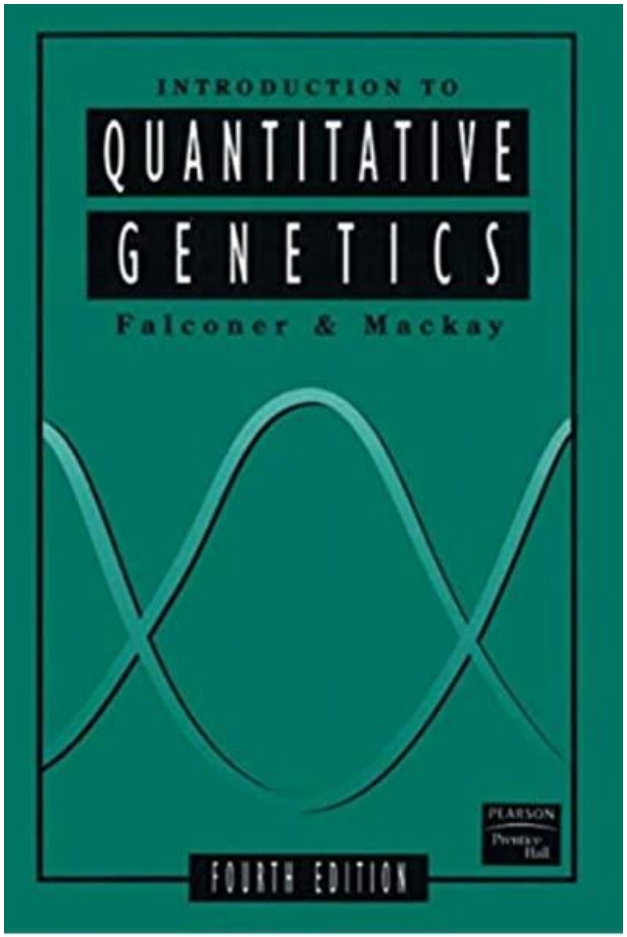
Introduction to quantitative genetics

KU LEUVEN

Further reading??



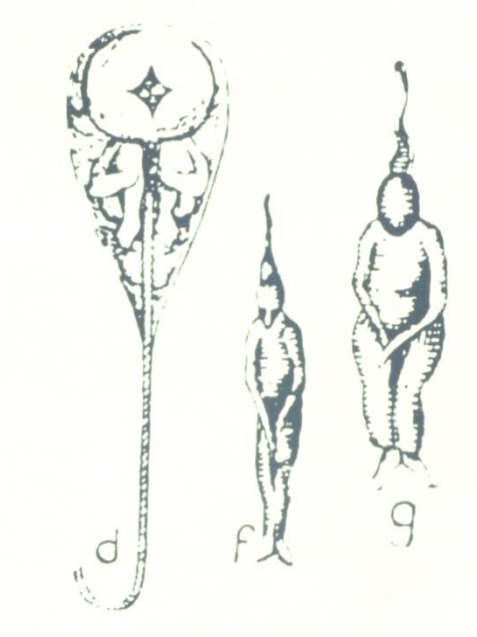
Why do children look like their parents?

KU LEUVEN

Theories related to inheritance

Preformation-theory

Spermists



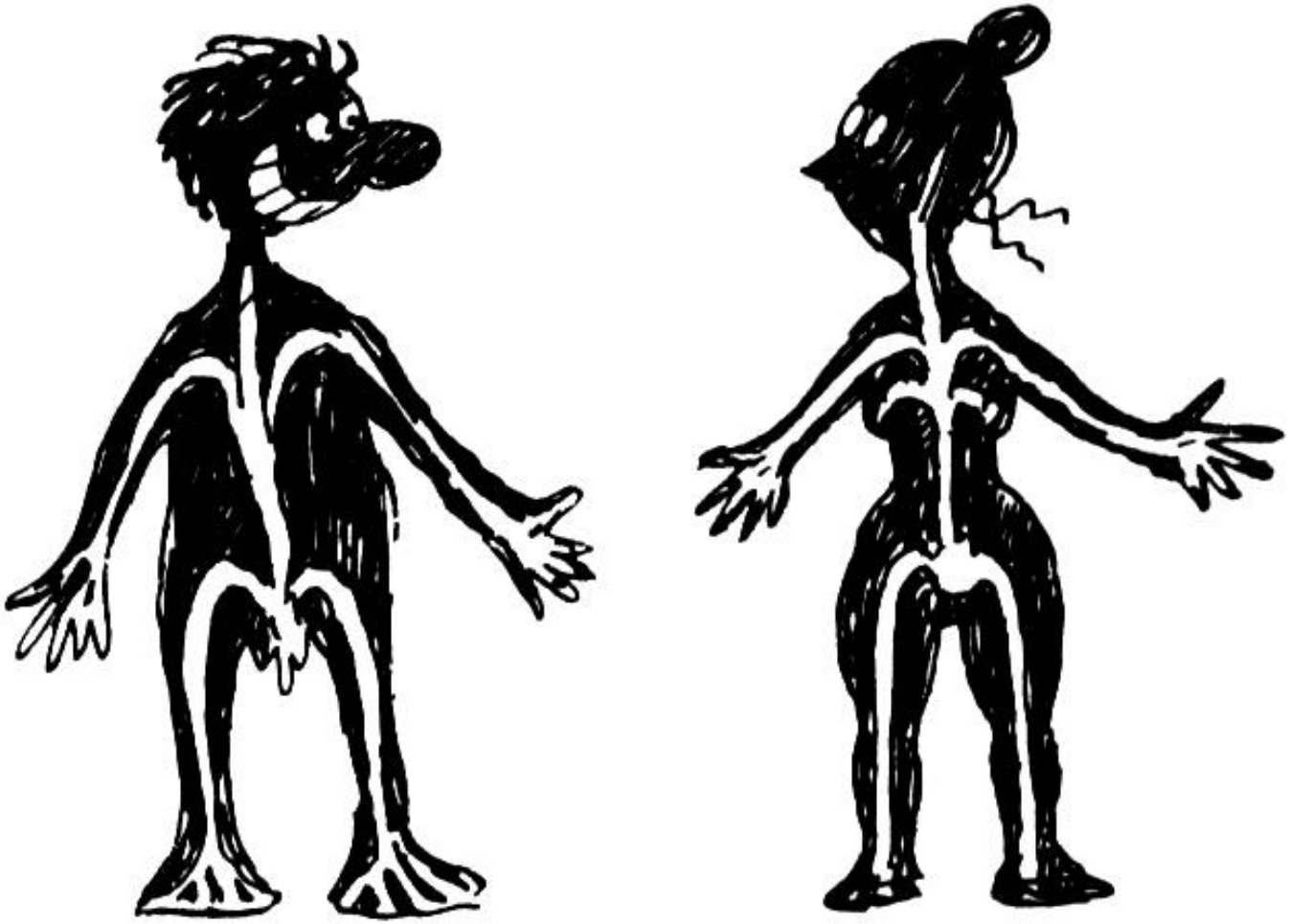
Van Leeuwenhoek
17th century

Ovists

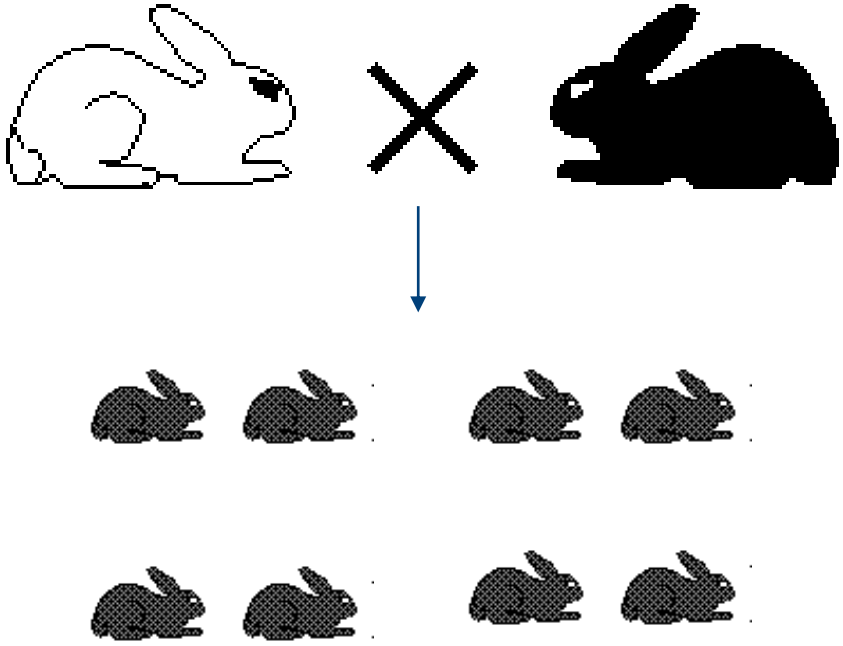
de Graaf
17th century

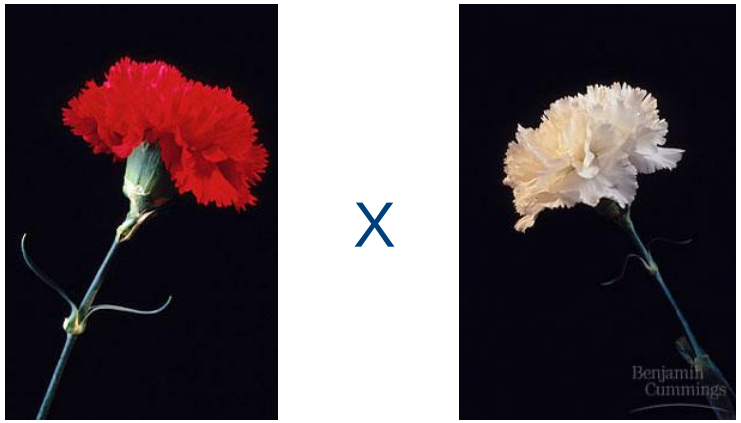
KU LEUVEN

19th century : ”pangenesis”



KU LEUVEN



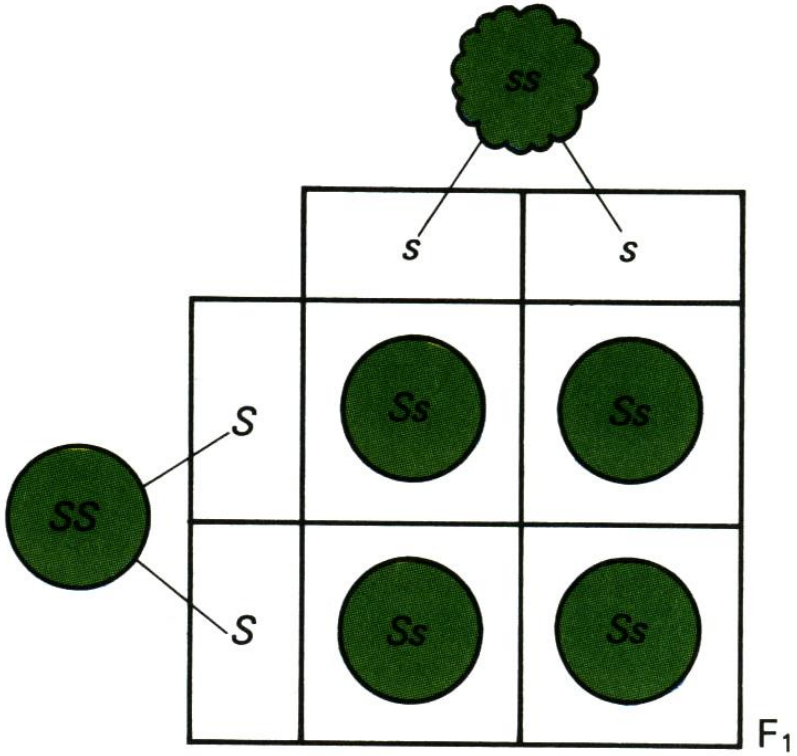

X
Benjamin
Cummings

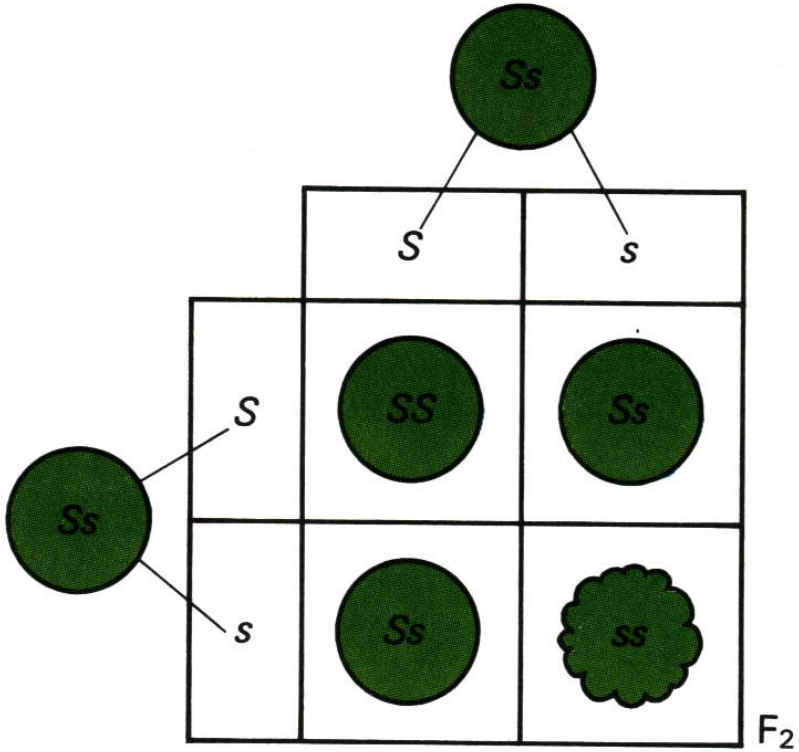
End 19th century – beginning 20th century

Gregor Mendel

KU LEUVEN

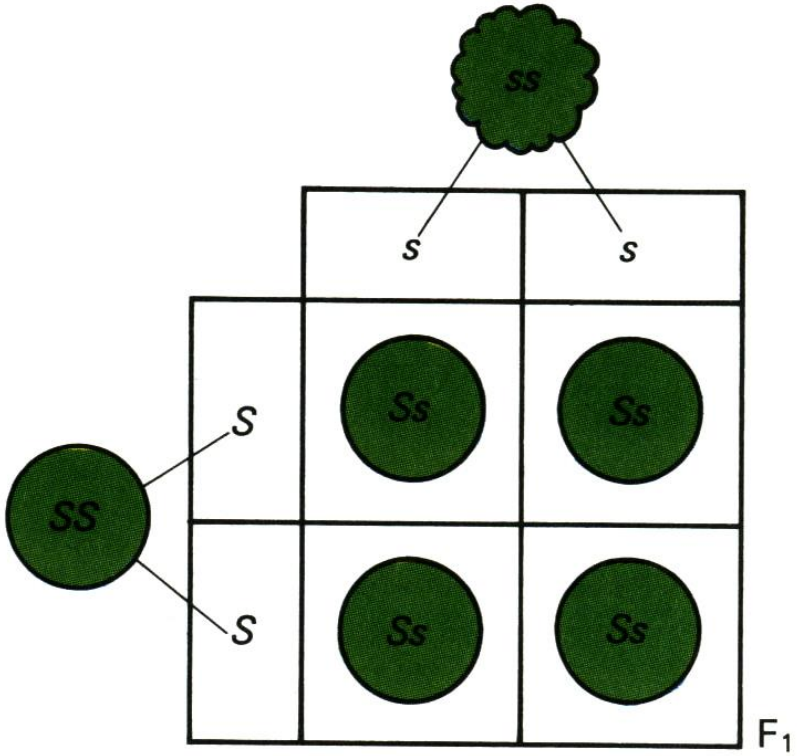
Mendel: no mixing

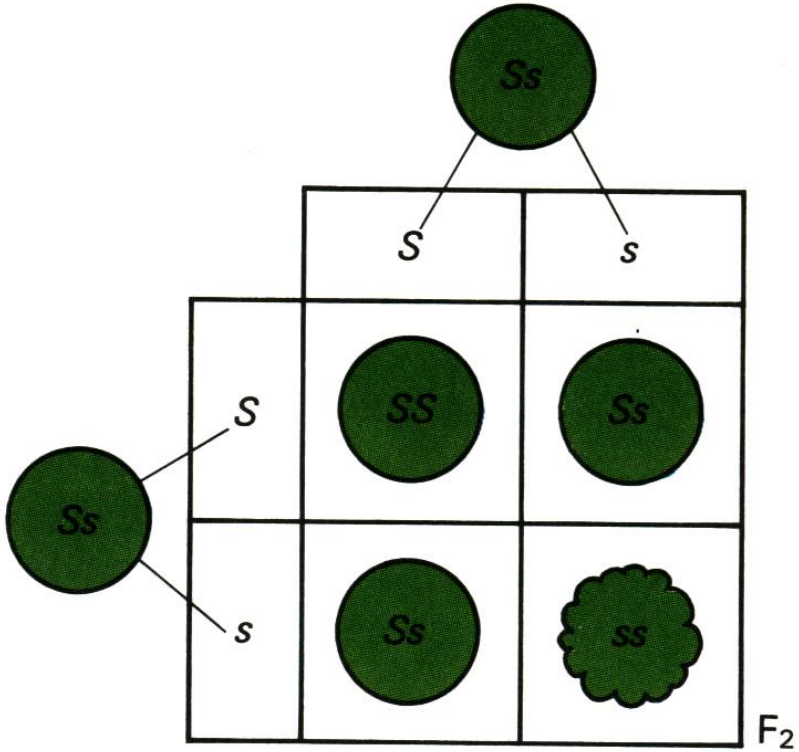




Mendel: no mixing

- S = round = ‘the boss’ = dominant
- s = wrinckled = recessive

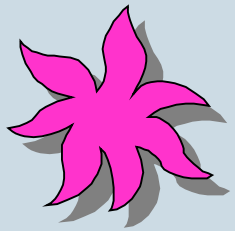
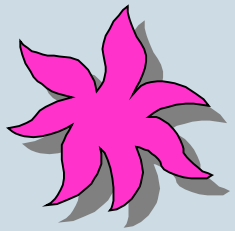
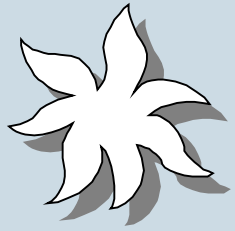
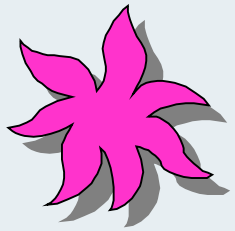
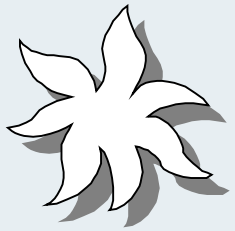
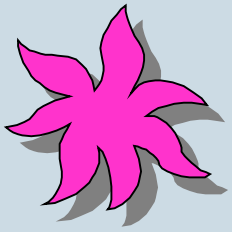
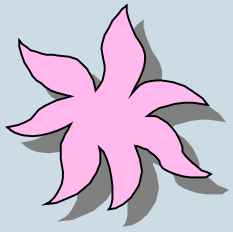
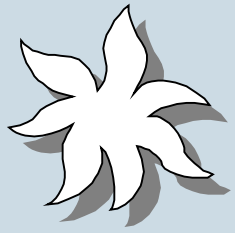




Separate, inheritable units (‘genes’)

Inheritance models in single gene traits

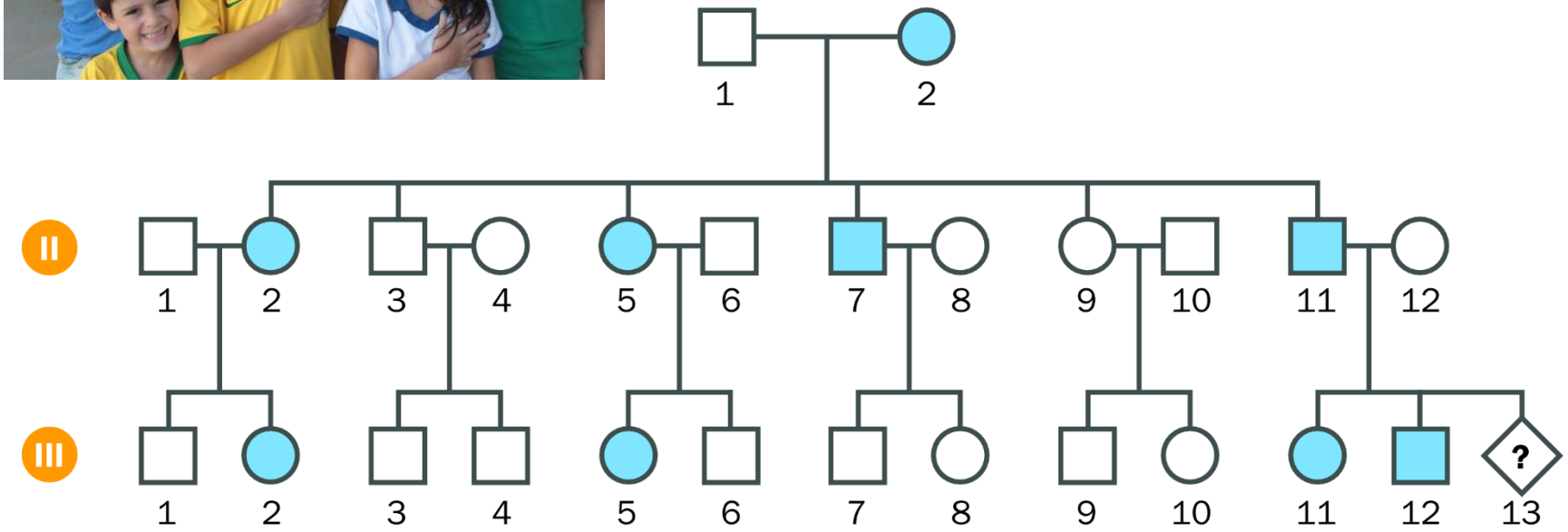
Genotype group

Model	AA	Aa	aa
A is dominant			
A is Recessive			
A is co-dominant			

KU LEUVEN

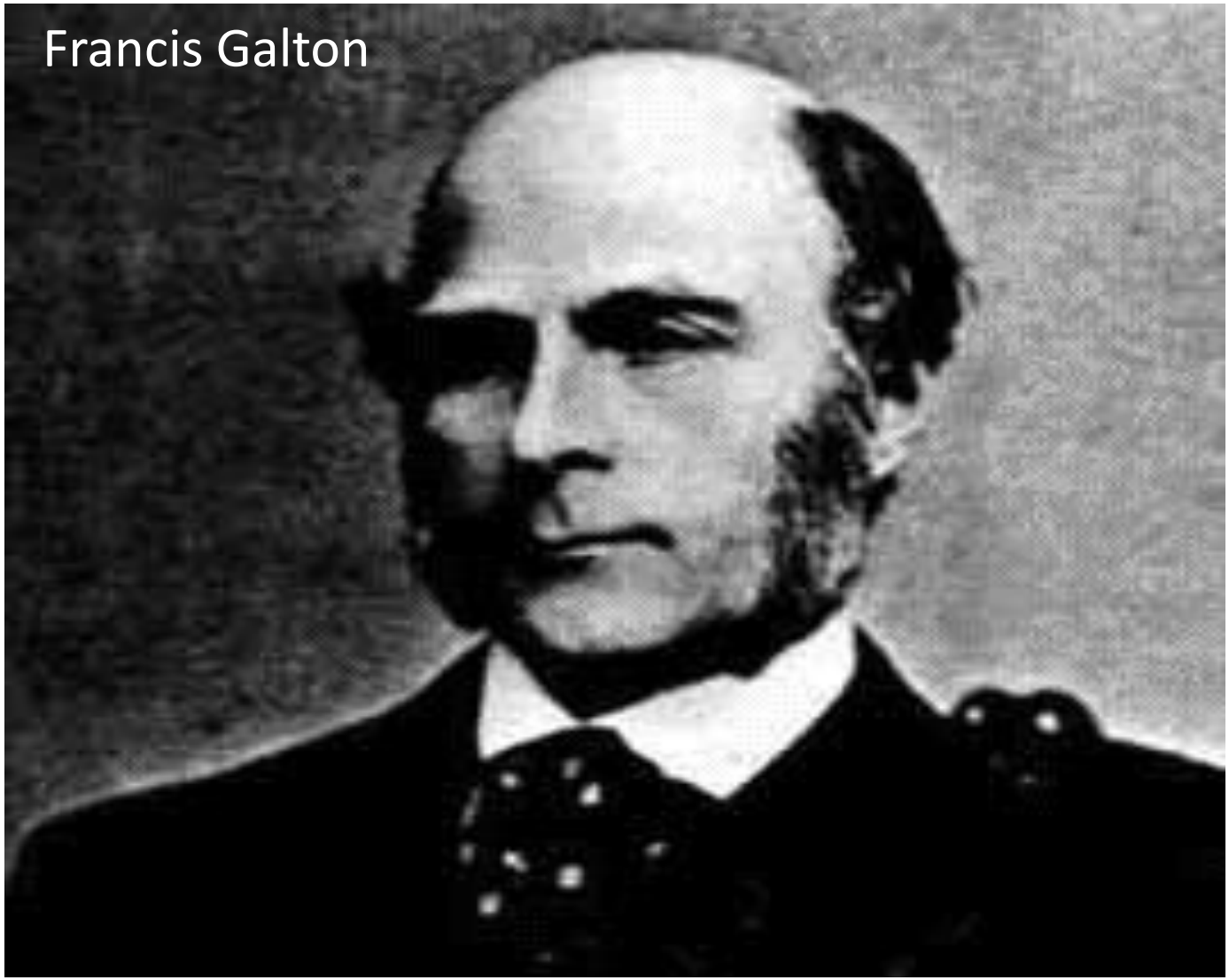
Adopted from: Manolio Teri

Mendelian genetics



Lower figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.3)

End 19th century – beginning 20th century



In the early 20th century there was controversy between proponents of Mendelian and quantitative models of inheritance

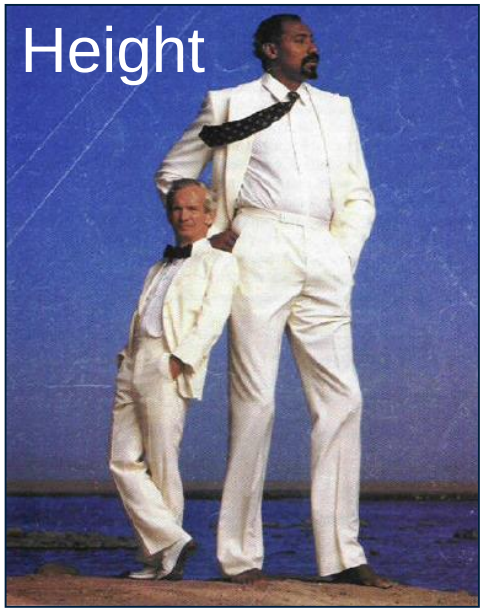
Francis Galton investigated family resemblances. He quantified observations (height, weight, intelligence, breathing capacity…) and applied statistical analyses to explain patterns of inheritance.

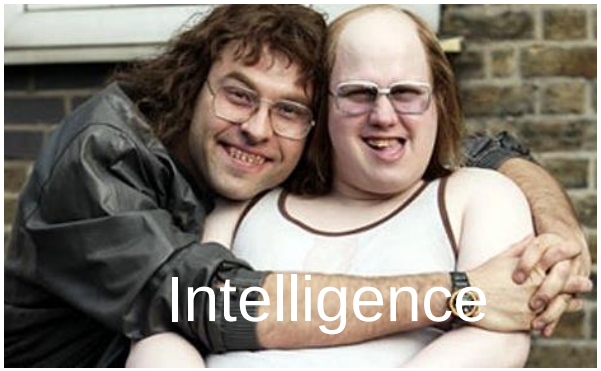
> **Biometry**: the statistical study of quantitative traits
>
> In one of the first applications of statistics, he compared physical attributes of parents and children, and established the degree of correlation between relatives.

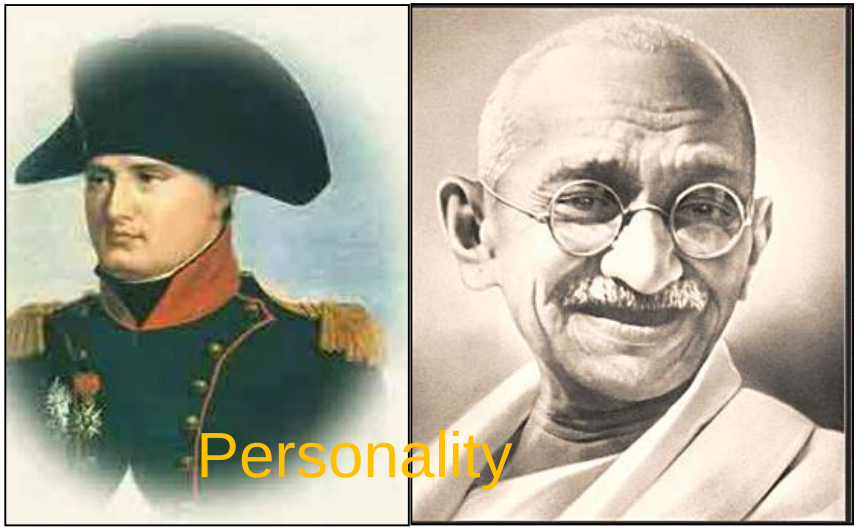
Biometricians accepted that a few rare abnormalities or curious traits might be inherited as Mendel described, but they pointed out that most of the characters/traits likely to be important in evolution (strength, fertility, body size…) were quantitative in nature, and thus not amenable to Mendelian analysis:

> We all have these characters, only to different degrees, so you cannot define their inheritance by drawing pedigrees and marking in the people who have them. Mendelian analysis requires dichotomous characters that you either have or do not have.

VARIATION IN NORMAL TRAITS







KU LEUVEN

Examples of quantitative traits

- Humans: blood pressure, cholesterol levels, height, weight, skin colour...
- Cattle: milk production, fat content of milk...
- Sheep: wool production, meat production, growth...
- Pigs: meat production, growth, piglets per litter...
- Laying hens: egg production
- Broiler chicken: growth, feed conversion ratio, breast muscle...
- Corn: yield
- ...

The **polygenic theory** explains how quantitative traits can be genetically determined

RA Fisher (1918)

Characters that are determined by a large number of independent Mendelian factors could display precisely the continuous nature, quantitative variation, and family correlations described by the biometricians.

DS Falconer

Extended this model to cover dichotomous non-Mendelian characters.

The **polygenic theory** explains how quantitative traits can be genetically determined

RA Fisher (1918)

Characters that are determined by a large number of independent Mendelian factors could display precisely the continuous nature, quantitative variation, and family correlations described by the biometricians.

DS Falconer

Extended this model to cover dichotomous non-Mendelian characters.

- Existence of monogenic vs multifactorial inheritance
- How does the polygenic theory work?

The **polygenic theory** explains how quantitative traits can be genetically determined

RA Fisher (1918)

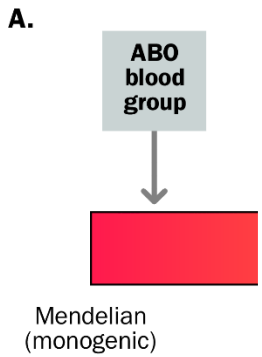
Characters that are determined by a large number of independent Mendelian factors could display precisely the continuous nature, quantitative variation, and family correlations described by the biometricians.

DS Falconer

Extended this model to cover dichotomous non-Mendelian characters.

- **Existence of monogenic vs multifactorial inheritance**
- How does the polygenic theory work?

The genetic spectrum

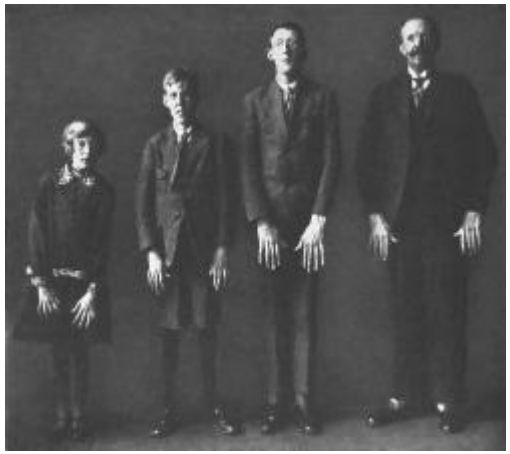


Trait / Disease

Gene1

! Sufficient and necessary "causal"

Mendelian characters are recognized by the characteristic pedigree patters they give



OMIM-database
(Online Mendelian Inheritance in Man)

Figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.1)

KU LEUVEN

The genetic spectrum

Most human genetic characters are not Mendelian

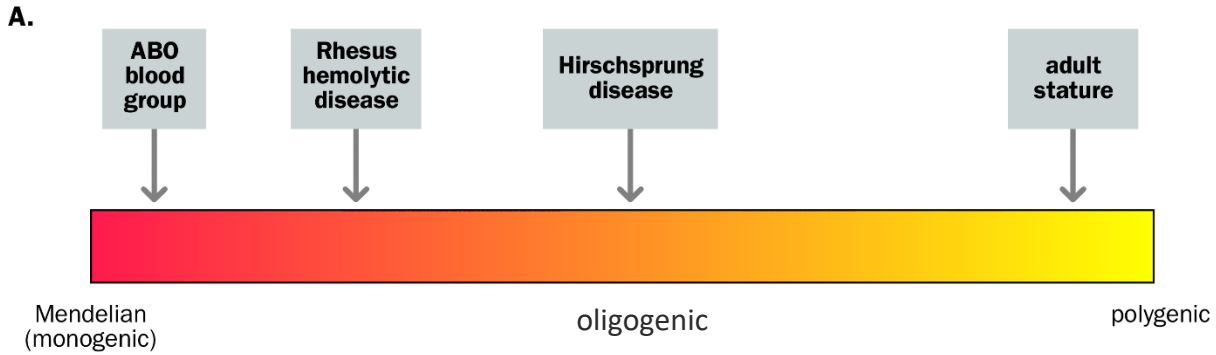


Figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.1)

KU LEUVEN

The genetic spectrum

Most human genetic characters are not Mendelian

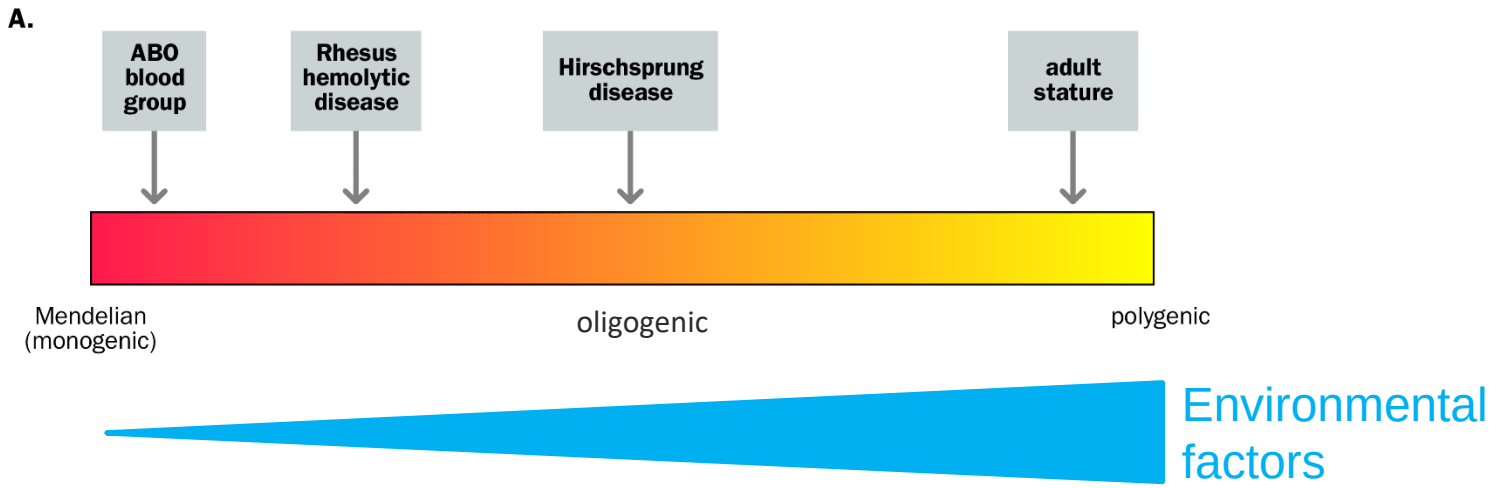
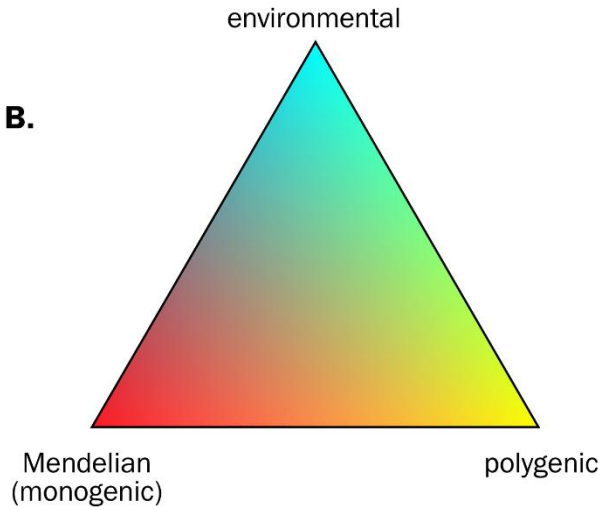


Figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.1)

The genetic spectrum



- Characters form a continuous spectrum, from perfectly Mendelian to truly polygenic.
- Superimposed on this there may be a greater or smaller effect of environmental factors.
- The overall etiology of any character can be represented as a point somewhere within the depicted triangle.

Figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.1)

The **polygenic theory** explains how quantitative traits can be genetically determined

RA Fisher (1918)

Characters that are determined by a large number of independent Mendelian factors could display precisely the continuous nature, quantitative variation, and family correlations described by the biometricians.

DS Falconer

Extended this model to cover dichotomous non-Mendelian characters.

- Existence of monogenic vs multifactorial inheritance
- **How does the polygenic theory work?**

The **polygenic theory** explains how quantitative traits can be genetically determined

RA Fisher

Characters that are determined by a large number of independent Mendelian factors could display precisely the continuous nature, quantitative variation, and family correlations described by the biometricians.

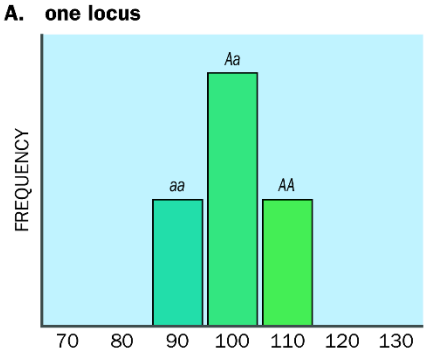
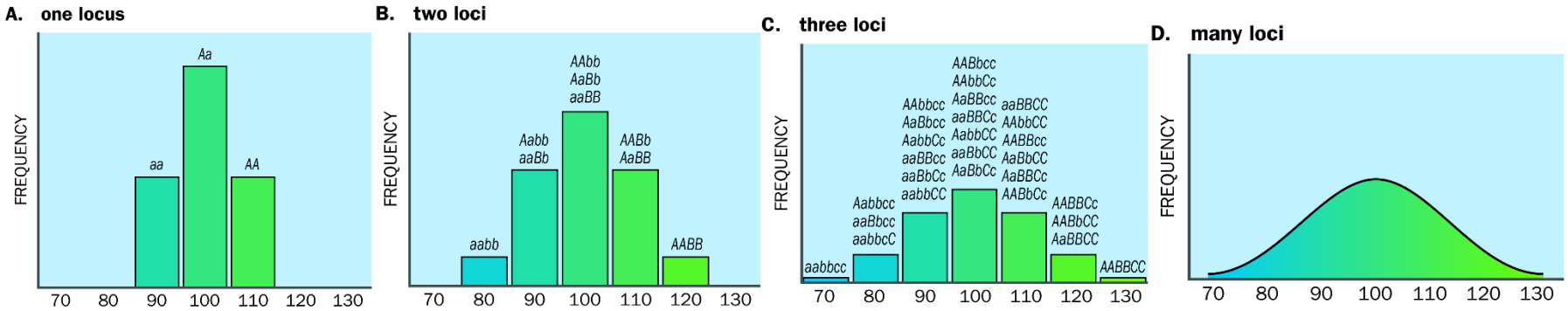


Figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.20)

The **polygenic theory** explains how quantitative traits can be genetically determined

RA Fisher

Characters that are determined by a large number of independent Mendelian factors could display precisely the continuous nature, quantitative variation, and family correlations described by the biometricians.



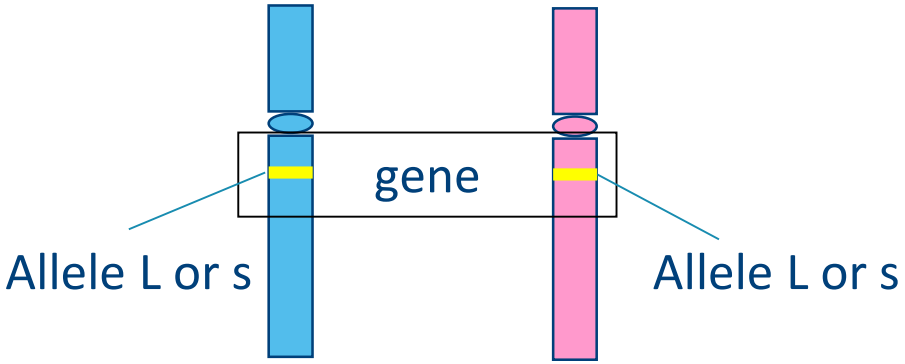
- As the number of loci increases, the distribution looks increasingly like a Gaussian curve. Adding little environmental variation would smooth out the three-locus distribution into a good Gaussian curve.

Figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.20)

KU LEUVEN

Example : height

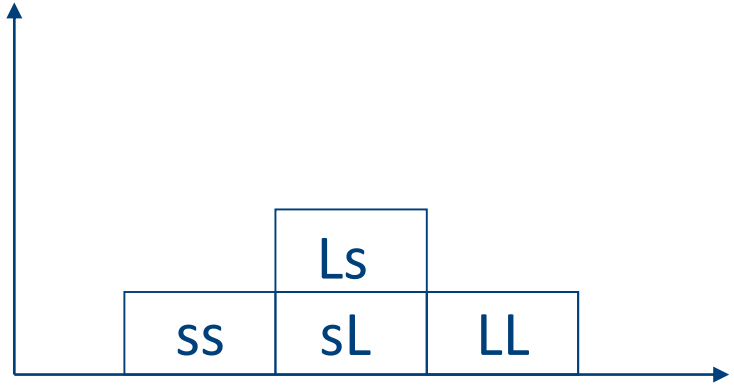
- One gene
  - allele L: large stature
  - allele s: short stature



inherited from father

inherited from mother

	L	s
L	LL	Ls
s	sL	ss



Short Medium Large

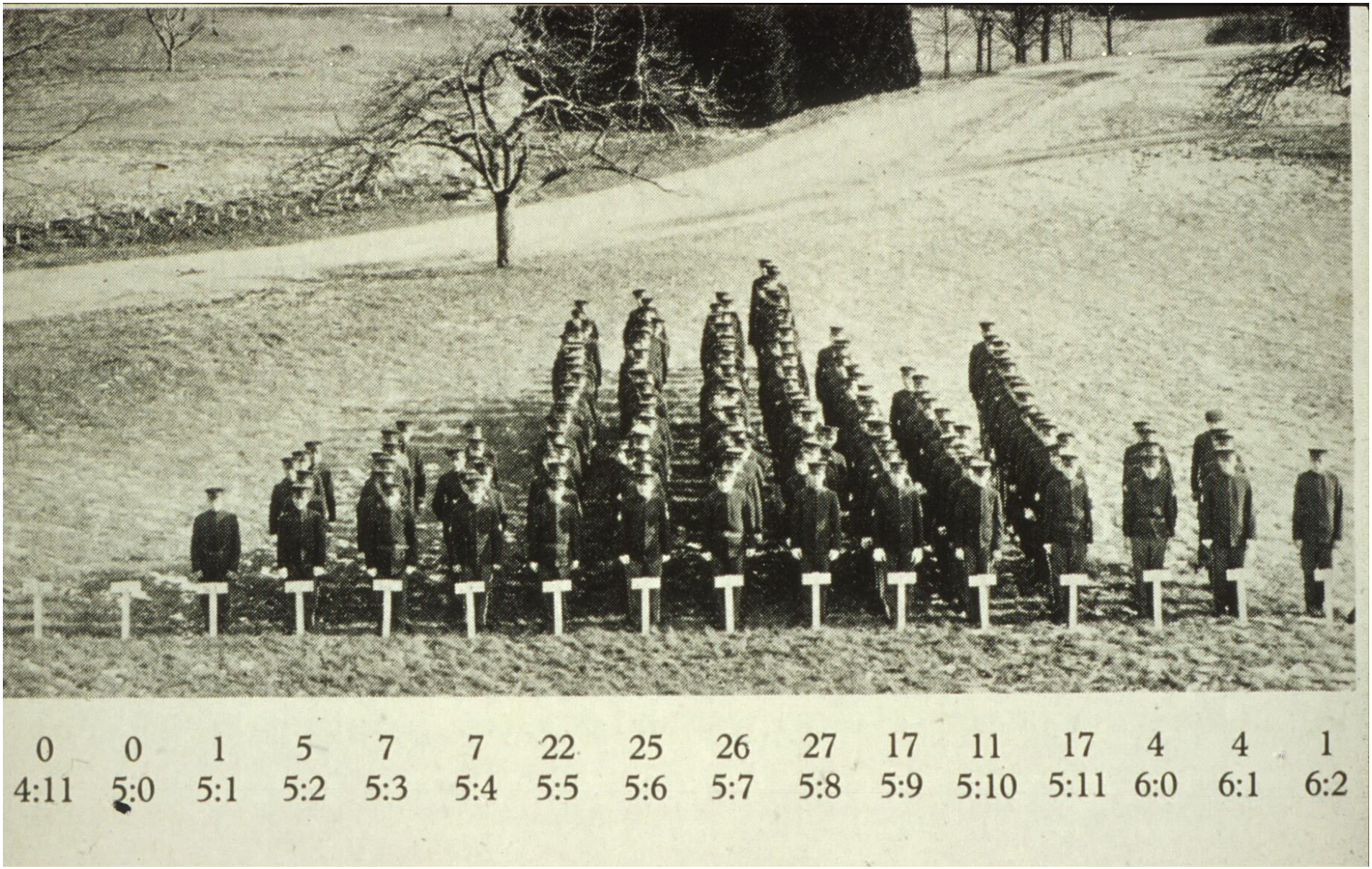
Example : height – two genes

- 2 genes - each 2 alleles, equally frequent
  - Gene 1: allele A, allele a
  - Gene 2: allele B, allele b
- Additive effect on growth

	AB	**Ab**	**aB**	**ab**
AB	AABB	AABb	AaBB	AaBb
Ab	AAbB	AAbb	AabB	Aabb
aB	aABB	aABb	aaBB	aaBb
ab	aAbB	aAbb	aabB	aabb

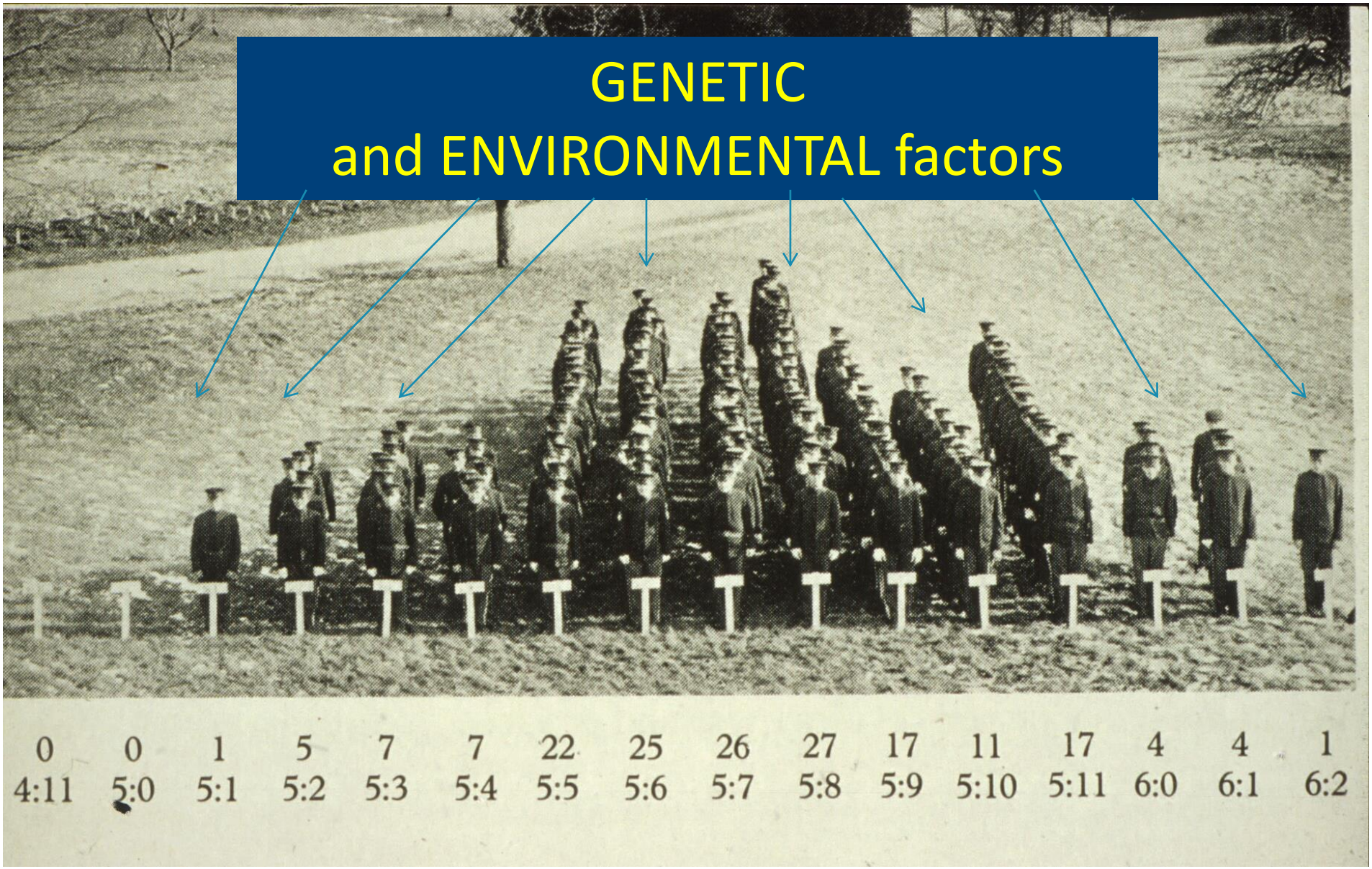
Etcetera...

KU LEUVEN



0	0	1	5	7	7	22	25	26	27	17	11	17	4	4	1
4:11	5:0	5:1	5:2	5:3	5:4	5:5	5:6	5:7	5:8	5:9	5:10	5:11	6:0	6:1	6:2

KU LEUVEN

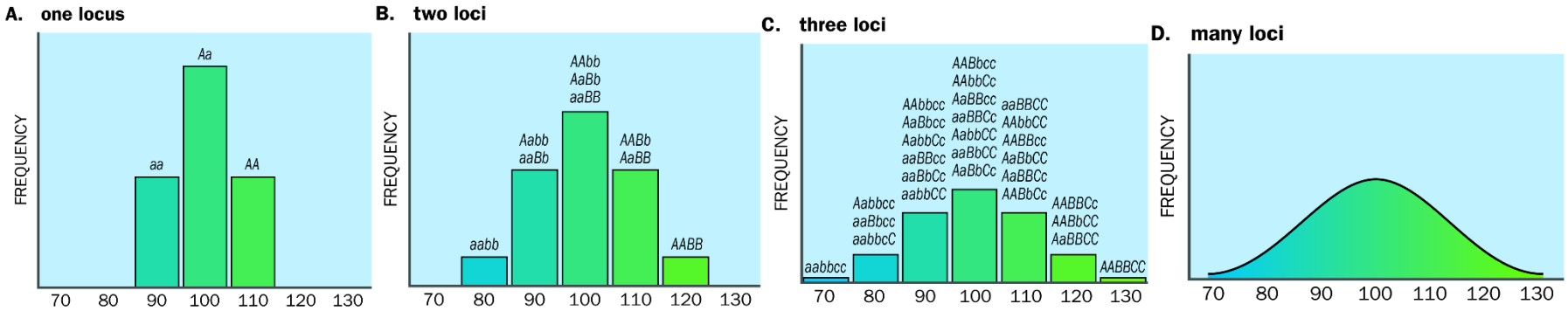

GENETIC
and ENVIRONMENTAL factors
0 0 1 5 7 7 22 25 26 27 17 11 17 4 4 1
4:11 5:0 5:1 5:2 5:3 5:4 5:5 5:6 5:7 5:8 5:9 5:10 5:11 6:0 6:1 6:2

KU LEUVEN

The **polygenic theory** explains how quantitative traits can be genetically determined

RA Fisher

Characters that are determined by a large number of independent Mendelian factors could display precisely the continuous nature, quantitative variation, and family correlations described by the biometricians.



- As the number of loci increases, the distribution looks increasingly like a Gaussian curve. Adding little environmental variation would smooth out the three-locus distribution into a good Gaussian curve.
- The simple one-to-one relationship genotype and phenotype disappears. Except for the extreme phenotypes, it is not possible to infer genotype from phenotype.

Figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.20)

From genes to quantitative traits – another example

- Example: 5 loci that equally determine the number of flowers of a plant – 2 alleles at each locus

+ + + + +	- - - - -
+ + + + +	- - - - -

- 3^5 = 243 possible genotypes
- Only 11 phenotypic classes (10,9,8...0) because a large number of genotypes (combination of alleles) will have the same phenotype
  - E.g. there is 1 phenotype with 10 flowers, but 51 possible genotypes with 5+ and 5- alleles

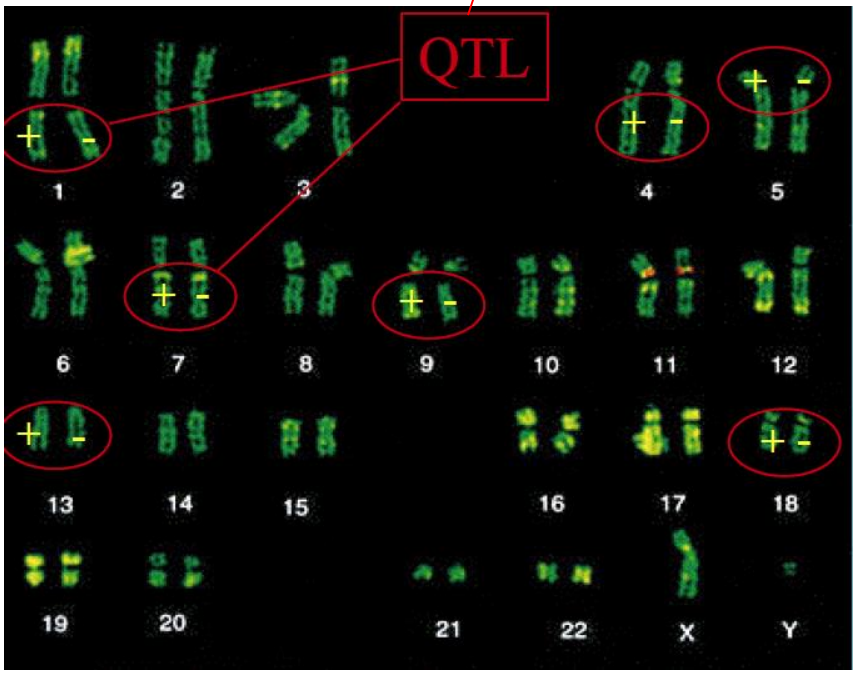
→ Different genotypes give the same phenotype:

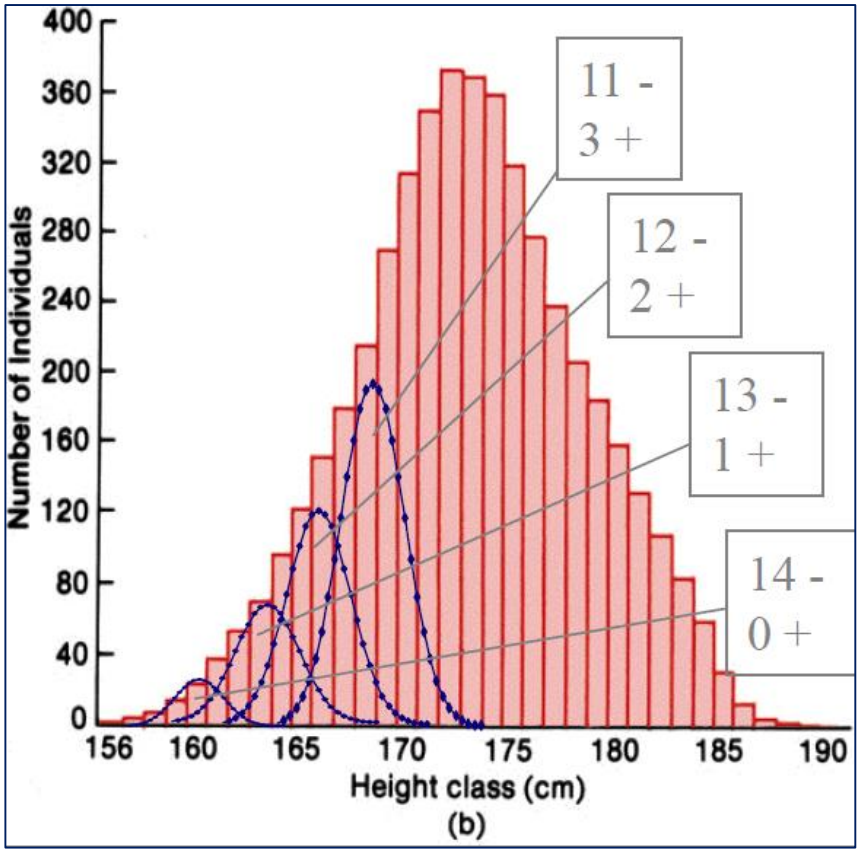
+ + + - +	+ + - - -
+ - - - -	+ + - - +

KU LEUVEN

Example: height

Quantitative trait loci

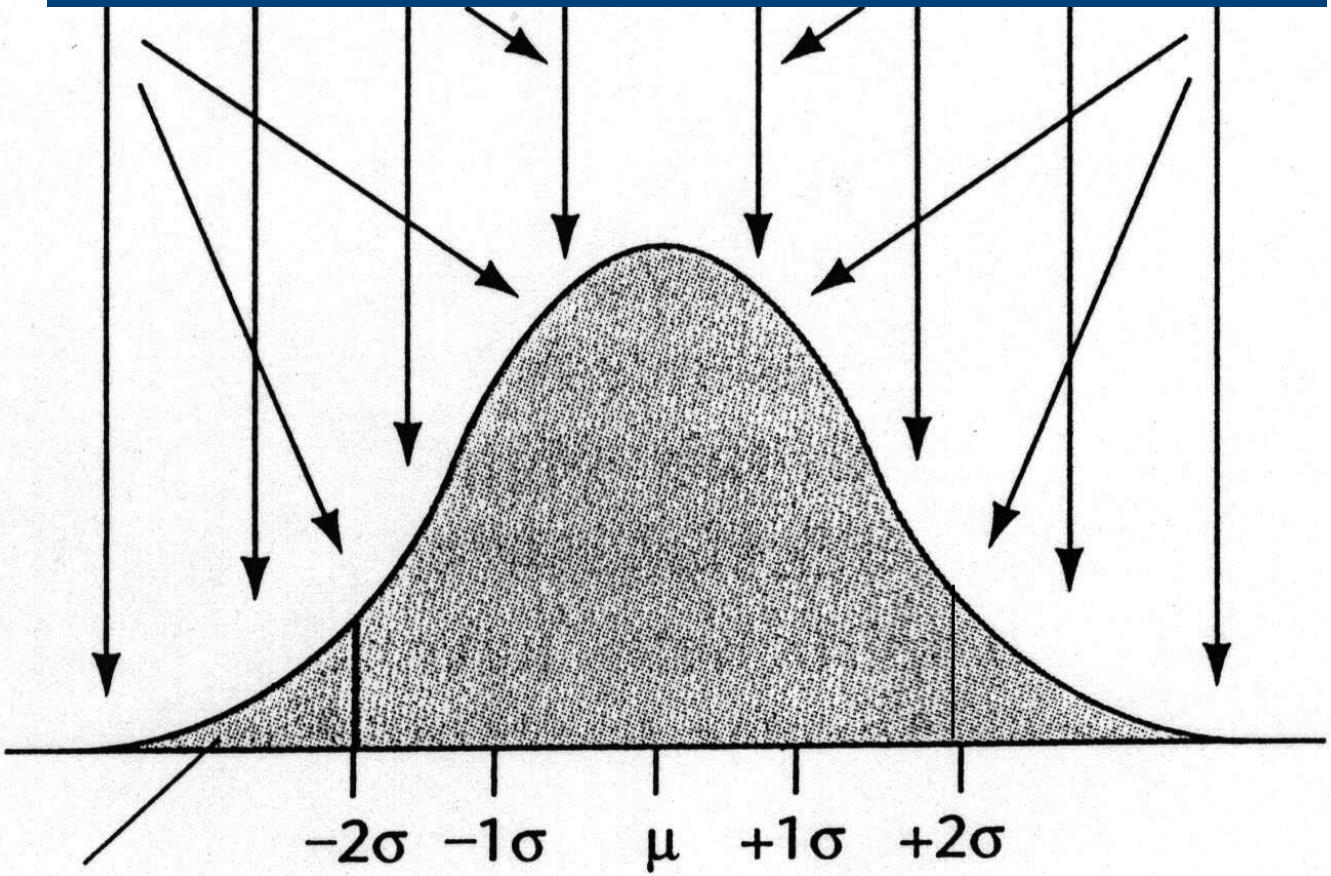




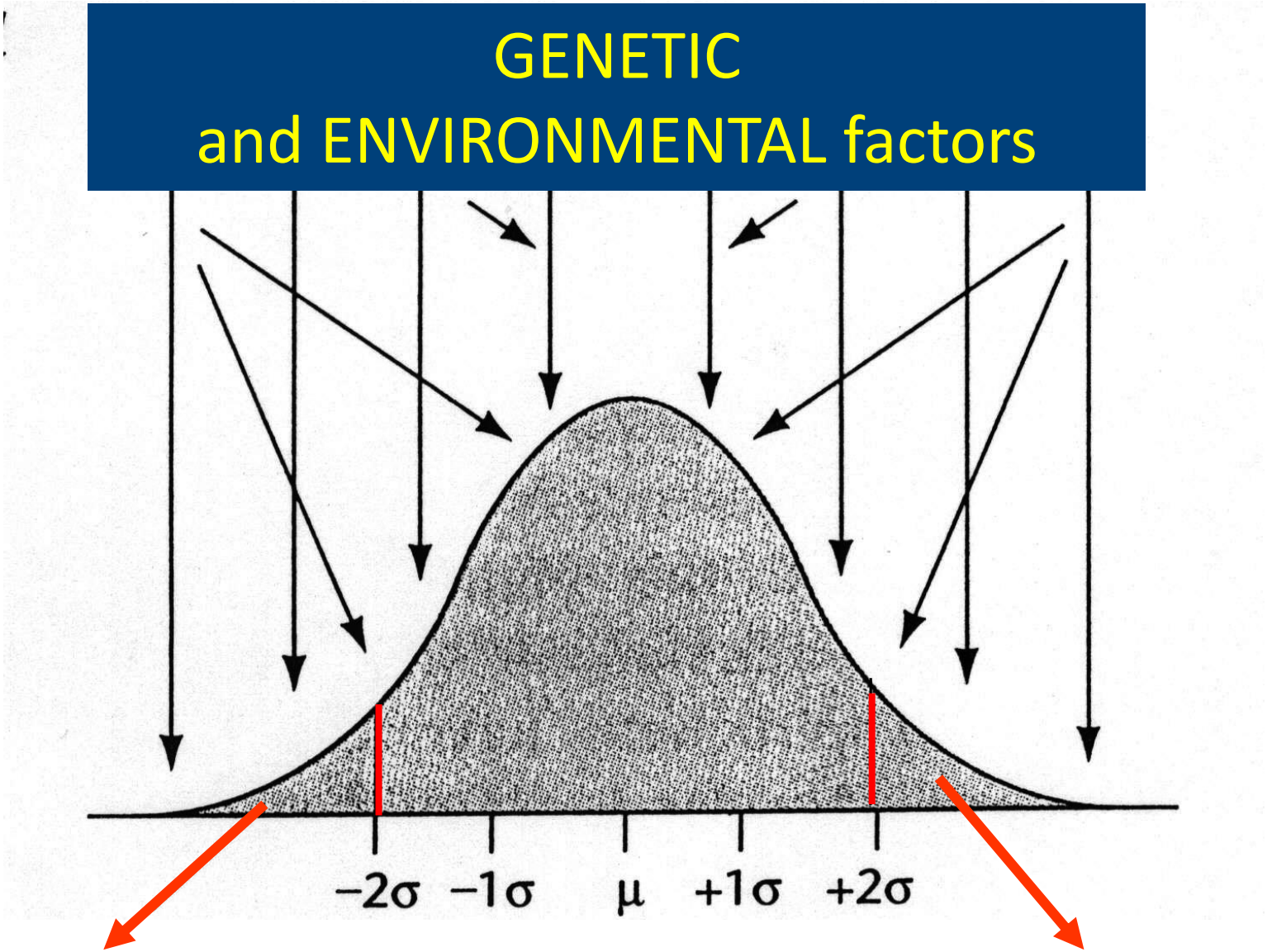
Traits whose variation is determined by
- a number of genes
- a number of environmental factors

KU LEUVEN

GENETIC
and ENVIRONMENTAL factors



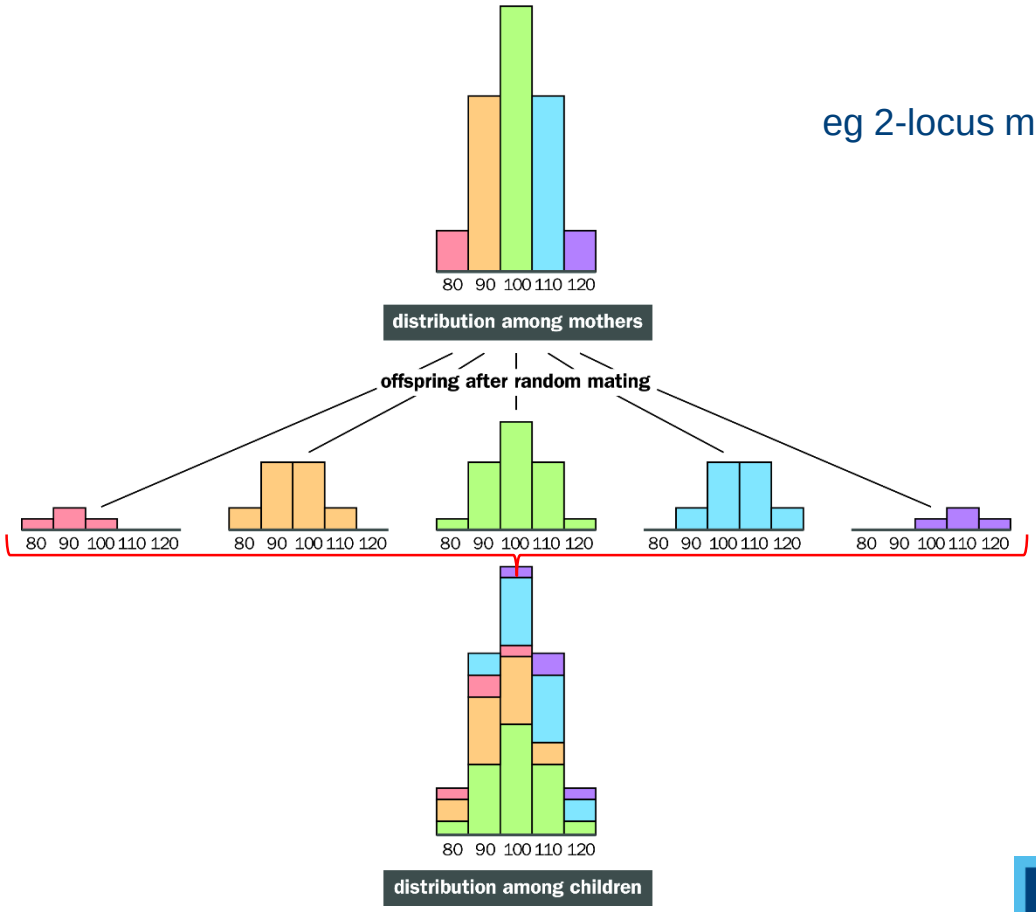
GENETIC
and ENVIRONMENTAL factors



Threshold for ‘abnormal’ (disease state)

Regression to the mean

For each class of mothers, the average IQ of their children is halfway between the mother's value and the population mean

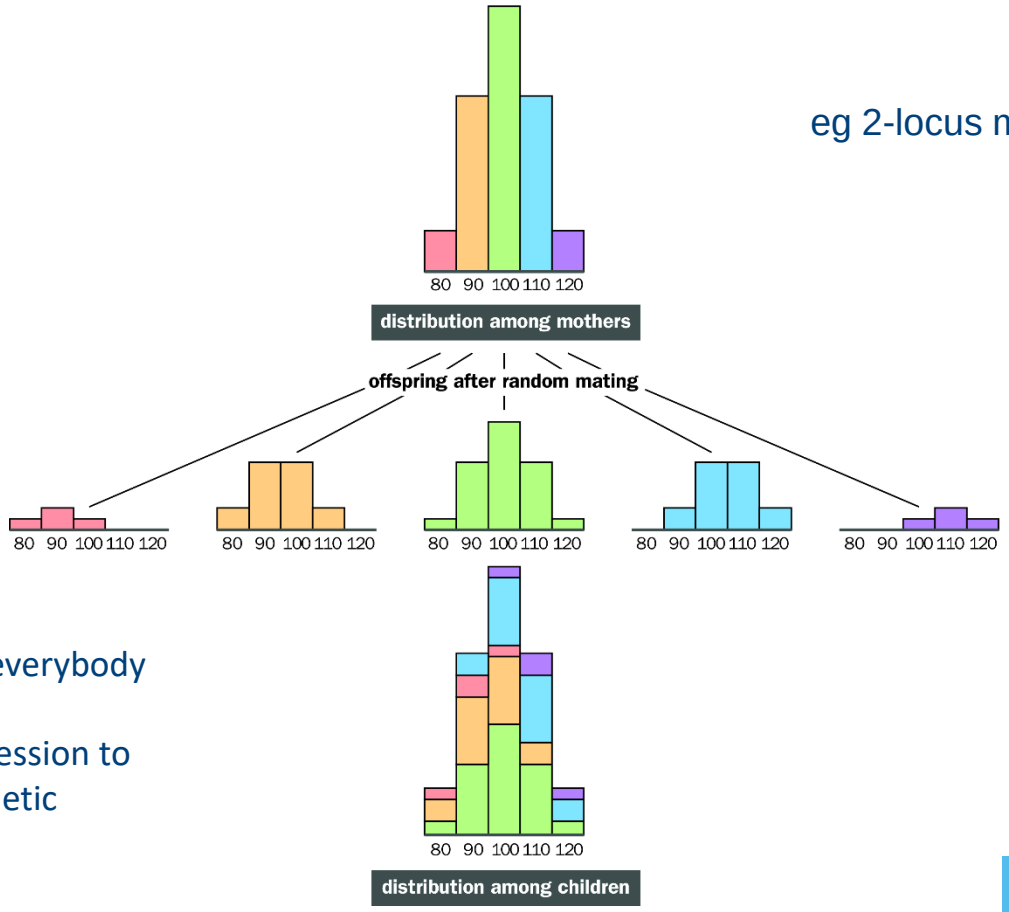


eg 2-locus model

Figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.21)

Regression to the mean

For each class of mothers, the average IQ of their children is halfway between the mother's value and the population mean



eg 2-locus model

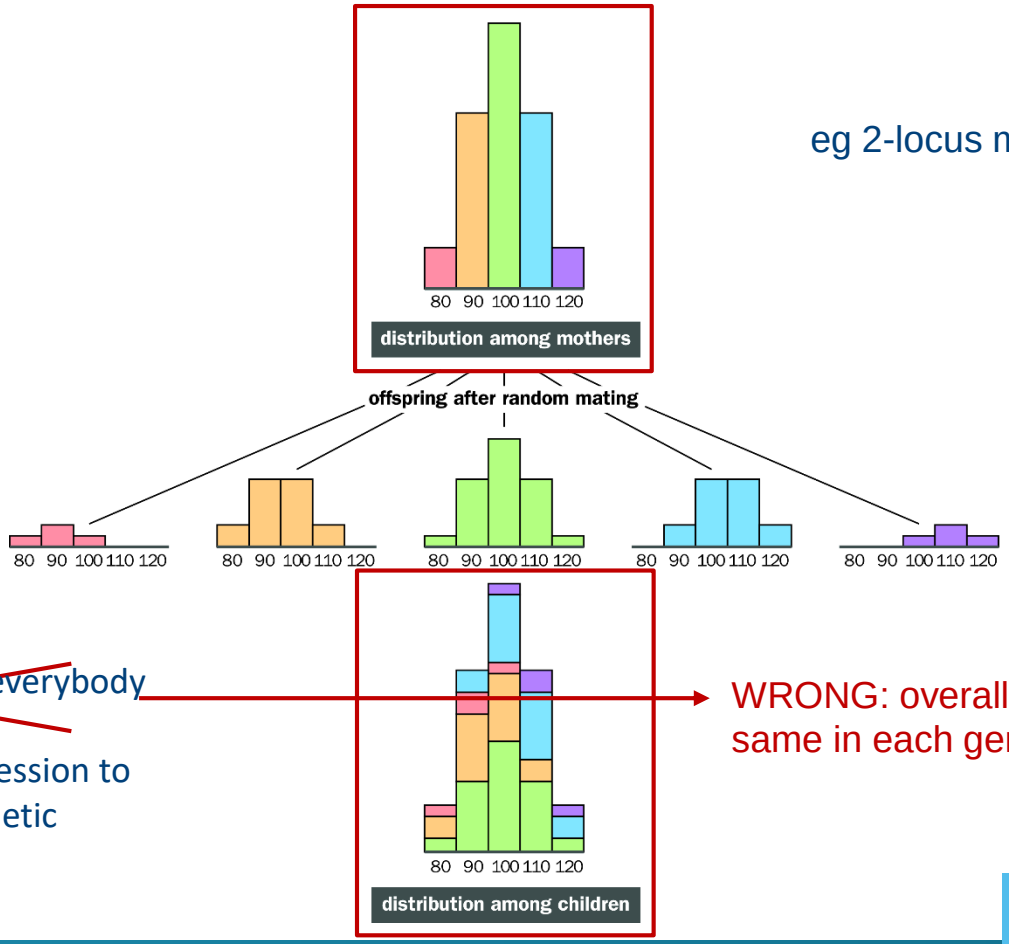
Misconceptions:

- After a few generations everybody will be exactly the same
- If a character shows regression to the mean, it must be genetic

Figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.21)

Regression to the mean

For each class of mothers, the average IQ of their children is halfway between the mother's value and the population mean



eg 2-locus model

Misconceptions:

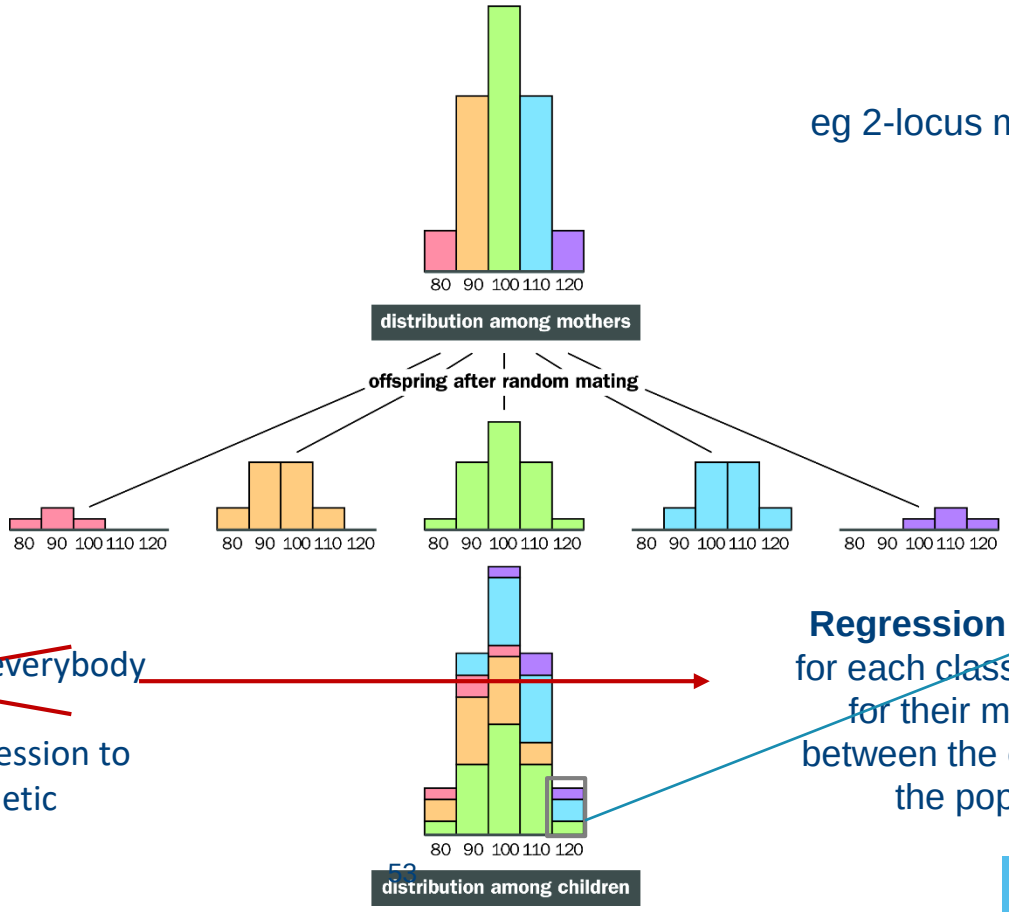
- ~~After a few generations everybody will be exactly the same~~
- If a character shows regression to the mean, it must be genetic

WRONG: overall distribution is the same in each generation

Figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.21)

Regression to the mean

For each class of mothers, the average IQ of their children is halfway between the mother's value and the population mean



eg 2-locus model

Misconceptions:

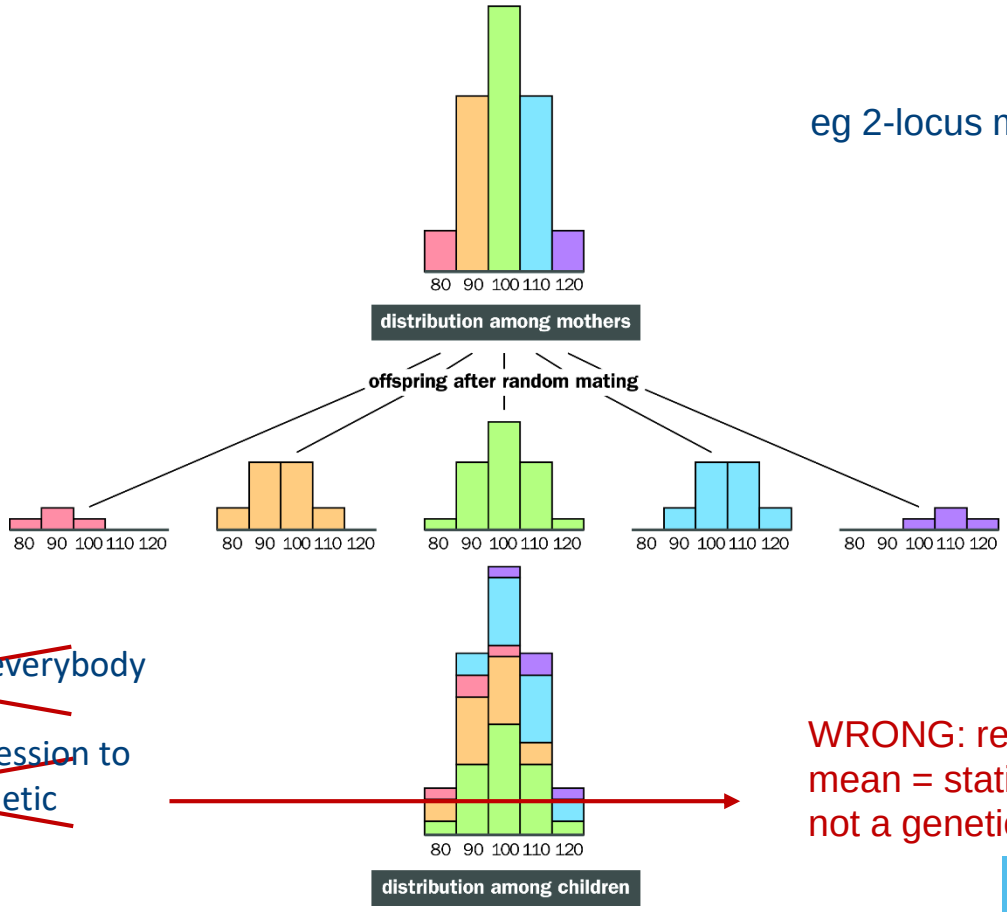
- ~~After a few generations everybody will be exactly the same~~
- If a character shows regression to the mean, it must be genetic

Regression works both ways: for each class of children, the avg for their mothers is halfway between the children's value and the population mean.

Figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.21)

Regression to the mean

For each class of mothers, the average IQ of their children is halfway between the mother's value and the population mean



eg 2-locus model

Misconceptions:

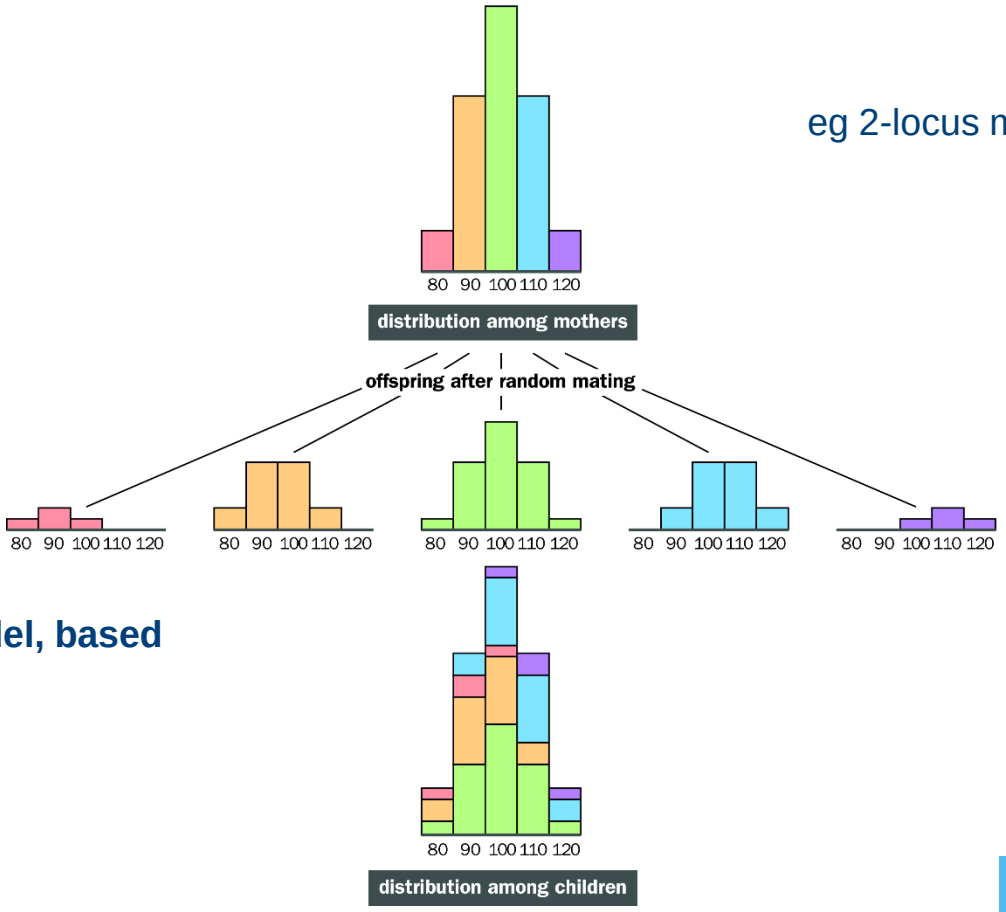
- ~~After a few generations everybody will be exactly the same~~
- ~~If a character shows regression to the mean, it must be genetic~~

WRONG: regression to the mean = statistical phenomenon, not a genetic mechanism.

Figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.21)

Regression to the mean

For each class of mothers, the average IQ of their children is halfway between the mother's value and the population mean



eg 2-locus model

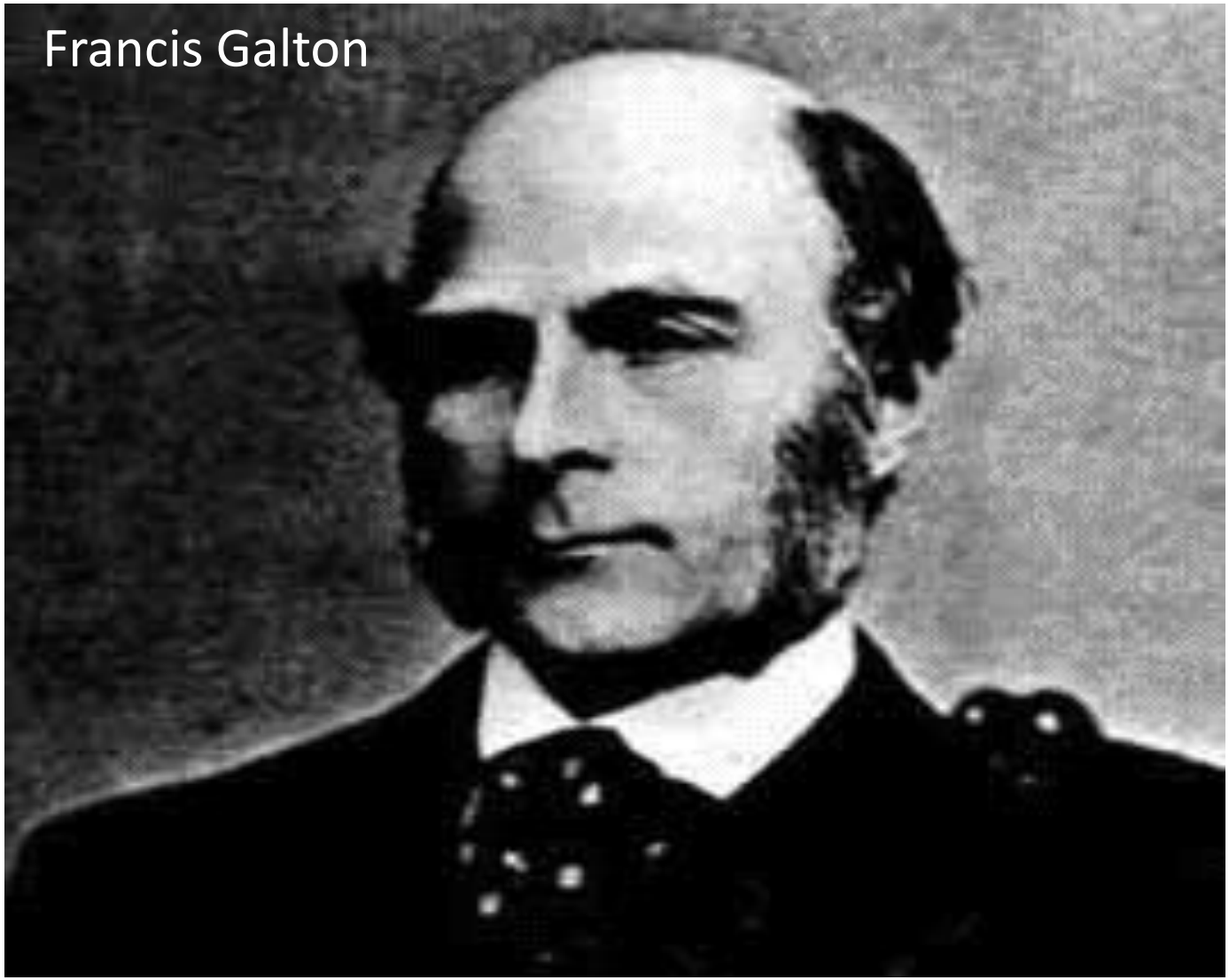
! This is a simplified model, based on a few assumptions:

- Random mating
- No dominance effects

Figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.21)

KU LEUVEN

End 19th century – beginning 20th century



Great-grandfathers : 0.5%

Great-uncles : 0.5%

Cousins : 1.5%

Grandfathers : 7.5%

Uncles : 4.5%

Fathers : 26%

100 ‘eminent’ individuals

Sons : 36%

Brothers : 23%

Grandsons : 9.5%

Nephew : 4.8%

Great-grandsons : 1.5%

Grandnephew : 2%

KU LEUVEN

The **polygenic theory** explains how quantitative traits can be genetically determined

RA Fisher (1918)

Characters that are determined by a large number of independent Mendelian factors could display precisely the continuous nature, quantitative variation, and family correlations described by the biometricians.

DS Falconer

Extended this model to cover dichotomous non-Mendelian characters.

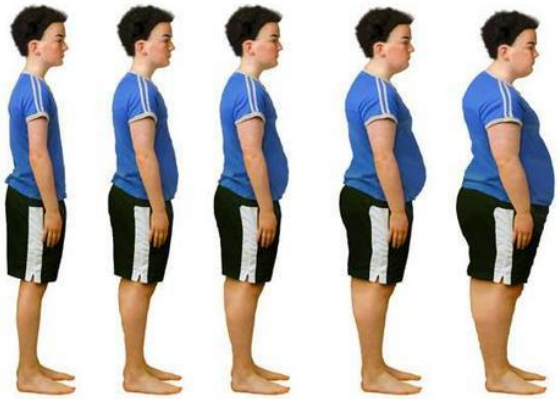
- Existence of monogenic vs multifactorial inheritance
- How does the polygenic theory work?

Quanlitative vs quantitative

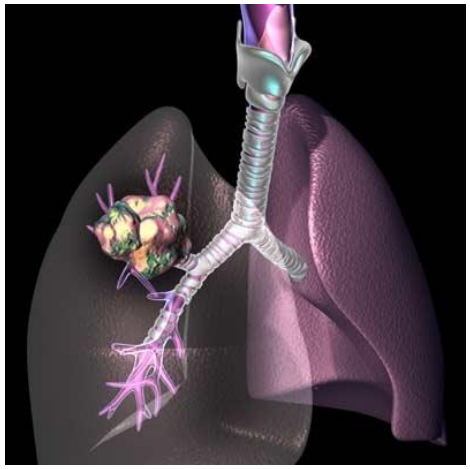
Qualitative	Quantitative
Sharply demarcated types with little connection by intermediates	Continuous gradation among individuals from one extreme to other
Effects of genes are large	Effects of genes are small
Single genes inherited in Mendelian ratios?	Usually many genes

Falconer and Mackay, *Quantitative Genetics* 1996.

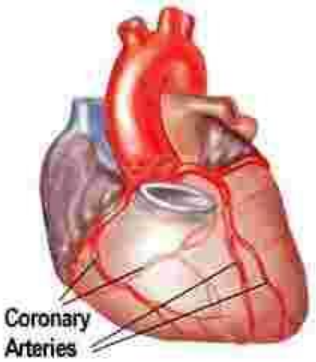
Not for all (seemingly) qualitative traits !



obesity

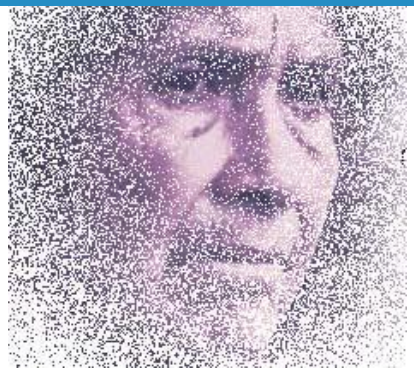


cancer



inherited heart conditions

And many others…



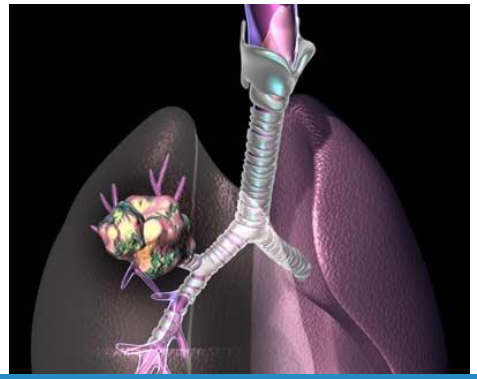
dementia

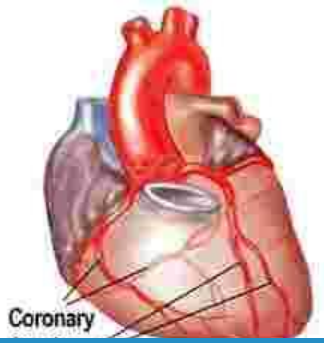
inflammatory bowel disease

cleft lip and/or cleft palate

KU LEUVEN

Not for all (seemingly) qualitative traits !





And many others…

Apparently phenotypic simplicity (dichotomous phenotype) can easily mask a complex genetic basis (genetic + environmental effects)

dementia

inflammatory bowel disease

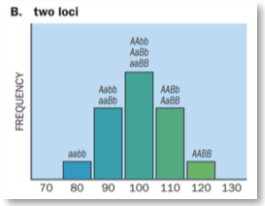
cleft lip and/or cleft palate

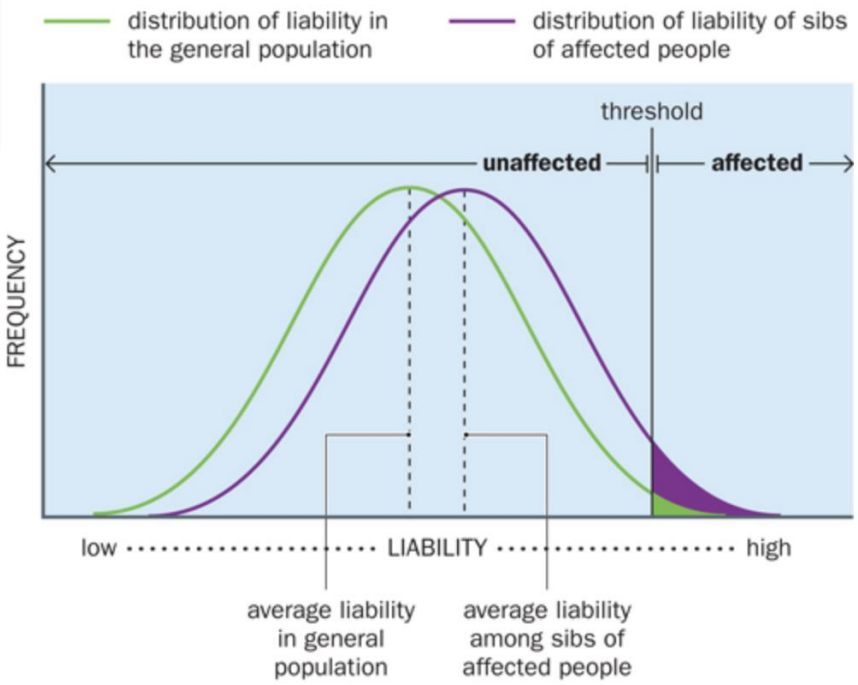
KU LEUVEN

The '**liability threshold model**' extended the polygenic theory to cover non-Mendelian dichotomous characters

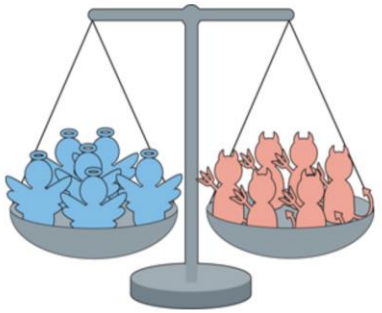
DS Falconer: Extended this model to cover dichotomous non-Mendelian characters.

The key concept is that even for a dichotomous character, there is an underlying continuous variable, namely susceptibility.





'Threshold'



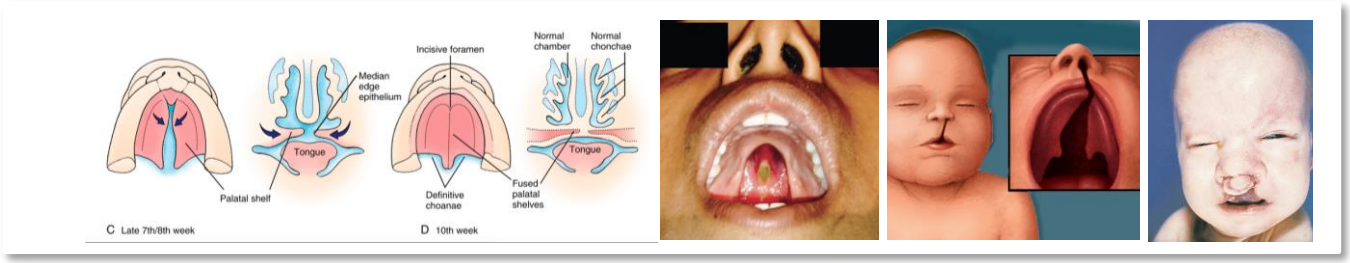
example : cleft lip/palate
(susceptibility is continuously distributed)

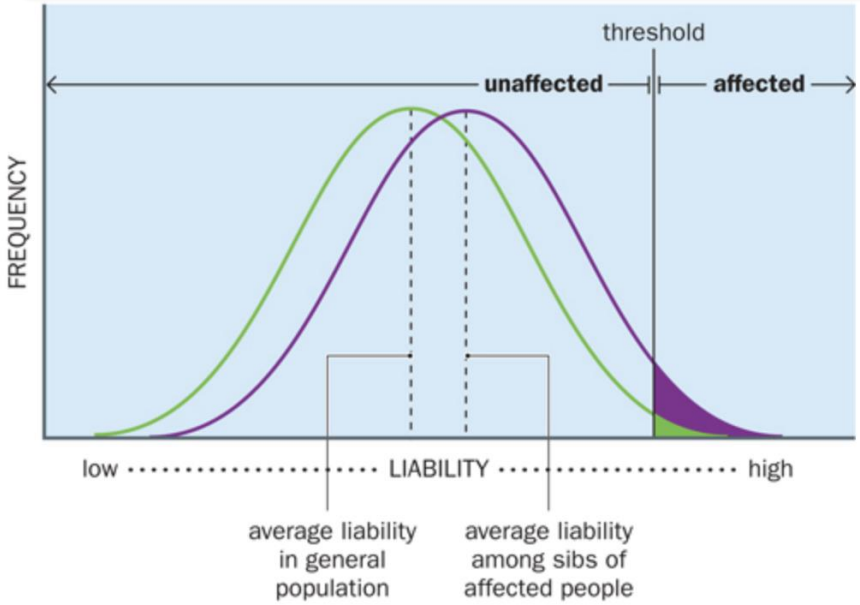
Liability – susceptibility – risk

Figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.22 + 5.23)

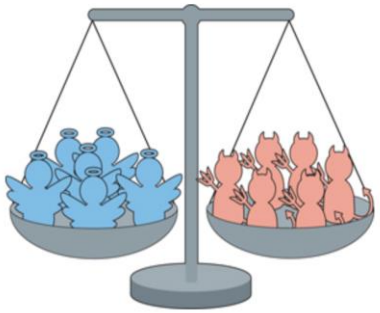
KU LEUVEN

The '**liability threshold model**' extended the polygenic theory to cover non-Mendelian dichotomous characters





'Threshold'



example : cleft lip/palate
(susceptibility is continuously distributed)

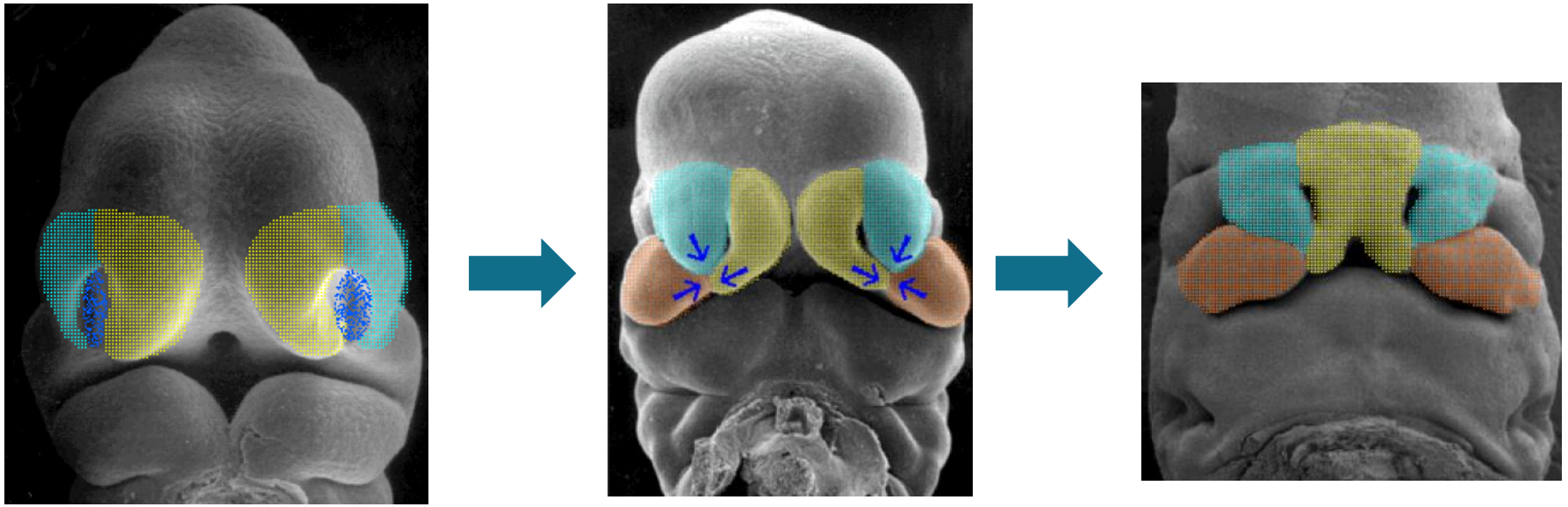
Liability: susceptibility – risk

Figure from: Human Molecular Genetics 5 (chapter 5, Figure 5.22 + 5.23)

KU LEUVEN

E.g. Cleft lip/palate

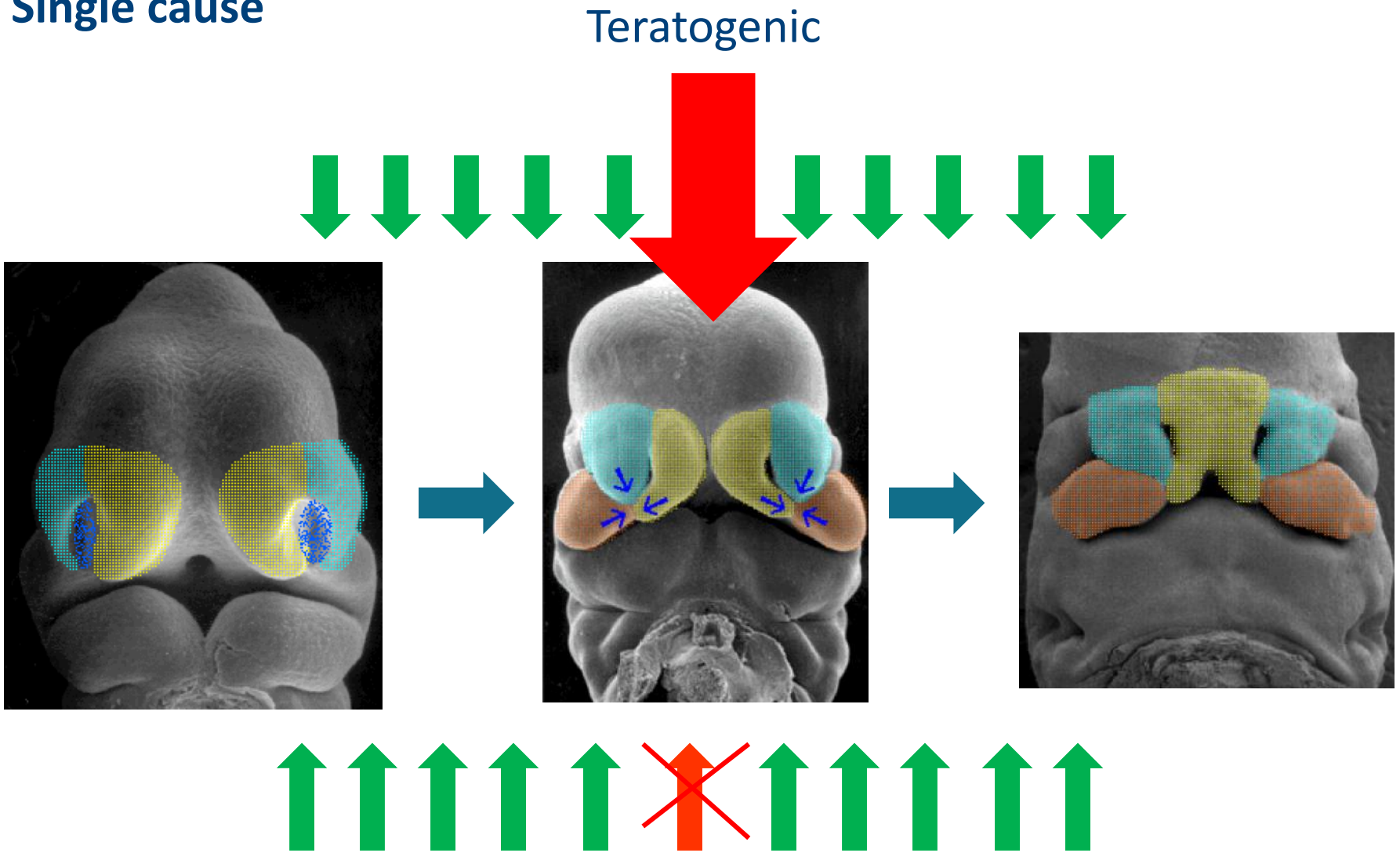
Many environmental factors
diet, O_2, vitamins, temperature …



100's of genes

KU LEUVEN

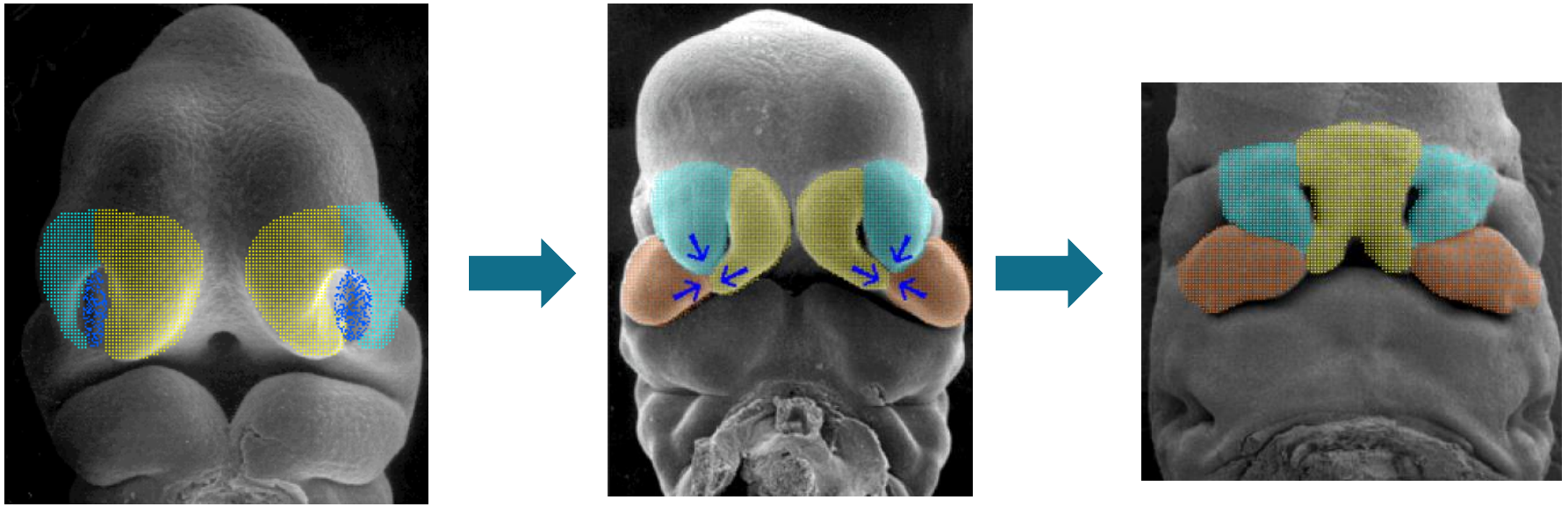
Single cause



Mutation in one gene
E.g. *IRF6* => °cleft lip

Multifactorial

Less favorable environmental conditions

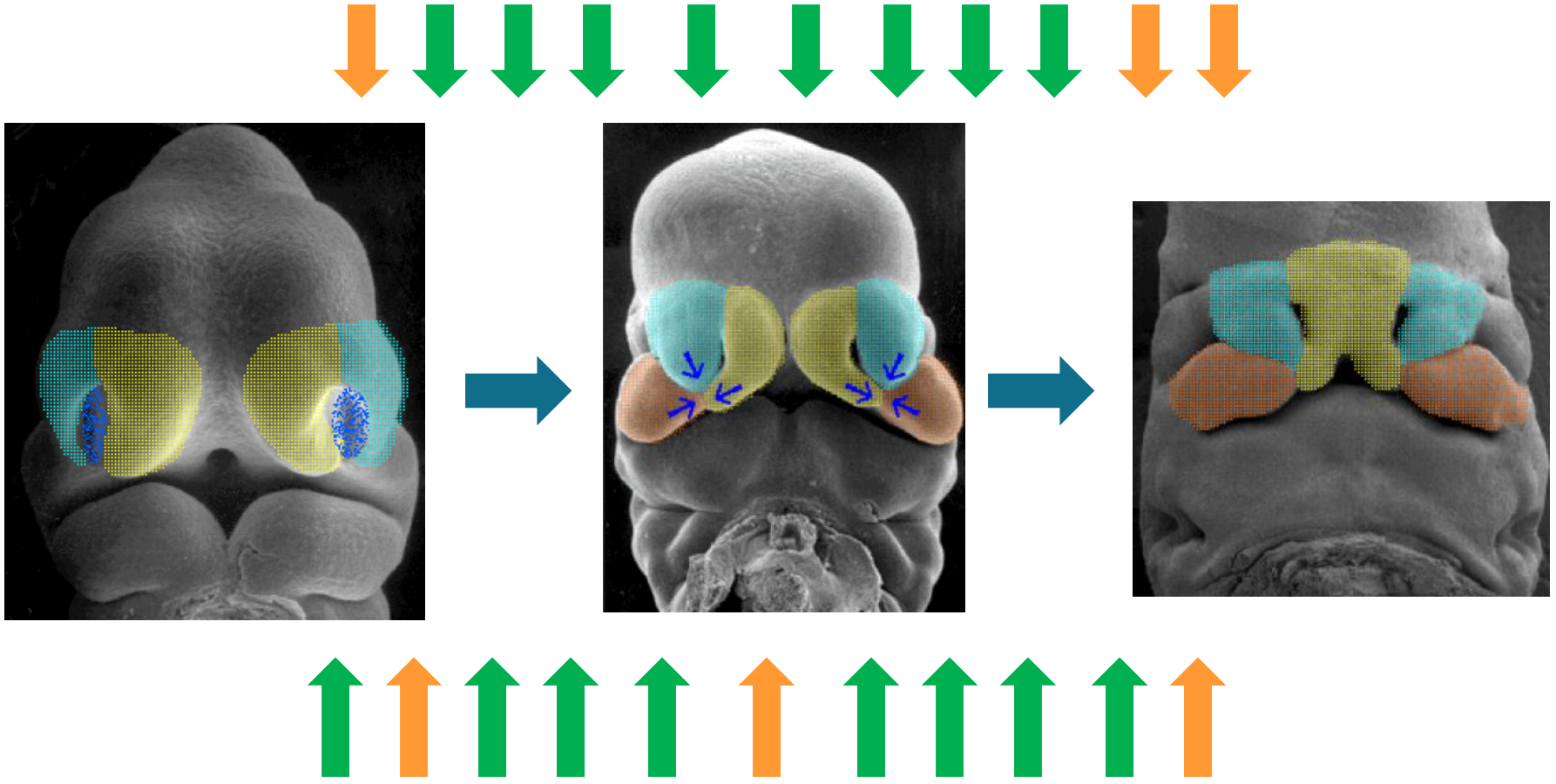


Variations in many genes

KU LEUVEN

Multifactorial

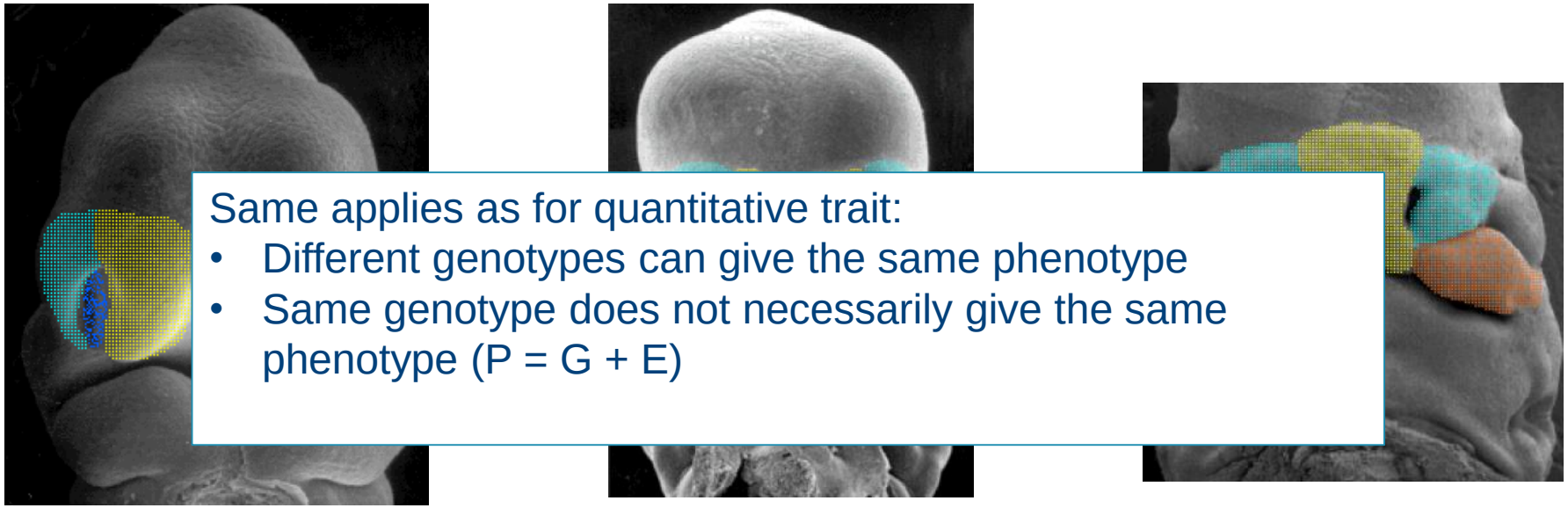
Less favorable environmental conditions



Variations in many genes

Multifactorial

Less favorable environmental conditions



Same applies as for quantitative trait:

- Different genotypes can give the same phenotype
- Same genotype does not necessarily give the same phenotype (P = G + E)

Variations in many genes

KU LEUVEN

Binary trait = genetics + environment

- Disease susceptibility locus with alleles A_1 and A_2 (↑ risk)
- Suppose $P(\text{disease}|A_2A_2) = 0.5$
  - penetrance of the genotype = 50%
  - Suppose for the other genotypes
    - $P(\text{disease}|A_1A_2) = 0.2$
    - $P(\text{disease}|A_1A_1) = 0.05$

	A1	A2
A1	A1A1	A1A2
A2	A2A1	A2A2

genotypes

- Presence of A_2 significantly increases the risk, but A_1A_1 individuals can sometimes display the disease because of exposure to adverse environmental factors = phenocopies
- If the A_2 allele is rare : most observed disease cases are environmental (resulting from A_1A_1) rather than genetically (from A_2*) in cause

Binary trait = genetics + environment

For example:

Suppose freq(A_1) = 0.9

- What is $P(A_2A_2 \mid \text{disease})$?

The population prevalence *K* (the frequency) of the disease is

$K = freq(disease)$
$= P(A_2A_2) * P(\text{disease}|A_2A_2) + P(A_1A_2) * P(\text{disease}|A_1A_2) + P(A_1A_1) * P(\text{disease}|A_1A_1)$
$= 0.01 * 0.5 + 2 * 0.1 * 0.9 * 0.2 + 0.81 * 0.05$
$= 0.0815$ or roughly 8% of the population shows the disease

Probabilists → Bayes' theorem

Binary trait = genetics + environment

- Application of Bayes' theorem (= probability of event, based on prior knowledge of conditions related to the event) :

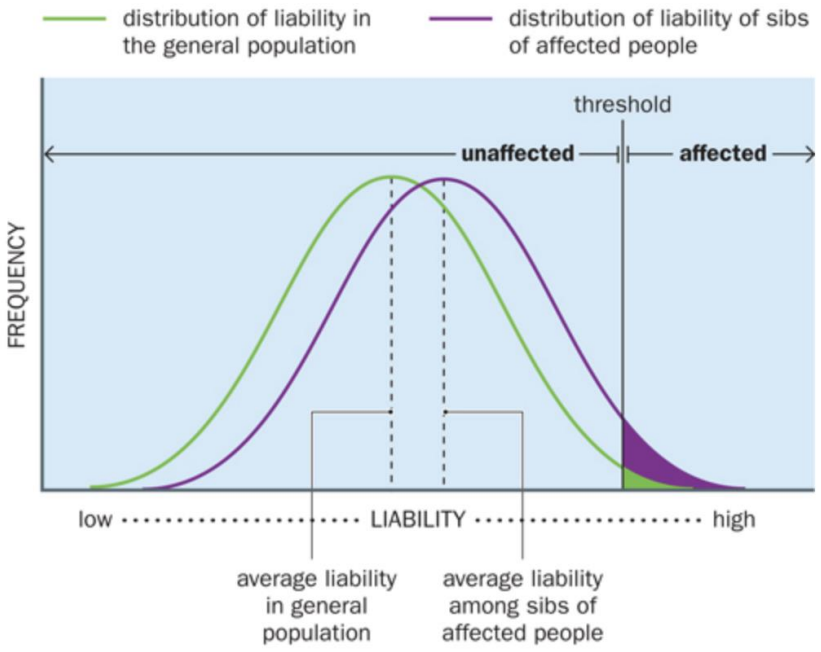
$$P(b|A) = \frac{P(A|b) * P(b)}{P(A)}$$

Probability of having b (genotype) given the observed outcome of correlated variable A (disease)

$$P(A_2A_2|disease) = \frac{P(disease|A_2A_2)*P(A_2A_2)}{P(disease)} = \frac{0.5 * 0.01}{0.0815} = 0.06$$

- If we pick a random individual showing the disease, there is only 6% chance he/she has the high risk genotype A_2A_2
- Likewise: $P(A_1A_2|disease) = 0.442$; $P(A_1A_1|disease) = 0.497$

The '**liability threshold model**' extended the polygenic theory to cover non-Mendelian dichotomous characters

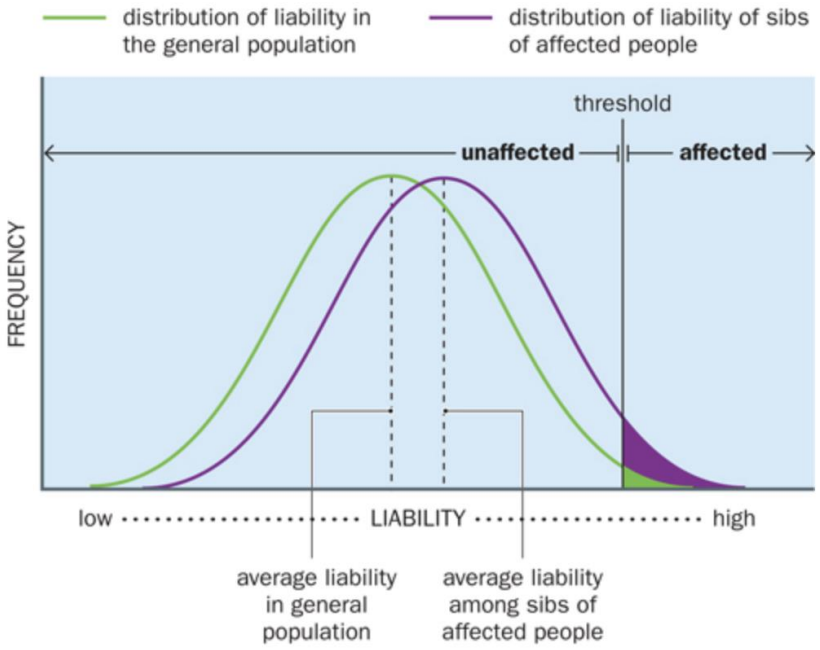


Affected individuals have an unfortunate combination of high-susceptibility alleles.

Their relatives who share genes (alleles) with them will also, on average, have an increased susceptibility. The divergence of the population mean will be dependent on the proportion of shared genes (alleles).

→ **Polygenic threshold characters (dichotomous characters) thus also run in families.**

In contrast to Mendelian conditions, the recurrence risk for polygenic conditions depend on previous history !



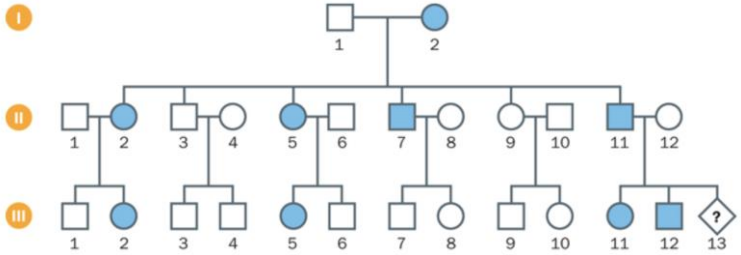
Parents who have several affected children may have just been unlucky, but on average they will have more high-risk alleles than parents with only one affected child:

- The treshold is fixed,
- But the average susceptibility, and hence recurrence risk, increases with increasing number of previous affected children

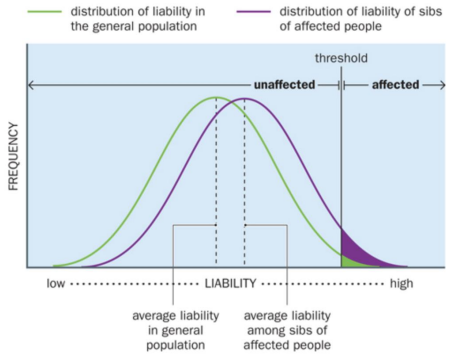
KU LEUVEN

Recurrence risk of a phenotype or character

Mendelian character: theoretical basis (based on pattern of inheritance) – 'chance has no memory'



Non-Mendelian character: empirical risks – 'bad luck in the past is a predictor of bad luck in the future'



KU LEUVEN

Non-mendelian characters (multifactorial)

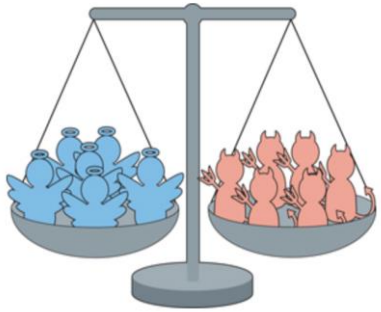
Quantitative trait (eg. height)

- Underlying loci = **Quantitative trait loci (QTLs)**: a locus/gene(variant) that contributes to the phenotype of a quantitative trait

Dichotomous trait: either you have it or you don't

- Underlying loci = **susceptibility loci/genes** : a locus/gene(variant) that influences your susceptibility for the phenotype

! One genetic variant is **not necessary nor sufficient** to determine the trait. A multitude of genetic and environmental factors increase or decrease the value of a quantitative trait, or the likelihood of manifesting the dichotomous trait.



Summary so far

- Many “traits” are expressed on a quantitative scale
- The underlying cause is the influence of many genes
- The polygenic model (infinitesimal model[1]) seems to work well, also for dichotomous traits

[1]developed by Fisher (1918) (“analysis of variance”)