Accuracies of different types of MAS-EBV in the French MAS Program

F.Guillaume, J. Tarres, D. Boichard, T. Druet, and S. Fritz

ALIMENTATION
AGRICULTURE
ENVIRONNEMENT

INTRODUCTION

- Dense maps provide
  - Better estimates of QTL locations
  - Better QTL transmission probabilities (*ie* close to 1 or 0) in a pedigree
  - LD information, *ie* across family (or across founder) PID information
- Therefore, they should improve MAS efficiency

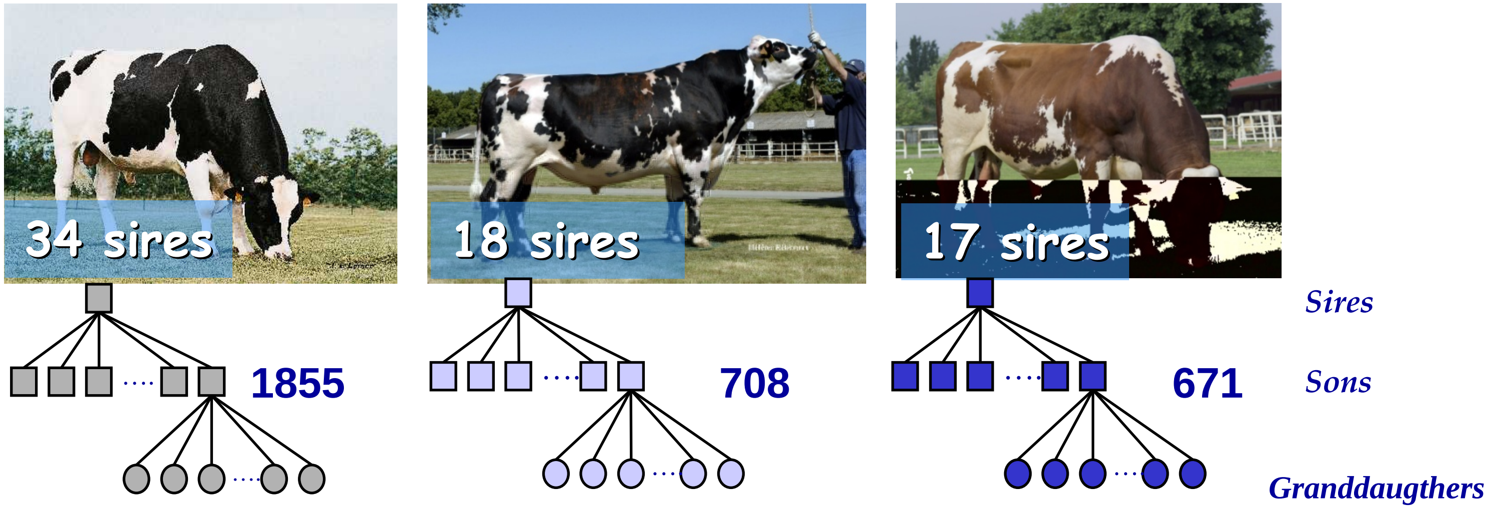
➔ Practical results from the French MAS experience

➔ MAS moving from low to dense marker information

Description of MAS in France

- Current MAS program (2001-2008)
  - Microsatellites markers – low density map :
    - 2-5 markers by QTL (45 markers for 14 QTLs)
    - 5-20 cM between markers
  - Linkage Equilibrium – within family estimation
- Situation in 2008
  - ~70 000 genotyped animal (Hol)
  - Confirmation of the QTL used in the evaluation

Evolution : QTL fine-mapping resource population



- ~3200 sires genotyped with the 54k SNP Illumina chip
- 15 traits
- Genotypes available since May

➔ Confirm and Fine Map QTL
➔ Identify Haplotypes in LD with QTL

Evolution of MAS-EBV

- From :

  a Fernando-Grossman model using 2 to 5 QTL per trait.

  QTL marked with few microsatellites markers
- To :

  an haplotype-based model using many QTL
  QTL followed with the 54k SNP chip

How large is the gain ?

Validation sample

- 468 Holstein sires receiving their first official proof in June 2008
- Information available in 2004:
  - phenotypes from relatives
  - microsatellite genotypes
- SNP information obtained in 2008 for this batch of males as well as for a population of older tested sires

➔ Comparison of DYD in 2008 with early prediction in 2004 based on microsatellites or SNP information.

Tested Models

Trait : milk yield ($h^2 = 0.30$)

- Fernando-Grossman model (*4 QTL*)

$$y_i = \mu + u_i + \sum_{j=1}^{n} (v_{ij}^p + v_{ij}^m) + e_i \quad \text{with} \quad v_i^p = p_i^p v_s^p + (1 - p_i^p) v_s^m + \varepsilon_i^p$$

- Haplotype based model (*4 QTL or 24 QTL*)

$$y_i = \mu + u_i + x_i' h + e_i$$

- Regression on SNPs (*23 SNP within the 4 QTL region*)

$$y_i = \mu + u_i + x_i' \beta + e_i$$

Results

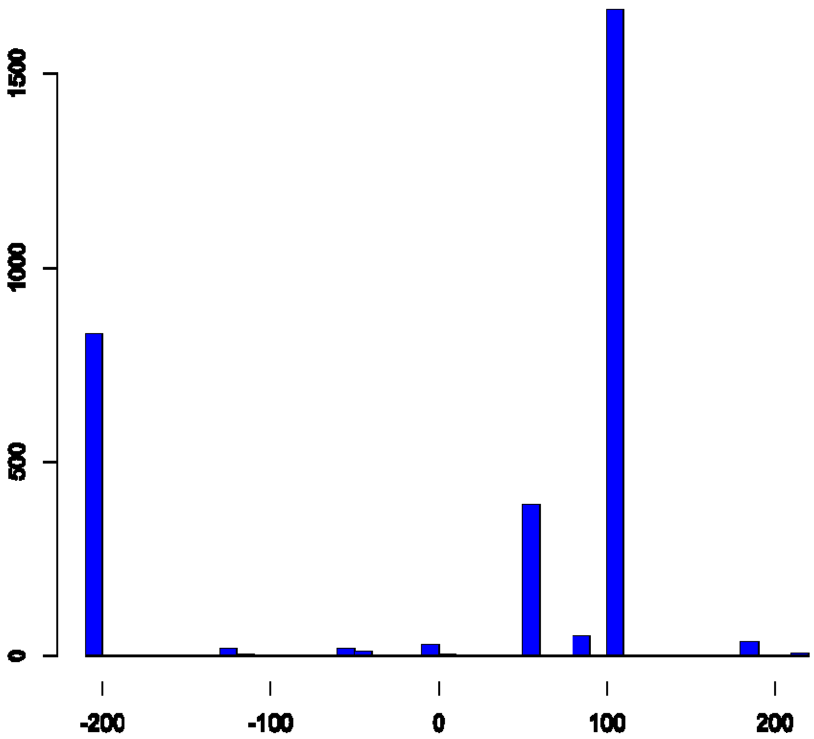
Correlation between milk yield DYD in 2008 and different types of MAS-EBV (2004 phenotypic situation)

N=468

Type	Correlation
Polygenic	0.406
FG-MAS	0.408
Haplotypes (4 QTL)	0.509
SNP regression (4 QTL)	0.462
Haplotypes (24 QTL)	0.560

Results

Distribution of estimates of haplotypes effects around DGAT1



- Haplotypes of 4 SNP
- Few haplotypes encountered in population
- 2 main haplotypes whose effects seems to correspond to DGAT1 mutation

Discussion : Fine Mapping results

- Work still in progress...
- Partial conclusion :
  - QTL previously used in the MAS program have been confirmed
  - QTL genetic variances previously overestimated
  - Many new QTL have been found
  - Correlations goes up to 0.56 with all already fine-mapped QTL (still incomplete)

Discussion

- Note that upper bound correlation lower than 1 (around 0.85) due to incomplete accuracy of DYD
- Correlations in agreement with the proportion of genetic variance explained by the QTLs
- Need to confirm validity and efficiency of the haplotype MAS approach with low heritability traits
- Work in progress to compare efficiency of this haplotype strategy with that of genomic selection

Conclusions

- Further works need to be done but results are promising
- Haplotype-based model provides a satisfactory EBV estimation at a low computation cost
- Haplotypes probably extract more information than single SNP
- Comparison with genomic selection is needed

Acknowledgements

This study is a collaboration between :

- Institut de l'élevage
- INRA
- UNCEIA
- Labogena

LABOGENA