

Genetica quantitativa

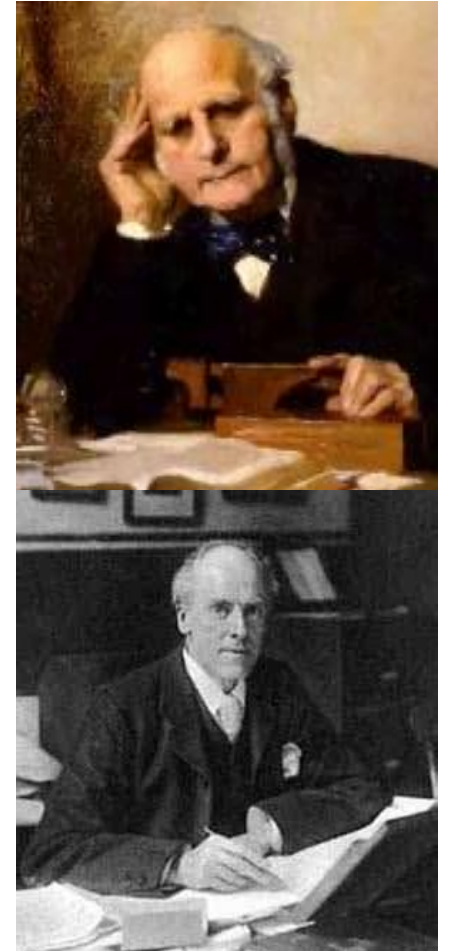
- Sommario:
- Galton e i tratti continui
- L'origine della genetica quantitativa
- Varianza del fenotipo e partizione della varianza
- Ereditabilità' sensu strictu, sensu amplo e sua stima
- Selezione naturale ed equazione dell'allevatore
- Dimostrazione della partizione della varianza genetica additive e di dominanza
- Genetica quantitative molecolare: QTLs, GWAS e Polygenic Risk scores

La genetica quantitativa e' lo studio dei caratteri continui

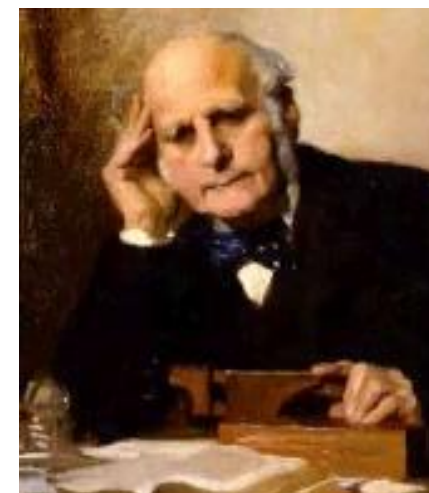
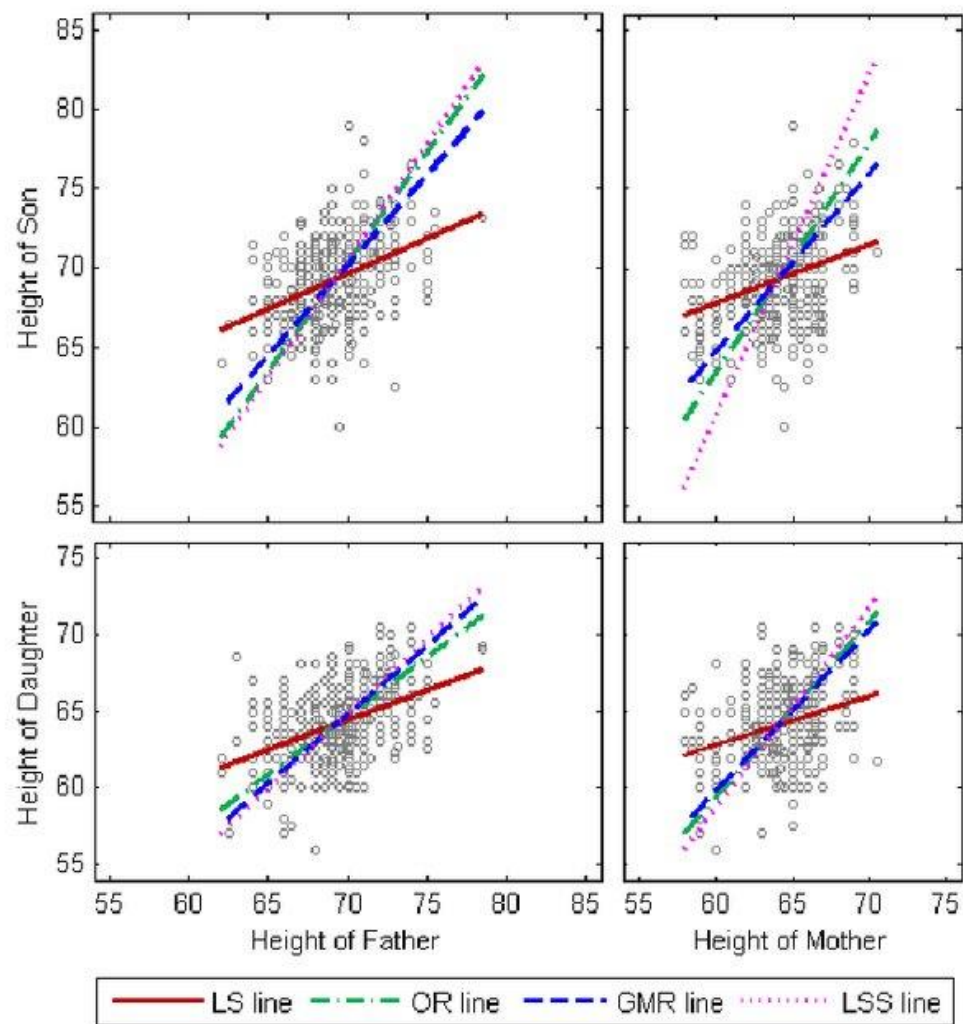
- La genetica quantitativa e' lo studio dei tratti continui e dei meccanismi che li determinano.
- I concetti principali sono stati sviluppati nel XX secolo, in risposta alla riscoperta della genetica mendeliana. La genetica mendeliana del 1900 infatti concentro' la sua attenzione su caratteri discreti (ad esempio, fiori rosa vs fiori bianchi, piselli lisci vs rugosi), in contrasto alle ricerche di Sir Francis Galton che tra il 1870 e il 1890 si concentro' su caratteri continui.

La scuola biometrica

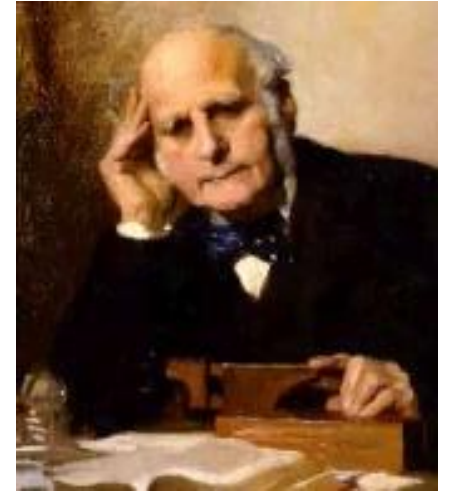
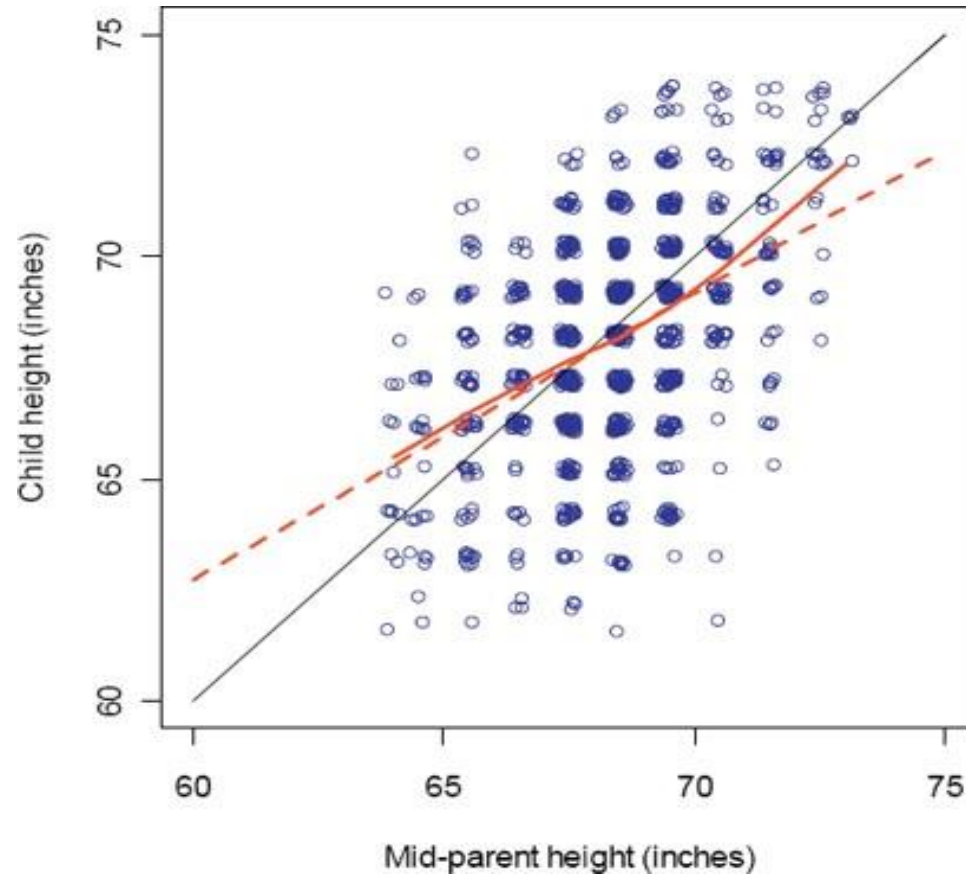
- Gli approcci metodologici di Galton sui tratti continui marcarono il fondamento della scuola Biometrica, che rappresenta il fondamento della statistica moderna (vedi Steve Stiglers nel libro *The history of statistics*).
- Karl Pearson fu profondamente ispirato da Galton, sviluppando vari metodi per l'analisi di caratteri quantitativi.



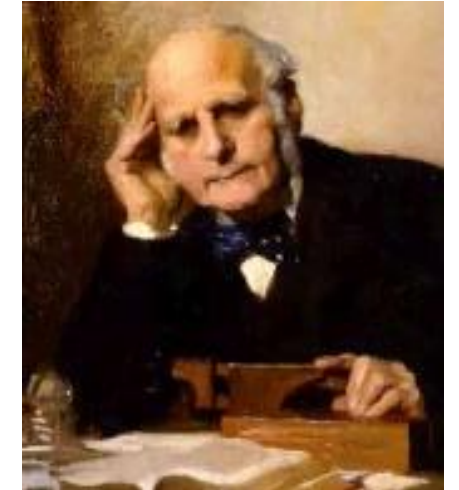
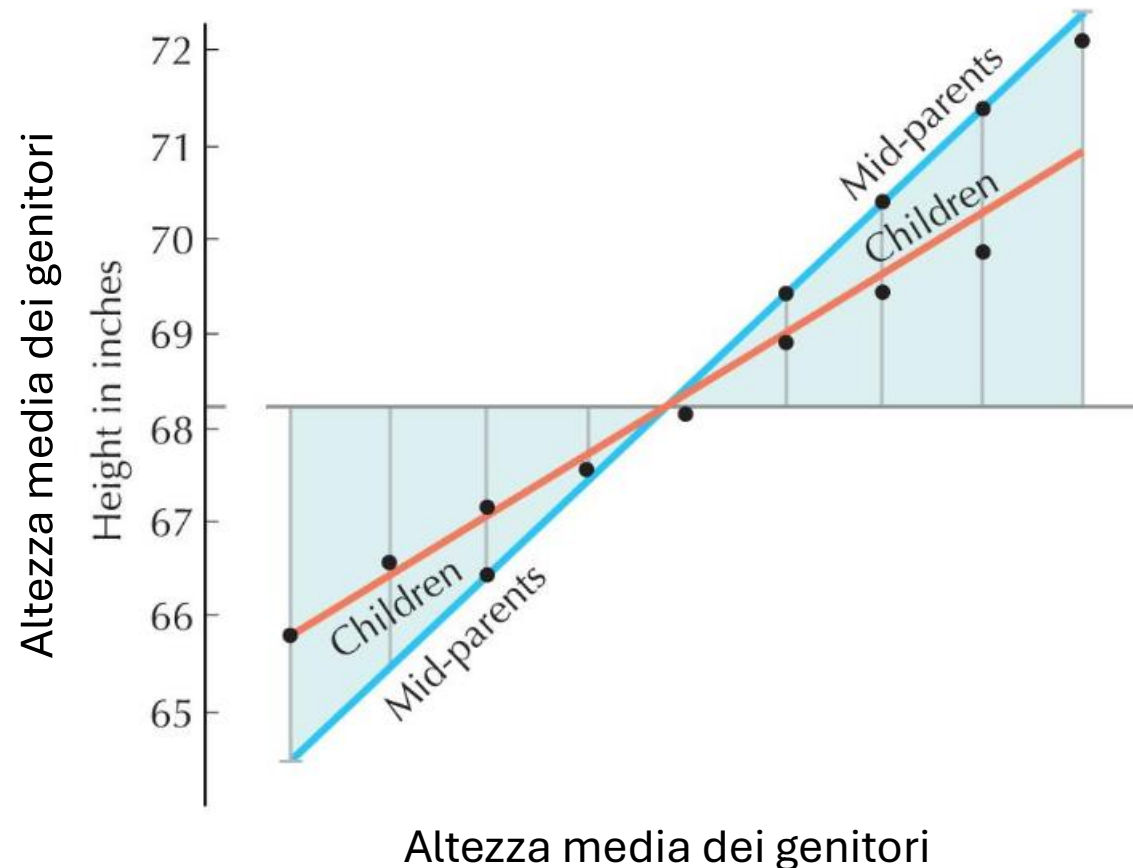
L' altezza dei figli correla con quella dei genitori



L' altezza dei figli correla con quella (media) dei genitori

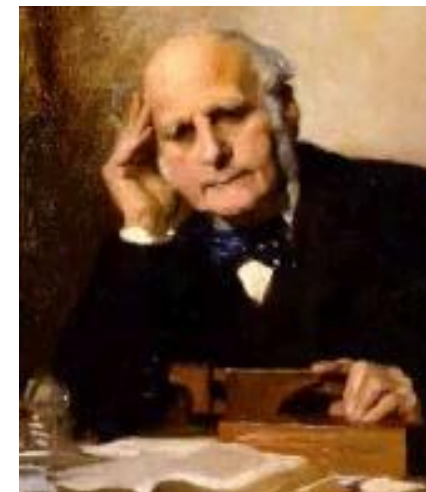
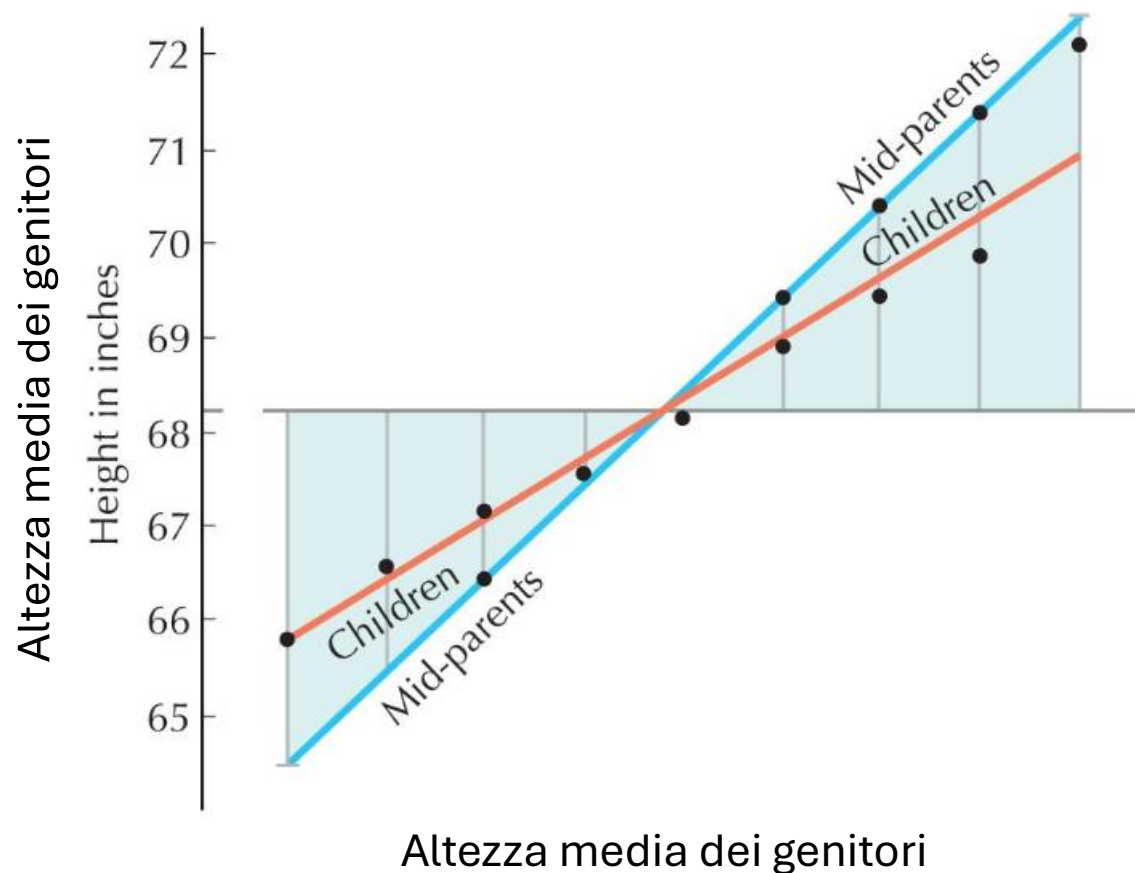


La “regressione verso la mediocrità”



Fenomeno statistico (non strettamente biologico) per il quale i valori estremi sono dovuti a combinazioni di fattori (ad esempio la combinazione di fattori genetici ed ambientali estremi) che mediamente non e' riprodotta nella generazione successiva

La “regressione verso la mediocrità”

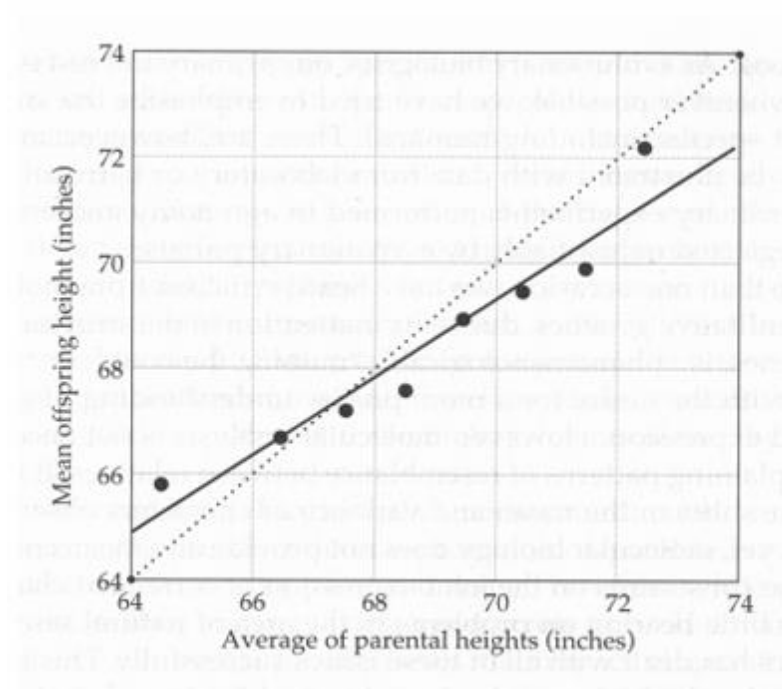


A cosa e' dovuto secondo voi?

Le origini della genetica quantitativa

- Ne seguì un dibattito acceso tra:
- I Mendeliani (guidati da William Bateson) che sostenevano che la variazione nei fenotipi continui fosse dovuta a poche mutazioni puntiformi con larghi effetti
- I Biometrici (guidati da Karl Pearson e W.F.R. Weldon) che vedevano l'evoluzione come il risultato di selezione naturale in azione su caratteristiche distribuite continuamente.
- Eventualmente il dibattito sfociò nella fusione della genetica e della Teoria dell'evoluzione di Charles Darwin, con i fondamenti della genetica quantitativa, sviluppati indipendentemente da Ronald Fisher (1918) e Sewall Wright (1921).

Il dibattito tra i Mendeliani e i Biometrici



Dobbiamo Vedere la natura soprattutto come il colore del fiore dei piselli o come l'altezza?

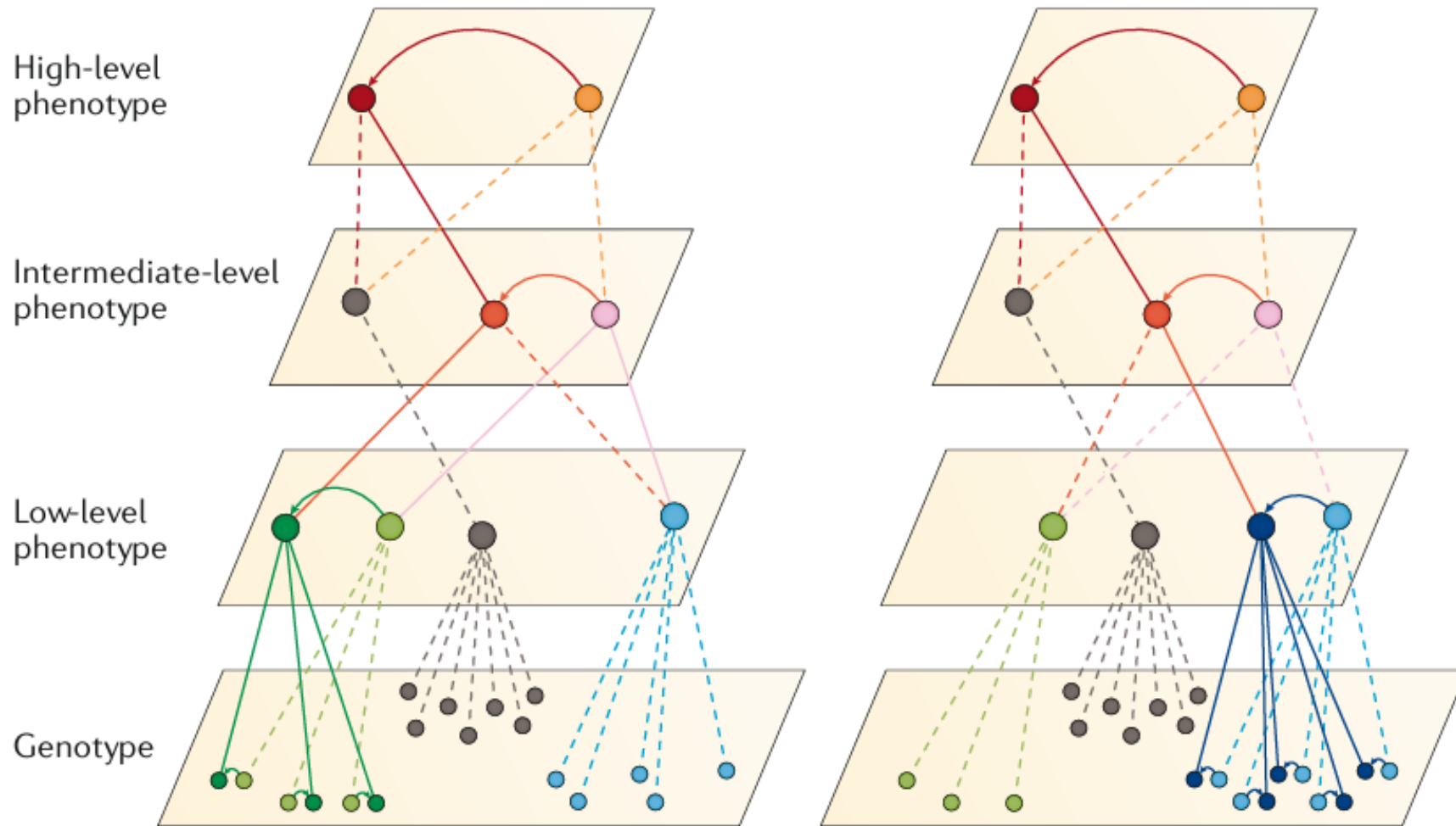
I tratti quantitativi hanno le stesse proprietà evolutive dei caratteri discreti?



Secondo voi, quali prevalgono in natura?

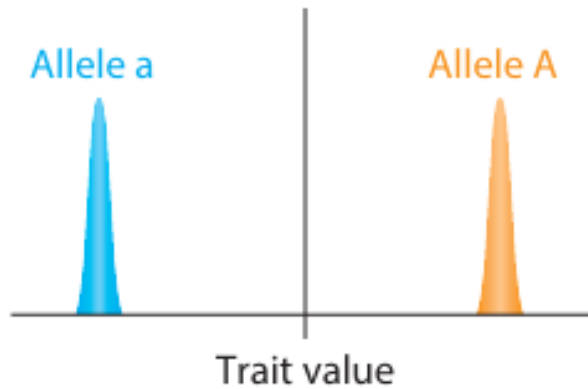


Fenotipi di alto livello sono piu' frequentemente multifattoriali

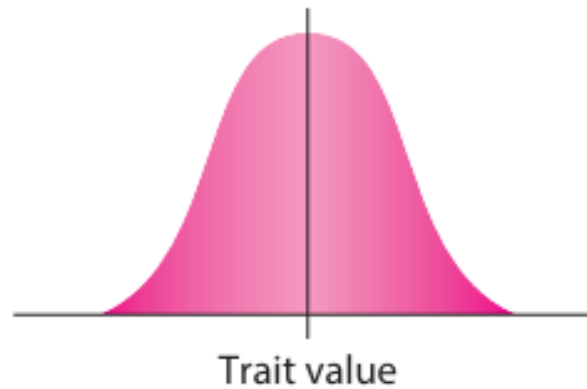


Tre tipologie di tratti

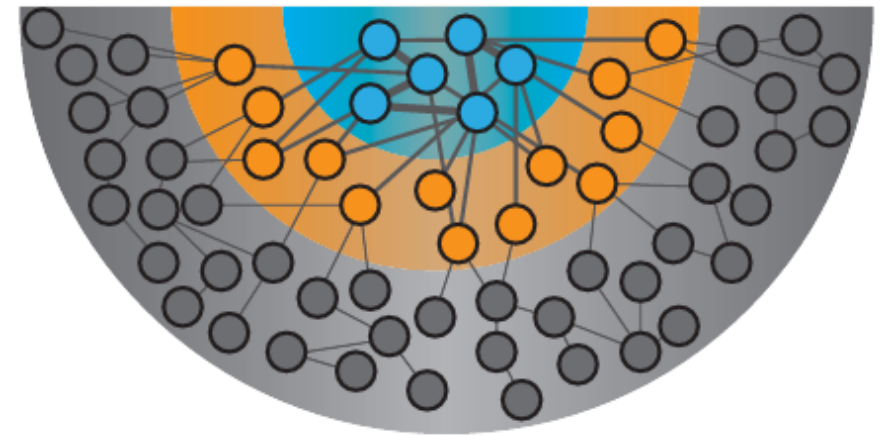
a **Monogenic trait**
controlled by a single locus






b **Polygenic trait**
controlled by many loci



Omnigenic trait
large numbers of haplotype blocks contain a causal variant

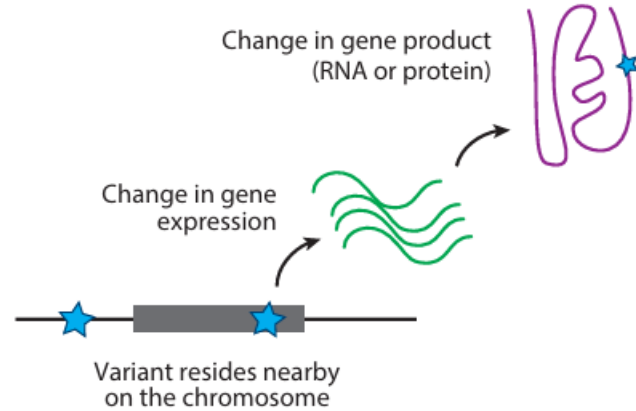


-  *cis*-Acting variants at core genes
-  *trans*-Acting variants affecting core genes
-  Variants affecting peripheral genes (and indirectly affecting core genes)

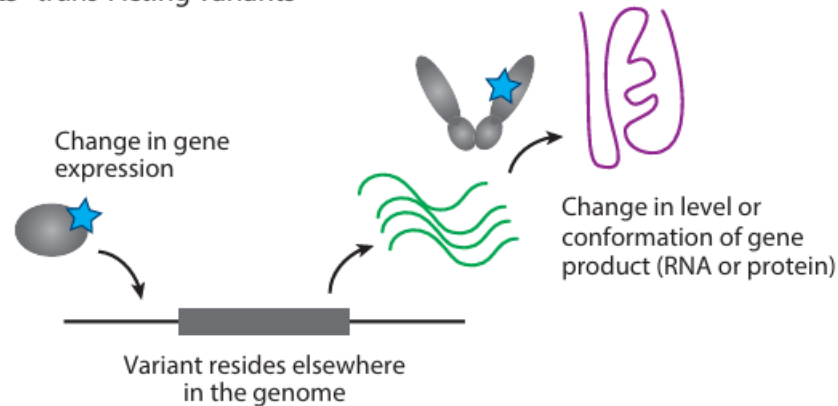
Nel modello omnigenico poche varianti contribuiscono molto, moltissime poco

Le varianti in *cis* hanno effetti spesso piu' forti e vengono segregate insieme, rispetto a quelle che agiscono in *trans*

a *cis*-Acting variants

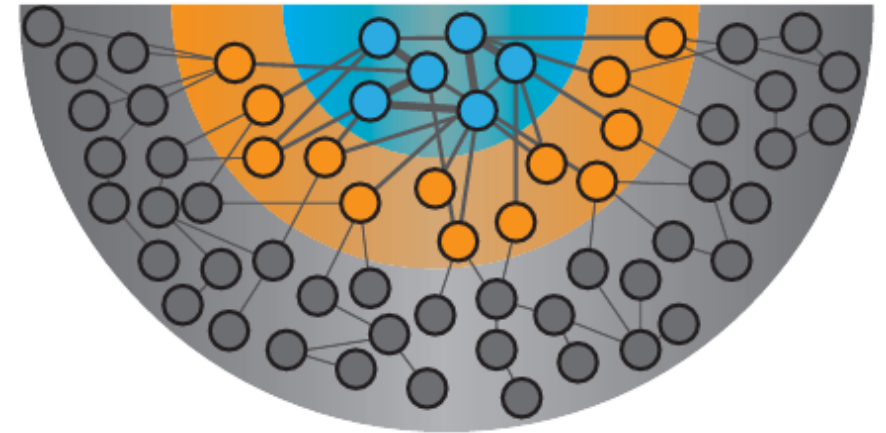


b *trans*-Acting variants



c

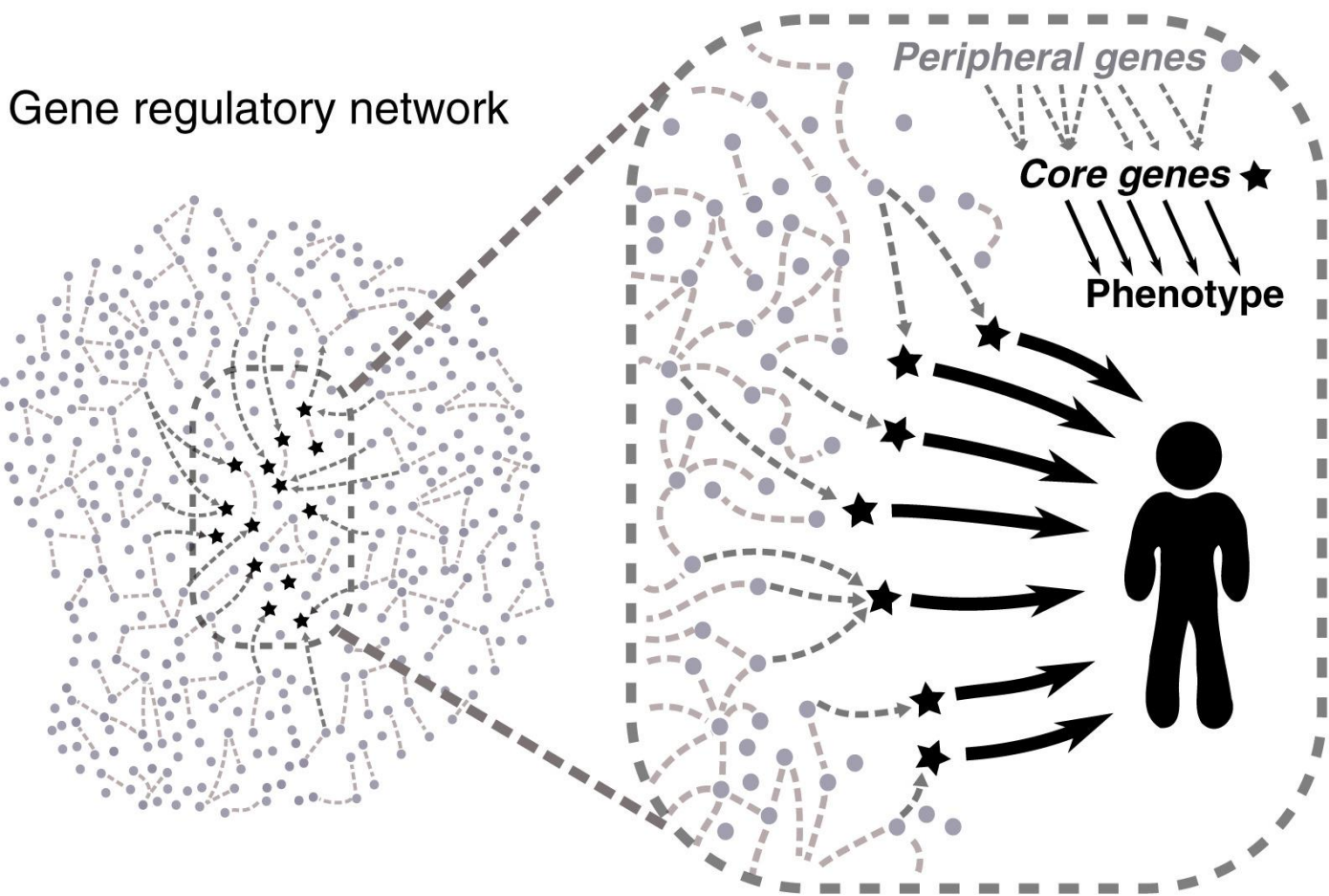
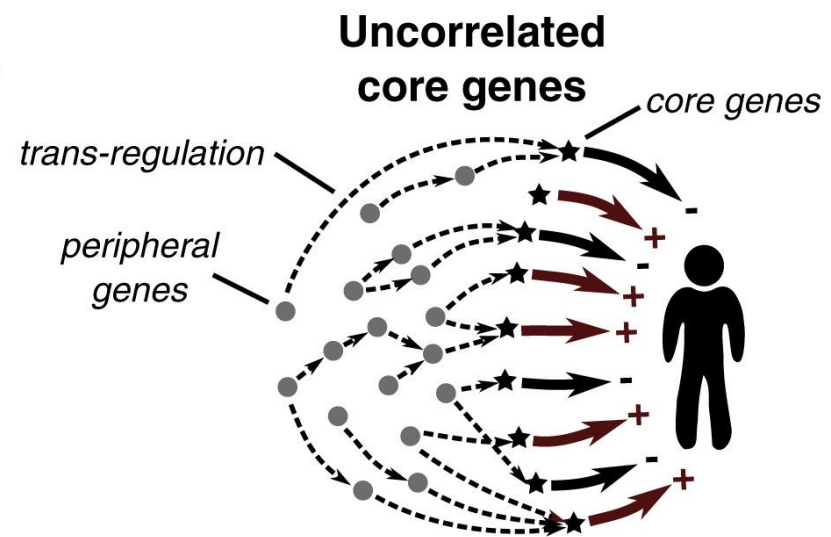
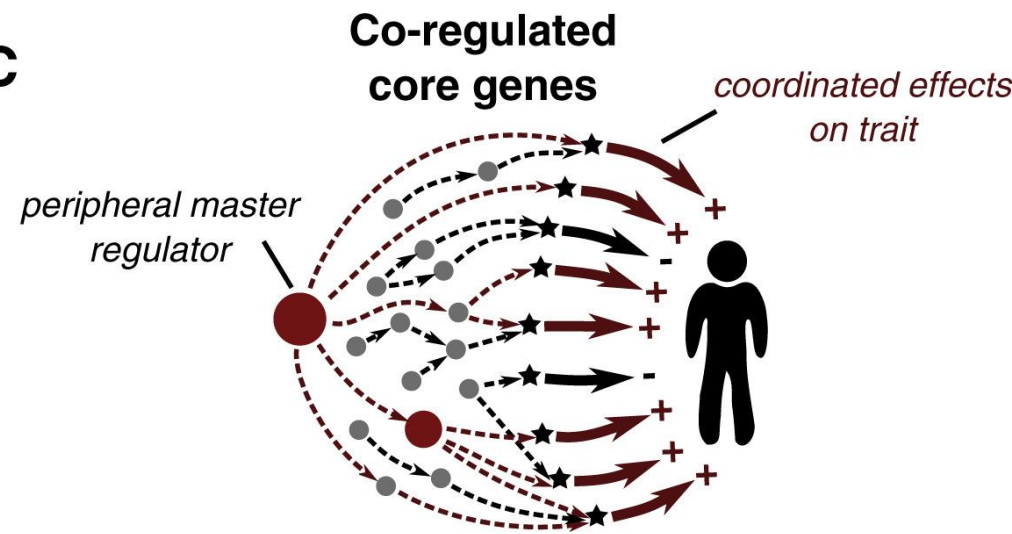
Omnigenic trait
large numbers of haplotype blocks contain a causal variant



- *cis*-Acting variants at core genes
- *trans*-Acting variants affecting core genes
- Variants affecting peripheral genes (and indirectly affecting core genes)

A

Gene regulatory network

**B****C**

Il modello di base della genetica quantitativa

- Uno degli obiettivi centrali della genetica quantitative e' la quantificazione della corrispondenza tra genotipo e fenotipo.
- E' ora generalmente accettato che un tratto fenotipico puo' essere influenzato da molte varianti genetiche.
- La genetica quantitativa e' spesso focalizzata tra separare la varianza di un fenotipo osservato tra component genetiche e non. Pertanto la genetica classica quantitativa si concentra su un semplice modello:
- **Fenotipo = Effetto del genotipo + Effetto dell'ambiente.**

La varianza di un fenotipo

$$V_P = \sum_{i=1}^N \frac{(x_i - \mu)^2}{N}$$

- Dove V_P e' la varianza osservata del fenotipo.
- x_i i valori del fenotipo per ognuno degli N individui i
- μ il valore medio del fenotipo

Partizione della varianza fenotipica

- Un approccio standard per suddividere la varianza fenotipica e' di modellizzare il valore osservato di un fenotipo per un individuo Y come la somma del contributo genetico di tutti i loci per quel tratto (che denomineremo il valore genotipico G) e il contributo dell'ambiente E . Quindi
- $Y=G+E$
- In assenza di covarianza tra ambiente e geni, lo scopo e' suddividere la varianza in Varianza Fenotipica V_P , Genotipica V_G ed Ambientale V_E .
- $V_P=V_G+V_E$
- Dove $V_G=V_A+V_D+V_I$ con Varianza Genetica Additiva V_A , Varianza dovuta alla dominanza V_D e alle interazioni epistatiche V_I

Partizione della varianza fenotipica

- Un approccio standard per suddividere la varianza fenotipica e' di modellizzare il valore osservato di un fenotipo per un individuo Y come la somma del contributo genetico di tutti i loci per quel tratto (che denomineremo il valore genotipico G) e il contributo dell'ambiente E . Quindi
- $Y=G+E$
- In assenza di covarianza tra ambiente e geni, lo scopo e' suddividere la varianza in Varianza Fenotipica V_P , Genotipica V_G ed Ambientale V_E .
- $V_P=V_G+V_E$

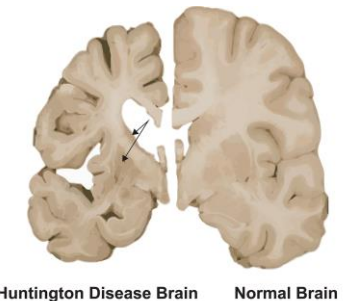
L'ereditabilità' misura quanto i fattori genetici spiegano della varianza fenotipica *in una popolazione*

$$h^2_b = V_G / V_P$$

L'ereditabilità' sensu amplo (o broad sense heritability) e' la proporzione di varianza fenotipica spiegata da quella genotipica

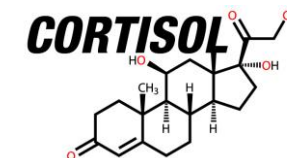
Alta ereditabilità':

La variabilità' nel tratto e' dovuto soprattutto a cause genetiche



Bassa ereditabilità':

La variabilità' nel tratto non e' dovuta ai geni



L'ereditabilita' misura quanto i fattori genetici spiegano della varianza fenotipica *in una popolazione*

$$h^2_b = V_G / V_P$$

L'ereditabilita' sensu amplo (o broad sense heritability) e' la proporzione di varianza fenotipica spiegata da quella genotipica

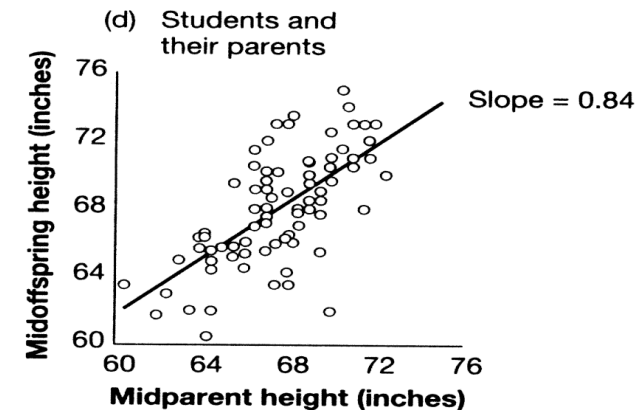
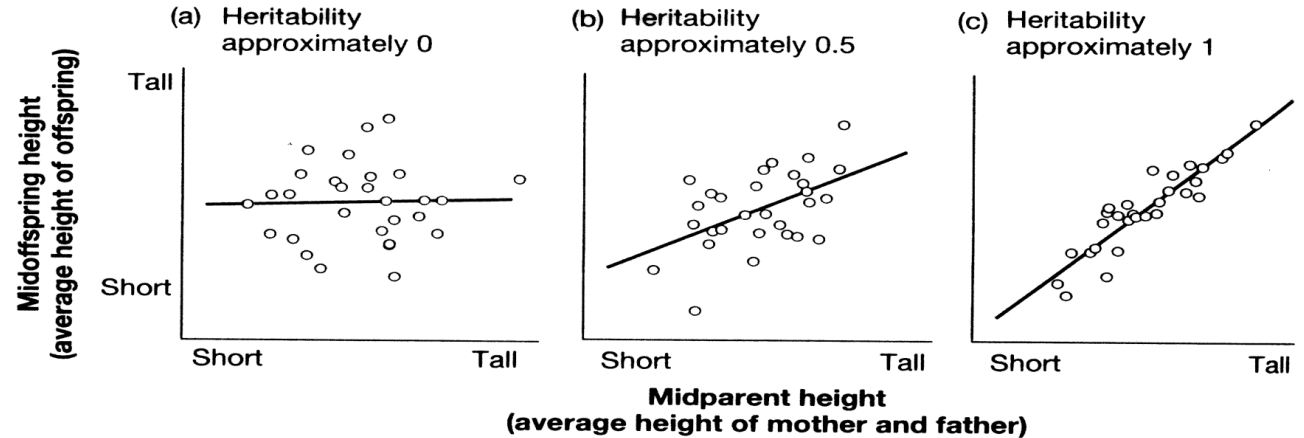
Come pensate possa essere stimata?



L'ereditabilita' misura quanto i fattori genetici spiegano della varianza fenotipica *in una popolazione*

$$h^2_b = V_G / V_P$$

L'ereditabilita' sensu amplo (o broad sense heritability) e' la proporzione di varianza fenotipica spiegata da quella genotipica



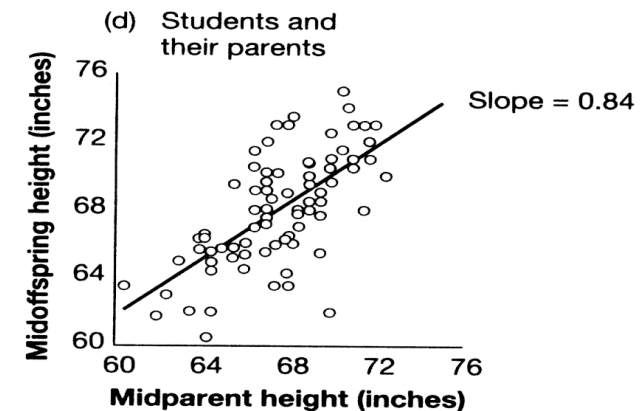
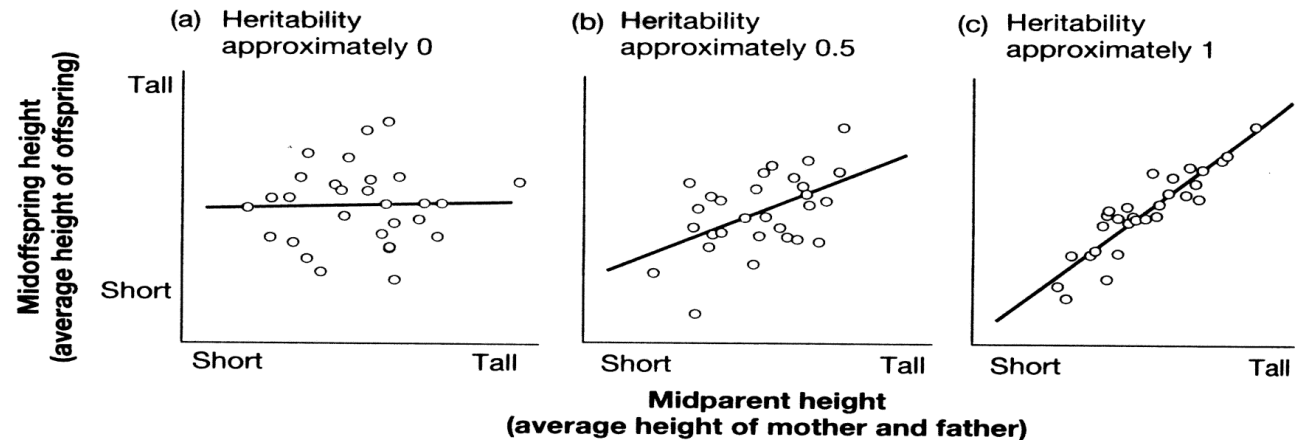
L'ereditabilità' misura quanto i fattori genetici spiegano della varianza fenotipica *in una popolazione*

$$h^2_b = V_G / V_P$$

L'ereditabilità' sensu amplo (o broad sense heritability) e' la proporzione di varianza fenotipica spiegata da quella genotipica

$$h^2_b = b/r = t/r$$

Dove b e' la pendenza della regressione (o t il coefficiente di correlazione) ed r e' il livello di parentela.



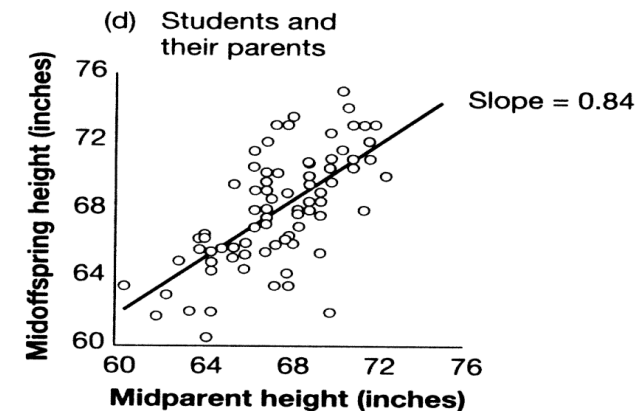
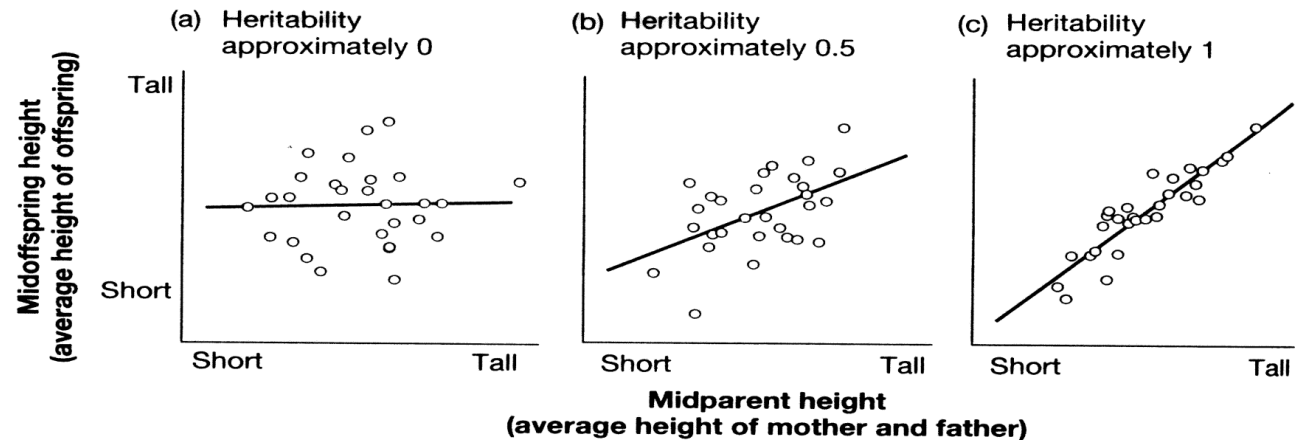
L'ereditabilità' misura quanto i fattori genetici spiegano della varianza fenotipica *in una popolazione*

$$h^2_b = V_G / V_P$$

L'ereditabilità' sensu amplo (o broad sense heritability) e' la proporzione di varianza fenotipica spiegata da quella genotipica

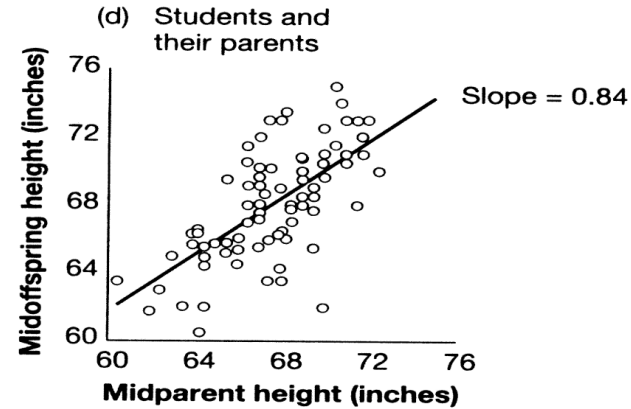
$$h^2_b = b/r = t/r$$

Dove b e' la pendenza della regressione (o t il coefficiente di correlazione) ed r e' il livello di parentela.



Stima dell'ereditabilità' a partire dalla regressione di parenti vicini

Quanta e' l'ereditabilità' in questo caso?

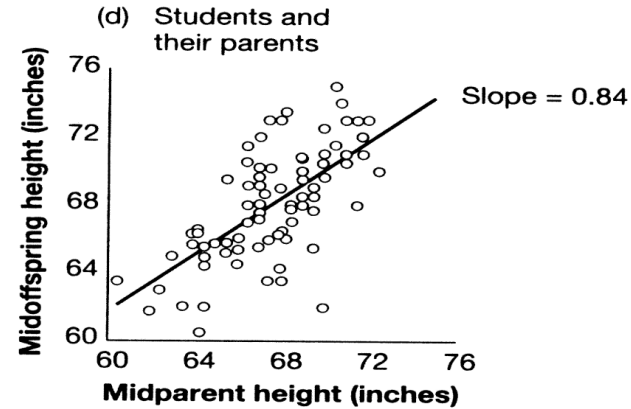


$$h^2_b = b/r = t/r$$

Dove b e' la pendenza della regressione (o t il coefficiente di correlazione) ed r e' il livello di parentela.

Stima dell'ereditabilità' a partire dalla regressione di parenti vicini

Quanta e' l'ereditabilità' in questo caso? →



$$h^2 = 0.84$$

E se la pendenza fosse stata 0.43?

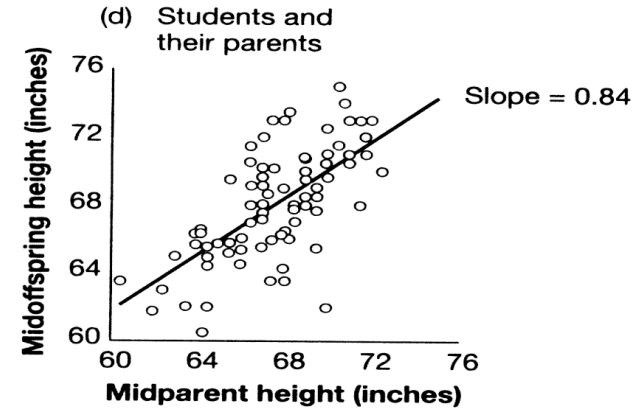
E se fosse stata 0.43 e calcolata con un solo genitore?

$$h^2_b = b/r = t/r$$

Dove b e' la pendenza della regressione (o t il coefficiente di correlazione) ed r e' il livello di parentela.

Stima dell'ereditabilità' a partire dalla regressione di parenti vicini

Quanta e' l'ereditabilità' in questo caso? →



$$h^2 = 0.84$$

E se la pendenza fosse stata 0.43?

$$h^2 = 0.43$$

E se fosse stata 0.43 e calcolata con un solo genitore?

$$h^2 = 0.86$$

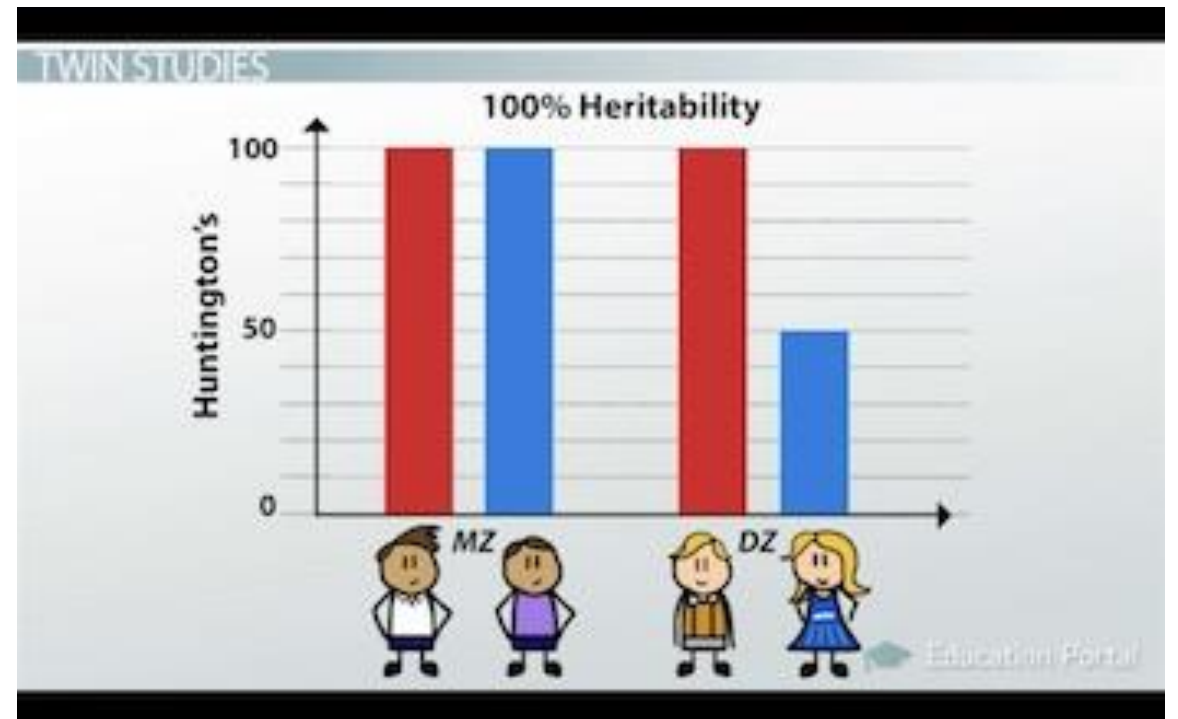
$$h^2_b = b/r = t/r$$

Dove b e' la pendenza della regressione (o t il coefficiente di correlazione) ed r e' il livello di parentela.

Stima dell'ereditabilità' a partire da confronto tra gemelli omozigoti ed eterozigoti

$$h^2_b = V_G / V_P$$

L'ereditabilità' sensu amplo (o broad sense heritability) e' la proporzione di varianza fenotipica spiegata da quella genotipica



L'ereditabilita' misura quanto i fattori genetici spiegano della varianza fenotipica *in una popolazione*

$$h^2_b = V_G / V_P$$

L'ereditabilita' sensu amplo (o broad sense heritability) e' la proporzione di varianza fenotipica spiegata da quella genotipica

Come pensate possa essere stimata?







QUANTITATIVE GENETICS IN THE WILD



EDITED BY ANNE **CHARMANTIER**,
DANY **GARANT** & LOESKE E. B. **KRUK**



Shared community effects and the non-genetic maternal environment shape cortisol levels in wild chimpanzees

Patrick J. Tkaczynski , Fabrizio Mafessoni , Cédric Girard-Buttoz, Liran Samuni, Corinne Y. Ackermann, Paweł Fedurek, Cristina Gomes, Catherine Hobaiter, Therese Löhrrich, Virgile Manin, Anna Preis, Prince D. Valé, Erin G. Wessling, Livia Wittiger, Zinta Zommers, Klaus Zuberbuehler, Linda Vigilant, Tobias Deschner, Roman M. Wittig & Catherine Crockford

Communications Biology 6, Article number: 565 (2023) | [Cite this article](#)

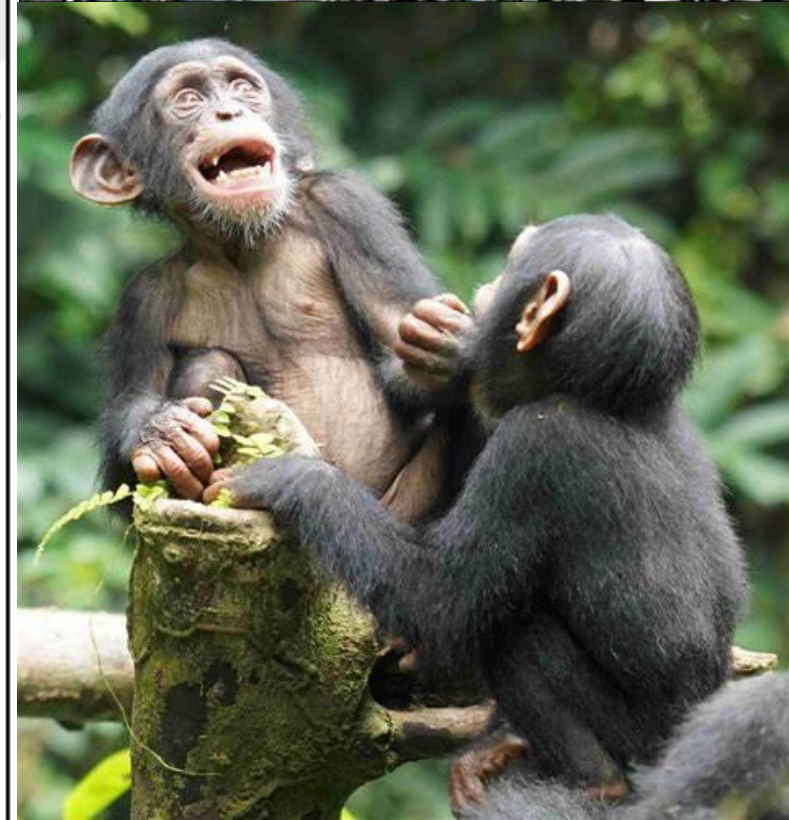


Table 1 Summary statistics for final dataset used in the study.

	N individuals	N samples	Mean (±SD) N of samples per subject
All	170	6123	36.02 (±48.17)
Adult males	48	3243	67.56 (±79.97)
Adult females	69	1742	23.86 (±19.72)
Immature males	37	545	17.03 (±16.58)
Immature females	32	593	15.95 (±11.09)
By pedigree			
Mothers with offspring in dataset	19	648	34.11 (±24.20)
Fathers with offspring in dataset	11	977	88.82 (±77.03)
Offspring with only their mother sampled	31	1335	43.06 (±51.26)
Offspring with only their father sampled	9	425	47.22 (±76.00)
Offspring with both parents sampled	33	1096	33.21 (41.42)
Individuals with only maternal half siblings in dataset	18	924	51.33 (±58.59)
Individuals with only paternal half siblings in dataset	28	699	24.96 (±22.04)
Individuals with full siblings in dataset	2	135	67.50 (±7.78)
Individuals with both maternal & paternal half siblings in dataset	31	1467	47.32 (±58.23)
Individuals without relations in dataset	62	1928	31.10 (±48.20)
By Population-group			
Taï-East	33	1531	46.39 (±72.61)
Taï-North	24	842	35.08 (±28.92)
Taï-South	51	2470	48.43 (±56.58)
Budongo-Sonso	45	1171	24.40 (±17.82)
Budongo-Waibira	17	109	7.79 (±3.34)

In total, 6123 urinary cortisol values from 170 individuals were included in the study. Note that certain individuals fall into several pedigree categories (e.g. an individual can be a father and have a maternal or paternal sibling), therefore, the number of individuals in pedigree categorization exceeds 170. The range of numbers of years of sampling of individuals in the dataset was 1–13 years, with a mean ± SD of 2.63 ± 3.01 years.

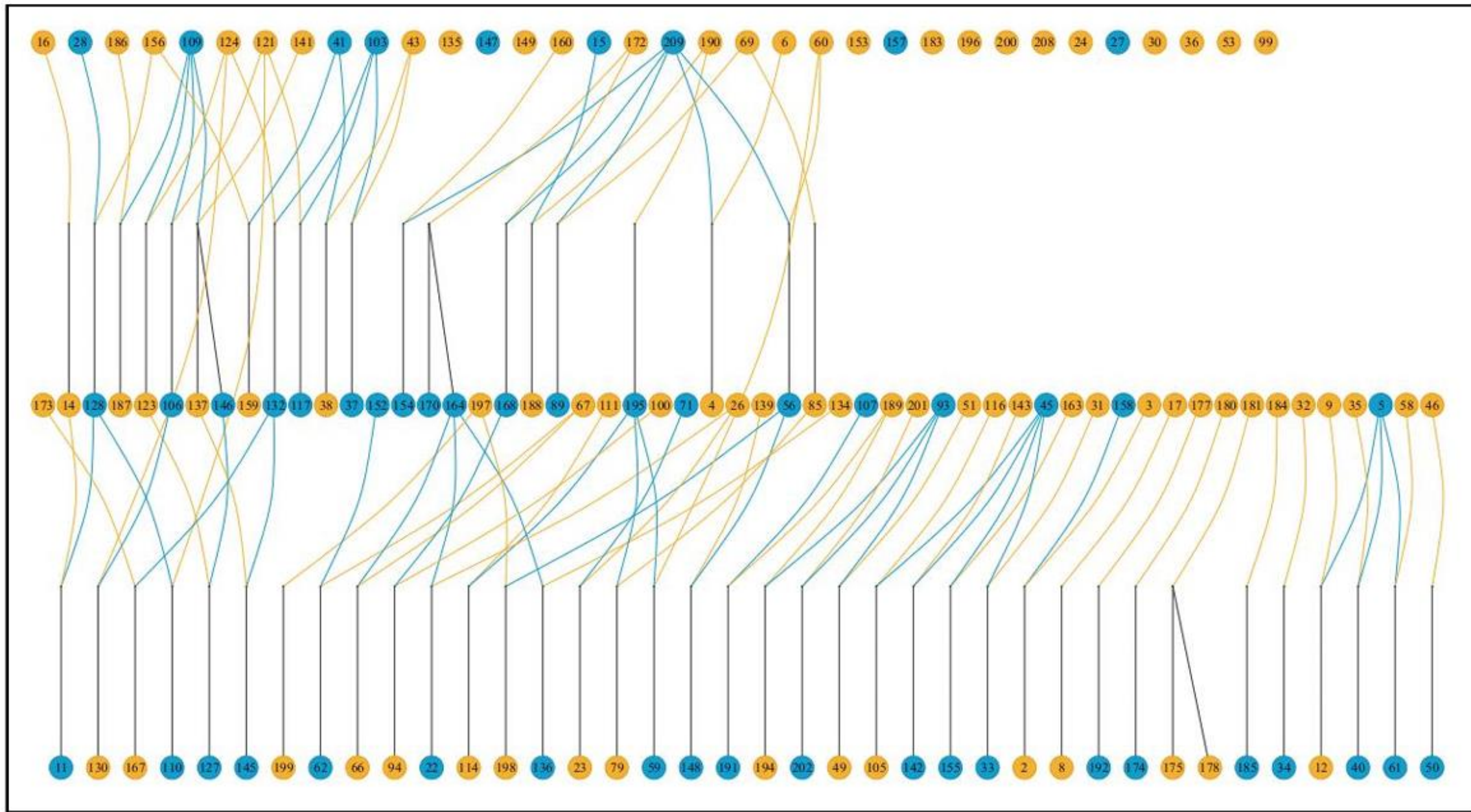
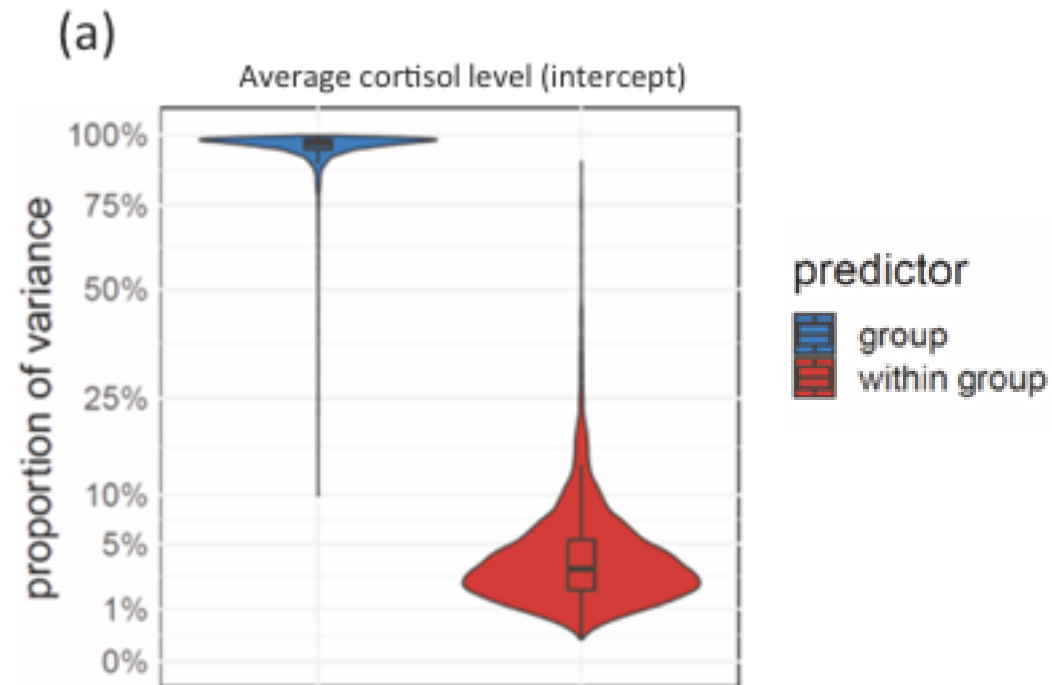


Figure S1: Pedigree of individuals from the Tai field site included within our analysis. Males are represented by the blue circles, females by the yellow circles. Note, this pedigree illustration includes the mothers and/or fathers of individuals with urinary cortisol values in our study, even if these mothers and/fathers were not sampled themselves. However, the offspring of individuals are only included if the offspring themselves had urinary cortisol values in the study.

Il cortisolo negli scimpanze' varia soprattutto in relazione al gruppo, non a fattori strettamente genetici



La varianza puo' essere scomposta nelle sue component genetiche e ambientali

- $V(\text{totale})=V(\text{entro-gruppi})+V(\text{tra-gruppi})$
- $V(\text{entro-gruppi})=V(\text{genetica})+V(\text{maternal})+V(\text{individuale})+V(\text{ambiente})$
- $h^2(\text{entro gruppi})= V(\text{genetica})/ V(\text{entro-gruppi})$
- $m^2(\text{entro gruppi})= V_{\text{materna}}/ V(\text{entro-gruppi})$

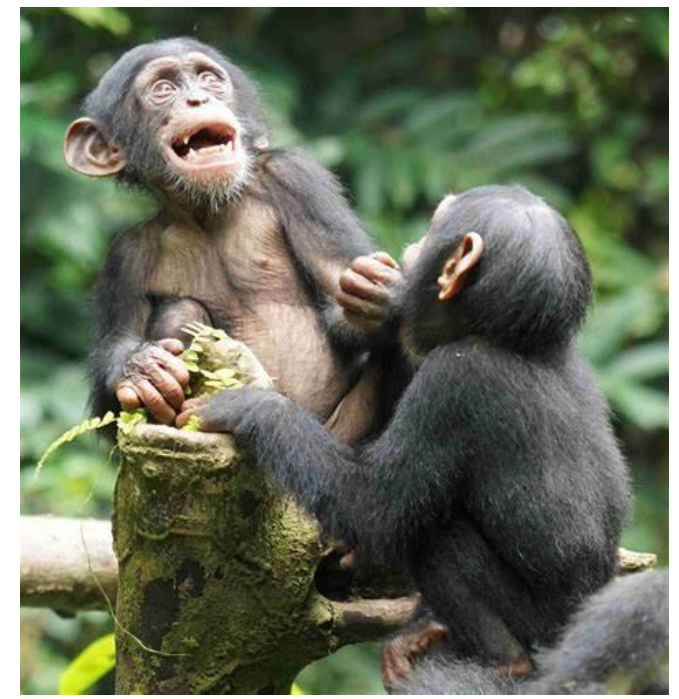
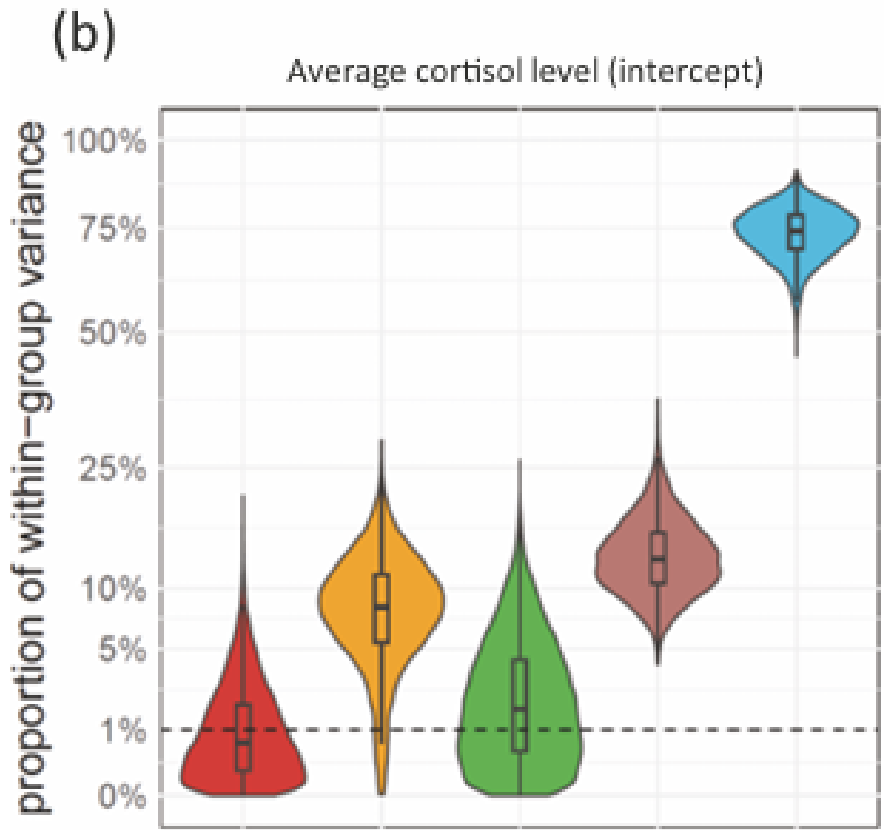
Table 4 Summary of genetic, maternal, shared community group-year, individual-year and individual effect estimates (with 95% credible intervals) on circadian cortisol responses in wild chimpanzees.

Coefficient	Intercept		Linear slope		Quadratic slope	
	Estimate	(ICI, uCI)	Estimate	(ICI, uCI)	Estimate	(ICI, uCI)
Proportion of variance in components of circadian urinary cortisol reaction norms						
Between-group effects (Shared community effects)	0.97	(0.87, 0.99)	0.63	(0.01, 0.98)	0.15	(0.00, 0.85)
Within group effects	0.03	(0.01, 0.13)	0.37	(0.02, 0.99)	0.85	(0.15, 1.00)
Proportion of within-group variance						
Genetic effects	0.01	(0.00, 0.05)	0.05	(0.00, 0.38)	0.04	(0.00, 0.31)
Maternal effects	0.08	(0.01, 0.16)	0.09	(0.00, 0.49)	0.03	(0.00, 0.29)
Group-year effects	0.74	(0.62, 0.83)	0.40	(0.01, 0.84)	0.36	(0.01, 0.84)
Individual identity effect	0.02	(0.00, 0.10)	0.05	(0.00, 0.41)	0.04	(0.00, 0.35)
ID-year effect	0.13	(0.08, 0.21)	0.17	(0.00, 0.70)	0.36	(0.01, 0.84)

Each coefficient represents a different component of the circadian cortisol response. Coefficients in bold had genetic/maternal/group_year/ID_year/individual identity effect contributions with 95% intervals distinguishable from 0%, i.e. either with a credible interval larger or equal to 1%, or significantly higher than 0 according to the dataset permutation (Fig. 2). Italicized coefficients are those with lower credible intervals larger or equal to 1% yet associated with a high degree of uncertainty (interval range exceeds 80%).

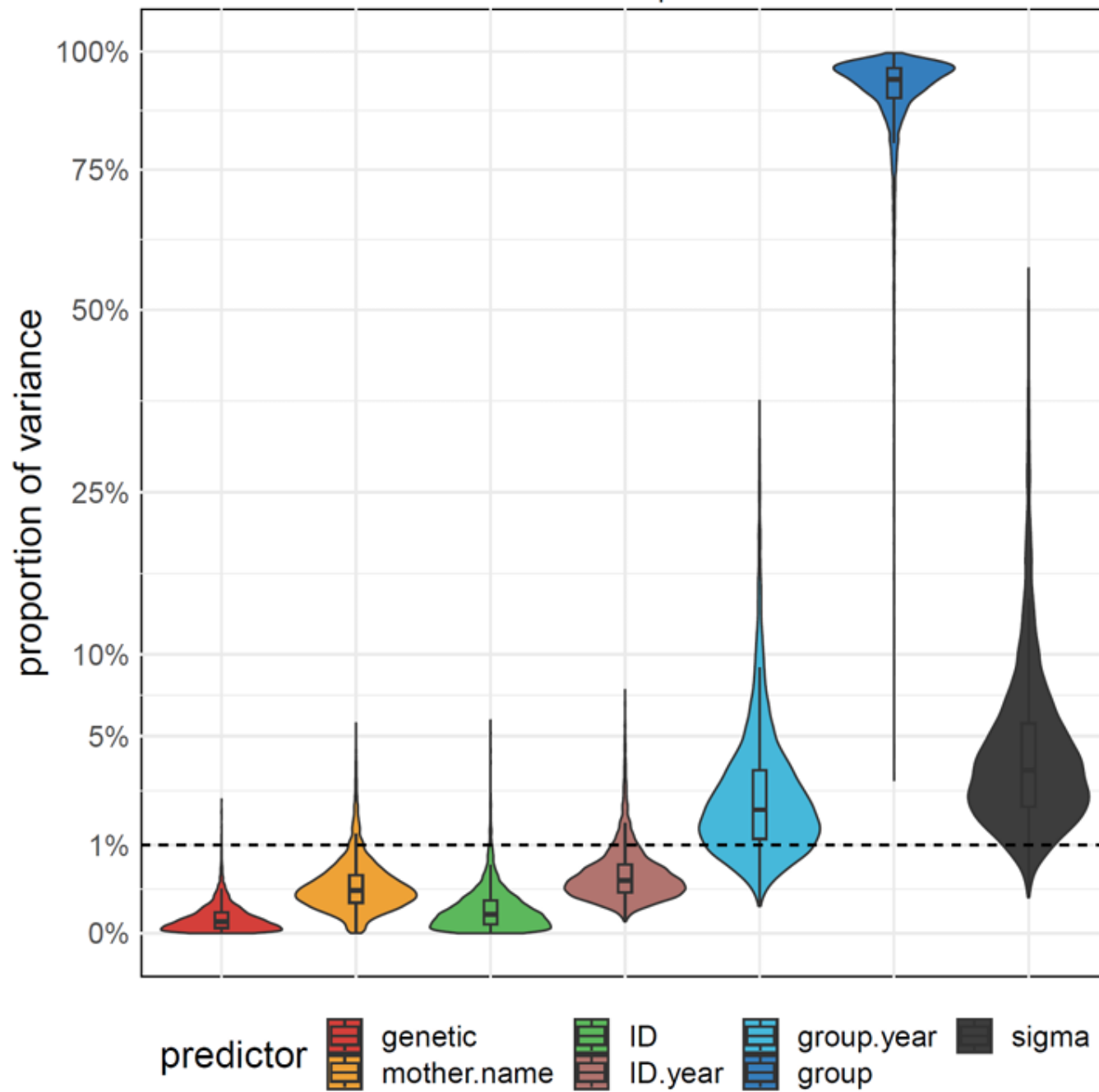
L'ambiente materno puo' confondersi con gli effetti genetici

- Negli scimpanze' e' piu' importante della componente genetica nel determinare i livelli di cortisolo



predictor  genetic  mother.name  ID  ID.year  group.year

Intercept



Necessita' di Metodi numerici (ad esempio Metodi Bayesiani come Markov Chain Monte Carlo) per la stima della varianza

3.4 Results and diagnostic of the MCMC output

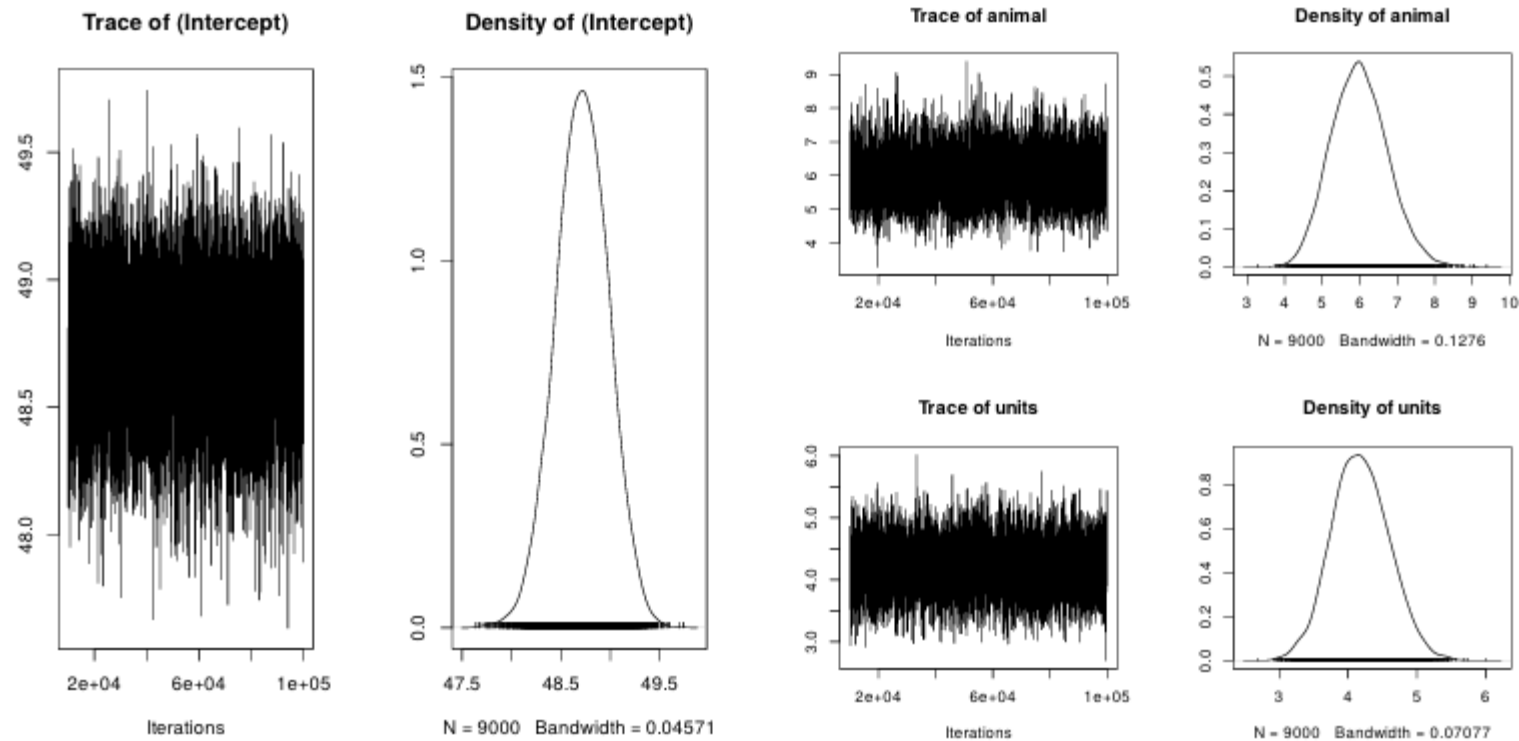
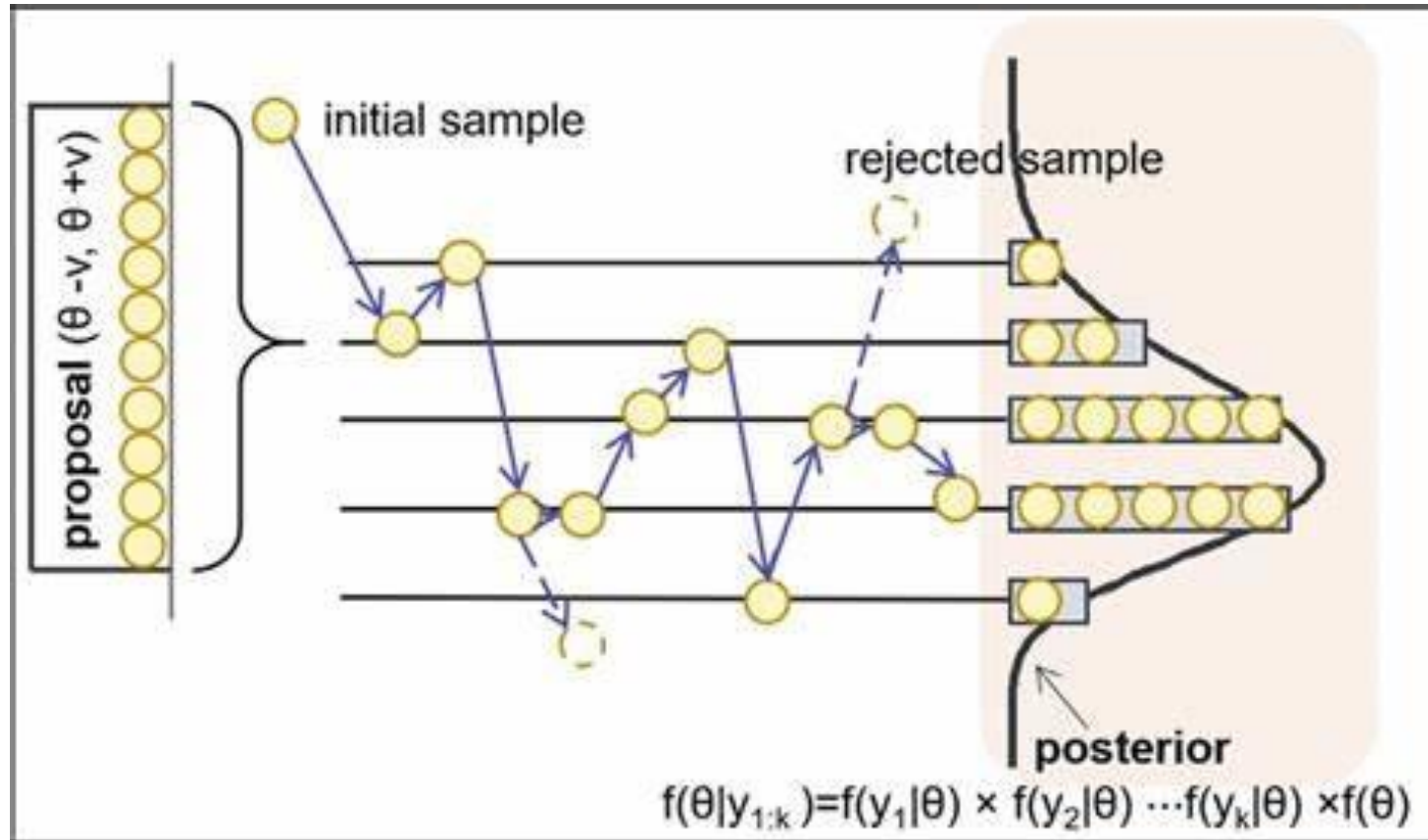


Figure 3: Trace of the intercept μ (or population mean, left) and the variances (right). The term animal refers to V_A and the term units refers to V_R .

Markov Chain Monte Carlo



L'ereditabilità misura quanto i fattori genetici spiegano della varianza fenotipica *in una popolazione*

$$h^2_b = V_G / V_P$$

L'ereditabilità sensu amplo (o broad sense heritability) è la proporzione di varianza fenotipica spiegata da quella genotipica

Quali sono i bias che questa misura può avere?



I bias delle stime di ereditabilita'

- Effetti parentali
- Accoppiamento assortativo
(assortative mating)
- Ambiente commune
- Comparazione tra
campioni/popolazioni diverse
(piu o meno rumore e varianza)

Quali sono i bias che
questa misura puo'
avere?



La varianza genetica puo' essere ulteriormente scomposta nella sua componente additiva

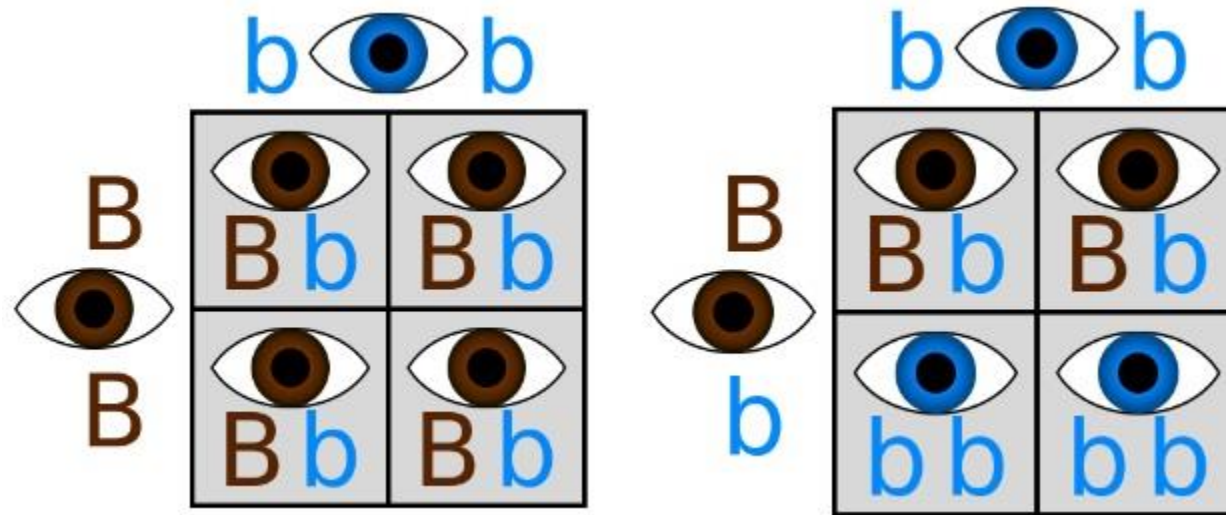
$$V_P = V_G + V_E$$

Dove

$$V_G = V_A + V_D + V_I$$

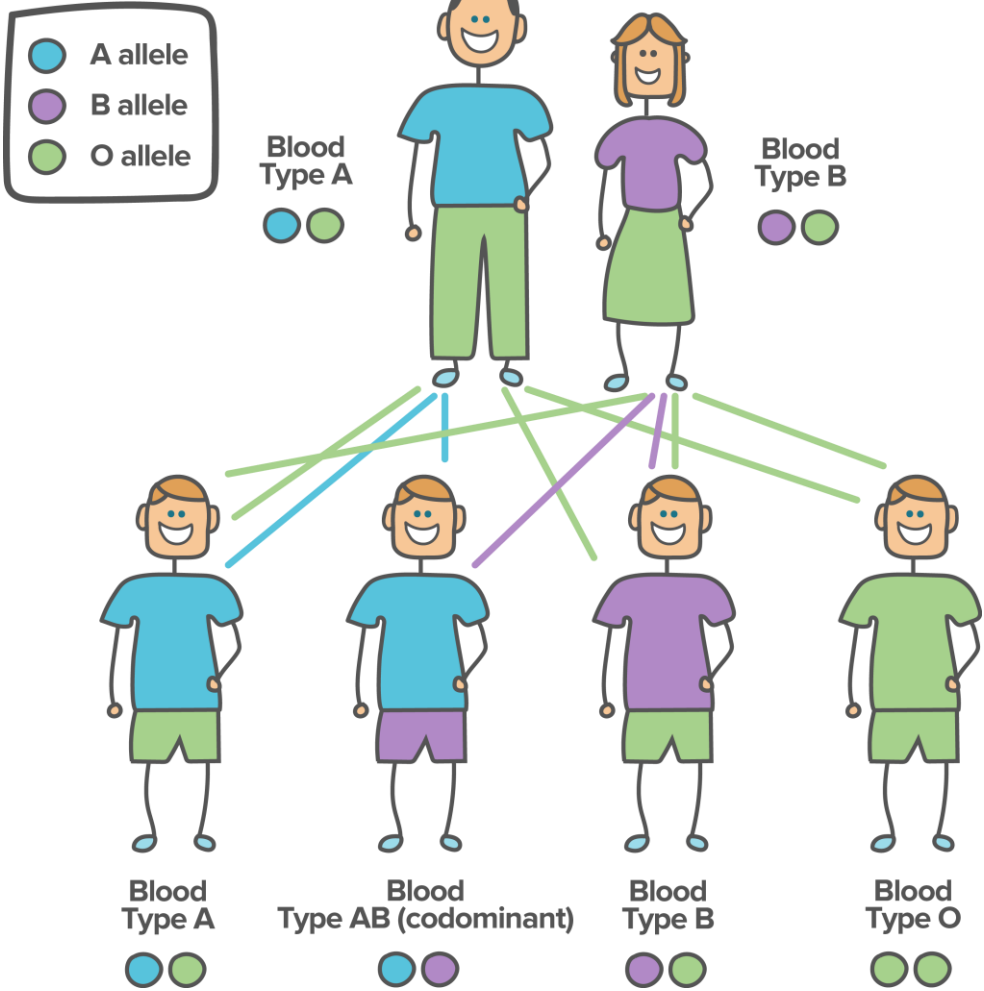
con Varianza Genetica Additiva V_A , Varianza dovuta alla dominanza V_D e alle interazioni epistatiche V_I

Dominanza

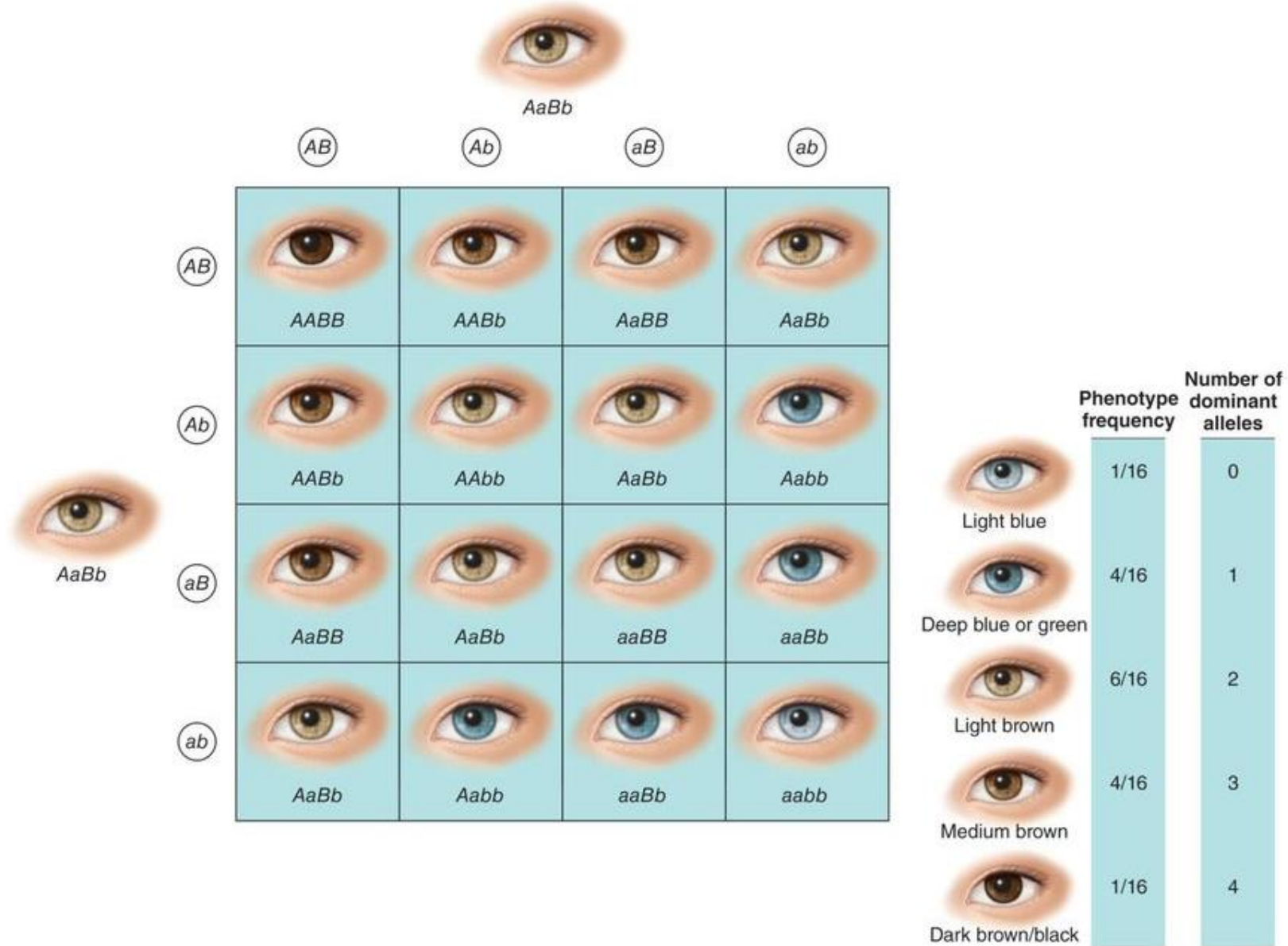


Blu e' recessivo, marrone e' dominante

Alleli codominanti

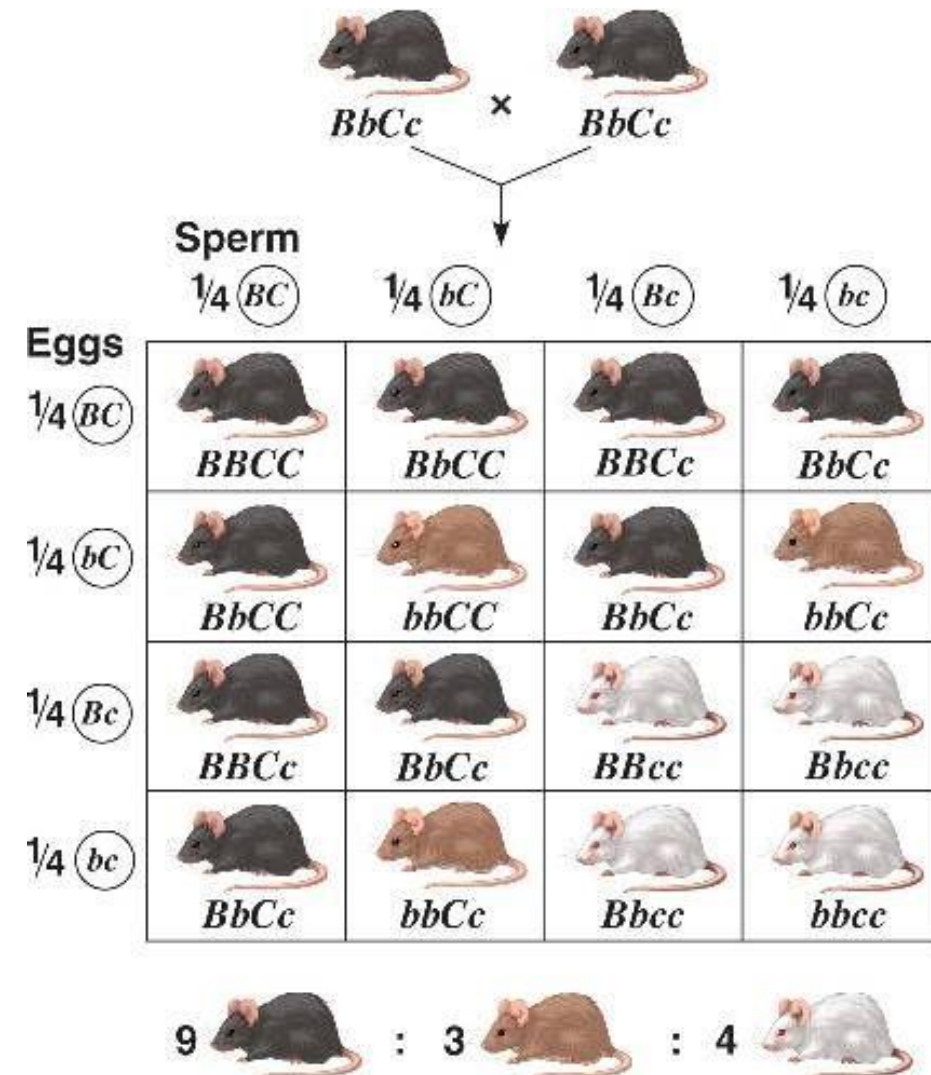


Dominanza incompleta ed epistasi

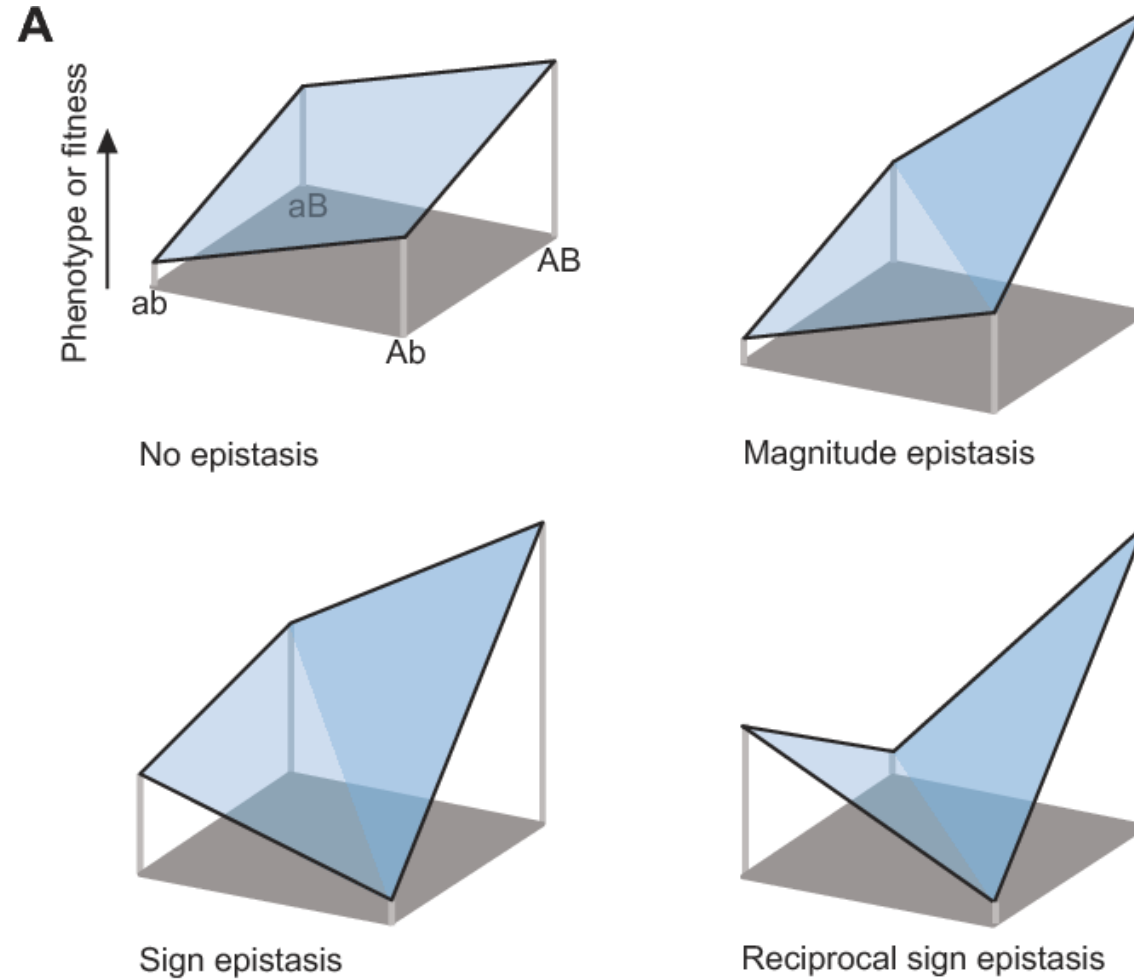


Epistasi

Combinazioni di alleli su loci diversi determinano effetti non-additivi/non-lineari su fenotipo o fitness



Esistono varie tipologie di epistasi



L'ereditabilità' misura quanto i fattori genetici spiegano della varianza fenotipica *in una popolazione*

$$h^2_b = V_G / V_P$$

L'ereditabilità' sensu amplo (o broad sense heritability) e' la proporzione di varianza fenotipica spiegata da quella genotipica

Alta ereditabilità':

La variabilità' nel tratto e' dovuto soprattutto a cause genetiche



Huntington Disease Brain Normal Brain

Bassa ereditabilità':

La variabilità' nel tratto non e' dovuta ai geni



L'ereditabilità misura quanto i fattori genetici spiegano della varianza fenotipica *in una popolazione*

$$h^2_b = V_G / V_P$$

L'*ereditabilità sensu amplo* (o broad sense heritability) è la proporzione di varianza fenotipica spiegata da quella genotipica

$$h^2_n = V_A / V_P$$

L'*ereditabilità sensu strictu* (o narrow sense heritability) è la proporzione di varianza fenotipica spiegata dalla varianza genetica additiva

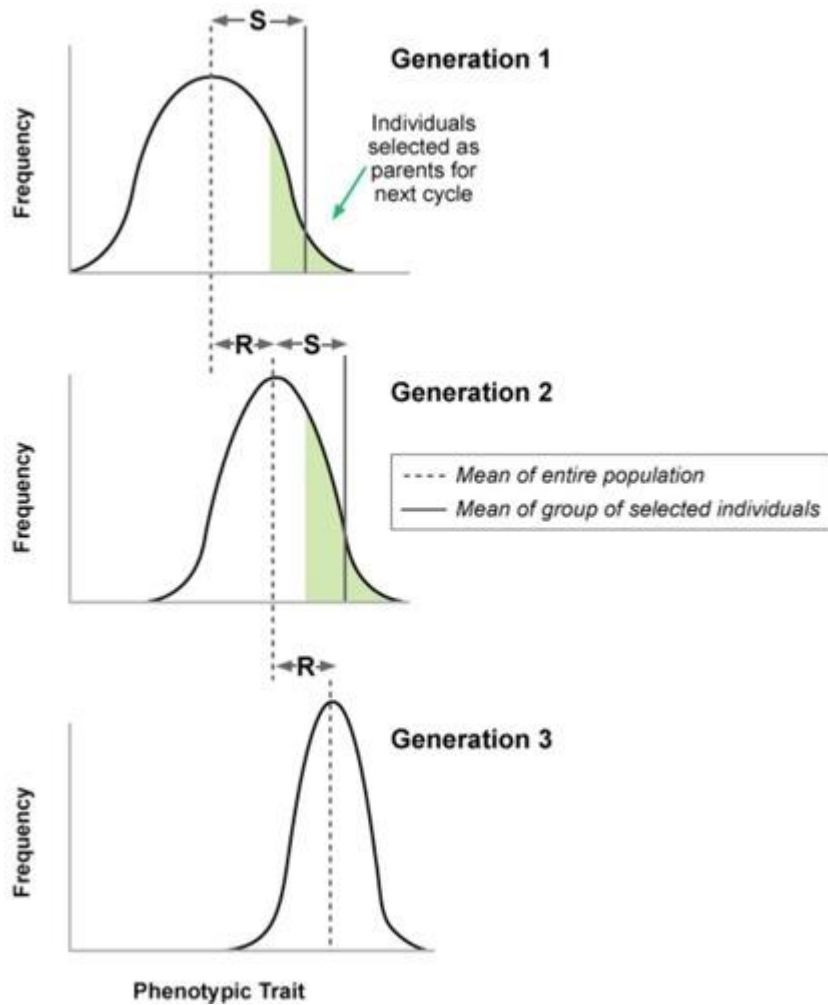
Tenendo presente che $V_P = V_G + V_E$ e $V_G = V_A + V_D + V_I$

La varianza genetica additiva e' il principale substrato della selezione naturale

$$\Delta Z = h^2_n S$$

- Dove ΔZ e' la differenza in un tratto Z in due generazioni successive, e S e' la forza della selezione (o differenziale di selezione) ed e' una misura dell'associazione tra il tratto e la selezione.
- Questa viene definita anche breeder's equation (**equazione dell'allevatore**).
- Questo accade perche' dominanza ed epistasi dipendono dalle combinazioni di geni e quindi non sono trasmesse efficientemente tra generazioni, mentre gli effetti additivi sono indipendenti e vengono passati di generazione in generazione.

L'equazione dell'allevatore



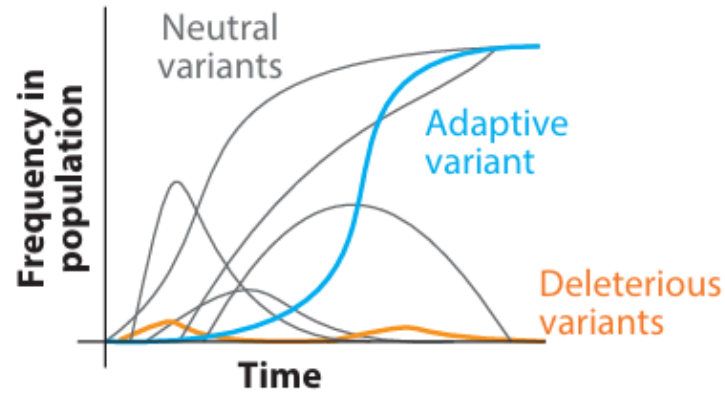
Sara' selezionato piu' efficacemente un tratto con bassa o alta ereditabilita'?

Con alta o bassa varianza genetica additiva?

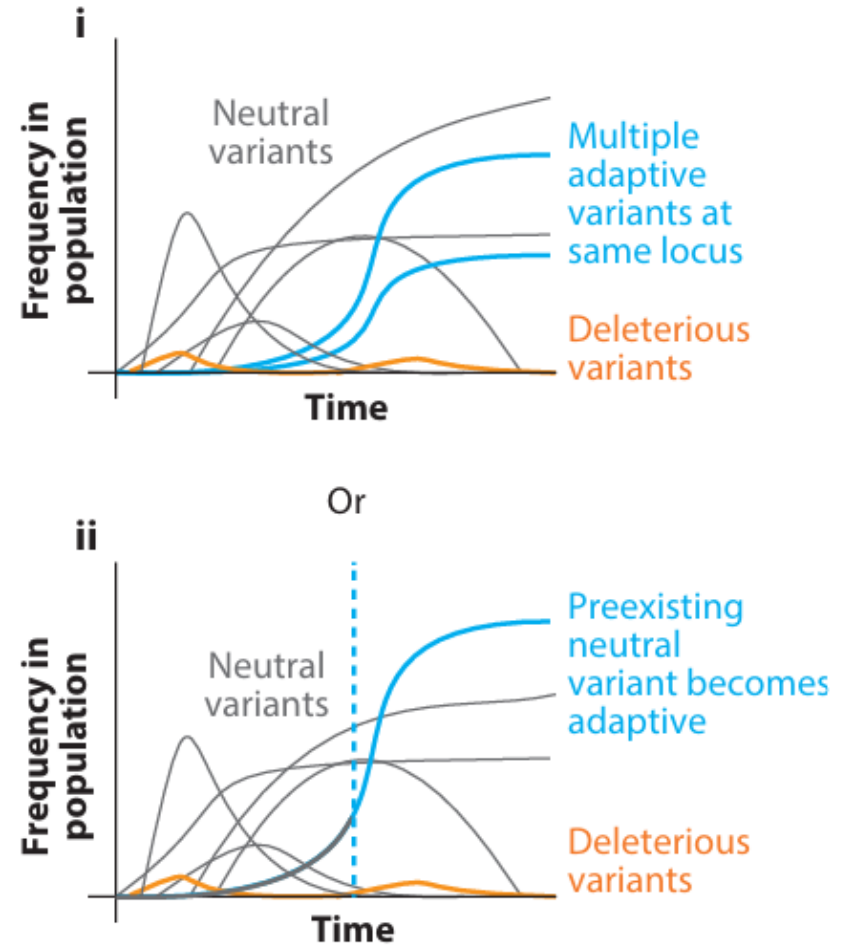


Gli alleli sotto selezione possono aumentare in una popolazione come **hard** o **soft sweeps**

a Selection and drift



b Soft sweeps



Gli hard sweep sono dovuti ad una nuova mutazione fortemente adattativa, al contrario dei soft-sweeps

La selezione naturale genera variazioni nelle frequenze alleliche di diverso tipo

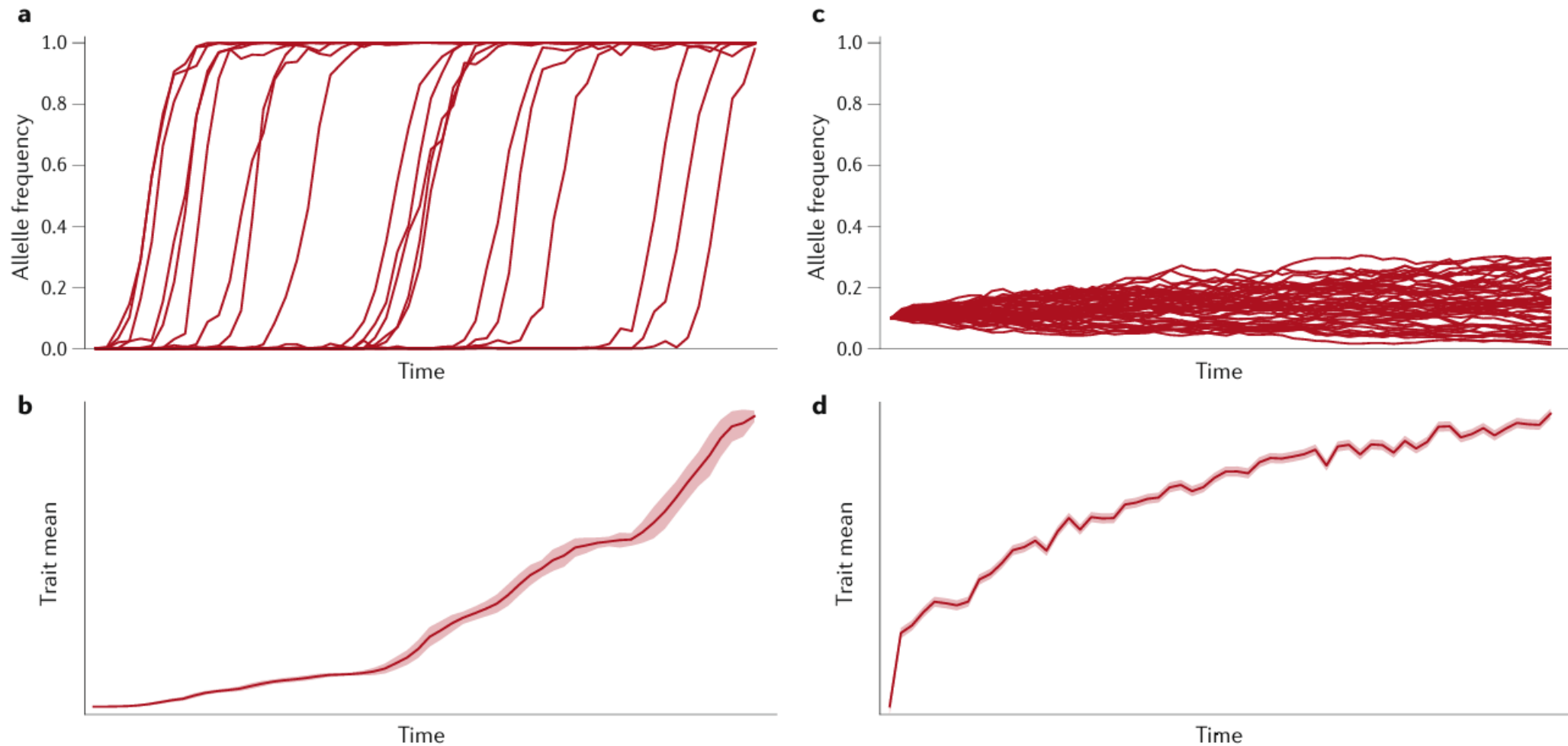
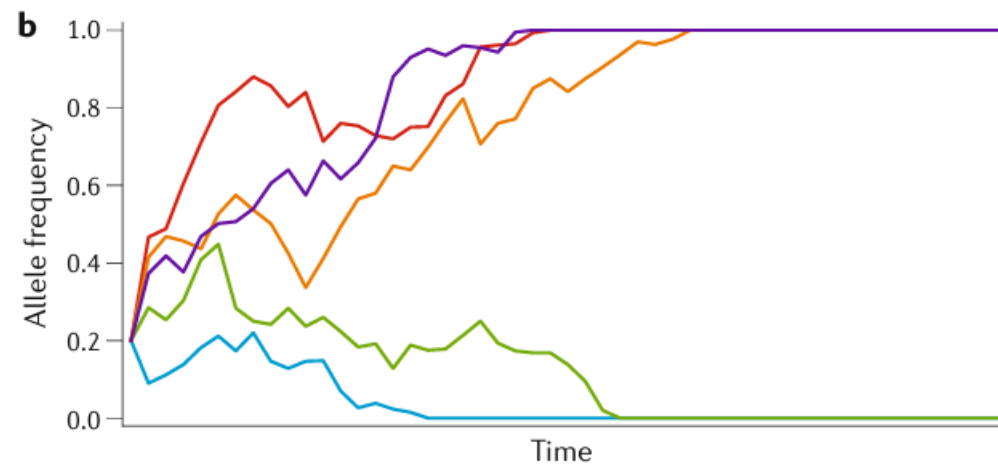
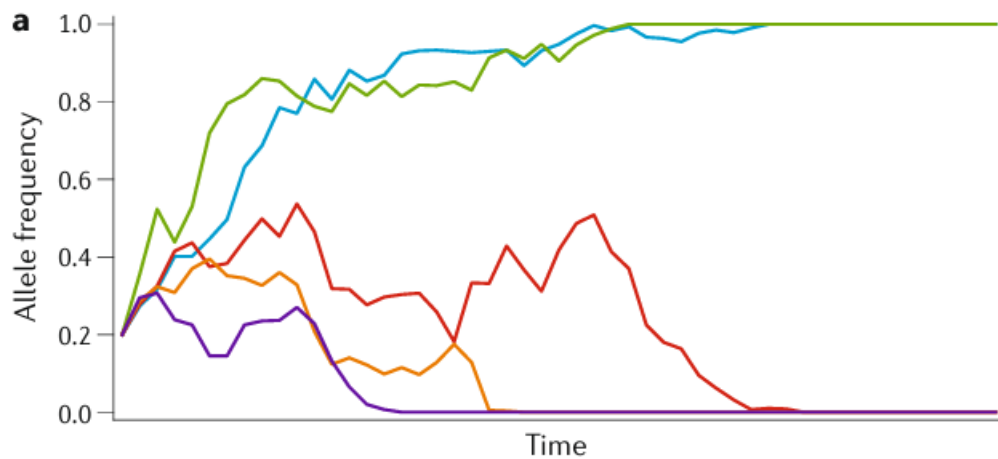


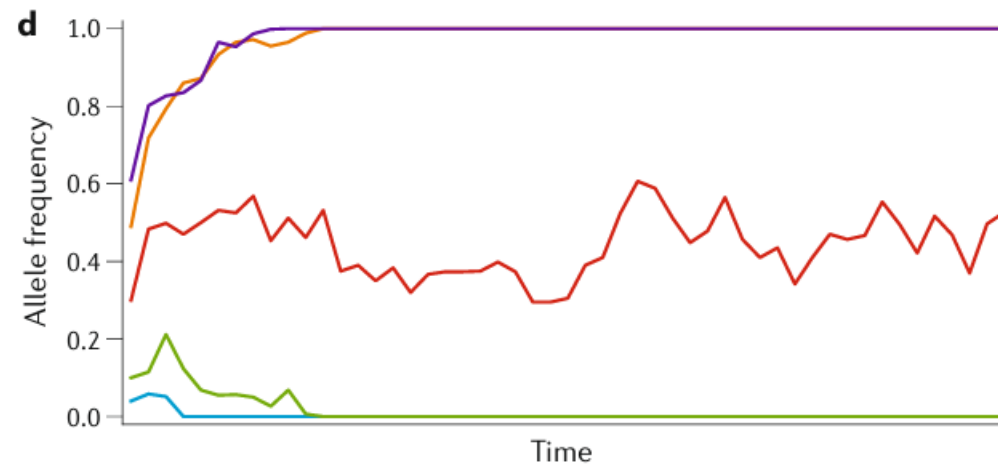
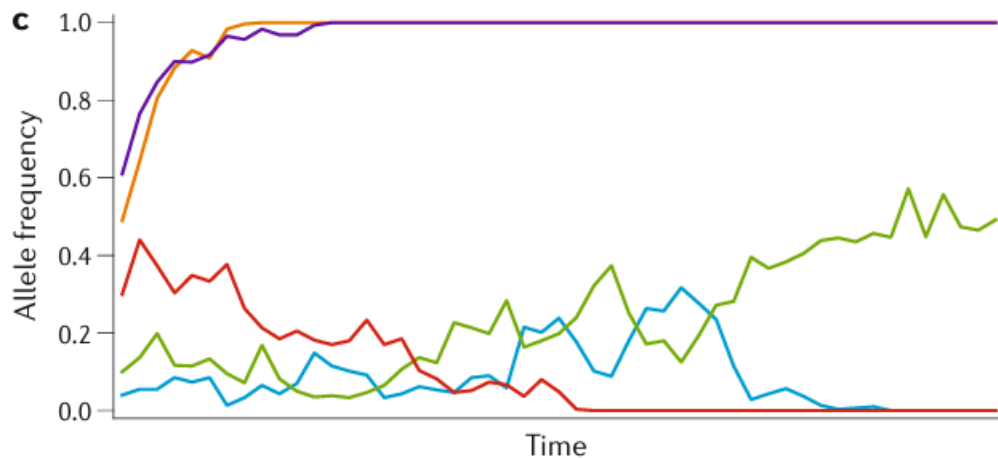
Fig. 1 | **Alternative adaptive paradigms: selective sweeps versus the infinitesimal model.** **a,b** | Selective sweeps. Beneficial mutations arise recurrently in a population and increase in frequency until fixation (part **a**). The consecutive fixation of beneficial alleles increases the fitness of the population (part **b**). **c,d** | The infinitesimal model. A large number of contributing alleles experience subtle frequency changes after a shift in the trait optimum (part **c**) that result in changes in the mean trait of the population (part **d**). For simulation parameters, see Supplementary information.

Gli effetti della selezione naturale sugli alleli dipendono dal loro coefficiente di selezione e dalla loro frequenza iniziale

Adaptation to the same trait optimum using different alleles

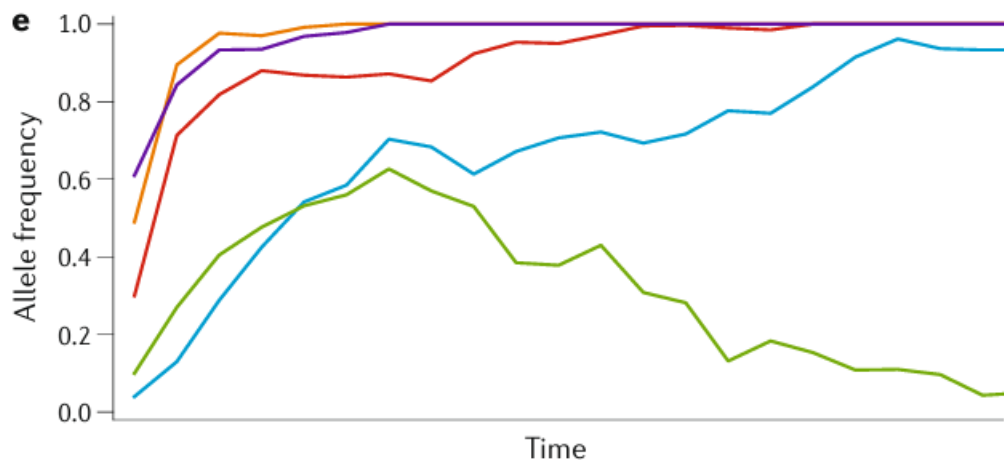


Repeatable adaptive response via high-frequency alleles

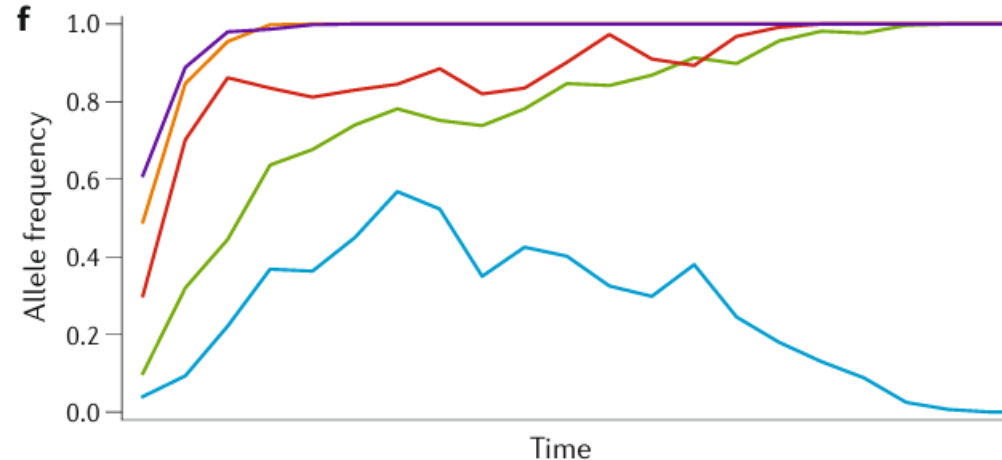


Gli effetti della selezione naturale sugli alleli dipendono dal loro coefficiente di selezione e dalla loro frequenza iniziale

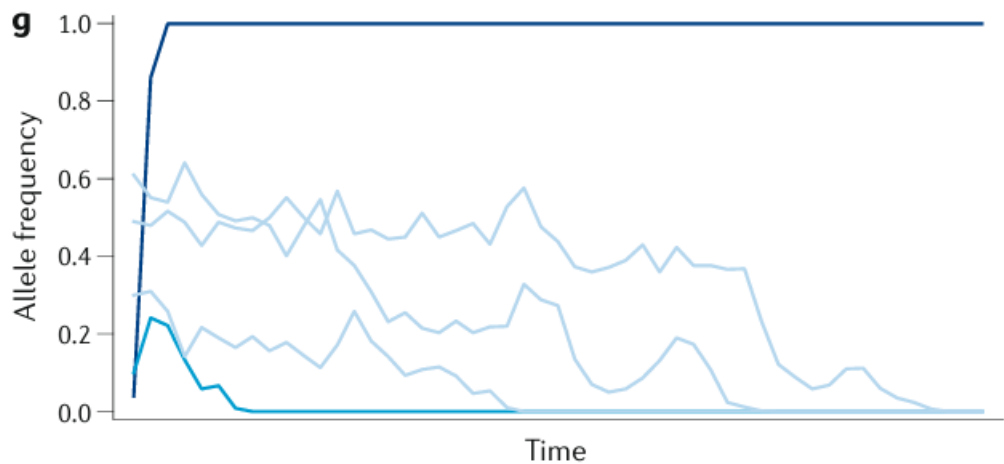
e Low-frequency alleles contribute to adaptation to a distant trait optimum



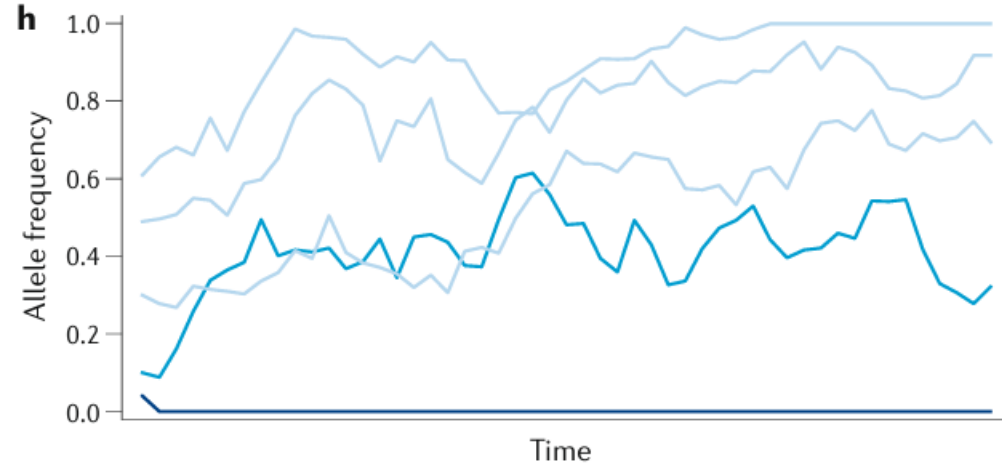
f



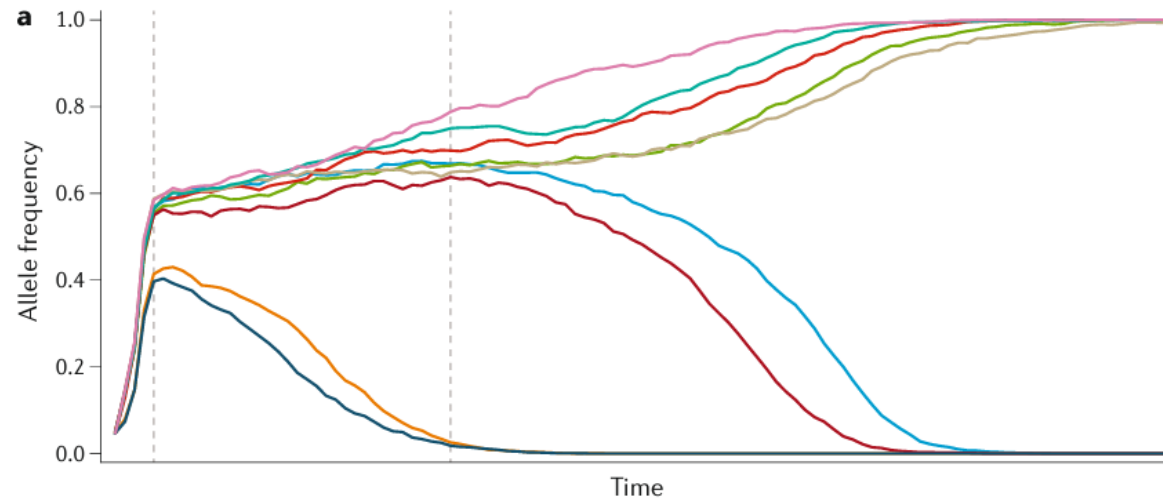
g Adaptation to the same trait optimum via sweeps or shifts



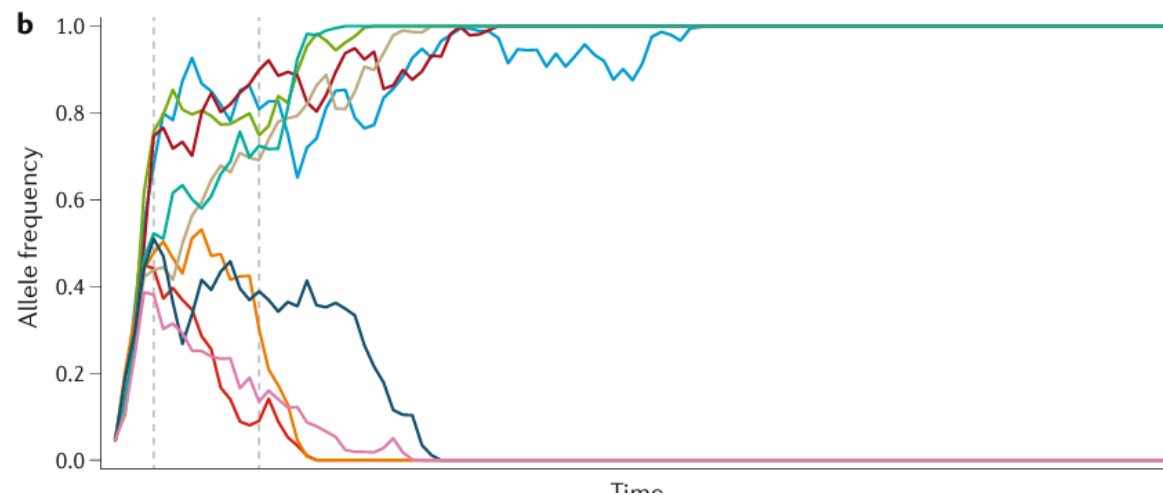
h



La dimensione di una popolazione determina l'intensità della selezione



1000 individui: lunga fase deterministica dovuta alla selezione



100 individui: la deriva genetica domina

Qual'è la differenza tra valori genotipici e valori genotipici additivi?

- Il **valore genotipico di un genotipo** è semplicemente il valore del fenotipo/tratto medio associato ad un certo genotipo.
- Il valore genotipico additivo è il valore del fenotipo/tratto medio associato ad un certo genotipo, se potessimo scomporlo dalle sue componenti di dominanza ed epistasi.
- E ora un po' di matematica (noiosa ma semplice) per mostrare che:
 - La differenza tra valore genotipico e valore genotipico additivo
 - La varianza genotipica si può scomporre in additiva e dovuta alla dominanza

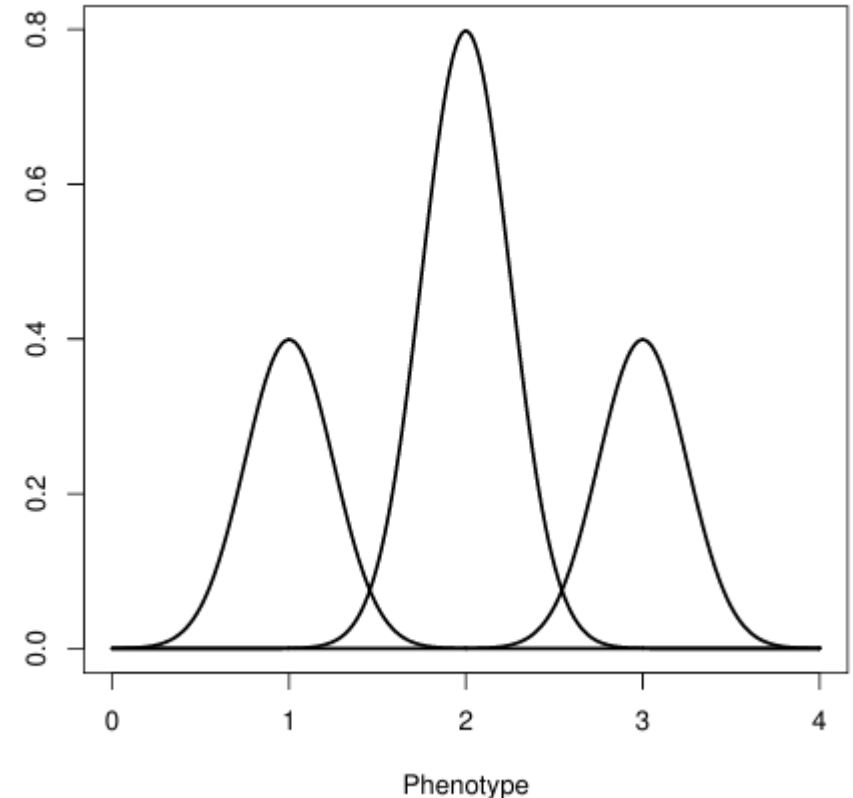
Qual'è la differenza tra valori genotipici e valori genotipici additivi?

Un esempio partendo dal valore genotipico.

Genotype	A_1A_1	A_1A_2	A_2A_2
Frequency	p^2	$2pq$	q^2
Genotypic value	x_{11}	x_{12}	x_{22}
Additive genotypic value	$2\alpha_1$	$\alpha_1 + \alpha_2$	$2\alpha_2$

Esempio:

In una popolazione ad uguale frequenza
Di A1 e A2 in equilibrio di Hardy-Weinberg:
A1A1 ha tratto medio 1,
A1A2 ha tratto medio 2,
A2A2 ha tratto medio 3,



Qual'è la differenza tra valori genotipici e valori genotipici additivi?

Un esempio partendo dal valore genotipico.

Genotype	A_1A_1	A_1A_2	A_2A_2
Frequency	p^2	$2pq$	q^2
Genotypic value	x_{11}	x_{12}	x_{22}
Additive genotypic value	$2\alpha_1$	$\alpha_1 + \alpha_2$	$2\alpha_2$



Esempio:

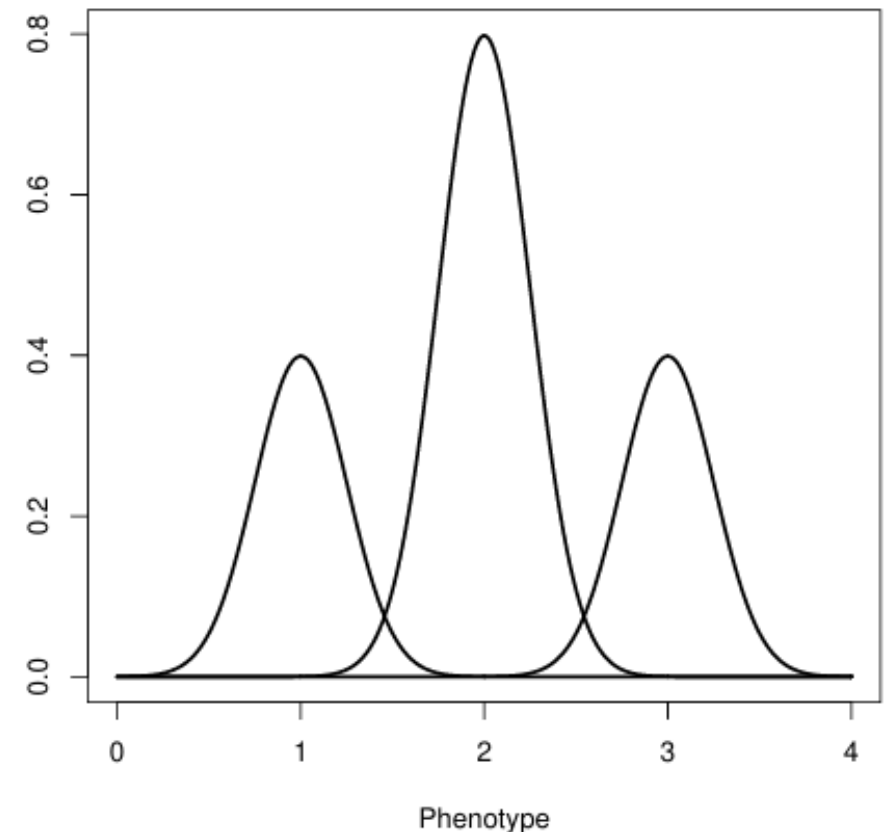
In una popolazione ad uguale frequenza

Di A_1 e A_2 in equilibrio di Hardy-Weinberg:

A_1A_1 ha tratto medio 1,

A_1A_2 ha tratto medio 2,

A_2A_2 ha tratto medio 3,



Qual'è la differenza tra valori genotipici e valori genotipici additivi?

Un esempio partendo dal valore genotipico.

Genotype	A_1A_1	A_1A_2	A_2A_2
Frequency	p^2	$2pq$	q^2
Genotypic value	x_{11}	x_{12}	x_{22}
Additive genotypic value	$2\alpha_1$	$\alpha_1 + \alpha_2$	$2\alpha_2$

Possiamo assumere che in generale valori genotipici additivi siano i più vicini possibili a quelli genotipici (e che abbiano effetti additivi), cercando quindi di minimizzare una funzione a

$$a = p^2[x_{11} - 2\alpha_1]^2 + 2pq[x_{12} - (\alpha_1 + \alpha_2)]^2 + q^2[x_{22} - 2\alpha_2]^2$$

Qual'è la differenza tra valori genotipici e valori genotipici additivi?

Un esempio partendo dal valore genotipico.

Genotype	A_1A_1	A_1A_2	A_2A_2
Frequency	p^2	$2pq$	q^2
Genotypic value	x_{11}	x_{12}	x_{22}
Additive genotypic value	$2\alpha_1$	$\alpha_1 + \alpha_2$	$2\alpha_2$

Possiamo assumere che in generale valori genotipici additivi siano i più vicini possibili a quelli genotipici (e che abbiano effetti additivi), cercando quindi di minimizzare una funzione a

$$a = p^2[x_{11} - 2\alpha_1]^2 + 2pq[x_{12} - (\alpha_1 + \alpha_2)]^2 + q^2[x_{22} - 2\alpha_2]^2$$

Possiamo trovare i valori del contributo genetico additivo che minimizzano a .

$$\begin{aligned}\frac{\partial a}{\partial \alpha_1} &= p^2\{2[x_{11} - 2\alpha_1][-2]\} + 2pq\{2[x_{12} - (\alpha_1 + \alpha_2)][-1]\} \\ &= -4p^2[x_{11} - 2\alpha_1] - 4pq[x_{12} - (\alpha_1 + \alpha_2)]\end{aligned}$$

$$\begin{aligned}\frac{\partial a}{\partial \alpha_2} &= q^2\{2[x_{22} - 2\alpha_2][-2]\} + 2pq\{2[x_{12} - (\alpha_1 + \alpha_2)][-1]\} \\ &= -4q^2[x_{22} - 2\alpha_2] - 4pq[x_{12} - (\alpha_1 + \alpha_2)]\end{aligned}$$

Qual'è la differenza tra valori genotipici e valori genotipici additivi?

Un esempio partendo dal valore genotipico.

Genotype	A_1A_1	A_1A_2	A_2A_2
Frequency	p^2	$2pq$	q^2
Genotypic value	x_{11}	x_{12}	x_{22}
Additive genotypic value	$2\alpha_1$	$\alpha_1 + \alpha_2$	$2\alpha_2$

Possiamo assumere che in generale valori genotipici additivi siano i più vicini possibili a quelli genotipici (e che abbiano effetti additivi), cercando quindi di minimizzare una funzione a

$$a = p^2[x_{11} - 2\alpha_1]^2 + 2pq[x_{12} - (\alpha_1 + \alpha_2)]^2 + q^2[x_{22} - 2\alpha_2]^2$$

Possiamo trovare i valori del contributo genetico additivo che minimizzano a , e cioè quei valori con derivate = 0.

$$p^2(x_{11} - 2\alpha_1) + pq(x_{12} - \alpha_1 - \alpha_2) = 0$$

$$q^2(x_{22} - 2\alpha_2) + pq(x_{12} - \alpha_1 - \alpha_2) = 0$$

..se proprio volete seguire i calcoli..

$$[p^2x_{11} + 2pqx_{12} + q^2x_{22}] - [p^2(2\alpha_1) + 2pq(\alpha_1 + \alpha_2) + q^2(2\alpha_2)] = 0$$

Dove il primo membro non e' altro che il tratto medio \bar{x} .

$$\begin{aligned}\bar{x} &= 2p^2\alpha_1 + 2pq(\alpha_1 + \alpha_2) + 2q^2\alpha_2 \\ &= 2p\alpha_1(p + q) + 2q\alpha_2(p + q) \\ &= 2(p\alpha_1 + q\alpha_2) \quad .\end{aligned}$$

$$p(x_{11} - 2\alpha_1) + q(x_{12} - \alpha_1 - \alpha_2) = 0$$

$$q(x_{22} - 2\alpha_2) + p(x_{12} - \alpha_1 - \alpha_2) = 0$$

$$\begin{aligned}px_{11} + qx_{12} &= 2p\alpha_1 + q\alpha_1 + q\alpha_2 \\ &= \alpha_1(p + q) + p\alpha_1 + q\alpha_2 \\ &= \alpha_1 + p\alpha_1 + q\alpha_2 \\ &= \alpha_1 + \bar{x}/2\end{aligned}$$

$$\begin{aligned}px_{12} + qx_{22} &= 2q\alpha_2 + p\alpha_1 + p\alpha_2 \\ &= \alpha_2(p + q) + p\alpha_1 + q\alpha_2 \\ &= \alpha_2 + p\alpha_1 + q\alpha_2 \\ &= \alpha_2 + \bar{x}/2\end{aligned}$$

..altrimenti..

$$\alpha_1 = px_{11} + qx_{12} - \bar{x}/2 \quad .$$

$$\alpha_2 = px_{12} + qx_{22} - \bar{x}/2 \quad .$$

Possiamo quindi dimostrare che V_G si può scomporre in V_A e V_D

$$V_g = p^2[x_{11} - \bar{x}]^2 + 2pq[x_{12} - \bar{x}]^2 + q^2[x_{22} - \bar{x}]^2 \quad \text{Dalla definizione di varianza}$$

Possiamo quindi dimostrare che V_G si puo' scomporre in V_A e V_D

$$V_g = p^2[x_{11} - \bar{x}]^2 + 2pq[x_{12} - \bar{x}]^2 + q^2[x_{22} - \bar{x}]^2$$

$$\begin{aligned} &= p^2[x_{11} - 2\alpha_1 + 2\alpha_1 - \bar{x}]^2 + 2pq[x_{12} - (\alpha_1 + \alpha_2) + (\alpha_1 + \alpha_2) - \bar{x}]^2 \\ &\quad + q^2[x_{22} - 2\alpha_2 + 2\alpha_2 - \bar{x}]^2 \\ &= p^2[x_{11} - 2\alpha_1]^2 + 2pq[x_{12} - (\alpha_1 + \alpha_2)]^2 + q^2[x_{22} - 2\alpha_2]^2 \\ &\quad + p^2[2\alpha_1 - \bar{x}]^2 + 2pq[(\alpha_1 + \alpha_2) - \bar{x}]^2 + q^2[2\alpha_2 - \bar{x}]^2 \\ &\quad + p^2[2(x_{11} - 2\alpha_1)(2\alpha_1 - \bar{x})] + 2pq[2(x_{12} - \{\alpha_1 + \alpha_2\})(\{\alpha_1 + \alpha_2\} - \bar{x})] \\ &\quad + q^2[2(x_{22} - 2\alpha_2)(2\alpha_2 - \bar{x})] \quad . \end{aligned}$$

Possiamo quindi dimostrare che V_G si puo' scomporre in V_A e V_D

$$V_g = p^2[x_{11} - \bar{x}]^2 + 2pq[x_{12} - \bar{x}]^2 + q^2[x_{22} - \bar{x}]^2$$

$$= p^2[x_{11} - 2\alpha_1 + 2\alpha_1 - \bar{x}]^2 + 2pq[x_{12} - (\alpha_1 + \alpha_2) + (\alpha_1 + \alpha_2) - \bar{x}]^2 + q^2[x_{22} - 2\alpha_2 + 2\alpha_2 - \bar{x}]^2$$

$$= \boxed{p^2[x_{11} - 2\alpha_1]^2 + 2pq[x_{12} - (\alpha_1 + \alpha_2)]^2 + q^2[x_{22} - 2\alpha_2]^2} = V_D$$

$$+ \boxed{p^2[2\alpha_1 - \bar{x}]^2 + 2pq[(\alpha_1 + \alpha_2) - \bar{x}]^2 + q^2[2\alpha_2 - \bar{x}]^2} = V_A$$

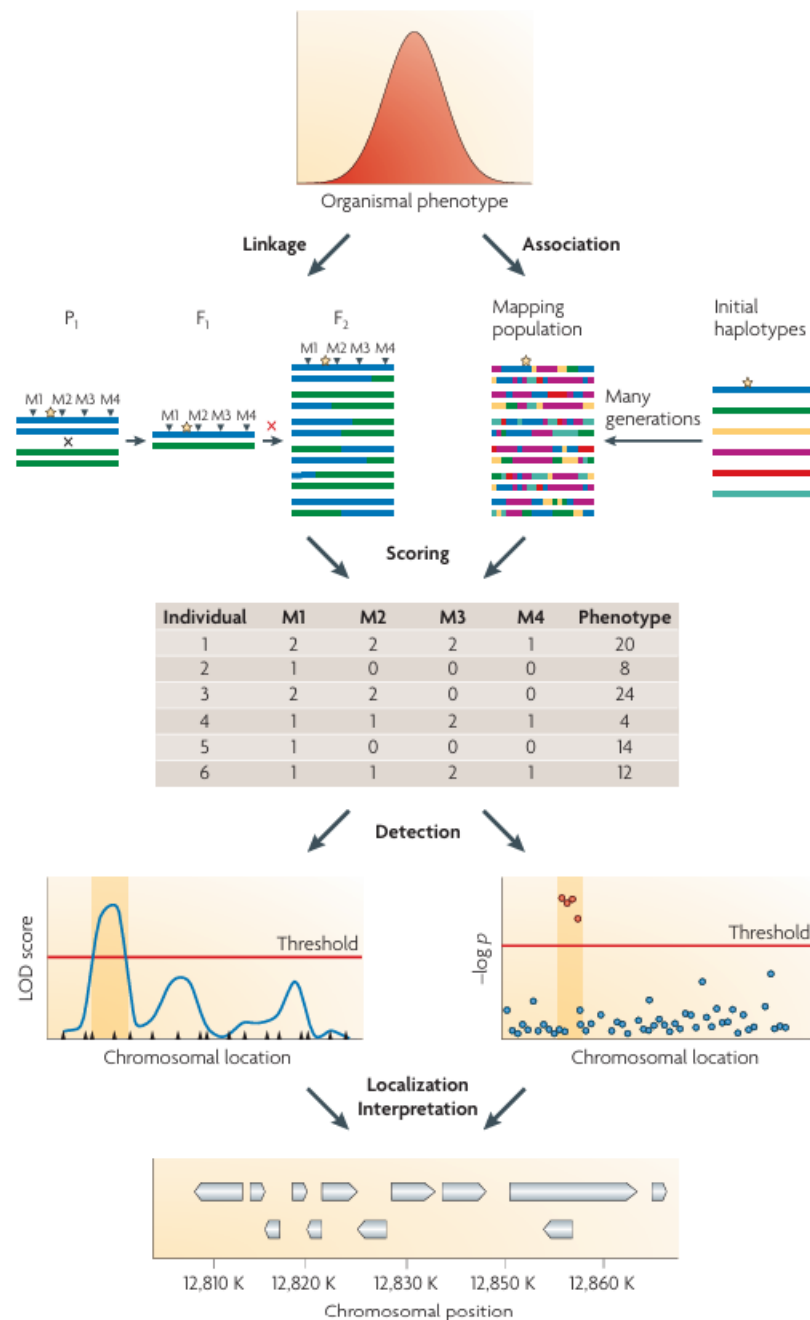
$$+ \boxed{p^2[2(x_{11} - 2\alpha_1)(2\alpha_1 - \bar{x})] + 2pq[2(x_{12} - \{\alpha_1 + \alpha_2\})(\{\alpha_1 + \alpha_2\} - \bar{x})] + q^2[2(x_{22} - 2\alpha_2)(2\alpha_2 - \bar{x})]} = 0$$

Genetica quantitativa

- Sommario:
- Galton e i tratti continui
- L'origine della genetica quantitativa
- Varianza del fenotipo e partizione della varianza
- Ereditabilità' sensu strictu, sensu amplo e sua stima
- Selezione naturale ed equazione dell'allevatore
- Dimostrazione della partizione della varianza genetica additive e di dominanza
- **Genetica quantitative molecolare: QTLs, GWAS e Polygenic Risk scores**

Studio dei tratti quantitative QTLs (quantitative trait locus)

- Dati una serie di marcatori genetici, i marcatori con più alta correlazione ad un tratto vengono identificati come QTLs
- Linkage disequilibrium viene utilizzato per la loro localizzazione cromosomica



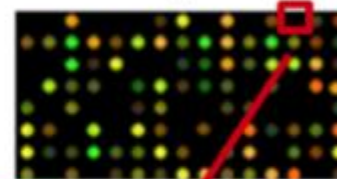
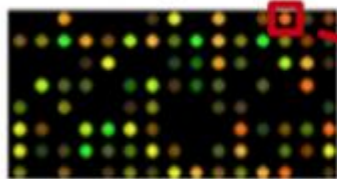
GWAS (Genome-wide association studies)

Casi

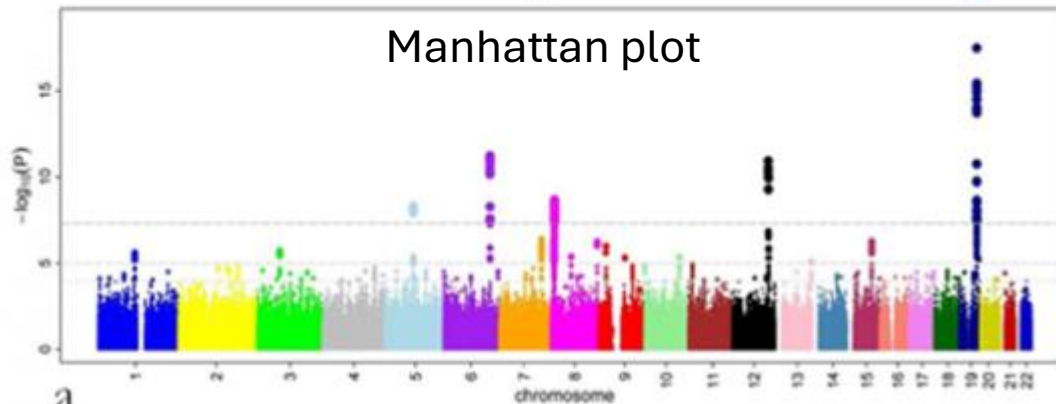
Controlli



Allele con frequenza molto piu alta nei casi che nei controlli



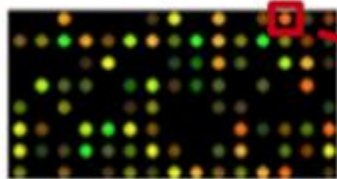
Modelli lineari (o lineari generalizzati) per testare l'associazione tra tratti e varianti genetiche



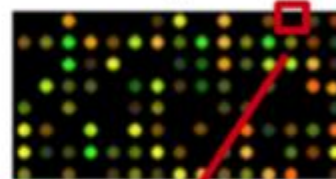
Forza dell'associazione (o p-value)

GWAS (Genome-wide association studies)

Casi

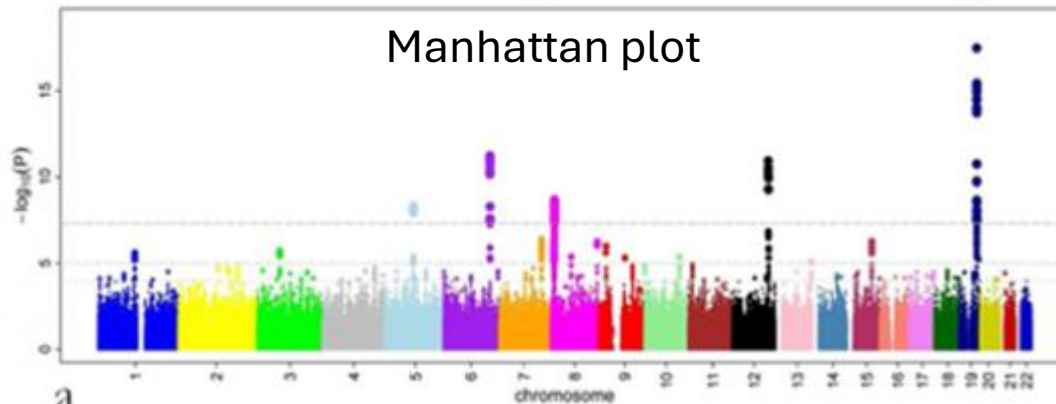


Controlli



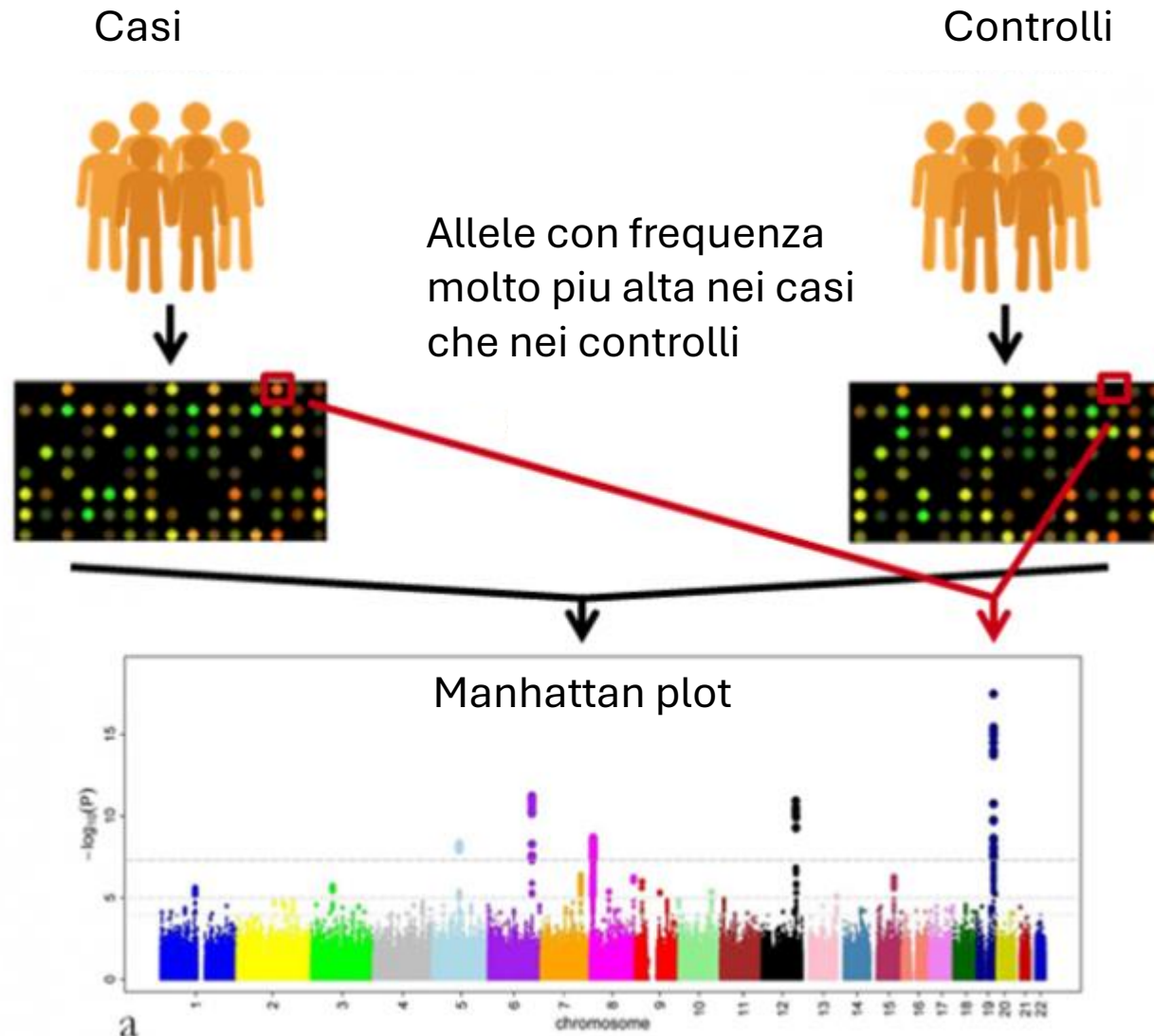
Allele con frequenza molto piu alta nei casi che nei controlli

Modelli lineari (o lineari generalizzati) per testare l'associazione tra tratti e varianti genetiche



Forza dell'associazione (o p-value)

GWAS (Genome-wide association studies)

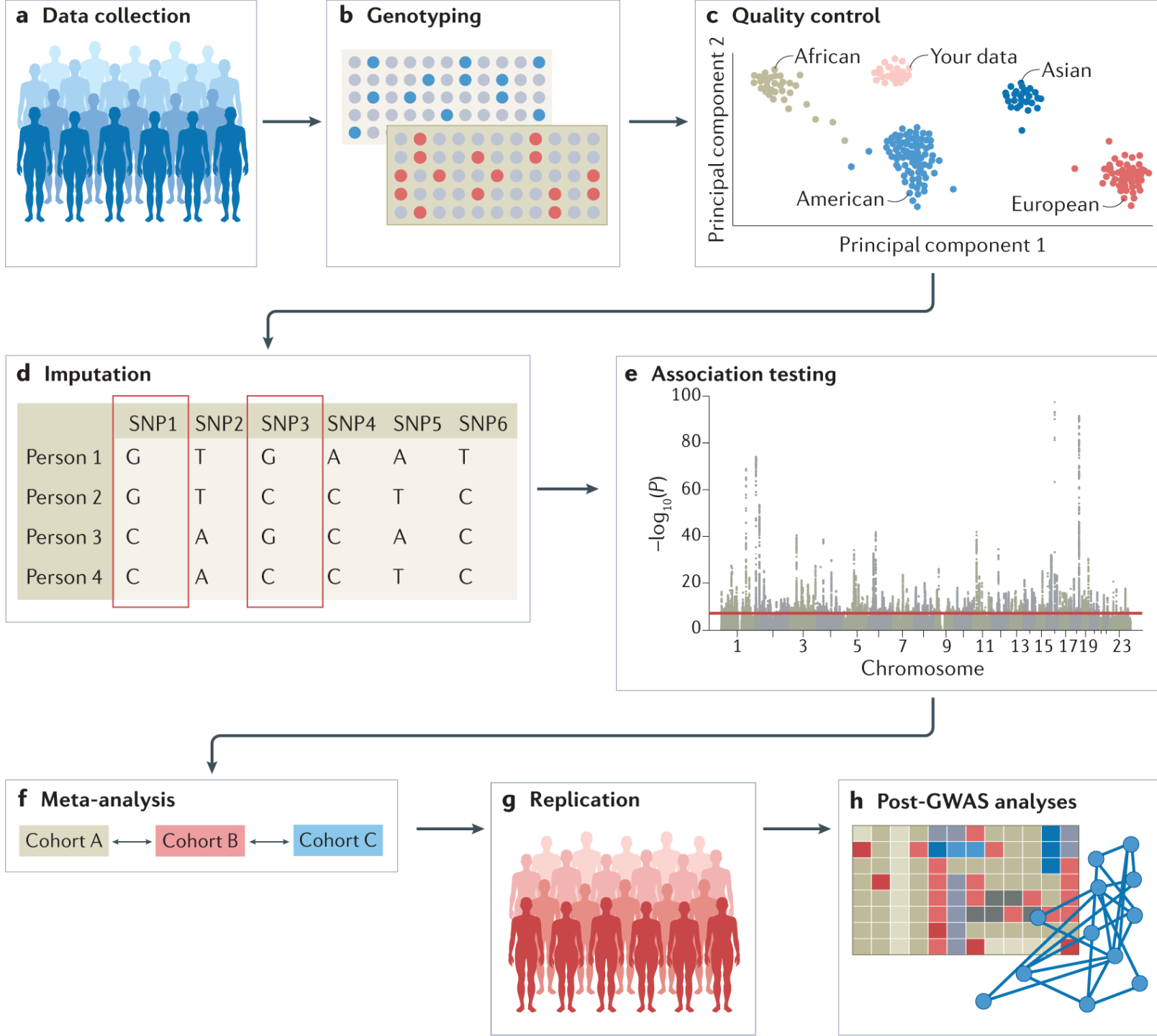


I modelli hanno la forma:

$$Y \sim W\alpha + X_s\beta_s + g + e$$

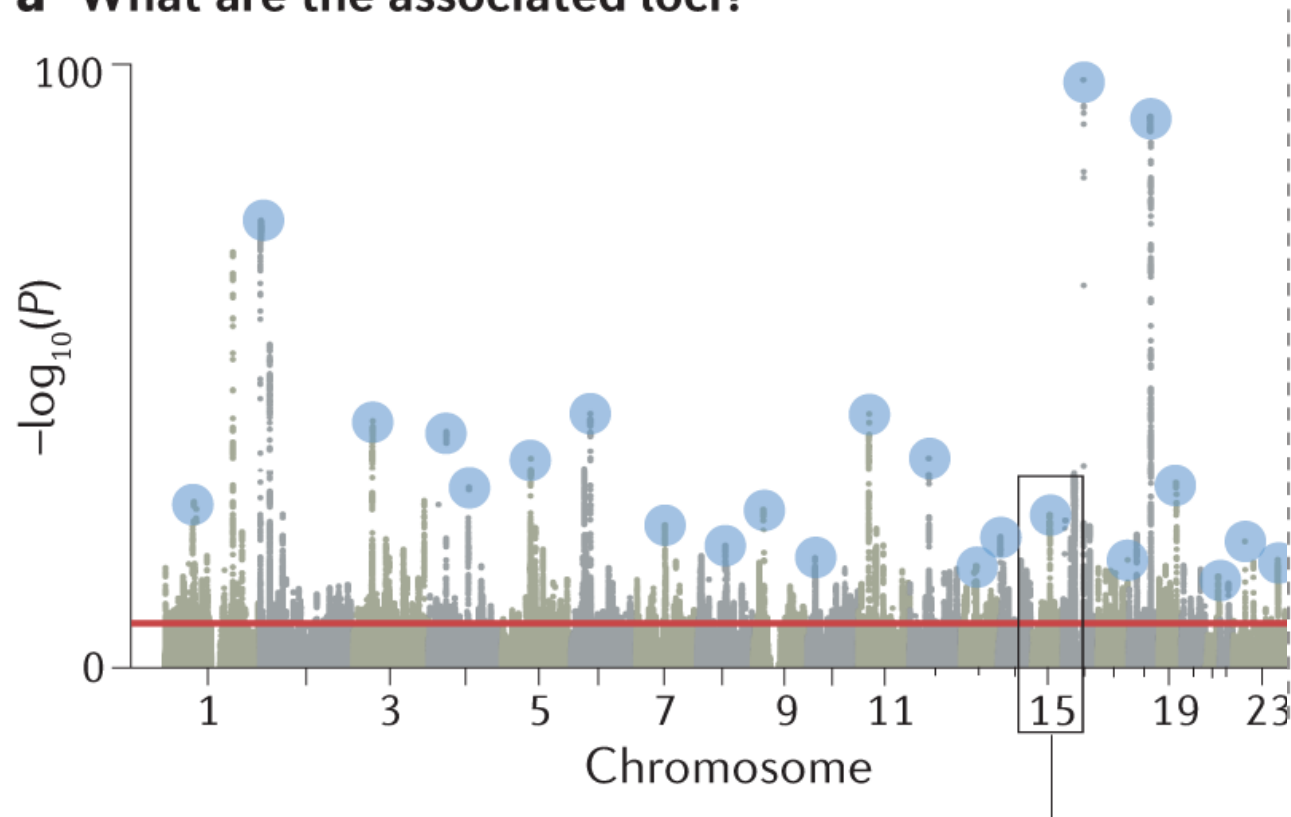
Valore del tratto fenotipico	Matrice di covariate e loro effetti	Effetti Poligenici (dipende da V_A)	Errore residuo (dipende da V_E)
------------------------------	-------------------------------------	--	------------------------------------

Effetto delle varianti s .
 X_s indica il "dosaggio" e β_s gli effetti (effect sizes) delle varianti

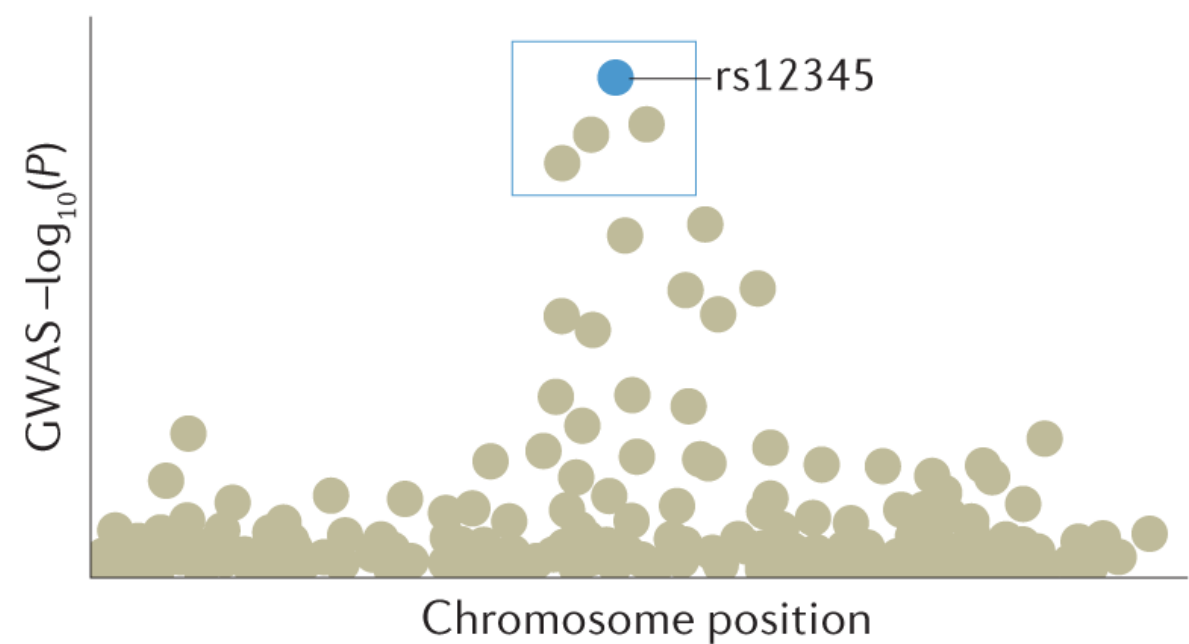


Non tutte le variant statisticamente associate sono causative del fenotipo

a What are the associated loci?

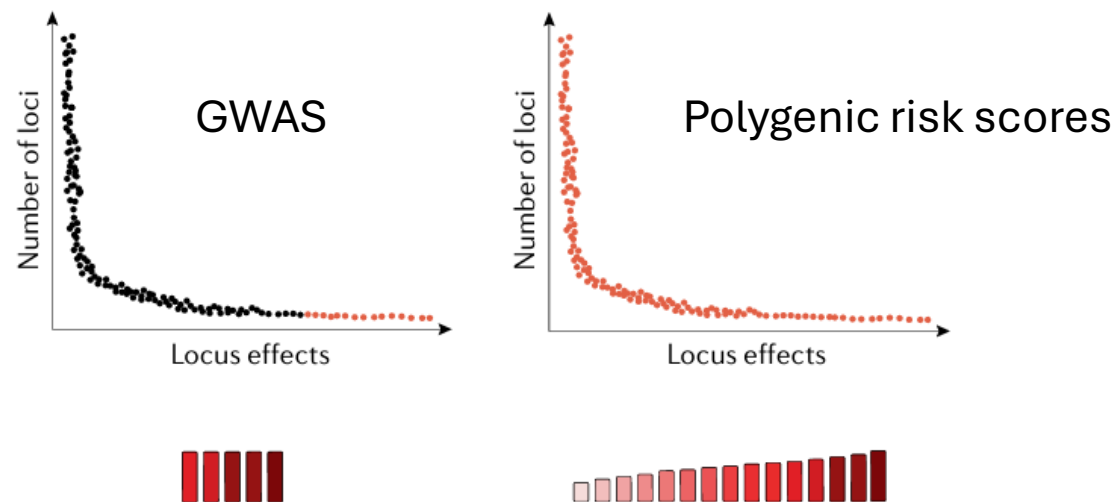


b What are the likely causal variants?



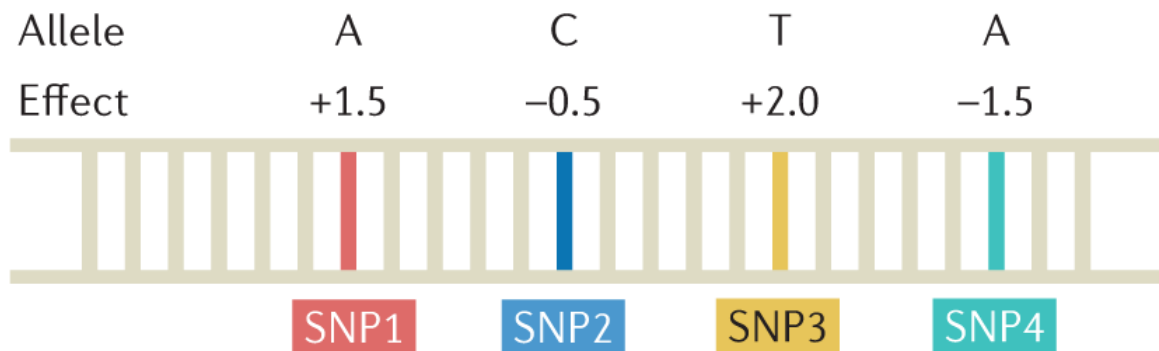
I Polygenic Risk Score

- I contributi dei singoli alleli dei GWAS viene composto in un indice per individuo
- *Polygenic risk score per l'individuo $j = \sum_i$ effetto dell'allele $i * \text{dosaggio di } i \text{ in } j$*
- Il dosaggio dell'allele i nell'individuo j non e' altro che il numero di copie in quell'individuo



Un semplice esempio di Polygenic Risk Score

① GWAS summary statistics



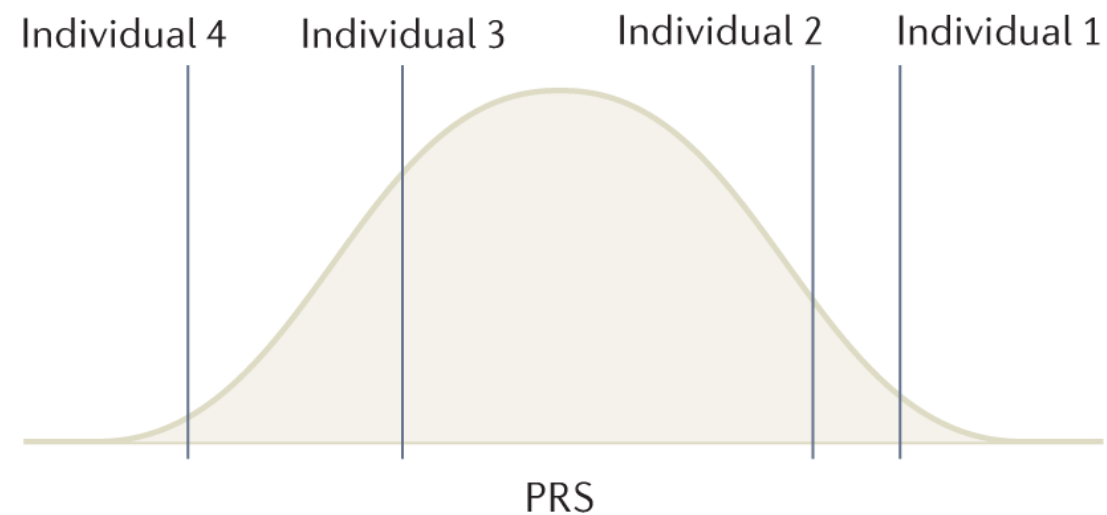
② Genotype data

	SNP1	SNP2	SNP3	SNP4
Individual 1	AT	CG	TT	CC
Individual 2	TA	GG	GT	CA
Individual 3	TT	CC	GT	CA
Individual 4	TT	CC	GG	AA

③ Polygenic risk score

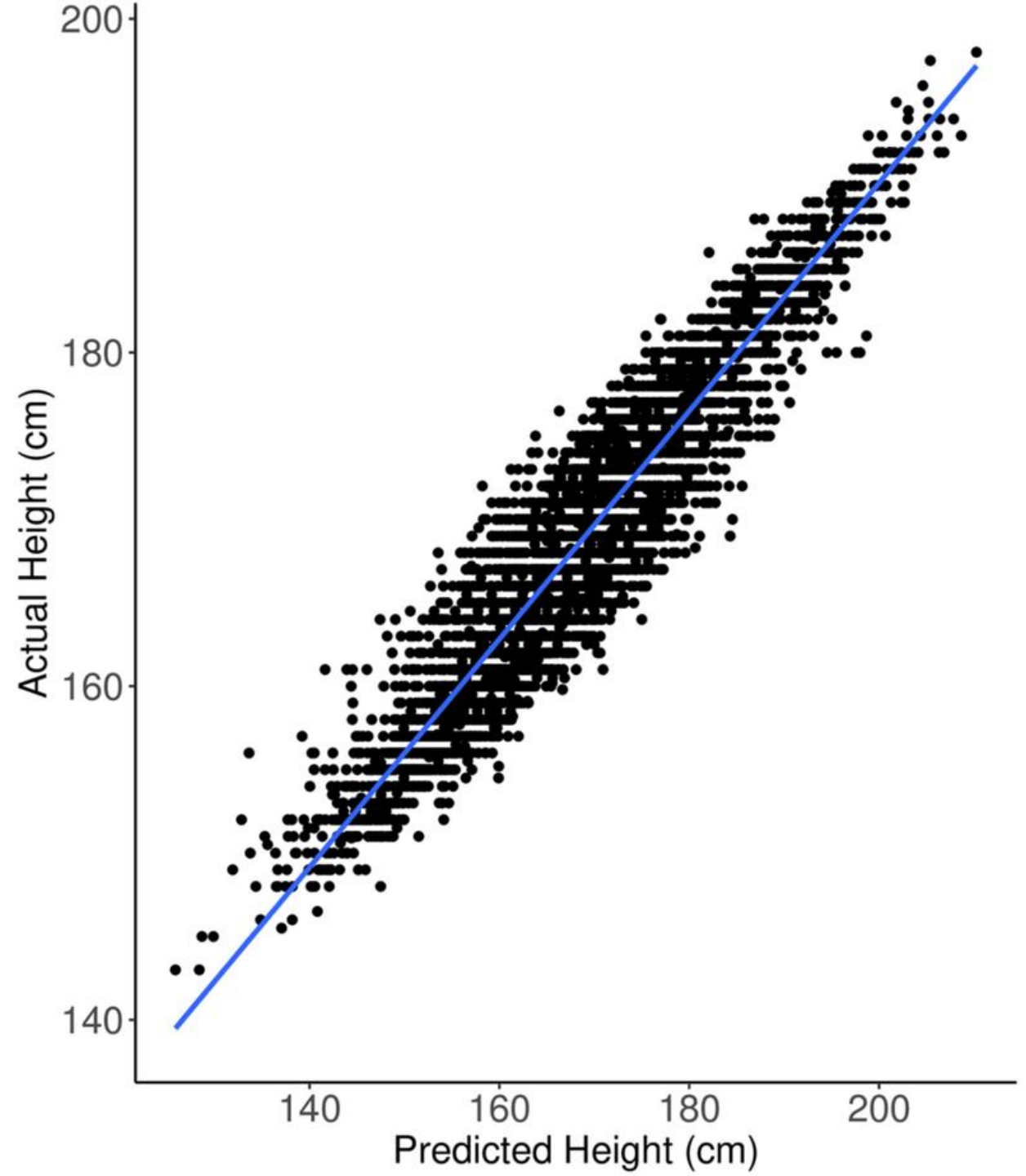
Individual 1	1.5	-	0.5	+	4.0	-	0.0	=	5.0
Individual 2	1.5	-	0.0	+	2.0	-	1.5	=	2.0
Individual 3	0.0	-	1.0	+	2.0	-	1.5	=	-0.5
Individual 4	0.0	-	1.0	+	0.0	-	3.0	=	-4.0

④ PRS distribution

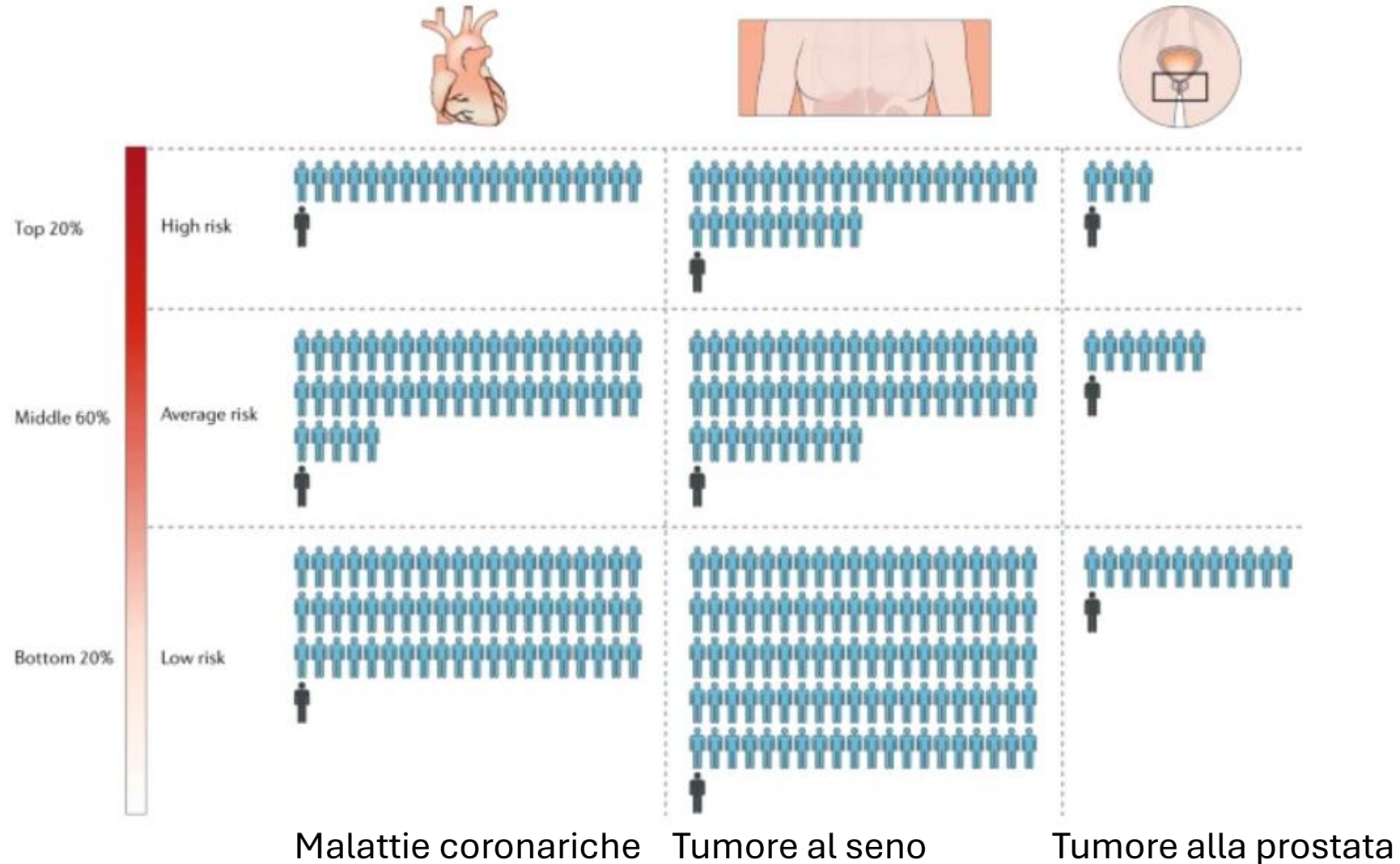


I Polygenic Risk Score

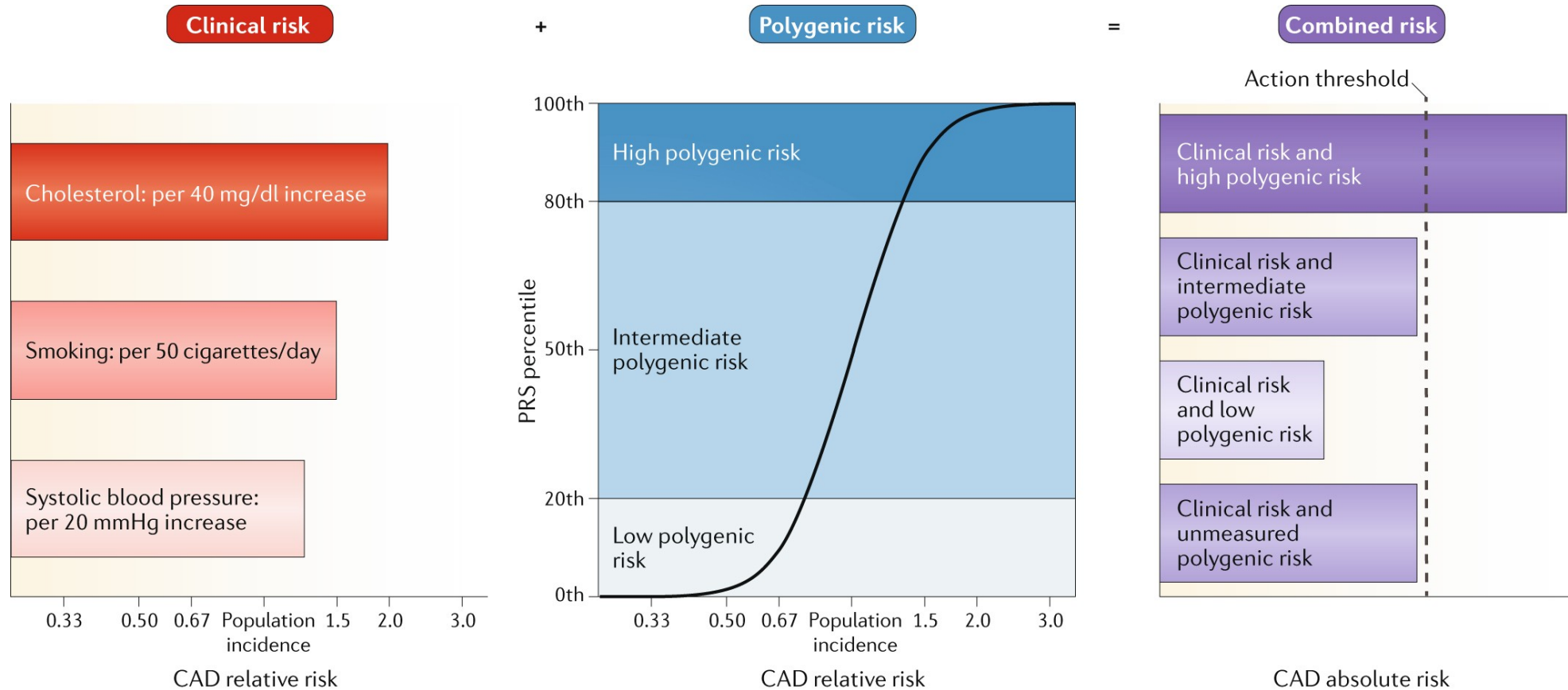
- Con campioni grandi permettono una predizione sostanzialmente accurata (anche se popolazione dipendente) di tratti quantitativi (ad alta ereditabilità, come ad esempio l'altezza, che ha $h^2 \approx 90\%$)



I PRS facilitano le diagnosi e la prevenzione



I PRS facilitano le diagnosi e la prevenzione



- Rischio per Coronary Artery Disease (CAD)

I PRS permettono la stratificazione della prevenzione



Risk stratification

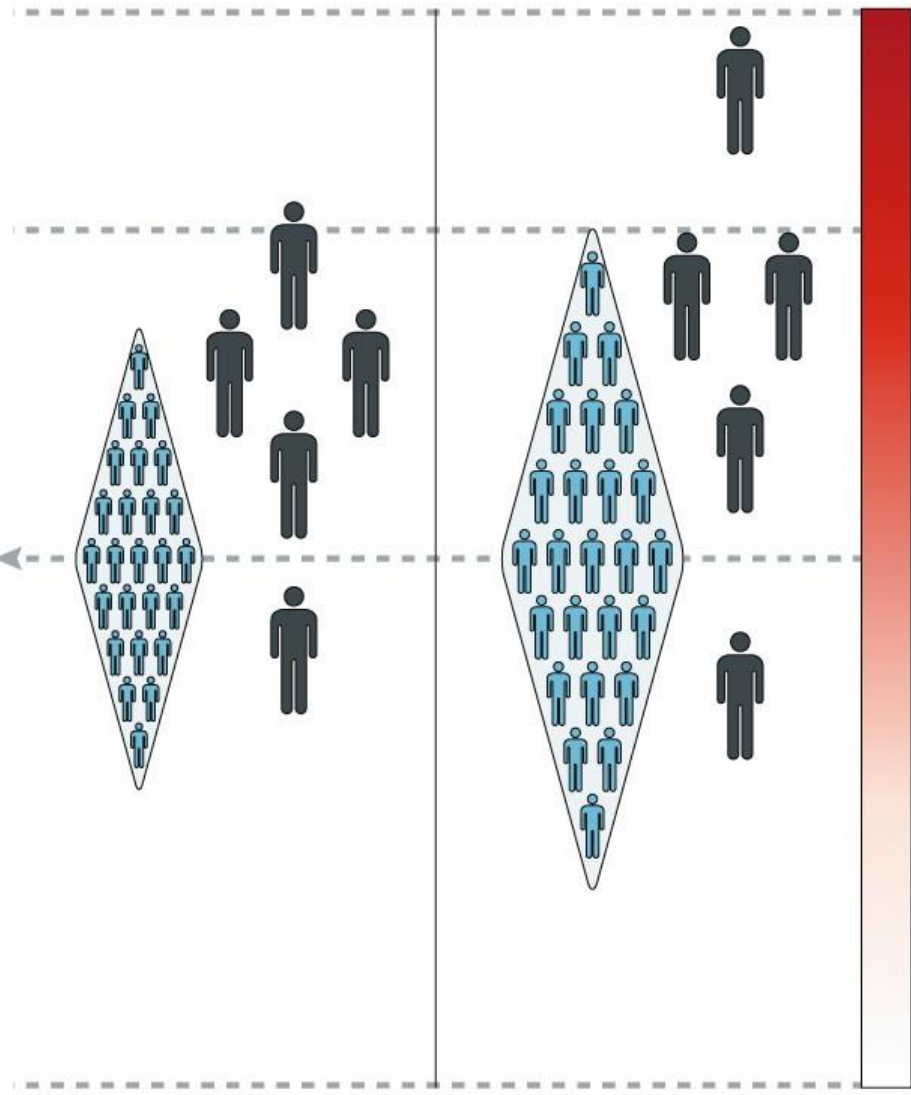
Model A

Model B

Risk prediction

Model A ≈ model B

 Diseased
 Healthy



Very high risk

High risk

Average risk

Low risk



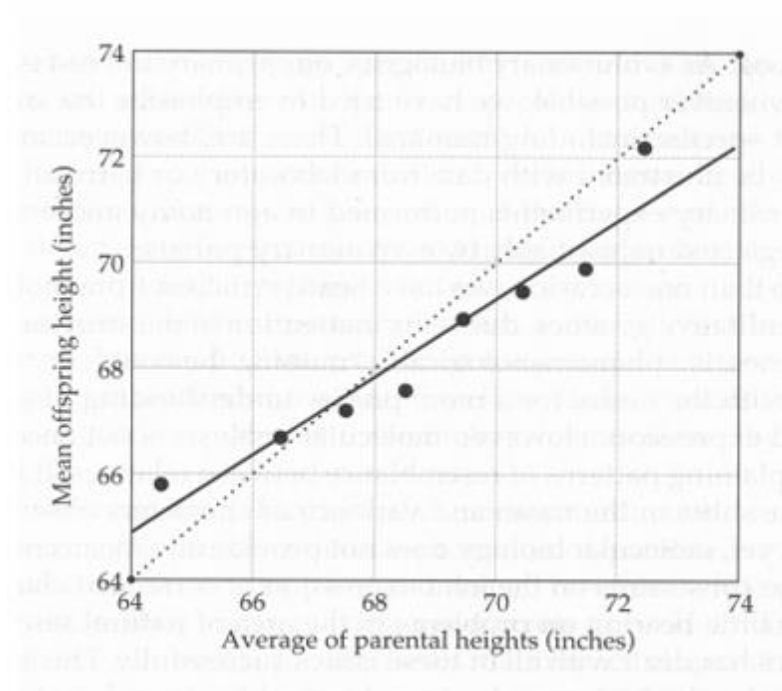
Invasive prevention

Accelerate noninvasive prevention

Standard prevention

Delay noninvasive prevention

Il dibattito tra i Mendeliani e i Biometrici



Dobbiamo Vedere la natura soprattutto come il colore del fiore dei piselli o come l'altezza?

I tratti quantitativi hanno le stesse proprietà evolutive dei caratteri discreti?



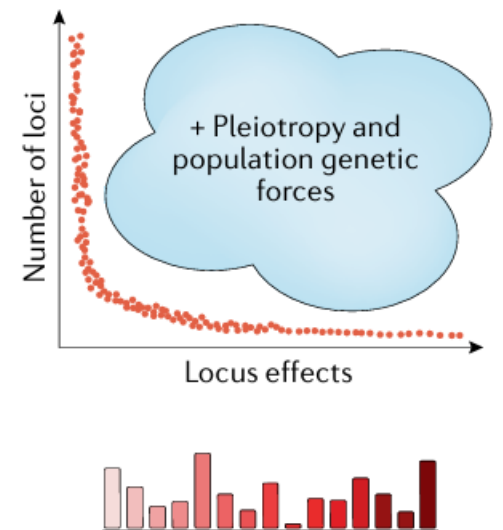
Secondo voi, quali prevalgono in natura?



La struttura tipica delle malattie continue e' di poche varianti importanti per i casi familiari contro effetti poligenici per il grosso della popolazione

....ma.....

- Portatori di *BRCA1* nel basso decile PRS per il tumore al seno hanno un rischio di 21% versus 39% di sviluppare il tumore prima dei 50 anni, e 56% contro 75% prima degli 80 anni.
- Rischio di cancro alla prostata prima degli 80 varia tra 7% e 26% per i portatori di *BRCA1* e tra 19% e 61% per i portatori di *BRCA2* nel 5th e nel 95th percentile dei PRS.



Genetica quantitativa

- Sommario:
- Galton e i tratti continui
- L'origine della genetica quantitativa
- Varianza del fenotipo e partizione della varianza
- Ereditabilità' sensu strictu, sensu amplo e sua stima
- Selezione naturale ed equazione dell'allevatore
- Dimostrazione della partizione della varianza genetica additive e di dominanza
- Genetica quantitative molecolare: QTLs, GWAS e Polygenic Risk scores