

HOS 6236 Molecular Marker Assisted Plant Breeding Fall 2017

Last Section:

QTL Mapping and Linkage mapping

Next Section:

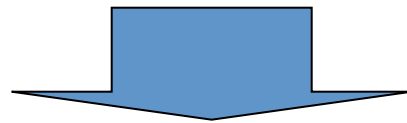
Genome Wide Association Studies (GWAS)

Todays Class:

Linkage Disequilibrium

Quantitative trait loci analysis:

- create a recombining population from two or more parents
- build a genetic map
- contrast the value of individuals in different marker allele classes
- significantly different phenotypic values identify markers linked to QTN

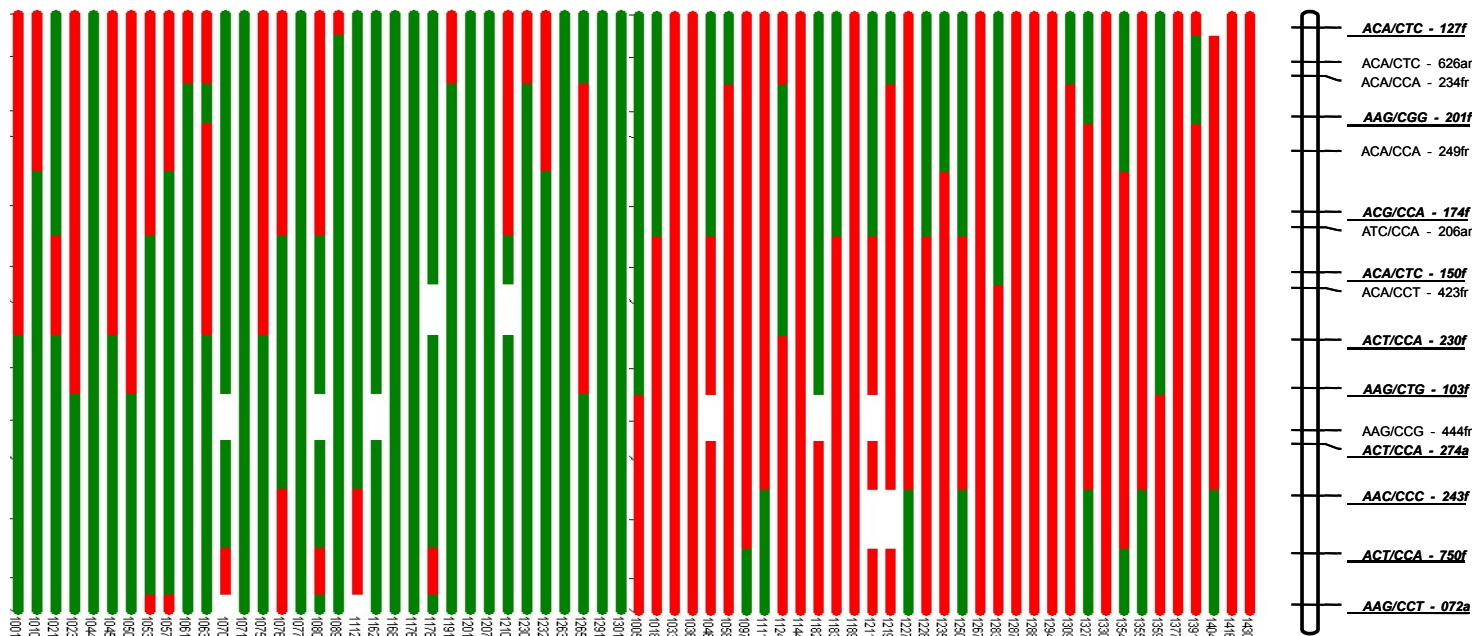


Distance between marker and QTN?

Unknown (~10-20 cM or 5-10 Mb)

Quantitative trait loci analysis

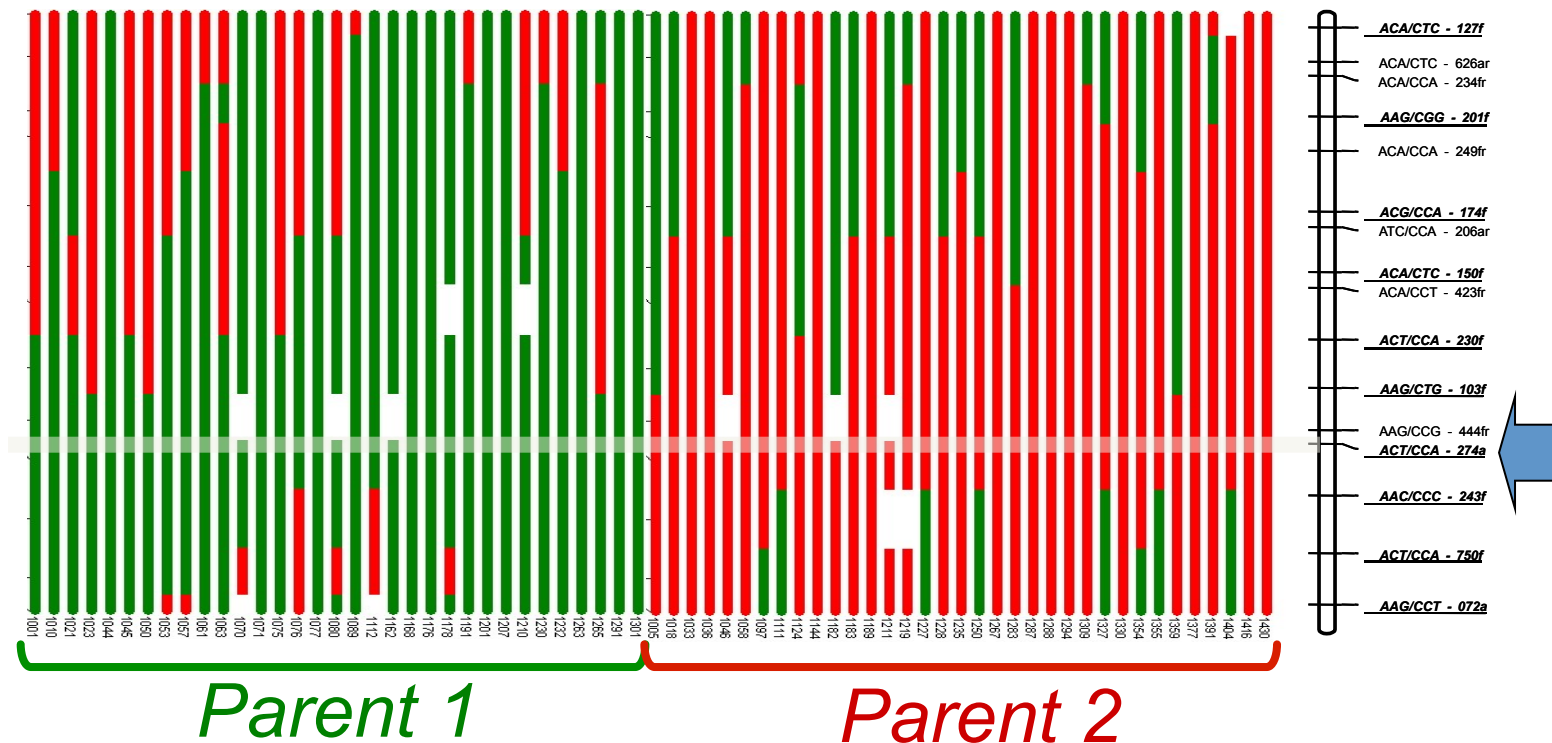
Genetic map:



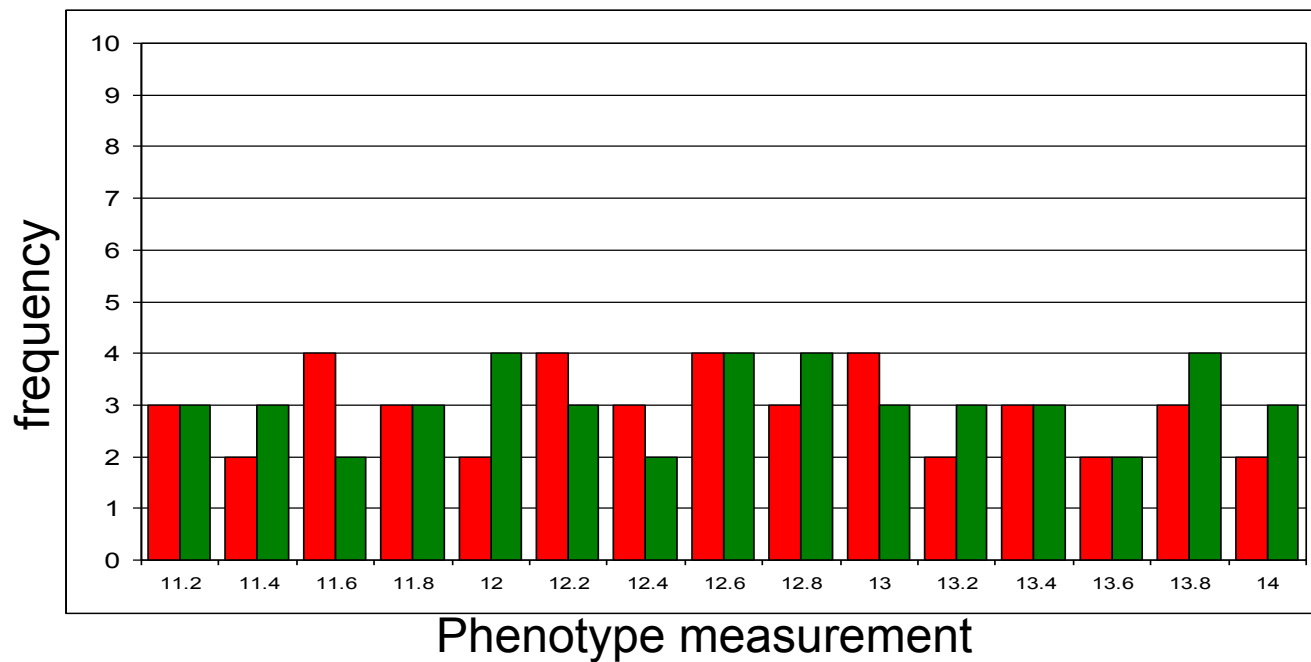
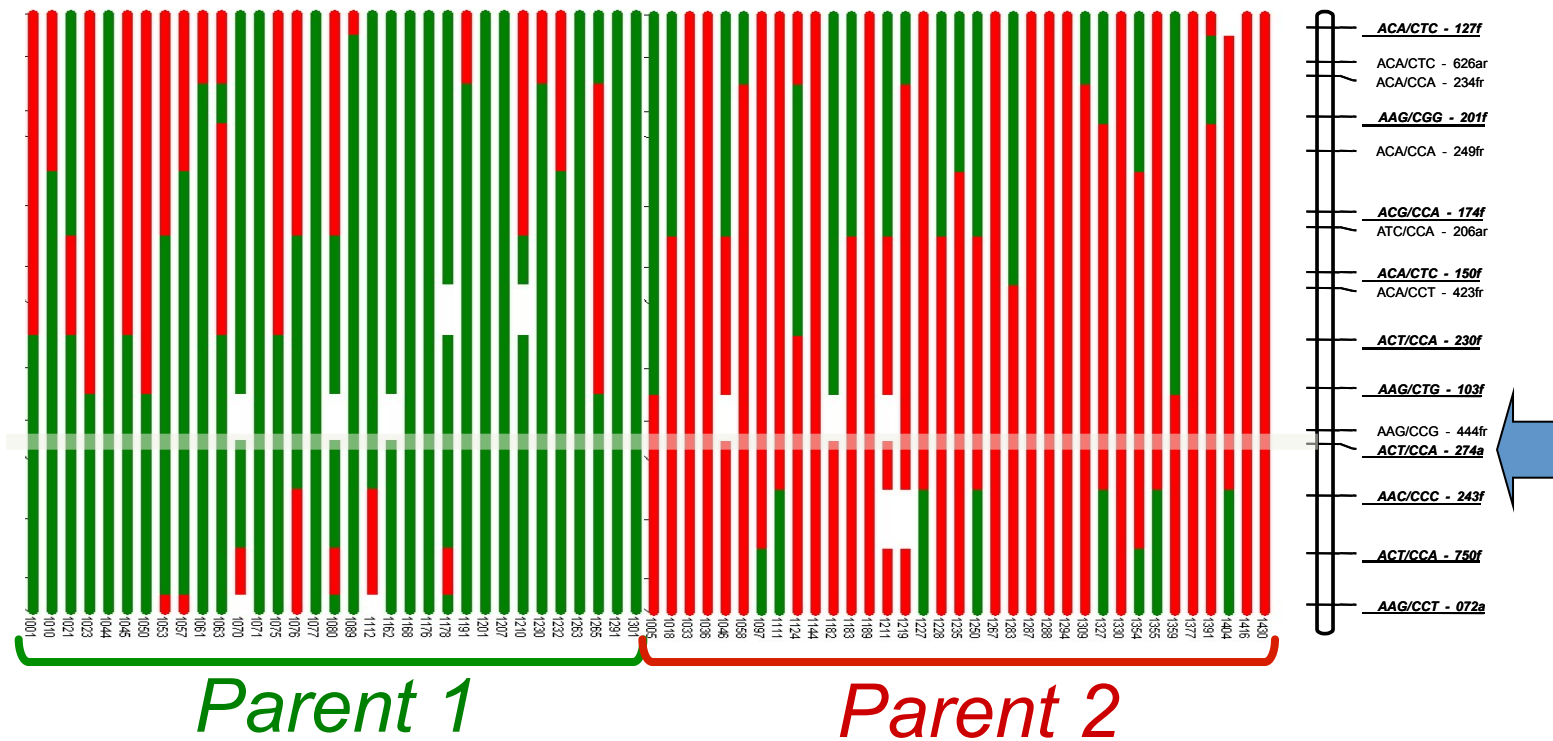
- Based on frequency of recombination between genetic markers inherited from the two parents (**Parent 1** and **Parent 2**), build genetic map
- distance between genetic markers is generally in the order of 1-20 cM

Quantitative trait loci analysis

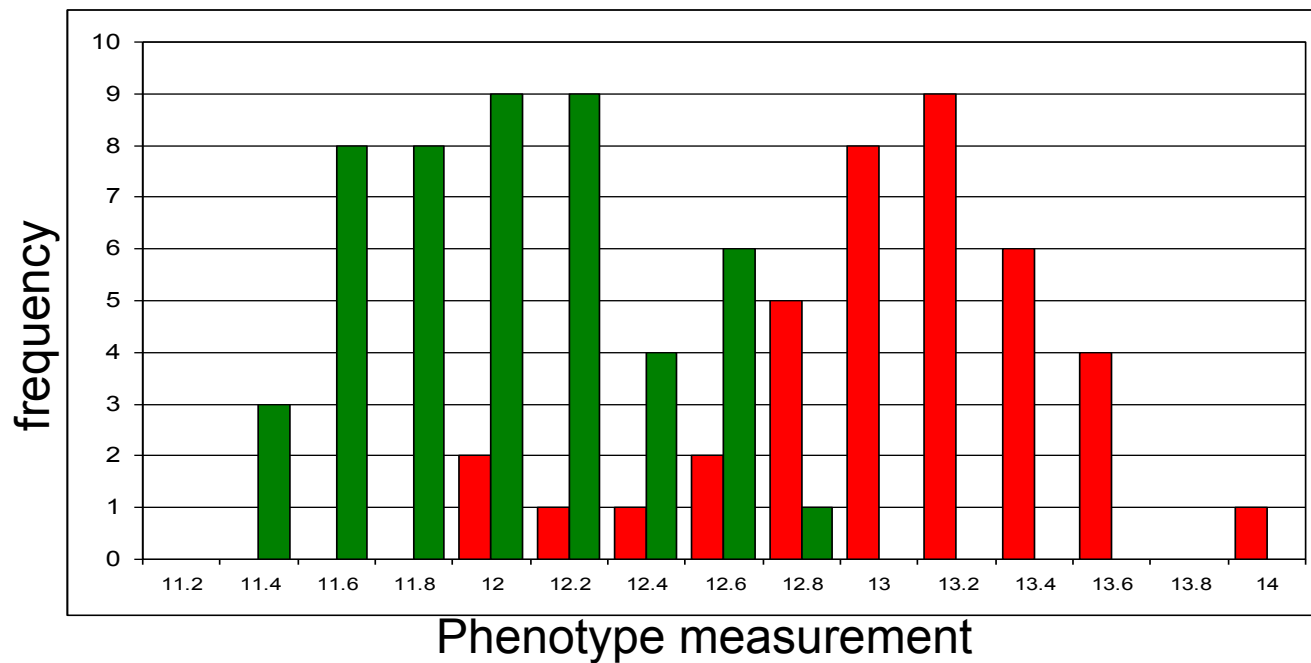
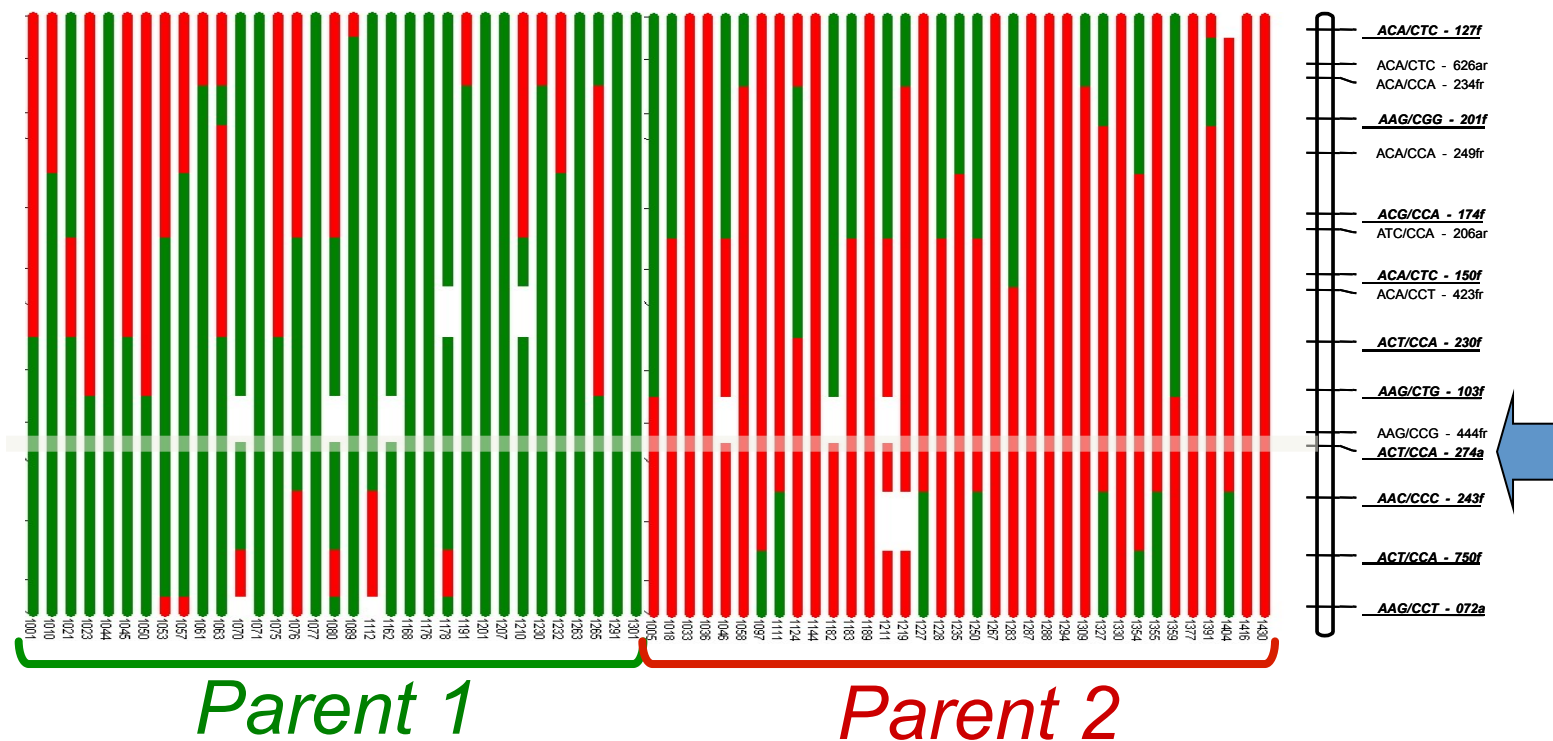
QTL identification:



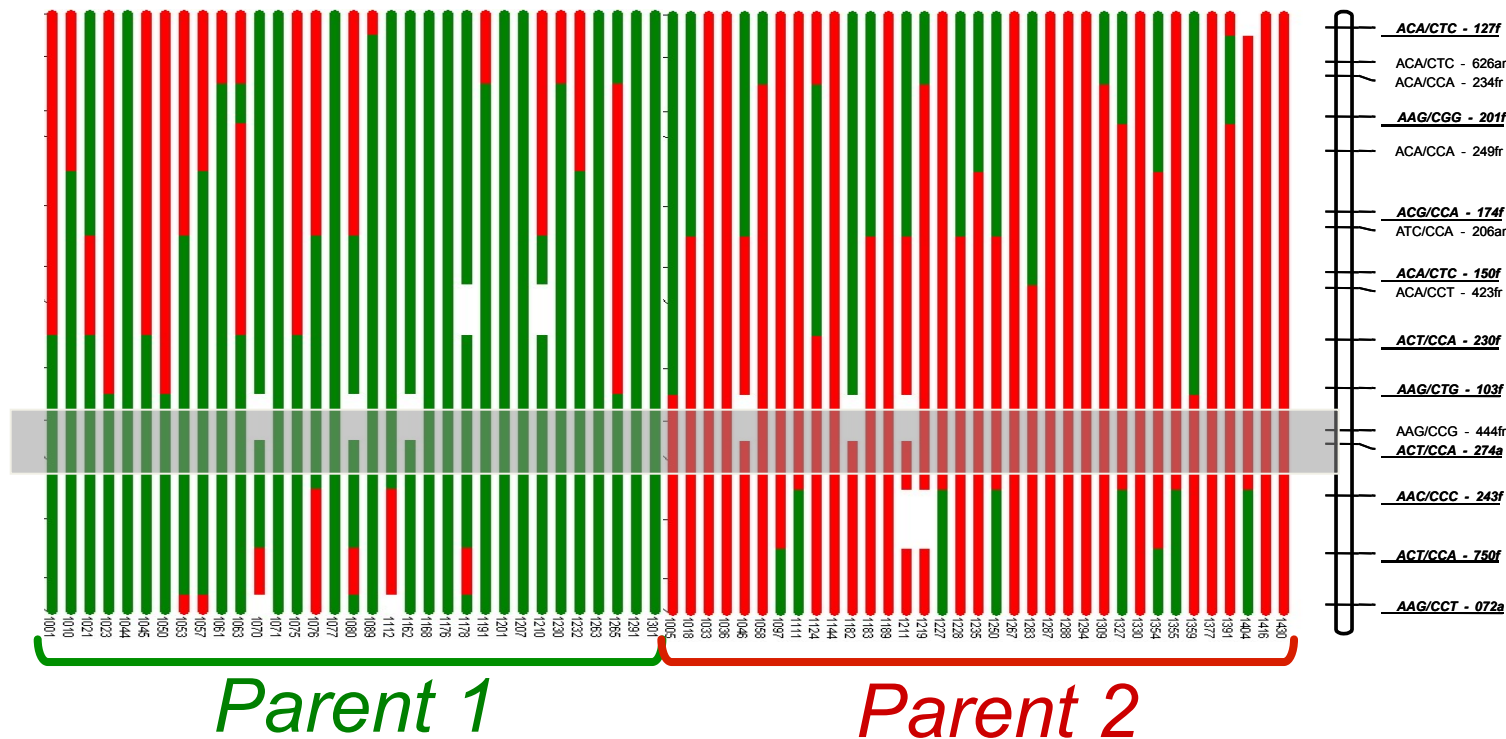
- Phenotype the progeny for trait(s) of interest
- Compare phenotypic means between allelic classes
- If statistically significant, there is evidence for a QTL



**TRAIT
A
NO QTL**



**TRAIT
B
QTL
Present**



- There is evidence that a QTL exists in a well defined genetic region
- The genetic region is broad, and most likely contains hundreds or thousands of genes
- Unless more laborious approaches are taken – i.e. fine-mapping and positional cloning - there is not obvious indication of what genes are in the interval, or are associated with the variation at the quantitative trait

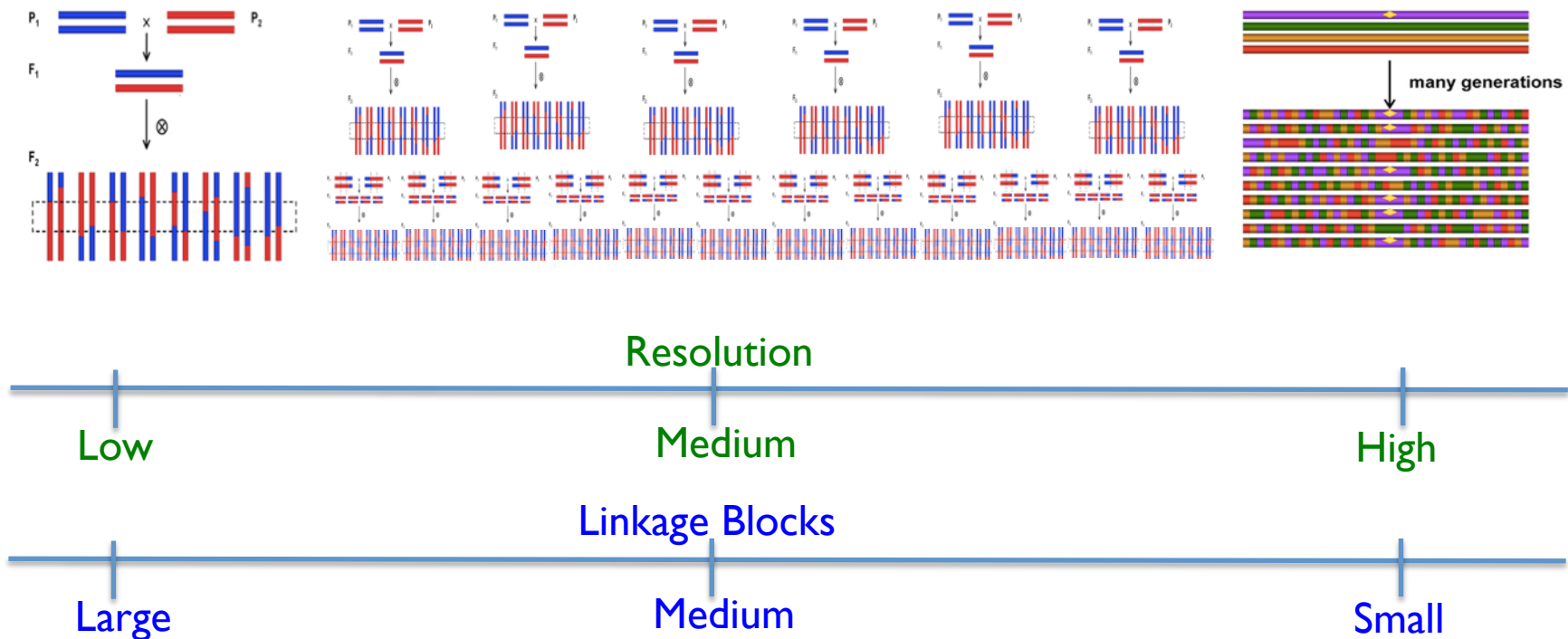
Marker Assisted Selection

Inference on markers based on linkage disequilibrium:

I - QTL analysis

II – Genomic selection

III – GWAS



Linkage Disequilibrium

- Two loci are in **linkage equilibrium**, if they are completely independently in each generation.
- If two genes are in **linkage disequilibrium**, it means that certain alleles of each gene are inherited together more often than would be expected by chance.

SNP1 [A / a]

SNP2 [B / b]

Allele Freq: $p(A)$

$p(B)$

Allele Freq: $p(a)$

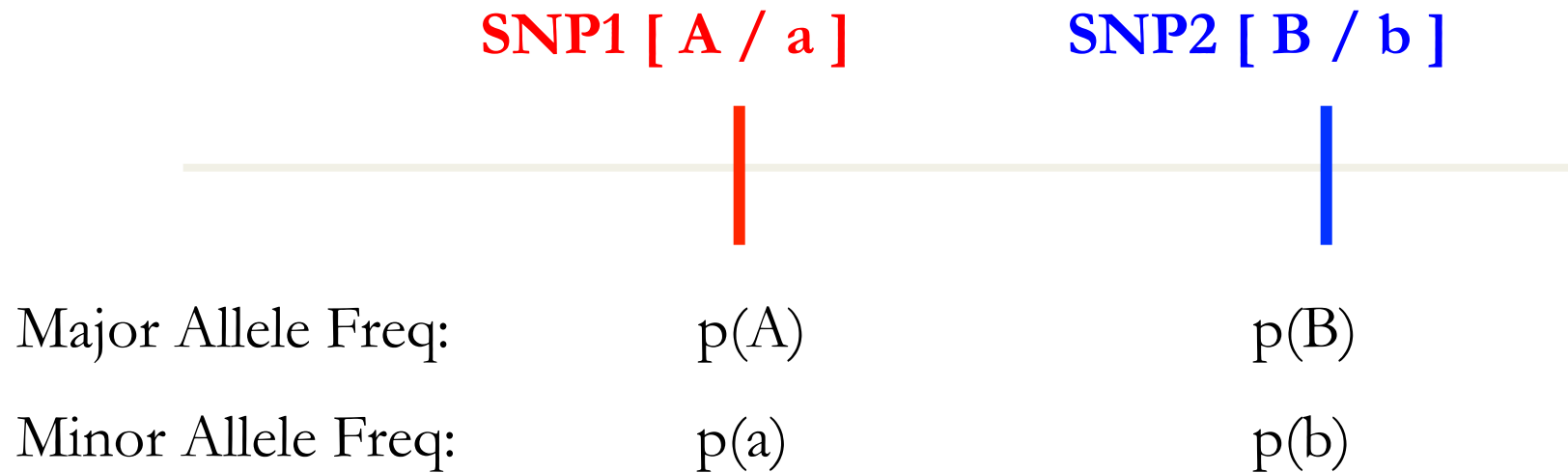
$p(b)$

Independently segregating SNPs:

Haplotype Frequency $p(ab) = p(a) \times p(b)$

LINKAGE DISEQUILIBRIUM

Haplotype Frequency $p(ab) \neq p(a) \times p(b)$



LINKAGE DISEQUILIBRIUM

Haplotype Frequency $p(ab) = p(a) p(b) + D$

(sign of D is generally arbitrary, unless comparing D values between populations or studies)

D : Lewontin's LD Parameter (Lewontin 1960)

D' (Lewontin, 1964)

$$D = P(A_1B_1) \times P(A_2B_2) - P(A_1B_2) \times P(A_2B_1)$$

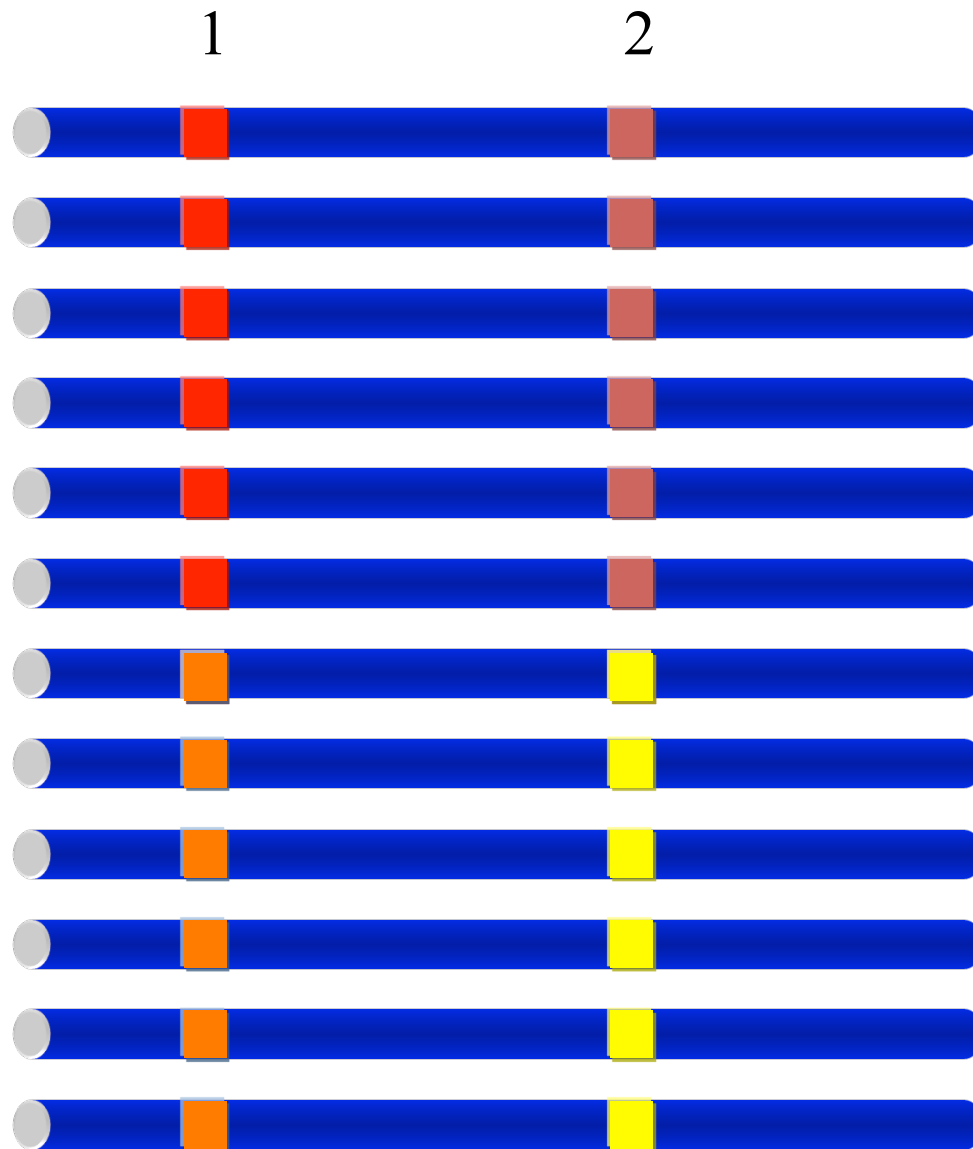
- D is not commonly used due to its difficult interpretation
- Quantifies disequilibrium as the difference between the observed frequency of a haplotype relative to the frequency expected if segregating was random

LD Parameter: r^2 (Hill & Robertson, 1968)

$$r^2 = \frac{D^2}{p(a)p(b)p(A)p(B)}$$

- Squared correlation coefficient varies 0 - 1
- Frequency dependent
- Better LD measure for allele correlation between markers
- Used extensively in disease gene or phenotype mapping through association testing

Complete Linkage Disequilibrium

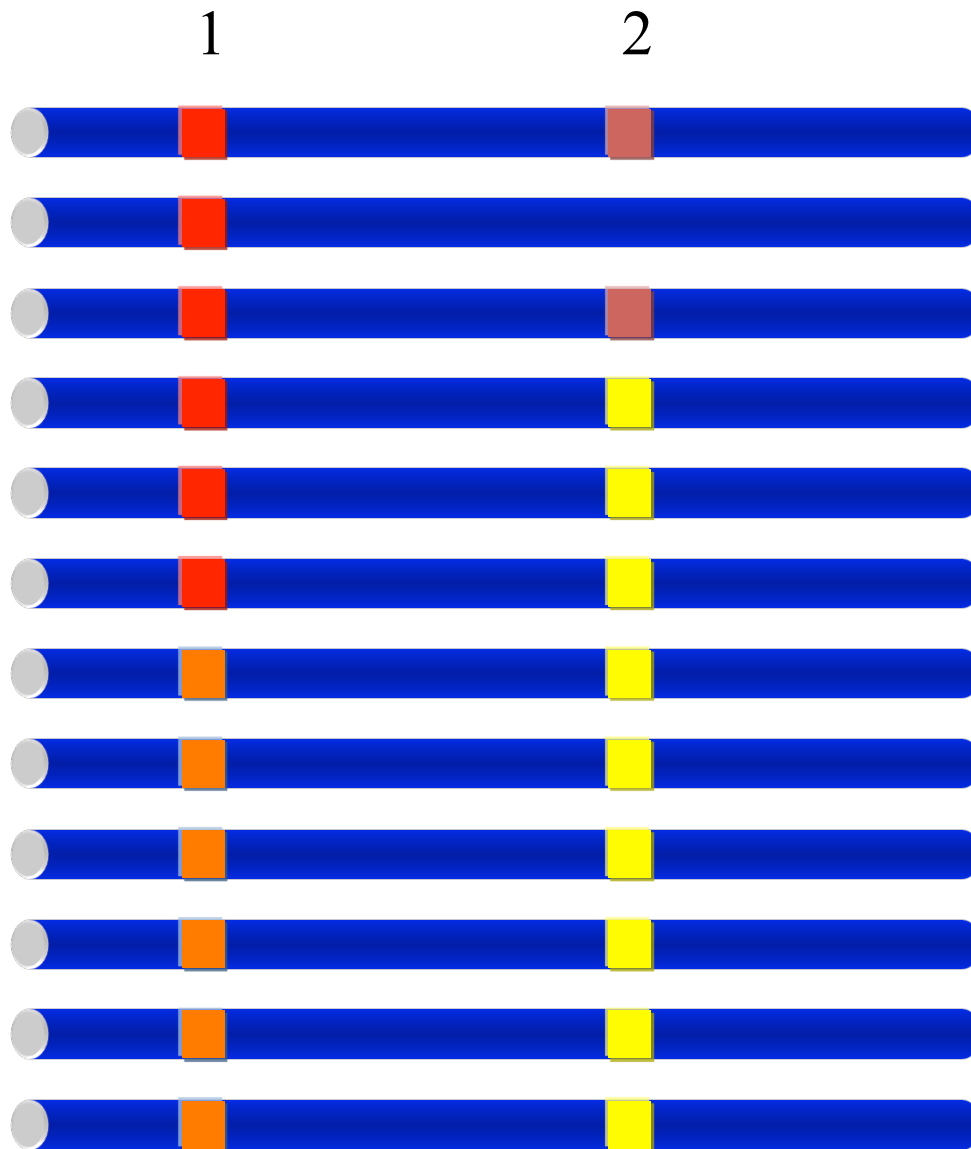






		Locus 1	
		Red	Orange
Locus 2	Red	6	
	Yellow		6

$$D = 0.25$$

$$r^2 = 1$$

Linkage Disequilibrium

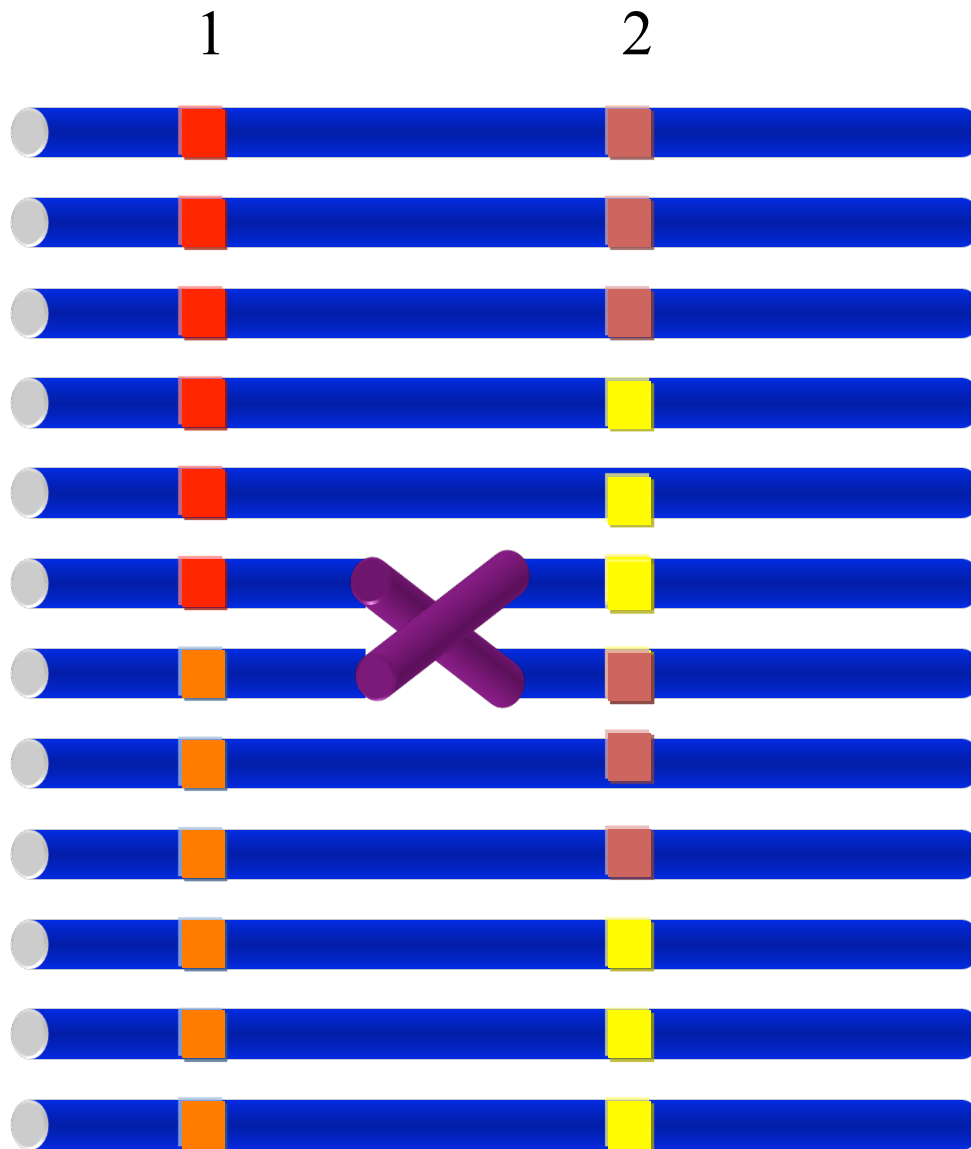


		Locus 1	
			
Locus 2		3	0
		3	6





$$D = 0.125$$

$$r^2 = 0.33$$

Linkage Equilibrium



Locus 1

		
	3	3
	3	3

Locus 2

$$D = 0$$

$$r^2 = 0$$

Why is LD relevant

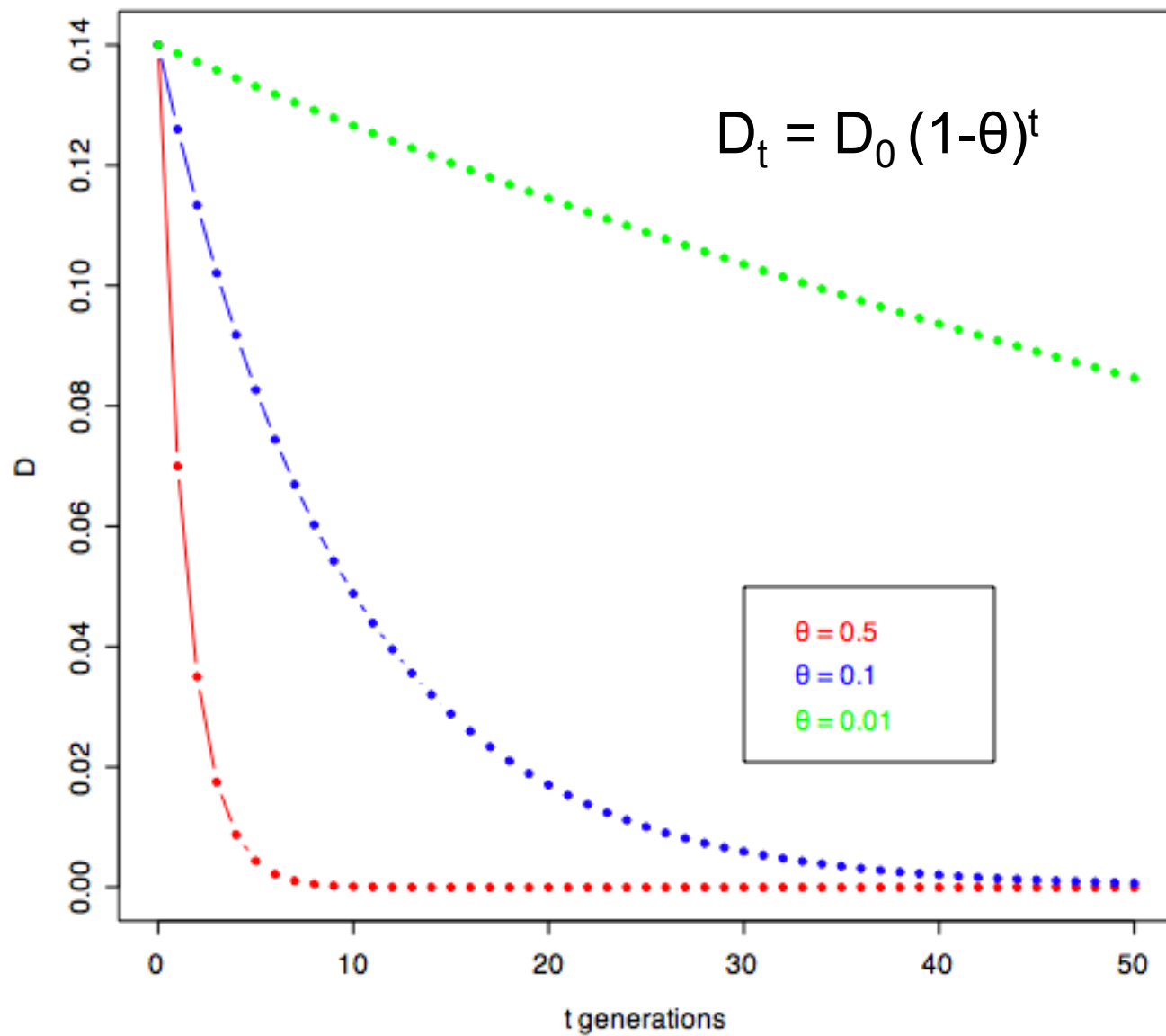
Why is LD relevant

- Measuring the extent of LD helps determine the number of markers needed for QTL mapping, association mapping and Genomic Selection
- When measuring LD we are measuring what would be the expected association between alleles at a marker locus (SNP, microsatellite) and alleles at a QTL
- The measure of LD, r^2 , corresponds to the proportion of the QTL variance that the marker can explain
- Key parameter determining the power of association mapping to detect a QTL

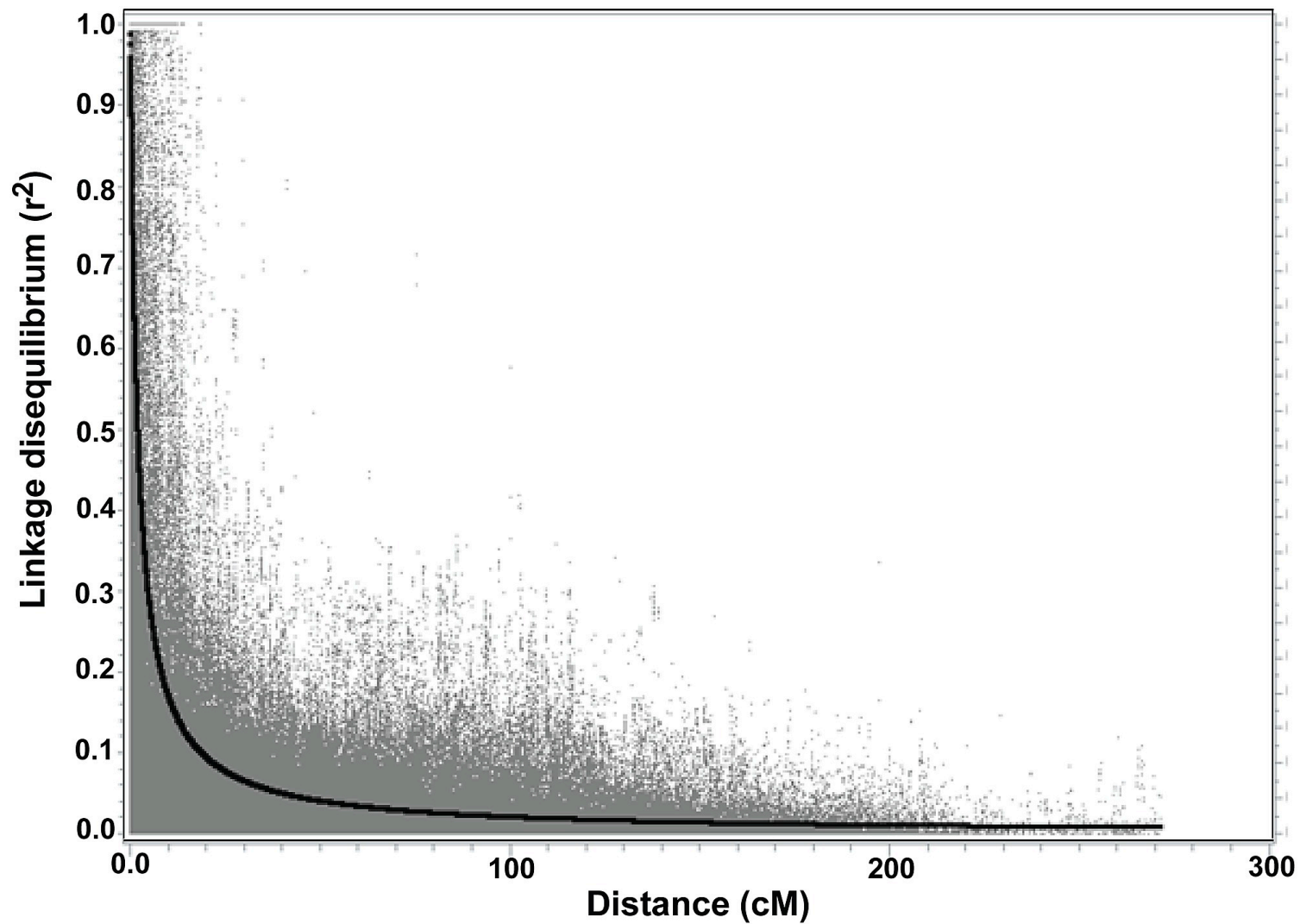
Creation of LD

- Easiest to understand when markers are physically linked
- Creation of LD
 - Mutation
 - Admixture
 - Inbreeding / non-random mating
 - Selection
 - Population bottleneck or stratification
 - Epistatic interaction
- LD can occur between *unlinked* markers
- LD decays over time (generations of interbreeding)
- LD decay is a function of recombination frequency

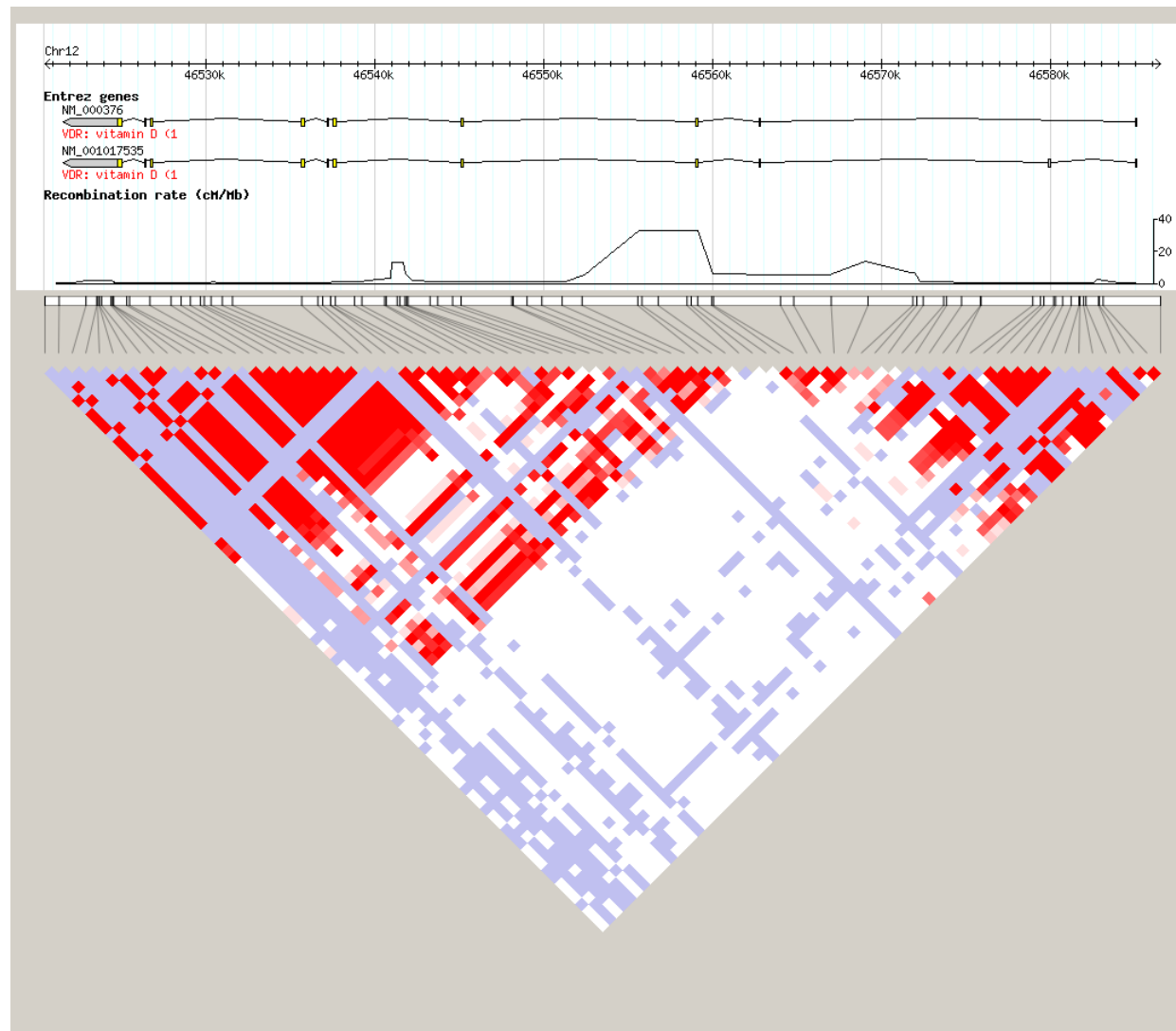
Linkage Disequilibrium (D)



Visualizing LD



Visualizing LD

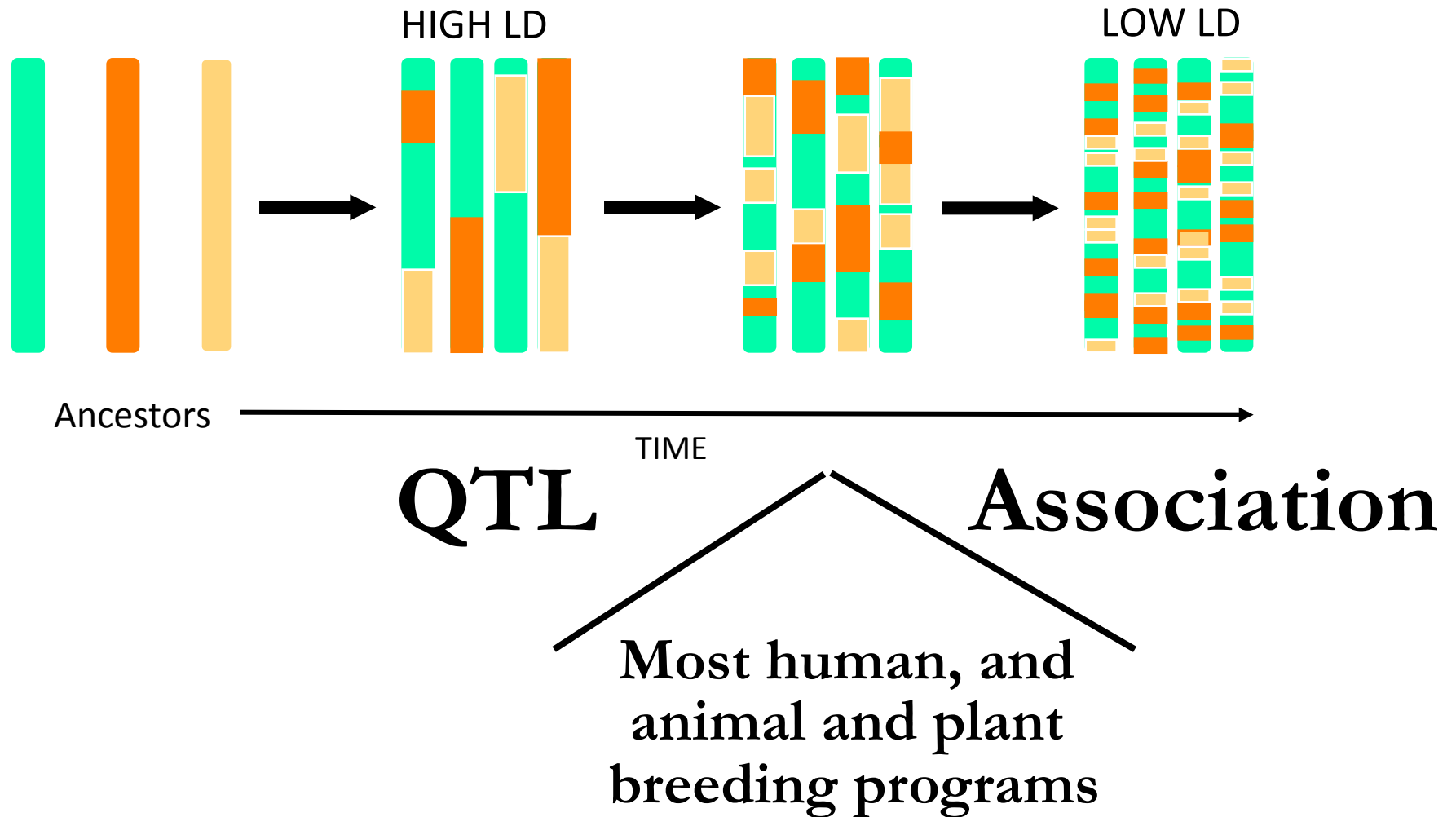


Population Effect on Linkage Disequilibrium in Maize

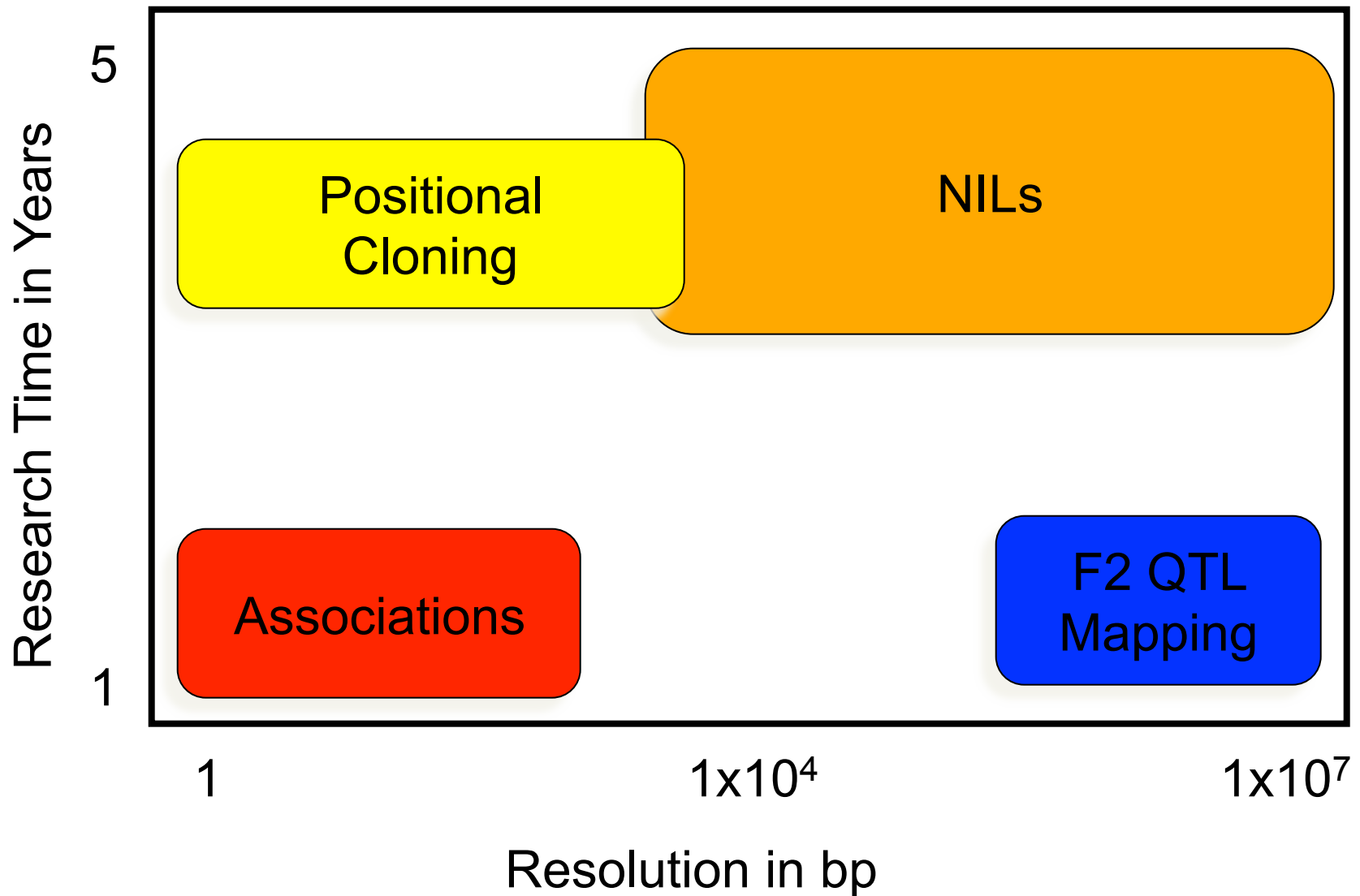
Investigator	Population Studied	Extent of LD
Gaut	Landraces	<1000 bp
Buckler	Diverse Inbreds	2000 bp
Rafalski	Elite Lines	100 kb

Reviewed in Flint-Garcia, S. A. et al. 2003. Annual Review of Plant Biology 54:357-374.

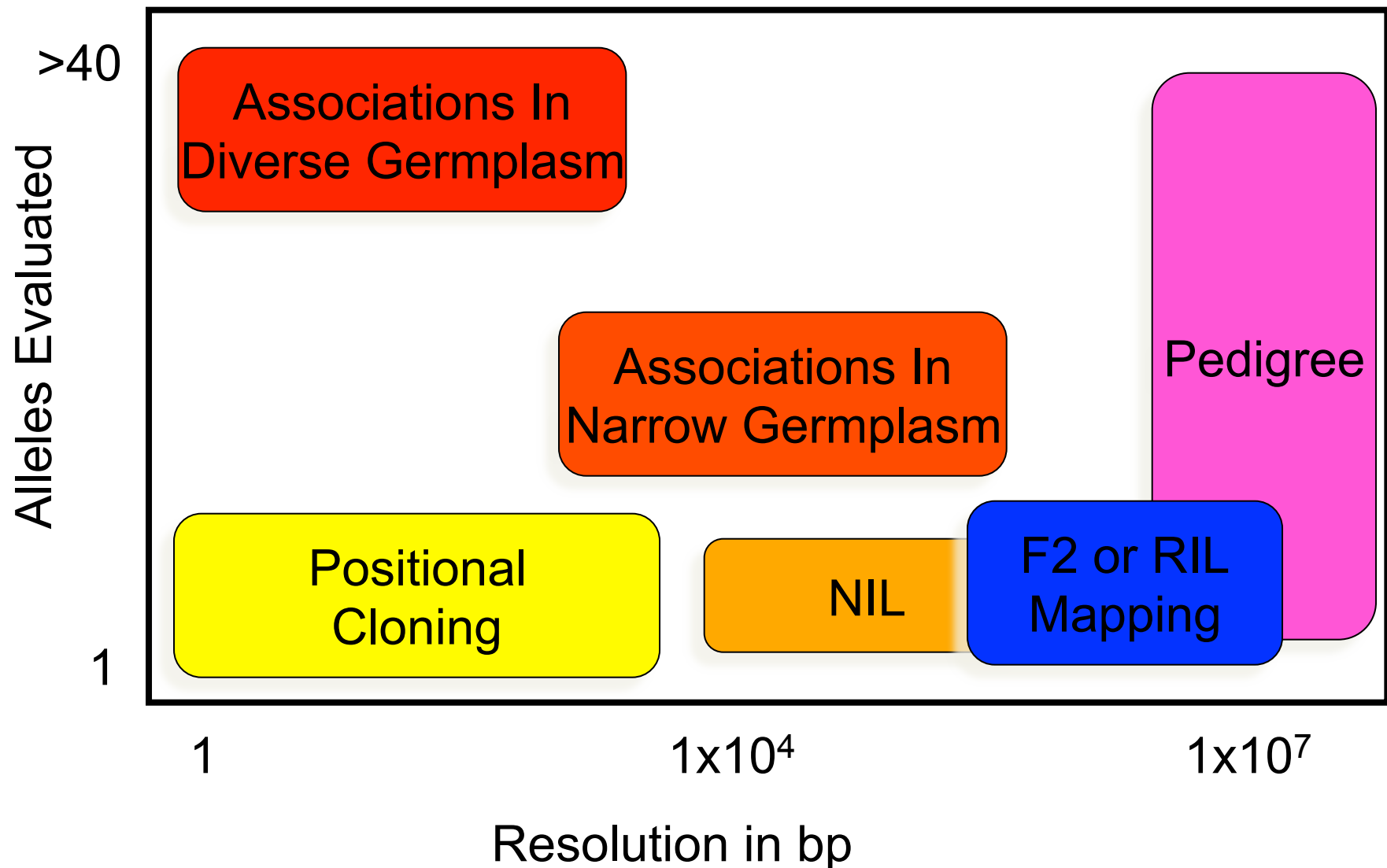
Most significant issue in QTL and association genetics:



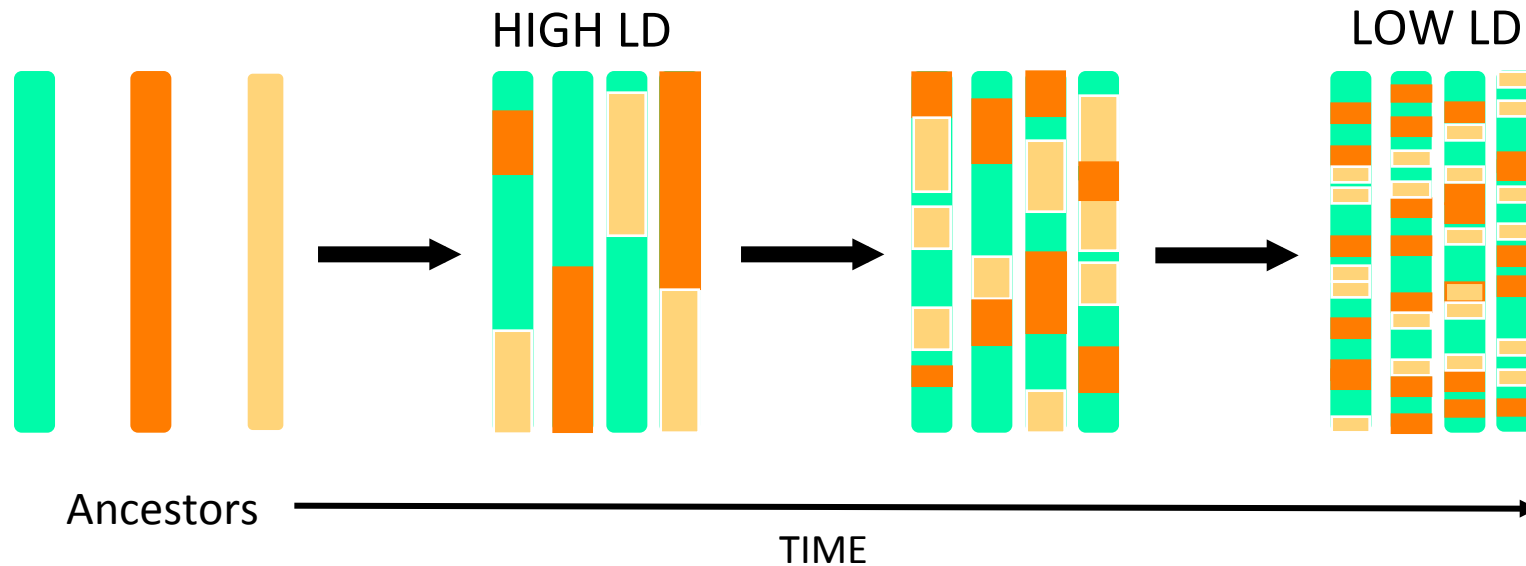
Dissecting A Quantitative Trait: Time Versus Resolution



Resolution Versus Allelic Range



Most significant issue in QTL and association genetics:



Effective population size (N_e): the number of breeding individuals in an idealized population showing similar properties as the population under consideration.

$$E(r^2) = 1/(4N_eL + 1)$$

(Sved equation 1971)