

# Next Generation Crop Improvement

Peter L. Morrell  
Agronomy & Plant Genetics - Minnesota

# REVIEWS

## Crop genomics: advances and applications

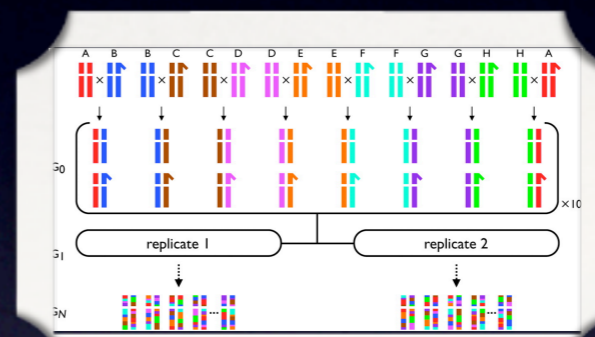
*Peter L. Morrell<sup>1</sup>, Edward S. Buckler<sup>2</sup> and Jeffrey Ross-Ibarra<sup>3</sup>*

# Soybean Genomics Strategic Plan

- Goal 1: Improve Utility of Genome Sequence
  - Improve bioinformatics resources - practical applications
- Goal 2: Translational Genomics - Optimize Breeding Efficiency

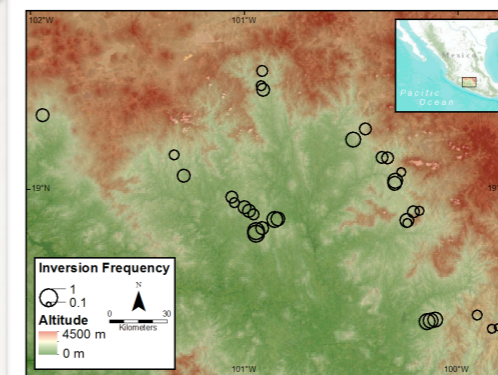
# Topics

- Next generation populations
- Selection against deleterious mutations
- Applications of genome-wide SNP data



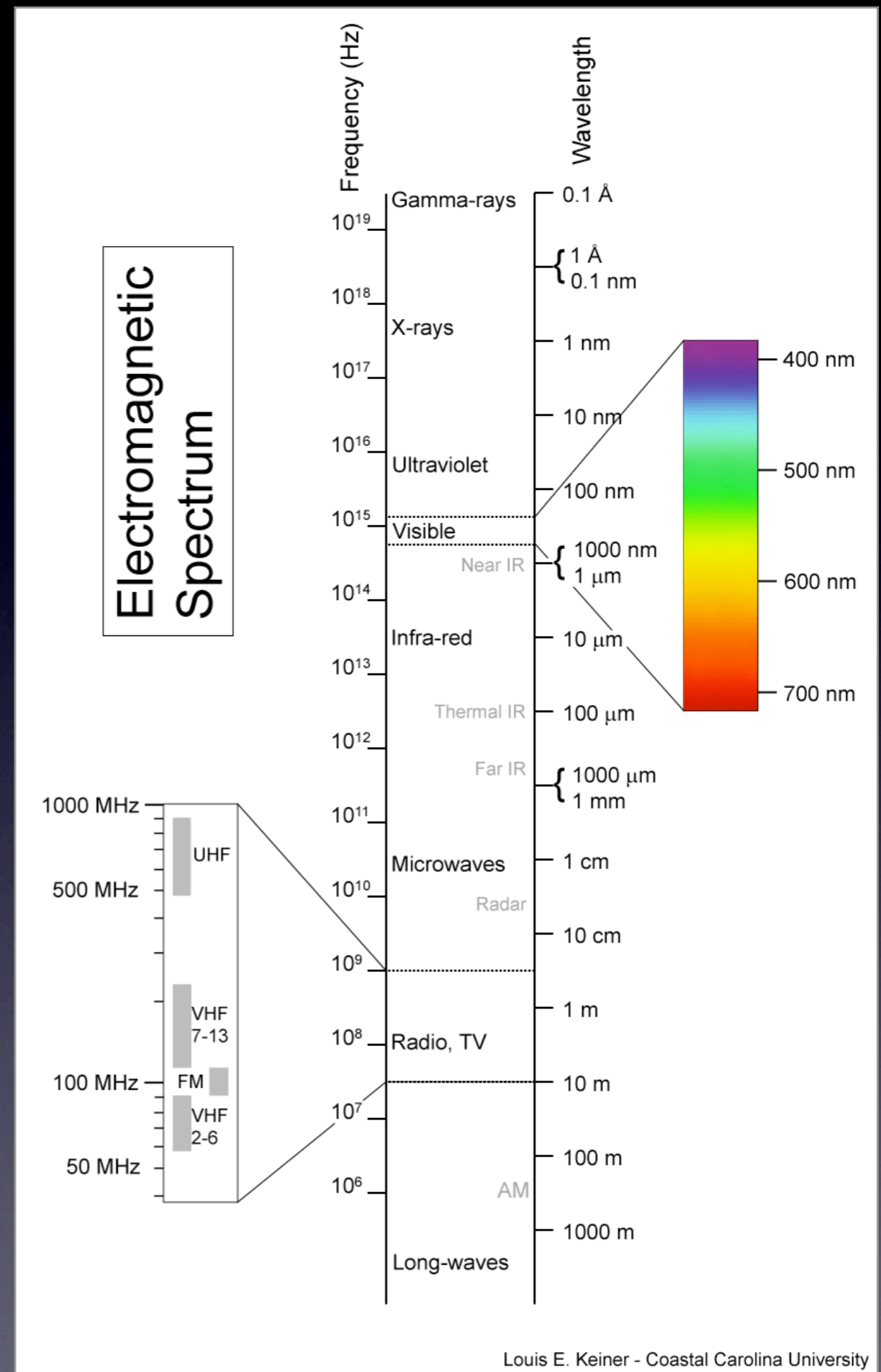
a

Rice	... AAC CAC CTT ...
Brachypodium	... AAC CAG CTC ...
Sorghum	... AAT CAT CTC ...
	Asn His Leu
Maize	... AAC GAT CTC ...
	Asn Asp Leu



# Observational Astronomy

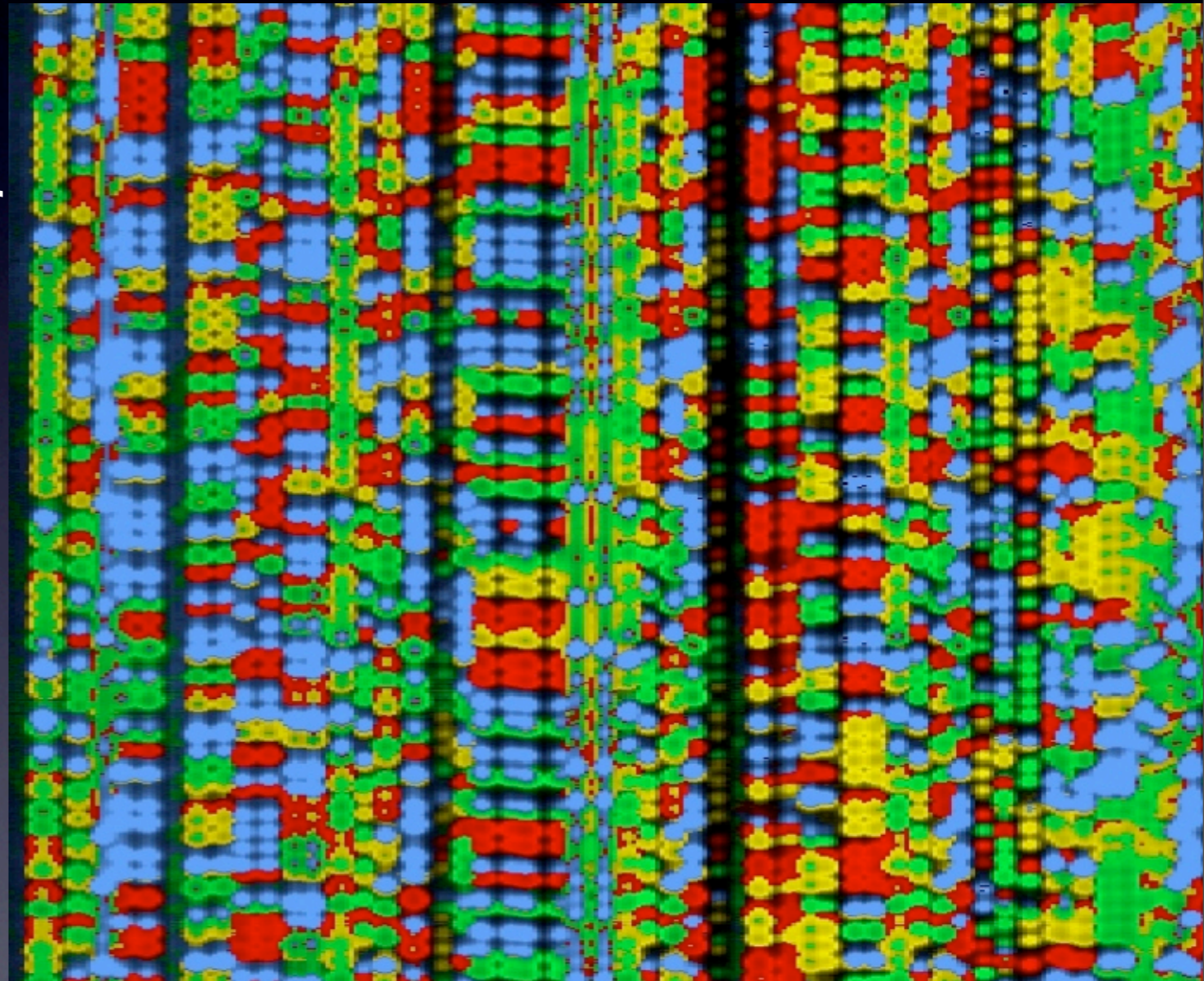
- Information about places we won't ever visit
- The visible portion of the electromagnetic spectrum is only a fraction of what exists
- DNA resequencing data latent with information about the past



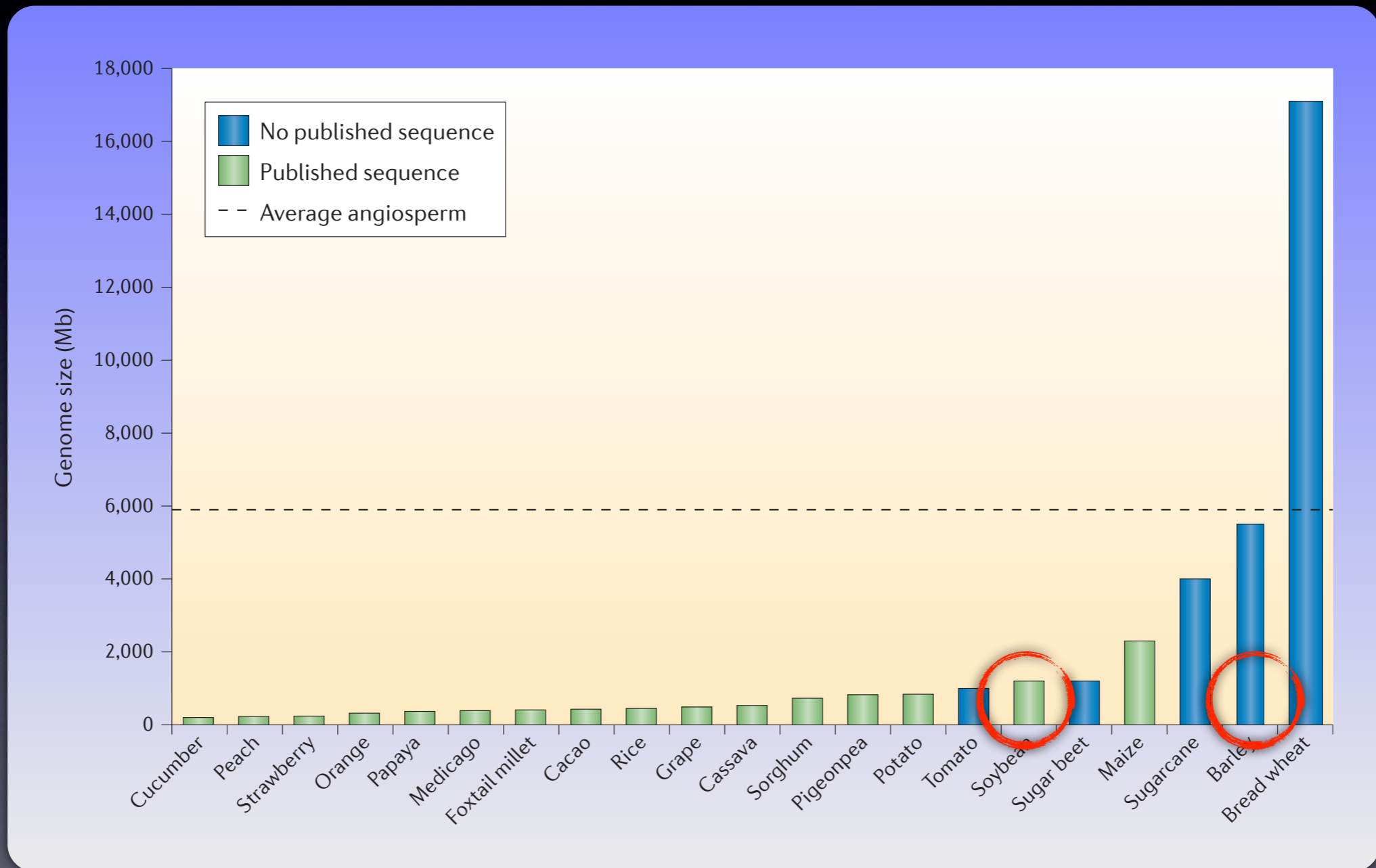
Louis E. Keiner - Coastal Carolina University

# DNA Resequencing

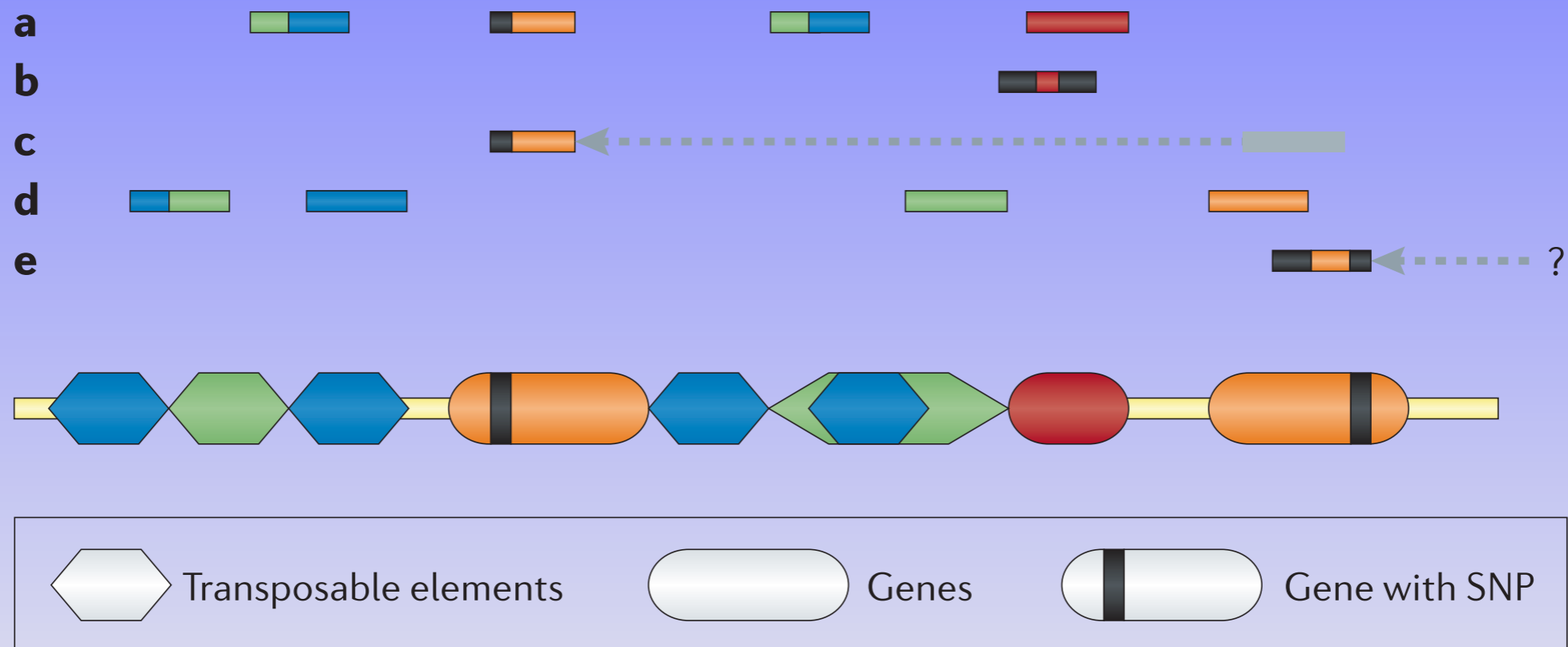
- Most direct measure of genetic diversity
- Can assay all heritable variation
- Can now be collected very rapidly



# Genome Size

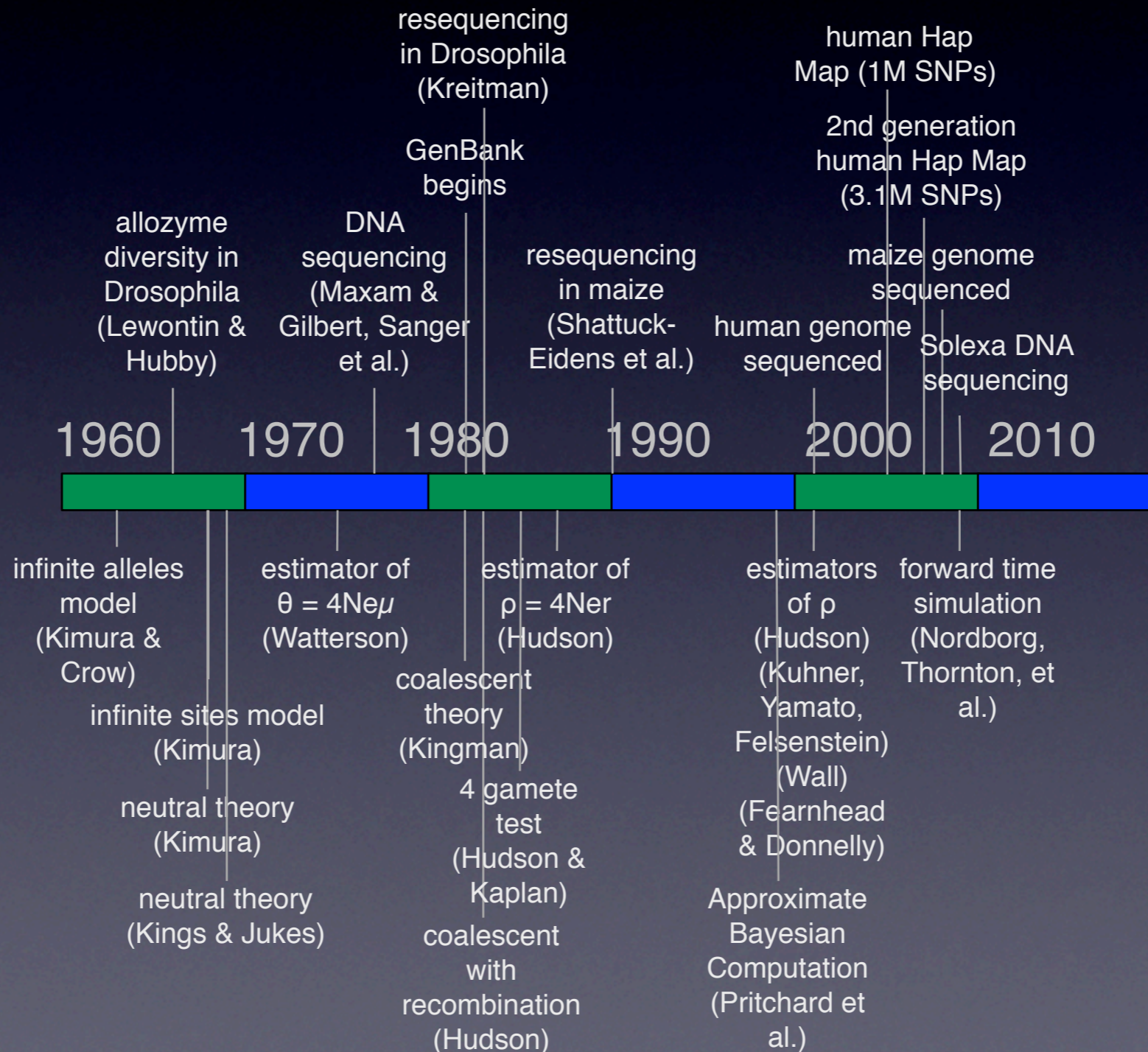


# Short Read Mapping

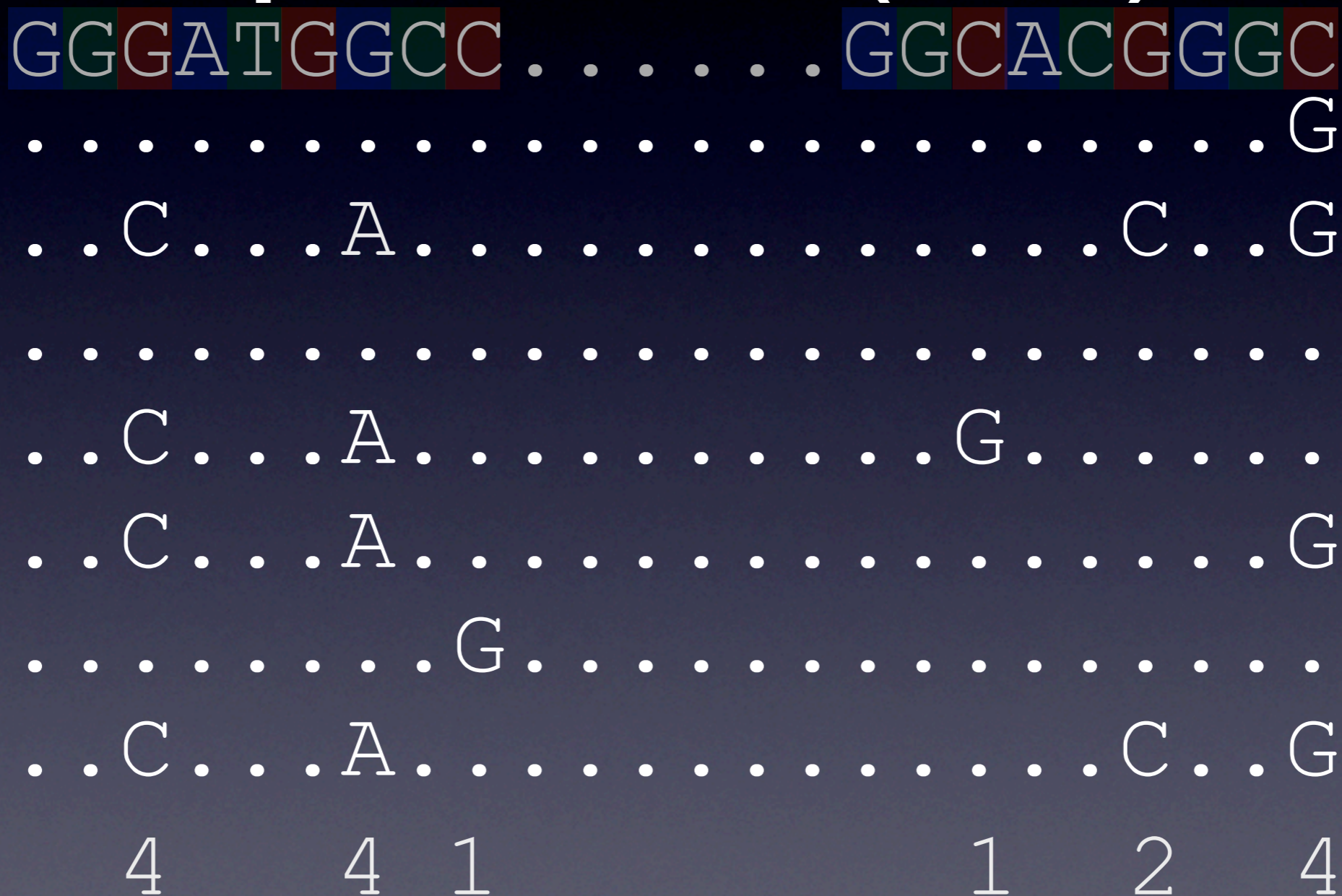


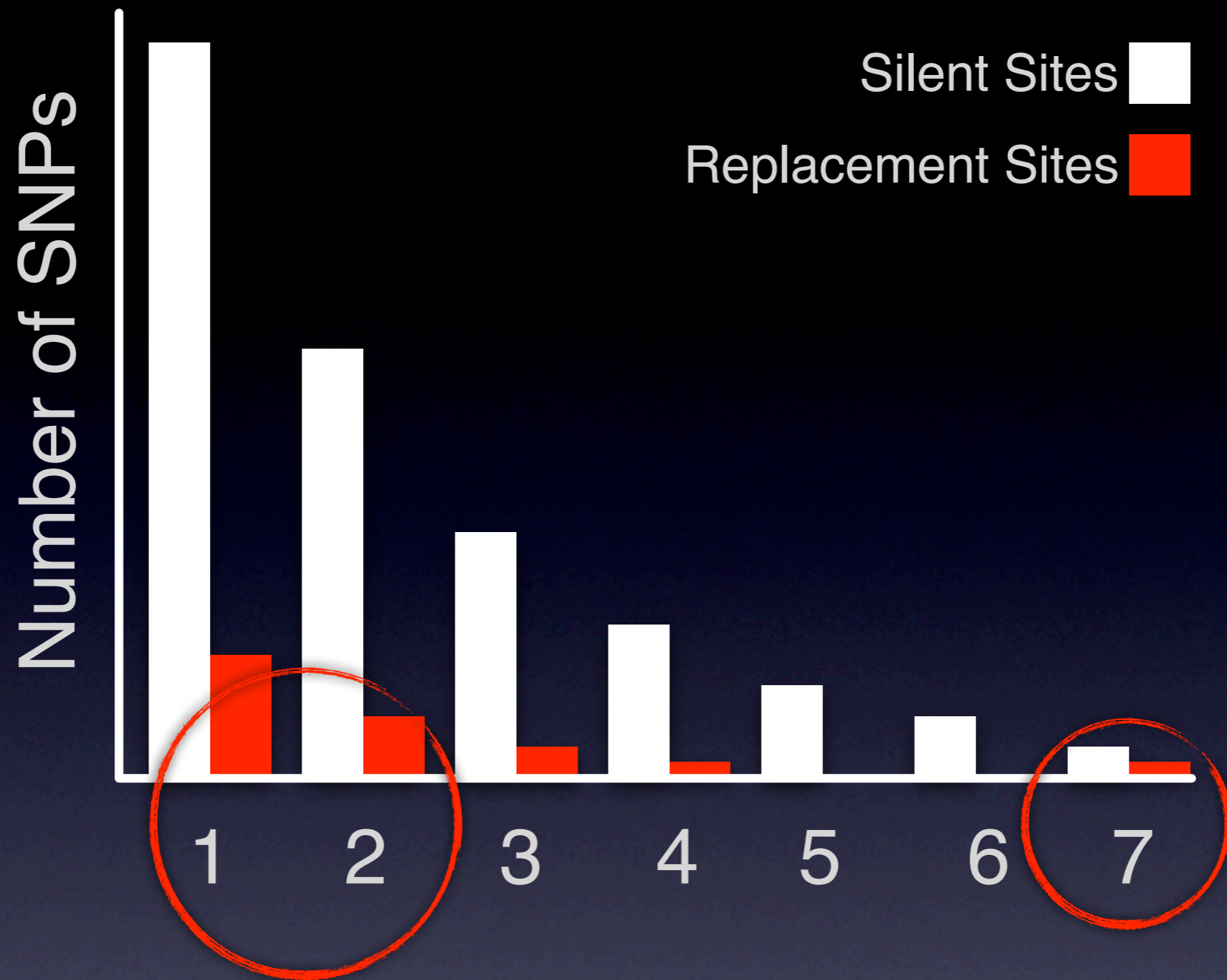


# Molecular Population Genetics



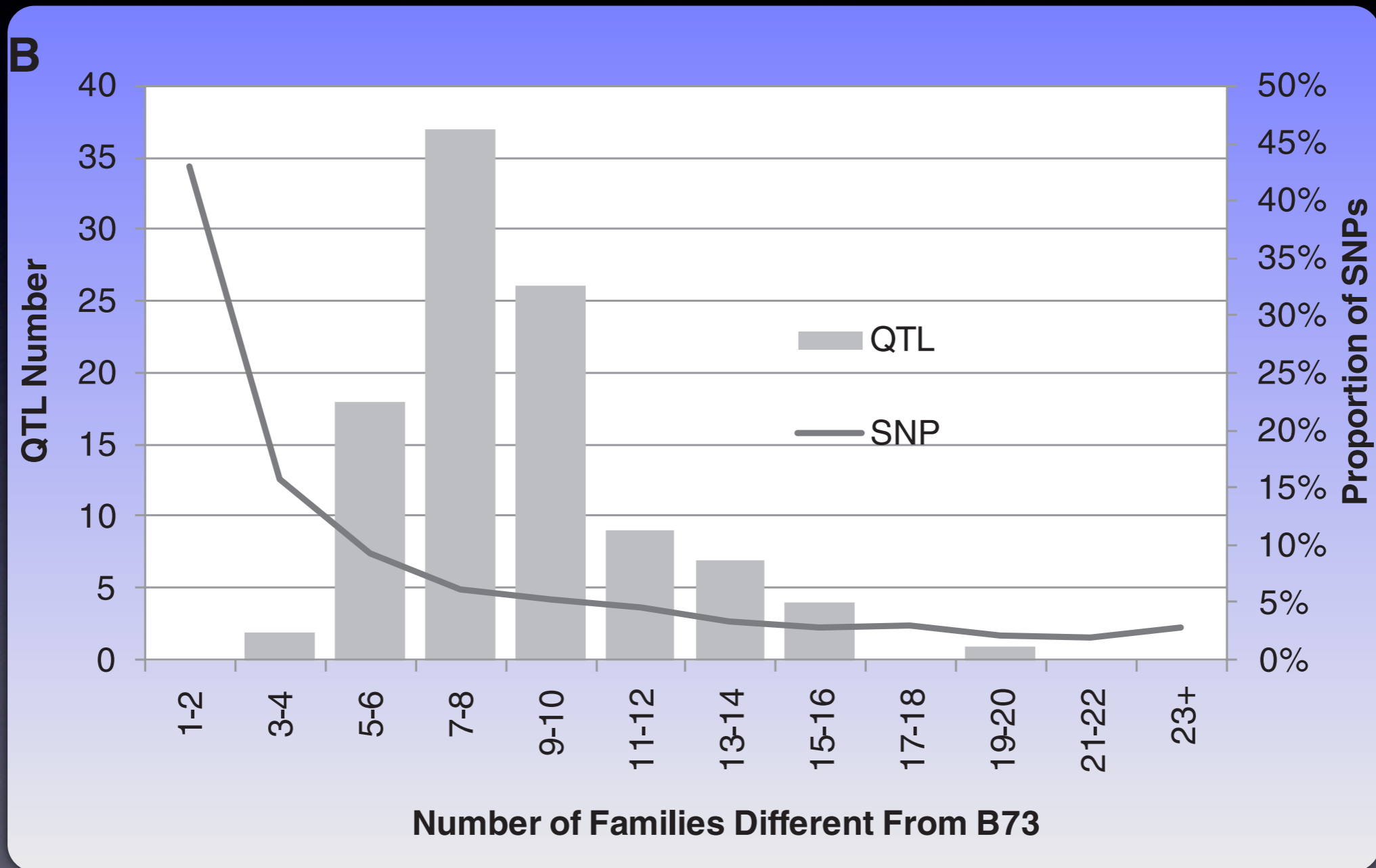
# Derived Site Frequency Spectrum (SFS)





# Derived SFS

# QTL & SNP Frequency



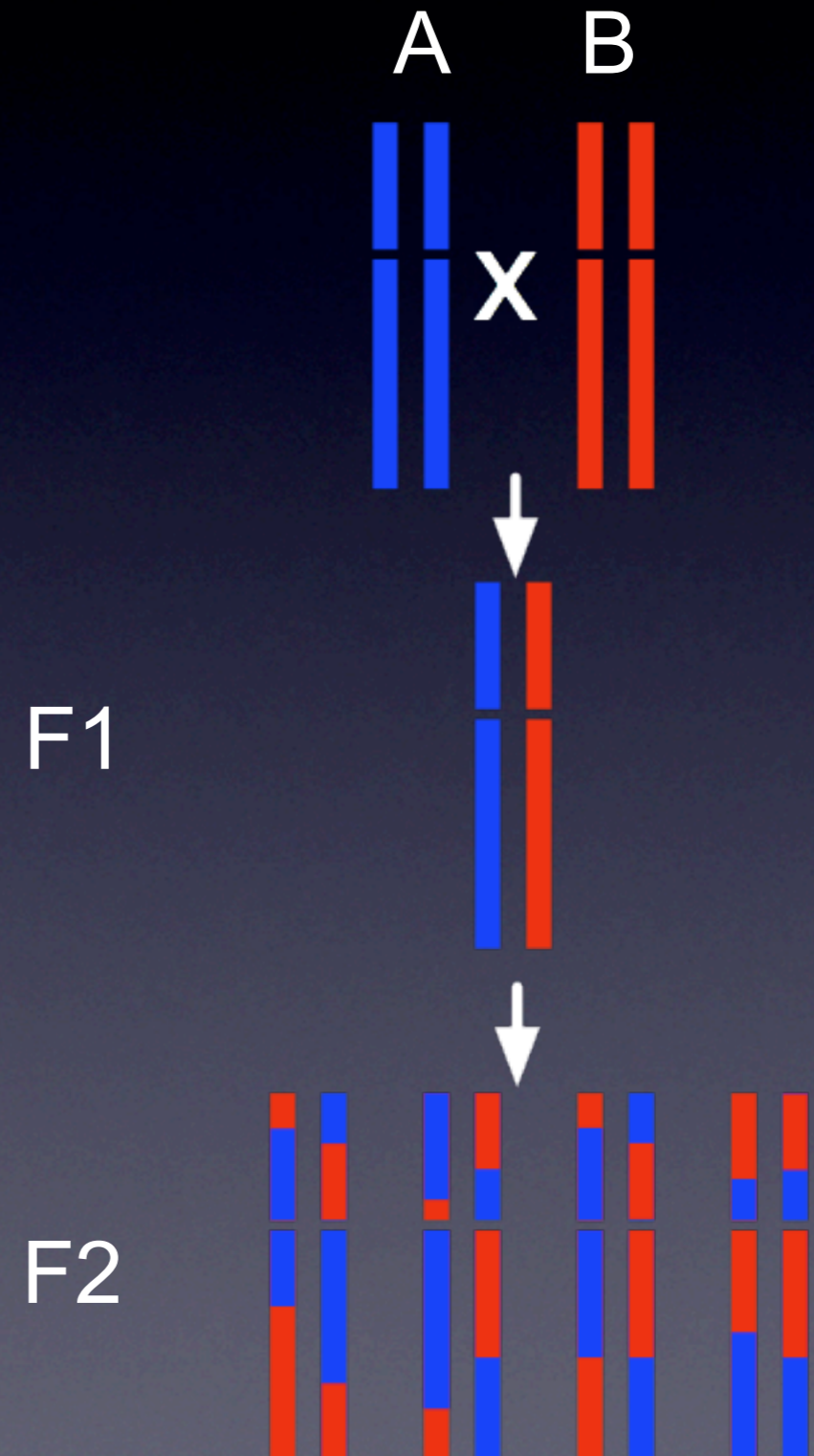
Buckler et al. 2009

# Next Generation Populations

- Where are mutations that control quantitative variation?
  - How large an effect do they have?
- What is their frequency in the population?

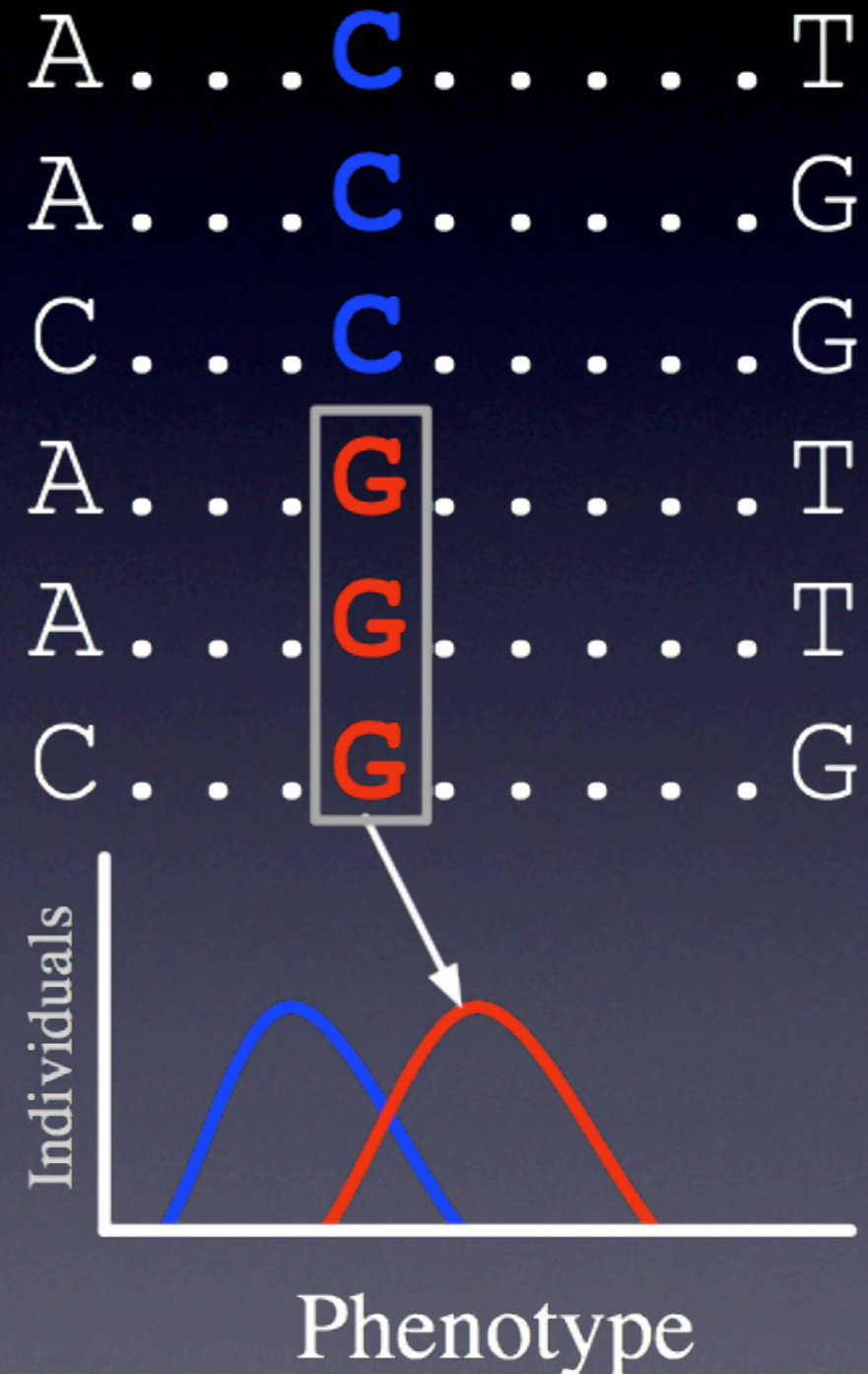
# QTL Mapping

- Quantitative Trait Locus mapping
- Progeny segregating for portions of the parental genomes
- Statistical test of association between genotype and phenotype



# Association Mapping

- Statistical association between SNPs and phenotype
- Generally assumes common disease (trait) common variant
- Fails under selection mutation balance
- Constant flux of rare causative mutations



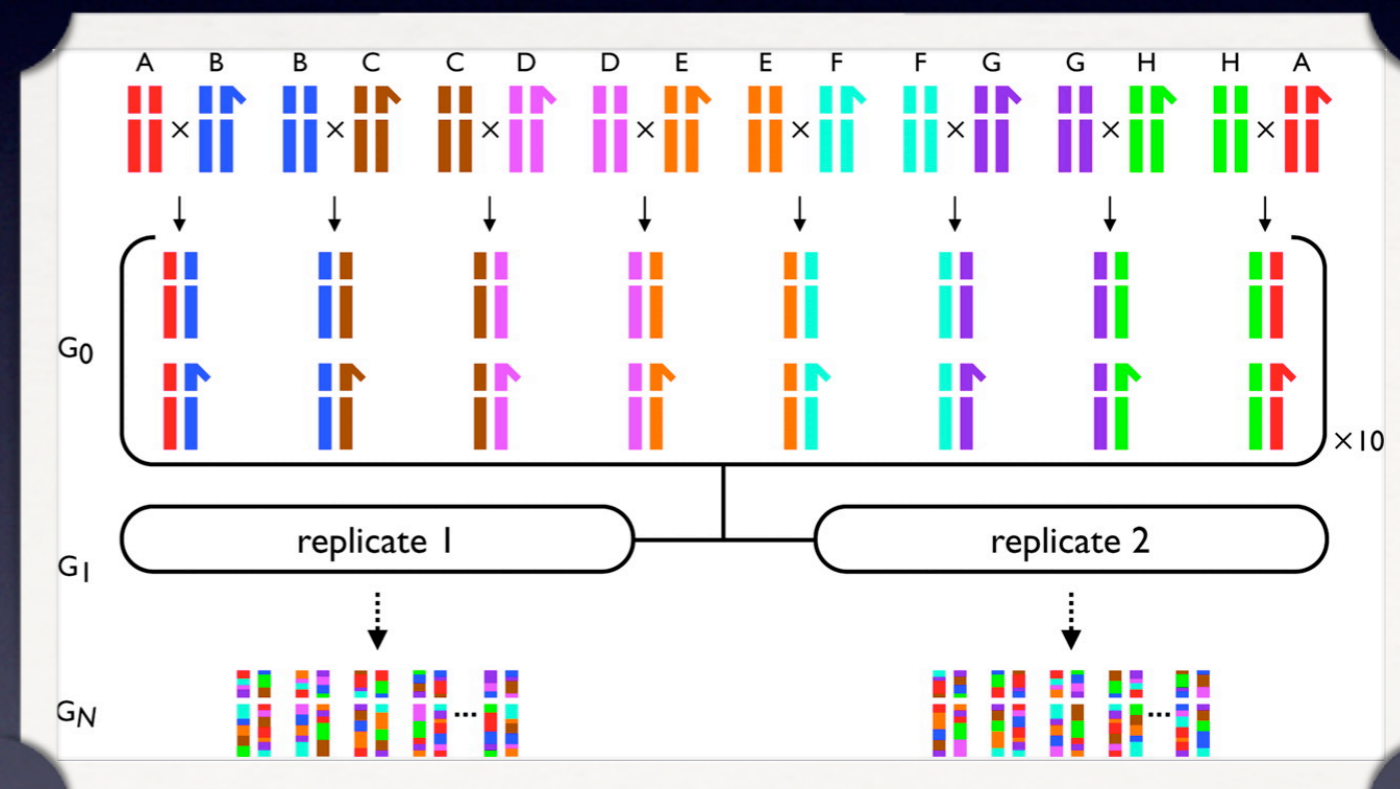
# Assumptions

- Future Resources - Next Generation
  - Reference genome sequences and fully resequence genomes for all major crops
  - $10^5$  -  $10^6$  SNPs will be available for inexpensive genotyping in most crops

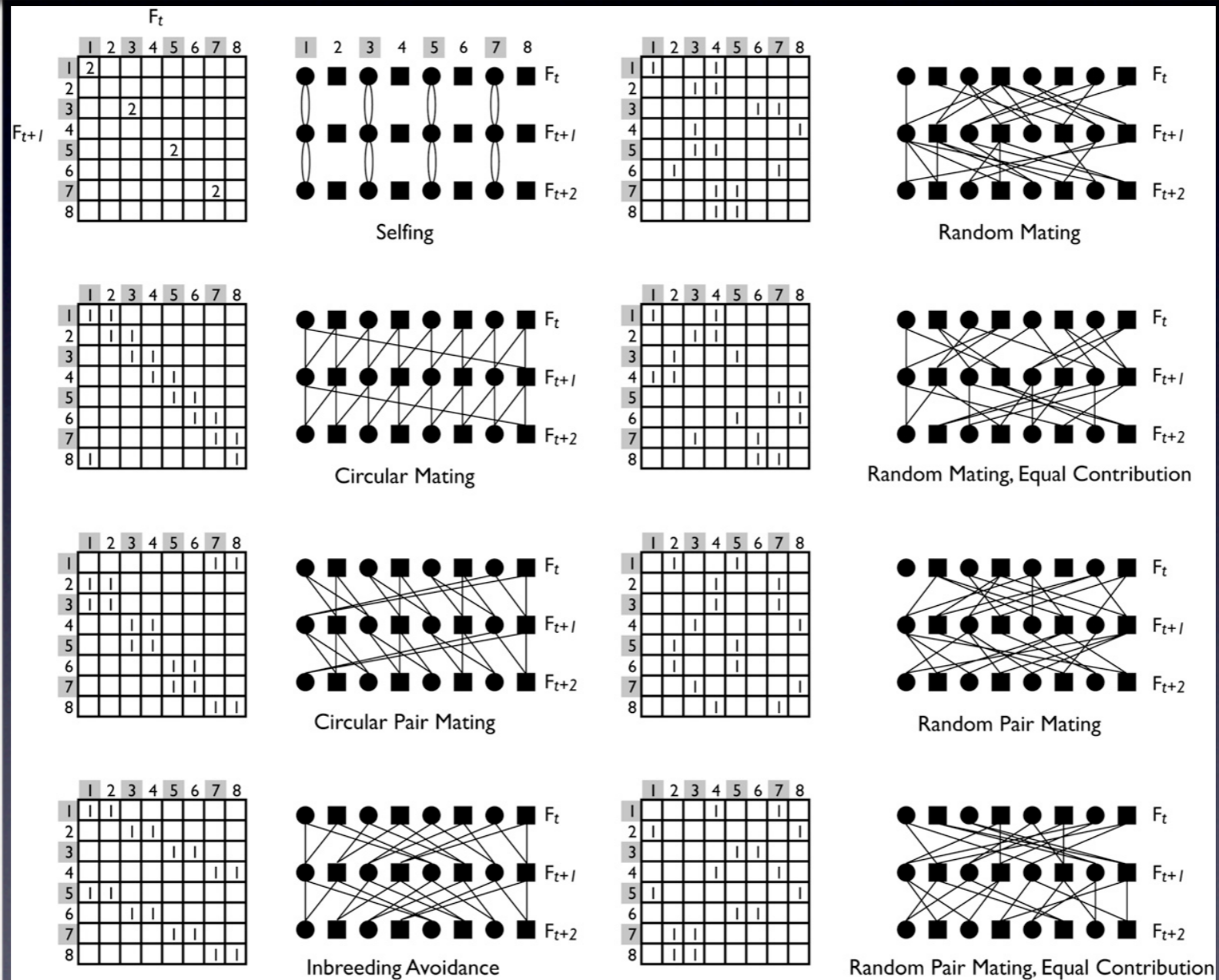


# Multiparent QTL

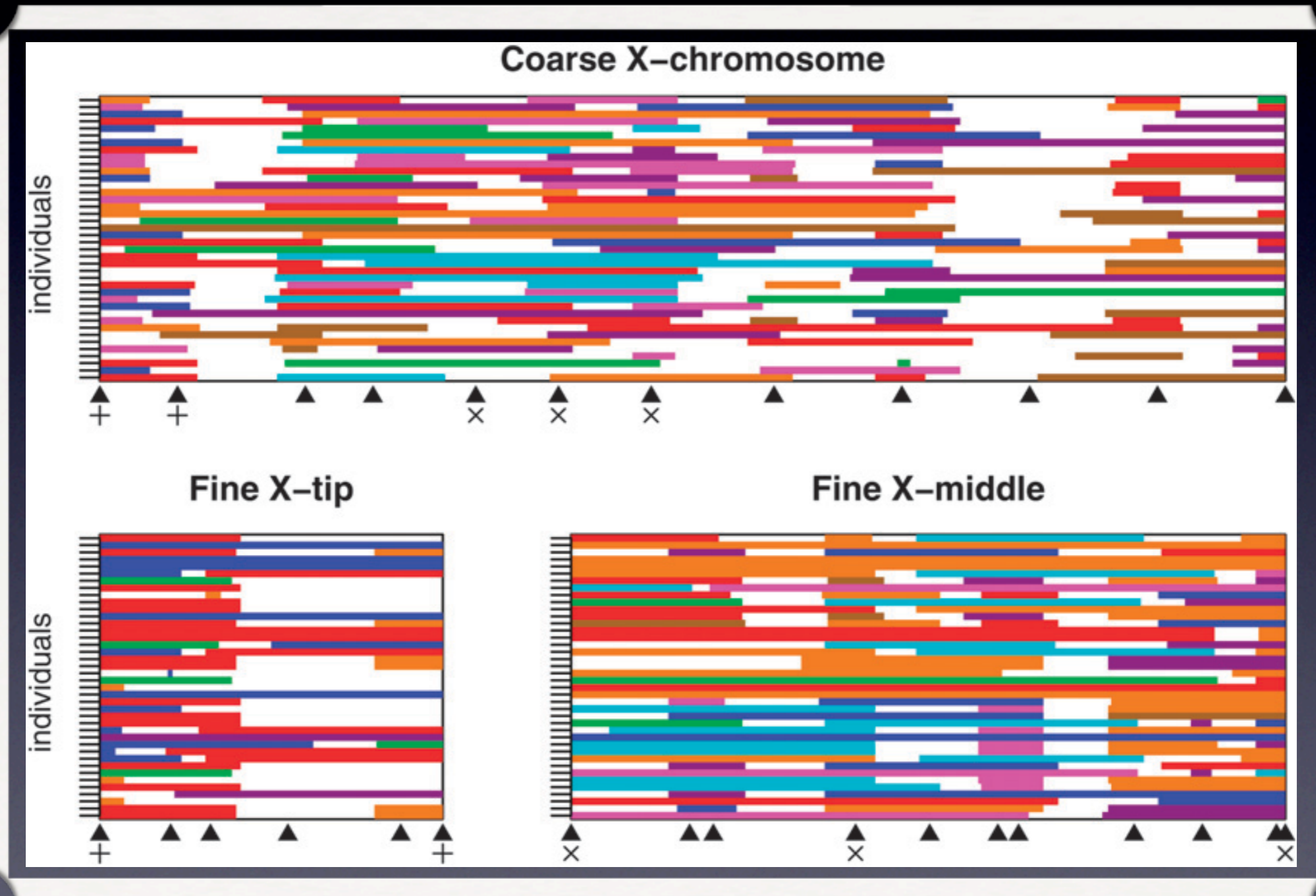
- Multiple parental lines; multiple generations
- Similar to barley composite crosses from 1920's
- *Arabidopsis* MAGIC populations
- *Drosophila* synthetic



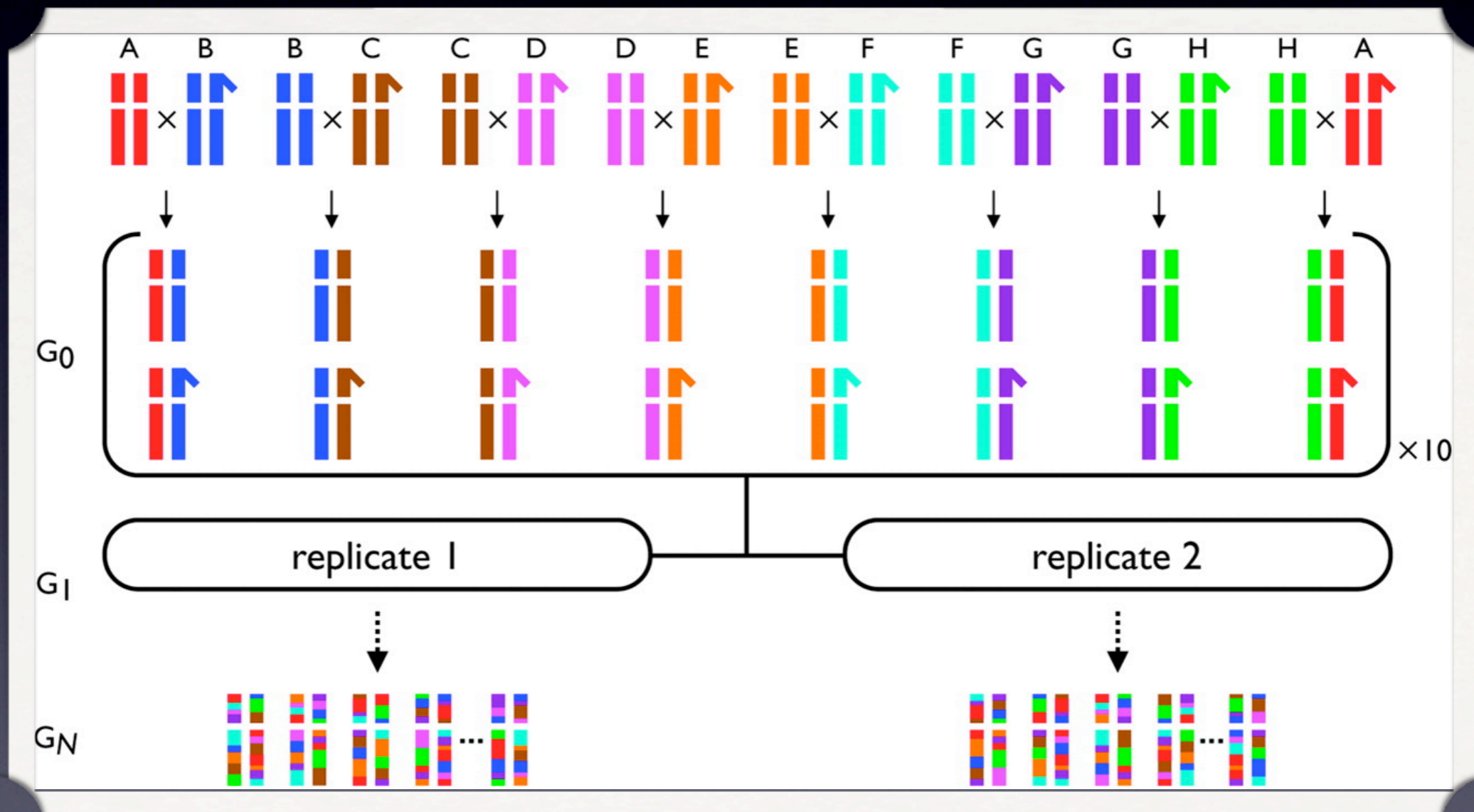
# Mating Designs



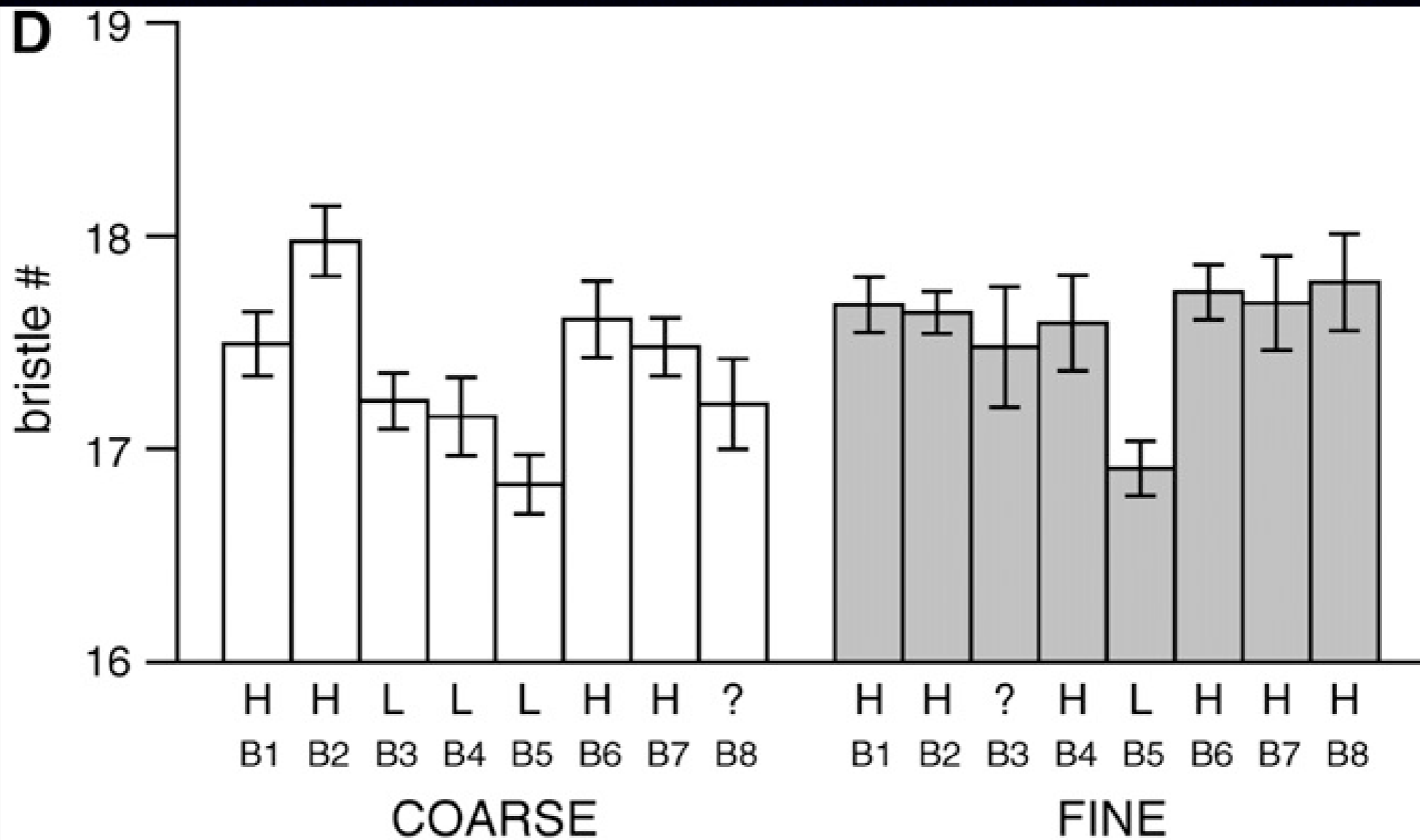
# Multiparent Population



# Multiparent QTL



# Allelic Effects



# Deleterious Mutations

- An mutation that reduces fitness or viability (yield)
- Premature stop codons, insertions/deletions, splice site variants
- Amino acid changes
- Cis-acting regulatory factors
- 100s to 1000s per genome in humans

**a**

Rice	... AAC CAC CTT ...
Brachypodium	... AAC CAG CTC ...
Sorghum	... AAT CAT CTC ...
	Asn His Leu
Maize	... AAC GAT CTC ...
	Asn Asp Leu

# Deleterious Mutations

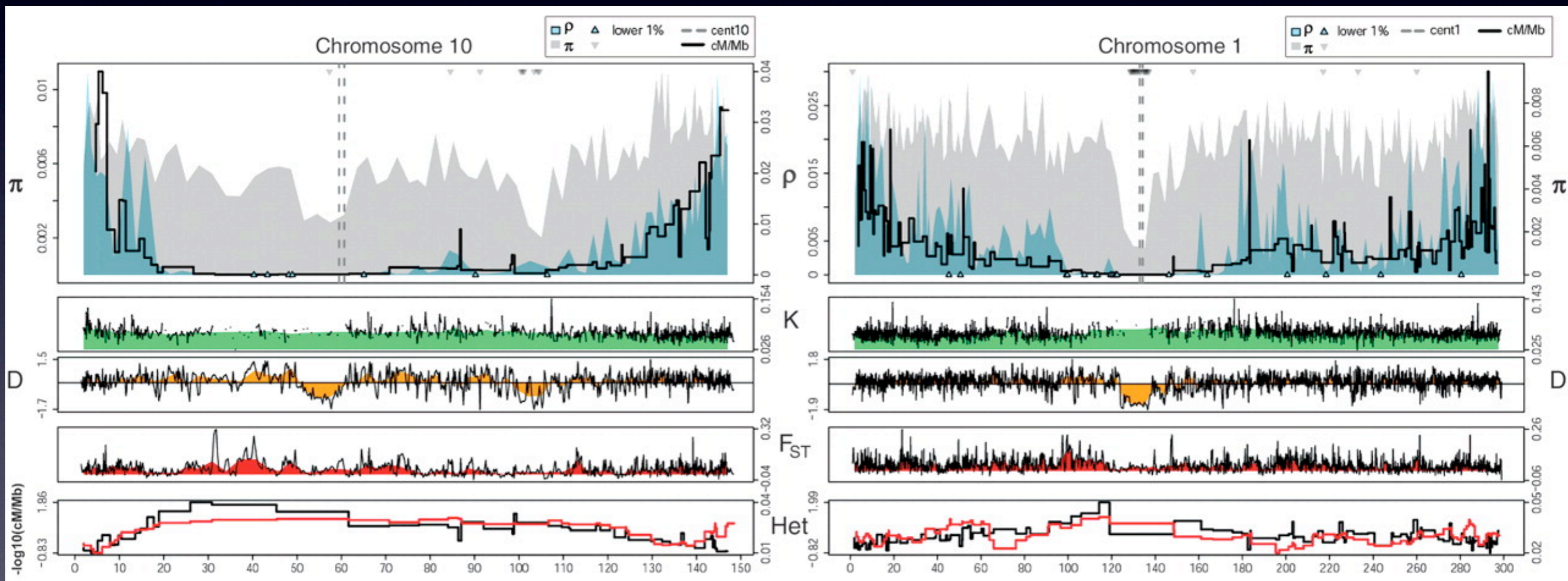
Rice	. . .	A A C	C A C	C T <b>T</b>	. . .
Brachypodium	. . .	A A C	C A <b>C</b>	C T C	. . .
Sorghum	. . .	A A <b>T</b>	C A T	C T C	. . .
		Asn	His	Leu	
Maize	. . .	A A C	<b>G</b> A T	C T C	. . .
		Asn	<b>Asp</b>	Leu	

# Deleterious vs. Lethal

- Lethal mutations - quickly purged upon inbreeding
- Deleterious mutations - modest frequency
  - Can reach higher frequencies due to linkage drag or 'allelic surfing'
  - Harbored in genomic regions with low recombination rate

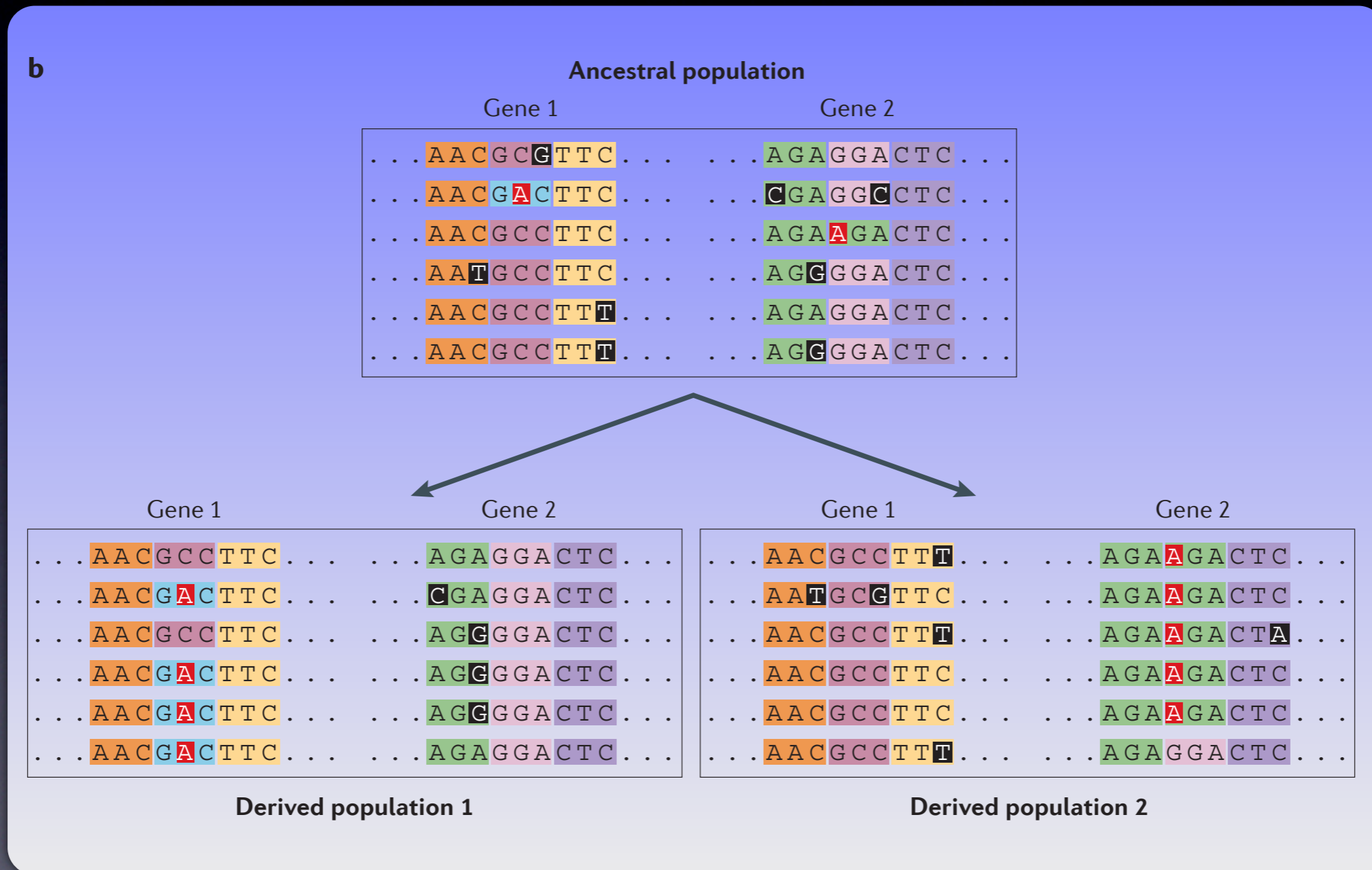


# Recombination Rate



Gore et al. 2009

# In Breeding Populations



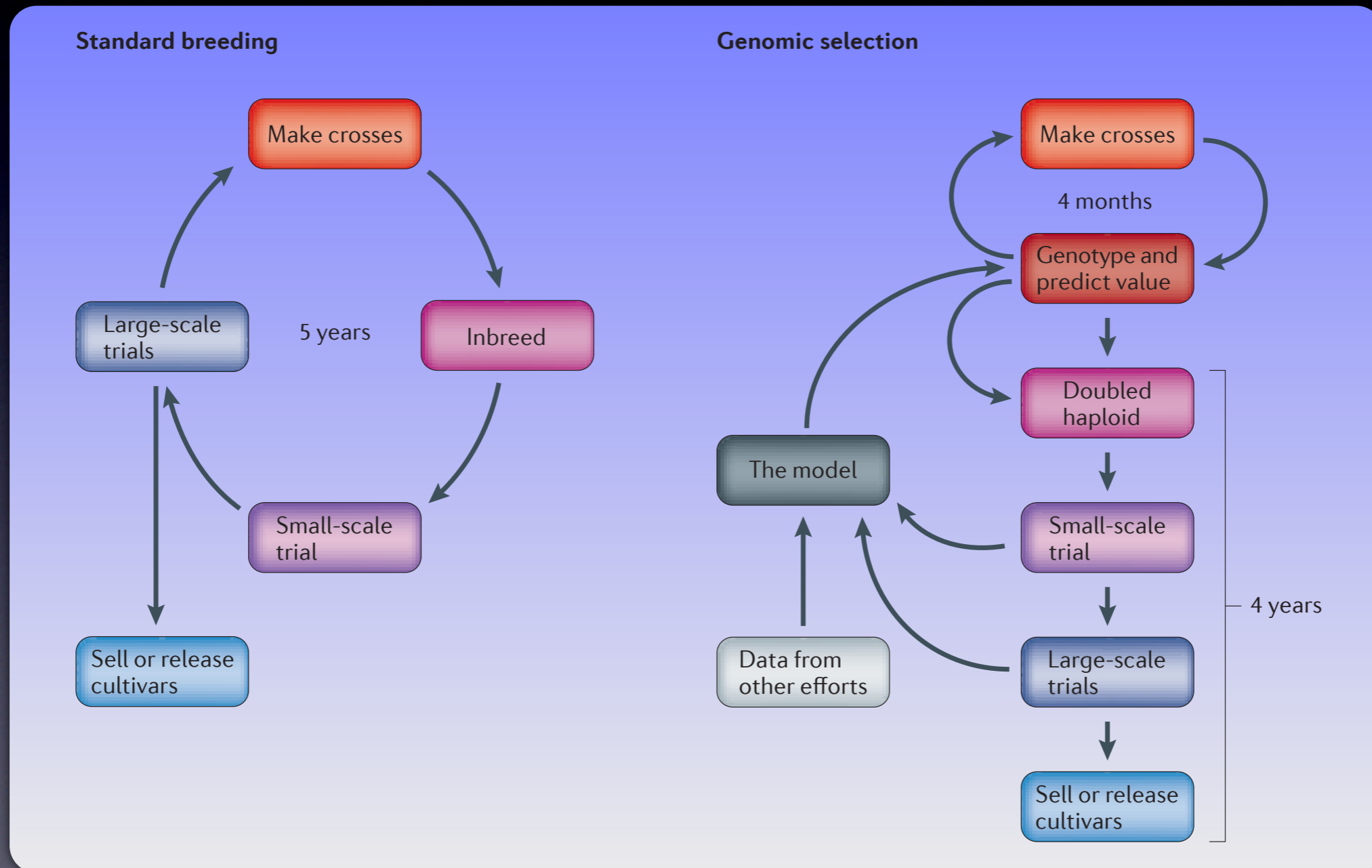
# Relevance to Crops

- The potential to purge deleterious mutations is related to population size
  - Many breeding programs have very small effective size
- Biggest diminution in fitness when deleterious mutations are homozygous
  - Barley, soybeans, and wheat...

# Applications

- ‘Reverse genomic selection’ against deleterious mutation could improve yield
- Deleterious changes in genes that aren’t directly associated to traits
- Targeted recombination - purge deleterious changes

# Genomic Selection

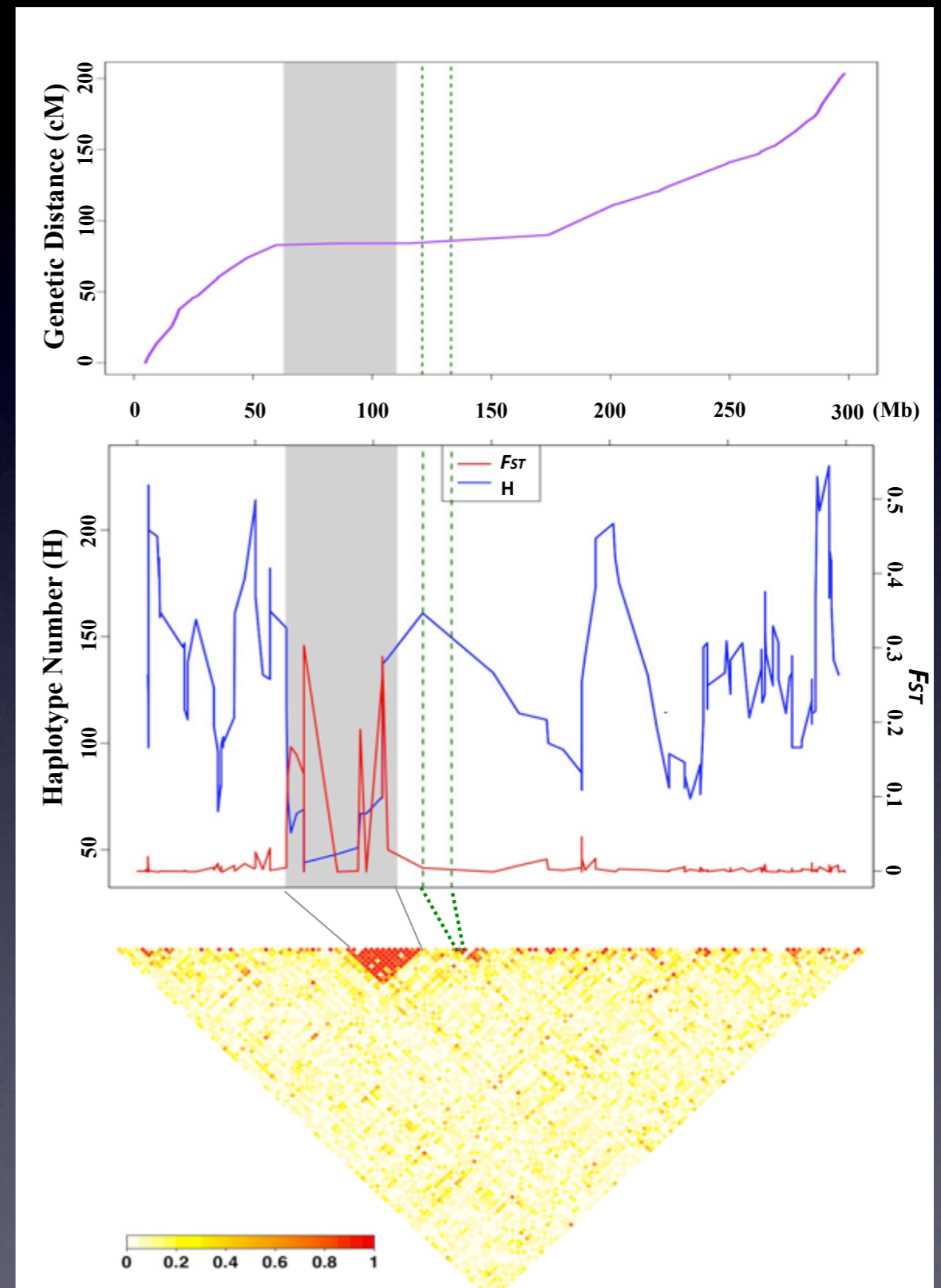


# Genome Level Analysis

- Extended linkage disequilibrium in teosinte
- Genetic provenance of climate adapted mutations in barley

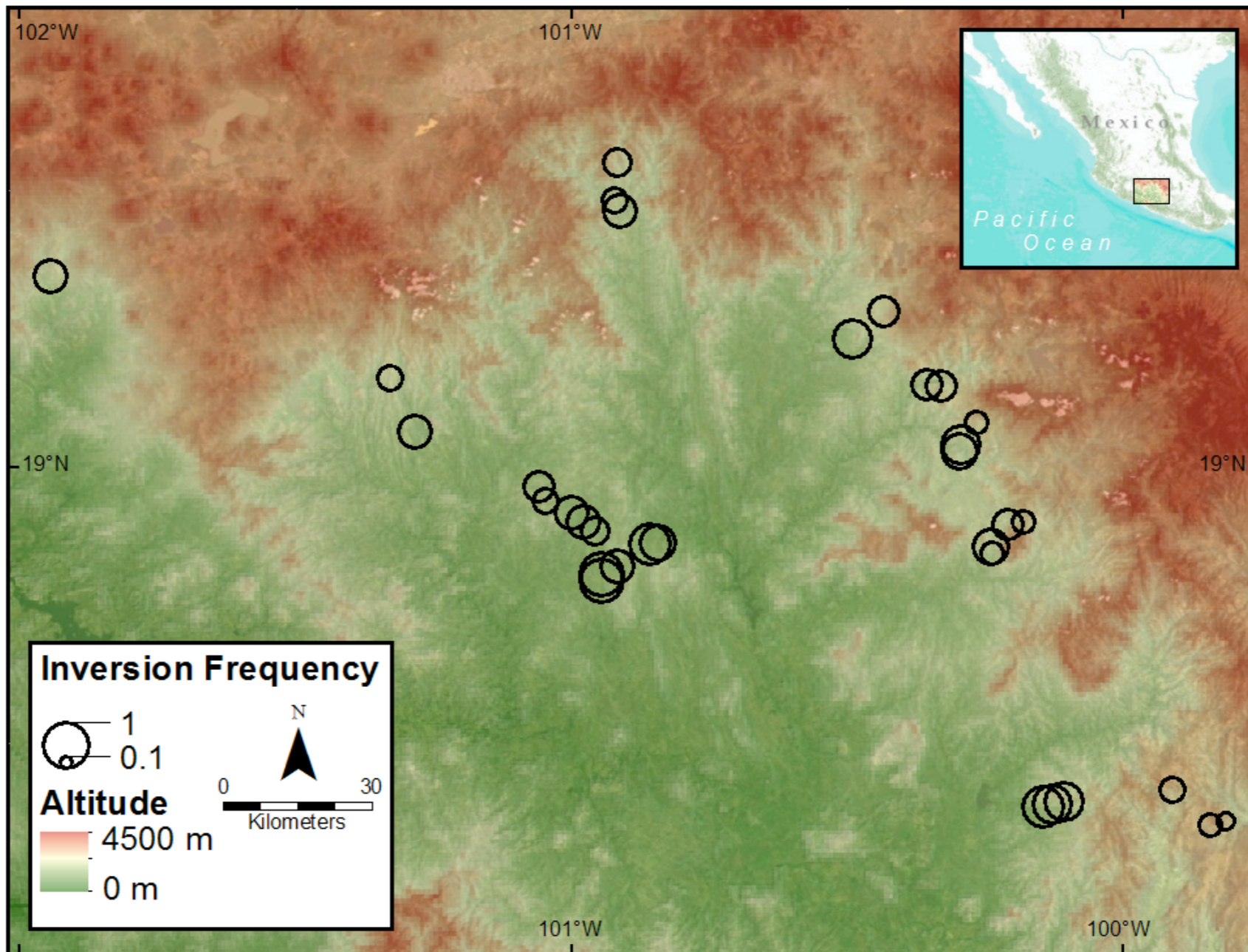
# Inversion Evidence

- Decreased recombination rate in inversion region
- Low haplotype numbers for the SNPs inside inversion
- High  $F_{ST}$  values for the SNPs inside inversion



Fang et al. (in review)

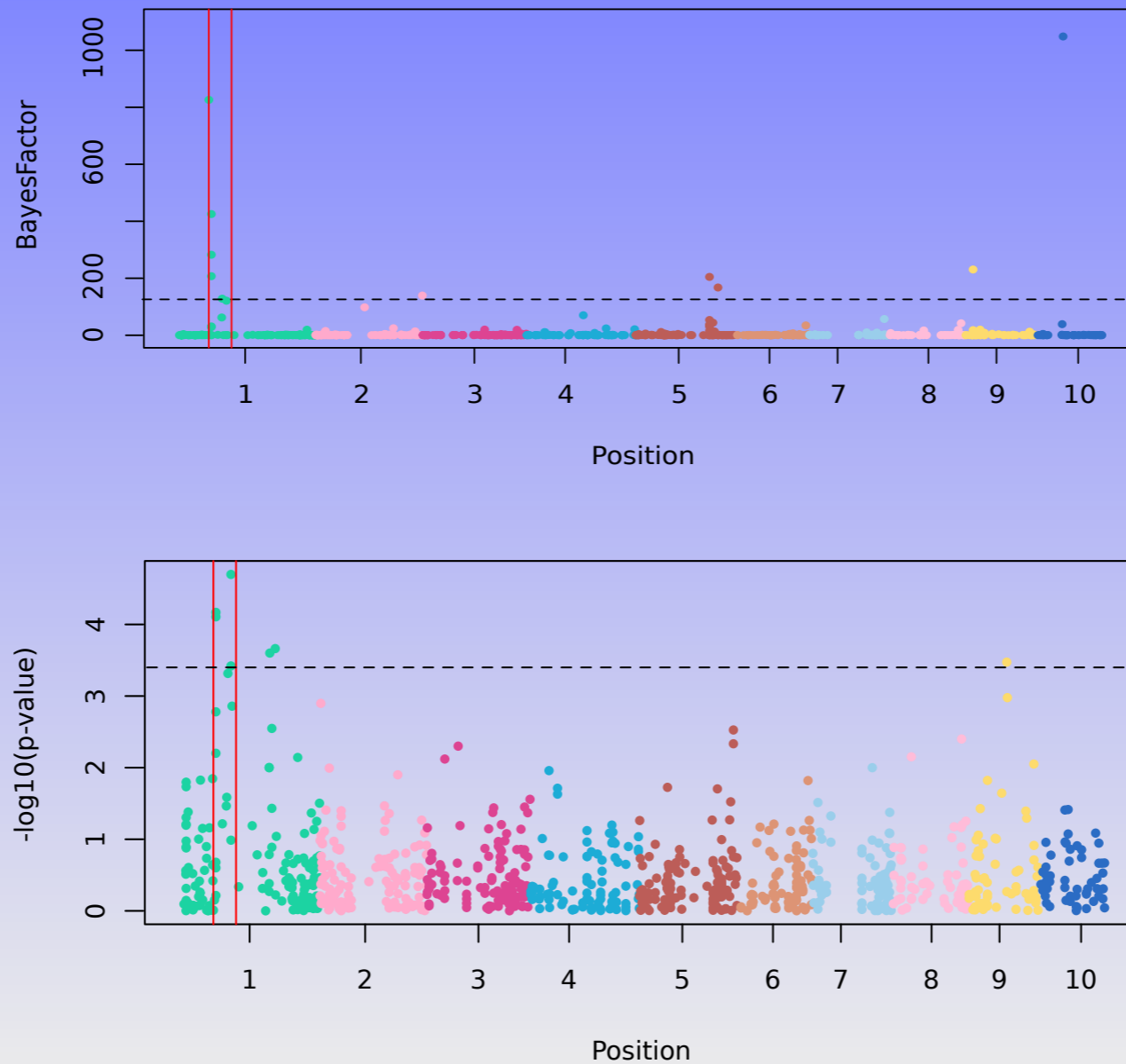
# Inversion Frequency



Fang et al. (in review)



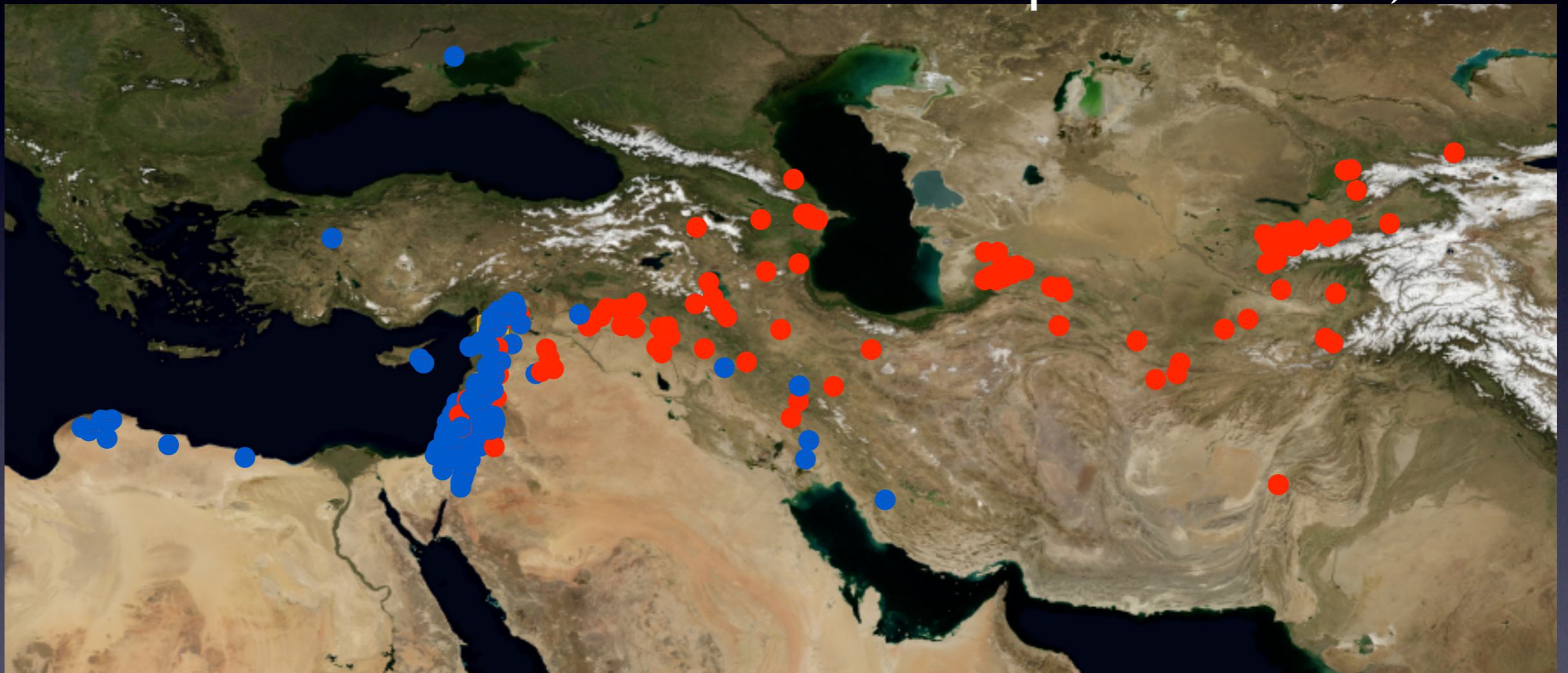
# Association - Altitude



Fang et al. (in review)

# Wild Barley Diversity Collection

Sample Sizes - 196, 116

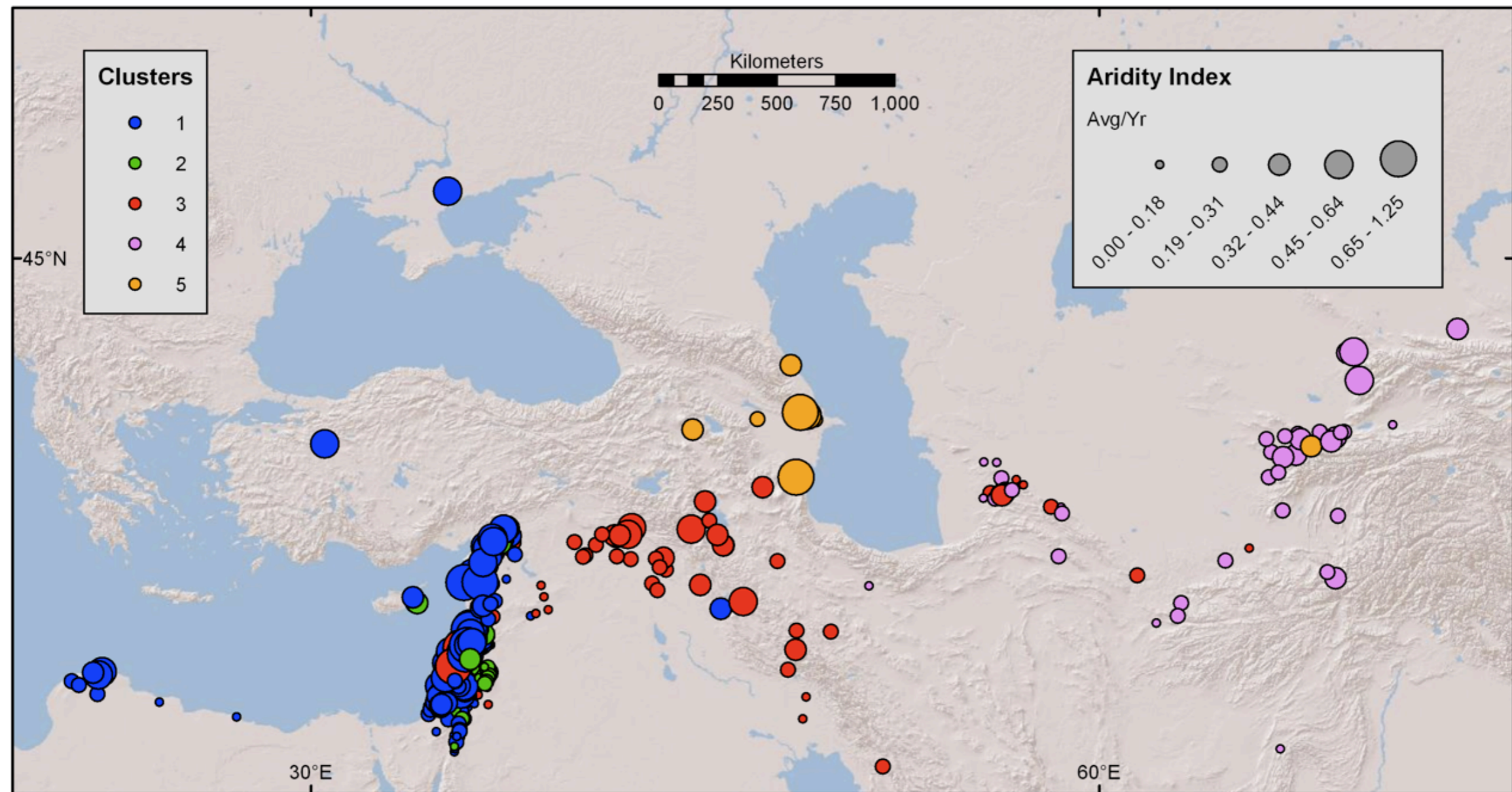


**Average**  
 **$F_{ST}$**

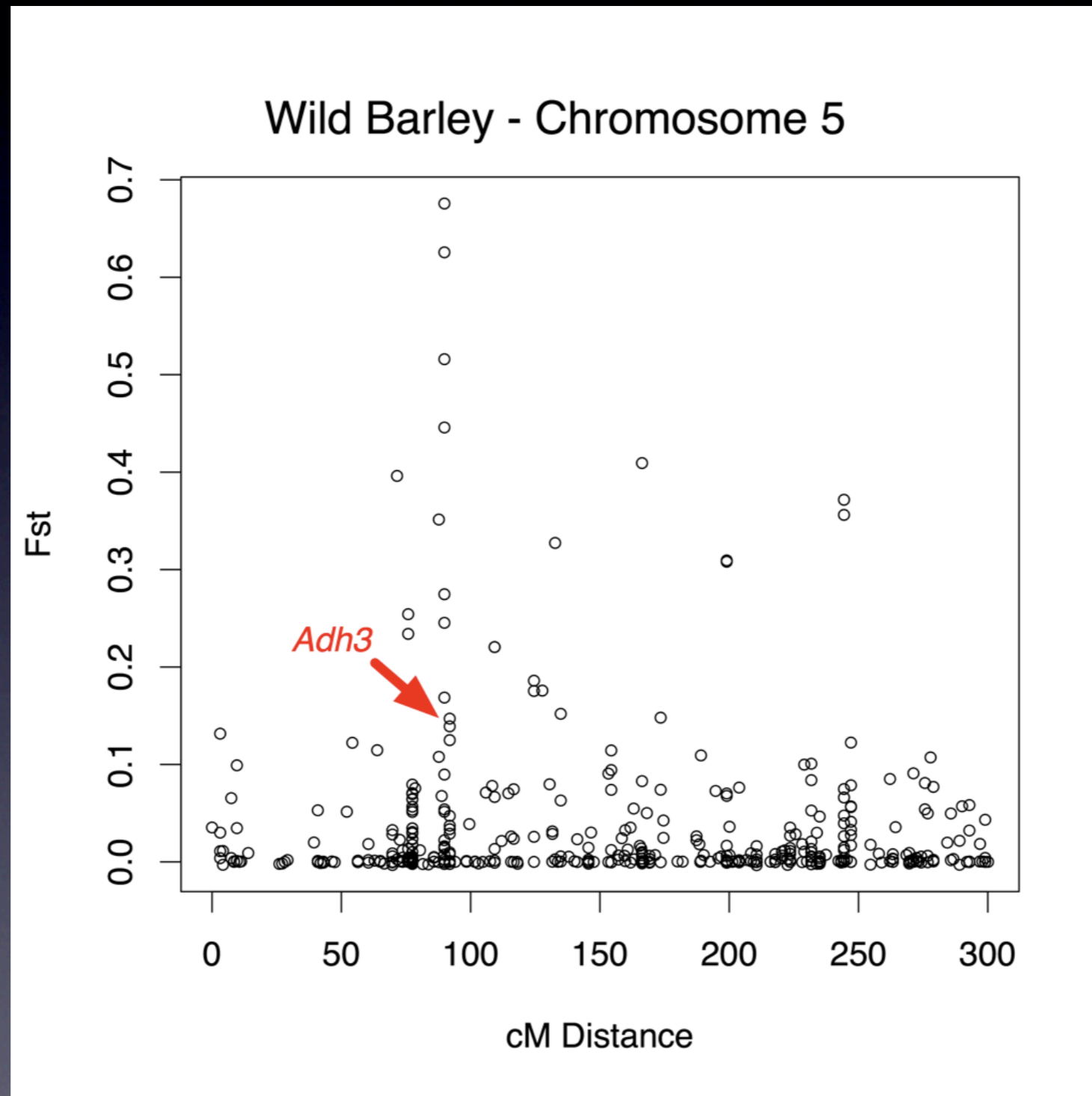
**BOPA1**  
**0.09**

**BOPA2**  
**0.13**

# Aridity - Western Asia



# $F_{ST}$ - Chromosomal



443 SNPs

# Opportunities

- Genomics era has just begun
  - Many patterns are easy to find
- SNP metadata 'data about data' is very important
  - Latent with information from places we can't visit!

# Acknowledgements

- Funding
  - USDA NIFA National Needs Fellowship
    - Soybean genetic load
  - USDA NIFA T-CAP
    - Barley
- Maize Inversions
  - Zhou Fang (UMN)
  - Tanja Pyhäjärvi (UC Davis)



See: <http://faculty.agronomy.cfans.umn.edu/pmorrell/>

# Gene Isolation

