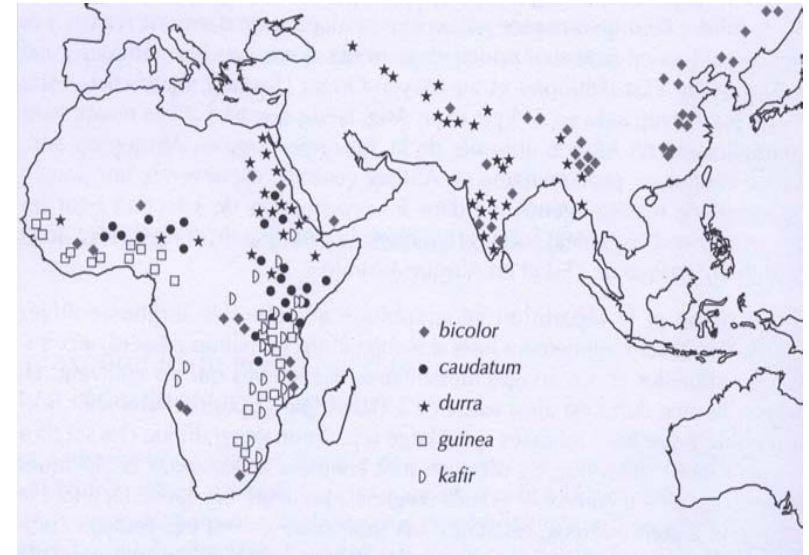
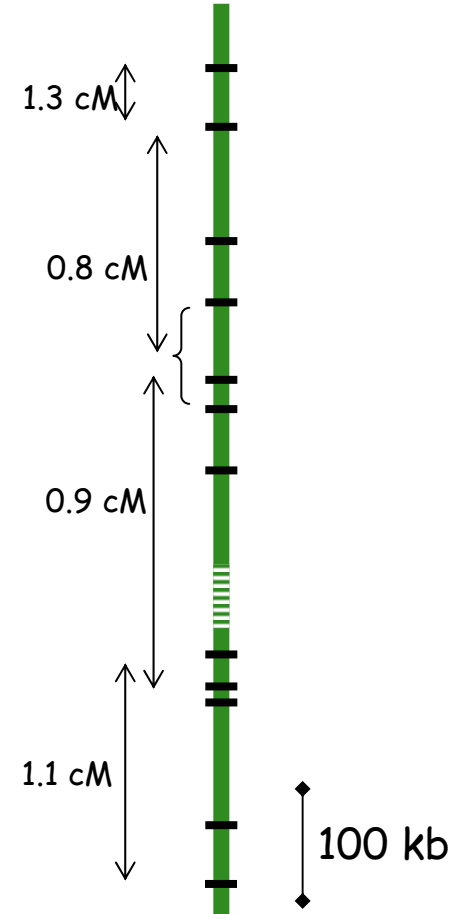
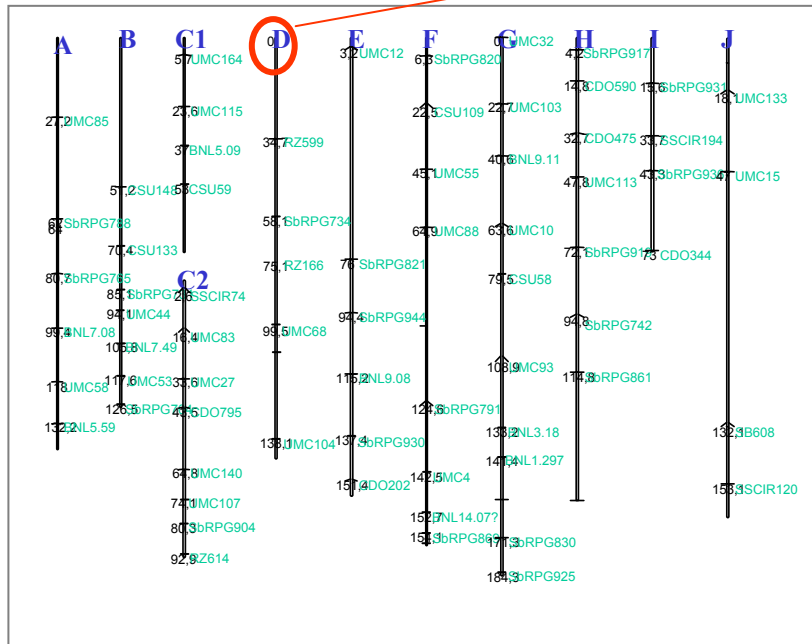


Sorghum (*Sorghum bicolor* L. Moench)

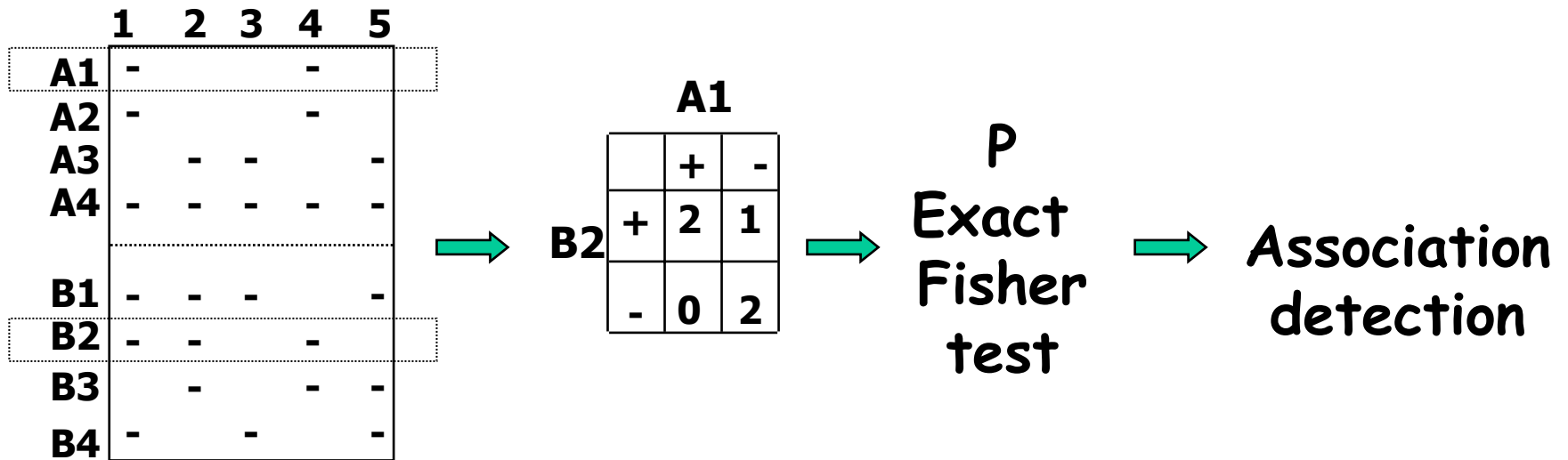


Core collection of 205 accessions

12 loci localised within
a contig of approx
4cM 600 kb

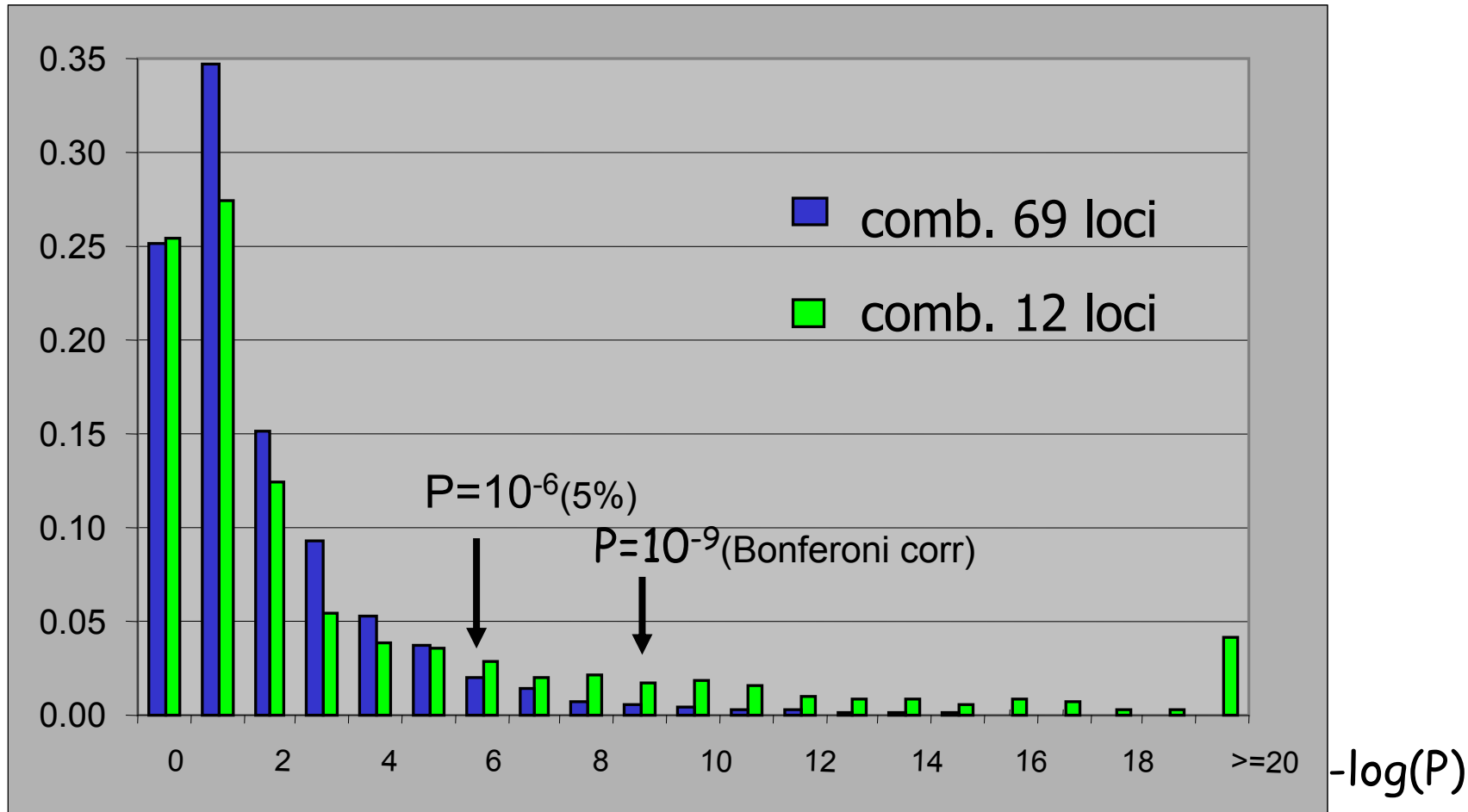


69 loci scattered over the genome



Use of Fisher statistics on
allele x allele contingency tables

Frequencies

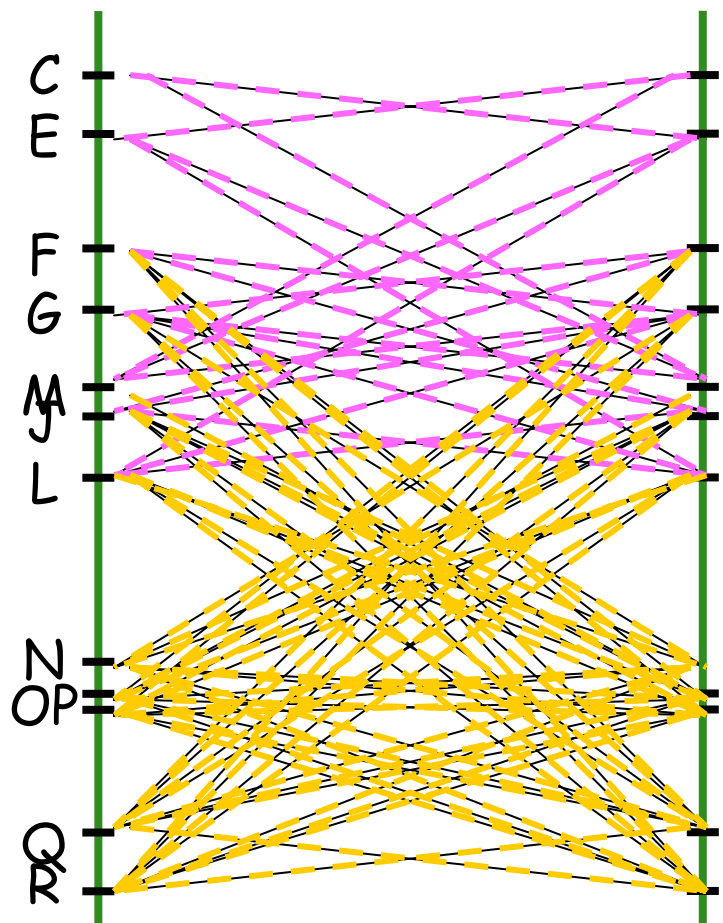


Exact Fisher test statistics

Comparison among between **distant loci** => P threshold determination

Comparison among **linked loci** => P threshold with Bonferoni correction

Global collection of 205 accessions



62 significant associations
among the alleles at the 12 markers

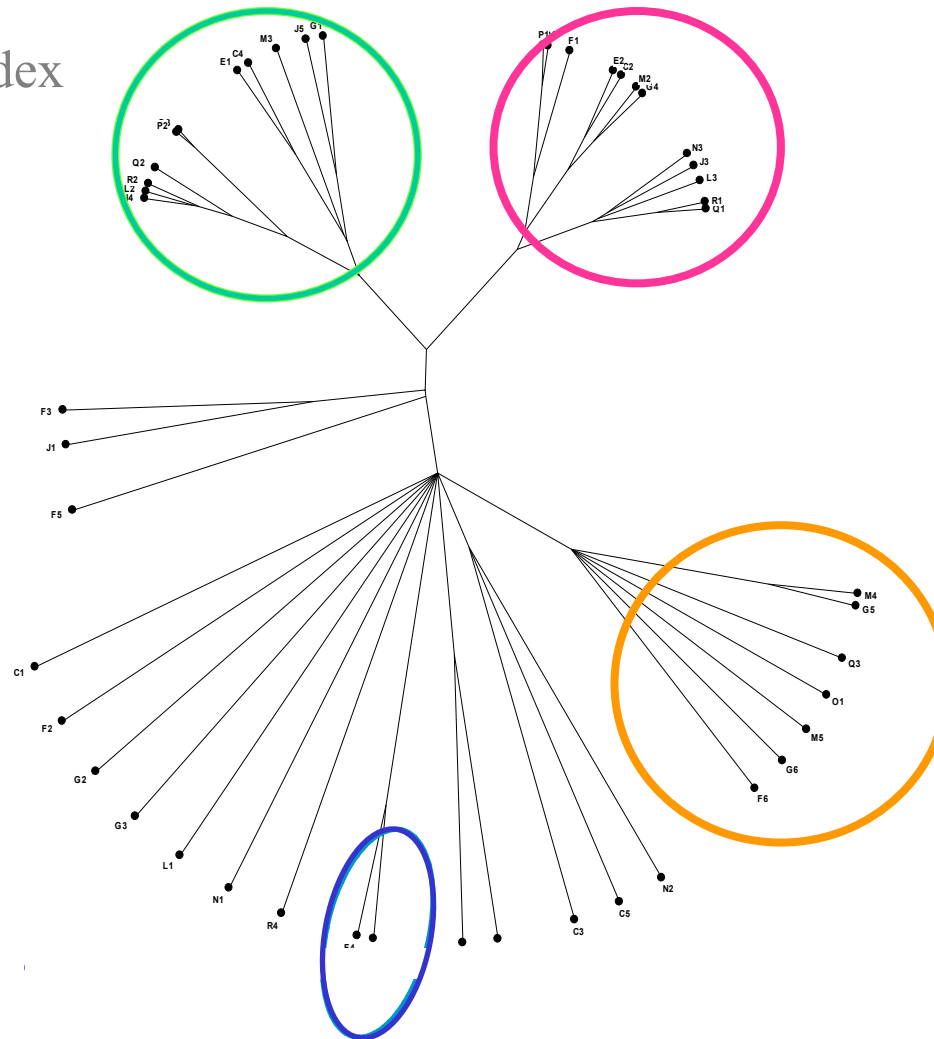


→ the 4cM/600kb
segment spans about the
resolution of LD in sorghum

Global collection of 205 accessions

Alleles at the 12 linked loci

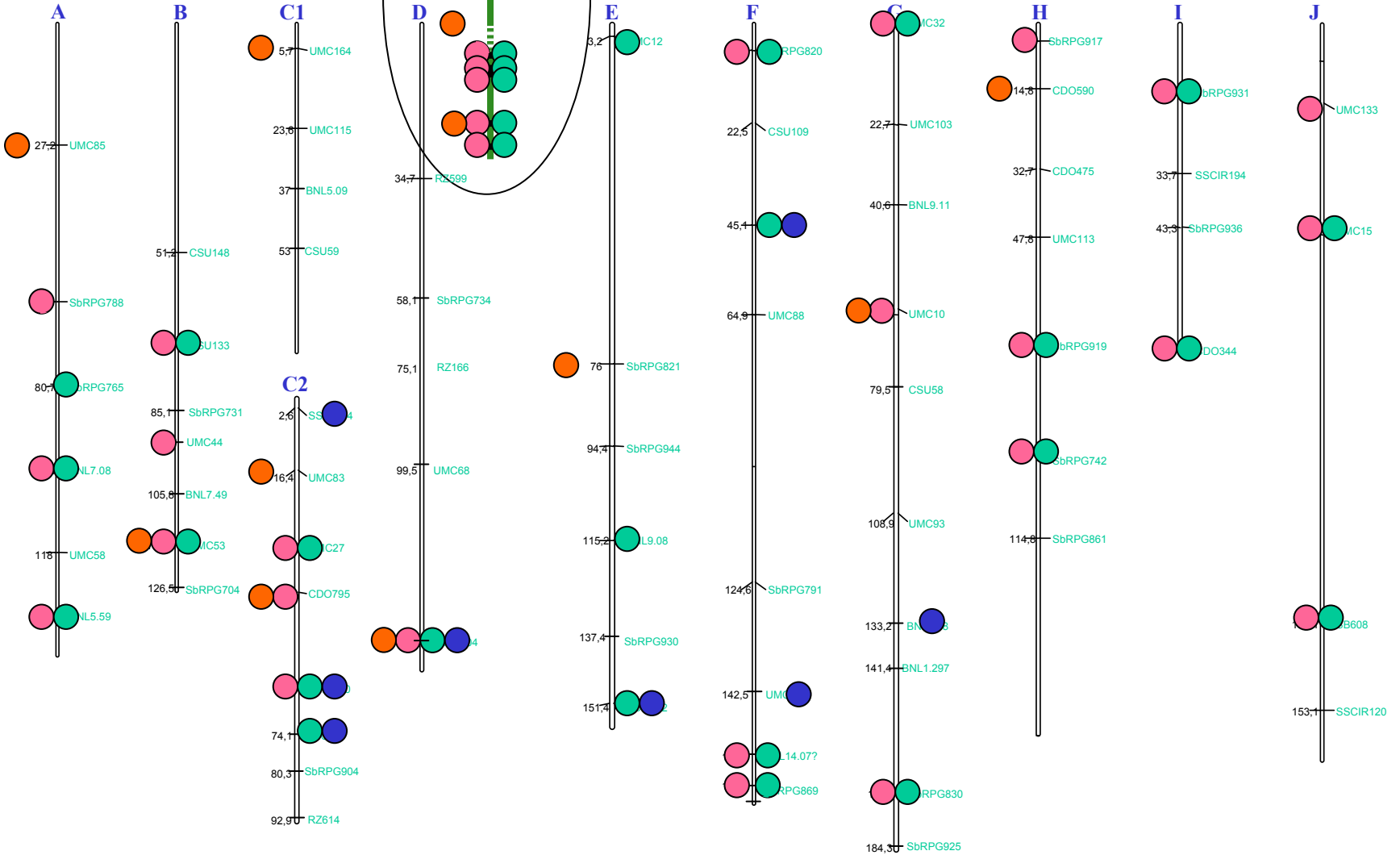
Dice dissimilarity index
NJ tree



➔ a multivariate analysis enables identification of 4 distinct haplotypes in the target region

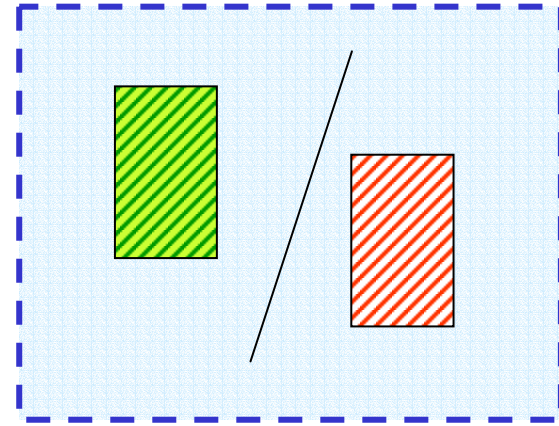
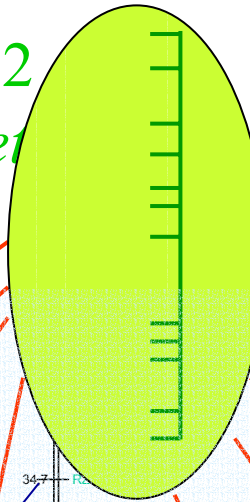
Global collection of 205 accessions

the alleles in the target region form
4 distinct haplotypes

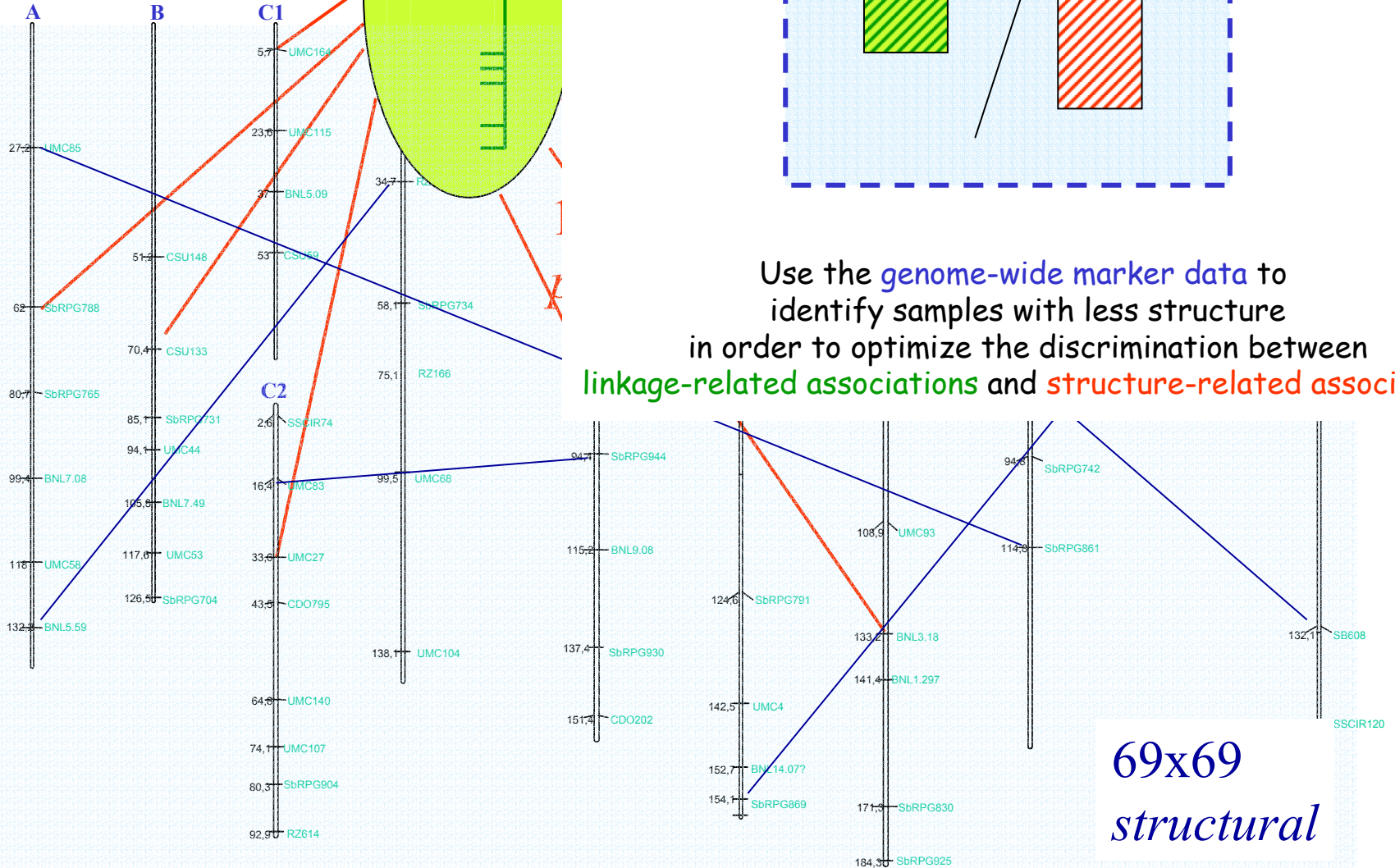


but exhibit associations with other alleles all over the genome

12x12
target



Use the genome-wide marker data to identify samples with less structure in order to optimize the discrimination between linkage-related associations and structure-related associations



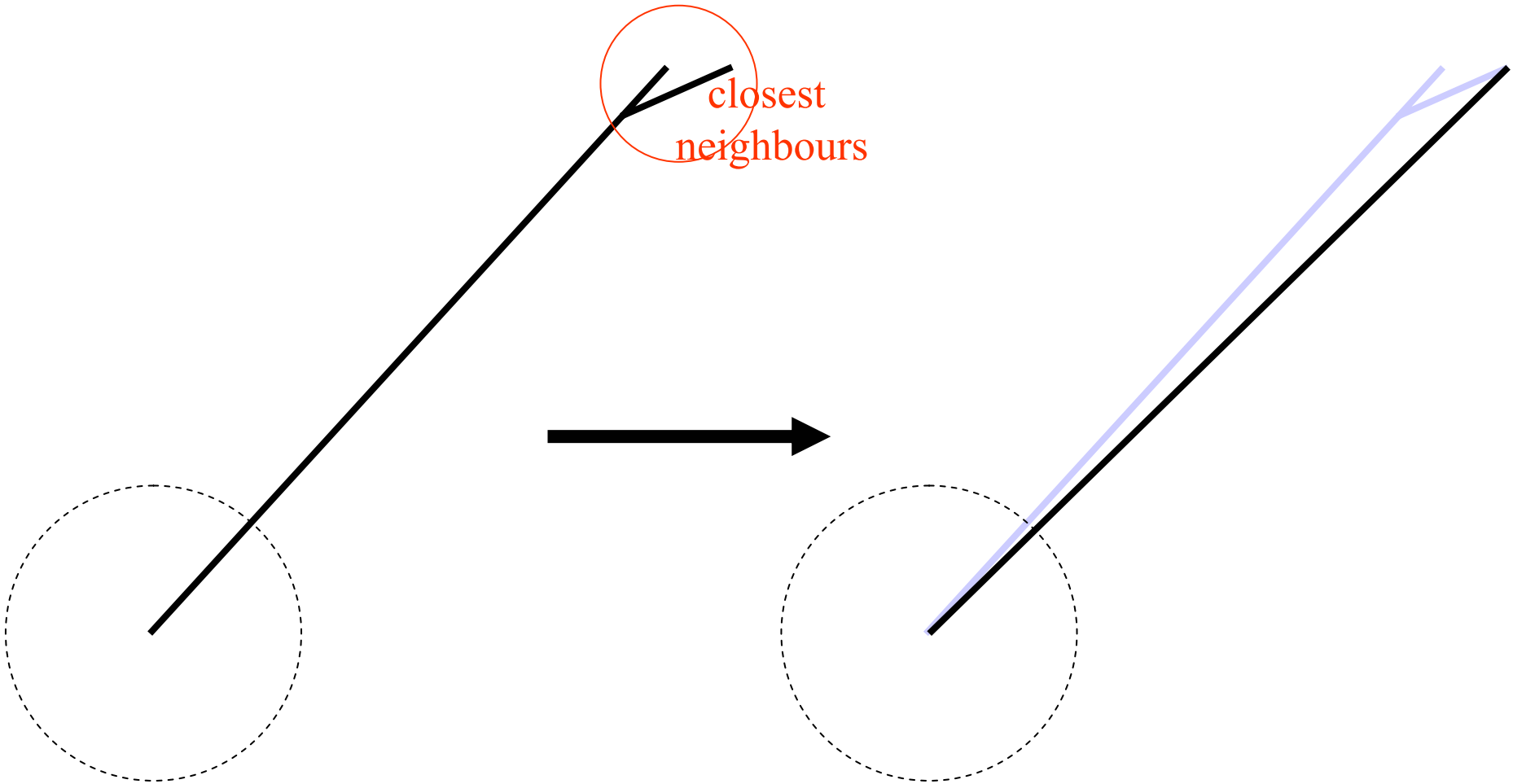
69x69
structural

Developing a sampling strategy to optimize resolution of linkage-related LD

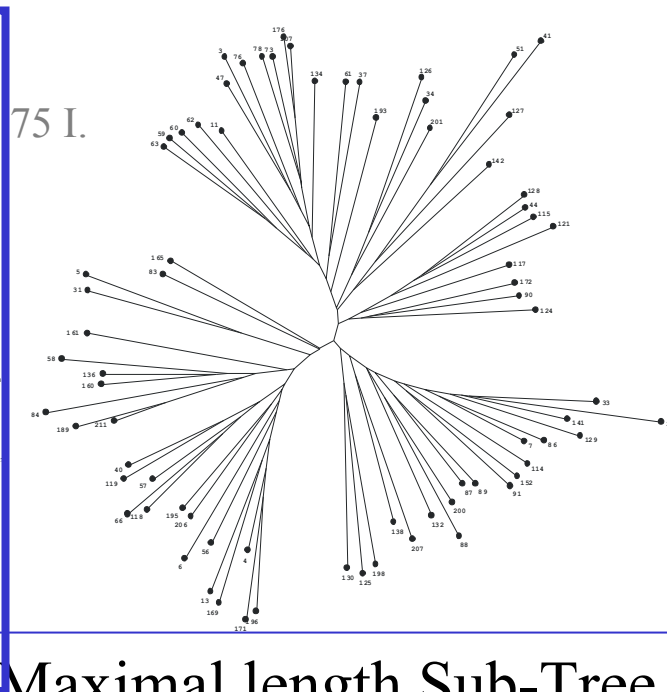
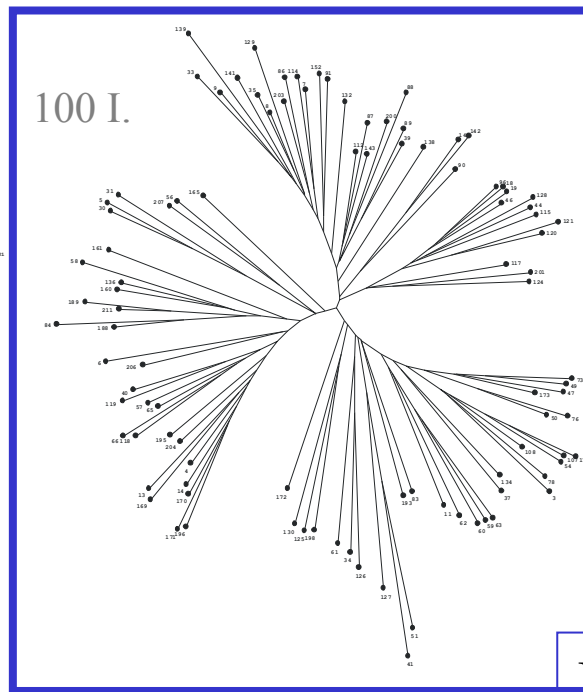
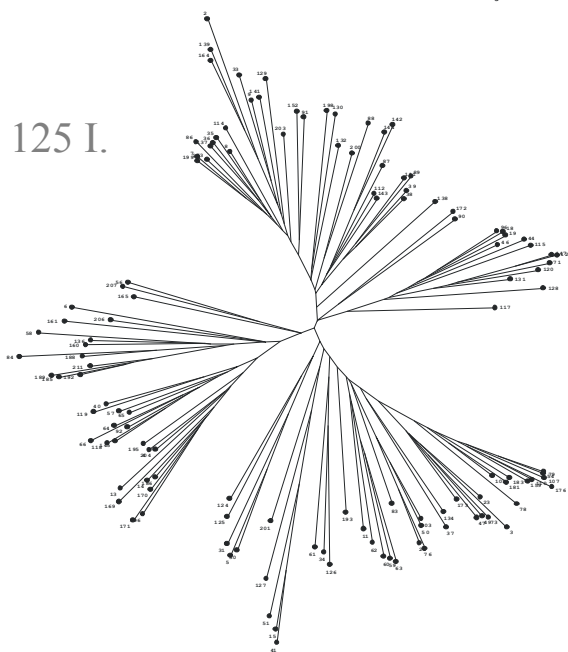
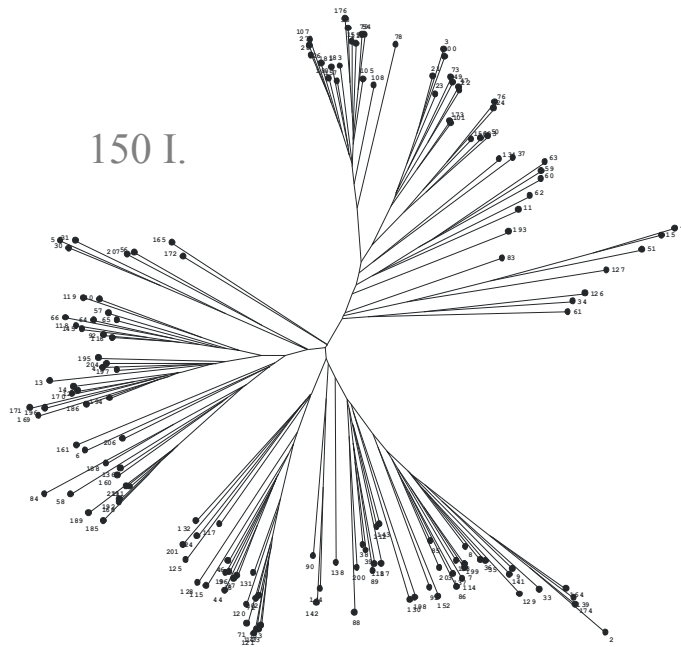
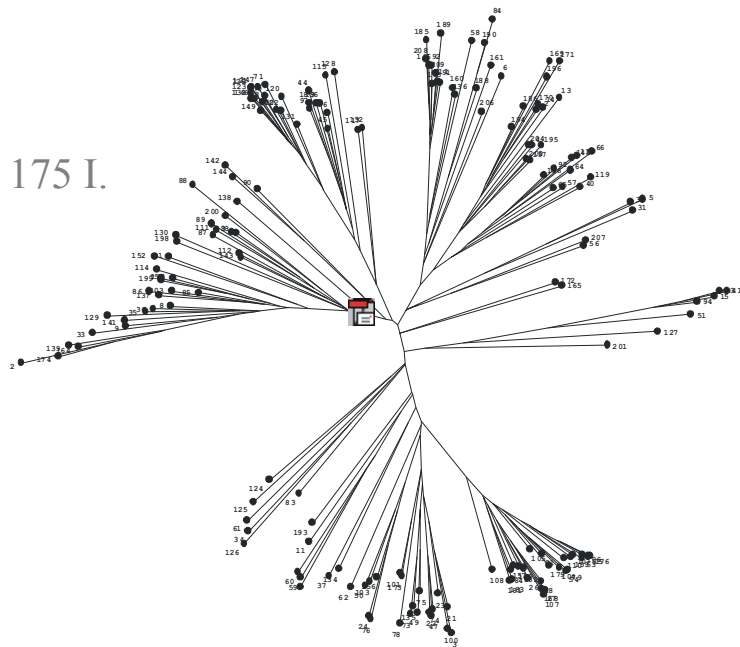
Sampling methods

Two sampling methods

- Maximum length sub-tree
- LD score : contribution of each accession to the global LD



Maximal length Sub-Tree



Maximal length Sub-Tree

LD-based sampling method

At each step:

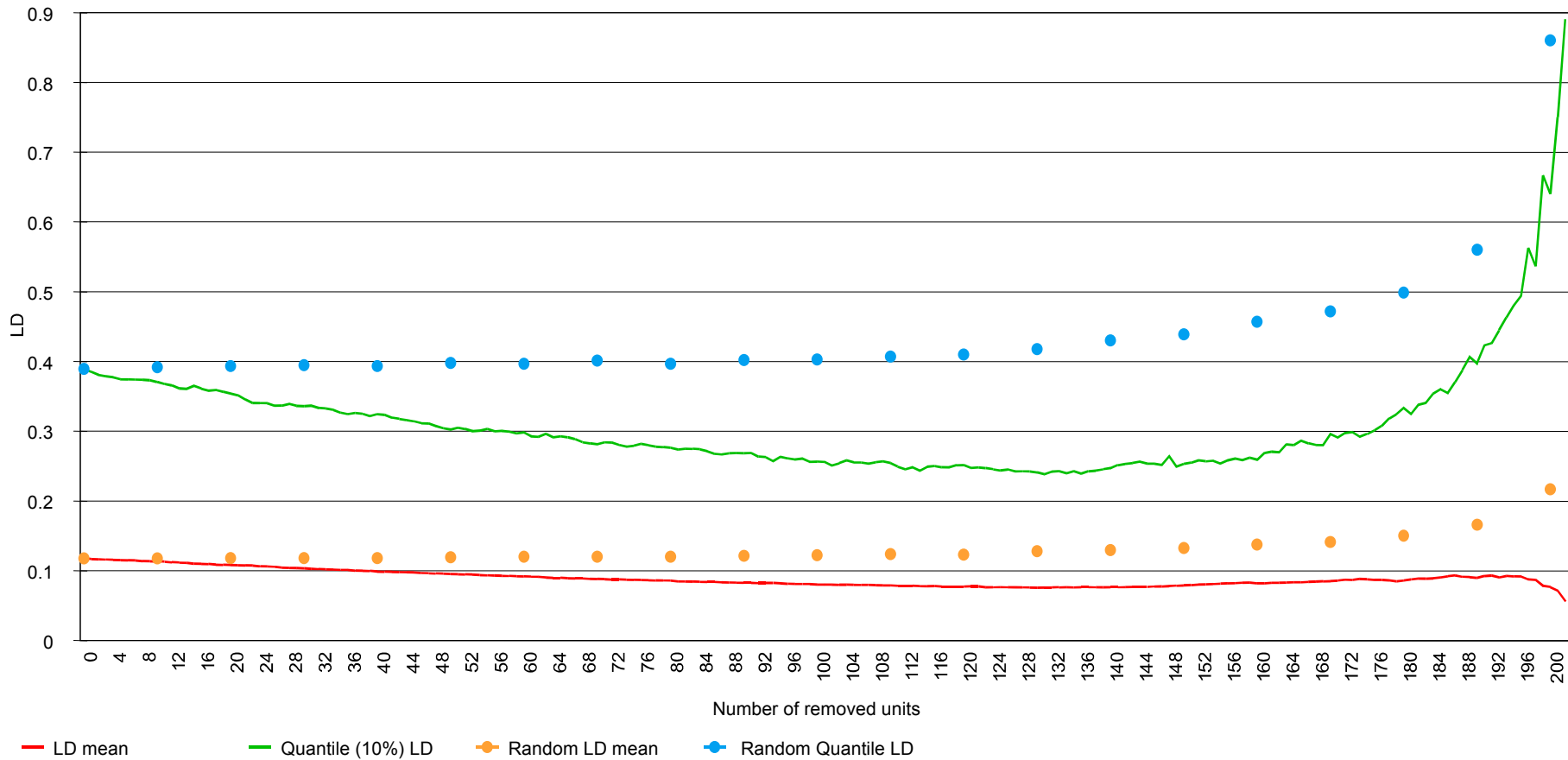
- the global LD is computed
as the total of all pairwise (locus x locus) LD
using the Lewontin LD parameter

$$Global\ LD = 1/[n(n-1)/2] \times \sum_{i=1}^k \sum_{j=1}^l p_i q_j D_{ij} \times 1/[N^2/4]$$

N = locus number

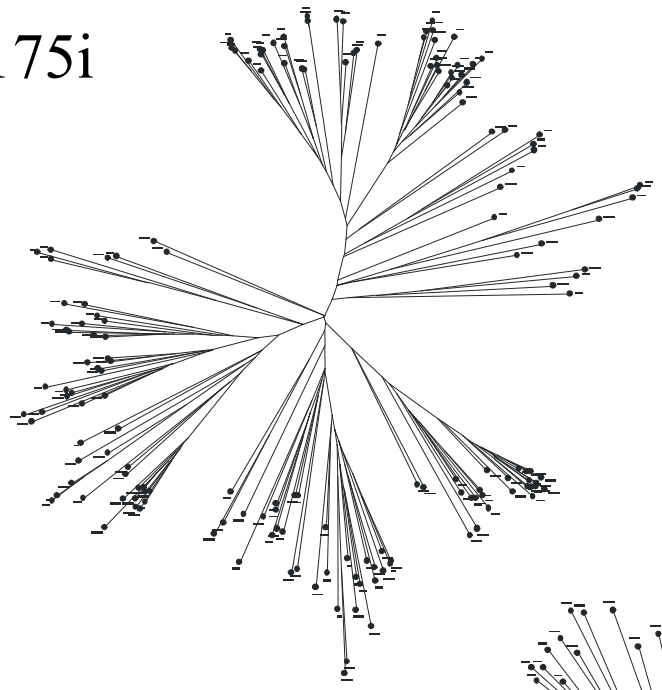
N = sample size

- the accession whose removal decreases most the LD is discarded

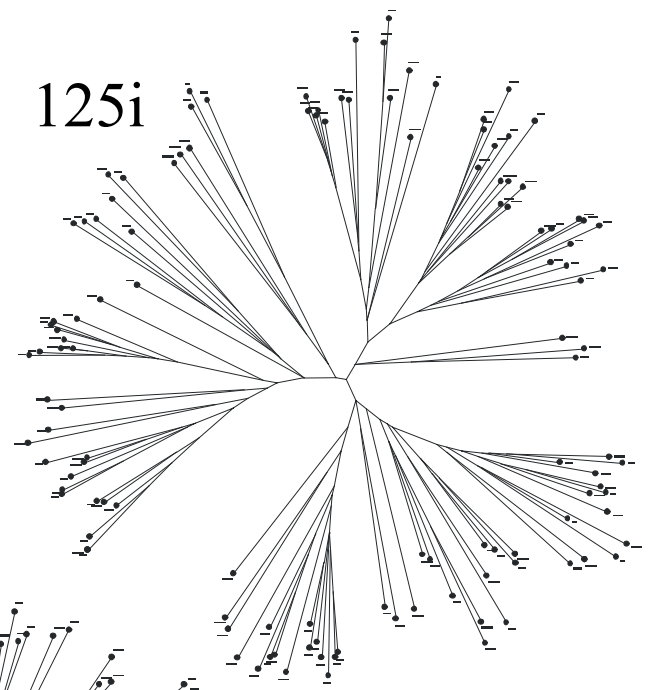


Score LD

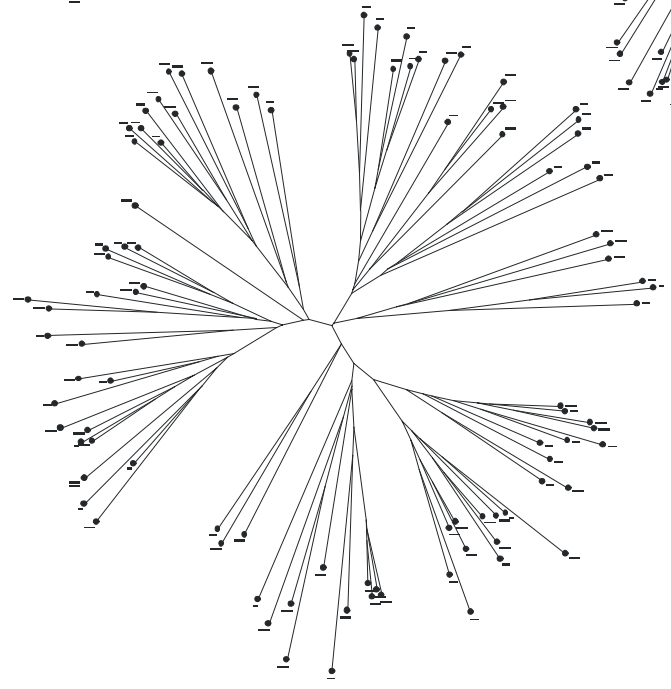
175i



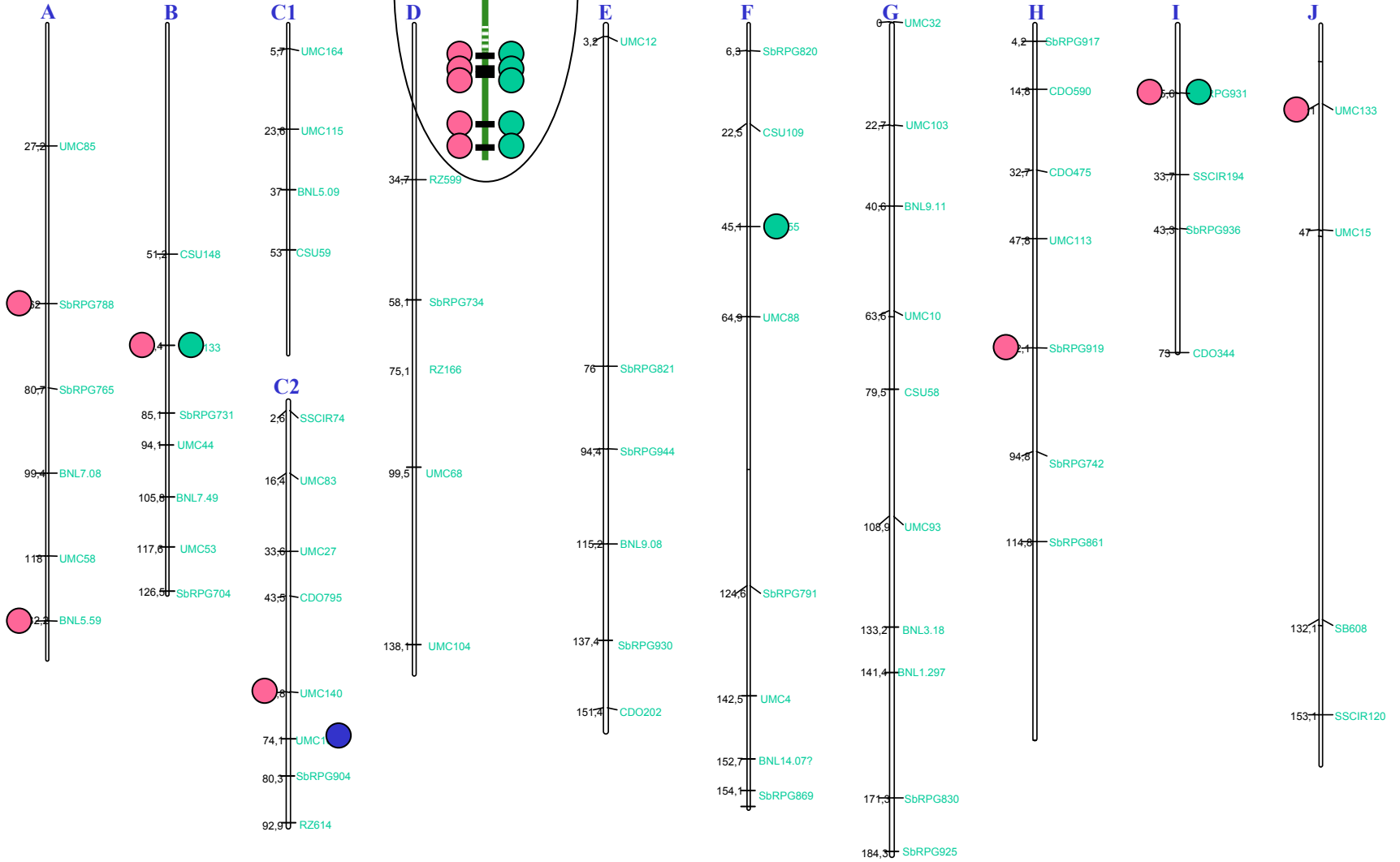
125i



100i



LD score



Issues

- Estimation of LD
- Multi allelism (>2)
- Rare alleles
- Evolution of allele diversity
- Mstrat, others
- Efficiency compared to application of Structure/Strat
- Imposing some choices
- Other population structures
etc

Acknowledgements

Cirad sugarcane group

Angélique D'Hont
Olivier Garsmeur

Cirad biometrics group

Xavier Perrier
Jean Pierre Jacquemoud- Collet

Cirad sorghum researcher

Jean François Rami