Searching for phenotypic causal networks of reproductive and body weight traits in Syrian hamsters

T. Okamura¹, M. Nishio¹, K. Ishii¹, M. Satoh² & O. Sasaki¹

¹ Institute of Livestock and Grassland Science, NARO, 2 Ikenodai Tsukuba, Ibaraki 3050901, Japan
okamut@affrc.go.jp (Corresponding Author)
² Graduate School of Agricultural Sciences, Tohoku University, 468-1 Aramaki Aza Aoba, Aoba-ku, Sendai 9800845, Japan

Summary

Inferring the underlying traits of phenotypic networks can be useful for predicting the behaviors of complex systems. It can also help in optimizing management practices and multi-trait selection strategies in livestock. In this study, we estimated the causal structure among the growth and reproductive traits of the Syrian hamsters. The total number of live sibs at 0, 1, and 21 days after birth (NL0, NL1, and NL21, respectively) and litter weight at weaning 21 days after birth (W21) were recorded. In addition, individual body weight was measured at an age of 8 weeks (BW). We used 894–901 phenotypes of the female hamsters. Causal structure among the phenotypes in each set was estimated using the inductive causation (IC) algorithm after adjusting the data for genetic effects. A standard multiple-trait model was fitted using Bayesian methods to obtain a posterior (co)variance matrix of the phenotypes conditional to unobservable additive genetic effects. We estimated the causal structure among the growth and reproductive traits with the IC algorithm and temporal and biological information. BW had a positive effect for NL0 (0.13 head/g), and NL0 had a positive effect for NL1 (0.78 head/heads). NL1 affected both NL21 (0.65 head/heads) and W21 (−5.73 g/heads), and NL21 affected W21 (22.48 g/heads). These observations suggested that the causal effect of NL1 on W21 acted both directly and indirectly via NL21. Inferring the causal structure among reproductive traits can help predict the extent of change in traits under artificial intervention such as cross-fostering. For improving W21, the aggregated breeding values that are composed of the direct breeding value of W21 and the indirect positive breeding value of NL21 will be more useful than the traditional breeding values. Further, we suggest that inferring the causal structure is useful for constructing a statistical model for breeding.

Keywords: causal structure, reproductive traits, Syrian hamster

Introduction

The relationships among economical traits have been studied traditionally based on multi-trait models used in animal breeding. However, these models are not sufficiently stable to predict how variables will change because of external interventions. Inferring phenotypic networks underlying the traits can be useful to predict the behaviors of complex systems and thus to optimize the management practices and multi-trait selection strategies in the livestock industry (Rosa et al., 2011). In particular, improving the reproductive traits in animals which can have multiple pregnancies such as pigs, requires the simultaneous improvement of traits
at birth and weaning, but few studies have delved into this matter. Therefore, in this study, we used Syrian hamsters as representative those animals and estimated the causal structure among their growth and reproductive traits. In addition, we propose that inferring the structure will be useful for breeding for a reproductive trait affected by the other reproductive traits.

Material and methods

Data

The base population of Syrian hamsters for this experiment was established from a four-way cross with four inbreeding strains that were unrelated each other. Two F1 crosses from four inbred lines were mated to each other, and 37 litters were produced. Thereafter, one generation was produced by paring 1–3 females with 1–3 by random mating, resulting in 106 litters (generation 0, G0). We randomly selected 205 males and 205 females at 9 weeks of age from this G0 population and mated one female per male for 4 days. After mating, females delivered alone, and 21 days after delivery, the sibs were weaned. We repeated these cycles 5 times, and finally, we obtained the 6th generation population.

The total number of live sibs at 0, 1, and 21 days after birth (NL0, NL1, and NL21, respectively) and litter weight at weaning 21 days after birth (W21) were recorded. In addition, individual body weight was measured at the age of 8 weeks (BW). We used only the data of the female individuals for analysis and eliminated individuals in which the difference between NL0 and NL21 was larger than 10. Finally, we used 894–901 phenotypes (Table 1). The total number of animals in pedigree was 7,053.

Multi-trait analysis

Genetic and residual (co)variance among traits were estimated by the Gibbs sampling method. The following statistical model was fitted:

\[
\begin{aligned}
\mathbf{y} &= \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \mathbf{e}, \\
\mathbf{b} &= \mathbf{G}_0\mathbf{b}_0, \\
\mathbf{u} &= \mathbf{A}\mathbf{u}, \\
\mathbf{e} &= \mathbf{R}\mathbf{u},
\end{aligned}
\]

where, \(\mathbf{y}\) is a vector of phenotypic values; \(\mathbf{X}\) is a known incidence matrix of \(\mathbf{b}\) on \(\mathbf{y}\); \(\mathbf{b}\) is a vector for the fixed effects of the generation; \(\mathbf{Z}\) is a known incidence matrix of \(\mathbf{u}\) on \(\mathbf{y}\); \(\mathbf{u}\) is a vector of random additive genetic effects; and \(\mathbf{e}\) is a vector of random residuals. \(\mathbf{G}_0\) is the additive genetic (co)variance matrix; \(\mathbf{A}\) is the additive genetic relationship matrix; \(\mathbf{R}\) is the residual (co)variance matrix; and \(\mathbf{I}_n\) is an identity matrix. GIBBS1F90 program was used to obtain the posterior distributions of parameters. One cycle of 1,000,000 iterations was run, with a 750,000 burn-in and thinning of 10 iterations.

Causal structure

The causal structure among traits was estimated following Valente et al. (2010). They used an algorithm called the inductive causation (IC) algorithm. We used the residual (co)variance of the multiple-trait model (MTM) to apply this algorithm to our data because correlated genetic effects are confounding factors. In the IC algorithm, the structure is estimated based on statistical decisions that are based on partial correlations between traits. These partial correlations were obtained using the posterior samples of the residual (co)variance of the
MTM. When 0 was not included in the highest posterior density (HPD) of a partial correlation, we declared the partial correlation non-null. In this study, the causal structure was estimated by the IC algorithm via the following three steps.

Step 1
Based on the partial correlations, it was statistically decided whether two traits were connected by an undirected edge. If the partial correlation conditioning the combination of all the other traits between two traits Y1 and Y2 differs from 0, it is decided that Y1 and Y2 are connected by an undirected edge (Y1-Y2).

Step 2
Based on the partial correlations, it was statistically decided whether an unshielded collider exists. When three traits Y1, Y2, and Y3 are connected at an undirected edge (Y1 - Y2 - Y3), if the partial correlations of two non-adjacent traits (Y1 and Y3) that have a common adjacent trait (Y2) in such an undirected graph are conditionally dependent on any possible set that includes the adjacent trait, the edges should be oriented towards the common adjacent trait; for example, Y1 \(\rightarrow\) Y2 \(\leftarrow\) Y3.

Step 3
When possible, the remaining undirected edges were oriented with the temporal and biological information of traits in a way that introduced no new unshielded colliders or cycles.

**Structural equation models**

The causal structure estimated by the IC algorithm was quantified by structural equation modelling (SEM), following Gianola & Sorensen (2004). The causal parents of a given trait were considered as covariates in the equations assigned to this trait (in the case of Y1 \(\rightarrow\) Y2, the covariate of Y2 is Y1) and a residual (co)variance matrix was imposed. Therefore, the following statistical model was fitted:

With the joint distribution of vector of \(u^*\) and \(e^*\), as:

where, \(\Lambda\) is a \(t \times t\) (with \(t\) equal to the number of traits) matrix with 0’s on the diagonal and with structural coefficients or 0’s on the off-diagonals. The causal structure defines which of the off-diagonal entries of \(\Lambda\) must be estimated and which ones are set to 0. \(G_{0}^*\) is the SEM additive genetic (co)variance matrix, and \(\Psi_0\) is a diagonal matrix with the SEM residual variances. The residual covariances between the traits in the SEM were assumed 0, conferring identifiability to the structure coefficients in the likelihood function. The software and these conditions were the same as those used for the multi-trait model.

**Results & discussion**

The undirected network is shown in Figure 1a based on the IC algorithm using the sampling of \(R\) obtained from the MTM. This structure was considered robust because we estimated it based on the 95% HPD interval. No unshielded colliders were detected from the data in Step 2. We oriented the undirected edges shown in Figure 1a according to temporal and biological
information. Thus, we estimated the causal structure as shown in Figure 1b. The temporal sequence followed by the phenotypic traits is (1) BW, (2) NL0, (3) NL1, and (4) NL21 and W21. We assumed that NL21 could affect W21 because this flow is reasonable for the biology. The causal structure shown in Figure 1b presents an unshielded collider that was not detected in Step 2, but we accepted this collider because the directed edges composing the collider were robust and reasonable for temporal and biological information. The causal structure shown in Figure 1b was quantified using SEM. The structure for matrix was obtained as follows:

The posterior means of the structural coefficients are shown in Figure 1c. These coefficients indicate that when the parent traits increase by 1 unit, the affected traits change by the degree of the coefficient. In particular, the causal effect of NL1 on W21 was both direct and indirect via NL21. The direct effect is negative (-5.73g/heads), which is considered to be intense competition for the effects of the mother, for example, milk. On the other hand, the indirect effect via NL21 is positive (22.48×0.65=14.61 g/heads), which is considered to be an increase in the number of individuals constituting W21. Inferring the causal structure among reproductive traits can help predict the extent of changes in the traits with artificial intervention such as cross-fostering.

Considering causal parent traits as covariates in the model for estimating breeding values of the affected trait can extract only the breeding value that directly affects the trait from amongst all the breeding values that affect the trait both directly and indirectly. To achieve the purpose of improving W21, the breeding values considering these causal structures W21 can be more useful for breeding than the traditional breeding values. We suggest that inferring the causal structure is useful for constructing a statistical model for breeding.

**List of References**


Table 1. Mean values and standard deviation for each trait.

<table>
<thead>
<tr>
<th>Trait</th>
<th>NL0</th>
<th>NL1</th>
<th>NL21</th>
<th>W21</th>
<th>BW</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>901</td>
<td>901</td>
<td>895</td>
<td>894</td>
<td>901</td>
</tr>
<tr>
<td>Mean</td>
<td>11.34</td>
<td>9.92</td>
<td>7.93</td>
<td>202.30</td>
<td>85.35</td>
</tr>
<tr>
<td>SD</td>
<td>2.56</td>
<td>2.69</td>
<td>2.58</td>
<td>69.22</td>
<td>9.06</td>
</tr>
</tbody>
</table>

1 NL0, NL1, and NL21 = total number of born alive at 0, 1, and 21 days after birth, respectively; W21 = litter weight at weaning 21 days after birth; BW = individual bodyweight at 8 weeks of age.

Figure 1. Graphs of the reproductive and bodyweight traits of the Syrian hamsters. The output of IC algorithm using 95%HPD (a); combining (a) with the temporal and biological information (b) and structural coefficients NL0, NL1, NL21=total number of born alive at 0, 1, and 21 days after birth, respectively (c); W21=litter weight at weaning 21 days after birth; BW=individual bodyweight at 8 weeks of age.