



INRA

LD and QTL mapping in selected populations

Florence YTOURNEL

Wroclaw, 11/21/2008

Linkage Disequilibrium (LD)

- Non random association of the alleles at two loci
- Population level
- Influenced by:
 - mutation
 - population admixture
 - random drift
 - **selection**

LD, selection and QTL fine mapping

- QTL (fine) mapping = study of the association between genetic polymorphisms and levels of performance
 - Use of markers in LD with the QTL
- Selection influences LD but not taken into account in fine mapping methods
- Can selection have an influence on fine mapping ?

Steps of the study

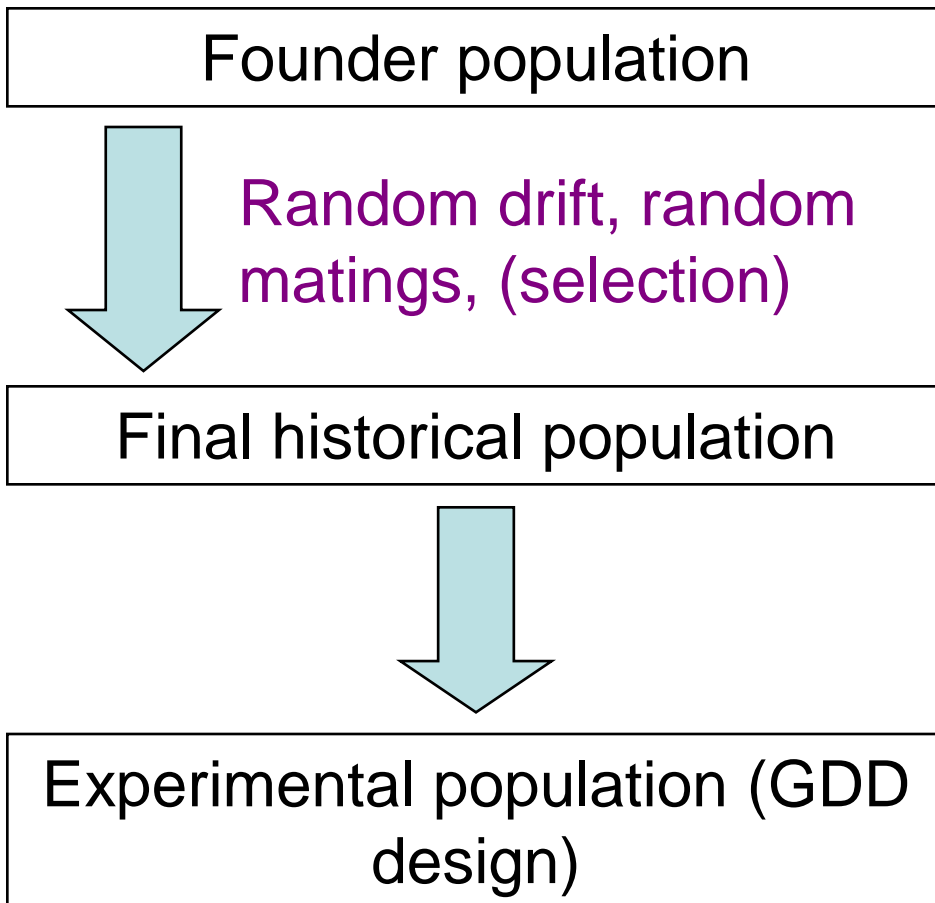
- I. Simulation of the data
- II. Location of the loci in maximum LD with the QTL
- III. IBD probabilities and selection
- IV. Influence of selection on fine mapping methods

Steps of the study

- I. Simulation of the data
- II. Location of the loci in maximum LD with the QTL
- III. IBD probabilities and selection
- IV. Influence of selection on fine mapping methods

I. Simulations

Principle and origin of the data



Population in Linkage Equilibrium (LE)

Data for the parts II et III

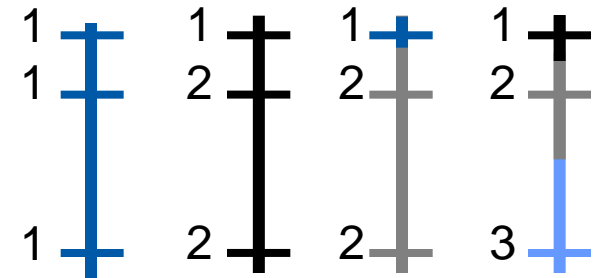
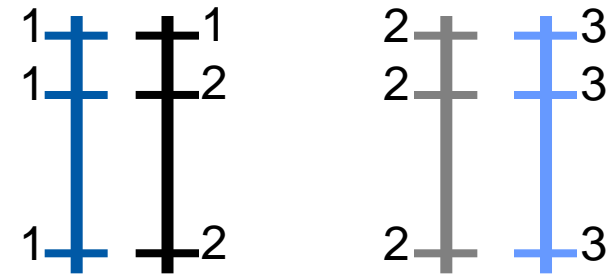
Data for part IV

I. Simulations

Simulated data

- Individuals haplotypes
 - Allele & haplotype frequencies
 - Founder origin of the marker
- Molecular information used: marker, 2- and 4-locus haplotypes
- Phenotypes of the individuals in the last generation
- Selection by truncation

$N=2 \rightarrow 4$ founder haplotypes



Steps of the study

- I. Simulation of the data
- II. Location of the loci in maximum LD with the QTL
- III. IBD probabilities and selection
- IV. Influence of selection on fine mapping methods

II. Loci in maximum LD with the QTL

Questioning

- Assumption of fine mapping : locus in maximum LD with the QTL is the closest to it
- How does selection affect that location ?

II. Loci in maximum LD with the QTL

Data

- N=150
- 100 or 80% of the individuals potential parents
- Genetic map : 19 cM, QTL in the middle, 5 alleles / locus

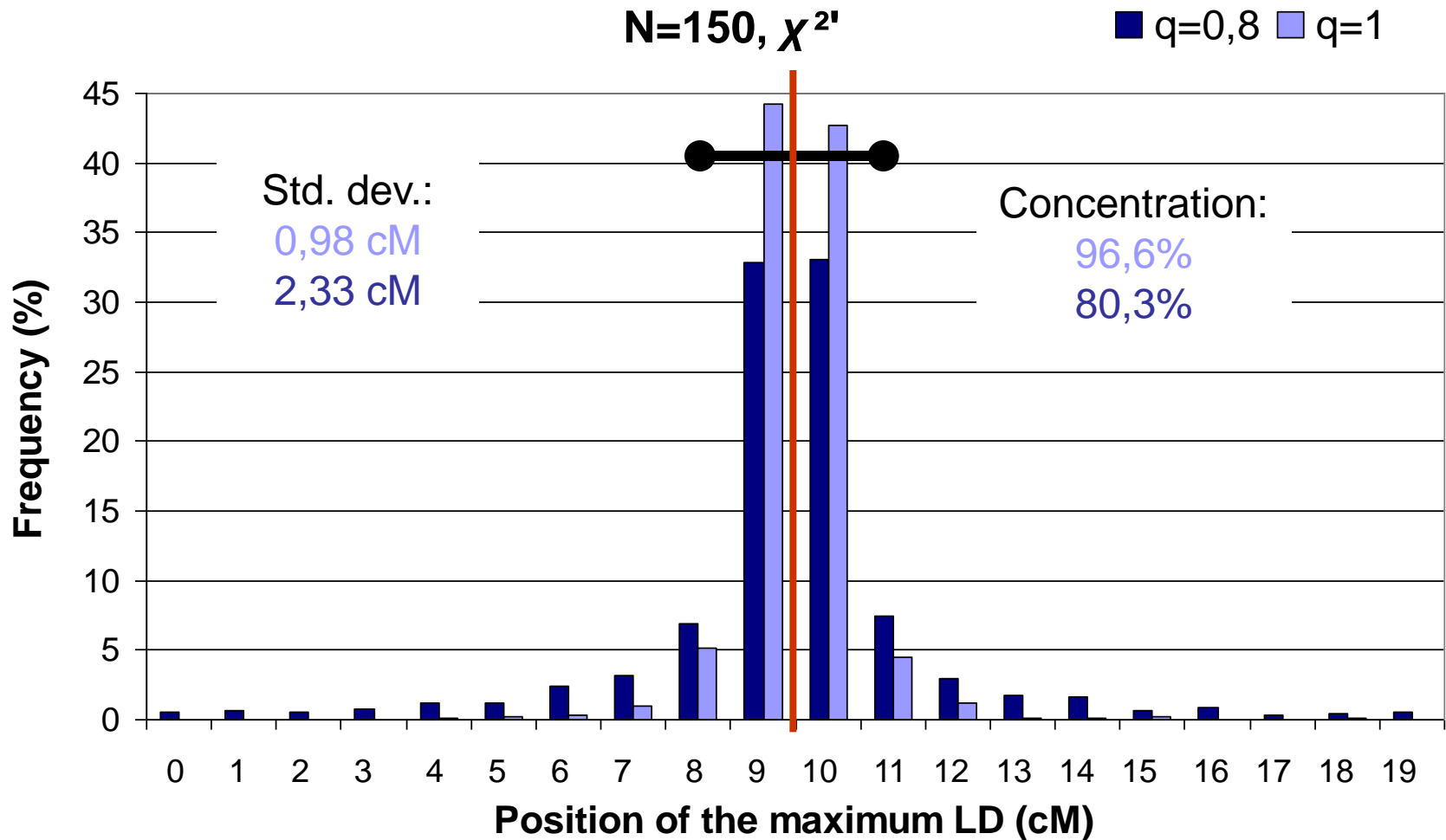
II. Loci in maximum LD with the QTL

Data and statistics

- Data :
 - Allele frequencies
 - LD (D' , χ^2')
 - Founder segment of the QTL
- Descriptive statistics of the distribution of the loci in max. LD with the QTL:
 - Average position and standard deviation
 - Proportion of loci contained in a defined interval around the QTL: « concentration »
 - Interval containing 95% of these loci

II. Loci in maximum LD with the QTL

Results



II. Loci in maximum LD with the QTL

Conclusions

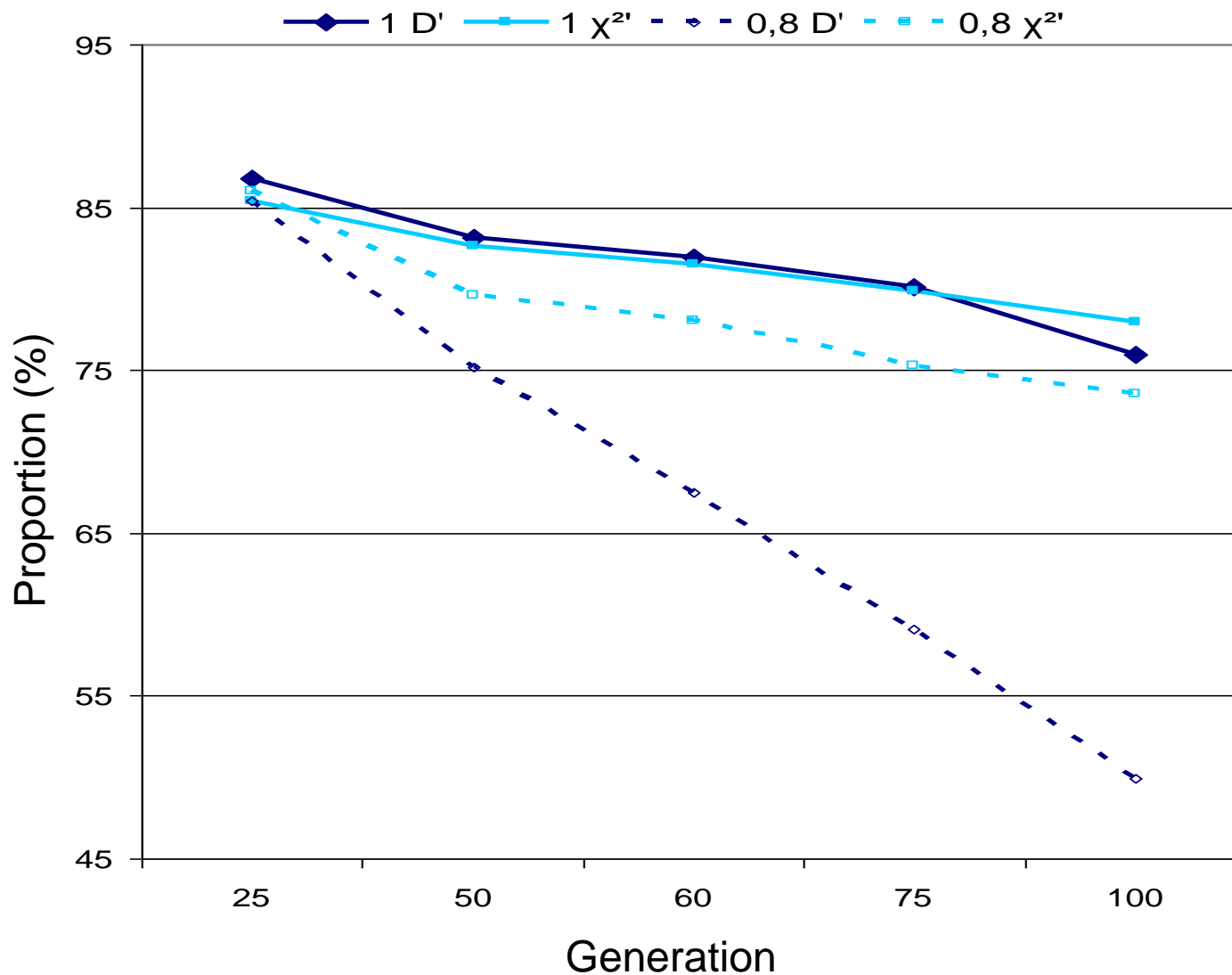
- Real populations : selected, small N_e , unknown age
 - Highly possible that locus in max. LD with the QTL is NOT the closest to it
 - But that locus remains quite close to it

Steps of the study

- I. Simulation of the data
- II. Location of the loci in maximum LD with the QTL
- III. IBD probabilities and selection**
- IV. Influence of selection on fine mapping methods

III. Selection and IBD probabilities

Locus in max. LD with QTL and QTL IBD segment



III. Selection and IBD probabilities

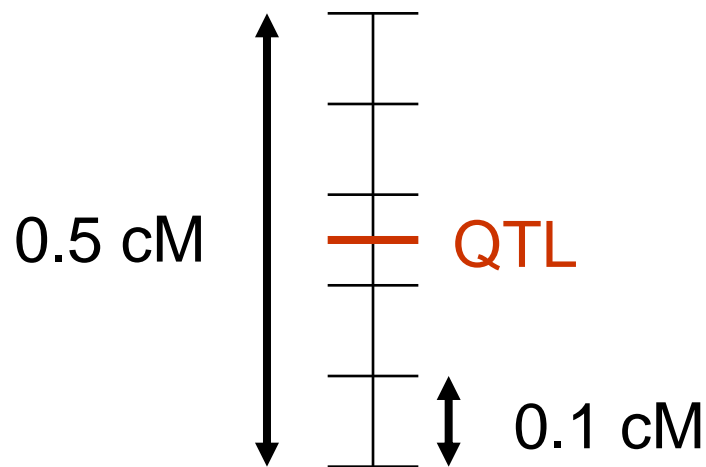
IBD probabilities

- Analytic method of Meuwissen and Goddard (2001) to compute IBD probabilities
 - IBD estimated according to the IBS status
 - No selection assumed
- Validity of these probabilities in selected populations?
- Clustering of these probabilities ?

III. Selection and IBD probabilities

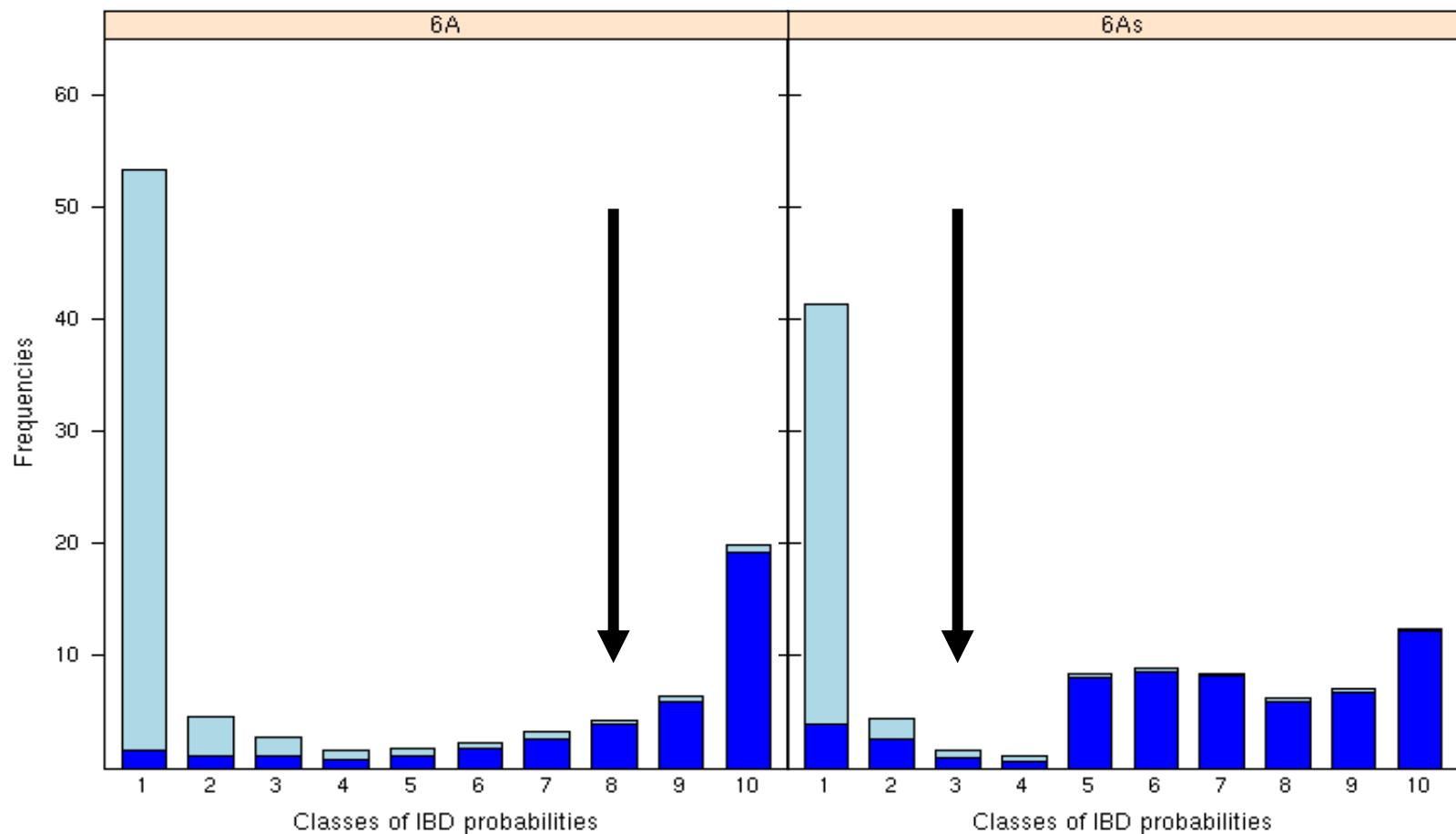
Data

- $N=100$
- 100 or 80% of the individuals as potential parents (20 last generations)



III. Selection and IBD probabilities

Distribution of the IBD probabilities



III. Selection and IBD probabilities

Conclusions

- IBD probabilities sensible to selection
- Possible to consider haplotypes showing middle-range IBD probabilities as IBD without increasing much the number of false inferences
 - Clustering according to the IBD status suggested by the IBD probabilities

Steps of the study

- I. Simulation of the data
- II. Location of the loci in maximum LD with the QTL
- III. IBD probabilities and selection
- IV. Influence of selection on fine mapping methods

IV. Selection and fine mapping methods

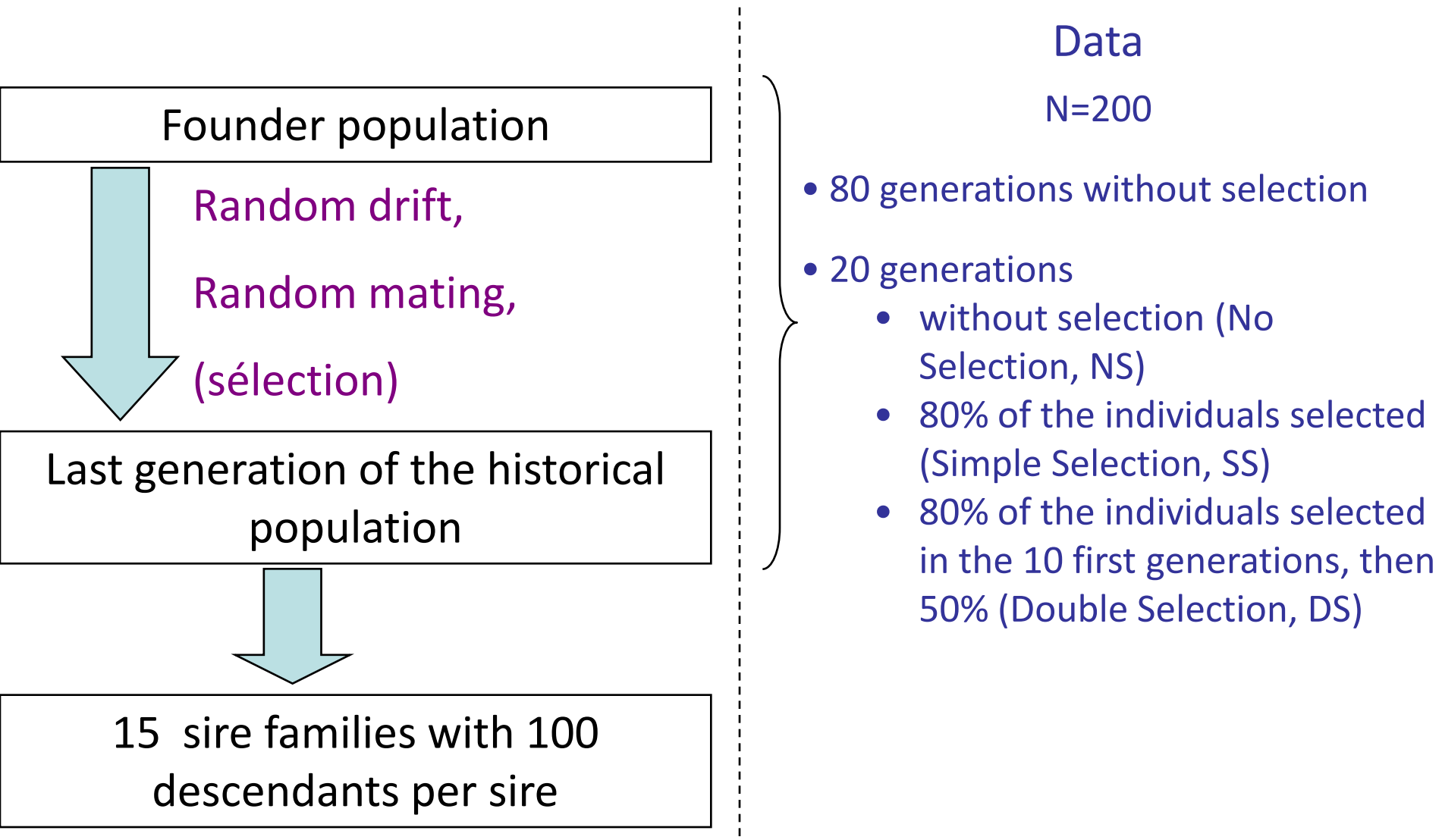
	Regression	HapIM	Terwilliger	LDLA	Cluster
Accounting for LD through ...	-	Decrease of LD with the distance		IBD probabilities	
Use of the pedigree	No	No		Yes	
Supposed origin of LD	-	Unique causal mutation		Initial population in LE	
Statistical method	Regression	Maximum likelihood		Variance components	Regression
Molecular information used	Single markers	Haplotypes of 2 markers		Haplotypes of 4 markers	

IV. Selection and fine mapping methods

- Clustering applied on IBD probabilities (Meuwissen and Goddard, 2001)
- No constant threshold for optimal clustering
 - choice of a fuzzy clustering algorithm

IV. Selection and fine mapping methods

Simulated populations

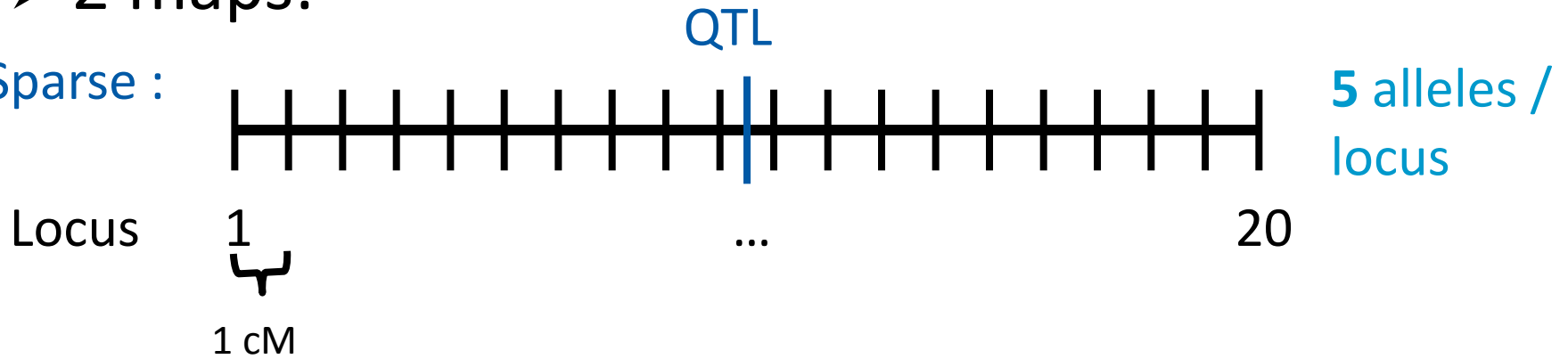


IV. Selection and fine mapping methods

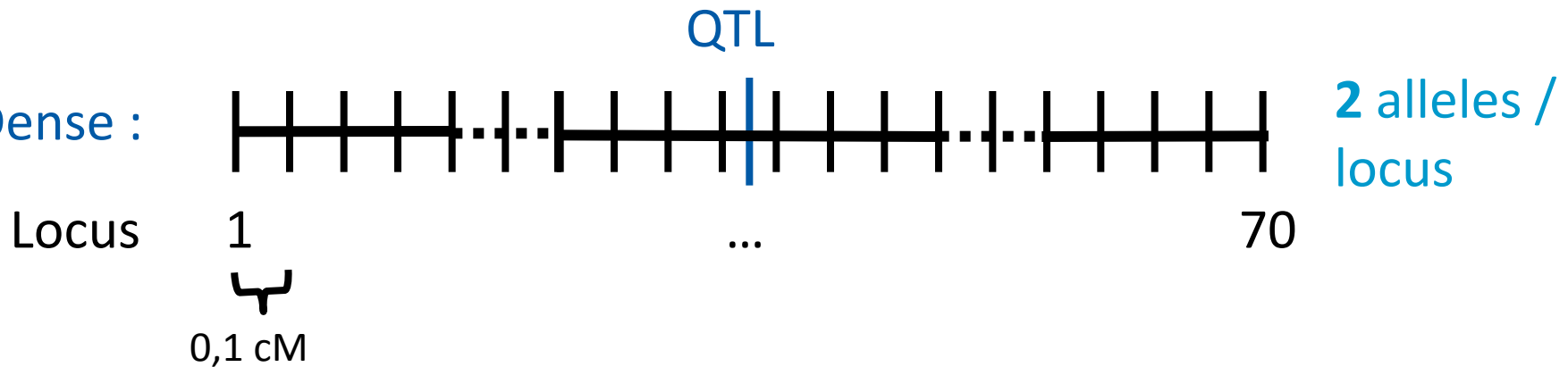
Genetic map

➤ 2 maps:

• Sparse :

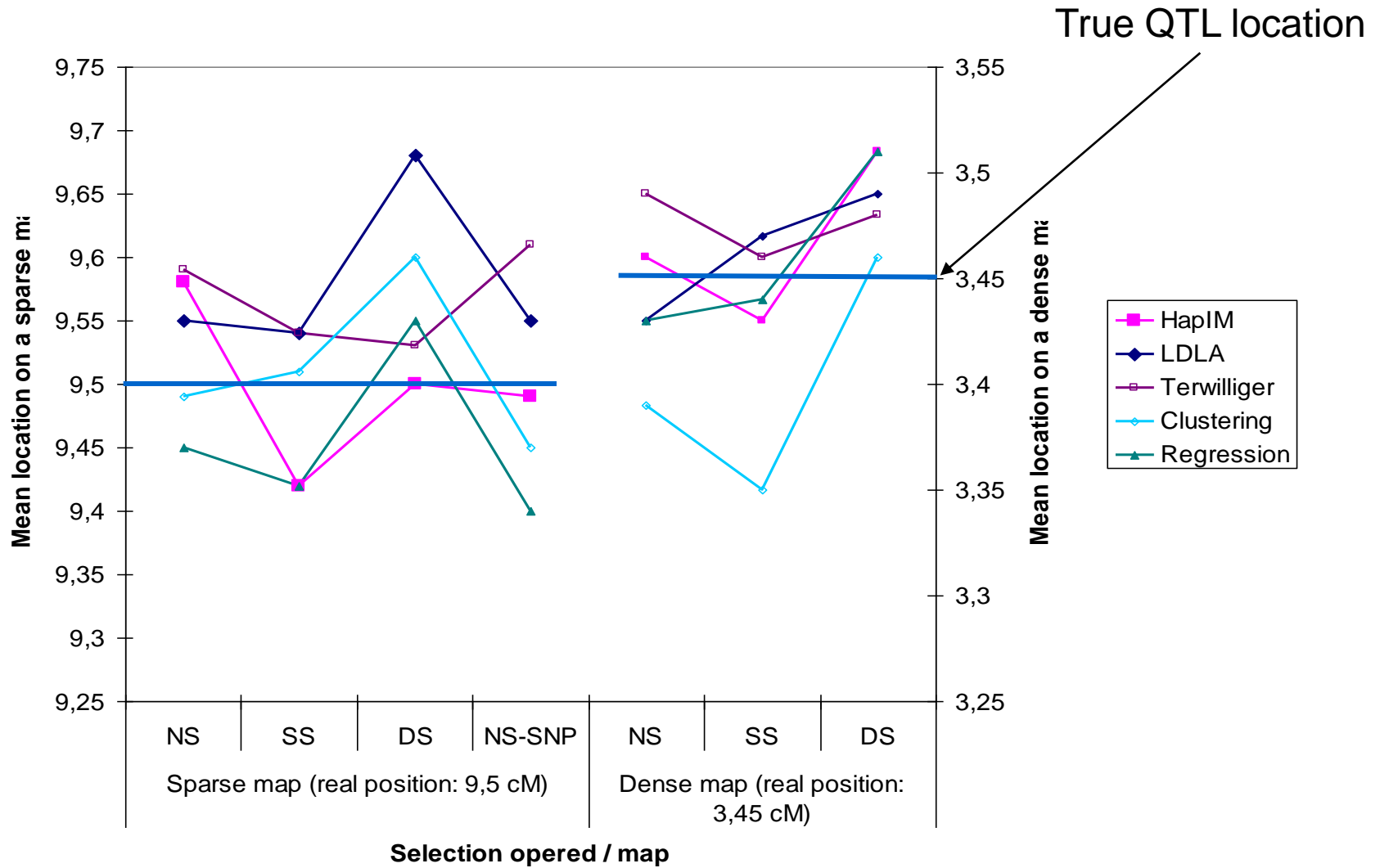


• Dense :



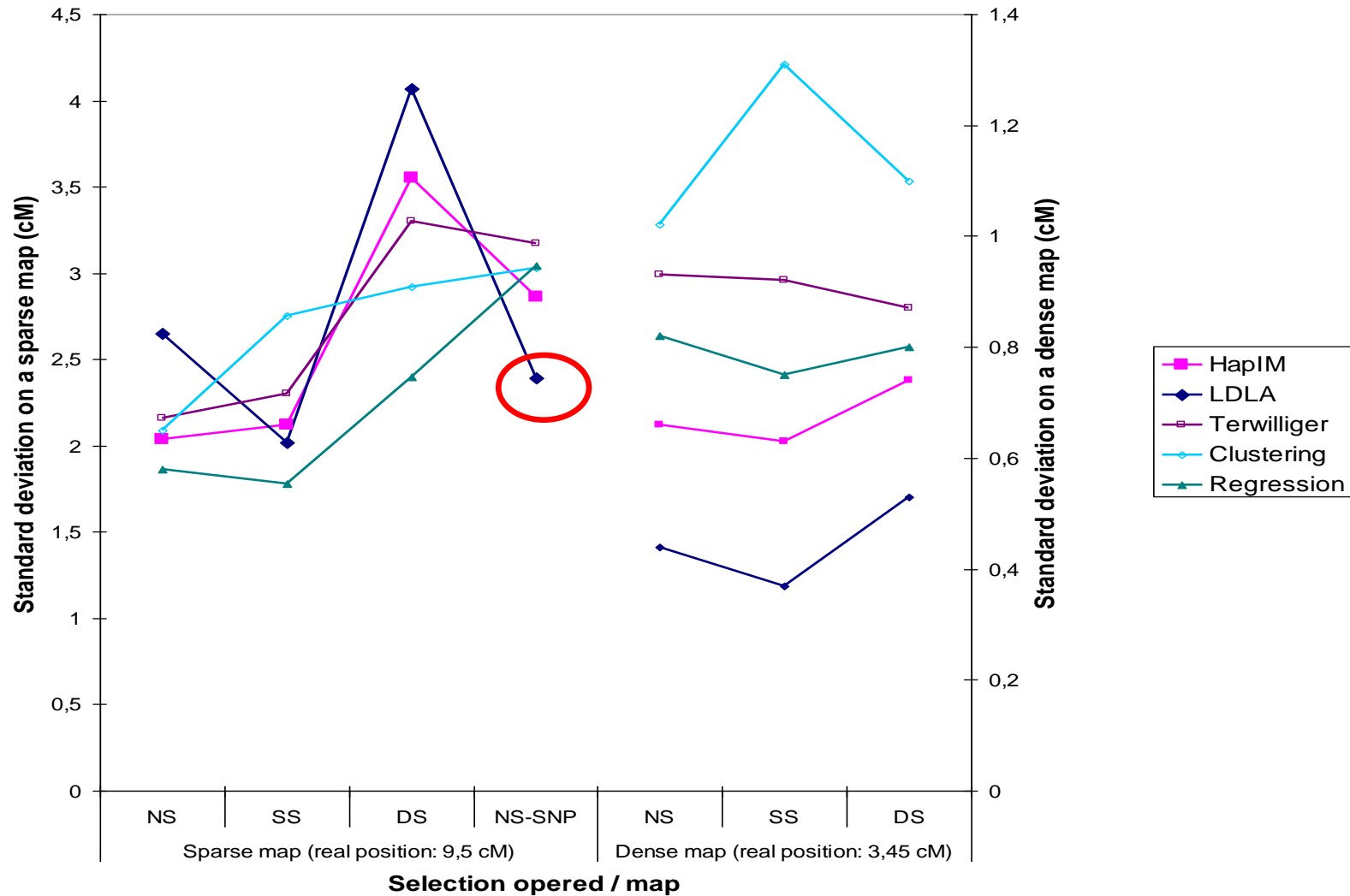
IV. Selection and fine mapping methods

Results: mean location



IV. Selection and fine mapping methods

Results: standard deviations



IV. Selection and fine mapping methods

Discussion

- On average, QTL correctly detected
- 10 generations of “strong” selection → deleterious impact (increase of variability of the location detected)
- Regression on single markers = best on sparse map covered by multi-allelic markers
- But LDLA method = best when SNP are used, whatever the map density
- This clustering does not have the qualities of full LDLA

Conclusions

- Selection & LD structure
 - ↑ distance (locus in max. LD with the QTL)
 - ↑ founder segment of the QTL
- Selection & IBD probabilities
 - Non IBD QTL correctly identified
 - Amelioration possible to identify IBD QTL

Conclusions

- Selection & fine mapping methods
 - Reduced accuracy
 - Possible explanations:
 - LD less centred on the QTL
 - LD evaluation affected
 - Best method depends on the kind of markers used
 - Haplotypes best with SNP
 - More informative
 - If length of the haplotype < length of founder segment of the QTL

Limits of the whole study

- Simulated data:
 - No comparison to real data
but mainly coherent with the literature
 - Simple hypothesis
 - No mutation

Thank you for your
attention !