

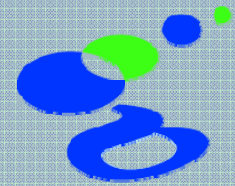
2009 ADSA-CSAS-ASAS Joint Annual Meeting
Montréal, July 12-16

Equivalent Mixed Model for Joint Genetic Evaluation Considering Molecular and Phenotypic Information

N. Gengler^{1,2}, and F. Colinet¹

¹ Gembloux Agricultural University, Animal Science Unit, Belgium

² National Fund for Scientific Research, Belgium



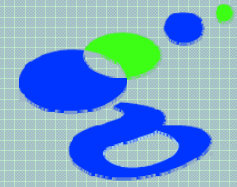
FNRS



SPW

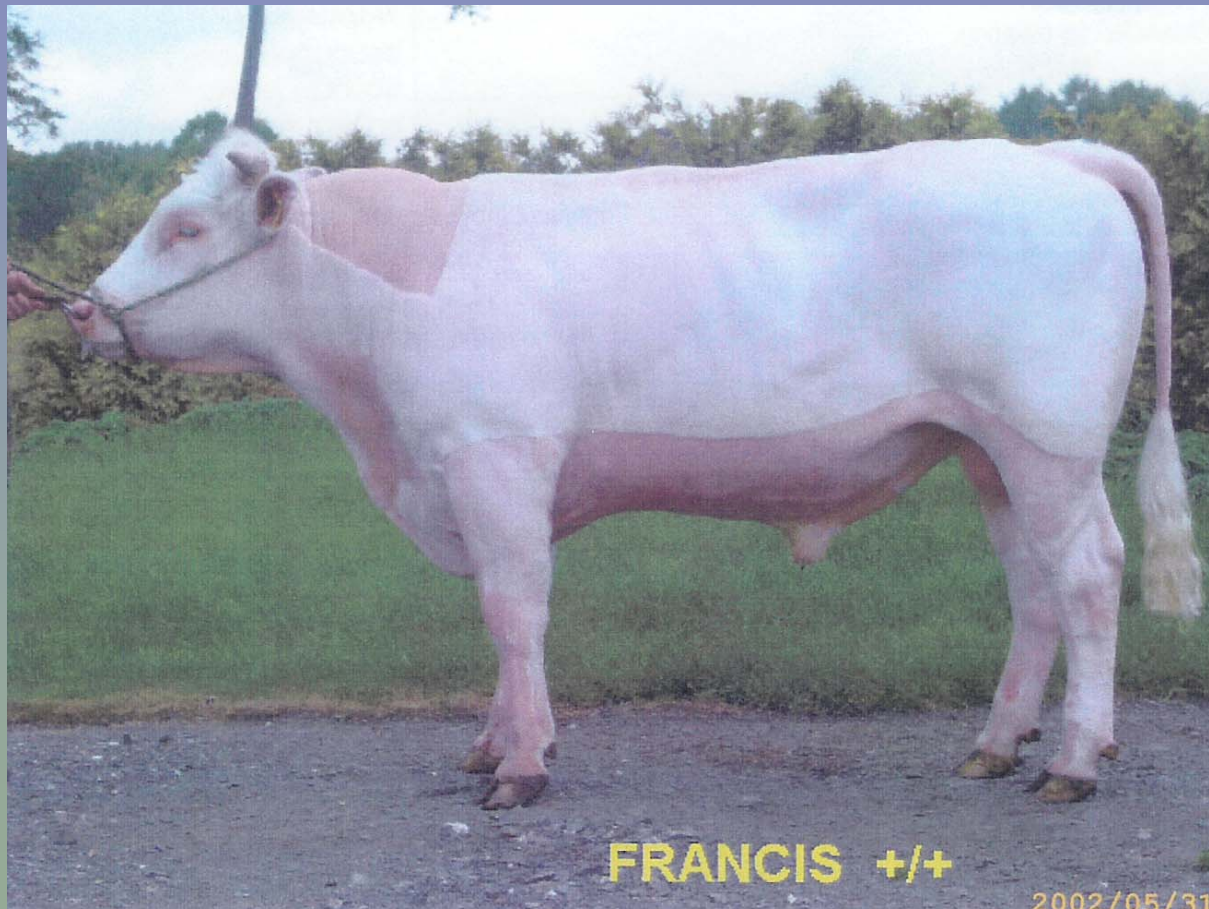
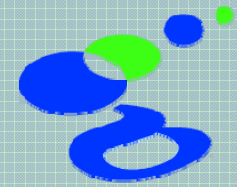
Service public de Wallonie

Dual Purpose Belgian Blue Breed (dp-BBB): Context of This Study

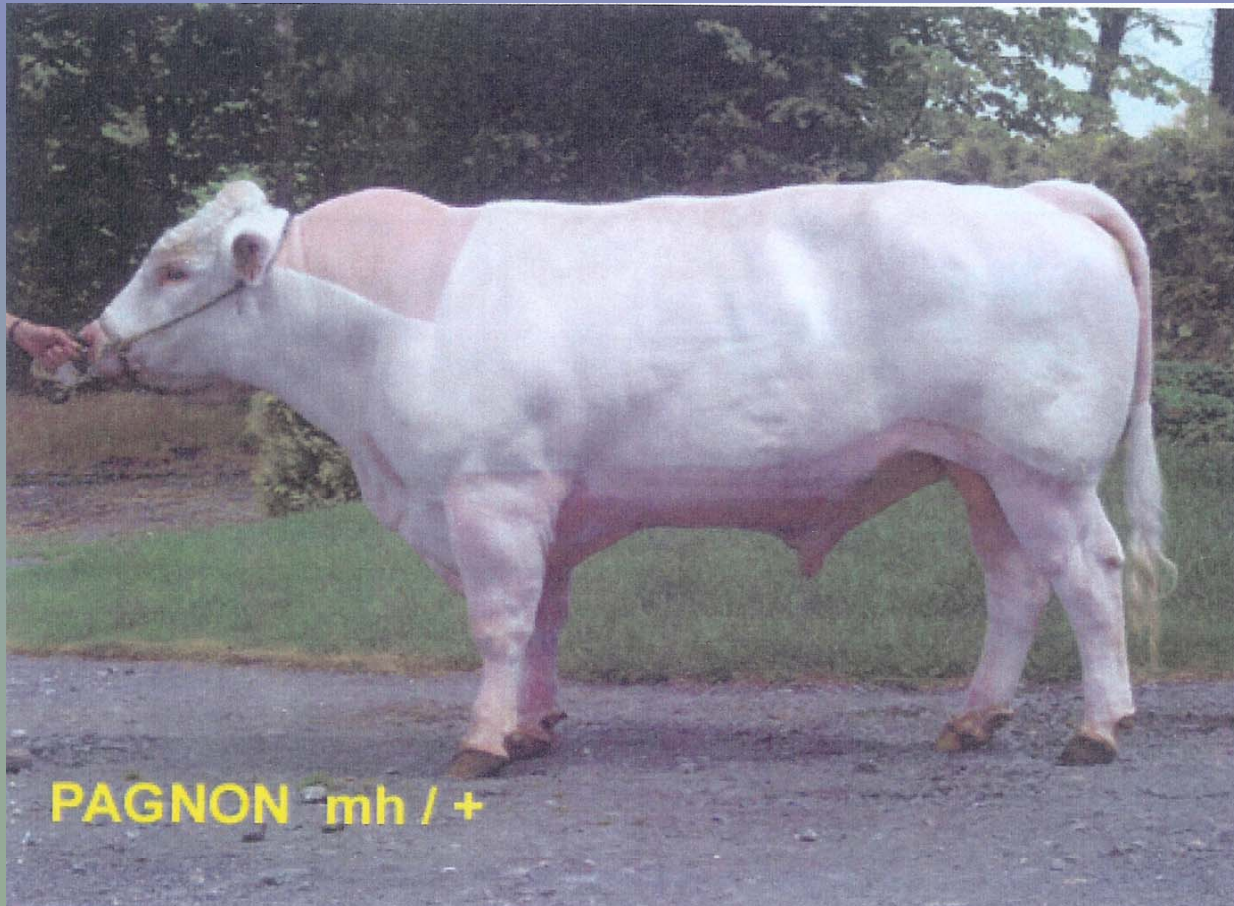
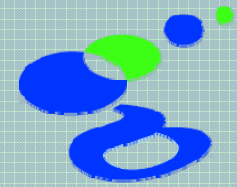


- Local breed in Belgium and Northern France
- Strong muscling and average 4000 kg milk
 - best cows over 6000 kg milk
- Show a myostatin (*mh*) mutation
- *mh* mutation frequent in Walloon dp-BBB
 - most recent females genotyped for *mh*
 - males are required to be genotyped for *mh*
- Leads to a special approach in the field to use (or better misuse) molecular information
 - three “types” of genotypes
 - ⇒ illustration of breed “types”

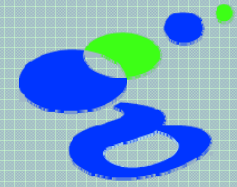
Three dp-BBB Types: +/-



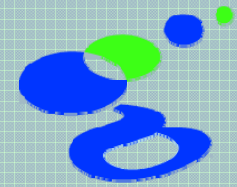
Three dp-BBB Types: mh/+



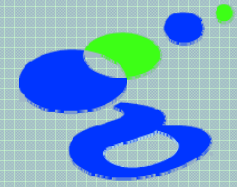
Three dp-BBB Types: mh/mh



Context of This Study: Dual Purpose Belgian Blue Breed



- Three types and their perception:
 - $mh/mh \Rightarrow$ beef
 - $+/+ \Rightarrow$ milk
 - $mh/+ \Rightarrow$ intermediate
- Large knowledge of genotypes
 - empirical use of *mh* locus for selection
 - makes dp-BBB rather exceptional compared to other cattle breeds
- Need for practical method to integrate molecular information
 - mixed inheritance model and beyond



Mixed Inheritance Model

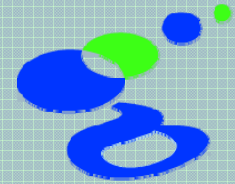
- Generic mixed inheritance model combining fixed gene effects \mathbf{g} and random polygenic \mathbf{u} effects:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{ZQg} + \mathbf{Zu} + \mathbf{e}$$

- Usual assumptions concerning distribution of random effects:

$$E \begin{bmatrix} \mathbf{u} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \text{ and } \text{Var} \begin{bmatrix} \mathbf{u} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix}$$

Equivalent Mixed Inheritance Model



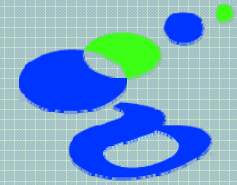
- Equivalent mixed inheritance model replacing fixed gene effects \mathbf{g} and random polygenic \mathbf{u} effects by a combined genetic effect \mathbf{a} :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{e} \quad \text{where } \mathbf{a} = \mathbf{Q}\mathbf{g} + \mathbf{u}$$

- Modified assumptions:

$$E \begin{bmatrix} \mathbf{a} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}\mathbf{g} \\ \mathbf{0} \end{bmatrix} \quad \text{and} \quad \text{Var} \begin{bmatrix} \mathbf{a} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix}$$

Associated Mixed Model Equations

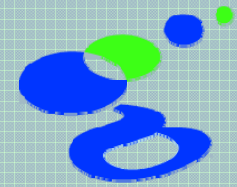


- Using a derivation similar to that of genetic groups (Quaas, J. Dairy Sci. 1988 71: 1338-1345)

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{0} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & -\mathbf{G}^{-1}\mathbf{Q} \\ \mathbf{0} & -\mathbf{Q}'\mathbf{G}^{-1} & \mathbf{Q}'\mathbf{G}^{-1}\mathbf{Q} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}\mathbf{y} \\ \mathbf{0} \end{bmatrix}$$

- Please note joint estimation of \mathbf{a} and \mathbf{g} !

Alternative Mixed Model Equations

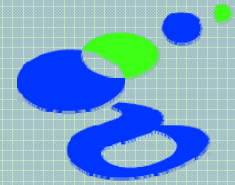


- Solving of whole system is equivalent of solving iteratively the following two systems of equations:

$$[\mathbf{Q}'\mathbf{G}^{-1}\mathbf{Q}]\hat{\mathbf{g}} = [\mathbf{Q}'\mathbf{G}^{-1}\hat{\mathbf{a}}] \quad (1)$$

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} + \mathbf{G}^{-1}\mathbf{Q}\hat{\mathbf{g}} \end{bmatrix} \quad (2)$$

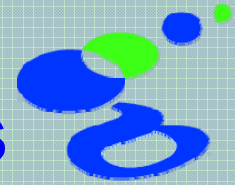
Interpretation of Alternative Equations



$$[\mathbf{Q}'\mathbf{G}^{-1}\mathbf{Q}]\hat{\mathbf{g}} = [\mathbf{Q}'\mathbf{G}^{-1}\hat{\mathbf{a}}] \quad (1)$$

- System of equations (1) extracts gene effects \mathbf{g} from current estimated combined genetic effects \mathbf{a}

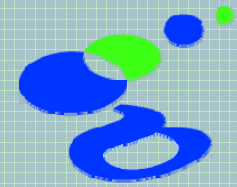
Interpretation of Alternative Equations



$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} + \mathbf{G}^{-1}\mathbf{Q}\hat{\mathbf{g}} \end{bmatrix} \quad (2)$$

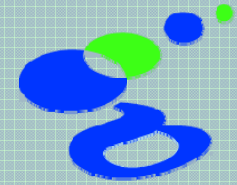
- System of equations (2) estimates combined genetic effects \mathbf{a} and fixed effects taking into account current estimates of gene effects \mathbf{g}

Not All Genotypes Known



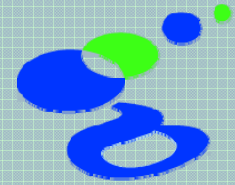
- Dimensions of \mathbf{Q} : $n \times m$
 - where n = animals and m = gene effects
- However dimensions of $\mathbf{Q}'\mathbf{G}^{-1}\mathbf{Q}$: $m \times m$
- \mathbf{Q} can contain « zero » lines
 - allows solving of (1) when only limited number of animals genotyped
 - only expectations for non genotyped animals change
 - gene effects can be estimated from limited known genotypes

Importance of \mathbf{G}^{-1}



- \mathbf{G} covariance structure among combined genetic effects \mathbf{a} (not $\hat{\mathbf{a}}$), weights them
- Solving (1) large analogy to solving a Generalized Linear Model
 - which splits \mathbf{a} into \mathbf{g} and random residual ε
 - model $\mathbf{a} = \mathbf{g} + \varepsilon$ with $\text{Var}(\mathbf{a}) = \mathbf{G}$
- Analogy can be used for simpler solving using equivalent BLUP

Random g Effects ?

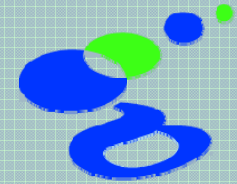


- Situations where estimating \mathbf{g} as fixed clearly sub-optimal (e.g., dense SNP panels)
- Modification of equations (1) to do random estimation after splitting variances $\mathbf{G} = \mathbf{Q}\mathbf{G}_g\mathbf{Q}' + \mathbf{G}_u$

$$\left[\mathbf{Q}'\mathbf{G}_u^{-1}\mathbf{Q} + \mathbf{G}_g^{-1} \right] \hat{\mathbf{g}} = \left[\mathbf{Q}'\mathbf{G}_u^{-1}\hat{\mathbf{a}} \right]$$

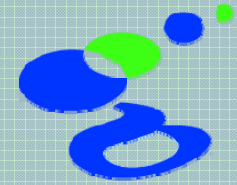
- Large similarity to genomic prediction equations
 - Most important difference use of current estimates of \mathbf{a} , also in an iterative manner
- Other modifications for (2)

Context: *mh* locus use in dp-BBB



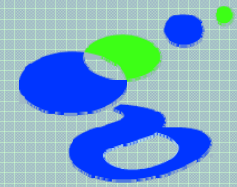
- Data used of official January 2009 genetic evaluations in the Walloon Region of Belgium
- Pedigree : 1,606,024 animals
- Data : 11,117,505 Test Day records (2009)
(10,019,460 TD before 2003)
- 689,057 cows with production records

Material and Methods



- *mh* genotypes available (and known by the breeders) for:
 - 123 dp-BBB bulls
 - 1,940 dp-BBB cows with production records
- Offspring of genotyped animals
 - 12,211 cows with production records
- Four evaluation runs using official Walloon random regression test-day model:
 - with data until 2003 or 2009
 - with (equivalent model “*mh*” suffix) or without (original model) modification for *mh* genotype in the model

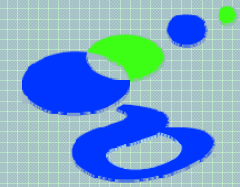
Allele Substitution Effect of mh Estimated with Equivalent Model



Trait	Based on data recorded before	
	2009 (11,117,505 TD)	2003 (10,019,460 TD)
Milk yield (kg/305 d of lact)	-158.7	-149.9
Fat yield (kg/305 d of lact)	-8.93	-8.23
Protein (kg/305 d of lact)	-5.64	-5.16

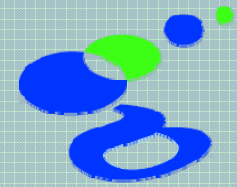
- Results in line with sparse literature

Correlations Between EBV for Milk Yield (MY): All Animals



Category	N	2009mh vs. 2009	2009mh vs. 2003mh	2009 vs. 2003
All evaluated	1,606,074	0.999	0.996	0.996
dp-BBB	18,433	0.992	0.992	0.992
Genotyped dp-BBB	2,001	0.998	0.985	0.984
Offspring of genotyped dp-BBB	12,211	0.999	0.993	0.993

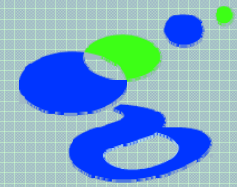
Correlations Between EBV for MY: Cows with No Test-Day Records Before 2003



Category	N	2009mh vs. 2009	2009mh vs. 2003mh	2009 vs. 2003
All evaluated	58,695	0.997	0.872	0.865
dp-BBB	1,873	0.996	0.917	0.909
Genotyped dp-BBB	427	0.996	0.864	0.843
Offspring of genotyped dp-BBB	743	0.995	0.895	0.879

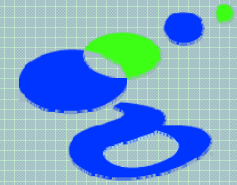
Coefficients of Regression Between EBV for MY:

Cows with No Test-Day Records Before 2003



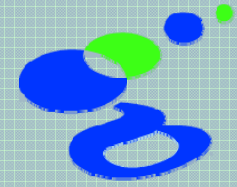
Category	N	2003mh	2003
		vs.	vs.
		2009mh	2009
All evaluated	58,695	0.785	0.767
dp-BBB	1,873	0.851	0.822
Genotyped dp-BBB	427	0.766	0.706
Offspring of genotyped dp-BBB	743	0.817	0.763

Conclusion



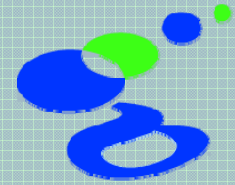
- Developed equivalent mixed inheritance model
 - straight forward derivation
 - no need that all animals being genotyped:
 - here only few thousand
- Estimate allele substitution effects with alternative model
 - results as expected (rather sparse literature)
- For all (great majority non genotyped) animals
 - minimal changes compared to original model
- For animals without records before 2003
 - rather large improvement in ability to predict realized EBV (2009) from PA (2003):
 - improved correlations
 - regression coefficients closer to 1

Implications



- Underlying mixed model can come from large range of models (any mixed model), e.g.:
 - random regression model (as used here)
 - different types of multi-trait models including MACE
- Basic hypothesis:
 - change of expectations
 - can be adapted for Genomic Prediction (random **g**)
- Estimating simultaneously SNP effects and combined genetic effects using this approach:
 - theoretically promising
 - obviously still large challenges ahead

Acknowledgements



- National Fund for Scientific Research
- Walloon Breeding Association (AWE)
- Walloon Regional Ministry of Agriculture
- You for your attention !