PhD Credit Seminar

ASSOCIATION MAPPING AND ITS ROLE IN CROP IMPROVEMENT

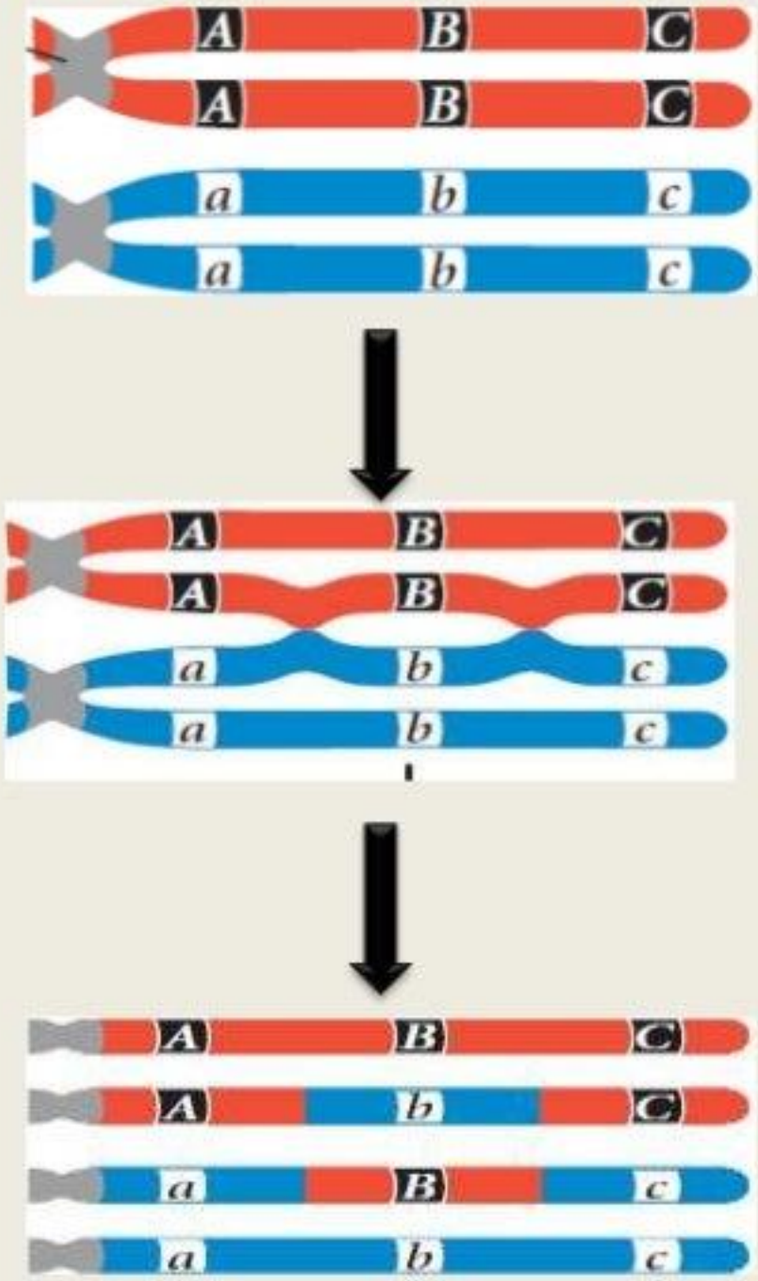
Bringing fresh air to mapping....

Presented by

WASEEM HUSSAIN
A-2011-40-006
Ph.D Student
(Crop Improvement)

Genetic mapping

Genes/ markers in order, indicating the relative genetic distances between them, and assigning them to their chromosome.



In 1913, the first individual to construct a (very small) genetic map was Alfred Sturtevant….: "

Distance = Recombination frequency=
No. of recombinants /Total progeny X 100

Suppose the recombination between loci A and B is 6%, that between loci Band C is 20%, and that between A and C 24%, then we can order the loci along the chromosome as…

Hartal and Jones 2010

Gene mapping by markers!

Easily detectable at genotypic level

The principles of mapping and linkage analyses remains the same way…..

Genes (marker or loci) segregarte via chromosome during meiosis thus allowing their analysis in progeny

- **Should be highly heritable**
- **Should be linked to gene of interest**
- **Must co-segregrate with the gene**

(Collard et al 2005 and 2008)

Why mapping genes!!!

Find molecular markers linked to target genesand then used In MAS

the markers can be used to transfer the gene of interest from a donor line to the target genotype

Two main approaches

Family-based Linkage mapping

Constructing linkage maps and conducting QTL analysis– to identify genomic regions associated with traits–is known as QTL mapping (also genetic,' 'gene' or 'genome' mapping/ gene tagging/ Family based mapping)

LD-based Association mapping

?

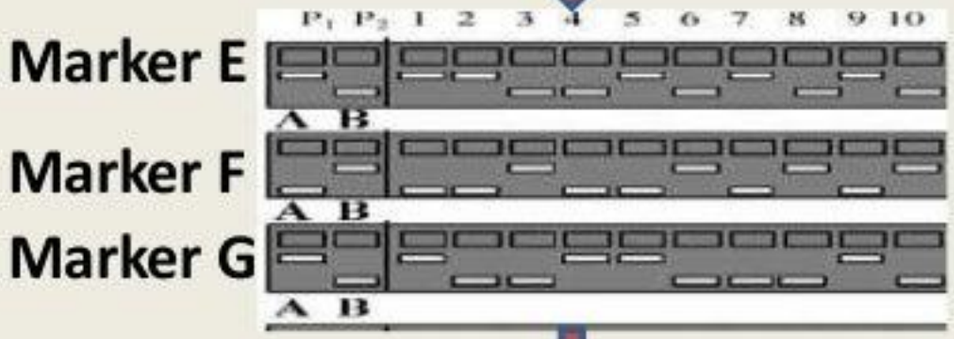
Mackay & Powell 2007

P_1 P_2

Mapping Population

DNA EXTRACTION AND PCR

Electrophoresis segregation analysis



RAPD

Marker E	AABBABABABB
Marker F	ABABBABBABB
Marker G	AABBAABBBAB

Genotypic data set

MAPMAKER/JOINMAP/ GMendel

Linkage map

Why QTL/ Family based mapping unsuccessful !!!!!!!!!

More than 11,000 articles published on QTL mapping in different species..... more than 360 articles relate to reports of over 1000 QTLs associated with various traits in maize....Despite....to date only a few QTLs have been commercially exploited (http://www.ncbi.nlm.nih.gov/pubmed 2011).

Successful only in case of major genes (Disease resistance genes...)

- **Small fraction of all genetic variation within a species,**
- **Only alleles at which the two parents differ can be detected**
- **Low map genetic resolution (10-20 CM)- due to limited recombination**
- **QTL often are not consistent across mapping populations**
- **Linked markers not suitable for unrelated genotypes....WHY?**

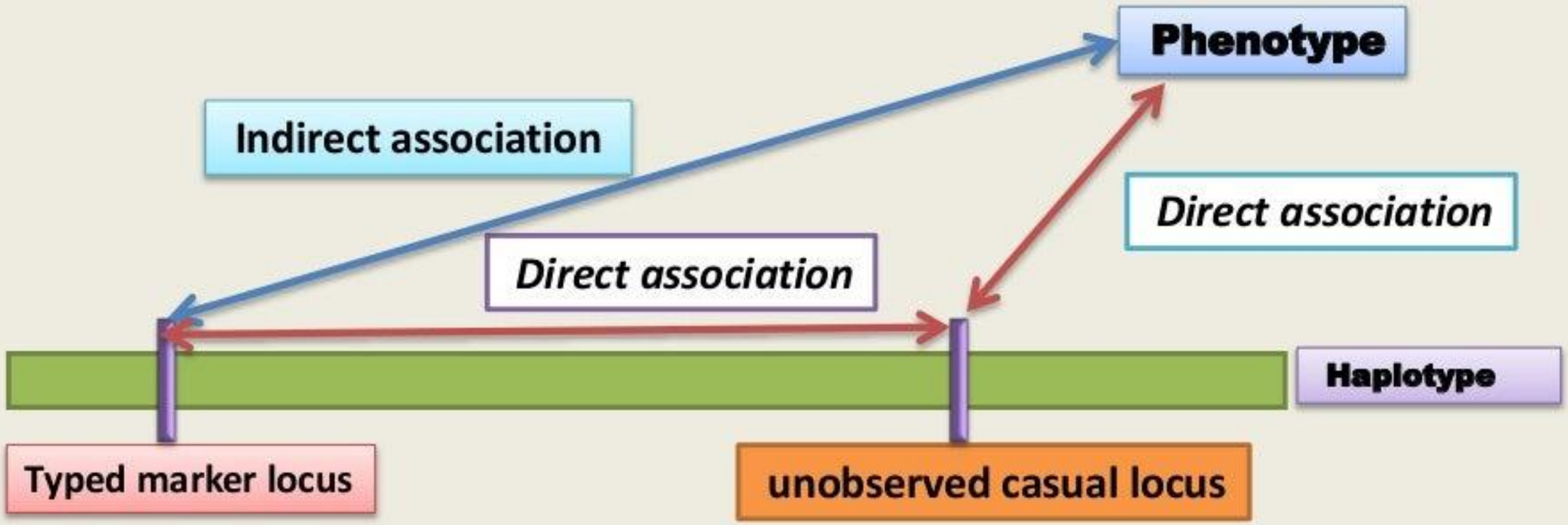
Ross-Ibarra et al. 2007 and Holland 2007

ASSOCIATION MAPPING

Could be answer and alternative to family based mapping

To dissect complex traits…

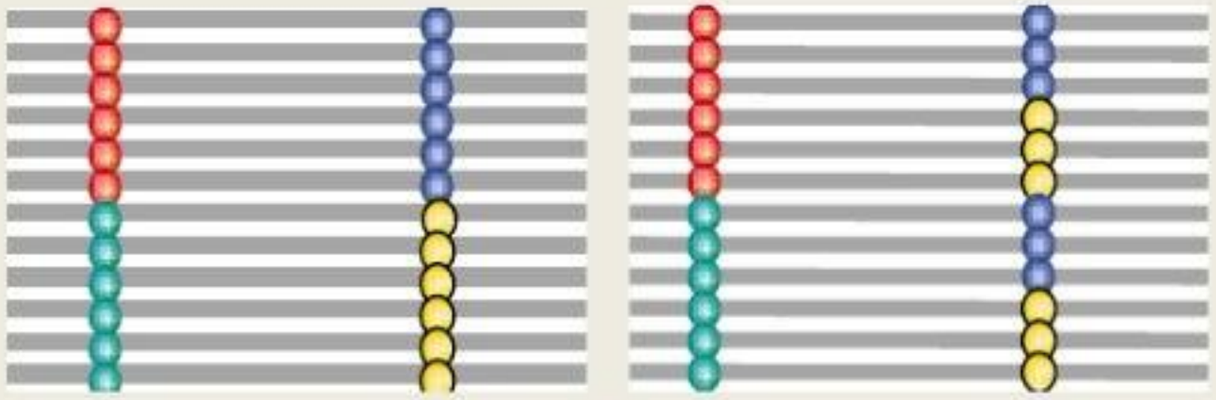
- **Also called linkage disequilibrium mapping**
- **A natural population survey to determine marker-trait associations using Linkage Disequilibrium**
- **The power depends on the strength of this correlation (i.e., on the degree of LD between the genotyped marker and the functional variant)**



Thornsberry et al.,2001

WHAT IS LINKAGE DISEQUILLIBRUM AND HOW...?

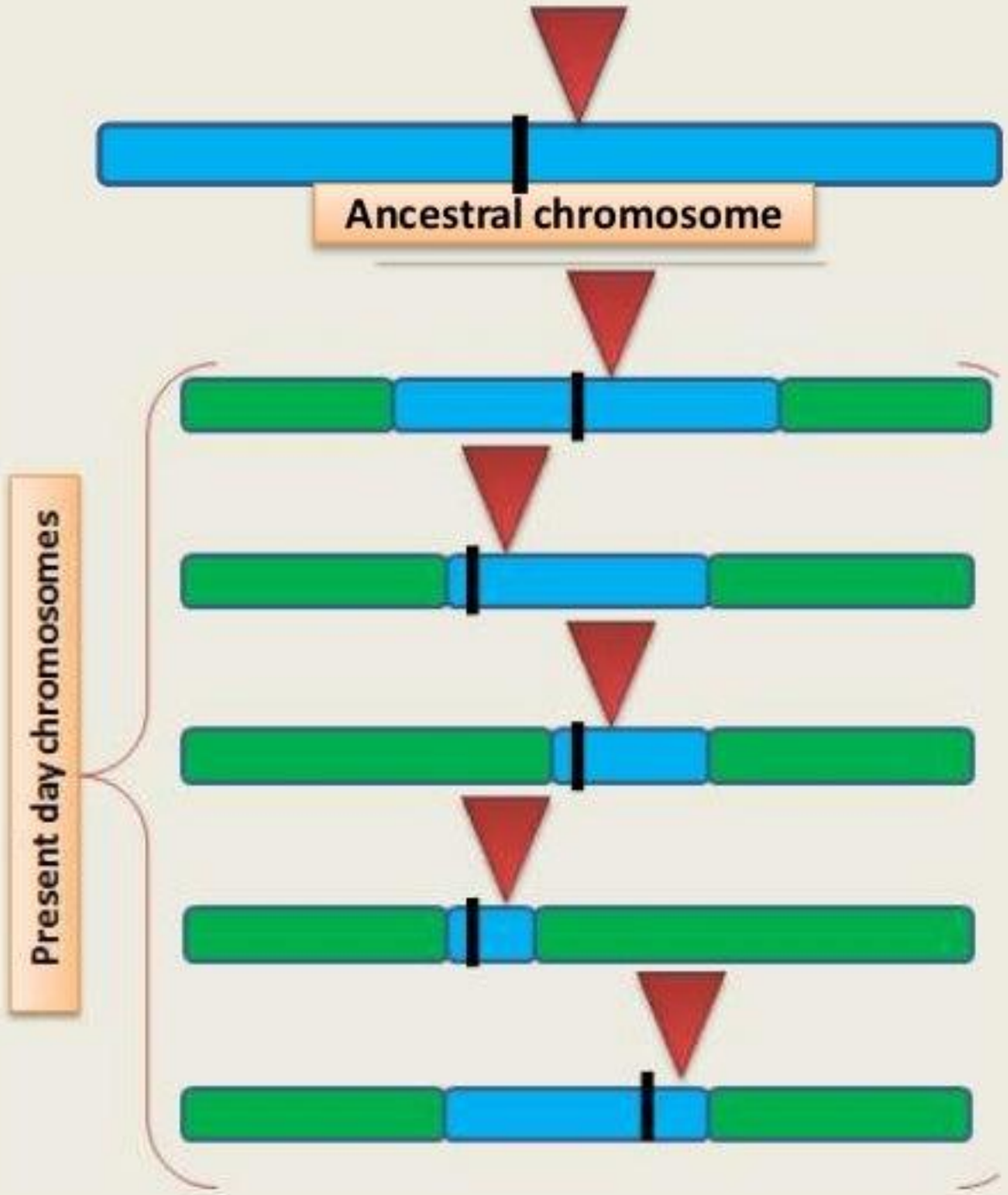
➢Non- random association of alleles at adjacent loci....



➢The closer two markers are, the stronger the LD

The resolution with which a QTL can be mapped is a function of how quickly LD decays over distance.

Linkage disequilibrium around an ancestral mutation



How to measure LD

Mathematically,

Linkage equilibrium can be described as PAB= PA x PB and

Linkage disequilibrium as PAB ≠ PA x PB

where A and B are alleles at two different loci, PAB is the frequency of haplotypes having both alleles at the two loci, PA and PB are the frequency of haplotypes having only A allele and B allele, respectively.

D ranges from 0 -1 (At equilibrium, D= 0)

D' and r2 **have been widely used to quantify LD**

$$D' = \frac{|D|}{D_{\max}}$$

$$\text{where} \quad D_{\max} = \min(p_A p_b, p_a p_B) \text{ if } D > 0;$$

$$D_{\max} = \min(p_A p_B, p_a p_b) \text{ if } D < 0$$

$$r^2 = \frac{D^2}{p_A p_a p_B p_b}$$

- ***D is that its range is determined*** by the allele frequency (0-1)
- The ***r square statistic*** is the squared value of the Pearson's corrélation coefficient
- In terms of identifying SNPs or haplotypes significantly associated with phenotypic trait variation, ***r2 is*** the most relevant LD measurement (0.1-0.2)

Lewontin 1964 and Hill and Robertson 1968)

Factors effecting Linkage disequilibrium and Association mapping

Germplasm

Longer in narrow-based germplasm than LD extent in broad-based germplasm

1 kb in landraces, (~ 2kb) in diverse inbred lines and can extend up to several hundred kb in commercial elite inbred lines

For higher resolution the mapping population ….unrelated diverse genotypes

Mating system

Selfing reduces opportunities for effective recombination…thus LD extends much further as compared to out crossing species

LD declines more rapidly in out-crossing plant species than self-pollinated plants, a higher resolution is expected, enabling more accurate fine mapping.

Nordborg 2013 and Zhang et al. 2013

Recombination

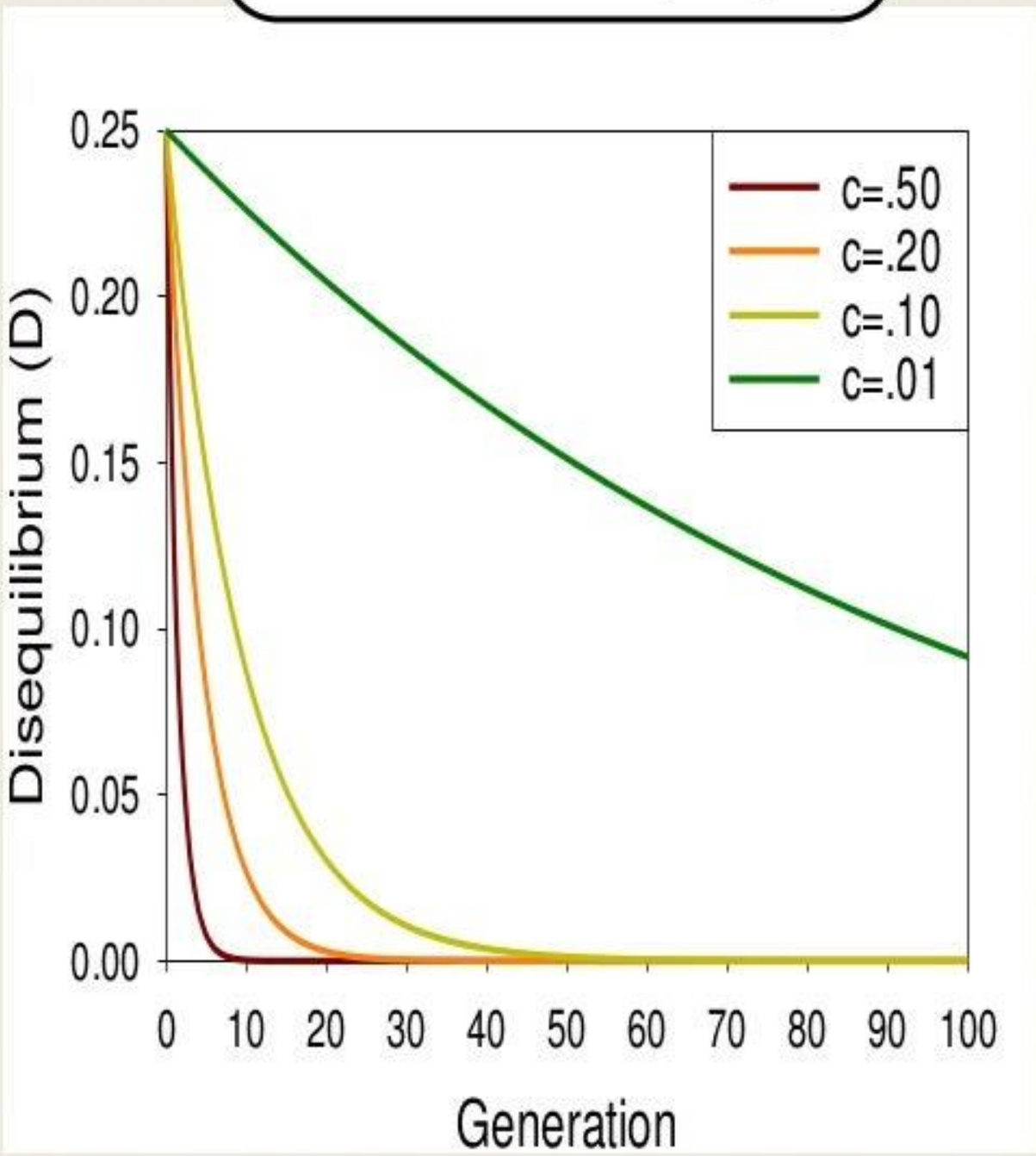
LD decays by one-half with each generation of random mating

$$D_{t+1} = (1-c)D_t$$

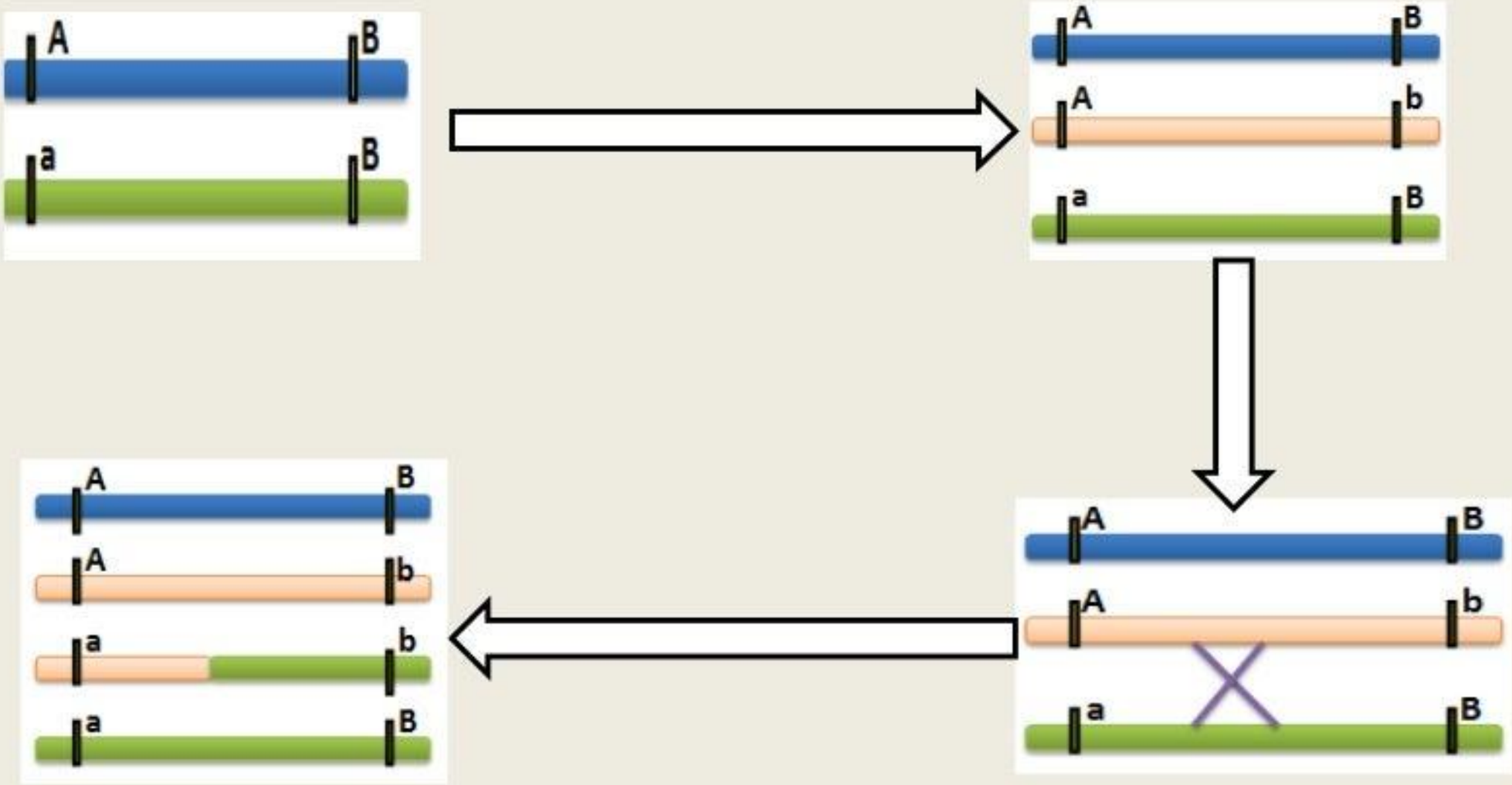
$$D_t = (1-c)^t D_0$$

Higher the recombination rate, LD decay (the rate of return to random association between two given alleles) occurs more rapidly even within few kilo bases

c = recombination frequency



Mutation



The creation of LD and erosion by recombination

Question is whether increased or decreased level of LD ?

Association mapping depends upon intensity of LD

LD level higher or lower depends on the objective of mapping study

The resolving power of LD mapping depends on how rapidly LD decays with genetic distance.

Extensive level of LD - low number of markers.....but lowers the mapping resolution (coarse mapping)...and Small stretched....more markers....fine mapping

In some populations, LD will decay so rapidly that they are best suited for fine mapping, whereas in others the decay might be so slow that whole genome scans are practical.

Inbred crops are theoretically less suitable for high-resolution association mapping because of their low level of molecular diversity and high overall genomic LD.

Advantages of Association mapping over conventional mapping

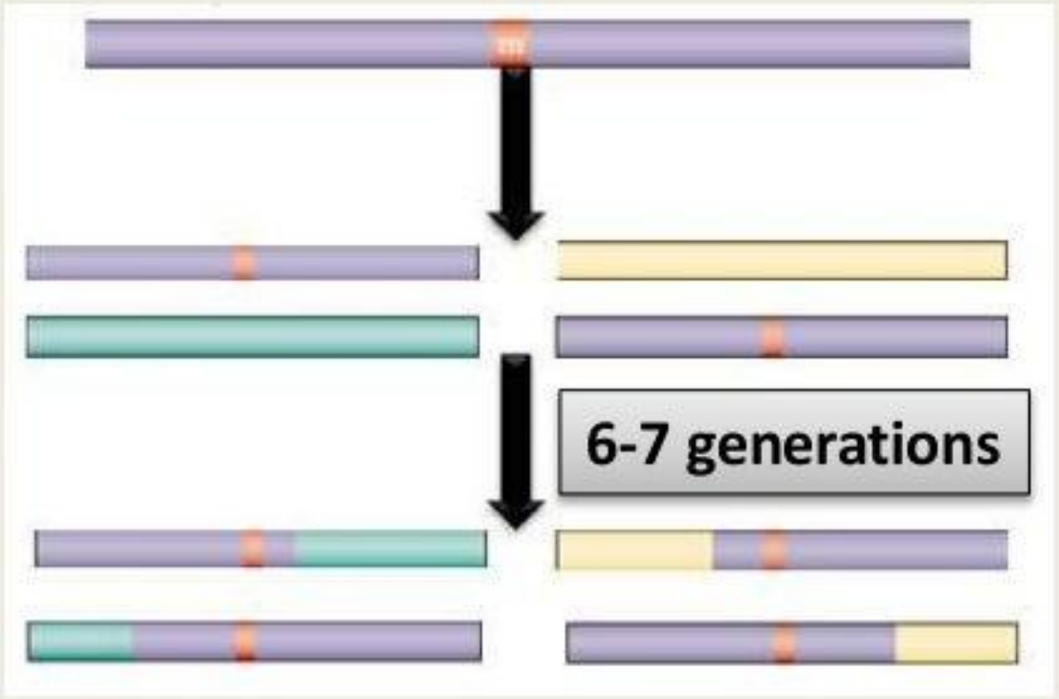
	Conventional	LD mapping
Mapping population	Biparental, structured	Natural/ breeding pool, not structured
Meiosis cycle	Few (6-7)	Several
QTL precision	Less	High –Great resolution
Trait variation	Explains between parents	Natural
LD break up	Less	more
Perennial crops	Not applicable	Effective
Markers	Specific	Diverse genotypes
Cost and ease	More cost and labour	Less cost and reduced time

Limitations

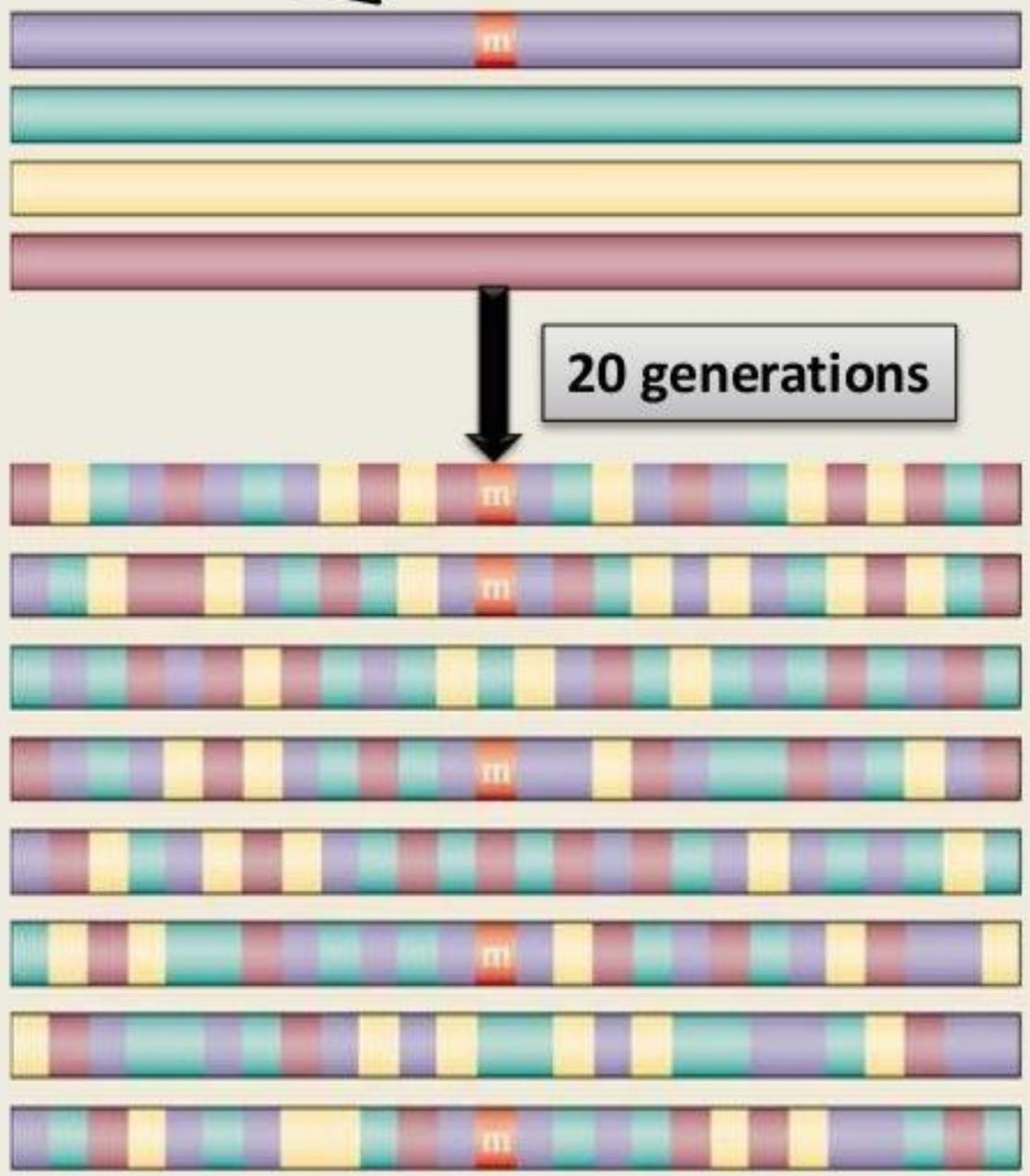
- *Needs statistical assessment to investigate the relatedness of the lines and the overall population structure*

Linkage vs Association mapping: How it leads to high resolution..

LINKAGE



ASSOCIATION MAPPING



Cardon and John,2001

Approaches for Association mapping

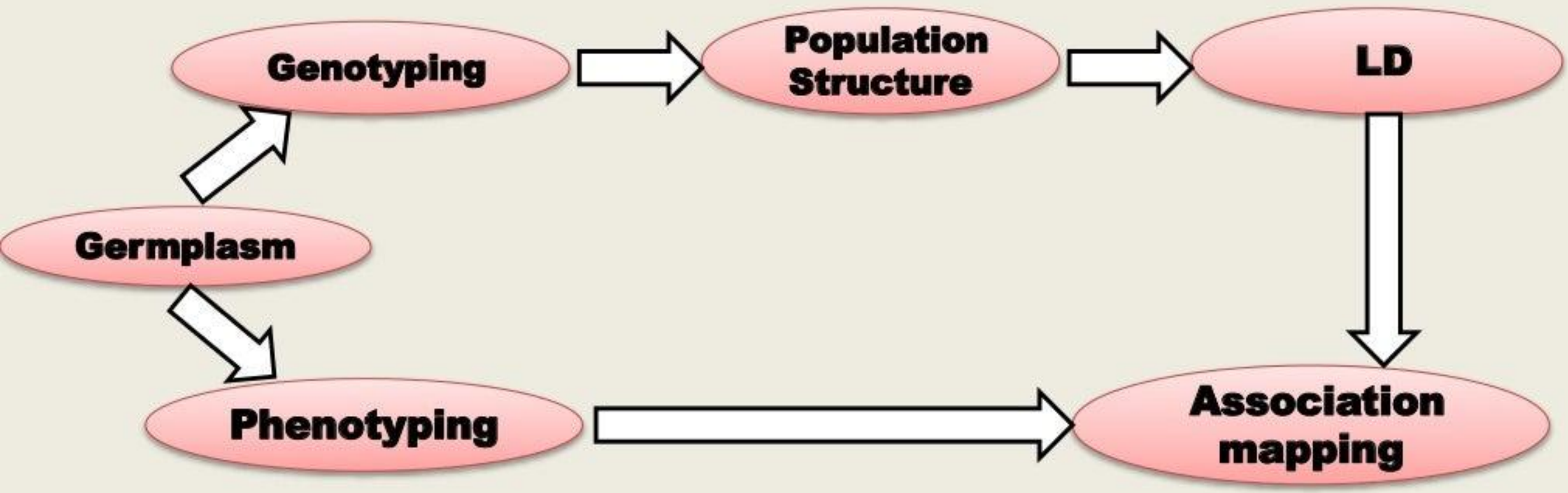
Candidate gene association mapping

Relates polymorphism in selected candidate genes

Genome wide association mapping

Genetic variation in whole genome to find signals

General procedure for association mapping



Advantages and drawbacks of methods for identifying the genetic basis of complex traits

Methods	Advantages	Drawbacks
Candidate gene association mapping (2002)	✓Relates sequence variation encoding either regulatory or functional product directly to trait of interest ✓Fine mapping	✓Requires detailed knowledge of the function, sequence and genetics of the trait under study
So why not just cover the entire genome with genetic markers?		
Genome wide Association (2005)	✓ Fine mapping (blind approach) ✓ Detection of common alleles	✓Reduced power to detect rare alleles ✓Genetic and allelic heterogeneity

letter

Dwarf8 polymorphisms associate with variation in flowering time

Jeffry M. Thornsberry[1], Major M. Goodman[2], John Doebley[3], Stephen Kresovich[4], Dahlia Nielsen[5], & Edward S. Buckler IV[1,6]

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 26 June 2012
doi: 10.3389/fpls.2012.00129

A candidate gene-based association study of tocopherol content and composition in rapeseed (*Brassica napus*)

Steffi Fritsche[1], Xingxing Wang[2], Jinquan Li[3], Benjamin Stich[3], Friedrich J. Kopisch-Obuch[1], Jessica Endrigkeit[1], Gunhild Leckband[4], Felix Dreyer[4], Wolfgang Friedt[5], Jinling Meng[2] and Christian Jung[1]*

[1] *Faculty of Agricultural and Nutritional Sciences, Plant Breeding Institute, Christian-Albrechts-University, Kiel, Germany*
[2] *National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, China*
[3] *Quantitative Crop Genetics, Max Planck Institute for Plant Breeding Research, Cologne, Germany*
[4] *Norddeutsche Pflanzenzucht Hans-Georg Lembke KG, Hohenlieth, Germany*
[5] *Faculty of Agricultural Sciences, Nutritional Sciences and Environmental Management, Institute of Agronomy and Plant Breeding I, Justus-Liebig-University, Giessen, Germany*

Journal of Experimental Botany, Vol. 62, No. 2, pp. 701–716, 2011
doi:10.1093/jxb/erq308 Advance Access publication 17 November, 2010
This paper is available online free of all access charges (see http://jxb.oxfordjournals.org/open_access.html for further details)

Journal of Experimental Botany
www.jxb.oxfordjournals.org

RESEARCH PAPER

Genetic association mapping identifies single nucleotide polymorphisms in genes that affect abscisic acid levels in maize floral tissues during drought

Tim L. Setter[1,*], Jianbing Yan[2], Marilyn Warburton[2,†], Jean-Marcel Ribaut[2,‡], Yunbi Xu[2], Mark Sawkins[2,§], Edward S. Buckler[1,3,4], Zhiwu Zhang[4] and Michael A. Gore[1,4,¶]

Journal of Experimental Botany Advance Access published February 5, 2013

Journal of Experimental Botany
doi:10.1093/jxb/ert018
This paper is available online free of all access charges (see http://jxb.oxfordjournals.org/open_access.html for further details)

Journal of Experimental Botany
www.jxb.oxfordjournals.org

RESEARCH PAPER

Association of candidate genes with drought tolerance traits in diverse perennial ryegrass accessions

Xiaoqing Yu[1], Guihua Bai[2], Shuwei Liu[3], Na Luo[4], Ying Wang[5], Douglas S. Richmond[6], Paula M. Pijut[7], Scott A. Jackson[8], Jianming Yu[9] and Yiwei Jiang[1,*]

Sehgal *et al. BMC Plant Biology* 2012, **12**:9
http://www.biomedcentral.com/1471-2229/12/9

BMC Plant Biology

RESEARCH ARTICLE Open Access

Integration of gene-based markers in a pearl millet genetic map for identification of candidate genes underlying drought tolerance quantitative trait loci

Deepmala Sehgal[1], Vengaldas Rajaram[2], Ian Peter Armstead[1], Vincent Vadez[2], Yash Pal Yadav[3], Charles Thomas Hash[2] and Rattan Singh Yadav[1*]

Theor Appl Genet (2011) 123:907–926
DOI 10.1007/s00122-011-1636-2

ORIGINAL PAPER

Deciphering the genetics of flowering time by an association study on candidate genes in bread wheat (*Triticum aestivum* L.)

Michel Rousset · Isabelle Bonnin · Carine Remoué · Matthieu Falque · Bénédicte Rhoné · Jean-Baptiste Veyrieras · Delphine Madur · Alain Murigneux · François Balfourier · Jacques Le Gouis · Sylvain Santoni · Isabelle Goldringer