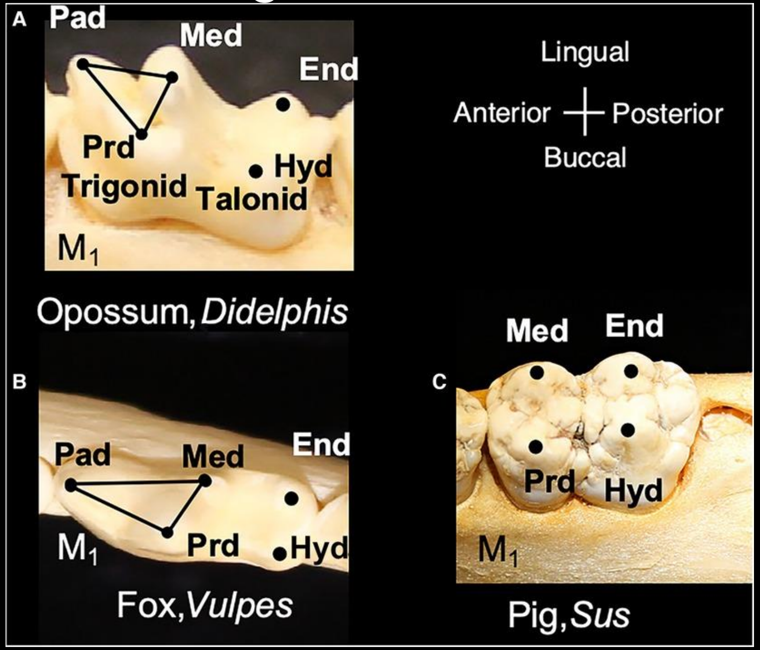
# MULTIVARIATE PCM

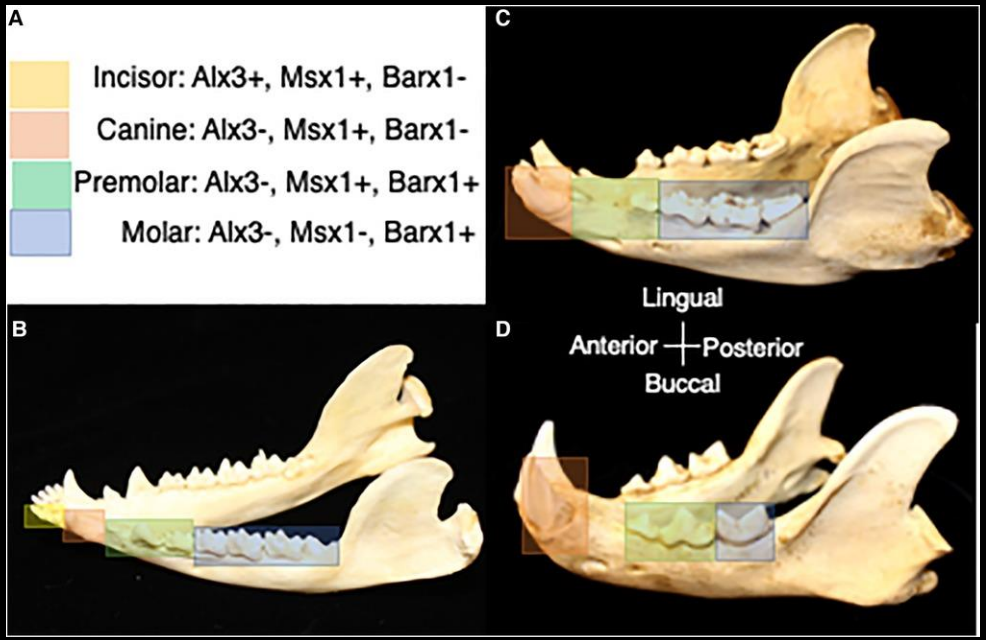
EQGW 2025

# Biological traits can be complex

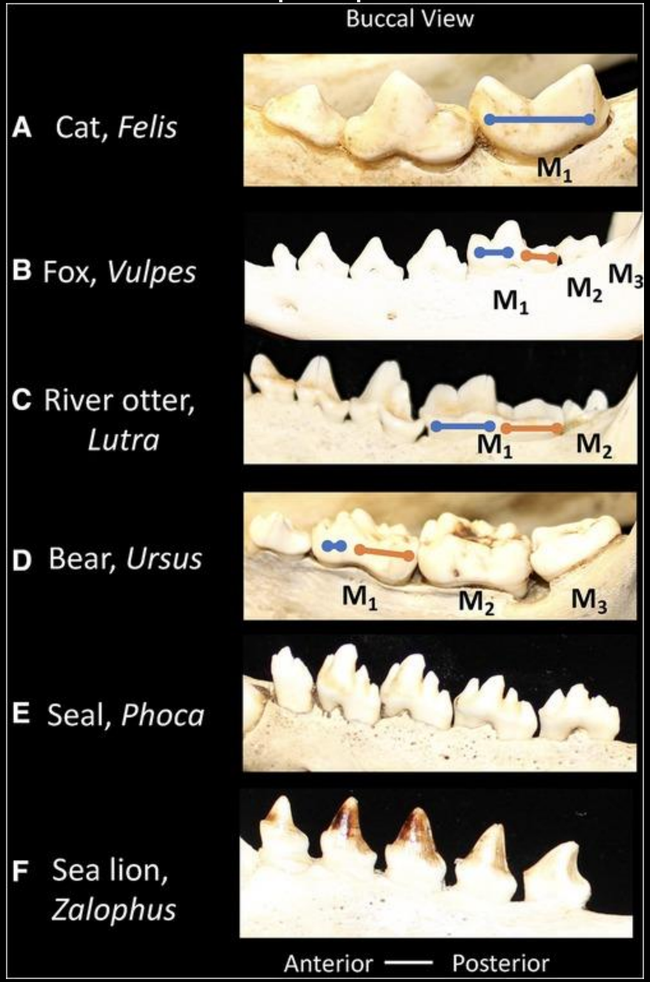
## Single tooth



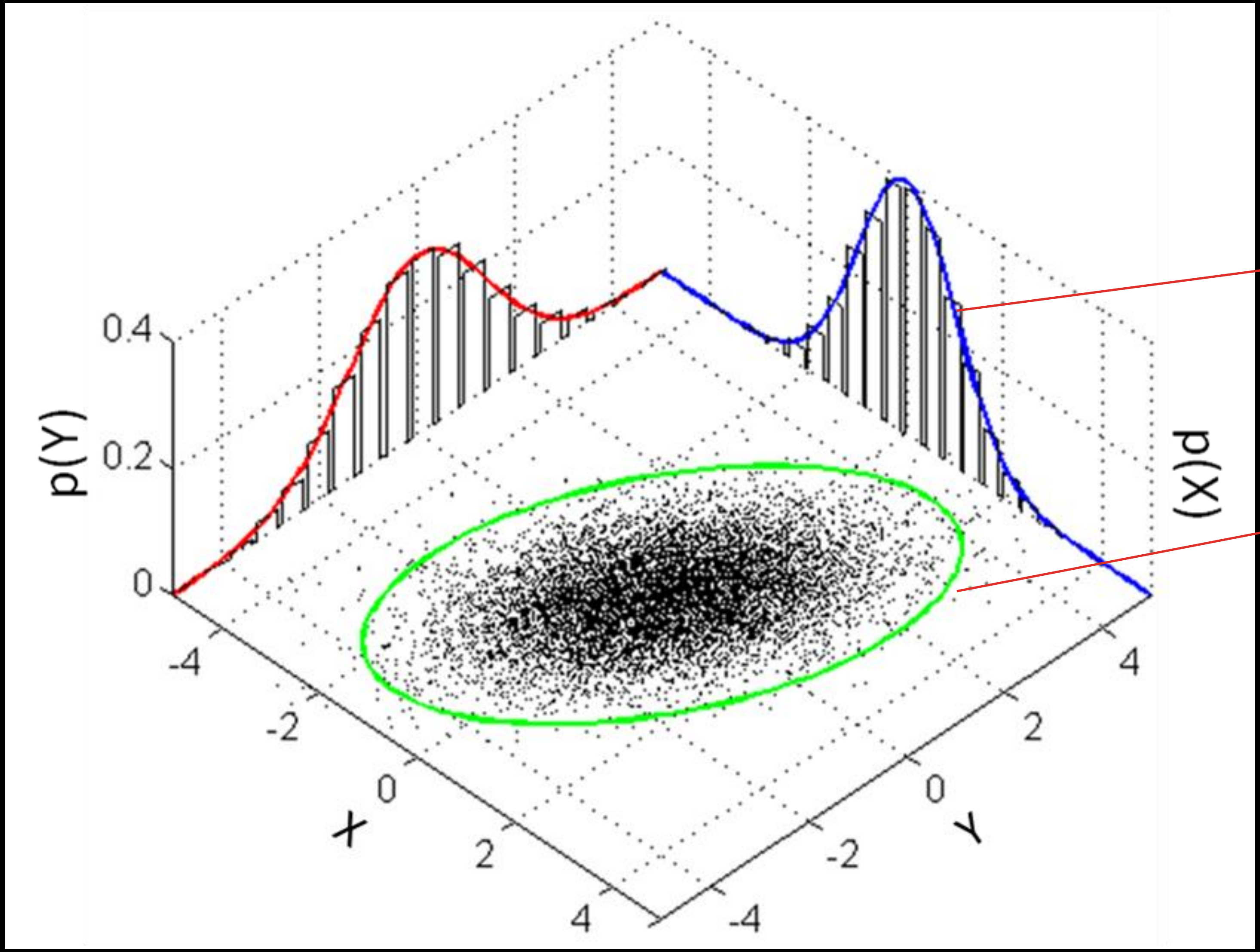

## Dental formulas




## Dentition proportions




Popowics & Mulamini 2023




Distribution of individual traits

Degree of trait association

$$\sigma^2 \rightarrow \Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}$$

# Univariate Case

$$dy = -\alpha(y(t) - \theta(t))dt + \sigma dW$$

?

?

?

Univariate Case

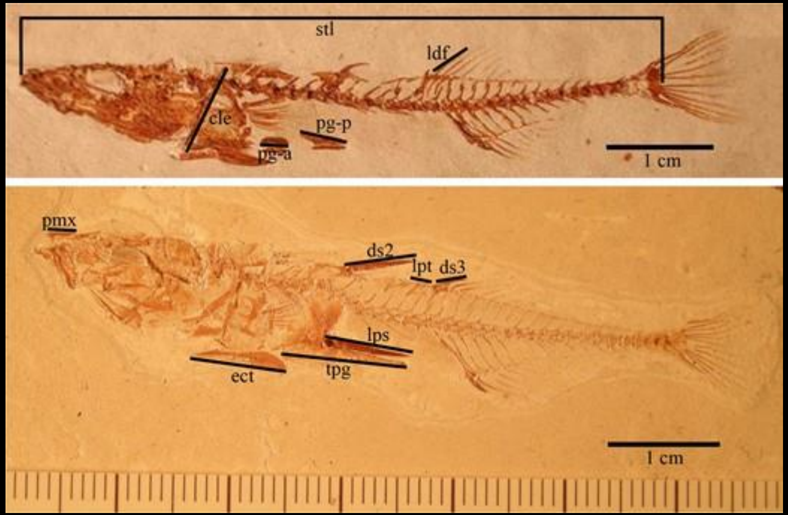
$$dy = -\alpha(y(t) - \theta(t))dt + \sigma dW$$

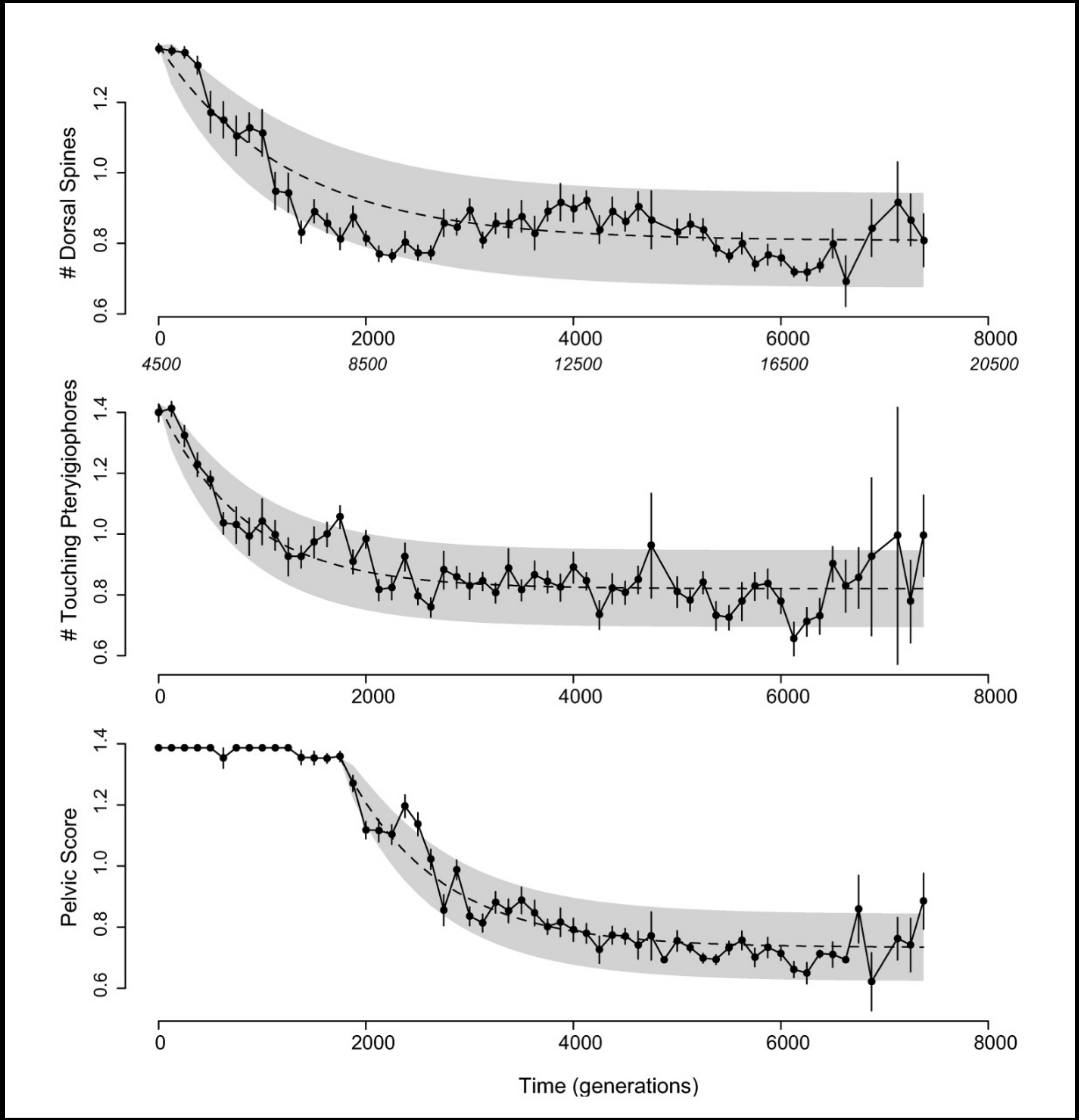
Strength of stabilizing selection

Adaptive peak

Rate of genetic drift

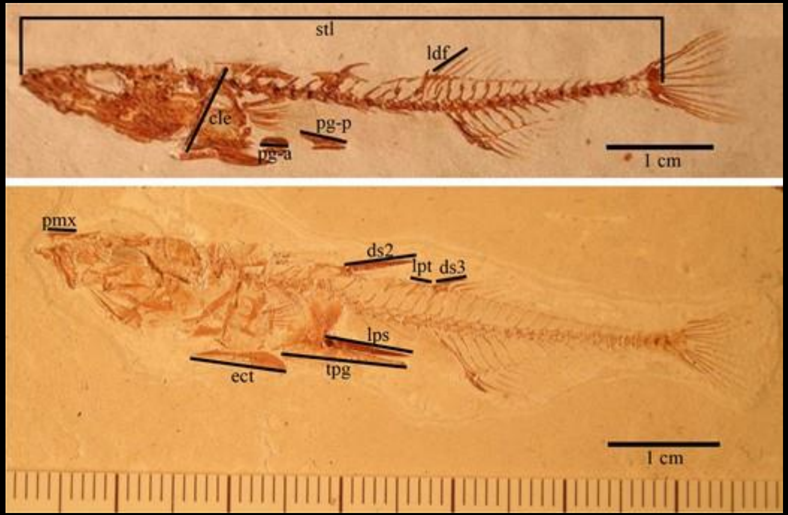
Microevolutionary interpretation

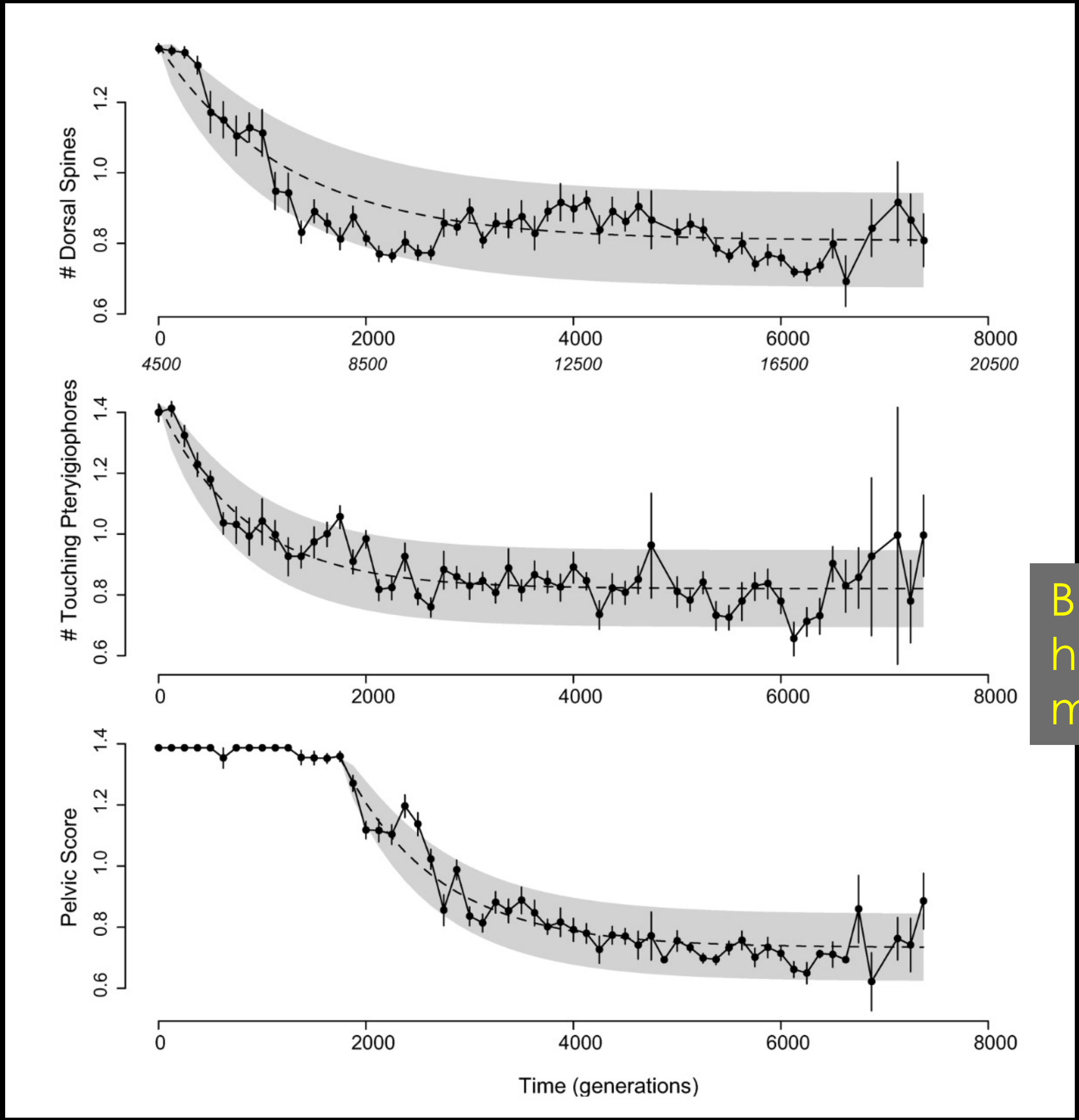



| Trait | Displacement | $\sigma^2_P$ | $N_e$ | $\omega^2$ | $t_{1/2}$ |
|---|---|---|---|---|---|
| No. of dorsal spines | −2.80 | 0.041 | 575–4023 | 5.0–35.2 | 853 |
| Pterygiophores | −2.13 | 0.081 | 851–5957 | 6.7–47.3 | 580 |
| Pelvic score | −2.57 | 0.059 | 889–6222 | 5.3–37.5 | 635 |

Hunt et al. 2008

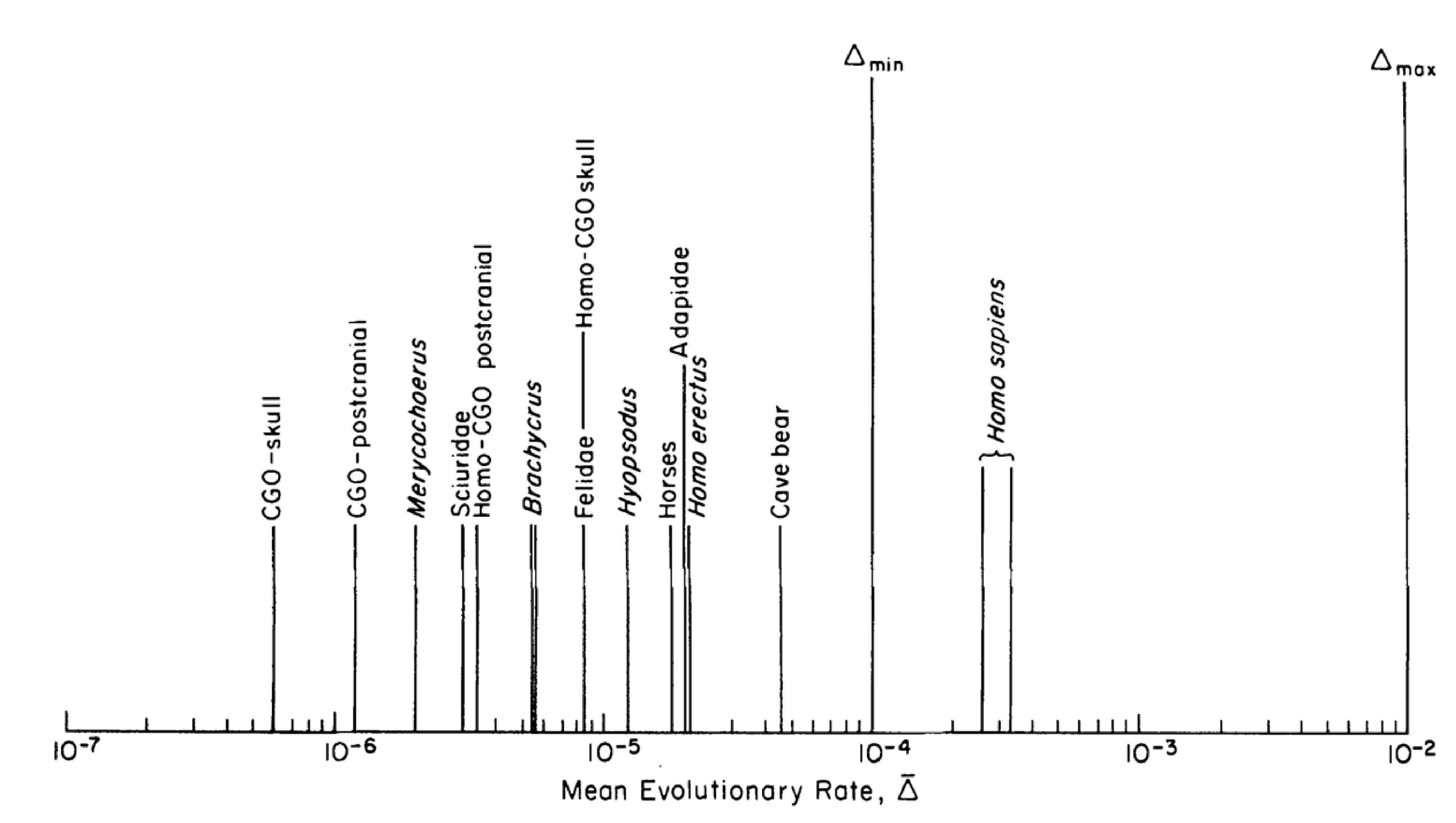



But this is not really happening in macroevolution, right?

| Trait | Displacement | $\sigma^2_P$ | $N_e$ | $\omega^2$ | $t_{1/2}$ |
|---|---|---|---|---|---|
| No. of dorsal spines | −2.80 | 0.041 | 575–4023 | 5.0–35.2 | 853 |
| Pterygiophores | −2.13 | 0.081 | 851–5957 | 6.7–47.3 | 580 |
| Pelvic score | −2.57 | 0.059 | 889–6222 | 5.3–37.5 | 635 |

Hunt et al. 2008


CGO-skull
CGO-postcranial
Merycochoerus
Sciuridae
Homo-CGO postcranial
Brachycrus
Felidae
Homo-CGO skull
Hyopsodus
Horses
Adapidae
Homo erectus
Cave bear
Δ min
Homo sapiens
Δ max
10-7
10-6
10-5
10-4
10-3
10-2
Mean Evolutionary Rate, Δ̄

Lynch, 1990

## Univariate Case

$$dy = -\alpha(y(t) - \theta(t))dt + \sigma dW$$

Rate of adaptation

Primary optima

Rate of stochastic evolution

All are scalars= single value

Hansen 1997

## Univariate Case

$$dy = -\alpha(y(t) - \theta(t))dt + \sigma dW$$

Rate of adaptation

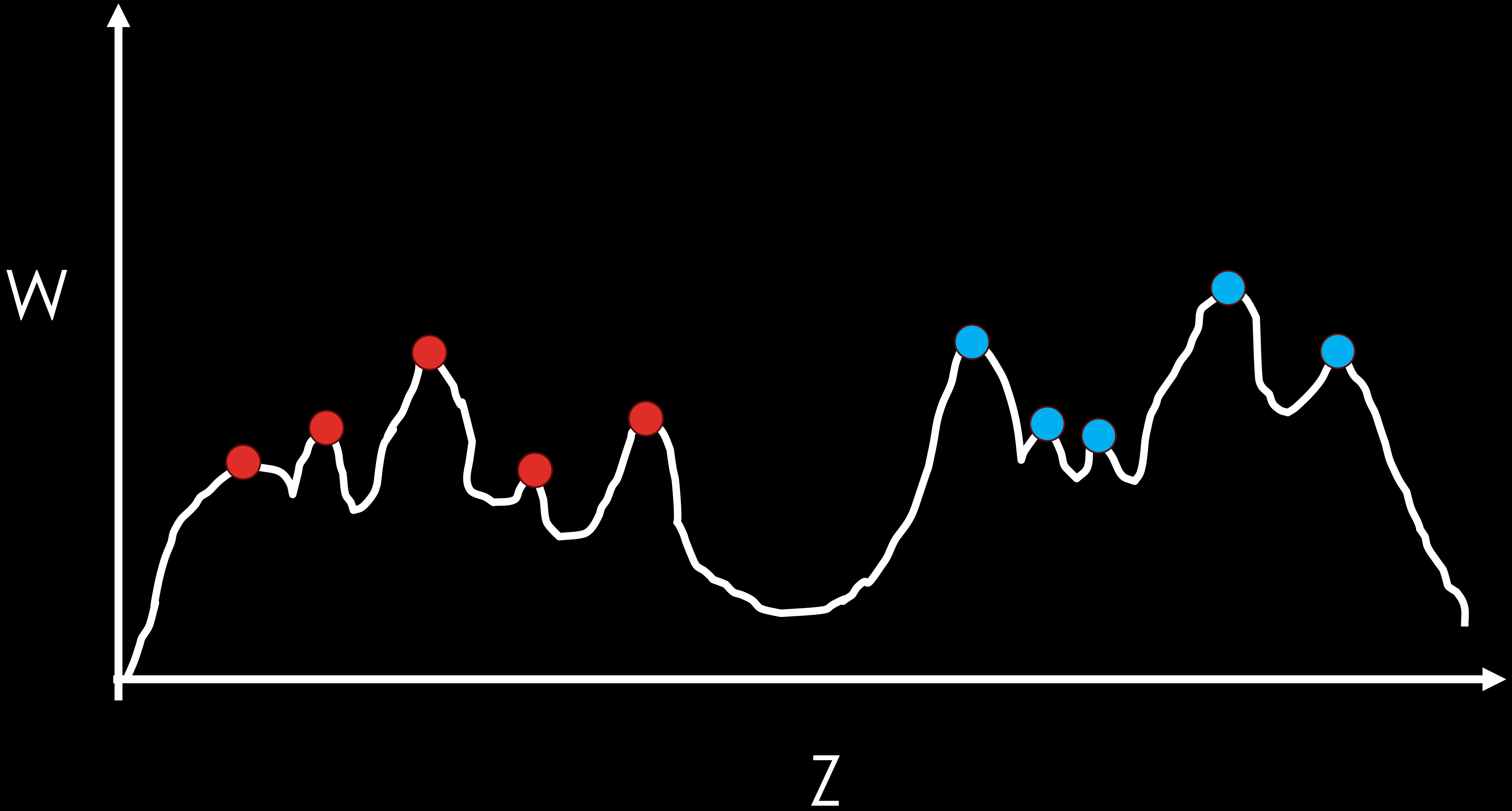
Primary optima

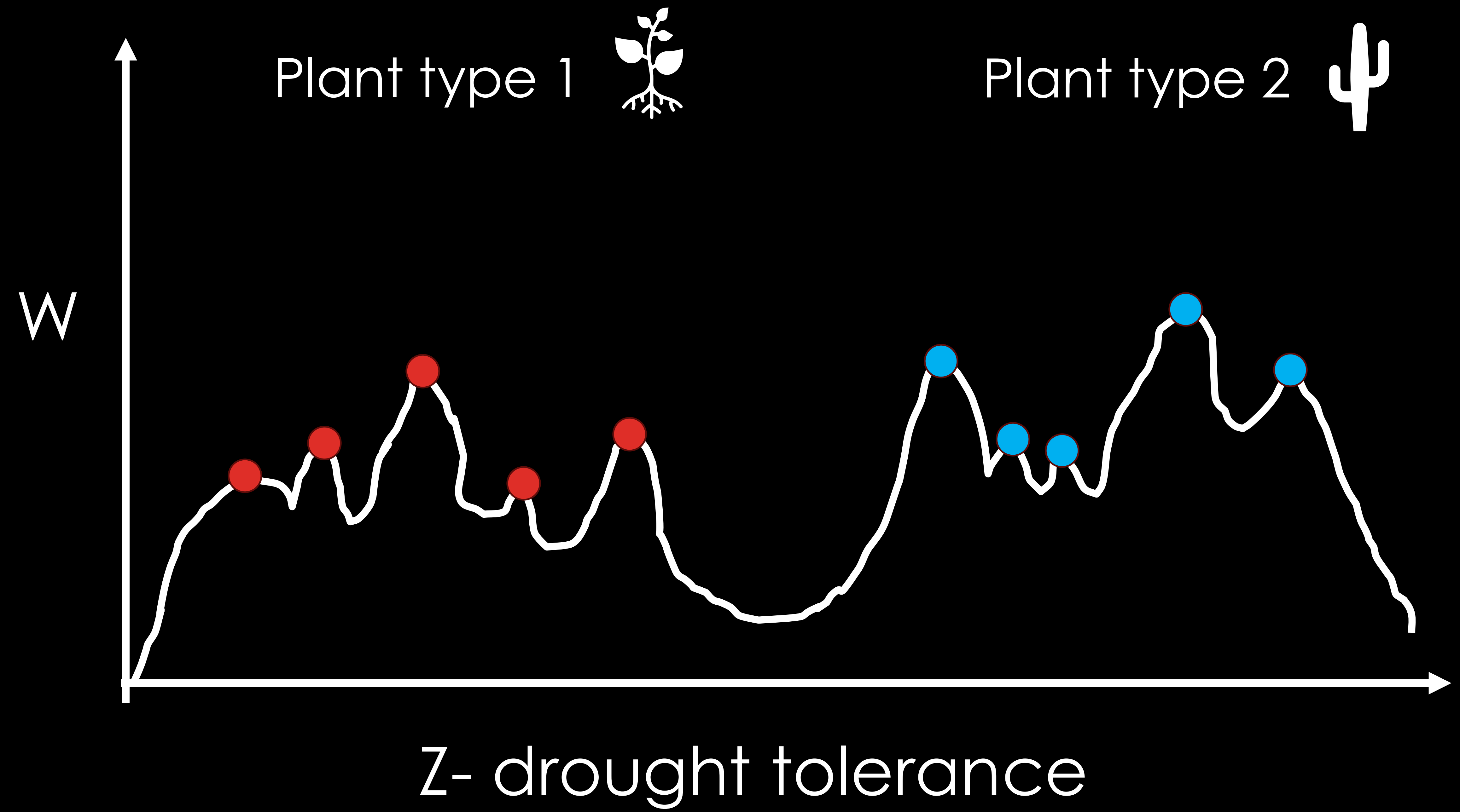
Rate of stochastic evolution

All are scalars= single value

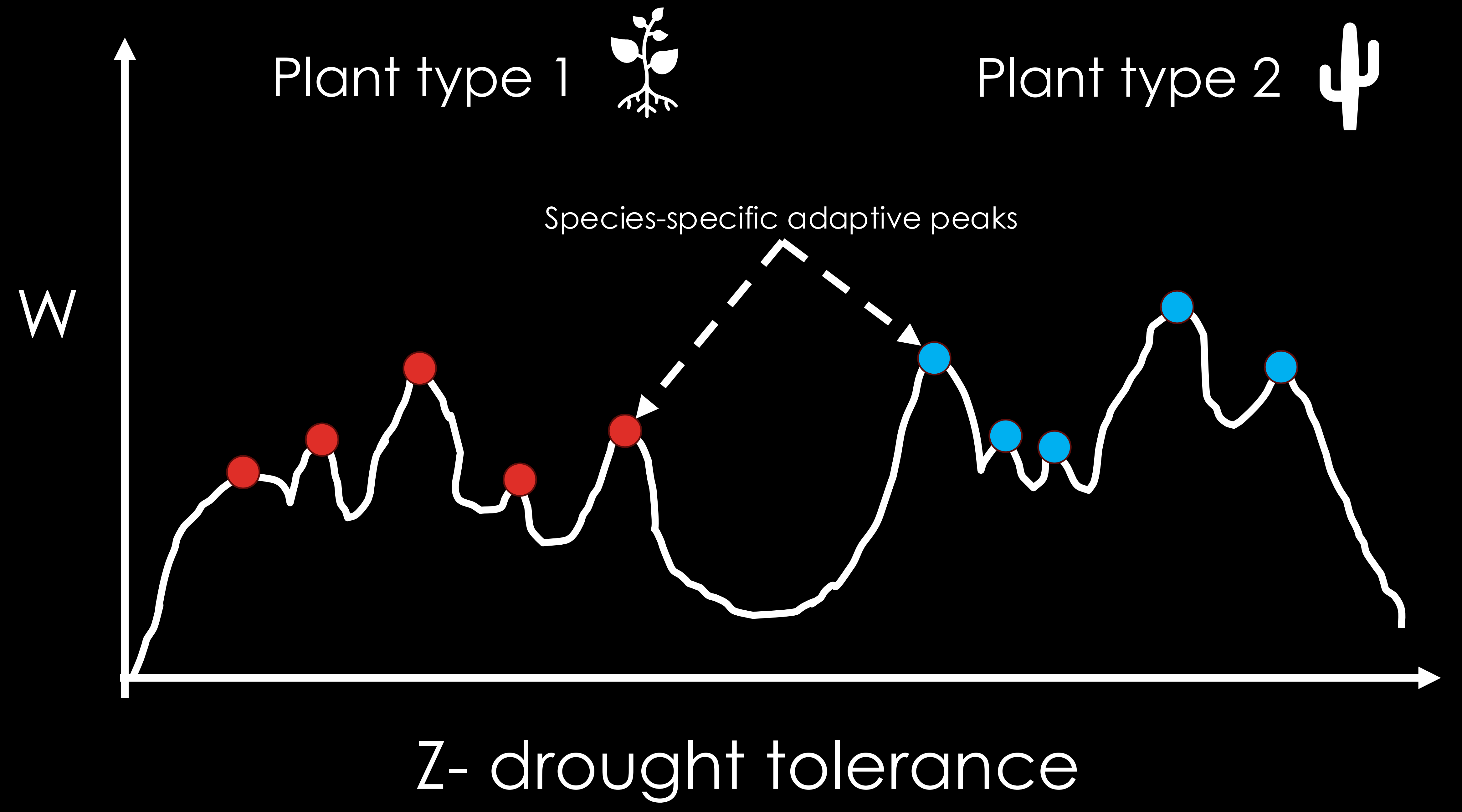
What is a "primary optima"?

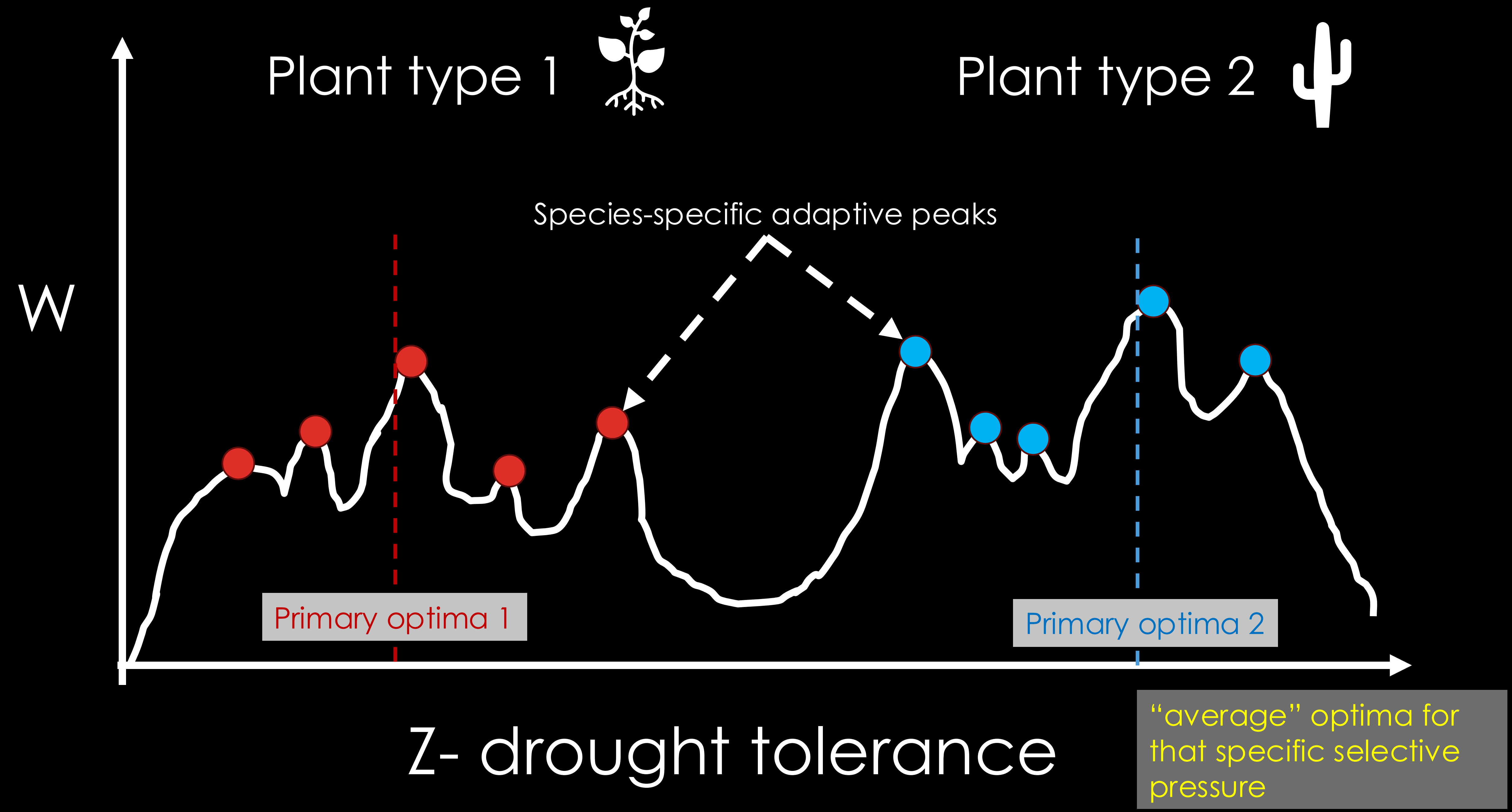
Hansen 1997


W
Z



Plant type 1
Plant type 2
W
Z- drought tolerance


Plant type 1
Plant type 2
Species-specific adaptive peaks
W
Z- drought tolerance


Plant type 1
Plant type 2
Species-specific adaptive peaks
W
Primary optima 1
Primary optima 2
Z- drought tolerance
"average" optima for that specific selective pressure

## Univariate Case

$$dy = -\alpha(y(t) - \theta(t))dt + \sigma dW$$

Rate of adaptation

Primary optima

Rate of stochastic evolution

All are scalars= single value

$$t_{1/2} = \frac{\log 2}{\alpha}$$

Phylogenetic half life

Hansen 1997

## Univariate Case

$$dy = -\alpha(y(t) - \theta(t))dt + \sigma dW$$

- $\alpha$ → Rate of adaptation
- $\theta(t)$ → Primary optima
- $\sigma$ → Rate of stochastic evolution

All are scalars= single value

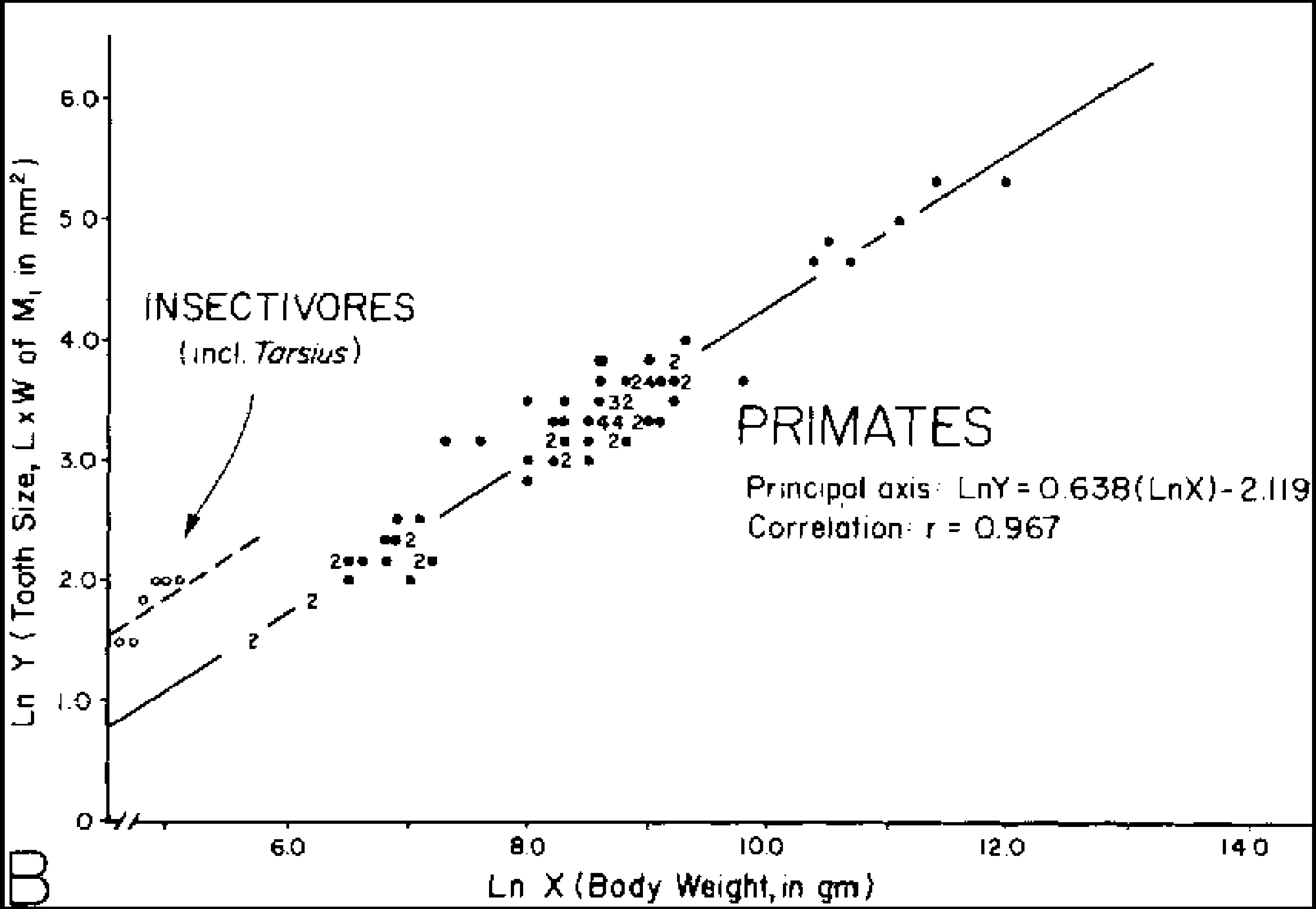
$$t_{1/2} = \frac{\log 2}{\alpha}$$

- $t_{1/2}$ → Phylogenetic half life

What is the simplest multivariate system you can imagine?

Hansen 1997

# Bivariate analysis




$y = a + bx$

slope

intercept

Gingrinch et al 1982

Univariate Case

$$dy = -\alpha(y(t) - \theta(t))dt + \sigma dW$$

Rate of adaptation

Primary optima

Rate of stochastic evolution

All are scalars= single value

Linear regression

$$y(x) = a + bx$$

How would you merge both equations?

## Univariate Case

$$dy = -\alpha(y(t) - \theta(t))dt + \sigma dW$$

Rate of adaptation

Primary optima

Rate of stochastic evolution

All are scalars= single value

## Regression Case

$$dy = -\alpha(y(t) - \theta(x))dt + \sigma_y dW$$

Rate of adaptation

Rate of stochastic evolution of y

Primary optima as a function of x

Linear regression

$$\theta(x) = a + bx$$

$$dx = \sigma_x dW$$

Rate of stochastic evolution of x

Hansen et al 2008

## Univariate Case

$$dy = -\alpha(y(t) - \theta(t))dt + \sigma dW$$

Rate of adaptation

Primary optima

Rate of stochastic evolution

All are scalars= single value

## Regression Case

$$dy = -\alpha(y(t) - \theta(x))dt + \sigma_y dW$$

Rate of adaptation

Rate of stochastic evolution of y

Primary optima as a function of x

Linear regression

$$\theta(x) = a + bx$$

$$dx = \sigma_x dW$$

Rate of stochastic evolution of x

The optimum for y is a function of x, which itself evolves stochastically

Hansen et al 2008

Regression Case

$$dy = -\alpha(y(t) - \theta(x))dt + \sigma_y dW$$

Rate of adaptation

Rate of stochastic evolution of y

Primary optima as a function of x

$$dx = \sigma_x dW$$

Rate of stochastic evolution of x

$$E[y|x] = k + \left(1 - \frac{1 - e^{-\alpha t}}{\alpha t}\right) bx$$

Hansen et al 2008

Regression Case

$$dy = -\alpha(y(t) - \theta(x))dt + \sigma_y dW$$

Rate of adaptation

Rate of stochastic evolution of y

Primary optima as a function of x

$$dx = \sigma_x dW$$

Rate of stochastic evolution of x

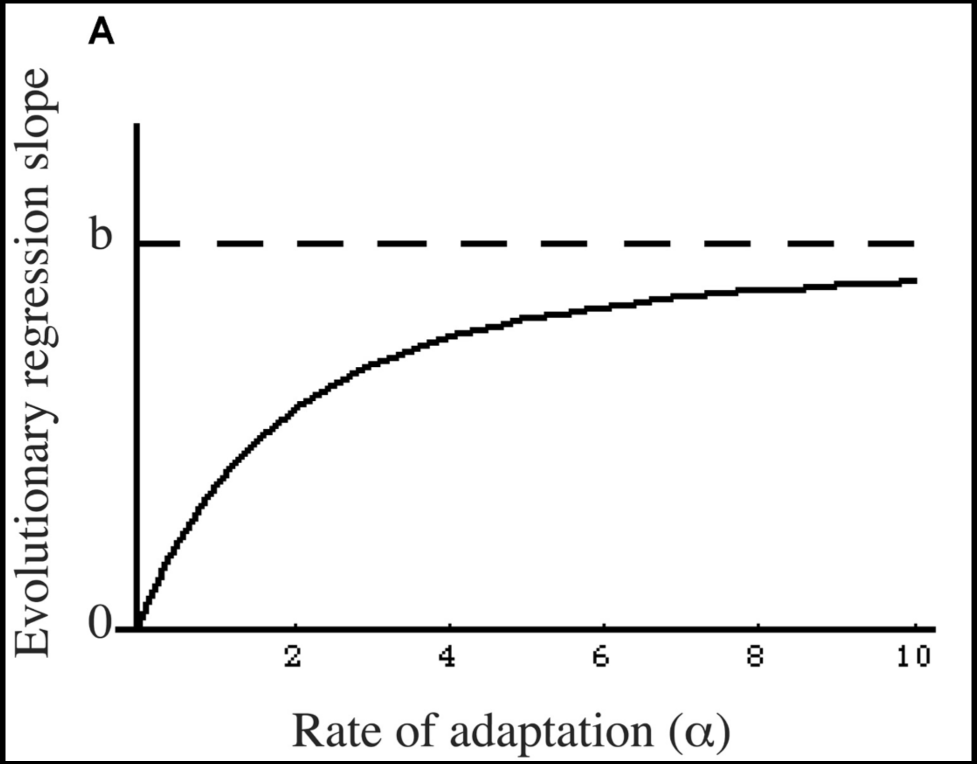
$$E[y|x] = k + \left(1 - \frac{1 - e^{-\alpha t}}{\alpha t}\right) bx$$

Optimal regression slope

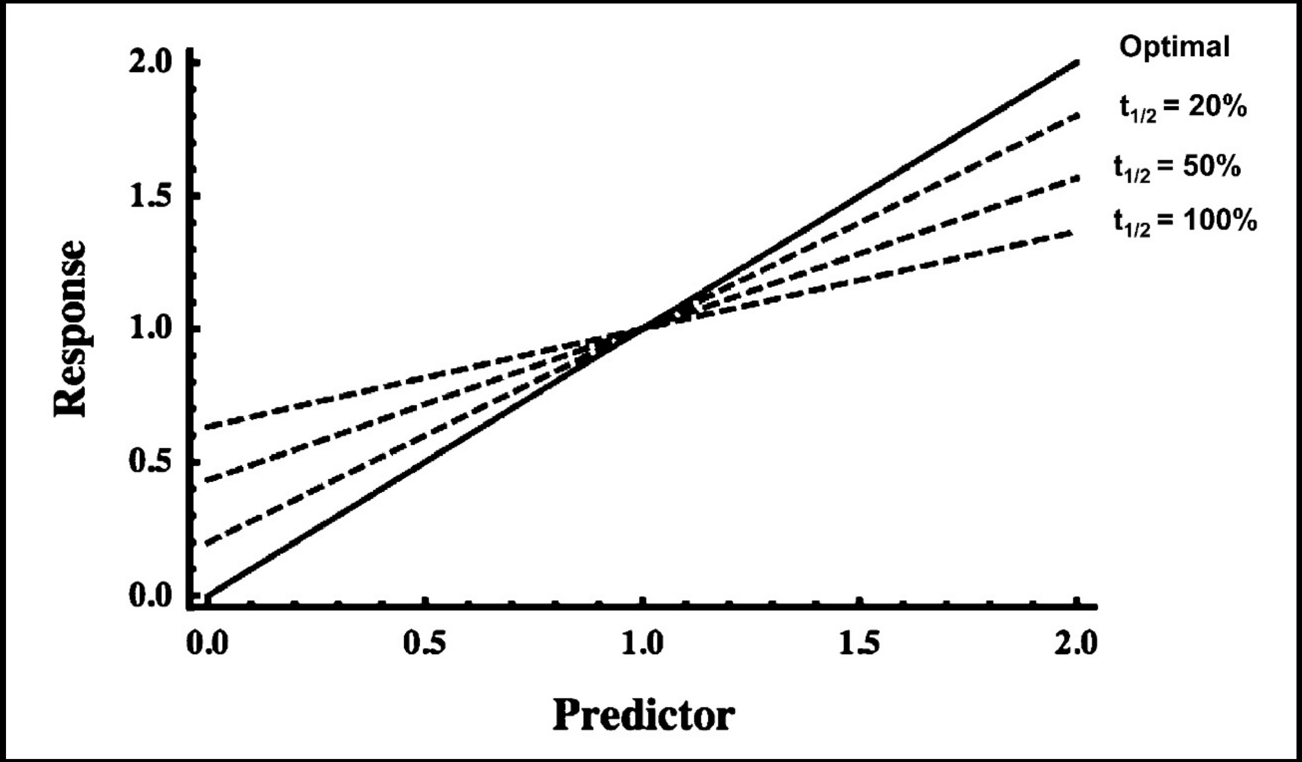
Phylogenetic correction factor

Hansen et al 2008

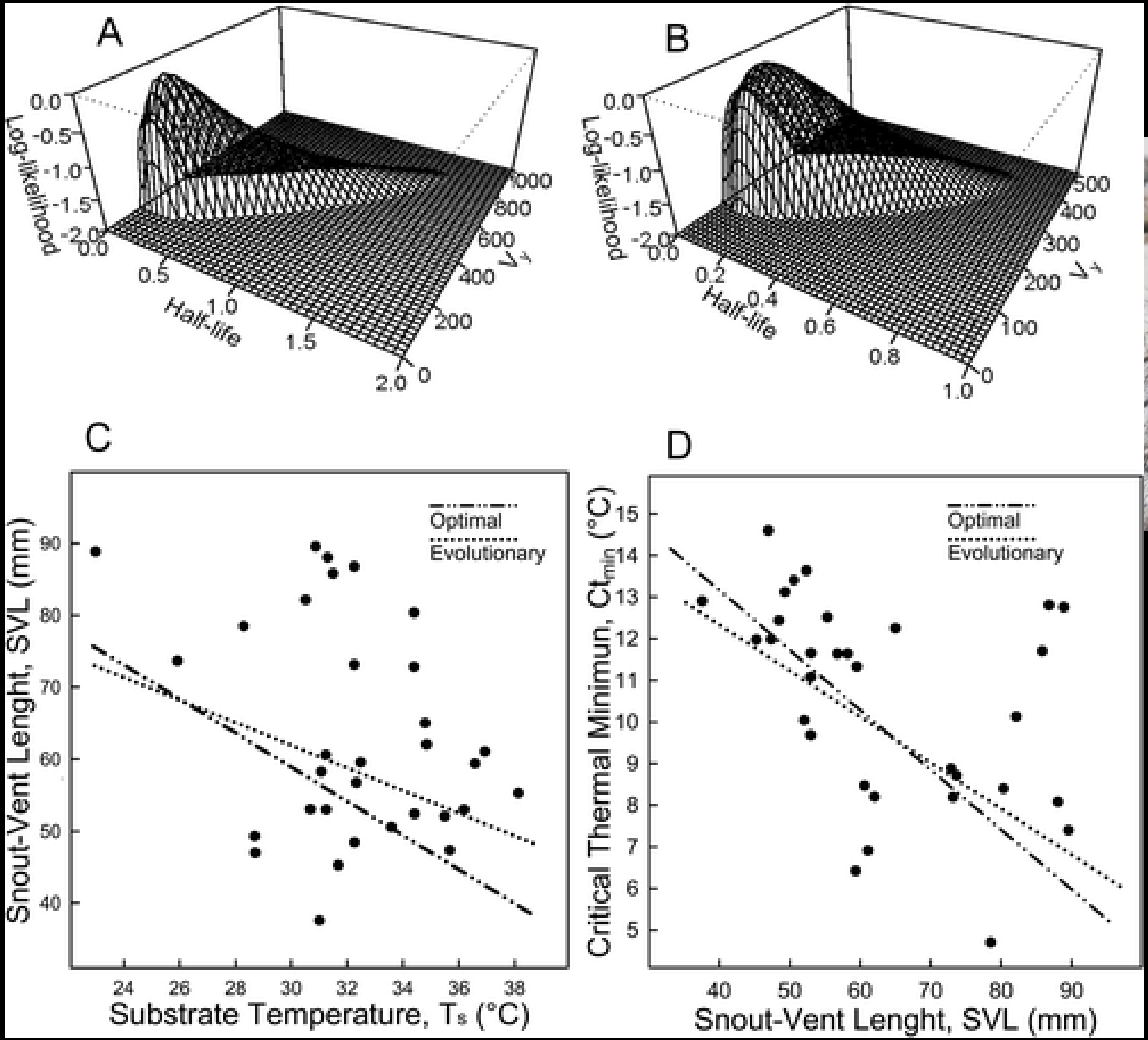
Hansen et al 2008




Hansen & Bartoszek 2012




The faster the rate of adaptation, the shallower the slope will be

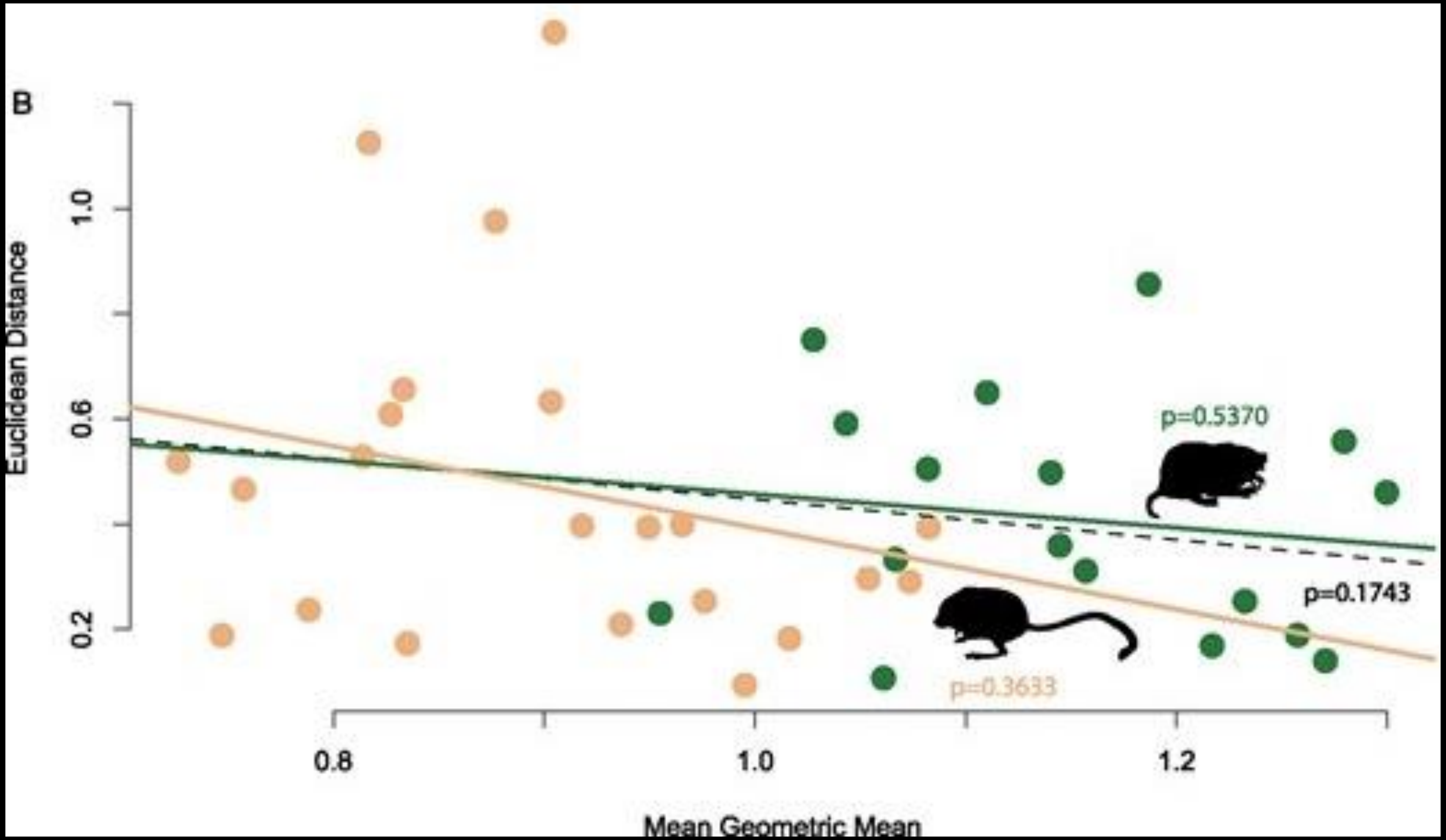



*Liolaemus*

What do you think this means?

Labra et al. 2009

Rensch's rule




Geomydae

Heteromydae

Calede&Brown, 2023

Regression Case

$$dy = -\alpha(y(t) - \theta(x))dt + \sigma_y dW$$

Rate of adaptation

Primary optima as a function of x

Rate of stochastic evolution of y

$$dx = \sigma_x dW$$

Rate of stochastic evolution of x

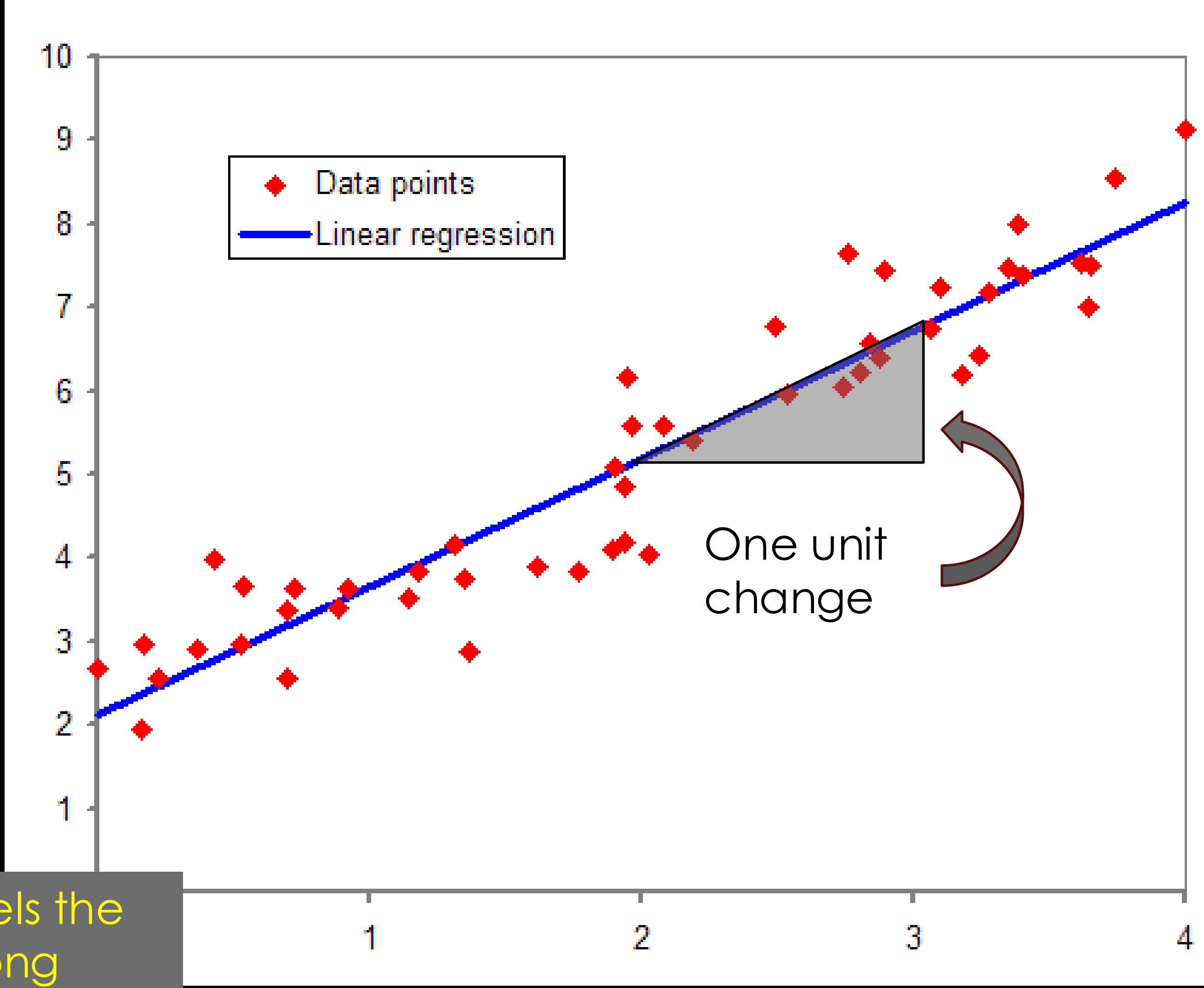
$$E[y|x] = k + \left(1 - \frac{1 - e^{-\alpha t}}{\alpha t}\right) bx$$

Phylogenetic correction factor

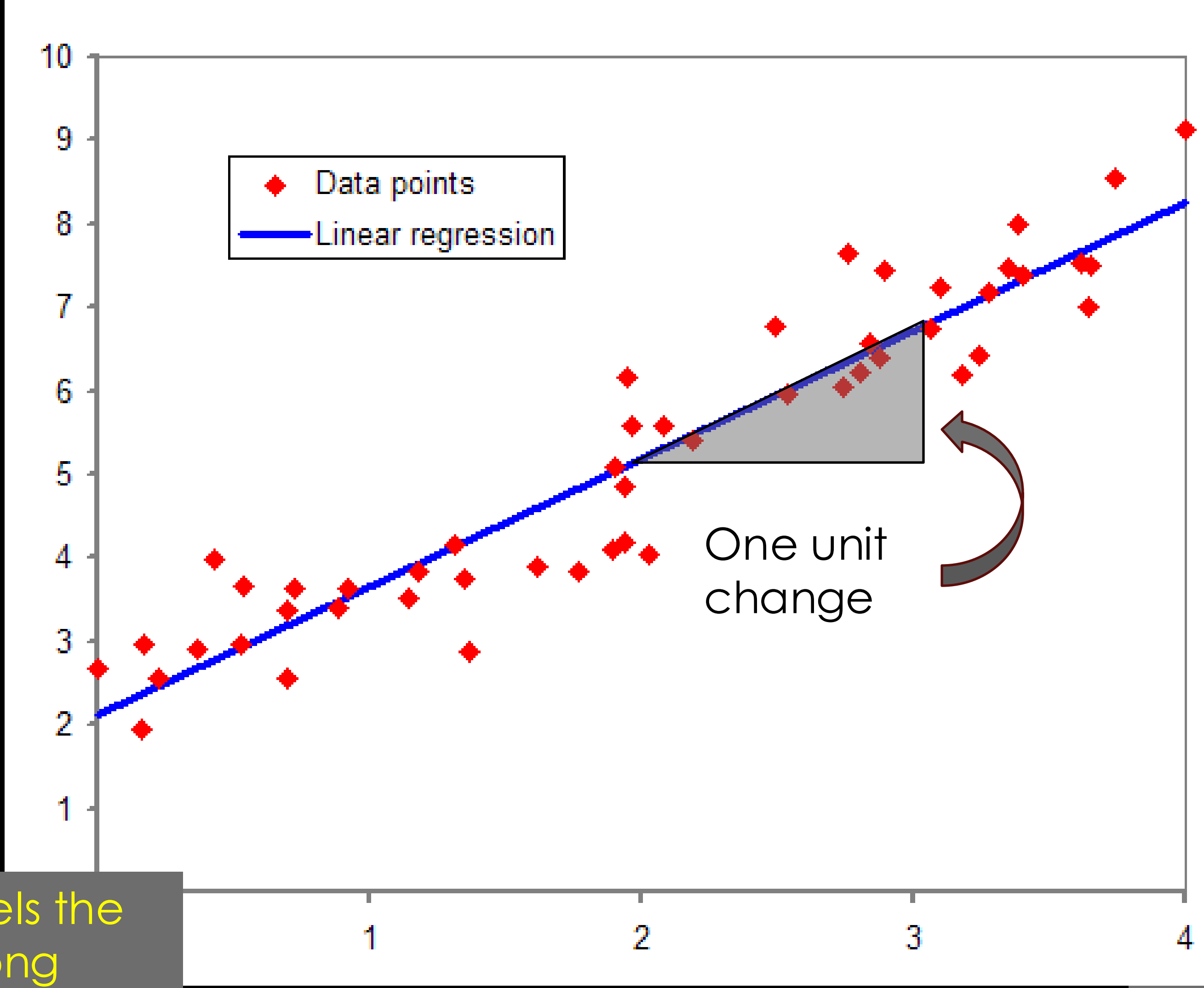
Optimal regression slope

This models the relationship among traits instantaneously

Hansen et al 2008




Regression models the relationship among traits instantaneously




Regression models the relationship among traits instantaneously

Alternatives?

# Gs are multivariate analogs of $\sigma_a^2$

$$G = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}$$

What about macroevolutionary parameters?

Univariate Case

$$dy = -\alpha(y(t) - \theta(t))dt + \sigma dW$$

Rate of adaptation

Primary optima

Rate of stochastic evolution

All are scalars= single value

## Univariate Case

$$dy = -\alpha(y(t) - \theta(t))dt + \sigma dW$$

- $\alpha$: Rate of adaptation
- $\theta(t)$: Primary optima
- $\sigma$: Rate of stochastic evolution

All are scalars= single value

## Multivariate Case

$$dy = -\boldsymbol{A}(\boldsymbol{y}(t) - \boldsymbol{\theta}(t))dt + \boldsymbol{\Sigma} d\boldsymbol{W}$$

- $\boldsymbol{A}$: Matrix of rate of adaptation
- $\boldsymbol{\theta}(t)$: Vector of primary optima
- $\boldsymbol{\Sigma}$: Rate matrix of stochastic evolution (Cholesky factor)

Multivariate Case

$$dy = -\boldsymbol{A}(\boldsymbol{y}(t) - \boldsymbol{\theta}(t))dt + \boldsymbol{\Sigma} d\boldsymbol{W}$$

Matrix of rate of adaptation

Vector of primary optima

Rate matrix of stochastic evolution (Cholesky factor)

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}$$

$\sigma_x^2 \rightarrow$ rate of stochastic evolution

$\sigma_{xy} \rightarrow$ rate of stochastic coevolution

symmetric

$$\boldsymbol{\theta} = \begin{matrix} \theta_x \\ \theta_y \end{matrix}$$

$\theta_i \rightarrow$ Primary optima for each trait i

Multivariate Case

$$dy = -\boldsymbol{A}(\boldsymbol{y}(t) - \boldsymbol{\theta}(t))dt + \boldsymbol{\Sigma} d\boldsymbol{W}$$

Matrix of rate of adaptation

Vector of primary optima

Rate matrix of stochastic evolution (Cholesky factor)

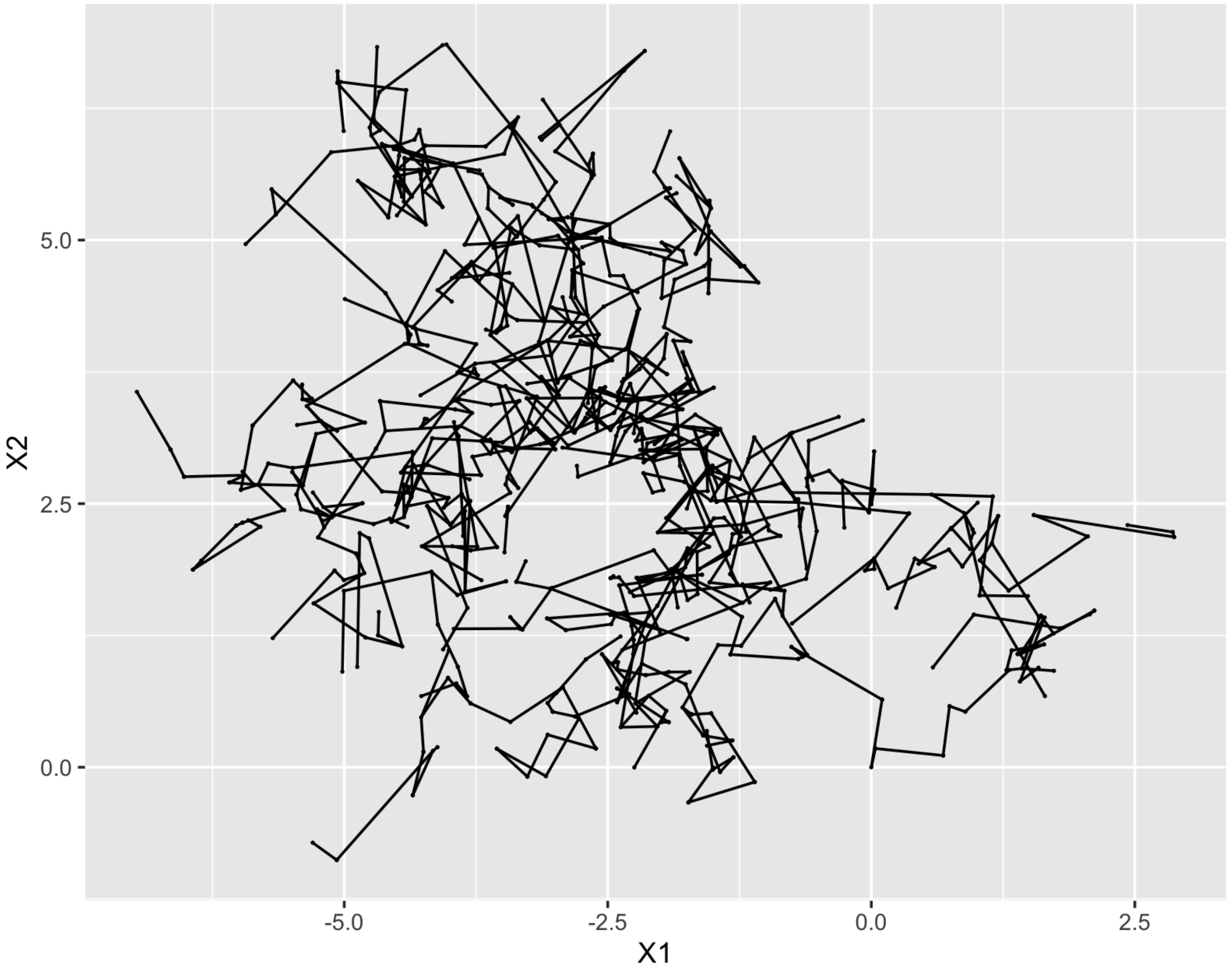
$$\boldsymbol{A} = \begin{bmatrix} \alpha_x & \alpha_{xy} \\ \alpha_{yx} & \alpha_y \end{bmatrix}$$

$\alpha_i \rightarrow$ rate of adaptation for trait i

$\alpha_{xy} \rightarrow$ rate of y tracking x

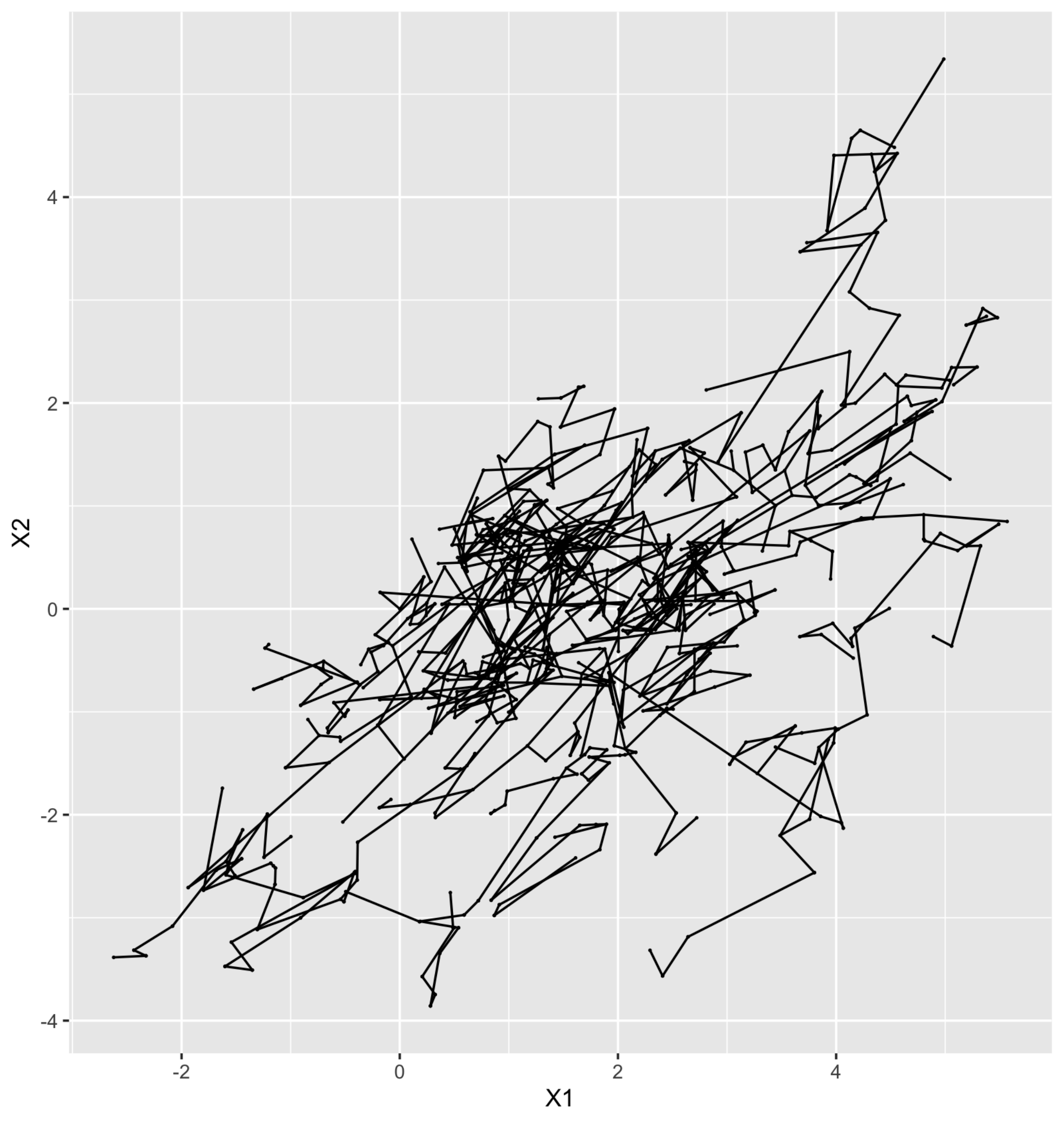
$\alpha_{yx} \rightarrow$ rate of x tracking y

Not-symmetric



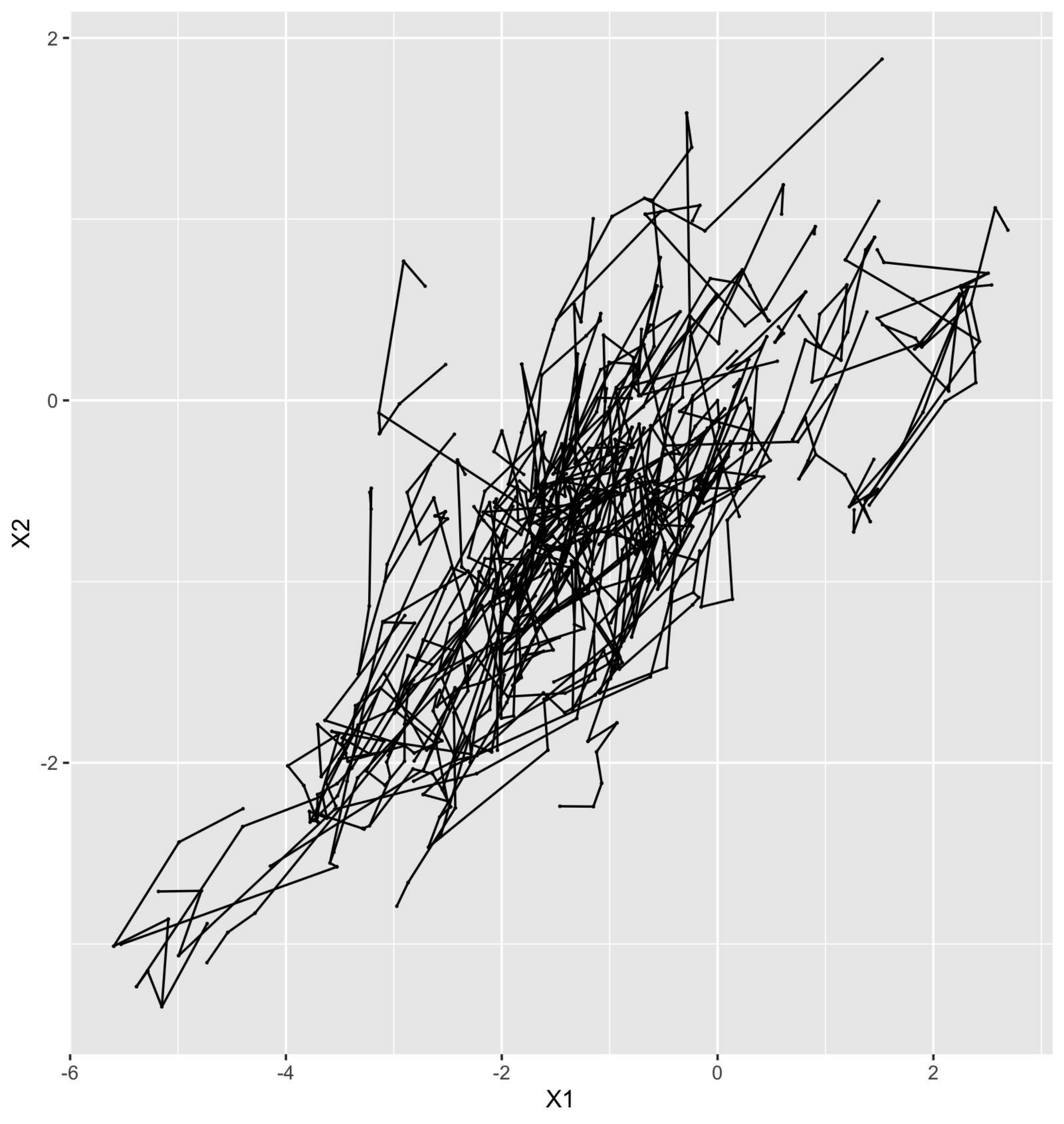

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\boldsymbol{A} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$



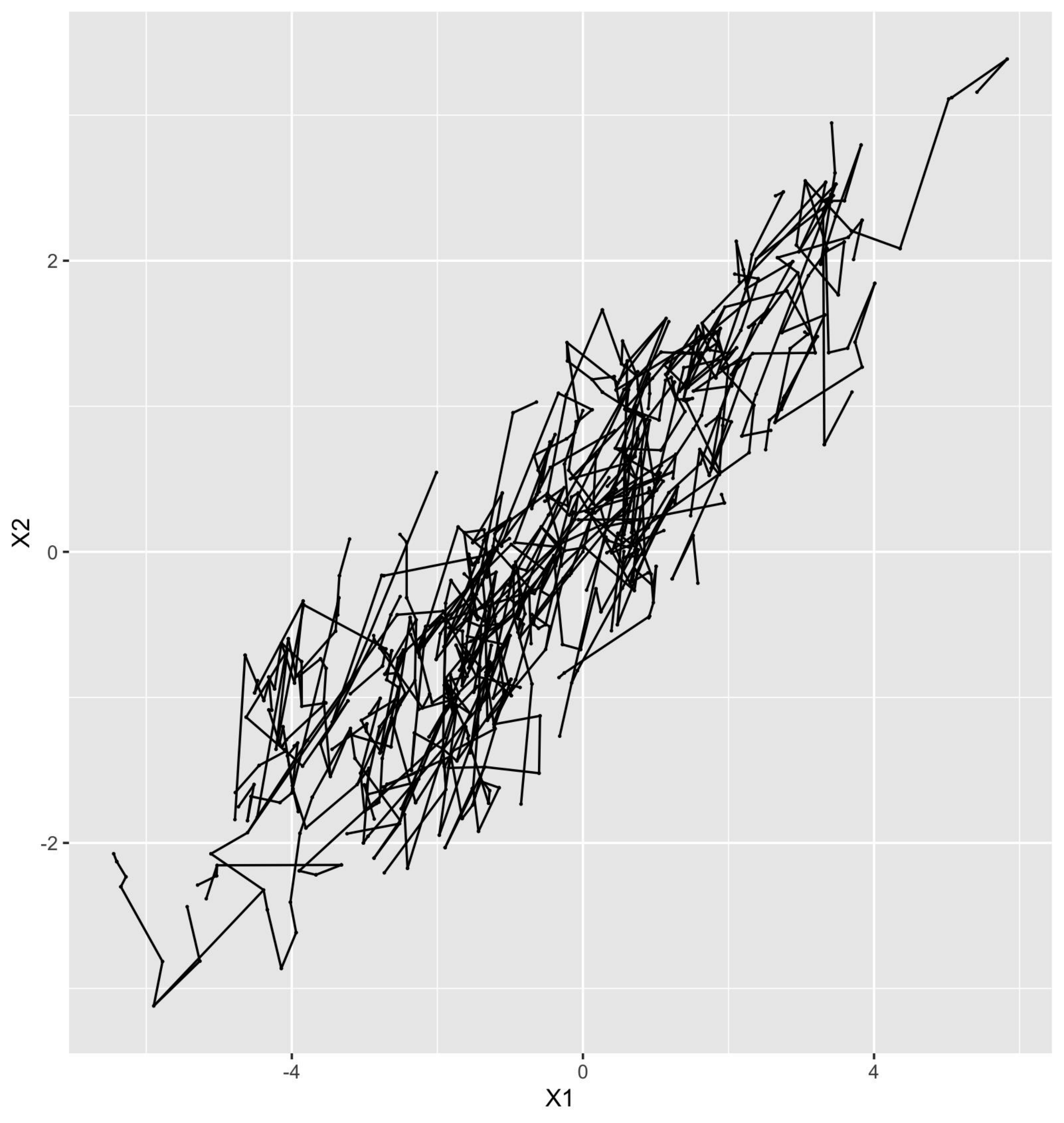

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix}$$

$$\boldsymbol{A} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$



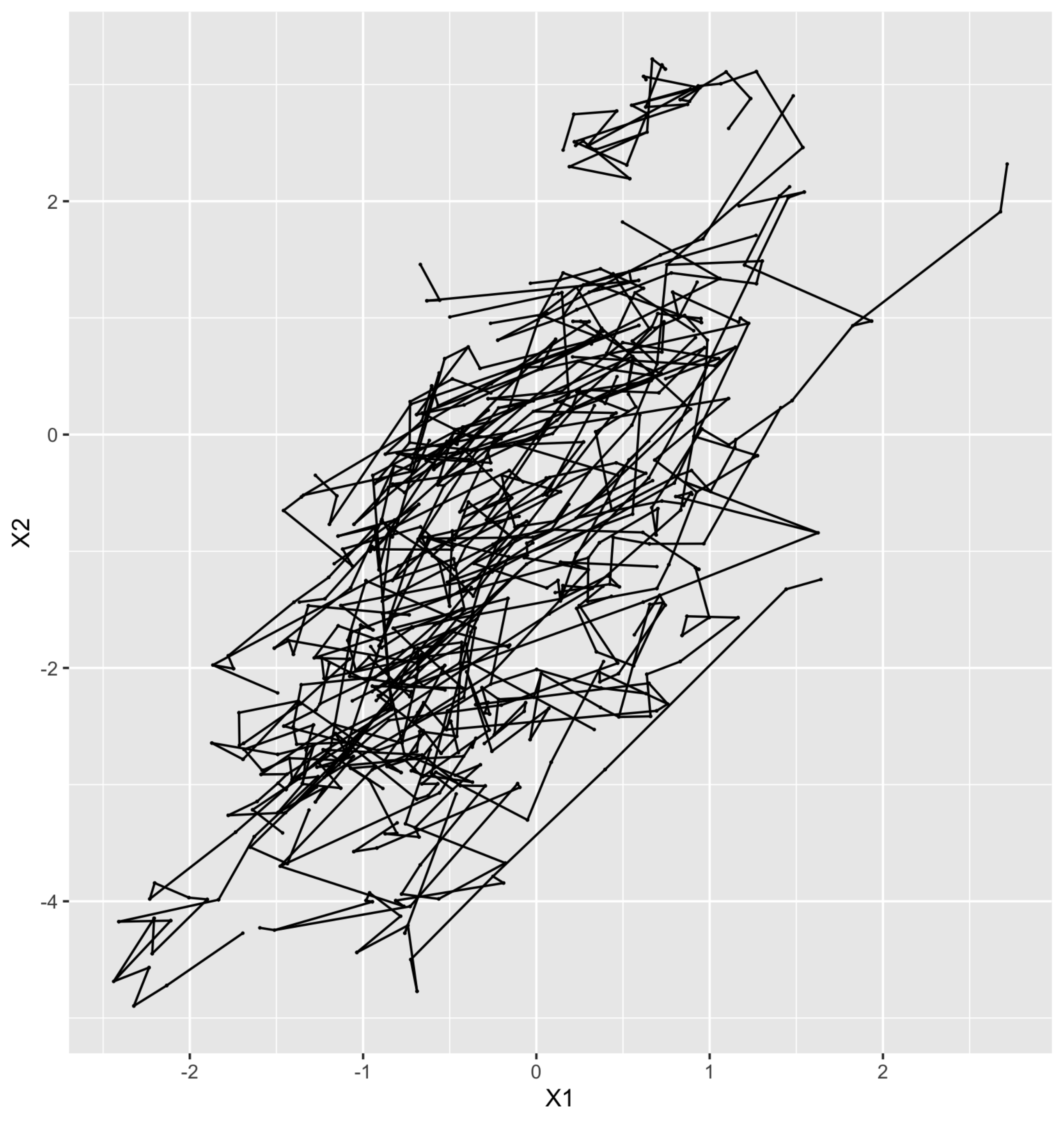

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix}$$

$$\boldsymbol{A} = \begin{bmatrix} 0.3 & 0 \\ 0 & 0.3 \end{bmatrix}$$



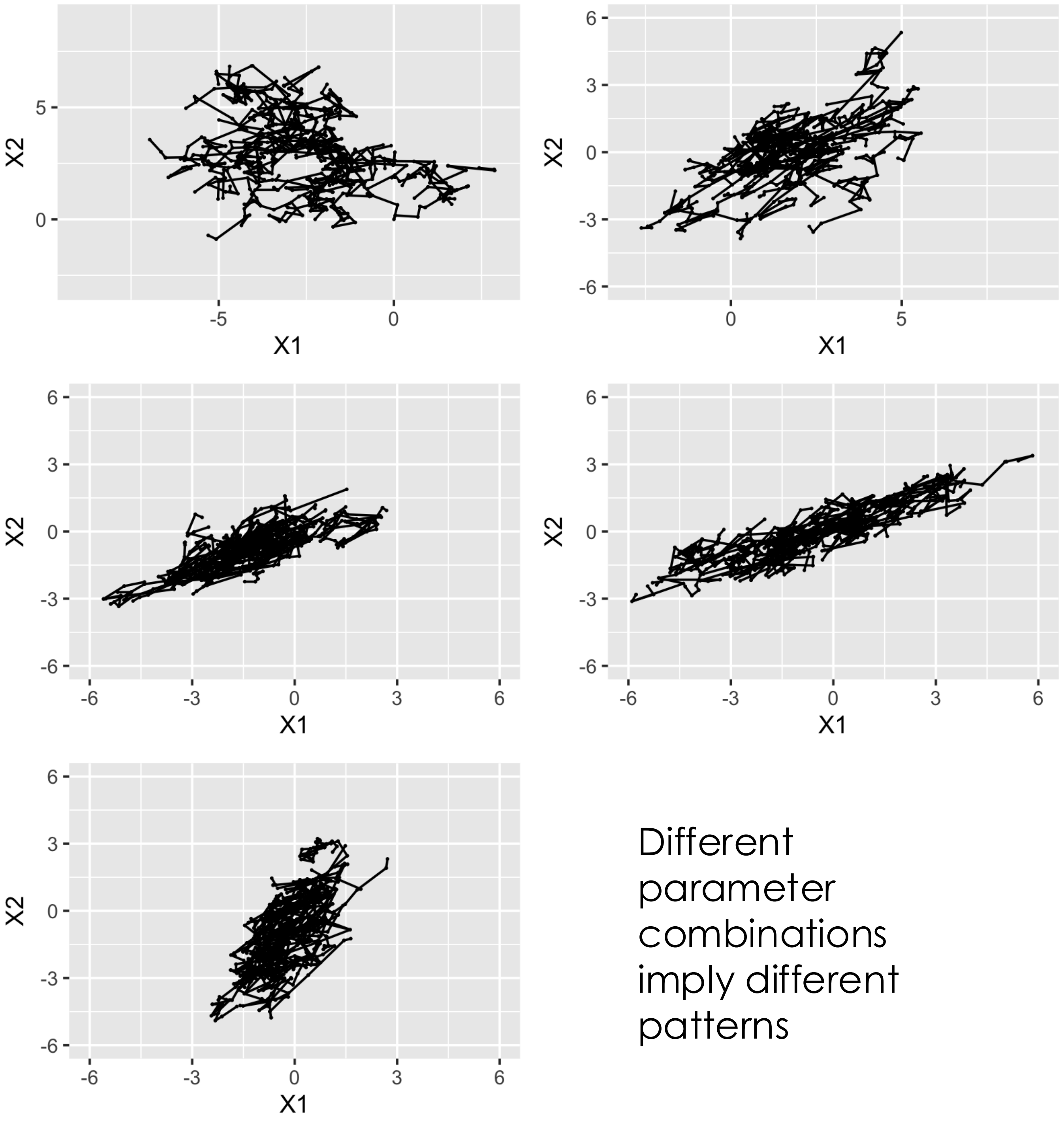

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix}$$

$$\boldsymbol{A} = \begin{bmatrix} 0.3 & -1 \\ -0.001 & 0.3 \end{bmatrix}$$




$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix}$$

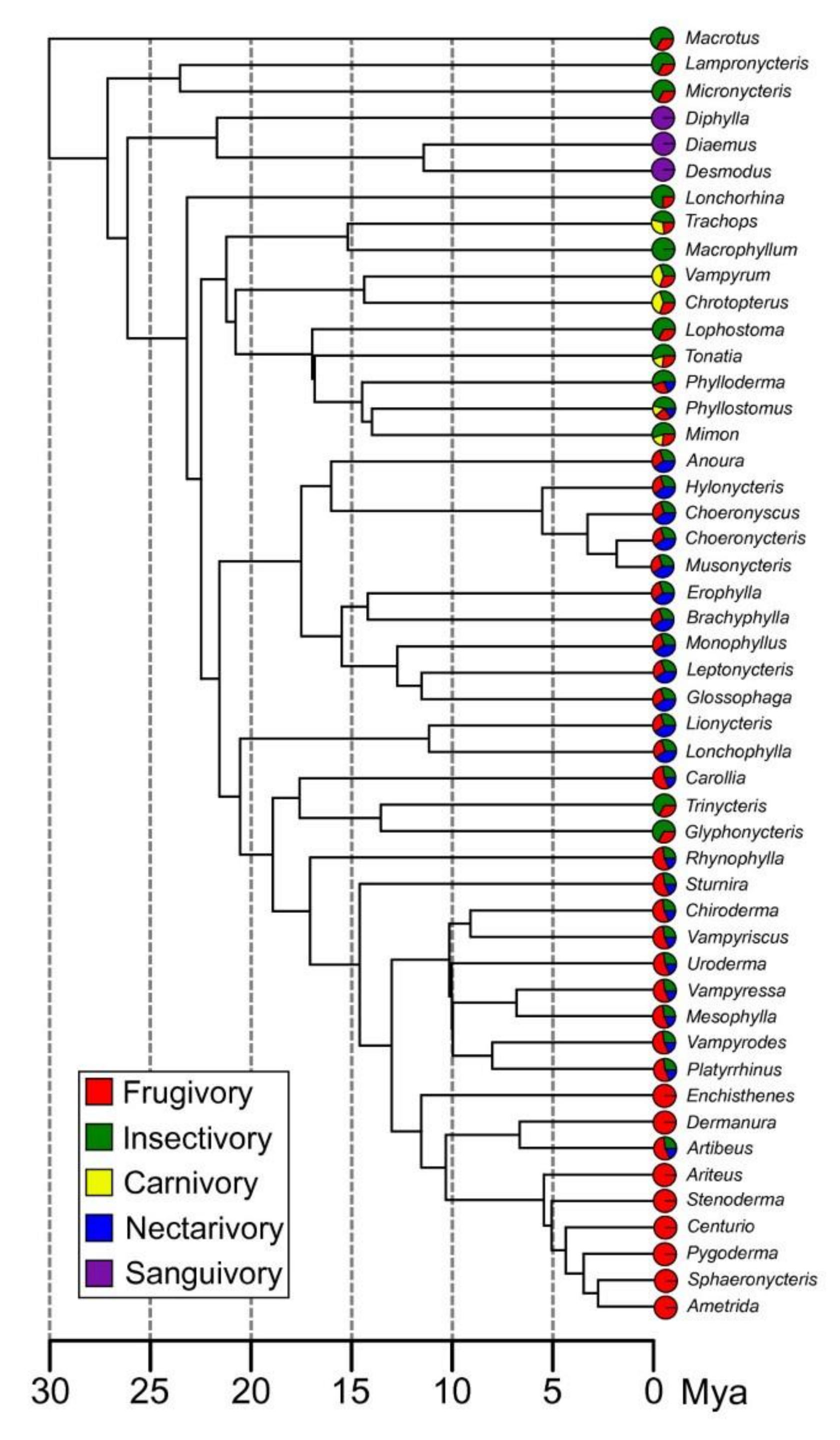
$$\boldsymbol{A} = \begin{bmatrix} 0.3 & -0.001 \\ -1 & 0.3 \end{bmatrix}$$

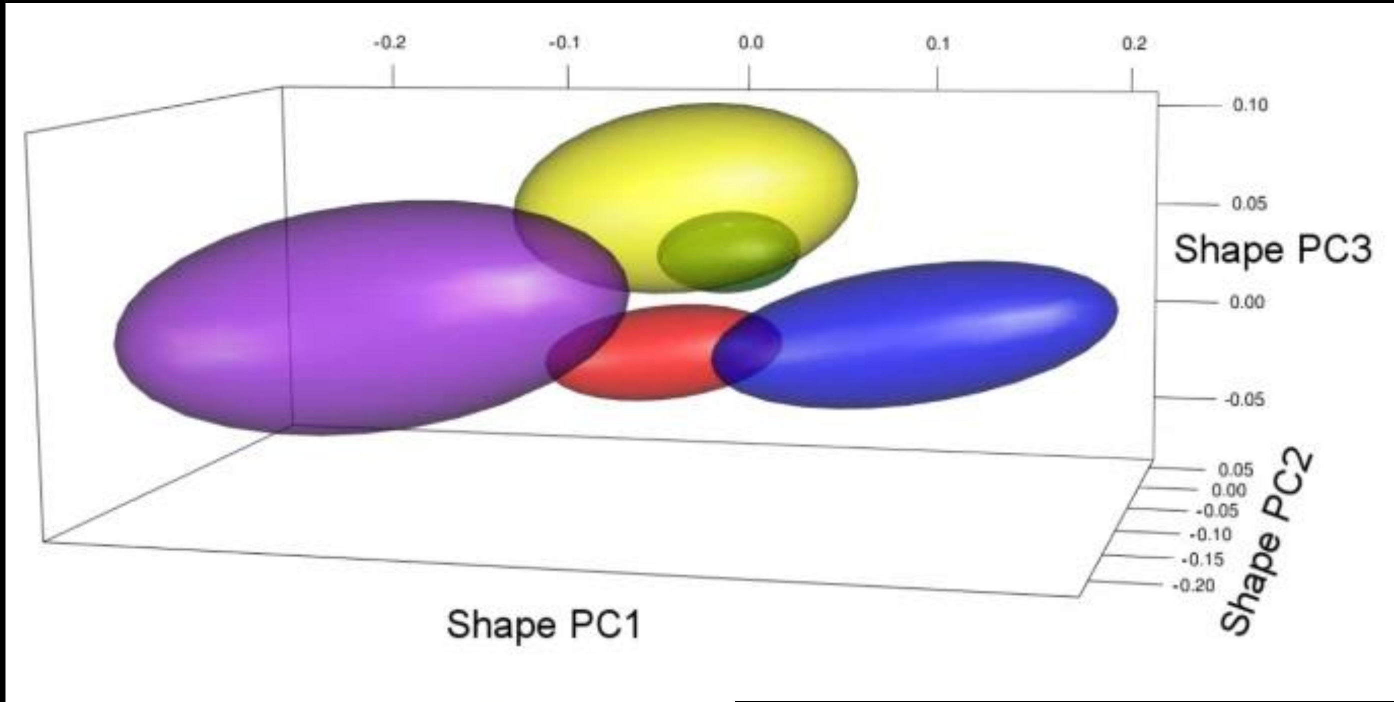


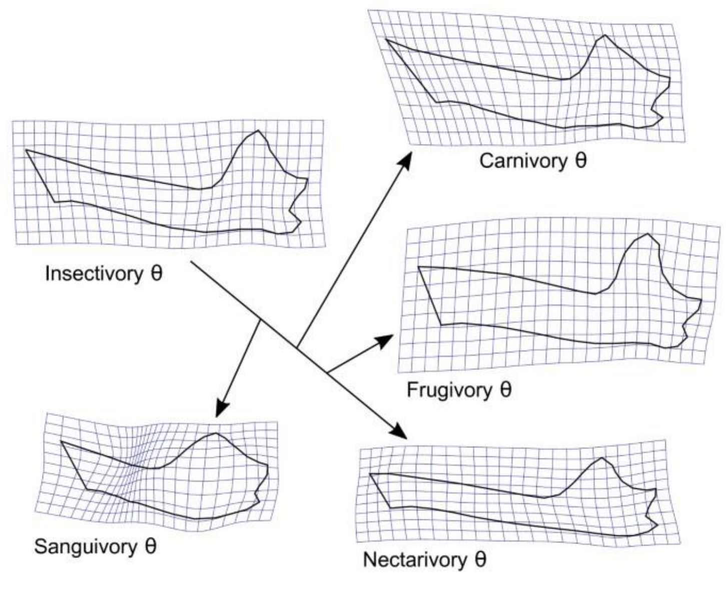

Different parameter combinations imply different patterns

We will fit some of these in the tutorial

# Multivariate approaches also can be fit into multiple regimes



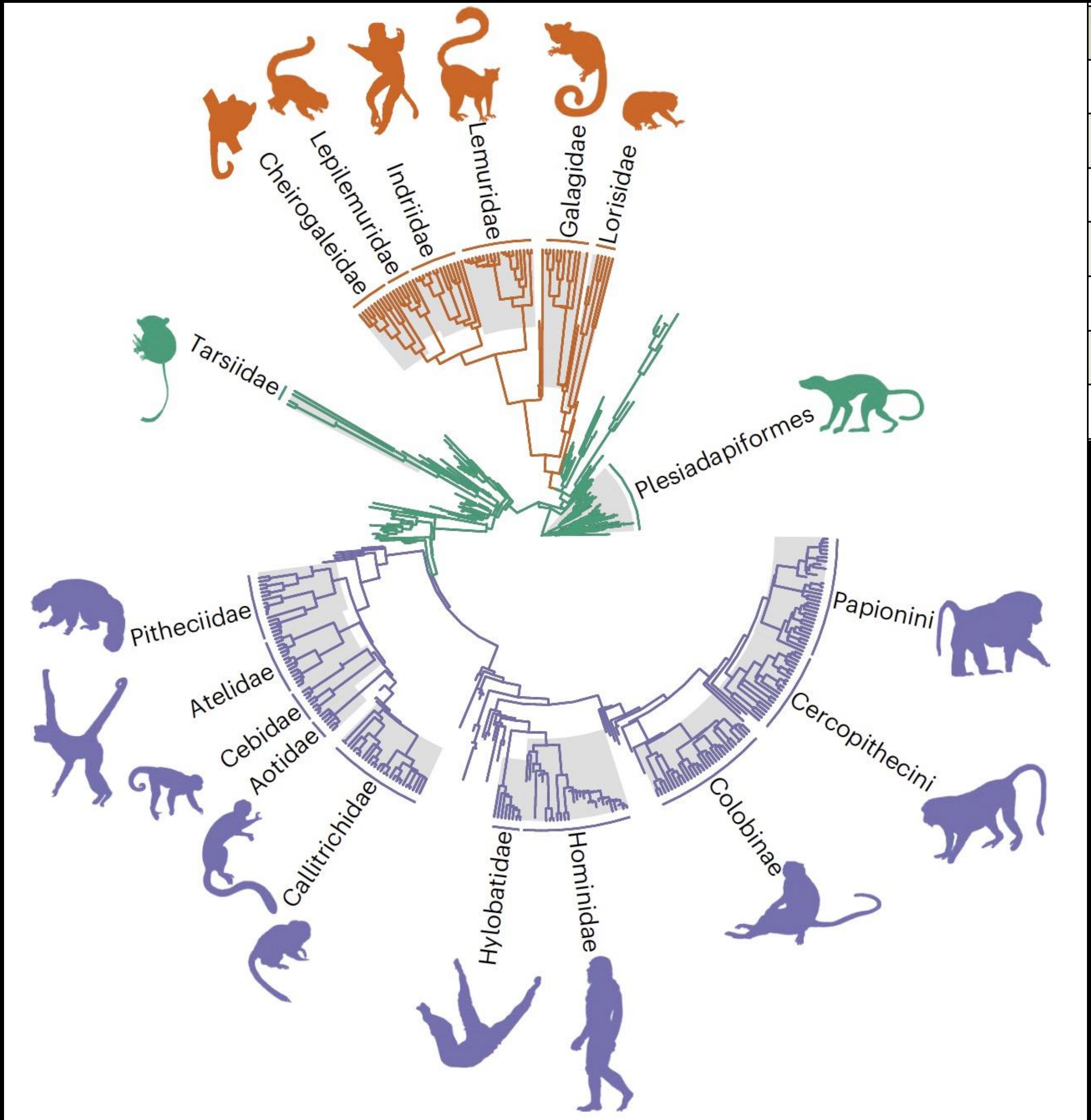




Monteiro & Nogueira 2011

| Package | Discrete | *Continuous fixed* | *Continuous random* | Regime detection |
|---|---|---|---|---|
| *L1OU (Khabbazian et al., 2016)* | X | | | X |
| *MVMORPH (Clavel et al., 2015)* | X | | | |
| *bayou (Uyeda & Harmon, 2014)* | X | X | | X |
| *OUCH (Butler & King, 2004)* | X | | | |
| *PHYLOGENETICEM (Bastide et al., 2017)* | X | | | |
| *SURFACE (Ingram & Mahler, 2013)* | X | | | X |
| *mvSLOUCH (Bartoszek et al., 2012)* | X | X | X | |
| *PCMfit (Mitov et al., 2019)* | X | X | X | X |

PCMfit

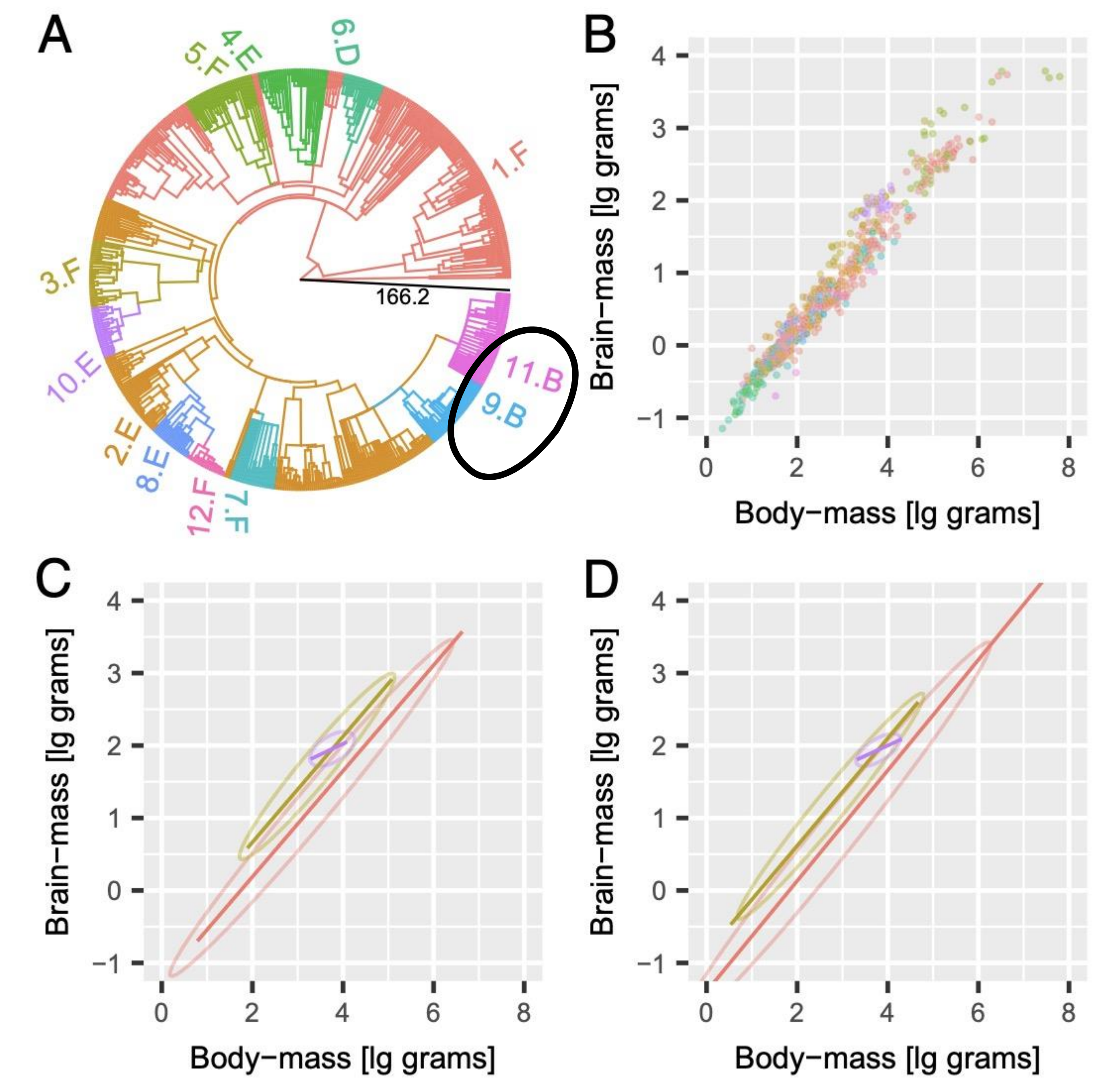



| Model[b] | $N_p$[c] | logLik[d] | BIC[e] |
|---|---|---|---|
| BM | 27 | 2,585.03 | −5,003.38 |
| OU | 54 | 2,684.57 | −5,035.75 |
| $BM_{\Sigma \propto P}$ | 7 | 2,108.57 | −4,173.93 |
| $OU_{\Sigma \propto P}$ | 34 | 2,174.45 | −4,138.99 |
| $BM_{\Sigma \propto G}$ | 7 | 1,480.63 | −2,918.04 |
| $OU_{\Sigma \propto G}$ | 34 | 1,510.57 | −2,811.23 |
| Three-regime BM[f] | 69 | 2,823.03 | **−5,484.50** |

Machado et al. 2023

PCMfit

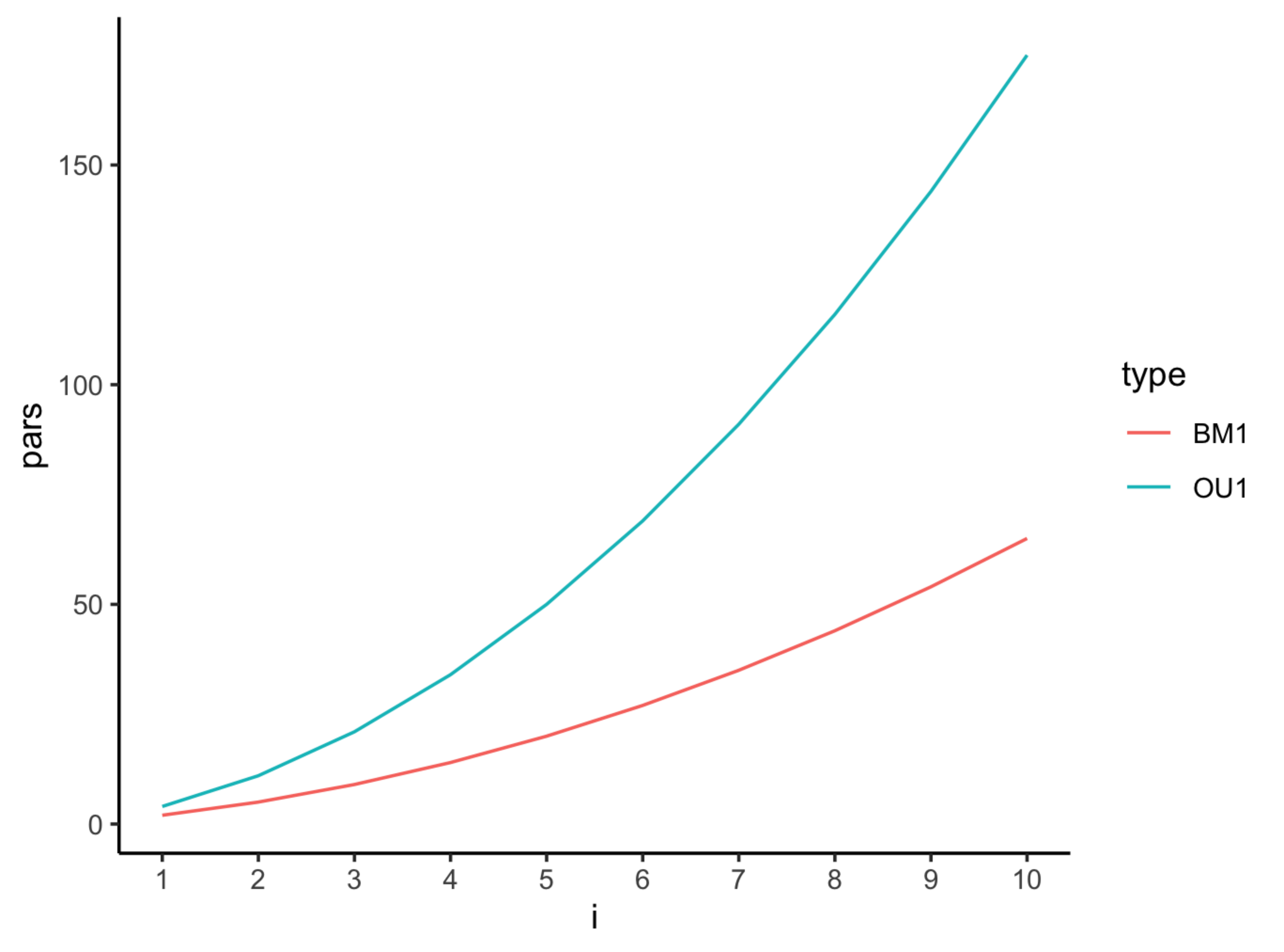
B-Brownian Motion
C-F- OU




Allows for a mixture of models in a single tree

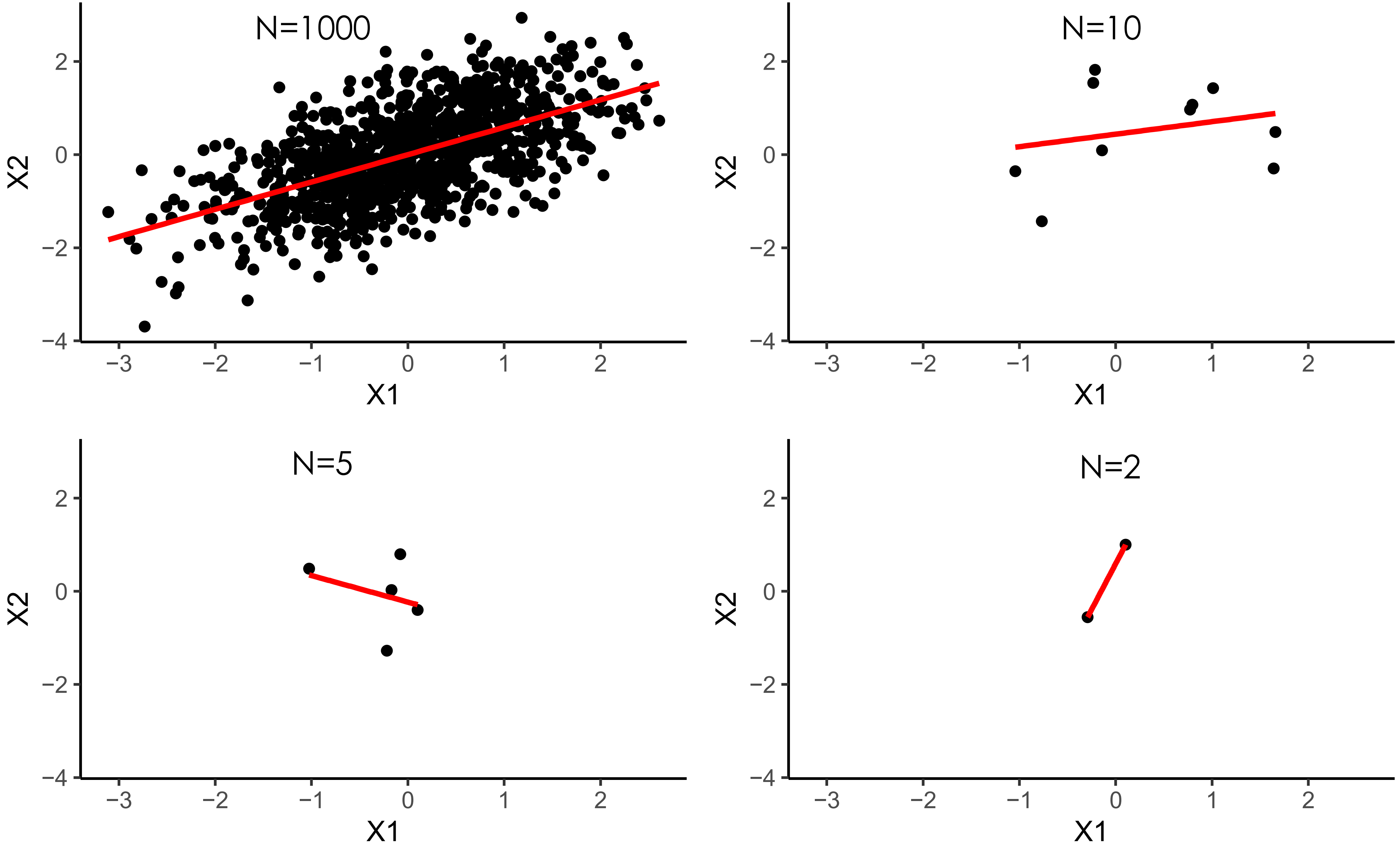
Mitov et al. 2019

# Potential drawbacks of multivariate methods




What should happen to the data?

Parameter number increases with dimensionality of the data


N=1000
N=10
N=5
N=2
X1
X2
−3
−2
−1
0
1
2
−4