

Genomic evaluation for both purebred and crossbred performance

Ole F. Christensen

Aarhus University, Center for Quantitative Genetics and Genomics

EAAP 2014, Copenhagen

Motivation (pigs)

- ▶ Genomic selection has been implemented for purebreds.
- ▶ Crossbreeding is predominant in the production.
- ▶ Ideally, genomic evaluation should include records on crossbreds, and breeding values for crossbred performance should be presented.

Aim here: develop methods

- ▶ Two-breed terminal crossbreeding
- ▶ Single-step method (pedigree and genomic)

Outline

- ▶ Genomic model for purebreds and crossbreds.
- ▶ Pedigree-based model for purebreds and crossbreds.
- ▶ Single-step method for purebreds and crossbreds
- ▶ Future work

Genomic model for purebreds and crossbreds

- ▶ Purebreds

$$\mathbf{y}_A = \mu_A \mathbf{1} + \mathbf{m}_A \boldsymbol{\beta}_A + \mathbf{e}_A,$$

$$\mathbf{y}_B = \mu_B \mathbf{1} + \mathbf{m}_B \boldsymbol{\beta}_B + \mathbf{e}_B,$$

where $\mathbf{m}_{ij}^A = 0, 1, 2$ for genotypes 11, 12, 22 (same for \mathbf{m}_{ij}^B).

- ▶ Crossbreds (Ibanez et al , 2009):

$$\mathbf{y}_{AB} = \mu_{AB} \mathbf{1} + \mathbf{q}_A \boldsymbol{\beta}_A^c + \mathbf{q}_B \boldsymbol{\beta}_B^c + \mathbf{e}_{AB}$$

where $\mathbf{q}_{ij}^A = 0, 1$ for haplotypes 1, 2 (same for \mathbf{q}_{ij}^B).

- ▶ Genotypes phased to breed of origin.

Genomic model for purebreds and crossbreds

$$\begin{aligned}\mathbf{y}_A &= \mu_A \mathbf{1} + \mathbf{a}_A + \mathbf{e}_A, \\ \mathbf{y}_B &= \mu_B \mathbf{1} + \mathbf{a}_B + \mathbf{e}_B, \\ \mathbf{y}_{AB} &= \mu_{AB} \mathbf{1} + \mathbf{c}_{AB}^A + \mathbf{c}_{AB}^B + \mathbf{e}_{AB}\end{aligned}$$

- ▶ BVs for purebred performance: $\mathbf{a}_A = \mathbf{m}_A \boldsymbol{\beta}_A$, $\mathbf{a}_B = \mathbf{m}_B \boldsymbol{\beta}_B$
- ▶ Partial genetic effects: $\mathbf{c}_{AB}^A = \mathbf{q}_A \boldsymbol{\beta}_A^c$, $\mathbf{c}_{AB}^B = \mathbf{q}_B \boldsymbol{\beta}_B^c$.
- ▶ BVs for crossbred performance: $\mathbf{c}_A = \mathbf{m}_A \boldsymbol{\beta}_A^c$, $\mathbf{c}_B = \mathbf{m}_B \boldsymbol{\beta}_B^c$
- ▶ Equiv. model: 2 Genomic breed specific "partial" relationship matrices, $\mathbf{G}^{(A)}$, $\mathbf{G}^{(B)}$.

Genomic breed \mathcal{A} partial relationship matrix

$$\mathbf{G}^{(\mathcal{A})} = \begin{bmatrix} \mathbf{G}_{\mathcal{A},\mathcal{A}}^{(\mathcal{A})} & \mathbf{G}_{\mathcal{A},\mathcal{AB}}^{(\mathcal{A})} \\ \mathbf{G}_{\mathcal{AB},\mathcal{A}}^{(\mathcal{A})} & \mathbf{G}_{\mathcal{AB},\mathcal{AB}}^{(\mathcal{A})} \end{bmatrix}$$

$$\mathbf{G}_{\mathcal{A},\mathcal{A}}^{(\mathcal{A})} = \frac{(\mathbf{m}^{\mathcal{A}} - 2\bar{\rho}^{\mathcal{A}}\mathbf{1}^{\text{T}})(\mathbf{m}^{\mathcal{A}} - 2\bar{\rho}^{\mathcal{A}}\mathbf{1}^{\text{T}})^{\text{T}}}{s^{(\mathcal{A})}},$$

$$\mathbf{G}_{\mathcal{A},\mathcal{AB}}^{(\mathcal{A})} = \frac{(\mathbf{m}^{\mathcal{A}} - 2\bar{\rho}^{\mathcal{A}}\mathbf{1}^{\text{T}})(\mathbf{q}^{\mathcal{A}} - \bar{\rho}^{\mathcal{A}}\mathbf{1}^{\text{T}})^{\text{T}}}{s^{(\mathcal{A})}},$$

$$\mathbf{G}_{\mathcal{AB},\mathcal{AB}}^{(\mathcal{A})} = \frac{(\mathbf{q}^{\mathcal{A}} - \bar{\rho}^{\mathcal{A}}\mathbf{1}^{\text{T}})(\mathbf{q}^{\mathcal{A}} - \bar{\rho}^{\mathcal{A}}\mathbf{1}^{\text{T}})^{\text{T}}}{s^{(\mathcal{A})}}.$$

- ▶ For breed \mathcal{A} : $\mathbf{m}_{ij}^{\mathcal{A}} = 0, 1, 2$ for genotypes 11, 12, 22
- ▶ For crossbreds: $\mathbf{q}_{ij}^{\mathcal{A}} = 0, 1$ for haplotypes 1, 2.
- ▶ Allele frequency: $\bar{\rho}_j$.

Pedigree-based model

(Wei and van der Werf, 1994)

$$\mathbf{y}_A = \mu_A \mathbf{1} + \mathbf{Z}_A \mathbf{a}_A + \mathbf{e}_A,$$

$$\mathbf{y}_B = \mu_B \mathbf{1} + \mathbf{Z}_B \mathbf{a}_B + \mathbf{e}_B,$$

$$\mathbf{y}_{AB} = \mu_{AB} \mathbf{1} + \mathbf{c}_{AB}^A + \mathbf{c}_{AB}^B + \mathbf{e}_{AB},$$

where

- ▶ BVs for purebred performance: \mathbf{a}_A , \mathbf{a}_B .
- ▶ Breed of origin specific genetic effects: \mathbf{c}_{AB}^A , \mathbf{c}_{AB}^B .
- ▶ BVs for crossbred performance: \mathbf{c}_A , \mathbf{c}_B .

Pedigree-based model

- ▶ BVs for crossbred performance \mathbf{c}_A are correlated with BVs for purebred performance \mathbf{a}_A (genetic correlation) and with partial genetic effects $\mathbf{c}_{AB}^{(A)}$ (partial pedigree relationships).
- ▶ Supports a breeding goal with both purebred and crossbred performance.
- ▶ Allows different genetic variances in two breeds.
- ▶ \mathbf{c}_A , \mathbf{c}_B captures dominance effects, different genetic background, $G \times E$

Genetic covariances

- ▶ Breed \mathcal{A} :

$$\text{Var} \begin{bmatrix} \mathbf{a}_{\mathcal{A}}^{(\mathcal{A})} \\ * \\ \mathbf{c}_{\mathcal{A}}^{(\mathcal{A})} \\ \mathbf{c}_{\mathcal{AB}}^{(\mathcal{A})} \end{bmatrix} = \boldsymbol{\Sigma}_g^{(\mathcal{A})} \otimes \mathbf{A}^{(\mathcal{A})}$$

where $\mathbf{A}^{(\mathcal{A})}$ is [partial relationship matrix](#) for breed \mathcal{A} (Garcia-Cortes and Toro, 2006).

- ▶ $\boldsymbol{\Sigma}_g^{(\mathcal{A})}$: 2×2 genetic variance-covariance matrix
- ▶ Breed \mathcal{B} is similar.
- ▶ Breed \mathcal{A} and breed \mathcal{B} specific genetic effects are independent.

Partial relationship matrix $\mathbf{A}^{(\mathcal{A})}$

Recursions (Garcia-Cortes, 1996) :

$$\mathbf{A}_{i,i'}^{(\mathcal{A})} = (\mathbf{A}_{f(i),i'}^{(\mathcal{A})} + \mathbf{A}_{m(i),i'}^{(\mathcal{A})})/2$$

$$\mathbf{A}_{i,i}^{(\mathcal{A})} = f_i^{\mathcal{A}} + (\mathbf{A}_{f(i),i'}^{(\mathcal{A})} + \mathbf{A}_{m(i),i'}^{(\mathcal{A})})/2$$

where $f_i^{\mathcal{A}} = 1, 0, 1/2$ for breed \mathcal{A} , breed \mathcal{B} , and crossbred \mathcal{AB} .

- ▶ Ignore breed \mathcal{B} animals in relationship matrix.
- ▶ Sparse inverse can be computed directly.
- ▶ Submatrix can be computed using Colleau (2002) algorithm.

Combined relationship matrix

(pedigree and genomic)

- ▶ Combined partial relationship matrix:

$$(\mathbf{H}^{(A)})^{-1} = \begin{bmatrix} (\mathbf{G}^{(A)})^{-1} - (\mathbf{A}_{11}^{(A)})^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + (\mathbf{A}^{(A)})^{-1}, \quad (1)$$

- ▶ $\mathbf{G}_a^{(A)}$: genomic partial relationship matrix
- ▶ Sparse inverse $(\mathbf{A}^{(A)})^{-1}$ computed directly.
- ▶ $\mathbf{A}_{11}^{(A)}$ computed using Colleau algorithm

Single-step method

- ▶ Partial relationship matrices based on pedigree are replaced by combined partial relationship matrices.
- ▶ Can be fitted using standard software for multi-trait genetic evaluation using $(\mathbf{H}^{(A)})^{-1}$ and $(\mathbf{H}^{(B)})^{-1}$ as input.
- ▶ Compatibility of partial **G** and **A**: Linear adjustment of partial **G** towards partial **A**.
- ▶ Assumption: marker-data for crossbred animals can be phased to breed of origin.
- ▶ Method is available!

Acknowledgements

- ▶ **Co-authors:** Per Madsen, Bjarne Nielsen, Guosheng Su
- ▶ **Project:** "Developing genomic selection for a pig breeding system based on crossbreeding"
- ▶ **Funding:** through "Green Development and Demonstration Programme" (GUDP) by Danish Ministry of Food, Agriculture and Fisheries, Pig Research Centre and Aarhus University.

Future work

- ▶ PhD student Tao Xiang will use the model on maternal trait in Landrace (L), Yorkshire (Y) and crossbred (LY).
- ▶ Questions of interest to investigate:
 - ▶ Genetic correlation between purebred and crossbred performance.
 - ▶ Accuracy of breeding values for crossbred performance.
- ▶ Extension to three-breed terminal crossbreeding is possible.