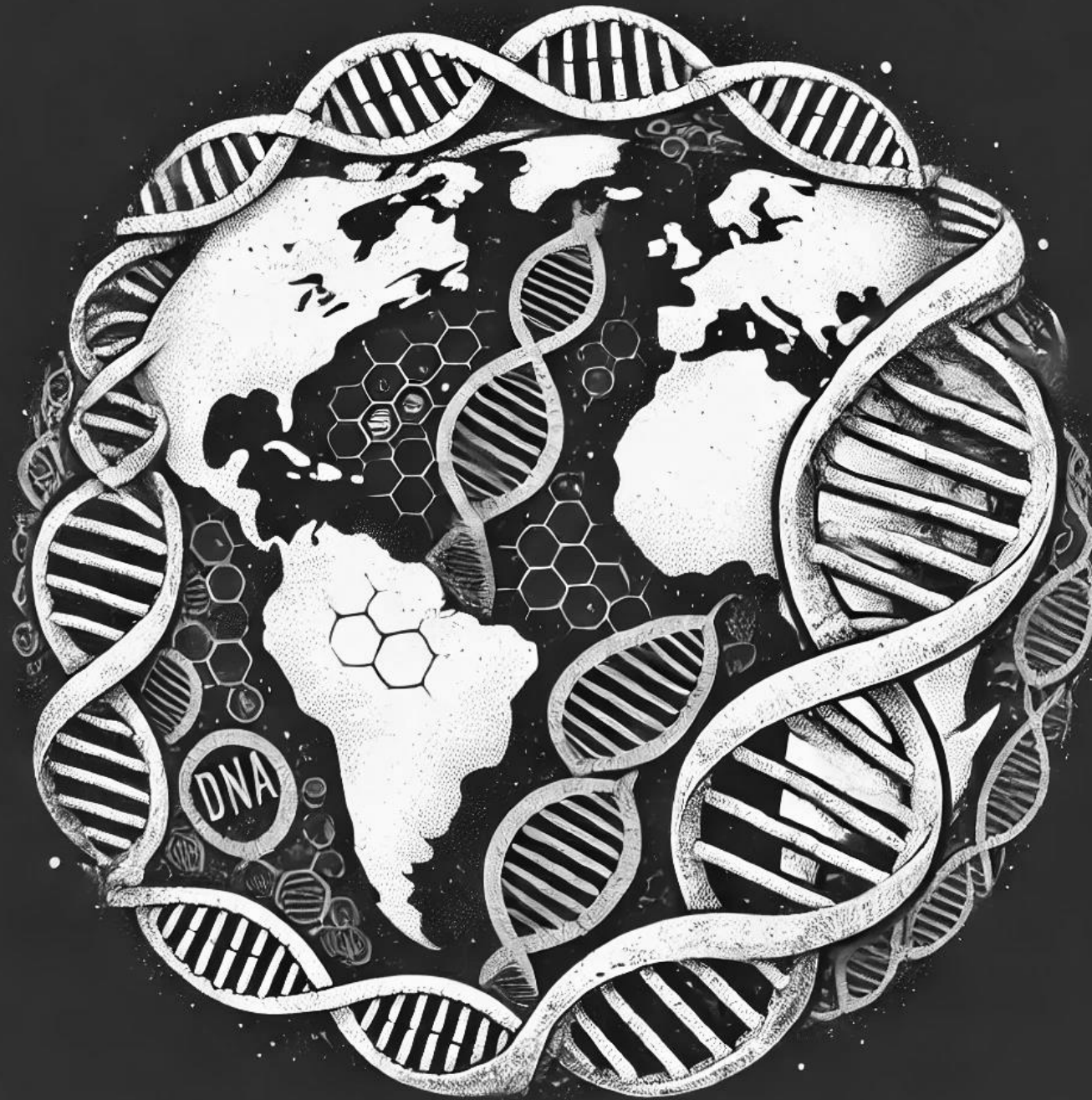




MASTER COURSE IN PERSONALISED MEDICINE POLYGENIC SCORES

Palle Duun Rohde

[palledr@hst.aau.dk | p.rohde@rn.dk]

<https://pdrohde.github.io/>

Associate Professor & Research Group Leader
Department of Health Science and Technology
Aalborg University

Clinical Academic
Department of Clinical Genetics
Aalborg University Hospital

THE PURPOSE OF TODAY

- ❖ Give an introduction to polygenic scores (PGS)
- ❖ How genetic variation shapes phenotypic diversity
- ❖ Provide an introduction to complex trait genetics
 - Monogenic vs multifactorial aetiology
- ❖ How we can utilize genomic data to elucidate molecular genetic aetiology underlying complex traits
 - Genome-wide association studies (GWAS)

Slides → <https://pdrohde.github.io/teachingmaterial/>



WHO AM I?

- ▶ BCs in biology from AU
- ▶ MSc in genetics from AU
- ▶ PhD in statistical and quantitative genetics from AU
- ▶ Research group leader aiming to understanding the genetic architecture of common complex diseases and translating genomic discoveries into improved prevention, diagnosis, and treatment
- ▶ A central research priority for the group is advancing statistical genomics through AI-driven methods.

Palle Duun Rohde



Associate Professor in Statistical and Complex Trait Genetics

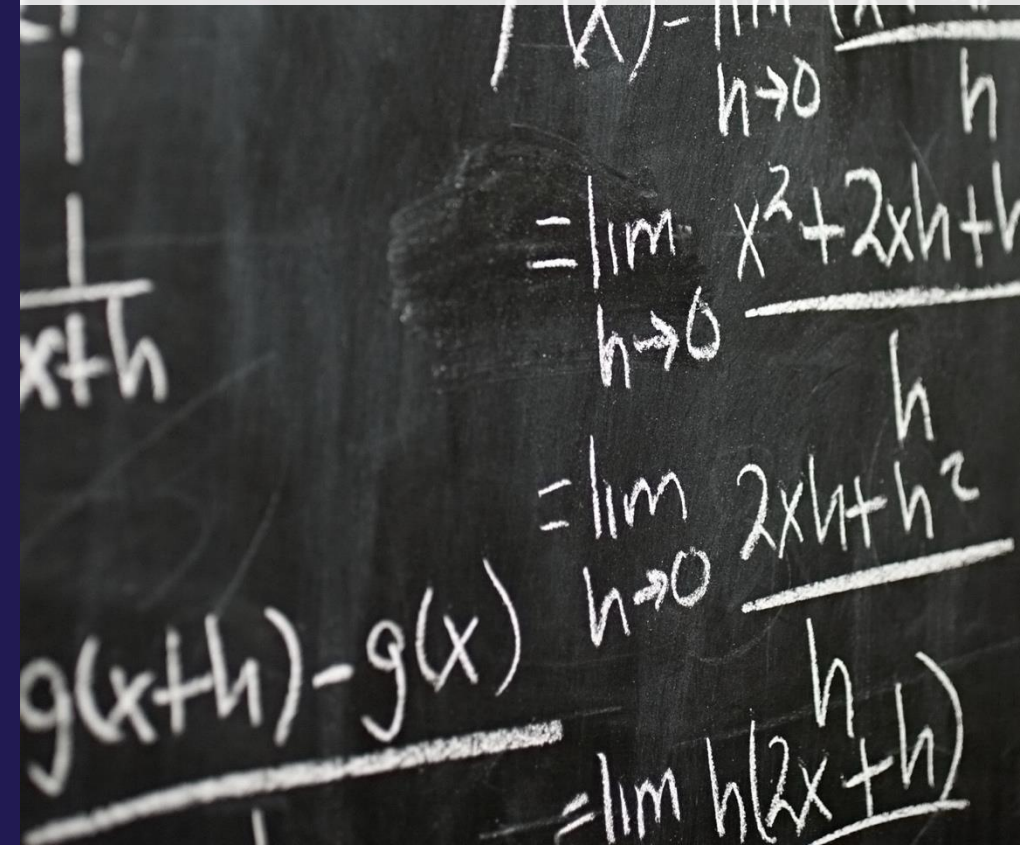
E-mail contact: palle@hst.aau.dk

I am an Associate Professor, and research group leader, at the [Genomic Medicine research group](#) at the [Department of Health Science and Technology, Aalborg University, Denmark](#), and hold a part-time position as clinical academic in theoretical genetics at the [Department of Clinical Genetics, Aalborg University Hospital](#).

Research

My full publication record can be found at the [Publications](#)-page at the top on this site, and my university research profile can be found [here](#).

Software Packages


$$\begin{aligned} f(x) &= \lim_{h \rightarrow 0} \frac{(x+h)^2}{h} \\ &= \lim_{h \rightarrow 0} \frac{x^2 + 2xh + h^2}{h} \\ &= \lim_{h \rightarrow 0} \frac{2xh + h^2}{h} \\ &= \lim_{h \rightarrow 0} h(2x + h) \end{aligned}$$

AGENDA

- 09:00 – 09:15** Welcome
- 09:15 – 09:45** Session 1: Genetic variation and heredity
- 09:45 – 10:15** Group work 1: Genetic variation and heredity
- 10:15 – 10:30** Break
- 10:30 – 11:00** Session 2: From SNP to GWAS to disease risk
- 11:00 – 11:40** Group work 2: Construct your own PGS
- 11:40 – 12:00** Common discussion and evaluation

Join at menti.com | use code **7798 1843**



What comes to mind when you hear *personalised medicine*?

accurate medicine
multiple variationer
kvalitet
patienten i fokus
hver patient sin behandli
targeted treatment
trakteret behandling
datedrevet beslutninger
prs
individualised treatment
individual risk
målrettet behandling
subgrupper
targetered behandling



9 / 10



PD

Menti

PM-genetik2025



Choose a slide to present

What is personalised medicine?

What do you associate with *genetic variation* in the context of *personalised medicine*?

Hvilken indsigt tager du med dig fra i dag?

0 questions
0 upvotes

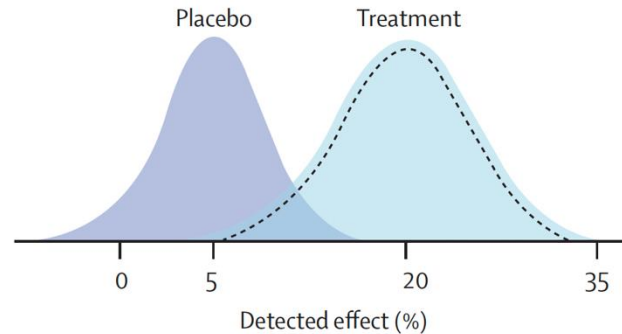
IMPLEMENTATION OF PRECISION MEDICINE

EPPOS [evidence-based precision personalised objective subjective]

Evidence-based Medicine

(1) Contemporary evidence-based medicine

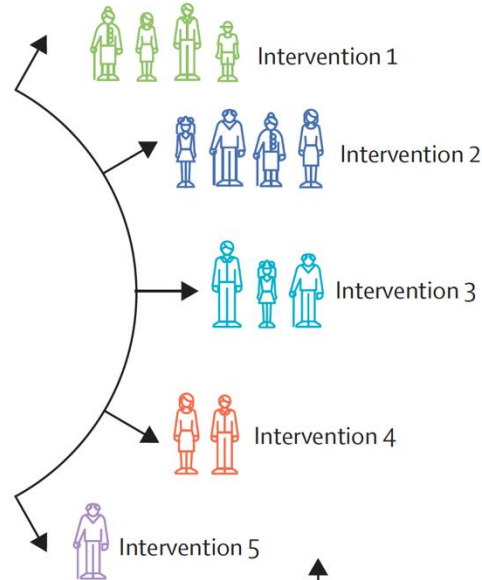
Estimate average risk or response using epidemiological and clinical trial cohorts



Precision Medicine

(2) Probability scoring and stratification

Maximise response and minimise risk using subclassification



Personalised Medicine

(3) Personalisation (objective)

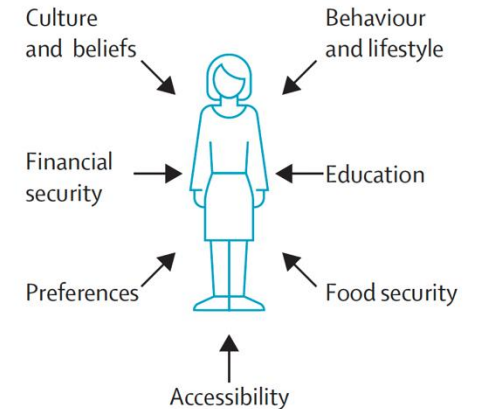
Monitor response to optimise dose, timing, and delivery



Individualised Medicine

(4) Personalisation (subjective)

Adapt intervention to fit the person's needs, capabilities, and preferences



AGENDA

- 09:00 – 09:15 Welcome
- 09:15 – 09:45** Session 1: Genetic variation and heredity
- 09:45 – 10:15 Group work 1: Genetic variation and heredity
- 10:15 – 10:30 Break
- 10:30 – 11:00 Session 2: From SNP to GWAS to disease risk
- 11:00 – 11:40 Group work 2: Construct your own PGS
- 11:40 – 12:00 Common discussion and evaluation

Join at menti.com | use code **7798 1843**



PD

**Menti**

PM-genetik2025

**Choose a slide to present**

What is personalised medicine?

What do you associate with *genetic variation* in the context of *personalised medicine*?

Hvilken indsigt tager du med dig fra i dag?



9 / 10



0 questions
0 upvotes

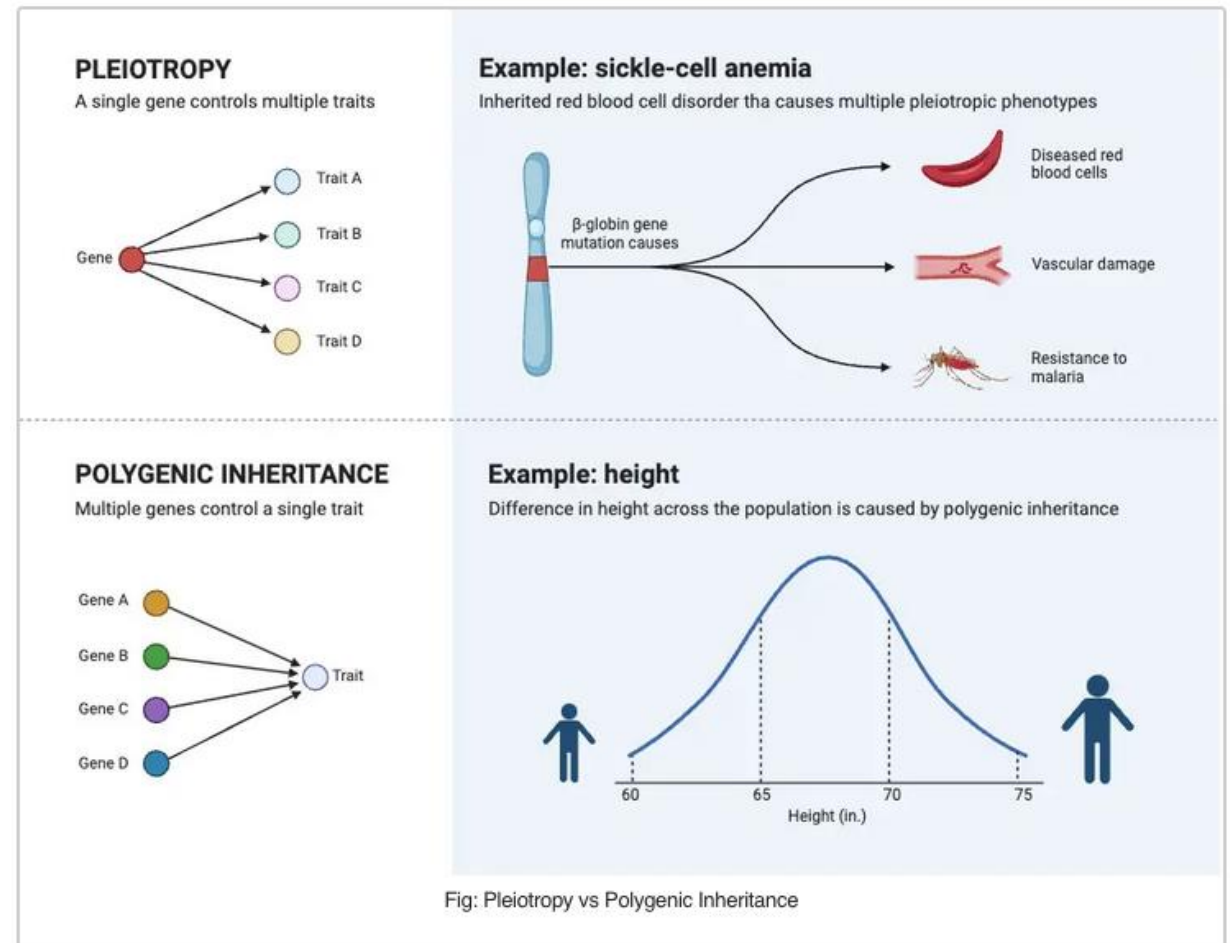
SESSION 1

- ▶ Complex traits
- ▶ Genetic variation
- ▶ Linkage Disequilibrium (LD)
- ▶ Genetic parameters



DIFFERENT MODE OF INHERITANCES

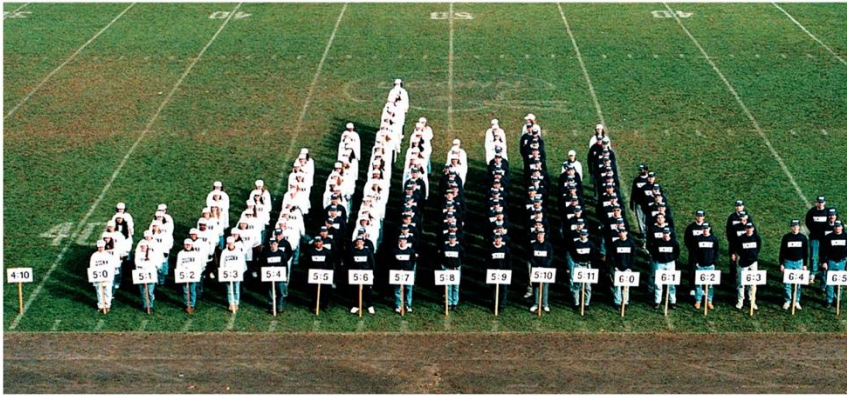
- ❖ Monogenic (single gene variant)
- ❖ Polygenic (many gene variants)
- ❖ Multifactorial (many gene variants plus environment exposures)



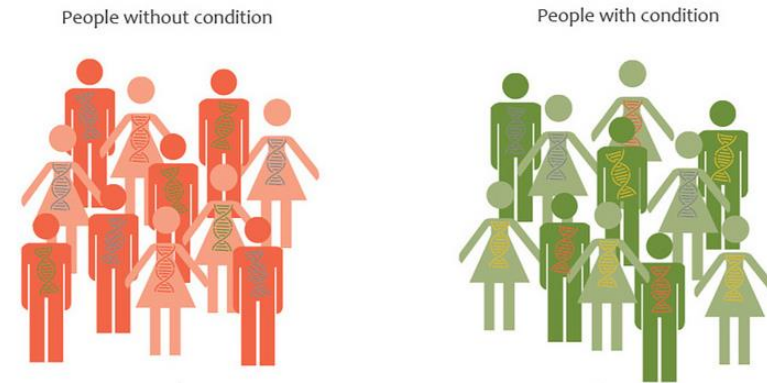
QUANTITATIVE TRAITS

IN DIFFERENT SHAPES

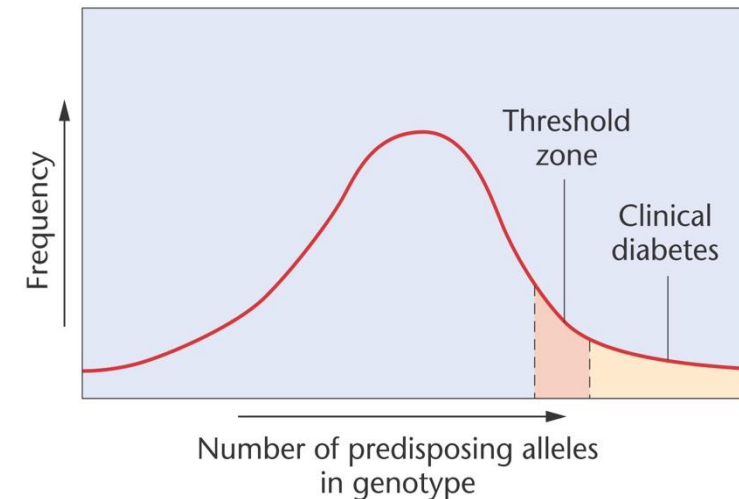
Continuous variation



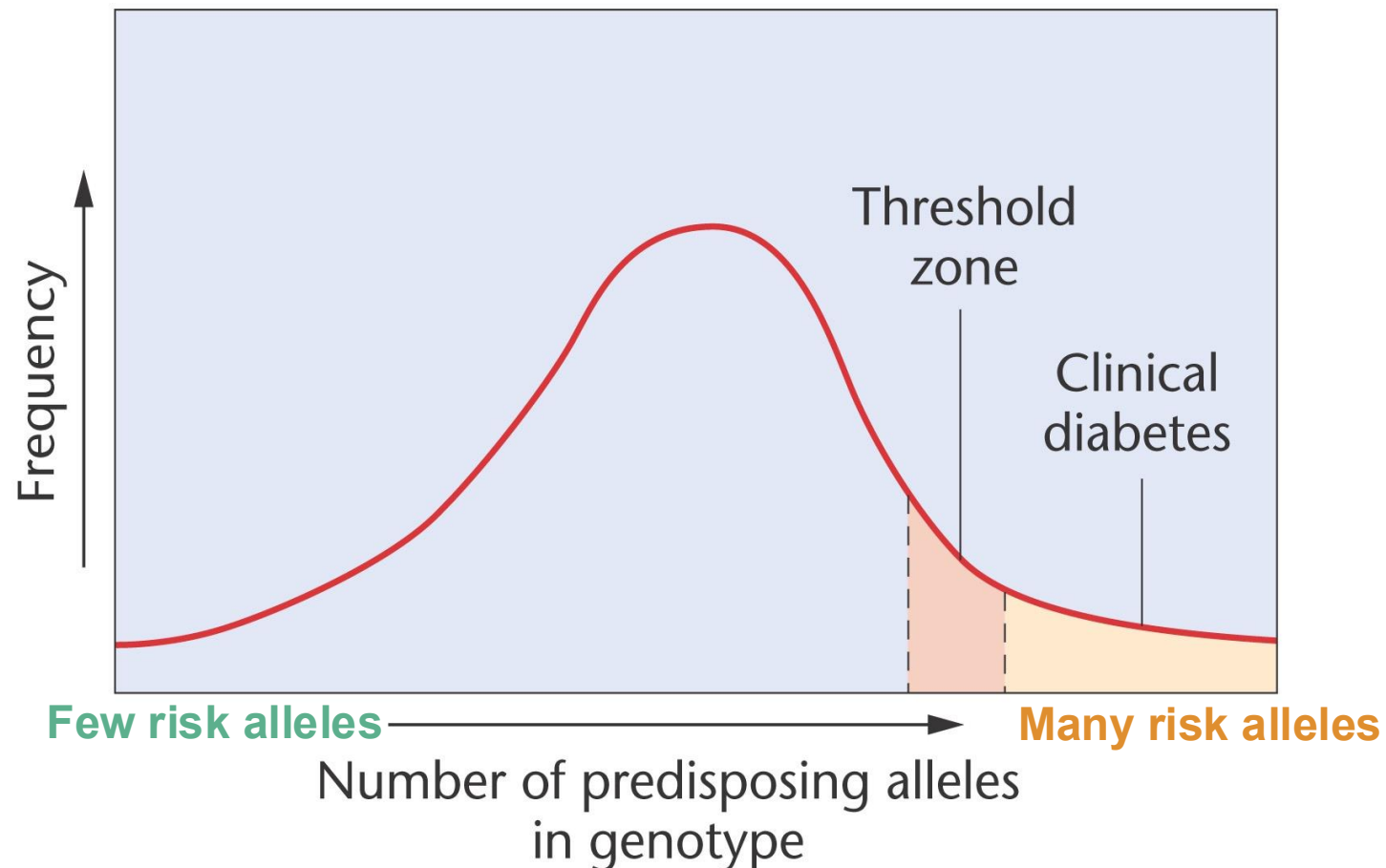
Threshold traits



Categorical variation



LIABILITY (THRESHOLD) MODEL



Liability model

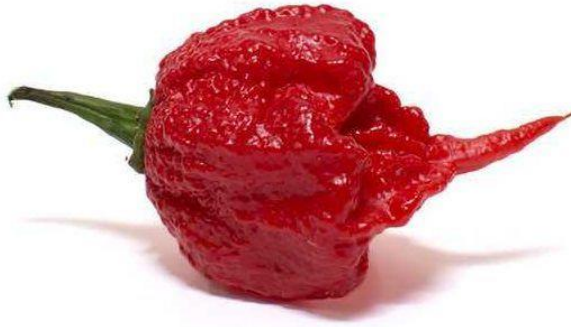
Only individuals with a liability over a certain threshold will become affected

The **sum** of many genetic variants with **small effect/risk**.

Each locus follow Mendelian inheritance pattern, although the trait does not

GENETIC VARIATION IS LIKE CHILI

Carolina reaper



Strong effect on phenotype

Habanero Lemon

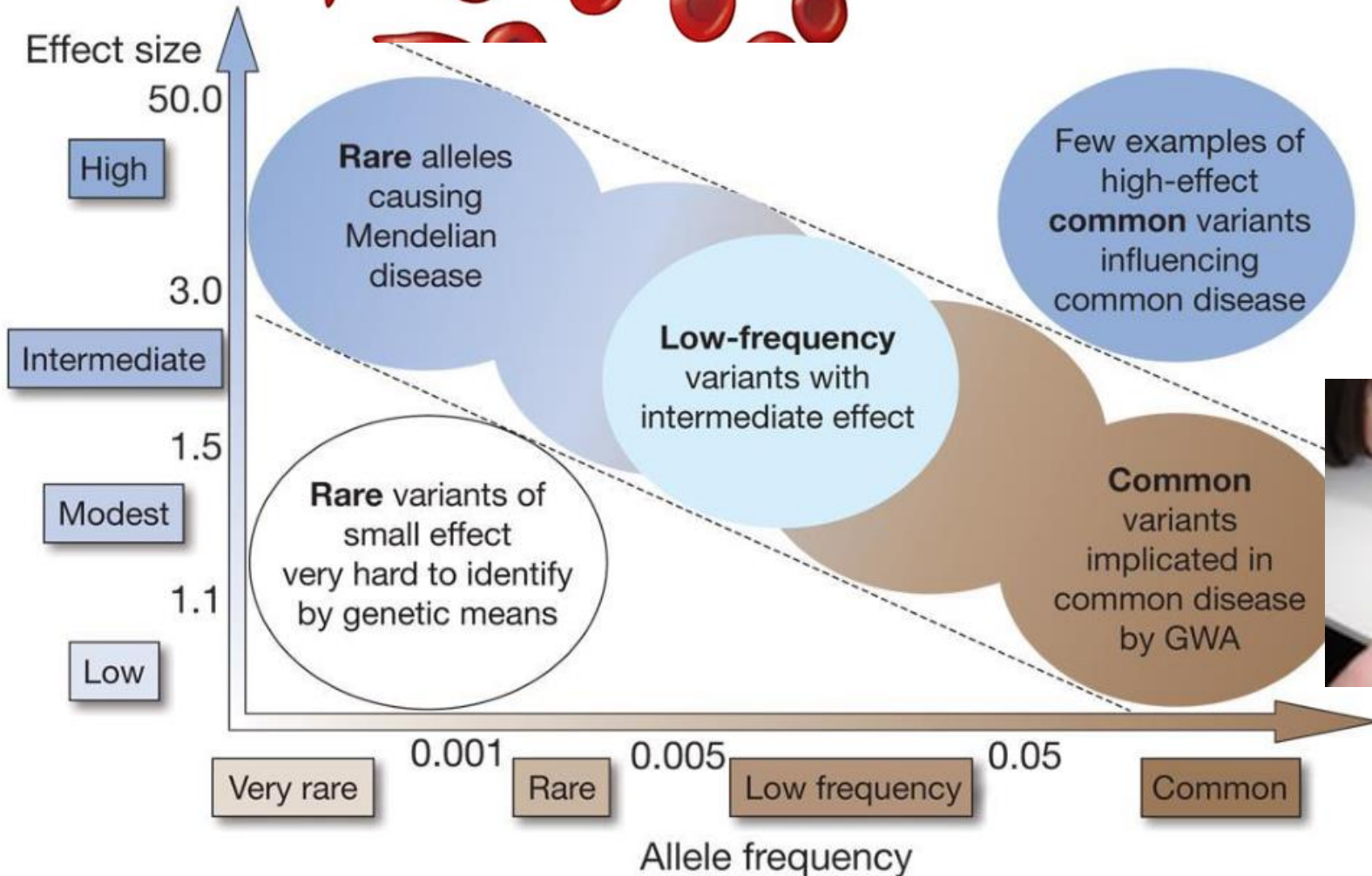
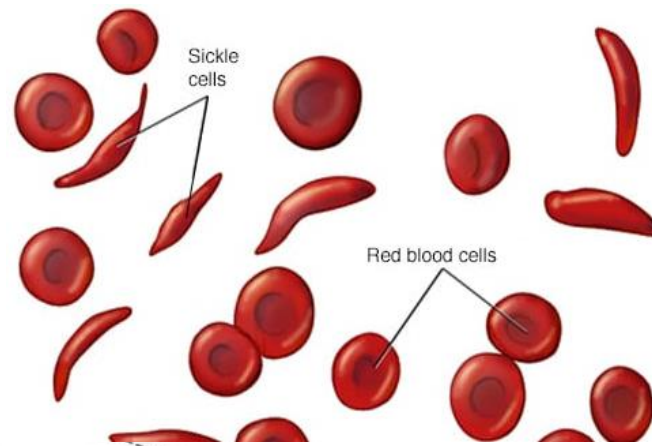


Moderate effect on phenotype

Bell pepper



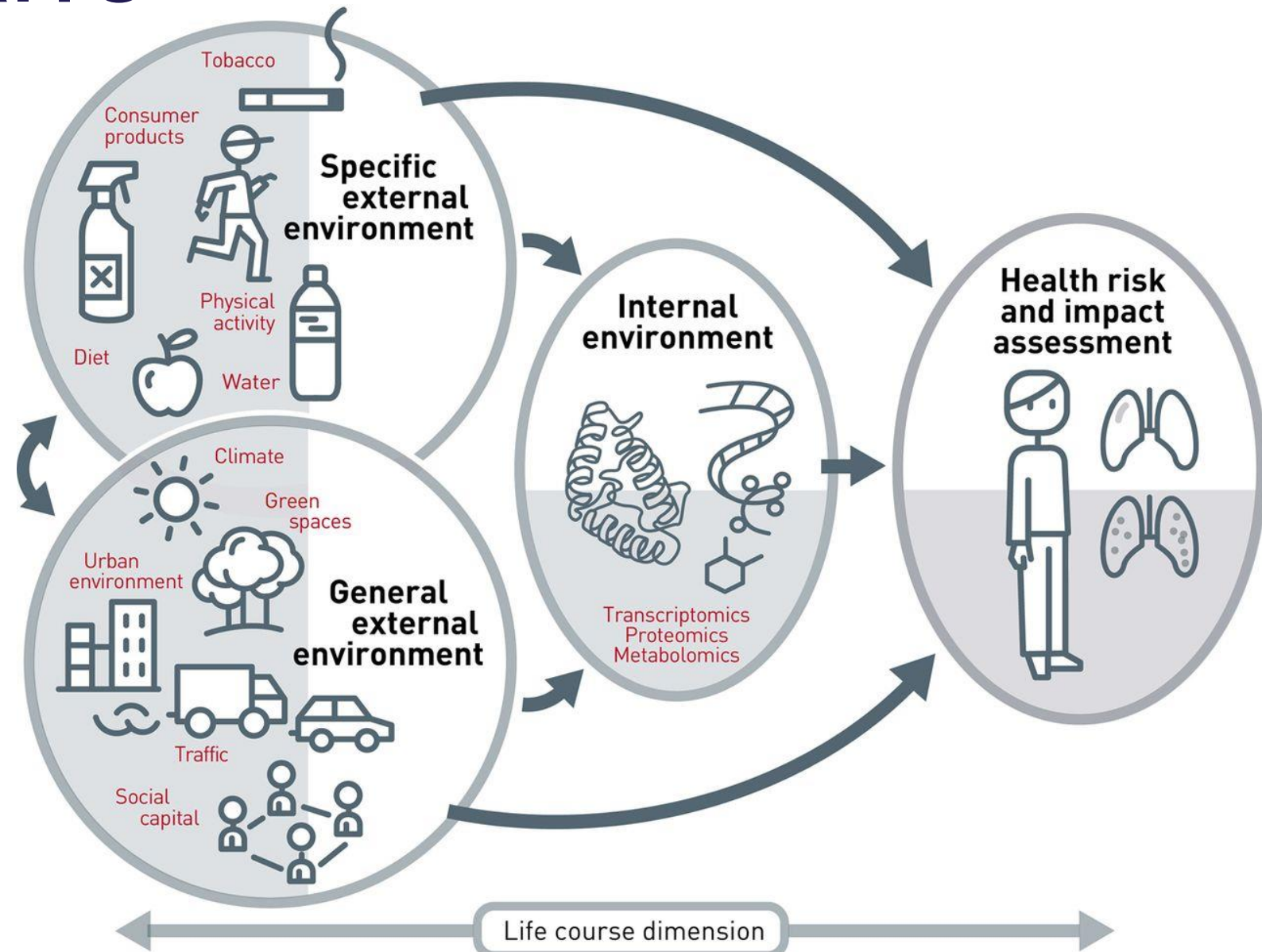
Weak effect on phenotype
(or none at all)



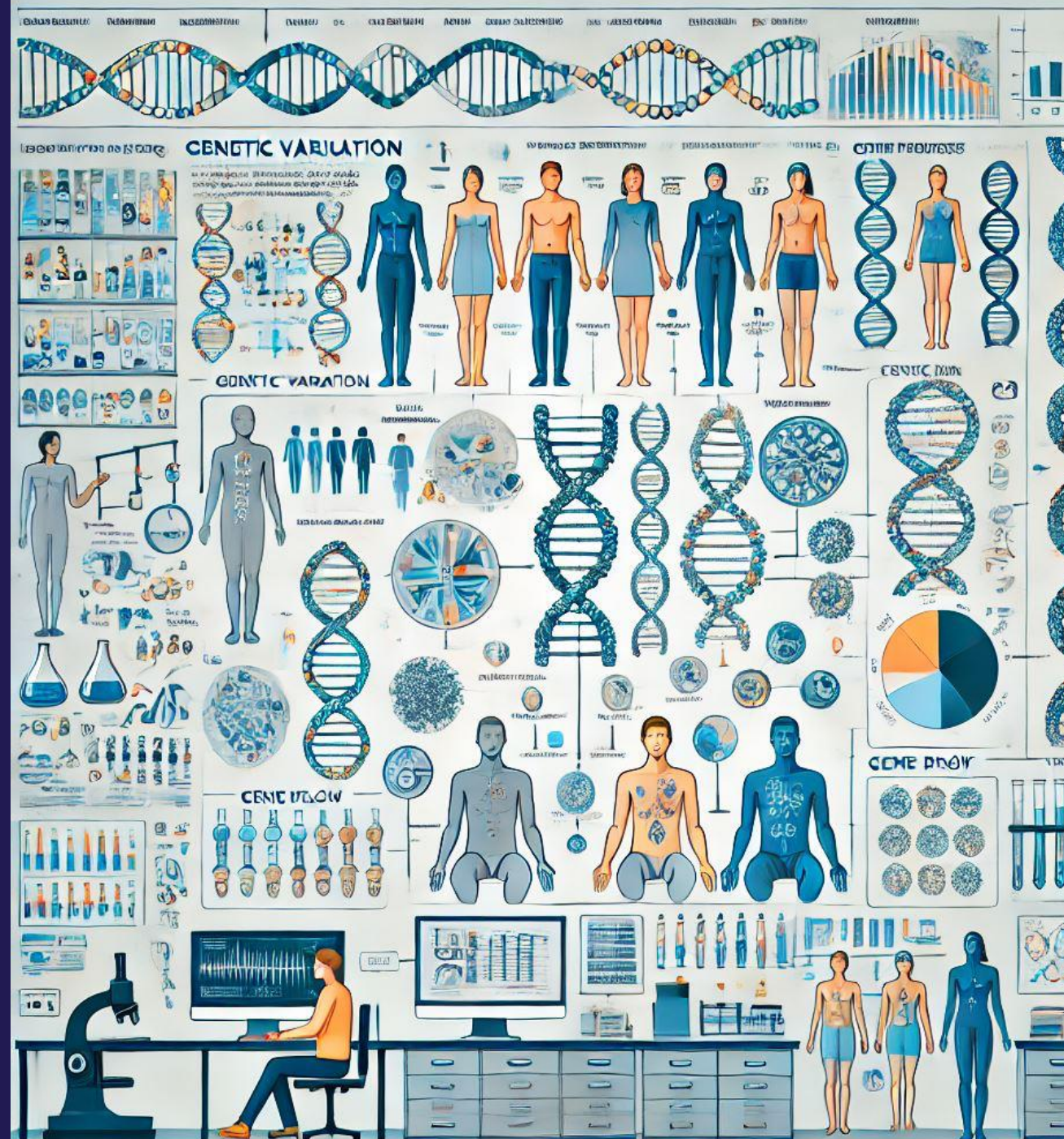
COMPLEX TRAITS are complex...

Complex traits have a **genetic component** and an **environmental component**

The relative genetic contribution is called **heritability**.



GENETIC VARIATION

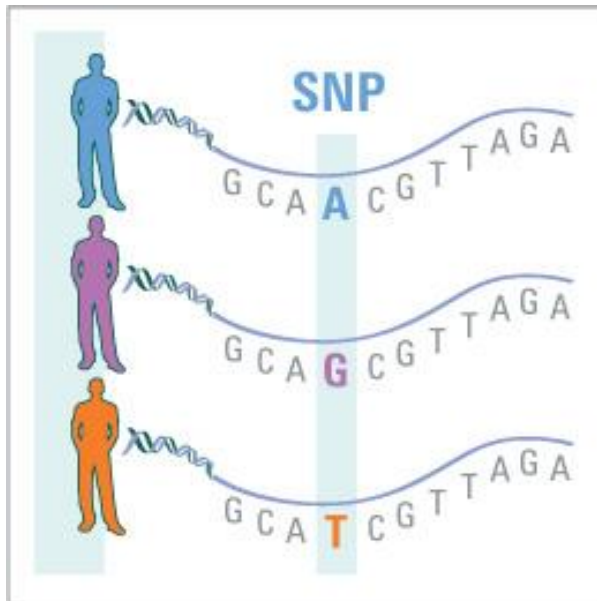


GENETIC VARIATION

SINGLE NUCLEOTIDE POLYMORPHISMS (SNPs)

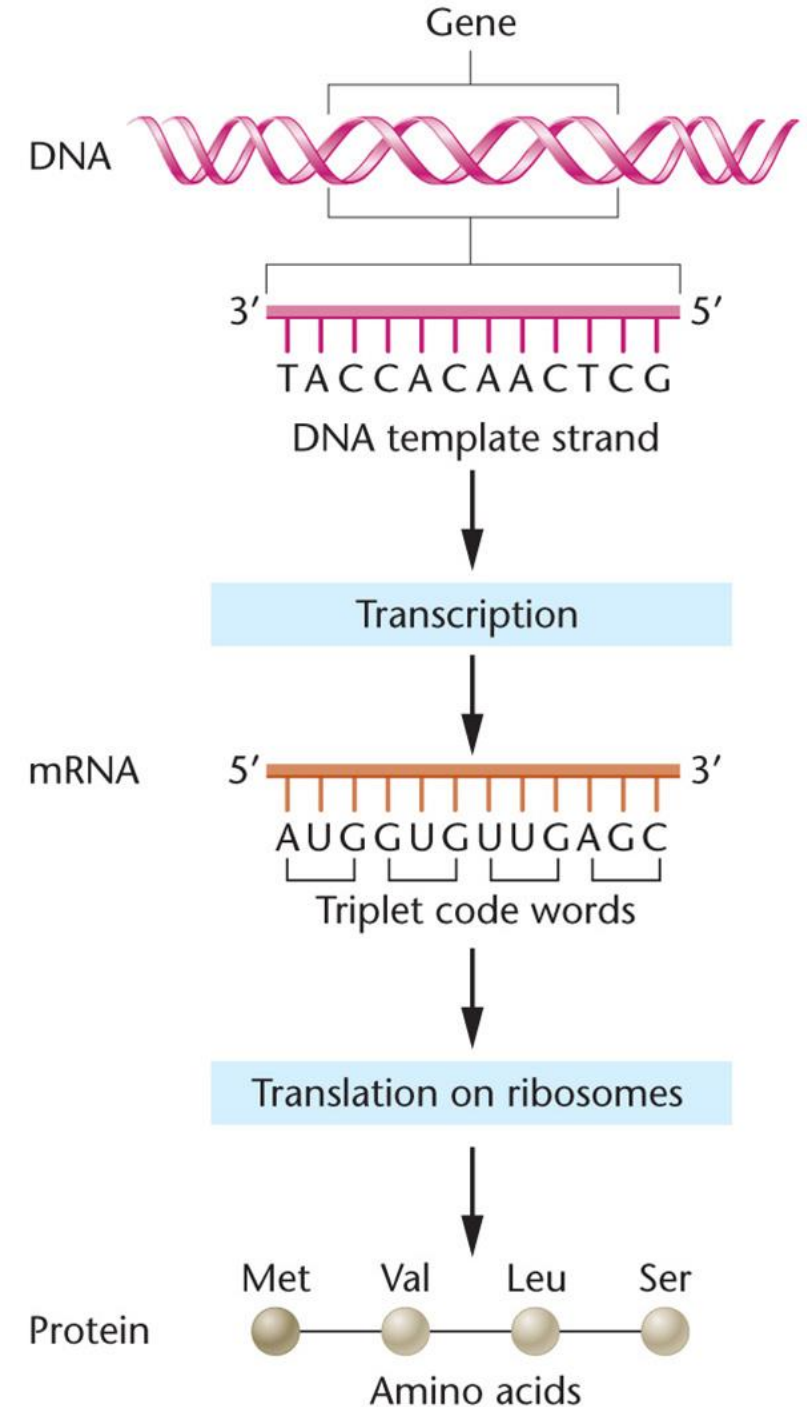
A common change in a single base pair; ~1/1000 bp

Accounts for ~90% of all variation in the human genome

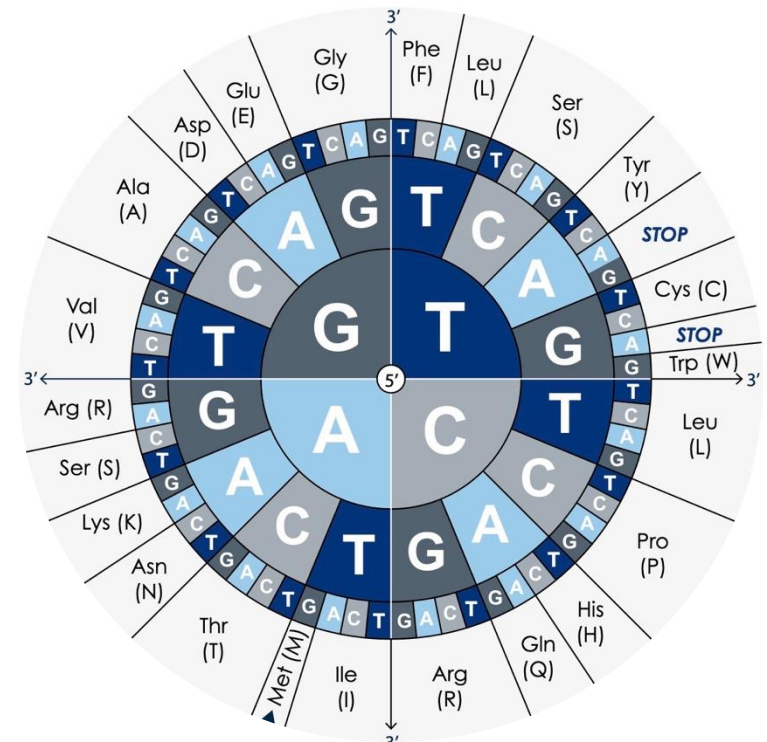
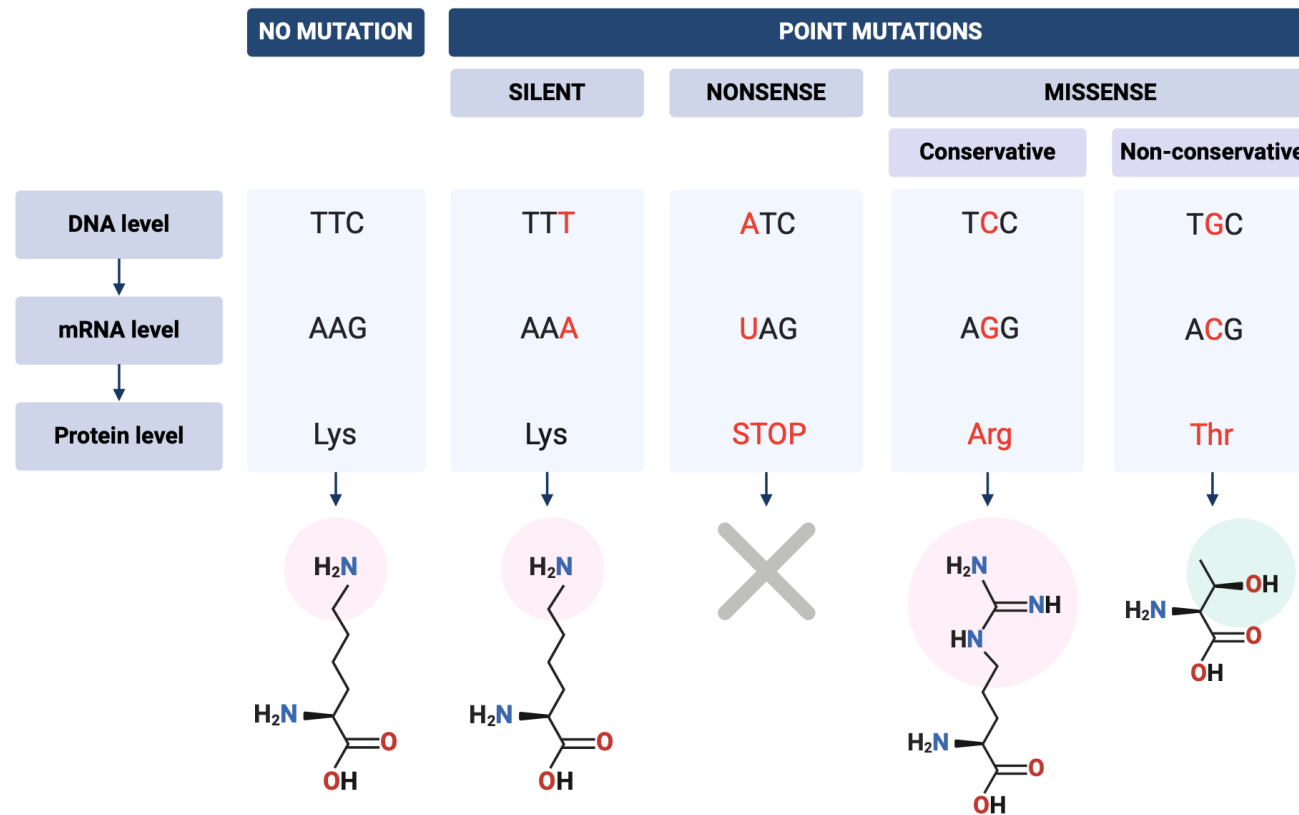


All (known) SNPs has a unique identifier
(independent of alleles)

rsXXX – Ref-SNP cluster ID number



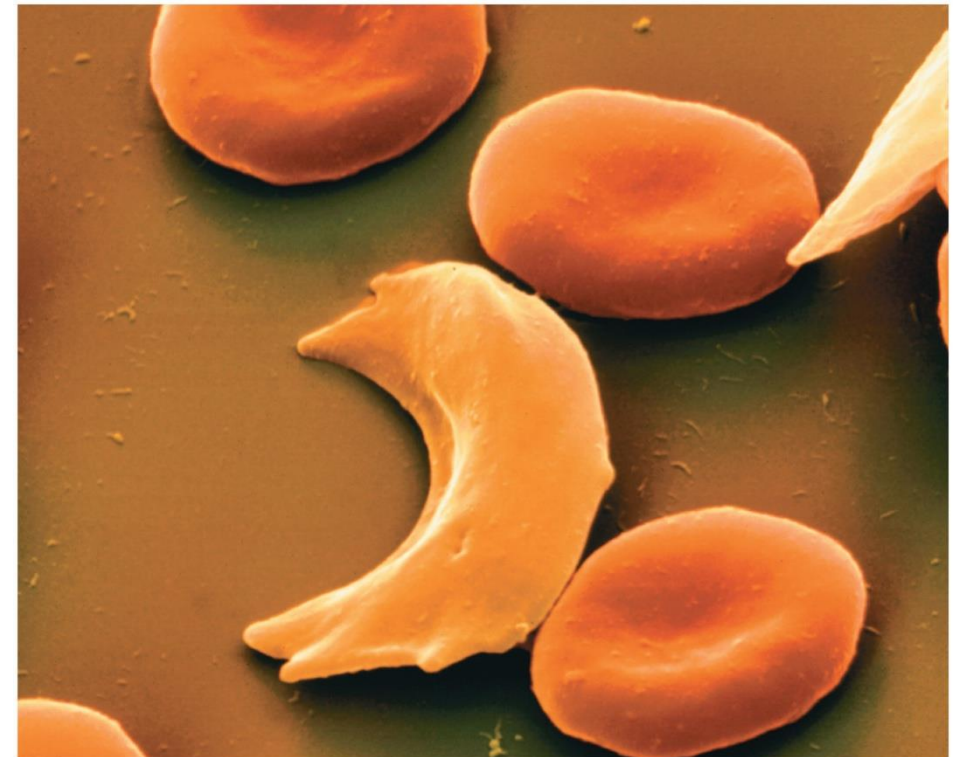
MUTATIONS GENERATE GENETIC VARIATION



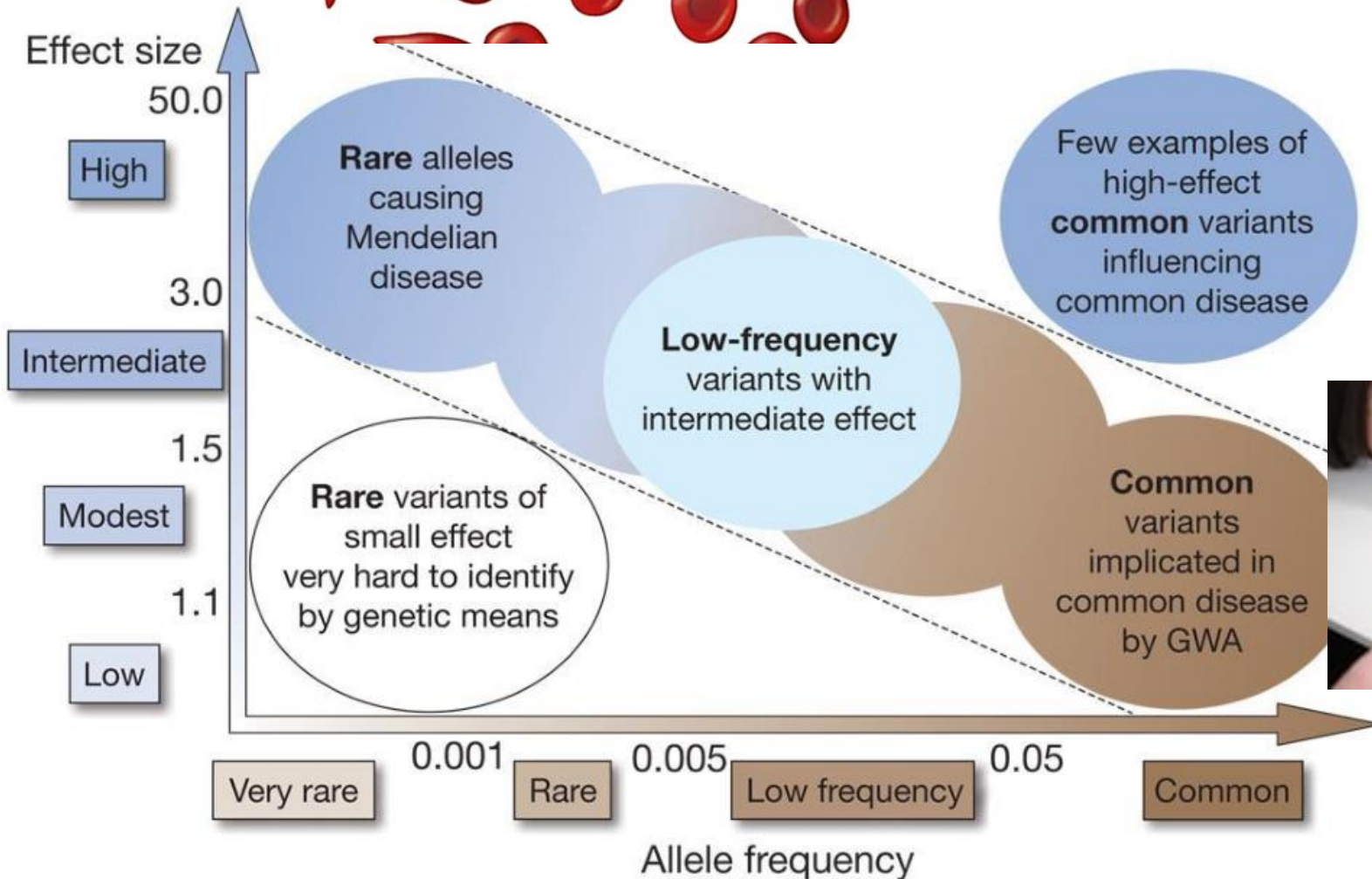
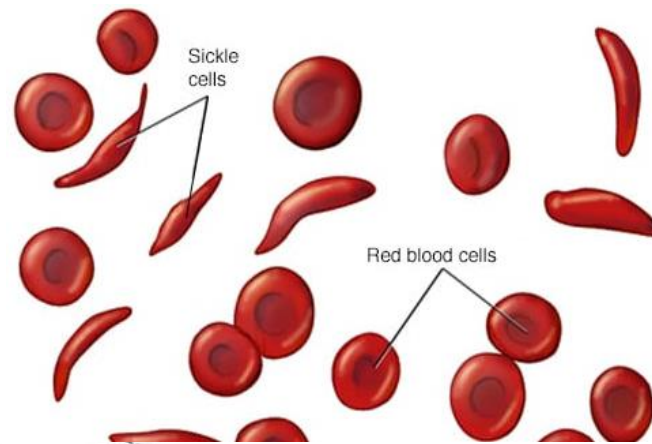
GENOTYPE TO PHENOTYPE

NORMAL β-GLOBIN				
DNA.....	TGA	GGA	CTC	CTC.....
mRNA.....	ACU	CCU	GAG	GAG.....
Amino acid.....	Thr	Pro	Glu	Glu.....
	4	5	6	7
MUTANT β-GLOBIN				
DNA.....	TGA	GGA	CAC	CTC.....
mRNA.....	ACU	CCU	GUG	GAG.....
Amino acid.....	Thr	Pro	Val	Glu.....
	4	5	6	7

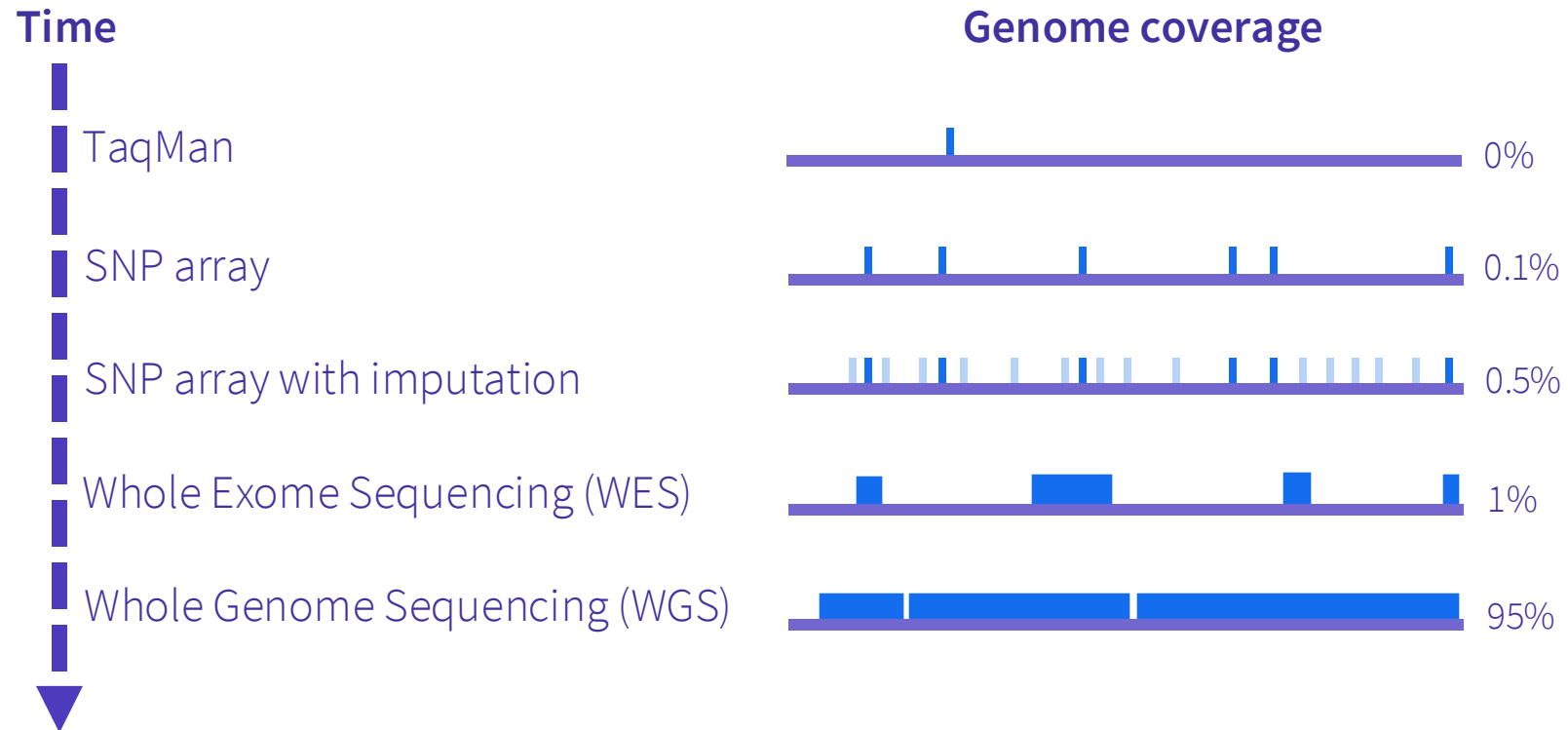
© 2019 Pearson Education, Inc.



© 2019 Pearson Education, Inc.



GENOMIC COVERAGE



GENOTYPING VS SEQUENCING

[illegible]

GENOTYPING VS SEQUENCING

Genotyping

AGCTGAAGGAGTGTGGCCATCTGCTCCACCTCCCGCGGACCCCTACTCTCAGGACCTCTGCAGCACCCAACTGGAAGTGGCCGCTGCAGACCCAAAGACGAGGGGCACGCGGGAGCCG GCAGCCCTAGTGAGCGGTGGAGATGTTGAGGTGGGAGGGT CACCC
AGGTGGGGT GAGGCT GGGGTAAGTAGCGAGTGAACGGCTTCGAAGCTCTGGGCCGCCAGGTGGACTAAGCAGGCGCTCTGTCTTCGCCGCCCAAGGTGGGCGTCTCTGAGGACTC CCGCCACACCT GACCCGAGACGCGCGCCAGCCTAGAAGCCTTCCC
CGACCCAGCGTAGGGCCGCGCGACTGGCGCGCAGGGGCGCGGGGAGGCTTGGCAACCCGGGGCGGGAGCAGCGCGGCAAGGCCGCTGCGCAGCGCCGCTCT GCGCAGGCGGCTCCGC CGCGGCGGAGGGATACGGCGACCATATAT ATATCGCGGGCGC
AGACTCGGCTCCGGCAGTGGTGCTGGGAGTGTCTGGGAGCGCGTCCGTTACTCTAGT CAGGCGGCGCGCAGCGGCGGCGCGCATAGCGCACAGCGCGCTTAGCAGCAGCAGCAGCAGCAGCGGCATCGGAGGTACCCCGCGCTGCAGCCCCGCGCTGGTG
CAGCCACCTCGCTCCCTCTGCTCTTCTCCCTTCGCTCGACCA TGGTAGGTCGGGAGT GGCAAATGCGGCGTAGCAGCTGCCGAATTTCTCCAGATTTCTAGTT GTT TTTGTTGTTTT TIGTTT GTT TTTGTTCT TGGAGGTTTTTCTTT TCTGAGTGTIACGCA GCA GCTG
CGCTTAAAGGAGGTTGCA TTTTGGA TTTGCA TCTCGGCGACCTCTGCCAGGGAGCTT CAT TTA TTTGTTCCCTTGGAGCTGGA CTTGGTCGT AGGCCGTCCACGGGCAAGGGGCTCCGGCCGCAA CTGCAGCGGGGTTTCTGCAT CCAATCCCCCTGCCCGGCCAGCCCCG
CACCCACTGCATCCACTAGCGCGCACCCGGGCTGCTGCAGCGCAGCGTT TCGGCTGGGAGCGGGCGGGGCGGCGGCACTAGACCCCGCCCGGCGCCCTCCCAACCCCGCTTCTCCG CGGCGCGCAAGGTGGCAGGTGCGGCGGGCAGTGGAATGAATGGGCT
GGA GCTGGCCGGTGGCGCACA TTTGTCGGCGCGGTGTTGAGGGCGCAGT CAGCGCCGCGCA CTTCCCACTTTGCGCGGCGCTGCTGGGCGCCCTCTCTCGT CGCTCT CCGCTCTTCTTCCCGGGGCGCGGCGCGGCGT GGGCTGGGA AGGAAGGAGCCGCGGAA
GGGTGGGGTTGGGGG CAGGAAGGCGAGGGGTTGGGGCGGAGAGGGCGGAA GCGGCGGCCGGGCCCTGCGCCGGGCGGGGCGGCGCTGCGCCGGGCGGGGCGGCGCTGT TCTGCGCTTTGCGAC CTTGCGGCCCCACCCAGTG CAGCAGTGCGGGCGGGCGTGAGC
CTCGGTGCA C CAGGA GGC ACT TCCCGCGGAGGCGCTGGGCT CGCGCTAAT TGGGCGGGGGGGGGGGGGCGGCGGGGAGGAGGAACTGGCGCGCGGCTTGTT TCCATT AGA GACGCA AAGTT TCTGCT CCGGGAGGA GGC GGC GCGGCCGCGGGCTCGT CCGCTGG
GGGAGCAGAAGCGGGTGGGAGGTGCGGGTGCC TTGGCTCAGCCCTGCTGCGCGGGGCGGGGGTGTTGA CCTCTTGCCGAGGAGGGGCGGCTCCAGACGCCGCT CGGGGGCGGCTT CCCCCACGCTGCCCGGGCACGCGCTGCCGGTCCCTCGCC
CGCGCCACTTCCAGTCCGCAGAGAGA TGCCCT CCACGT TCTGCTTT CTCTGCAGCTCTAGATTGCCAGATGCAGTGTGCGCT CGCTGGGTGTGT TTTCCACAGCCCTTCTCTCGGCG TG CAGGGCTGACATCAACGAGTTCGTTGTTGGCGGGTGGGAGATG
GTTCCCGCAGGGTT CTG GTACACTTTTGCCCCAAGGTTAGCGCATTTGGGGGAGGAGGTTTTCGTT GTCGAGAAAGCCCTCTAA CAA GAGAGTTCTGTAGCGAG GTGGGACTGTTCTCCCATAGAGTGACAGT TTTCT TGCAGGTGTGGCA
GCGCTTCT GTT GTA CAA GACAGATGTGCTT GCGGTT ACGTAAATCATCTGTCTCCGTCATTT AAAGAAAGCCAA TTTTATGTGATT GAGGTAGAAAGAAAGATCCGTTTATAA TTTGTA AA AAAAA TTTTCACCAGAAATCAATATATTGGAACACCATTCCTA TGTGTTAA
GTTTTCACTAA GAGTAT AAAC TTCATCAGCTT TCTATTAGGACTTATTTT GTAATTGGCTTCTTAGGCATCCTT CTTAAAGAGAAATCCACGTAGCTCTCTT GAGGTCTCGAGTT CCGTCGGCTGGAGGCAGAGTT CAGTGGAGACCAAATAATGCAGGTGAATTACCTTCG
TGGCCATTACTGCTCCAACGAAGTGT GTT TATTAAGAACAGT TCTTA TGT CAT TCT TAAGGTAGGTAGGGT TAATACTCTCCAGCAAATTTAGTAGATACTTTTGCAGAAAAGAGAGGAGTA TATATA GTT TGA TAATTA TTGTGT AGT TTTCTGTGT ACTTAATTTTGCAGTT
TTGTAAACATTCATT TGT AAGATGGTACCATTT TTTCTGGCTTCTGAATCATAGGA TAGTTTGACCCAGGGCATTAGCCA TTGTAATGGTAGGCTTTT AAAAA TAACTGCT AATTTAAAGGA TTGAAAGCATTTGTTACATGGAAATGAAGTTGGTGGGTACCCAGTTGCTG
TAT CTT TATTTT TTCTACTTAAT TTTTCTATAAAATGGATATAAAAGCCTGT TAATCCAACCAATGCCATTATGTAAAGAGTTTGGAGATTCTCAGGGGCTGAGCAGTGCAGAGT GC GCTGAAAGCCTGCCCTGGATGAGATCTTATCTCGGCTGTGATGGCAGTGG
CAGTGGGCTGGGTCCTTGTGAGTGGAAAGGGGACTGCGGTGTCCATGGTGCAGTAGGTGGCGCTCTTCTGCTTAGAGCCTGCCGCCACTGCA GCTGCTGCCAAGGGGCTTCTGCCACTAG AGGTGCCATTTTTCACATGATGAACCTAGCCTAGTTAGA TCGCAGAGCA
AGCTGT AAGCCATGGGCCAGAAAAGAAAAGTGAAGTGAGCAGATGTTGCTACTTCTT GTAATCCTT TGT TAAATAGCATAAGGAGTTTTTCTTATCTATT TACTTTCAATAAGTACCGTGTCTACAGGT TTCAAAGGA TTTTAAAGTGA TTTTGTGAAA GAT CACAATATTAA
AAGTATAACTGGAAAACCTAT GTT GAAATCAACCAACA TGT CGTGA CTGAAT GATAACCTT TCTTTCTCATATA GGCTGATCA GCTGACCGAAGAACAGATTGCTGGTAA GTT GACAACCTCAAGGAGTCCCGAAGGCCAGAACTAGGCACTGACTCAGTTTGTGTGAC
TCCTCT GTT CCTCCCCGTACAGTCTGGGCAGT TTTCTAAGAATT TATTTAAATAAGAACAGTAAGCAGAAACACTGAGGTGAGATGTATTCTTGCCA GTA CTTTATAGATGAGGT GAAAGGAA GTAAAA CTAAGGATGCCCATGT TAAACTCTGAGAGATTTGACCATGTTTC
ACAATGTGCAAAGTTTGCGTATGA TTAATT GTA CTGAGCCTGCTACTCAGCGGT TTAGTTTACAAT TCT TATG CCAATGGGT CTT TCA GTAATCTGCCACGAAAGCTTGTGCTCGTATCTCTAAAA TAAATGGAAATGGGT GAA TATGAGTGT TAGGACCACCTGT AGTAAITGGGAA
GAAAGTTACATT AGT TAAACTGTGTCAGGCTGGTCTTAACCTCTGGGCTCAAGCAATCTCTGCTCAGCCTCTGAGTAGCTGGGA CTA CAGGCA TGT GCCACCACGCTCTGGCAGATT TTAGCT TTTTAA TATTCTGGAGGACTGT TTTGAGACT GTT TCT CGTTAGGA
AACCAGGAATGCTTCTGAAATATTCTAAAA GTCATGTGGAGAGAGTTT ACCTGGGAATGT ACA TTTCTAGTAACCATTTTATTTGTTATGAAA CAAGGGATT TTTATGGCT TTAGAAATGTAA CA GGAAGGGATTGAAGGGGGGCATGTGA CCAATCTTGT CAGATTGGA TTTA
GTCCCTTGAACCTGGGAGGCGAGGGTTGTAGTGAGCTGAGATTGCACCACTGCACTCCAATCTCGGTGACAGAGCGAGACTCCATTGTTAAAAA AAAAAAAGAGTTGATTAGGACTAATT TAAGCATGT TCCAGCTTAGCCGCTTGAAACCTTTGGGAATATTGTGGTG
TGTGGCACTGTTTAT TGGGAGCAGTGTGCTTTATGGGCTGCTGTATGAAGGCCAGTCCAAAGGAGACTATTGTGGTCAATTATTCAGTAGATAAA GACCACTCTGAT ACGTTGCACAACCTGAATGGCTGGCT TGGCAAGCCCCGCGCAAGTGTGATT GTGACTGGGTTG
GATAAA GACATT GAT TCT AACGGGTCAACTTTGTTTTAGAAAT CAA GGAAGCCTTCCCTATTTGATAAGAT TGGCGATGGCACCAT CACAACAAAGAACTTGGAACTGGAACCTGCTCAT GAGGTCACTGGGTGAGA ACCCAACAGAGCTGAATTG CAGGATATGATCAATGAAGTG
GATGCTGATGGTAAGAGCTTTAAACCATGAATGAGG CCAATTGTGTGAATT CAA GTT CAGACATGTTACAGGATTGCTTTTCAGGTCCCAGAGCA AAGCAAATGTGCAAGATCCTTTCTG TGGTTGCCCGAGGGCCATTGA CAA TAAAA TAGAAGATGATGGGCTTGC
GTCCATCTGCTTAGTGTCTA GAATGT TTTCTGCATGGGATCACTATTGTTTTCTCTGCTTGGTGCGACCTAGAGCTCAAAATCTATTTTT TTTTT TTTTGGAGAGCGGAGTCTCGCCCTGT CGCCCAAGCTGGAGTGGCACTGGCGGATCTCGGCTACTGCAACCTCTGCCTC
TTGGGTTCCAGCGATTCTCTGCGTCAGCCTCTGAGTAGCTGGAATTACAGGCGTGTGTGCGCACGCCAGCTTAGTGTTT TGTATCTTATAGTAGAGATGGGGTTTACCATGT TGGCCA GGTCTGTCTCAAACCTCTGACCTCGT GAT CCGCCCTCCCGGCTCCCAAAGT GCTG
GGATTACAGGCGTGAACCACTGCTCTTGCCGAGCTCAAAGCTTTTATCAATGCCCCATGAGTCTGCATGAGTCTT GAGGGGGGAGGTGAAATTAATAGCCATAGAAAGTGTCTTTAAACA ACTTACTGTGTTTAAAGA GGAGGAGGAACCCCAAGTGAAGT AGGTGAC
GAGCACTCTTAGAAGTTACATAAAGTGA GTA CAGTGTGAGCTGATGTGTTTGTGTCGAGAGGAGCTGTAGGT TGGAGGCGGATGTTGTGGTGACTCCAGGGGATA GTTTG CAGAACCTAACCGAAAGGAAGCTGT AAGGTG CAGGGCCAGGGCAGGAA CCAAGCA
TAA CCGTATAGCGGTCTGTACTCTGTTCTCTGACTCTACAGCAGCGGACAACAGAACTTTGATTGCTGATTTCA TCA GTAAGCAGGCTTGAAGCACACTTCCCCACCCCTAAAAA AAAAAACAGCTATTT TGGTAAATCCTATATATA TTCTAATGTACTGTATGACAGTATAG
AACATGATT TTTAAAGA TGA GTT GGGAGGAGAAAA GGATAAAGAAAAA TAAAGAAGCAATTAAGAA TAAACAATT CGGATCTAGATTTTACTTCT AGA TGA TTGACTCGAGGGTGGTGTAGTAAATATCGCTTGTCTGGT CACAACATTTTGGCAGCAGAGCTT TGAATTA
GGTTCTTTGACAAGCCTTCAGCA CGTTAGAGTGGT TTTCACTAATAGTGT TTTGGAAGAAAAAGGTTGTCATAGTTCTCTAGTTTGTCTAAGATGATCAGTACCCAGGAACGTGGAGTAACTT CTTCTTGT TTTGTTGGAGCCCCGGGAATCTGTGCCTGGGAGGGGGA GAA GT
CTGTAGGCTCTTGATTGTGTGGAAGAAGGAGAAGTTGTGCCAGGCTACAGAACTCTGT GTT TGCAGTGAAGAAACAGGATGGTACCTGACCTTCTGCA TGGCTGTGA GATAGCTTA AATA ATTTCTTTT GTT TTTGATGAATAT GAA CAATATCTTAAATTTT GAGGCTA
AAAAAGTCTTGAAGGATCCCTGAGGTATTTTCTTT GAAAGGTACTGGTGAAATAGAGTAACCTAA CTAAGGGT TTTTCTTTCTAA TTTTATTTCCATTTAGTTTCAATGACACTGTAGTCT GGAGTGTGT CTTTGGGGGATTTCACTCTT AGT TTTAAAGAGGAGTTTGTG
AGTACTGGCCGTAGAACA GATTGT TCTGACAGTTCCCTAAGT GTT ACT AGTCTGAGCTGTGAGAATGCTCTGAGCTTTTCCCTTAATGGGAAATAAGATACTGAGTTGGAAGAAACAGGTGGCTAACCATCATAGCTGGCCAAGAAATGATCTGGAGAA GACTTGGTAA
GACTTCATGGCCCATGCATGGCAT AACAGATCAATGTTCTCTCTATAA TCTTTTCTCTGAAACACTTTATACACTTAACCTGCAGCTCAGTTCTAGGCTTTTGTGT TACTGCTGTCACTAACCAAGGCAGAGTGA GACCTGAGT GATTTCCTTAACCTAGGGATGGCA GT
CGGGGGCGCTTTCTTCCCTCGGAAGTGGAAAATTCAGCCTGCGGAGTGGTGTATGCTATTTTCTCTTGAACGTACAGCCCTTCATGACCTTCCATGGGCTTGAATCCAGATGTGCAGTTTCTTTGTATAA TTAATACTATCTGGGCACTGATGATGAGTTTGAATATGT
GAAATTGCCCTGTGAAGTGT

GENOTYPING VS SEQUENCING

[illegible]

Genotyping

WES

GENOTYPING VS SEQUENCING

Genotyping

WES

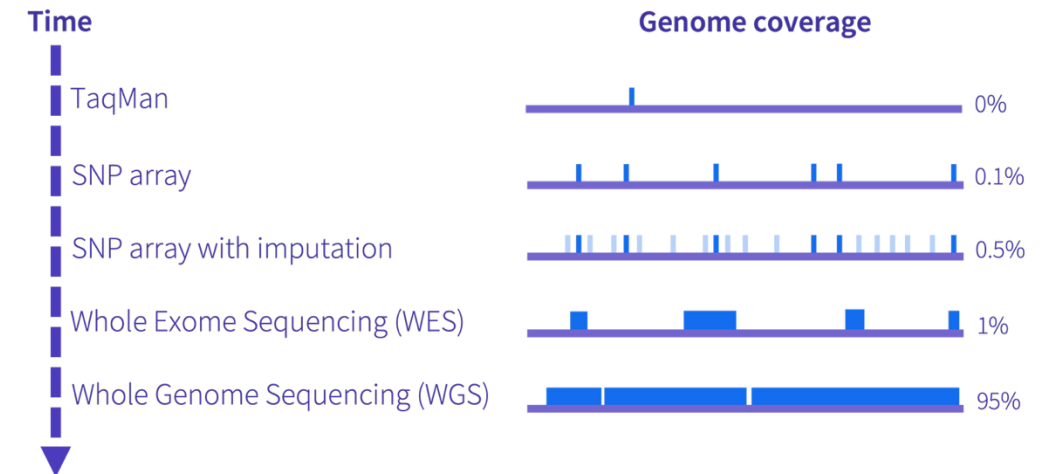
WGS

[illegible]

WHICH TECHNOLOGY?

The choice of technology for detecting single nucleotide polymorphisms (SNPs) depends upon the application.

For GWAS / PGS we use ‘*SNP array with imputation*’

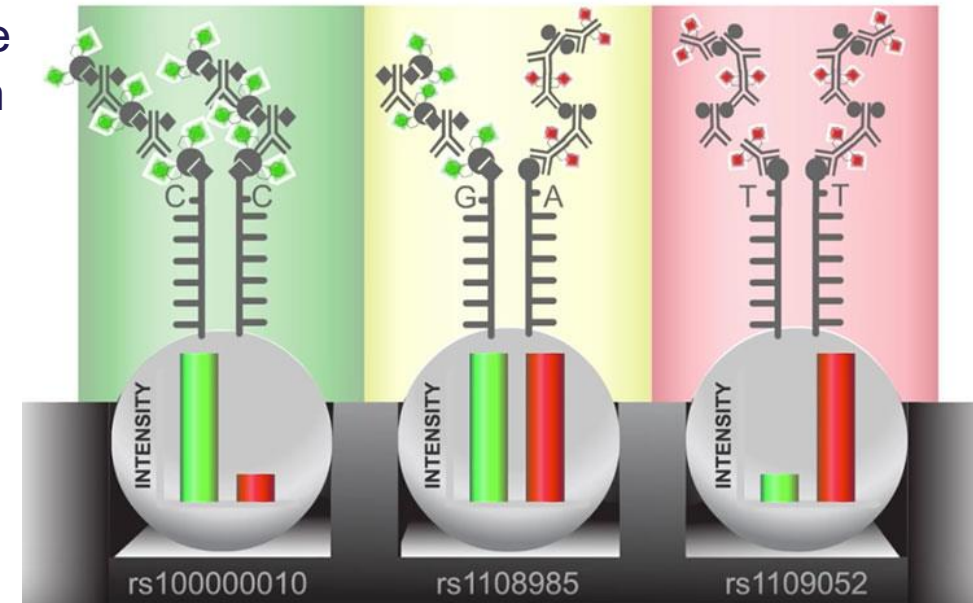


GENOTYPING

For GWAS/PGS we don't need to analyse all 3 billion positions in the human genome.

Instead, we can measure about 500,000 to 1 million genetic variants that act as markers across the genome.

Because nearby variants are often inherited together, we can use these markers to predict (“impute”) millions of additional common variants — up to about 10 million in total.

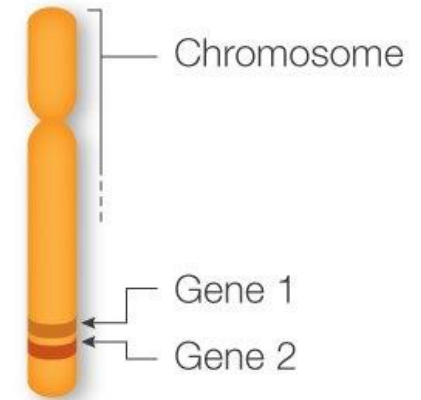


WHY DON'T WE NEED TO TEST EVERY VARIANT?

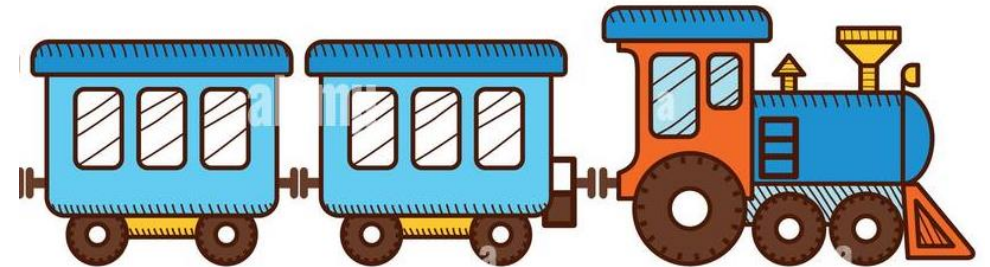
The human genome has about 3 billion positions.

But many variants are inherited together.

Variants that tend to be passed on together are said to be in linkage disequilibrium (LD).

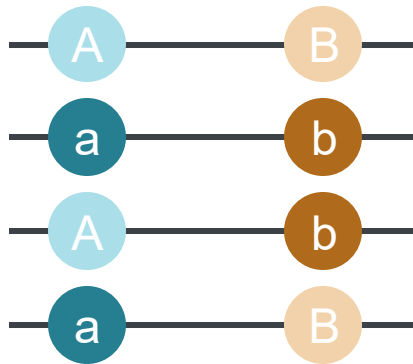


Linked loci
(Physical proximity)



Linked train wagons

WHAT IS LD?



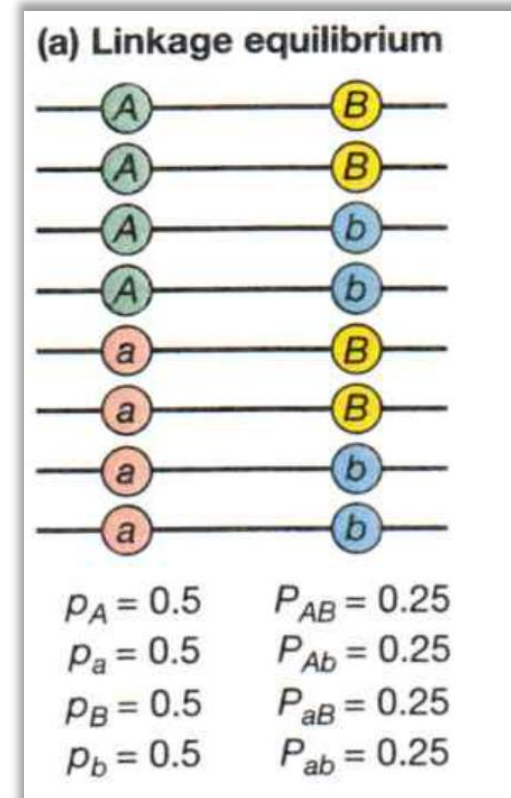
If there is random relationship among alleles at the two loci then the frequency of the haplotypes will be the product of the frequencies of the two alleles:

$$P(AB) = P(A) \times P(B)$$

$$P(Ab) = P(A) \times P(b)$$

$$P(aB) = P(a) \times P(B)$$

$$P(ab) = P(a) \times P(b)$$

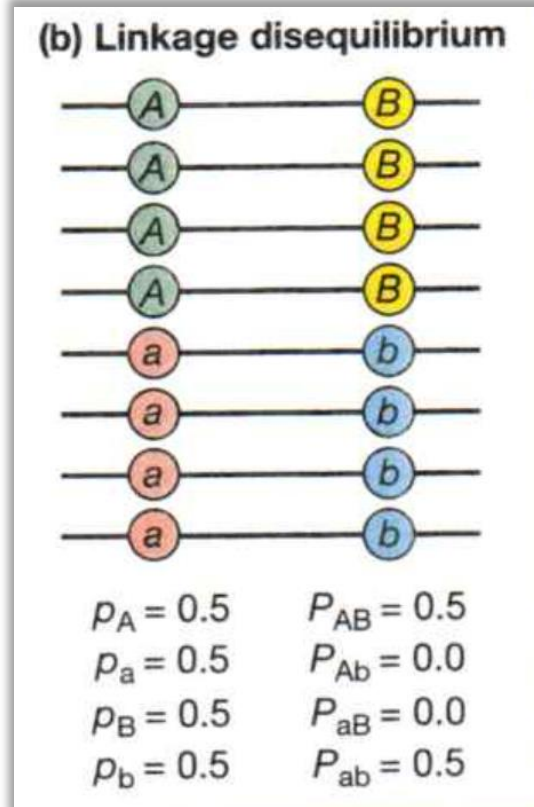


WHAT IS LD?

LD is just a way of describing how strongly two variants are correlated. In practice, we use r^2 to say how well one variant predicts another

$r^2 = 1 \rightarrow$ perfect correlation (completely linked)

$r^2 = 0 \rightarrow$ no correlation (completely independent)

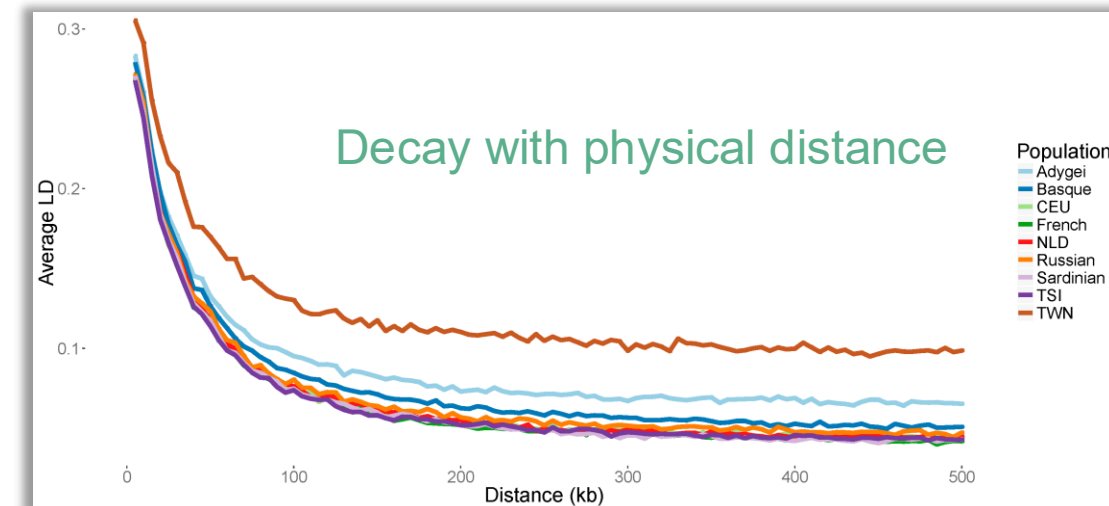


When the association between alleles at two loci is **non-random** they are said to be in **linkage disequilibrium**

The degree of LD can be measure in several ways – the simples one is:

$$D = P_{AB} - P_A P_B$$

If $D=0$, no LD, if $D>0$ LD



IMPUTATION

USING HAPLOTYPES

The true haplotypes

A	T	C
G	C	A

This individual has inherited a chromosome with alleles A-T-C from one parent, and G-C-A from the other parent

We observe only the genotypes

A/G T/C C/A

Genotype data does not carry information about the haplotypes.

We do not know whether A at SNP1 is coming from the same parent as T or C at SNP2

Different haplotypes

A	C	A
A	C	C
A	T	A
A	C	A
G	C	A
G	C	C
G	T	A
G	T	C

Phasing = estimate the most likely haplotypes

IMPUTATION

Genotypes

Study sample

...A...A...A...
...G...C...A...

Reference haplotypes

CGAGATCTCCCTTCTTCTGTGC
CGAGATCTCCCGACCTCATGG
CCAAGCTCTTTTCTTCTGTGC
CGAAGCTCTTTTCTTCTGTGC
CGAGACTCTCCGACCTTATGC
TGGGATCTCCCGACCTCATGG
CGAGATCTCCCGACCTTGTGC
CGAGACTCTTTTCTTTGTAC
CGAGACTCTCCGACCTCGTGC
CGAAGCTCTTTTCTTCTGTGC

From a sequencing study

Study sample

...A...A...A...
...G...C...A...

Reference haplotypes

CGAGATCTCCCTTCTTCTGTGC
CGAGATCTCCCGACCTCATGG
CCAAGCTCTTTTCTTCTGTGC
CGAAGCTCTTTTCTTCTGTGC
CGAGACTCTCCGACCTTATGC
TGGGATCTCCCGACCTCATGG
CGAGATCTCCCGACCTTGTGC
CGAGACTCTTTTCTTTGTAC
CGAGACTCTCCGACCTCGTGC
CGAAGCTCTTTTCTTCTGTGC

Study sample

cgagAtctcccgAcctcAtgg
cgaaGctcttttCtttcAtgg

Reference haplotypes

CGGCCCCCGGCAATTTTTTTT
CGAGATCTCCCGACCTCATGG
CCAAGCTCTTTTCTTCTGTGC
CGAAGCTCTTTTCTTCTGTGC
CGAGACTCTCCGACCTTATGC
TGGGATCTCCCGACCTCATGG
CGAGATCTCCCGACCTTGTGC
CGAGACTCTTTTCTTTGTAC
CGAGACTCTCCGACCTCGTGC
CGAAGCTCTTTTCTTCTGTGC

GENETIC PARAMETERS

- ▶ Separate V_G and V_E
- ▶ Heritability
- ▶ Genetic correlation



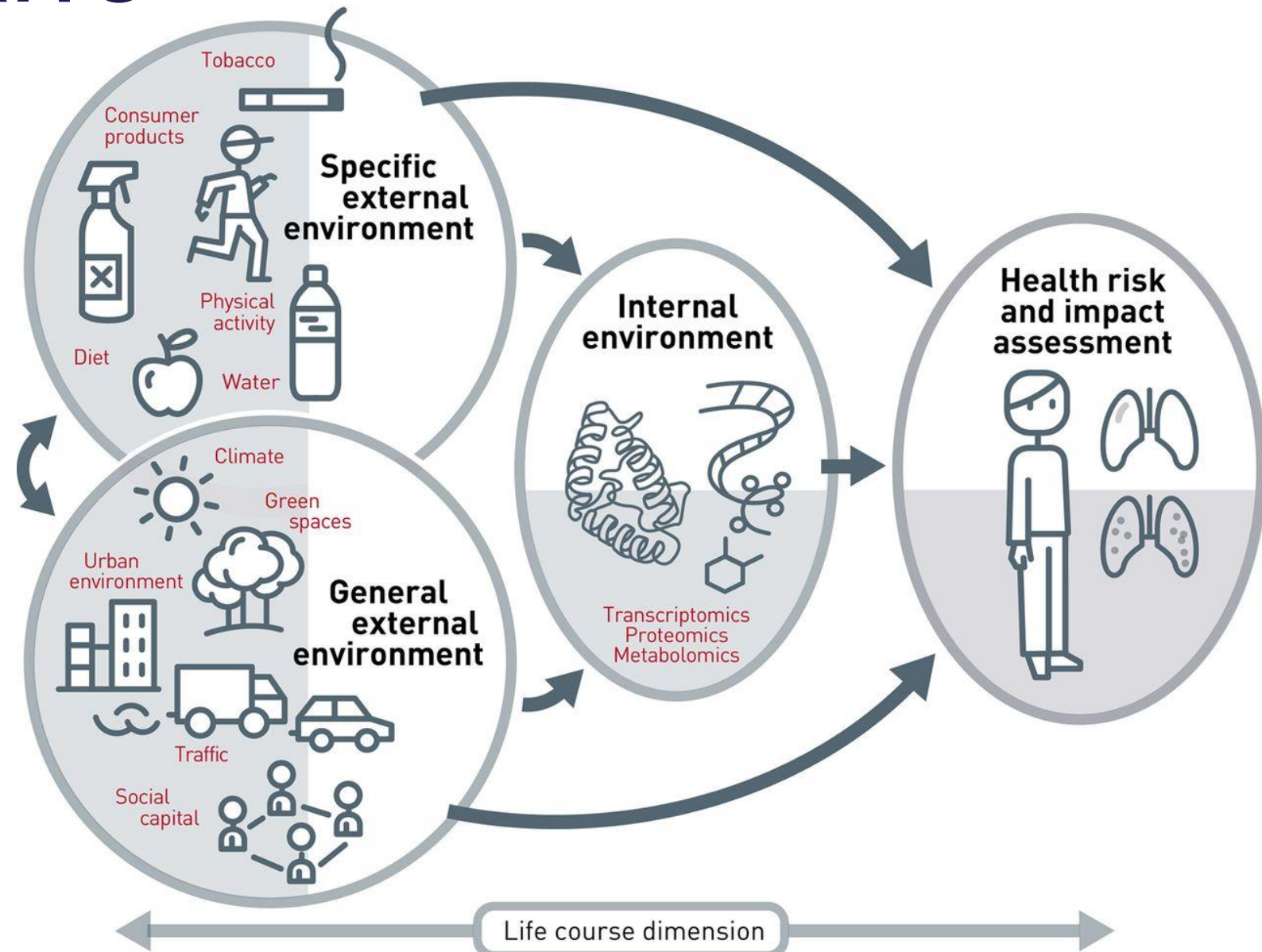
COMPLEX TRAITS are complex...

Complex traits have a **genetic component** and an **environmental component**

The relative genetic contribution is called **heritability**.

$$V_G / (V_G + V_E)$$

Thus, set a limits for
the genetic predictor



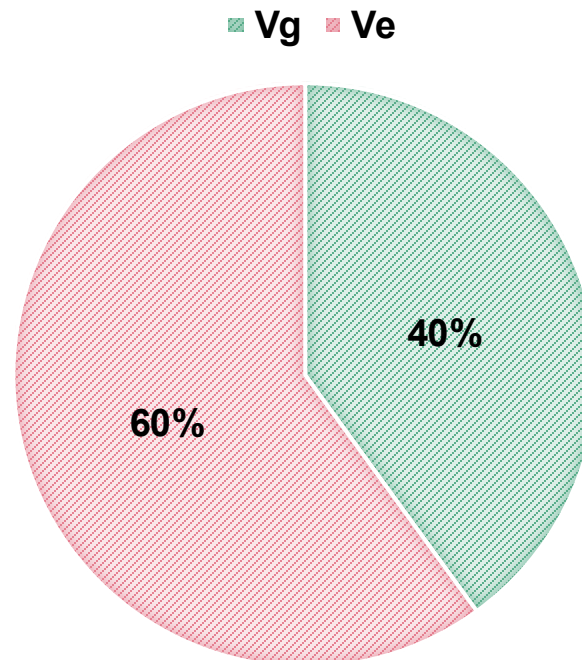
MULTIFACTORIAL TRAITS

Multifactorial traits = polygenic effect + environmental effect

many genes/alleles

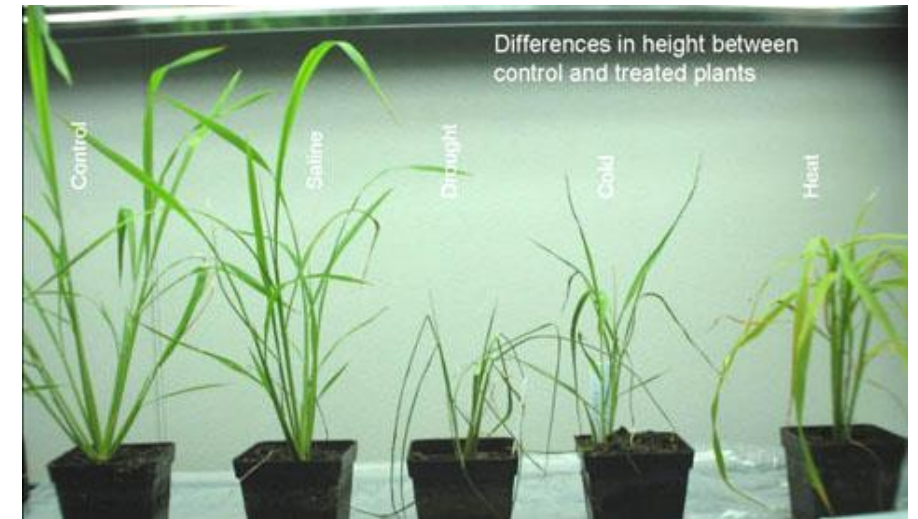
$$V_P = V_G + V_E$$

$$H^2 = V_G / (V_G + V_E)$$



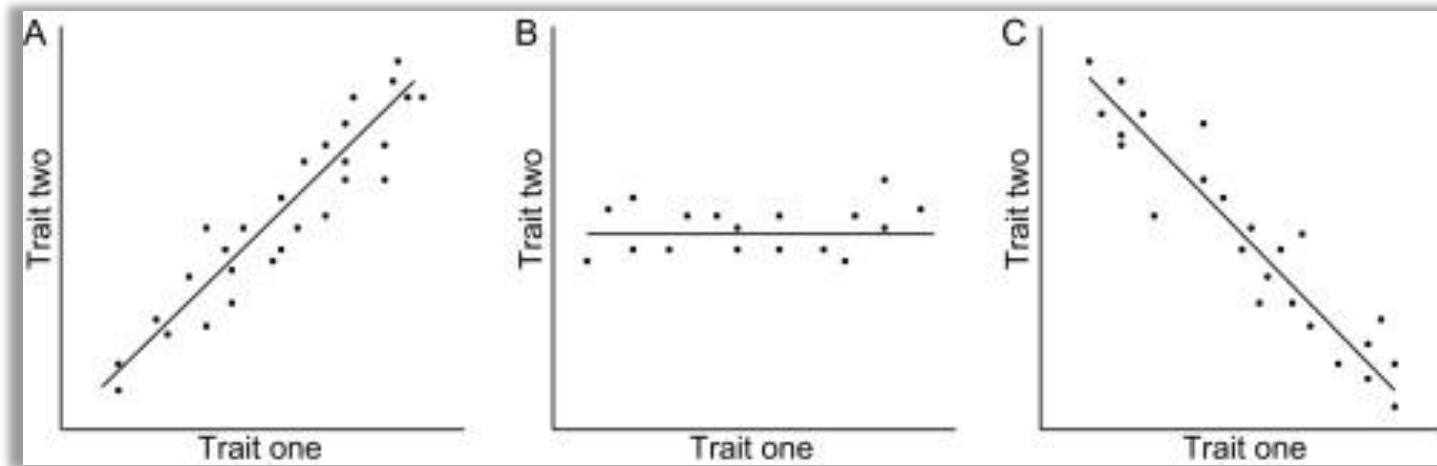
External effects that modulates the phenotypic value.

... or things that we cannot explain



GENETIC CORRELATION

- The **genetic correlation** refers to the genetic link between two traits, which can help elucidate the shared biological pathways and/or causal relationships between them.



Genetic correlation quantifies how much two traits share in their genetic basis — not because of environment or chance, but because *the same DNA variants affect both*.

SESSION 1

► Complex traits

- *Most diseases are influenced by many genetic and environmental factors — understanding their complexity helps us predict risk more accurately*

► Genetic variation

- *Individual differences in DNA underlie why people respond differently to diseases and treatments.*

► Haplotypes and LD

- *Patterns of correlated variants (LD) allow us to identify genetic regions associated with disease risk even when the causal variant is unknown.*

► Genetic parameters

- *Measures like heritability and genetic correlation help quantify how much of disease risk is due to genetics and how traits are biologically connected.*



GROUP WORK

PART 1 (15 min)

Make 3 groups

- ☐ Each group work with one exercise
- ☐ Explain the concepts & discuss how these concepts relates to identification of genetic risk variants

PART 2 (15 min)

We make 3 new groups with 1 person from each of the previous groups

- ☐ Present key messages from each exercise





BREAK

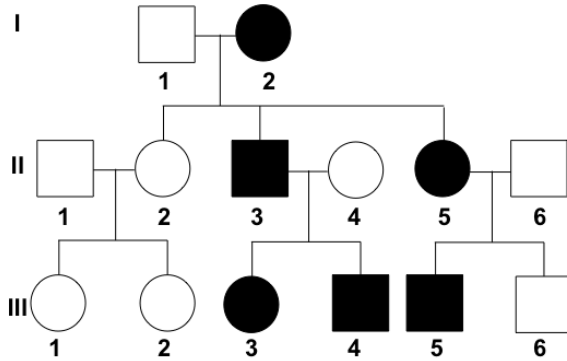
AGENDA

- 09:00 – 09:15 Welcome
- 09:15 – 09:45 Session 1: Genetic variation and heredity
- 09:45 – 10:15 Group work 1: Genetic variation and heredity
- 10:15 – 10:30 Break
- 10:30 – 11:00** Session 2: From SNP to GWAS to disease risk
- 11:00 – 11:40 Group work 2: Construct your own PGS
- 11:40 – 12:00 Common discussion and evaluation

SESSION 2

- ▶ Genetic associations
- ▶ Polygenic score

Monogenic disorders



Mutation



Each genetic variant is **both** necessary and sufficient

Linkage analysis
(pedigree)

Common complex disorders



Each genetic variant is **neither** necessary nor sufficient

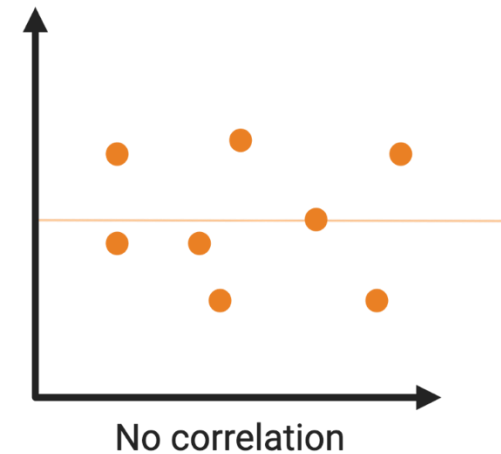
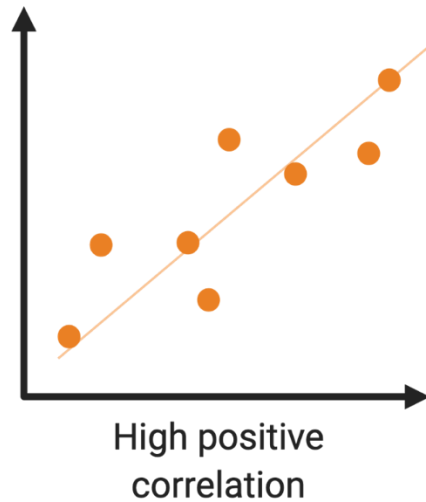
Association study
(unrelated population)



What does association mean

ASSOCIATION

An association defines a relationship between two entity objects based on common attributes.

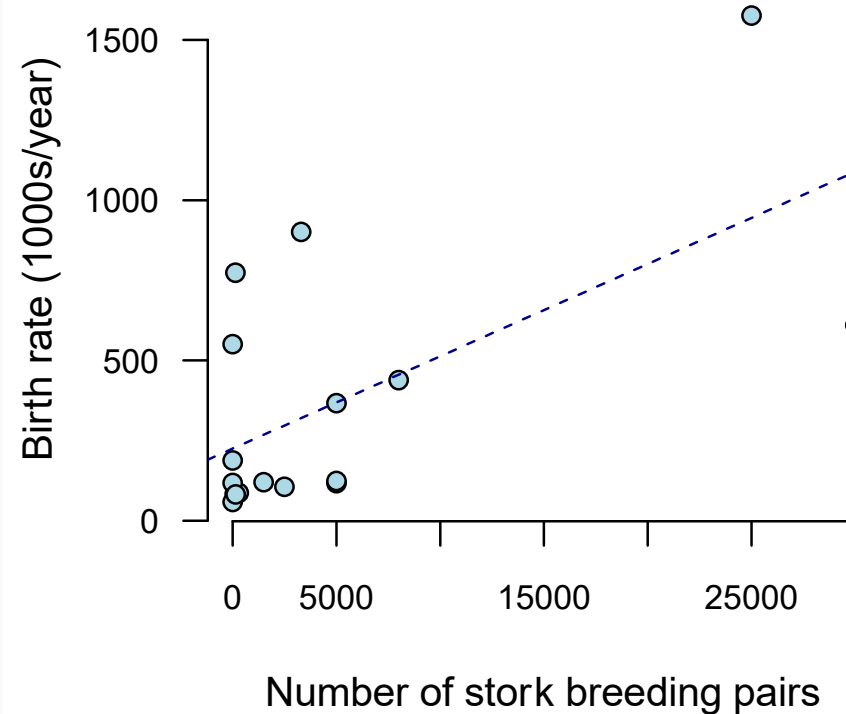


Is there an association between **exposure** and **outcome**?

ASSOCIATION *NOT CAUSATION*

Country	Area (km ²)	Storks (pairs)	Humans (10 ⁶)	Birth rate (10 ³ /yr)
Albania	28,750	100	3.2	83
Austria	83,860	300	7.6	87
Belgium	30,520	1	9.9	118
Bulgaria	111,000	5000	9.0	117
Denmark	43,100	9	5.1	59
France	544,000	140	56	774
Germany	357,000	3300	78	901
Greece	132,000	2500	10	106
Holland	41,900	4	15	188
Hungary	93,000	5000	11	124
Italy	301,280	5	57	551
Poland	312,680	30,000	38	610
Portugal	92,390	1500	10	120
Romania	237,500	5000	23	367
Spain	504,750	8000	39	439
Switzerland	41,290	150	6.7	82
Turkey	779,450	25,000	56	1576

Table 1. Geographic, human and stork data for 17 European countries



Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.250e+02	9.356e+01	2.405	0.0295 *
Storks	2.879e-02	9.402e-03	3.063	0.0079 **

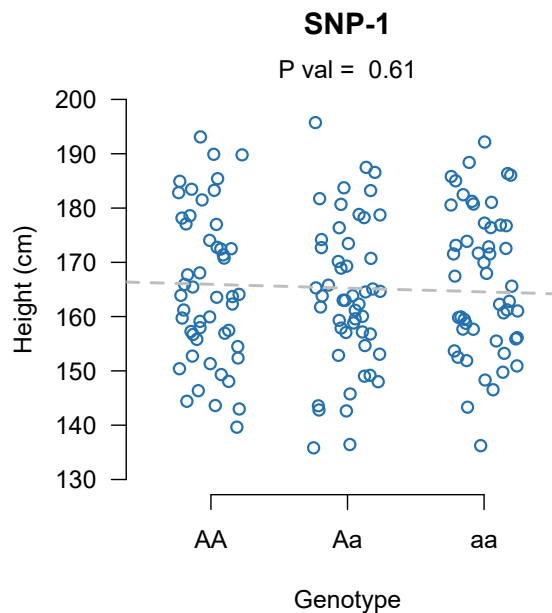
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1



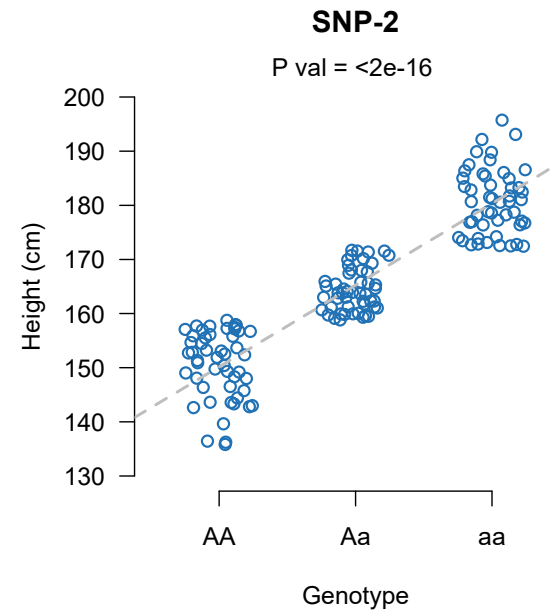


IDENTIFY RISK VARIANTS

GENOME-WIDE ASSOCIATION STUDY (GWAS)

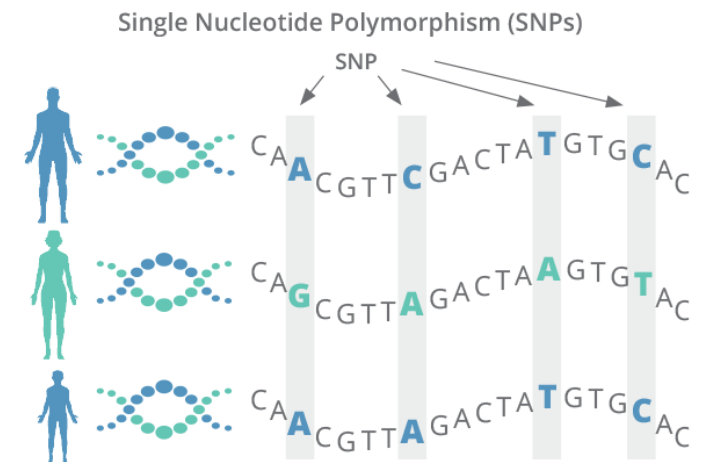
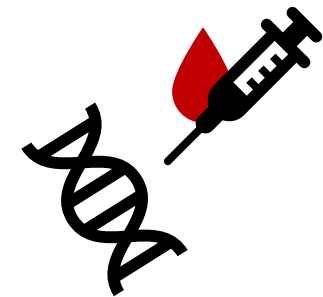
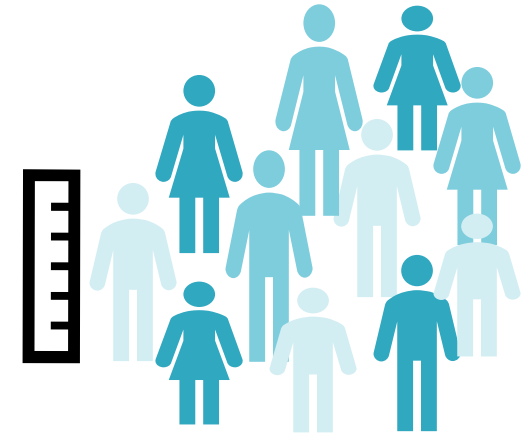


No association



Association

Systematic hypothesis-free scanning of all
common genetic variants



MANHATTAN PLOT

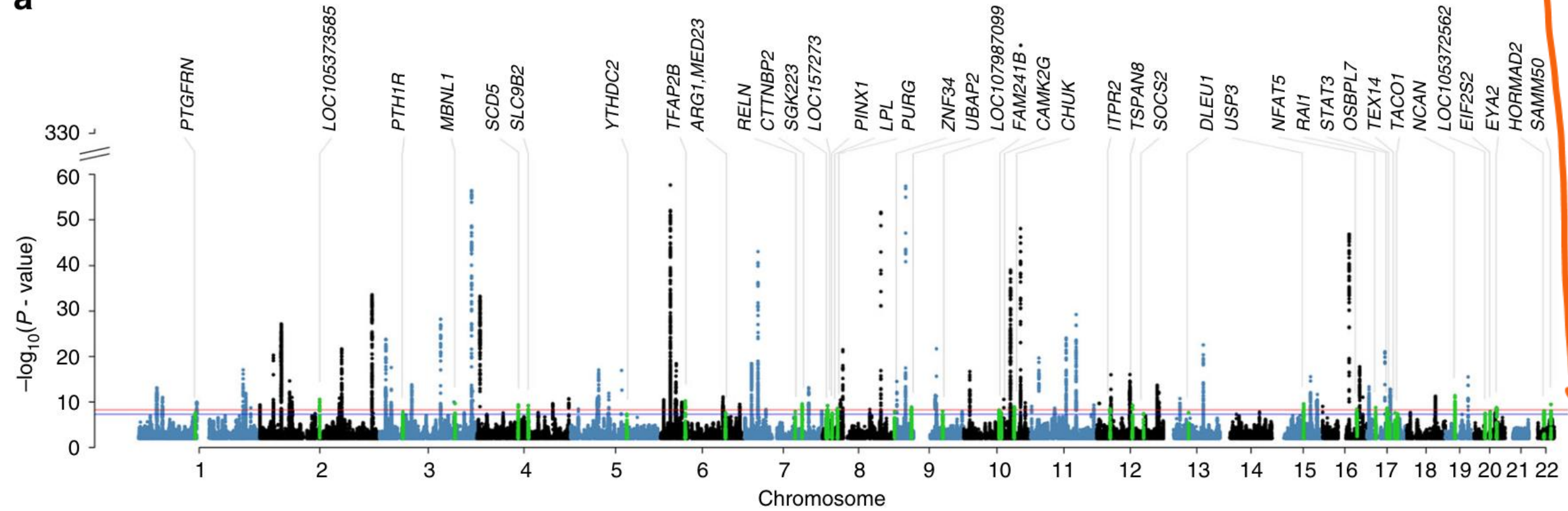
For each variant, plot the $-\log_{10}(\text{P-value})$ as function of chromosomal position.

$P=0.05 \rightarrow -\log_{10}(0.05) = 1.3$

$P=0.001 \rightarrow -\log_{10}(0.001) = 3$

$P=0.000000005 \rightarrow -\log_{10}(0.000000005) = 8.3$

a



HYPOTHESIS TESTING

	Null hypothesis (H_0) is true	Null hypothesis (H_0) is false
Reject null hypothesis (H_0)	Type I error α False positive	Correct outcome True positive
Accept null hypothesis (H_0)	Correct outcome True negative	Type II error β False negative

The probability (P) of making a type I error is denoted by α ; we reject the null hypothesis if the inferred P value is less than the significance level ($\alpha=0.05$). I.e., the probability of rejecting the null hypothesis when should be accepted.

Why multiple testing correction? If we test 500,000 SNPs, then by chance we expect 25.000 SNPs to be significant (if $\alpha=0.05$) $\rightarrow$ i.e., **25.000 false-positive associations**.

One solution is to correct for number of tests performed; Bonferroni correction;

$$\text{Corrected } P\text{-value} = P \times n_{\text{tests}} \leq 0.05 \text{ OR } \frac{\alpha}{n_{\text{tests}}} = \text{new significance threshold}$$

MANHATTAN PLOT

For each variant, plot the $-\log_{10}(\text{P-value})$ as function of chromosomal position.

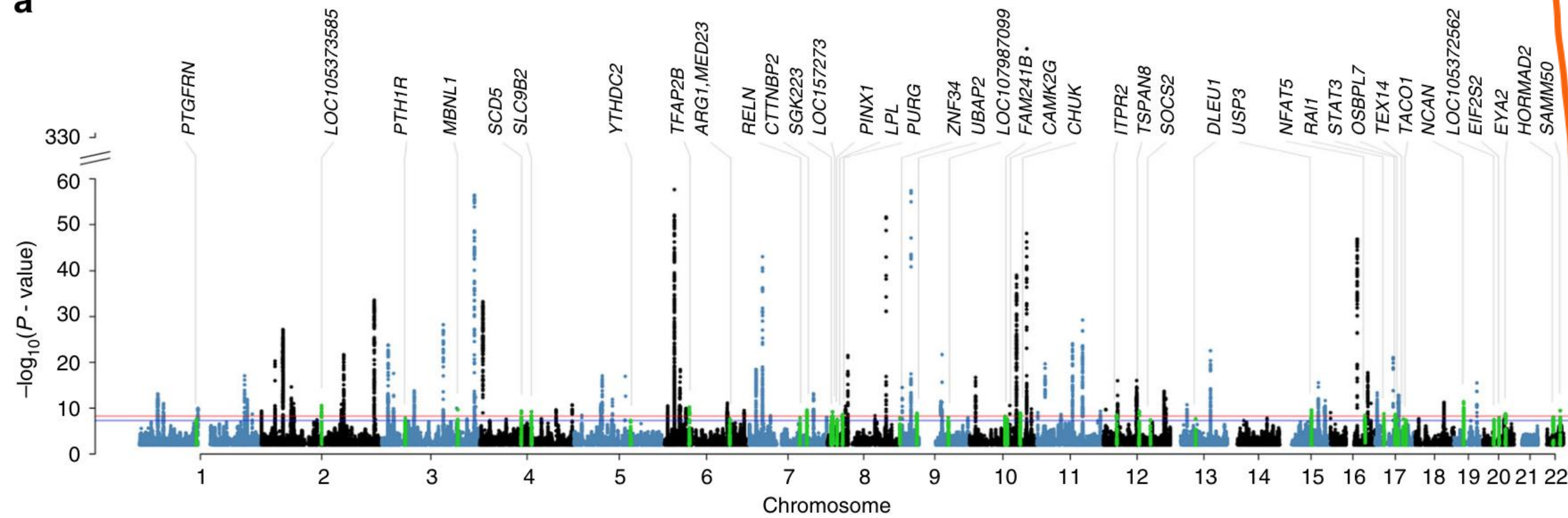
Genome-wide significance level?
-adjust for no. of independent statistical tests
(1,000,000 independent genomic regions [LD])

$P=0.05 \rightarrow -\log_{10}(0.05) = 1.3$

$P=0.001 \rightarrow -\log_{10}(0.001) = 3$

$P=0.000000005 \rightarrow -\log_{10}(0.000000005) = 8.3$

a

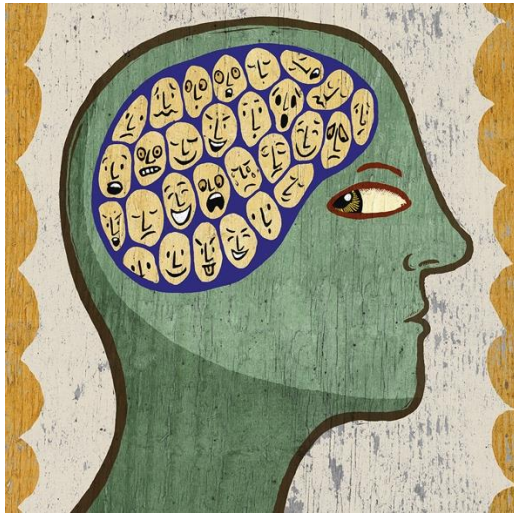


POWER IS EVERYTHING IN GWAS

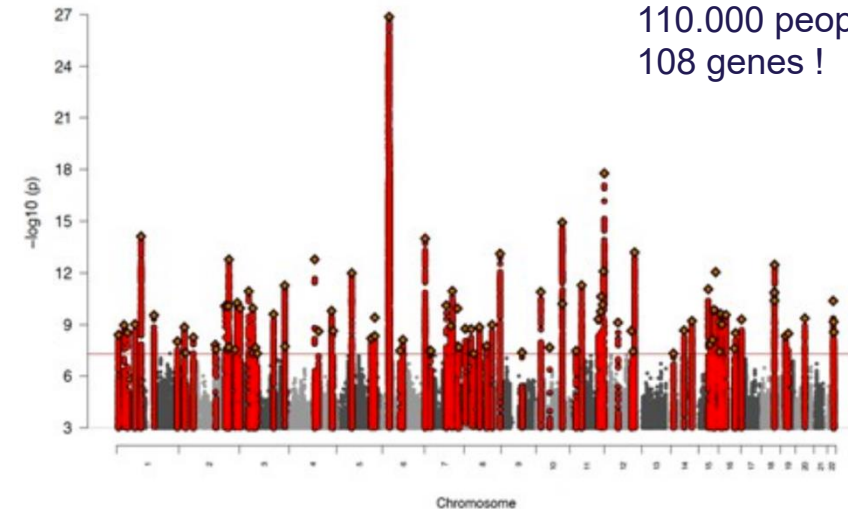
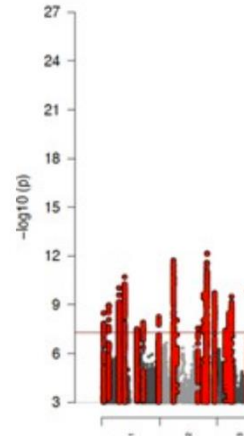
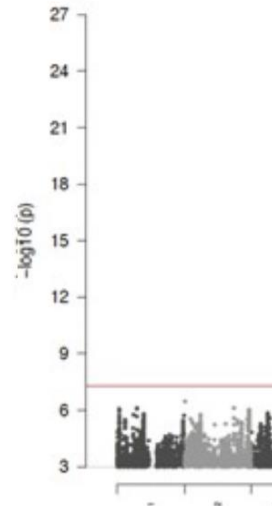
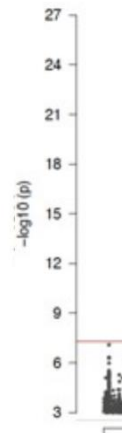
the probability of detecting an effect, if there is a true effect present to detect

depends among other things on sample size

Schizophrenia



Has high heritability ($h^2 = 0.85$)
The population prevalence is 1%
Emerge in late teens
Molecular aetiology is unknown

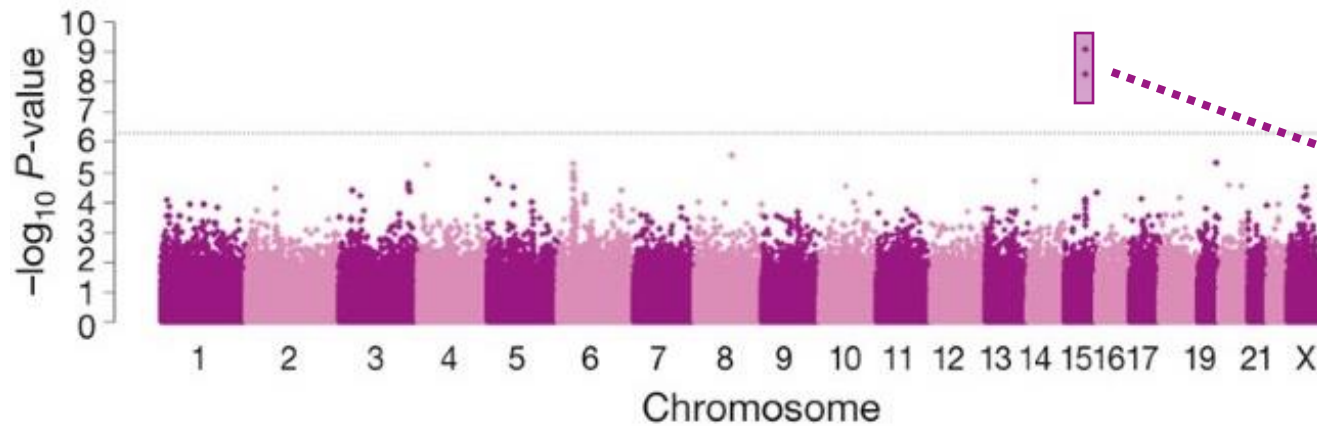


110.000 people
108 genes !

320.000 people
287 genes !

LETTERS

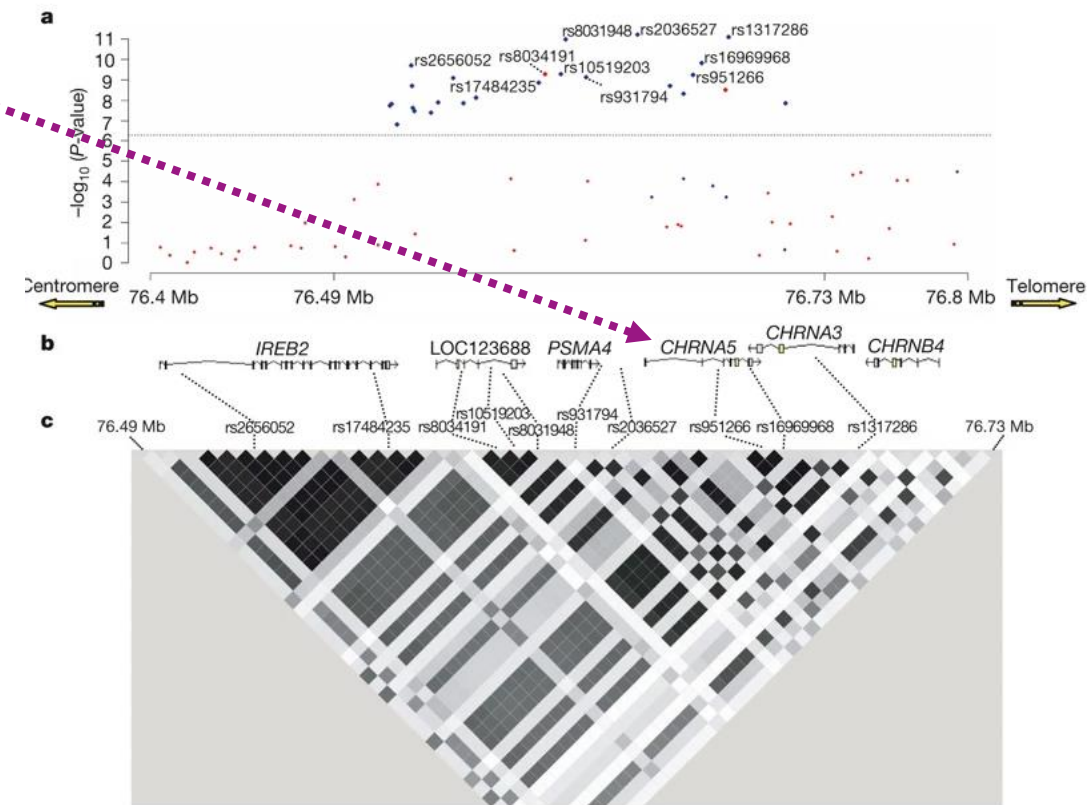
A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25



A genetic study of lung cancer (LC, 1989 cases and 2625 controls) found a nicotine receptor (*CHRNA3/5*) to be associated with the risk of developing lung cancer.

Does that then mean that *CHRNA3/5* is a risk loci for LC?

Risk loci for nicotine addition → addicted to smoking → increase LC risk



PREDICTING DISEASE RISK FROM GENETIC DATA

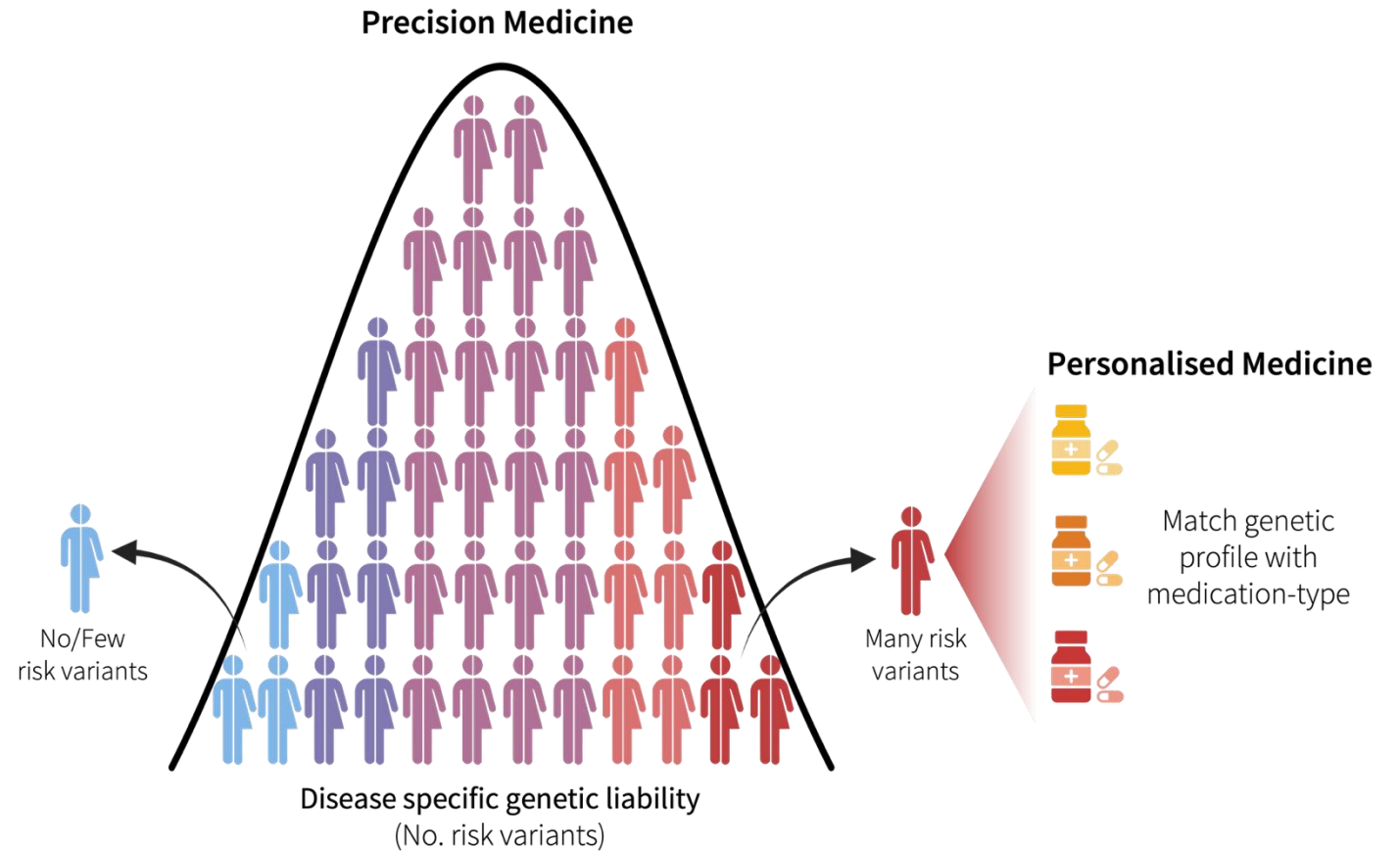
A “polygenic score” is one way by which people can learn about their risk of developing a disease, based on the total number of changes (i.e., SNPs) related to the disease (NHI)



DIFFERENT NAMES

BUT THE SAME

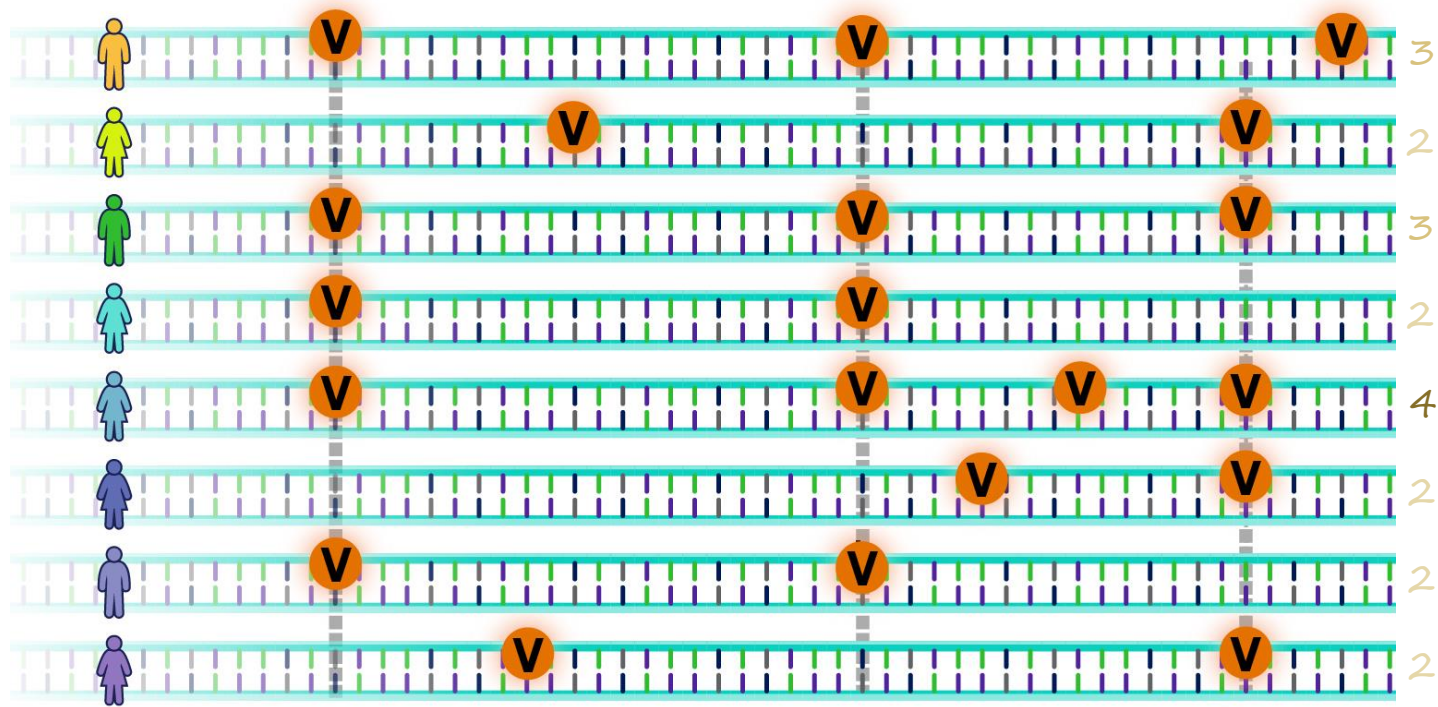
- ▶ Polygenic risk score (PRS)
- ▶ Polygenic score (PGS)
- ▶ Genetic score (GS)
- ▶ Genetic risk score (GRS)
- ▶ Genetic value
- ▶ Genetic liability
- ▶ Breeding value
- ▶ ...



THE INHERENT DISEASE RISK



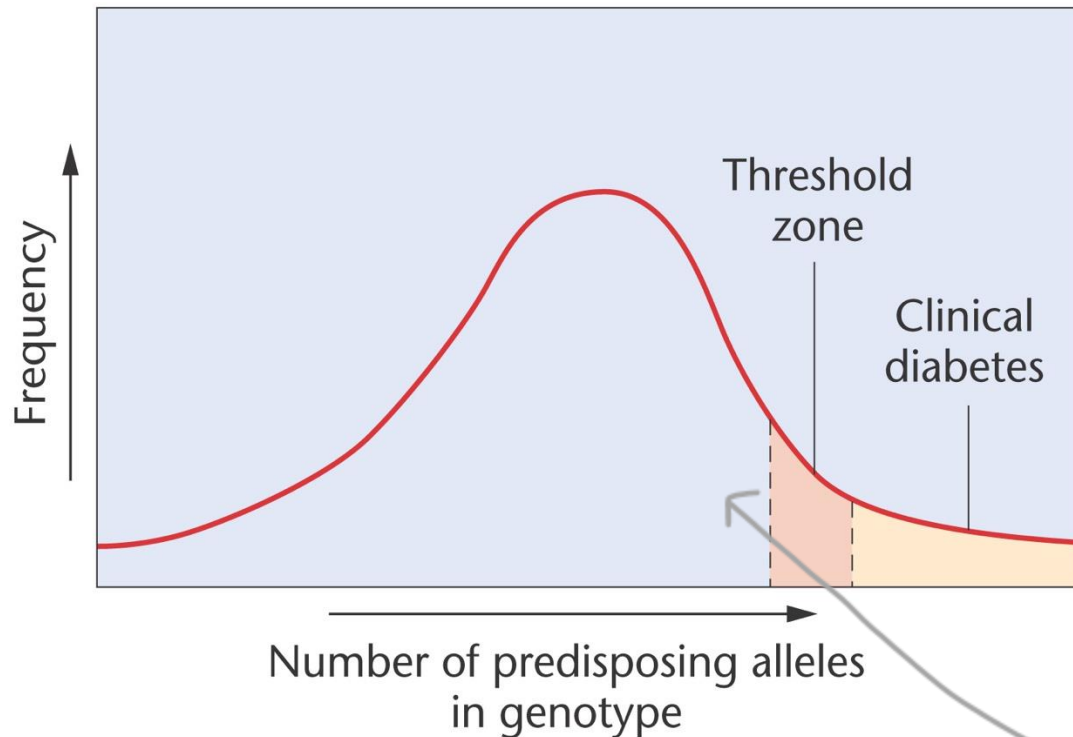
THE INHERENT DISEASE RISK



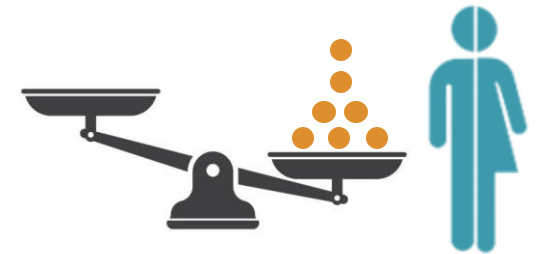
Variation in the
number of risk
variants
→
Polygenic score
(PGS)

WHAT IS A PGS?

A RELATIVE RISK



SNP1	A	G
SNP2	T	T
SNP3	A	C
SNP4	T	A
SNP5	C	C



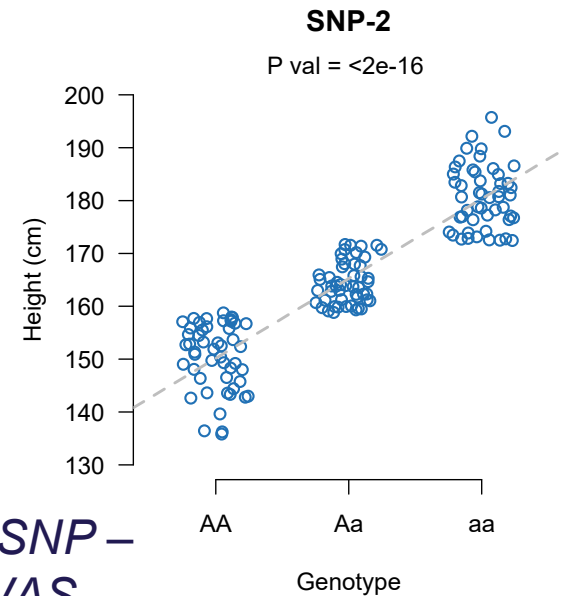
PGS = 7

How do we find the 'orange' alleles?

WHAT IS A PGS?

“A PGS combines information from large numbers of markers across the genome (hundreds to millions) to give a single numerical score for an individual’s risk for developing a specific disease on the basis of the DNA variants they have inherited.”

b is the slope (effect size) from regression



The effect size of the SNP – obtain from the GWAS

$$PGS = \sum X_i b_i$$

The genotype of the individual for SNP i (0, 1, 2 – counting the number of the alternative allele)

AA = 0

Aa = 1

aa = 2

$$PGS = \sum X_i b_i$$

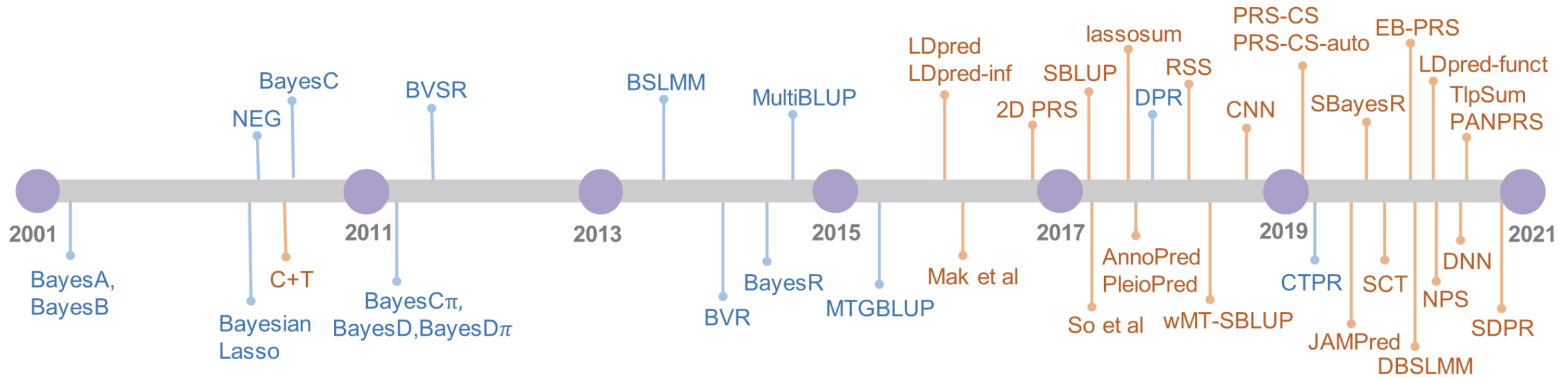
HOW TO COMPUTE A (simple) PGS?

SNPs	Adams Genotypes	Ref allele	Alt allele
SNP-1	TC	T	C
SNP-2	GG	G	T
SNP-3	CC	A	C
SNP-4	TG	T	G
SNP-5	AA	A	G



A LARGE PALETTE OF PGS METHODS

$$PGS = \sum X_i b_i$$



WHY DIFFERENT PGS SCORING METHODS?

Complex traits have different underlying genetic architectures

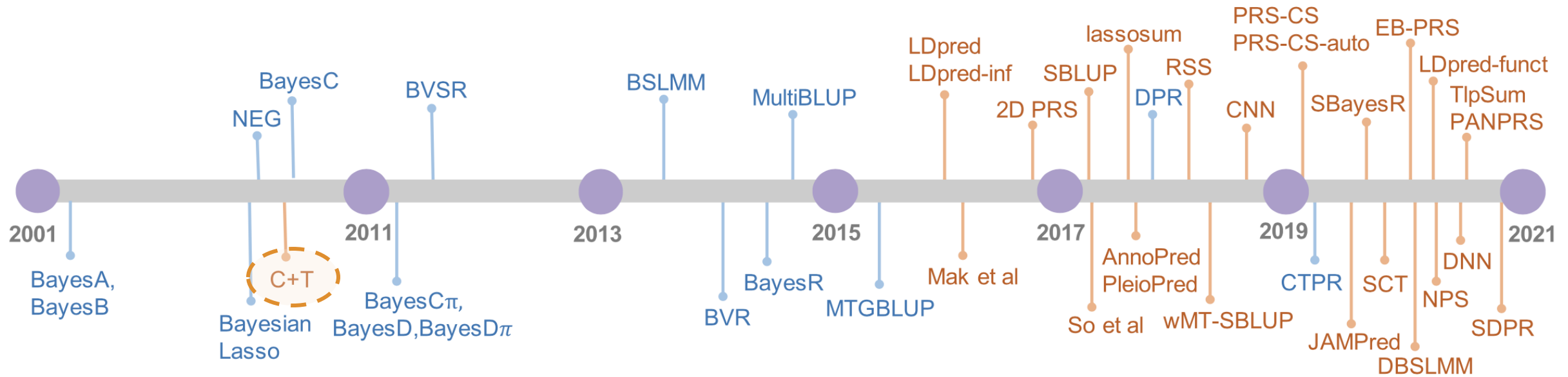
- ❖ some are influenced by <100 genetic loci
- ❖ some are influenced by >1000 genetic loci
- ❖ some loci have very small effects
- ❖ some loci have moderate effects
- ❖ correlation structure among loci (linkage disequilibrium)

The different scoring algorithms attempts to account for this.



A LARGE PALETTE OF PGS METHODS

$$PGS = \sum X_i b_i$$



CLUMPING AND THRESHOLDING (C+T)

0: Set LD (=0.8) and P values (0.01)

SNP	b	p
1	0.21	0.005
2	0.22	0.0048
3	0.25	0.0003
4	0.1	0.04
5	0.05	0.15
6	0.02	0.49
7	0.03	0.87
8	0.12	0.003
9	0.14	0.0034
10	0.18	0.0004
11	0.21	0.00003
12	0.12	0.15
13	0.14	0.12
14	0.03	0.84
15	0.02	0.32

1: Sort by P -value

SNP	b	p
11	0.21	0.00003
3	0.25	0.0003
10	0.18	0.0004
8	0.12	0.003
9	0.14	0.0034
2	0.22	0.0048
1	0.21	0.005
4	0.1	0.04
13	0.14	0.12
5	0.05	0.15
12	0.12	0.15
15	0.02	0.32
6	0.02	0.49
14	0.03	0.84
7	0.03	0.87

2: Compute LD and select variants based of thresholds

SNP	b	p	r^2
11	0.21	0.00003	
3	0.25	0.0003	0.96
10	0.18	0.0004	0.93
8	0.12	0.003	0.88
9	0.14	0.0034	0.74
2	0.22	0.0048	0.4
1	0.21	0.005	0.03
4	0.1	0.04	0.04
13	0.14	0.12	0.05
5	0.05	0.15	0.03
12	0.12	0.15	0.04
15	0.02	0.32	0.01
6	0.02	0.49	0.01
14	0.03	0.84	0.01
7	0.03	0.87	0.01

1st variant in LD-pair

Have $LD > r^2$ – ignore those

CLUMPING AND THRESHOLDING (C+T)

0: Set LD (=0.8) and P values (0.01)

SNP	b	p
1	0.21	0.005
2	0.22	0.0048
3	0.25	0.0003
4	0.1	0.04
5	0.05	0.15
6	0.02	0.49
7	0.03	0.87
8	0.12	0.003
9	0.14	0.0034
10	0.18	0.0004
11	0.21	0.00003
12	0.12	0.15
13	0.14	0.12
14	0.03	0.84
15	0.02	0.32

1: Sort by P-value

SNP	b	p
11	0.21	0.00003
3	0.25	0.0003
10	0.18	0.0004
8	0.12	0.003
9	0.14	0.0034
2	0.22	0.0048
1	0.21	0.005
4	0.1	0.04
13	0.14	0.12
5	0.05	0.15
12	0.12	0.15
15	0.02	0.32
6	0.02	0.49
14	0.03	0.84
7	0.03	0.87

2: Compute LD and select variants based of thresholds

SNP	b	p	r ²
11	0.21	0.00003	
3	0.25	0.0003	
10	0.18	0.0004	
8	0.12	0.003	
9	0.14	0.0034	1st variant in LD-pair
2	0.22	0.0048	0.98
1	0.21	0.005	0.96
4	0.1	0.04	0.96
13	0.14	0.12	0.52
5	0.05	0.15	0.34
12	0.12	0.15	0.10
15	0.02	0.32	0.04
6	0.02	0.49	0.01
14	0.03	0.84	0.01
7	0.03	0.87	0.01

Have LD>r² – ignore those

CLUMPING AND THRESHOLDING (C+T)

0: Set LD (=0.8) and P values (0.01)

SNP	b	p
1	0.21	0.005
2	0.22	0.0048
3	0.25	0.0003
4	0.1	0.04
5	0.05	0.15
6	0.02	0.49
7	0.03	0.87
8	0.12	0.003
9	0.14	0.0034
10	0.18	0.0004
11	0.21	0.00003
12	0.12	0.15
13	0.14	0.12
14	0.03	0.84
15	0.02	0.32

1: Sort by P-value

SNP	b	p
11	0.21	0.00003
3	0.25	0.0003
10	0.18	0.0004
8	0.12	0.003
9	0.14	0.0034
2	0.22	0.0048
1	0.21	0.005
4	0.1	0.04
13	0.14	0.12
5	0.05	0.15
12	0.12	0.15
15	0.02	0.32
6	0.02	0.49
14	0.03	0.84
7	0.03	0.87

2: Compute LD and select variants based of thresholds

SNP	b	p	r ²
11	0.21	0.00003	
3	0.25	0.0003	
10	0.18	0.0004	
8	0.12	0.003	
9	0.14	0.0034	
2	0.22	0.0048	
1	0.21	0.005	
4	0.1	0.04	
13	0.14	0.12	1st variant in LD-pair
5	0.05	0.15	0.86
12	0.12	0.15	0.82
15	0.02	0.32	0.81
6	0.02	0.49	0.85
14	0.03	0.84	0.85
7	0.03	0.87	0.81

Have LD>r² – ignore those

CLUMPING AND THRESHOLDING (C+T)

0: Set LD (=0.8) and P values (0.01)

SNP	b	p
1	0.21	0.005
2	0.22	0.0048
3	0.25	0.0003
4	0.1	0.04
5	0.05	0.15
6	0.02	0.49
7	0.03	0.87
8	0.12	0.003
9	0.14	0.0034
10	0.18	0.0004
11	0.21	0.00003
12	0.12	0.15
13	0.14	0.12
14	0.03	0.84
15	0.02	0.32

1: Sort by P-value

SNP	b	p
11	0.21	0.00003
3	0.25	0.0003
10	0.18	0.0004
8	0.12	0.003
9	0.14	0.0034
2	0.22	0.0048
1	0.21	0.005
4	0.1	0.04
13	0.14	0.12
5	0.05	0.15
12	0.12	0.15
15	0.02	0.32
6	0.02	0.49
14	0.03	0.84
7	0.03	0.87

2: Compute LD and select variants based on LD

SNP	b	p	r ²
11	0.21	0.00003	
3	0.25	0.0003	
10	0.18	0.0004	
8	0.12	0.003	
9	0.14	0.0034	
2	0.22	0.0048	
1	0.21	0.005	
4	0.1	0.04	
13	0.14	0.12	
5	0.05	0.15	
12	0.12	0.15	
15	0.02	0.32	
6	0.02	0.49	
14	0.03	0.84	
7	0.03	0.87	

3: Compute PGS based on effect sizes (b) and P -values

$$PGS = \sum X_i b_i$$

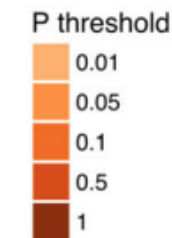
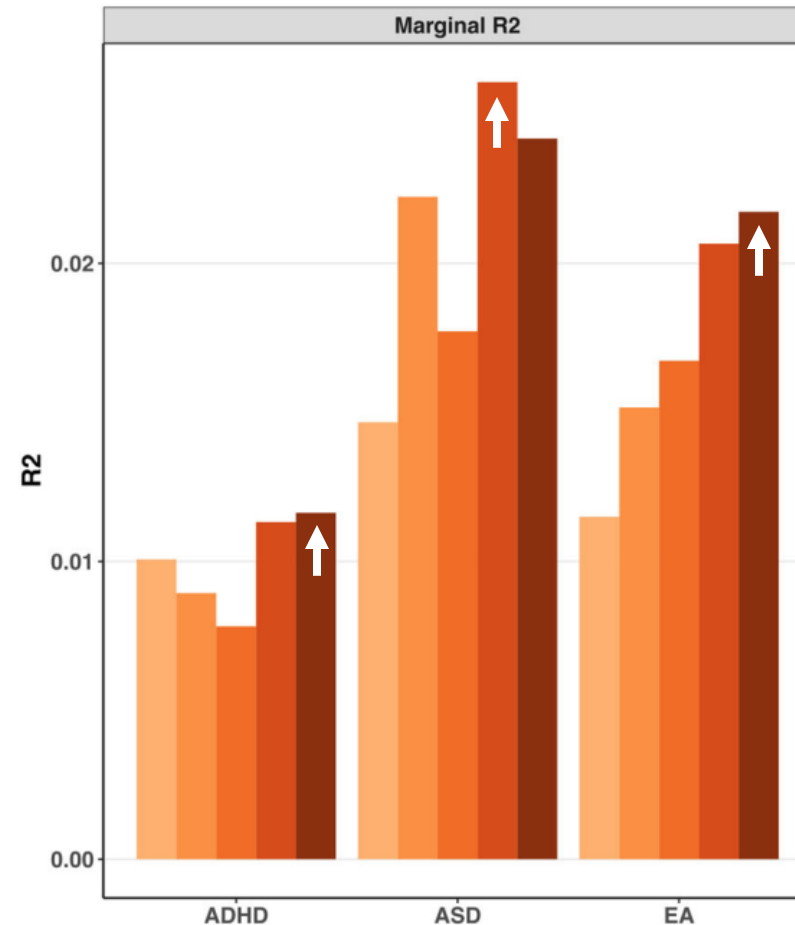
$$= X_{11} \times 0.21 + X_9 \times 0.14$$

CLUMPING AND THRESHOLDING (C+T)

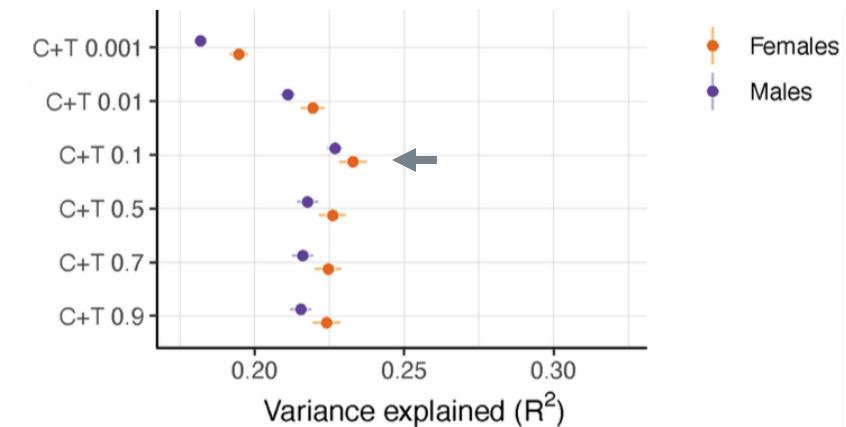
Repeat for other P -value cutoffs (and LD values)

How does the PGS associate with the disease

$$y_{\text{trait}} = \text{PGS} + \varepsilon$$



Finding optimal r^2 and P -cutoff and apply in second cohort



CLUMPING AND THRESHOLDING (C+T)

Clumping and Thresholding (C+T) and Stacked C+T (SCT)

We compute C+T scores for each chromosome separately and for several parameters:

- Threshold on imputation INFO score INFO_T within $\{0.3, 0.6, 0.9, 0.95\}$.
- Squared correlation threshold of clumping r_c^2 within $\{0.01, 0.05, 0.1, 0.2, 0.5, 0.8, 0.95\}$.
- Base size of clumping window within $\{50, 100, 200, 500\}$. The window size w_c is then computed as the base size divided by r_c^2 . For example, for $r_c^2 = 0.2$, we test values of w_c within $\{250, 500, 1000, 2500\}$ (in kb). This is motivated by the fact that linkage disequilibrium is inversely proportional to genetic distance between variants.¹¹
- A sequence of 50 thresholds on p values between the least and the most significant p values, equally spaced on a log-log scale.

Thus, for individual i , chromosome k , and the four hyper-parameters INFO_T , r_c^2 , w_c , and p_T , we compute C+T predictions

$$V_i^{(k)}(\text{INFO}_T, r_c^2, w_c, p_T) = \sum_{\substack{j \in S_{\text{clumping}}(k, \text{INFO}_T, r_c^2, w_c) \\ p_j < p_T}} \hat{\beta}_j \cdot G_{i,j},$$

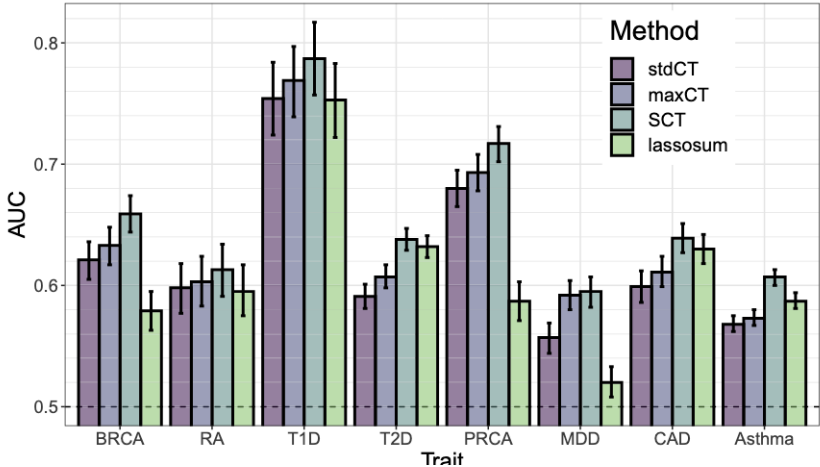


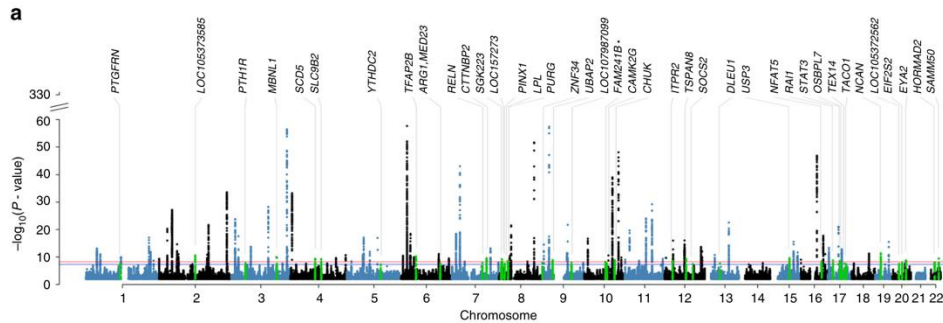
Table 2. Optimal Choices of C+T Parameters

Trait	w_c	r_c^2	INFO_T	p_T
Breast cancer (BRCA)	2,500	0.2	0.95	2.2×10^{-4}
Rheumatoid arthritis (RA)	200	0.5	0.95	7.5×10^{-2}
Type 1 diabetes (T1D)	10K–50K	0.01	0.90	2.6×10^{-5}
Type 2 diabetes (T2D)	625	0.8	0.95	1.1×10^{-2}
Prostate cancer (PRCA)	10K–50K	0.01	0.90	4.2×10^{-6}
Depression (MDD)	625	0.8	0.95	1.0×10^{-1}
Coronary artery disease (CAD)	526	0.95	0.95	3.5×10^{-2}
Asthma	2,500	0.2	0.90	2.2×10^{-4}

Choice of C+T parameters is based on the maximum AUC in the training set. Hyper-parameters of C+T are the squared correlation threshold r_c^2 and the window size w_c of clumping, the p value threshold p_T and the threshold on the quality of imputation INFO_T . Choosing optimal hyper-parameters for C+T use 63%–90% of the individuals reported in Table 1. Resulting predictions of maxCT in the test set are reported in Figure 2.

WHAT DO YOU NEED?

1. A large well-powered GWAS for your trait of interest

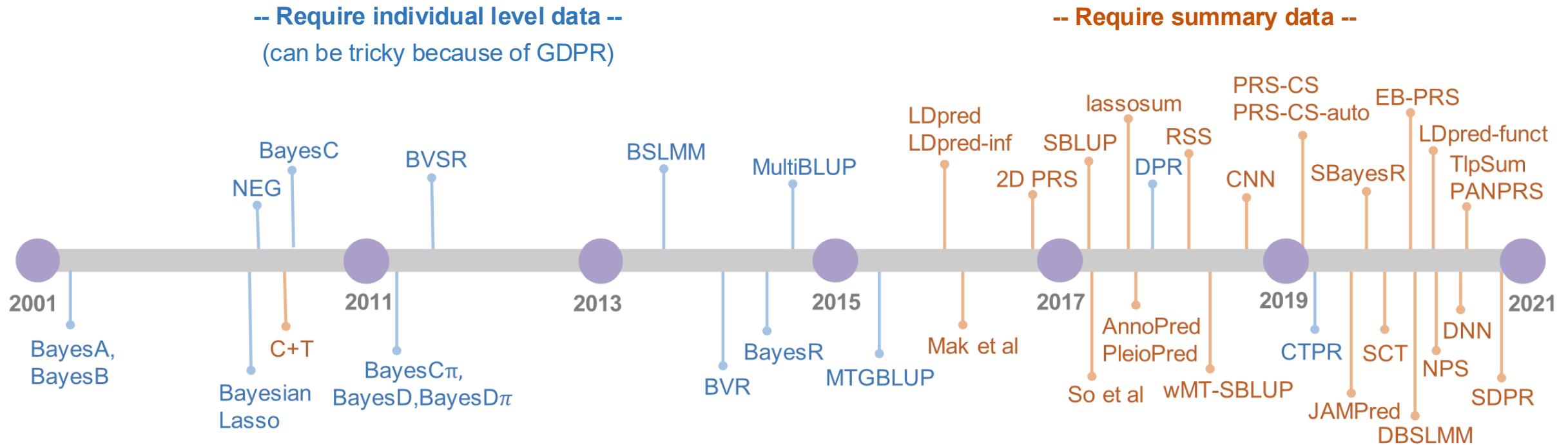


2. An independent cohort that has been genotyped

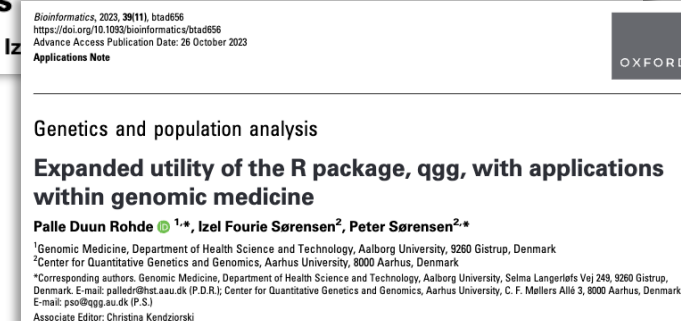
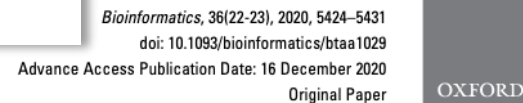
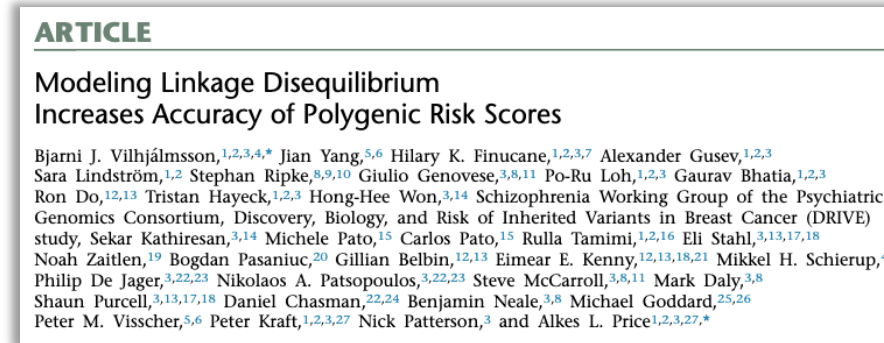
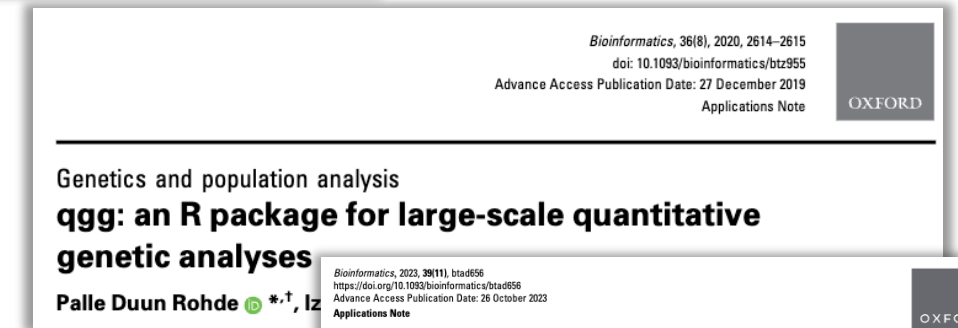
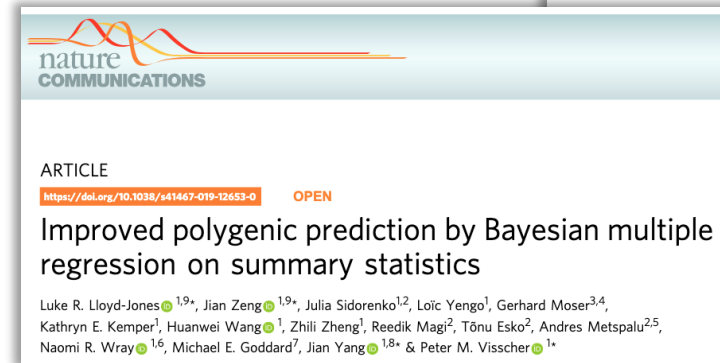
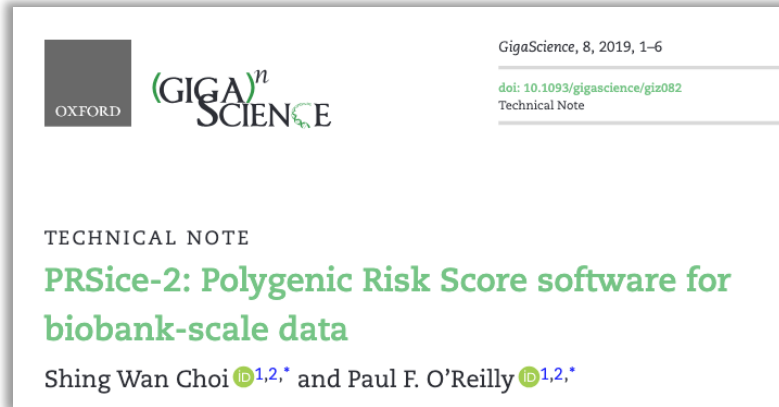


(3. That some individuals in the cohort has the phenotype)

A LARGE PALETTE OF PGS METHODS



COMMONLY USED METHODS



SESSION 2

► Genetic associations

- *By linking specific genetic variants to traits or diseases, we can identify biological pathways and potential drug targets.*

► Polygenic score

- *Combining information from thousands of variants allows us to estimate an individual's genetic risk — a key step toward personalised prevention and treatment.*



GROUP WORK



AGENDA

- 09:00 – 09:15 Welcome
- 09:15 – 09:45 Session 1: Genetic variation and heredity
- 09:45 – 10:15 Group work 1: Genetic variation and heredity
- 10:15 – 10:30 Break
- 10:30 – 11:00 Session 2: From SNP to GWAS to disease risk
- 11:00 – 11:40 Group work 2: Construct your own PGS
- 11:40 – 12:00 Common discussion and evaluation**

Join at menti.com | use code **7798 1843**

Hvilken indsigt tager du med dig fra i dag?

No questions from the audience!

Incoming questions will show up here so that you can answer them one by one.



PD

**Menti**

PM-genetik2025

**Choose a slide to present**

What is personalised medicine?

What do you associate with *genetic variation* in the context of *personalised medicine*?

Hvilken indsigt tager du med dig fra i dag?

0 questions
0 upvotes



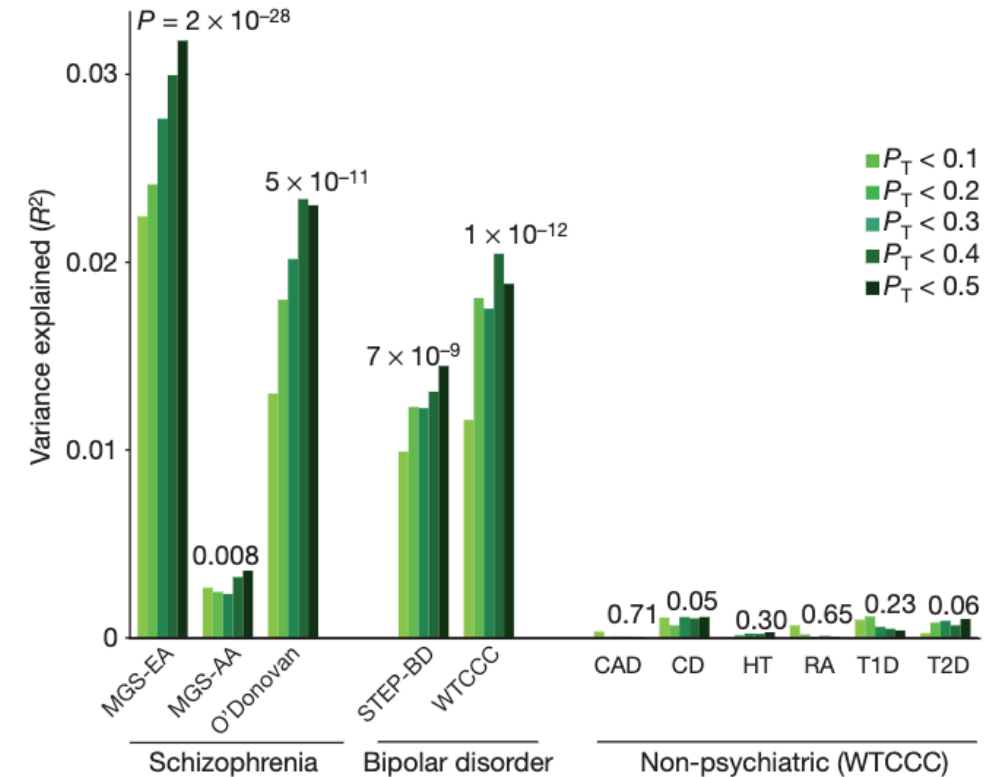
EVALUATE POLYGENIC PROFILES

$$y = Xb + Zc + e$$

y = phenotype; X = PGS; Z = covariates

Compare variance explained from the full model (with X+covariates) compared to a reduced model (covariates only)

Variance explained (R^2) for quantitative traits, and Nagelkerke's R^2 for binary traits (**however, NagR2 is biased with disease prevalence!**)



The International Schizophrenia Consortium (2009) Nature, 460

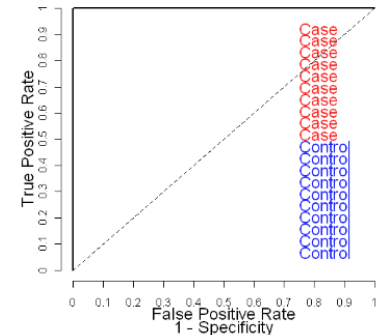
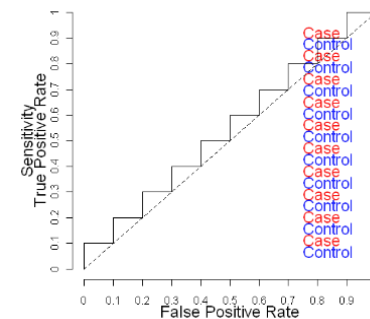
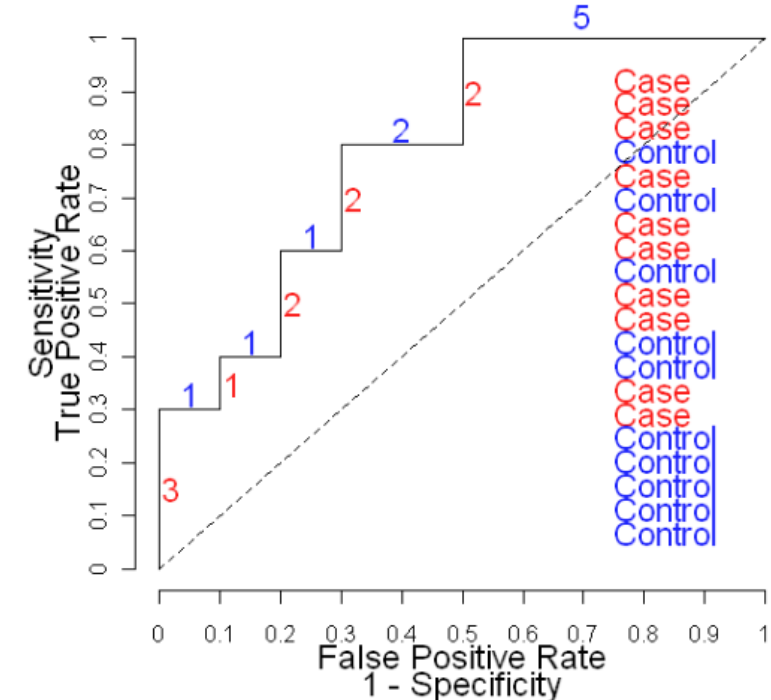
EVALUATE POLYGENIC PROFILES

Area Under Receiver Operator Characteristic Curve (AUC)

Well established measure of validity of tests for classifier diseased vs non-diseased individuals

- Nice property – independent to proportion of cases and controls in sample
- Range 0.5 to 1
- 0.5 the score has no predictive value
- Probability that a randomly selected case has a score higher than a randomly selected control

Rank individuals on score from highest ranked to lowest



EVALUATE POLYGENIC PROFILES

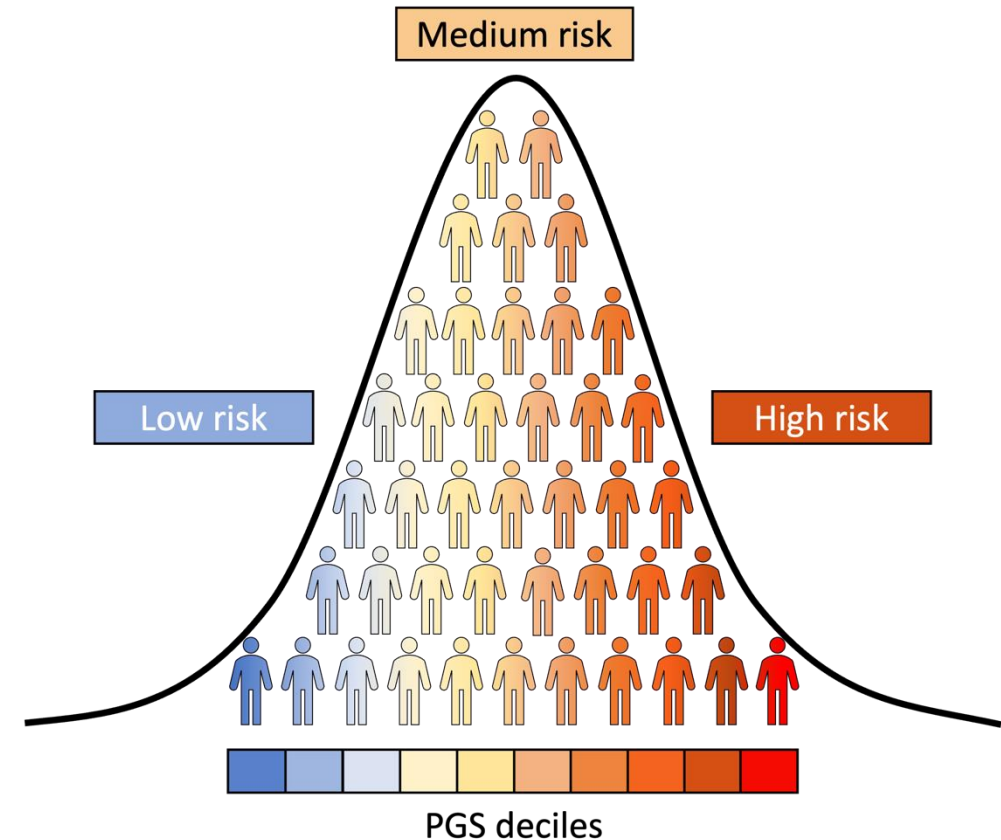
Odds ratio (OR)

Cut the distribution into deciles

Each decile will include both cases and controls

Odds of being a case in each decile

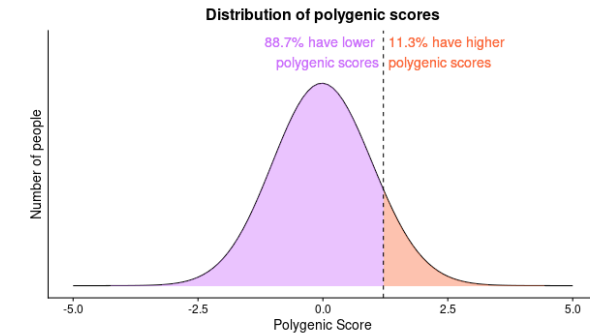
Odds ratio for each decile compared to the first decile



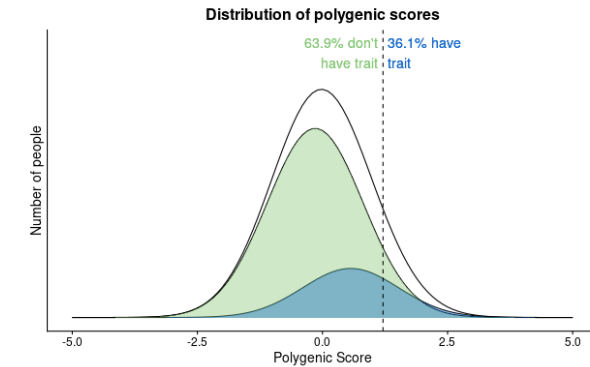
INTERPRETABILITY & RISK COMMUNICATION

- The risk associated with the PGS is a relative risk
 - People have suggested methods to convert relative PGS risk to absolute risks (https://opain.github.io/GenoPred/PRS_to_Abs_tool.html)
- The relative risk may sound high, but the absolute risk is low
- Hard to use meaningfully in clinical decisions without baseline risk
- Effective for population stratification, less so for individual prediction
- Lack of population transferability

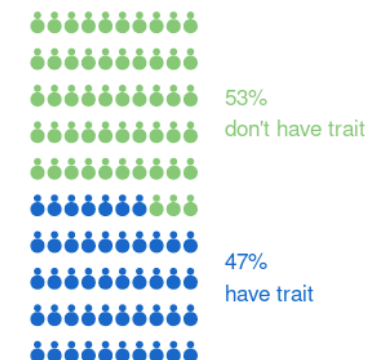
Relative Risk



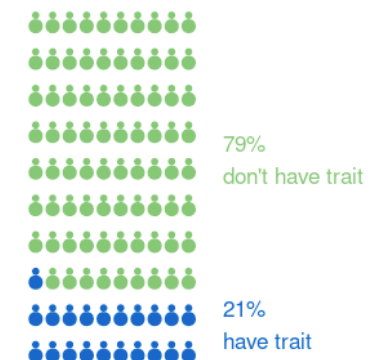
Absolute Risk



Of people with your genetics,



In the general population,



LACK OF TRANSFERABILITY

