

Exercise: *Effect of the IL6R gene on IL-6R concentration*

Behav Genet (2014) 44:368–382
DOI 10.1007/s10519-014-9656-8

ORIGINAL RESEARCH



The Contribution of the Functional *IL6R* Polymorphism rs2228145, eQTLs and Other Genome-Wide SNPs to the Heritability of Plasma sIL-6R Levels

Jenny van Dongen • Rick Jansen • Dirk Smit • Jouke-Jan Hottenga • Hamdi Mbarek •
Gonneke Willemsen • Cornelis Klufft • AAGC Collaborators • Brenda W. J. Penninx •
Manuel A. Ferreira • Dorret I. Boomsma • Eco J. C. de Geus

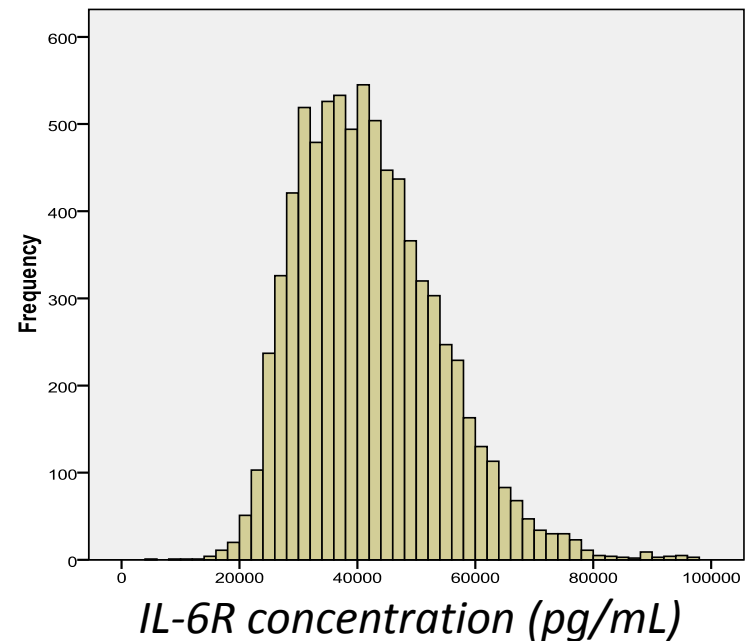
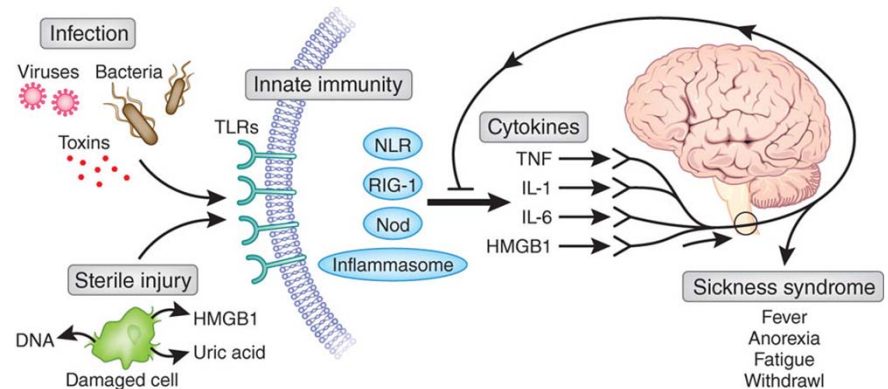
Analyses:

- 1 Heritability: twins – families
- 2 SNP heritability: GCTA
- 3 linkage (IBD from Merlin / linkage QTDT)
- 4 GWA (Plink)
- 5 Linkage & association (QTDT)
- 6 Gene expression (RNA expression data; maybe Friday)

Exercise: Apply the biometrical model to the empirical results

Effect of the IL6R gene on IL-6R concentration

- **Interleukin-6 receptor (IL-6R)** is involved in immune system
- We measured IL-6R concentration in blood in ~5000 individuals
- IL-6R concentration in blood is a **quantitative trait**



Genetics → IL-6R concentration → common disease

- IL-6R protein is encoded by the ***IL6R*** gene (chromosome 1)
- *IL6R* gene important for **several common diseases**
- A SNP (single nucleotide polymorphism) in the *IL6R* gene (rs2228145) is associated with:
 - Asthma¹
 - Coronary heart disease²
 - Type 1 diabetes³

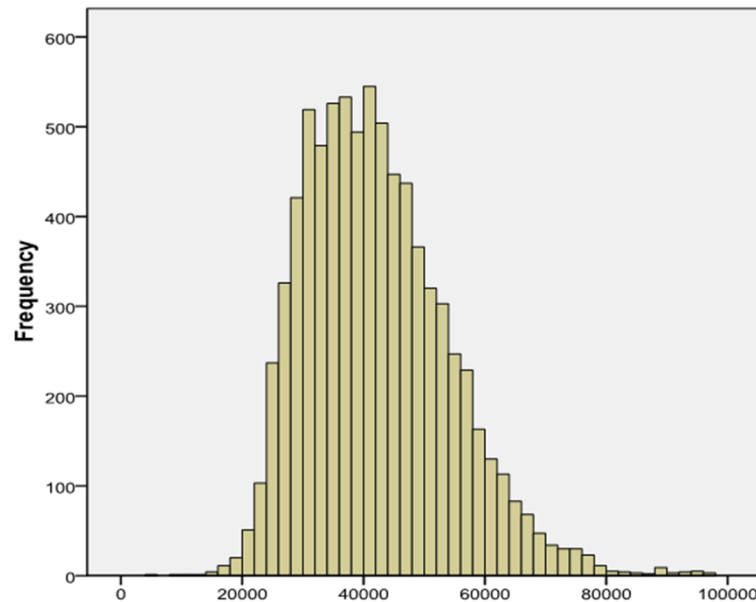
¹Ferreira M.A. *et al* Lancet 2011

²*IL6R* consortium Lancet 2012

³Ferreira R.C. *et al* PLoS Genetics 2013

Genetics → IL-6R concentration

- Functional studies show that the SNP affects IL-6R production
 - Rs2228145: **Large effect** on sIL-6R level
 - **how much variance does this SNP explain?**
 - **sIL-6R level affected by more genetic variants or only this SNP?**



Mean=4.17
Variance=1.35

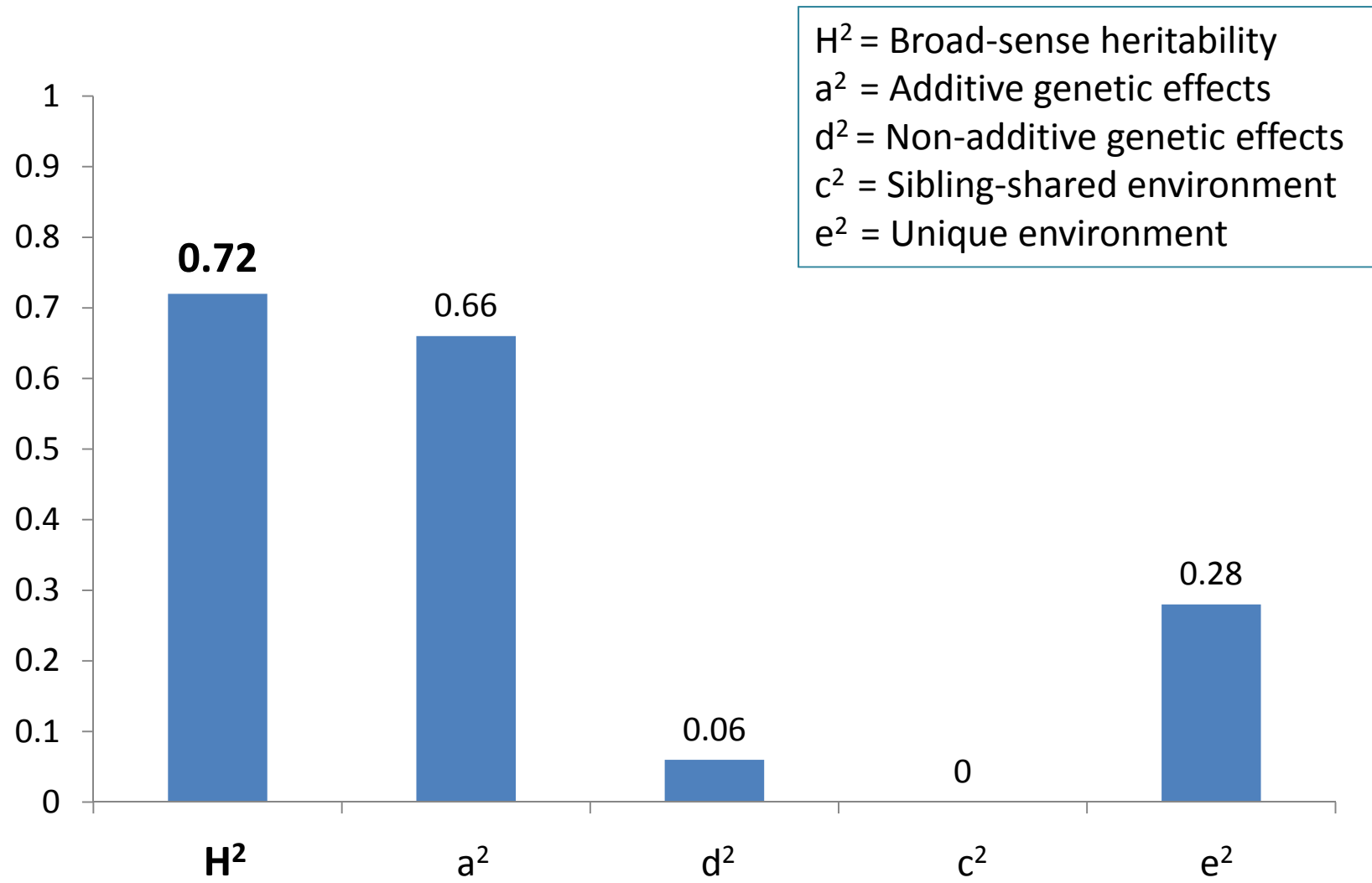
*IL-6R concentration (pg/mL)
in 5000 people*

Aim: How much variance in IL-6R levels is explained by the known SNP and how much by other genetic variants?

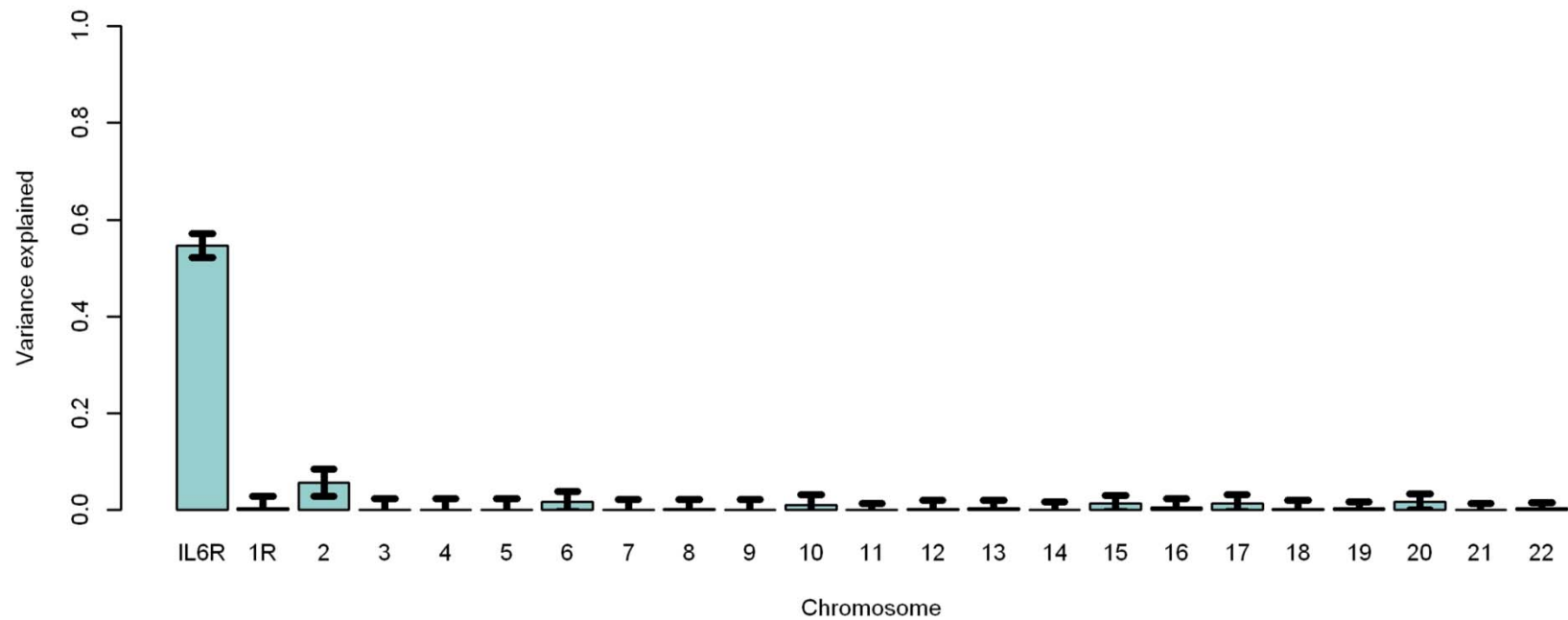
Methods

- We measured IL-6R concentration in ~5000 twins & parents & siblings
- We estimated **Heritability**: Variance of sIL-6R level explained by total genetic effects
- We measured genome-wide SNP genotypes of the same subjects:
 - How much variance is explained by **all SNPs in the genome** (Genomewide-complex trait analysis, GCTA)
 - How much variance is explained by **all genetic variation in the *IL6R* gene** (linkage analysis)
 - How much variance is explained by **the SNP rs2228145**

Heritability of sIL-6R level



Variance explained by chromosome-wide SNPs

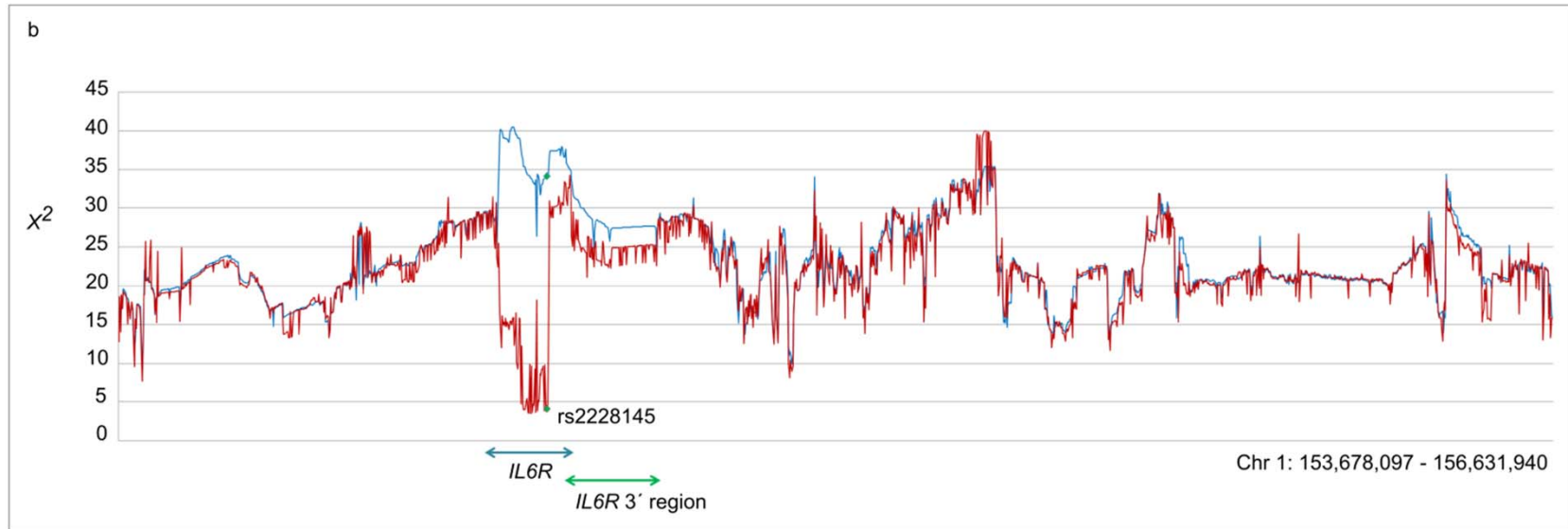


SNPs in the *IL6R* gene on Chromosome 1 (+/- 10MB): **54.7 %** (SE=2.5%)

Combined linkage and association analysis

— Chi-squared from linkage test

— Chi-squared from linkage test – while modeling association for individual SNPs



Chr1: (q21.3) 

IL6R region:

1. Variance explained by linkage (V_A/V_{total}): **69 %**
2. Variance explained by linkage after correction for rs2228145: **19%**

Exercise: *Effect of the IL6R gene on IL-6R concentration*

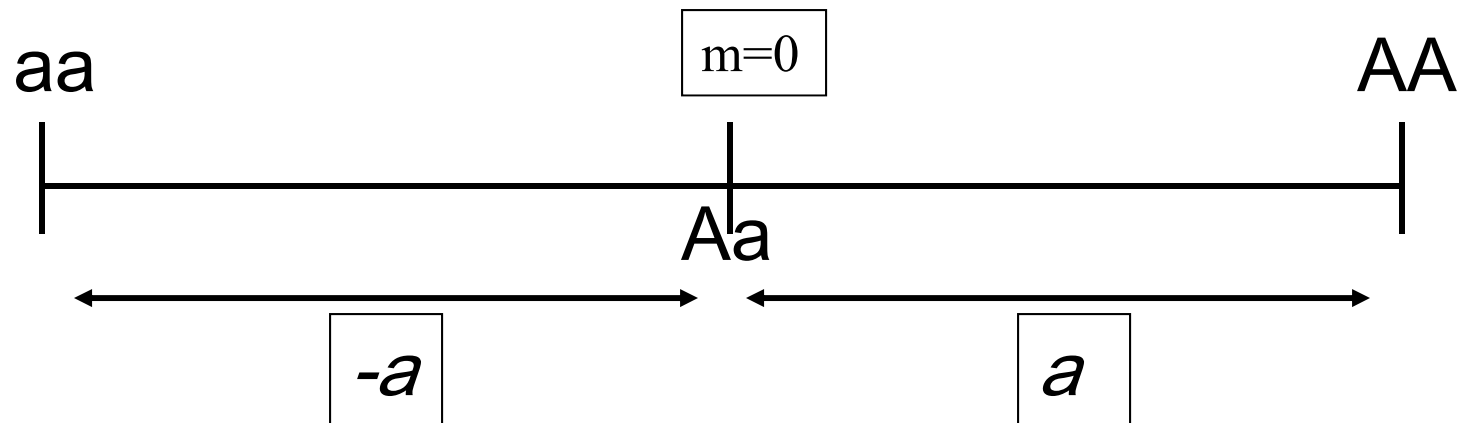
INFORMATION

- The SNP (single nucleotide polymorphism) has 2 alleles:
 - Minor allele: C, frequency: $p=0.39$
 - Major Allele: A, frequency: $q=0.61$
- Mean IL-6R concentration of each genotype:
 - CC: 5.698 (10^{-8} g/mL)
 - CA: 4.418 (10^{-8} g/mL)
 - AA: 3.238 (10^{-8} g/mL)
- Total Variance of IL-6R concentration=1.35

QUESTIONS (Falconer & MacKay; 1996: Introduction to quantitative genetics)

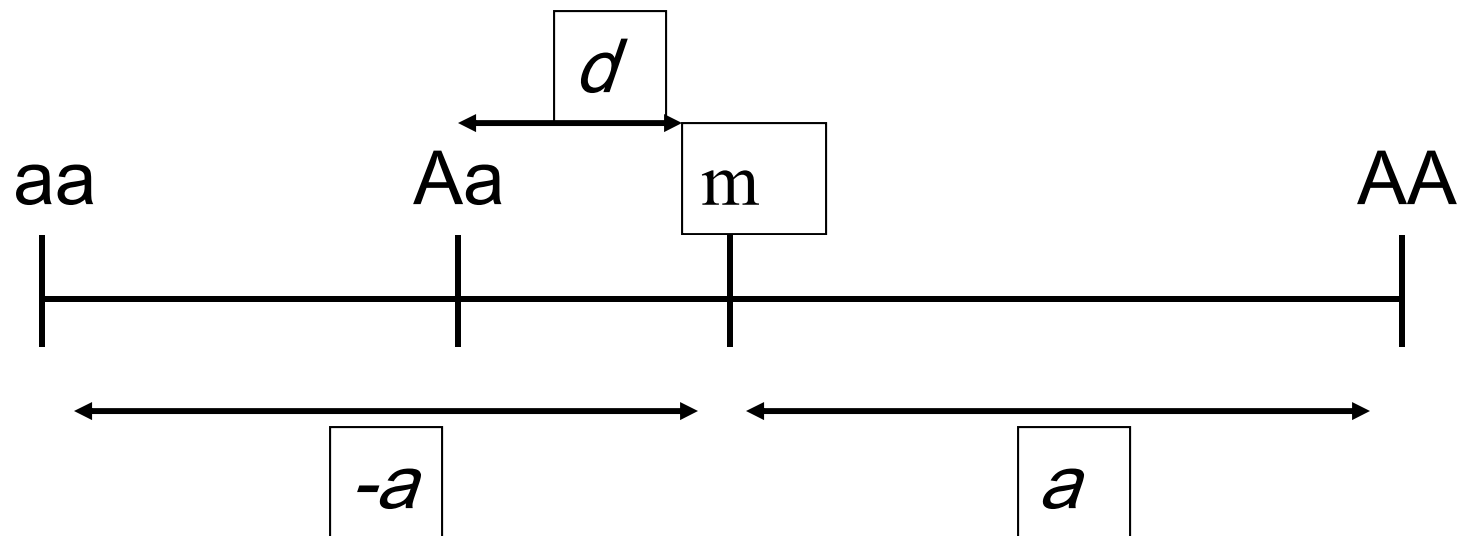
1. Calculate genotypic values (a and d) (page 109)
2. [Calculate the average effect of the alleles (page 113)]
3. Calculate the genotype frequencies (page 7)
4. Calculate the mean IL6-R concentration in the population (page 110)
5. Calculate how much of the variance is explained by this SNP
(*Variance= Sum of squared deviations from the mean*)
6. Calculate heritability

Model: gene with 2 alleles A and a
and 3 genotypes AA, Aa and aa



The difference on a quantitative scale between AA and aa is $2a$.
The middle (m) is zero and the value of Aa is 0 (no dominance).

Model: gene with 2 alleles A and a
and 3 genotypes AA, Aa and aa



The deviation from m (middle) of the heterozygote Aa is d :
partial dominance.

Genotype (i)	AA	Aa	aa
Frequency (f)	p^2	$2pq$	q^2
Genotypic effect (x)	a	d	-a

Mean?

Genotype (i)	AA	Aa	aa
Frequency (f)	p^2	$2pq$	q^2
Genotypic effect (x)	a	d	-a
$f * x$	$p^2 a$	$2pqd$	$-q^2 a$

$$\text{mean: } p^2 a + 2pqd - q^2 a =$$

(recall $p+q = 1$)

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

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

$$\text{Mean} = a(p-q) + 2pqd$$

$a(p-q)$: attributable to homozygotes

$2pqd$: attributable to heterozygotes

Genotype (i)	AA	Aa	aa
Frequency (f)	p^2	$2pq$	q^2
Genotypic effect (x)	a	d	-a
$f * x$	$p^2 a$	$2pqd$	$- q^2 a$

mean: $p^2 a + 2pqd - q^2 a = a(p-q) + 2pqd$

Variation: $2pq[a+d(q-p)]^2 + (2pqd)^2$

Population variation depends on 'a' (difference between homozygote individuals), 'd' (deviation of heterozygote persons from zero) and on allele frequency (p & q).

Average effect

(associated with genes and not with genotypes)

The average effect of a gene (allele) is the mean deviation from the population mean of individuals which received that gene from one parent, the gene received from the other parent having come *at random* from the population.

Falconer (p112): The concept of average effect is not easy to grasp.

See additional slides

Average effect is related to genotypic values a and d

$$q [a + d (q - p)] = \alpha_1$$

$$-p [a + d (q - p)] = \alpha_2$$

Average effect of gene substitution is $\alpha_1 - \alpha_2 = \alpha$. This is the difference between the average effect of the 2 alleles:

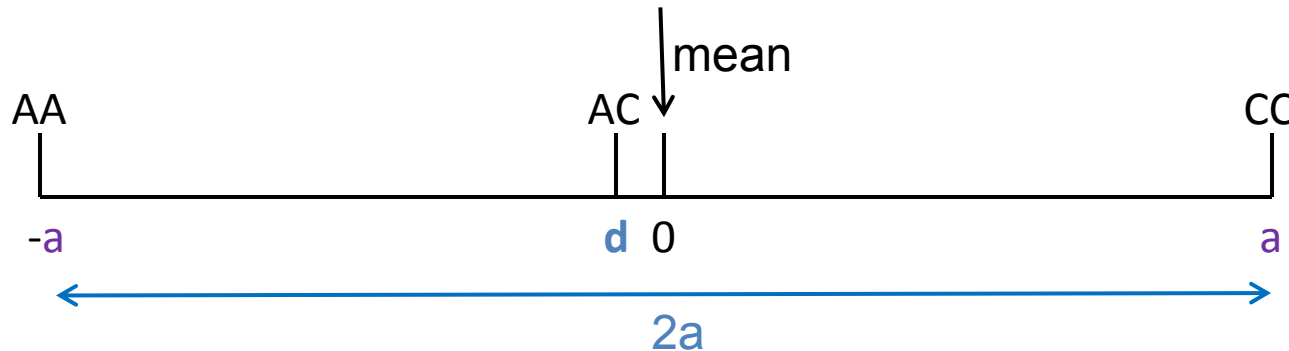
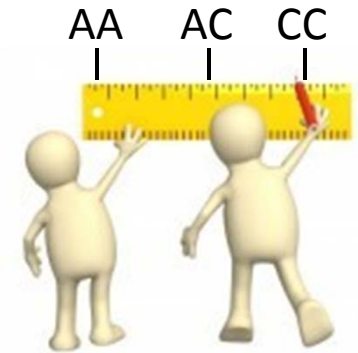
$$\alpha = a + d(q-p)$$

Mean IL-6R concentration of each genotype:

CC: 5.698 / CA: 4.418 / AA: 3.238 (10^{-8} g/mL)

Total Variance of IL-6R concentration=1.35

Frequencies: C, frequency: $p=0.39$ / A, frequency: $q=0.61$



QUESTIONS (Falconer & MacKay; 1996: Introduction to quantitative genetics)

1. Calculate genotypic values (a and d) (page 109)
2. Calculate the average effect of the alleles (page 113)
3. Calculate the genotype frequencies (page 7)
4. Calculate the mean IL6-R concentration in the population (page 110)
5. Calculate how much of the variance is explained by this SNP
(*Variance = Sum of squared deviations from the mean*)
6. Calculate heritability

Key concepts

- α_i = average effect of allele i
 - Property of a single allele in a particular population (depends on genetic background of the population)
- A = Additive Genetic Value (A)
 - A = sum (over all loci) of breeding values
 - Fraction of G that parents pass along to their offspring
 - Property of an Individual in a particular population
- $\text{Var}(A)$ = additive genetic variance
 - Variance in additive genetic values
 - Property of a population

Average effect (Falconer p112)

The concept of average effect is not easy to grasp.

It has to do with the transmission of value from parent to offspring: parents transmit genes (alleles) and not genotypes; thus a measure of value is needed that refers to genes and not to genotypes.

Average effect (Falconer p112)

= Value that refers to genes and not to genotypes.

Average effects depend on genotypic values a and d and also on the gene frequencies (and are therefore properties of populations as well as of the genes concerned).

Average effect

The average effect of a gene (allele) is the mean deviation from the population mean of individuals which received that gene from one parent, the gene received from the other parent having come *at random* from the population.

Let a number of gametes all carrying A1 unite at random with gametes from the population; then the mean of the genotypes produced deviates from the population mean by an amount which is the average effect of A1.

Average effect is related to genotypic values a and d

$$q [a + d (q - p)] = \alpha_1$$

$$-p [a + d (q - p)] = \alpha_2$$

Average effect of gene substitution is $\alpha_1 - \alpha_2 = \alpha$

This is the difference between the average effect of the 2 alleles:

$$\alpha = a + d(q-p)$$

Average effect is related to genotypic values a and d

$$\alpha = \alpha_1 - \alpha_2 = q [a + d (q - p)] - \{-p [a + d (q - p)]\} =$$

$$qa + qd(q-p) - \{-pa - pd(q-p)\} =$$

$$qa + qd(q-p) + pa + pd(q-p) =$$

$$a(q+p) + d(q-p)(p+q) = a + d(q-p)$$

Alternative approach to obtaining alpha
(based mean value of genotypes produced)

$$(pa + qd) - (-qa + pd) =$$

$$pa + qa + qd - pd =$$

$$a(p + q) + d(q - p) =$$

$$a + d(q - p)$$

Breeding value

Value associated with the genes carried by an individual and transmitted to its offspring (the value of an individual judged by the mean value of its progeny (over multiple matings)).

The breeding value is sometimes referred to as the 'additive genotype' (A).

Breeding value

The average effects of the parents' genes determine the mean genotypic value of its progeny (NB breeding value is thus population dependent).

Genotype	Breeding value (=A)
A1A1	$2 \alpha_1 = 2q\alpha$
A1A2	$\alpha_1 + \alpha_2 = (q-p)\alpha$
A2A2	$2 \alpha_2 = -2p\alpha$

Genotype (i)	AA	Aa	aa
Frequency (f)	p^2	$2pq$	q^2
A	$2q\alpha$	$(q-p)\alpha$	$-2p\alpha$
D	$-2q^2d$	$2pqd$	$-2p^2d$

See: table 7.3

Variation ?

Due to A: V_a (= variance of breeding values)

Due to D: V_d (=variance of dominance deviations)

Genotype (i)	AA	Aa	aa
Frequency (f)	p^2	$2pq$	q^2
A	$2q\alpha$	$(q-p)\alpha$	$-2p\alpha$
D	$-2q^2d$	$2pqd$	$-2p^2d$

$$V_a = \sum f_i A_i^2 \quad \text{and} \quad V_d = \sum f_i D_i^2$$

Derive as function of frequency and α and d

Genotype (i)	AA	Aa	aa
Frequency (f)	p^2	$2pq$	q^2
A	$2q\alpha$	$(q-p)\alpha$	$-2p\alpha$
D	$-2q^2d$	$2pqd$	$-2p^2d$

$$V_a = \sum f_i A_i^2 =$$

$$p^2(2q\alpha)^2 + 2pq [(q-p)\alpha]^2 + q^2 (-2p\alpha)^2 =$$

$$4p^2q^2\alpha^2 + 2pq\alpha^2 [q^2 - 2pq + p^2] + 4p^2q^2\alpha^2 =$$

$$2pq\alpha^2 (2pq + q^2 - 2pq + p^2 + 2pq) =$$

$$2pq\alpha^2$$

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

Thus, dominance is part of V_a (unless $p = q$)

Genotype (i)	AA	Aa	aa
Frequency (f)	p^2	$2pq$	q^2
A	$2q\alpha$	$(q-p)\alpha$	$-2p\alpha$
D	$-2q^2d$	$2pqd$	$-2p^2d$

$$V_d = \sum f_i D_i^2 =$$

$$p^2(-2q^2d)^2 + 2pq(2qpd)^2 + q^2(-2p^2d)^2 =$$

$$4p^2q^4d^2 + 8p^3q^3d^2 + 4p^4q^2d^2 =$$

$$4p^2q^2d^2 (p^2 + 2pq + q^2) =$$

$$4p^2q^2d^2 =$$

$$(2pqd)^2$$

Genotype (i)	AA	Aa	aa
Frequency (f)	p^2	$2pq$	q^2
Genotypic effect (x)	a	d	-a
$f * x$	$p^2 a$	$2pqd$	$- q^2 a$

mean: $p^2 a + 2pqd - q^2 a = a(p-q) + 2pqd$

Variation: $2pq[a+d(q-p)]^2 + (2pqd)^2$

Population variation depends on 'a' (difference between homozygote individuals), 'd' (deviation of heterozygote persons from zero) and on allele frequency (p & q).