

The genetic dissection of complex traits

Karl W Broman

Biostatistics & Medical Informatics
University of Wisconsin – Madison

www.biostat.wisc.edu/~kbroman

The goal

Identify genes that contribute to complex diseases

Complex disease = one that's hard to figure out

Many genes + environment + other stuff

The genetic approach

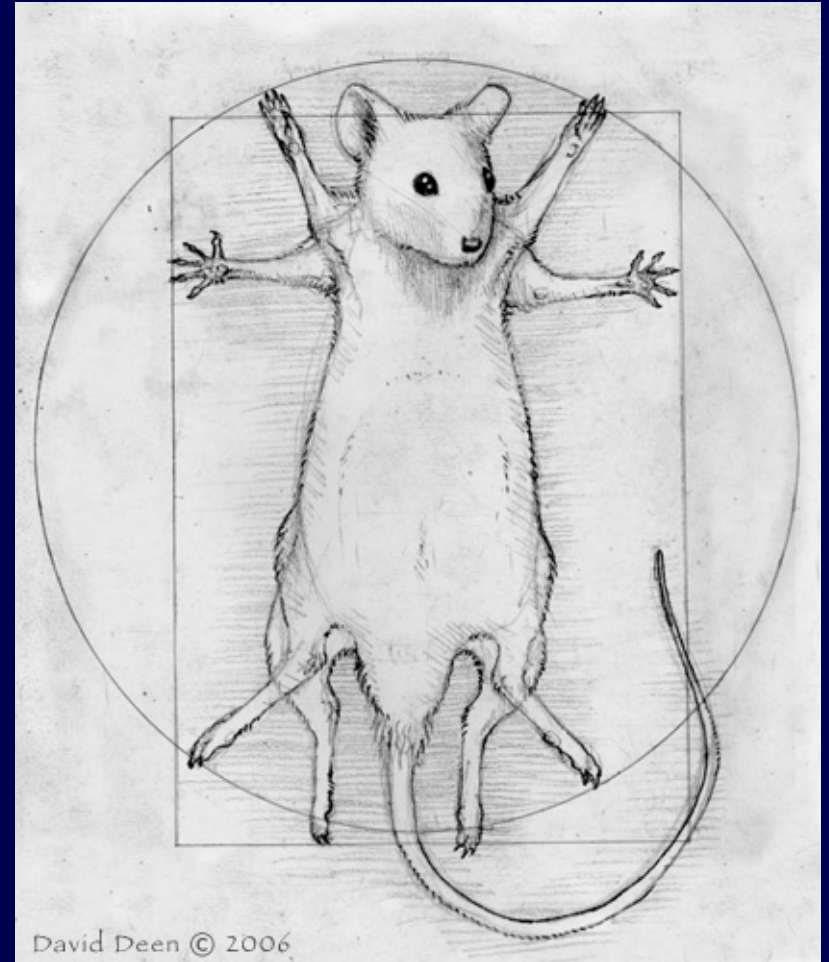
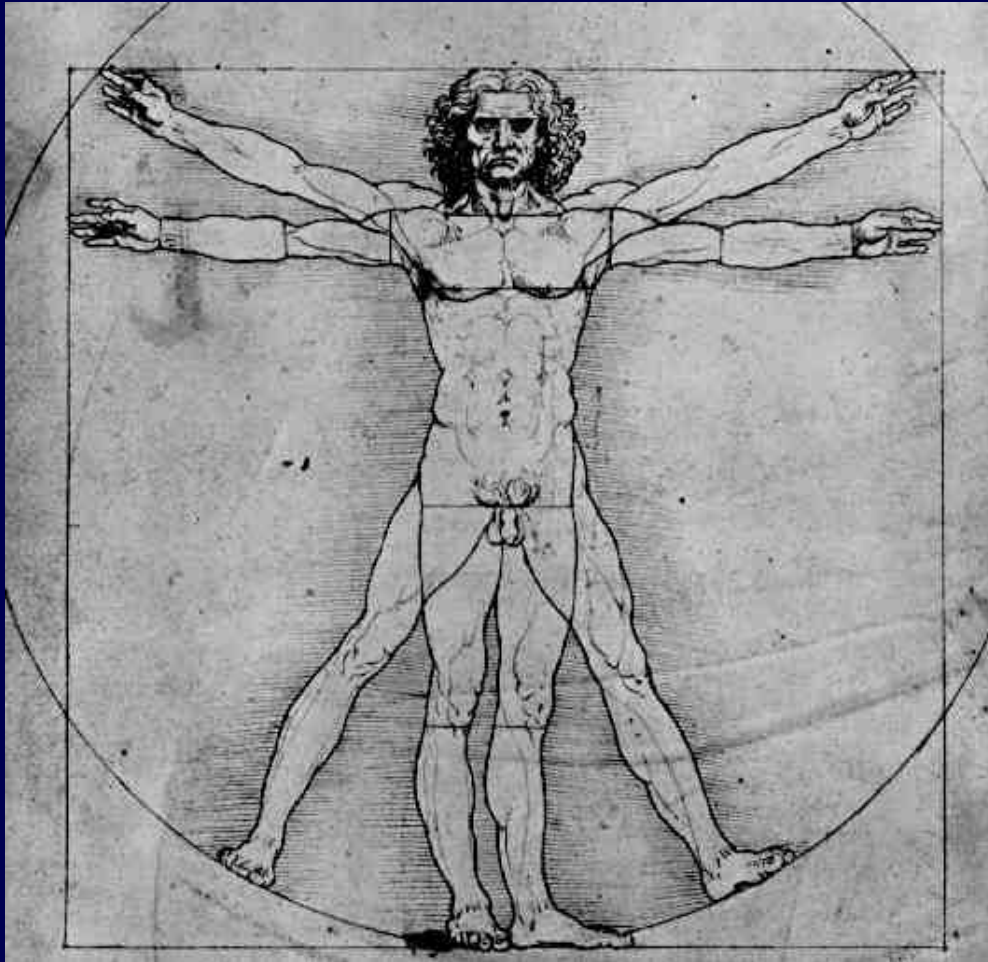
- Start with the trait; find genes that influence it
 - Allelic differences at the gene result in phenotypic differences
- **Value:** Need not know anything in advance
- Goal
 - Understand the disease etiology (pathways/mechanisms)
 - Prediction/prevention
 - Identify possible drug targets

Approaches

- Experimental crosses in model organisms
- Mutagenesis in model organisms
- Association analysis with inbred strains
- Linkage analysis in human pedigrees
 - A few large pedigrees
 - Many small families (e.g., sib pairs)
- Association analysis in human populations
 - Candidate genes vs. whole genome



Human vs mouse



www.daviddeen.com

Mutagenesis

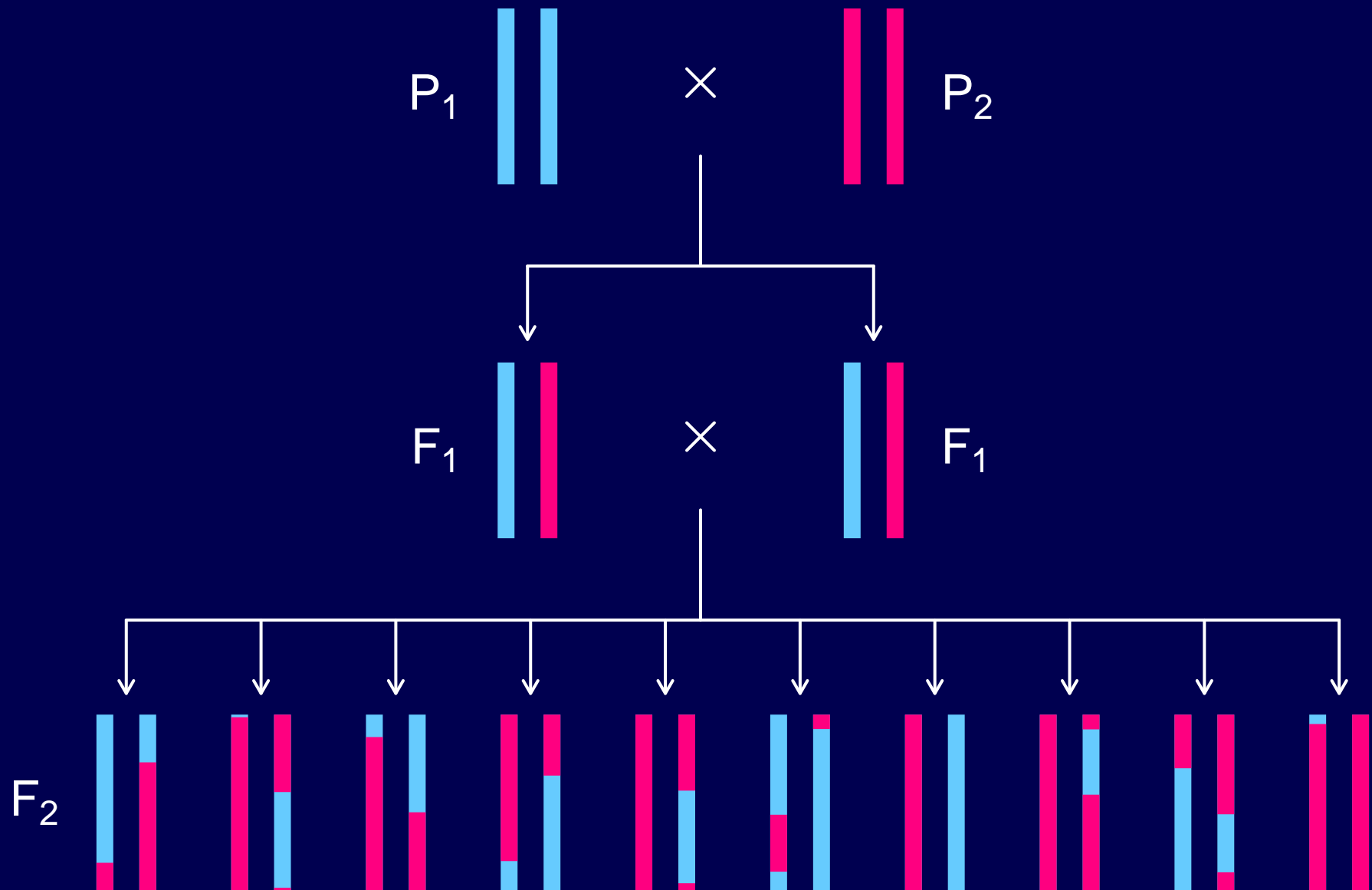
Advantages

- + Can find things
- + At the gene

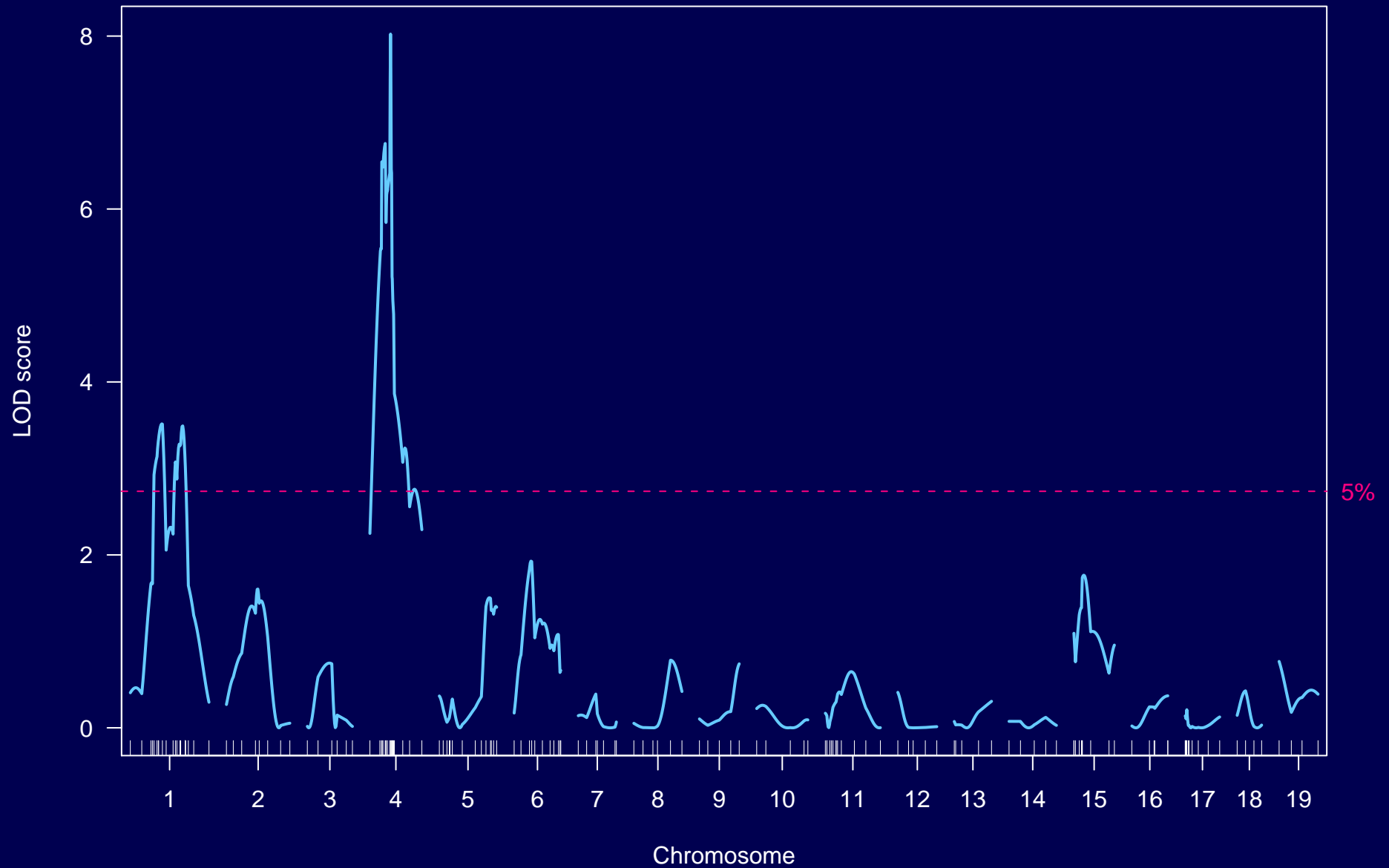
Disadvantages

- Need cheap phenotype screen
- Mutations must have large effect
- Genes found may be irrelevant
- Still need to map the mutation
- Recessive mutations are hard to see

Intercross



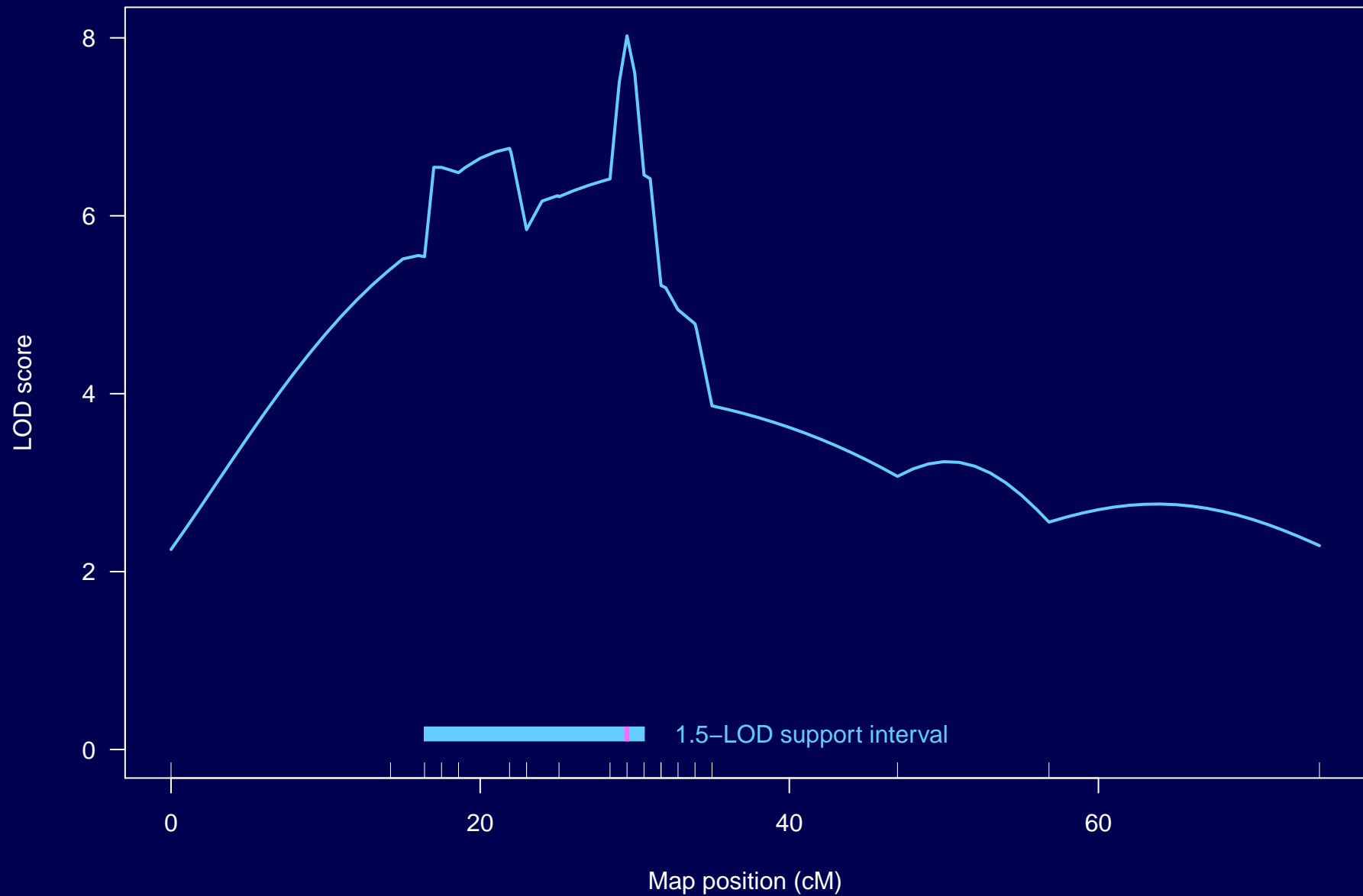
LOD curves



Joys of QTL mapping

- Genotype \rightarrow phenotype
- Recombination is cool
- Simple correlation structure
- Lots of opportunity for collaboration
- Lots of open problems
- Not many competitors

Chromosome 4



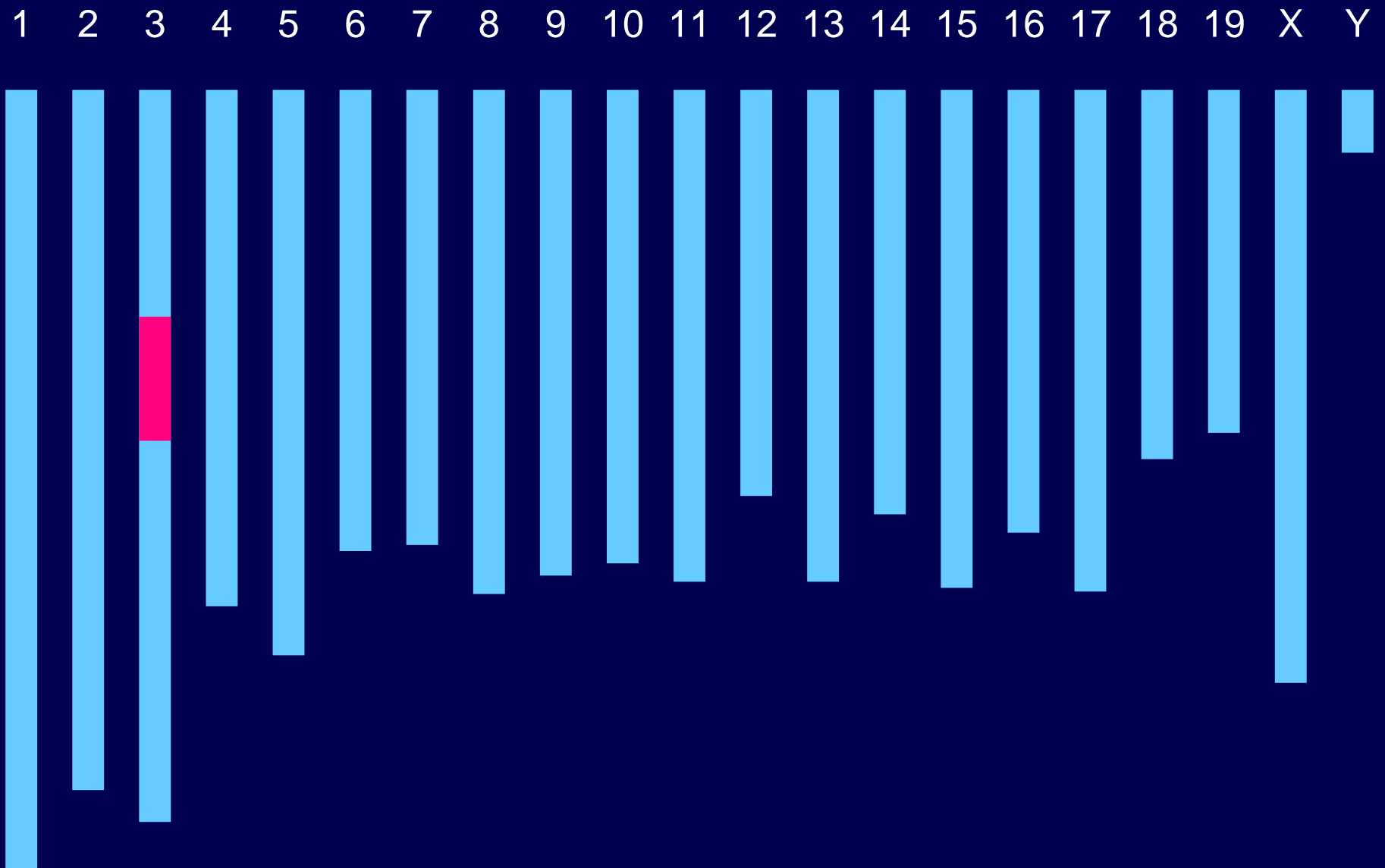
Traditional approach

F₂/BC → QTL → congenic → subcongenics

-
- coding/regulatory polymorphism
 - expression/function difference
 - knock-in / transgenic
 - knock-out
 - homology to other species

- Issues:
- Large QTL regions
 - Time consuming and expensive

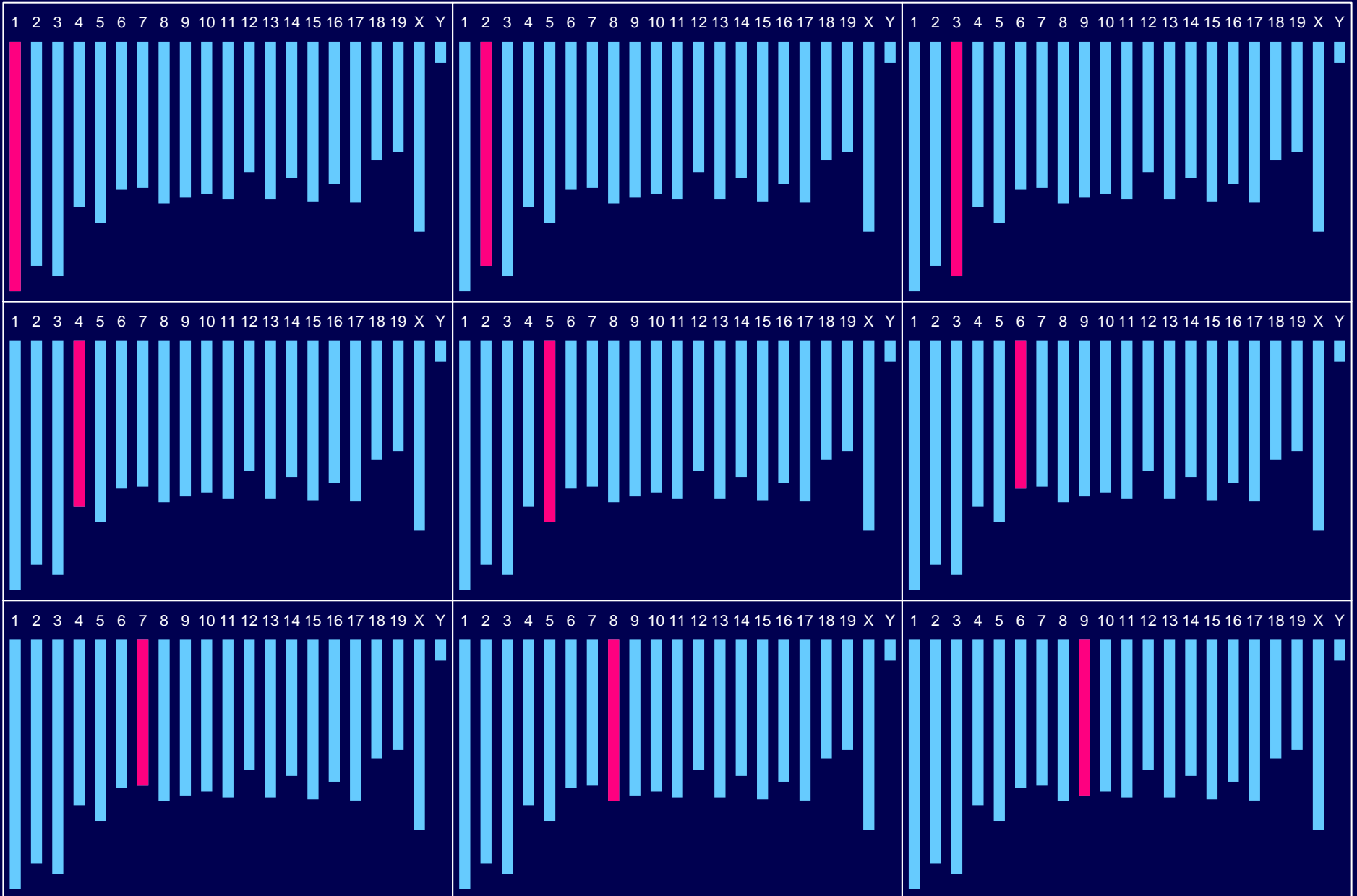
A congenic line



More modern approaches

- Gene expression data
 - and proteins, metabolites, epigenetic marks
 - and eQTL analyses
- Consomics (aka chromosome substitution strains)
- Panels of congenics
- Multiple crosses + haplotype analysis
- Cross-species comparisons
- Advanced intercross lines / Heterogenous Stock
- Recombinant inbred lines / Collaborative Cross
- Association mapping with inbred lines

Consomics



Consomics

Advantages

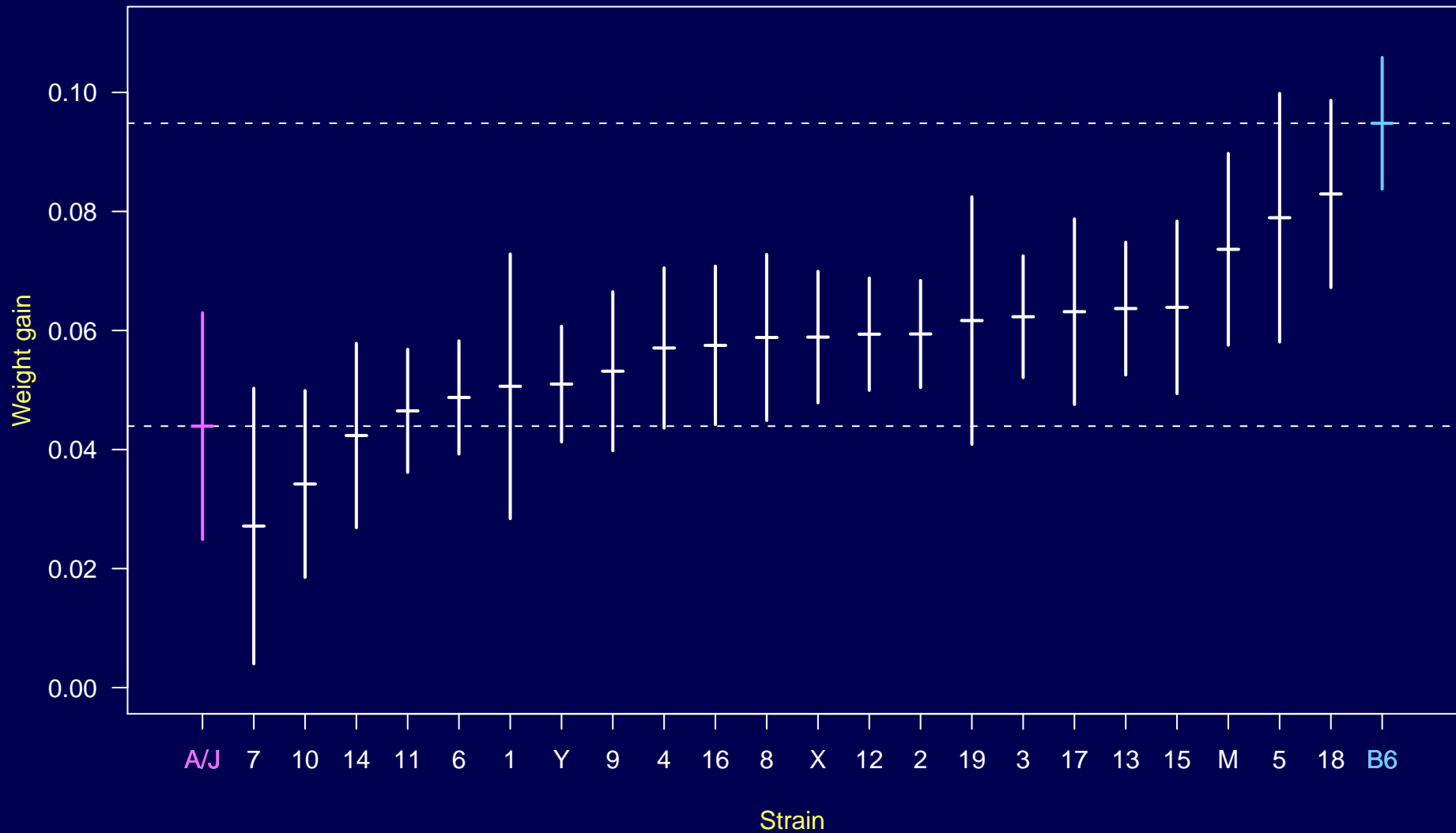
- + Just phenotyping can get you to the chromosomes
- + Eliminate the effects of other QTL
- + Easy to create congenics

Disadvantages

- Time-consuming, expensive to create
- Lots of phenotyping required
- Cannot see interactions

B.A on low fat diet

C57BL/6J – Chr^{A/J} / NaJ

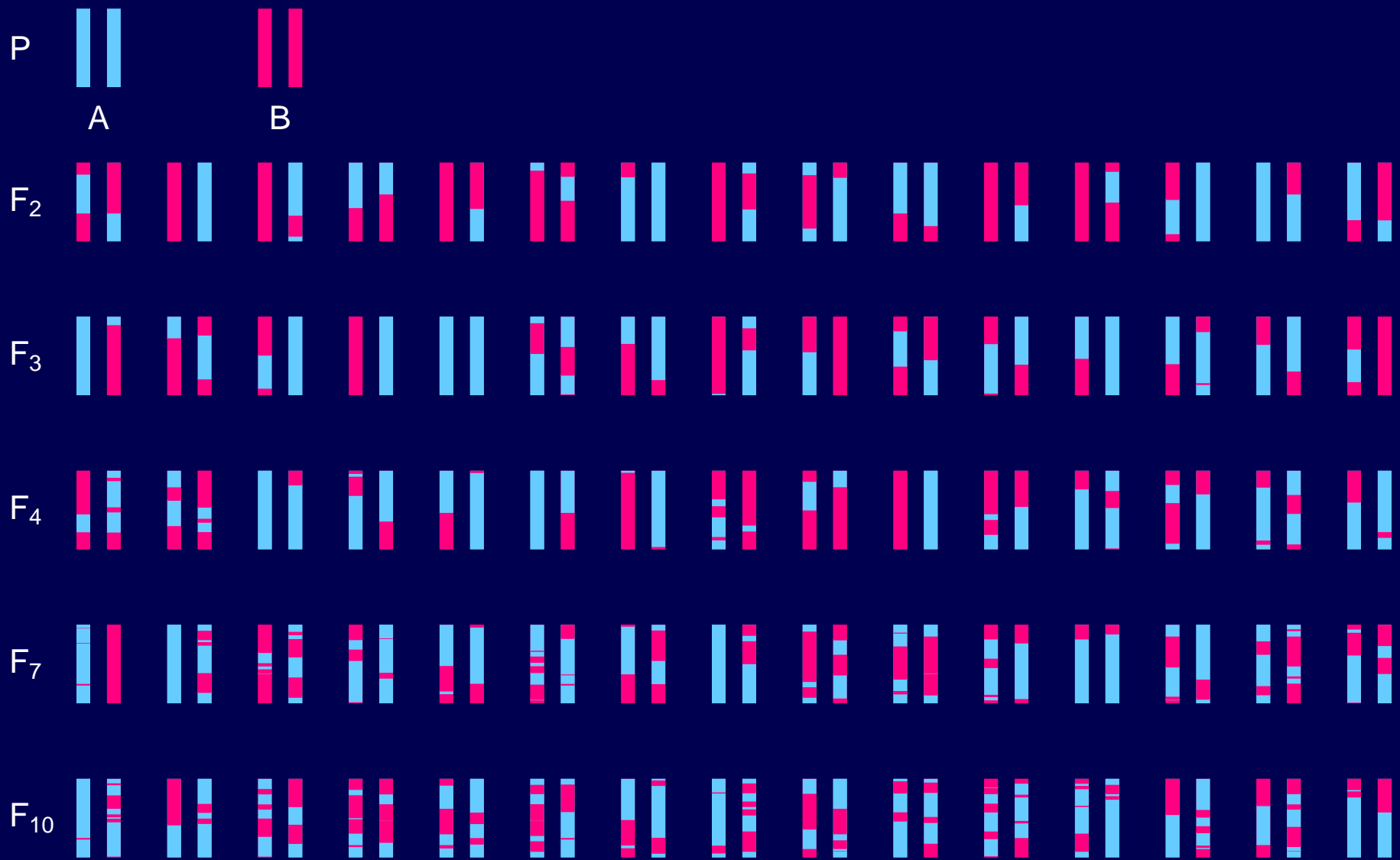


Shao et al., PNAS, 105:19910–19914, 2008

More modern approaches

- Gene expression data
 - and proteins, metabolites, epigenetic marks
 - and eQTL analyses
- Consomics (aka chromosome substitution strains)
- Panels of congenics
- Multiple crosses + haplotype analysis
- Cross-species comparisons
- Advanced intercross lines / Heterogenous Stock
- Recombinant inbred lines / Collaborative Cross
- Association mapping with inbred lines

Advanced intercross lines



AIL

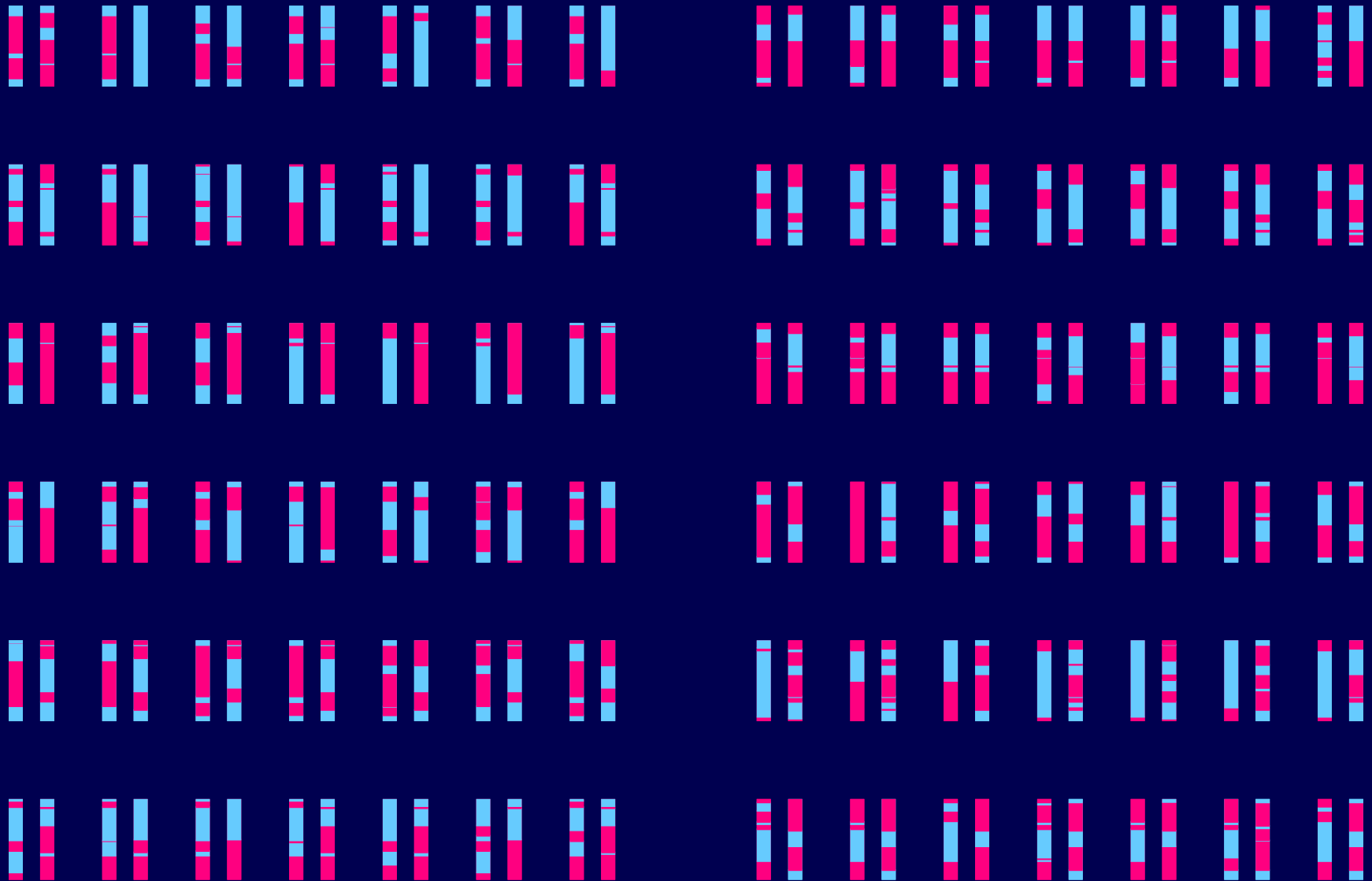
Advantages

- + Many more breakpoints \implies more precise mapping
- + Straightforward to create

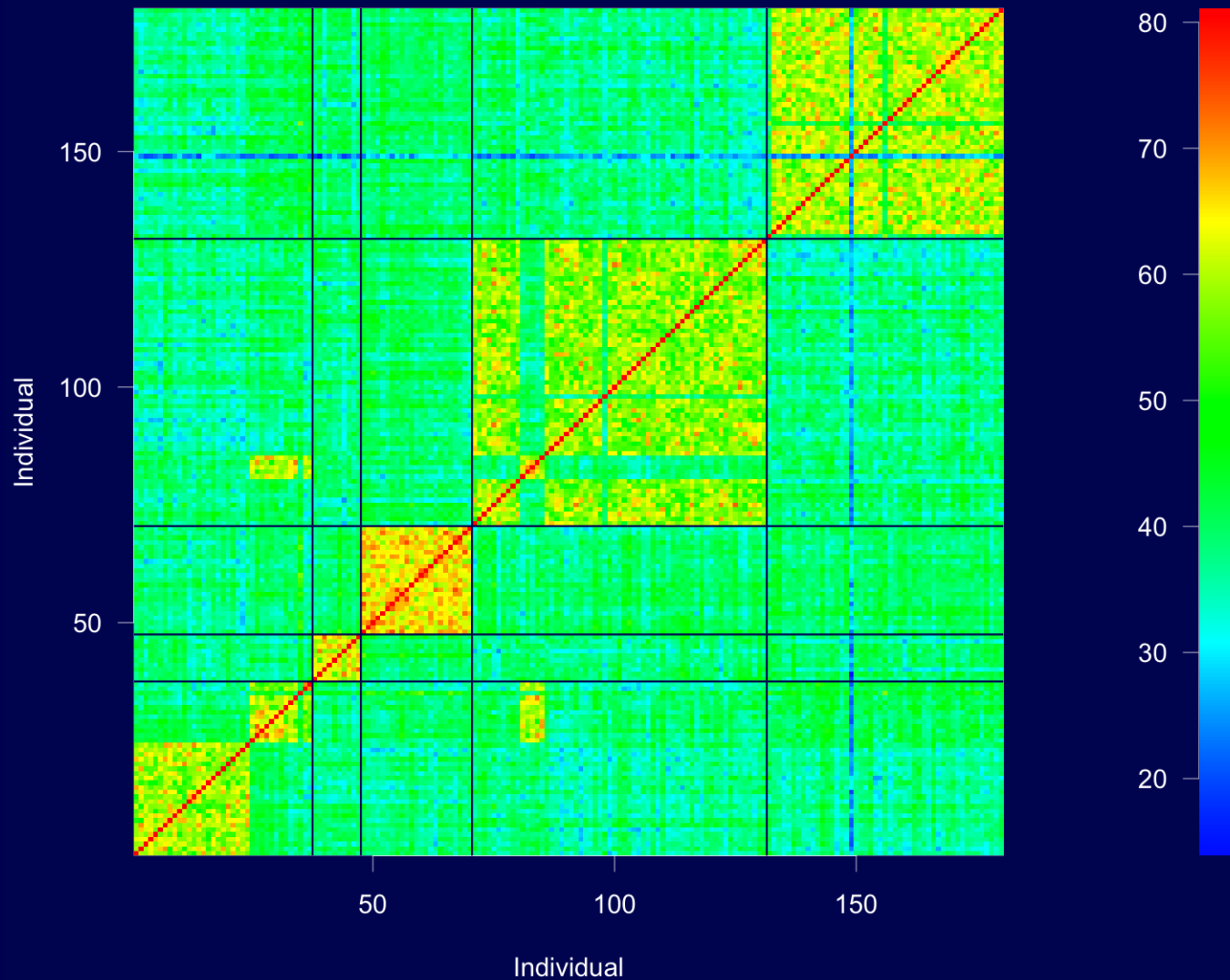
Disadvantages

- Time and cost
- Each individual genetically distinct
- Useful largely for fine-mapping known loci
- Relationships among individuals in final generation

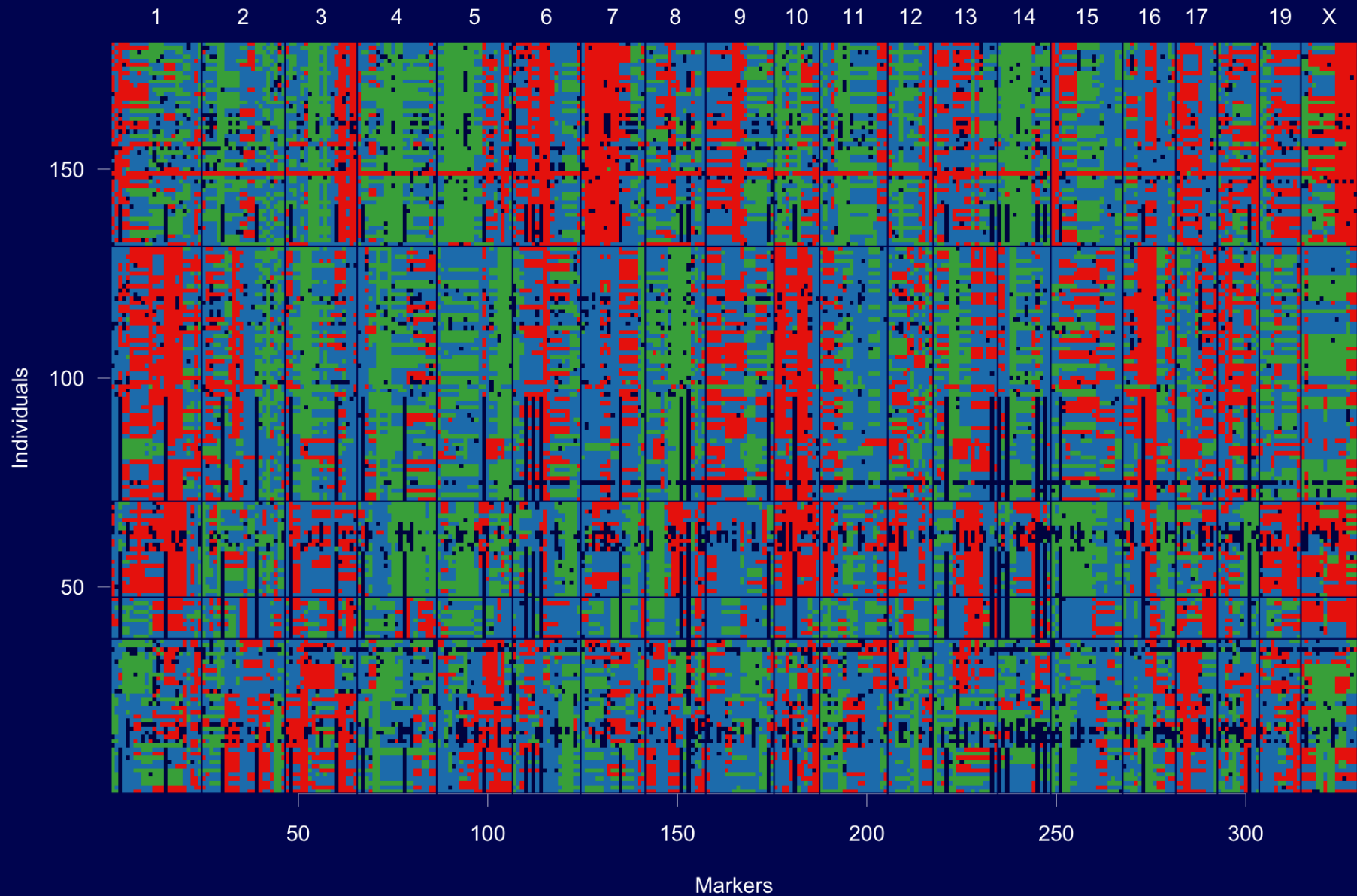
Sibships at F_8



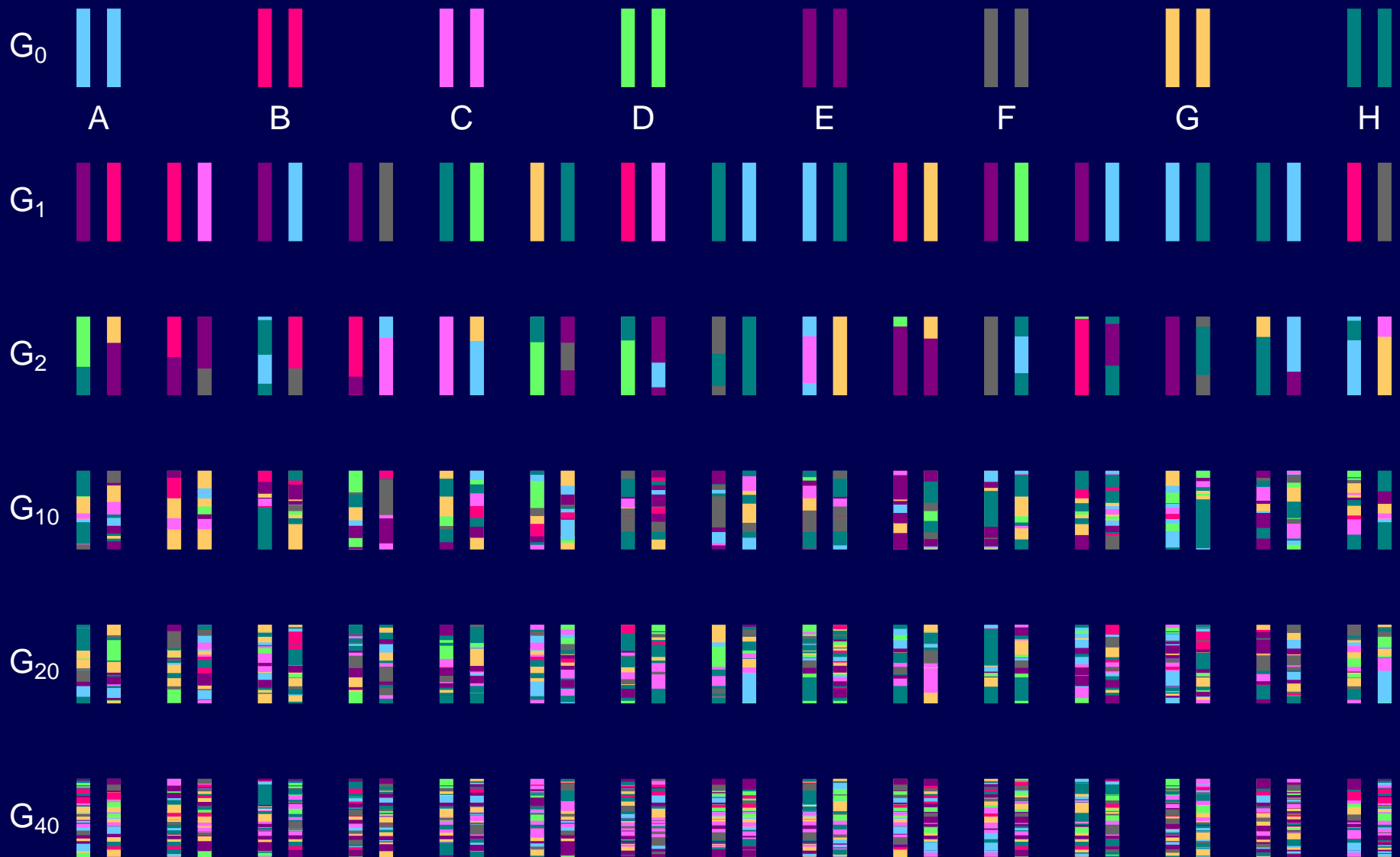
AIL: percent matching genotypes



AIL: genotype data



Heterogeneous stock



HS

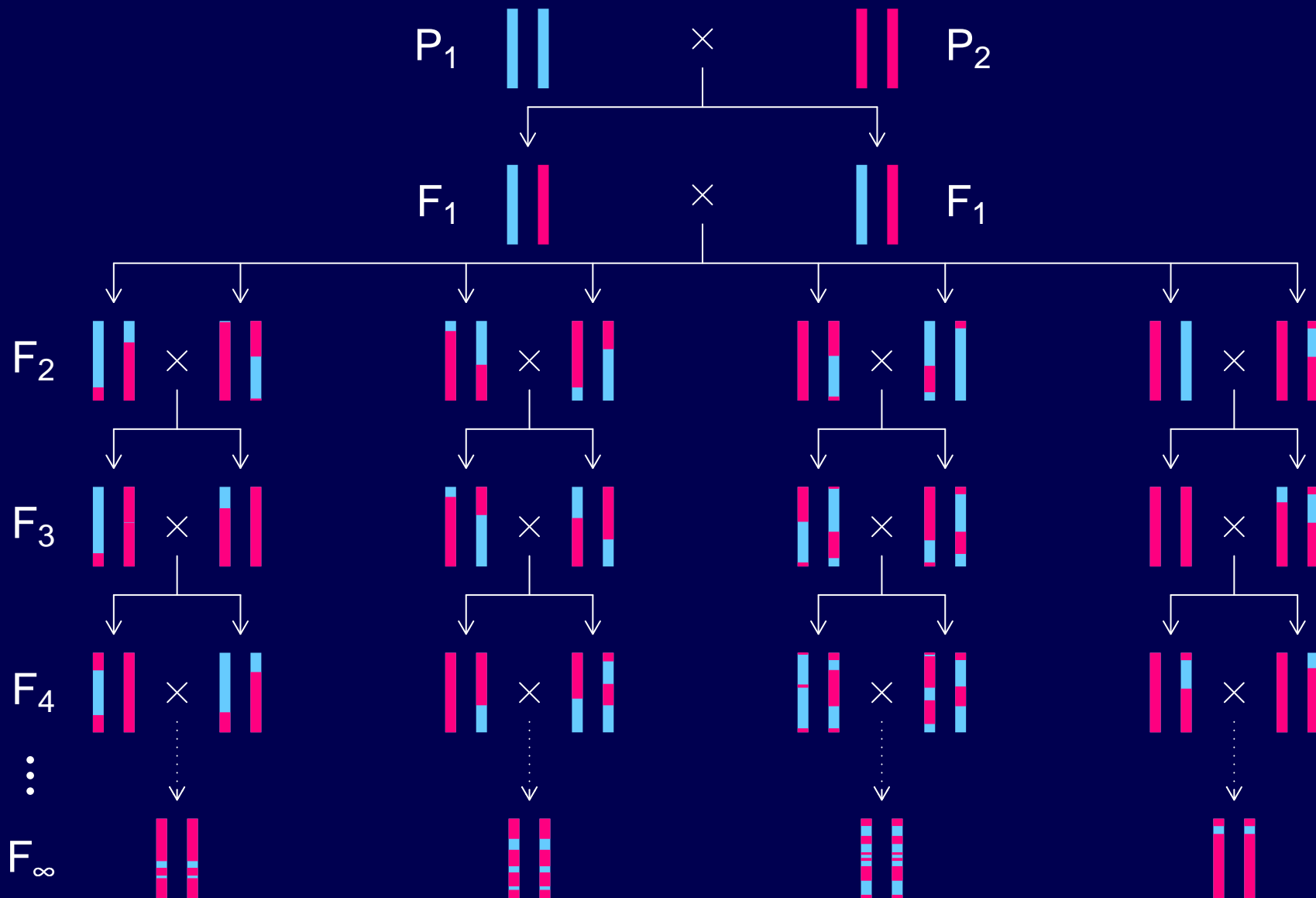
Advantages

- + Super-dense breakpoints
- + Many alleles
- + Heterozygous

Disadvantages

- Must be satisfied with what is available
- Inbreeding: loss of alleles
- Each individual unique
- Like AIL, maybe best for fine-mapping known loci
- Like AIL, relationships at last generations

Recombinant inbred lines



RIL

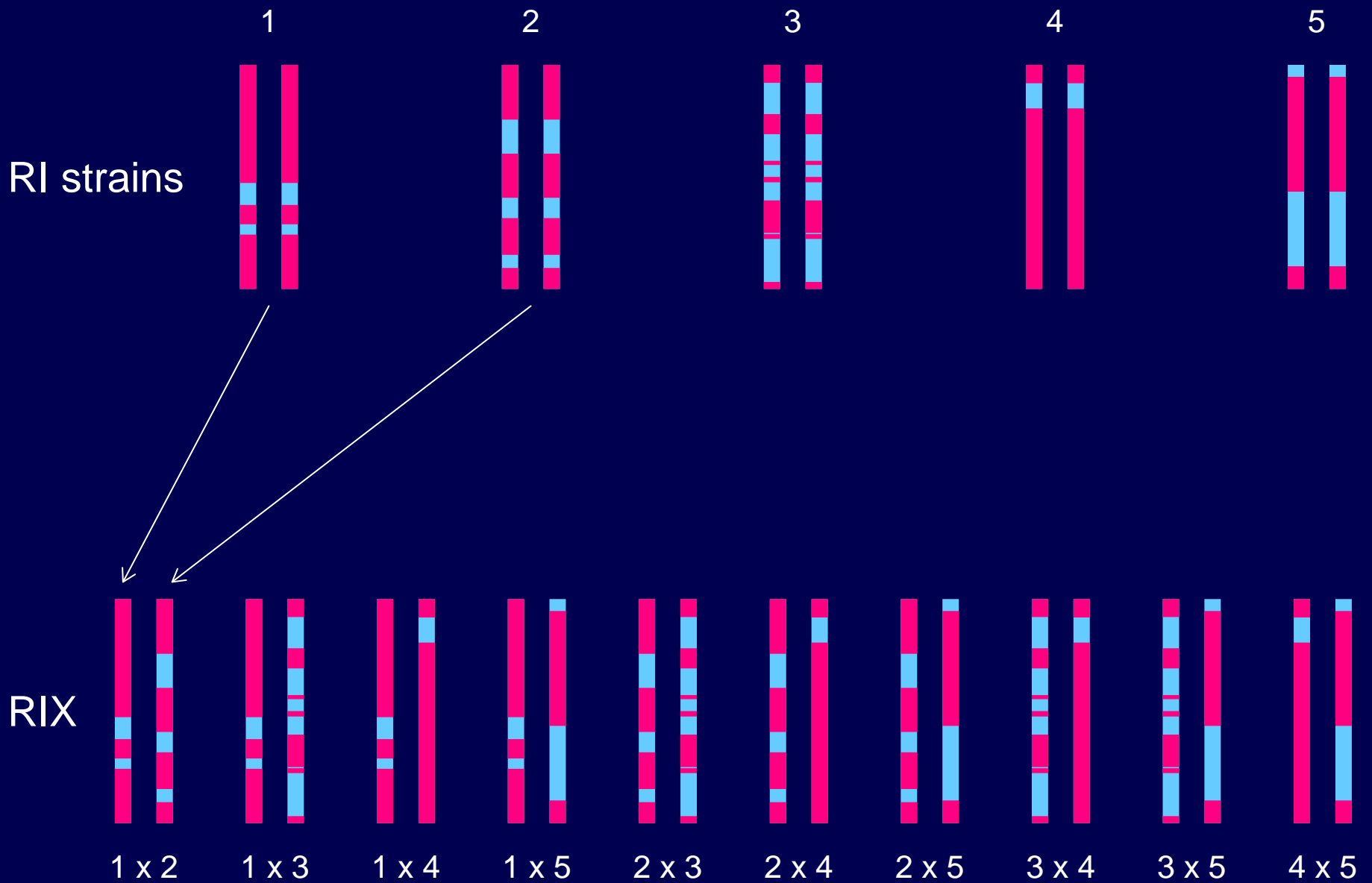
Advantages

- + High density of breakpoints
- + Just genotype once
- + Phenotype multiple individuals to reduce environmental/individual variation
- + Multiple phenotypes on the same genomes
- + Longitudinal phenotypes
- + Genotype \times environment interactions

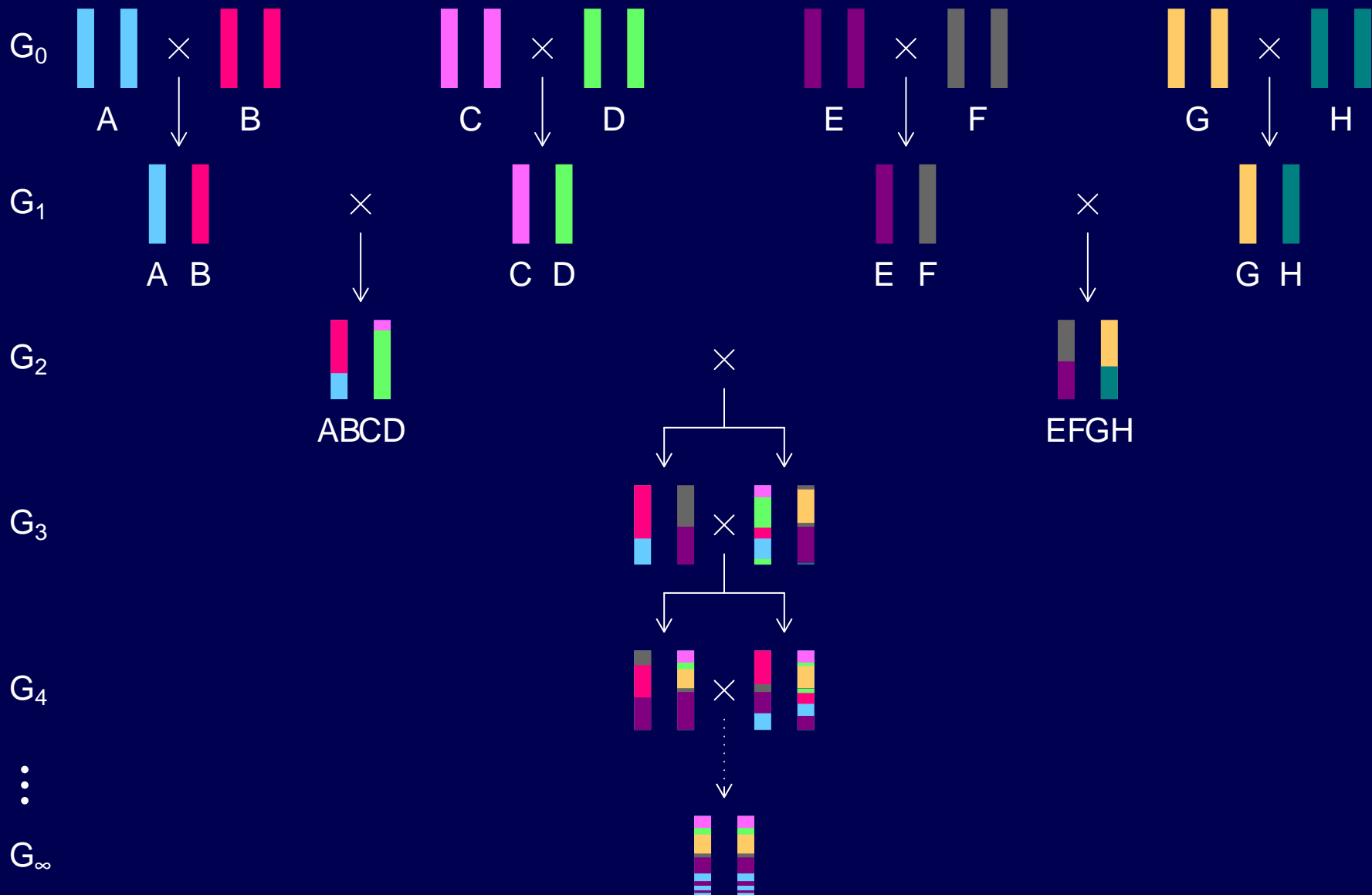
Disadvantages

- Time-consuming, expensive to create
- Available panels generally too small
- Only homozygotes

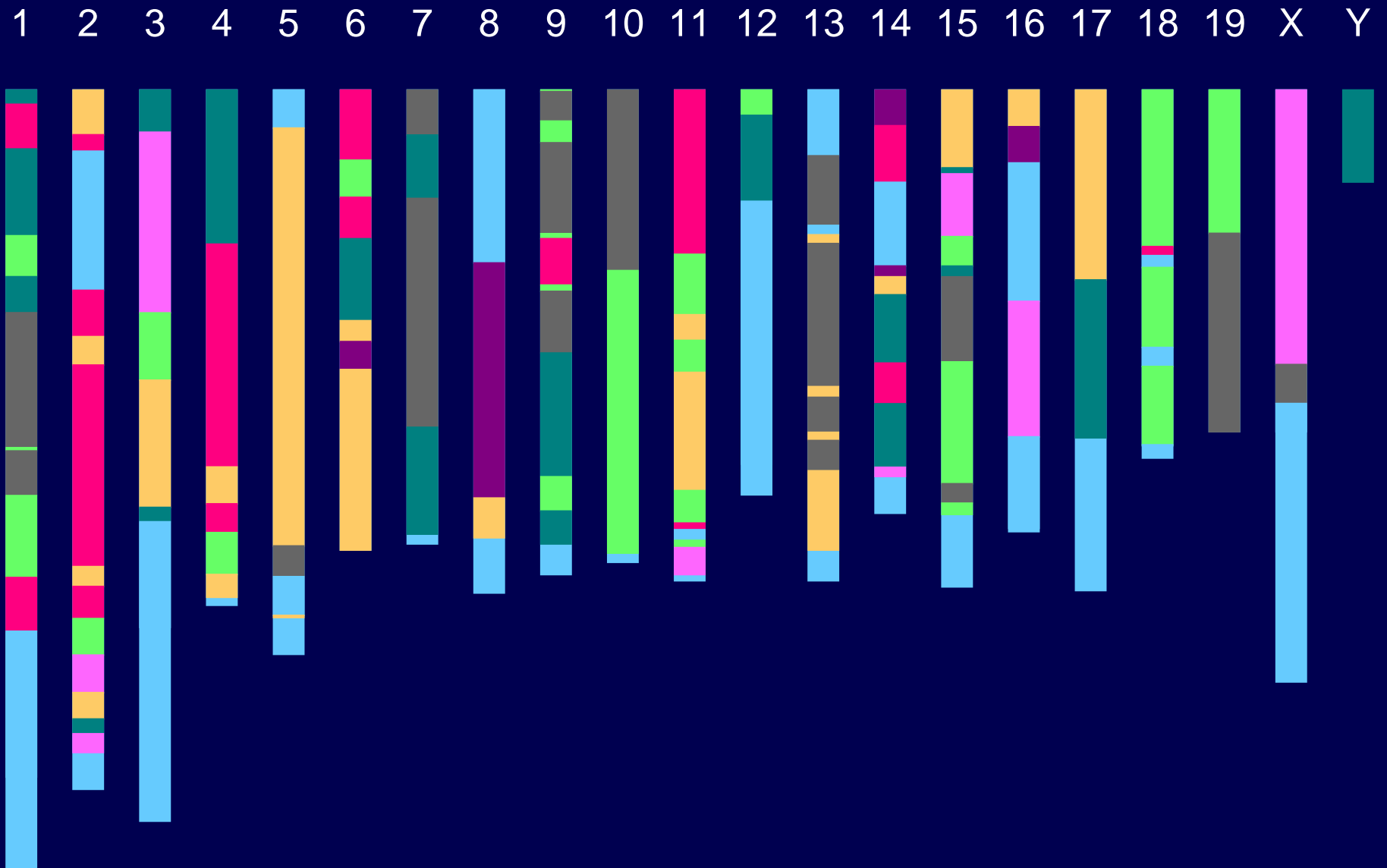
RIX



Collaborative Cross



A CC genome



Association mapping

- Phenotype available inbred strains
- Make use of available SNP data
- Need to account for the correlations among strains
- Likely want to work with haplotypes rather than just individual SNPs
- Be careful about wild-derived strains

Association mapping

Advantages

- + Once you've done a strain survey, no further data needed
- + Potentially very high resolution

Disadvantages

- All the usual problems with association mapping
- Power is unpredictable
- How to account for relationships among strains?

CC vs HS vs association mapping

These approaches have many similarities.

Key differences:

- CC, HS: pattern of association along chromosomes **by design**
- HS: each individual unique

Summary

- QTL mapping in mice is fun and useful
- But we need to be able to get to gene
- There are lots of new strategies to speed things along
- Numerous interesting statistical problems remain
- Need to consider human GWAS