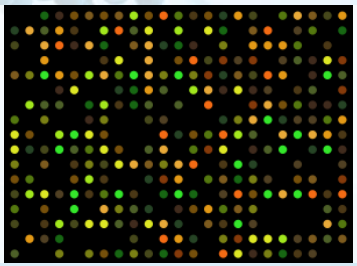
Alternative haplotype construction methods for genomic evaluation

D. Jónás[1,2,3], V. Ducrocq[1], M. N. Fouilloux[4], P. Croiseau[1]

[1]INRA, UMR1313 GABI, Jouy-en-Josas, France
[2]AgroParisTech, Paris, France
[3]ALLICE, Paris, France
[4]Idele, UMR1313 GABI, Jouy en Josas, France



Ecole Doctorale ABIES

Allice ÉLEVAGE | INNOVATION | SERVICE

AgroParisTech INSTITUT DES SCIENCES ET INDUSTRIES DU VIVANT ET DE L'ENVIRONNEMENT PARIS INSTITUTE OF TECHNOLOGY FOR LIFE, FOOD AND ENVIRONMENTAL SCIENCES

3G UMT

INRA SCIENCE & IMPACT

Introduction

- Genomic selection
  - Improved breeding value estimation by exploiting knowledge on **DNA sequence variation**
  - Bi-allelic **SNP** markers
  - Requirement: large reference populations
- An alternative to SNP: **haplotypes**
  - A combination of N SNP
  - **Multi-allelic genetic marker**
  - Which haplotypes should be used?

Objectives

- Develop a method to *a priori* **construct** the **"best" haplotypes** for genomic selection
  - How to define the "best" haplotype?
    - Maximize the number of alleles
    - Consider the distribution of **allele frequencies**



Allice
ÉLEVAGE | INNOVATION | SERVICE

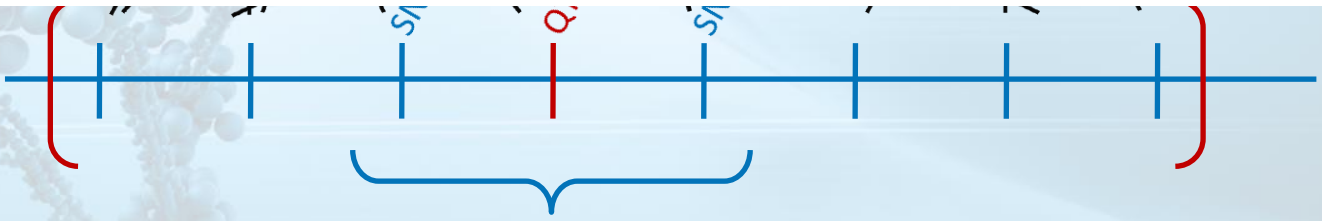
Materials and Methods

- Haplotype construction
  - Defining a combination of consecutive SNP

An allele effect was considered "predictable", if it had a sufficiently high allele frequency

- Tested haplotype sizes: 3-5 SNP
- Window size: 10 SNP

- Two, slightly different criteria
  - Termed as Criterion-A and Criterion-B
  - Difference: Criterion-B includes a constraint on the equilibrium of allele frequencies



AgroParisTech
INSTITUT DES SCIENCES ET INDUSTRIES DU VIVANT ET DE L'ENVIRONNEMENT
PARIS INSTITUTE OF TECHNOLOGY FOR LIFE, FOOD AND ENVIRONMENTAL SCIENCES

Materials and Methods

- The dataset
  - Dairy cattle breed (Montbéliarde)
  - n=2,235 individuals (Training: 1,666; Validation: 569)
  - 5 dairy cattle production traits

    Milk quantity, fat yield, fat content, protein yield, protein content
  - 43,801 SNP from the 50K chip
  - Assumed *a priori* information (QTL-SNP)
    - Represent **approximate** QTL positions
    - The most significant 1-, 3- and 6 thousand QTL-SNP

INRA
SCIENCE & IMPACT

Materials and Methods

- Analyses based on different sources of genomic information (**Bayes-C**)
  - Only QTL-SNP
  - Haplotypes built from:
    - Flanking markers
    - Criterion-A
    - Criterion-B
- Comparison of the results
  - **Correlation** between phenotypes (DYD) and GEBV

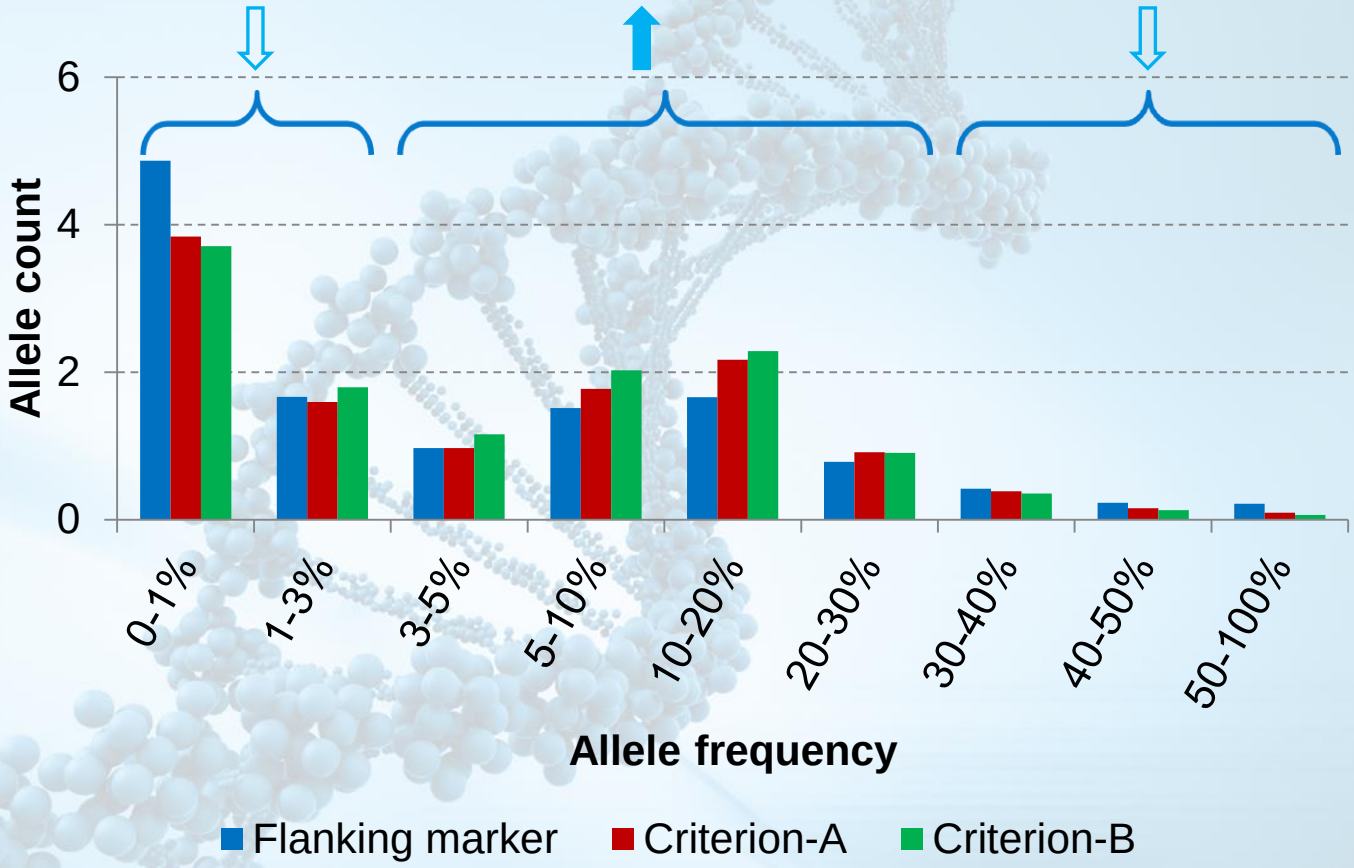
DYD-GEBV correlations observed with the flanking marker haplotypes

Number of QTL-SNP	QTL-SNP	Flanking marker haplotype		
		HS=3	HS=4	HS=5
1K	0.480	0.491	0.492	0.488
3K	0.499	0.523	0.526	0.528
6K	0.512	0.534	0.538	0.541
Optimal	0.512	0.534	0.538	0.542

- **Optimal number of QTL-SNP**: for each trait separately, the number of haplotypes leading to the highest correlation coefficient is considered
- Average correlations of the 5 production traits are shown

Allice
ÉLEVAGE | INNOVATION | SERVICE

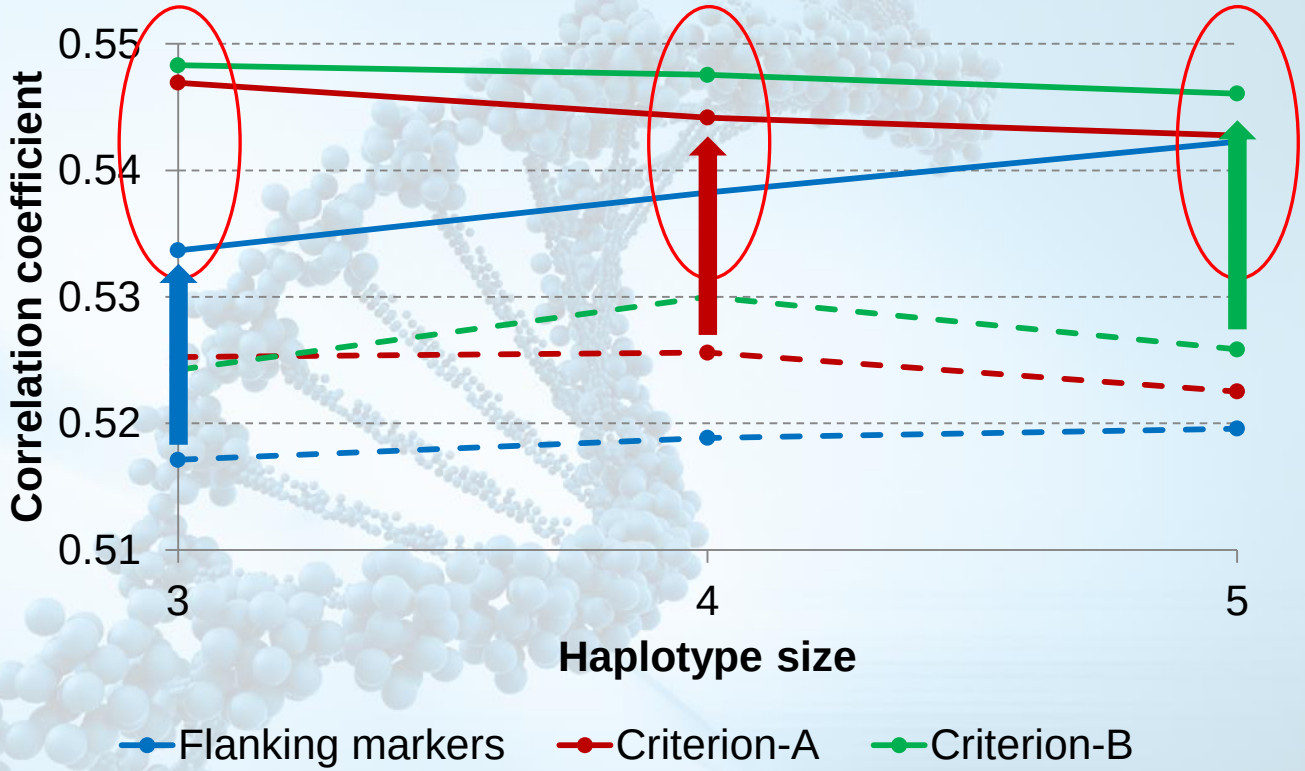
Average number of alleles per allele frequency group



- Frequency of the rare alleles decreased with either of the 2 proposed cirteria
- Similar results were observed with haplotypes of 3 and 5 SNP

AgroParisTech
INSTITUT DES SCIENCES ET INDUSTRIES DU VIVANT ET DE L'ENVIRONNEMENT
PARIS INSTITUTE OF TECHNOLOGY FOR LIFE, FOOD AND ENVIRONMENTAL SCIENCES

Correlations observed with the flanking- and selected SNP



- Haplotypes built from the selected markers were superior in genomic selection
- Gain with the haplotype selection methods decreased as the haplotype size increased

INRA SCIENCE & IMPACT

Conclusions

- **Selection of SNP is beneficial** to build haplotypes for genomic selection
  - Reduced number of markers in the model
  - Larger number of alleles with proper frequencies
  - **Higher DYD-GEBV correlations**
- Decreasing gain with the increase of haplotype size
- Similar results with the **HD-chip**
- Similar results with **other breeds**

Ecole Doctorale ABIES

Thank you for your kind attention!

Ecole Doctorale ABIES

Allice ÉLEVAGE | INNOVATION | SERVICE

AgroParisTech INSTITUT DES SCIENCES ET INDUSTRIES DU VIVANT ET DE L'ENVIRONNEMENT PARIS INSTITUTE OF TECHNOLOGY FOR LIFE, FOOD AND ENVIRONMENTAL SCIENCES

INRA SCIENCE & IMPACT