



State of the Art of Genomics for Selection

Didier Boichard

*UMR1313 Animal Genetics and integrative Biology,
F-78350 Jouy en Josas, France*



Whole Genome Sequence of Farm Animals

- After Human (2001) and Mice, WGS of the chicken (2004), the dog (2005), bovine (2006), horse (2007), pig (2009), ...
- Entirely in the public domain
- Sequencing of different individuals => polymorphisms discovery :
 - 3.2 million bovine polymorphisms in dbSNP
 - probably >10 million known today
- New technologies for genotyping and sequencing

SNP : Single Nucleotide Polymorphism

DNA Variation of one base

..GAATCTTATGCTATACATAATTATATACTAAT**C**GGGTATTGTTCTTAT..

..GAATCTTATGCTATACATAATTATATACTAAT**A**GGGTATTGTTCTTAT..

↑
SNP

Genotyping chips

- Miniaturized device for the simultaneous genotyping of many SNP
- From few dozens up to several million SNP
- Two main technology providers, Illumina and Affymetrix
- Illumina products in cattle
 - 3000 (6500=LD), **54 000=50k**, 777 000=HD

FIGURE 1: BOVINESNP50 BEADCHIP



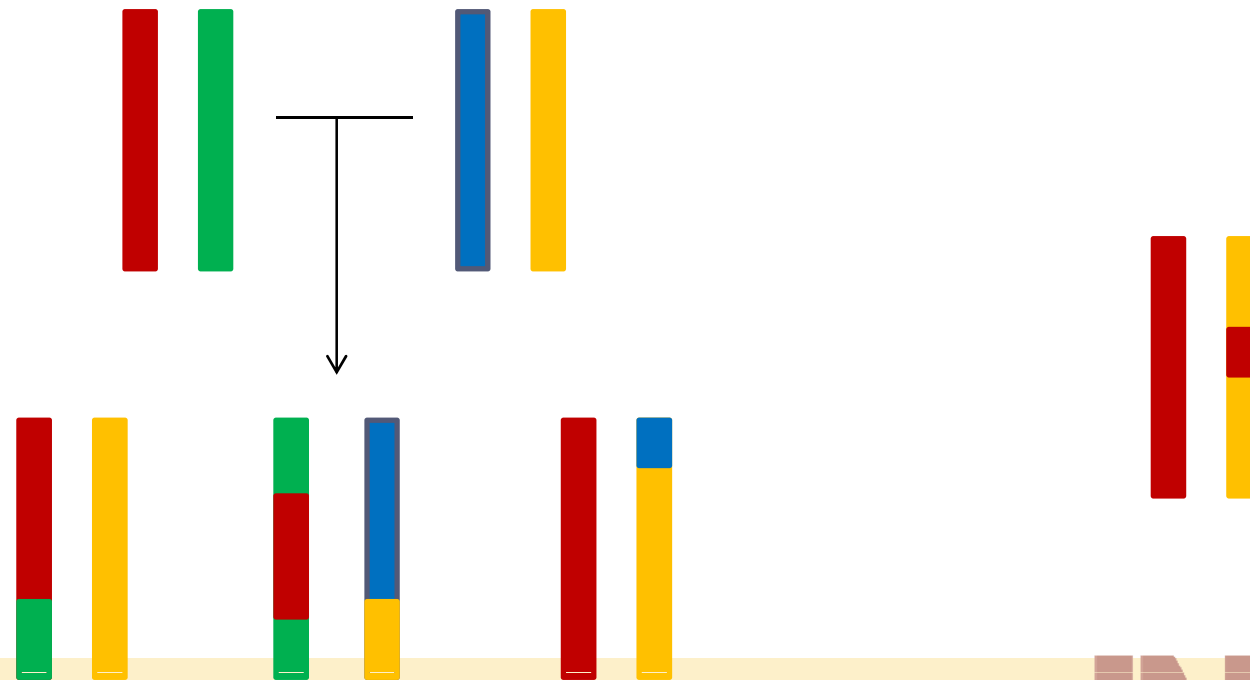
The BovineSNP50 BeadChip features more than 54,000 evenly-spaced SNPs across the entire bovine genome.

Exemples of information provided by markers

1) Trace transmissions

2) Measure relationships

3) Measure inbreeding



FOOD AND NUTRITION

AGRICULTURE

ENVIRONMENT

INRA

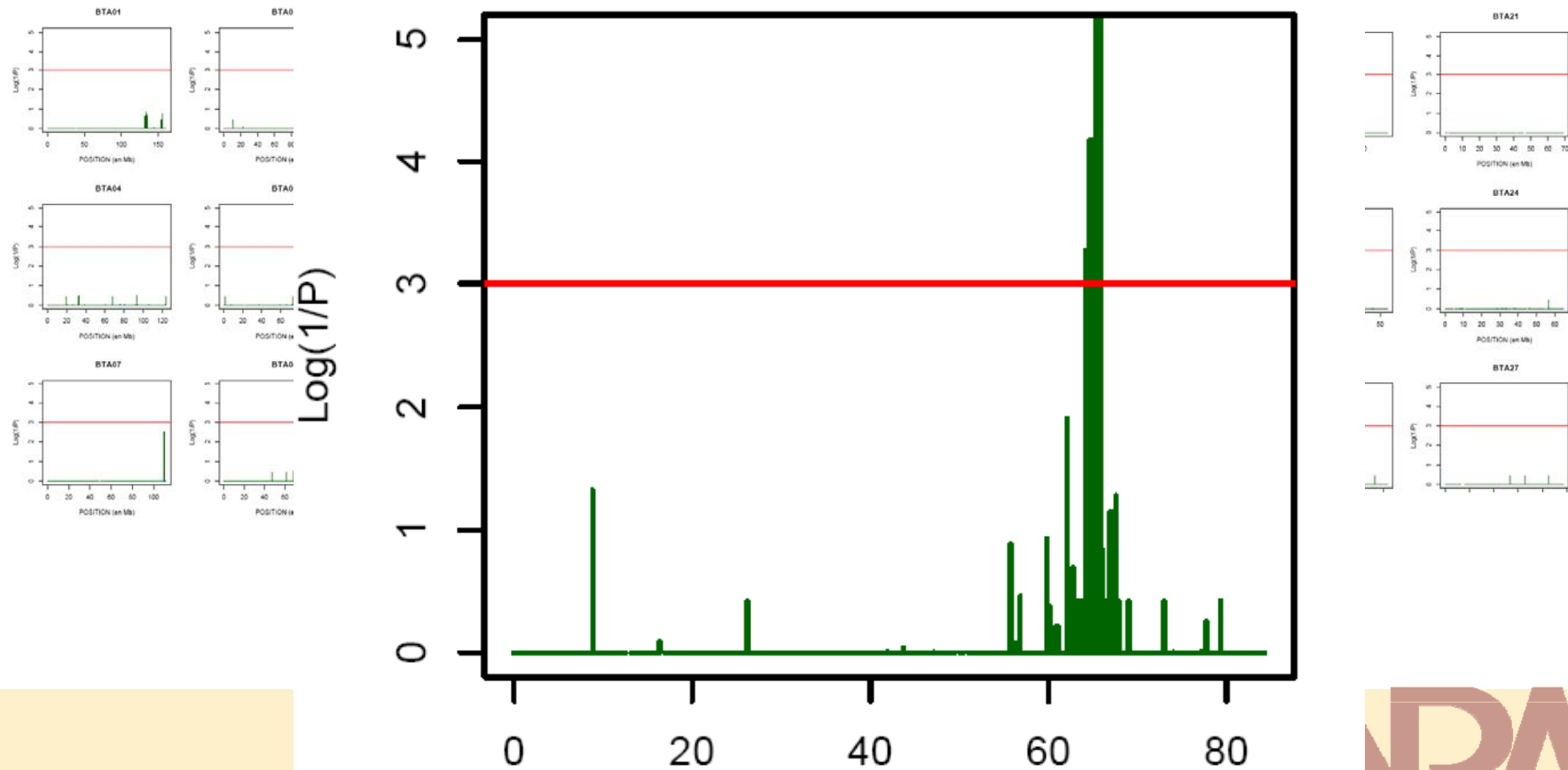
Example of application : mapping recessive defects

- A defect is rare and originates from one unique mutation
- It is recessive, therefore the affected animals carry two copies of the mutation
- Affected animals are also homozygous for the DNA segment surrounding the mutation



Results based on a few affected animals

BTA13



Genomic selection

Selection based on the prediction of breeding values from the information of dense markers covering the whole genome

Copyright © 2001 by the Genetics Society of America

Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps

T. H. E. Meuwissen,* B. J. Hayes[†] and M. E. Goddard^{†,‡}

**Research Institute of Animal Science and Health, 8200 AB Lelystad, The Netherlands, [†]Victorian Institute of Animal Science, Attwood 3049, Victoria, Australia and [‡]Institute of Land and Food Resources, University of Melbourne, Parkville 3052, Victoria, Australia*

Manuscript received August 17, 2000

Accepted for publication January 17, 2001

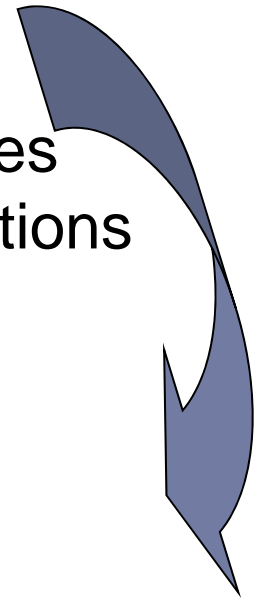
How it works ?

- Reference Population

- Population with both phenotypes and genotypes
- Analysis of the genotype – phenotype associations
- Estimation of marker effects

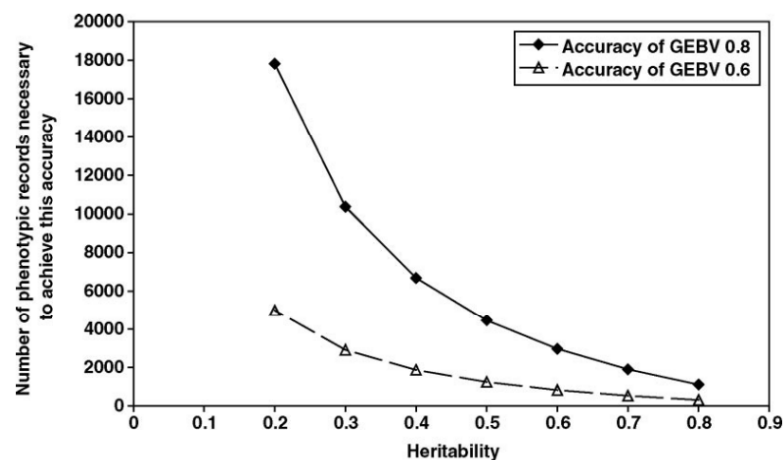
- Population of candidats for selection

- Population with the same associations
- Genotypes
- Prediction of the breeding value by using marker effects estimated in the reference population



Factor of variation of GS efficiency

- Two big factors :
 - Accuracy of SNP effect estimation
 - size of reference population
 - heritability
 - LD between markers and QTL
 - marker density
 - effective size of the population => number of « indépendants » segments
 - Relationship between the candidates and the reference population
- Statistical Methods



Hayes et al, 2009)

A number of methods used

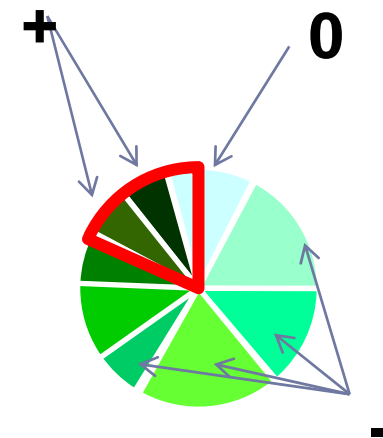
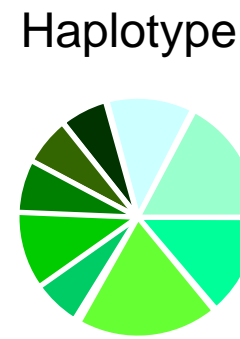
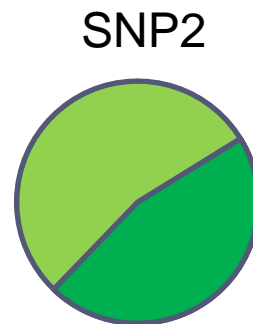
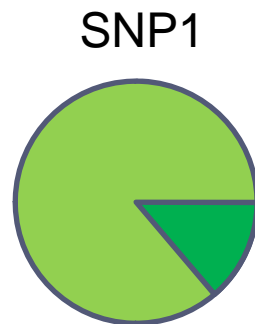
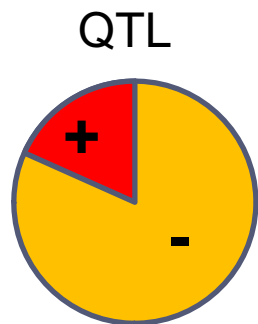
1. **G-BLUP** : in the conventional BLUP, replace the pedigree based relationships by the marker based relationships
2. **Bayesian Methods (Bayes B, C, R ...)** : tries to find the SNP in association with QTL and to give a zero value to most SNP without effect

Comparison of methods

- **Marker Density**
 - low => little differences between GBLUP and Bayesian methods, accuracy low to moderate
 - high => saturation of efficiency of GBLUP, whereas Bayesian approaches increase in accuracy
- **Genetic Determinism**
 - polygenic : some advantage to GBLUP
 - at least partially oligogenic : advantage to Bayesian methods

SNP or Haplotypes ?

- Most work with individual SNP
- Loss of efficiency due to incomplete Linkage Disequilibrium
- Personal point of view : a haplotype with 8-15 alleles is much more informative



SNP or Haplotypes ?

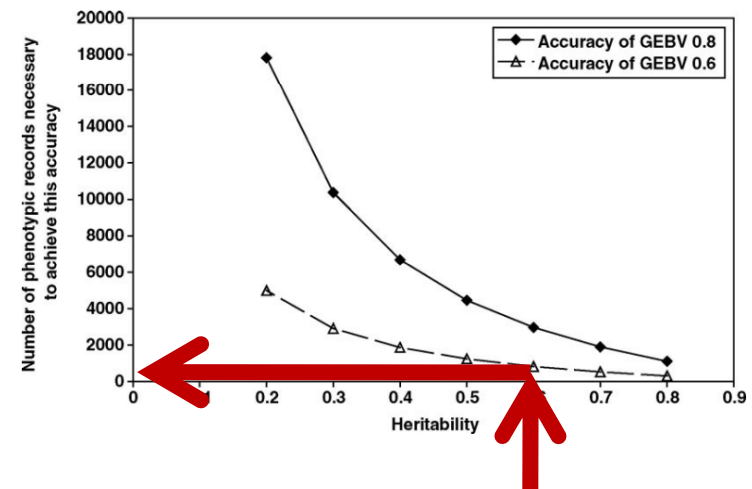
- In France, a method based on haplotypes of 3-6 markers
 - 300-700 regions targeted on the genome, for each trait
 - QTL-BLUP, including a residual polygenic effect

$$y_i = \mu + u_i + \sum_j (h_{ij1} + h_{ij2}) + e_i$$

	Milk	Protein	Fat	Protein content	Fat content	Fertility
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
QTL-BLUP	0.60	0.57	0.66	0.73	0.81	0.39

Reference populations

- Individuals with performances or EBV
- Focus on progeny tested bulls, because of their high reliability
- Probably more females in the future
- Figures in France
 - 18,300 in Holstein
(EuroGenomics Consortium)
 - 1,800 in Montbéliarde
 - 1,300 in Normande



Major consequences of genomic selection

- High reliability ($R^2 = 0.5$ to 0.7)
- At a early age, before any performance of the candidate
- For all traits (depends only on the reference population)
=> More balanced genetic trend

A fantastic opportunity to improve functional traits

⇒ Use of bulls without progeny test

note they will get progeny based EBV, but later
=> Maintaining performance recording is essential !!!

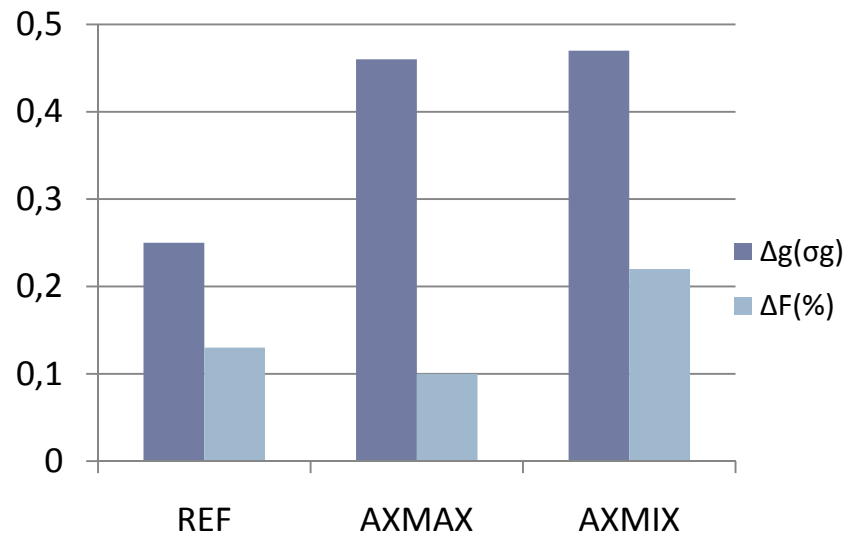
Massive use for females

- Same reliability ($R^2 = 0.5$ to 0.7) for females as for males
- New possibility for within herd selection, for customized breeding objective
 - use of a wide range of males
 - selection of the best cows + increased prolificacy
 - embryo transfer, sexed semen
- A large proportion of genotyped cows if the cost is reasonable
- A low cost implies 1) large volumes, 2) a low chip (3 -> 6k)

Major consequences of genomic selection

- **A nearly doubled potential genetic trend**
 - **due to a reduced generation interval, combined with a good accuracy and an increased selection intensity**
- **A more balanced genetic trend**
 - **due to a rather homogeneous reliability across traits and EBV available for all animals**
 - **due to an increase in weight for functional traits in the breeding objective (no increased selection pressure for production)**
- **Possibly, a lower inbreeding trend, if many young bulls are used**

Changes in breeding practices



(Colleau et al, 2009)

- **REF** : GS for preselection, and progeny test
- **AXMAX** : only young bulls, every young bull also bull sire
- **AXMIX** : 50% AI by young bulls, 50% by older bulls with progeny information

⇒ Stop progeny test

⇒ Don't use bulls when they have progeny information (in competition with their sons and even grandsons...)

The French situation in 2010

669 young bulls marketed in 2010

Breed	Bull category	Number	Doses par bull
Montbéliarde	Young	141	1600
	Progeny tested	35	14000
Normande	Young	161	1200
	Progeny tested	25	12700
Holstein	Young	367	2330
	Progeny tested	107	15800

(Institut de l'Élevage)

The French situation in 2010

Mean EBV (genetic standard deviations)

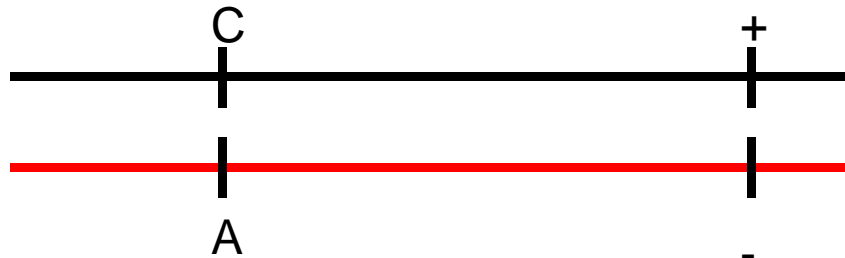
Breed	Bull category	Total merit	Dairy traits	SCC	Fertility	Longevity	Type
Montbéliarde	Young	1,8	1,4	0,3	-0,1	0,7	0,8
	Progeny tested	1,7	1,4	0,2	0,1	0,3	0,9
Normande	Young	1,5	1,5	0,2	0,1	0,6	0,3
	Progeny tested	1,8	1,4	0,8	0	0,4	0,6
Holstein	Young	2,7	1,8	0,6	0,2	1,2	1,8
	Progeny tested	2,2	1,8	0,5	-0,1	0,3	1,3

(INRA - Institut de l'Élevage)

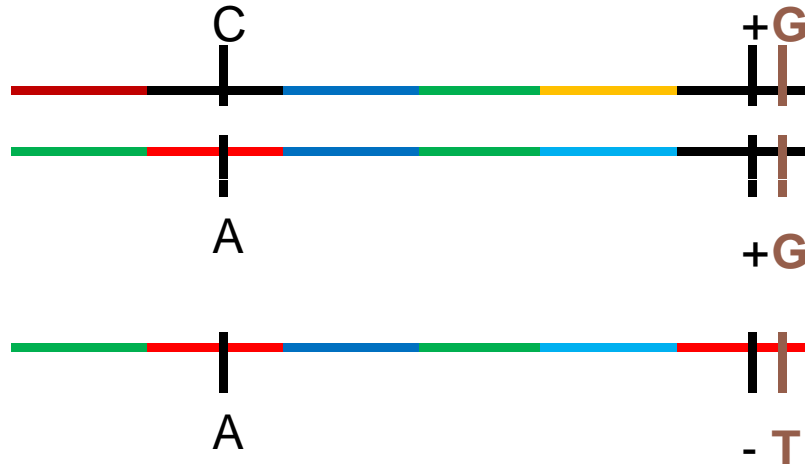
A challenge for the short term : combine breeds

- Share reference populations
- Share cost, solution for smaller breeds, increase overall efficiency, maintain solidarity, only solution for new traits difficult to record
- The trick : adapt the marker density to across breed LD, by using a High Density chip
 - ⇒ A third kind of population for imputation, in addition to candidates and reference populations

A challenge for the short term : combine breeds



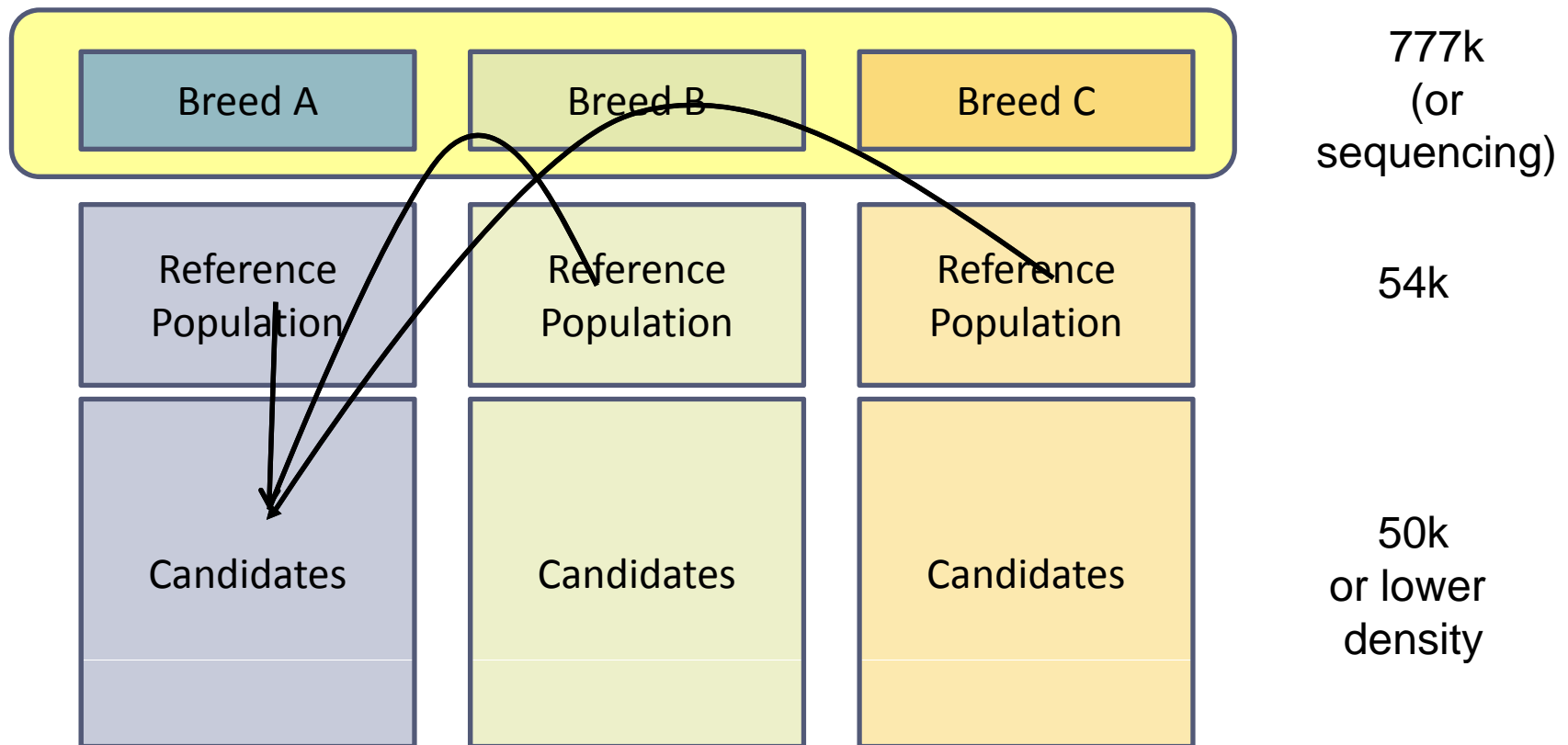
Within breed :
Long segments (300-400kb)
Medium marker density



Across breeds :
Short segments (10-20kb)
High marker density

Multi breed evaluation

Gembal project : 5000 animals genotyped in HD (beef and dairy)



And about sequencing ?

- Very rapid technological developments !
- Sequencing corresponds to the complete genotyping of all mutations (5 to 10 million ?)
- Use with the same principles as HD chip : sequence a limited number of animals, impute missing information in the rest of the population
- What evolution if sequencing is as cheap as genotyping ?

Another challenge : select for new traits

- Generate the corresponding reference populations
- Several (tens of) thousand animals
- Female population, with own performances
- Taking advantage of the large scale genotyping
- Economic model : who pays for these data ?
- New consortia : breeding company – performance recording organizations - farmers

Another challenge : select for new traits

- New phenotypes :
 - animal health : metabolic diseases, trimming data, paratuberculosis
 - milk quality : fatty acids, individual proteins
 - carcass quality : data from slaughterhouse
 - meat quality,
 - environmental footprint (methane emission), feed efficiency
 - heat detection,
 - behaviour....

In conclusion

- A revolution for dairy cattle !!!
- Potential genetic trend nearly doubled !
- Need to revisit completely the management of selection
 - => large scale genotyping for a strong selection pressure
 - => stop progeny test, use of many young bulls and bull sires
- An opportunity for functional traits, with a good reliability in spite of their low heritability
- An opportunity for new traits (milk composition, health...), as far as the phenotypes are collected for several thousand of cows
- A possibility of within herd cow selection and the only way to replace the reference populations

Collaborators and sponsors

INRA

J.J. Colleau
P. Croiseau
T. Druet ⁽¹⁾
V. Ducrocq
A. Eggen ⁽²⁾
F. Guillaume ⁽³⁾
D. Boichard

- (1) Now at Liège
(2) Now at Illumina
(3) INRA & Institut de L'Elevage

UNCEIA

A. Baur
S. Fritz
C. Hoze
L. Journaux

LABOGENA

M.N. Rossignol
L. Genestout
M.Y. Boscher

