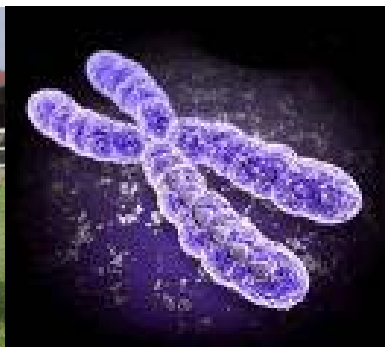


Description of the French genomic Marker Assisted by Selection program

P. Croiseau, C. Hoze, S. Fritz, F. Guillaume,
C. Colombani, A. Legarra, A. Baur, C. Robert-Granié,
D. Boichard and V. Ducrocq



Use of the MA-BLUP in France

- A strong background with MA-BLUP/QTL-BLUP for national genomic evaluations
- This strategy was tested in Holstein and provided better results than others genomic selection approaches

Correlation between DYD _{obs} and GEBV	Milk	Protein	Fat	Protein %	Fat %	Conception rate
pedigree-based BLUP	0.38	0.44	0.40	0.47	0.44	0.29
GBLUP	0.56	0.55	0.59	0.73	0.72	0.35
PLS	0.53	0.55	0.58	0.71	0.70	0.33
Elastic-Net	0.57	0.57	0.63	0.75	0.80	0.34
French BLUP-QTL	0.60	0.57	0.7	0.73	0.81	0.39

**WCGALP, Leipzig 2010*



QTL-BLUP

(Fernando and Grossman, 1989)

$$y_i = \sum_{j=1}^{QTL} (h_{ij}^{sire} + h_{ij}^{dam}) + u_i + e_i$$

- Y_i → performances for individual i
- h_{ij} → gametic effect from sire and dam for the QTL j
- u_i → polygenic effect for individual i
- e_i → residual effect for individual i

Particularities of the QTL-BLUP

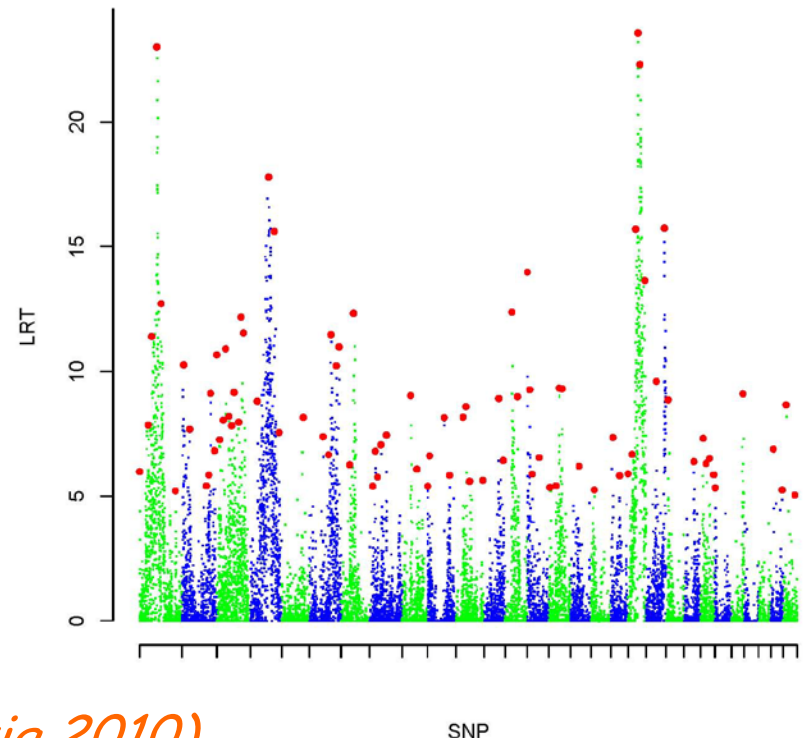
- Haplotypes are taken into account
 - LD between QTL and markers is increased using haplotypes instead of SNP
 - Since the causal mutations are probably rarely genotyped, the use of haplotypes should improve prediction equations
- Since the same list of QTL is used for successive national genomic evaluations
 - GEBV stability over time
- Since a restricted number of QTL is included in the model, such an evaluation is :
 - less computationally demanding
 - adapted to higher density chips once QTL are chosen

Which list of haplotypes for the MA-BLUP?

1. LRT peaks from a QTL detection (LDLA, Meuwissen and Goddard)

2 criteria to define a LRT peak:

- ✓ Have the highest LRT value in a window of SNP (1 or 2 cM)
- ✓ Have a LRT value higher than a threshold (3 or 5)



(Croiseau et al., Leibzig 2010)

Which list of haplotypes for the QTL-BLUP?

1. LRT peaks from a QTL detection (LDLA, Meuwissen and Goddard)
2. Use of a genomic selection approach based on a variable selection method
 - Elastic-Net (EN); sparse Partial Least Squares (sPLS)
 - EN is a linear combination of Ridge Regression and LASSO
 - sPLS* is the variable selection version of the PLS (citation)
 - From the set of SNP obtained using EN or sPLS, the SNP which are in the same cM were grouped in haplotypes

(Colombani et al., WCGALP 2011)*



DATA

- This strategy worked well in Holstein
- In this study we wanted to test it in a smaller reference population (Montbéliarde)
 - 1392 animals in Montbéliarde breed
 - Training : 1170 individuals
 - Validation: 222 individuals
 - 5 traits
 - Rear udder width
 - height at sacrum
 - Somatic Cell Counts
 - milk yield
 - protein yield
 - Performances were DYD (Daughter Yield Deviation)

Analysis

- Weighted correlation between observed DYD and DGV are calculated
 - ✓ Weight= Equivalent Daughter Contribution (EDC)
- QTL-BLUP were performed using QTL lists defined by
 - LDLA
 - Elastic-Net
 - sPLS
- Comparison with pedigree-based BLUP, GBLUP, EN and sPLS
- For LDLA, results for the best definition of LRT peaks are shown
- For EN and sPLS, results of the best combination of parameters are shown

Correlation between DYD_{obs} and DGV

	pedigree-based BLUP	GBLUP	EN	sPLS	QTL-BLUP		
					LDLA	EN	sPLS
somatic cell count	0.50	0.59	0.57	0.47	0.57	0.55	0.49
rear udder width	0.39	0.55	0.54	0.48	0.50	0.50	0.46
milk	0.27	0.41	0.44	0.41	0.45	0.46	0.43
protein yield	0.28	0.43	0.47	0.42	0.47	0.49	0.39
height at sacrum	0.41	0.54	0.54	0.49	0.53	0.55	0.45
<i>mean over the 5 traits</i>	<i>0.37</i>	<i>0.50</i>	<i>0.51</i>	<i>0.45</i>	<i>0.50</i>	<i>0.51</i>	<i>0.44</i>

French QTL-BLUP

- In the QTL-BLUP based on LDLA, QTL variances were estimated
- In the QTL-BLUP based on EN list, all the QTL have the same variance
 - ✓ Total QTL variance was 60% in all cases (40% for the polygenic component)
 - ✓ Proportion selected to improve the slope of regression
- In the French QTL-BLUP
 - ✓ QTL come from the combine LDLA + EN list

Correlation between DYD_{obs} and GEBV



	pedigree-based BLUP	GBLUP	EN	sPLS	BLUP-QTL			French QTL-BLUP
					LDLA	EN	sPLS	
somatic cell count	0.50	0.59	0.57	0.47	0.57	0.55	0.49	0.56
rear udder width	0.39	0.55	0.54	0.48	0.50	0.50	0.46	0.53
milk	0.27	0.41	0.44	0.41	0.45	0.46	0.43	0.42
protein yield	0.28	0.43	0.47	0.42	0.47	0.49	0.39	0.45
height at sacrum	0.41	0.54	0.54	0.49	0.53	0.55	0.45	0.53
<i>mean over the 5 traits</i>	<i>0.37</i>	<i>0.50</i>	<i>0.51</i>	<i>0.45</i>	<i>0.50</i>	<i>0.51</i>	<i>0.44</i>	<i>0.50</i>

Number of SNP/QTL used in the prediction equation

	pedigree-based BLUP	GBLUP	EN	BLUP-QTL				French BLUP-QTL	
			SNP number	LDLA		EN		SNP	QTL
				SNP	QTL	SNP	QTL		
somatic cell count	-	38490	13687	2350	470	940	312	1568	392
rear udder width	-	38490	19957	2750	550	652	400	1920	480
milk	-	38490	25713	3025	605	1007	628	2832	708
protein yield	-	38490	14879	3145	629	1211	479	2236	559
height at sacrum	-	38490	22703	2065	413	784	344	1696	424
<i>mean over the 5 traits</i>	-	38490	19388	2667	533	919	433	2050	582

Conclusion

- Optimal correlations were obtained using
 - ✓ EN
 - ✓ QTL-BLUP with a list of QTL defined using EN
- French QTL-BLUP brings more benefits for Holstein than for Montbéliarde
 - ✓ Maybe due to the reference population size which is smaller
 - haplotype effects estimation is more difficult for rare variants
 - We need to work on haplotype clustering to avoid this problem
 - ✓ Results should be improved with HD chip
- French QTL-BLUP maintained the stability of GEBV
 - ✓ When the same list of QTL is used over successive genomic evaluations



Acknowledgment

LABOGENA for genotyping

ANR and APISGEN for funding AMASGEN program