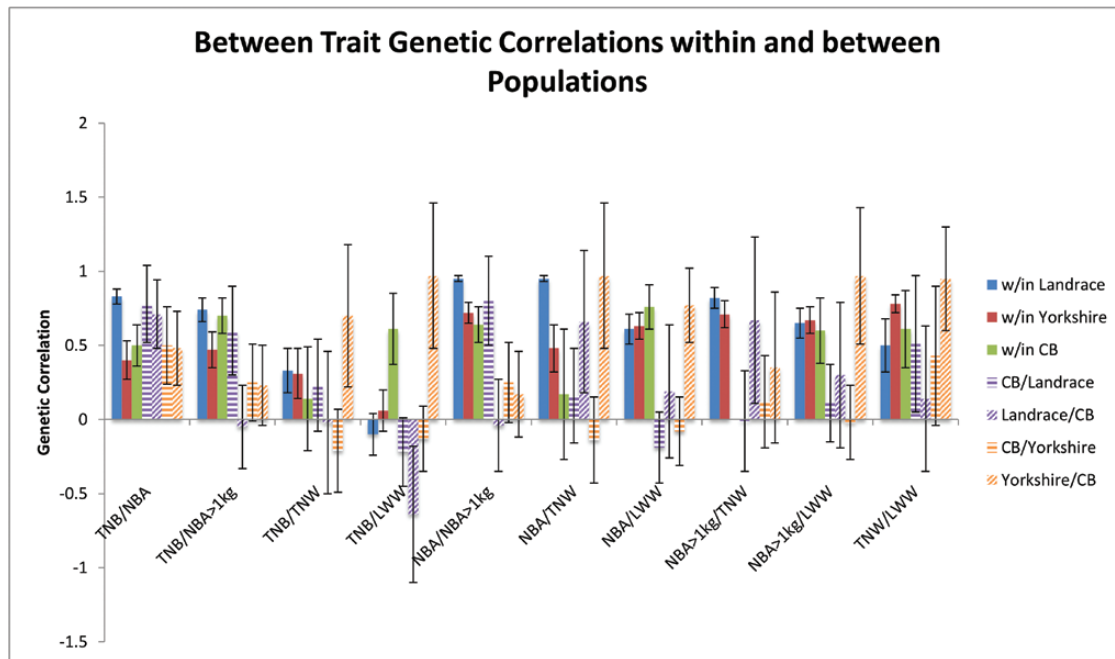


Figure 1. Between trait genetic correlations within and between populations between trait genetic correlations within and between populations (1 purebred and crossbred). TNB: Total Number Born, NBA: Number Born Alive, NBA > 1 kg: Number Born Alive > 1 kg, LWW: Litter Weight at Weaning (kg), TNW: Total Number Weaned.

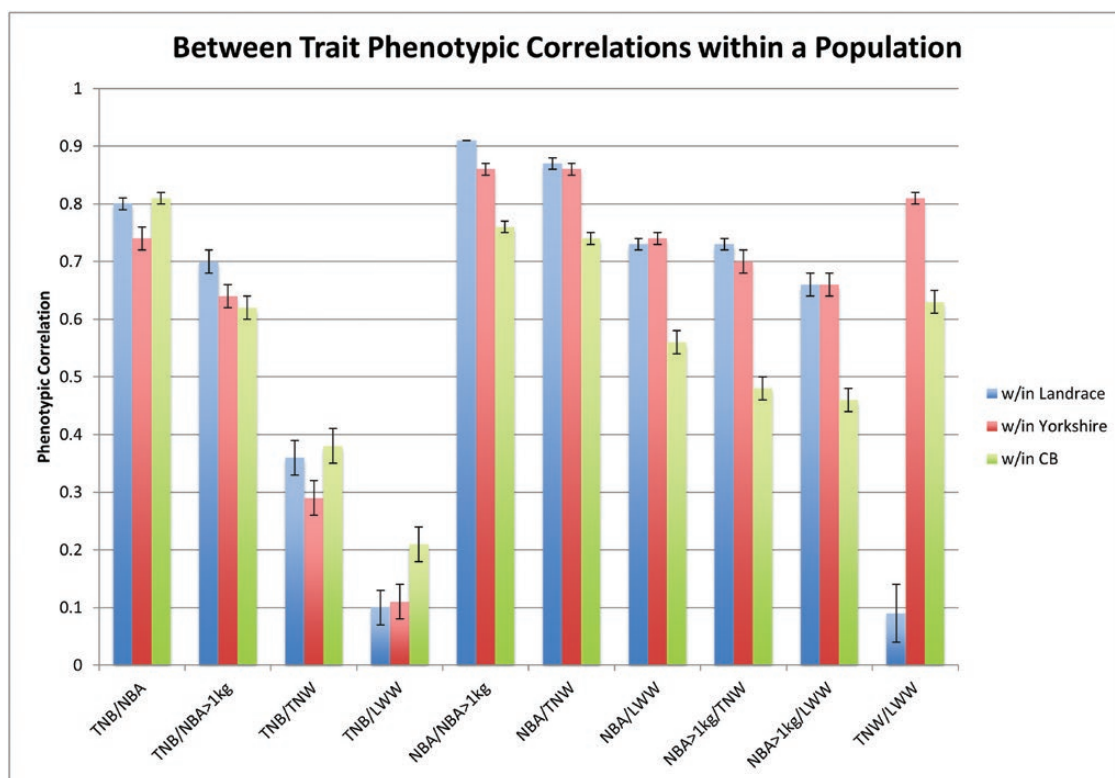


Figure 2. Between trait phenotypic correlations within a population between trait phenotypic correlations within a population. TNB: Total Number Born, NBA: Number Born Alive, NBA > 1 kg: Number Born Alive > 1 kg, LWW: Litter Weight at Weaning (kg), TNW: Total Number Weaned.

Estimates of  $r_{pc}$  for the same phenotype between the PB and CB populations ( $r_{pc}$ , Table 2) ranged from moderate to high (0.52 to 0.91), depending on the trait and populations compared, except for an estimate of 0.24 for NBA > 1 kg between CB and Landrace, with many estimates having large standard errors. Estimates of  $r_{pc}$  differed greatly for the two PB populations, although not significantly, and there was no general trend other than estimates being smaller for Yorkshire than for Landrace except for NBA > 1 kg.

Estimates of  $r_{pc}$  for different phenotypes between the PB and CB populations ranged from highly negative to highly positive for Landrace with CB (-0.64 to 0.8) and from lowly negative to highly positive for Yorkshire with CB (-0.21 to 0.97; Figure 1) but with moderate to high standard errors. For Landrace, the largest estimates of  $r_{pc}$  involved the NBA and NBA > 1 kg phenotypes, whereas for Yorkshire, estimates of  $r_{pc}$  that involved CB LWW were consistently high.

## Discussion

This study investigated estimates of variance components, heritabilities, repeatabilities, and genetic and phenotypic correlations for sow reproduction phenotypes for two PB populations (Landrace and Yorkshire) and their CB population. The ultimate goal of this project is to determine whether the use of CB phenotypes in genetic evaluation of PB selection candidates can improve the rate of genetic improvement of the PB populations for CB performance. Incorporating CB information in multiple-trait PB selection programs requires knowledge of genetic parameters between PB and CB populations, both for the same and different phenotypes. Especially the latter are not widely available in the literature. Substantial numbers of phenotypes were available on PB and CB animals, as well as SNP genotypes, which improves the accuracy of genetic parameter estimates. The ssGBLUP method was used to capitalize on both pedigree and genomic information.

Estimates of variance components and heritabilities were similar for two-population analyses (CB plus one PB population) compared to using all three populations together. Nevertheless, given adequate computational resources, it is recommended that analyses are conducted across the three populations to make full use of all information. The use of metafounders to better align genomic and pedigree relationships and allow relationships between PB populations was also investigated but was found to have negligible effects on estimates of variance components and genetic correlations.

The CB population had similar estimates of repeatability for LWW and TNW than the PB populations but noticeably lower estimates of heritability, a result of a larger sow permanent environmental effect variance in the CB population. This may be related to differences in management and environment in the nucleus vs. the multiplier herds; for example, CB sows had on average more than 3 records, whereas PB sows had on average less than 2 records. There may also be additional factors that were not accounted for, such as non-additive genetic effects that resulted in the greater permanent environmental effects for CB sows. Determining what these are may improve estimates of heritability for LWW and TNW, which are the only two traits that could be affected by adjustments in how piglets are managed during lactation and are important sow productivity traits.

The PB and CB populations differed in two ways in how piglets were managed, namely, cross-fostering between sows and early weaning of piglets. Cross-fostering, as a routine management practice, was discouraged in the PB nucleus and was used when sows were not capable of providing sufficient milk for their litter. In the CB herd, cross-fostering was also not routinely practiced.

Instead, the two largest piglets were routinely weaned about 7 d earlier than the rest of the piglets for more than 85% of the litters. Phenotypes for LWW and TNW on CB litters included early-weaned piglets. Both early weaning and cross fostering may bias LWW and TNW phenotypes and may result in them not being entirely accurate measures of a sow's maternal capability. To account for the potential effects of cross-fostering and early weaning, the number of net fostered pigs and the number of early weaned piglets (only for CB phenotypes) were fitted as fixed covariates for TNW and LWW. There was no adjustment for the timing of early weaning and weight at early weaning was considered as the final weaning weight for those pigs. To obtain unbiased estimates of genetic parameters for LWW and TNW as measures of a sow's maternal ability, net fostered and early weaning were fit as covariates to account for management practices.

With estimates of repeatability being nearly double the estimates of heritability for all traits except one, permanent environmental effects of the sow were found to have a large effect on the reproduction phenotypes analyzed here. Estimates of repeatability of NBA and NBA > 1 kg were higher in the CB than in both PB populations, although estimates of heritability were similar between the three populations. This may be indicative of differences in management. Identification of these management differences and their incorporation in the analysis models could alleviate some disparity between repeatability and heritability. Incorporation of non-additive genetic effects may also help us to bridge the gap between repeatability and heritability.

An estimate of  $r_{pc}$  of 0.75 for a given trait has been suggested to be the threshold above which addition of CB phenotypes in genetic evaluation of PB selection candidates does not materially improve the selection response in CB animals compared to just using PB phenotypes (van Grevenhof et al., 2015; Iversen et al., 2017; Wientjes et al., 2020). Our findings show that none of the five traits fully meet this minimum threshold. Across populations, each trait had at least one  $r_{pc}$  less than 0.75 (Table 2). NBA > 1 kg and TNW would benefit from the inclusion of CB data for both PB populations, whereas TNB, NBA, and LWW may benefit from the inclusion of CB data for PB Yorkshire selection. The standard errors on these estimates of  $r_{pc}$ , however, put substantial uncertainty on these assessments. As such, it would be beneficial to continue recording data from the CB population to improve estimates of  $r_{pc}$  and increase the accuracy of selection for CB performance in the two PB lines.

Of particular interest are the high estimates of  $r_{pc}$  between different phenotypes in a PB and the CB population, such as between LWW in the CB and TNB in Yorkshire, which was 0.97, though with a large standard error (0.49) (Figure 1). In contrast, the estimate of  $r_{pc}$  for LWW for CB and Yorkshire was only 0.6 (+0.3). These results indicate that recording of CB traits may benefit more from across phenotype evaluations than from utilizing the same phenotype from the PB and CB populations. This may be because of differences in management practices between the nucleus and CB herds, which may lead to TNB in PB to be a better genetic indicator of LWW in the CB herd than PB records on LWW. Multiple-trait genetic evaluation methods, with multiple traits in the PB and the CB population, would allow these between phenotype  $r_{pc}$  to be capitalized on. The large standard errors associated with these  $r_{pc}$ , however, do require careful consideration. Additional data and records may help us to improve these  $r_{pc}$  estimates.

## Conclusions

This study estimated variance components, heritabilities, and genetic and phenotypic correlations for five economically important maternal phenotypes in pigs. It examined the

differences and similarities of these phenotypes in two PB populations and their CB offspring to determine if there is a potential benefit to using CB data for genetic evaluation of PB selection candidates in order to increase genetic improvement for CB performance. Heritability estimates ranged from low to moderate across phenotypes and populations, whereas estimates of  $r_{pc}$  between the CB and the two purebred populations ranged from low to high. This range of  $r_{pc}$  estimates indicates that there is the potential for improved accuracy of selection by using CB in addition to PB records, with all five phenotypes having at least one PB population that would benefit from the addition of CB data based on estimates of  $r_{pc} < 0.75$ . Standard errors of estimates were, however, generally large. Furthermore, estimates of  $r_{pc}$  between different phenotypes across the PB and CB populations indicated that selection for improvement of a given CB phenotype may be more effective when based on another, more highly correlated phenotype in the PB population, which can be accommodated by multi-trait genetic evaluation. Availability of additional populations and/or more extensive phenotyping and genotyping of the CB population is needed to further refine the estimates of heritability and  $r_{pc}$  to better understand which traits could benefit from the inclusion of CB data in PB selection.

## Acknowledgments

The authors acknowledge the efforts of the staff at the farms in animal care and data collection. The authors acknowledge funding support received from Genesus Inc., the Government of Canada through Genome Canada and Genome Alberta under the Genomic Applications Partnership Program (GAPP). Genesus Inc. received advisory services and research and development funding for this project through the National Research Council of Canada Industrial Research Assistance Program (NRC IRAP). This study was partially supported by USDA-NIFA grant 2015-67015-22947.

## Supplementary Data

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

## Conflict of interest statement

The authors declare no real or perceived conflicts of interest.

## Literature Cited

Bijma, P., and J. A. M. van Arendonk. 1998. Maximizing genetic gain for the sire line of a crossbreeding scheme utilizing both purebred and crossbred information. *J. Anim. Sci.* 66:529–542. doi: [10.1017/S13577298000970X](https://doi.org/10.1017/S13577298000970X)

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. 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).

Esfandyari, H., A. C. Sørensen, and P. Bijma. 2015. 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).

Gilmour, A. R., B. J. Gogel, B. R. Cullis, S. J. Welham, and R. Thompson. *ASReml user guide release 4.1 structural specification*. VSN International Ltd. Hemel Hempstead. 2015

Iversen, M. W., Ø. Nordbø, E. Gjerlaug-Enger, E. Grindflek, M. S. Lopes, and T. H. E. Meuwissen. 2017. Including crossbred pigs in the genomic relationship matrix through utilization of both linkage disequilibrium and linkage analysis. *J. Anim. Sci.* 95:5197–5207. doi: [10.2527/jas2017.1705](https://doi.org/10.2527/jas2017.1705).

Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. *J. Dairy Sci.* 92:4656–4663. doi: [10.3168/jds.2009-2061](https://doi.org/10.3168/jds.2009-2061).

Legarra, A., O. F. Christensen, I. Aguilar, and I. Misztal. 2014. Single Step, a general approach for genomic selection. *Livest. Sci.* 166:54–65. doi: [10.1016/j.livsci.2014.04.029](https://doi.org/10.1016/j.livsci.2014.04.029)

Lopes, M. S., H. Bovenhuis, A. M. Hidalgo, J. A. M. van Arendonk, E. F. Knol, and J. W. M. Bastiaansen. 2017. Genomic selection for crossbred performance accounting for breed-specific effects. *Genet. Sel. Evol.* 49:51. doi: [10.1186/s12711-017-0328-z](https://doi.org/10.1186/s12711-017-0328-z).

Misztal, I., S. Tsuruta, D. Lourenco, Y. Masuda, I. Aguilar, A. Legarra, and Z. Vitezica. 2018. *Manual for BLUPF90 family of programs*. Univ. Georg: Athens.

Olfert, E. D., B. M. Cross, and A. A. McWilliam. 1993. *Guide to the care and use of experimental animals*. Canadian Council on Animal Care: Ottawa (ON).

Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. *BMC Genomics* 15:478. doi: [10.1186/1471-2164-15-478](https://doi.org/10.1186/1471-2164-15-478).

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).

van Grevenhof, E. M., J. Vandenplas, and M. P. L. Calus. 2019. Genomic prediction for crossbred performance using metafounders. *J. Anim. Sci.* 97:548–558. doi: [10.1093/jas/sky433](https://doi.org/10.1093/jas/sky433).

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).

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. 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)

Wientjes, Y. C. J., P. Bijma, and M. P. L. Calus. 2020. Optimizing genomic breeding program designs to improve crossbred performance. *Genet. Sel. Evol.* 52:65. doi: [10.1186/s12711-020-00573-3](https://doi.org/10.1186/s12711-020-00573-3)

Wientjes, Y. C. J., P. Bijma, J. Vandenplas, and M. P. L. Calus. 2017. Multi-population genomic relationships for estimating current genetic variances within and genetic correlations between populations. *Genetics* 207:503–515. doi: [10.1534/genetics.117.300152](https://doi.org/10.1534/genetics.117.300152).

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).

Xiang, T., O. F. Christensen, and A. Legarra. 2017. Technical note: Genomic evaluation for crossbred performance in a single-step approach with metafounders. *J. Anim. Sci.* 95:1472–1480. doi: [10.2527/jas.2016.1155](https://doi.org/10.2527/jas.2016.1155).

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).