Genomic prediction in pigs using a single-step method

(S 58 No 10164)

Christensen O.F.[1], Madsen P.[1], Nielsen B.[2], Ostersen, T.[2] and Su G.[1]

1. Aarhus University, Dept. Molecular Biology and Genetics
2. Danish Agriculture & Food Council, Pig Research Centre

EAAP 2011, Stavanger

Genomic prediction in pigs

□ Danish pig breeding

□ Single-step method:

    ▷ Misztal, Legarra, Aguilar + coworkers, Christensen and Lund.

    ▷ Relationship matrix from both markers and pedigree.

□ Aim here: compare single-step for daily gain and feed conversion ratio in Danish Duroc pigs.

Duroc data

- Daily gain 30-100 kg (all animals), feed conversion ratio ($< 10\%$ of animals).

- Genotyping: Illumina 60k SNP-chip. 3,554 animals. 25,720 SNP after editing.

- Pedigree: nearly all great-grandparents known, no import.

- Genotyped animals : relationship is 0.298, inbreeding is 0.145.

- Model: start-weight, herd-week-section (fixed); pen, genetic (random). For daily gain, in addition: sex (fixed), litter (random).

Training and validation data sets

□ Split by 1st October 2008.

□ Number of animals

	training	validation
Daily gain	313,068	41,771
Feed conv.	23,628	5,323
Genotyped	2,001	1,553

□ Comparison based on $\mathrm{Cor}(GBV, y_a)$ where $y_a = \hat{g} + \hat{e}$ adjusted phenotype. Bias: check slope in regression of y_a on GBV is 1.

Single-step method

☐ Model:
$$y = X\beta + u + g + e$$

where genetic effects

$$g \sim N(0, \sigma_g^2 H)$$

☐ Sparse inverse

$$H^{-1} = \begin{bmatrix} G_w^{-1} - A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} + A^{-1}.$$

where

$$G_w = (1 - w)G + wA_{11}.$$

with G genomic matrix and A pedigree matrix.

☐ Polygenic weight: here $w = 0.2$.

Genomic Relationship matrix

☐ Genomic relationship matrix G based on SNPs:

$$G = (m - \bar{m})(m - \bar{m})^T / s$$

where

$$m_{ij} = \begin{cases} -1 & \text{homozygote 11} \\ 0 & \text{heterozygote 12} \\ 1 & \text{homozygote 22} \end{cases}$$

☐ $\bar{m}_j = 2(p_j - 1)$ is observed average.

☐ $s = 2 \sum_j p_j (1 - p_j)$

☐ For Duroc pig data, observed G has

$$\text{Avg-diag}(G) = 0.99 \quad \text{Avg-offdiag}(G) \approx 0$$

Adjusted G

☐ Combined relationship matrix

$$H = \begin{bmatrix} G & GA_{11}^{-1}A_{12} \\ A_{21}A_{11}^{-1}G & A_{22} + A_{21}A_{11}^{-1}(G - A_{11})A_{11}^{-1}A_{12} \end{bmatrix}.$$

☐ Need: adjust G to be compatible to A_{11}.

☐ Use $G_a = G/\tilde{s} + \alpha$ instead of G, where $\tilde{s}$ and α solve system equations

$$\text{Avg-diag}(G)/\tilde{s} + \alpha = \text{Avg-diag}(A_{11}) = 1.145.$$

$$\text{Avg-offdiag}(G)/\tilde{s} + \alpha = \text{Avg-offdiag}(A_{11}) = 0.298$$

☐ Idea based on Forni et al. (2011) and Vitezica et al. (2011).

Results – daily gain

Method	$\mathrm{Cor}(GBV, y_a)$	regr. (bias)
Ped	0.183	1.021
G 1-step	0.242	1.011
G_a 1-step	0.251	1.036

Correlations different ? (Hotelling-Williams t-test)

Method	G 1-step	G_a 1-step
Ped	$< 1e - 60$	$< 1e - 60$
G 1-step	$\cdot$	$< 1e - 60$

Results – daily gain

☐ Single step with adjusted G is significantly better.

☐ No bias for single-step with adjusted G.

☐ Single step gives predictions for all animals.

☐ Single-step is better than pedigree method, both for genotyped animals and non-genotyped animals.

Results – feed conversion

Method	$\mathrm{Cor}(GBV, y_a)$	regr. (bias)
Ped	0.105	1.012
G 1-step	0.147	0.911
G_a 1-step	0.156	0.992

All correlations are significantly different !

☐ Single step with adjusted G is significantly better.

☐ No bias for single-step with adjusted G.

Bivariate – preliminary results

(only adjusted G is presented)

Random effect correlations:

genetic	pen	residual
-0.39	-0.99	-0.48

Predictions:

Trait	$\mathrm{Cor}(GBV, y_a)$-Biv	$\mathrm{Cor}(GBV, y_a)$-Univ	signif ?
Daily gain	0.250	0.251	yes
Fedd conv.	0.184	0.156	yes

Take-home messages

☐ **Conclusion:** Single-step method needs an adjusted G.

☐ **Discussion:** Does single-step need further improvement ?

☐ **Personal opinions:**

    ▷ Single-step is conceptually simpler.

    ▷ Single-step is easier to use in genetic evaluation system.

    ▷ Comparisons based on correlation/reliability should include p-values !