Introduction to Biometrical Genetics
{in the classical twin design}

Conor Dolan et al.
Vrije Universiteit

Boulder 2018 Workshop

Slide acknowledgement: Manuel Ferreira, Pak Sham, Shaun Purcell, Sarah Medland, and Sophie van der Sluis
Slides: 3-7 .... what’s it all about? individual differences
Slides: 8-15 ..... how to quantify individual differences
Slides: 16-18 .... genetic terminology, QTL

Slides: 19-27 ... mean & variance as function of QTL
Slides: 28- 30... interpretation of variance components in regression
Slides: 31- 43 ... covariance as function of QTL and IBD

Slides: 44-48 ... intro to practical (there is no practical!)
“Having 5 fingers genetically determined”

“DNA includes a blueprint to build a hand”
Behavior genetic research is concerned with relating individual differences in phenotypes to individual differences at the genetic level and individual differences in environmental influences.
Phenotype: continuously varying, genetically complex, e.g. (ideally) normally distributed e.g., binary (dichotomous, 0-1 coded) phenotype (based on continuous phenotype; liability threshold model).

The phenotype is a **quantitative** trait, a **metric** trait, a **complex** trait.
Genetically complex:

Individual differences in the phenotype are subject to the effects of many genes of small effects, a.k.a. polygenes, minor genes. How many? Hundreds (Educational Attainment, Height) … Thousands….?

Phenotypic individual differences are attributable to genetic individual differences in a large number of polygenes, a.k.a. QTLs (quantitative trait loci).

Polygenicity implies phenotypic continuous distributions (Nick Martin’s intro talk)
People differ phenotypically

Q. How to quantify individual differences?

The variance: $s^2$, $\sigma^2$, $\sigma^2_X$, $\text{var}(X)$, $V_X$

mean ($X$)

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i$$

variance ($X$)

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2$$

$x_i$ is the phenotypic value of person $i$ ($i=1,\ldots,N$)
We need the covariance: express the phenotypic relatedness among family members.

Formula to find the mean for $X$

$$\mu_x = \frac{\sum_{i=1}^{n} x_i}{n}$$

Formula to find the mean for $Y$

$$\mu_y = \frac{\sum_{i=1}^{n} y_i}{n}$$

Formula to find covariance of $X$ & $Y$

$$cov(X, Y) = \frac{\sum_{i=1}^{n} (x_i - \mu_x)(y_i - \mu_y)}{(n - 1)}$$
Means, Variances and Covariances

\[ \mu = E(X) = \sum_i x_i f(x_i) \]

\[ \text{Var}(X) = E(X - \mu)^2 \]

\[ = \sum_i (x_i - \mu)^2 f(x_i) \]

\[ \text{Cov}(X, Y) = E(X - \mu_X)(Y - \mu_Y) \]

\[ = \sum_i (x_i - \mu_X)(y_i - \mu_Y)f(x_i, y_i) \]

\[ \mu = \frac{1}{N} \sum_{i=1}^{N} x_i \]

\[ \sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2 \]
1,1,2,2,3,4,5,5,6,6

mean = \frac{(1+1+2+2+3+4+5+5+6+6)}{10} = \frac{36}{12} = 3.5

f(1) = \frac{2}{10} = .2 
0.2*1 +
f(2) = \frac{2}{10} = .2 
0.2*2 +
f(3) = \frac{1}{10} = .1 
0.1*3 +
f(4) = \frac{1}{10} = .1 
0.1*4 +
f(5) = \frac{2}{10} = .2 
0.2*5 +
f(6) = \frac{2}{10} = .2 
0.2*6

\text{---------}

3.5

# in R
x=c(1,1,2,2,3,4,5,5,6,6)
mean(x)

\mu = \frac{\sum_{i=1}^{N} x_i}{N} = E(X) = \sum_{i} x_i f \left( x_i \right)
1,1,2,2,2,3,4,5,5,5,6,6

mean = 3.5

\[ f(1) = \frac{2}{10} = 0.2 \]
\[ 0.2 \times (1 - 3.5)^2 + \]

\[ f(2) = \frac{2}{10} = 0.2 \]
\[ 0.2 \times (2 - 3.5)^2 + \]

\[ f(3) = \frac{1}{10} = 0.1 \]
\[ 0.1 \times (3 - 3.5)^2 + \]

\[ f(4) = \frac{1}{10} = 0.1 \]
\[ 0.1 \times (4 - 3.5)^2 + \]

\[ f(5) = \frac{2}{10} = 0.2 \]
\[ 0.2 \times (5 - 3.5)^2 + \]

\[ f(6) = \frac{2}{10} = 0.2 \]
\[ 0.2 \times (6 - 3.5)^2 \]

-----------------

\[ \text{variance} = 3.45 \]

\[ \text{stdev} = \sqrt{\text{variance}} \]
\[ \text{stdev} = \sqrt{3.45} = 1.857 \]

\[ \mu = E(X) = \sum x_i f(x_i) \]
\[ \text{Var}(X) = E(X - \mu)^2 \]
\[ = \sum (x_i - \mu)^2 f(x_i) \]

# in R
\[
x=c(1,1,2,2,3,4,5,5,6,6)
var(x)
\]
covariance

\[ \text{Cov}(X,Y) = E(X - \mu_X)(Y - \mu_Y) \]

\[ \sum_i (x_i - \mu_X)(y_i - \mu_Y)f(x_i, y_i) \]

correlation

\[ \text{Cor}(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)*\text{var}(Y)}} = \frac{\text{Cov}(X,Y)}{\text{stdev}(X)*\text{stdev}(Y)} \]

\[ \text{Cor}(X,Y) \text{ is – stand-alone - interpretable} \]

\[ \text{MZ covariance is 291.... uninterpretable} \]
\[ \text{MZ correlation is .80 .... interpretable} \]
To what extent, and how, are *individual differences* in genetic makeup, and *individual differences* in environmental factors, related to *phenotypic (observed) individual differences*?

To what extent, and how, do *individual differences* in genetic makeup, and *individual differences* in environmental factors, explain *phenotypic (observed) variance*?

\[
\sigma_x^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N - 1}.
\]

\[
\text{Var}(X) = E(X - \mu)^2 = \sum_i (x_i - \mu)^2 f(x_i)
\]
terminology

• **QTL Quantative trait locus**: a sequence of DNA base pairs (may be a SNP: single base pair).

• **Locus**: the site of the specific QTL on a chromosome (22 pairs + XY). Humans are diploid (22 pairs autosomal chromosomes + sex chromosomes XY or XX).

• **Allele**: an alternative form of a gene at a locus.

• **Genotype**: the combination of alleles at a particular locus.

• **Phenotype**: an observed characteristic, which displays individual differences (in part due to genetic differences).
chromosome 9
location 9q34.2

Mendel’s law of segregation
The FNBP1L gene has been associated with intelligence in two studies:


This gene is on chromosome 1 (1p22.1), and it comprises 106531 bases (106.5Kb). Within this gene the SNP rs236330 specifically is associated with intelligence.
1. **Allele frequencies (QTL: diallelic autosomal; e.g., SNP rs236330)**

- A single locus, with two alleles
  - Biallelic a.k.a. diallelic
  - in GWAS: Single nucleotide polymorphism, SNP

- Alleles \( A \) and \( a \)
  - Frequency of \( A \) is \( p \)
  - Frequency of \( a \) is \( q = 1 - p \)

- Every individual inherits two alleles
  - A genotype is the combination of the two alleles
  - e.g. \( AA, aa \) (the homozygotes) or \( Aa \) (the heterozygote)

.........................genotype frequencies?
Biometrical model for single biallelic QTL

- **Biallelic locus**
  - Genotypes: *AA, Aa, aa*
  - Genotype frequencies: *p<sup>2</sup>, 2pq, q<sup>2</sup>*

<table>
<thead>
<tr>
<th>Father's gametes</th>
<th>Mother's gametes (egg)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>A</em> (<em>p</em>)</td>
<td><em>A</em> (<em>p</em>)</td>
</tr>
<tr>
<td><em>A</em> (<em>p</em>)</td>
<td><em>Aa</em> (<em>pq</em>)</td>
</tr>
<tr>
<td><em>a</em> (<em>q</em>)</td>
<td><em>aA</em> (<em>qp</em>)</td>
</tr>
<tr>
<td><em>a</em> (<em>q</em>)</td>
<td><em>aa</em> (<em>q&lt;sup&gt;2&lt;/sup&gt;</em>)</td>
</tr>
</tbody>
</table>

**Hardy-Weinberg Equilibrium** frequencies

\[
P(AA) = p^2 \\
P(Aa) = 2pq \\
P(aa) = q^2
\]

\[p^2 + 2pq + q^2 = 1\]
Phenotype level: contribution to continuous variation

Biometric Model

\[ aa \quad Aa \quad AA \]

\[ \mu - a \quad \mu + d \quad \mu + a \]

- \( a \) represents the genotypic effect
- \( d \) represents the environmental deviation

Genotypic effect

Means within each genotype (aa, Aa, AA)

...conditional on genotype

Take all aa individuals and calculate their mean phenotypic value:

\[ \mu - a \] (the phenotypic mean conditional on genotype aa)
### Biometrical model for single biallelic QTL

#### 1. Contribution of the QTL to the Mean

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>AA</th>
<th>Aa</th>
<th>aa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect, $x$</td>
<td>$\mu + a$</td>
<td>$\mu + d$</td>
<td>$\mu - a$</td>
</tr>
<tr>
<td>Frequencies, $f(x)$</td>
<td>$p^2$</td>
<td>$2pq$</td>
<td>$q^2$</td>
</tr>
</tbody>
</table>

$$(\mu+a)(p^2) + (\mu+d)(2pq) + (\mu-a)(q^2) = \mu + a(p^2) + d(2pq) - a(q^2) = \mu + a(p-q) + 2pqd. \ (pop \ pheno \ mean)$$

$$\mu = E(X) = \sum_i x_i f(x_i)$$

**contribution of the QTL to the population phenotypic mean**

$m = a(p-q) + 2pqd$
2. Contribution of the QTL to the Variance ($X$)

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>AA</th>
<th>Aa</th>
<th>aa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect ($x$)</td>
<td>$\mu + a$</td>
<td>$\mu + d$</td>
<td>$\mu - a$</td>
</tr>
<tr>
<td>Frequencies, $f(x)$</td>
<td>$p^2$</td>
<td>$2pq$</td>
<td>$q^2$</td>
</tr>
</tbody>
</table>

$$s_{QTL}^2 = (a-m)^2 p^2 + (d-m)^2 2pq + (-a-m)^2 q^2$$

$$m = a(p-q) + 2pqd$$

$$Var(X) = E(X - \mu)^2 = \sum_i (x_i - \mu)^2 f(x_i)$$
Biometrical model for single biallelic QTL

\[ s^2_{QTL} = (a-m)^2 p^2 + (d-m)^2 2pq + (-a-m)^2 q^2 \]

\[ = 2pq[a+(q-p)d]^2 + (2pqd)^2 \]

\[ = s^2_{QTL(A)} + s^2_{QTL(D)} \]

Additive or linear effects give rise to variance component

\[ s^2_{QTL(A)} = 2pq[a+(q-p)d]^2 \]

Dominance or within local allelic interaction effects give rise to variance component

\[ s^2_{QTL(D)} = (2pqd)^2 \]
Biometrical model for single biallelic QTL

\[ s_{QTL}^2 = (a-m)^2 p^2 + (d-m)^2 2pq + (-a-m)^2 q^2 \]

\[ = 2pq[a+(q-p)d]^2 + (2pqd)^2 \]

\[ = s_{QTL(A)}^2 + s_{QTL(D)}^2 \]

**Additive effects:** \( s_{QTL(A)}^2 = 2* pq[a]^2 \)

**Dominance effects:** \( s_{QTL(D)}^2 = 0 \)
Biometrical model for single biallelic QTL

\[
S^2_{QTL} = (a-m)^2p^2 + (d-m)^22pq + (-a-m)^2q^2
\]

\[
= 2pq[a+(q-p)d]^2 + (2pqd)^2
\]

\[
= S^2_{QTL(A)} + S^2_{QTL(D)}
\]

**Additive effects:** \( s^2_{QTL(A)} = 2* pq[a+(q-p)d]^2 \)

**Dominance effects:** \( s^2_{QTL(D)} = (2pqd)^2 \)
Suppose we measure the QTL and the phenotype and regress X on QTL. The scatterplot of the data (aa coded -1; Aa coded 0; AA coded 1).

In the following slides we look at the regression lines only (not plotting the residuals – just to avoid clutter).

we ask: *how much of the phenotypic variance is explained by the predictor (QTL)*
Explained variance:

\[ s^2_{QTL(A)} = 2^* pq[a]^2 \]

\[ s^2_{QTL(D)} = 0 \]
$$s^2_{QTL(A)} = 2pq[a + (q-p)d]^2$$

**Not explained**

$$s^2_{QTL(D)} = (2pqd)^2$$

**Explained variance (blue line):**

$$s^2_{QTL(A)} = 2pq[a + (q-p)d]^2$$
$s^2_{QTL(A)}$ always greater than zero

$s^2_{QTL(D)}$ can be zero (additive model $d=0$)
Biometrical model for single biallelic QTL

3. Contribution of the QTL to the Cov $(X, Y)$

$$Cov(X, Y) = \sum_{i} (x_i - \mu_X) (y_i - \mu_Y) f(x_i, y_i)$$

<table>
<thead>
<tr>
<th></th>
<th>AA $(a-m)$</th>
<th>Aa $(d-m)$</th>
<th>aa $(-a-m)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA $(a-m)$</td>
<td>$(a-m)^2$</td>
<td>$(a-m)(d-m)$</td>
<td>$(a-m)(-a-m)$</td>
</tr>
<tr>
<td>Aa $(d-m)$</td>
<td>$(a-m)(d-m)$</td>
<td>$(d-m)^2$</td>
<td>$(d-m)(-a-m)$</td>
</tr>
<tr>
<td>aa $(-a-m)$</td>
<td>$(a-m)(-a-m)$</td>
<td>$(d-m)(-a-m)$</td>
<td>$(-a-m)^2$</td>
</tr>
</tbody>
</table>

$m = a(p-q) + 2pqd$

What about the $f(x_i, y_i)$?
Biometrical model for single biallelic QTL

3A. Contribution of the QTL to the Cov \((X, Y)\) – MZ twins

\[
\text{Cov}(X, Y) = \sum_i (x_i - \mu_X)(y_i - \mu_Y)f(x_i, y_i)
\]

<table>
<thead>
<tr>
<th></th>
<th>AA (a-m)</th>
<th>Aa (d-m)</th>
<th>aa (-a-m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA (a-m)</td>
<td>(p^2(a-m)^2)</td>
<td>0 (a-m) (d-m)</td>
<td>0 (a-m) (-a-m)</td>
</tr>
<tr>
<td>Aa (d-m)</td>
<td>0 (a-m) (d-m)</td>
<td>2pq (d-m)^2</td>
<td>0 (d-m) (-a-m)</td>
</tr>
<tr>
<td>aa (-a-m)</td>
<td>0 (a-m) (-a-m)</td>
<td>0 (d-m) (-a-m)</td>
<td>q^2 (-a-m)^2</td>
</tr>
</tbody>
</table>

\[
\text{Covar}(X_i, X_j) = (a-m)^2p^2 + (d-m)^22pq + (-a-m)^2q^2
\]

\[
= 2pq[a+(q-p)d]^2 + (2pqd)^2 = s^2_{QTL(A)} + s^2_{QTL(D)}
\]
Biometrical model for single biallelic QTL

3B. Contribution of the QTL to the Cov (X, Y) – Parent-Offspring

\[
\text{Cov}(X, Y) = \sum_i (x_i - \mu_X)(y_i - \mu_Y)f(x_i, y_i)
\]

<table>
<thead>
<tr>
<th></th>
<th>AA (a-m)</th>
<th>Aa (d-m)</th>
<th>aa (-a-m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA (a-m)</td>
<td>(p^3(a-m)^2)</td>
<td>(p^2q(a-m)(d-m))</td>
<td>0 (a-m) (-a-m)</td>
</tr>
<tr>
<td>Aa (d-m)</td>
<td>(p^2q(a-m)(d-m))</td>
<td>(pq(d-m)^2)</td>
<td>(pq^2(d-m)(-a-m))</td>
</tr>
<tr>
<td>aa (-a-m)</td>
<td>0 (a-m) (-a-m)</td>
<td>(pq^2(d-m)(-a-m))</td>
<td>(q^3(-a-m)^2)</td>
</tr>
</tbody>
</table>
given an AA parent, an AA offspring can come from either AA x AA or AA x Aa parental mating types

\[ \text{AA x AA} \quad \text{will occur} \quad p^2 \times p^2 = p^4 \]
and have AA offspring \( \text{Prob}(\text{AA}) = 1 \)

\[ \text{AA x Aa} \quad \text{will occur} \quad p^2 \times 2pq = 2p^3q \]
and have AA offspring \( \text{Prob}(\text{AA}) = 0.5 \)
and have Aa offspring \( \text{Prob}(\text{Aa}) = 0.5 \)

Therefore, \( \text{P(AA parent & AA offspring)} = p^4 + 0.5 \times 2 \times p^3q \)
\[ = p^3(p+q) \]
\[ = p^3 \]
So can be complicated, but can also be simple ….

<table>
<thead>
<tr>
<th>Parent</th>
<th>AA (a-m)</th>
<th>Aa (d-m)</th>
<th>aa (-a-m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA (a-m)</td>
<td>$p^3(a-m)^2$</td>
<td>$p^2q(a-m)(d-m)$</td>
<td>0 (a-m)(-a-m)</td>
</tr>
<tr>
<td>Aa (d-m)</td>
<td>$p^2q(a-m)(d-m)$</td>
<td>$pq(d-m)^2$</td>
<td>$pq^2(d-m)(-a-m)$</td>
</tr>
<tr>
<td>aa (-a-m)</td>
<td>0 (a-m)(-a-m)</td>
<td>$pq^2(d-m)(-a-m)$</td>
<td>$q^3(-a-m)^2$</td>
</tr>
</tbody>
</table>

why zero probability?
### Biometrical model for single biallelic QTL

#### 3B. Contribution of the QTL to the Cov \((X, Y)\) – Parent-Offspring

<table>
<thead>
<tr>
<th></th>
<th>AA ((a-m))</th>
<th>Aa ((d-m))</th>
<th>aa ((-a-m))</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA ((a-m))</td>
<td>(p^3(a-m)^2)</td>
<td>(p^2q(a-m)(d-m))</td>
<td>0 ((a-m)(-a-m))</td>
</tr>
<tr>
<td>Aa ((d-m))</td>
<td>(p^2q(a-m)(d-m))</td>
<td>(pq(d-m)^2)</td>
<td>(pq^2(d-m)(-a-m))</td>
</tr>
<tr>
<td>aa ((-a-m))</td>
<td>0 ((a-m)(-a-m))</td>
<td>(pq^2(d-m)(-a-m))</td>
<td>(q^3(-a-m)^2)</td>
</tr>
</tbody>
</table>

\[
\text{Cov} (X_i, X_j) = (a-m)^2p^3 + \ldots + (-a-m)^2q^3
\]

\[
= pq[a+(q-p)d]^2 = \frac{1}{2} S_{QTL(A)}^2
\]
Biometrical model for single biallelic QTL

3C. Contribution of the QTL to the Cov \((X, Y)\) – Unrelated individuals

<table>
<thead>
<tr>
<th></th>
<th>(p^2)</th>
<th>(2pq)</th>
<th>(q^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(AA) ((a-m))</td>
<td>(p^2) ((a-m))</td>
<td>(p^2) ((a-m))</td>
<td>(q^2) ((a-m))</td>
</tr>
<tr>
<td>(Aa) ((d-m))</td>
<td>(2pq) ((d-m))</td>
<td>(2pq) ((d-m))</td>
<td>(q^2) ((d-m))</td>
</tr>
<tr>
<td>(aa) ((-a-m))</td>
<td>(q^2) ((-a-m))</td>
<td>(q^2) ((-a-m))</td>
<td>(q^2) ((-a-m))</td>
</tr>
</tbody>
</table>

\[
\text{Cov} \,(X_i, X_j) = (a-m)^2 p^4 + \ldots + (-a-m)^2 q^4
\]

= 0
Follow same method for full sibs and DZ twins
Derive genotype frequencies ....

<table>
<thead>
<tr>
<th>s1</th>
<th>s2</th>
<th>eff</th>
<th>eff</th>
<th>frequency ( (p(A)=p, p(a)=1-p) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>AA</td>
<td>a</td>
<td>a</td>
<td>( p^{**4}+p^{**3}q+p^{**2}q^{**2}/4 )</td>
</tr>
<tr>
<td>aa</td>
<td>aa</td>
<td>-a</td>
<td>-a</td>
<td>( p^{**2}q^{**2}/4+p*q^{**3}+q^{**4} )</td>
</tr>
<tr>
<td>Aa</td>
<td>Aa</td>
<td>d</td>
<td>d</td>
<td>( p^{**3}q+3p^{**2}q^{**2}+p*q^{**3} )</td>
</tr>
<tr>
<td>AA</td>
<td>Aa</td>
<td>a</td>
<td>d</td>
<td>( p^{**3}q+p^{**2}q^{**2}/2 )</td>
</tr>
<tr>
<td>Aa</td>
<td>AA</td>
<td>d</td>
<td>a</td>
<td>( p^{**3}q+p^{**2}q^{**2}/2 )</td>
</tr>
<tr>
<td>Aa</td>
<td>aa</td>
<td>d</td>
<td>-a</td>
<td>( p^{**2}q^{**2}/2+p*q^{**3} )</td>
</tr>
<tr>
<td>aa</td>
<td>Aa</td>
<td>-a</td>
<td>d</td>
<td>( p^{**2}q^{**2}/2+p*q^{**3} )</td>
</tr>
<tr>
<td>AA</td>
<td>aa</td>
<td>a</td>
<td>-a</td>
<td>( p^{**2}q^{**2}/4 )</td>
</tr>
<tr>
<td>aa</td>
<td>AA</td>
<td>-a</td>
<td>a</td>
<td>( p^{**2}q^{**2}/4 )</td>
</tr>
</tbody>
</table>
Biometrical model for single biallelic QTL

3B. Contribution of the QTL to the Cov \((X, Y)\) – DZ twins

<table>
<thead>
<tr>
<th></th>
<th>AA (a-m)</th>
<th>Aa (d-m)</th>
<th>aa (-a-m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA (a-m)</td>
<td>(r1(a-m)^2)</td>
<td>(r4 (a-m) (d-m))</td>
<td>(r6(a-m)(-a-m))</td>
</tr>
<tr>
<td>Aa (d-m)</td>
<td>(r4 (a-m) (d-m))</td>
<td>(r2 (d-m)^2)</td>
<td>(r5 (d-m)(-a-m))</td>
</tr>
<tr>
<td>aa (-a-m)</td>
<td>(r6(a-m)(-a-m))</td>
<td>(r5 (d-m)(-a-m))</td>
<td>(r3 (-a-m)^2)</td>
</tr>
</tbody>
</table>

\[
\text{Cov} (X_i, X_j) = (a-m)^2 r1 + \ldots + (-a-m)^2 r3
\]

\[
= \frac{1}{2} 2pq[a+(q-p)d]^2 + \frac{1}{4}(2pqd)^2 = \frac{1}{2} S^2_{QTL(A)} + \frac{1}{4} S^2_{QTL(D)}
\]
And Now For Something Completely Different

John Cleese ... A famous British person
Random segregation and identity-by-descent (IBD) in sibpairs
**IDENTITY BY DESCENT (IBD) MZs**

<table>
<thead>
<tr>
<th></th>
<th>Sib 1</th>
<th>Sib 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>100% MZ sibs share BOTH parental alleles</td>
<td>IBD = 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>0 sibs share ONE parental allele</td>
<td>IBD = 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0 sibs share NO parental alleles</td>
<td>IBD = 0</td>
<td></td>
</tr>
</tbody>
</table>

- Red fields represent red alleles.
- Yellow fields represent yellow alleles.
- Pink fields represent pink alleles.
### Identity by Descent (IBD) DZs

#### Sib 1

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

#### Sib 2

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

- **4/16 = 1/4 sibs share BOTH parental alleles**  \( \text{IBD} = 2 \)
- **8/16 = 1/2 sibs share ONE parental allele**  \( \text{IBD} = 1 \)
- **4/16 = 1/4 sibs share NO parental alleles**  \( \text{IBD} = 0 \)
What about parent offspring? many alleles do they share IBD?
Biometrical model for single biallelic QTL

3D. Contribution of the QTL to the Cov \((X, Y)\) – DZ twins and full sibs

<table>
<thead>
<tr>
<th># identical alleles inherited from parents</th>
<th>2 (father)</th>
<th>1 (mother)</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>MZ twins</td>
<td>¼ (2 alleles)</td>
<td>+</td>
<td>½ (1 allele)</td>
</tr>
</tbody>
</table>

\[
DZ \text{ Cov} (X_i, X_j) = \frac{1}{4} \text{Cov(MZ)} + \frac{1}{2} \text{Cov(P-O)} + \frac{1}{4} \text{Cov(Unrel)}
\]

\[
= \frac{1}{4}(s_{QTL(A)}^2 + s_{QTL(D)}^2) + \frac{1}{2}(\frac{1}{2} s_{QTL(A)}^2) + \frac{1}{4} (0)
\]

\[
= \frac{1}{2} s_{QTL(A)}^2 + \frac{1}{4} s_{QTL(D)}^2
\]
Biometrical model predicts contribution of a QTL to the mean, variance and covariances of a trait (discarding environmental effects)

**1 QTL**

\[
\text{Var} (X) = s^2_{QTL(A)} + s^2_{QTL(D)}
\]

\[
\text{Cov} (MZ) = s^2_{QTL(A)} + s^2_{QTL(D)}
\]

\[
\text{Cov} (DZ) = \frac{1}{2} s^2_{QTL(A)} + \frac{1}{4} s^2_{QTL(D)}
\]

\[
\text{Cov} (P-O) = \frac{1}{2} s^2_{QTL(A)}
\]

**Multiple QTL**

\[
\text{Var} (X) = \sum (s^2_{QTL(A)}) + \sum (s^2_{QTL(D)}) = V_A + V_D
\]

\[
\text{Cov} (MZ) = \sum (s^2_{QTL(A)}) + \sum (s^2_{QTL(D)}) = V_A + V_D
\]

\[
\text{Cov} (DZ) = \sum (\frac{1}{2} s^2_{QTL(A)}) + \sum (\frac{1}{4} s^2_{QTL(D)}) = \frac{1}{2} V_A + \frac{1}{4} V_D
\]

\[
\text{Cov} (P-O) = \sum (\frac{1}{2} s^2_{QTL(A)}) = \frac{1}{2} V_A
\]
Contributions of $V_A$ and $V_D$ to covariances between relatives

<table>
<thead>
<tr>
<th>Relationship</th>
<th>contribution of QTL to covariance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$V_A$</td>
</tr>
<tr>
<td>Sibling (DZ twin)</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>MZ twin</td>
<td>1</td>
</tr>
<tr>
<td>Half-sibling</td>
<td>$\frac{1}{4}$</td>
</tr>
<tr>
<td>First cousin</td>
<td>$\frac{1}{8}$</td>
</tr>
<tr>
<td>Parent-offspring</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>Avuncular</td>
<td>$\frac{1}{4}$</td>
</tr>
<tr>
<td>Grand-parent</td>
<td>$\frac{1}{8}$</td>
</tr>
<tr>
<td>Unrelated</td>
<td>0</td>
</tr>
</tbody>
</table>

average proportion of alleles shared identically by descent

proportion of IBD=2

These proportions tell us how much of $V_A \& V_D$ contribute to the phenotypic covariance among family members (useful info in extended twin design / extended pedigrees)
Slide acknowledgement: Manuel Ferreira, Pak Sham, Shaun Purcell, Sarah Medland, and Sophie van der Sluis

see also Manuel AR Ferreira’s
http://slidegur.com/doc/4322268/biometrical-genetics