Principles of mapping complex traits

Most variable traits people study are controlled by many genes

- Fruit size
- Human height
- Bristle numbers in Drosophila
- Many diseases
  - Cancers
  - Diabetes
  - Schizophrenia

Fictional (simplified) example: 6 genes for women's "height"

Person:	1	2	3	4	5	6	7
Gene 1	**AA**	**aa**	**Aa**	**Aa**	**aa**	**Aa**	**AA**
Gene 2	**Bb**	**Bb**	**BB**	**Bb**	**Bb**	**bb**	**BB**
Gene 3	**CC**	**cc**	**Cc**	**CC**	**Cc**	**Cc**	**CC**
Gene 4	**Dd**	**Dd**	**Dd**	**Dd**	**Dd**	**DD**	**DD**
Gene 5	**Ee**	**ee**	**Ee**	**EE**	**Ee**	**Ee**	**Ee**
Gene 6	**ff**	**ff**	**ff**	**Ff**	**FF**	**Ff**	**Ff**
Height	**5'7"**	**5'2"**	**5'6"**	**5'8"**	**5'6"**	**5'6"**	**5'10"**

Height in inches = 5' 0" + number capital letter alleles
Hence, range 5' 0" – 6' 0"

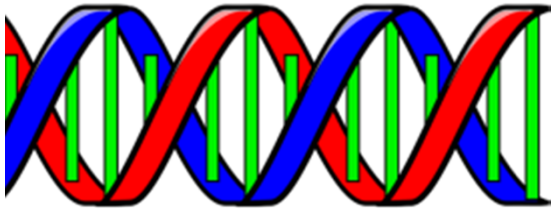
How does one map the genes causing such differences?

- Concepts very similar to mapping simple, single gene traits
- BUT many genes are contributing
- How do we find these genes?
  Two general approaches (again)
  - Mapping difference in crosses/ pedigrees
  - Mapping variation within populations (e.g., GWAS)

Points to emphasize

- We often KNOW the locations of the *markers* being used ahead of time
  - Identified from genome sequences

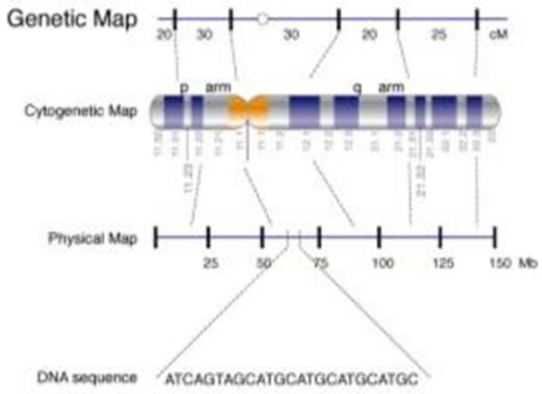
- We DO NOT KNOW the locations of the *genes causing traits/ diseases* ahead of time
  - We're using their linkage to markers to find them!



We map complex traits to QTLs

- **QTL** = "quantitative trait locus"
  - A locus with allelic variation that influences a phenotype

- *Its exact location is not known*, but we infer its approximate location from association with marker genotypes



- Often detect **many** QTLs affecting single traits

Fundamental underlying principle: Association of marker genotype with phenotype

- Markers near ("linked to") gene(s) causing different phenotypes will show an association with that phenotype
- If there are many genes, and/ or if the effect is "complicated", the association may not be very strong

Illustrative example

- Cross 1 tall corn (5’) to 1 short corn (1’)
- F1s are intermediate- 3’ on average
- Cross F1s together to get F2
- F2 range ~1’ to ~5’, bell curve around 3’
  - Assuming the height variation is genetic, how many genes are involved? Could it be 1 gene?

Illustrative example (cont'd)

- … but it IS at least partially genetic, so there are genes with alleles affecting height
- Get genotype for 2 markers (capital letter is from tall)
- Look for association between genotype and average phenotype for that genotype
  - **AA**: 3.5', **Aa**: 3', **aa**: 2.5'
  - **ZZ**: 3', **Zz**: 3', **zz**: 3'
  - Is either marker associated with phenotype?

Illustrative example (cont'd)

- Why wasn't the association of alleles at the A marker gene more complete? For example, why didn't all (or average) AA individuals have height 5'?

Issues with which to contend

- Multiple genes affecting phenotype
- Interactions among genes (“epistasis”)
- Etc.
- … all that complications we talked about last time!
- **Recombination between marker and gene affecting phenotype**



Issues with which to contend

- **Recombination between marker and gene affecting phenotype**
  - If you have NO recombination between marker and trait gene, then marker genotype predicts trait phenotype very well (maybe perfectly)
  - If you have A LITTLE recombination, marker genotype may still be "associated" with trait
  - If LOTS of recombination, no association



How do we pinpoint gene locations amidst this madness?

- Crude localization:
  - Just see association of marker alleles to phenotype
  - Can say "there's a gene 'near there' affecting trait."

- Fine localization:
  - Look at increasing association of neighboring markers…

Fine localization requires examining associations of multiple linked markers with trait

- Identify “stronger” associations in some markers than others
  - What does this mean?

Fine localization requires examining associations of multiple linked markers with trait

- Identify "stronger" associations in some markers than others
  - Bigger difference between average phenotypes associated with genotypes

Genotype	Ave Phenotype	Genotype	Ave Phenotype
AA	**3.5**	**BB**	**3.8**
Aa	**3**	**Bb**	**3**
aa	**2.5**	**bb**	**2.2**

Fine localization requires examining associations of multiple linked markers with trait

- Identify "stronger" associations in some markers than others
  - Bigger difference between average phenotypes associated with genotypes

Genotype	Ave Phenotype	Genotype	Ave Phenotype
AA	**3.5**	**BB**	**3.8**
Aa	**3**	**Bb**	**3**
aa	**2.5**	**bb**	**2.2**

Fine localization requires examining associations of multiple linked markers with trait

- Identify "stronger" associations in some markers than others
  - Bigger difference between average phenotypes associated with genotypes

- **Associations will be stronger when there's less recombination between a marker and a causative gene ("QTL")**

PREDICTION:

- If look at many linked markers, you should be able to pinpoint the location of a QTL by where the association is strongest
  - OR by where you'd "PREDICT" it to be strongest

- Can follow "trajectory of association strength" to infer location of the QTL
  - We'll do this in next video

Image Credits, Unit 7-1

- Chromosomes, © Alexandr Mitiuc, all rights reserved, www.photoxpress.com
- Drosophila bristles, © 2006 Karl Magnacca, CC by-SA 2.5, en.wikipedia.org.
- Brain lasers, © ktsdesign, all rights reserved, www.photoxpress.com