Mapping a *simple* genetic trait relative to genetic markers in a population

Can we map diseases without controlled crosses/ pedigrees?

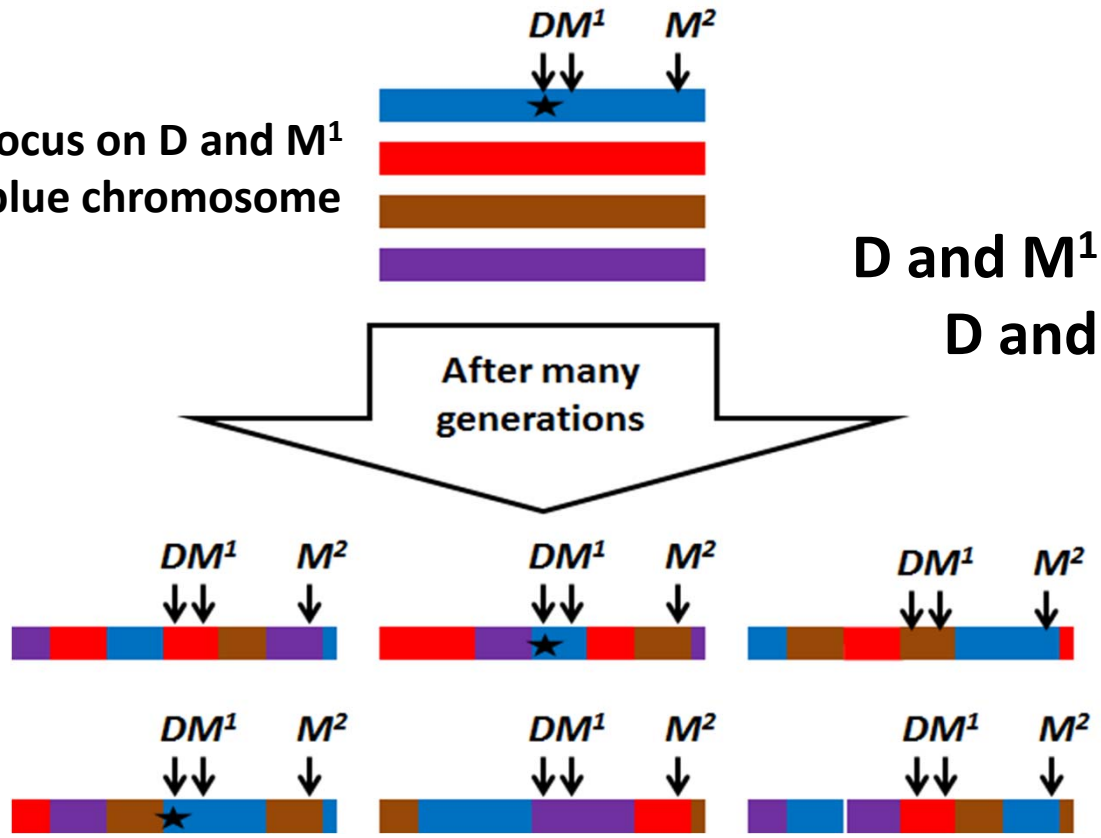
- Not socially acceptable to force people to mate with others, especially ones who have diseases...
- Can't always find enough subjects who happened to breed in an informative way.
- Another route?

Use “population data” over time

- Human generation time historical average ~20 years
- When looking at very close genes, we saw low probability of exchange in ONE generation
  - BUT, over the thousands of years of history, there has been A LOT of recombination
    - Even 0.1% exchange in 1 generation has 99.3% probability of exchange in 5000 generations
  - Most neighboring genes shuffled, and even areas within genes are sometimes shuffled

Start “population” with 4 chromosomes

We’ll focus on D and M^1 in the blue chromosome

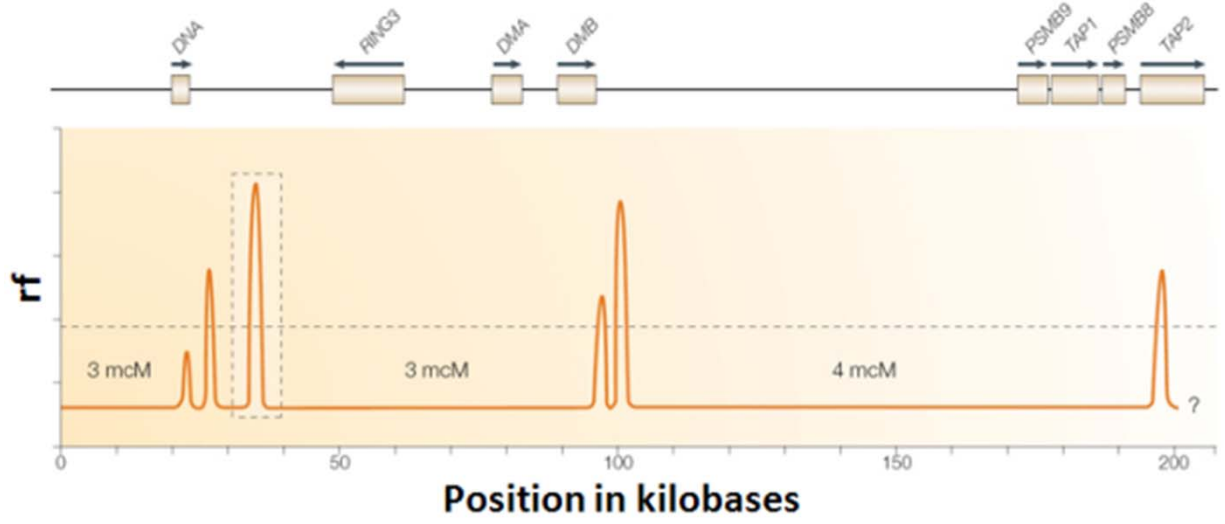


D and M^1 allele stay associated

D and M^2 allele shuffled

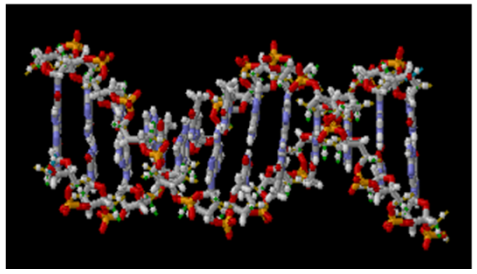
… AND, recombination is not homogeneous when look at a very fine scale

- Tends to occur in “hotspots” every few thousand bases; rest of genome is ~0 rf

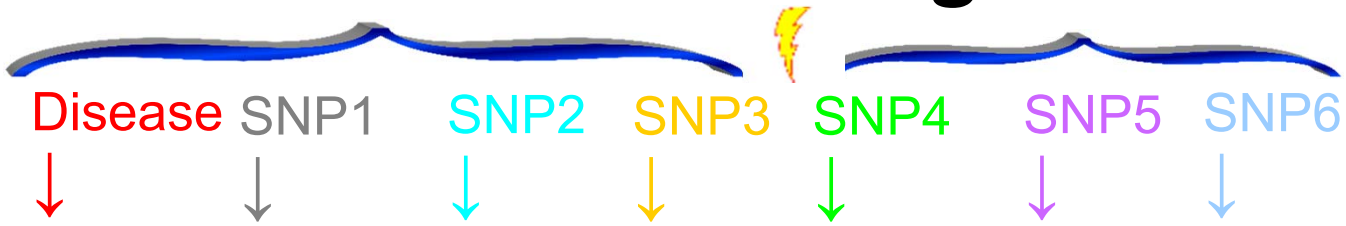


What are the implications?

- Shuffling occurs between the “windows” every few thousand bp
- Virtually no shuffling occurs WITHIN windows
  - Said to be in “linkage disequilibrium” (LD)
- Some windows contain “disease genes”
- We can leverage these features to find disease genes!



One Marker May Rarely Recombine From Its Close Neighbors



CAGATCGCTGGATGAATCGCATCTGTAAGC

CGGATTGCTGCATGGATCGCATCTGTAAGC

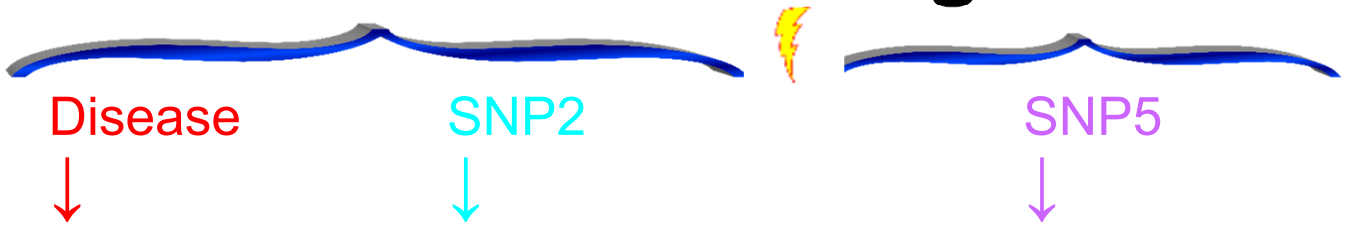
CGGATCGCTGGATGAATCGCATCTGTAAGC

CAGATCGCTGGATGAATCCCATCAGTACGC

CGGATTGCTGCATGGATCCCATCAGTACGC

CGGATTGCTGCATGGATCCCATCAGTACGC

One Marker May Rarely Recombine From Its Close Neighbors



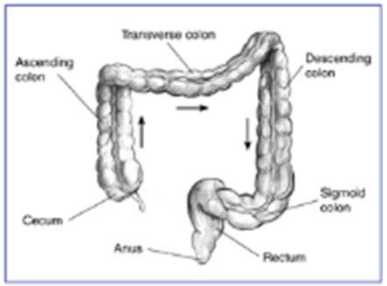
PREDICTION:

- Disease gene mapping is associating a "genotype" (marker allele at one location) with a "phenotype" (disease)

- **If a marker is very close to the disease-causing gene, individuals having one allele will be more likely to have the disease than individuals having the other allele**
  - **The marker is in "LD" with the disease gene**

- Does this mean the marker gene or SNP causes the disease?

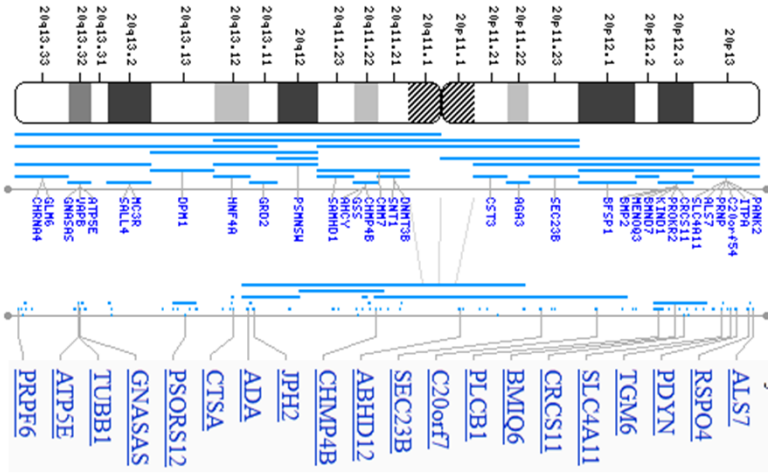
Example



- Sample the population for alleles at 2 markers and for incidence of irritable bowel syndrome (IBS)
  - For simplicity, let's ignore heterozygotes
- SNP 1: AA- 100 individuals, 20 w/ IBS
  aa- 200 individuals, 40 w/ IBS
- SNP 2: BB- 50 individuals, 45 w/ IBS
  bb- 250 individuals, 15 w/ IBS

but the genome is so big…

- ~3 billion bases in the genome
- If there's a hotspot every ~3000bp, how many markers would you need to study to find disease genes?
  - About how many "windows" are there?



Technology helps!

- Not a problem! We have "microarray chips" that can tell us genotypes at >1 million markers at once from spit!

- Companies will do this for you at low cost, and tell you susceptibility to many mapped diseases!

23andMe

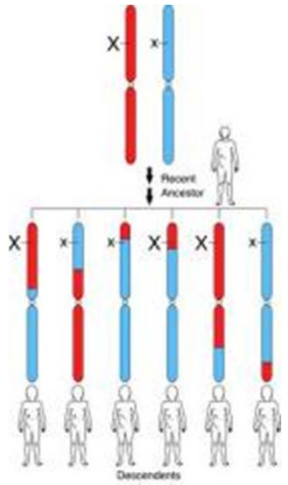
Example for you to try

- Sampled 1000 people: 950 healthy; 50 w/ CF
- Marker1: **AA-** 600 people, 28 w/ CF
  aa- 400 people, 22 w/ CF
- Marker2: **BB-** 750 people, 39 w/ CF
  bb- 250 people, 11 w/ CF
- Marker3: **CC-** 100 people, 45 w/ CF
  cc- 900 people, 5 w/ CF
- Marker4: **DD-** 800 people, 42 w/ CF
  dd- 200 people, 8 w/ CF

So…

- CC genotype at Marker 3 causes CF, right?
- Why do some people with “cc” still have CF?
  - **MANY answers** (remember title of italicized word in video title)

Distinctions between association mapping and cross/ pedigree mapping

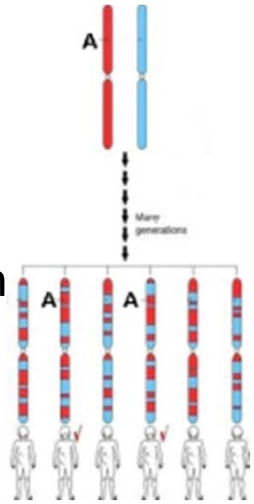


Cross

Maps in known families
Resolution ~2 million bp
1 gen of recombination
Works even if mutation rare

Population

Maps across population
Resolution ~3000bp
MANY gen of recombination
Fails when mutation rare



For reference, two sides of hotspot probably are <0.01cM apart.

Image Credits, Unit 5-3

- Chromosomes, © Alexandr Mitiuc, all rights reserved, www.photoxpress.com
- Recombination graph, © 2004 Liisa Kauppi, Nature Reviews Genetics 5, 413-424 (June 2004), all rights reserved, reprinted with permission
- Microarray, © 2006 Schutz, CC by-SA 3.0, en.wikipedia.org.
- Human diversity, © Kurhan, all rights reserved, www.photoxpress.com
- Haplotype, © Images courtesy of Hua Zhang and The International HapMap Consortium. The International HapMap Project. Nature 426, 789-796 (2003).
- Colon, Image courtesy of NIH.