MTM

Why multivariate models?

- MT models are useful for traits where the difference between genetic and residual correlations are large (e.g. greater than 0.5 difference)
- Where one trait has a much higher heritability than the other trait.
- EBVs for traits with low heritability tend to gain more in accuracy than EBVs for traits with high heritability, although all traits benefit to some degree from the simultaneous analysis.

Why multivariate models?

- Another use of MT models is for traits that occur at different times in the life of the animal, such that culling of animals results in fewer observations on animals for traits that occur later in life compared to those at the younger ages.
- Consequently, animals which have observations later in life tend to have been selected based on their performance for earlier traits. Thus, analysis of later life traits by themselves could suffer from the effects of **culling bias**, and the resulting EBV could lead to errors in selecting future parents.

Multivariate models

- MT models do not offer great increases in accuracy for cases where
  - heritabilities of traits are similar in magnitude,
  - both genetic and residual correlations are relatively the same magnitude and sign,
  - every animal is measured for all traits.

MT

- However, if culling bias may exist, then an MT analysis should be performed even if the parameters are similar.
- An MT analysis relies on the accuracy of the genetic and residual correlations that are assumed.
- If the parameter estimates are greatly different from the underlying, unknown true values, then an MT analysis could do as much harm as it might do good.

Drawback of multivariate models

- Computer programs are more complicated, require more memory and disk storage for MT analyses.
- Verification of results might be more complicated.
- These have to be balanced against the benefits of an MT analysis.
- If culling bias is the main concern, then an MT model must be used regardless of the costs or no analysis should be done at all, except for the traits not affected by culling bias.

Some theory

Decomposition of genetic variances and covariances

$\sigma^2_{Gx} = \sigma^2_{Ax} + \sigma^2_{Dx} + \sigma^2_{AA\,x} + \sigma^2_{AD\,x} + \sigma^2_{DD\,x} + \ldots$

$\sigma^2_{Gy} = \sigma^2_{Ay} + \sigma^2_{Dy} + \sigma^2_{AAy} + \sigma^2_{ADy} + \sigma^2_{DDy} + \ldots$

$\sigma_{G(x,y)} = \sigma_{A(x,y)} + \sigma_{D(x,y)} + \sigma_{AA(x,y)} + \sigma_{AD(x,y)} + \sigma_{DD(x,y)} + \ldots$

we can decompose phenotypic correlation (r_p) to its genetic and environmental components.

$$r_p \; ? \; \frac{\upsilon_{Gxy} - \upsilon_{Exy}}{\sqrt{\upsilon^2_{Gx}\upsilon^2_{Gy}}\sqrt{\upsilon^2_{Ex}\upsilon^2_{Ey}}}$$

where $\sigma_G(x,y)$ is genetic covariance, $\sigma_E(x,y)$ is environmental covariance, σ^2 _subscripts are genetic or environmental variances for traits x and y.

Direct Response to Selection

- Genetic change due to selection for that trait

Correlated Response to Selection

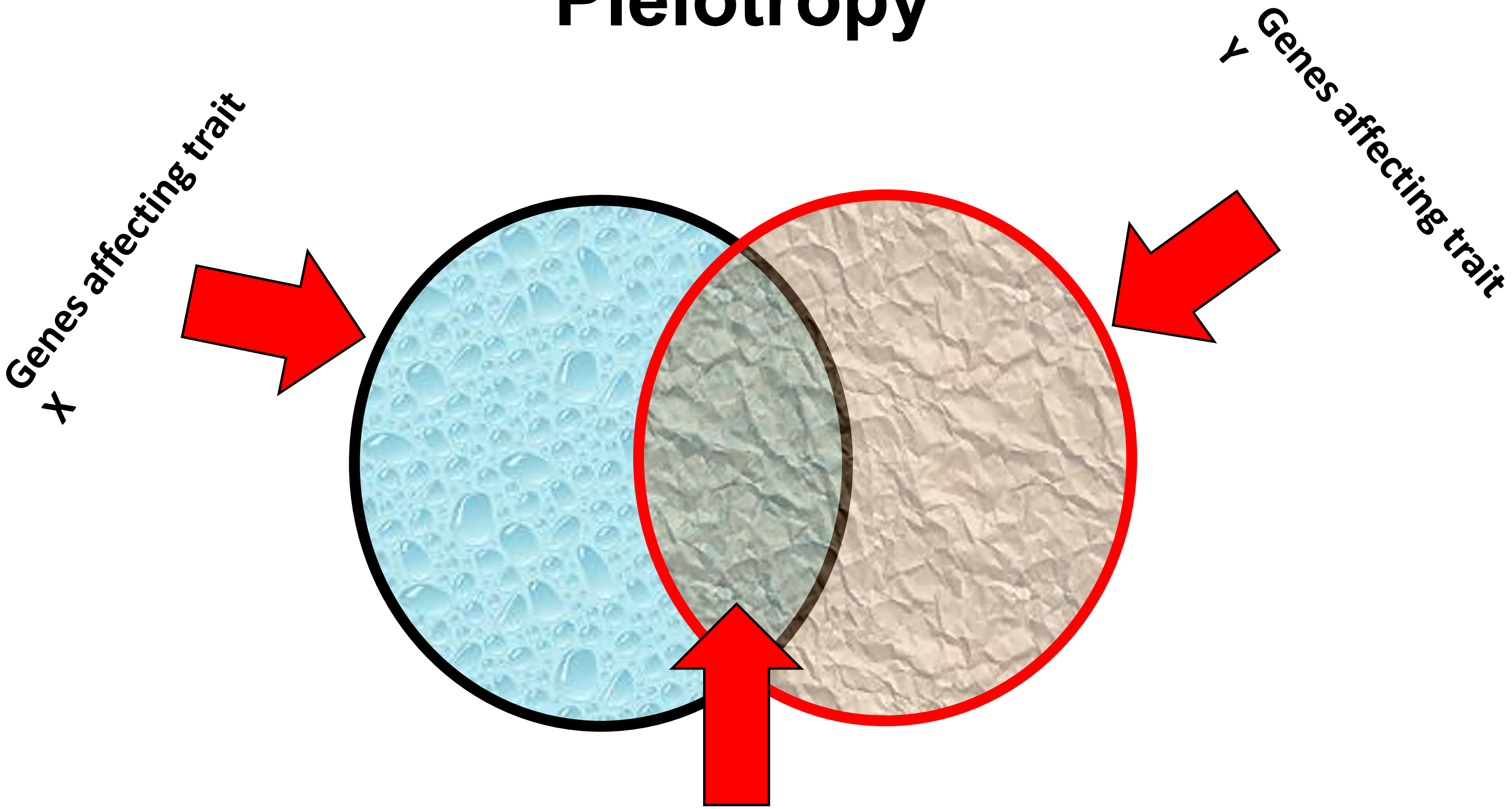
- Genetic change in one or more traits resulting from selection for another

Causes of Correlated Response to Selection

- Linkage
  - This is only temporary because of recombination and Mendelian segregation

- Pleiotropy
  - A single gene affecting more than one trait
  - This is the major cause of correlated response to selection

Pleiotropy



Genes affecting both traits

Genetic Correlations

$$\left(r_{BV_X,BV_Y}\right)$$

- A measure of the strength of the relationship between breeding values for one trait and breeding value for another trait
- A measure of pleiotropy

Phenotypic Correlations

$$\left(r_{P_X, P_Y}\right)$$

- A measure of the strength of the relationship between performance in one trait and performance in another trait

Environmental Correlations

$$\left(r_{E_X, E_Y}\right)$$

- A measure of the strength of the relationship between environmental effects on one trait and environmental effects on another trait

Correlated response to selection

How much of phenotypic response of **y** when we select x phenotype

$$\Delta BV_{Y|X} \Big/ t = \frac{r_{BV_X,BV_Y}\, r_{BV_X,\widehat{BV_Y}}\, i\sigma_{BV_Y}}{L}$$

Correlated response to selection

$$\Delta BV_{Y|X}/t = r_{BV_X,BV_Y}\left(\frac{\sigma_{BV_Y}}{\sigma_{BV_X}}\right)\left(\Delta BV_X/t\right)$$

$$\Delta BV_{Y|X}/t = \frac{r_{BV_X,BV_Y} h_X h_Y i_X \sigma_{P_Y}}{L}$$

$r_{xy}h_xh_y$ is called co-heritability of y and x

LMM for MT

ID	Height	Yield
1	87	0.52
2	84	0.48
3	75	0.45
4	90	0.69
5	79	0.74

The general form of multivariate model is

$$\mathbf{Y}_{nxd} = \mathbf{X}_{nxp+1}\mathbf{B}_{q+1xp} + \mathbf{E}_{nxd}$$

Where n is the number of rows (observations); d is the number of dependent variables; p is a set of predictors. The **X** design matrix is consisting of n x $p+1$ dimension, where p is the number of predictors and 1 is for the intercept. The **B** is the matrix of coefficients to be estimated with $q+1$ x p dimension. The rows of **B** are predictor variables and the columns are response variables.

LMM for MT

y_{11} - individual 1, trait 1
y_{12} - individual 1, trait 2
y_{21} - individual 2, trait 1
...

$$\mathbf{y} = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \\ y_{31} \\ y_{32} \\ y_{41} \\ y_{42} \\ y_{51} \\ y_{52} \end{bmatrix} = \begin{bmatrix} 87 \\ 0.52 \\ 84 \\ 0.48 \\ 75 \\ 0.45 \\ 90 \\ 0.69 \\ 79 \\ 0.74 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad \mathbf{u} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \quad \mathbf{e} = \begin{bmatrix} e_{11} \\ e_{12} \\ e_{211} \\ e_{22} \\ e_{31} \\ e_{32} \\ e_{41} \\ e_{42} \\ e_{51} \\ e_{52} \end{bmatrix}$$

For ease of calculations, the vectors in **Y**, **B** and **E** are organized in a different way (vectors instead of matrices).

The LMM for multivariate models

The linear mixed model in compact matrix form is as follows.

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e}$$

It looks the same as analyzing one trait at time. However design matrices are different. The same model can be written in a different way to show design matrices associated with effects

$$y_n = (I_2 \otimes X)\, b + (I_2 \otimes \mathbf{Z_f})\, \mathbf{u_f} + \mathbf{e^*}$$

Where ***y*** is the vector of traits with *n* rows,
e*is residuals of two traits,
I_2 is the identity matrix with n x 2 dimensions,
n is the number of pairs of observations for the traits,
b is the vector of fixed effects, such as intercepts for two traits,
Z_f is the design matrix for random effect, and
u is vector of random effects.

LMM for MT

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}$$

$$\mathbf{G} = \begin{bmatrix} \sigma^2_{\mathrm{1a}} & \sigma_{a21} \\ \sigma_{a12} & \sigma^2_{\mathrm{2a}} \end{bmatrix} \otimes \mathbf{I} \qquad \mathbf{R} = \mathbf{I} \otimes \begin{bmatrix} \sigma^2_{\mathrm{1e}} & \sigma_{e21} \\ \sigma_{e12} & \sigma^2_{\mathrm{2e}} \end{bmatrix}$$

In the **G** matrix, variances of two traits are given in the diagonal ($\sigma^2{}_{1a}$ and $\sigma^2{}_{2a}$ and the covariances or the correlations are given in the off-diagonals (σ_{12a}). **I** is the identity matrix.

If genetic effects are related (have common parents or grandparents), then we can substitute the identity matrix **I** with the numerator relationships matrix **A**.

Similarly, the diagonal elements of **R** are the residual variances of two traits; the off-diagonal elements are either covariances or correlations.

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \text{ and } \operatorname{var}(u) = G = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \qquad G_0 = \begin{pmatrix} \sigma_{g11}^2 & \sigma_{g12} \\ \sigma_{g21} & \sigma_{g22}^2 \end{pmatrix} \qquad G^{-1} = \begin{pmatrix} G^{11} & G^{12} \\ G^{21} & G^{22} \end{pmatrix} \text{ where } G^{ij} = g^{ij}A^{-1},$$

$$R^{-1} = \begin{bmatrix} Ir^{11} & Ir^{12} \\ Ir^{21} & Ir^{22} \end{bmatrix} \qquad \mathbf{R} = \begin{bmatrix} \sigma(e_1, e_1) & \sigma(e_1, e_2) \\ \sigma(e_2, e_1) & \sigma(e_2, e_2) \end{bmatrix} = \begin{bmatrix} I r_{11} & Ir_{12} \\ I r_{21} & I r_{22} \end{bmatrix}$$

Precision of correlations

Estimation of genetic correlations requires many individuals each with a large number of progenies. The standard error of correlation should be used with caution, because we do not know its distribution. Although with the increased power of computing, resampling methods have become available to obtain approximate distributions. One of two formulas given below can be used to calculate an approximate standard error of genetic correlation.

1) An approximate standard error of genetic correlation (Falconer and Mackay 1996):

$$\mathrm{SE}(r_G)\ ?\ \frac{1/\ r_G^2}{\sqrt{2}}\sqrt{\frac{\mathrm{SE}(h_x^2)\mathrm{SE}(h_y^2)}{h_x^2 h_y^2}}$$

Where, $\mathrm{SE}(h^2_x)$ and $\mathrm{SE}(h^2_x)$ are standard errors of narrow-sense heritabilities of trait x and y, respectively. The h^2_x and h^2_x are narrow sense heritabilities of two traits.

Precision of correlations

2) Delta method is considered a better way to estimate variance of ratios for unknown distributions (Lynch and Walsh 1998). See Appendix 1 in Lynch and Walsh for more details about the Delta method.

$$\mathrm{var}(r)\ ?\ r^2 \left\{ \frac{\mathrm{var}(\sigma_x^2)}{4\sigma_x^{2^2}} - \frac{\mathrm{var}(\sigma_y^2)}{4\sigma_y^{2^2}} - \frac{\mathrm{var}(\sigma_{xy})}{\sigma_{xy}^2} - \frac{2\mathrm{cov}(\sigma_x^2)\sigma_y^2}{4\sigma_x^2\sigma_y^2} \,/\, \frac{2\mathrm{cov}(\sigma_x^2)\sigma_{xy}}{2\sigma_x^2\sigma_{xy}} \,/\, \frac{2\mathrm{cov}(\sigma_{xy})\sigma_y^2}{2\sigma_{xy}\sigma_y^2} \right\}$$

var(σ^2_x) = Variance of variance for trait x,
σ^2_x = Variance of trait x,
var(σ^2_y) = Variance of variance for trait y,
σ^2_y = Variance of trait y,
var(σ_{xy}) = Variance of covariance of traits x and y,
σ^2_{xy} = Squared covariance of traits x and y,
cov(σ^2_x)σ^2_y = Covariance between the variances of x and y
cov(σ^2_x)σ_{xy} = Covariance between the variance of x and the covariance of xy
σ_{xy} = Covariance between x and y
cov(σ_{xy})σ^2_y = Covariance between the covariance of xy and the variance of b

Implementation of MT analysis

- Good starting values will save a lot of trouble!
- Analyze traits independently first to obtain variances for each. You may use these variances as starting values for multivariate analysis.
- Multivariate models may not converge for small data sets. You need a lot of genetic entries each with a lot of progeny. Even if the model converges the reliability of correlations would be questionable.

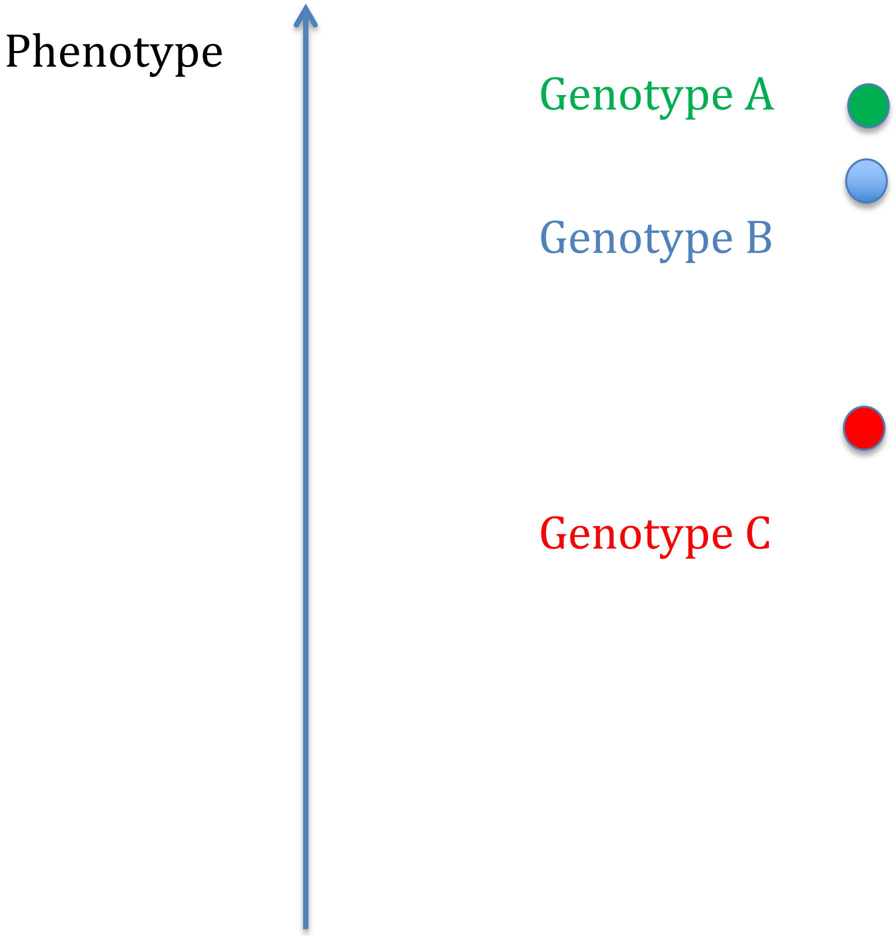
NC STATE UNIVERSITY

Modeling genotype by environment interactions

F. Tiezzi

Department of Animal Science
NCSU

The problem

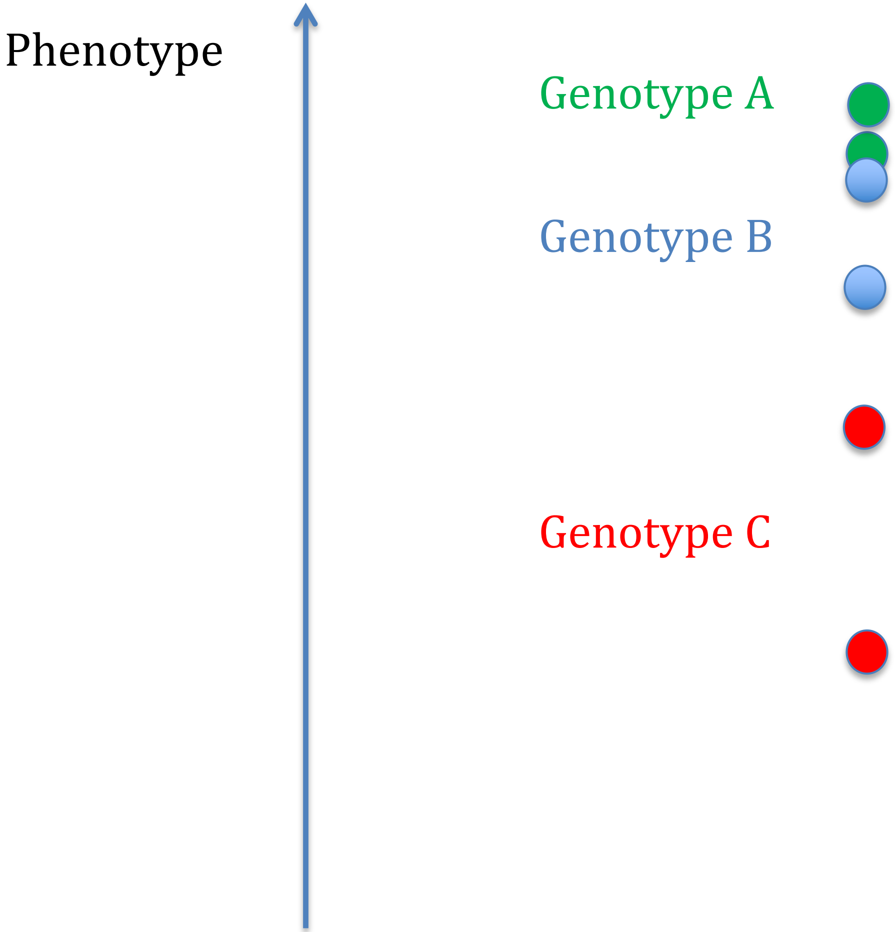


Phenotype

Genotype A

Genotype B

Genotype C

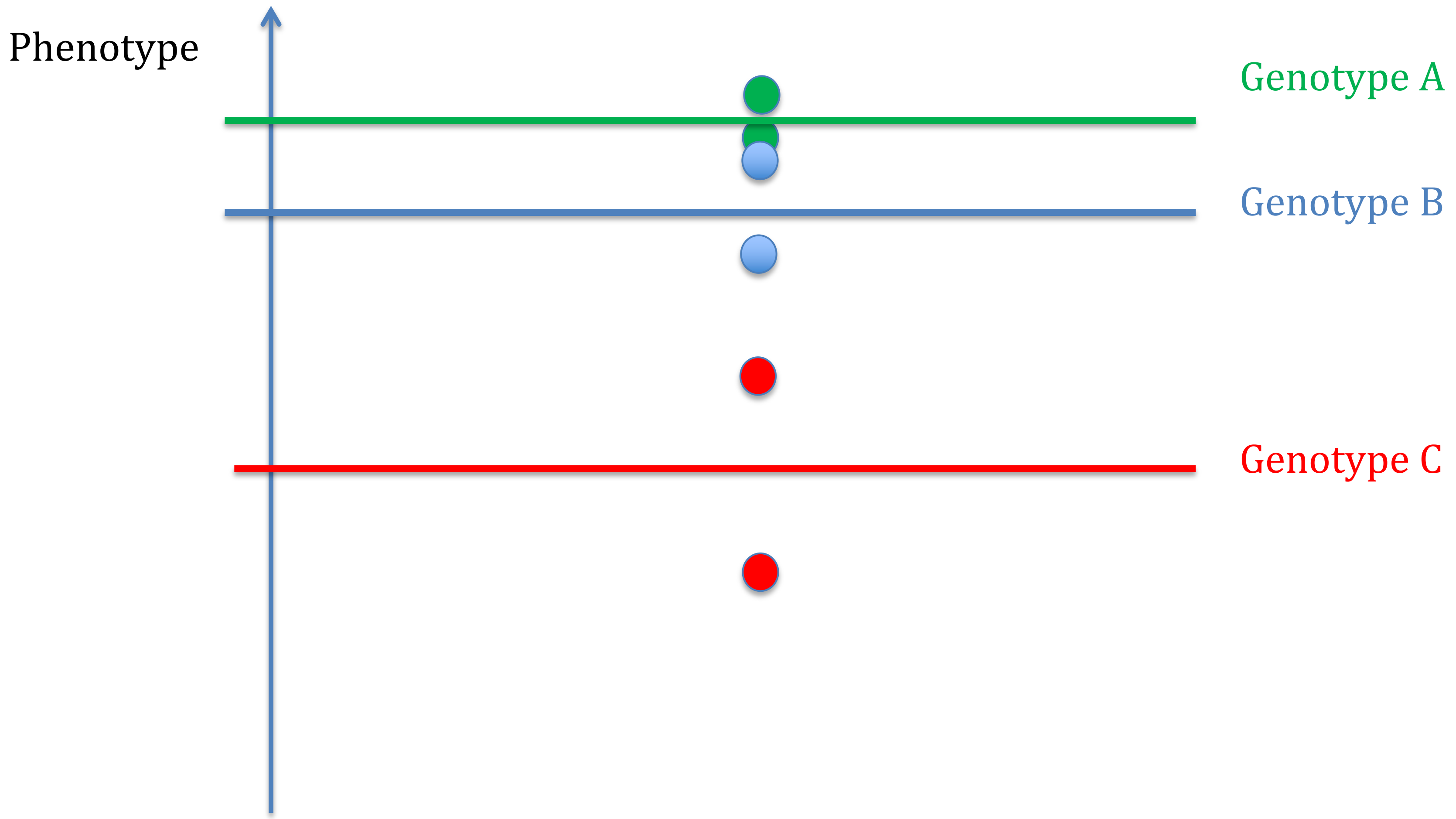


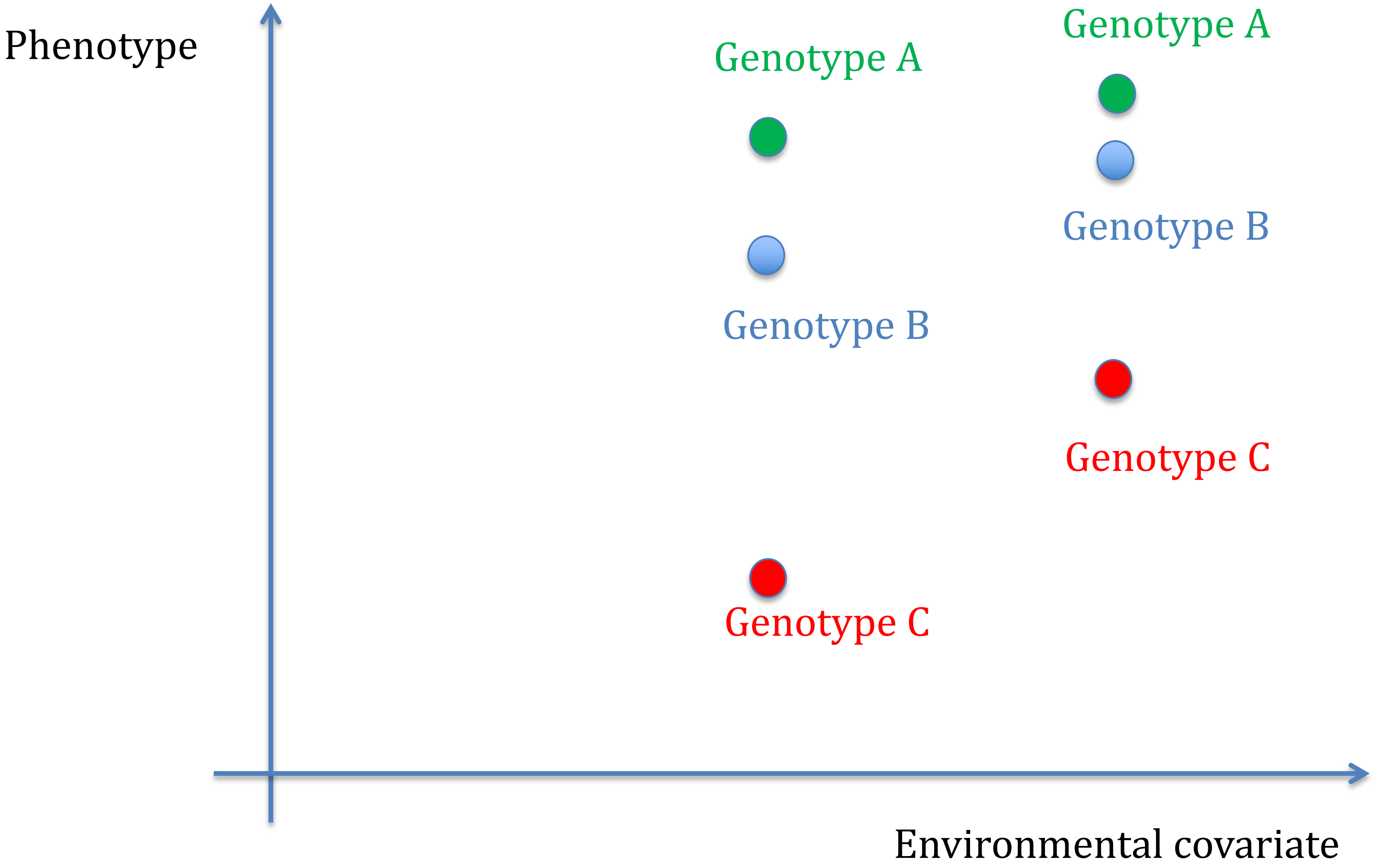
Phenotype

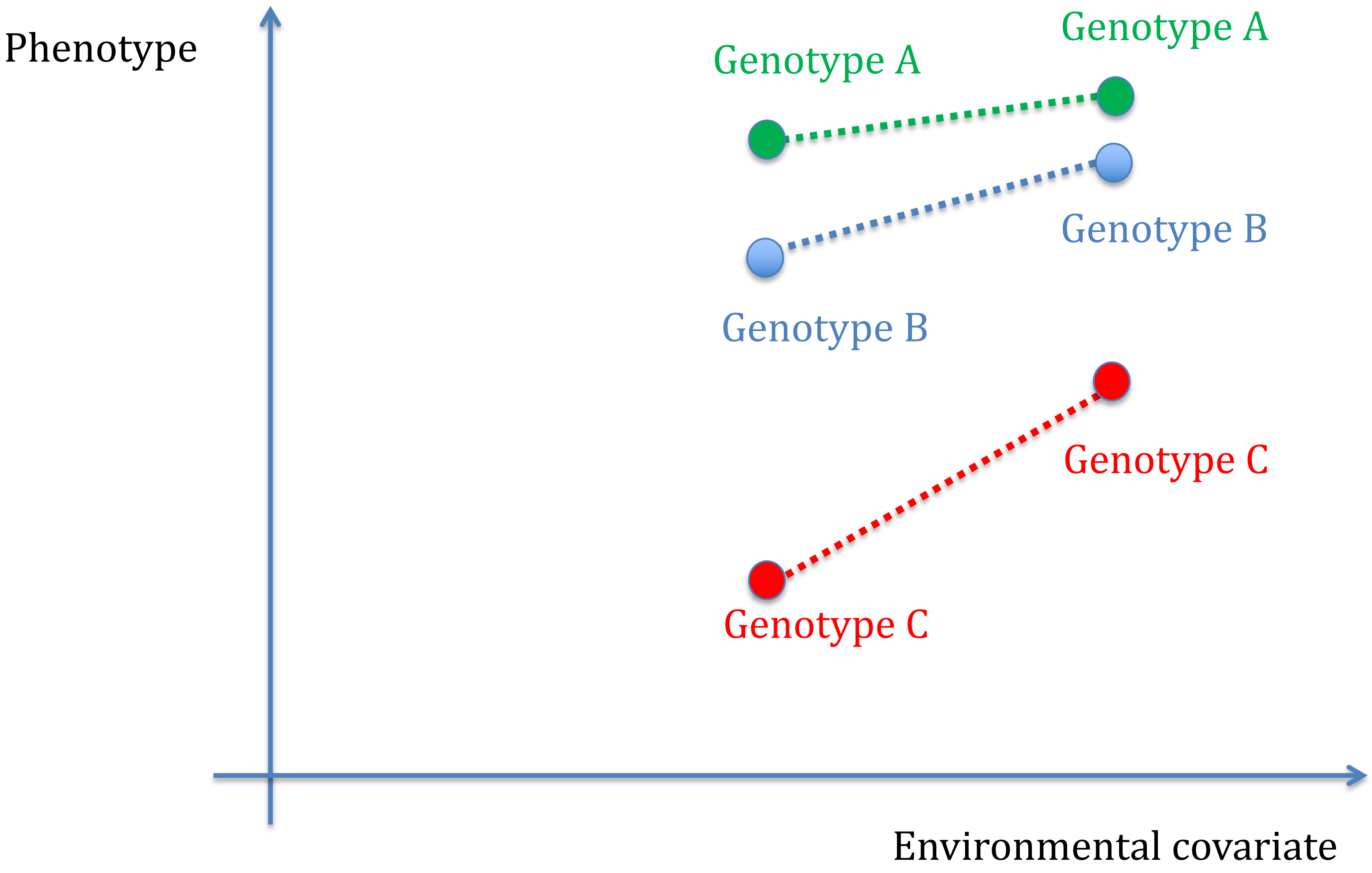
Genotype A

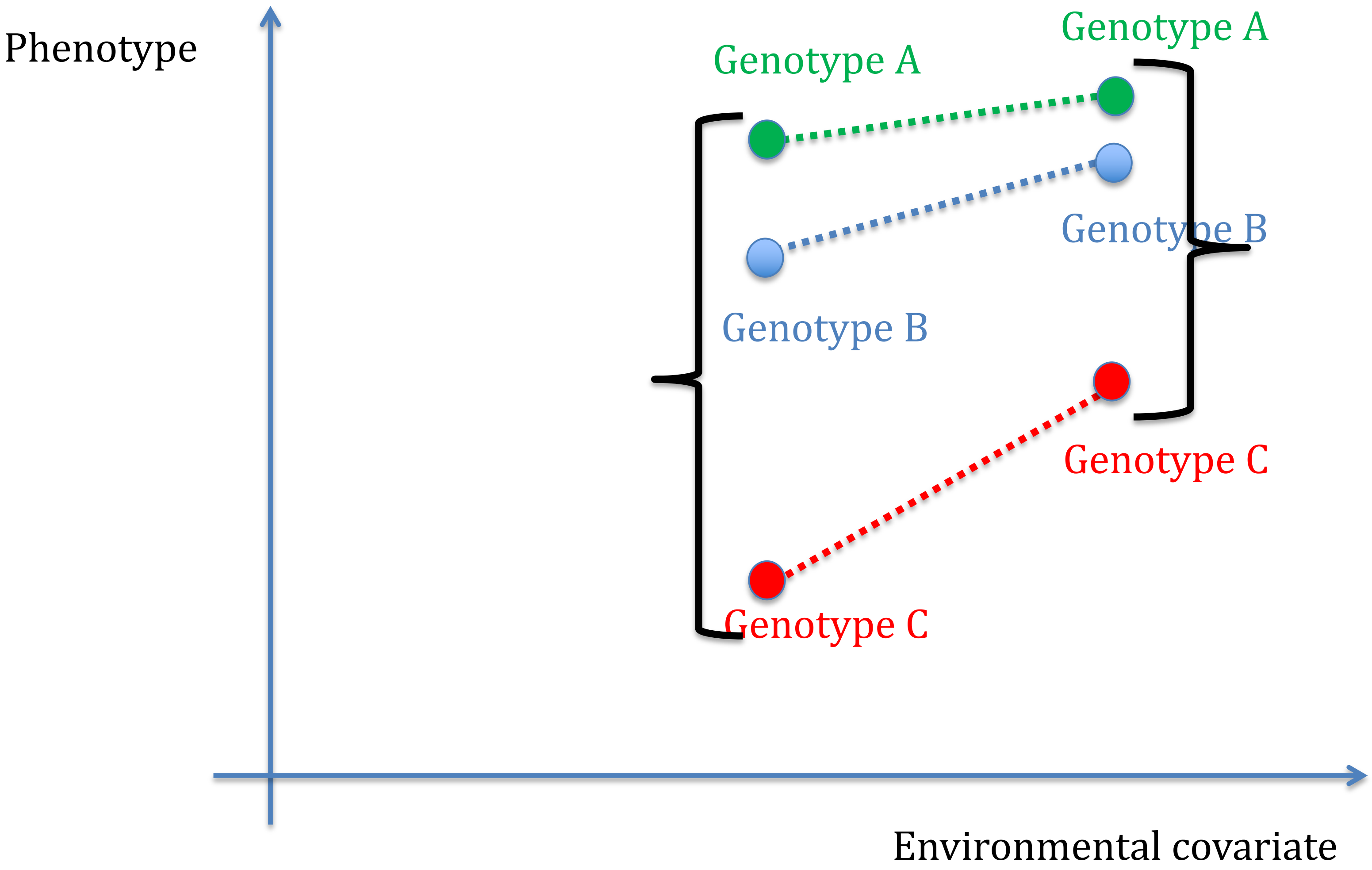
Genotype B

Genotype C

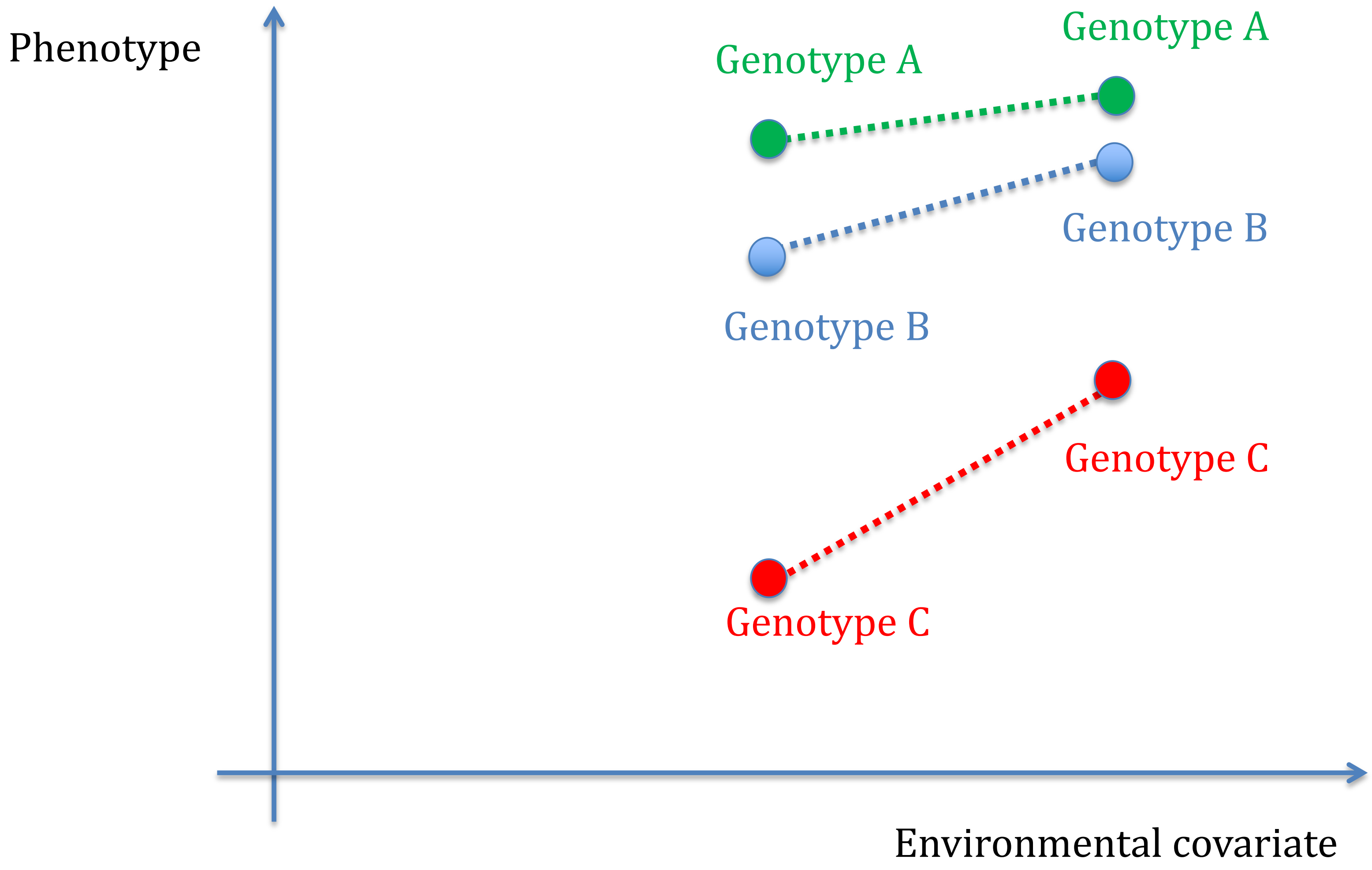

Phenotype
Genotype A
Genotype B
Genotype C

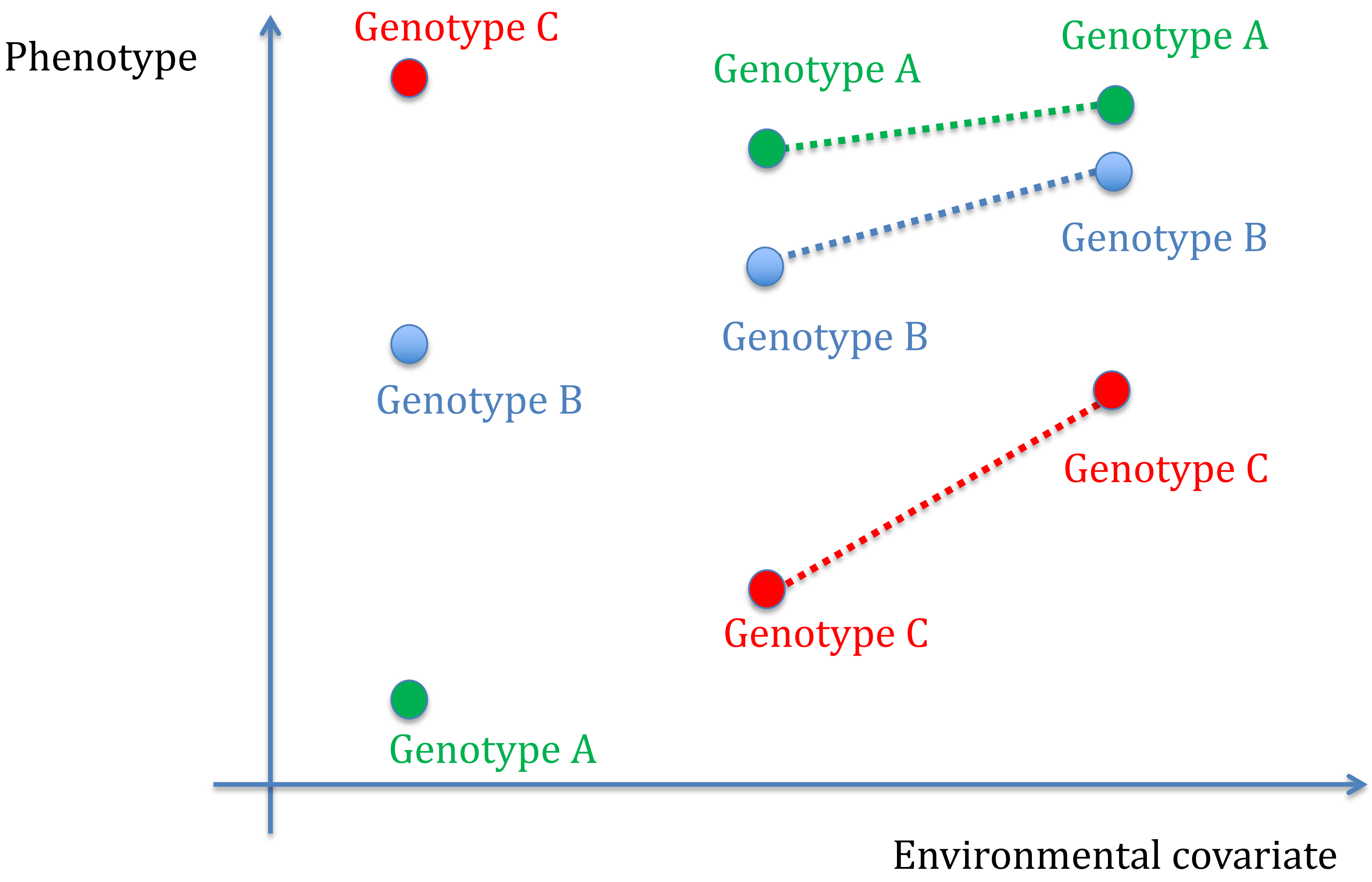

Phenotype
Genotype A
Genotype A
Genotype B
Genotype B
Genotype C
Genotype C
Environmental covariate

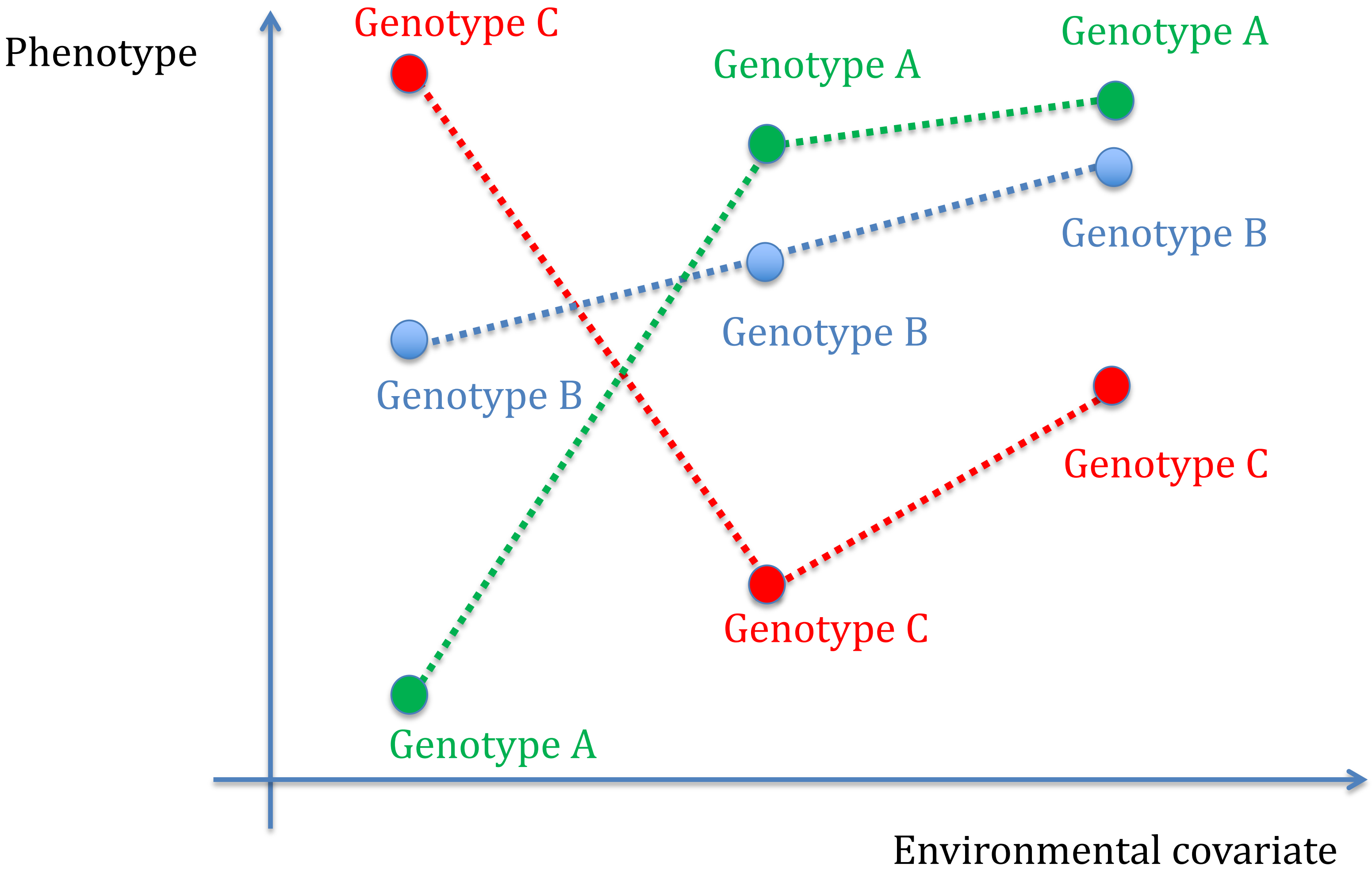

Phenotype
Genotype A
Genotype A
Genotype B
Genotype B
Genotype C
Genotype C
Environmental covariate



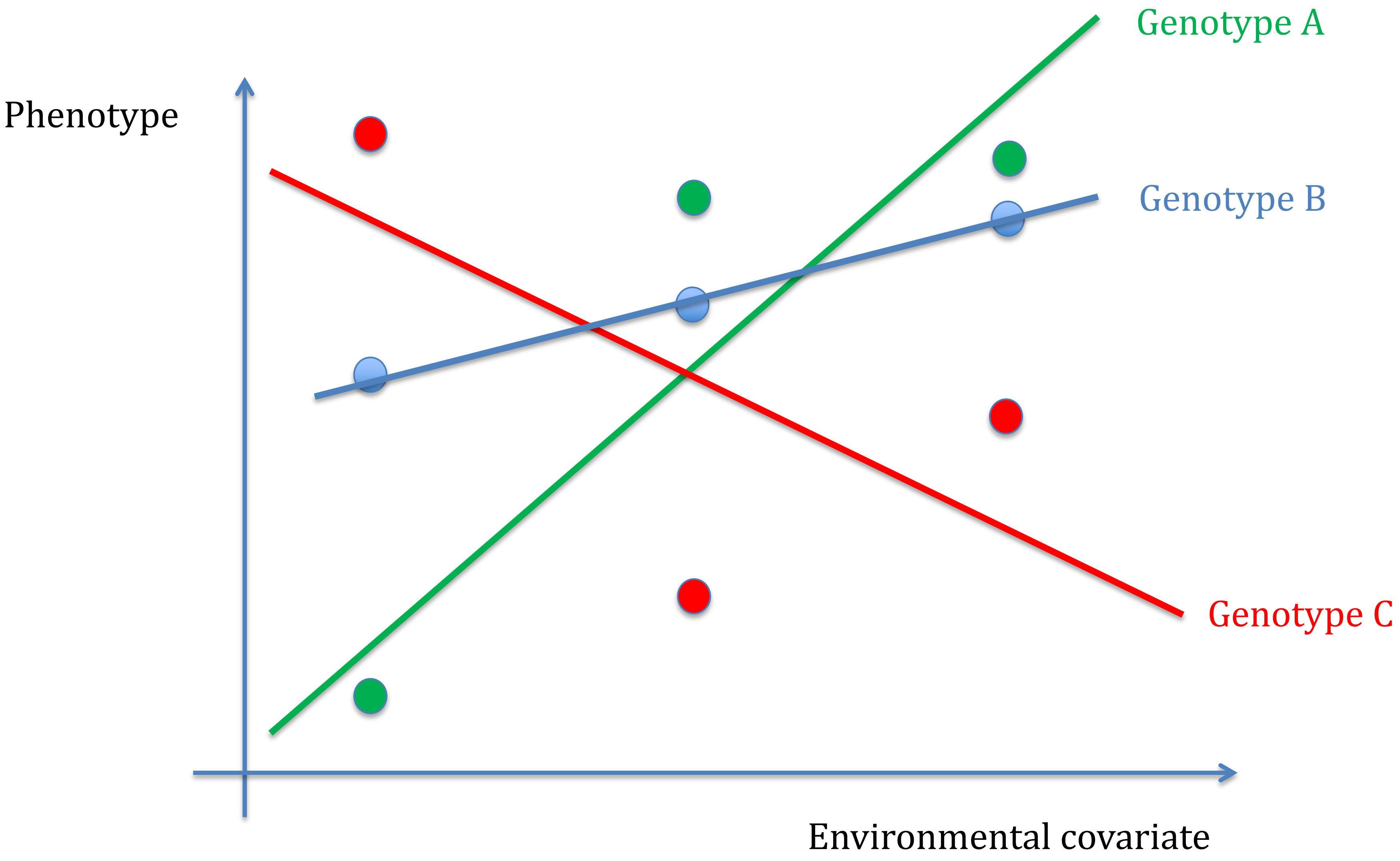
- Just different (environmental) plasticity?


Phenotype
Genotype A
Genotype A
Genotype B
Genotype B
Genotype C
Genotype C
Environmental covariate

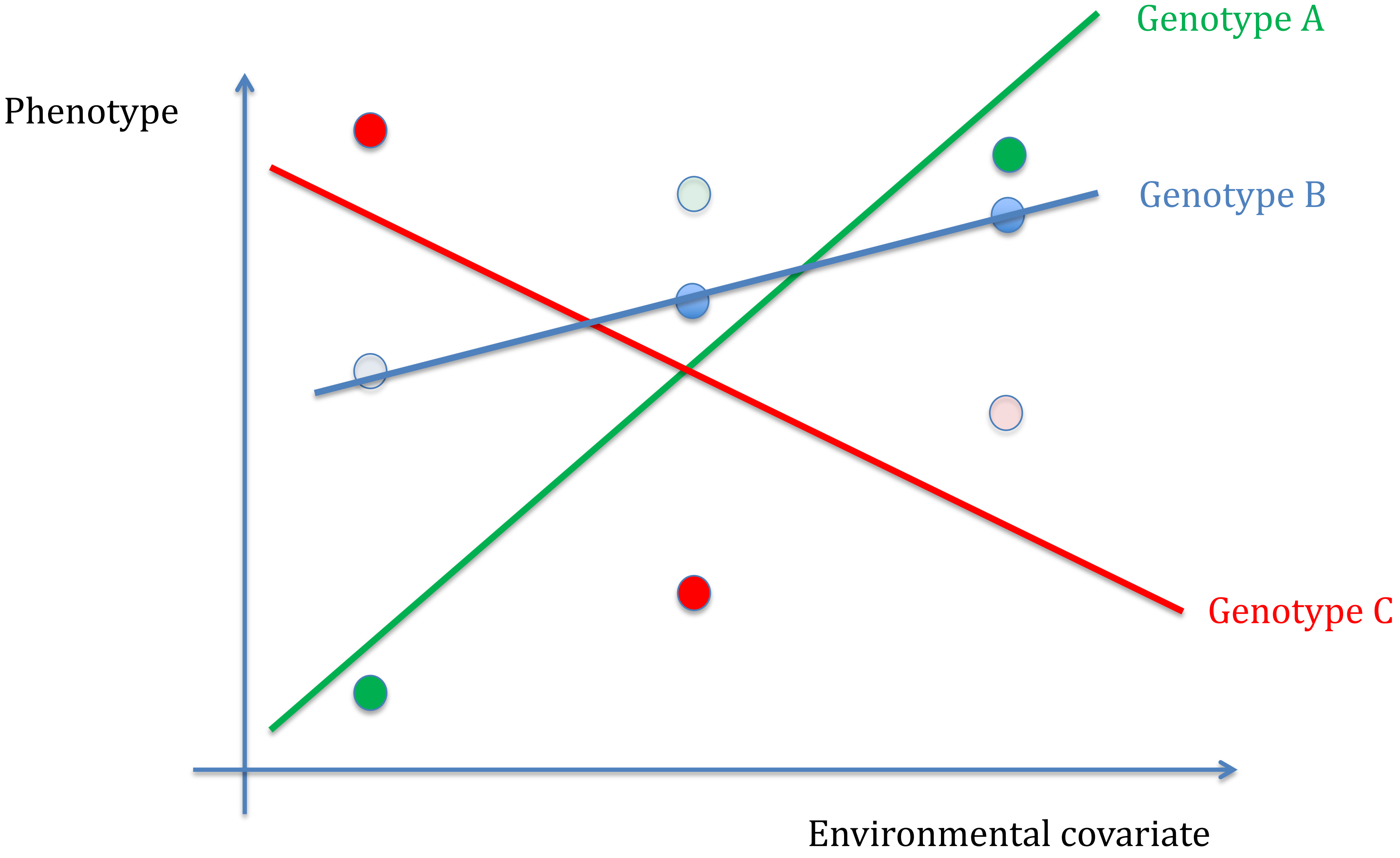

Phenotype
Genotype C
Genotype A
Genotype A
Genotype B
Genotype B
Genotype B
Genotype C
Genotype C
Genotype A
Environmental covariate



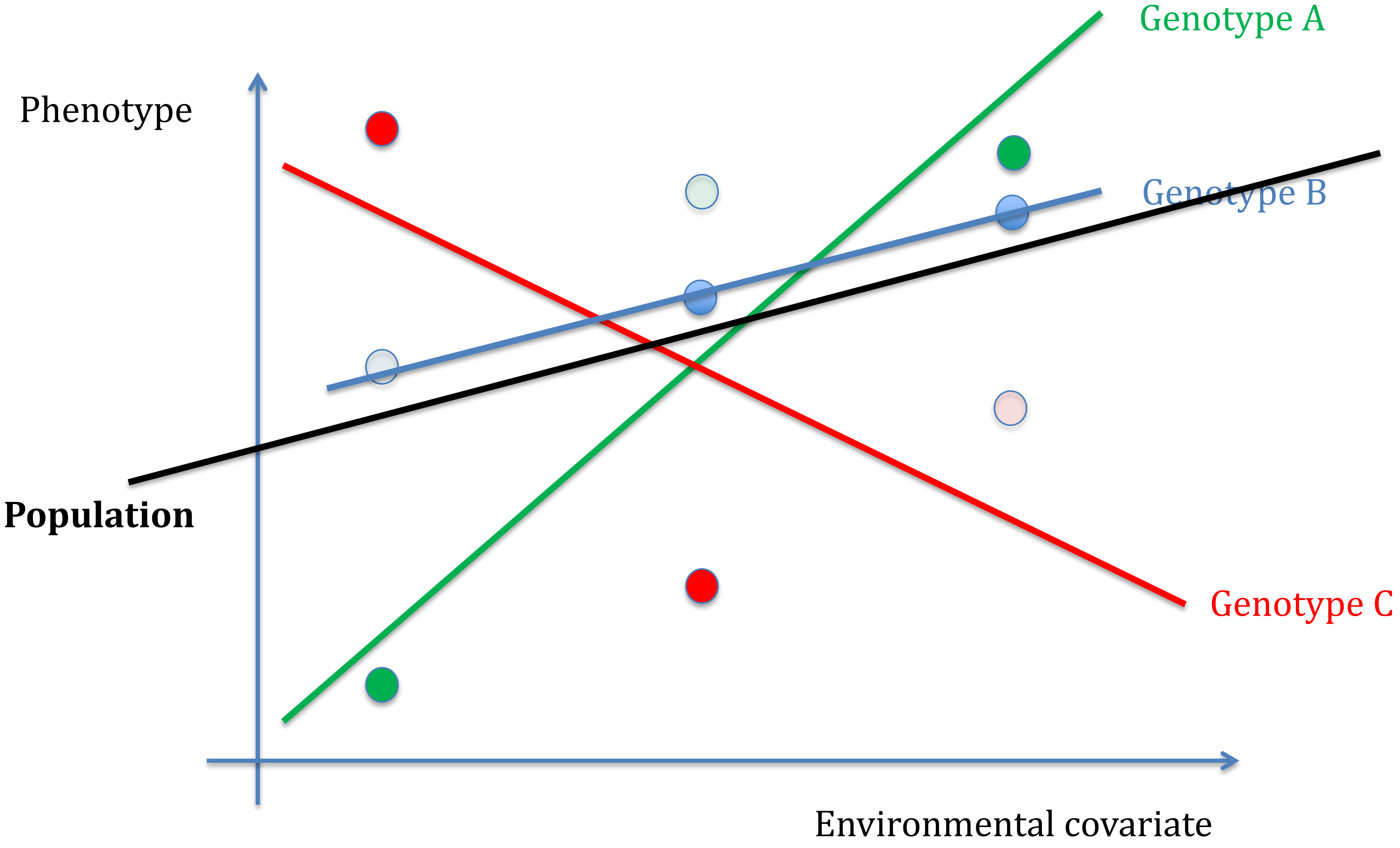
- Genotype by Environment interaction



- Genotype by Environment interaction: **reaction norms**

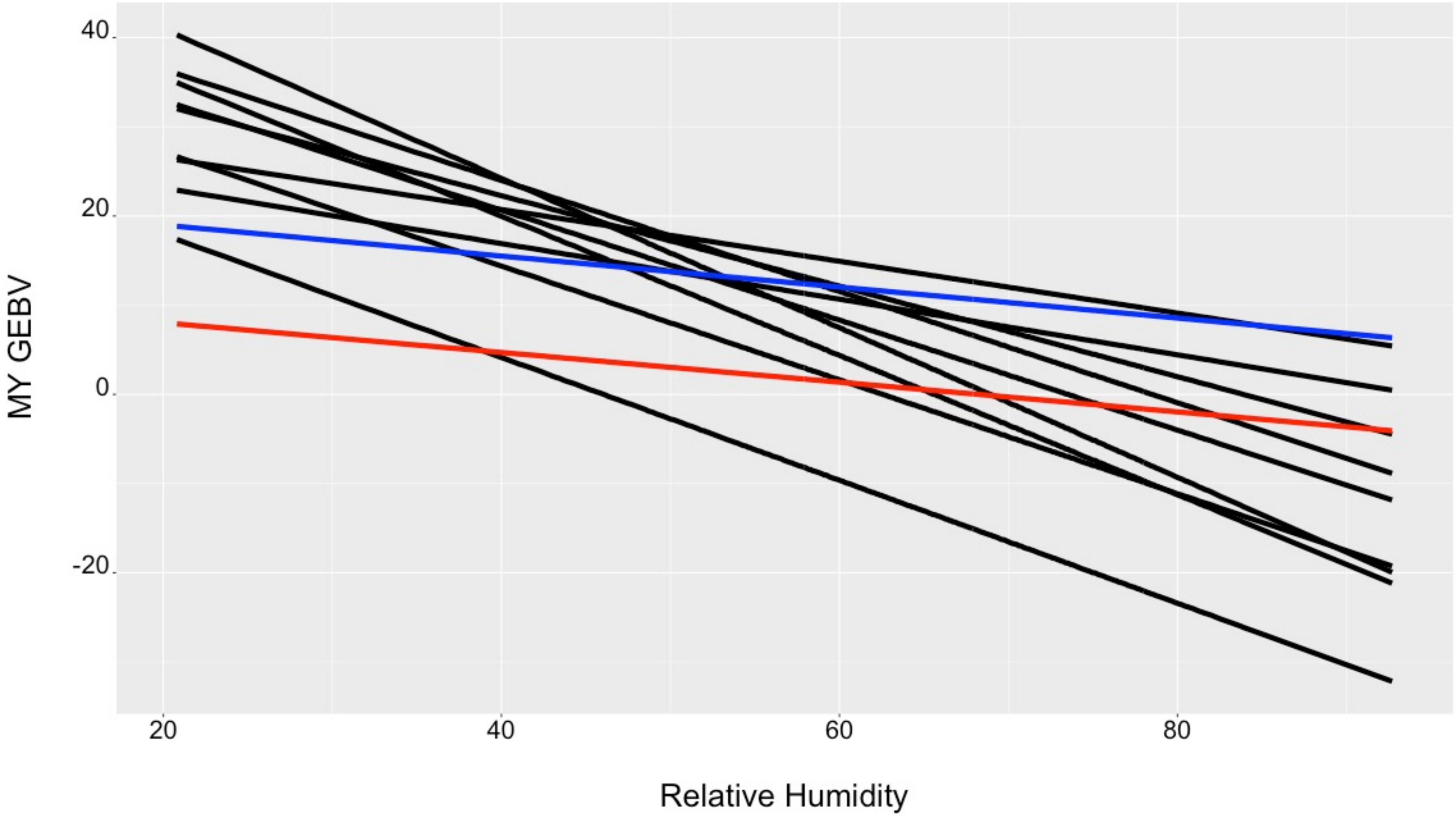


- Genotype by Environment interaction: **reaction norms**



- Genotype by Environment interaction: **reaction norms**

Reaction norms for most represented sires

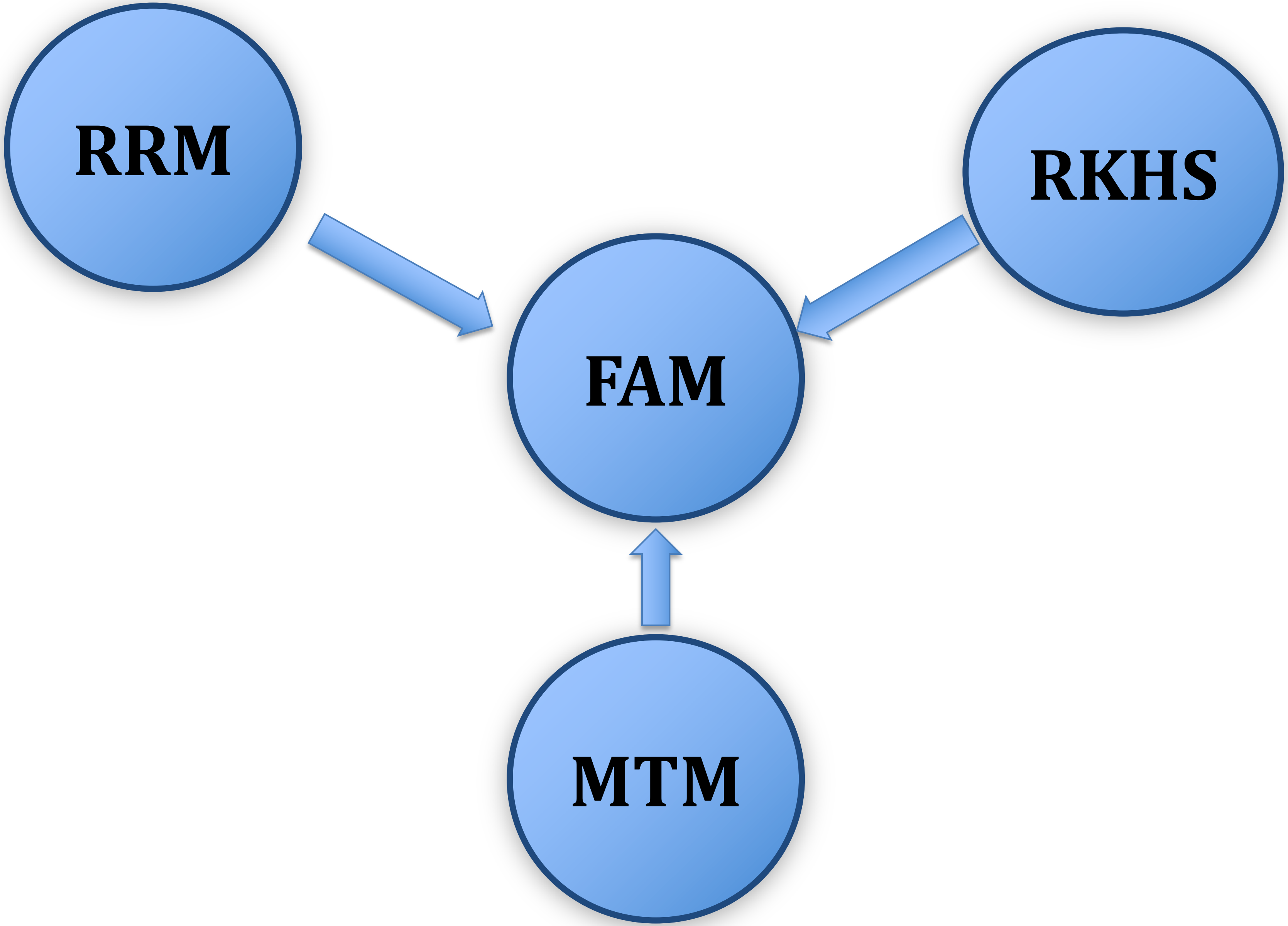


MTM

RRM

RKHS

MTM



RRM

RKHS

FAM

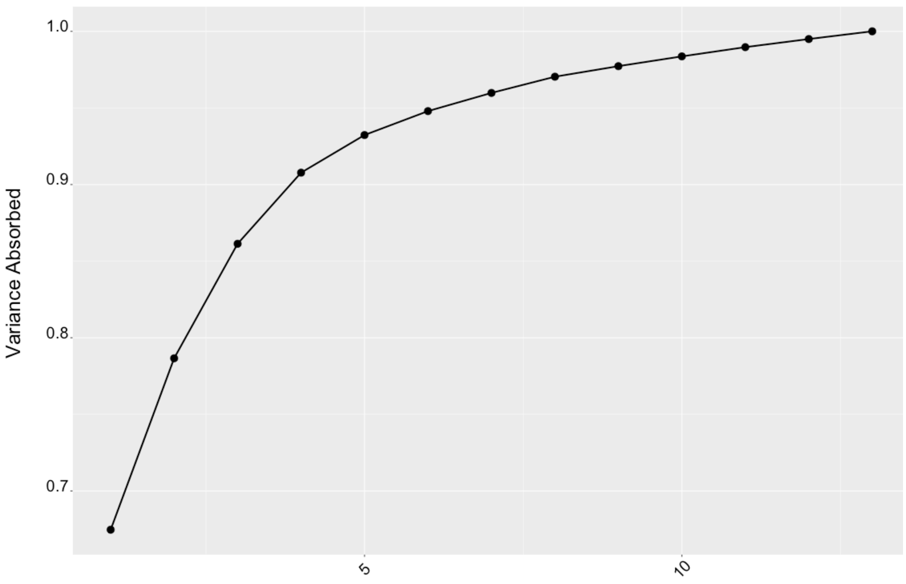
MTM

The multiple trait model

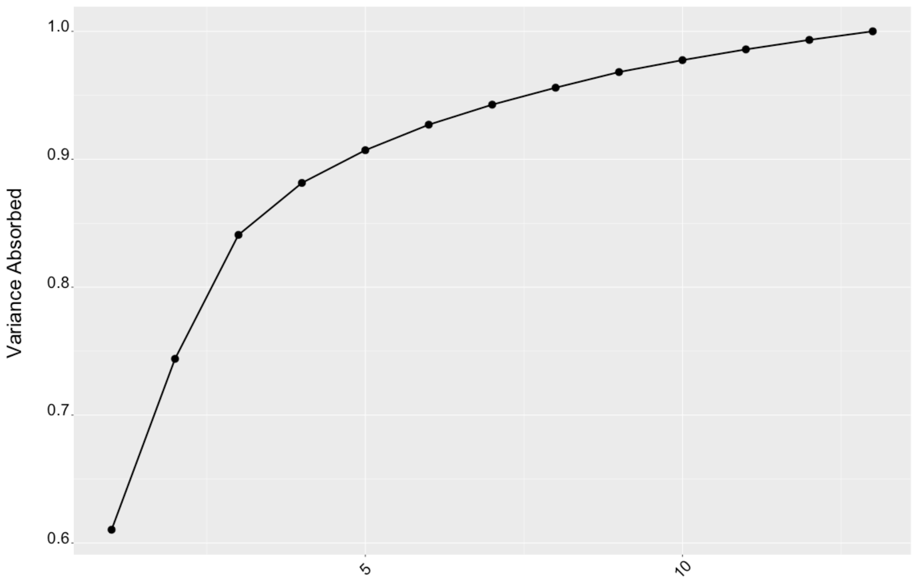
Genotype	Phenotype	Environment
A	7.36	a
B	6.51	a
A	4.06	a
B	6.52	a
A	4.18	a
B	6.88	a
A	10.62	b
B	12.86	b
A	11.41	b
B	11.18	b
A	13.59	b
B	10.91	b
A	2.24	c
B	1.45	c
A	3.83	c
B	2.25	c
A	3.59	c
B	3.33	c

Genotype	Phenotype	Environment	Trait
A	7.36	a	1
B	6.51	a	1
A	4.06	a	1
B	6.52	a	1
A	4.18	a	1
B	6.88	a	1
A	10.62	b	2
B	12.86	b	2
A	11.41	b	2
B	11.18	b	2
A	13.59	b	2
B	10.91	b	2
A	2.24	c	3
B	1.45	c	3
A	3.83	c	3
B	2.25	c	3
A	3.59	c	3
B	3.33	c	3

height	loc_1	loc_2	loc_3	loc_4	loc_5	loc_6	loc_7	loc_8	loc_9	loc_10	loc_11	loc_12	loc_13
loc_1	.	0.42	0.44	0.58	0.13	0.51	0.51	0.46	0.49	0.45	0.68	0.54	0.54
loc_2	0.42	.	0.74	0.81	0.57	0.57	0.83	0.81	0.67	0.86	0.4	0.64	0.64
loc_3	0.44	0.74	.	0.82	0.51	0.71	0.82	0.85	0.81	0.81	0.63	0.64	0.64
loc_4	0.58	0.81	0.82	.	0.62	0.61	0.82	0.84	0.75	0.8	0.6	0.71	0.71
loc_5	0.13	0.57	0.51	0.62	.	0.12	0.41	0.52	0.28	0.47	0.12	0.14	0.15
loc_6	0.51	0.57	0.71	0.61	0.12	.	0.75	0.68	0.77	0.69	0.7	0.66	0.66
loc_7	0.51	0.83	0.82	0.82	0.41	0.75	.	0.89	0.85	0.86	0.64	0.76	0.76
loc_8	0.46	0.81	0.85	0.84	0.52	0.68	0.89	.	0.82	0.84	0.61	0.68	0.68
loc_9	0.49	0.67	0.81	0.75	0.28	0.77	0.85	0.82	.	0.77	0.66	0.76	0.76
loc_10	0.45	0.86	0.81	0.8	0.47	0.69	0.86	0.84	0.77	.	0.51	0.69	0.69
loc_11	0.68	0.4	0.63	0.6	0.12	0.7	0.64	0.61	0.66	0.51	.	0.71	0.71
loc_12	0.54	0.64	0.64	0.71	0.14	0.66	0.76	0.68	0.76	0.69	0.71	.	0.95
loc_13	0.54	0.64	0.64	0.71	0.15	0.66	0.76	0.68	0.76	0.69	0.71	0.95	.

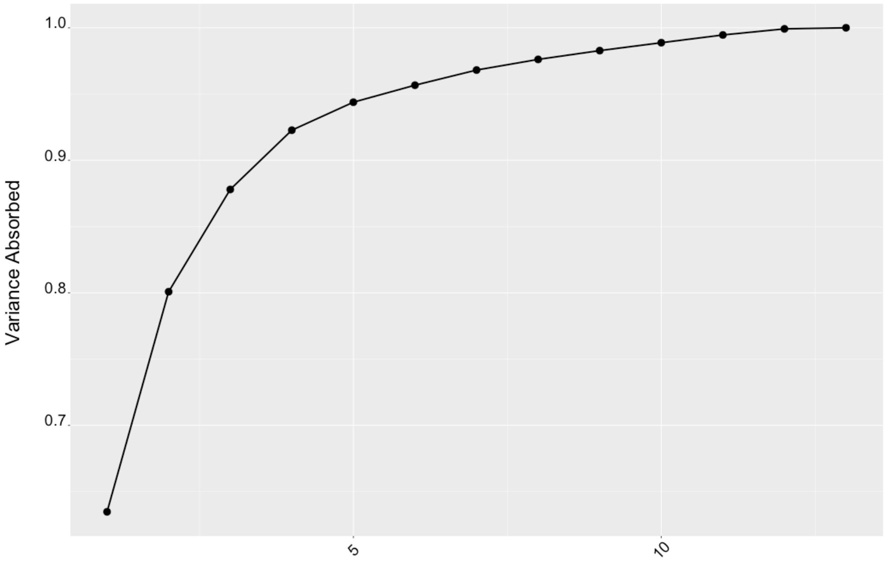


height



diameter

Eigenvalue decomposition
of the (co)variance matrix



volume

Genotype	Phenotype	Environment	Trait	Env Cov
A	7.36	a	1	-1
B	6.51	a	1	-1
A	4.06	a	1	-1
B	6.52	a	1	-1
A	4.18	a	1	-1
B	6.88	a	1	-1
A	10.62	b	1	5
B	12.86	b	1	5
A	11.41	b	1	5
B	11.18	b	1	5
A	13.59	b	1	5
B	10.91	b	1	5
A	2.24	c	1	-4
B	1.45	c	1	-4
A	3.83	c	1	-4
B	2.25	c	1	-4
A	3.59	c	1	-4
B	3.33	c	1	-4

Incidence Matrix **X**: covariate

Genotype	Env Cov
A	-1
B	-1
A	-1
B	-1
A	-1
B	-1
A	5
B	5
A	5
B	5
A	5
B	5
A	-4
B	-4
A	-4
B	-4
A	-4
B	-4

-1
-1
-1
-1
-1
-1
5
5
5
5
5
5
-4
-4
-4
-4
-4
-4

Genotype	Env Cov
A	-1
B	-1
A	-1
B	-1
A	-1
B	-1
A	5
B	5
A	5
B	5
A	5
B	5
A	-4
B	-4
A	-4
B	-4
A	-4
B	-4

Incidence Matrix **Z**:

additive genetic intercept

1	0
0	1
1	0
0	1
1	0
0	1
1	0
0	1
1	0
0	1
1	0
0	1
1	0
0	1
1	0
0	1
1	0
0	1

Incidence Matrix **Z**:

additive
genetic
slope

Genotype	Env Cov
A	-1
B	-1
A	-1
B	-1
A	-1
B	-1
A	5
B	5
A	5
B	5
A	5
B	5
A	-4
B	-4
A	-4
B	-4
A	-4
B	-4

-1	0
0	-1
-1	0
0	-1
-1	0
0	-1
5	0
0	5
5	0
0	5
5	0
0	5
-4	0
0	-4
-4	0
0	-4
-4	0
0	-4

Genotype	Env Cov
A	-1
B	-1
A	-1
B	-1
A	-1
B	-1
A	5
B	5
A	5
B	5
A	5
B	5
A	-4
B	-4
A	-4
B	-4
A	-4
B	-4

X
-1
-1
-1
-1
-1
-1
5
5
5
5
5
5
-4
-4
-4
-4
-4
-4

$Z_{intercept}$	
1	0
0	1
1	0
0	1
1	0
0	1
1	0
0	1
1	0
0	1
1	0
0	1
1	0
0	1
1	0
0	1
1	0
0	1

Z_{slope}	
-1	0
0	-1
-1	0
0	-1
-1	0
0	-1
5	0
0	5
5	0
0	5
5	0
0	5
-4	0
0	-4
-4	0
0	-4
-4	0
0	-4

location	dam	testLat	testTemp	rep	ht	dbh	vol
L1	8LKKNK18RW	-1.068474	-0.2096745	1	7.62	12.19	32.93
L1	N21WTSHLDC	-1.068474	-0.2096745	1	7.47	12.19	32.3
L1	2Q3JTJYKJ5	-1.068474	-0.2096745	1	6.86	10.67	23.01
L1	9952X5QYCW	-1.068474	-0.2096745	1	8.38	15.24	55.92
L1	25S2KSLMOD	-1.068474	-0.2096745	1	7.32	11.68	29.16
L1	ASFRGUPAX1	-1.068474	-0.2096745	1	7.32	11.94	30.43
L1	5RKHPBRZYT	-1.068474	-0.2096745	1	7.32	12.45	33
L1	J35M737LYL	-1.068474	-0.2096745	1	7.16	11.43	27.37
L1	NGZXWQXZ6X	-1.068474	-0.2096745	1	6.1	9.4	16.18
…	…	…	…	…	…	…	…

E

Latitude	Leg_0	Leg_1
-1.9829	0.7071	-1.2247
-1.3445	0.7071	-0.8372
-1.0685	0.7071	-0.6696
-0.9404	0.7071	-0.5919
-0.694	0.7071	-0.4423
-0.3804	0.7071	-0.252
-0.3485	0.7071	-0.2327
-0.1188	0.7071	-0.0932
0.1636	0.7071	0.0782
0.4869	0.7071	0.2744
0.6098	0.7071	0.349
1.4584	0.7071	0.8641
2.0525	0.7071	1.2247

Height regressed on latitude

$$y_{ijk} = b_1 E_{i1} + b_2 E_{i2} + a_{j1} E_{i1} + a_{j2} E_{i2} + e_{ijk}$$

$$\mathbf{a} \mid \mathbf{G} \sim N(0, \mathbf{G} \otimes \mathbf{A})$$

$$\mathbf{G} = \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix}$$

Height regressed on latitude

$$\mathbf{b} = \begin{bmatrix} 11.159 \\ 1.518 \end{bmatrix}$$

$$\mathbf{G} = \begin{bmatrix} 0.073 & 0.0001 \\ 0.0001 & 0.0097636877 \end{bmatrix}$$

dam	a_0	a_1
1C42G8GL8F	0.041	-0.013
1LVIVK3E3T	0.225	-0.004
22EQY58P1F	-0.138	-0.015
2JLGOZIC4Z	0.127	0.110
…	…	…

$$\mathbf{b} = \begin{bmatrix} 11.159 \\ 1.518 \end{bmatrix}$$

dam	a_0	a_1
1C42G8GL8F	**0.041**	**-0.013**
1LVIVK3E3T	0.225	-0.004
22EQY58P1F	-0.138	-0.015
2JLGOZIC4Z	0.127	0.110
…	…	…

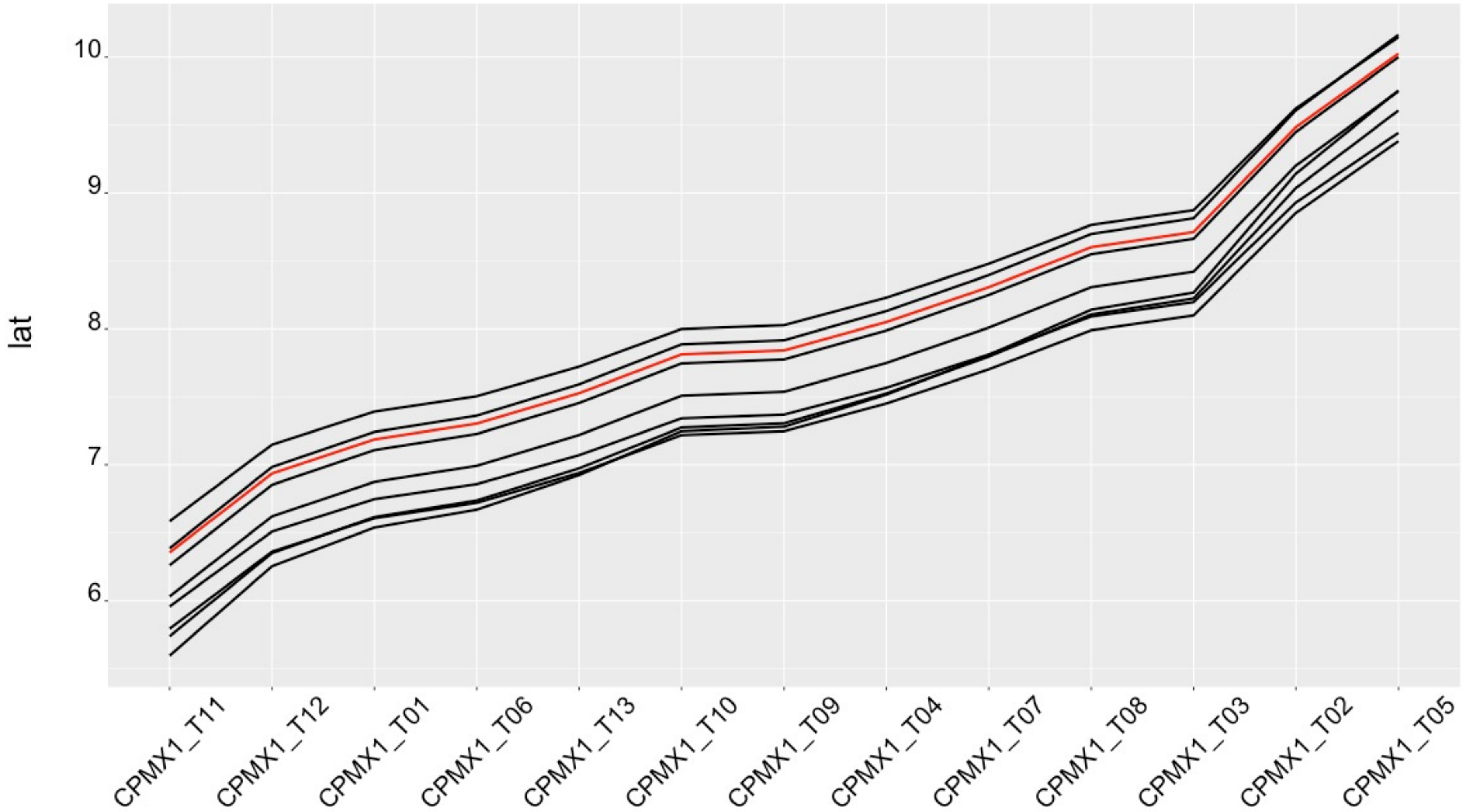
Latitude	**Leg_0**	**Leg_1**
-1.9829	**0.7071**	**-1.2247**
-1.3445	0.7071	-0.8372

`E%*%t(a)`

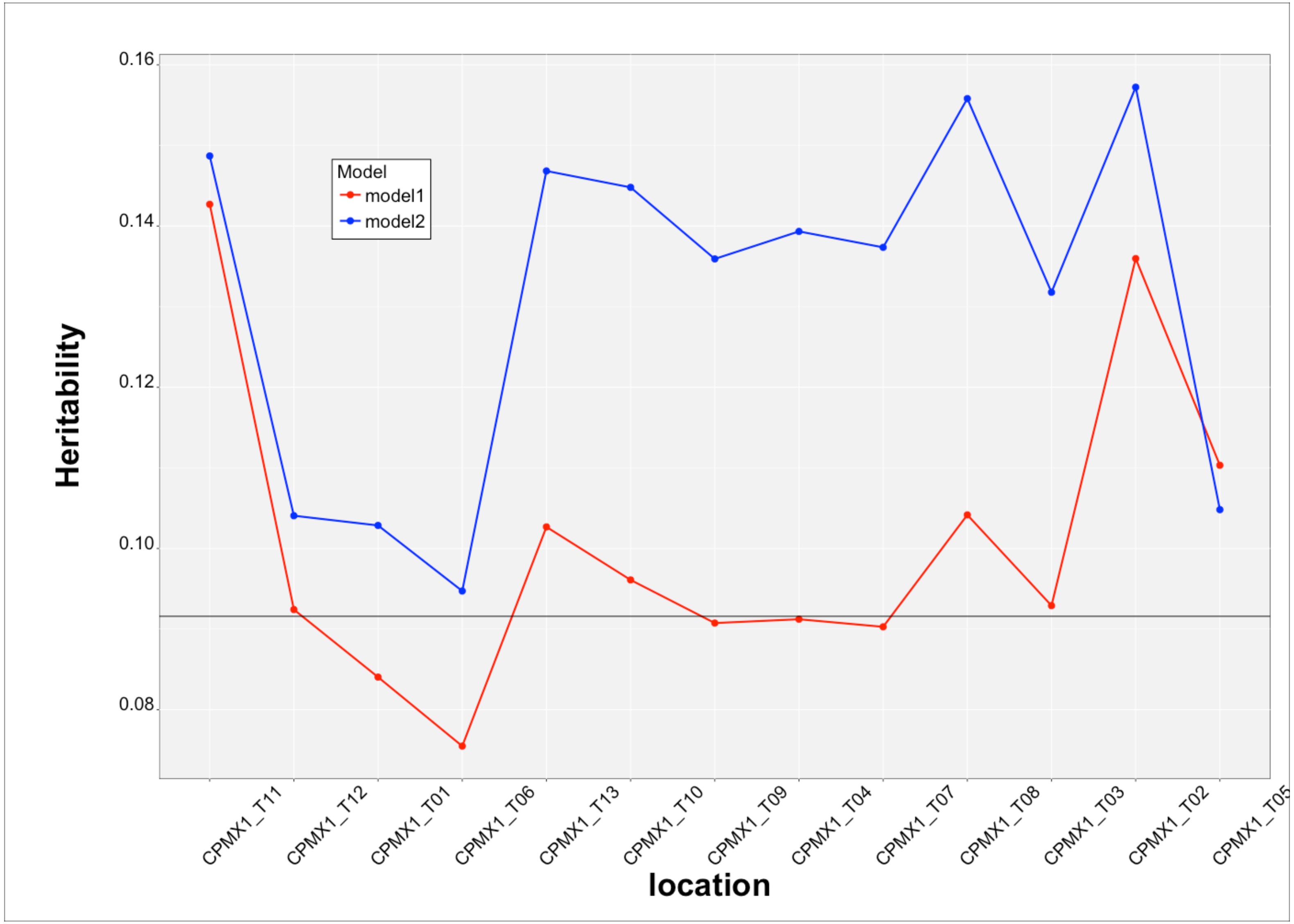
$$y_{ij} = 11.159 * 0.7071 + 1.518 * -1.2247 + 0.041 * 0.7071 + -0.013 * -1.2247$$

NC STATE UNIVERSITY

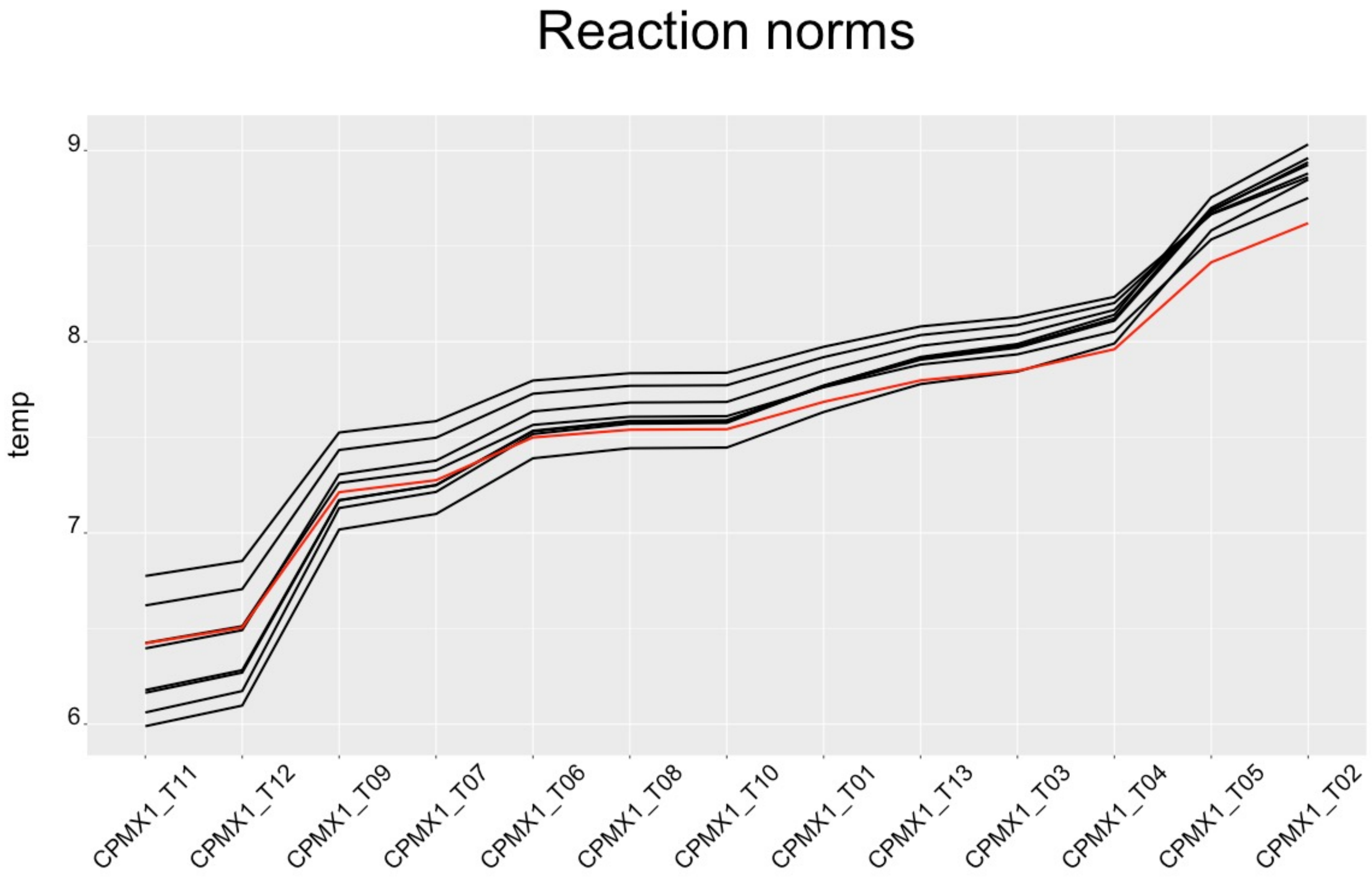
Reaction norms



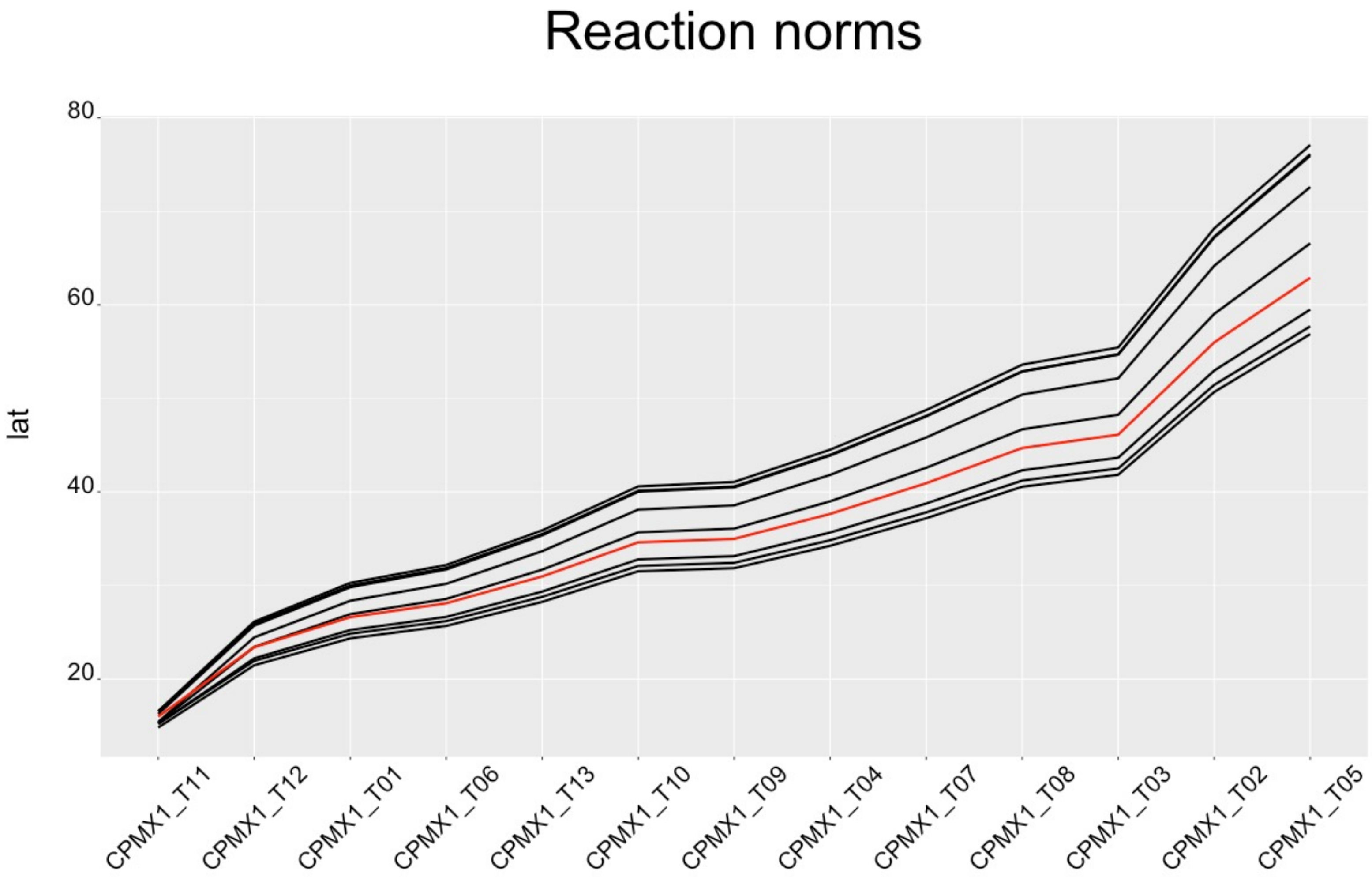
`diag(E%*%G%*%t(E))`


Heritability
0.16
0.14
0.12
0.10
0.08
Model
model1
model2
CPMX1_T11
CPMX1_T12
CPMX1_T01
CPMX1_T06
CPMX1_T13
CPMX1_T10
CPMX1_T09
CPMX1_T04
CPMX1_T07
CPMX1_T08
CPMX1_T03
CPMX1_T02
CPMX1_T05
location

Height on temperature



Volume on latitude



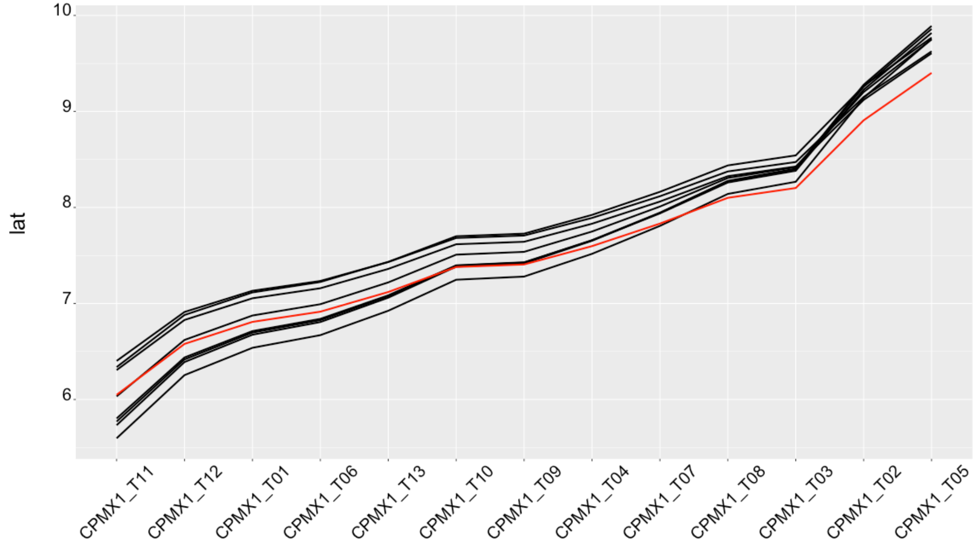
Height regressed on latitude

$$\mathbf{b} = \begin{bmatrix} 11.159 \\ 1.518 \\ -0.654 \end{bmatrix} \qquad \mathbf{G} = \begin{bmatrix} 0.0733 & -0.0006 & -0.0073 \\ -0.0006 & 0.0099 & 0.0002 \\ -0.0073 & 0.0002 & 0.0025 \end{bmatrix}$$

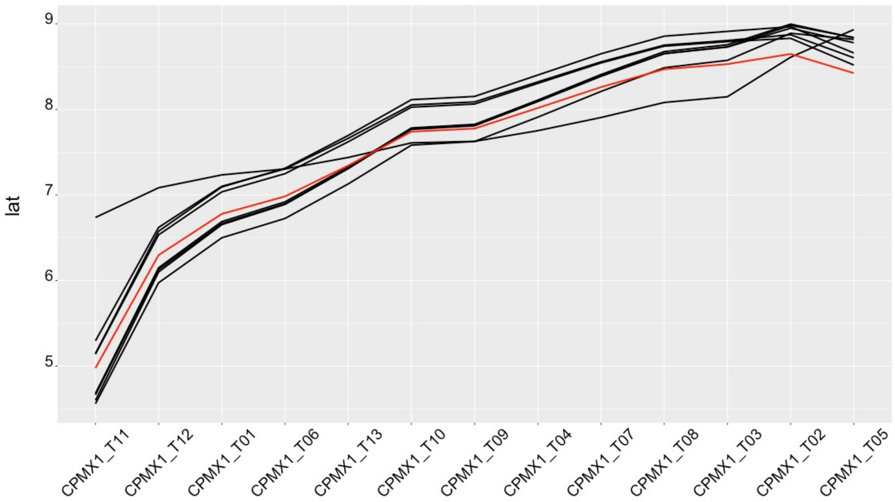
dam	a_0	a_1	a_2
1C42G8GL8F	11.13	1.55	-0.67
1LVIVK3E3T	11.31	1.65	-0.69
22EQY58P1F	10.95	1.56	-0.67
2JLGOZIC4Z	11.18	1.38	-0.68
…	…	…	…

Height on latitude

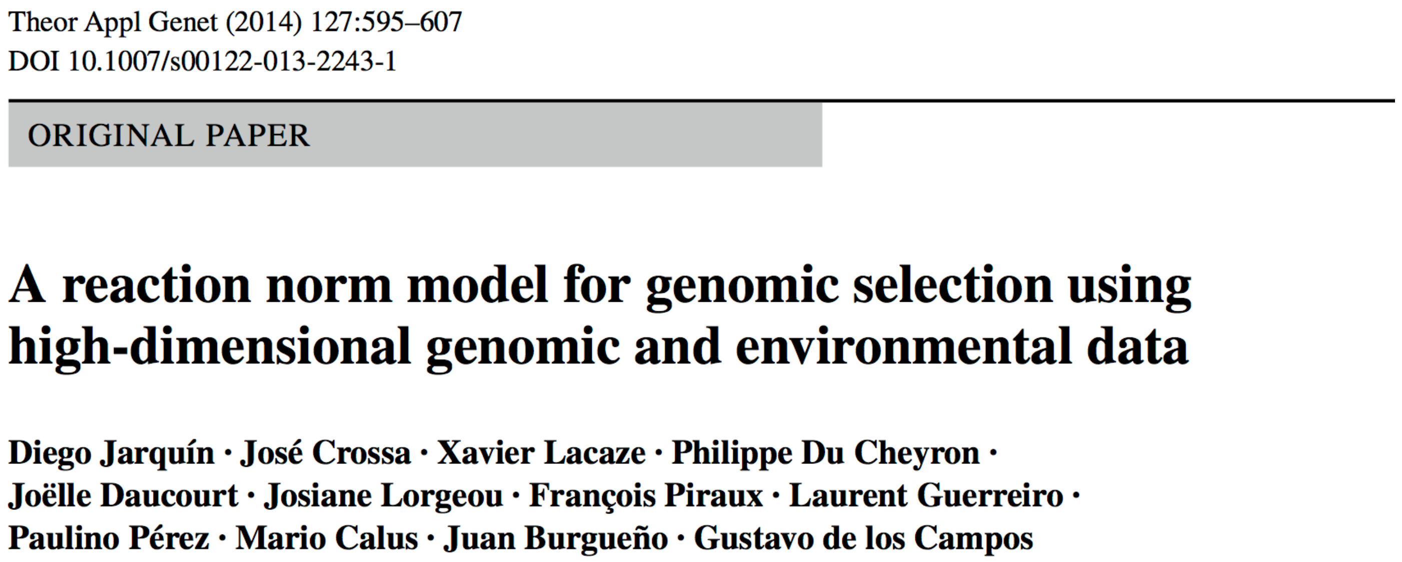
Reaction norms



Reaction norms



Reproducting Kernel Hilbert Spaces regression

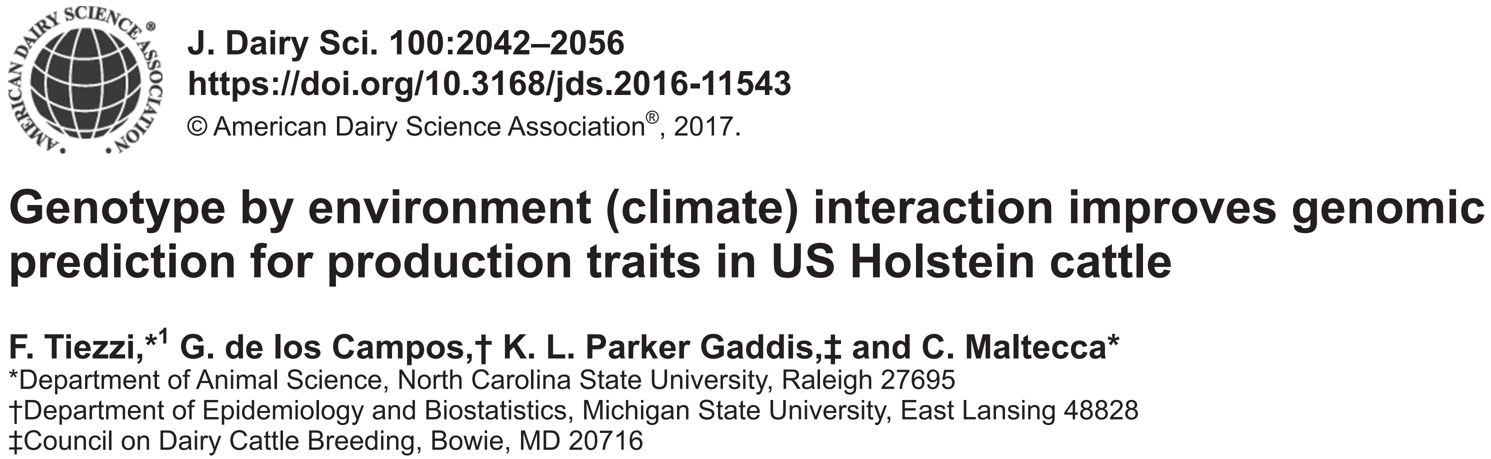


Theor Appl Genet (2014) 127:595–607
DOI 10.1007/s00122-013-2243-1

ORIGINAL PAPER

A reaction norm model for genomic selection using high-dimensional genomic and environmental data

**Diego Jarquín · José Crossa · Xavier Lacaze · Philippe Du Cheyron ·
Joëlle Daucourt · Josiane Lorgeou · François Piraux · Laurent Guerreiro ·
Paulino Pérez · Mario Calus · Juan Burgueño · Gustavo de los Campos**



J. Dairy Sci. 100:2042–2056
https://doi.org/10.3168/jds.2016-11543

Genotype by environment (climate) interaction improves genomic prediction for production traits in US Holstein cattle

F. Tiezzi,*[1] G. de los Campos,† K. L. Parker Gaddis,‡ and C. Maltecca*
*Department of Animal Science, North Carolina State University, Raleigh 27695
†Department of Epidemiology and Biostatistics, Michigan State University, East Lansing 48828
‡Council on Dairy Cattle Breeding, Bowie, MD 20716

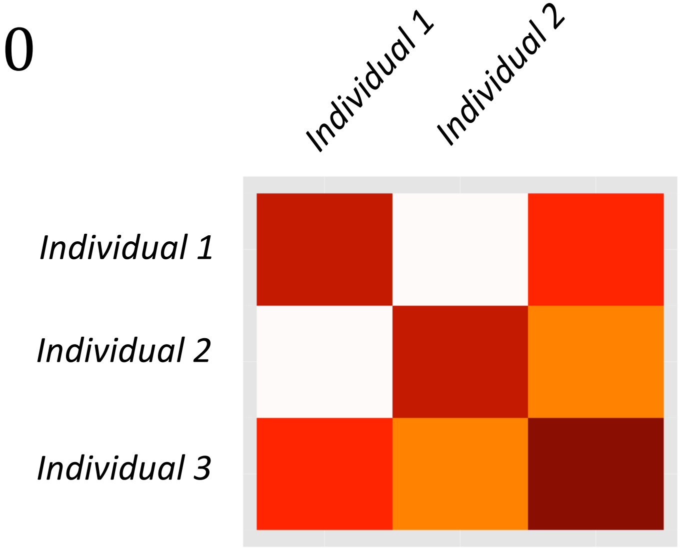
The Genomic Relationship Matrix

AA = 2

Aa = 1

aa = 0

$$\mathbf{M} = \begin{array}{cccccc} \text{Locus 1} & \text{Locus 2} & \text{Locus 3} & \text{Locus 4} & \text{Locus 5} & \\ 1 & \boxed{2} & 2 & 0 & \boxed{1} & \text{Individual 1} \\ \boxed{0} & 0 & 2 & \boxed{1} & \boxed{1} & \text{Individual 2} \\ \boxed{0} & \boxed{2} & 2 & \boxed{1} & 0 & \text{Individual 3} \end{array}$$



$$\mathbf{G} = \frac{[\mathbf{M}-\mathbf{P}][\mathbf{M}-\mathbf{P}]'}{\sum 2p(1-p)}$$

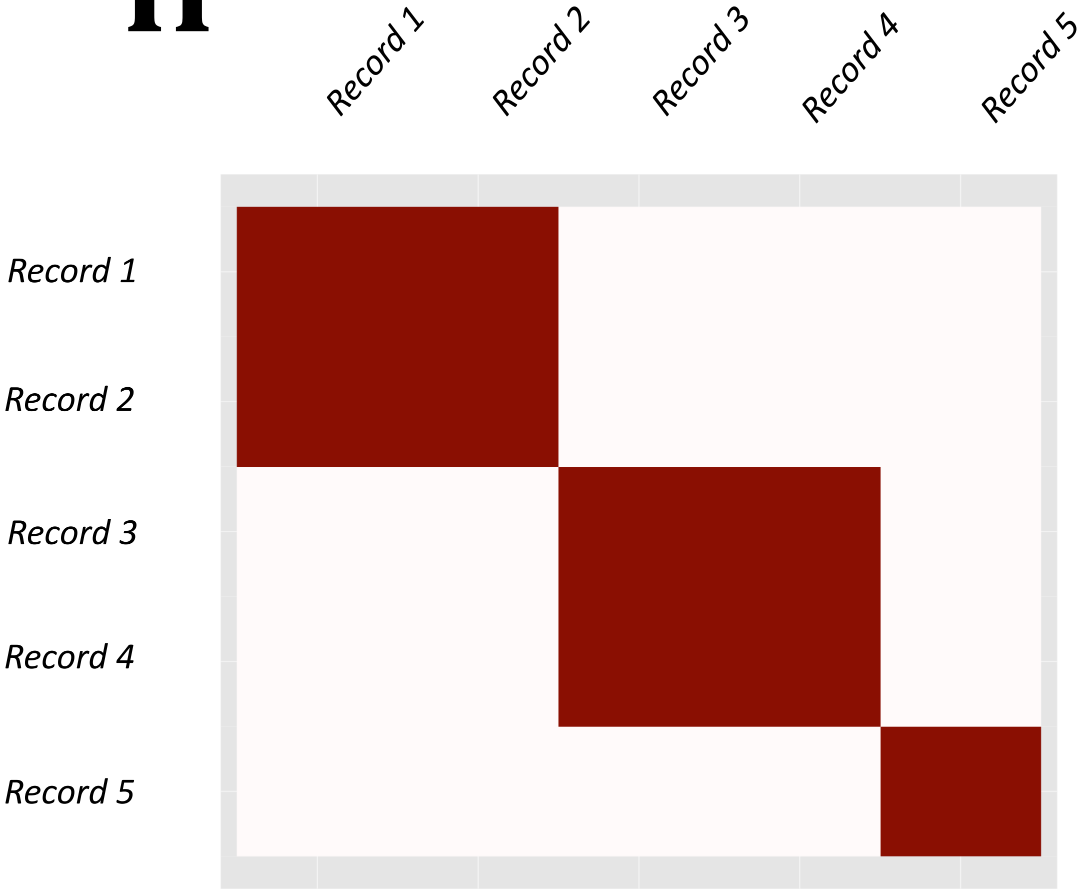
$$\mathrm{Cov}(\mathbf{g}) = \mathbf{G}\sigma_g^2$$

Nejati-Javaremi et al., 1997; VanRaden, 2008

The ‘Environmental’ Relationship Matrix

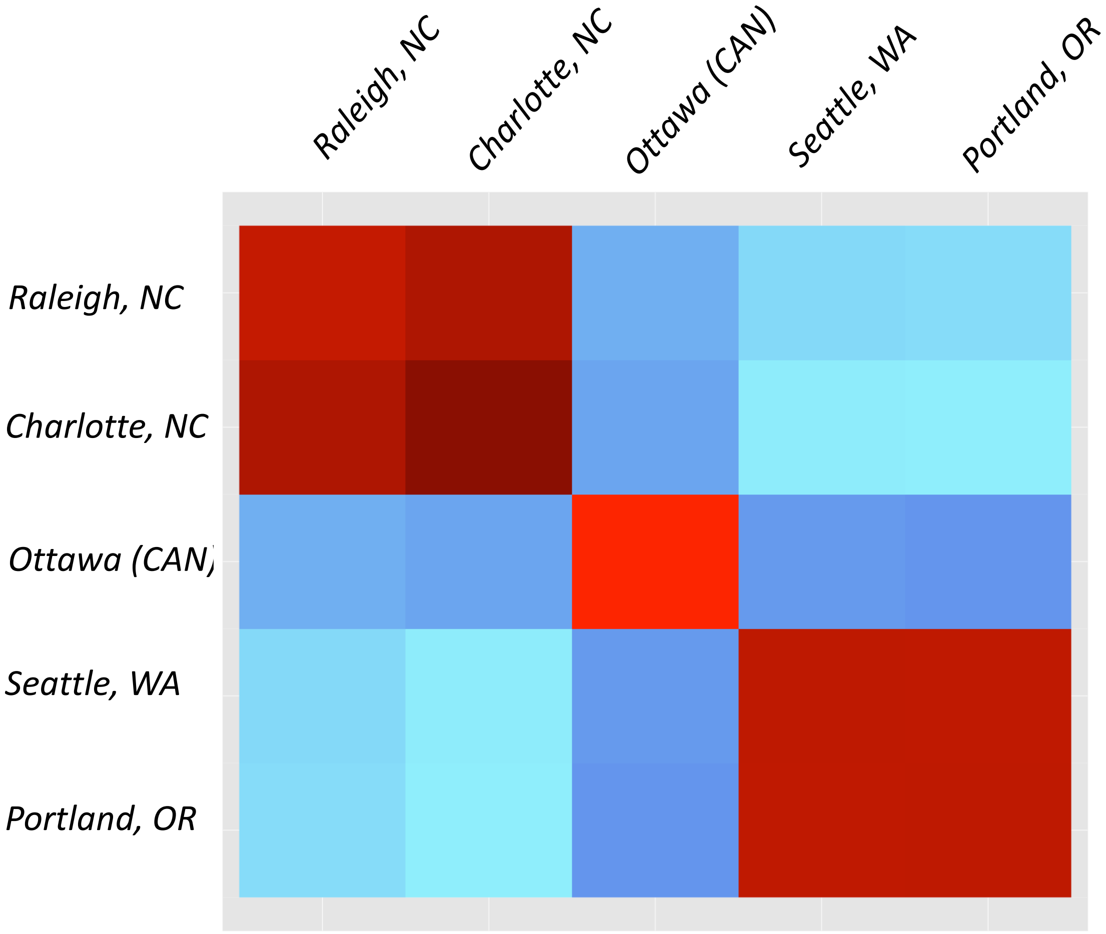
$$\mathbf{F} = \begin{array}{ccc|l} \textit{Herd 1} & \textit{Herd 2} & \textit{Herd 3} & \\ 1 & 0 & 0 & \textit{Record 1} \\ 1 & 0 & 0 & \textit{Record 2} \\ 0 & 1 & 0 & \textit{Record 3} \\ 0 & 1 & 0 & \textit{Record 4} \\ 0 & 0 & 1 & \textit{Record 5} \end{array}$$

H



The ‘Geographical’ Relationship Matrix

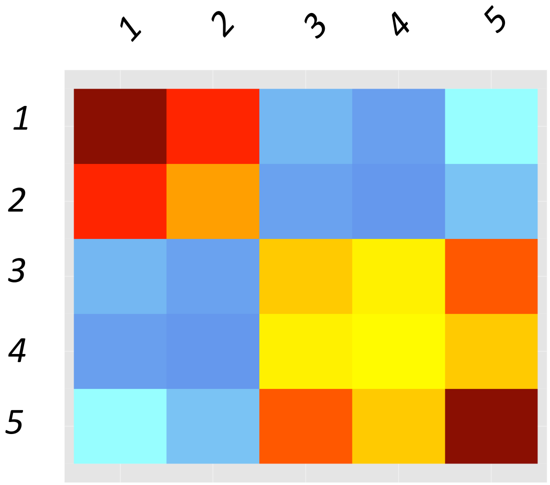
$$\mathbf{L} = \begin{bmatrix} 35.46 & 78.38 \\ 35.13 & 80.50 \\ 45.25 & 75.41 \\ 47.37 & 122.19 \\ 45.31 & 122.40 \end{bmatrix}$$



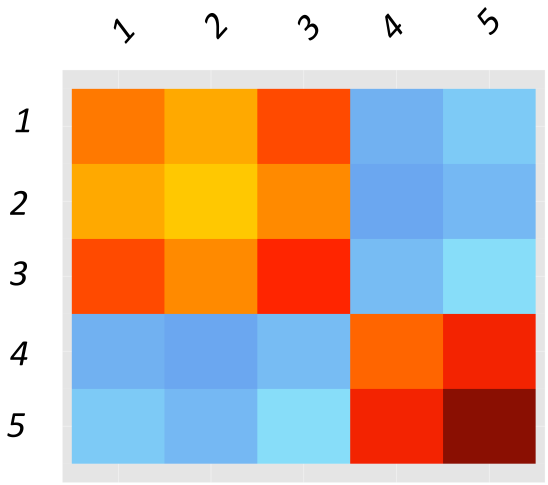
Hadamard product and the GxE

$$\begin{bmatrix} 2 & \cdots & 2 \\ \vdots & \ddots & \vdots \\ 2 & \cdots & 2 \end{bmatrix} \circ \begin{bmatrix} 3 & \cdots & 3 \\ \vdots & \ddots & \vdots \\ 3 & \cdots & 3 \end{bmatrix} = \begin{bmatrix} 6 & \cdots & 6 \\ \vdots & \ddots & \vdots \\ 6 & \cdots & 6 \end{bmatrix}$$

Hadamard product and the GxE



$\circ$



Hadamard product and the GxE