Review of quantitative genetics

...

Evangelina López de Maturana & Oscar González-Recio

Something you need to carefully look at, or that may impair your GWP

Something to do, or that optimizes your GWP

Don’t. Discourage to use this.

Smart tip. Something that makes the trick.

Advanced. Something to dive in.

Challenges

What you need to know from this lecture

Basic concepts

How genes and environment modulate the phenotype

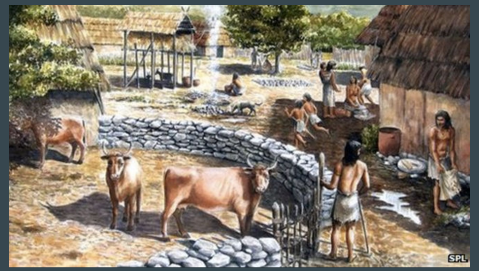
Difference between gene and marker

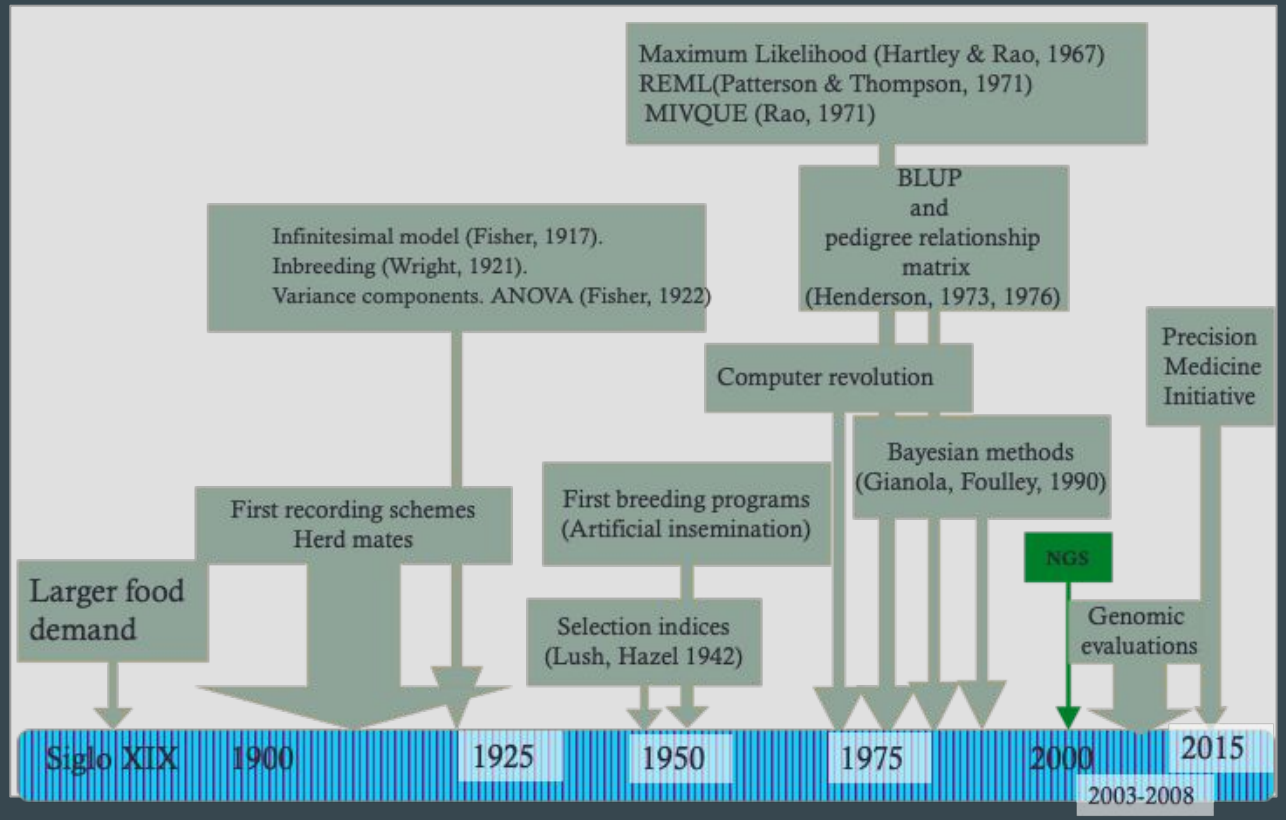
What is genomic heritability and how it affects to GWP

Interpret what a GWP implies

Understand the importance of genetic and genomic variance

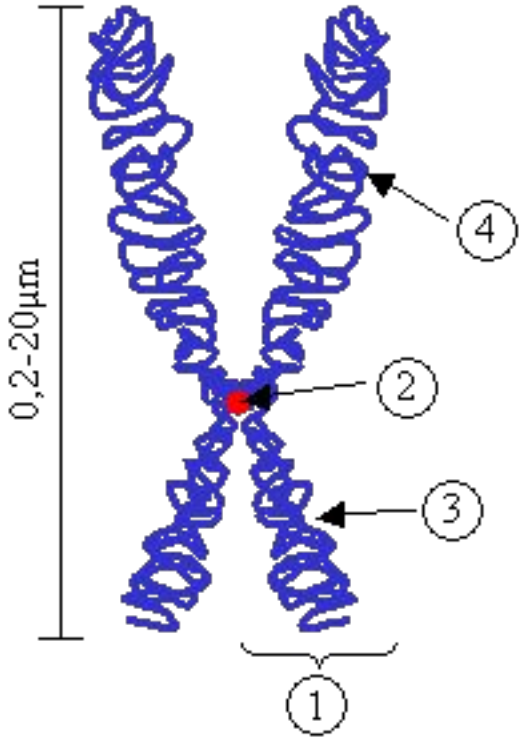
A bit of history





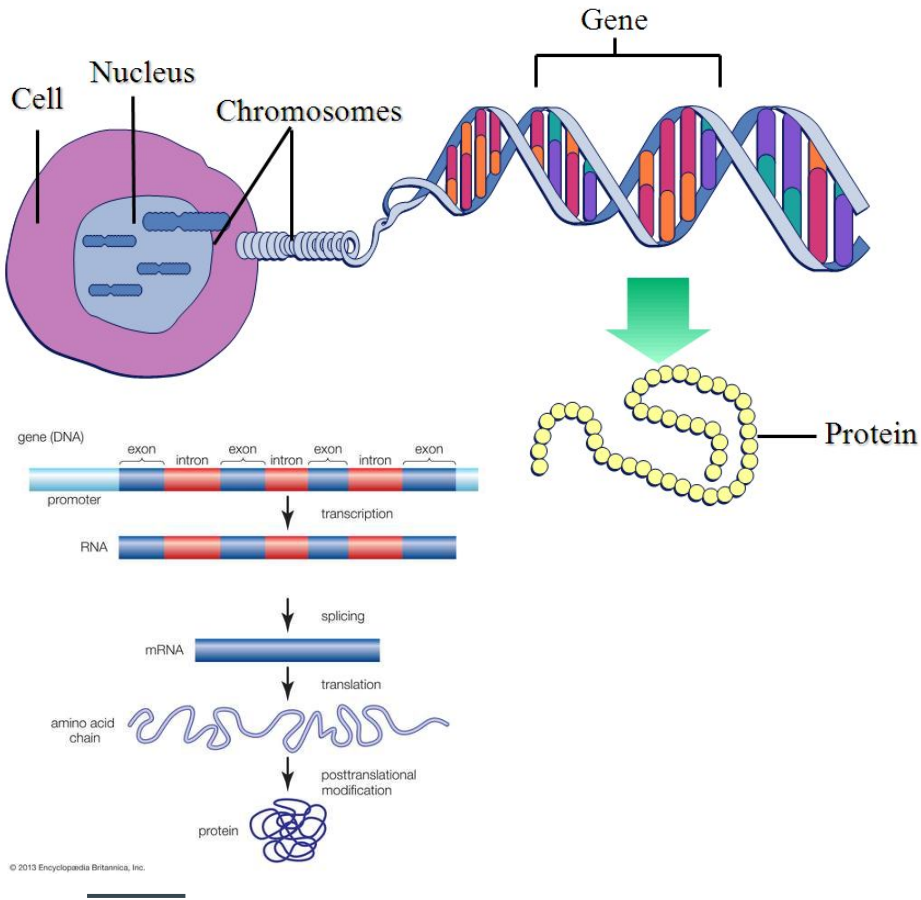
Locus, loci

A specific physical location of a gene, DNA sequence or genetic marker on a chromosome; like a genetic street address



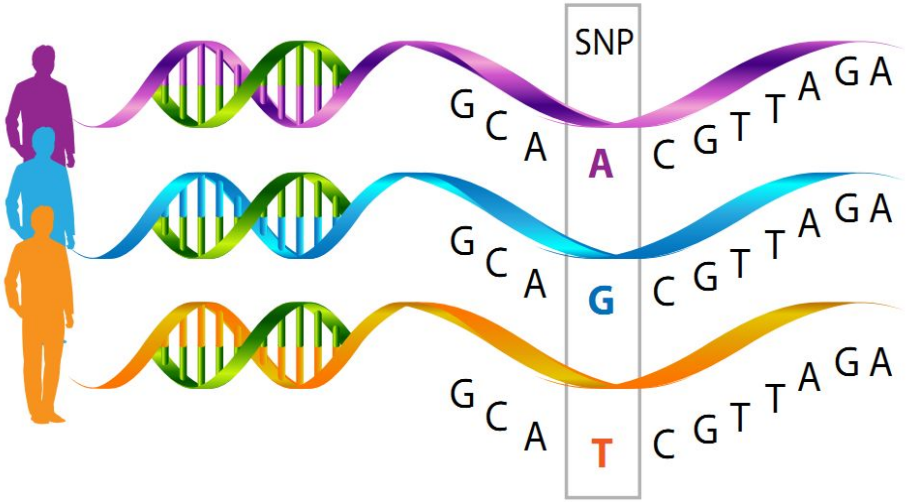
Gene

Gene, unit of hereditary information that occupies a fixed position (locus) on a chromosome. Genes achieve their effects by directing the synthesis of proteins.



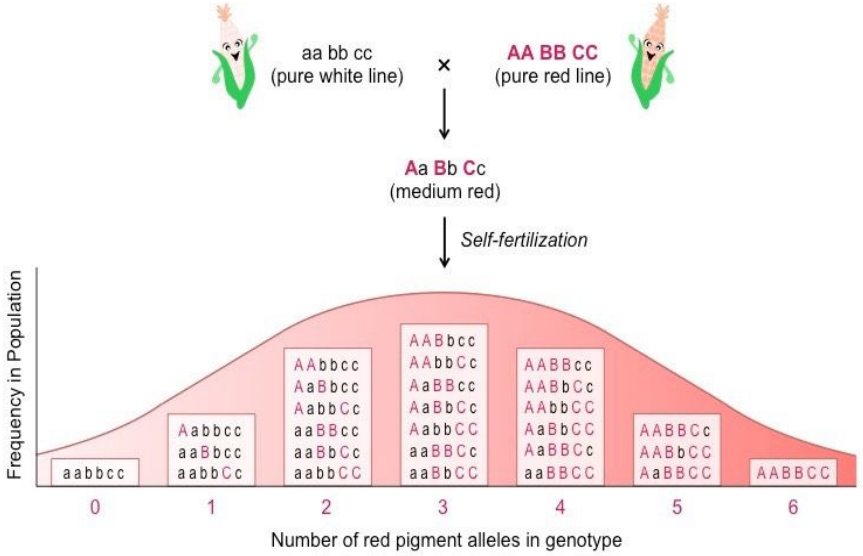
Genetic marker

DNA sequence with a known location on a chromosome that can be used to identify individuals or species. It can be described as a variation (which may arise due to mutation or alteration in the genomic loci) that can be observed



Infinitesimal model

A quantitative trait is influenced by an infinitely large number of genes, each of which makes an infinitely small (infinitesimal) effect, as well as by environmental factors. Random sampling of alleles at each gene produces a continuous, normally distributed phenotype in the population (at least around the average of that of the individual's parents).



Genome-wide prediction

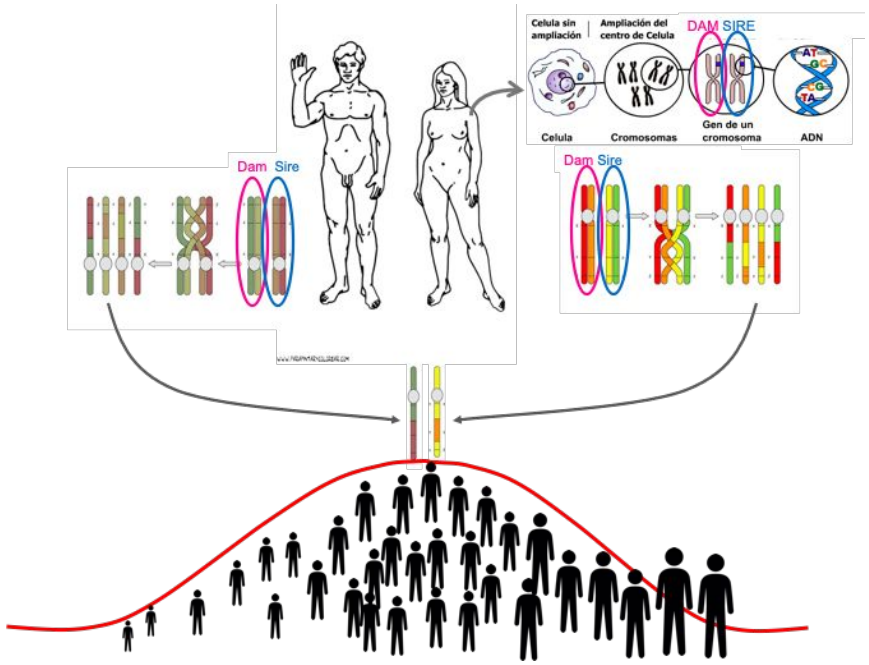
Pedigree index

Parent average

$$1/2\ EBV_{sire} + 1/2\ EBV_{dam}$$

Mendelian effect

Deviation from the expected parent average



$$a_i = \frac{1}{2}a_p + \frac{1}{2}a_m + \phi_i$$

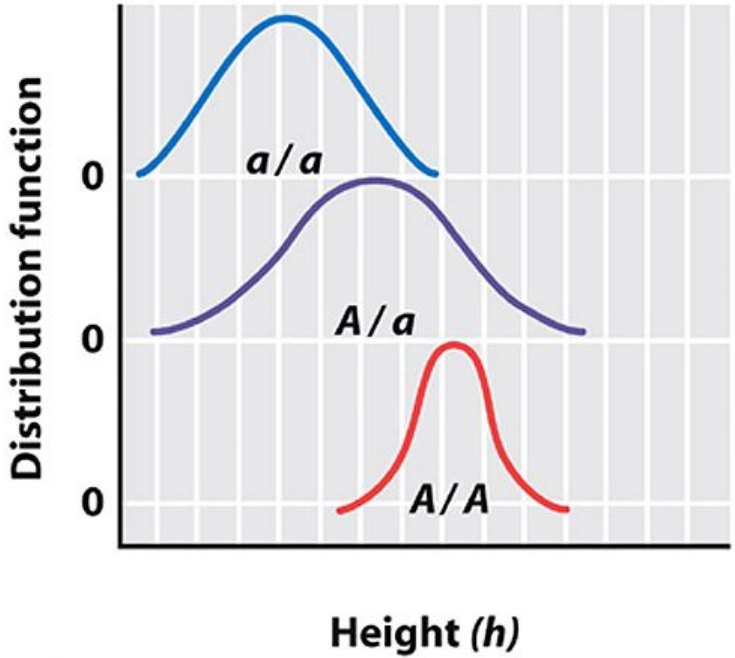
$$\phi_i \sim N(0, \sigma_a^2 / d_i)$$

Genome-wide prediction

Phenotype decomposition

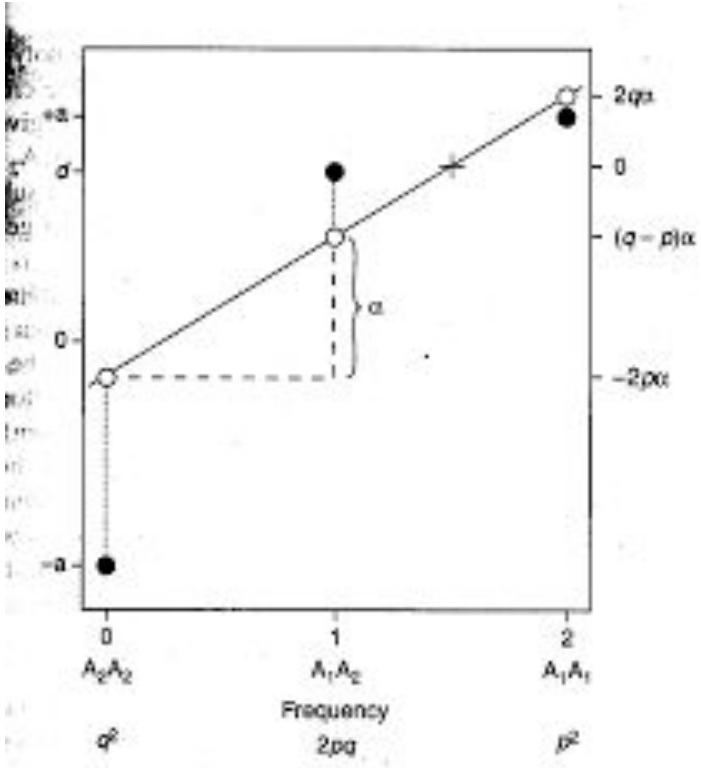
Phenotype is affected by genetic (additive +dominance +epistasis), environment and their interactions.

$$P = G + E$$



⚠ Allele substitution effect

The effect that the presence of a copy of an allele has on the phenotype (regarding the reference allele).



f(A)= mean(Aa)-mean(aa)

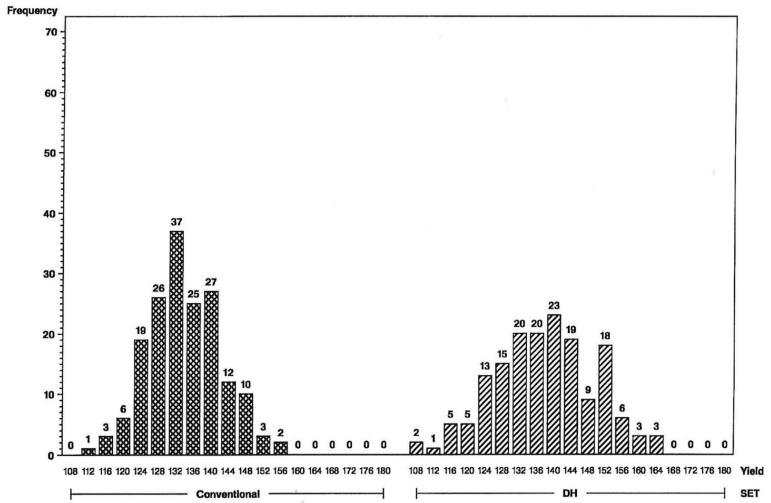
Marker variance

Phenotype deviation from the mean phenotype caused by the inheritance of a particular allele from parental and this allele's independent effect on the specific phenotype

Notation	Variance component	Genotype coding
V_A	$2pq[a + d(p - q)]^2$	$x_A \in \{0, 1, 2\}$
V_D	$(2pqd)^2$	$x_D \in \{0, 2p, 2(p - q)\}$
V'_D	$\frac{4pq^2}{1+q}(a + dq)^2$	$x'_D \in \{0, 2, 2\}$
V'_A	$\frac{2p^2q}{1+q}(a - d)^2$	$x'_A \in \{0, \frac{1-q}{1+q}, \frac{-2q}{1+q}\}$
V''_{AA}	computed numerically	$x''_{AA} \in (x_{A,1} - 1)(x_{A,2} - 1)$

Genetic variance

Phenotype deviation from the mean phenotype caused by the combination of alleles inherited from parentals and these alleles independent effects on the specific phenotype



Genome-wide prediction

Heritability

The amount of phenotypic (observable) variation in a population that is attributable to individual genetic differences

“Narrow and broad sense”

$$H^2 = \frac{V_g}{V_g + V_e}$$

Genomic variance

The amount of variance explained by marker effects

$$
\begin{aligned}
Var(\beta' x_i) &= \beta' Cov(x_i, x_i') \beta \\
&= \beta' \Sigma_x \beta \\
&= \alpha' \Sigma_{zx} \Sigma_x^{-1} \Sigma_x \Sigma_x^{-1} \Sigma_{xz} \alpha \\
&= \alpha' \Sigma_{zx} \Sigma_x^{-1} \Sigma_{xz} \alpha
\end{aligned}
$$

Genomic heritability

The proportion of variance of a trait that can be explained (in the population) by a linear regression on a set of markers

$$h_g^2 = \frac{\sigma_g^2}{\sigma_y^2} = \frac{\sigma_a^2}{\sigma_y^2}\frac{\sigma_g^2}{\sigma_a^2} = h^2\frac{\sigma_g^2}{\sigma_a^2}$$

$h^2_g <= h^2$

Missing heritability

The problem of missing heritability, that is to say the gap between heritability estimates from genotype data and heritability estimates from twin data

The case of the missing heritability

When scientists opened up the human genome, they expected to find the genetic components of common traits and diseases. But they were nowhere to be seen. Brendan Maher shines a light on six places where the missing loot could be stashed away.

Pleiotropy

the phenomenon in which a single locus affects two or more apparently unrelated phenotypic traits

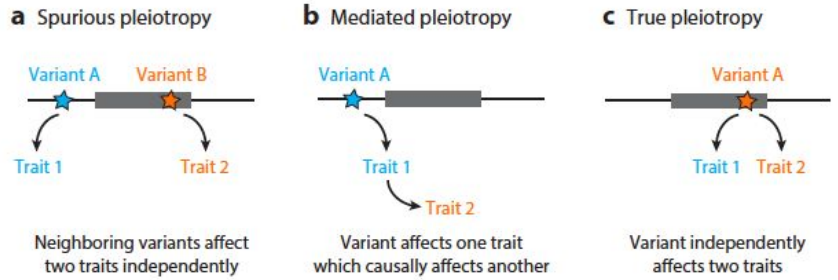
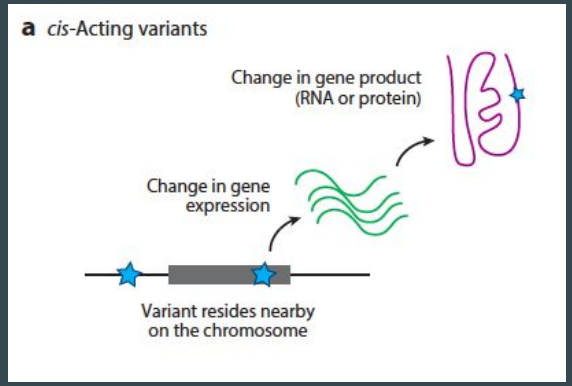


Figure 4

Diagrams illustrating (*a*) spurious pleiotropy, in which two neighboring, separately causal variants (*blue* and *orange stars*) are mistakenly inferred to be pleiotropic because they cannot be statistically distinguished; (*b*) mediated pleiotropy, in which a variant is statistically associated with two traits because it has a causal effect on one trait that in turn causally impacts another; and (*c*) true pleiotropy, in which a single unambiguous causal variant is separately biologically causal for two independent traits.



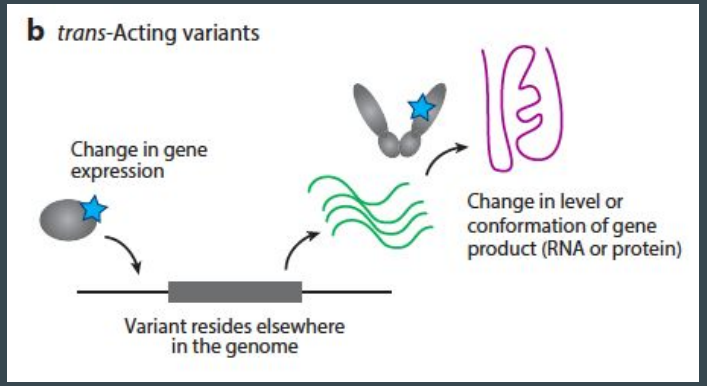
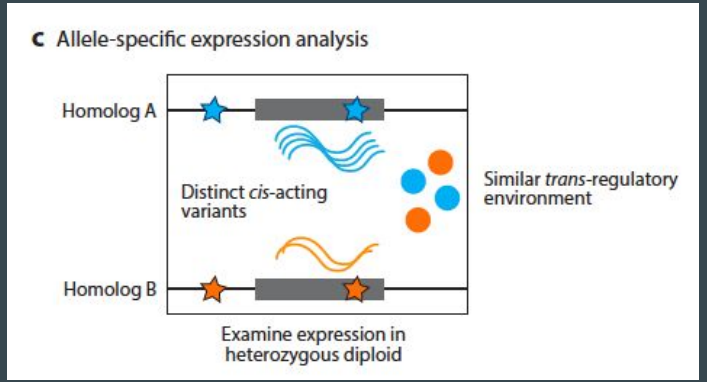


Figure 2

(*a*) *cis*-Acting variants that impact the expression of a gene immediately proximal on the chromosome. (*b*) *trans*-Acting variants that impact a gene product originating from a distal genetic locus. (*c*) Schematic of an experimental design to measure allele-specific mRNA levels. Due to the presence of both parental alleles in the F_0 heterozygote, *cis*-acting regulatory activity is inferred from differential expression of the messenger RNA attributable to one of the two homologous loci. This is because both homologs exist in an essentially equivalent *trans*-regulatory environment; any difference in abundance must therefore be due to a nearby *cis*-acting variant.



Jakobson and Jarosz, 2020

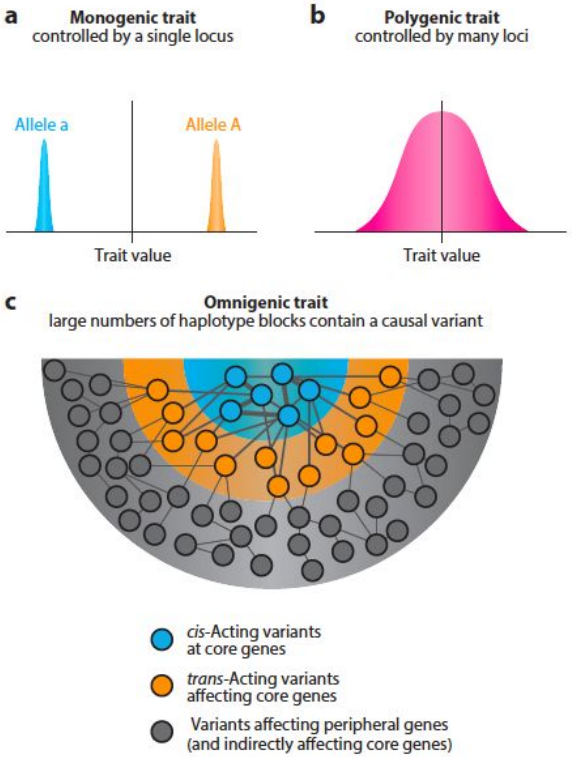
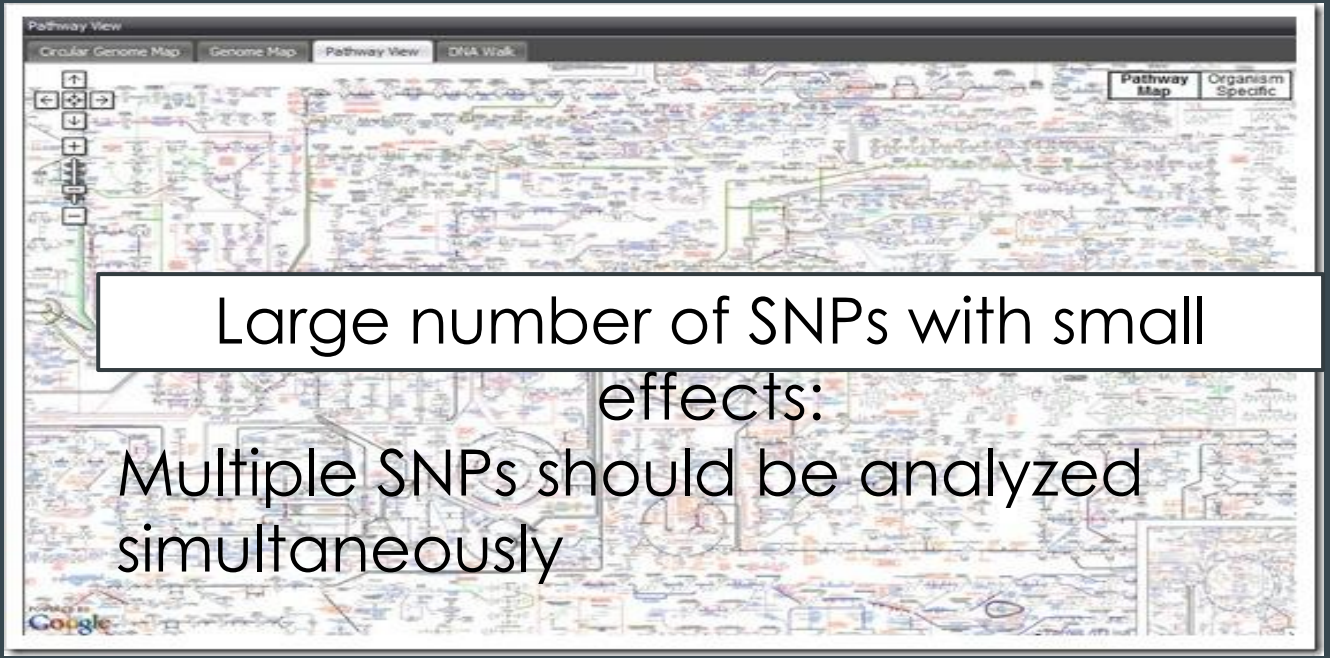


Figure 1

(*a*) A bimodal trait distribution for a monogenic trait controlled by a single genetic locus, as compared to (*b*) a continuous trait distribution for a polygenic trait controlled by many genetic loci. (*c*) Schematic of one possible architecture for an omnigenic trait, in which several large-effect *cis*-acting and many smaller-effect *trans*-acting variants modulate a set of core genes, as does a much larger ensemble of *cis*- and *trans*-acting variants impacting peripheral genes that only indirectly modulate the phenotype.

Genome-wide prediction



Large number of SNPs with small effects:

Multiple SNPs should be analyzed simultaneously

RECAP

Gene vs Marker

Assume gaussian distribution on phenotypes (... subsequently residuals)

Why variance is important

Inference is different from prediction.

Genetic architecture challenge