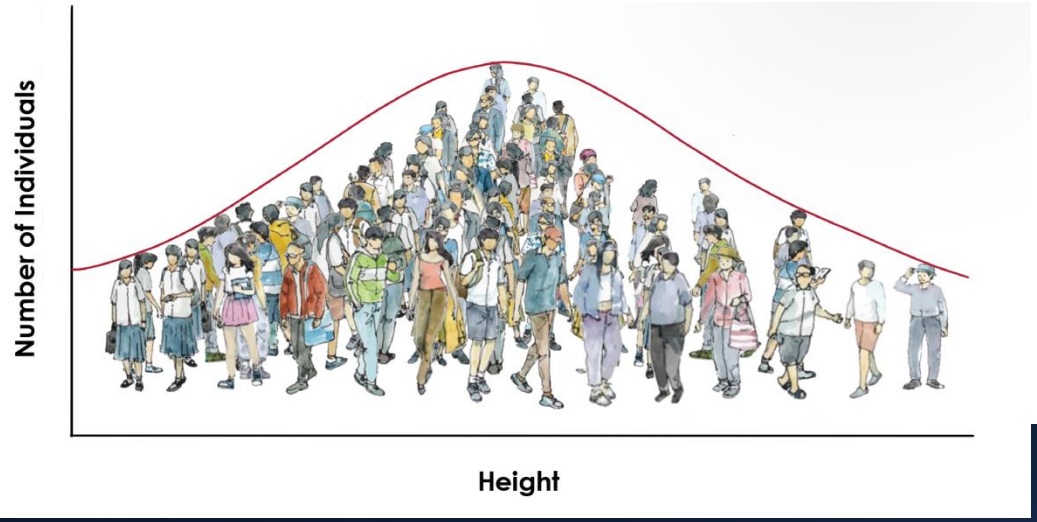
Introduction to Multivariate Quantitative Genetics



EQGW 2025– Mountain Lake

Jacqueline Sztepanacz

University of Toronto

Genetic covariances are important for evolution

- selection on one trait will lead to a correlated response in the other
- can accelerate evolution
- can slow evolution
- can prevent evolution?

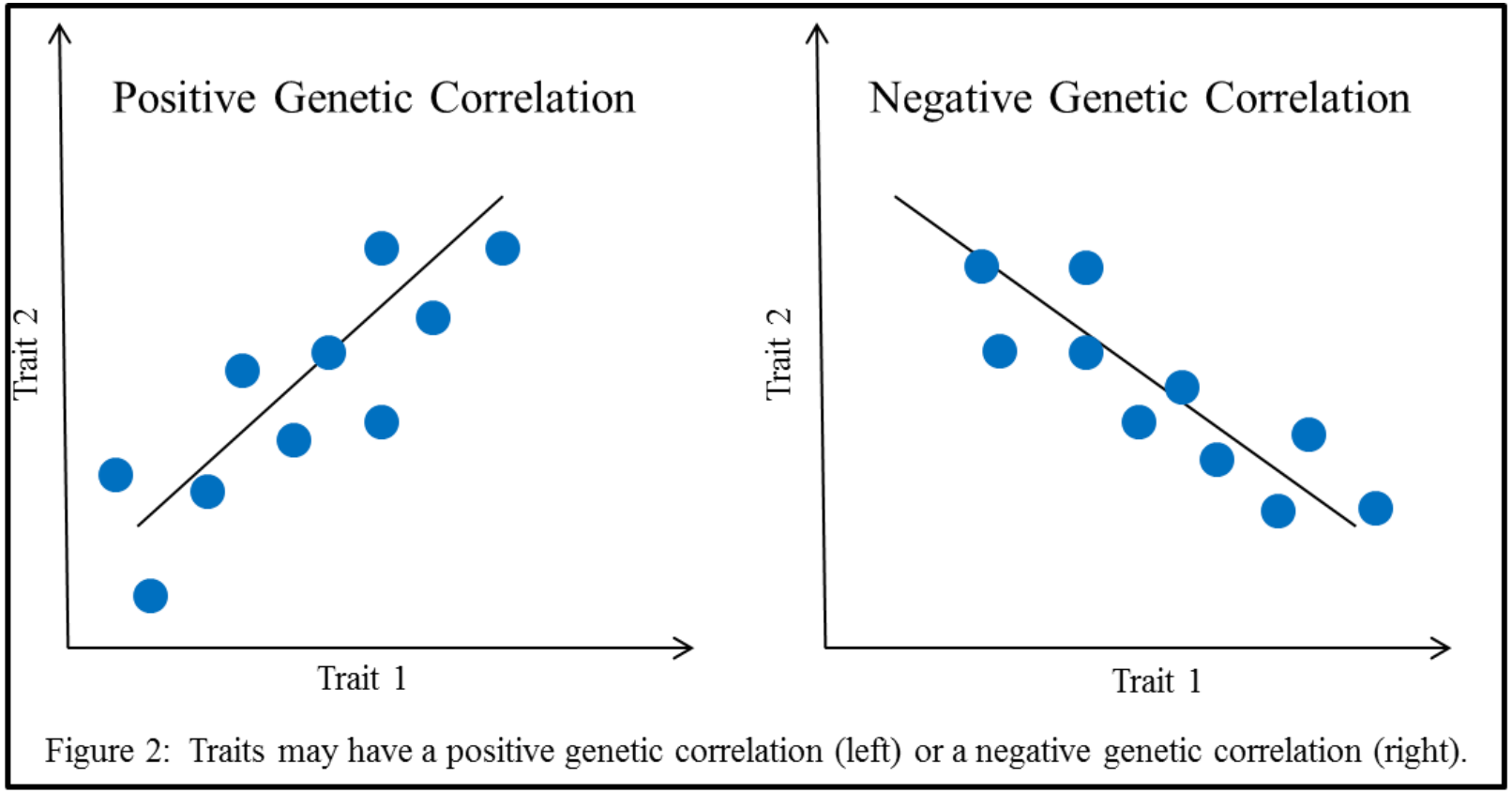


Figure 2: Traits may have a positive genetic correlation (left) or a negative genetic correlation (right).

Let’s us answer important and interesting questions…

- can this population adapt to ecological change, or will it go extinct?
- is the evolutionary response in my favourite trait constrained?
- do diseases commonly co-occur? which ones?
- what is the degree of integration of a phenotype (or how many independent genetic dimensions underly an organism)
- will my award-winning racehorse have good sons and good daughters?

Topics we will cover:

- **Multivariate quantitative genetics**

1. Pleiotropy & Genetic correlations
2. The G matrix
3. Genetic constraints

- **Selection**

1. Empirical methods to estimate selection
2. Empirical results

Key Take Aways:

- genetic variation is unevenly distributed across multivariate trait combinations because of pleiotropy
- the uneven distribution of genetic variance can lead to evolutionary constraints
- we can estimate selection on multiple traits using linear or quadratic regression approaches

Some historical context…

- Quantitative genetics wasn't a major focus in evolutionary research until the 1970's/1980's

- Evolutionary quantitative genetics happened in 2 steps

1. Lande's papers in the late 70's

2. Operational framework to estimate selection in natural populations (eg. Lande and Arnold 1983)

Breeder’s vs Lande equation

Breeder’s equation

Lande equation

Topics we will cover:

- Multivariate quantitative genetics

1. **Pleiotropy & Genetic correlations**
2. The G matrix
3. Genetic constraints

- Selection

1. Empirical methods to estimate selection
2. Empirical results

Phenotypic correlation between traits

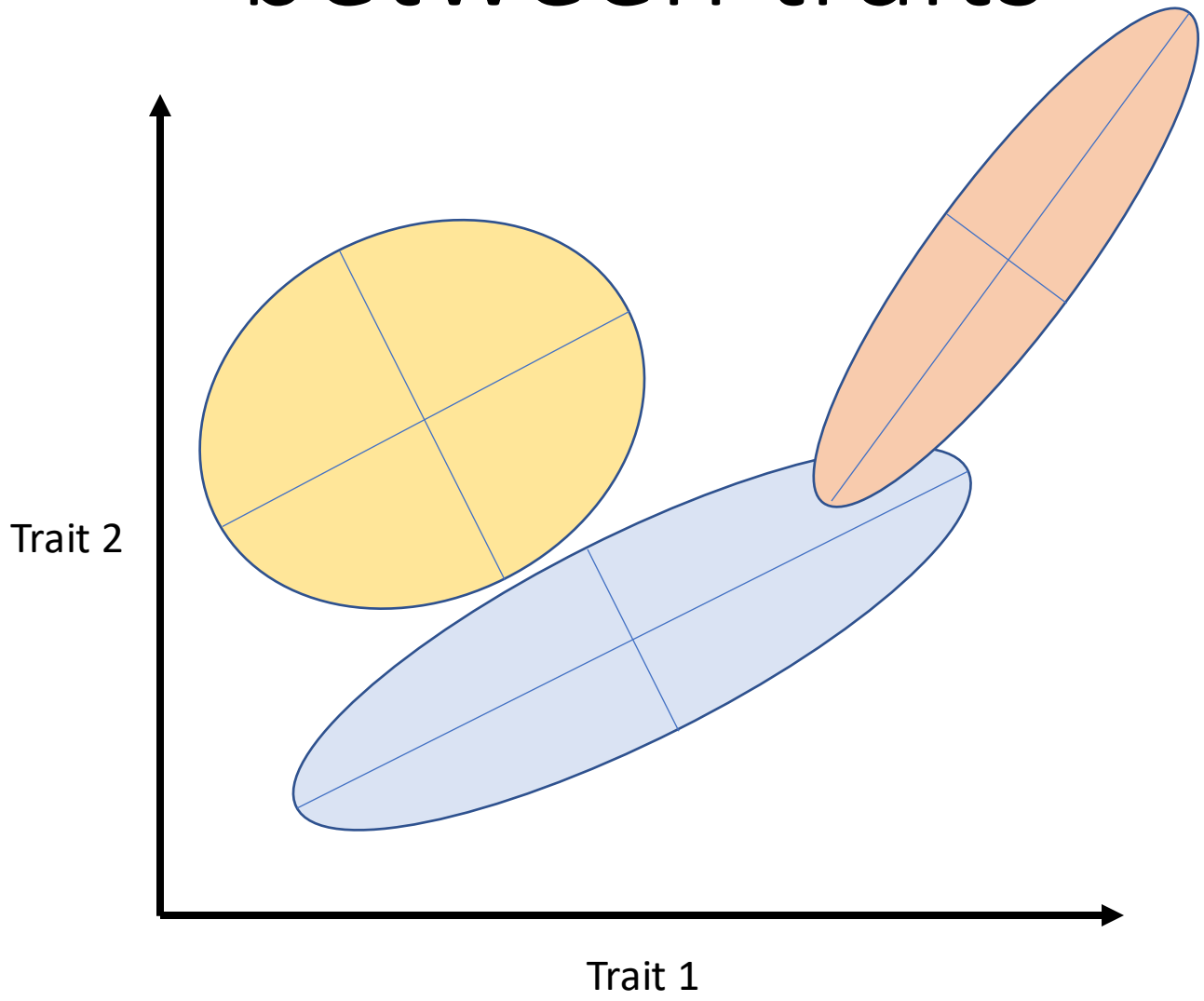
Breeders believe that long limbs are almost always accompanied by an elongated head. Some instances of correlation are quite whimsical; thus cats with blue eyes are invariably deaf; … Hairless dogs have imperfect teeth; long-haired and coarse-haired animals are apt to have, as is asserted, long or many horns; pigeons with feathered feet have skin between their outer toes; pigeons with short beaks have small feet, and those with long beaks large feet. Hence, if man goes on selecting, and thus augmenting, any peculiarity, he will almost certainly unconsciously modify other parts of the structure, owing to the mysterious laws of the correlation of growth.

— *Charles Darwin*, *The Origin of Species*, *1859*

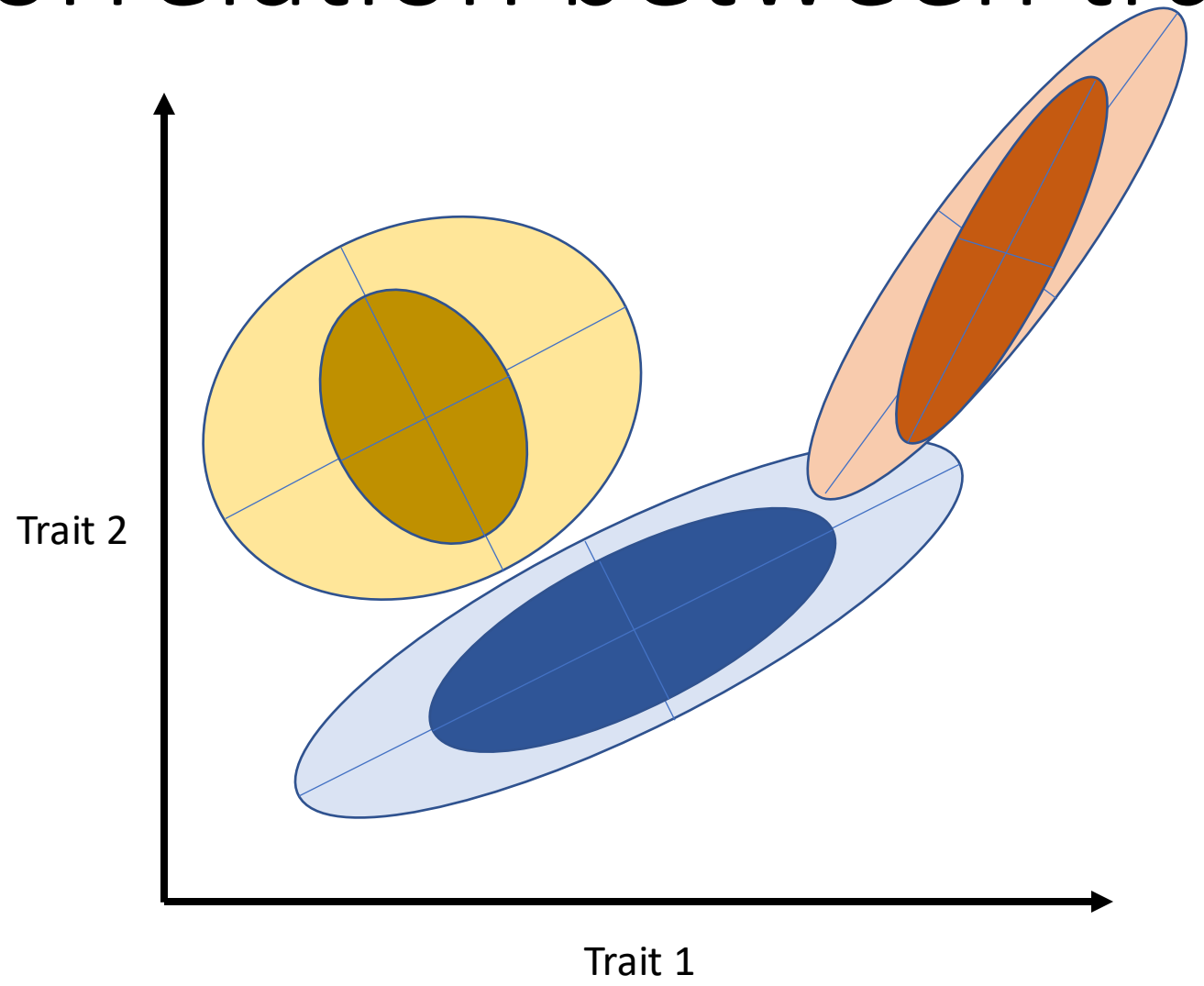
Phenotypic correlation between traits

- phenotypic correlations can be caused by environmental factors
- variation in resource availability can lead to a positive correlation between the size of all appendages
- environmental cue to initiate the allocation of resources to reproduction causes a curtailment in growth
- Can also cause correlations between traits and fitness

The environmental correlation between traits



The genetic and environmental correlation between traits



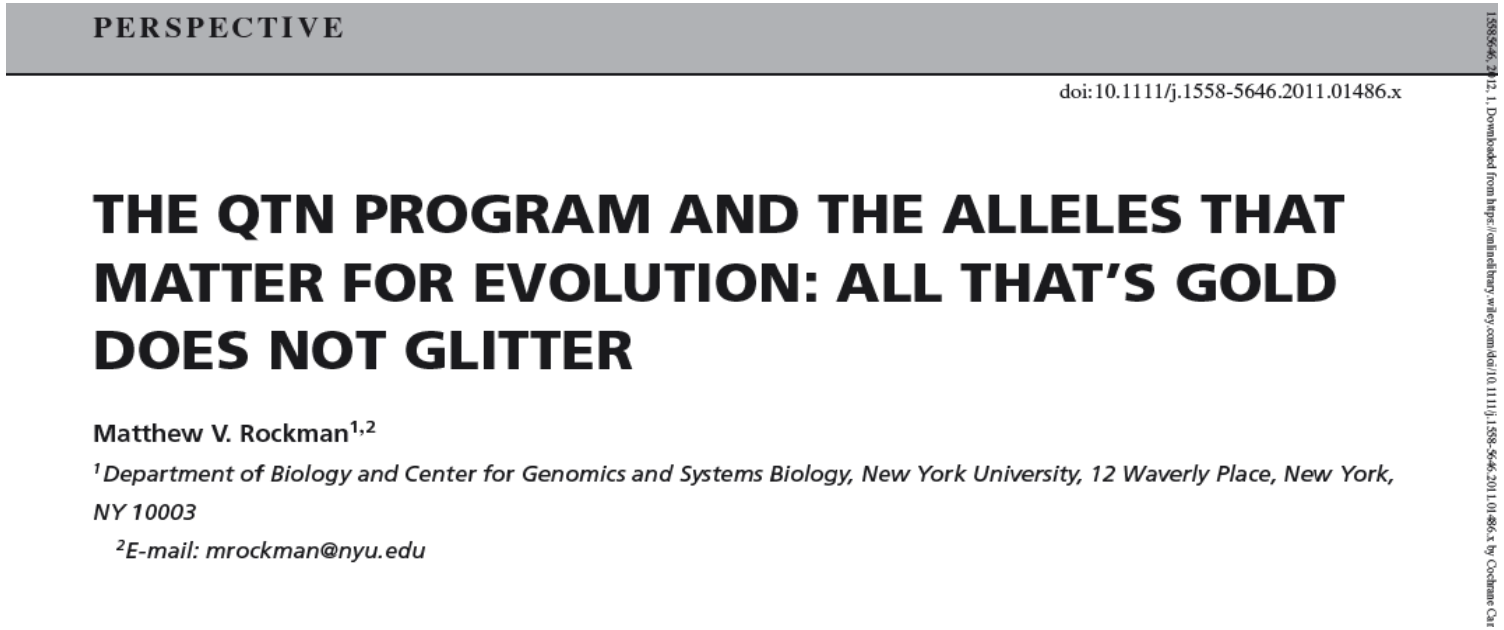
Unintended effects from breeding- double muscling

- some beef cattle show extraordinary muscle
- caused by mutations in myostatin genes
- has been selected for in Belgian blue cattle- they produce 20% more lean edible meat than other cattle
- leads to problems with stress tolerance, fertility, and calf viability

Unintended effects from breeding- Super-chickens

- artificial selection for egg laying- individual level selection
- over time they produced fewer eggs- pleiotropic side effect of aggression
- they pecked each other to death

QTL / GWAS- promised to find many major effect loci



PERSPECTIVE

doi:10.1111/j.1558-5646.2011.01486.x

THE QTN PROGRAM AND THE ALLELES THAT MATTER FOR EVOLUTION: ALL THAT'S GOLD DOES NOT GLITTER

Matthew V. Rockman[1,2]

[1]*Department of Biology and Center for Genomics and Systems Biology, New York University, 12 Waverly Place, New York, NY 10003*

[2]*E-mail: mrockman@nyu.edu*

- most traits have a polygenic genetic architecture

- if there are many more phenotypes than genotypes (and there are- the genome is finite and the phenome is not) AND most traits are affected by many genes, then most genes must affect many traits

Omnigenic model

- human gene regulatory networks are so interconnected that thousands of individual genes contribute at least slightly to the phenotype (infinitesimal model)

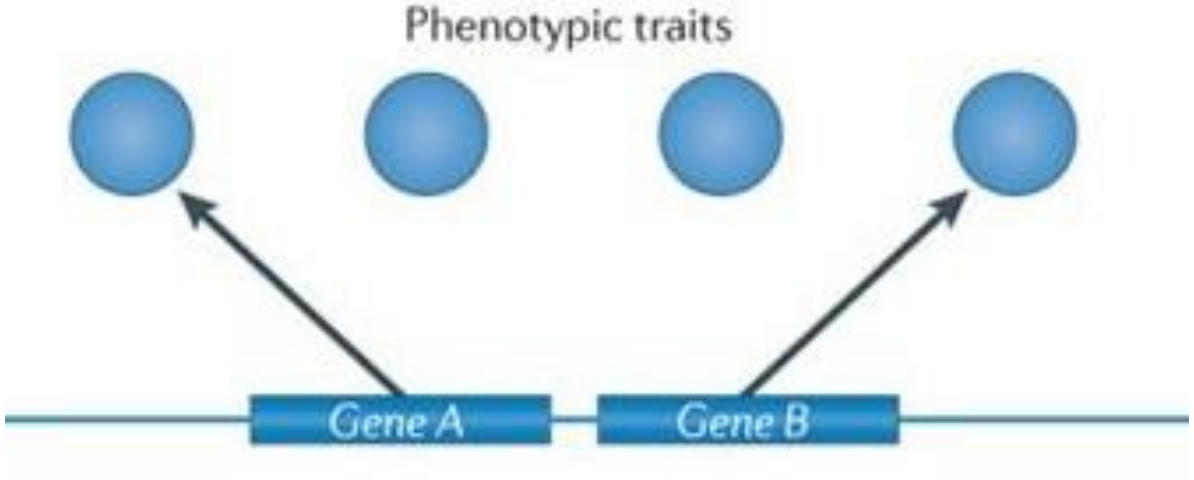
- variation in one part of the genome can have indirect effects on any other trait (universal pleiotropy)

- “peripheral” genes far outnumber core genes and contribute much more to a trait’s heritability

LD causes genetic correlations

- linkage disequilibrium is a measure of whether an allele at one locus is found more often with an allele at another locus

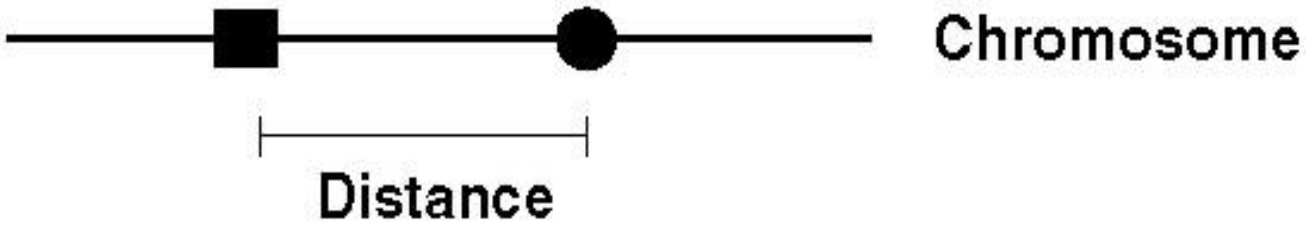
- can be caused by physical linkage



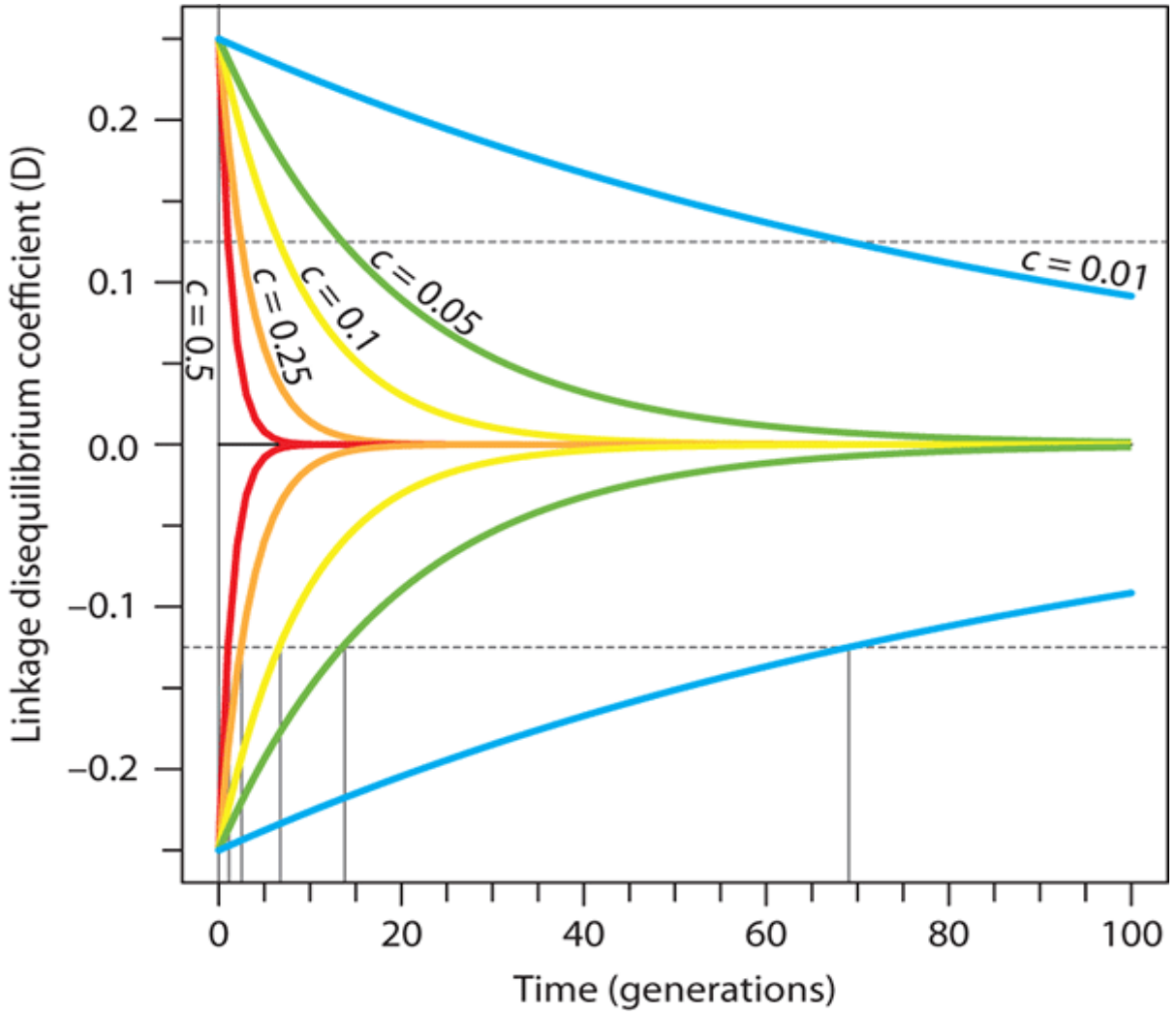
Wagner, G., Zhang, J. The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. *Nat Rev Genet* **12**, 204–213 (2011). https://doi.org/10.1038/nrg2949

LD causes genetic correlations

- over time LD caused by physical linkage will decay due to recombination



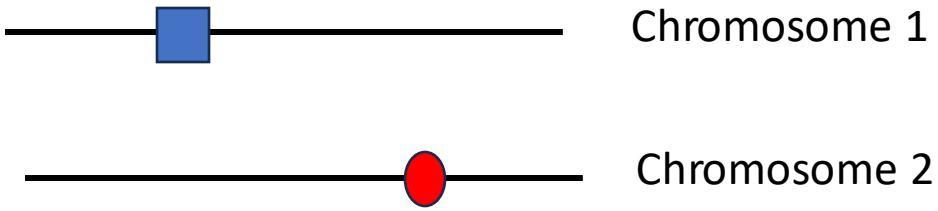
Decay of LD with time



6 Recombination and linkage disequilibrium in evolutionary signatures in A Primer of Molecular Population Genetics, https://doi.org/10.1093/oso/9780198838944.003.0006

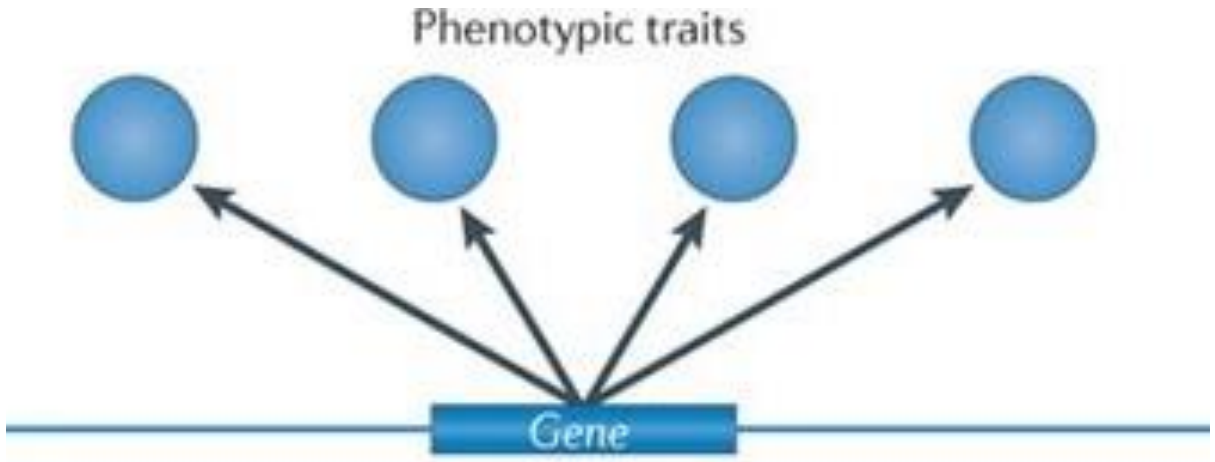
LD causes genetic correlations

- can be caused by ‘statistical linkage’
- selection can maintain LD (eg. non-random mating; covariance between traits and preference; others?)
- One generation of random mating will restore linkage equilibrium



Pleiotropy causes genetic correlations

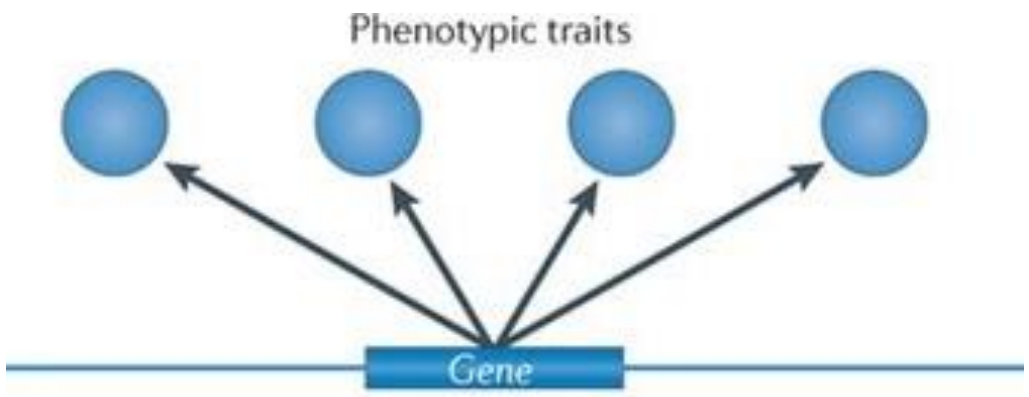
- pleiotropy occurs when a gene/allele affects more than one trait

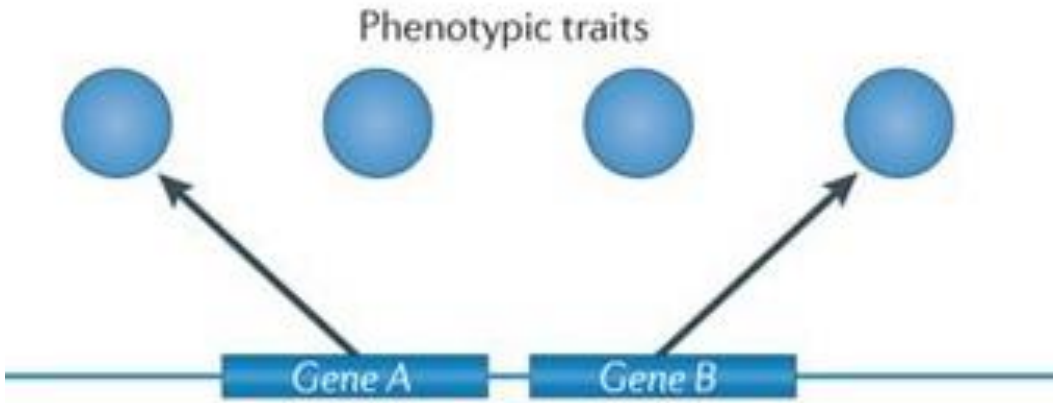


Wagner, G., Zhang, J. The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. *Nat Rev Genet* **12**, 204–213 (2011). https://doi.org/10.1038/nrg2949

Pleiotropy causes genetic correlations

- defining pleiotropy (at least intuitively) is easy- measuring it is not!
- how to distinguish between a pleiotropic mutation and two closely linked mutations?

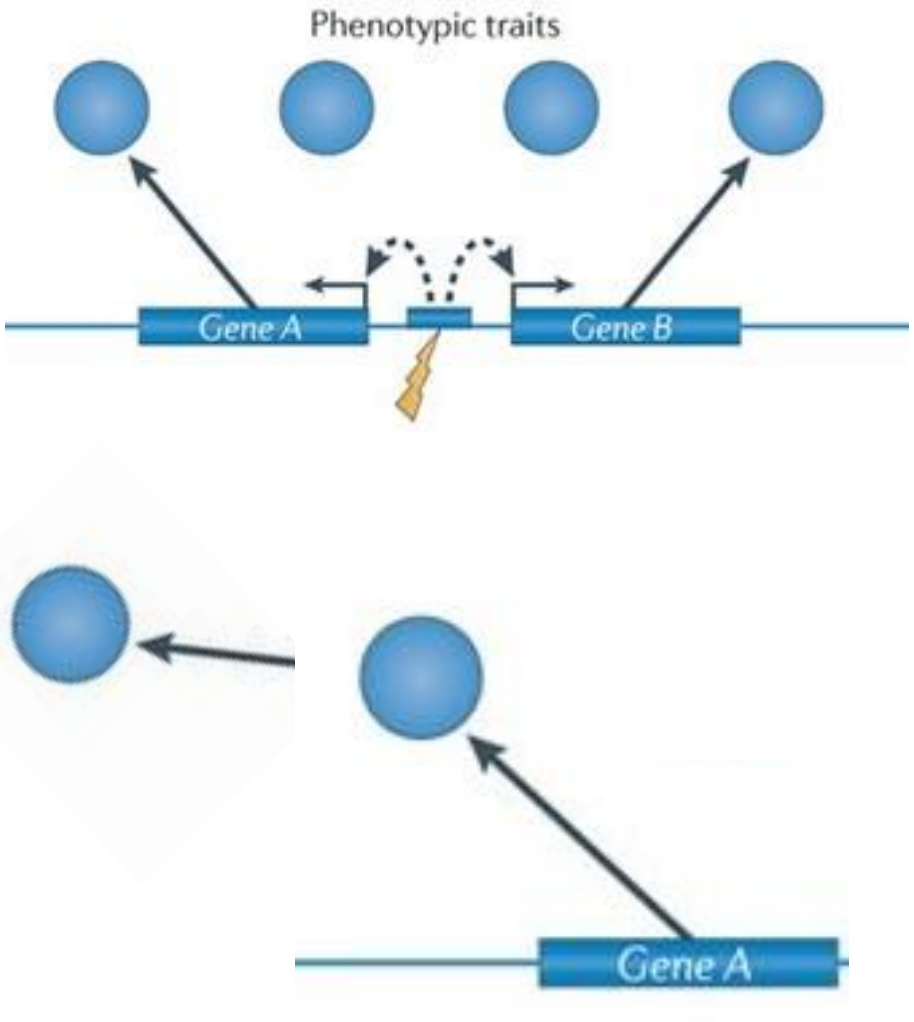


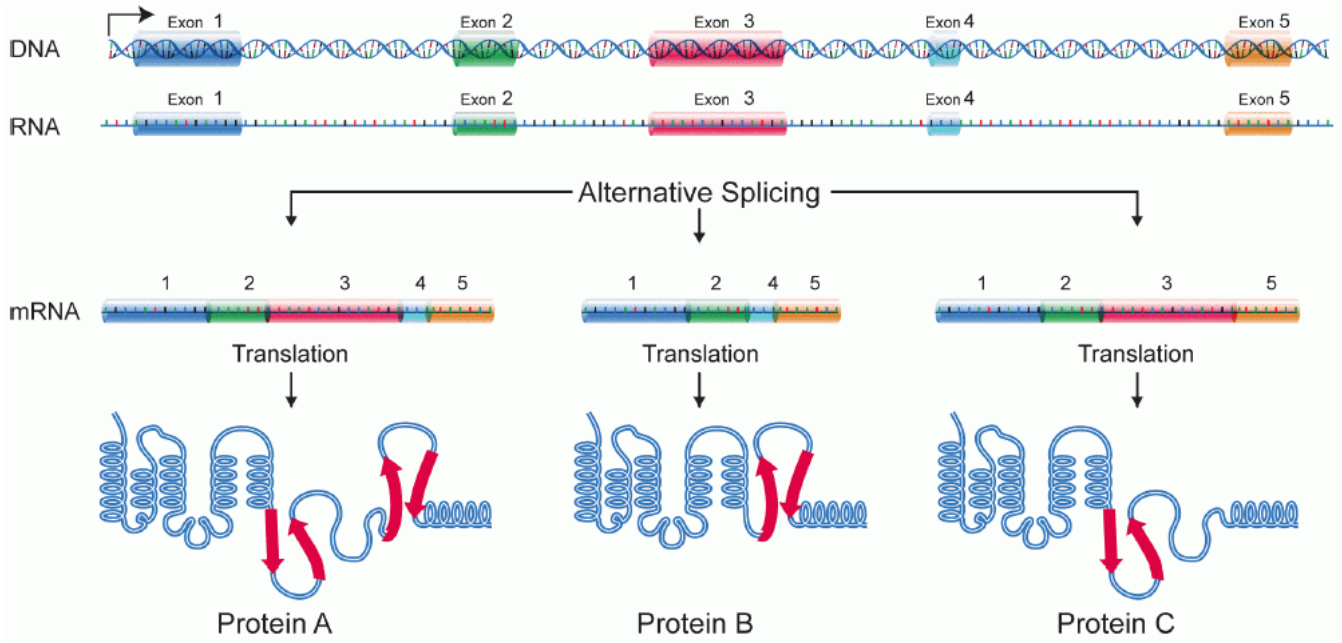


The additive genetic covariance/correlation between traits

- let’s assume that most genetic correlation is caused by pleiotropy….seems like a reasonable assumption

Is pleiotropy a property of a mutation or a gene? Which of these are pleiotropy?





Is pleiotropy a property of a mutation or a gene? Which of these are pleiotropy?

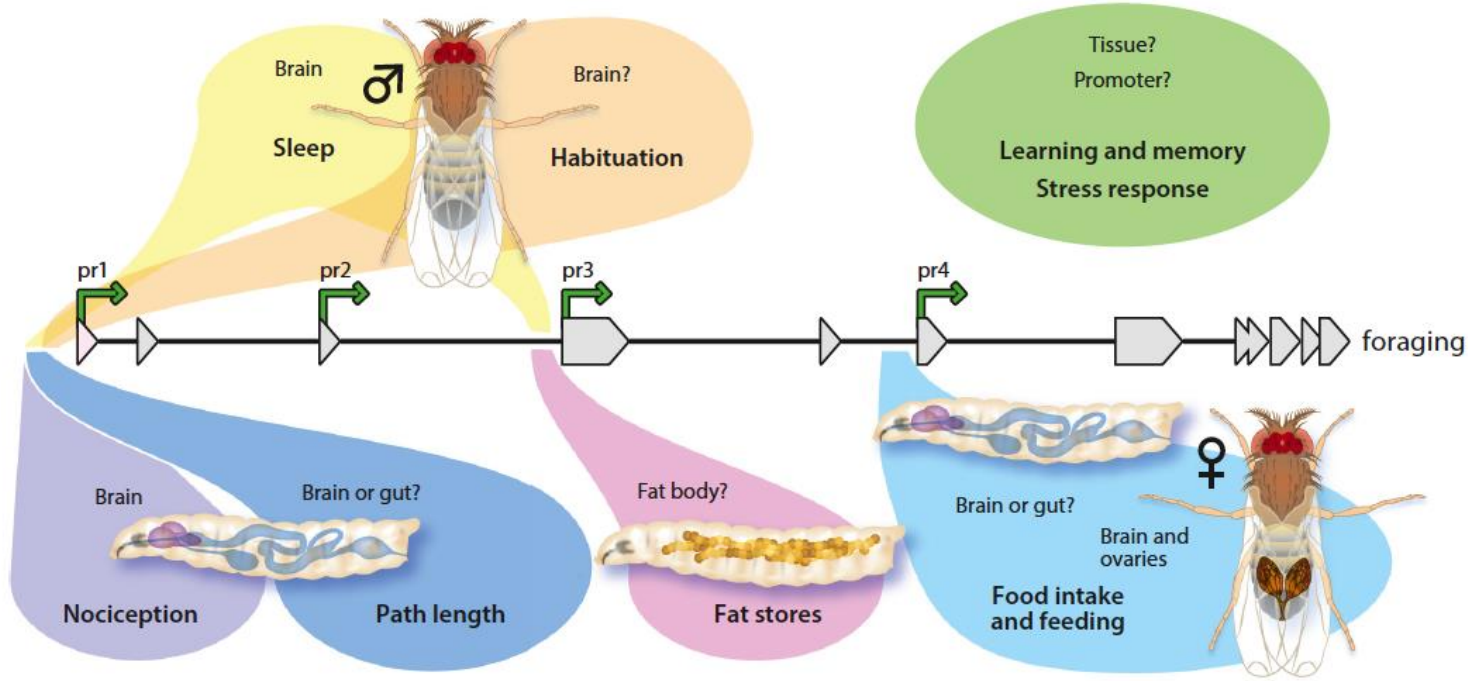


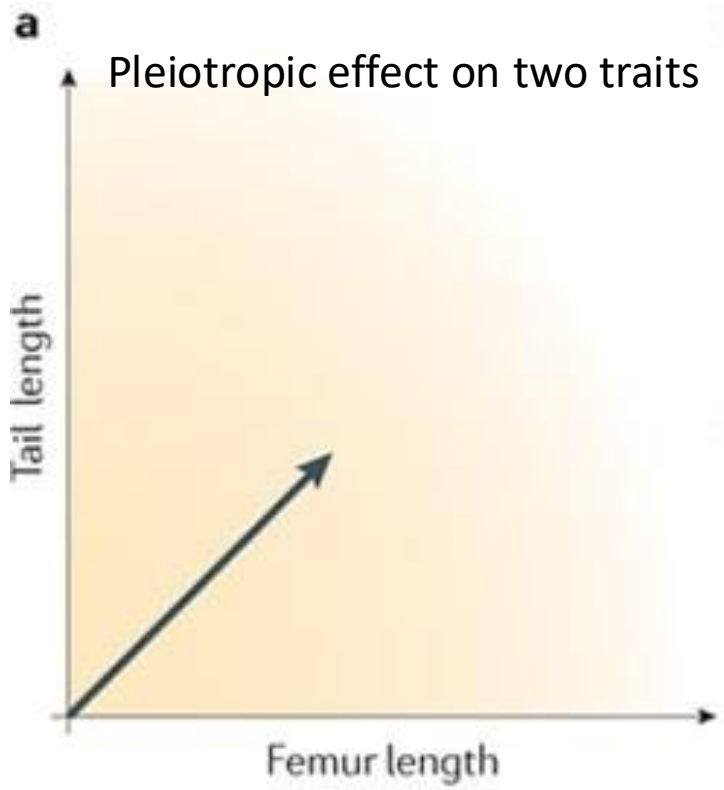
Figure 2

Pleiotropic effects of *foraging*'s promoters. In *Drosophila*, each of the *foraging* gene's promoters regulates distinct behavioral phenotypes in a variety of tissues. Promoter 1 regulates larval nociception (via expression in a neuronal circuit) (21), larval path length (tissue uncertain) (2), and adult sleep (mushroom bodies) (27). Promoter 3 affects fat stores in larvae (tissue uncertain) (2). Habituation is regulated by promoter 1, 3, or both (in olfactory receptor neurons and mushroom bodies) (29). Promoter 4 regulates feeding behavior (adult female brain or ovaries) (5) and in larvae (tissue uncertain) (2). Promoter 2 has so far not been associated with a phenotype; different forms of learning and memory and stress response regulated by *foraging* have not yet been associated with specific promoters or their associated tissues. Abbreviation: pr, promoter.

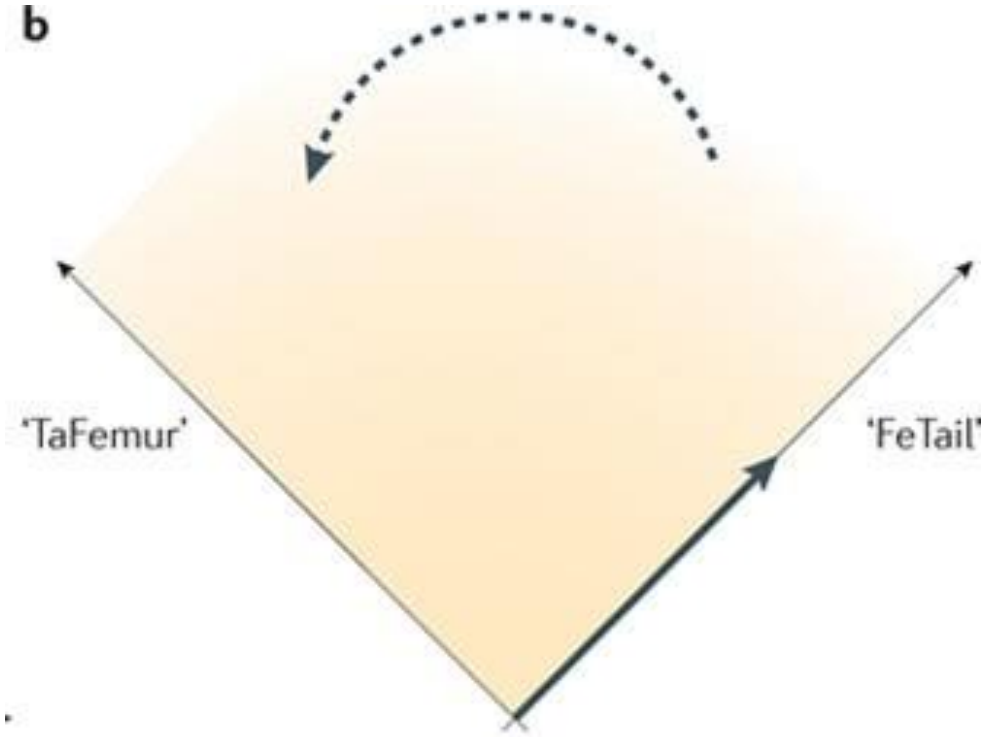
Anreiter I, Sokolowski MB. 2019. The *foraging* gene and its behavioral effects: pleiotropy and plasticity. *Annual Review of Genetics* 53:373-392. DOI

The degree of pleiotropy depends on the definition of a trait

Pleiotropic effect on two traits



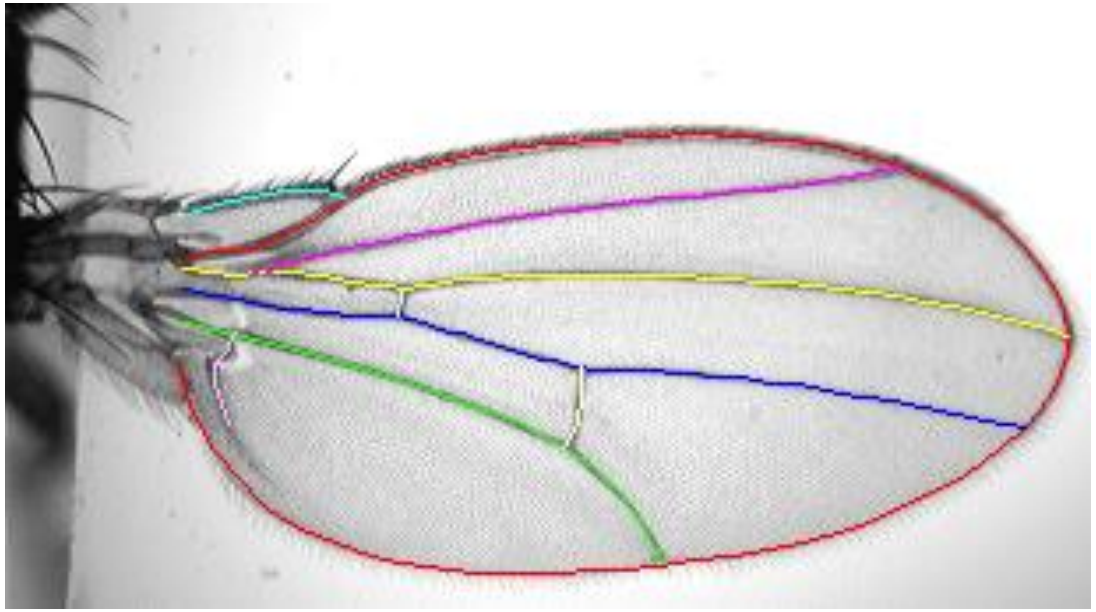
what if we define the trait as the um of tail length and femur length ('FeTail')?



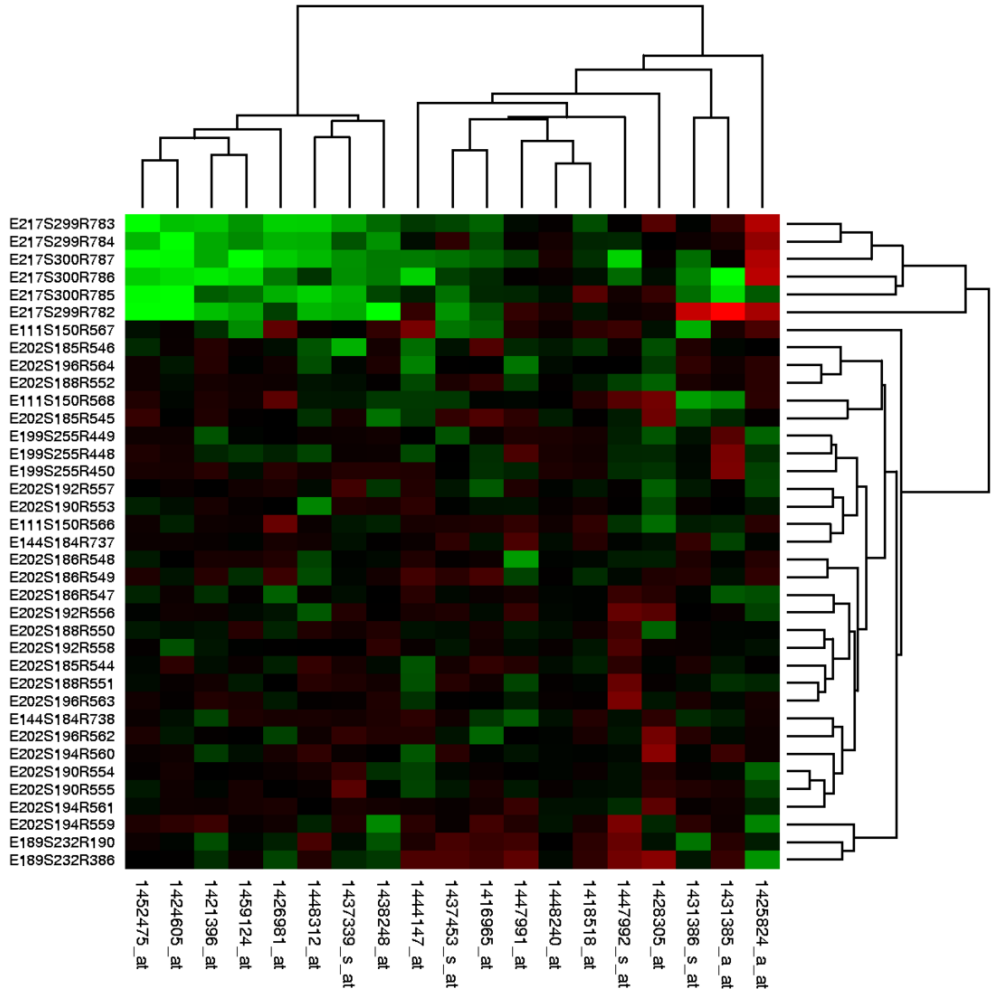
Wagner, G., Zhang, J. The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. *Nat Rev Genet* **12**, 204–213 (2011). https://doi.org/10.1038/nrg2949

The degree of pleiotropy depends on the definition of a trait

morphometrics



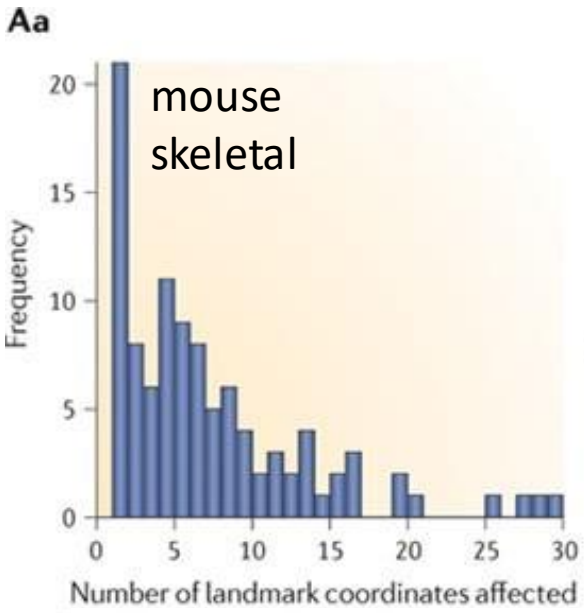
gene expression

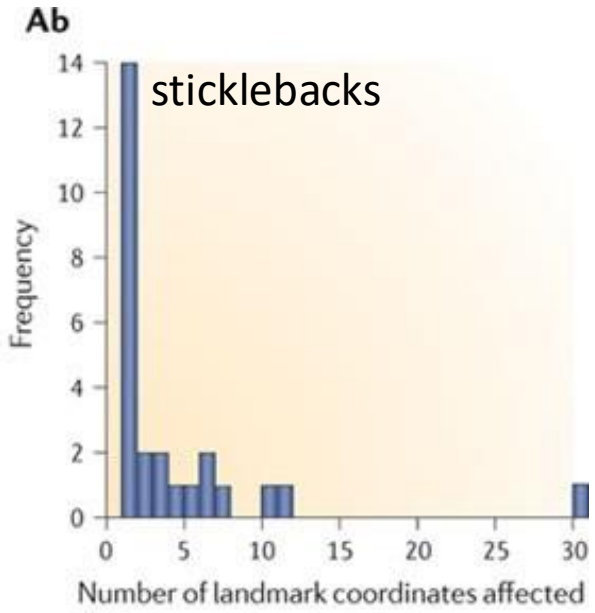


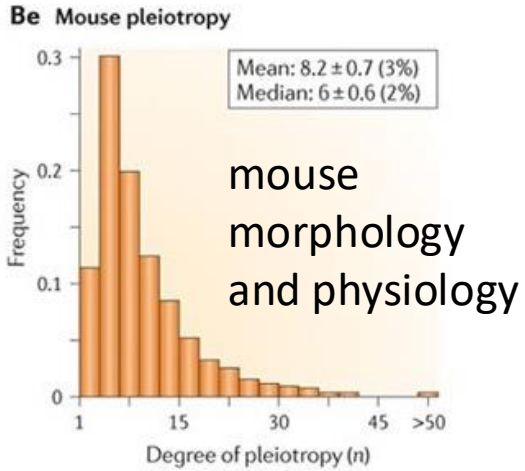
Identifying/Quantifying pleiotropy

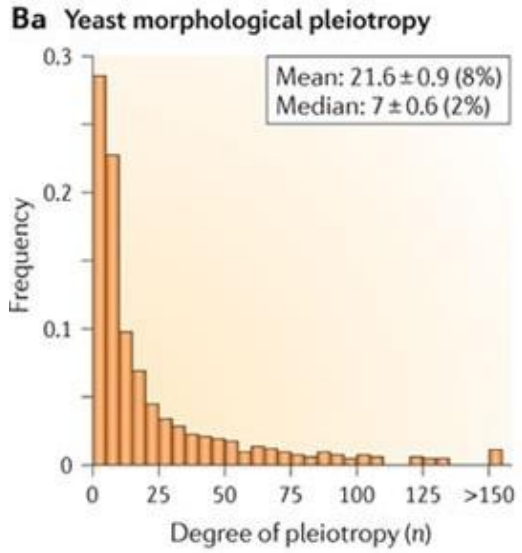
- through gene knockdown studies
- through GWAS
- by studying patterns of genetic variation in multiple dimensions

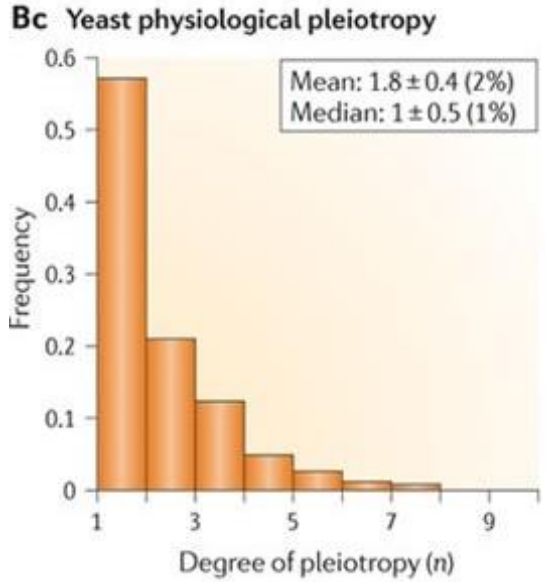
Gene-knockdown studies

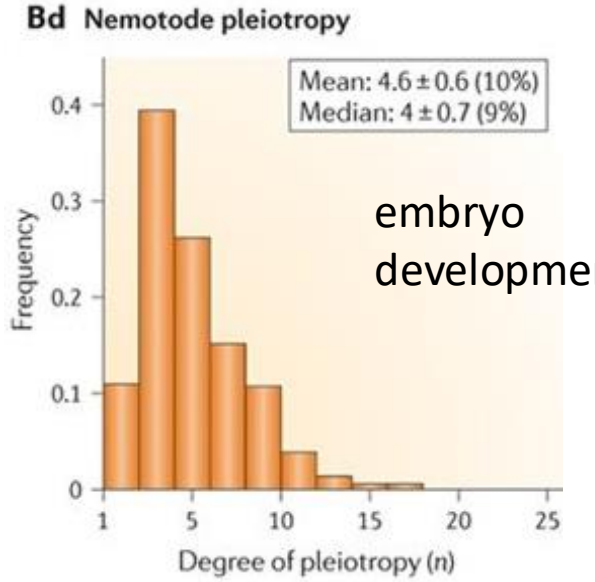










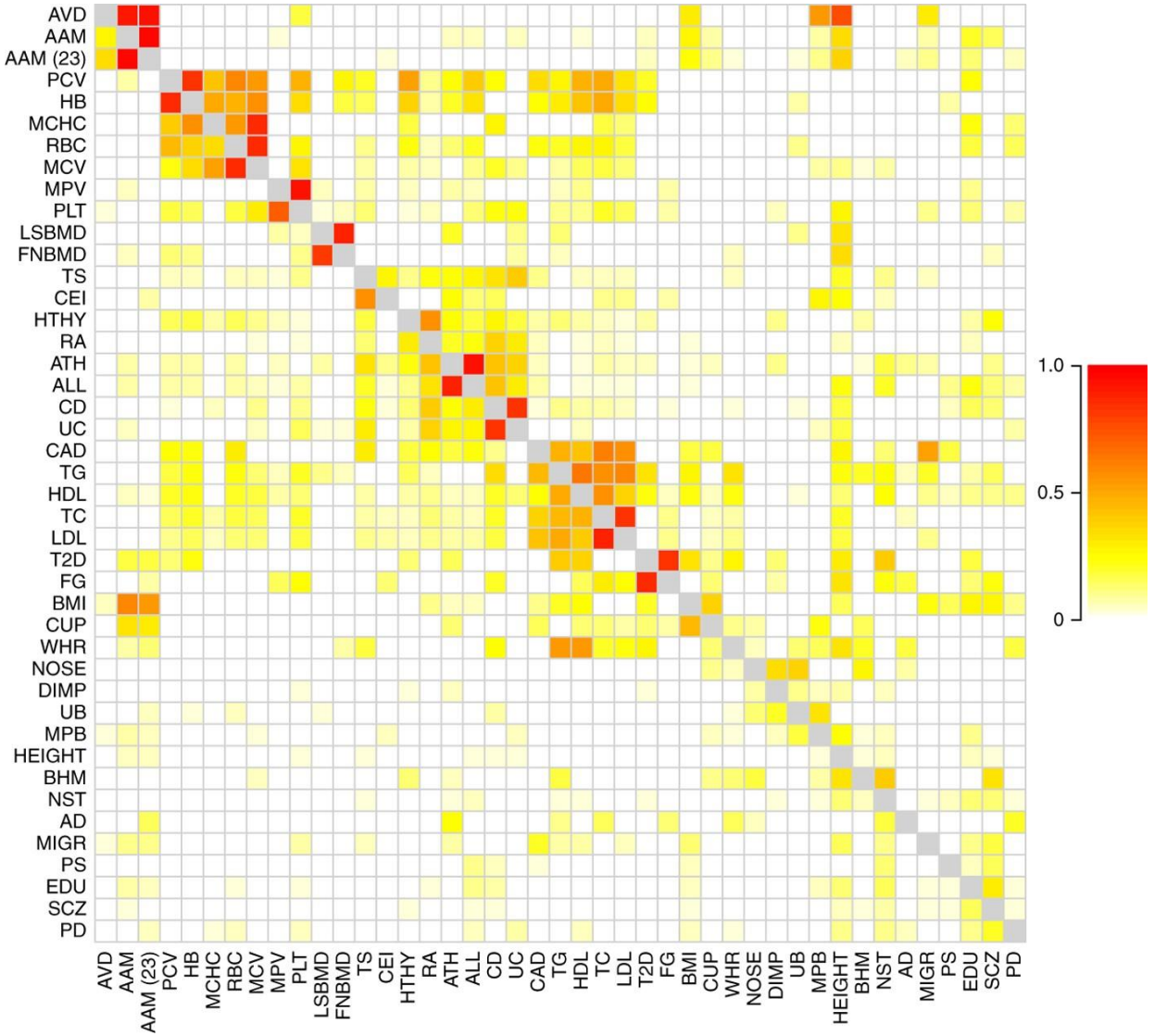


GWAS studies

- find the SNPs that are significantly associated with each of your favourite traits
- count the SNPs that are significantly associated with each pair
- power issues for most studies

GWAS studies

Heatmap shows the proportion of SNPs that are significantly associated with both traits

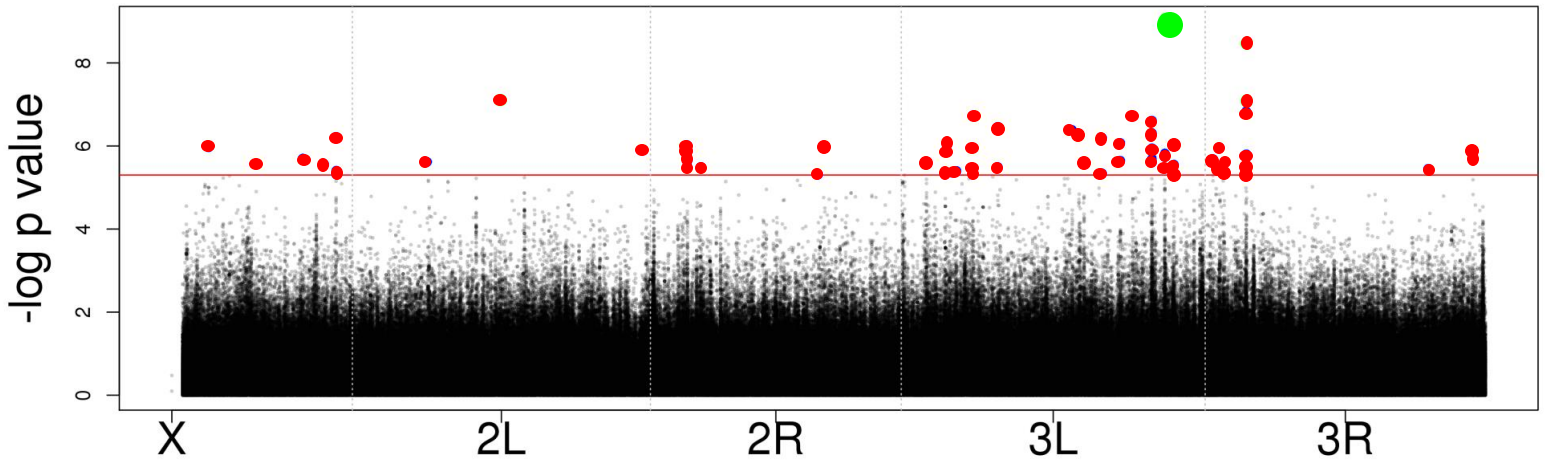


Phenotype
Neurological phenotypes
Alzheimer disease
Migraine
Schizophrenia
Anthropometric and social traits
Beighton hypermobility
Breast size
Body mass index
Educational attainment
Height
Male-pattern baldness
Nearsightedness
Nose size
Waist–hip ratio
Unibrow
Immune-related traits
Any allergies
Asthma
Childhood ear infections
Crohn's disease
Hypothyroidism
Rheumatoid arthritis
Metabolic phenotypes
Age at menarche
Age at menarche (23andMe)
Triglycerides
Total cholesterol
Hematopoietic traits
Hemoglobin
Mean cell hemoglobin concentration
Platelet count
Mean platelet volume

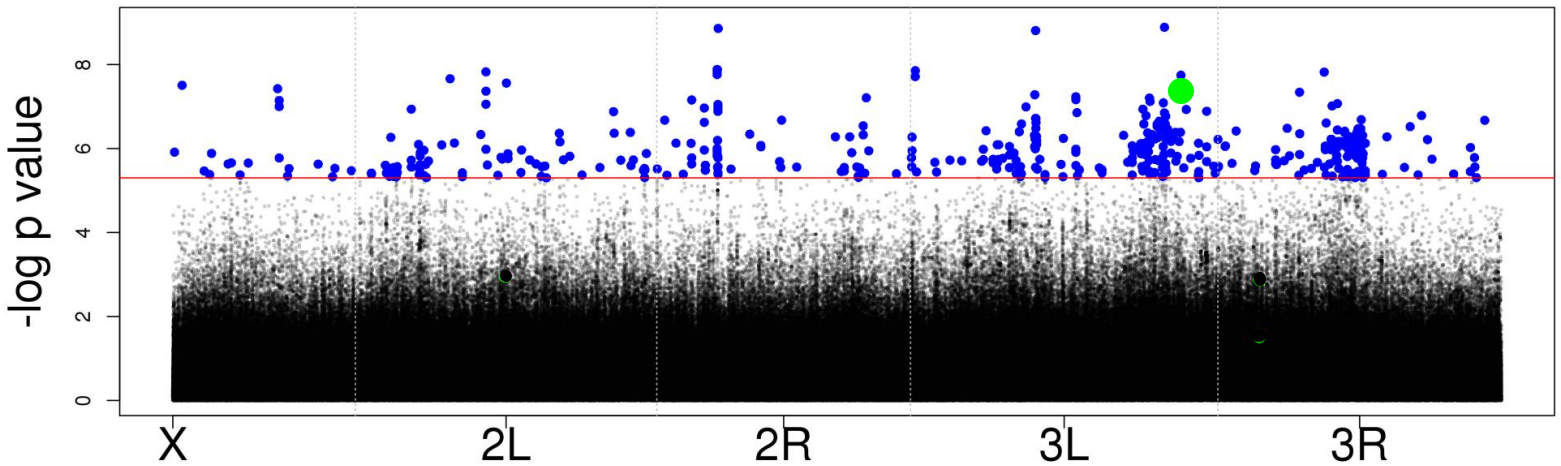
Pickrell, J., Berisa, T., Liu, J. *et al.* Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet* **48**, 709–717 (2016). https://doi.org/10.1038/ng.3570

GWAS studies

Univariate – PC1



Multivariate – 59 dimensions



Only one SNP in common in two analyses

Pitchers, W., Nye, J., Márquez, E.J., Kowalski, A., Dworkin, I. and Houle, D., 2019. A multivariate genome-wide association study of wing shape in Drosophila melanogaster. *Genetics*, *211*(4), pp.1429-1447.

Questions?

Topics we will cover:

- Multivariate quantitative genetics

1. Pleiotropy & Genetic correlations
2. **The G matrix**
3. Genetic constraints

- Selection

1. Empirical methods to estimate selection
2. Empirical results

How to estimate genetic variance using relatedness information

- There are a number of factors that can lead to similarity among relatives *eg.* Common environment, maternal effects, GENETICS

- If phenotypic variation has a genetic basis, then relatives will appear more similar than non-relatives, and the closer the relative the more similar they will appear

- We can use information about the covariance between relatives to partition phenotypic variation into genetic and non-genetic components (and different types of genetic components

How to estimate genetic variance using relatedness information

- Relatives are more likely to share alleles than non-relatives

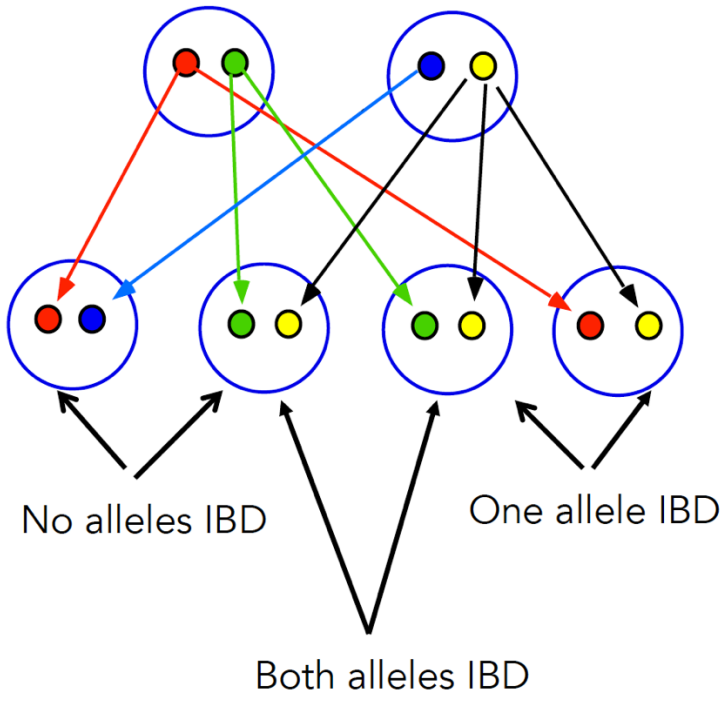
- You share 50% of your genes with your mother or father
- Siblings share 50% of their genes

- What you really mean is that you share alleles with your relatives that are IDENTICAL BY DESCENT (IBD)
- IBD means the same alleles can be traced to a common ancestor

How to estimate genetic variance using relatedness information

- IBD means that a gene is a direct descendent of a specific gene carried by some ancestral individual.
- different than identical/alike by state (IBS/AIS) which means the allele is the same but they have descended from different copies in the reference population

How to estimate genetic variance using relatedness information



$$cov(z, z') = \underline{\underline{r V_A}} + \underline{\underline{u V_D}}.$$

focal individual

relative

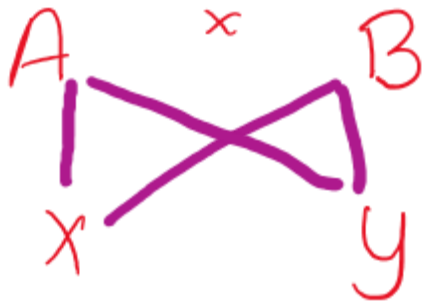
probability that one allele is IBD

probability that both alleles are IBD

Coefficients of Coancestry and Relatedness

- We need to figure out these weights for any type of relatives

- Path counting– identifies the path linking individuals, lets you calculate the probability that their alleles are IBD (coefficient of coancestry)



parents

offspring

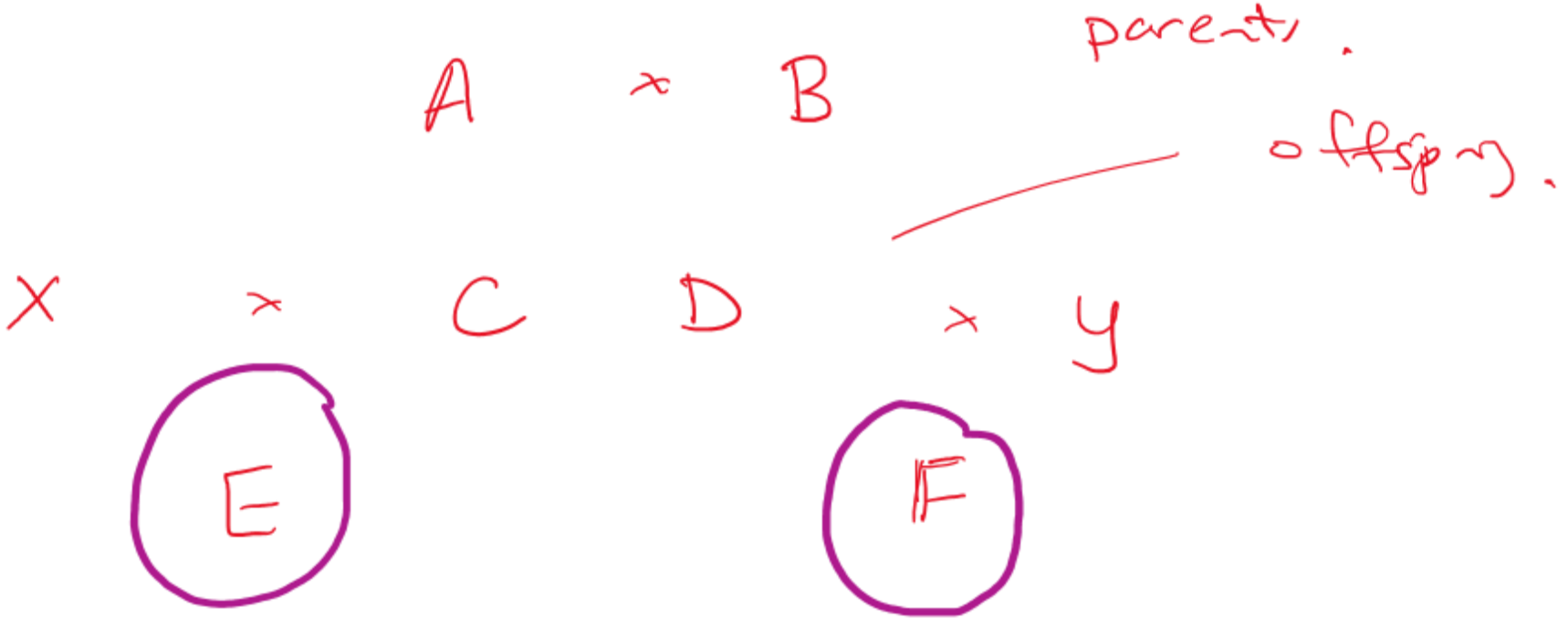
$xAy \qquad xBy$

$$\theta_{xy} = \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^3 = \frac{1}{4}$$

- Coefficient of relatedness is 2X the coefficient of coancestry because it takes into account that either pair of alleles can be shared

Coefficients of Coancestry and Relatedness

- We need to figure out these weights for any type of relatives



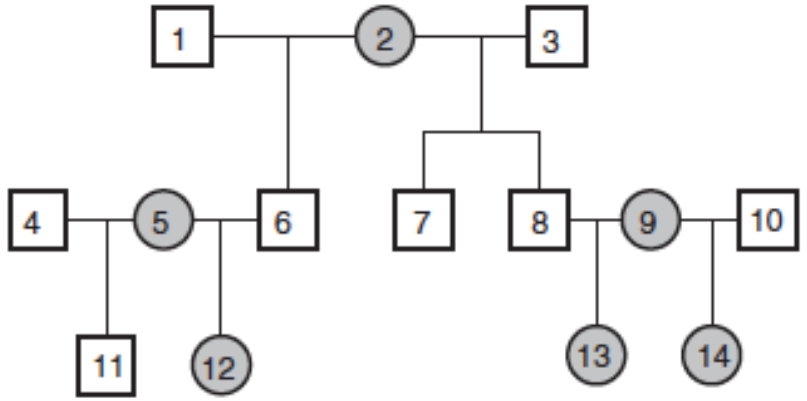
ECADF
ECBDF

$\theta_{EF} = \left(\frac{1}{2}\right)^5 + \left(\frac{1}{2}\right)^5 = \frac{1}{16}$

Coefficient of relatedness $= \frac{1}{8}$

How to estimate genetic variance

(a)

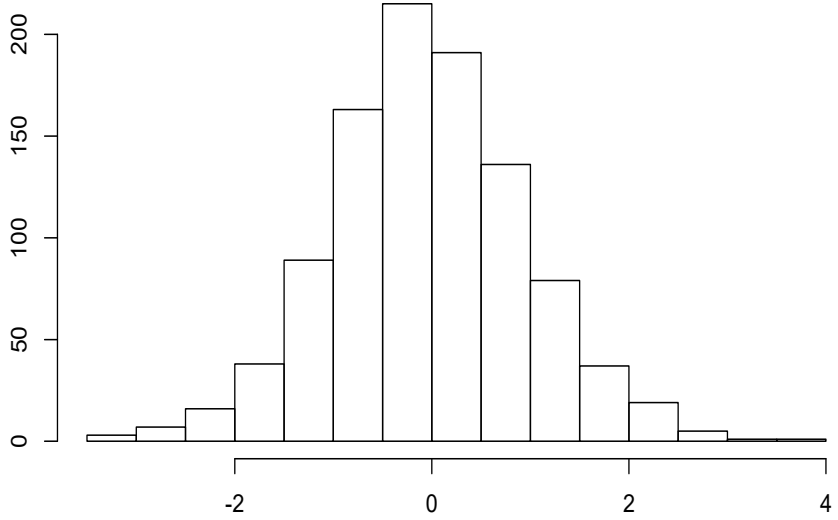


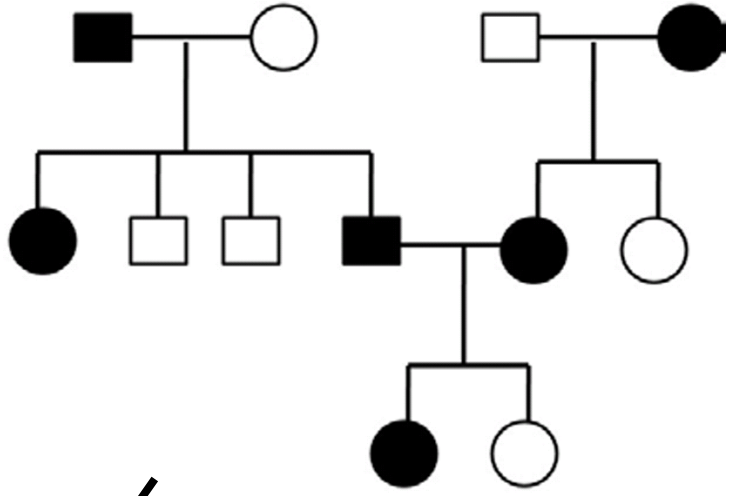
(b)

$$\mathbf{A} = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0.25 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0.5 & 0.5 & 0.5 & 0 & 0 & 0 & 0.25 & 0.25 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0.25 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 \\
0.5 & 0.5 & 0 & 0 & 0 & 1 & 0.25 & 0.25 & 0 & 0 & 0 & 0.5 & 0.125 & 0 \\
0 & 0.5 & 0.5 & 0 & 0 & 0.25 & 1 & 0.5 & 0 & 0 & 0 & 0.125 & 0.25 & 0 \\
0 & 0.5 & 0.5 & 0 & 0 & 0.25 & 0.5 & 1 & 0 & 0 & 0 & 0.125 & 0.5 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0.5 & 0.5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0.5 \\
0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 1 & 0.25 & 0 & 0 \\
0.25 & 0.25 & 0 & 0 & 0.5 & 0.5 & 0.125 & 0.125 & 0 & 0 & 0.25 & 1 & 0.0625 & 0 \\
0 & 0.25 & 0.25 & 0 & 0 & 0.125 & 0.25 & 0.5 & 0.5 & 0 & 0 & 0.0625 & 1 & 0.25 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0.25 & 1
\end{pmatrix}$$

Alastair J. Wilson, et al (2010) An ecologist's guide to the animal model. *Journal of Animal Ecology*, **79**, 13–26.

How to estimate genetic variance





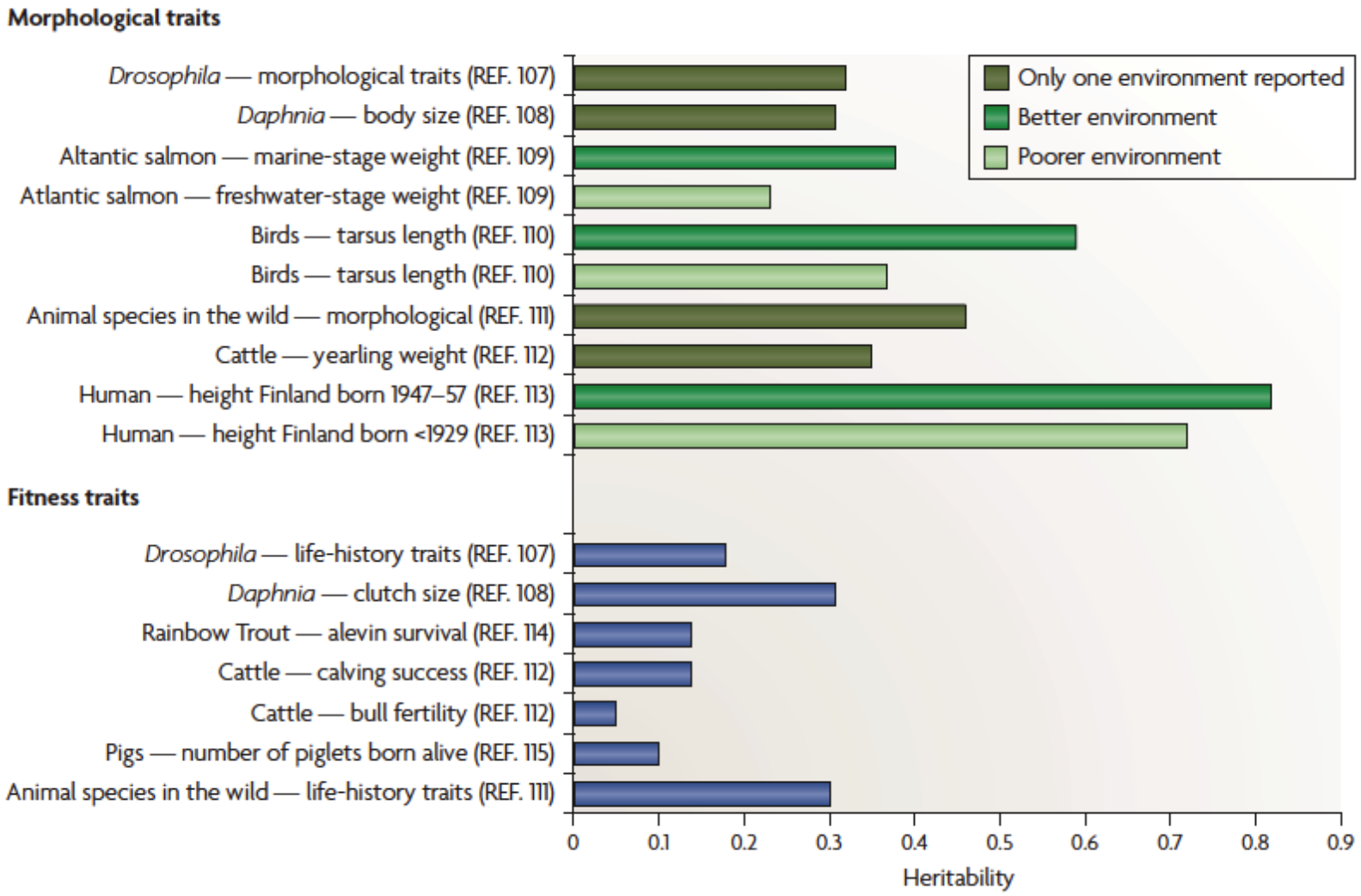
$$\mathbf{y}_k = \mathbf{X}\beta + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{d} + \mathbf{I}\mathbf{e}$$

Additive genetic variance: V_A

Heritability

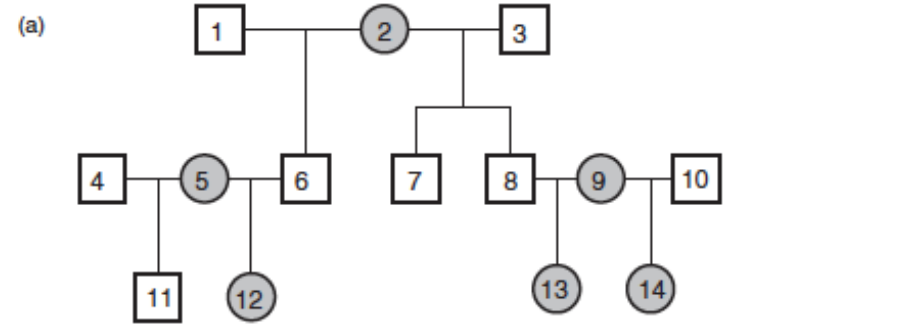
- Most traits have heritability between 20-60%

- Heritability is higher for morphological traits than for life-history or fitness related traits

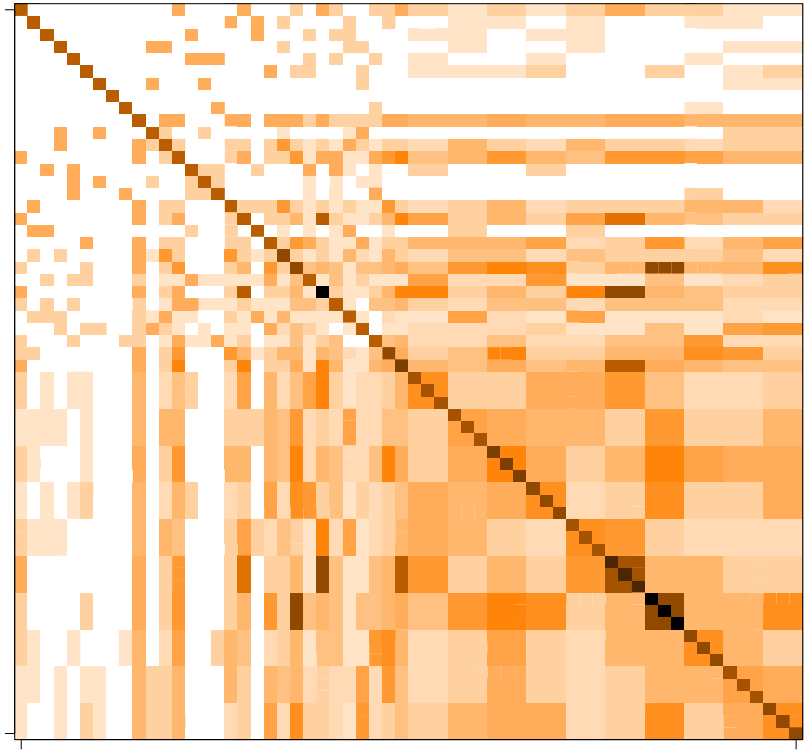


Hoffmann, A.A., Merilä, J. and Kristensen, T.N., 2016. Heritability and evolvability of fitness and nonfitness traits: lessons from livestock. *Evolution*, *70*(8), pp.1770-1779

How to estimate genetic variance



1	0	0	0	0	0.5	0	0	0	0	0	0.25	0	0
0	1	0	0	0	0.5	0.5	0.5	0	0	0	0.25	0.25	0
0	0	1	0	0	0	0.5	0.5	0	0	0	0	0.25	0
0	0	0	1	0	0	0	0	0	0	0.5	0	0	0
0	0	0	0	1	0	0	0	0	0	0.5	0.5	0	0
0.5	0.5	0	0	0	1	0.25	0.25	0	0	0	0.5	0.125	0
0	0.5	0.5	0	0	0.25	1	0.5	0	0	0	0.125	0.25	0
0	0.5	0.5	0	0	0.25	0.5	1	0	0	0	0.125	0.5	0
0	0	0	0	0	0	0	0	1	0	0	0	0.5	0.5
0	0	0	0	0	0	0	0	0	1	0	0	0	0.5
0	0	0	0.5	0.5	0	0	0	0	0	1	0.25	0	0
0.25	0.25	0	0	0.5	0.5	0.125	0.125	0	0	0.25	1	0.0625	0
0	0.25	0.25	0	0	0.125	0.25	0.5	0.5	0	0	0.0625	1	0.25
0	0	0	0	0	0	0	0	0.5	0.5	0	0	0.25	1



Alastair J. Wilson, et al (2010) An ecologist's guide to the animal model. *Journal of Animal Ecology*, **79**, 13–26.

How to estimate genetic variance using genomic information

- Genomic information (sequencing data) can also be used to estimate relatedness

- Treat identity by state (IBS, AIS) as identity by descent (IBD)

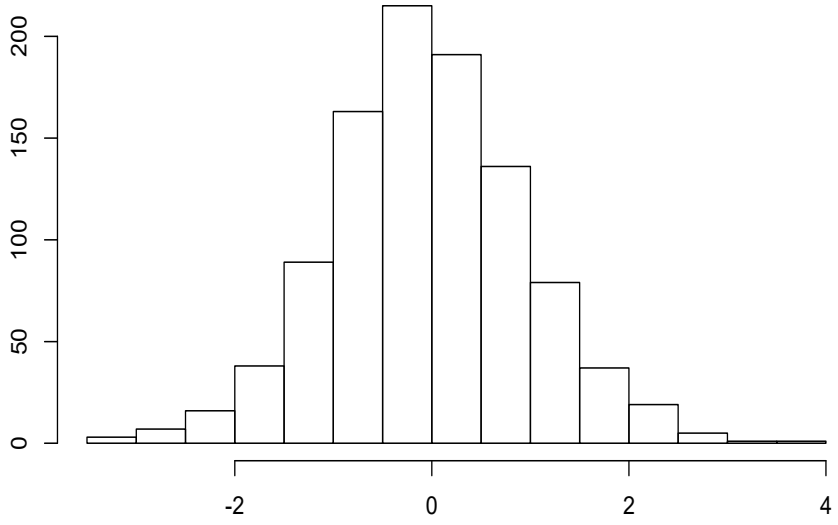
Locus	Individual 1	Individual 2
1	00	10
2	10	00
3	01	01
4	11	11
5	11	01
6	01	10
7	10	10

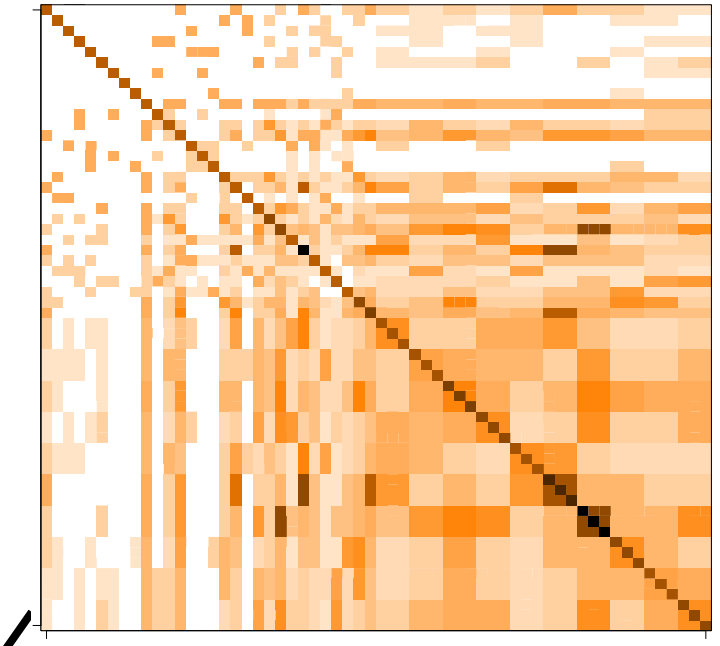
Genomic Relatedness Matrices

$$GRM_{ADD} = \frac{WW'}{trace(WW') / n}$$

W is a marker matrix
n is the number of individuals

How to estimate genetic variance





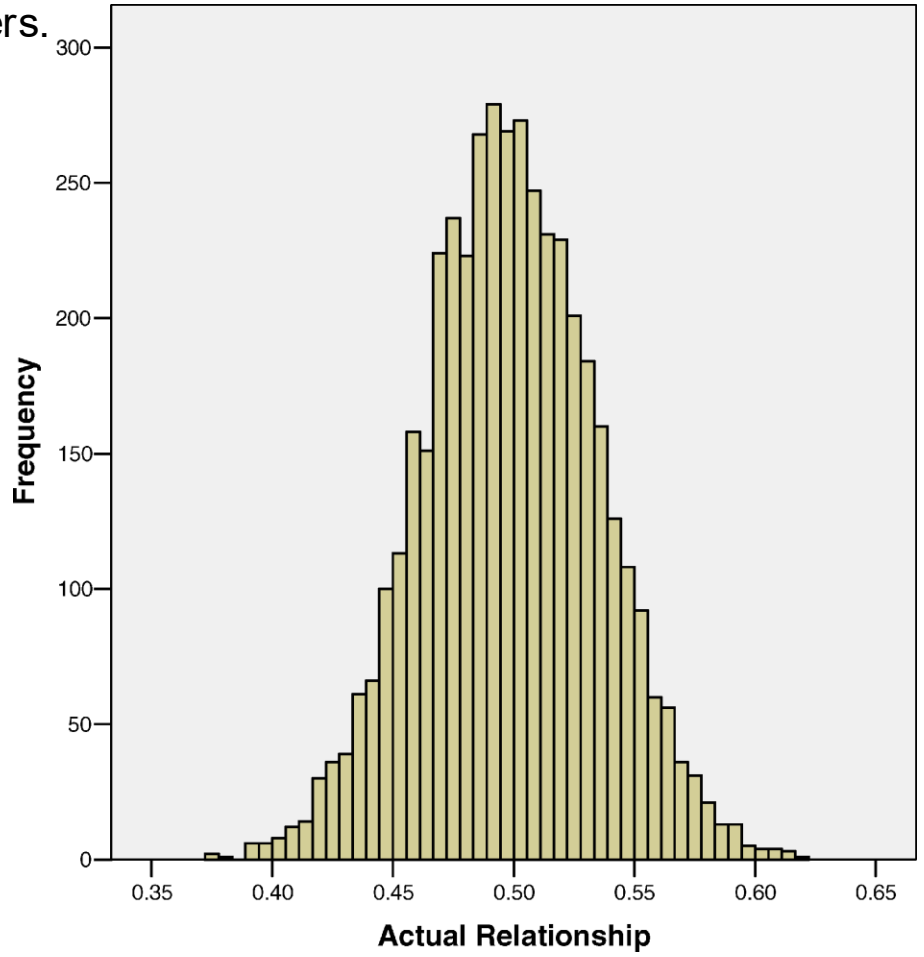
$$\mathbf{y}_k = \mathbf{X}\beta + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{d} + \mathbf{I}\mathbf{e}$$

Additive genetic variance: V_A

How to estimate genetic variance using relatedness information

- Coefficients of relatedness can be greater than 1 with inbreeding, non-random mating etc
- Also remember these are expected values!
- Genomic relatedness can differ quite dramatically from the expected values

Histogram of the genome-wide additive genetic relationships of full-sib pairs estimated from genetic markers.



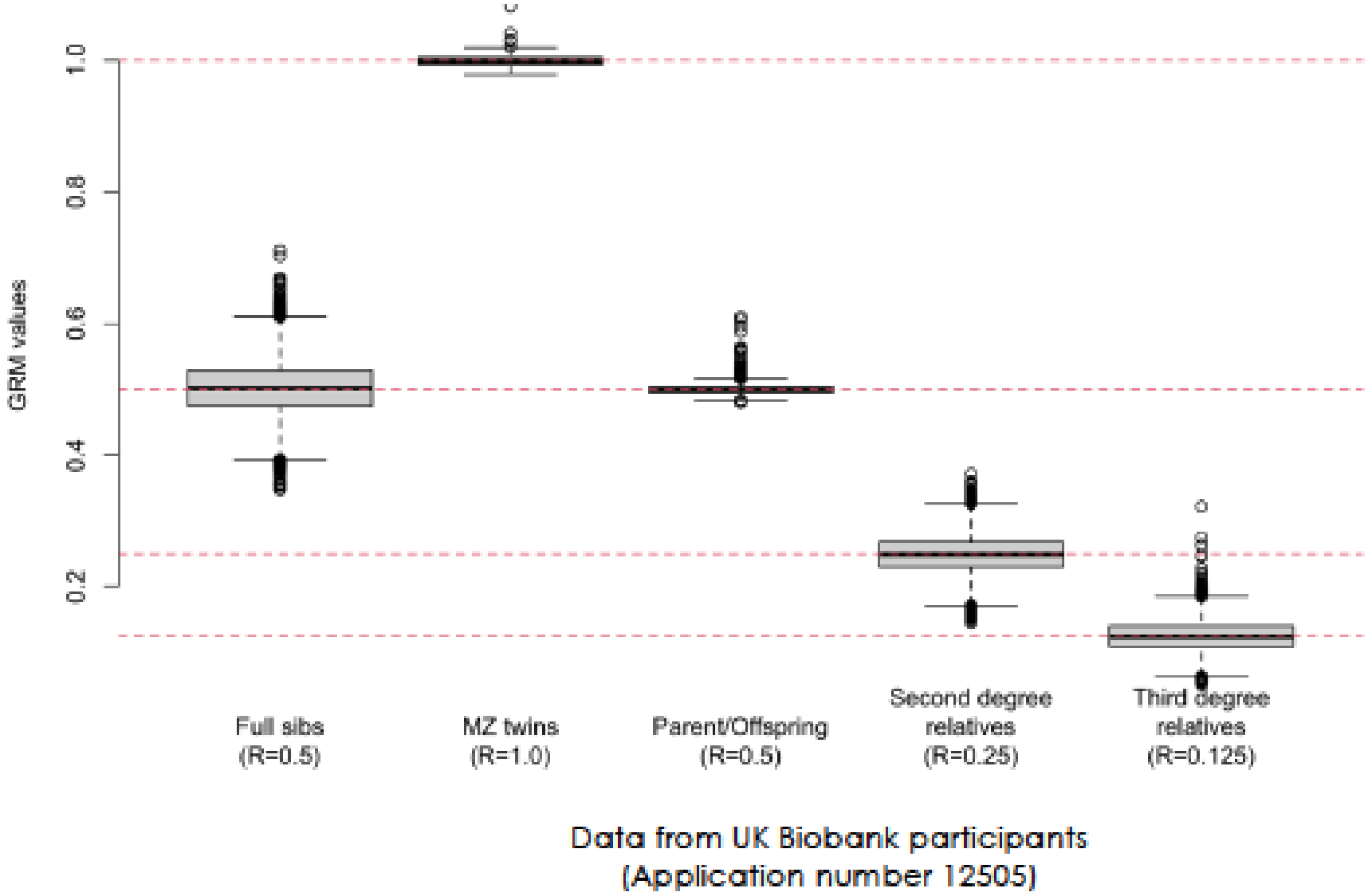
Visscher PM, Medland SE, Ferreira MAR, Morley KI, Zhu G, Cornes BK, et al. (2006) Assumption-Free Estimation of Heritability from Genome-Wide Identity-by-Descent Sharing between Full Siblings. PLoS Genet 2(3): e41.

Difference between pedigree and GRM

- Pedigree is based on IBD
- GRM is based on IBS

- Pedigree contains expected values of relatedness
- GRM contains actual values of relatedness which can differ from expected values due to segregation

Difference between pedigree and GRM



Data from UK Biobank participants
(Application number 12505)

Difference between pedigree and GRM

- Pedigree relationship matrix estimates genetic variance for the group of unrelated founders in the pedigree (ie. base population)

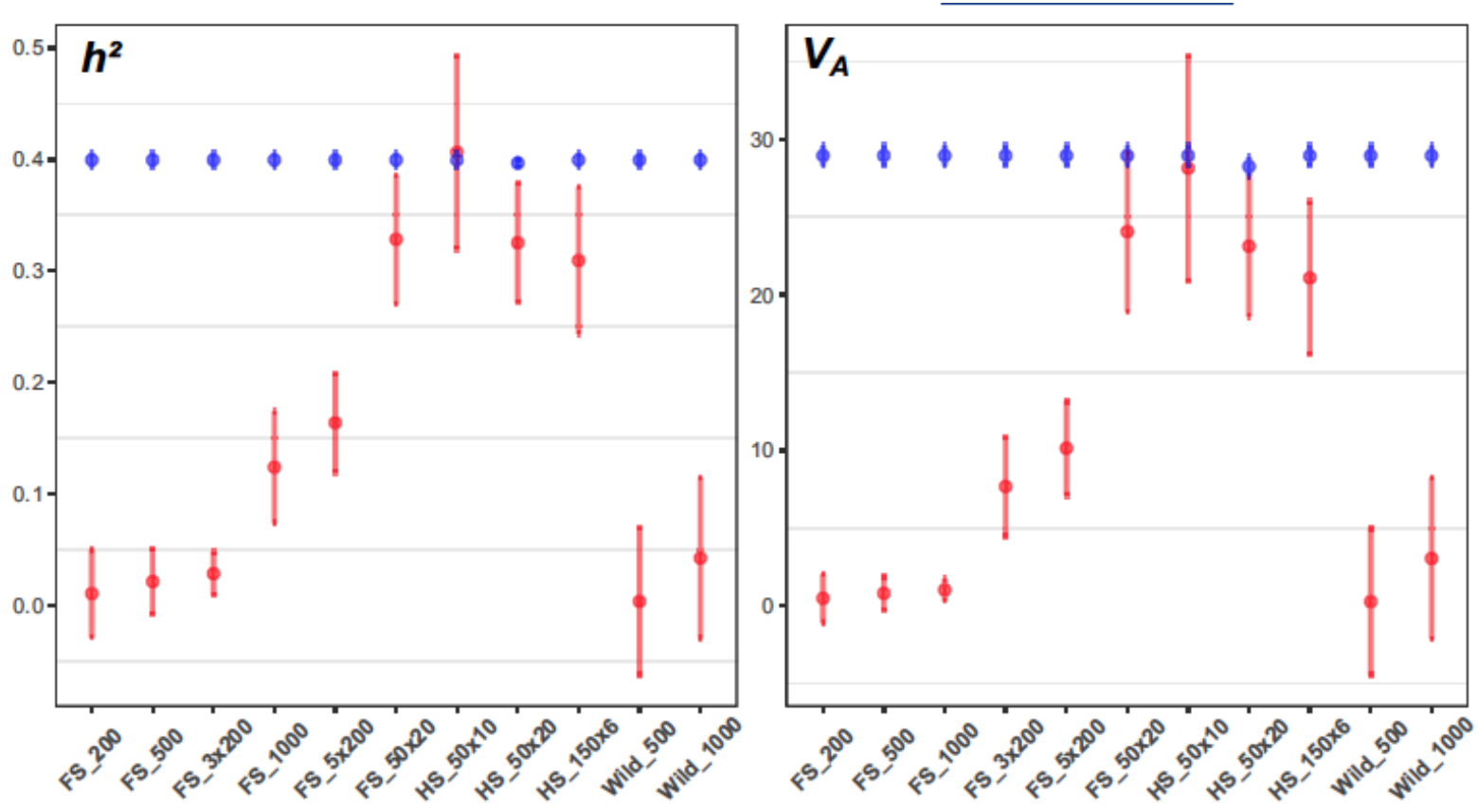
- GRM estimates genetic variance among the set of genotyped individuals

Difference between pedigree and GRM

- Power for estimating genetic variance comes in part from the variance in relatedness among individuals

- Low relatedness can lead to biased estimates of additive variance

Difference between pedigree and GRM



Fraimout, A., Guillaume, F., Li, Z., Sillanpää, M.J., Rastas, P. and Merilä, J., 2024. Dissecting the genetic architecture of quantitative traits using genome-wide identity-by-descent sharing. *Molecular Ecology*, *33*(6), p.e17299

Difference between pedigree and GRM

- Power for estimating genetic variance comes in part from the variance in relatedness among individuals
- Low relatedness can lead to biased estimates of additive variance
- Which SNPs are included in GRM can also lead to biased estimates of variance

Missing heritability

Variants affecting human height:

2008: ~**12** SNPs explain ~**2%** variance[1]
2008: ~**30** SNPs explain ~**4%** variance[2]
2010: ~**180** SNPs explain ~**10%** variance[3]
2011: ~**200** SNPs explain ~**10%** variance[4]
2014: ~**700** SNPs explain ~**20%** variance[5]

….
….
….

2022: ~**12,111** SNPs explain ~**50%** variance

[1] Lettre, G. *et al*. (2008) *Nat. Genet.* **40**, 584–591; [2] Gudbjartsson *et al* . (2008) *Nat. Genet.* **40**, 609-615; [3]Allen *et al* (2010) *Nature* **467**, 832–838; [4] Zhang G, et al. (2011) *PLoS ONE* **6(12)**: e29475; [5] Wood, A. *et al*. (2014) *Nat. Genet.* **46**, 1173-1186;

Missing heritability for height has been found

Article

A saturated map of common genetic variants associated with human height

https://doi.org/10.1038/s41586-022-05275-y

Received: 19 December 2021

Accepted: 24 August 2022

Published online: 12 October 2022

Open access

Check for updates

Common single-nucleotide polymorphisms (SNPs) are predicted to collectively explain 40–50% of phenotypic variation in human height, but identifying the specific variants and associated regions requires huge sample sizes[1]. Here, using data from a genome-wide association study of 5.4 million individuals of diverse ancestries, we show that 12,111 independent SNPs that are significantly associated with height account for nearly all of the common SNP-based heritability. These SNPs are clustered within 7,209 non-overlapping genomic segments with a mean size of around 90 kb, covering about 21% of the genome. The density of independent associations varies across the genome and the regions of increased density are enriched for biologically

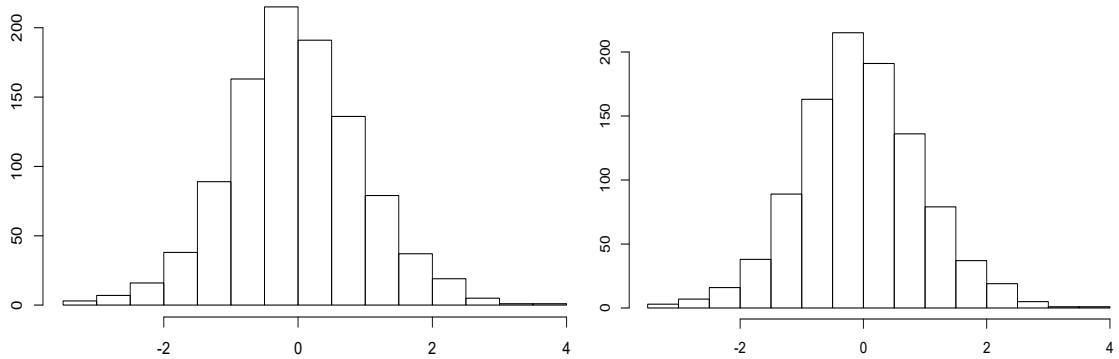
- data from ~5.4 million people
- identified 12,111 genetic variants affecting height that cover ~21% of the genome
- together explain 50% of the phenotypic variation in height

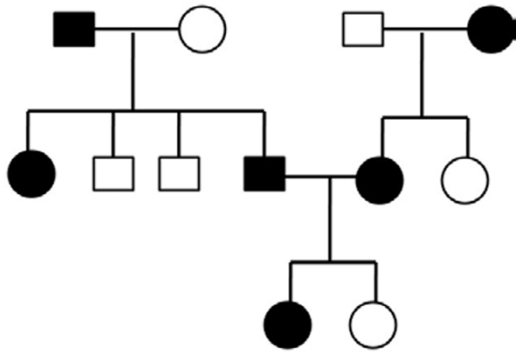
Missing heritability

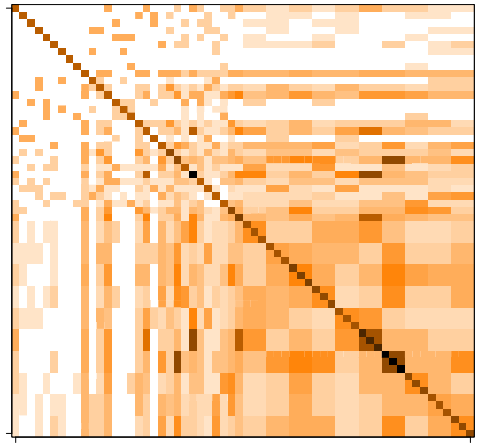
$$h^2_{\mathrm{GWAS}} \leq h^2_{\mathrm{SNP}} \leq h^2$$

h^2-h^2_{GWAS} is often denoted the "missing" heritability (e.g., 5% vs 80%).
h^2_{SNP}-h^2_{GWAS} is often denoted the "hidden/hiding" heritability.
h^2-h^2_{SNP} is denoted the (still) missing heritability.

Estimating genetic covariance: G-matrix







Genetic variance in each trait

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{d} + \mathbf{I}\mathbf{e}$$

$$\begin{bmatrix} \mathrm{var}_{11} & \mathrm{cov}_{12} & \cdots & \mathrm{cov}_{1n} \\ \mathrm{cov}_{21} & \mathrm{var}_{22} & \cdots & \mathrm{cov}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}_{n1} & \mathrm{cov}_{n2} & \cdots & \mathrm{var}_{nn} \end{bmatrix}$$

Use a linear model to partition phenotypic variation

Genetic covariance between the traits

Software to estimate genetic variance

Table 2. A list of some available software packages that can be used to run animal models, with details of whether the software is freely available, the method of statistical inference implemented (REML: restricted maximum likelihood; MCMC: Markov Chain Monte Carlo) and on-line sources of further information. This is not an exhaustive list and merely reflects the software the authors are familiar with

Software	Free to download/ use?	Inference	Notes/Website	Documentation
ASReml	No	REML	Owned and licensed by VSN International Ltd http://www.vsni.co.uk/software/asreml/	http://www.vsni.co.uk/resources/doc/asreml2/UserGuide.pdf
ASReml-R	No	REML	Commercially available R interface for ASReml http://www.vsni.co.uk/software/asreml/	http://www.vsni.co.uk/resources/doc/asreml-R.pdf
DMU	Yes	REML or MCMC	http://www.dmu.agrsci.dk/	http://www.dmu.agrsci.dk/dmuv6_guide-R4-6-7.pdf
MCMCglmm	Yes	MCMC	R package http://cran.r-project.org/web/packages/MCMCglmm/index.html	http://cran.r-project.org/web/packages/MCMCglmm/MCMCglmm.pdf
WOMBAT	Yes	REML	Replaces DFREML http://agbu.une.edu.au/~kmeyer/wombat.html	http://agbu.une.edu.au/~kmeyer/WOMBAT/WWW/manual.html
VCE	Yes	REML or MCMC	http://vce.tzv.fal.de/software	ftp://ftp.tzv.fal.de/pub/latest_vce/doc/vce6-manual-3-1-A4.pdf

- GCTA
- GREML
- + others

Alastair J. Wilson, et al (2010) An ecologist's guide to the animal model. *Journal of Animal Ecology*, **79**, 13–26.

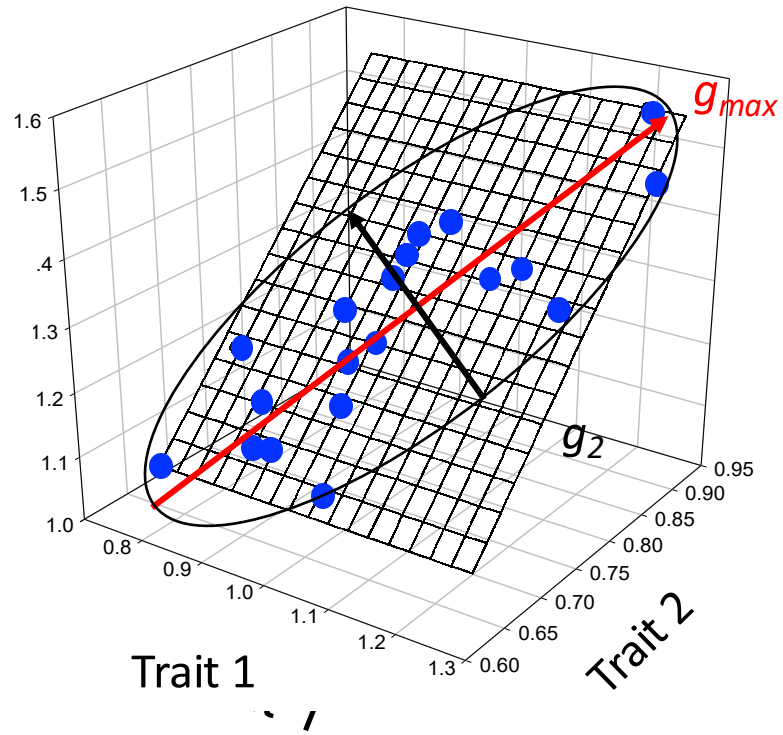
Estimating genetic covariance: G-matrix

	Trait 1	Trait 2	Trait 3
Trait 1	1	-0.625	0.294
Trait 2	-0.625	1	0.563
Trait 3	0.294	0.563	1

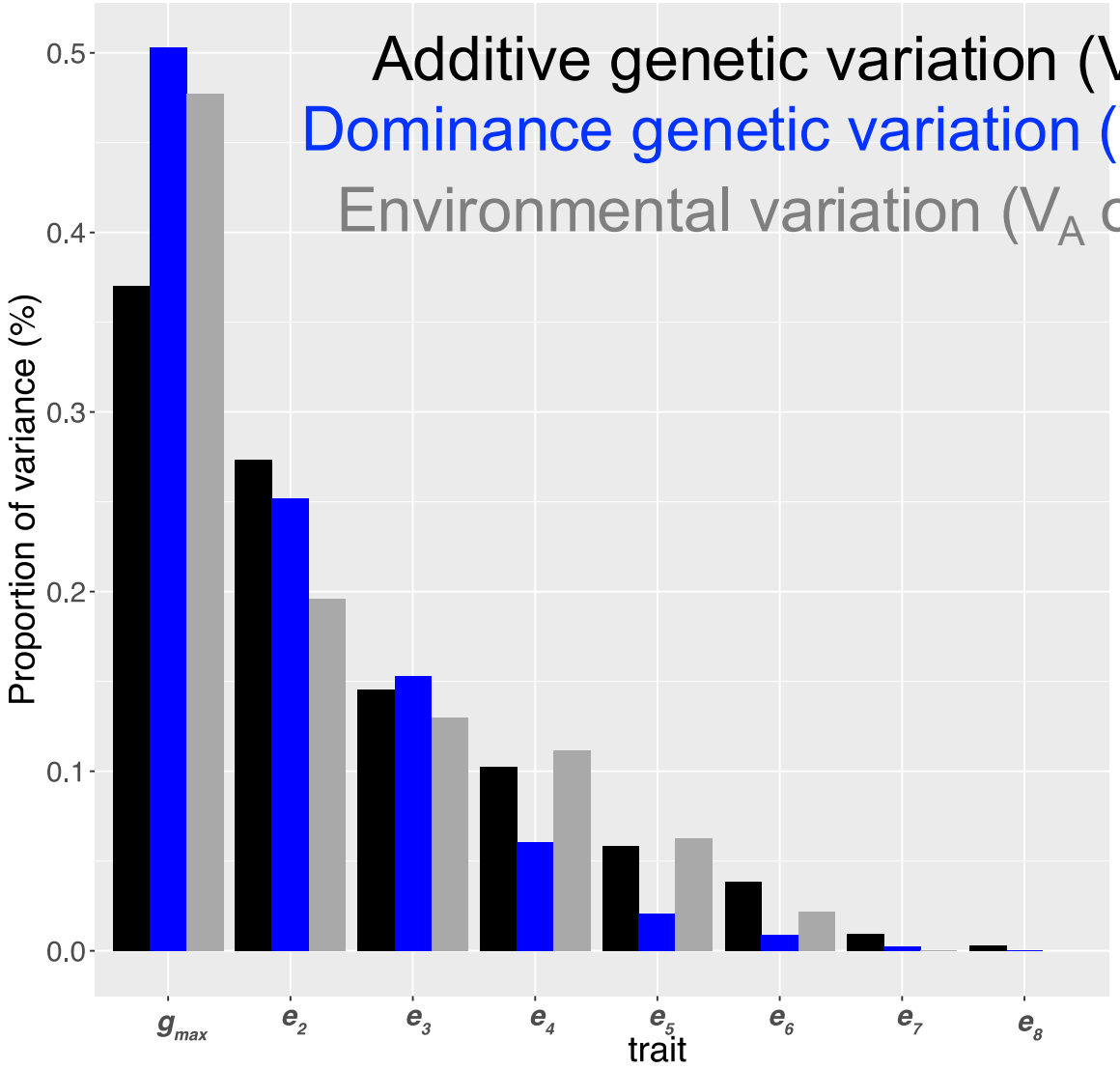
Eigenanalysis

	Eigenvalues	Eigenvectors		
g_{max}	1.708	-0.509	0.764	0.396
g_2	1.292	0.656	0.048	0.753
g_{min}	0	-0.556	-0.643	0.525

No genetic variance



Genetic variation tends to be concentrated in certain trait combinations



- almost every trait we go out and measure has additive genetic variation

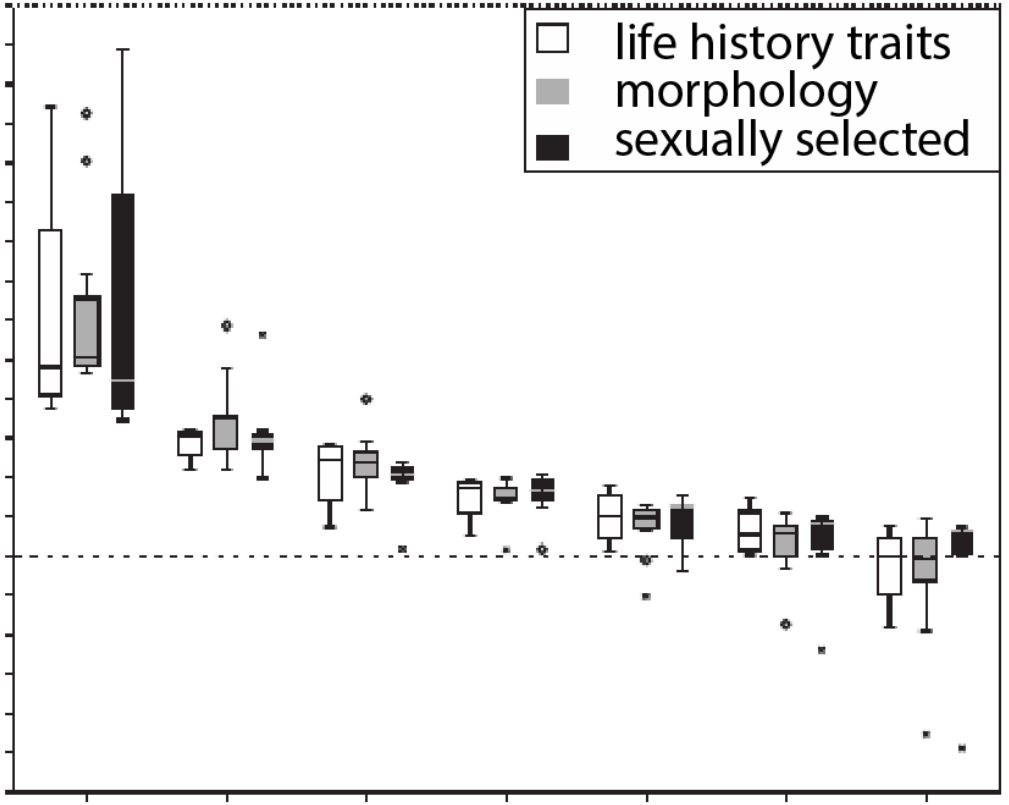
- the genetic variation in a set of traits is often restricted to a few multivariate combinations of those traits

Sztepanacz and Blows (2015) *Genetics* 200: 371-384

Sztepanacz *et al* (2017) *Genetics* 206: 2185-2198

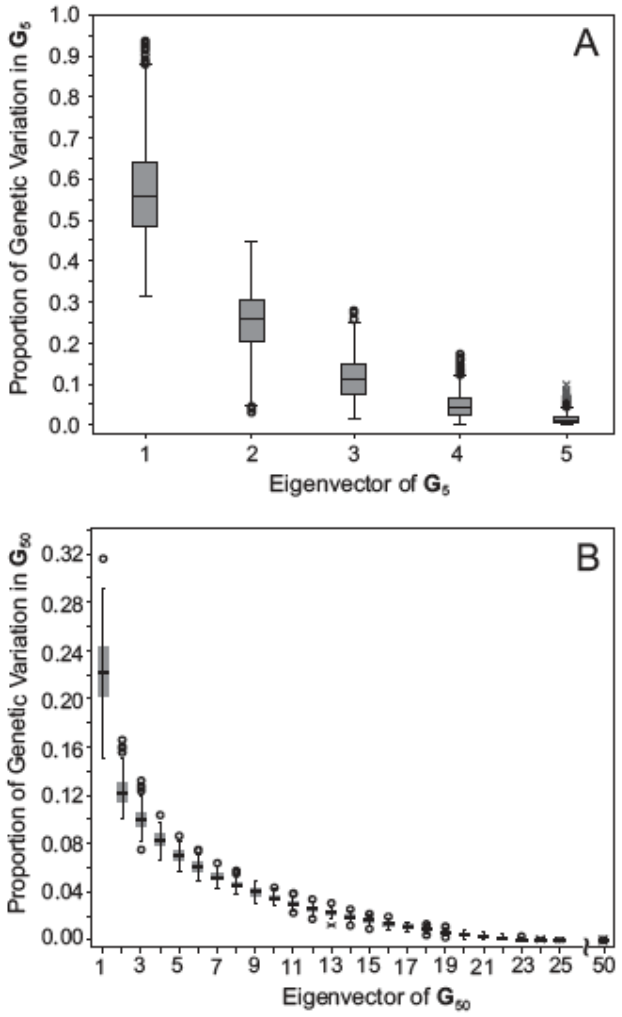
Genetic variance in unevenly distributed across **G**

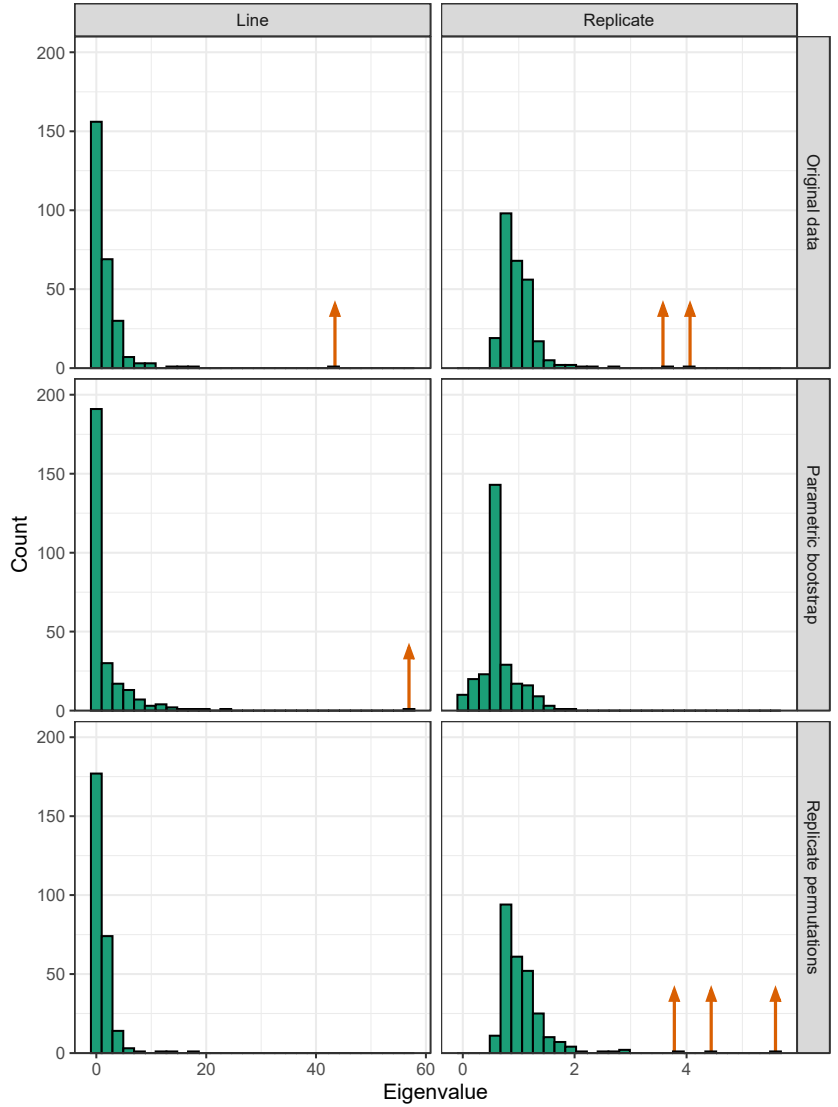
- a few traits have most genetic variation
- many traits have little genetic variation
- suggests there are few independent genetic dimensions underlying organisms



The typical distribution of eigenvalues from genetic covariance matrices.
From McGuigan and Blows 2015 Mol Ecol

Genetic variance in unevenly distributed across **G**





Blows, M.W., Allen, S.L., Collet, J.M., Chenoweth, S.F. and McGuigan, K., 2015. The phenome-wide distribution of genetic variance. *The American Naturalist*, *186*(1), pp.15-30.

Pavlyshyn, D., Johnstone, I.M. and Sztepanacz, J.L., 2022. Comparison of REML methods for the study of phenome-wide genetic variation. *arXiv preprint arXiv:2210.11709*.

Questions?

Topics we will cover:

- Multivariate quantitative genetics

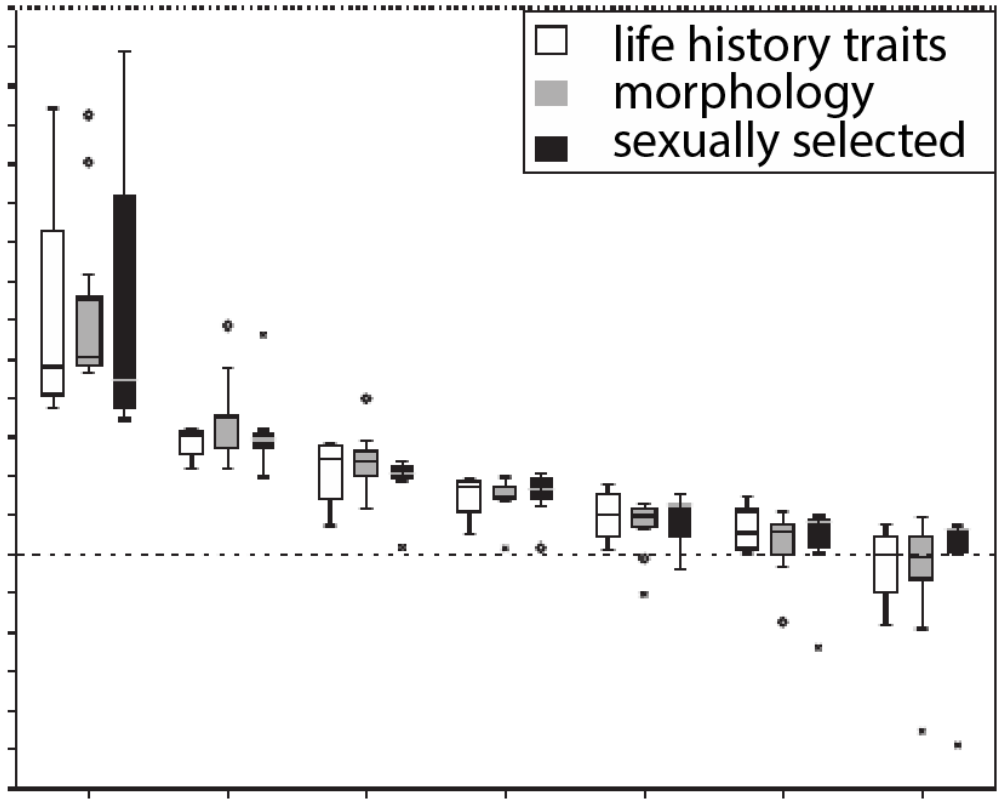
1. Pleiotropy & Genetic correlations
2. The G matrix
3. **Genetic constraints**

- Selection

1. Empirical methods to estimate selection
2. Empirical results

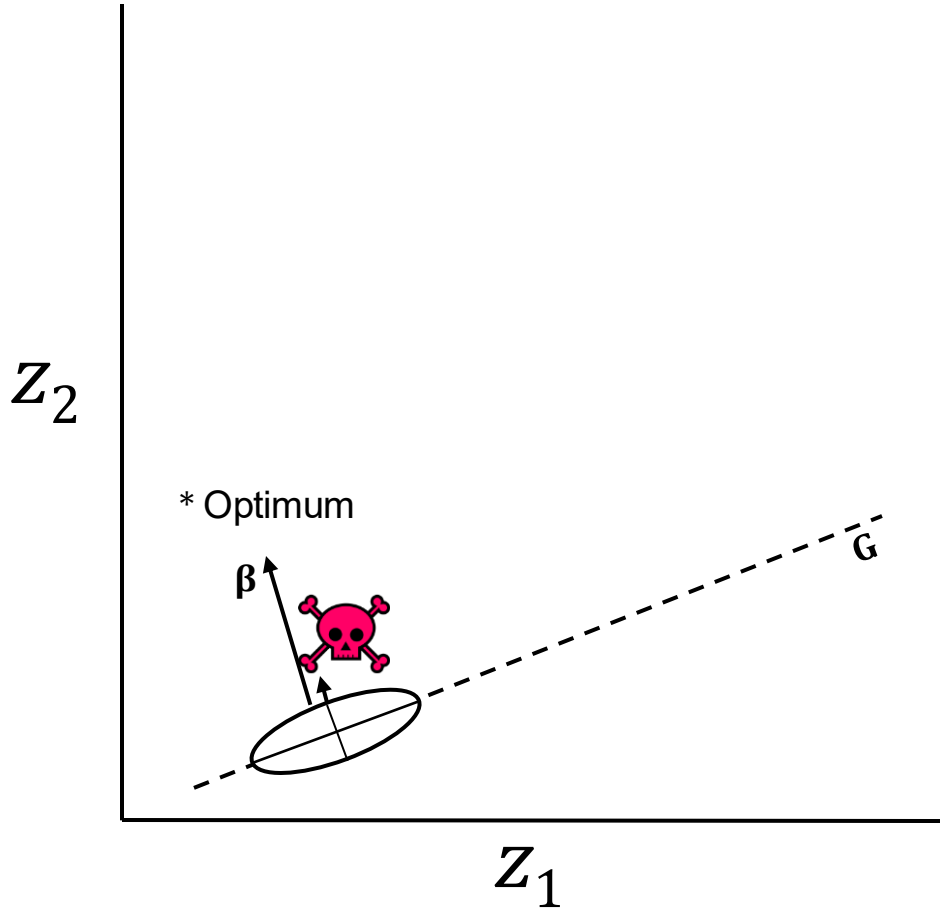
Uneven distribution of genetic variance can lead to evolutionary constraints

- nearly-null subspace of genetic variation (Mezey and Houle 2005)
- qualitative vs. quantitative constraints
- quantitative constraints can become qualitative because of demography (Gomulkiewicz and Houle 2009)



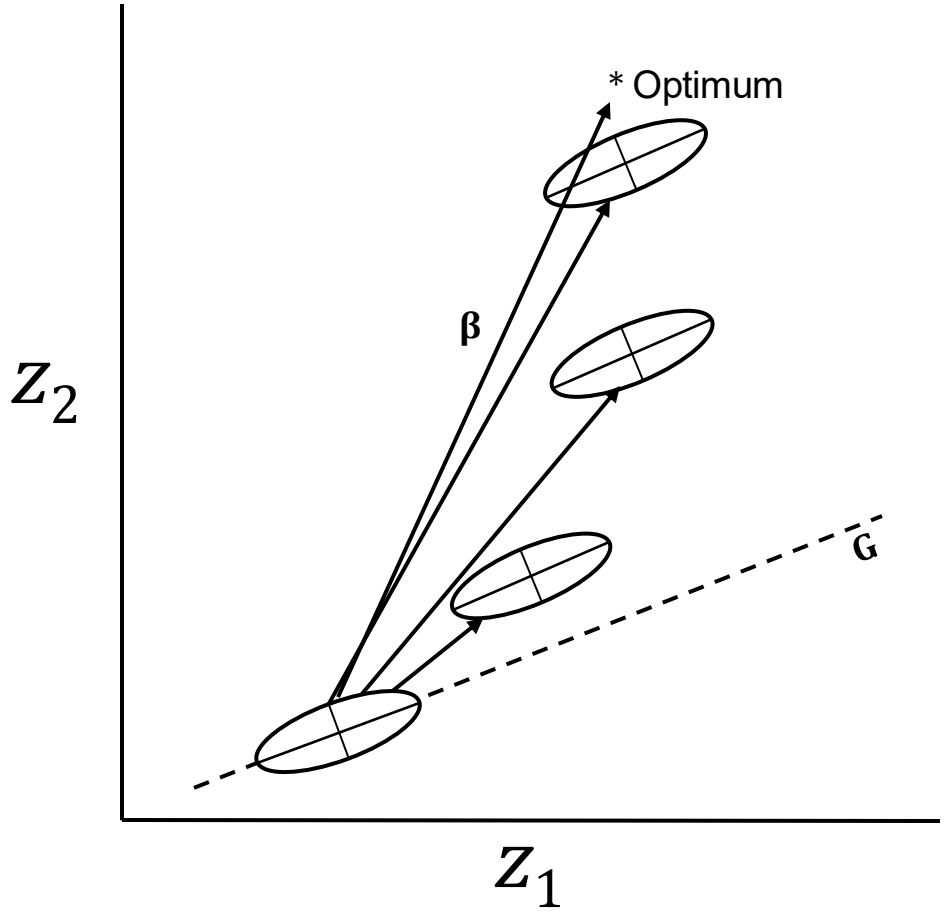
The typical distribution of eigenvalues from genetic covariance matrices.
From McGuigan and Blows 2015 Mol Ecol

Uneven distribution of genetic variance can lead to evolutionary constraints



Schluter 1996

Uneven distribution of genetic variance can lead to evolutionary constraints



Schluter 1996

How can we quantify genetic constraints

- maximum evolvability
- total genetic variance
- average evolvability
- effective number of dimensions
- eigenvalue variance
- eigenvalue evenness
- number of 0 eigenvalues

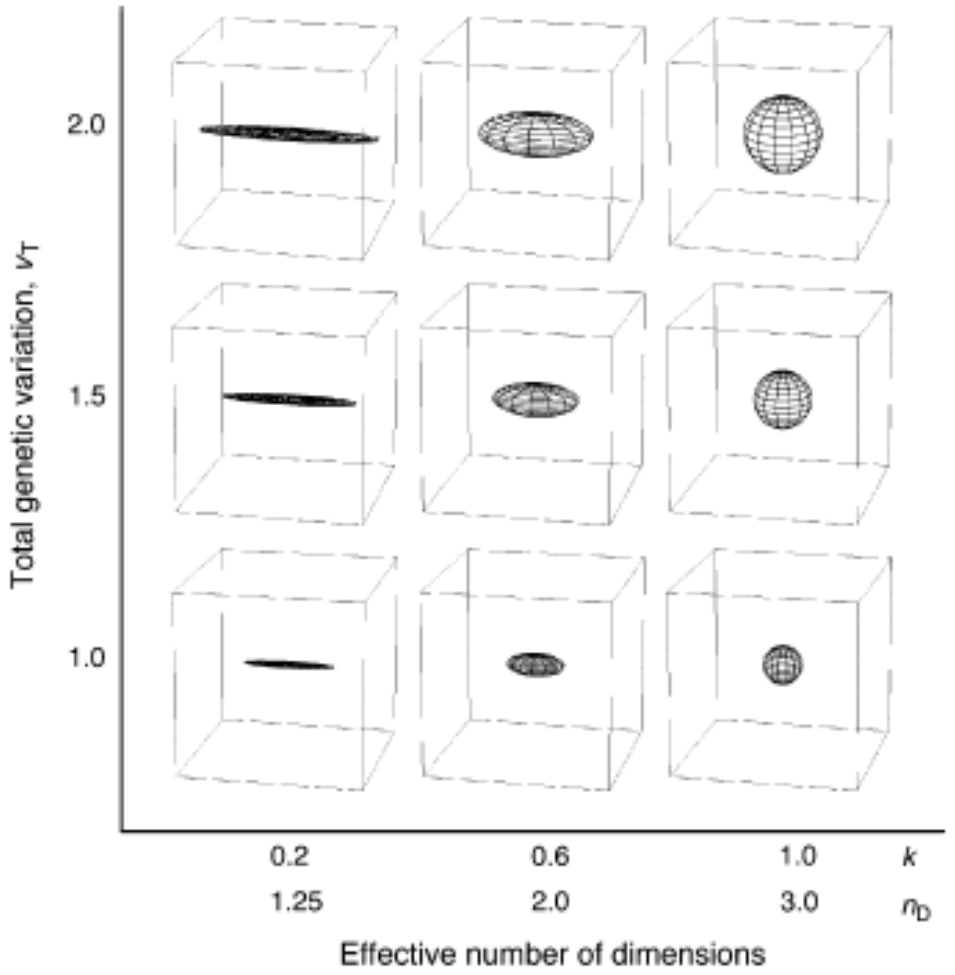


Fig. 1 Patterns of multivariate variation with $n = 3$ traits as a function of the effective number of dimensions (n_D) and the total genetic variation (v_T)

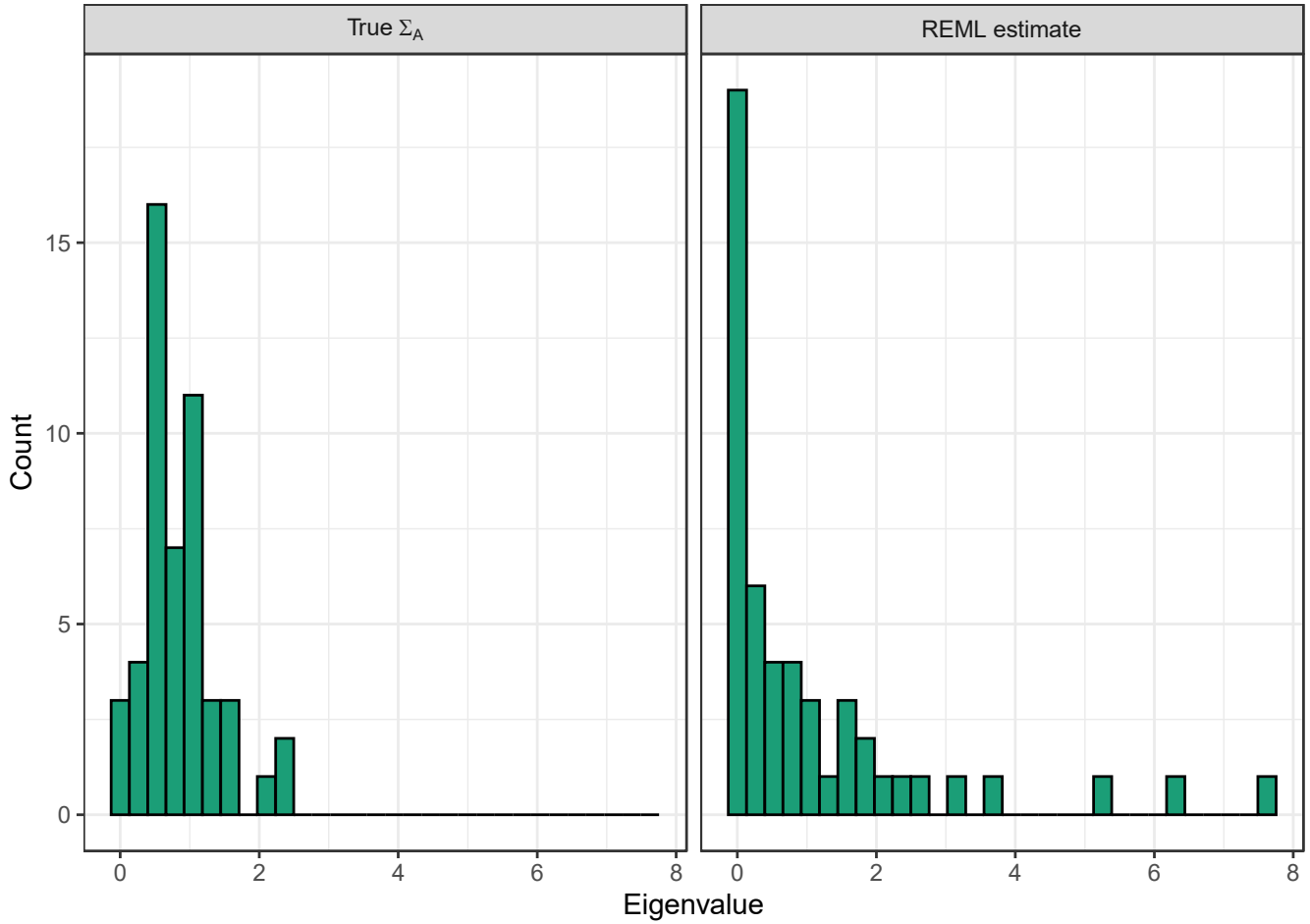
The curse of dimensionality

- to estimate the genetic covariance between two traits we need to estimate three parameters at the genetic level
- we need lots more data to estimate genetic covariances than genetic variances
- estimates of genetic covariances often have high standard errors
- Systematic bias in estimation of eigenvalues

Systematic biases in the estimation of eigenvalues

- with MANOVA or REML leading eigenvalues are overestimated and trailing eigenvalues are underestimated

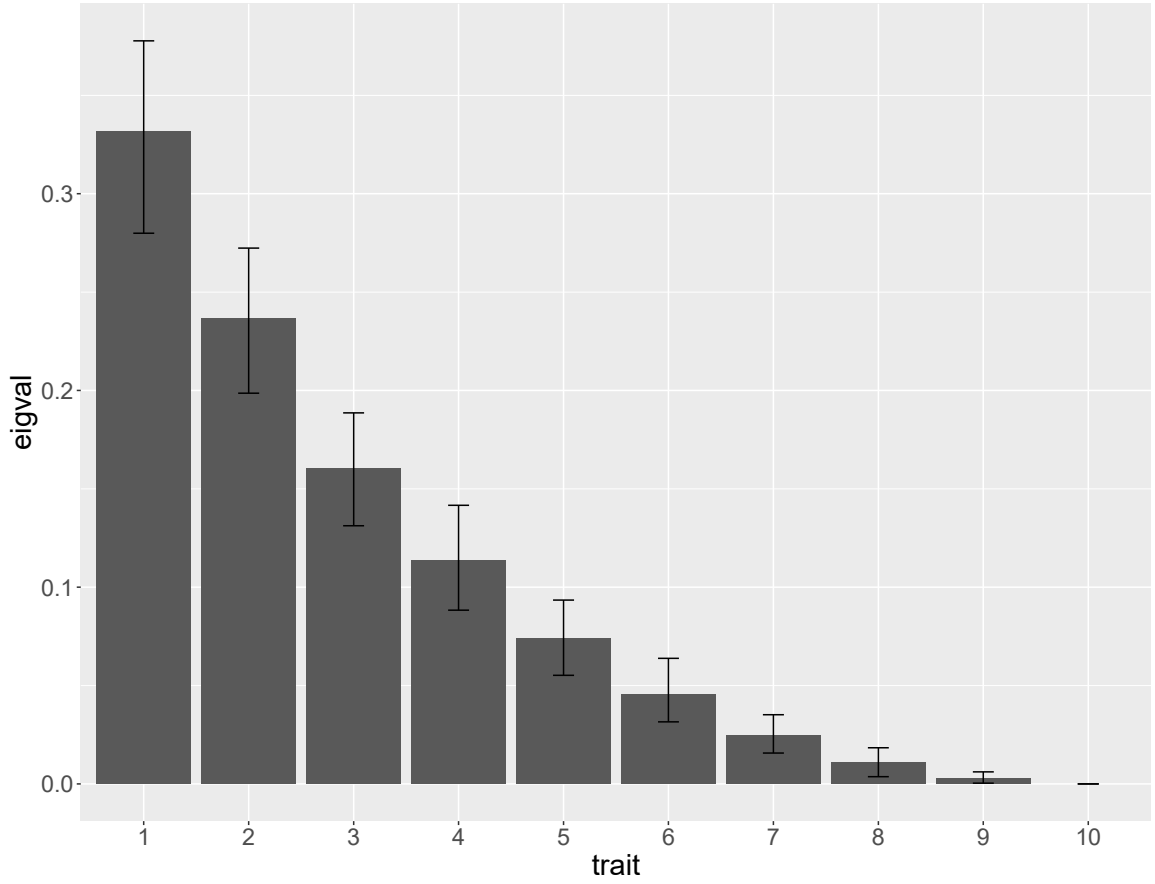
- the uneven distribution of genetic variance we saw earlier (that we interpret to arise as a consequence of pleiotropy) is the same pattern as statistical bias

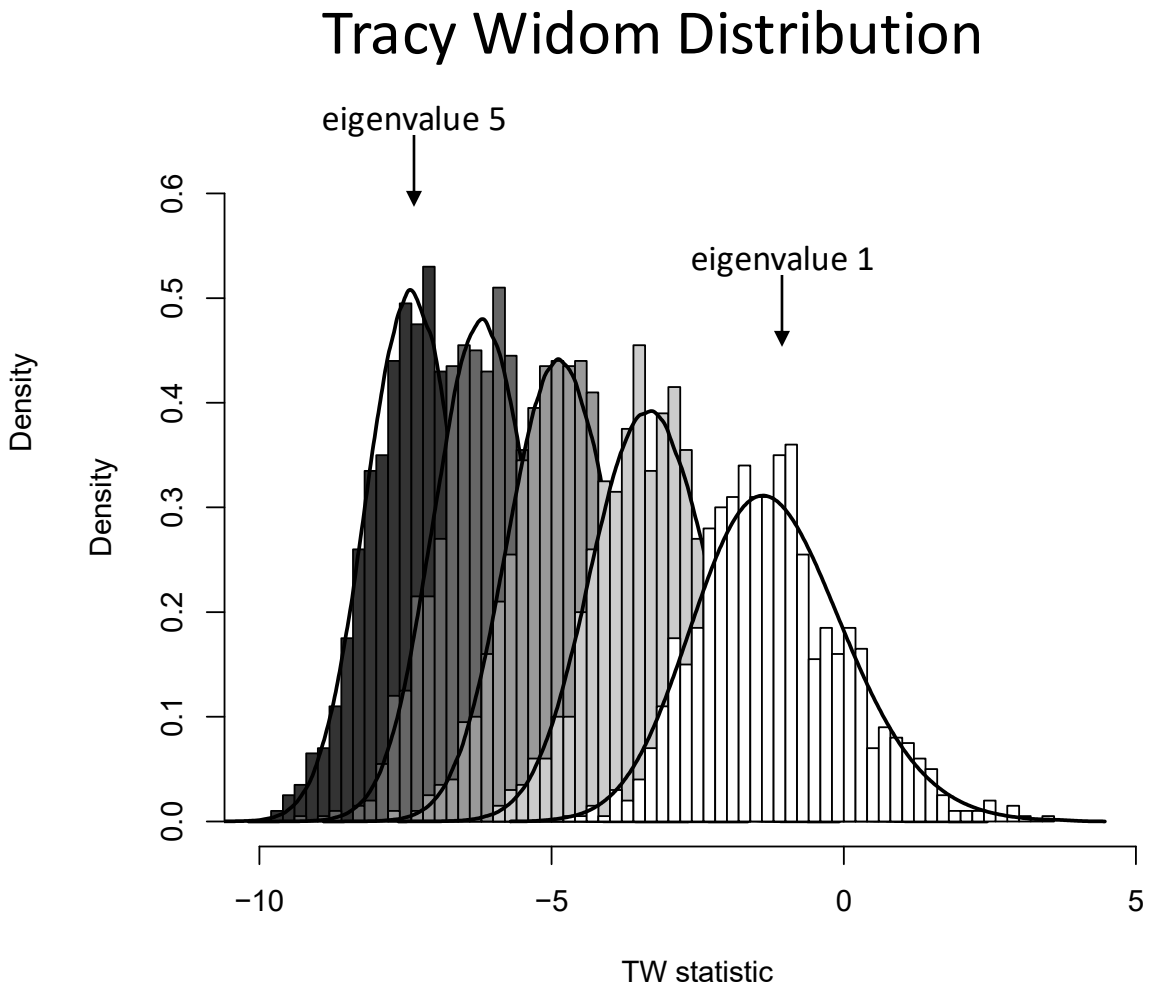


Differences between eigenvalues and estimated eigenvalues

- we now know that the overdispersion of estimated eigenvalues from MANOVA and REML is qualitatively similar to the bias observed for sample covariance matrices

Eigenvalues of sample covariance matrices follow the Tracy Widom distribution





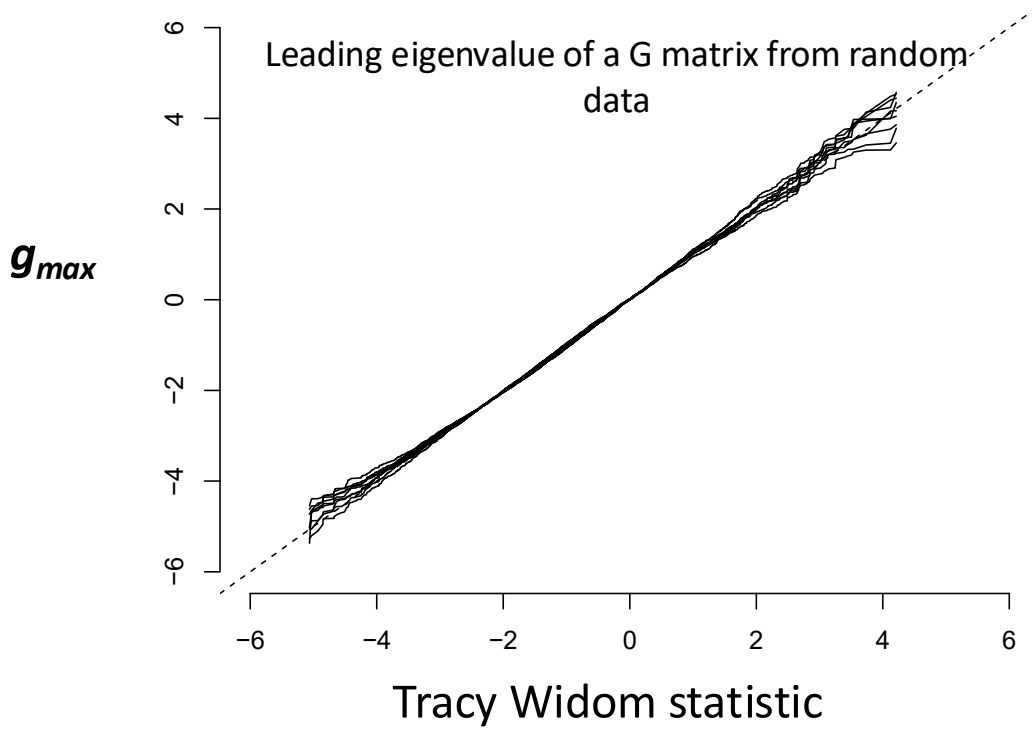
Differences between eigenvalues and estimated eigenvalues

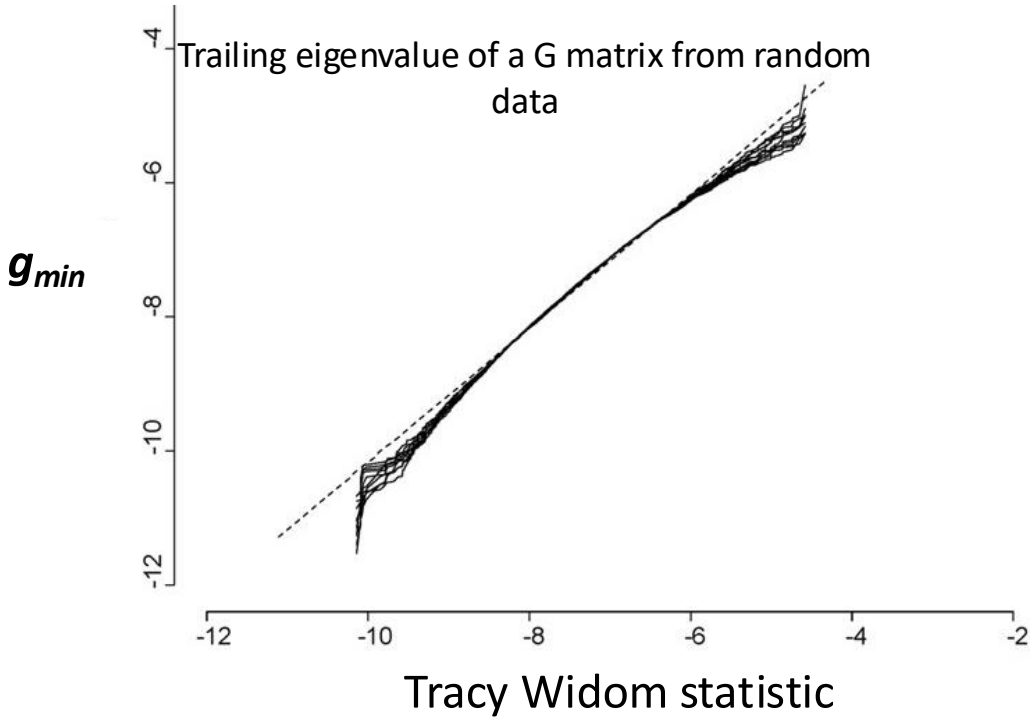
- we now know that the overdispersion of estimated eigenvalues from MANOVA and REML is qualitatively similar to the bias observed for sample covariance matrices

- for sample covariance matrices there are analytical solutions to this bias

- for **G** matrices there are not (at least not yet)

Eigenvalues of G matrices also follow Tracy Widom



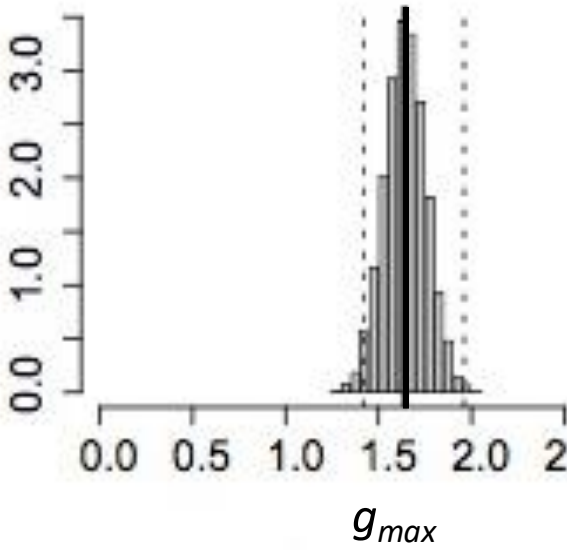


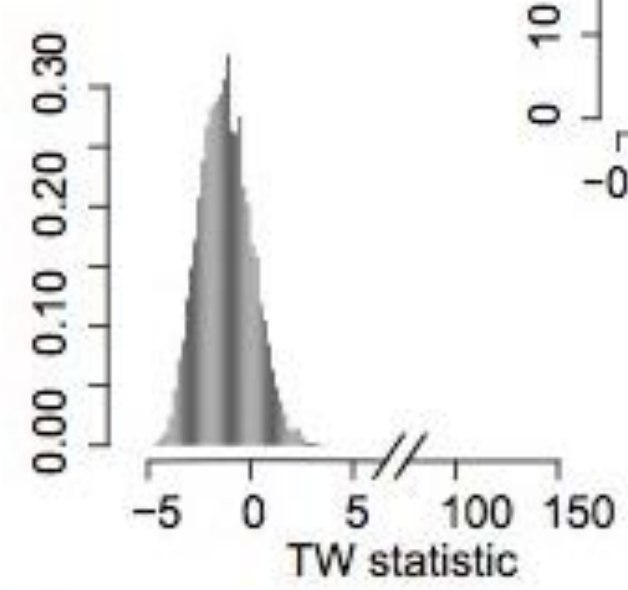
- we can use the Tracy Widom distribution as a null for the leading eigenvalues of **G**

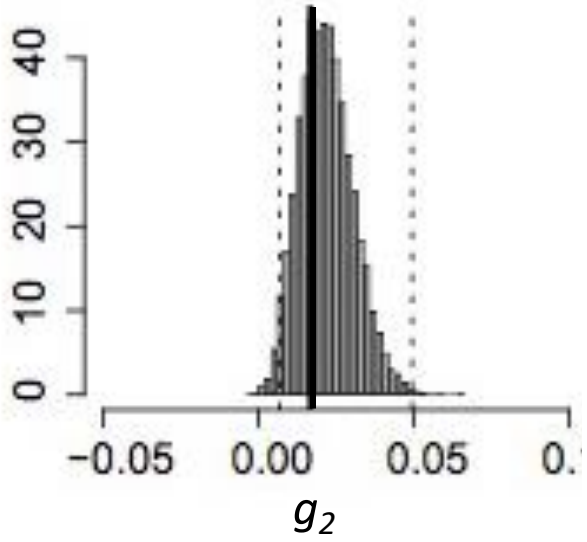
Sztepanacz, J.L. and Blows, M. W. (2017). Accounting for sampling error in genetic eigenvalues using random matrix theory. ***Genetics.*** 206: 1271-1284

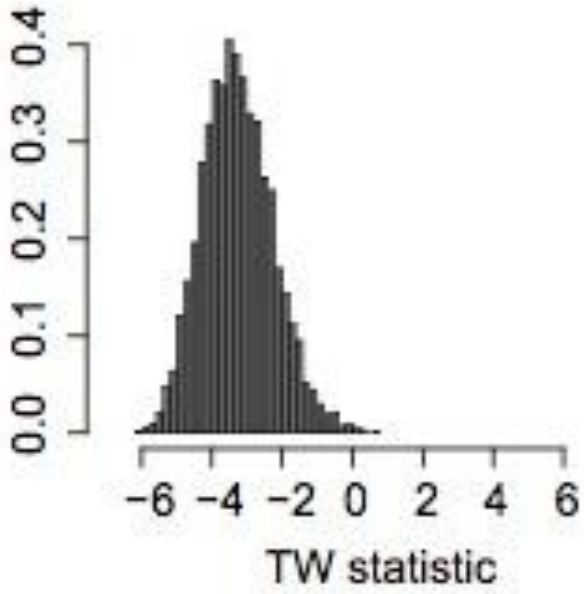
TW works as a null distribution

- simulate a **G** matrix with one phenotypic dimension that has genetic variance
- compare estimated eigenvalues to the null Tracy Widom distribution





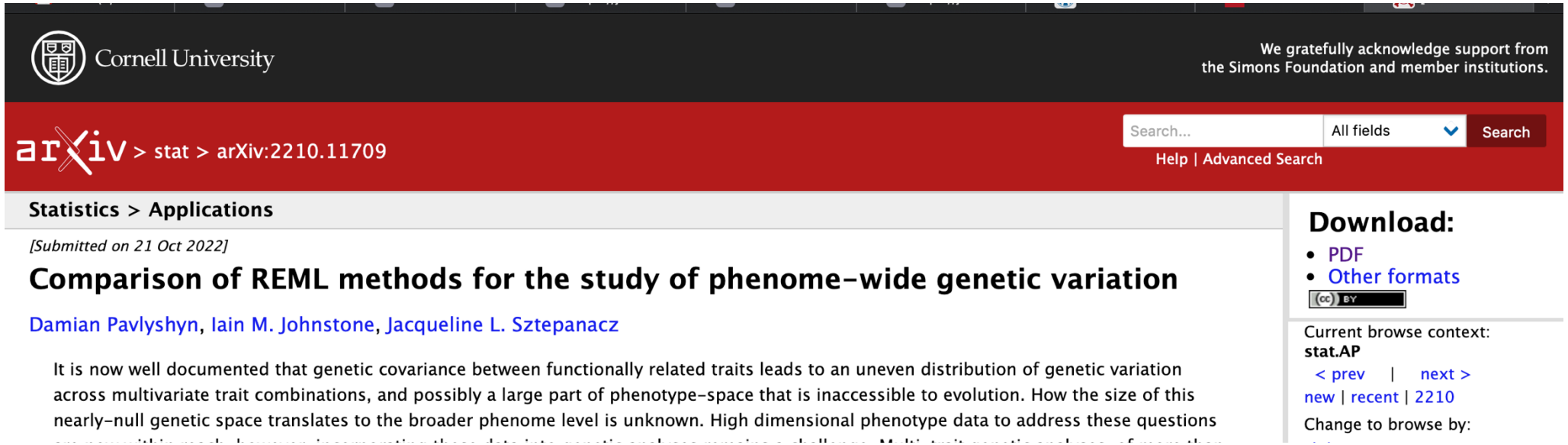




We are now expanding that work…

- that was for a specific design with few traits

- we are now expanding to more complex data structure and many traits- phenomes



These patterns are all correlative

- studying correlative patterns is useful to detecting and inferring regions of phenotype space that might experience evolutionary constraints

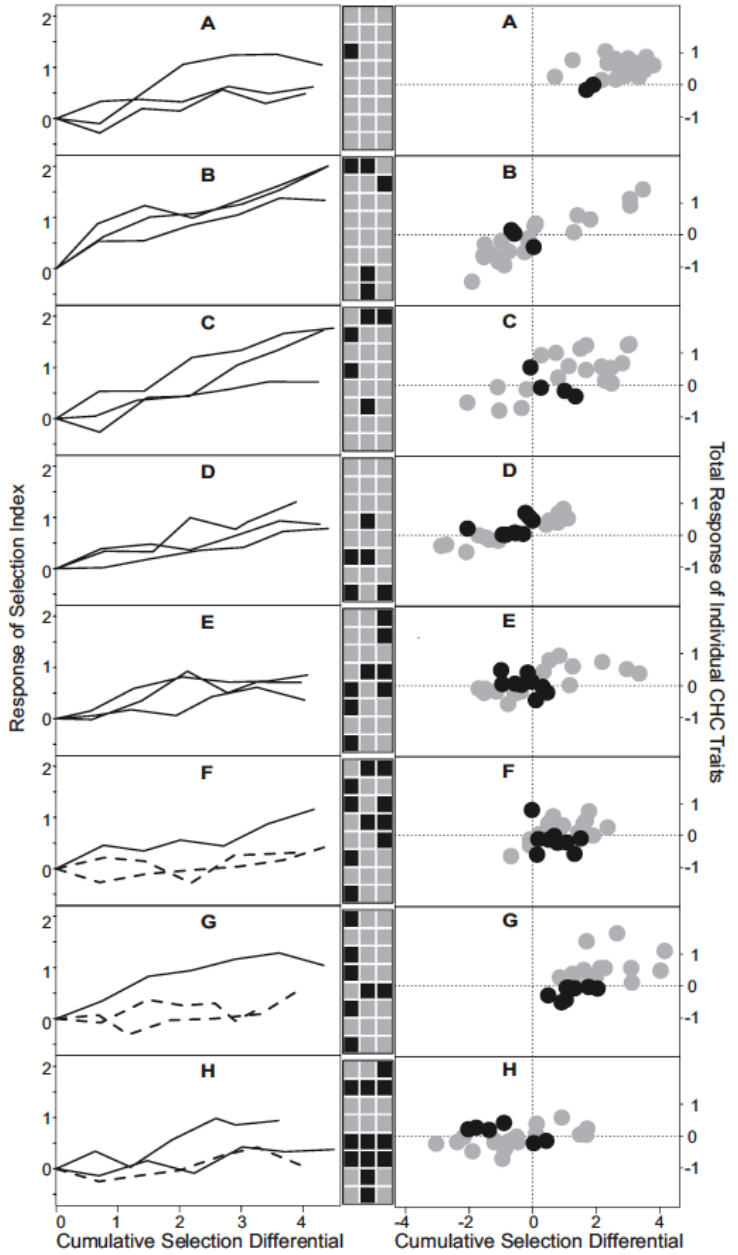
- it’s the only feasible way to study evolutionary constraints on a large-scale

- but, we also need manipulative evidence

Empirical test of nearly null subspace

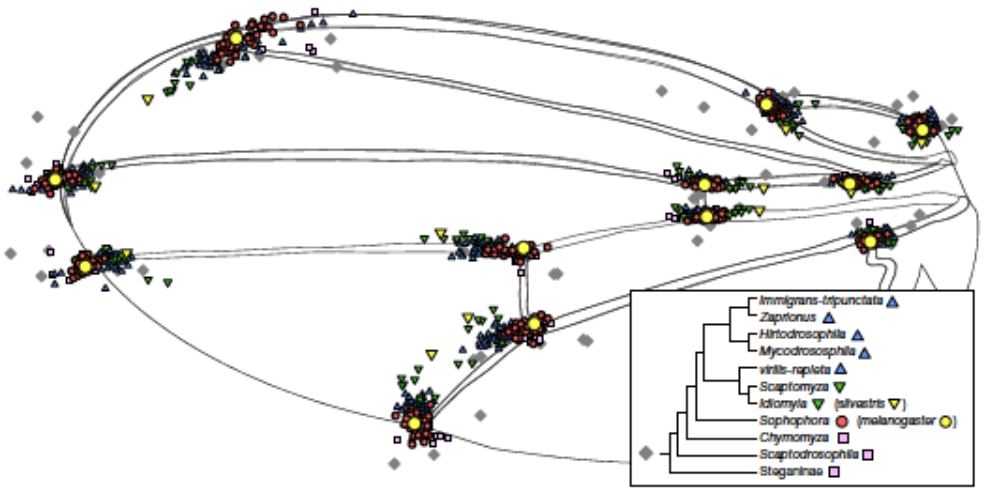
Table 1: Cuticular hydrocarbon (CHC) selection gradients and estimated genetic variance (V_A) in and heritabilities (h^2) of CHC and selection index traits

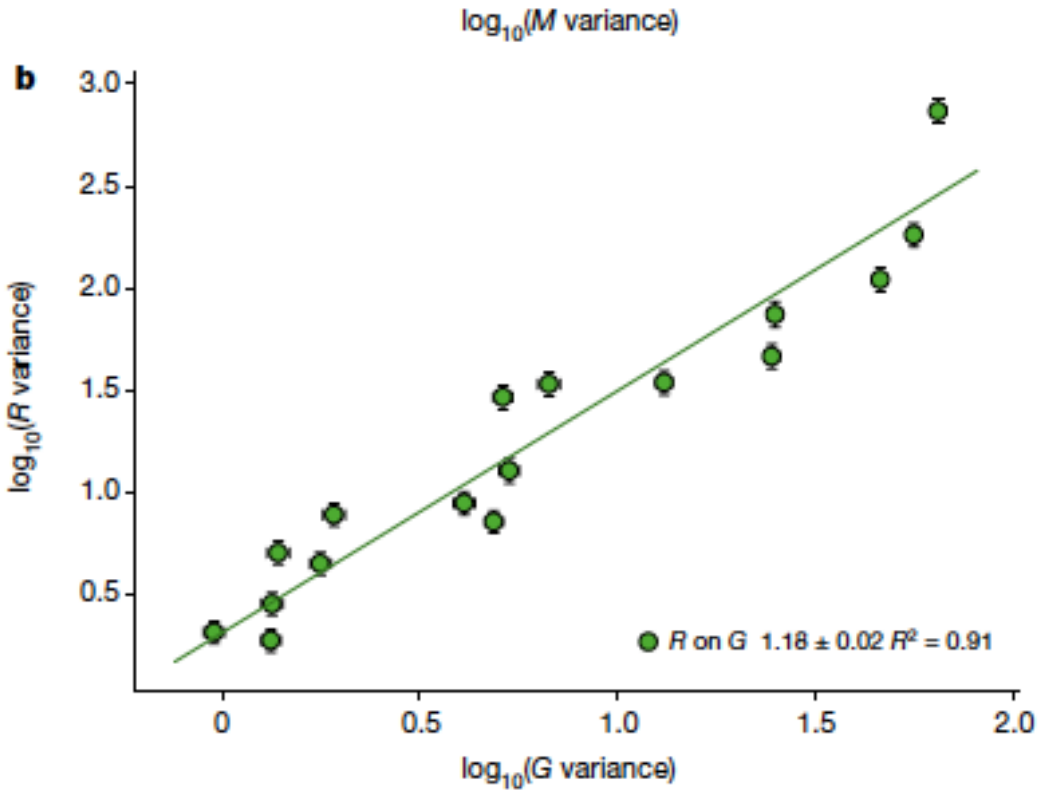
			Selection gradients							
	Base V_A	Base h^2	A	B	C	D	E	F	G	H
Base V_A			.725	.483	.295	.134	.176	.054	.172	.059
Base h^2			.115	.399	.273	.251	.208	.273	.128	.199
Realized h^2			.206	.427	.326	.237	.197	.121	.155	.197
Z,Z-5,9-$C_{25:2}$	.069	.111	.146	.028	−.132	.063	−.159	−.301	.342	−.851
Z-9-$C_{25:1}$	.191	.128	.078	−.063	.003	.190	−.148	.276	.873	.306
Z-9-$C_{26:1}$	.164	.160	.131	−.285	.387	−.800	−.293	.128	.074	−.067
2-Me-C_{26}	.482	.255	.380	.791	.108	−.222	−.020	−.333	.097	.219
Z,Z-5,9-$C_{27:2}$	.195	.175	.417	−.161	.437	.493	−.541	−.088	−.246	.068
2-Me-C_{28}	.258	.178	.467	.216	−.240	−.002	.029	.769	−.183	−.227
Z,Z-5,9-$C_{29:2}$	.349	.147	.358	−.197	.480	.096	.756	−.044	.107	−.095
2-Me-C_{30}	.391	.213	.538	−.420	−.585	−.137	.041	−.325	−.027	.256



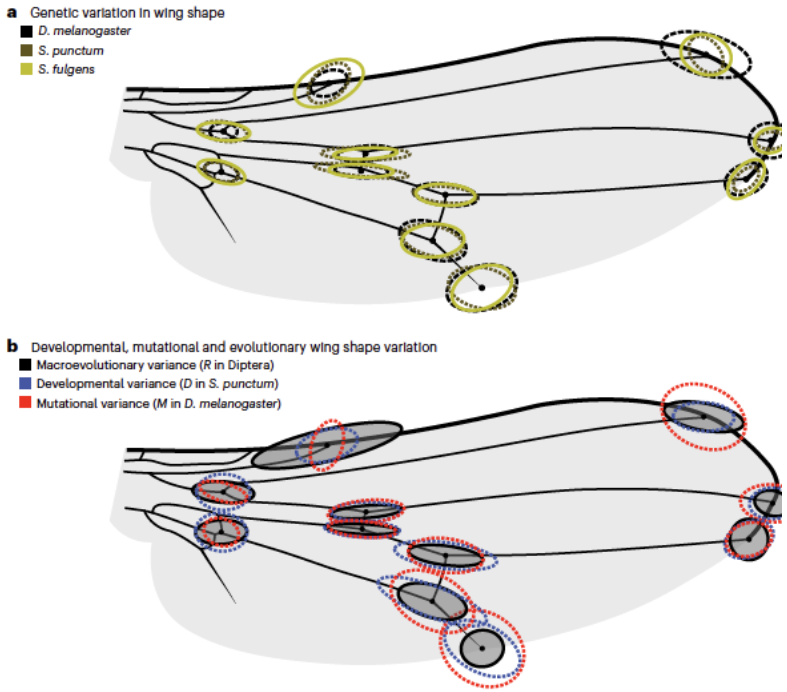
Hine, E., McGuigan, K. and Blows, M.W., 2014. Evolutionary constraints in high-dimensional trait sets. *The American Naturalist*, *184*(1), pp.119-131.

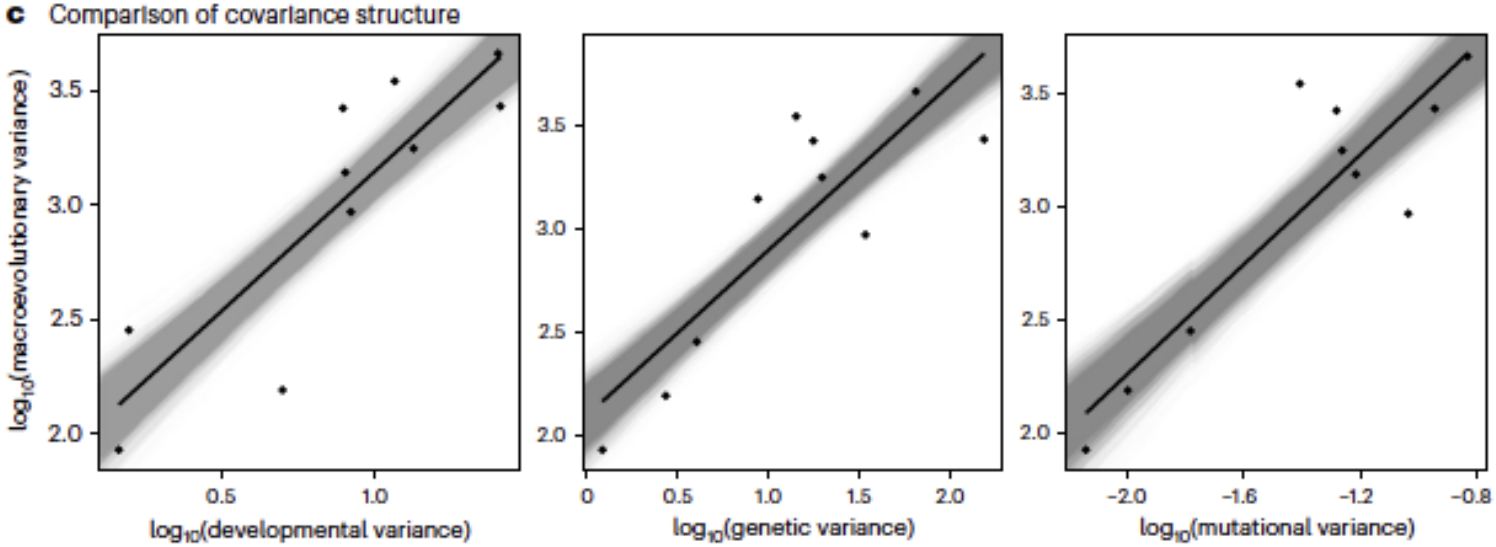
Do genetic constraints predict evolutionary response?



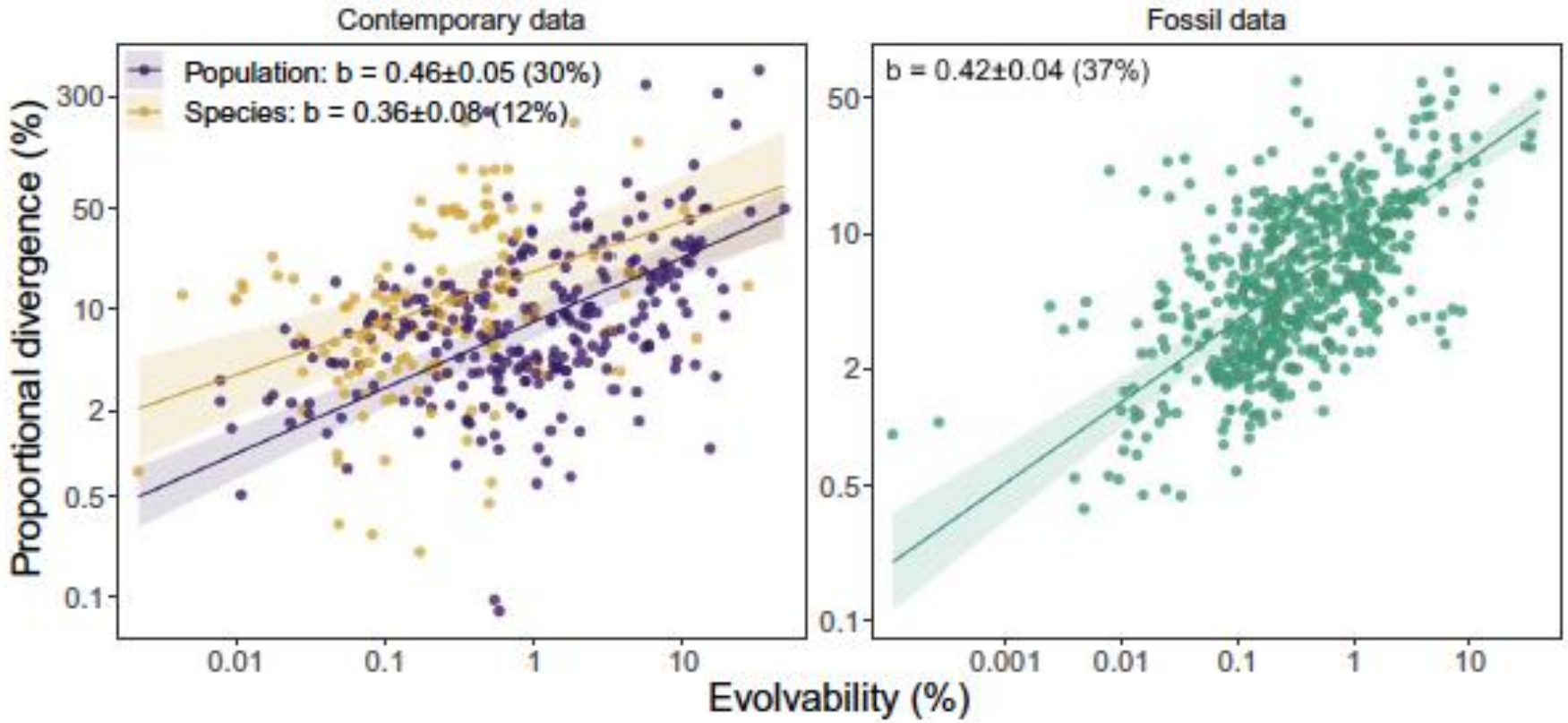


Do genetic constraints predict evolutionary response?





Do genetic constraints predict evolutionary response?



Questions?

Genetic correlations also occur between the sexes

- many traits are expressed in both males and females
- often, we might expect that breeding values for a trait in males and females are not exactly equal
- this means the cross-sex genetic correlation will be less than unity (maybe even negative)
- particularly for traits that experience intralocus sexual conflict

Genetic correlations also occur between the sexes

- the cross-sex genetic correlation is represented as r_{mf}
- cross sex genetic correlations determine the correlated response to selection on one sex in the other sex

Between-sex G

Genetic covariance between traits in males

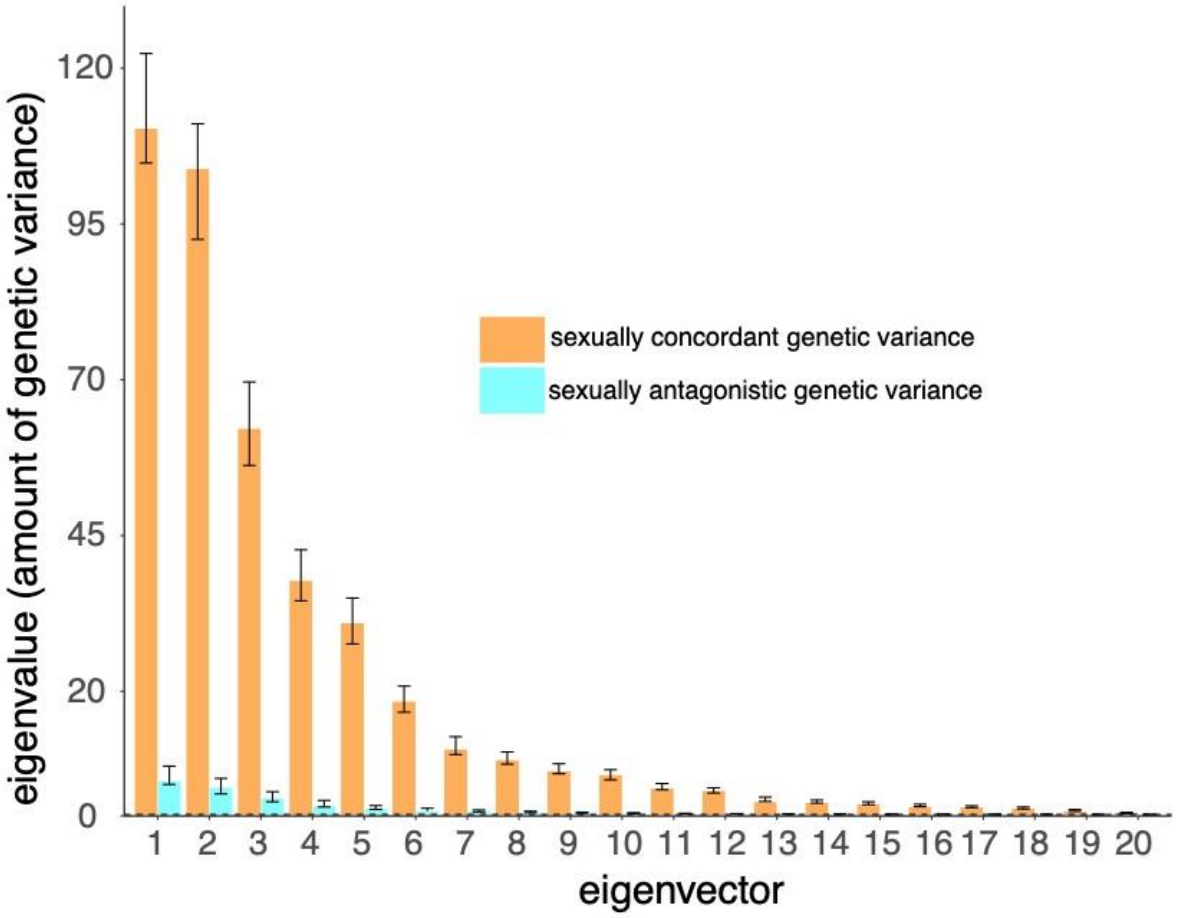
$$\mathbf{G}_{mf} = \begin{bmatrix} \mathbf{G}_m & \mathbf{B}' \\ \mathbf{B} & \mathbf{G}_f \end{bmatrix}$$

Cross-sex covariance

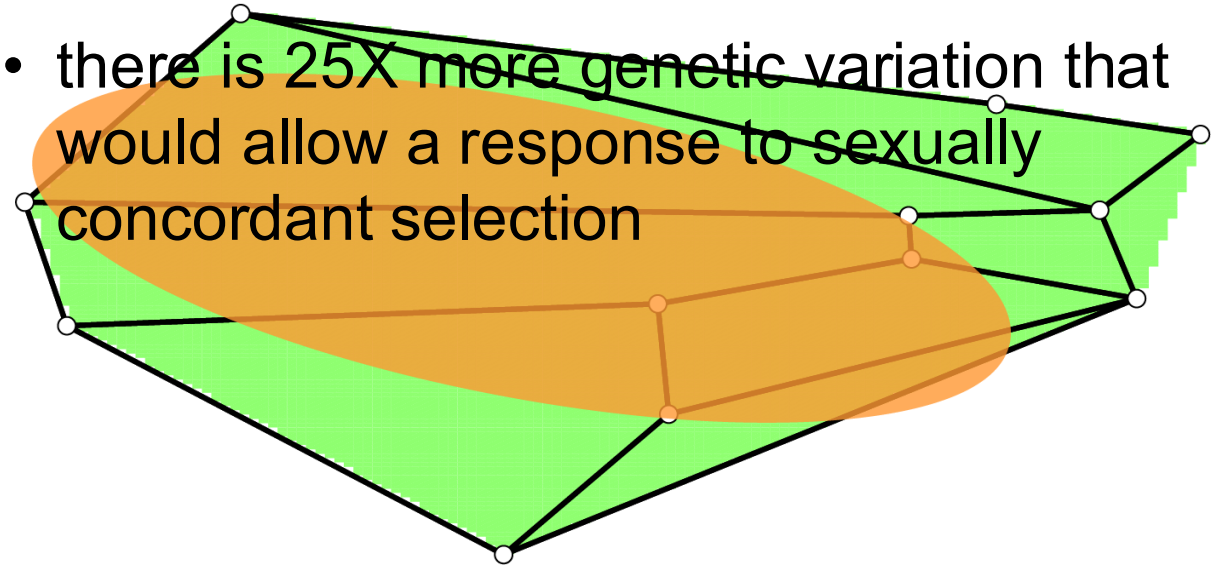
Genetic covariance between traits in females

Lande (1980) *Evolution*, 34*:* 292-305

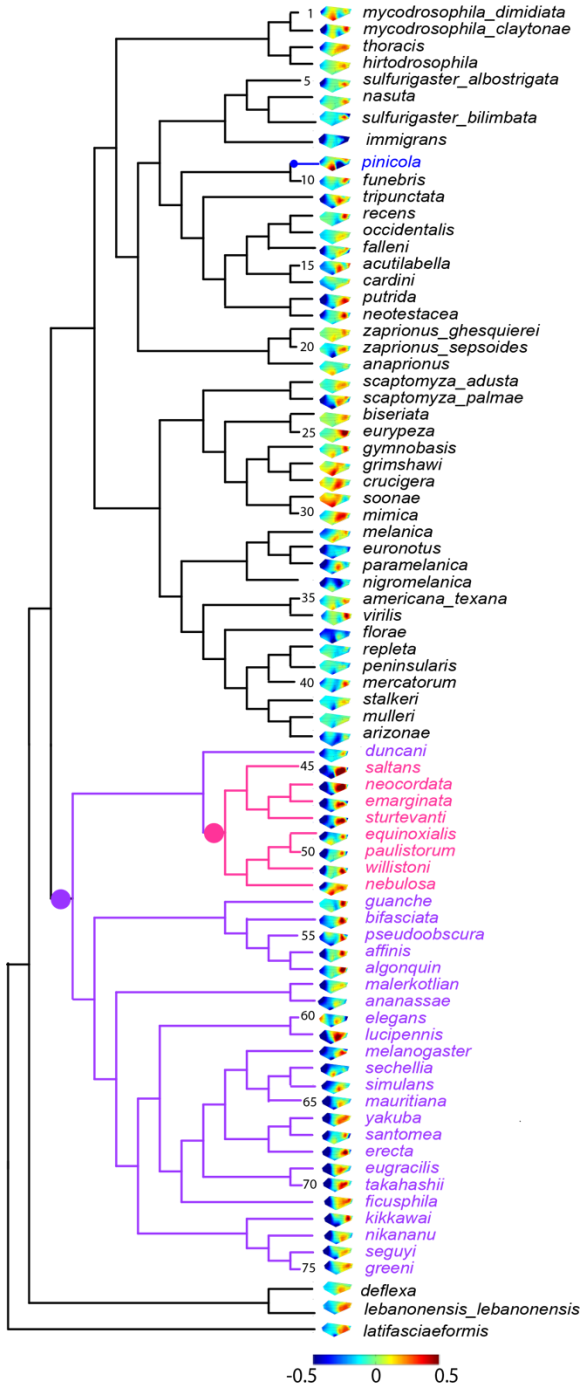
More sexually concordant genetic variation in *Drosophila* wings



- there is 25X more genetic variation that would allow a response to sexually concordant selection

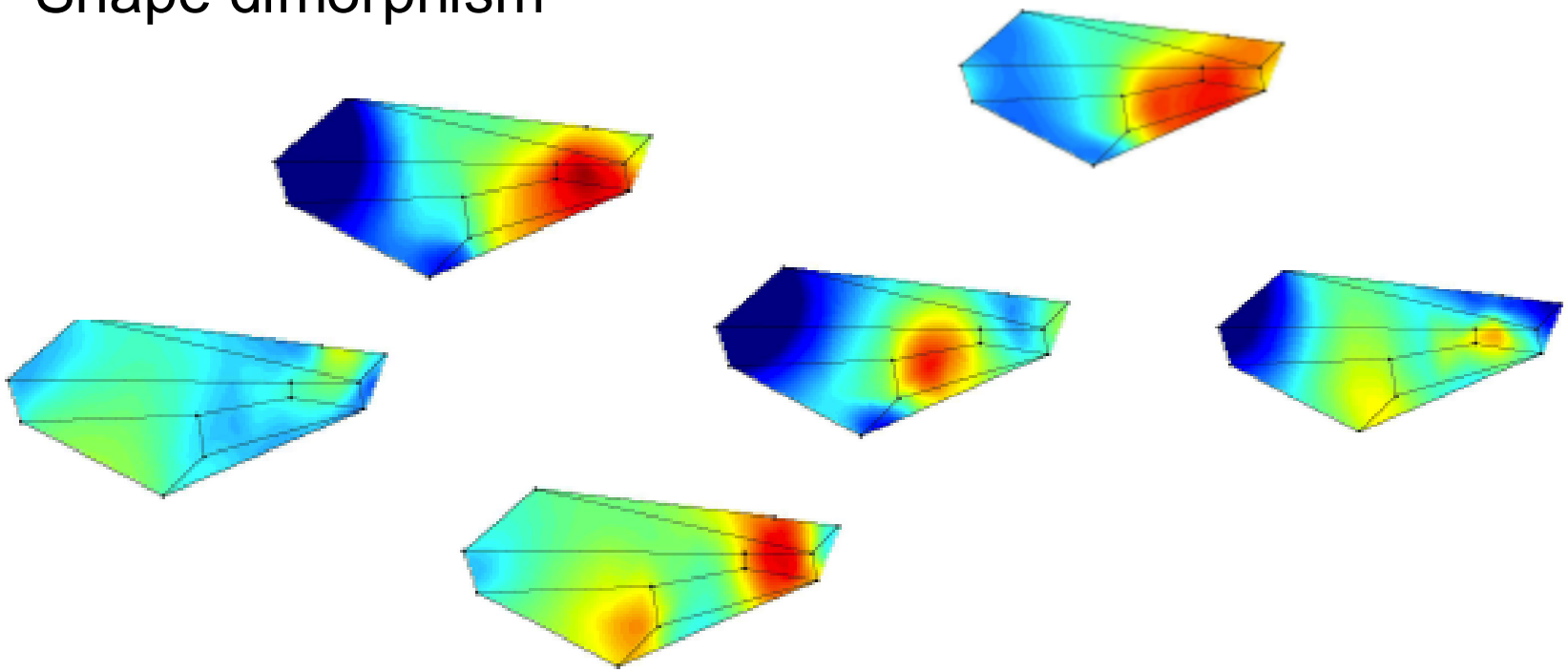


Sztepanacz and Houle (2019) *Evolution* 73: 1617-1633



In most species males have longer thinner wings than females

Shape dimorphism



Sztepanacz and Houle (2021) *Evolution* 75: 1117-1131

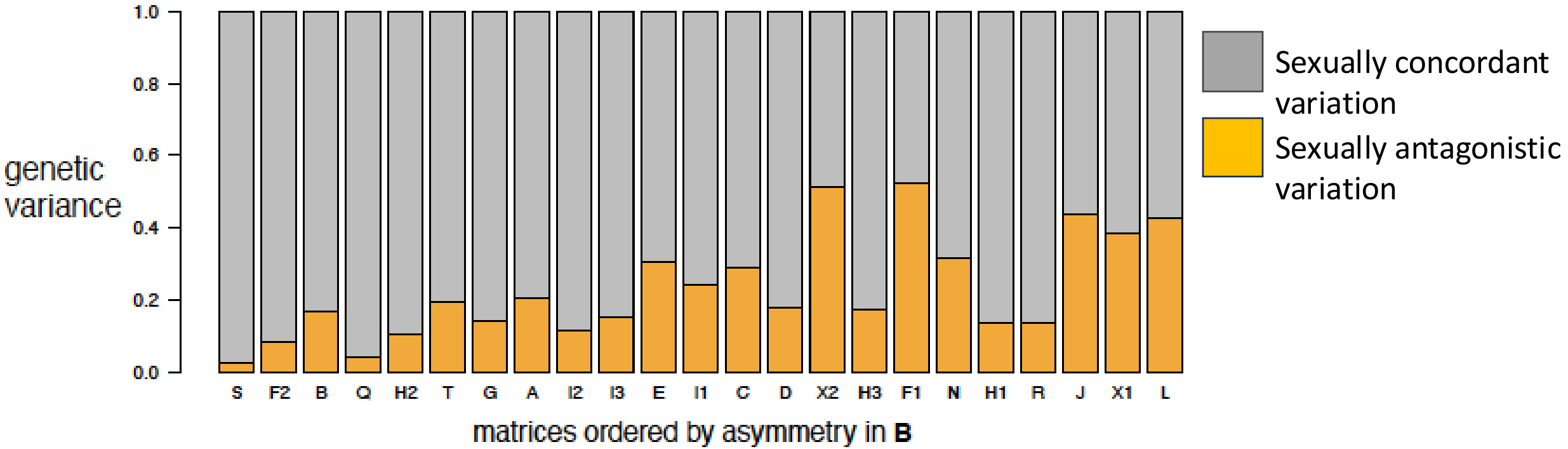
Literature Survey of $\mathbf{G}_{mf}$

Taxon		Traits		Statistical approach		Environment	
Invertebrates	10	Morphology	8	Frequentist	13	Laboratory	14
Plants	1	Behaviour	2	Bayesian	4	Field	3
Vertebrates	6	Life-history	3				
		Physiology	3				
		Transcriptomics	1				

- 23 estimates of $\mathbf{G}_{mf}$ from 17 studies

Videlier and Sztepanacz 2025 AmNat

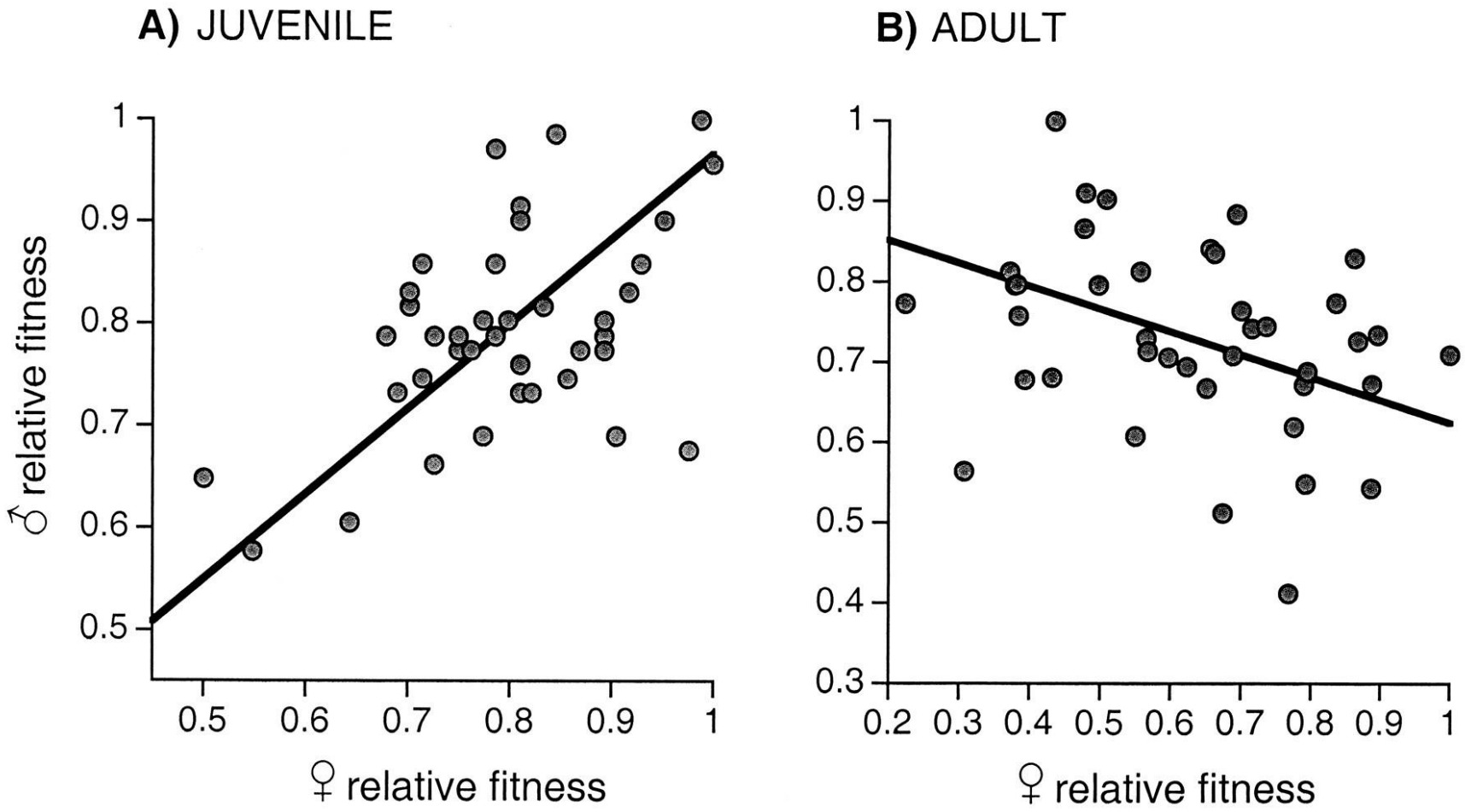
Most genetic variance is sexually concordant



- On average 77% of the genetic variation is sexually concordant

Videlier and Sztepanacz 2025 AmNat

Genetic correlations also occur between life-stages (and sexes, and all combinations!)



Questions?