Sequence-based association study of resistance to paratuberculosis in Holstein and Normande cattle

Sanchez M.P., Guatteo R., Davergne A., Grohs C., Taussat S., Saout J, Blanquefort P., Delafosse A., Joly A., Fourichon C., Boichard D.

UMR BioEpAR UMR GABI UMR ISP

Context

Paratuberculosis generates large economic losses
and affects animal welfare
Early infection but late expression after a latency period

The pathogen is Mycobacterium avium ssp paratuberculosis
(or **MAP**)

No treatment, no authorized vaccine,
limited efficiency of prophylaxis

=> A better **genetic resistance**
would contribute to control the disease

Objectives

Genomic study of confirmed host phenotypes to identify genomic regions (**QTL**) affecting susceptibility to MAP

in two dairy cattle breeds

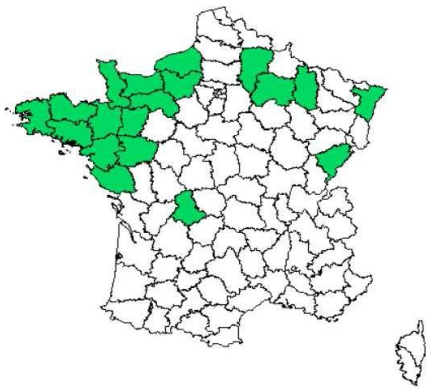
Holstein

Normande

➔ Better understand the **genetic determinism** of the susceptibility to MAP

➔ Build a first reference population to pave the way for **genomic selection** against the disease

Material & Methods



Herds with detected cases enrolled in surveillance program

regular blood **ELISA** + fecal **PCR** tests

Selection of cows in these herds:

Confirmed Clinical cases

Or

Subclinical cases And Non infected controls

1) With 2 ELISA & 2 PCR tests distant from > 8 months
2) With **clear** and **concordant** negative or positive results, excluding intermediate ones
3) Negative cows required to be at least 60 months old, negative at all tests, and **born in the same herd and in the same month** as affected cows

Material & methods: accurate phenotypes

Selection of cows in these herds

Cows with uncertain phenotypes removed

Additional analyses to confirm phenotypes
1 ELISA & 1 PCR test
(same kit, same lab)

Cows with discordant phenotypes removed

50k genotypes

Quality control, Crossbreds, parentage incompatibility

	Non infected	Subclinical	Clinical	Total
Holstein	**838**	**577**	**229**	**1644**
Normande	**233**	**347**	**69**	**649**

Material & Methods

4 traits :

- **0 = non infected vs 1 = clinical and subclinical**
- **0 = non infected, 1 = subclinical, 2 = clinical**
- **PCR CT**
- **Elisa S/P score**

Material & Methods

Genome Wide Association Study

*** with GCTA software**
*** within breed**
*** Relationship stucture accounted for through a genomic matrix computed with 630k markers of the HD**

Imputation to whole genome sequence

Bioinformatic inference of whole genome sequence of each cow

? ? A ? ? G ? ? ? C ? ? ? T ? ?
? ? A ? ? C ? ? ? A ? ? ? A ? ?

A C A T T G A C A C A C A T A G
G C A A A C A C G A G G A A A A

Reference Population

50K

Step 1

Within breed
FImpute (Sargolzaei et al., 2014)

546 NOR
776 HOL

HD: 777K

Etape 2

Within breed
Minimac (Howie et al., 2012)

1000BG RUN6
= 2,333 animals
including
44 NOR + 544 HOL

Sequence

=> 649 NOR + 1644 HOL with imputed whole genome sequence
27 million variants (including >8 million imputed with R^2>0.3)

Results: heritability estimates

In this sample, h^2 is high

$h^2 = 50\%$

$h^2 = 57\%$

Higher than previous estimates

Possible reasons :

$\Rightarrow$ Accurate phenotypes

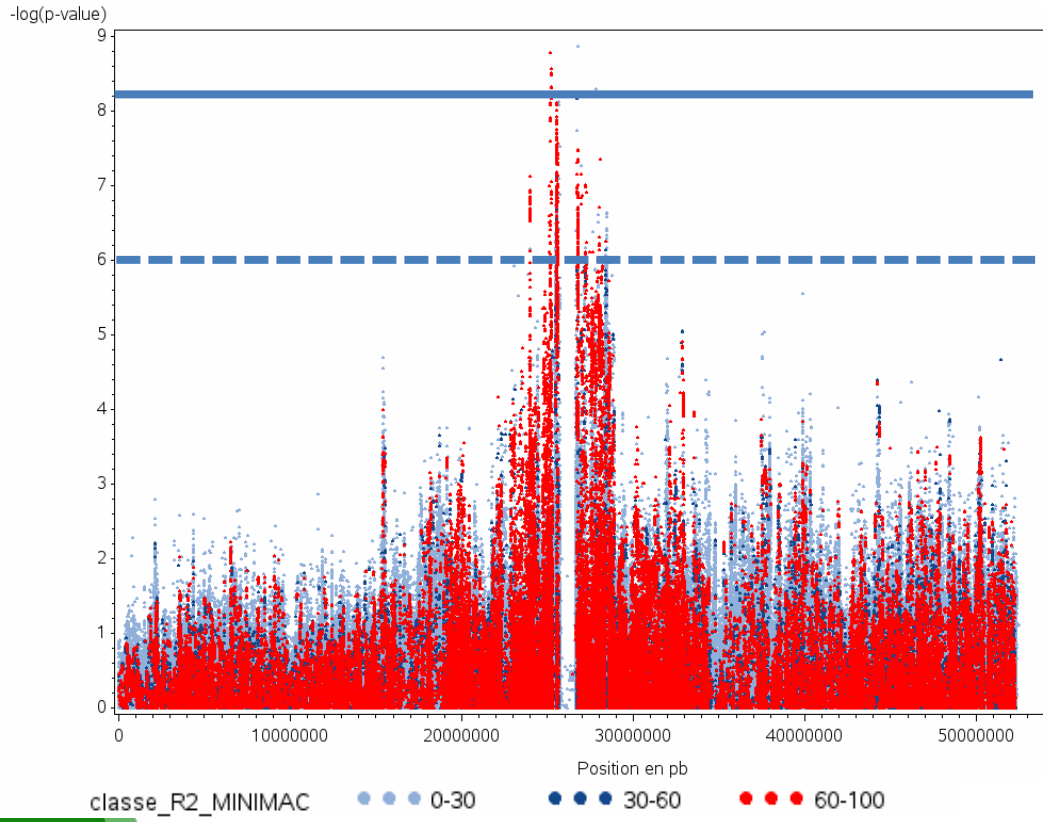
$\Rightarrow$ Effect of phenotype selection

$\Rightarrow$ Balanced design

Nevertheless, **genetic variation**
of susceptibility to MAP appears to be **large**

Chromosome 23

Holstein



10 variants with –logP > 8.2

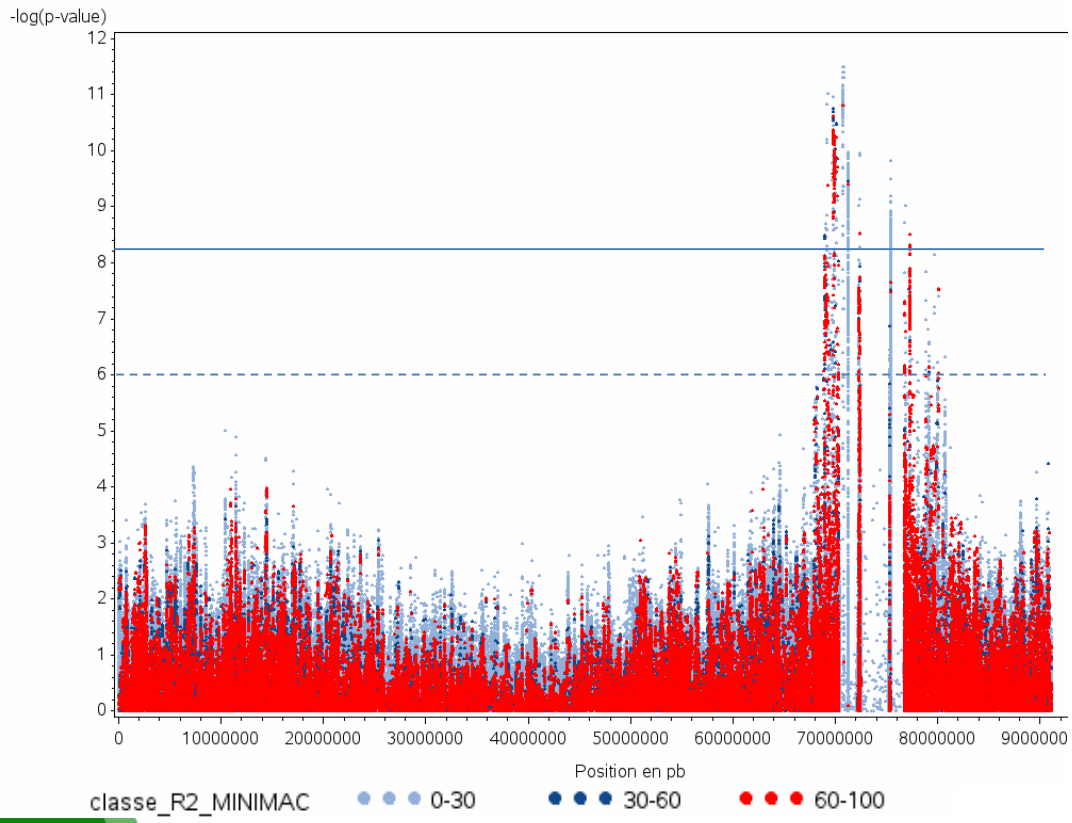
In the **MHC** region

(**ELOVL5** gene, among others**)**

MHC region significant in both breeds

Chromosome 12

Holstein



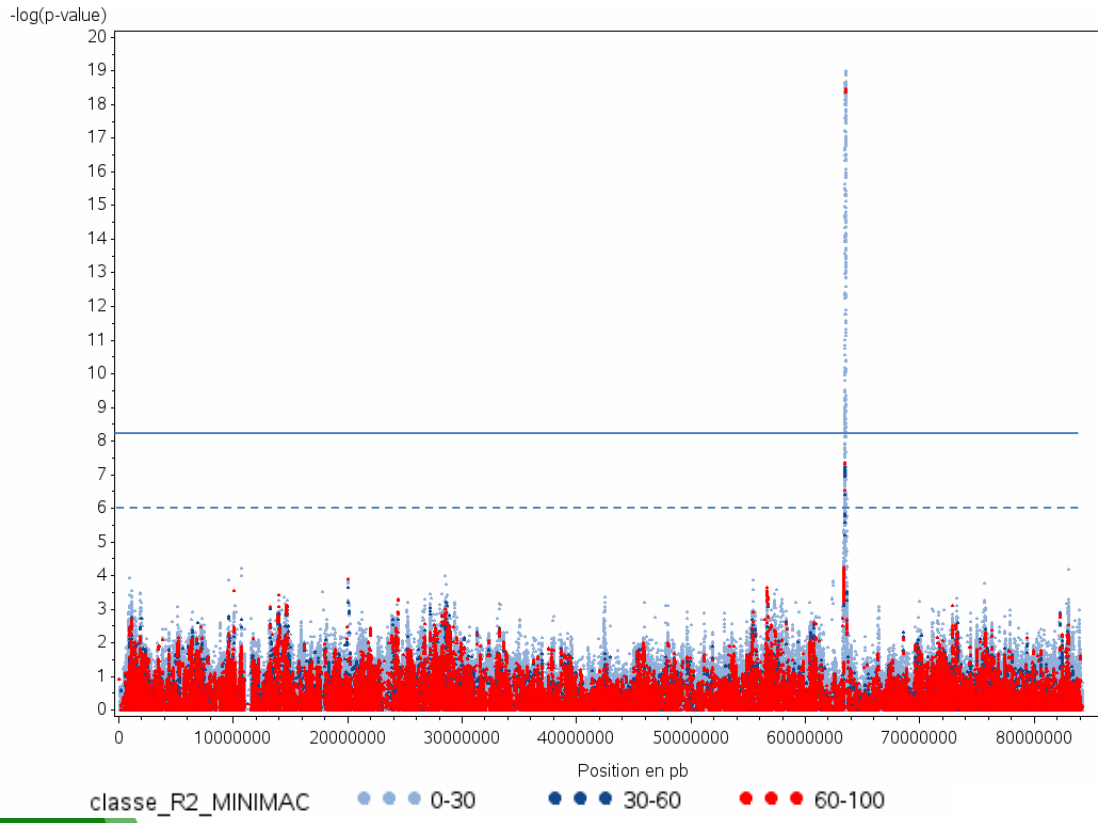
273 variants with –logP > 8.2

One candidate gene : **ABCC4**, already mentioned by Minozzi et al (2012)

<10 candidate variants in intronic regions

Chromosome 13

Holstein



21 variants (with R^2>0.3) with –logP > 8.2

QTL shared across breeds

The two best variants are intergenic in the region of two genes known to be involved in intestine cell morphology in mice

Results with other phenotypes

* 0-1-2

=> Increased significance of results

=> It can be probably concluded that

Clinical cases are more sensitive than subclinical

* CT PCR or Elisa S/P values alone

=> less significant

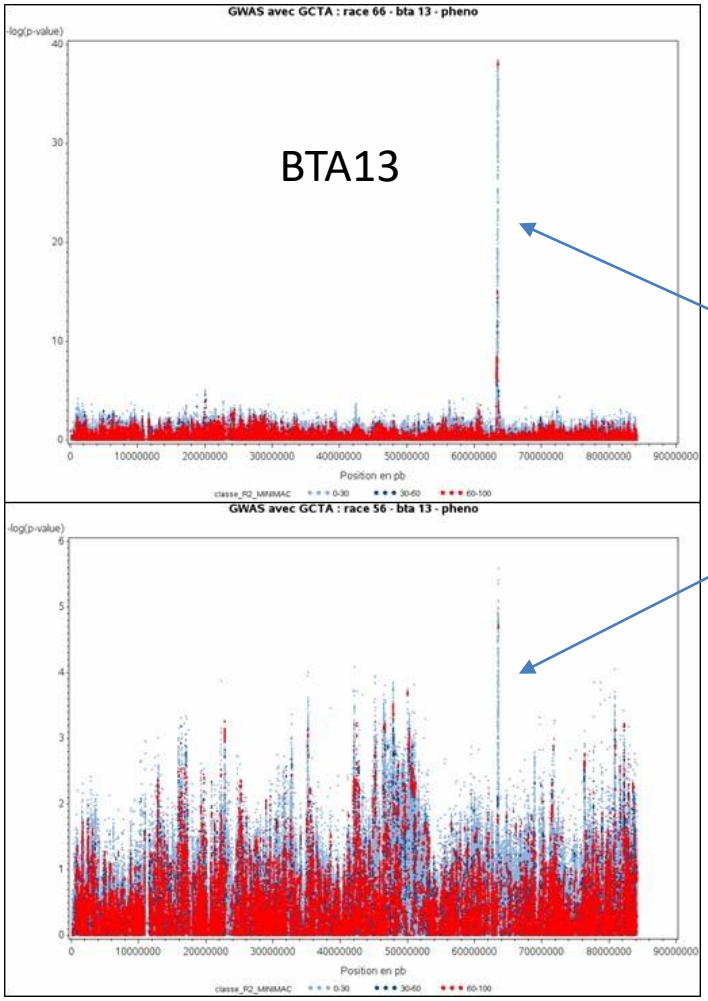
(but less data, no value for clinical cases)

Comparison across breeds

Little overlap across breeds, but limited power in Normande breed
Common QTL on BTA23 and probably BTA13

Holstein

Normande



First conclusions

Strong genetic determinism, therefore good theoretical possibilities of selection

Not monogenic, complex genetic determinism

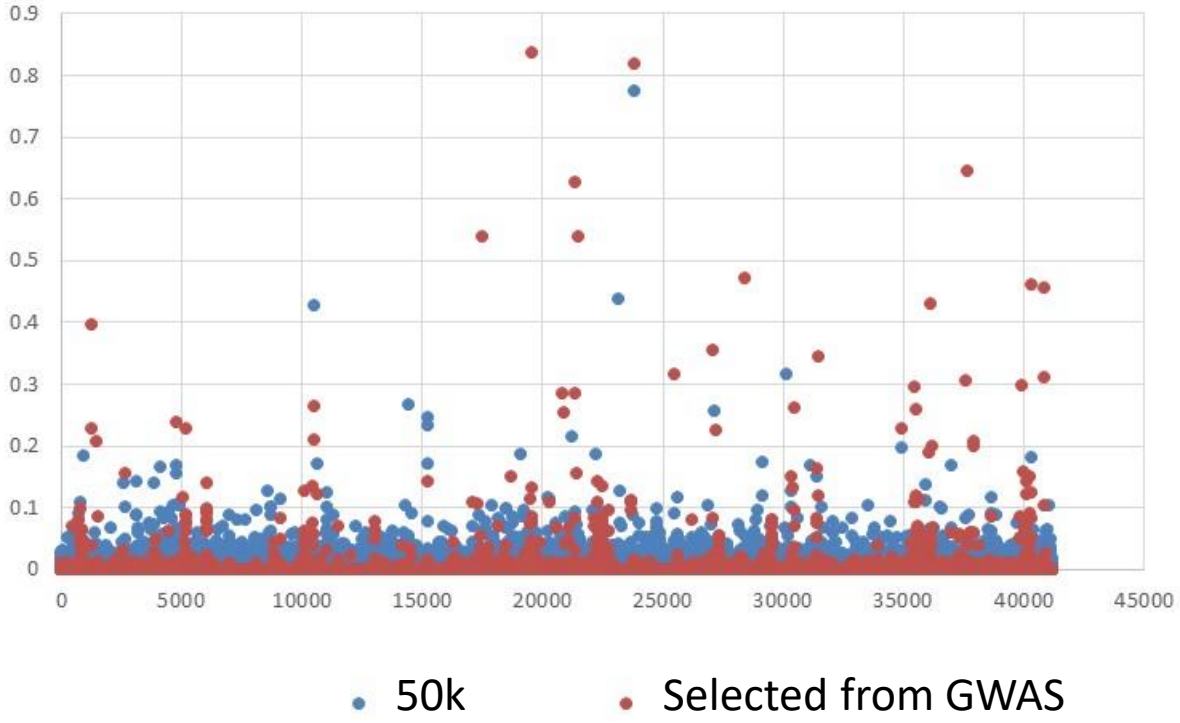
=> Genomic selection

But strong QTL : the method must account for them

Choice of markers and method

- Markers for the 50k chip
- + 1813 variants selected from seq. GWAS
  - with –log(p-value)>4 in at least one breed x trait analysis (8 analyses)
  - Only the most significant variant / 20kb interval
- BayesC to give more importance to predictive variants
- In Holstein only
- Results shown for 0-1 trait

Inclusion probability



Selected variants are very important for prediction !

Accuracy

➢ Correlation between predicted genomic values and phenotypes

Trait	Correlation
Non infected / infected (0/1)	**0.58**
Non infected / subclinical/ Clinical (0/1/2)	**0.57**
Elisa	**0.54**
PCR	**-0.54**

(Holstein, 10-fold cross-validation)

BUT : Optimistic because variants were selected on the same population, an independent sample is required for validation

However, first results are very promising

Perspectives

- Identify / validate the causal variants for 2 QTLs
- Increase the size of the reference population for genomic prediction validation, and then practical implementation
- Extend to other breeds

This study was funded by **INRA** (GISA- **Picsar**),
APIS-GENE and **GDS France (Paradigm)**

Thank you for your attention

UMR BioEpAR UMR GABI UMR ISP