

Molecular marker-based technologies for crop improvement

Gregory G. Brown
Department of Biology
McGill University

1st GCN Workshop
July 10-11, 2008
Montreal



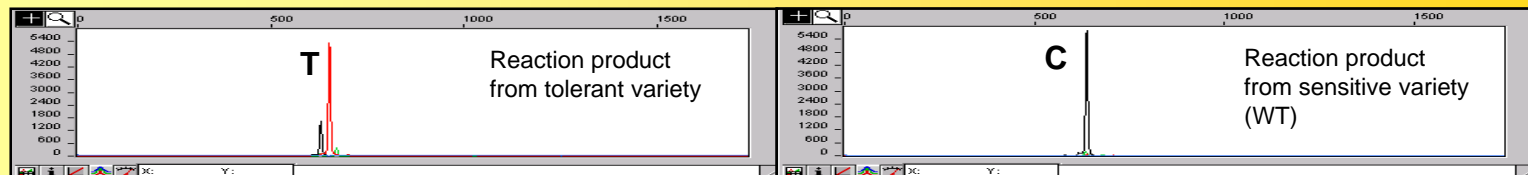
- Molecular markers allow DNA polymorphisms – DNA sequence differences between individuals in a population (or species) – to be detected
- Marker-based technologies – allow us to more effectively exploit natural variation in crop plants
- Bridge the gap between traditional breeding and the “post-genomic world”
- Non-biased gene discovery – “forward genetics”

Single Nucleotide Polymorphism (SNP)

- **SNP** markers exploit single nucleotide variations in the DNAs of individuals.
- The example below illustrates the detection of a SNP in a canola gene that is associated with herbicide tolerance.
- In this case, the two alleles are differentiated through the incorporation of differently labeled chain terminators (C for wild type or WT, T for herbicide tolerant at the 3' of a specific primer).

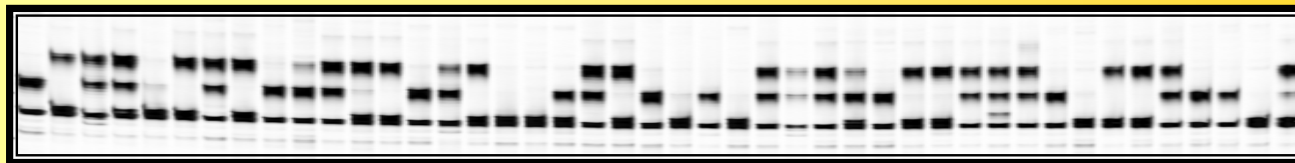
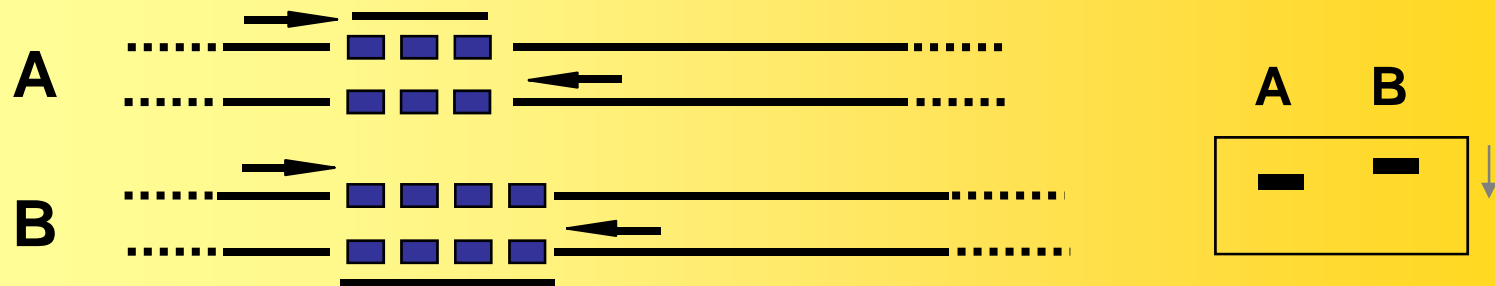
```
sensitive      Ile Pro Ser Gly Gly Thr Phe Lys Asp Val
                ATC CCA AGT GGT GGC ACT TTC AAA GAT GTA
                3' CA CCA CCG TGA AAG TTT CTA C 5' PM1 low primer

tolerant       Ile Pro Asn Gly Gly Thr Phe Lys Asp Val
                ATC CCA AAT GGT GGC ACT TTC AAA GAT GTA
                3' TA CCA CCG TGA AAG TTT CTA C 5' PM1 low primer
```

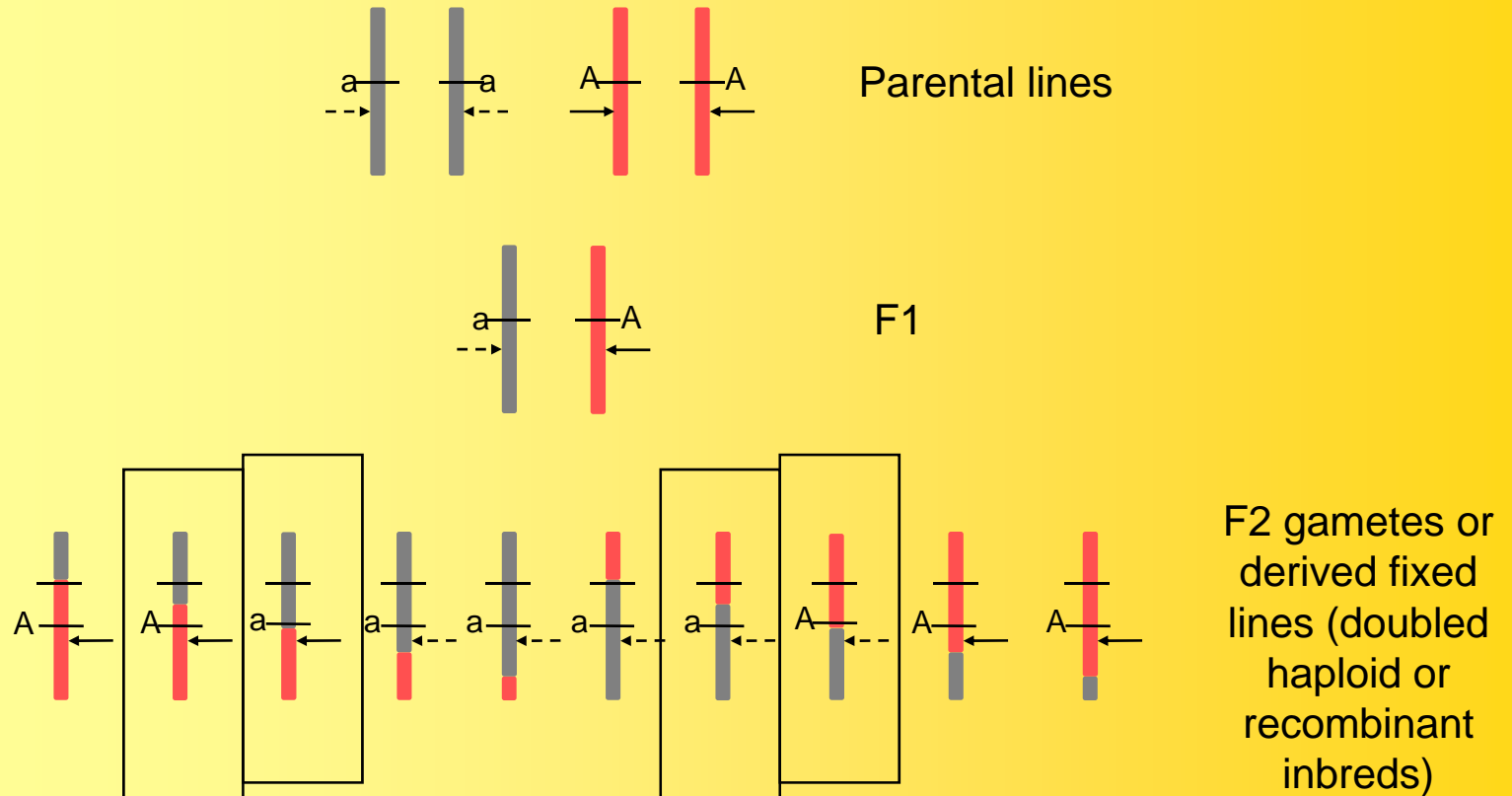


Simple sequence repeat (SSR)

- **SSR or Microsatellite** polymorphism represents variation in the number of repeat units of di-, tri- and tetranucleotides (blue boxes)
- The polymorphism is expressed as different-sized PCR products generated from primers complementary to sequences flanking the repeat.
- An example of the genetic segregation of an SSR marker in a cross between different wheat varieties is illustrated.

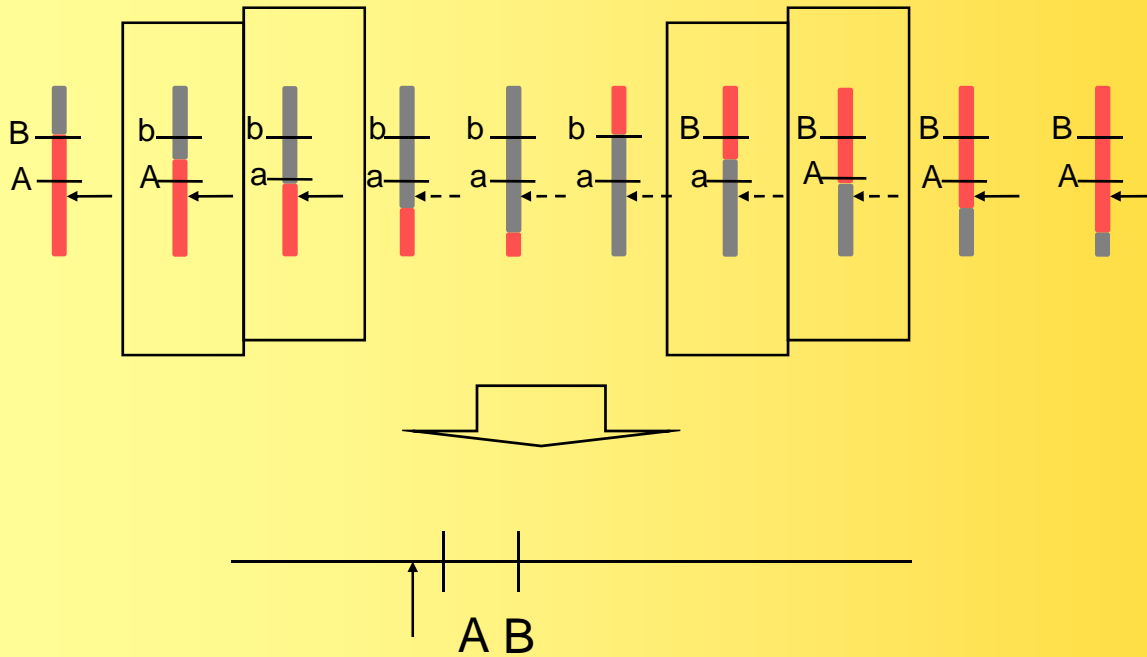


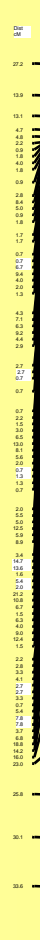
Genetic mapping



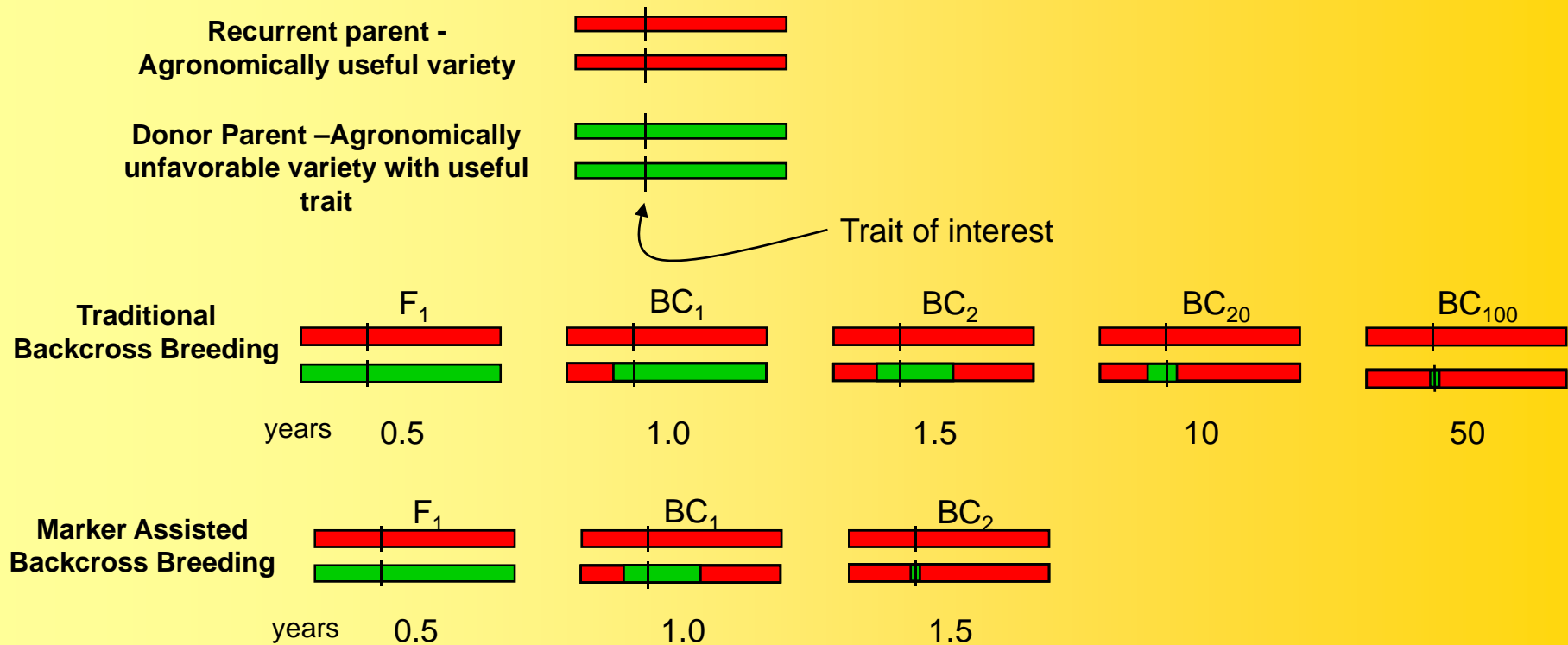
Frequency of recombination events between marker and trait = genetic distance (cM)

Genetic mapping

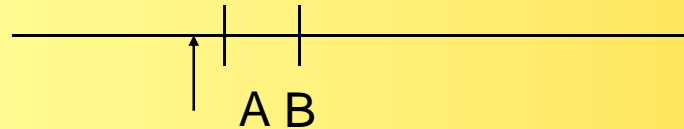




Marker assisted backcrossing



Markers as tools for plant genome analysis and gene discovery

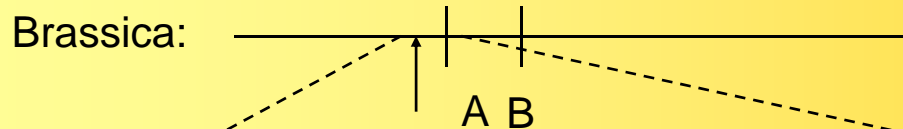


Map resolution increased by more markers and larger populations

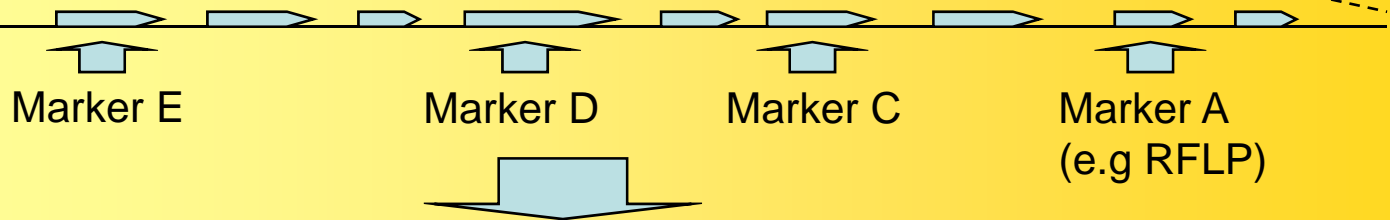
Different plant genomes share extended regions of organizational co-linearity (synteny)



Sequenced plant genomes provide a key information source for finding markers in specific genomic region



Arabidopsis

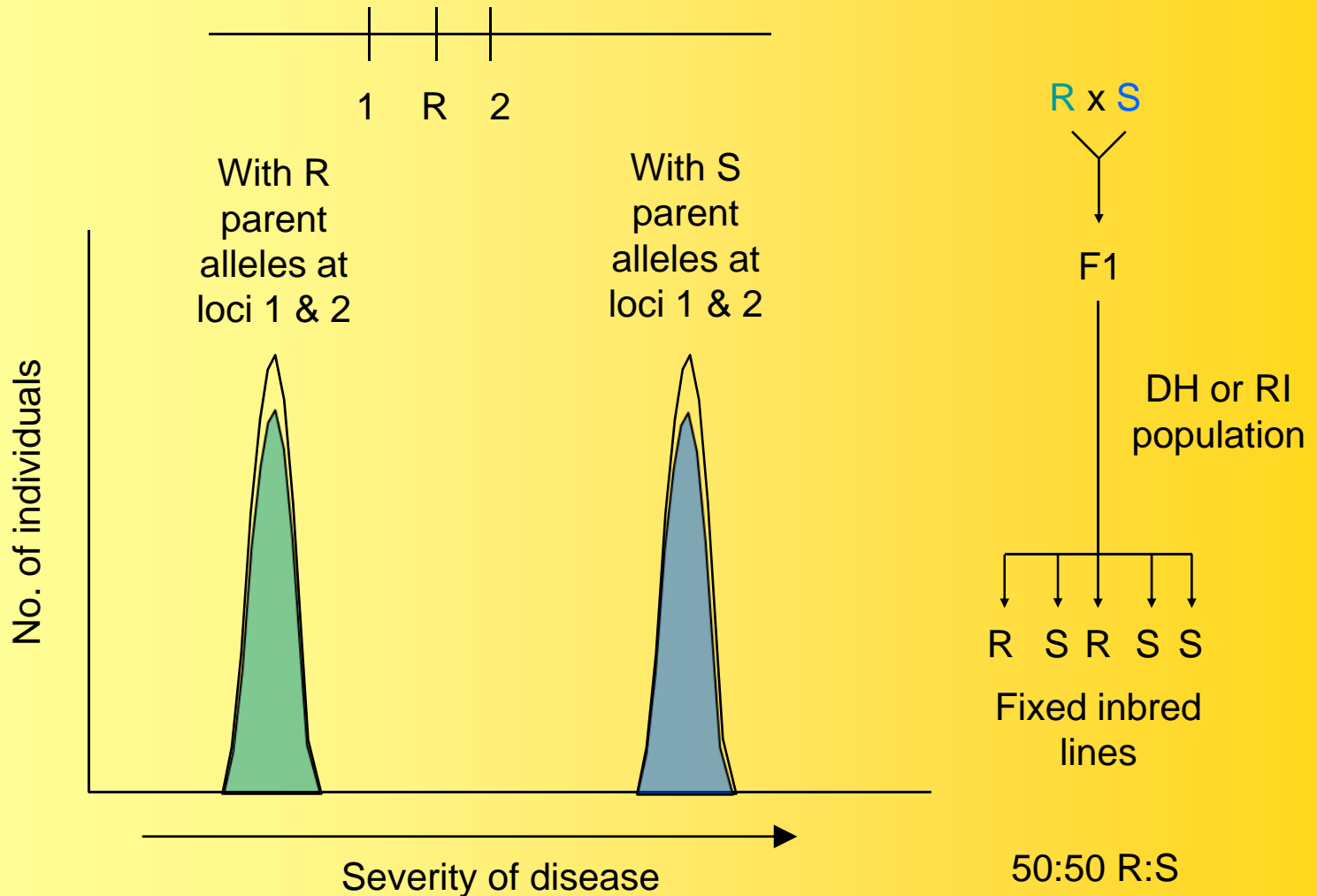


Using Markers to Analyze Complex Traits

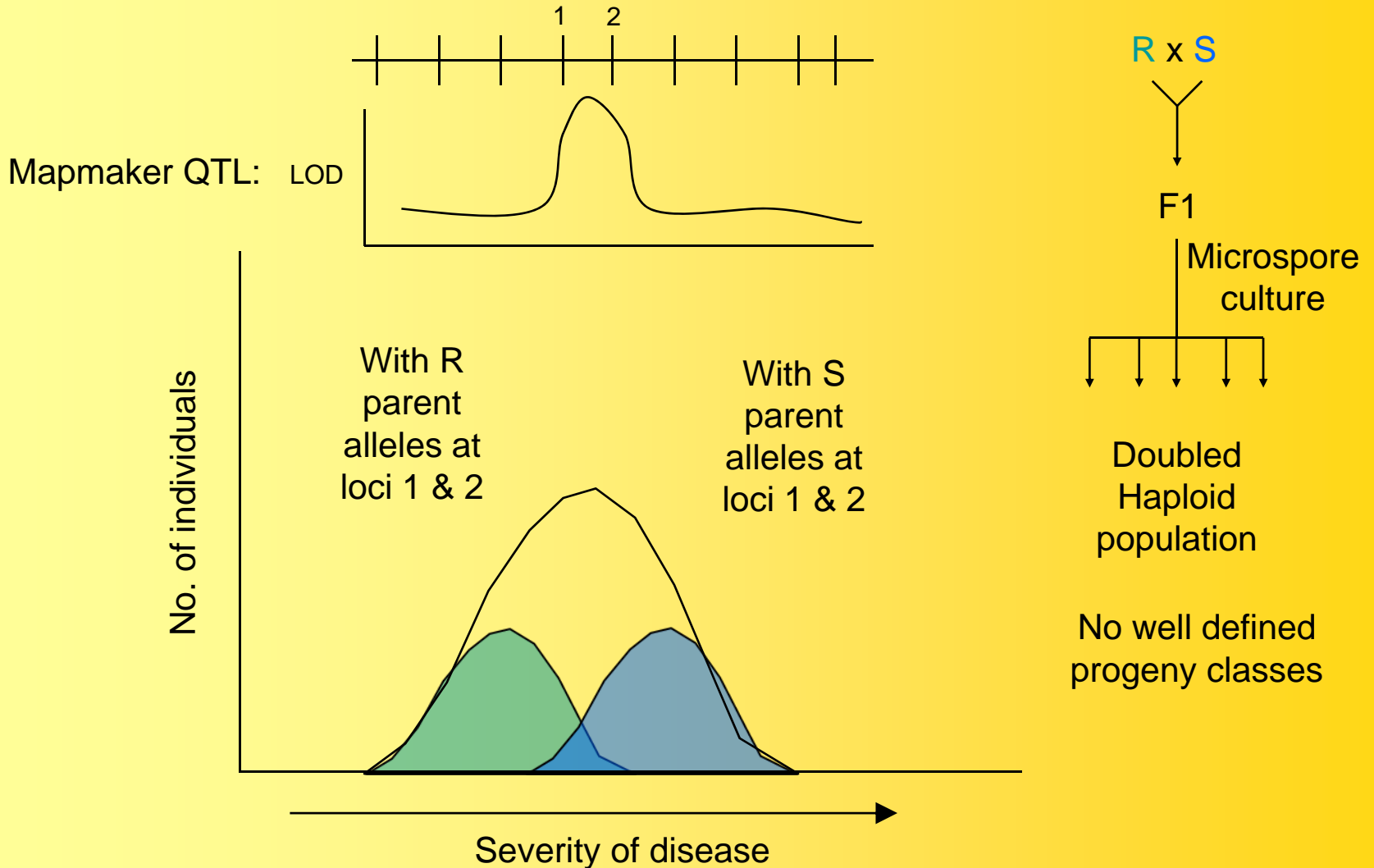
Quantitative Trait Loci (QTL)

- A genetic interval containing a gene that contributes to a complex, quantitative trait
- Most important agronomic characters are quantitative
- Both positive (e.g. nutrient, water use efficiency) and negative (e.g.lodging) traits are quantitative
- Number of QTL detected will vary depending on trait, species, genetic cross

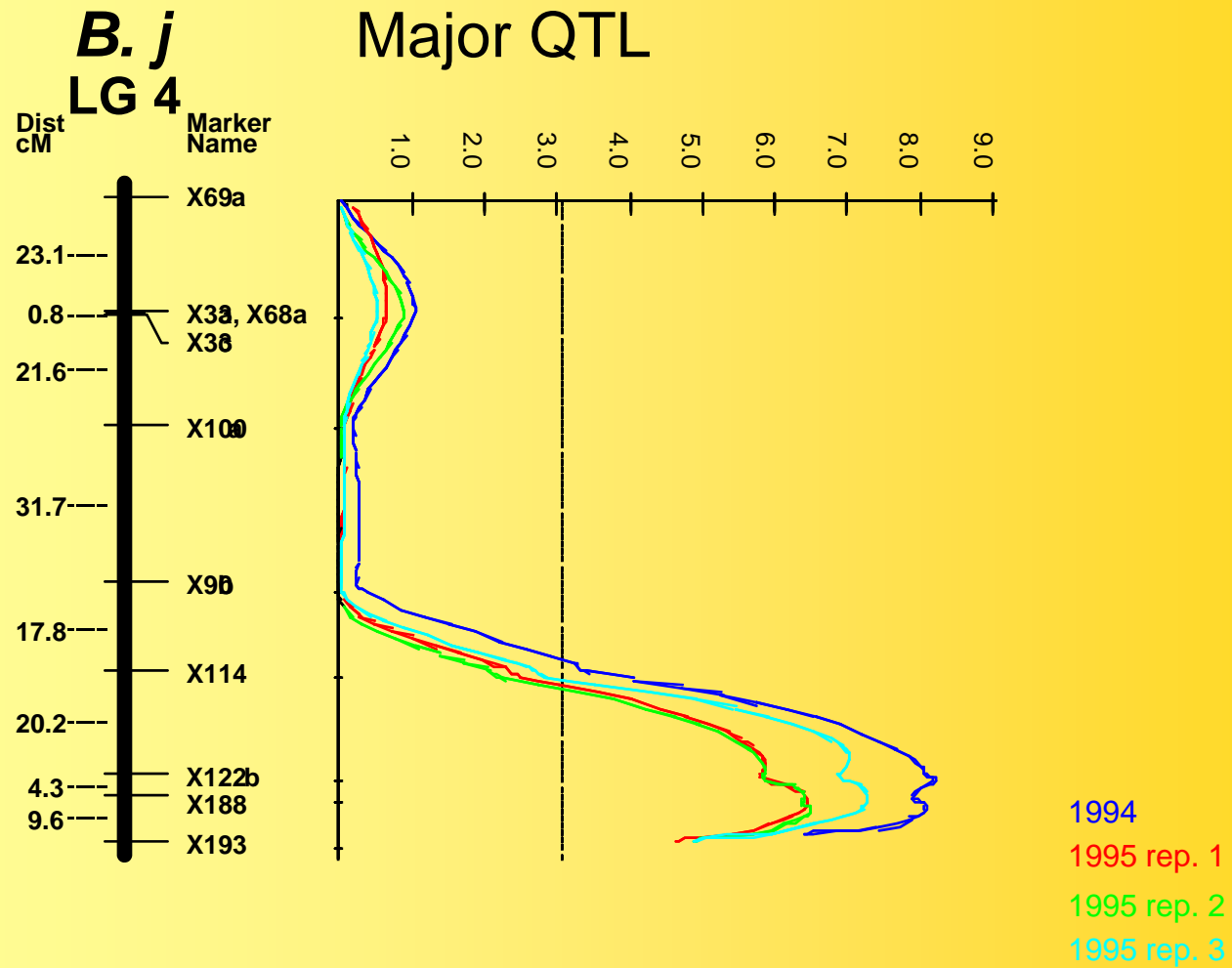
A simple (binary) trait



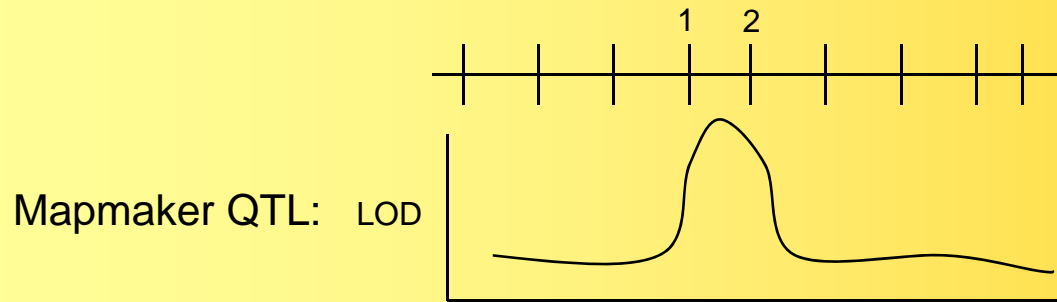
A complex (quantitative) trait



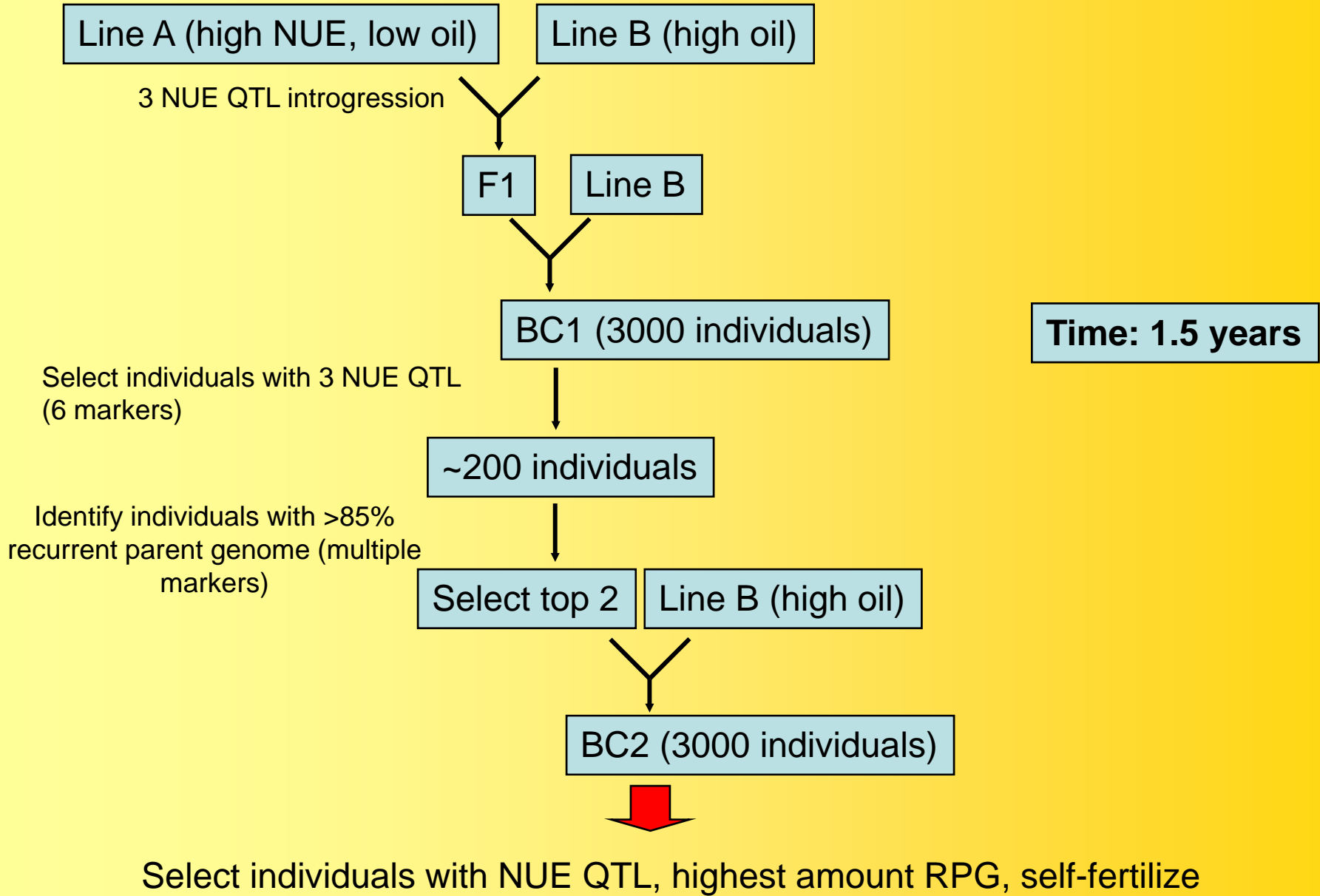
QTL analysis



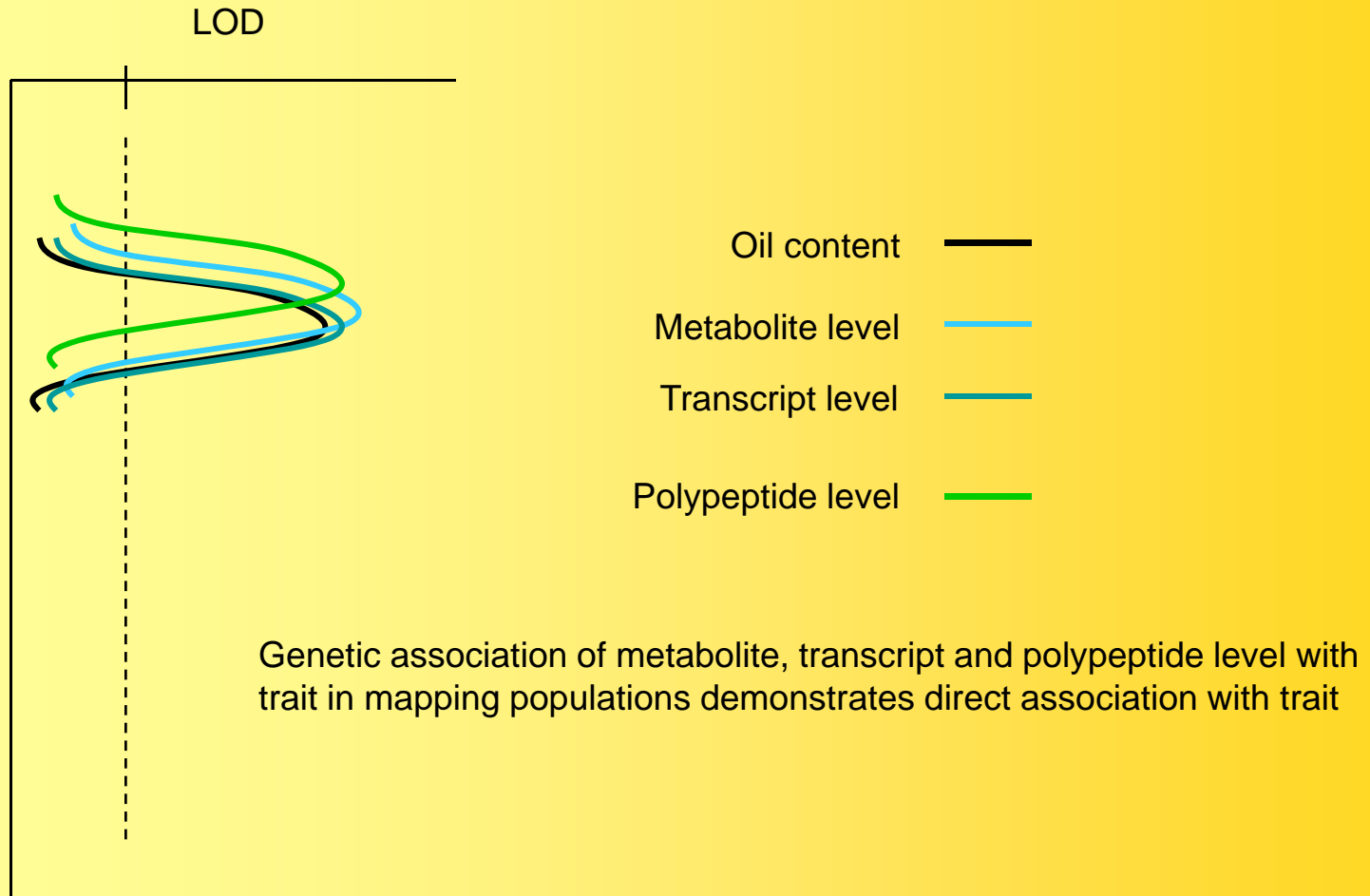
QTL can be used directly to speed-up breeding for complex traits



- Use markers 1 and 2 to select for QTL
- Use markers spaced throughout genome to select for individuals with highest RPG

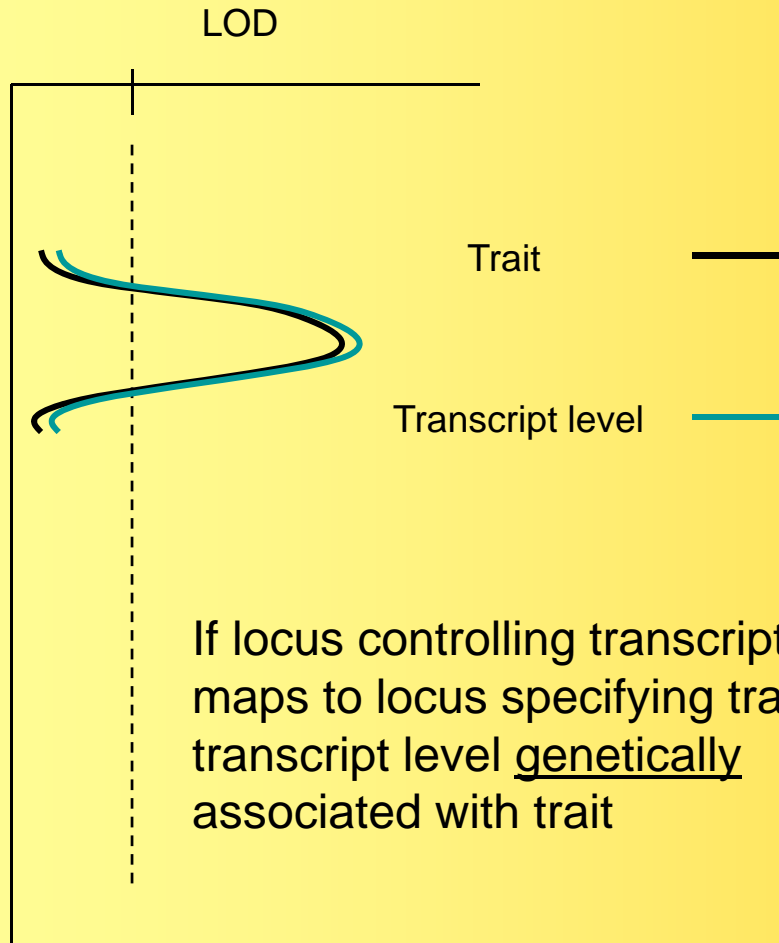


QTL analysis



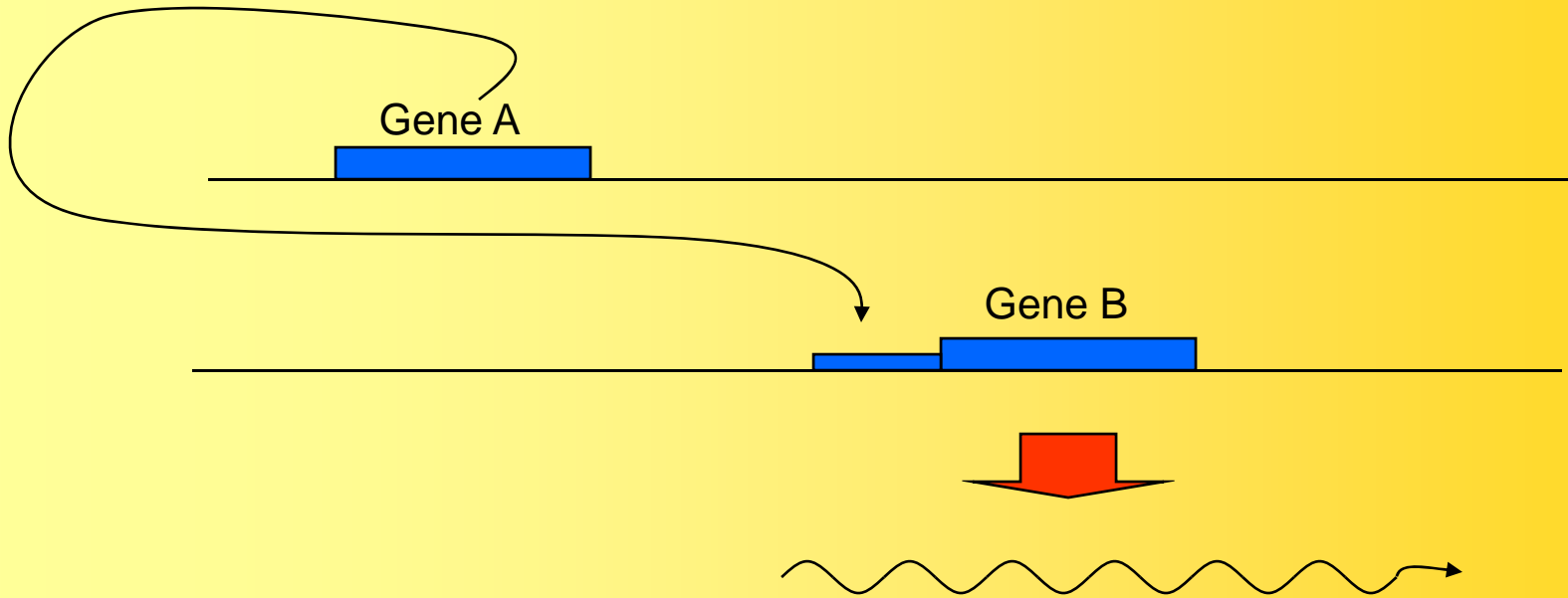
eQTL analysis

eQTL: transcript level = quantitative trait

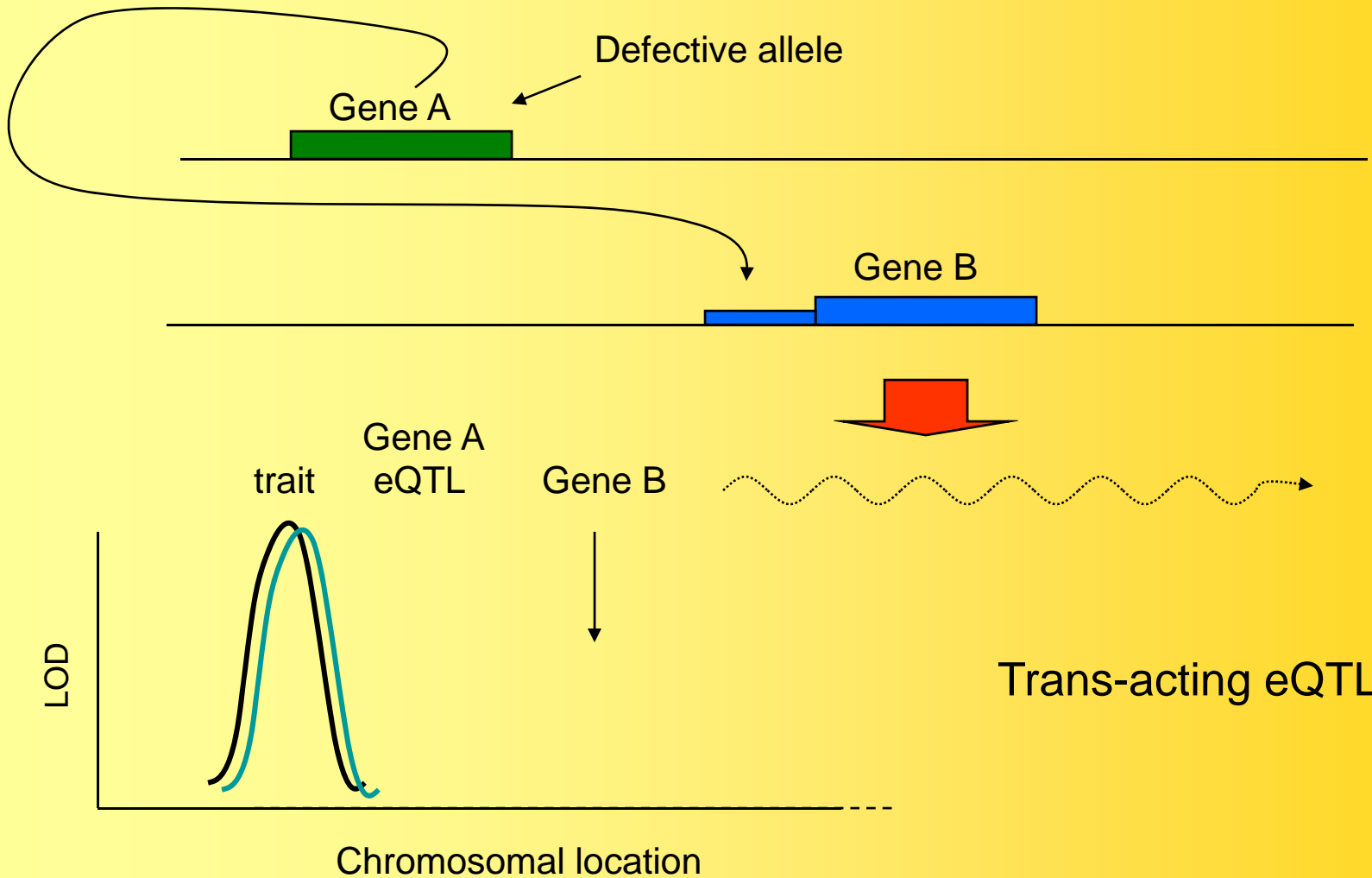


If locus controlling transcript maps to locus specifying trait, transcript level genetically associated with trait

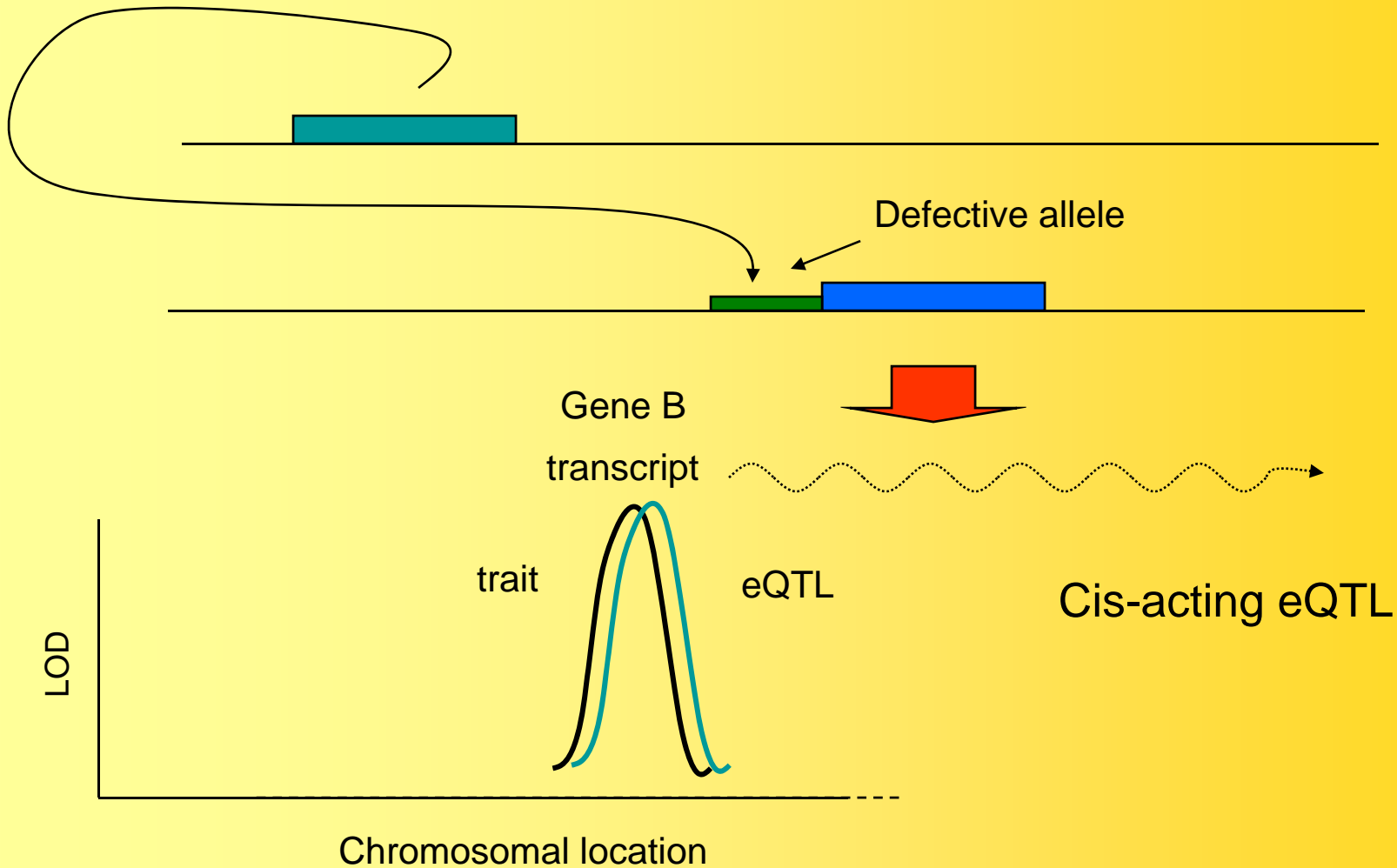
A simple gene regulatory system



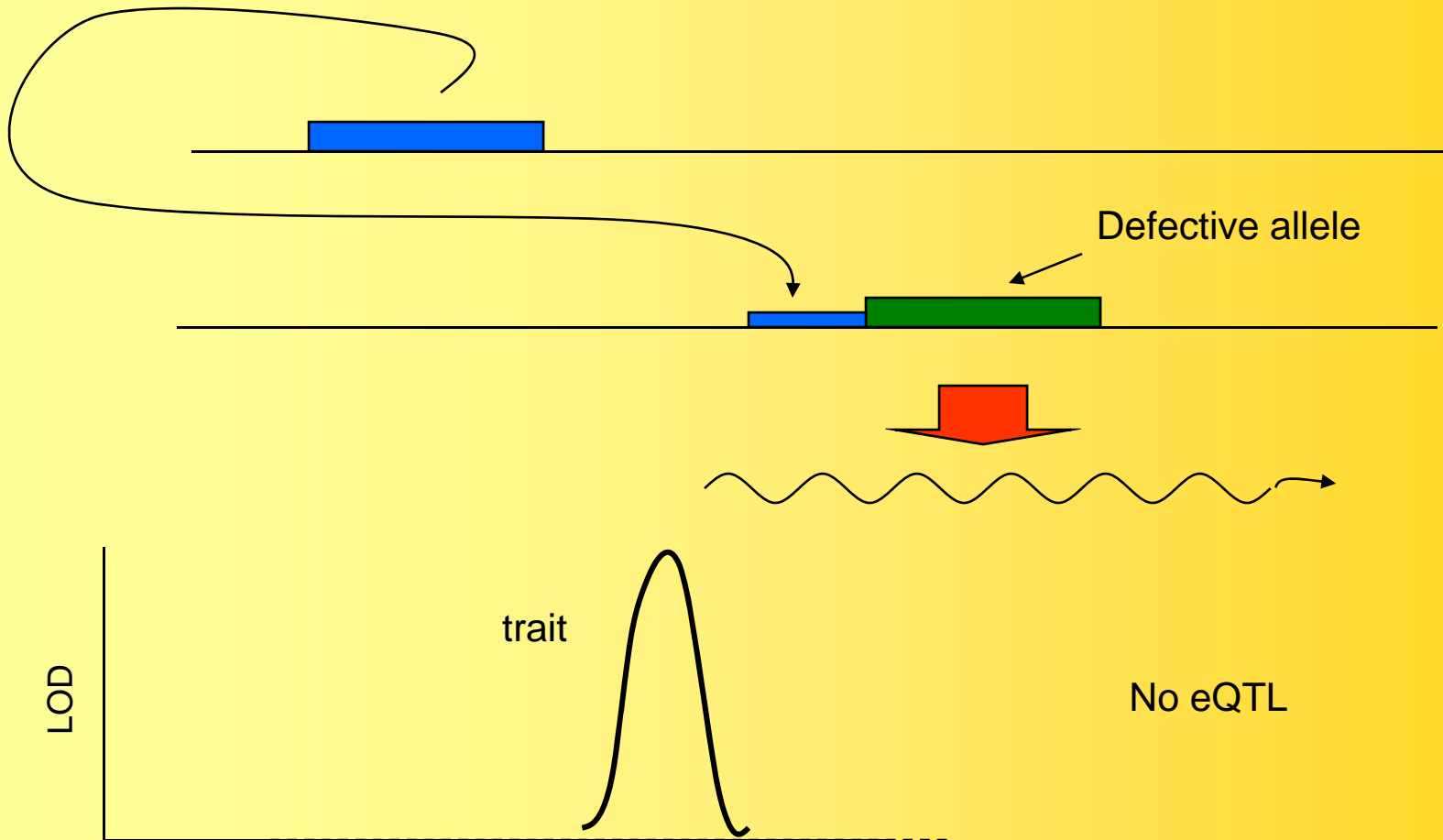
A simple gene regulatory system



A simple gene regulatory system

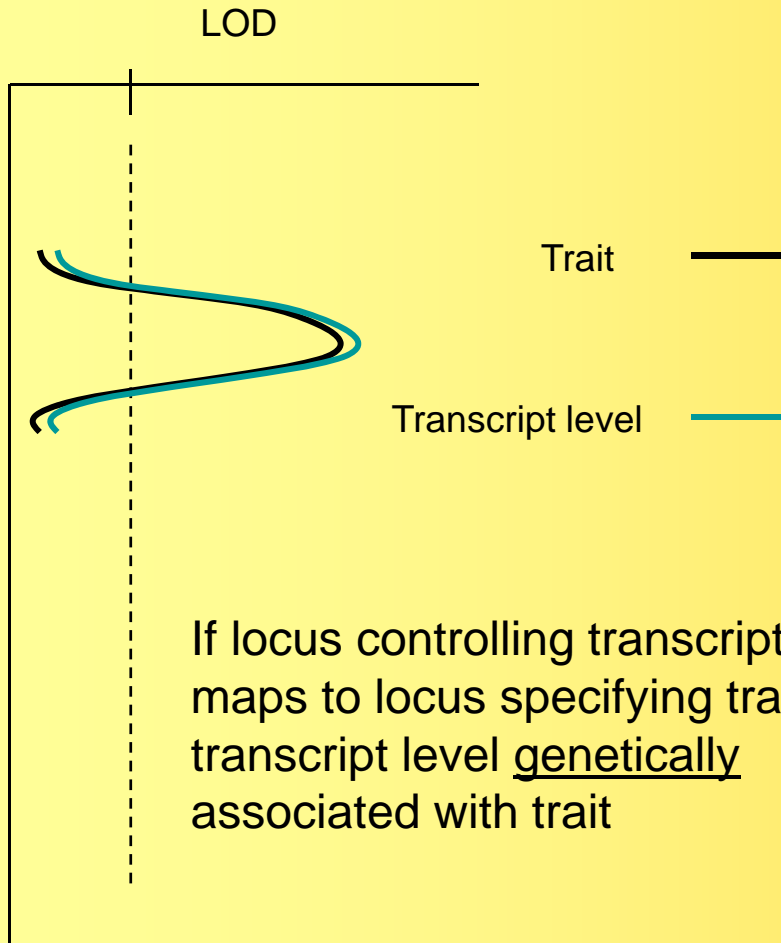


A simple gene regulatory system



eQTL analysis

eQTL: transcript level = quantitative trait



If locus controlling transcript maps to locus specifying trait, transcript level genetically associated with trait

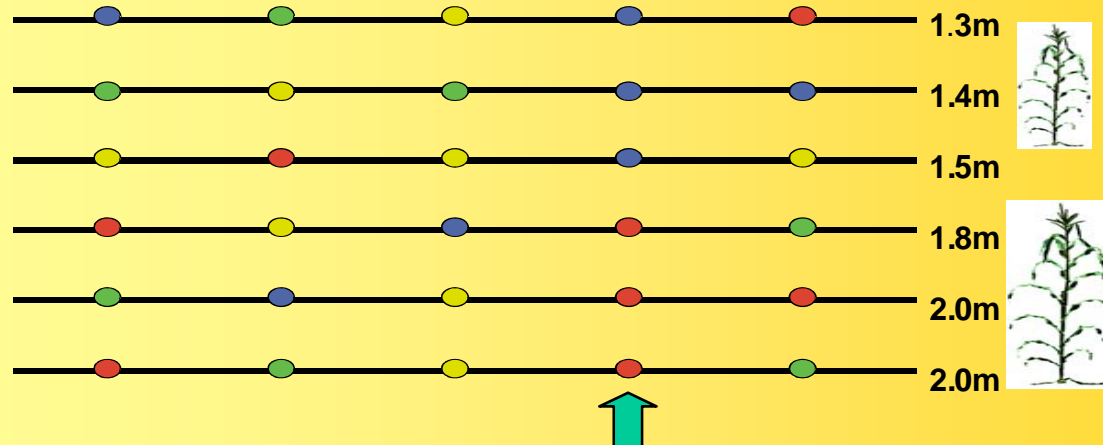
If map location of eQTL different than location of transcript's gene, eQTL is trans-acting

If map location of eQTL is same as location of transcript's gene, then eQTL is cis-acting (trait likely specified by gene encoding transcript)

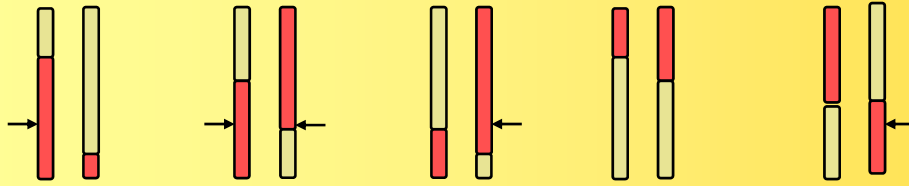
Association mapping

Principle:

- Evaluate whether DNA marker / nucleotide polymorphisms associate with phenotype
- Natural populations



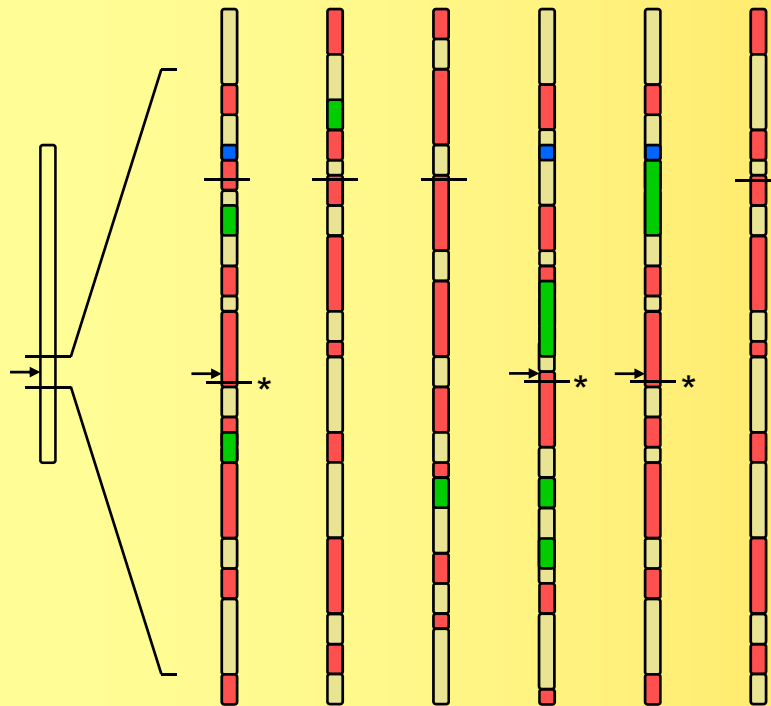
Adapted from Ed Buckler



F2 individuals



Many rounds of random mating



+ admixture

+ mutation

Tightly linked alleles may show consistent association with phenotype

Alleles at unlinked loci do not

The more tightly linked two genes or markers are, the more frequently they will be found together within individuals

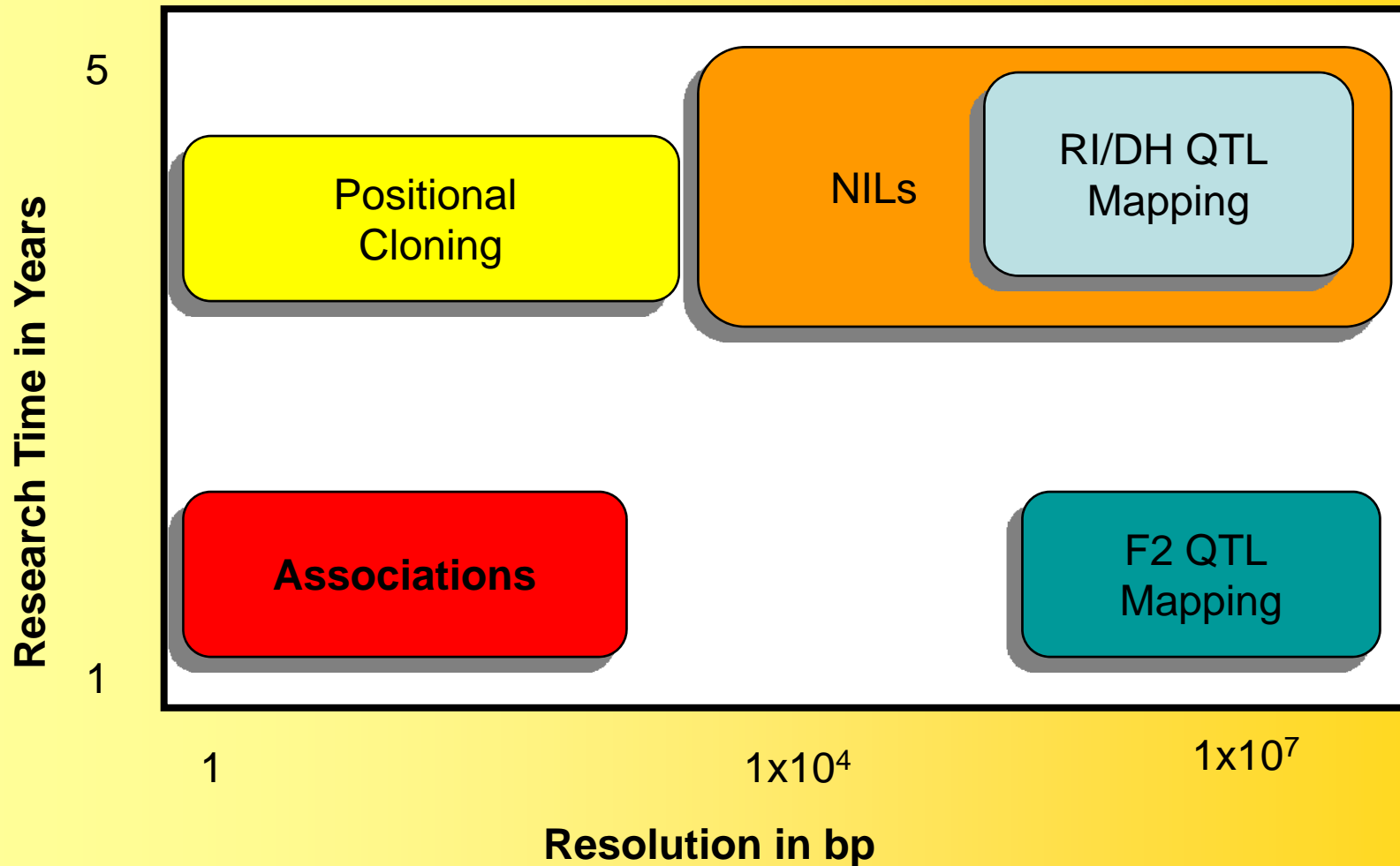
Linkage Disequilibrium

- The non-random association of alleles at different loci
- “The correlation between polymorphisms caused by their shared history of mutation and recombination”
- Linkage, selection, mutation, introduction of new genotypes all increase levels of LD

Population Effect on Linkage Disequilibrium in Maize

Investigator	Population Studied	Extent of LD
Gaut	Landraces	<500 bp
Buckler	Diverse Inbreds	1500 bp
Rafalski	Elite Lines	100 kb? (6 kb euchromatin?)

Dissecting A Quantitative Trait: Time vs Resolution



Association mapping

- Competitive advantages of association mapping:
 - Fast, economical lead gene identification from list of candidates
 - Variable range in mapping resolution depending on populations employed – allows for range of applications
 - Genome-wide scans can allow gene identification when only the target phenotype is known
- Limitations of association mapping:
 - Results can be confounded by population structure

Thank you

Acknowledgements:
DNA LandMarks, Inc.
A BASF Plant Science Company