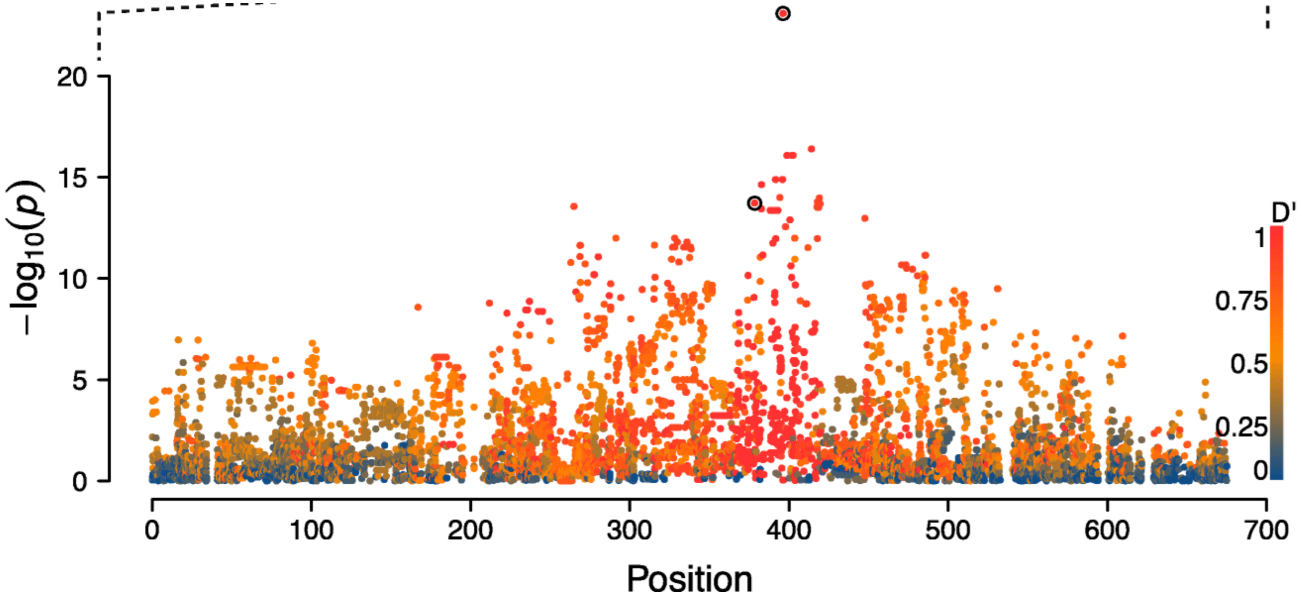
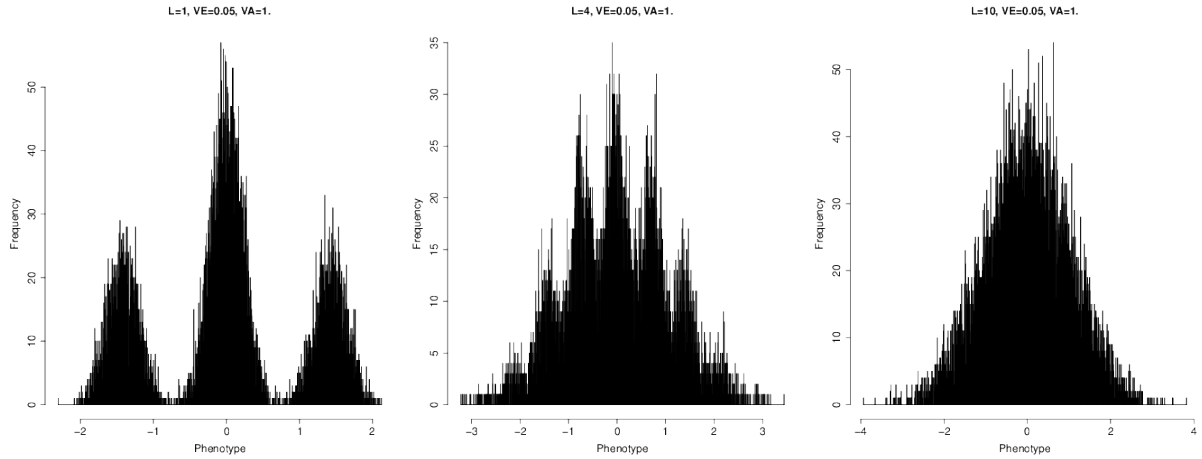
Coop, Chapter 7: Intro.-7.0.2

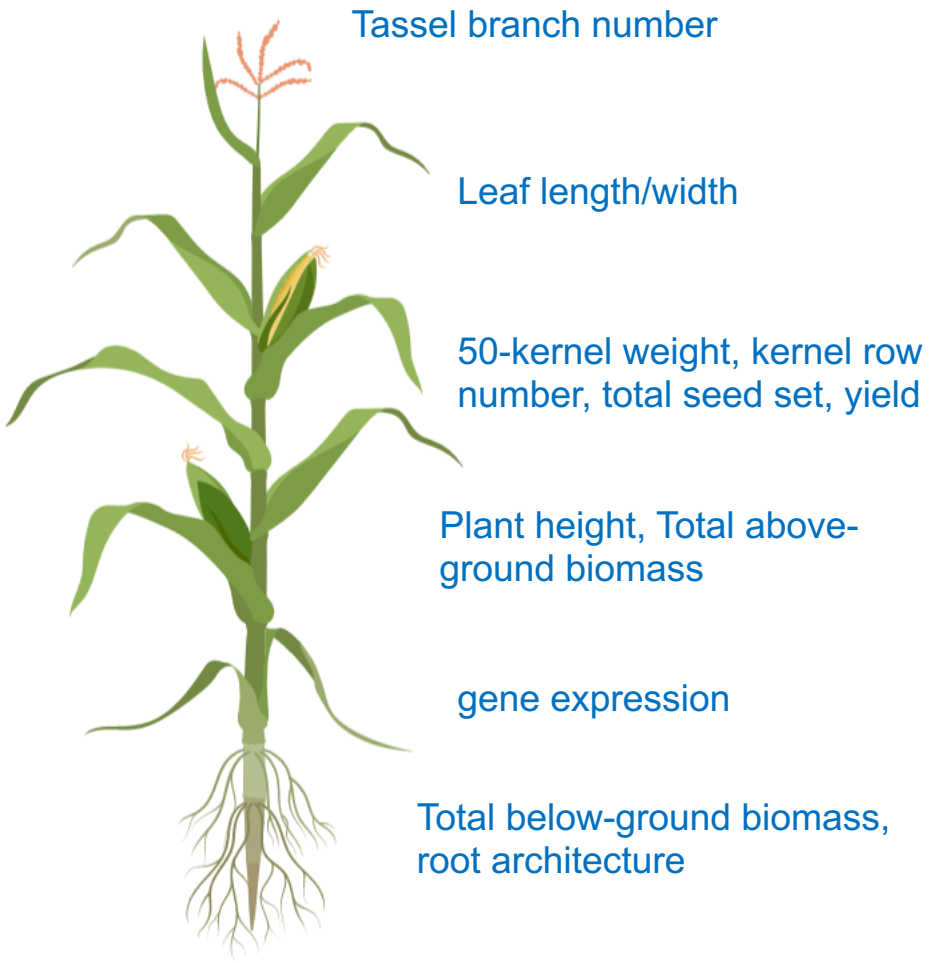
Phenotypic variation and the resemblance between relatives





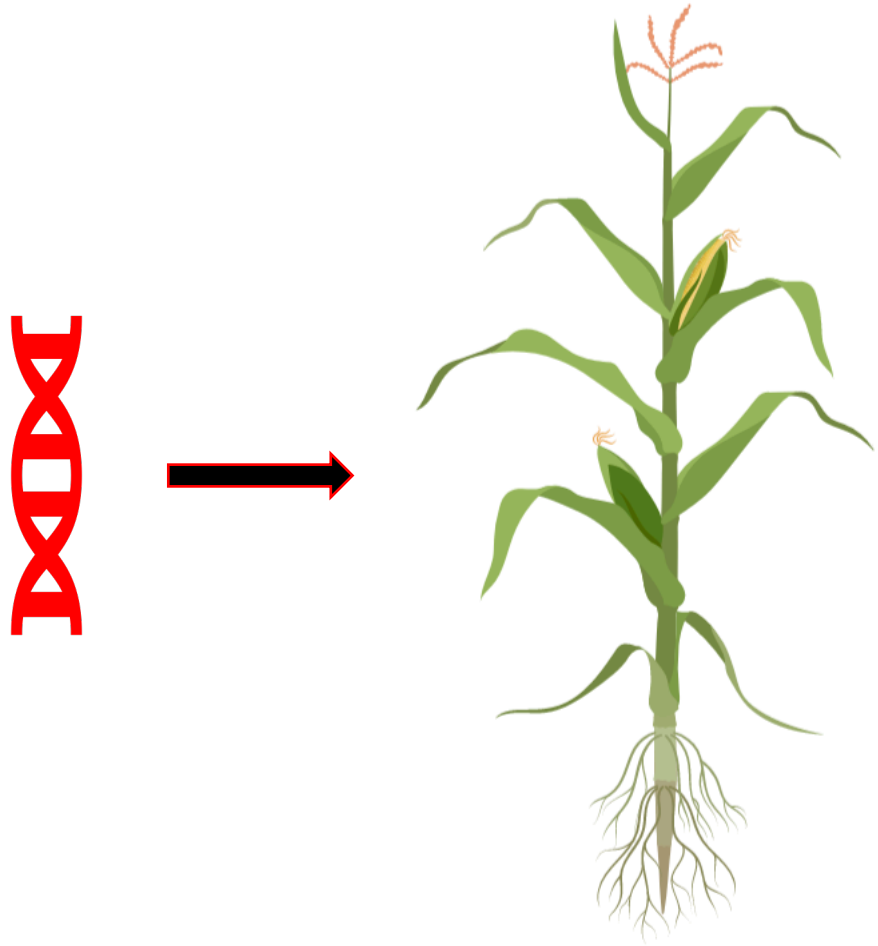
Introduction

- To this point, we've been primarily focused on the genotypes of individuals--the DNA sequence that is decided when gametes fuse to form a zygote
- The phenotype is any measurable aspect of an organism
- Phenotype is the outcome of information encoded by the genome filtering through complex developmental and physiological processes and interaction with environment



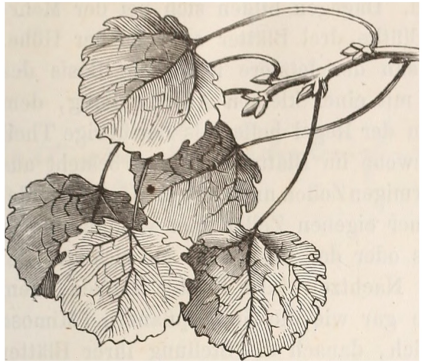
Introduction

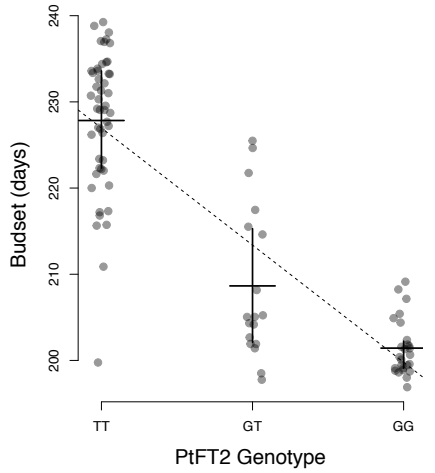
- A central focus in biology is trying to understand the path from genotype to phenotype, often referred to as the genotype-phenotype map
- In this chapter we will focus on understanding how phenotypic variation in a population is the product of variation at the genotypic level
- For example, we can calculate the mean phenotypic value of a species for each genotype at a particular locus



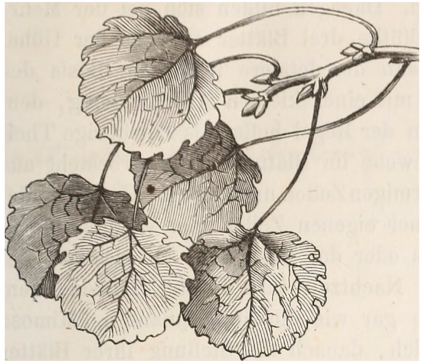
Introduction

- For example, Wang and colleagues (2018), looked at the association of a phenotype known as “budset” in European Aspen with genotypes at various loci across the genome
- Timing of budset is very important for local adaptation since trees at higher latitude must flower earlier to reproduce prior to frost
- The strongest association with budset timing occurs with a SNP in the *PtFT2* gene, a locus associated with flowering time across many plant species



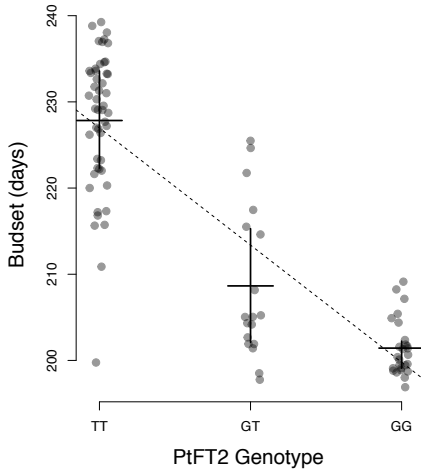


Introduction



- As seen in the lower right plot, we can determine the relationship between genotype and phenotype by fitting a linear regression of phenotypes across the genotypes at a particular SNP:

$$X \sim \mu + a_l G_l \qquad (7.1)$$

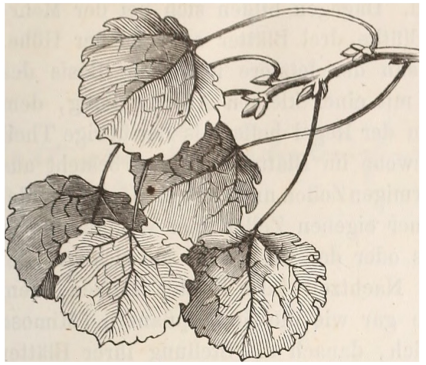


- X is a vector of individuals' phenotypes, G_l is a vector of genotypes at a particular locus, taking the value of 0, 1, or 2 depending on whether an individuals is a homozygote (0 or 2) or heterozygote (1)

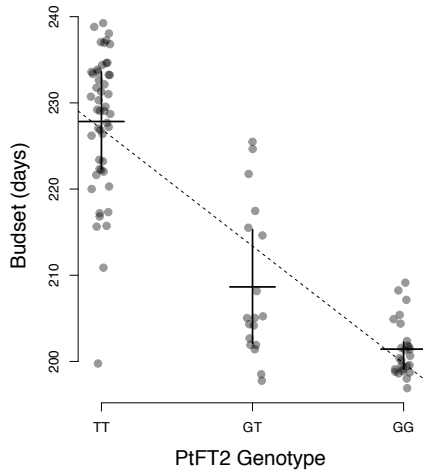
- μ is the phenotypic mean

Introduction

$$X \sim \mu + a_l G_l \qquad (7.1)$$

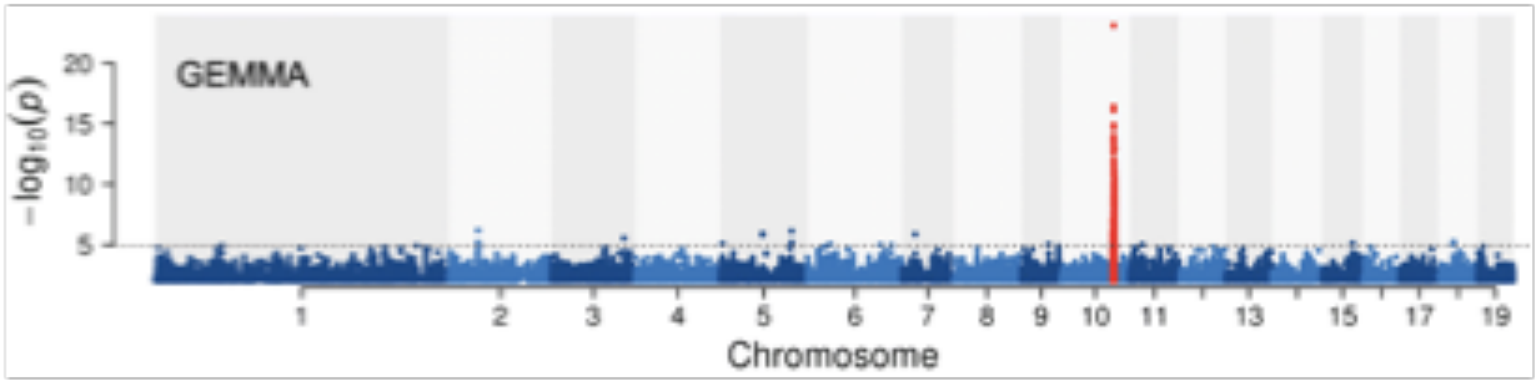


- The slope of the regression line (a_l) can be interpreted as the average effect of substituting a copy of one allele for the alternative allele
- In the Aspen example, the slope is -13.6 days
- Each copy of the G allele at *PtFT2* causes budset to occur 13.6 days earlier



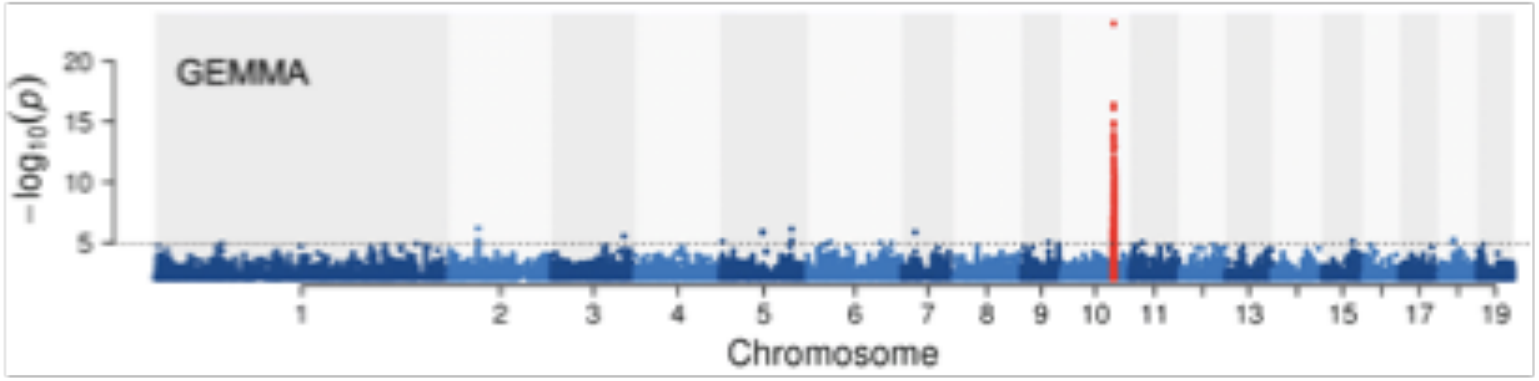
Introduction

- We can calculate a p-value for our regression and do this SNP-by-SNP across the genome, an analysis known as a Genome-Wide Association Study (GWAS)
- The logarithm of p-values is often plotted for each SNP across the entire genome in a so-called “Manhattan plot”
- Below is the Manhattan plot from the Wang *et al.* study, clearly showing the *PtFT2* locus as an outlier (colored in red)



Introduction

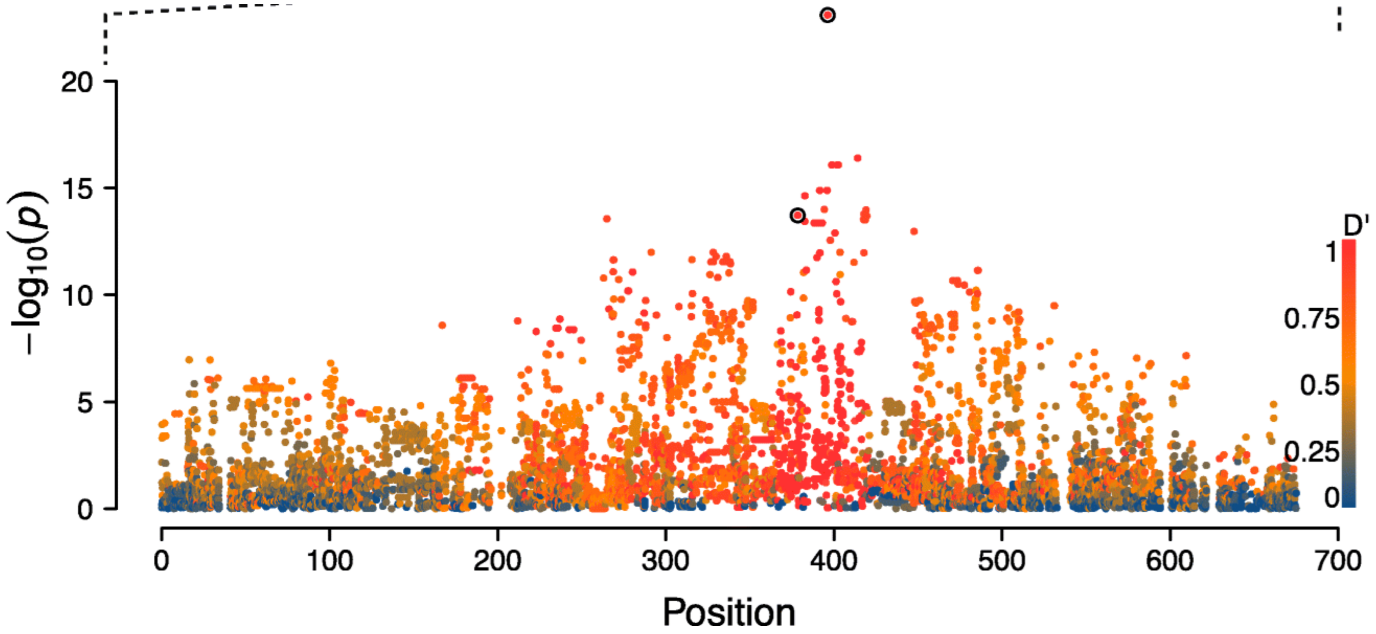
- In the plot below, each SNP is a different dot on the Manhattan plot
- Note how many SNPs nearby the causal SNP at *PtFT2* also have significant p-values (above the threshold indicated by the dotted line)
- These SNPs do not directly affect budset, but they are associated with the causal locus because of linkage disequilibrium (not enough recombination has occurred between these close loci to break up associations)



Introduction

- Below is a zoomed-in plot of p-values of SNPs right around the *PtFT2* locus, with SNPs colored based on their value of *D'* (remember, this is a measure of linkage disequilibrium)

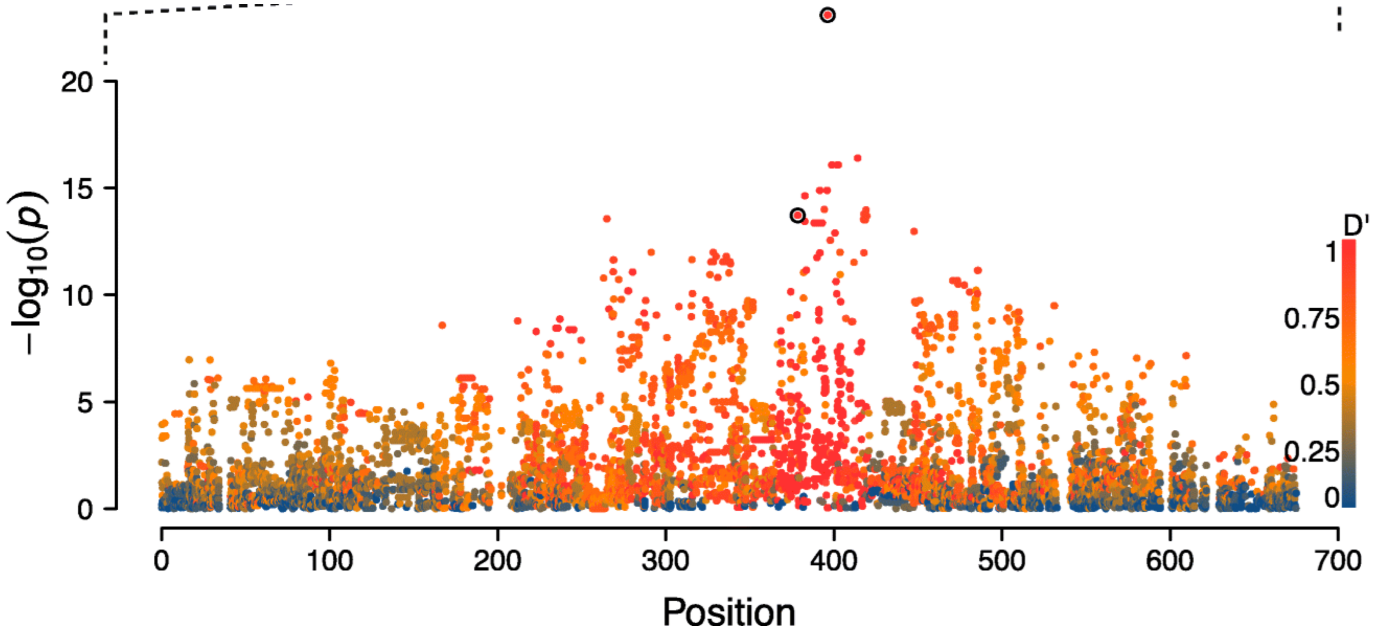
- As SNPs are further and further away from the causal locus, LD and the significance of p-values decay



Introduction

- We must be careful to understand that *PtFT2* is merely a gene that is associated with budset in this particular sample under these particular environmental conditions; *PtFT2* is not generally a gene ***for*** budset

- In other samples and/or environments or through mutant screens, we may find other loci associated with budset

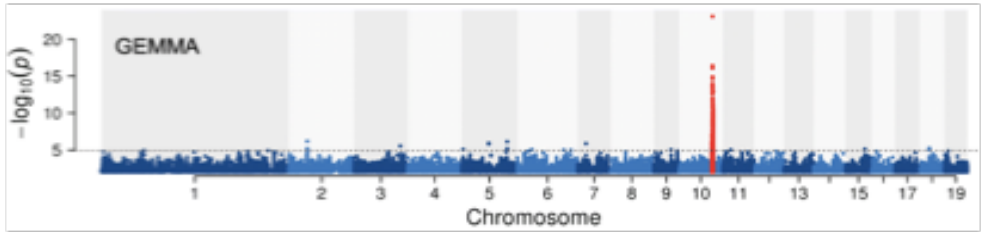


Introduction

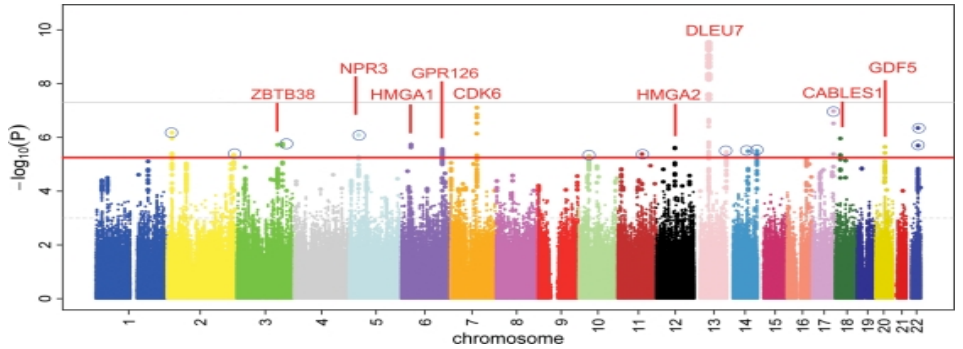
- Also note how the genetic basis of budset in this particular study appears to have a relatively simple genetic basis; (*PtFT2*) is the only clear outlier and has a large effect on the trait

- Other traits, like human height, have a much more complex genetic basis, with small functional effects being detected across hundreds of loci; these are referred to as "polygenic" traits

Budset



Human Height



Estrada *et al.* 2009

7.0.1 A simple additive model of a trait

- To develop a simple model to connect genotype and phenotype, let's imagine that a particular trait is controlled by L loci that act in an additive manner (*i.e.*, we can sum their effects)

- We can sum effects across loci using the following equation:

$$\mathbb{E}(X_i|G_{i,1},\cdots,G_{i,L}) = \mu + X_{A,i} = \mu + \sum_{l=1}^{L} G_{i,l}a_l \qquad (7.2)$$

- Where μ is the mean phenotype in our population and $X_{A,i}$ is deviation from the mean based on genotype, which is determined by the sum of effects (a_l) across loci

7.0.1 A simple additive model of a trait

- To make this simple model more realistic, we can incorporate the effects of environment on a particular trait:

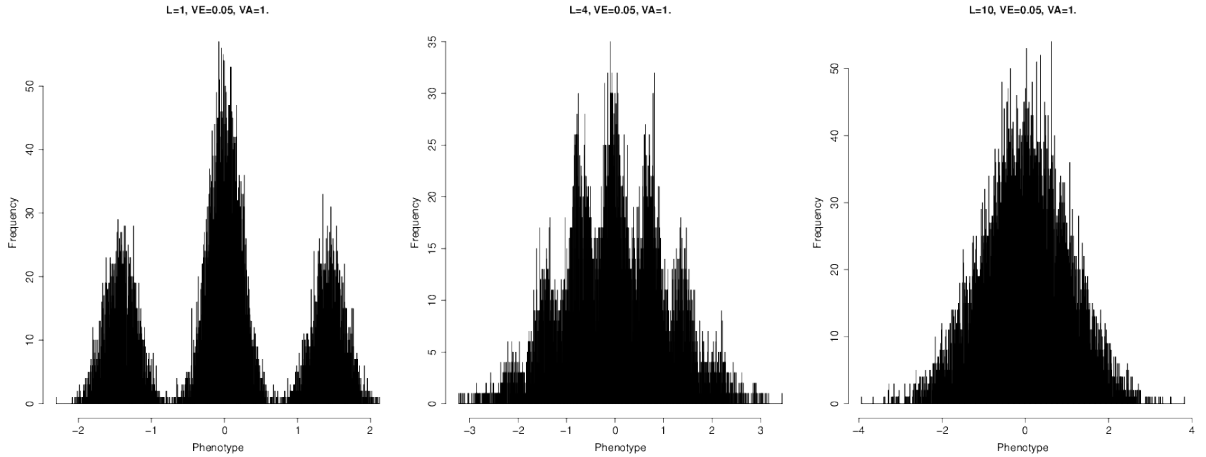
$$X_i = \mu + X_{A,i} + X_{E,i} \qquad (7.3)$$

- Here, $X_{E,i}$ is the deviation of an individual away from the phenotypic mean due to the environment

- This includes more tractable aspects of the environment (*e.g.*, temperature, precipitation) as well as more stochastic noise (pests, disease, predation, storms)

7.0.1 A simple additive model of a trait

- Given the complexity of environment, its effects will usually follow a normal distribution across individuals
- If enough loci contribute to a trait, the genotypic effects will be normal as well
- Combined, the contributions of genotype and environment lead to a normal distribution for the value of a trait across a population
- Bear in mind that we are ignoring the interactions of alleles at a locus (dominant vs. recessive) and the interactions of alleles across the genome (epistasis)



7.0.2 Additive genetic variance and heritability

- Focusing just on additive effects (ignoring non-additive for now), we'll begin to build models including variance

- Additive genetic variance (V_A), is the phenotypic variance due to the additive effects of segregating genetic variation

- Our phenotypic variance across individuals can be written as:

$$V = Var(X_A) + Var(X_E) = V_A + V_E \qquad (7.5)$$

- Which includes an environmental component (but no interaction between genotype and environment)

7.0.2 Additive genetic variance and heritability

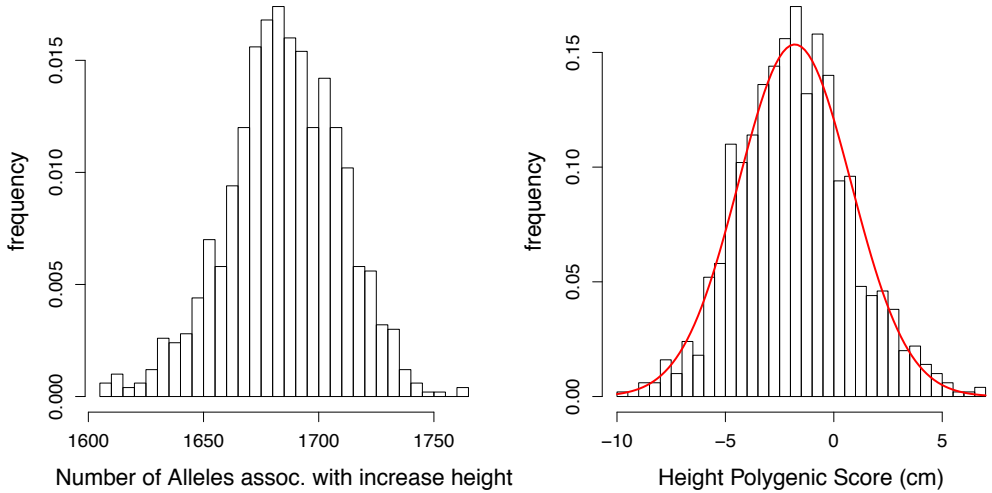
- Additive genetic variance (V_A) can be written as:

$$V_A = \sum_{l=1}^{L} Var(G_{i,l} a_l) \qquad (7.6)$$

- Where $Var(G_{i,lal})$ is the contribution of locus l to the additive variance across individuals

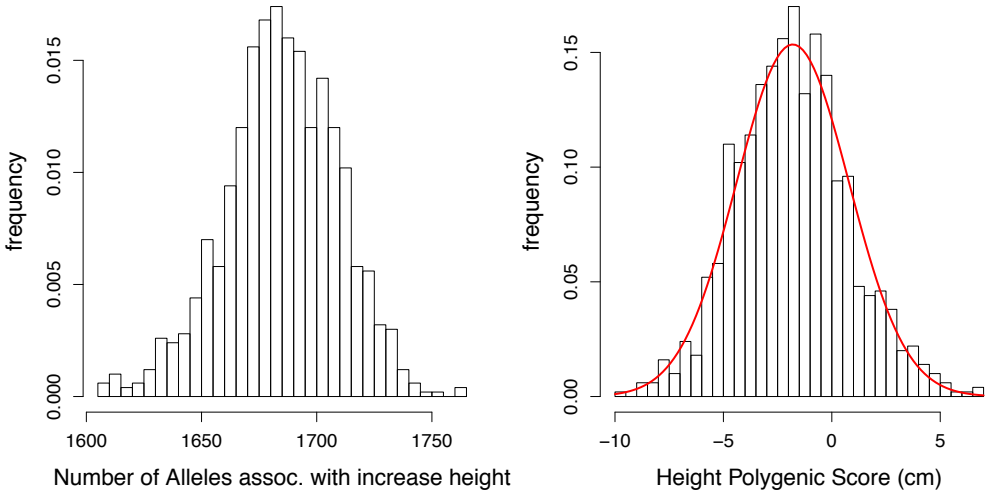
7.0.2 Additive genetic variance and heritability

- Given the findings of various GWAS that have identified candidates underlying traits, we can start to represent how loci contribute to a polygenic trait
- Estimated effect size of each locus can be used in a weighted sum based on an individual's genotype
- The weighted sum is the individual's polygenic score



7.0.2 Additive genetic variance and heritability

- Take for example 1700 SNPs that are associated with variation in human height
- The effect of each SNP is very small (median = 0.07cm)
- Plotted to the right is the distribution of the number of height-increasing alleles in 1000 humans and their polygenic scores
- These roughly follow a normal distribution due to the number of loci
- Those with higher polygenic scores are taller



7.0.2 Additive genetic variance and heritability

- For considering the evolution of phenotypes, it will be helpful to think about what proportion of phenotypic variance is due to genetic differences alone (*i.e.*, not due to the environment)

- If a trait had no genetic basis, no matter how much selection changed the mean phenotype within a generation, the trait would not change over multiple generations

- The proportion of phenotypic variance that is genetic is known as heritability (h^2):

$$h^2 = \frac{Var(X_A)}{V} = \frac{V_A}{V} \qquad (7.8)$$

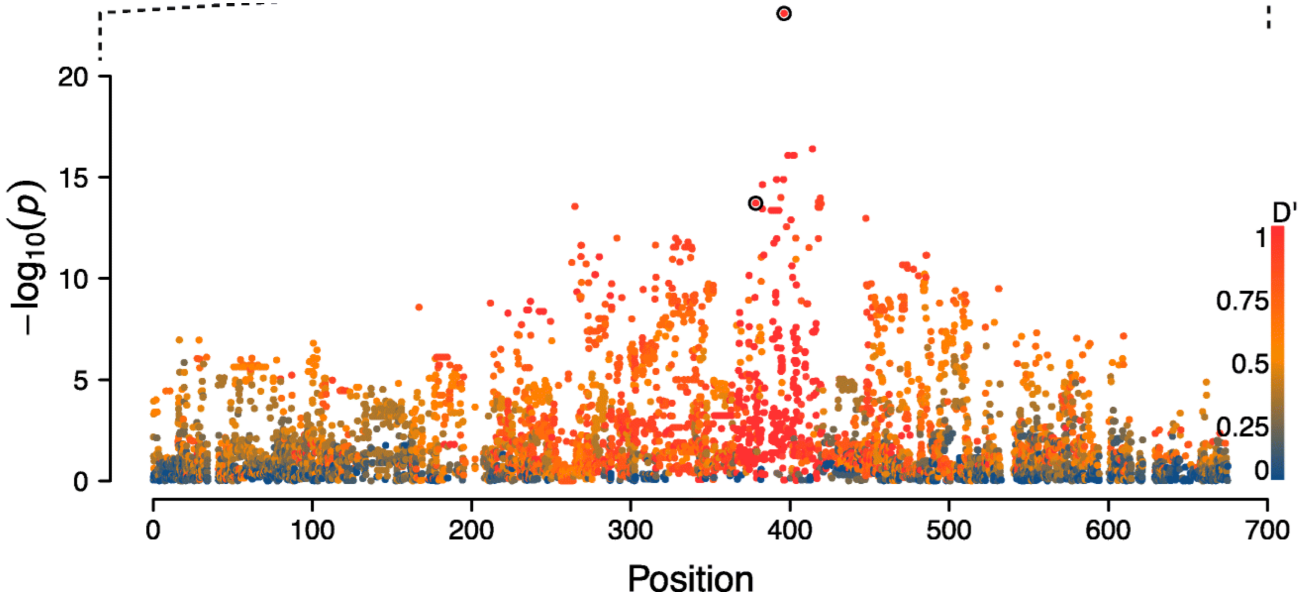
7.0.2 Additive genetic variance and heritability

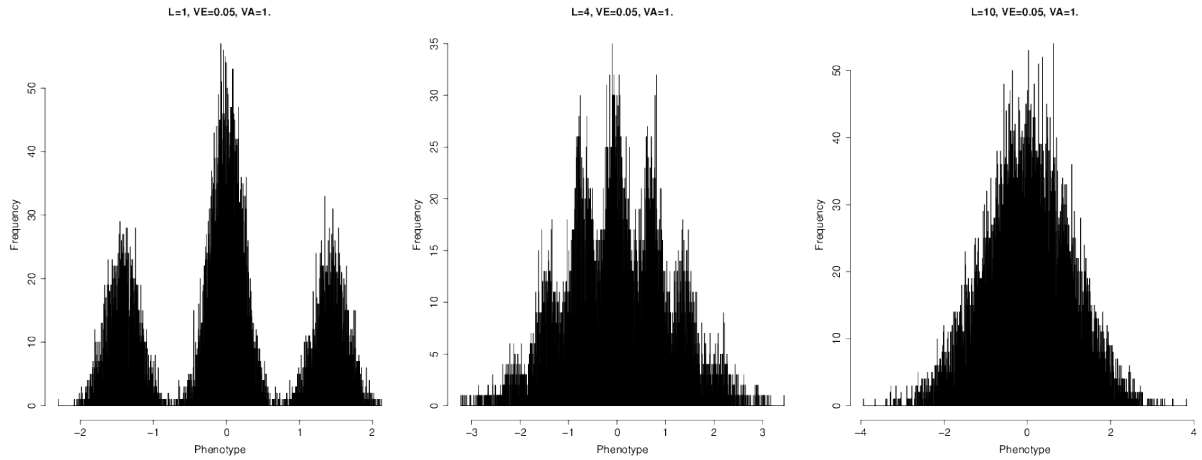
$$h^2 = \frac{Var(X_A)}{V} = \frac{V_A}{V} \tag{7.8}$$

- Since we are assuming this is all additive variance, we call this the *narrow sense heritability*
- If we were also considering dominance and epistasis (the total proportion of phenotypic variance attributable to genetic factors), this would be the *broad sense heritability*
- Since phenotypic variance will change based on sample and environment, we cannot generalize heritability across other samples or environments

Coop, Chapter 7: 7.0.3

Phenotypic variation and the resemblance between relatives



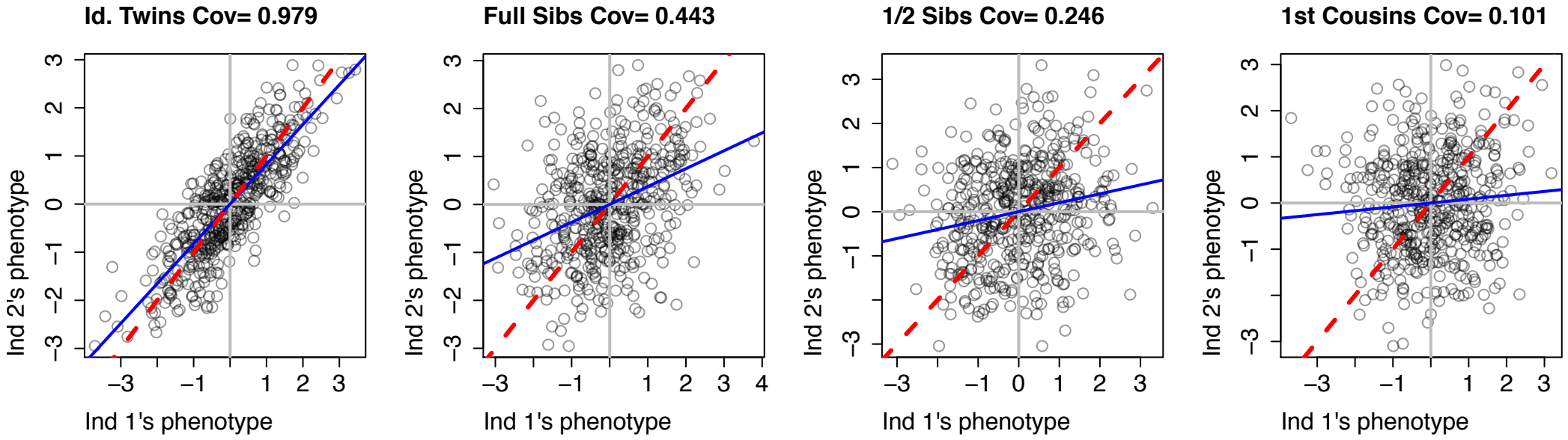


7.0.3 The covariance between relatives

- Resemblance between relatives (*e.g.,* twins) has long been of interest to quantitative geneticists
- In fact, similarity in phenotype can be used to estimate heritability and the covariance of traits
- For example, do tall women tend to have tall sisters?
- If there is a genetic component to traits, we would expect these patterns of decreasing covariance: identical twins > full sibs > half sibs > first cousins

7.0.3 The covariance between relatives

- Simulation of covariance between individuals across a range of genetic relatedness



7.0.3 The covariance between relatives

- Under a simple additive model of phenotypes, we can write covariance as:

$$Cov(X_1, X_2) = Cov\left((X_{1M} + X_{1P} + X_{1E}), ((X_{2M} + X_{2P} + X_{2E})\right) \tag{7.9}$$

- Where we are looking at the covariance in phenotypes of two individuals (X_1, X_2) and breaking this into components of maternal (M) alleles, paternal (P) alleles, and the environment (E)

- When expanding this, we can ignore terms of covariance between environments of individuals $(Cov(X_{1E}, X_{2E}) = 0)$ and covariance between the environment of one individual and the genetic variation in the other individual $(Cov(X_{1E}, (X_{2M} + X_{2P}))$

7.0.3 The covariance between relatives

- After ignoring these covariances that involve environment, we are left with:

$$Cov(X_1, X_2) = Cov((X_{1M}, X_{2M}) + Cov(X_{1M}, X_{2P}) + Cov(X_{1P}, X_{2M}) + Cov(X_{1P}, X_{2P}) \quad (7.10)$$

- Now our assessment of covariance in phenotype between individuals is just based on covariance in their maternal and paternal allelic effects
- Based on what we know about relatedness, we can simplify this equation
- We'll start with a few specific cases before developing this more generally

7.0.3 The covariance between relatives

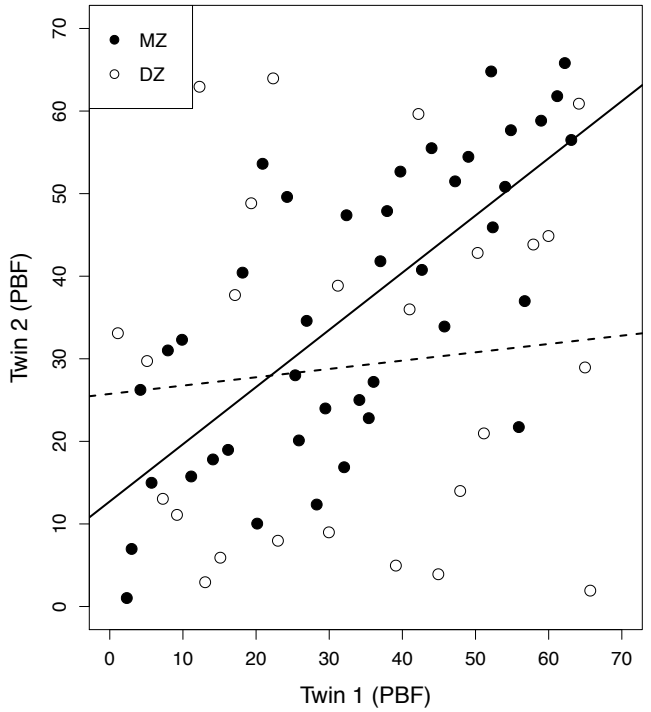
Identical Twins:

- Twins share their maternal and paternal alleles, identical by descent ($X_{1M} = X_{2M}$ and $X_{1P} = X_{2P}$)

- If we assume that the twins' mother and father are unrelated, the covariance between their alleles is 0: ($Cov(X_{1P}, X_{2M}) = Cov(X_{1M}, X_{2P}) = 0$)

- By utilizing these assumptions, equation 7.10 simplifies to:

$$Cov(X_1, X_2) = Cov((X_{1M}, X_{2M}) + Cov(X_{1P}, X_{2P}) = 2Var(X_{1M}) = V_A \quad (7.11)$$



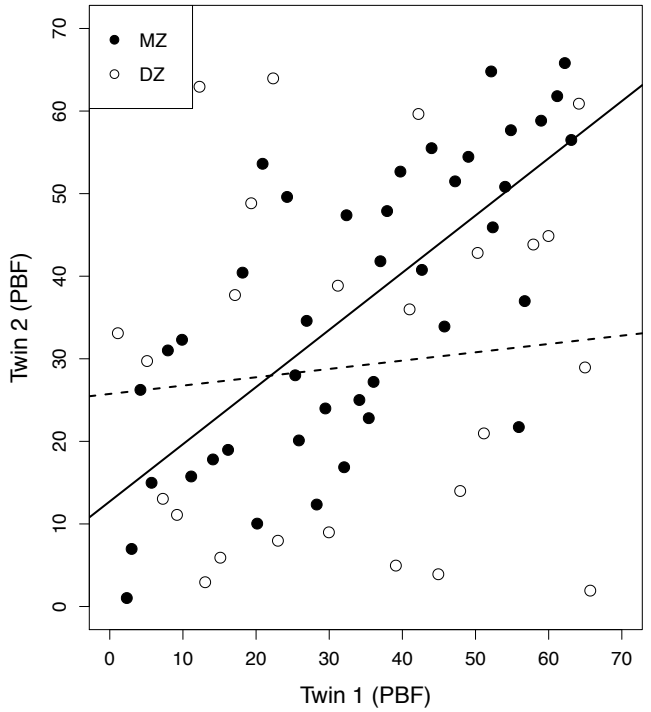
7.0.3 The covariance between relatives

Identical Twins:

- This tells us that that covariance in traits between monozygotic twins is equal to the additive genetic variance (V_A)
- This allows for a very simple calculation of the narrow sense heritability in which we divide our covariance in traits from monozygotic twins by the trait variance:

$$h^2 = \frac{Cov(\mathrm{MZ}_1, \mathrm{MZ}_2)}{V} = \rho_{\mathrm{MZ}} \qquad (7.12)$$

- Where ρ_{MZ} is just the correlation in traits across monozygotic twins
- This excludes the shared environment of twins and non-additive effects



7.0.3 The covariance between relatives

Parent and Child:

- Children resemble their parents because they inherit their genome from their parents
- Assuming a mother and father are unrelated, children have 1 allele that is IBD from their mother and one allele that is IBD with their father (r_1 = 1 and $r_0 = r_2 = 0$)
- For example, let's assume that a mother (ind 1) transmits her own paternal allele to her child (ind 2) so $X_{P1} = X_{M2}$
- In this case, $Cov(X_{P1}, X_{M2}) = Var(X_{P1}) = \frac{1}{2}V_A$
- This means that the covariance in traits between parent and child ($Cov(X_1, X_2)$) is equal to $\frac{1}{2}V_A$

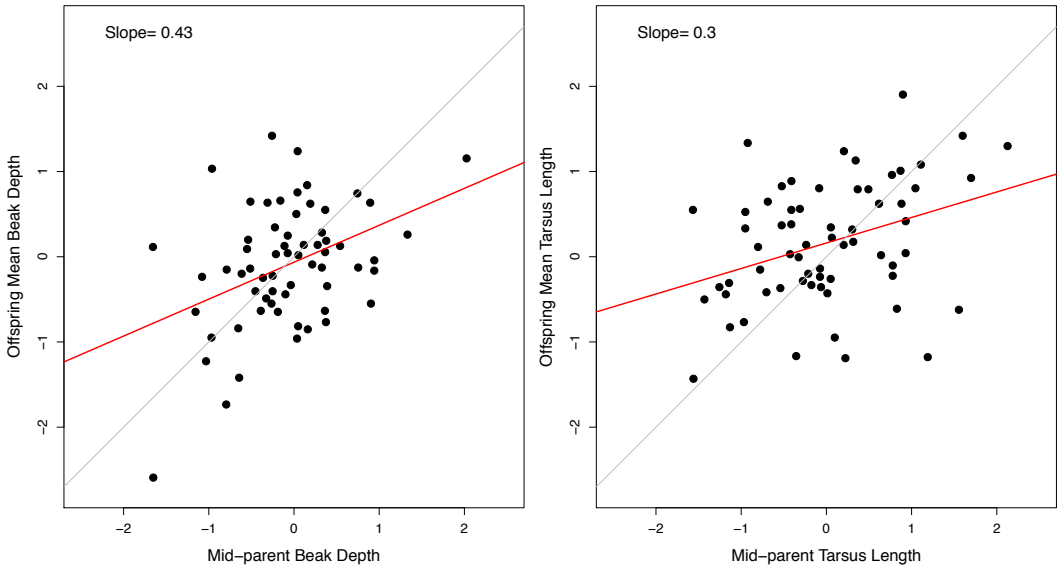
7.0.3 The covariance between relatives

Parent and Child:

- We can use this result to calculate the narrow sense heritability through the regression of the child's phenotype on the parental mid-point phenotype, which is the average of mother and father's phenotypes

- The slope of this regression provides an estimate of h^2:

$$\beta_{\text{mid,kid}} = \frac{Cov(X_{\text{kid}}, X_{\text{mid}})}{Var(X_{\text{mid}})} = \frac{V_A}{V} = h^2 \qquad (7.13)$$



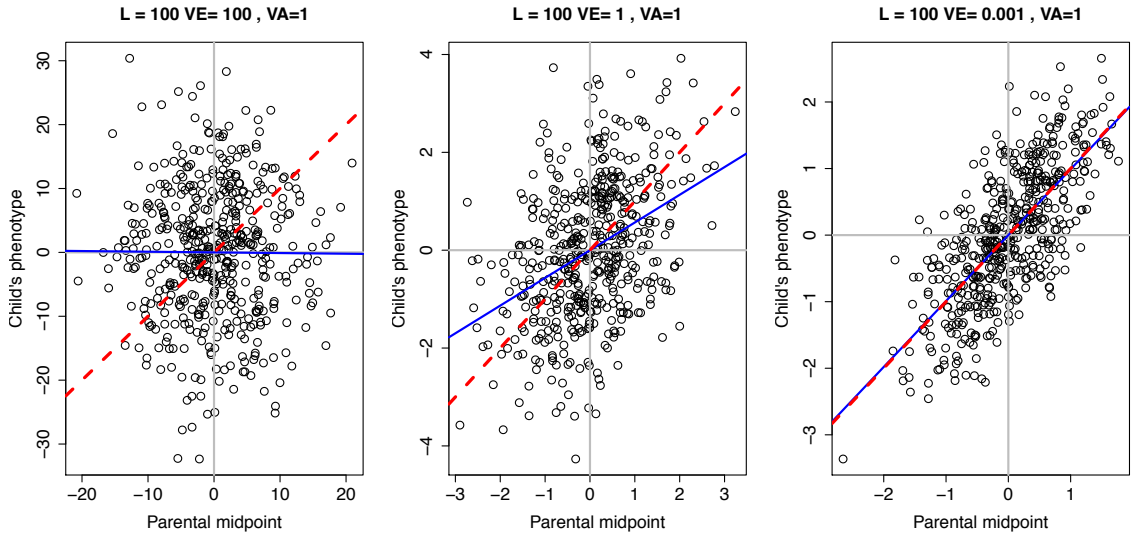
7.0.3 The covariance between relatives

Parent and Child:

- If much of the phenotypic variation is due to additive genetic variation ($h^2 \approx 1$), then children will resemble their parents

- If much of the phenotypic variation is environmental ($h^2 \approx 0$) and children and parents do not share an environment, then children will not resemble their parents

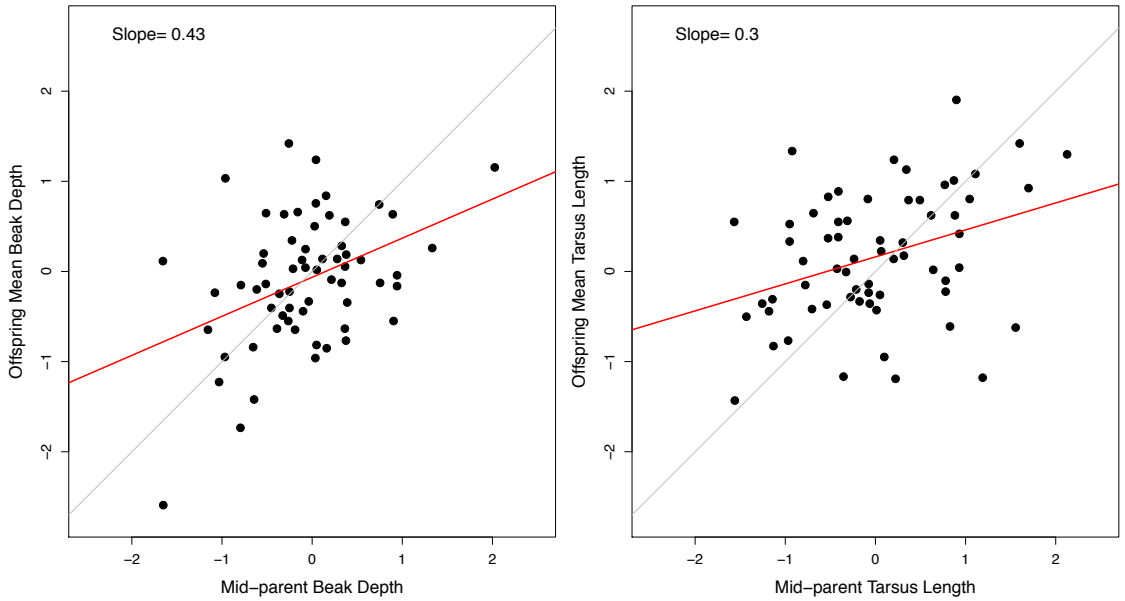
Decreasing Environmental Variance



7.0.3 The covariance between relatives

Parent and Child:

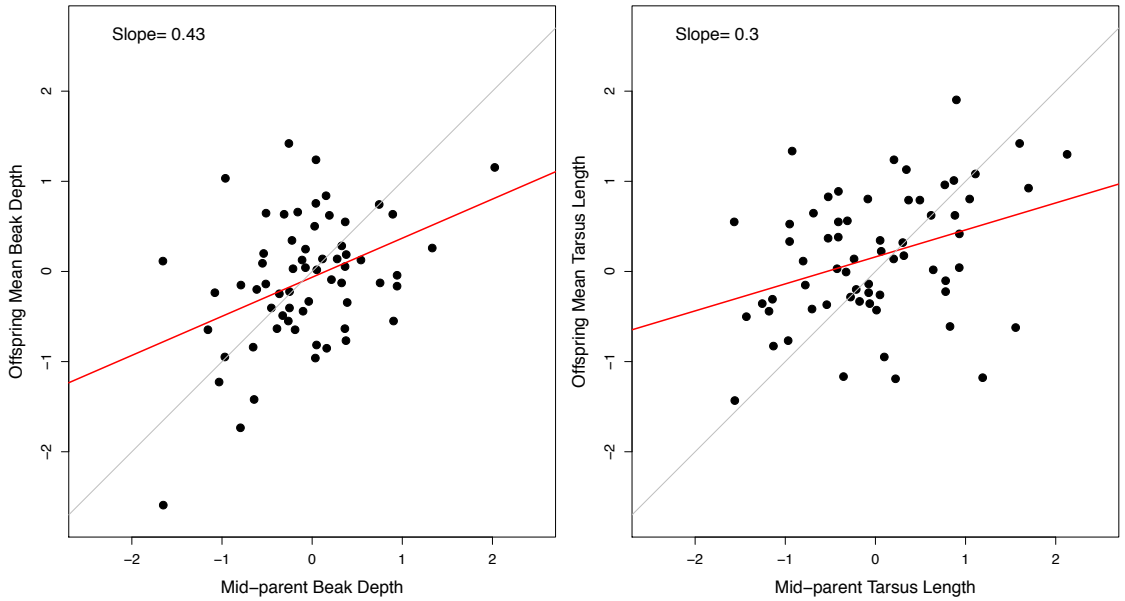
- Going back to the song sparrow data, we can see that $h^2 = 0.43$ for beak depth and 0.3 for tarsus length based on the slope of the regression
- This tells us, for example, that 30% of the difference in tarsus length is attributable to additive genetic differences across individuals
- Smith and Zach (1979) who published these results, also estimated h^2 based on the slope using each parent and found this to be about half of h^2 based on the midpoint between parents



7.0.3 The covariance between relatives

Parent and Child:

- This regression approach does not take into account the shared environment between parents and offspring, and therefore likely inflates our estimate of heritability
- For plants, we can grow our samples across different environments to disentangle genetics and environment and better understand heritability
- In the song sparrow experiment, they tried cross-fostering offspring (moving eggs to new nests) and found that heritability was still high, so the effects of environment were minimal



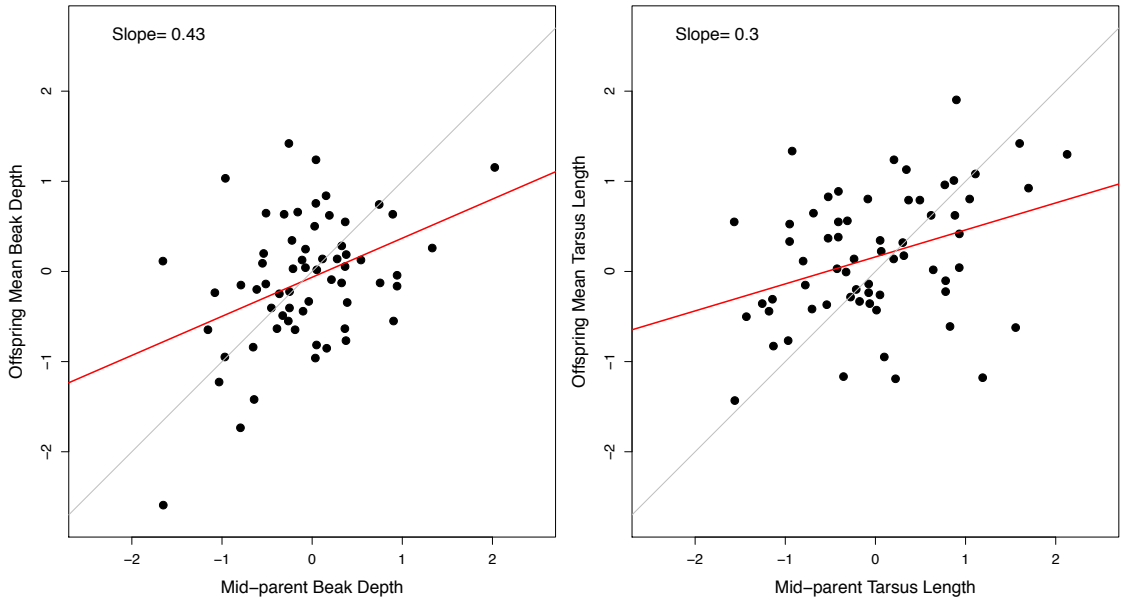
7.0.3 The covariance between relatives

Parent and Child:

- While the simple approximation of heritability based on parent/child regression is somewhat flawed, it provides useful intuition on whether we can predict offspring phenotype based on parents

- If the slope is close to 0, we have little hope of predicting offspring phenotypes based on the parents; whereas a slope close to 1 indicates we have good predictive power

- Natural selection will be most efficient when children closely resemble their parents

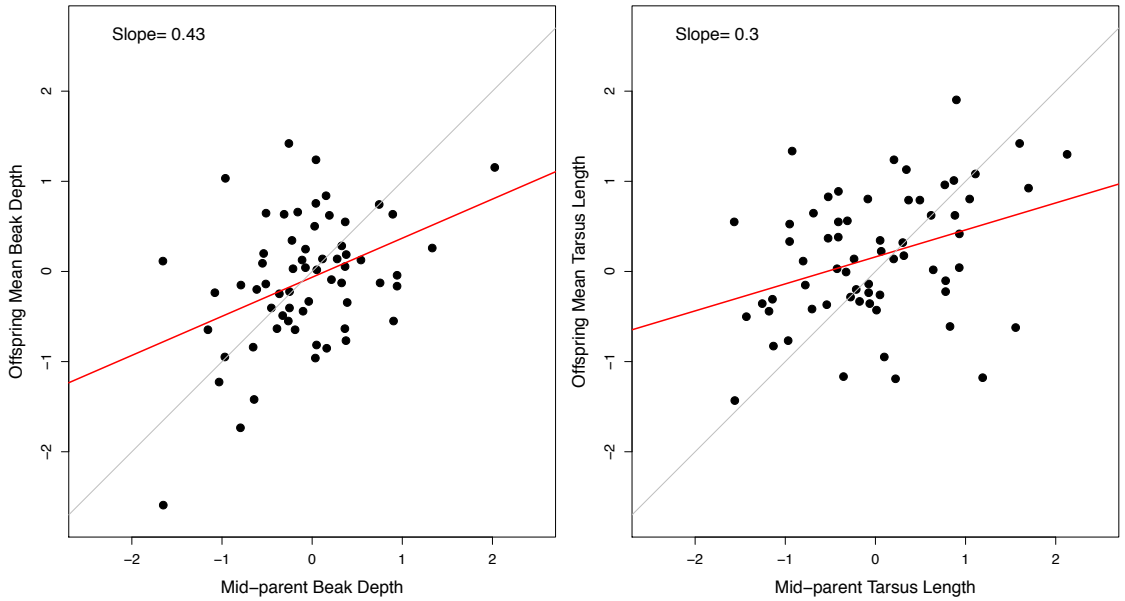


7.0.3 The covariance between relatives

Parent and Child:

- Finally, with this method, if we want to predict a child's phenotype based on those of the parent, we take the mean phenotype of the population and add the difference between our parental mean phenotype and the population mean multiplied by the narrow sense heritability:

$$\mathbb{E}(X_{kid}|X_{mum}, X_{dad}) = \mu + \beta_{mid,kid}(X_{mid} - \mu) = \mu + h^2(X_{mid} - \mu) \tag{7.14}$$



7.0.3 The covariance between relatives

General pairs of relatives:

- In considering twins and parent-child relationships, it is clear that an understanding of covariance in phenotypes across relatives requires knowledge of the number of alleles they share as IBD
- If relatives 1 & 2 share 0 alleles IBD, then: $Cov\big((X_{1M} + X_{1P}), (X_{2M} + X_{2P})\big) = 0$
- 1 allele IBD, then: $Cov\big((X_{1M} + X_{1P}), (X_{2M} + X_{2P})\big) = Var(X_{1M}) = \frac{1}{2}V_A$
- 2 alleles IBD, then: $Cov\big((X_{1M} + X_{1P}), (X_{2M} + X_{2P})\big) = V_A$
- The general equation, therefore, for any pair of relatives becomes:

$$Cov(X_1, X_2) = r_0 \times 0 + r_1 \frac{1}{2} V_A + r_2 V_A = 2F_{1,2} V_A \qquad (7.15)$$

7.0.3 The covariance between relatives

General pairs of relatives:

- In summary, under a simple additive model of the genetic basis of a phenotype, to measure the narrow sense heritability we need to:

1. Measure the covariance between pairs of relatives (assuming that we can remove the effect of shared environmental noise)

2. Use the covariance between relatives to calculate V_A

3. Divide this by the total phenotypic variance to get h^2

7.0.3 The covariance between relatives

Question 3. **A)** In polygynous red-winged blackbird populations (i.e. males mate with several females), paternal half-sibs can be identified. Suppose that the covariance of tarsus lengths among half-sibs is 0.25 cm^2 and that the total phenotypic variance is 4 cm^2. Use these data to estimate h^2 for tarsus length in this population.

$$Cov(X_1, X_2) = r_0 \times 0 + r_1 \frac{1}{2} V_A + r_2 V_A = 2F_{1,2} V_A \qquad (7.15)$$

$Cov(X_1, X_2) = 0.25\text{cm}^2$ $\quad V_P = 4\text{cm}^2$ $\quad h^2 = V_A/V_P$

For half sibs, $r_1 = \frac{1}{2}$ and $r_0 = r_2 = 0$, therefore, $Cov(X_1, X_2) = \frac{V_A}{4}$

$$0.25\text{cm}^2 = \frac{V_A}{4}$$

$$1\text{cm}^2 = V_A$$

$$h^2 = V_A/V_P$$

$$h^2 = 1\text{cm}^2/4\text{cm}^2$$

$$h^2 = ¼$$

7.0.3 The covariance between relatives

The Animal Model:

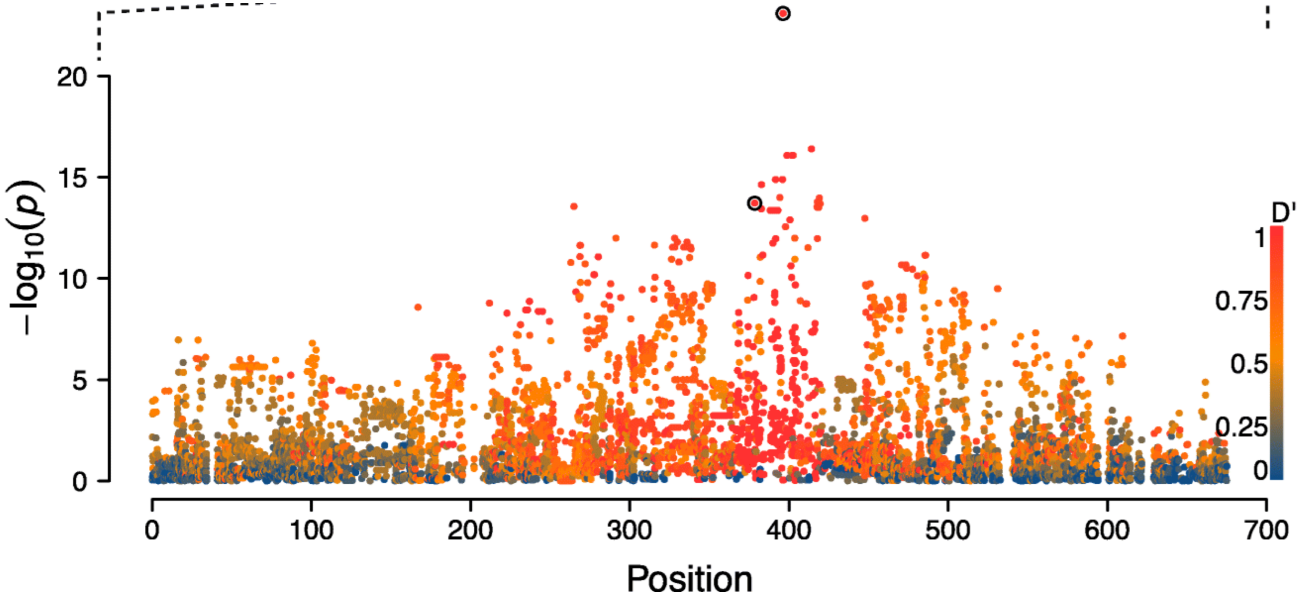
- Our models up until now have assumed a particular relationship between a set of individuals, but often we will be working with a population with range of relationships across individuals (*e.g.,* think of breeding programs with complex pedigrees)

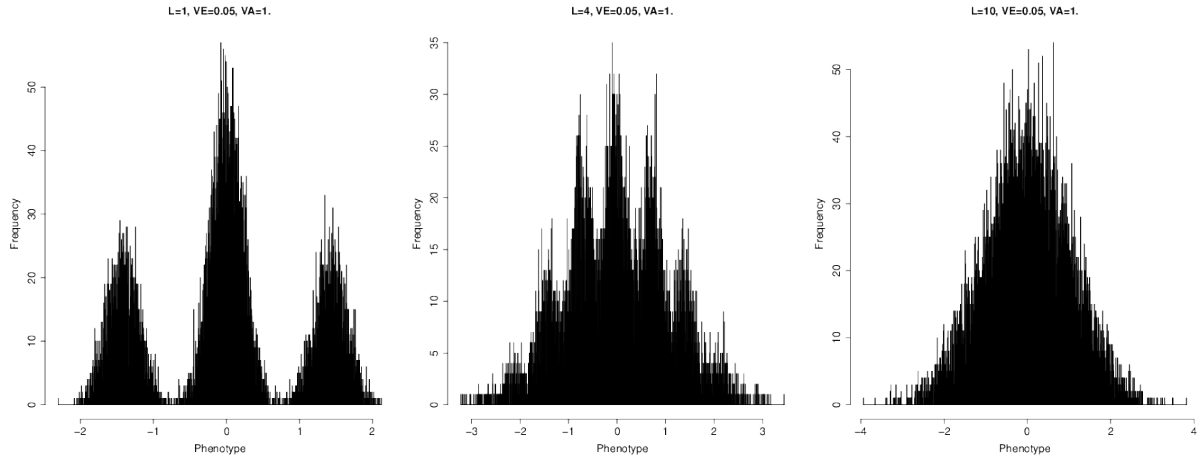
- More complex “mixed” models have been developed that can accommodate distributions of additive genetic and environmental variance.

- These approaches are widely used in modern quantitative genetics to estimate genetic variances and heritabilities

Coop, Chapter 7: 7.1-7.1.1

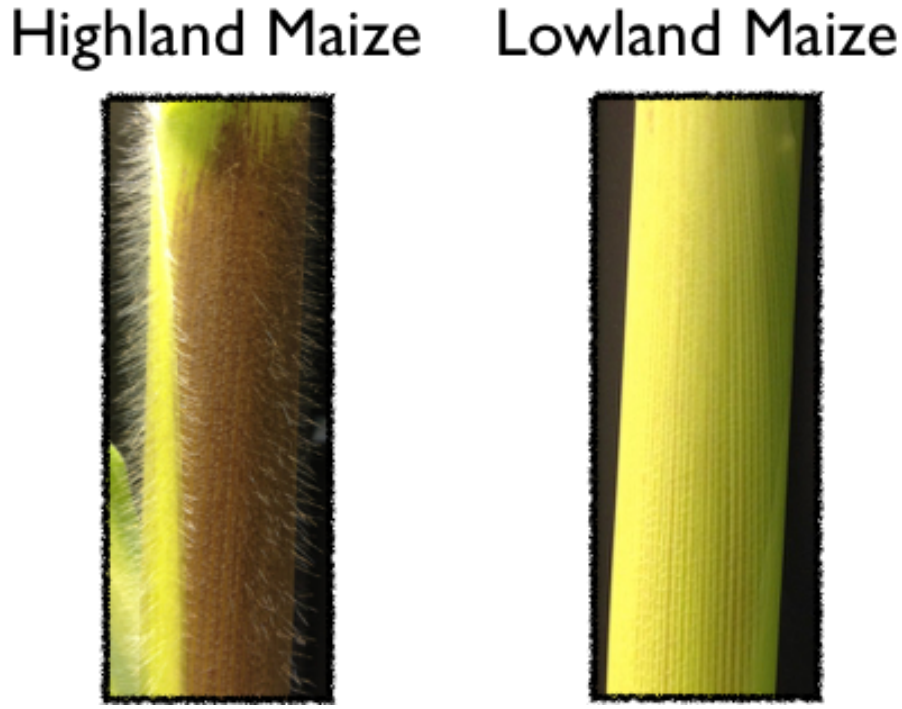
Phenotypic variation and the resemblance between relatives





7.1 Multiple Traits

- Traits often covary, due to environment (*e.g.*, high elevation in maize) or due to underlying genetic covariance between traits
- Genetic covariance can be due to pleiotropy (one gene affecting multiple traits), for example, genes that affect skin pigmentation can also affect hair color
- Genes that are linked may also result in covariation of the traits they encode
- Assortative mating may drive covariation as well



7.1 Multiple Traits

- Let's consider two traits (*e.g.,* leg and nose length, but we'll call them 1 & 2) in an individual i: $X_{1,i}$ and $X_{2,i}$
- We can write these as:

$$\begin{aligned} X_{1,i} &= \mu_1 + X_{1,A,i} + X_{1,E,i} \\ X_{2,i} &= \mu_2 + X_{2,A,i} + X_{2,E,i} \end{aligned} \tag{7.17}$$

- Where we have our trait means (μ), genetic effects (X_A), and environmental effects (X_E)
- We can, once again, describe variance for phenotype (V_1, V_2), environment ($V_{1,E}$, $V_{2,E}$), and the additive genetic variance ($V_{1,A}$, $V_{2,A}$) for both traits

7.1 Multiple Traits

- With two traits, however, we also need to think about covariance in phenotype ($V_{1,2} = Cov(X_1, X_2)$), environmentally induced covariance ($V_{E,1,2} = Cov(X_{1,E}, X_{2,E})$), and additive genetic covariance ($V_{A,1,2} = Cov(X_{1,A}, X_{2,A})$)

- For covariance due to pleiotropy, for example, let's consider two traits that are affected by L SNPs; if the additive effect of an allele at the i^{th} SNP is $\alpha_{i,1}$ and $\alpha_{i,2}$ on our two traits, then additive covariance between traits is:

$$V_{A,1,2} = \sum_{i=1}^{L} 2\alpha_{i,1}\alpha_{i,2}p_i(1-p_i) \qquad (7.18)$$

- This describes genetic covariance between traits, because loci that affect trait 1 also affect trait 2

7.1 Multiple Traits

- While we won't develop this notation further at the moment, variances and covariances of traits can be written in matrix notation across any number of traits:

$$\mathbf{V} = \begin{pmatrix} V_1 & V_{1,2} \\ V_{1,2} & V_2 \end{pmatrix} \tag{7.19}$$

and

$$\mathbf{G} = \begin{pmatrix} V_{1,A} & V_{A,1,2} \\ V_{A,1,2} & V_{2,A} \end{pmatrix} \tag{7.20}$$

- Just as we did with single traits, we can estimate covariance between traits in different relatives:

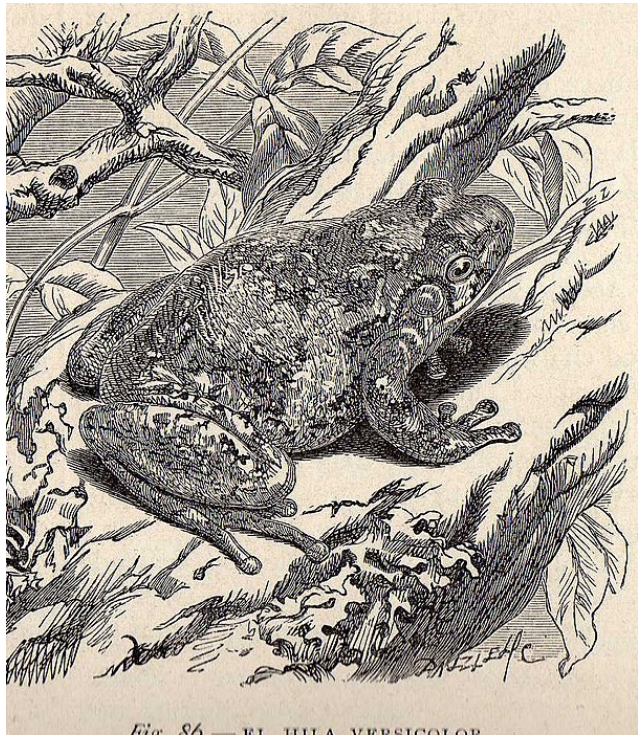
$$Cov(X_{1,i}, X_{2,j}) = 2F_{i,j}V_{A,1,2} \tag{7.21}$$

7.1 Multiple Traits

- As an example of phenotypic and genetic covariance, green treefrog males make calls to attract mates that vary based on trill length and frequency

- Females prefer both many trills and calls, but this is energetically costly

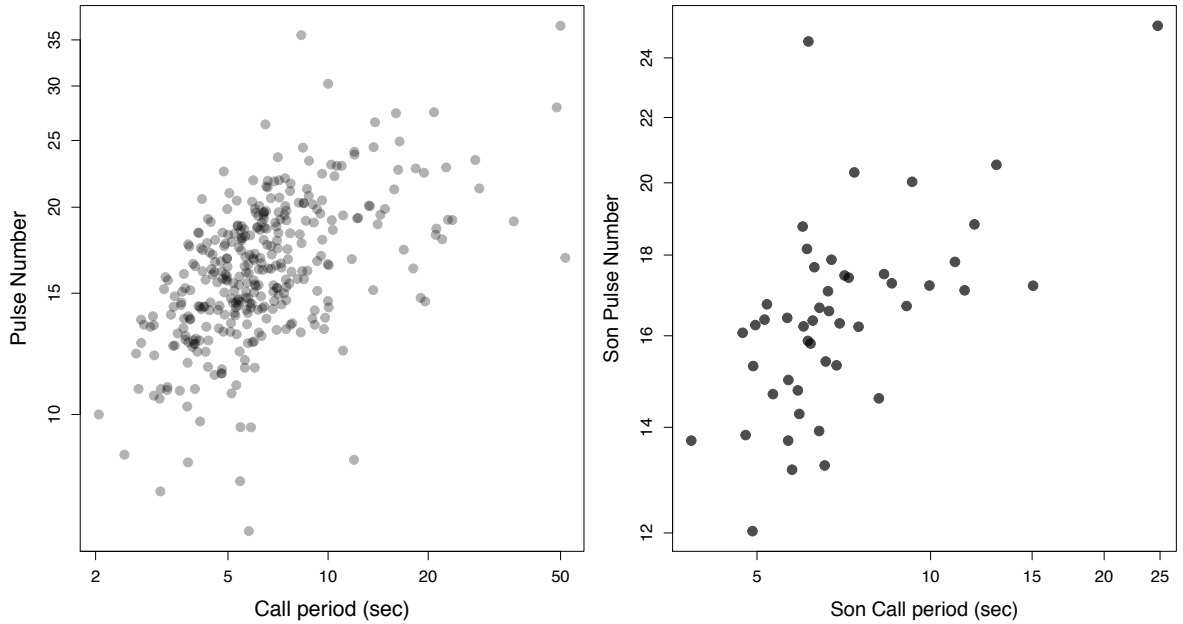
- There may be a trade-off between these traits, so Welch and colleagues (2014) explored whether covariance between traits could be detected



7.1 Multiple Traits

- Males appeared to either take the strategy of 1) more trills, or 2) more calls, and a positive correlation was observed between pulse number and call period

- This phenotypic covariance is associated with underlying genetic covariance between the traits, because the same trait correlation was observed in offspring



7.1 Multiple Traits

- We can further assess additive genetic covariance by calculating the statistic known as the “additive genetic correlation” between phenotypes:

$$r_g = \frac{V_{A,1,2}}{\sqrt{V_{A,1}V_{A,2}}} \qquad (7.22)$$

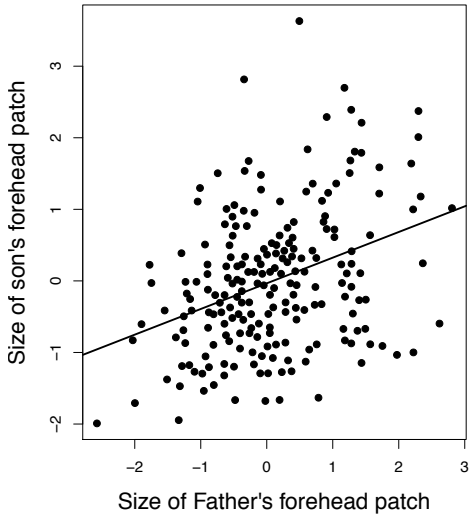
- As a reminder $V_{A,1}$ and $V_{A,2}$ are the additive genetic variance for each trait individually and this statistic will determine the extent to which they are correlated

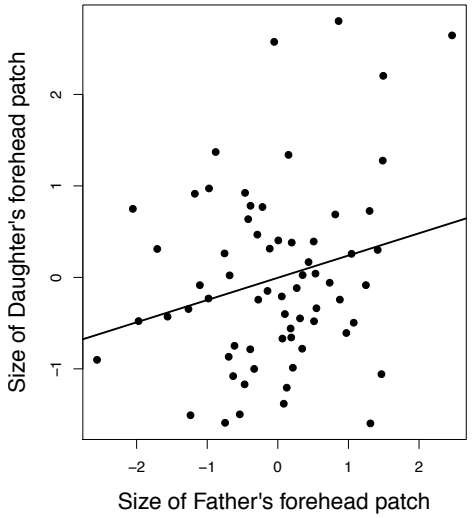
7.1 Multiple Traits

- One important application of genetic covariance in evolutionary genetics is where this breaks down, for example, with sexually antagonistic selection and the evolution of sexual dimorphism

- Potti and Canal (2011) studied forehead patch sizes in male and female Pied fly-catchers

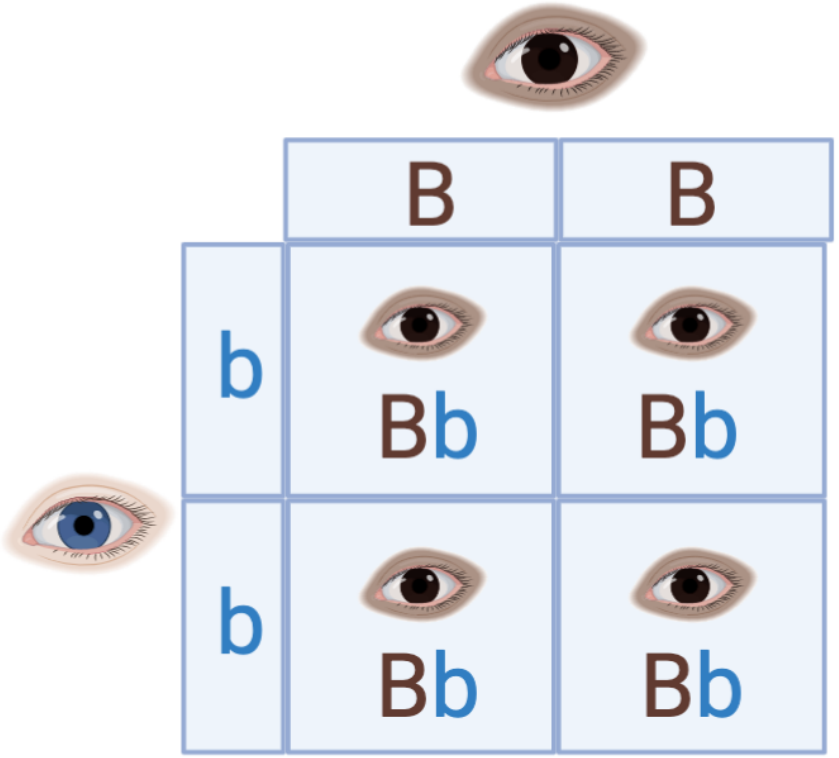
- Stronger covariance is seen between fathers and their sons than between fathers and their daughters





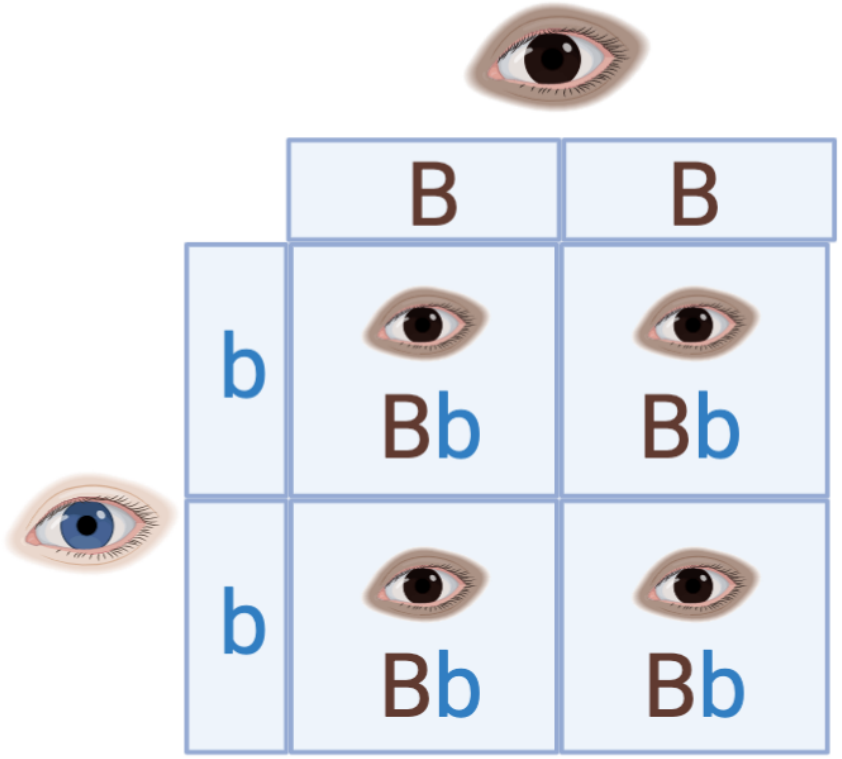
7.1.1 Non-additive variation

- Up until now, we have been discussing additive genetic variation, where we can sum the effects of alleles and loci to estimate phenotype
- Non-additivity can occur when there is dominance of an allele at a locus or when there are interactions across loci (epistasis)
- These complications are incorporated into quantitative genetic models by partitioning genetic variance across multiple components



7.1.1 Non-additive variation

- To create a model for the dominance component of genetic variance, we need to think about how alleles are transmitted across generations
- A parent transmits one allele to its offspring and the other allele is sampled at random from the population
- If your mom transmits allele 1 (frequency = p) to you, then with probability p you will be a 11 homozygote and with probability q you will be a 12 heterozygote



7.1.1 Non-additive variation

- More generally, let's consider a biallelic locus l with frequency p for allele 1 and genotypes 0, 1, and 2 based on copies of allele 1
- The mean phenotypes across environments and genetic backgrounds for each of these genotypes are $\overline{X}_{l,0}$, $\overline{X}_{l,1}$, and $\overline{X}_{l,2}$
- We can "mean center (MC)" these phenotypes by subtracting the mean phenotype across all genotypes: $\overline{X}'_{l,0} = \overline{X}_{l,0} - \mu$
- We can think about the average MC phenotype of an individual who received allele 1 from one parent as being averages of homozygotes (receiving another allele 1) and heterozygotes (receiving allele 2) weighted by the probability of those genotypes:

$$a_{\ell,1} = p\overline{X}'_{\ell,2} + q\overline{X}'_{\ell,1}, \qquad (7.23)$$

7.1.1 Non-additive variation

- Next, we can consider the average phenotype of offspring based on the number of copies of allele 1 that they have:

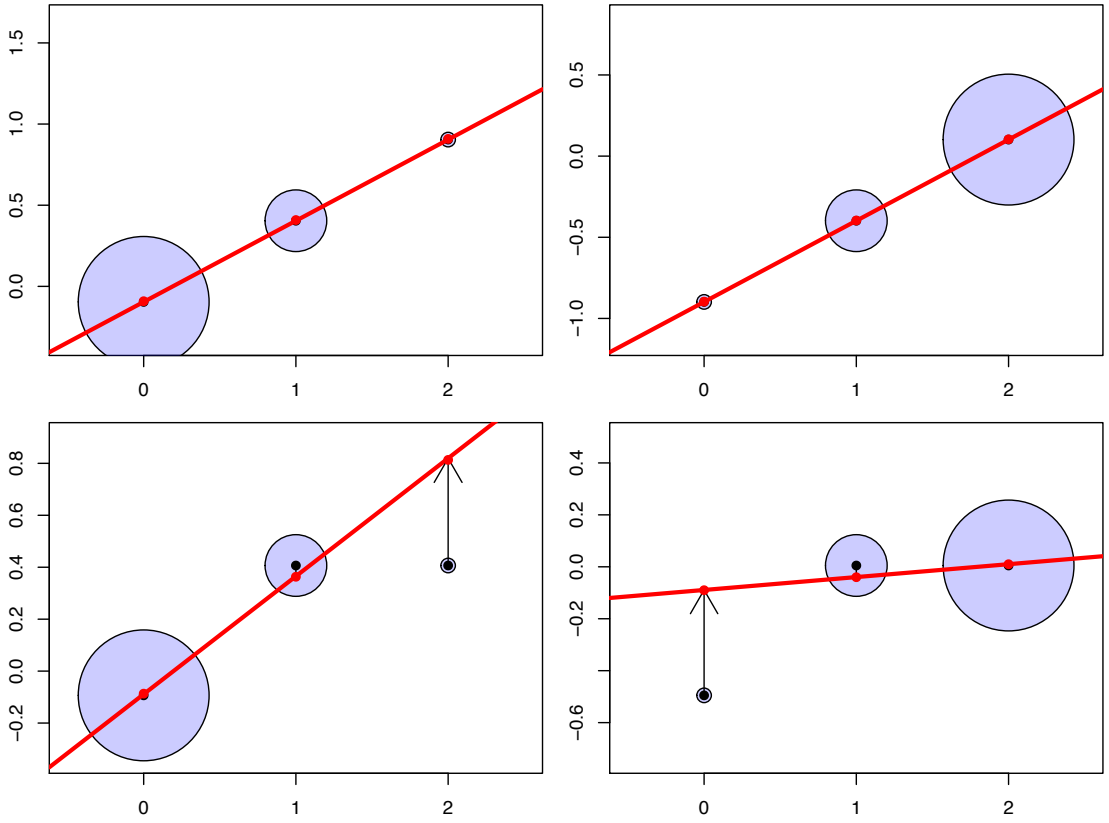
genotype:	0,	1,	2.
additive genetic value:	$a_{\ell,2} + a_{\ell,2}$,	$a_{\ell,1} + a_{\ell,2}$,	$a_{\ell,1} + a_{\ell,1}$

- These are the additive MC genetic values (also known as the breeding values)

- Here, we are considering only additive contributions of alleles in each genotype

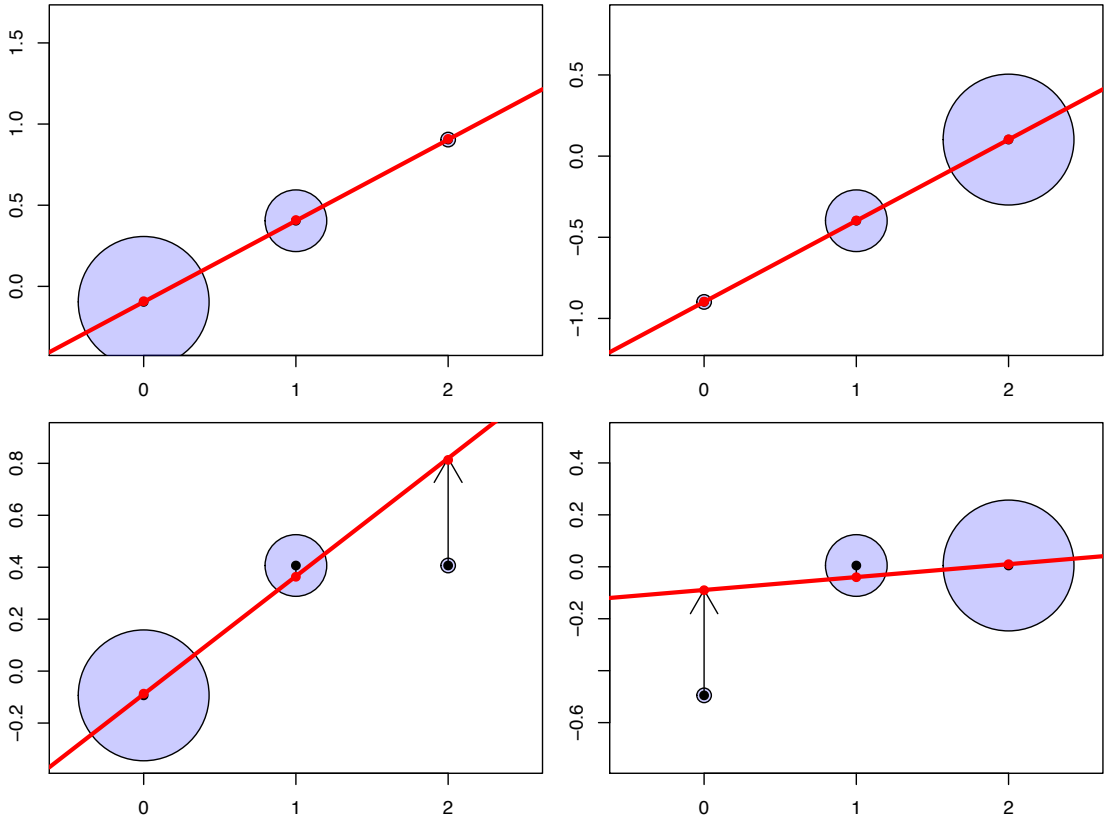
7.1.1 Non-additive variation

- To further illustrate dominance, let's consider an additive locus (top row) and a fully dominant allele (bottom row)
- The MC phenotype from each genotype is represented by a black dot with purple circles illustrating frequency of genotypes in the population
- The additive genetic value is indicated by red dots for each genotype
- The red line is a regression through the additive genotype and phenotype



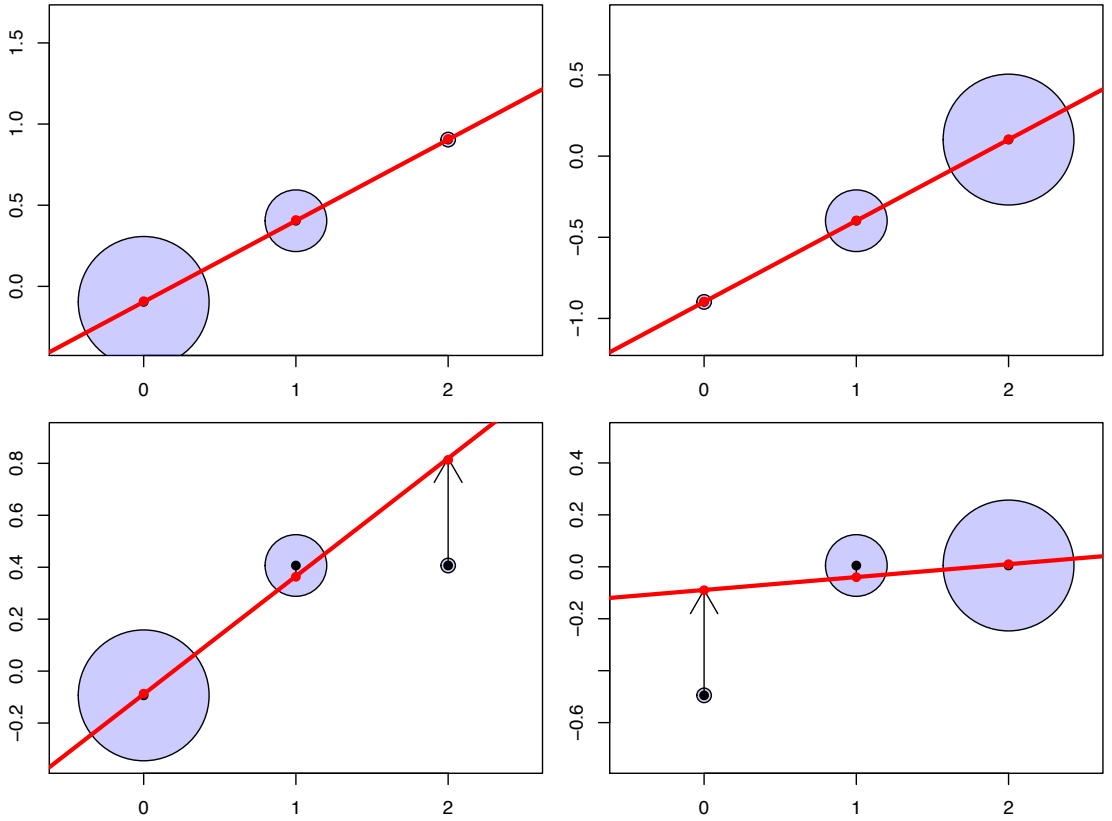
7.1.1 Non-additive variation

- In the additive case, the additive values of genotypes coincide with the MC phenotypes across all genotypic classes (0, 1, and 2)
- The regression line also goes through all three genotypic classes
- In the dominant case, the additive genetic values differ from phenotypic means and are closer to the observed values that are more common in the population



7.1.1 Non-additive variation

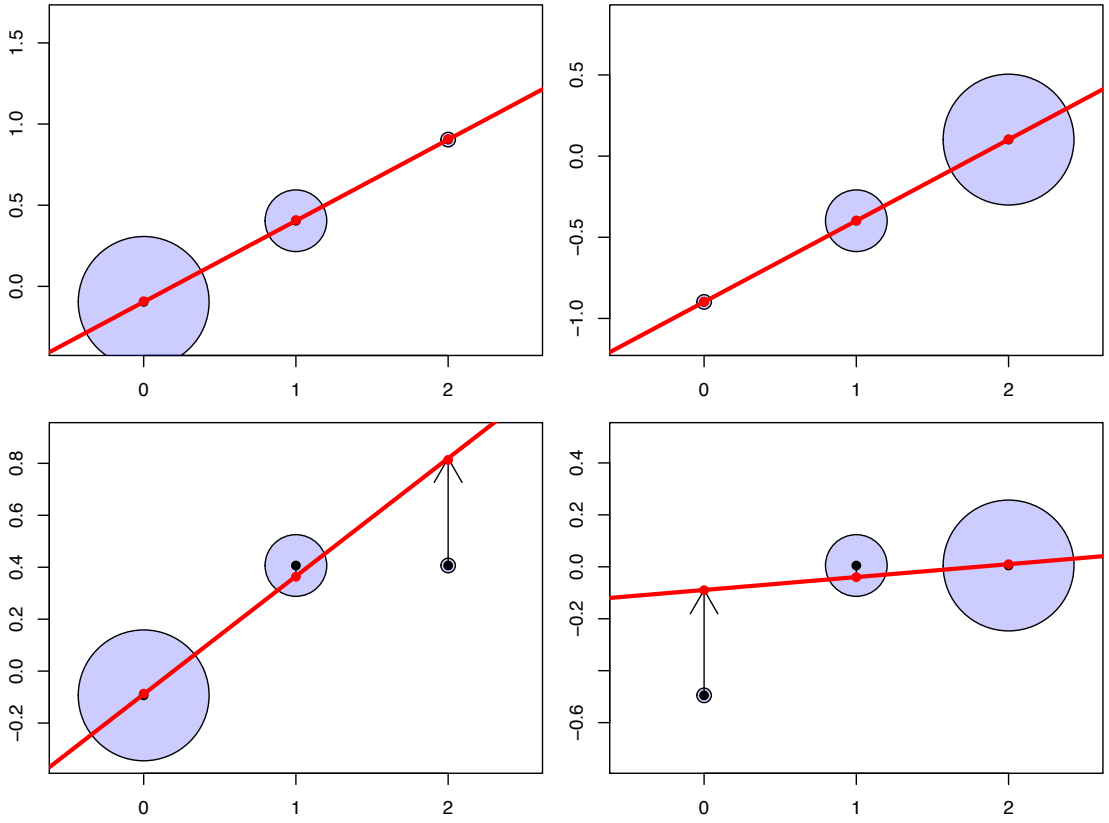
- When there is a difference in the additive effect of two alleles, we can measure this as the effect of swapping an allele 1 for an allele 2:
  - $\alpha_l = \alpha_{l,2} - \alpha_{l,1}$
- This is the same as the regression of phenotype against genotype (the red lines in the figure)
- When there is dominance, our α_l (or slope of the red line) is affected by allele frequency



7.1.1 Non-additive variation

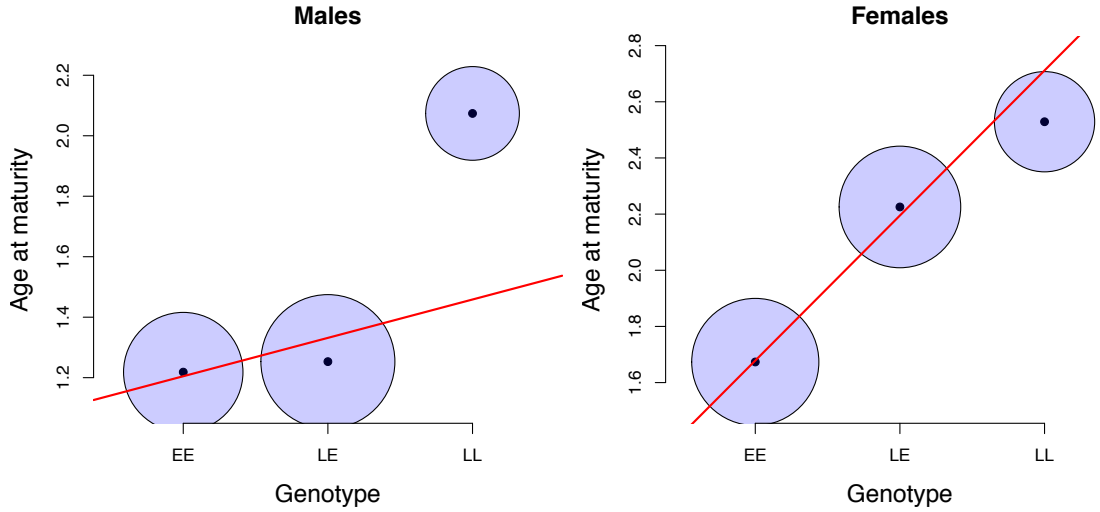
- When the dominant allele (allele 1) is rare, the slope will be greater, because the effect on average phenotypes will be stronger when an allele 2 homozygous genotype becomes a heterozygote

- When the dominant allele is common, the slope will be less because swapping out an allele 2 for an allele 1 will mainly make heterozygotes into 11 homozygotes which have the same phenotype



7.1.1 Non-additive variation

- The genetics of the age of sexual maturity in salmon is a good example of dominance in nature
- A single allele of large effect is found near the VGLL3 gene
- In males, the L allele which causes late maturity is recessive to the E early allele, but causes a 1 year difference in sexual maturity when homozygous
- In females, the L allele is not recessive, showing a much more additive trend



7.1.1 Non-additive variation

- The variance in the population phenotype due to additive breeding values at a locus is:

$$\begin{aligned} V_{A,\ell} &= p^2(2a_{\ell,2})^2 + 2pq(a_{\ell,1} + a_{\ell,1})^2 + q^2(2a_{\ell,0})^2 \\ &= 2(pa_{\ell,1}^2 + qa_{\ell,2}^2) \\ &= 2pq\alpha_\ell^2 \end{aligned} \tag{7.25}$$

- And can be summarized across all loci as:

$$V_A = \sum_{\ell=1}^{L} V_{A,\ell} = \sum_{\ell=1}^{L} 2p_\ell q_\ell \alpha_\ell^2. \tag{7.26}$$

7.1.1 Non-additive variation

- Having partitioned the additive component of variance, we can now consider the dominance component of variance which is defined as “the population variance among genotypes at a locus due to their deviation from additivity”

- For example, the deviation for the heterozygote genotype would be:

$$d_{\ell,1} = \overline{X}'_{\ell,1} - (a_{\ell,1} + a_{\ell,2}) \qquad (7.27)$$

- And dominance variance at a locus is this genotype-frequency-weighted sum of squared dominance deviations

$$V_{D,\ell} = p^2 d^2_{\ell,0} + 2pq d^2_{\ell,1} + q^2 d^2_{\ell,2}. \qquad (7.28)$$

7.1.1 Non-additive variation

- Across all loci in the genome this dominance variance is written as:

$$V_D = \sum_{\ell=1}^{L} V_{D,\ell}. \qquad (7.29)$$

- And now we have partitioned all genetic variance and can write:

$$V_G = V_A + V_D. \qquad (7.30)$$

7.1.1 Non-additive variation

- We can now revisit our narrow sense heritability as $h^2 = \frac{V_A}{V_P} = V_A/(V_G + V_E)$, which is the proportion of phenotypic variance explained by additive genetic variance

- We can also now write an equation for total proportion of phenotypic variance explained by genetic differences among individuals as the broad sense heritability: $H^2 = V_G/(V_G + V_E)$

7.1.1 Non-additive variation

- When dominance is present in loci that affect a trait of interest, we need to modify our equation summarizing phenotypic covariance among relatives to:

$$Cov(X_1, X_2) = 2F_{1,2}V_A + r_2V_D \qquad (7.31)$$

- The dominance variance acts on loci where both alleles are identical by descent, increasing their covariance beyond what might be expected due to additivity

Relationship (i,j)*	$Cov(X_i, X_j)$
parent–child	$1/2V_A$
full siblings	$1/2V_A + 1/4V_D$
identical (monzygotic) twins	$V_A + V_D$
1^{st} cousins	$1/8V_A$

7.1.1 Non-additive variation

Relationship (i,j)*	$Cov(X_i, X_j)$
parent–child	$1/2 V_A$
full siblings	$1/2 V_A + 1/4 V_D$
identical (monzygotic) twins	$V_A + V_D$
1^{st} cousins	$1/8 V_A$

- The fact that parent-child covariance is only a factor of additive genetic variance is an important outcome and reflects the fact that alleles are random draws from the population

- In the short-term this holds, but as allele frequencies change in a population (*e.g.*, dominant alleles become more common), the additive genetic variance will change

7.1.1 Non-additive variation

- Finally, epistasis, as a deviation from additivity is only briefly considered
- Essentially, residual variance between two-locus genotypes after accounting for additive and dominant deviations at each locus individually can be attributed to epistasis
- Epistasis, though, can be particularly difficult to quantify and additive genetic variance alone is typically quite good at predicting short-term evolution