

Inferring phenotypic causal structure among farrowing and weaning traits in pigs

Toshihiro Okamura¹  | Kazuo Ishii¹ | Motohide Nishio¹ | Guilherme J. M. Rosa^{2,3}  | Masahiro Satoh⁴  | Osamu Sasaki¹ 

¹Institute of Livestock and Grassland Science, NARO, Tsukuba, Ibaraki, Japan

²Department of Animal Sciences, University of Wisconsin-Madison, Madison, WI, USA

³Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, WI, USA

⁴Graduate School of Agricultural Sciences, Tohoku University, Aoba, Sendai, Japan

Correspondence

Toshihiro Okamura, Institute of Livestock and Grassland Science, NARO, Tsukuba, Ibaraki 305-0074, Japan.
Email: okamat@affrc.go.jp

Funding information

Ministry of Agriculture, Forestry and Fisheries of Japan

Abstract

Direct selection for litter size or weight at weaning in pigs is often hindered by external interventions such as cross-fostering. The objective of this study was to infer the causal structure among phenotypes of reproductive traits in pigs to enable subsequent direct selection for these traits. Examined traits included: number born alive (NBA), litter size on day 21 (LS21), and litter weight on day 21 (LW21). The study included 6,240 litters from 1,673 Landrace dams and 5,393 litters from 1,484 Large White dams. The inductive causation (IC) algorithm was used to infer the causal structure, which was then fitted to a structural equation model (SEM) to estimate causal coefficients and genetic parameters. Based on the IC algorithm and temporal and biological information, the causal structure among traits was identified as: $NBA \rightarrow LS21 \rightarrow LW21$ and $NBA \rightarrow LW21$. Owing to the causal effect of NBA on LS21 and LW21, the genetic, permanent environmental, and residual variances of LS21 and LW21 were much lower in the SEM than in the multiple-trait model for both breeds. Given the strong effect of NBA on LS21 and LW21, the SEM and causal information might assist with selective breeding for LS21 and LW21 when cross-fostering occurs.

KEYWORDS

inductive causation, phenotypic causal structure, pig, reproduction, structural equation model

1 | INTRODUCTION

In the swine industry, the number of weaning piglets and their total weight are particularly important traits. Direct selection for these traits is often restricted in practice due to external interventions such as cross-fostering, whereby piglets are transferred between sows to equalize litter size. As a result, the number of nursing piglets at a sow can be different from the number of farrowing piglets. The purpose of cross-fostering is to reduce piglet mortality at preweaning (Straw, Dewey, & Burgi, 1998). This management technique, however, can make it difficult to adequately estimate genetic parameters for litter

size and weight at weaning (Su, Lund, & Sorensen, 2007). Therefore, in actual pig breeding, weaning is improved by selecting for total number born (Sorensen, Vernersen, & Andersen, 2000), number born alive (NBA) (Holl & Robinson, 2003), and litter size on day 5 based on the biological dam (Nielsen, Su, Lund, & Madsen, 2013; Su et al., 2007). These traits are not or only slightly affected by cross-fostering because they are generally evaluated before or just after it. However, the genetic correlation between total number born and mortality remains unfavorable, and selection for higher total number born will increase the number of stillborn piglets. Putz, Tiezzi, Maltecca, Gray, and Knauer (2015) suggested that a higher amount of cross-fostering

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Animal Science Journal* published by John Wiley & Sons Australia, Ltd on behalf of Japanese Society of Animal Science

reduced the genetic correlation between litter sizes in different days after farrowing. In the absence of cross-fostering, selection could rely directly on the number of weaning piglets and their total weight. However, the beneficial effect of cross-fostering on piglets' survival increases selection intensity and makes selection more efficient because more candidates of a specific sire and dam can be obtained.

Gianola and Sorensen (2004) adapted the structural equation model (SEM) to mixed-effects models in quantitative genetics to convey causal relationships among traits. Valente, Rosa, Gianola, Wu, and Weigel (2013) suggested that SEM made it possible to predict the effects of external interventions. Importantly, uncovering this information among reproductive traits in pigs might be useful for adequately evaluating weaning traits even in the case of cross-fostering. For such evaluation the causal structure among preweaning and weaning traits should be determined first, but previous studies about causal structure among reproductive traits in pig have not focused on weaning traits (Chitakasempornkul et al., 2019; Varona, Sorensen, & Thompson, 2007).

The objective of this study was to infer the causal structure and estimate causal coefficients among phenotypes of farrowing and weaning traits in two breeds of pigs. This study is the first to describe a phenotypic causal structure among farrowing and weaning reproductive traits in pigs.

2 | MATERIALS AND METHODS

2.1 | Ethical statement

Approval of Animal Care and Use Committee was not required for this study because the data were acquired from an existing database.

2.2 | Data

Data from Landrace and Large White populations from two farms for the years 2001–2017 were provided by CIMCO Corporation. All matings were performed by artificial insemination. Landrace, Large White, and Duroc sire breeds were used for matings. The NBA was recorded 1 day after farrowing and it took into account the number of piglets that seemed to be alive at farrowing but were actually dead. Records suspected of cross-fostering or with missing values were not included in the analyses. Overall, the final dataset comprised of 6,240 litters from

1,673 Landrace dams and 5,393 litters from 1,484 Large White dams. Parity varied from one to eight. Pedigree data for Landrace and Large White sows included reproductive data on 2,102 dams and 1,849 sires.

Reproductive traits were NBA, litter size on day 21 (LS21), and litter weight on day 21 (LW21) after farrowing. Descriptive statistics for each trait are presented in Table 1.

2.3 | Statistical analyses

A multiple-trait animal model (MTM) was used for initial analysis. The following MTM was considered:

$$y = X\beta + Zu + Wc + e,$$

where y is a vector of observations; β is a vector of systematic effects, including farrowing year (17 levels), farrowing month (12 levels), parity (8 levels), farm (2 levels), and mating sire breed (3 levels); u is a vector of random additive genetic effects; c is a vector of permanent environmental effects on the dams; e is a vector of random residuals; and X , Z , and W are known incidence matrices.

The joint distribution of the random vectors u , c , and e was given by:

$$\begin{pmatrix} u \\ c \\ e \end{pmatrix} = N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} G_0 \otimes A & 0 & 0 \\ 0 & C_0 \otimes I & 0 \\ 0 & 0 & R_0 \otimes I \end{pmatrix} \right],$$

where G_0 is the additive genetic (co)variance matrix, A is the additive (numerator) genetic relationship matrix, C_0 is the permanent environmental (co)variance matrix, I is an identity matrix with suitable dimensions, and R_0 is the residual (co)variance matrix. Such (co)variance matrices can be expressed as:

$$G_0 = \begin{pmatrix} \sigma_{u_1}^2 & \sigma_{u_1 u_2} & \sigma_{u_1 u_3} \\ \sigma_{u_1 u_2} & \sigma_{u_2}^2 & \sigma_{u_2 u_3} \\ \sigma_{u_1 u_3} & \sigma_{u_2 u_3} & \sigma_{u_3}^2 \end{pmatrix}, \quad C_0 = \begin{pmatrix} \sigma_{c_1}^2 & \sigma_{c_1 c_2} & \sigma_{c_1 c_3} \\ \sigma_{c_1 c_2} & \sigma_{c_2}^2 & \sigma_{c_2 c_3} \\ \sigma_{c_1 c_3} & \sigma_{c_2 c_3} & \sigma_{c_3}^2 \end{pmatrix},$$

$$R_0 = \begin{pmatrix} \sigma_{e_1}^2 & \sigma_{e_1 e_2} & \sigma_{e_1 e_3} \\ \sigma_{e_1 e_2} & \sigma_{e_2}^2 & \sigma_{e_2 e_3} \\ \sigma_{e_1 e_3} & \sigma_{e_2 e_3} & \sigma_{e_3}^2 \end{pmatrix}$$

Trait	Landrace			Large White		
	NBA	LS21	LW21	NBA	LS21	LW21
N	6,240	6,240	6,240	5,393	5,393	5,393
Mean	10.81	9.25	52.49	10.40	9.15	53.33
SD	2.90	2.47	12.30	2.71	2.39	12.94
Minimum	1.00	1.00	4.00	1.00	1.00	4.60
Maximum	20.00	17.00	93.20	20.00	16.00	88.70

TABLE 1 Descriptive statistics of reproductive traits

Abbreviations: LS21, litter size on day 21; LW21, litter weight on day 21; NBA, number born alive.

where $\sigma_{u_i}^2$ is the additive genetic variance of trait i , $\sigma_{u_i u_j}$ is the additive genetic covariance between traits i and j , $\sigma_{c_i}^2$ is the permanent environmental variance of trait i , $\sigma_{c_i c_j}$ is the permanent environmental covariance between traits i and j , $\sigma_{e_i}^2$ is the residual variance of trait i , $\sigma_{e_i e_j}$ is the residual covariance between traits i and j , and i and j for 1, 2, and 3 represent NBA, LS21, and LW21, respectively.

The program GIBBS2F90 (Misztal et al., 2002) was used to fit the model within a Bayesian approach and to implement Gibbs sampling in order to obtain samples from posterior distributions of genetic, permanent environmental, and residual (co)variances. A total of 1,000,000 samples were generated; however, 500,000 were discarded as a conservative burn-in and the remaining samples were thinned every 10 iteration, resulting in a total of 50,000 samples for posterior analysis. The results of Geweke's diagnostic (Geweke, 1992) and effective sample size derived by the program POSTGIBBSF90 (Misztal et al., 2002) were used to assess convergence with their recommended criteria.

2.4 | The inductive causation algorithm

The inductive causation (IC) algorithm was applied to the residual (co)variances obtained from MTM analysis to infer a potential causal structure among the three traits, as already proposed by Valente, Rosa, de los Campos, Gianola, and Silva (2010). The residual (co)variances obtained from the MTM provide information from the joint distribution of phenotypic traits conditional on genetic effects. This method corrects for confounding effects when traits are genetically correlated (Rosa et al., 2011; Valente et al., 2010). The IC algorithm performs a series of statistical decisions based on partial correlations (ρ) between traits and consists of the following three steps (Pearl, 2000):

Step 1: Based on ρ , a statistical decision is made as to whether two traits are connected by an undirected edge. If ρ conditioning of the combination of all other traits between traits Y1 and Y2 differs from 0, Y1 and Y2 are connected by an undirected edge (Y1-Y2).

Step 2: Based on ρ , a statistical decision is made about the existence of an unshielded collider. When three traits Y1, Y2, and Y3 are connected with undirected edges to form a trio such as Y1-Y2-Y3, whereby two nonadjacent traits (Y1 and Y3) have a common adjacent trait (Y2), and Y1 and Y3 are conditionally dependent on any possible set that includes the adjacent trait Y2, the edges should be oriented toward the common adjacent trait; for example, Y1 → Y2 ← Y3, then Y2 is considered an unshielded collider. If Y1 depends on Y3 or Y3 depends on Y1 (i.e., Y3 → Y1 or Y1 → Y3) in this case, Y2 is considered a shielded collider, which the IC algorithm cannot detect because Y1 and Y3 are always conditionally dependent on any possible set.

Step 3: When possible, the remaining undirected edges are oriented so that no new unshielded colliders or cycles are introduced.

Statistical decisions regarding whether to declare ρ as null or not were based on from 75% to 95% highest posterior density (HPD) intervals, in 5% increments. If the intervals contained the value 0,

the correlation was declared null. The analysis was performed using an R (R Development Core Team, 2009) script written by Valente and Rosa (2013).

2.5 | SEM analysis

The SEM was fitted to the causal network inferred by the IC algorithm. The model can be described as:

$$y = (\Lambda \otimes I) y + X\beta^* + Zu^* + Wc^* + e^*,$$

with the joint distribution of vectors u^* , c^* , and e^* as:

$$\begin{pmatrix} u^* \\ c^* \\ e^* \end{pmatrix} = N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} G_0^* \otimes A & 0 & 0 \\ 0 & \Gamma_0^* \otimes I & 0 \\ 0 & 0 & \Psi_0^* \otimes I \end{pmatrix} \right]$$

where vectors y , β^* , u^* , c^* , e^* , X , Z , A , and I have a similar meaning as described above for the MTM. However, here, these vectors represent systematic and random effects directly affecting each trait, that is, effects that are not mediated by other phenotypic traits (Gianola & Sorensen, 2004; Rosa et al., 2011; Valente et al., 2013). Additionally, Λ is a 3×3 matrix with 0 on the diagonal and with structural coefficients (linear effects between pairs of traits) or 0 on the off-diagonals, G_0^* is the SEM additive genetic (co)variance matrix (it describes variance and covariances of direct genetic effects), Γ_0^* is a diagonal matrix with the SEM permanent environmental variances, and Ψ_0^* is a diagonal matrix with the SEM residual variances. These permanent environmental and residual covariances were assumed to be 0 in the SEM. Such direct SEM (co)variance matrices can be expressed as:

$$G_0^* = \begin{pmatrix} \sigma_{u_1}^2 & \sigma_{u_1 u_2} & \sigma_{u_1 u_3} \\ \sigma_{u_1 u_2} & \sigma_{u_2}^2 & \sigma_{u_2 u_3} \\ \sigma_{u_1 u_3} & \sigma_{u_2 u_3} & \sigma_{u_3}^2 \end{pmatrix}, \quad \Gamma_0^* = \begin{pmatrix} \sigma_{c_1}^2 & 0 & 0 \\ 0 & \sigma_{c_2}^2 & 0 \\ 0 & 0 & \sigma_{c_3}^2 \end{pmatrix},$$

$$\Psi_0^* = \begin{pmatrix} \sigma_{e_1}^2 & 0 & 0 \\ \sigma_{e_1 e_2} & \sigma_{e_2}^2 & 0 \\ 0 & 0 & \sigma_{e_3}^2 \end{pmatrix}.$$

where $\sigma_{u_i}^2$ is the additive genetic variance of trait i , $\sigma_{u_i u_j}$ is the additive genetic covariance between traits i and j , $\sigma_{c_i}^2$ is the permanent environmental variance of trait i , $\sigma_{e_i}^2$ is the residual variance of trait i , and i and j for 1, 2, and 3 represent NBA, LS21, and LW21, respectively.

This model was fitted to estimate genetic parameters and causal coefficients representing causal effects with the assumed causal structure mentioned above. The covariances of permanent environmental and residual effect were conducted as diagonal to achieve parameter identifiability. Importantly, in the SEM, the causal parents of a given trait (e.g., Y1 is the causal parent of Y2 in Y1 → Y2) are included as covariates in the SEM assigned to that trait. The program

GIBBS2F90 was used to fit this model and to obtain posterior samples for SEM parameters; a Gibbs sampling strategy similar to that used for the MTM was employed here as well.

3 | RESULTS

The posterior means and 95% HPD intervals describing \mathbf{G}_0 , \mathbf{C}_0 , and \mathbf{R}_0 obtained with the MTM for each breed are listed in Table 2. No major differences were observed between Landrace and Large White pigs in terms of genetic and permanent environmental (co) variances. However, the residual variance of NBA was larger in Landrace than in Large White pigs with no 95% HPD interval overlap observed between breeds. In contrast, residual variance of LW21 and covariance between LS21 and LW21 were greater in Large White than Landrace.

Based on the posterior distribution of \mathbf{R}_0 , ρ between traits was estimated and applied in Step 1 for causal structure search using different HPD interval contents. None of the HPD intervals contained 0, indicating that the three traits were connected by an undirected edge (Figure 1). In our case, there was no unshielded collider because none of the traits were “unshielded”. Therefore, Step 2 of the IC algorithm was not able to indicate any link. This suggested the IC

algorithm could not detect any directed edges among these traits. Then, temporal prior knowledge about the relationship between NBA and LS21 or LW21 was applied to specify directions, that is, $\text{NBA} \rightarrow \text{LS21}$ and $\text{NBA} \rightarrow \text{LW21}$, as NBA is expressed before LS21 or LW21. We also applied prior biological knowledge about the relationship between LS21 and LW21, that is, $\text{LS21} \rightarrow \text{LW21}$, meaning that litter size affects total litter weight because the vice versa is generally difficult to be explained by a biological system. Hence, we assumed the causal structure shown in Figure 2, which describes the additive effect of genetic, permanent environmental, and residual parameters.

The dispersion parameters and their structural coefficients were fitted into the SEM with the assumed structure (Figure 2) and are presented in Tables 3 and 4. A comparison of dispersion parameters for each trait between MTM (Table 2) and SEM (Table 3) revealed that all variances for LS21 and LW21 (but not NBA) in the SEM were lower than in the MTM. These differences were found in both breeds although some residual variances differed significantly between them. Here, λ_{ij} denotes a structural coefficient from the j th to the i th trait; this means that when the j th trait increases by 1 unit, the i th trait increases by λ_{ij} units. The direct effect of NBA on LS21 ($\lambda_{\text{LS21, NBA}}$) was positive. The indirect effect of NBA on LW21 via LS21 ($\lambda_{\text{LS21, NBA}} \times \lambda_{\text{LW21, LS21}}$) was also positive, but the direct effect

TABLE 2 Posterior means and 95% highest posterior density (HPD) intervals for the dispersion parameters pertaining to the multiple-trait animal model

Parameter	Landrace			Large White		
	Posterior mean	95% HPD interval		Posterior mean	95% HPD interval	
		Low	High		Low	High
$\sigma_{u_1}^2$	0.98	0.60	1.38	0.87	0.48	1.28
$\sigma_{u_1 u_2}$	0.66	0.38	0.97	0.62	0.30	0.94
$\sigma_{u_1 u_3}$	1.26	-0.01	2.53	1.69	0.09	3.33
$\sigma_{u_2}^2$	0.63	0.37	0.89	0.57	0.29	0.84
$\sigma_{u_2 u_3}$	1.74	0.66	2.85	2.37	0.96	3.84
$\sigma_{u_3}^2$	19.90	13.39	26.81	29.76	19.51	40.19
$\sigma_{c_1}^2$	0.50	0.23	0.79	0.68	0.39	1.00
$\sigma_{c_1 c_2}$	0.42	0.19	0.64	0.54	0.30	0.79
$\sigma_{c_1 c_3}$	1.62	0.63	2.57	2.16	0.92	3.36
$\sigma_{c_2}^2$	0.40	0.20	0.61	0.45	0.24	0.67
$\sigma_{c_2 c_3}$	2.01	1.16	2.91	2.12	1.02	3.19
$\sigma_{c_3}^2$	14.59	9.42	19.77	17.24	9.97	24.54
$\sigma_{e_1}^2$	5.85	5.62	6.09	4.92	4.70	5.14
$\sigma_{e_1 e_2}$	4.06	3.88	4.25	3.70	3.51	3.88
$\sigma_{e_1 e_3}$	9.67	8.98	10.36	10.53	9.75	11.35
$\sigma_{e_2}^2$	4.59	4.40	4.77	4.37	4.18	4.56
$\sigma_{e_2 e_3}$	13.30	12.63	13.98	15.25	14.39	16.06
$\sigma_{e_3}^2$	83.54	80.14	86.96	107.28	102.40	112.00

Note: $\sigma_{c_i}^2$, permanent environmental variance of trait i ; $\sigma_{c_i c_j}$, permanent environmental covariance between traits i and j ; $\sigma_{e_i}^2$, residual variance of trait i ; $\sigma_{e_i e_j}$, residual covariance between traits i and j ; $\sigma_{a_i}^2$, additive genetic variance of trait i ; $\sigma_{a_i a_j}$, additive genetic covariance between traits i and j , and i and j for 1, 2, and 3 represent number born alive, litter size on day 21, and litter weight on day 21, respectively.

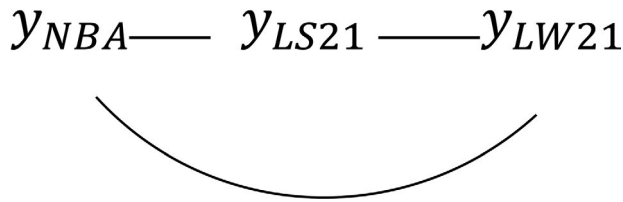


FIGURE 1 Undirected graph detected by the inductive causation algorithm with 95% HPD intervals. LS21, litter size on day 21; LW21, litter weight on day 21; NBA, number born alive

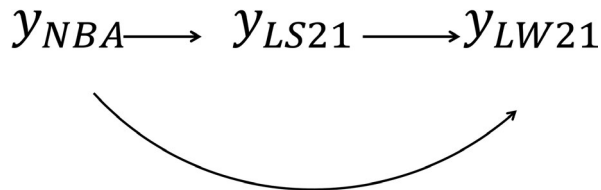


FIGURE 2 Directed graph assumed based on Figure 1 plus temporal and biological information. LS21, litter size on day 21; LW21, litter weight on day 21; NBA, number born alive

of NBA on LW21 ($\lambda_{LW21, NBA}$) was negative. Finally, the total effect of NBA on LW21 ($\lambda_{LW21, NBA} + \lambda_{LS21, NBA} \times \lambda_{LW21, LS21}$) was positive. Although all coefficients were significantly different between both breeds, their signs were always in the same direction.

4 | DISCUSSION

The objective of this study was to investigate the causal structure and estimate the causal coefficients among preweaning and weaning

traits in pigs, which could then be used for direct selection based on litter size and litter weight at weaning when cross-fostering is employed. Although it was impossible to artificially change the empirical NBA values, we assumed that this artificial change could be done by cross-fostering just after farrowing and NBA assessment. This assumption allowed us to determine the causal structure among NBA, LS21, and LW21, as well as estimate the causal coefficients.

The IC algorithm was based on 95% HPD intervals and returned a completely undirected structure. This result indicated that the structure was very reliable because the level for statistical decision was high and all causal coefficients were still statistically different from 0. Although many studies have reported a genetic relationship among reproductive traits (Chen, Baas, Mabry, Koehler, & Dekkers, 2003; Putz et al., 2015; Roehe & Kennedy, 1995), only a few have described phenotypic causal structures among reproductive traits in pigs. For example, Varona et al. (2007) observed a negative causal relationship between litter size and average piglet weight at birth in Large White animals. Chitakasempornkul et al. (2019) used the IC algorithm to infer a causal phenotypic network among reproductive traits at birth and found a negative relationship between NBA and average piglet body weight, with the latter negatively affecting the total number born during subsequent gestation in gilts. However, they did not focus on weaning traits and this study is the first to describe a phenotypic causal structure among farrowing and weaning reproductive traits in pigs. In addition, the structure described here was common in both Landrace and Large White breeds, suggesting that it may be stable among different breeds in spite of differences in magnitude of causal coefficients. In fact, these differences indicate that such coefficients may need to be estimated for each population or breed.

TABLE 3 Posterior means and 95% highest posterior density (HPD) intervals for the dispersion parameters pertaining to the structural equation model

Parameter	Landrace			Large White		
	Posterior mean	95% HPD interval		Posterior mean	95% HPD interval	
		Low	High		Low	High
$\sigma_{u_1}^2$	0.93	0.55	1.34	0.82	0.43	1.21
$\sigma_{u_1 u_2}$	0.03	-0.07	0.12	-0.02	-0.11	0.07
$\sigma_{u_1 u_3}$	0.26	-0.39	0.91	0.56	-0.21	1.37
$\sigma_{u_2}^2$	0.20	0.13	0.27	0.13	0.08	0.18
$\sigma_{u_2 u_3}$	0.66	0.35	0.97	0.80	0.46	1.13
$\sigma_{u_3}^2$	14.54	10.72	18.45	18.44	13.17	23.85
$\sigma_{c_1}^2$	0.53	0.23	0.82	0.71	0.42	1.02
$\sigma_{c_2}^2$	0.07	0.02	0.12	0.06	0.02	0.10
$\sigma_{c_3}^2$	5.65	3.11	8.11	7.65	4.28	11.09
$\sigma_{e_1}^2$	5.85	5.62	6.09	4.92	4.70	5.13
$\sigma_{e_2}^2$	1.76	1.69	1.83	1.58	1.51	1.64
$\sigma_{e_3}^2$	42.99	41.21	44.69	50.98	48.73	53.19

Note: $\sigma_{c_i}^2$ permanent environmental variance of trait i ; $\sigma_{e_i}^2$ residual variance of trait i ; $\sigma_{u_i}^2$ additive genetic variance of trait i ; $\sigma_{u_i u_j}$, additive genetic covariance between traits i and j , and i and j for 1, 2, and 3 represent number born alive, litter size on day 21, and litter weight on day 21, respectively.

TABLE 4 Posterior means and 95% highest posterior density (HPD) intervals for the causal coefficients pertaining to the structural equation model

Parameter	Landrace			Large White		
	Posterior mean	95% HPD interval		Posterior mean	95% HPD interval	
		Low	High		Low	High
$\lambda_{LS21,NBA}$	0.70	0.68	0.71	0.75	0.74	0.77
$\lambda_{LW21,NBA}$	-0.96	-1.08	-0.84	-1.33	-1.49	-1.17
$\lambda_{LW21,LS21}$	3.78	3.65	3.92	4.65	4.48	4.82
$\lambda_{LW21,LS21} \times \lambda_{LS21,NBA}$	2.64	2.53	2.75	3.50	3.34	3.65
$\lambda_{LW21,NBA} + \lambda_{LW21,LS21} \times \lambda_{LS21,NBA}$	1.68	1.58	1.77	2.16	2.04	2.29

Note: $\lambda_{i,j}$, structural coefficient of j on i ; LS21, litter size on day 21; LW21, litter weight on day 21; NBA, number born alive.

In our structure, NBA could directly affect LS21 and LW21, and indirectly affect LW21 via LS21. This means that LS21 was not only affected directly by u_{LS21}^* but also indirectly by u_{NBA}^* . LW21, on the other hand, was affected directly by u_{LW21}^* and indirectly by both u_{NBA}^* and u_{LS21}^* . We believe that u_{NBA}^* might represent genes affecting uterine size or ovulation rate, u_{LS21}^* may include genetic effects determining the piglets' survival (e.g., quality and quantity of colostrum or piglet crushing), and u_{LW21}^* may represent total volume and quality of milk produced (Valente et al., 2013). Cross-fostering can affect the estimation of genetic effects associated with these traits in the MTM because direct effects cannot be distinguished and can vary with the number of nursing piglets. If the number of nursing piglets is artificially changed by cross-fostering, these direct genetic effects can be accurately estimated by the SEM because in this case direct genetic effects are independent of the number of nursing piglets. Therefore, inferring a causal structure can benefit direct selection based on LS21 and LW21 when cross-fostering is applied.

The genetic, permanent environmental, and residual variances of LS21 and LW21 in the SEM were smaller than in the MTM in both breeds because components of these effects were different in the two models. In the MTM, the effects represent overall effects, which include all direct and indirect (i.e., mediated by other phenotypic traits) effects on each trait. In contrast, the effects represent only direct effects in the SEM (i.e., not mediated by other traits in the causal structure) (Valente et al., 2013). The reduction in variance of downstream traits conditioning upstream traits was in agreement with the results observed for bovine milk fatty acid (Bouwman, Valente, Janss, Bovenhuis, & Rosa, 2014) and bovine meat quality (Inoue et al., 2016). The substantial reduction in variances of LS21 and LW21 indicated that these two traits were strongly affected by NBA. Accordingly, an artificial change (i.e., an external intervention) in the number of nursing piglets by cross-fostering could strongly affect the phenotype and estimated breeding values of LS21 and LW21. In addition, its usefulness in breeding for LS21 and LW21 with cross-fostering on computer simulation or real data should be confirmed in the future.

In general, residual covariance matrix is constructed as diagonal to achieve parameter identifiability. The permanent environmental and residual covariance matrices were set as diagonal in the SEM of this study. Although this restriction was strong it was necessary to identify the

parameters, previous studies used the same method to identify them (Bouwman et al., 2014; Inoue et al., 2016; Rosa et al., 2011; Valente et al., 2010). However, when the covariances among permanent environmental effects or among residual effects would not be 0, the causal effects would be over- or underestimated. Such covariances can be taken account by the statistical model with proper fixed and random effects, but we cannot find whether the model was sufficient. Therefore, it is noted that the parameter could have resulted in bias in this study.

In conclusion, we describe here the causal structure among NBA, LS21, and LW21 based on the IC algorithm and temporal and biological information. The causal structure allowed for the estimation of causal coefficients. A comparison of dispersion parameters for LS21 and LW21 between the MTM and SEM indicated that the phenotype arising from LS21 and LW21 traits could be strongly affected by the NBA trait. This finding suggests that cross-fostering could have a big impact on LS21 and LW21, and thus the causal information might be useful for direct selection based on LS21 and LW21 when cross-fostering is employed.

ACKNOWLEDGMENTS

We wish to thank the CIMCO Corporation for providing the data. This work was supported by a grant from the Ministry of Agriculture, Forestry and Fisheries of Japan (Development of Breeding Technology for Animal Life Production).

ORCID

Toshihiro Okamura  <https://orcid.org/0000-0002-4691-4436>
 Guilherme J. M. Rosa  <https://orcid.org/0000-0001-9172-6461>
 Masahiro Satoh  <https://orcid.org/0000-0002-9198-6666>
 Osamu Sasaki  <https://orcid.org/0000-0002-0304-2172>

REFERENCES

- Bouwman, A. C., Valente, B. D., Janss, L. L. G., Bovenhuis, H., & Rosa, G. J. M. (2014). Exploring causal networks of bovine milk fatty acids in a multivariate mixed model context. *Genetics Selection Evolution*, 46, 2. <https://doi.org/10.1186/1297-9686-46-2>
- Chen, P., Baas, T. J., Mabry, J. W., Koehler, K. J., & Dekkers, J. C. M. (2003). Genetic parameters and trends for litter traits in US Yorkshire, Duroc, Hampshire, and Landrace pigs. *Journal of Animal Science*, 81, 46–53. <https://doi.org/10.2527/2003.81146x>

- Chitakasempornkul, K., Meneget, M. B., Rosa, G. J. M., Lopes, F. B., Jager, A., Goncalves, M. A. D., ... Bello, N. M. (2019). Investigating causal biological relationships between reproductive performance traits in high-performing gilts and sows. *Journal of Animal Science*, *97*, 2385–2401. <https://doi.org/10.1093/jas/skz115>
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In J. M. Bernardo, J. O. Berger, A. P. Dawid, & A. F. M. Smith (Eds.), *Bayesian statistics 4* (pp. 169–193). Oxford, UK: Oxford University Press.
- Gianola, D., & Sorensen, D. (2004). Quantitative genetic models for describing simultaneous and recursive relationships between phenotypes. *Genetics*, *167*, 1407–1424. <https://doi.org/10.1534/genetics.103.025734>
- Holl, J. W., & Robinson, O. W. (2003). Results from nine generations of selection for increased litter size in swine. *Journal of Animal Science*, *81*, 624–629. <https://doi.org/10.2527/2003.813624x>
- Inoue, K., Valente, B. D., Shoji, N., Honda, T., Oyama, K., & Rosa, G. J. M. (2016). Inferring phenotypic causal structures among meat quality traits and the application of a structural equation model in Japanese Black cattle. *Journal of Animal Science*, *94*, 4133–4142. <https://doi.org/10.2527/jas2016-0554>
- Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T., & Lee, D. (2002). BLUPF90 and related programs (BGF90). In *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France. [CD-ROM] Comm. no. 28.07.
- Nielsen, B., Su, G., Lund, M. S., & Madsen, P. (2013). Selection for increased number of piglets at d 5 after farrowing has increased litter size and reduced piglet mortality. *Journal of Animal Science*, *91*, 2575–2582. <https://doi.org/10.2527/jas.2012-5990>
- Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge, UK: Cambridge University Press.
- Putz, A. M., Tiezzi, F., Maltecca, C., Gray, K. A., & Knauer, M. T. (2015). Variance component estimates for alternative litter size traits in swine. *Journal of Animal Science*, *93*, 5153–5163. <https://doi.org/10.2527/jas.2015-9416>
- R Development Core Team. (2009). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Roehe, R., & Kennedy, B. W. (1995). Estimation of genetic parameters for litter size in Canadian Yorkshire and Landrace swine with each parity of farrowing treated as a different trait. *Journal of Animal Science*, *73*, 2959–2970. <https://doi.org/10.2527/1995.73102959x>
- Rosa, G. J. M., Valente, B. D., de los Campos, G., Wu, X.-L., Gianola, D., & Silva, M. A. (2011). Inferring causal phenotype networks using structural equation models. *Genetics Selection Evolution*, *43*, 6. <https://doi.org/10.1186/1297-9686-43-6>
- Sorensen, D., Vernersen, A., & Andersen, S. (2000). Bayesian analysis of response to selection: A case study using litter size in Danish Yorkshire pigs. *Genetics*, *156*, 283–295.
- Straw, B. E., Dewey, C. E., & Burgi, E. J. (1998). Patterns of crossfostering and piglet mortality on commercial US and Canadian swine farms. *Preventive Veterinary Medicine*, *33*, 83–89. [https://doi.org/10.1016/s0167-5877\(97\)00051-2](https://doi.org/10.1016/s0167-5877(97)00051-2)
- Su, G., Lund, M. S., & Sorensen, D. (2007). Selection for litter size at day five to improve litter size at weaning and piglet survival rate. *Journal of Animal Science*, *85*, 1385–1392. <https://doi.org/10.2527/jas.2006-631>
- Valente, B. D., & Rosa, G. J. M. (2013). Mixed effects structural equation models and phenotypic causal networks. In C. Gondro, J. van der Werf, & B. Hayes (Eds.), *Genome-wide association studies and genomic prediction* (pp. 449–464). New York, NY: Humana Press. https://doi.org/10.1007/987-1-62703-447-0_21
- Valente, B. D., Rosa, G. J. M., de los Campos, G., Gianola, D., & Silva, M. A. (2010). Searching for recursive causal structures in multivariate quantitative genetics mixed models. *Genetics*, *185*, 633–644. <https://doi.org/10.1534/genetics.109.112979>
- Valente, B. D., Rosa, G. J. M., Gianola, D., Wu, X. L., & Weigel, K. (2013). Is structural equation modeling advantageous for the genetic improvement of multiple traits? *Genetics*, *194*, 561–572. <https://doi.org/10.1534/genetics.113.151209>
- Varona, L., Sorensen, D., & Thompson, R. (2007). Analysis of litter size and average litter weight in pigs using a recursive model. *Genetics*, *177*, 1791–1799. <https://doi.org/10.1534/genetics.107.077818>

How to cite this article: Okamura T, Ishii K, Nishio M, Rosa GJM, Satoh M, Sasaki O. Inferring phenotypic causal structure among farrowing and weaning traits in pigs. *Anim Sci J*. 2020;91:e13369. <https://doi.org/10.1111/asj.13369>