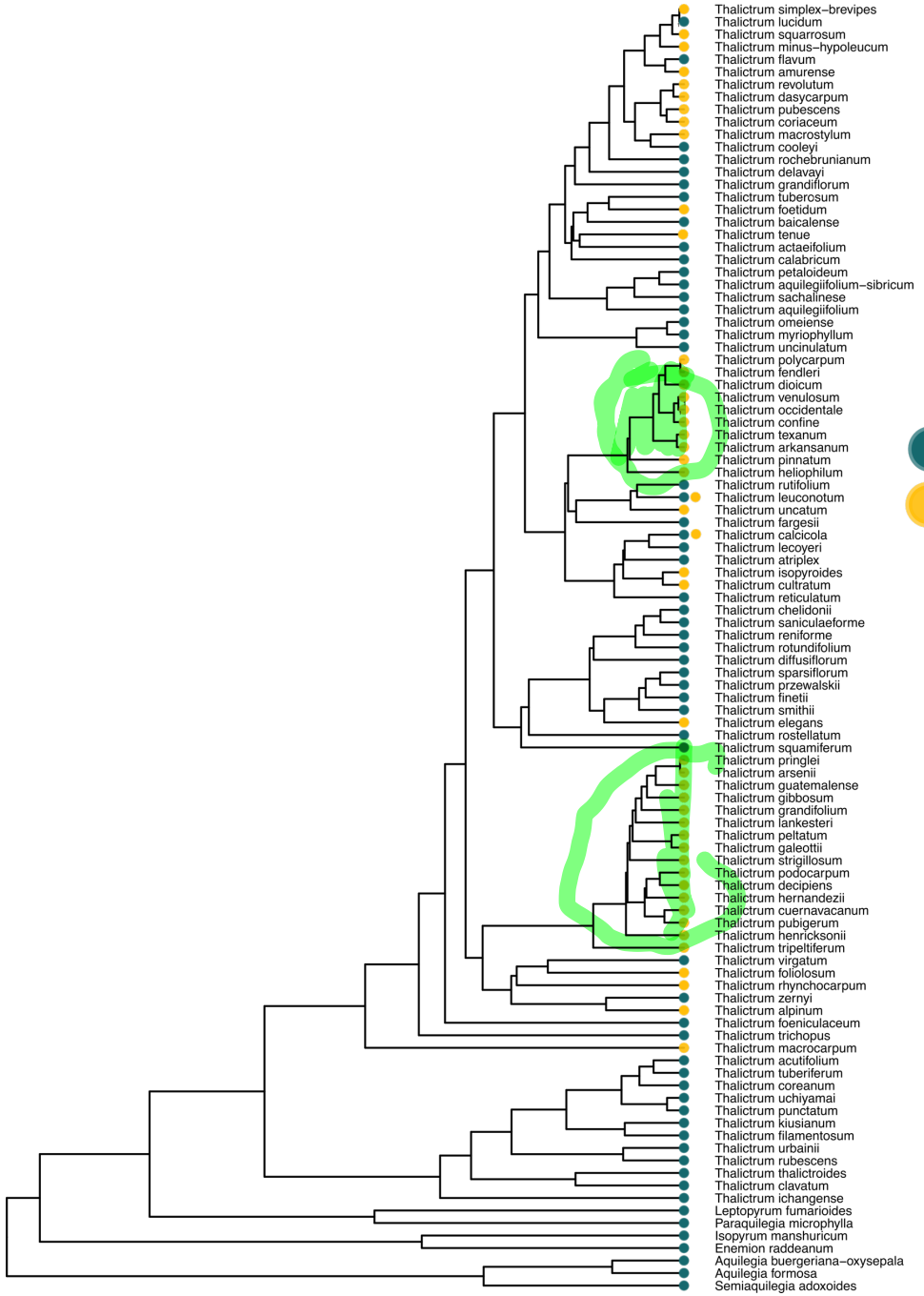
Discrete trait models

Comparing models and correlating traits

BIO508- Rosana Zenil-Ferguson



- Always plot your data!
- Check for nestedness of trait!
- Transitions are the effective sample size for these analyses

Continuous-Time Markov Chains (CTMC)

$$\{X(t), t \geq 0\}$$

Stochastic models that follow change in time with an associated **probability**

$X(t)$= phenotype (trait) value at time t

t= millions of years or expected number of substitutions

Discrete models of evolution

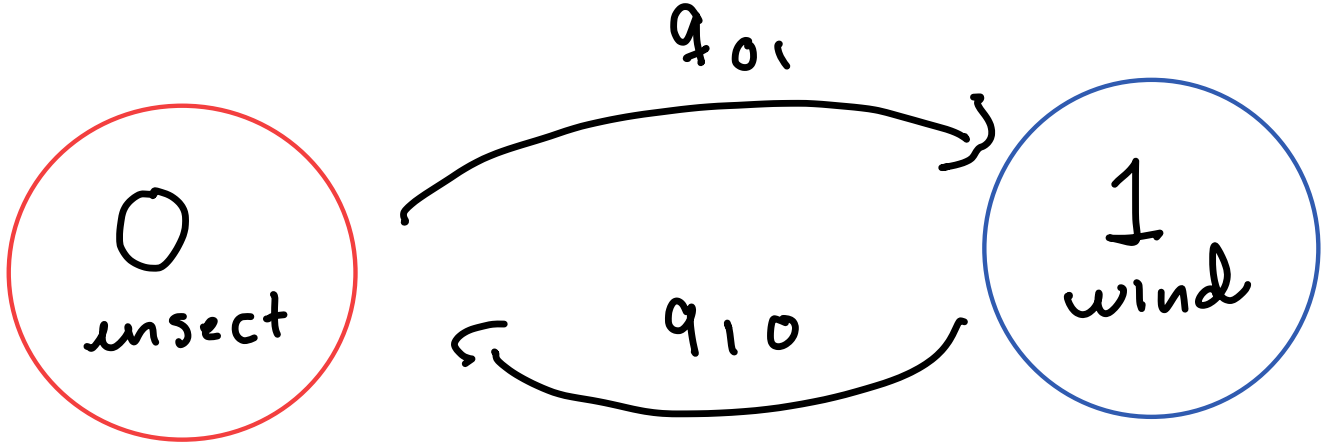
$X(t)$= phenotype (trait) value at time t

$X(t)$= Insect (0), Wind (1)

We use conditional probabilities. For example:

$$P(X(t) = 1 | X(0) = 0)$$

Drawing the FULL model



$$Q = \begin{array}{c} \text{from} \\ \downarrow \\ 0 \\ 1 \end{array} \begin{array}{c} \text{to} \rightarrow \\ \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} -q_{01} & q_{01} \\ q_{10} & -q_{10} \end{pmatrix} \end{array}$$

Full-model

$$\begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$$

Fit MK in R

model

q_{01} is parameter 1

q_{10} is parameter 2

Estimation : Full model

$q_{01}^{MLE} = 0.05$ Change per lineage per 1 my (Events)

E / L / 1 my

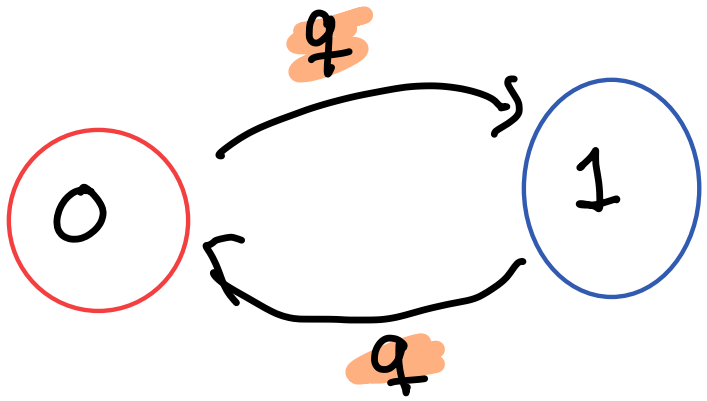
$q_{10}^{MLE} = 0.13$

Estimates MLE = Maximum Likelihood Estimate

log Like : −47.78

logarithm (natural) of the likelihood

Drawing a Reduced model



Null hypothesis
there is no difference between rates
(a.k.a rates are equal)

$$q = q_{01} = q_{10}$$

$$Q = \begin{pmatrix} -q & q \\ q & -q \end{pmatrix}$$

reduced_model

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$q^{MLE} = 0.05$ E/L/My

log like = -49.08

Comparing nested models

- Likelihood Ratio Test
- Null hypothesis: H_0: $q_{01} = q_{10}$ → Reject or Fail to reject
- Test statistic: $LRT = -2 \times \log\left(\frac{Likelihood\ (Reduced\ model)}{Likelihood\ (Full\ model)}\right)$
- $LTR \sim \chi^2$ with degrees of freedom k= #parameters full model- #parameters reduced model

$$LRT = -2 \log\left(\frac{\text{Likelihood (Reduced model)}}{\text{Likelihood (Full model)}}\right)$$

$$\log\left(\frac{a}{b}\right) = \log(a) - \log(b)$$

$$LRT = -2\left(\text{log lik(Reduced)} - \text{log Lik (Full)}\right)$$

$$= -2\left(-49.0\% - (-47.7)\right)$$

$$= -2(-1.38)$$

$$\boxed{LRT = 2.78}$$

test statistic

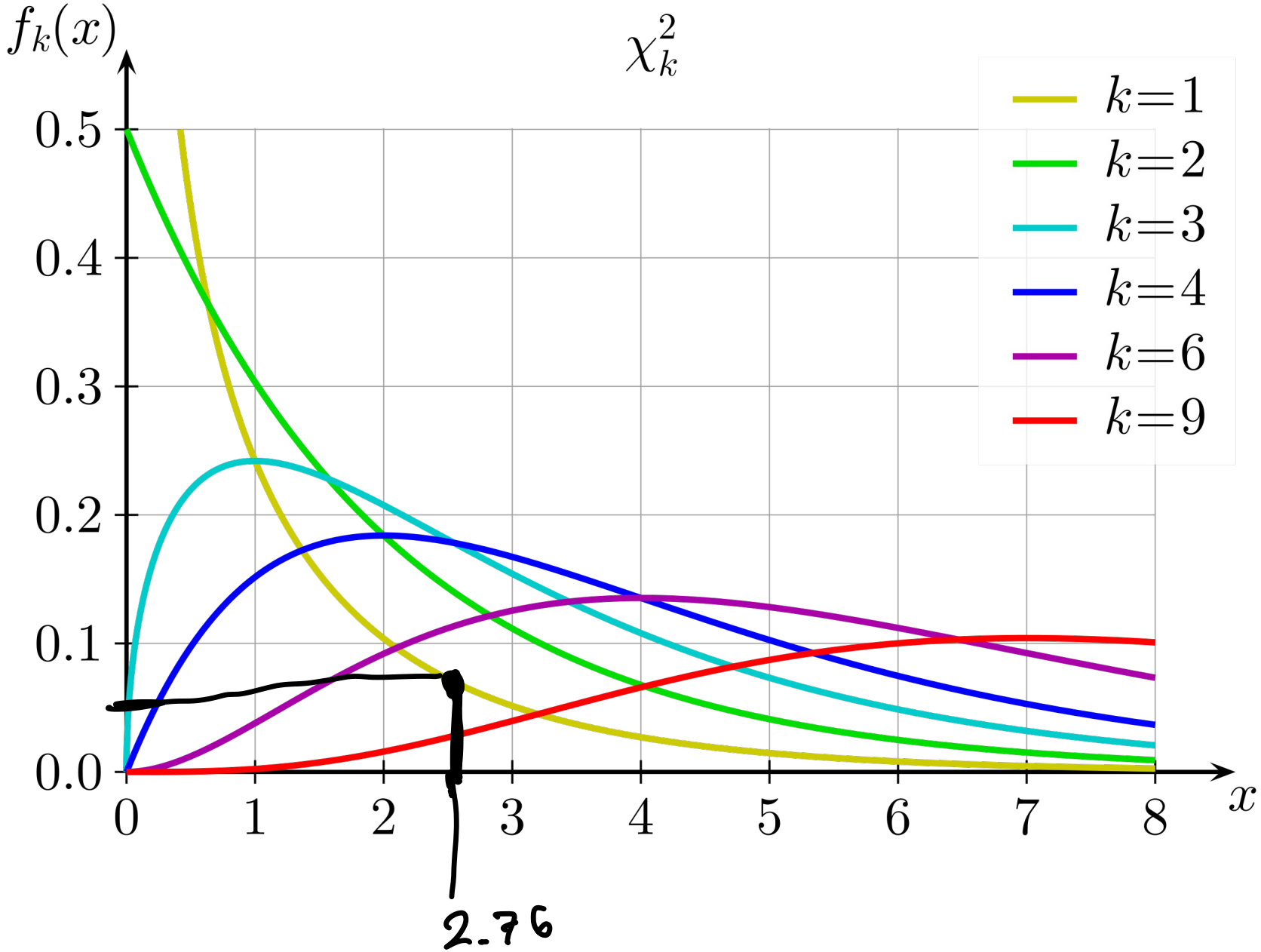
$\sim$ probability χ^2

degrees of freedom

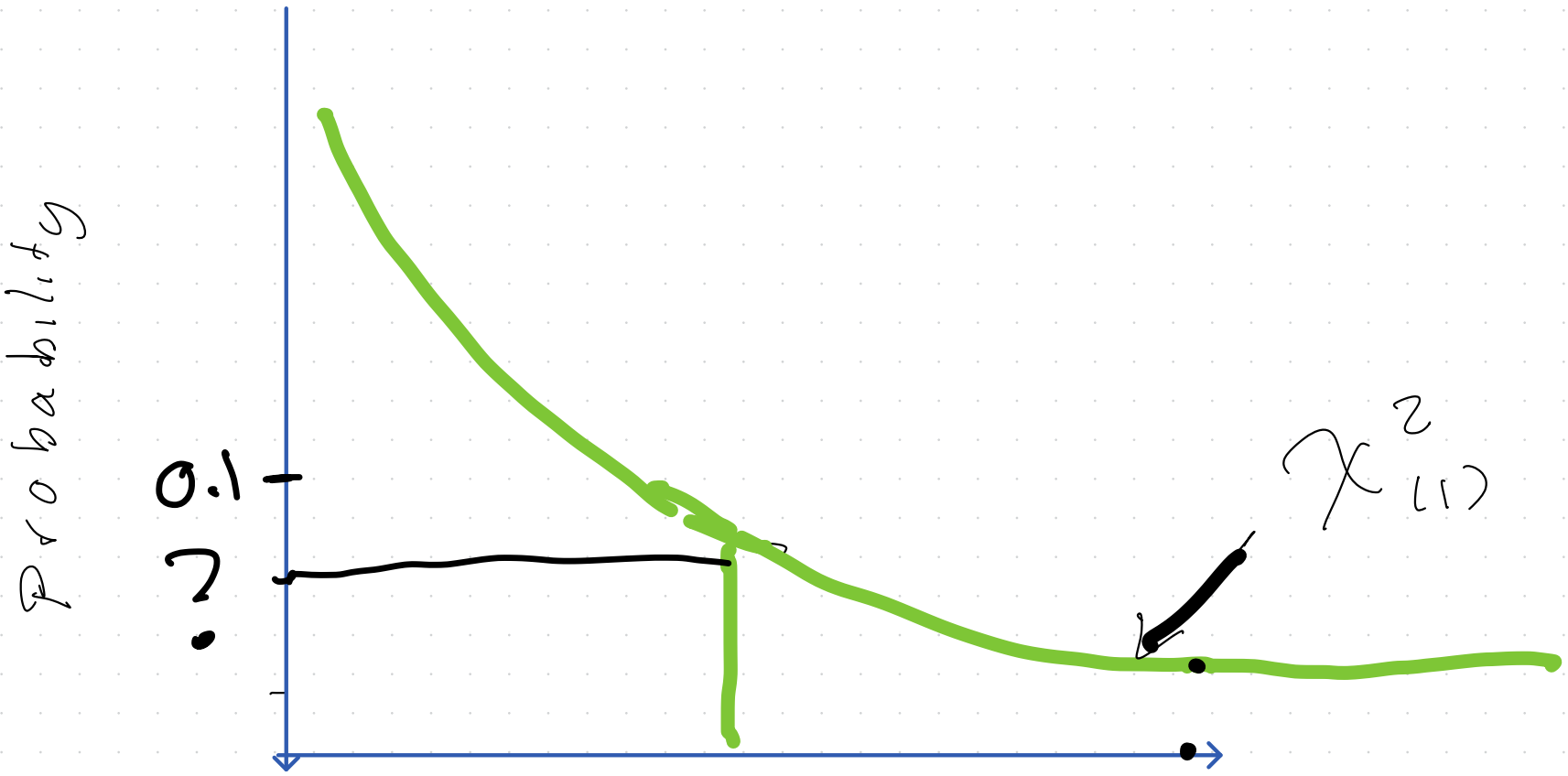
$k = (2 - 1) = 1$

2 ← parameters Full

1 ← parameters reduced



χ^2 distribution with
k degrees of freedom

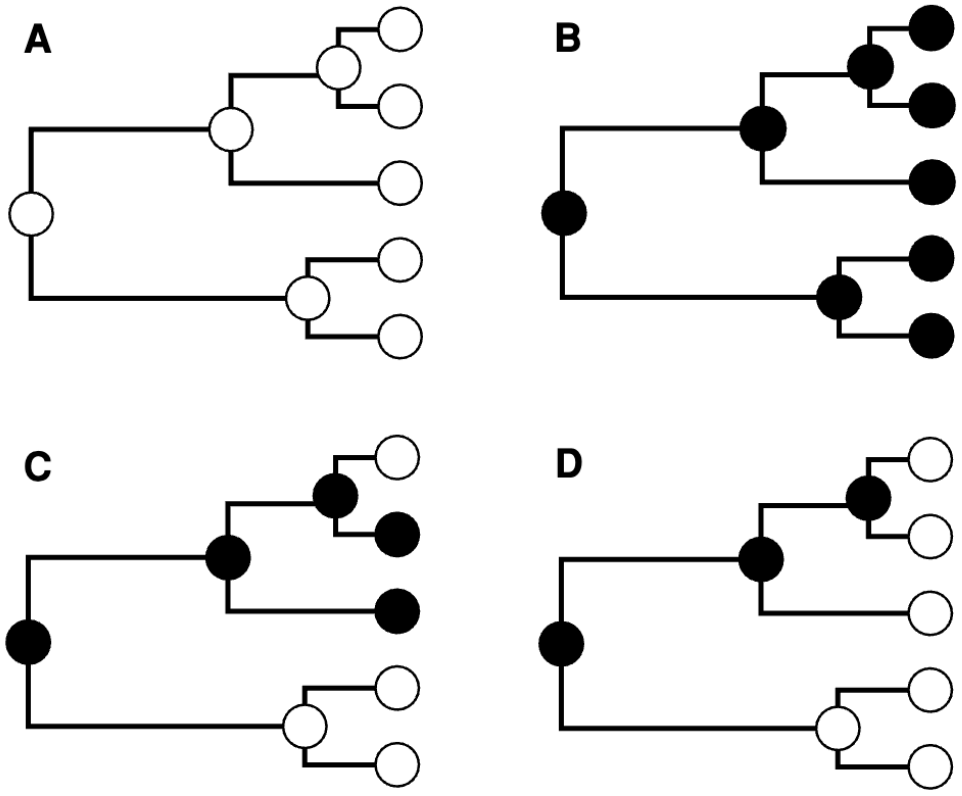


LTR = 2.76

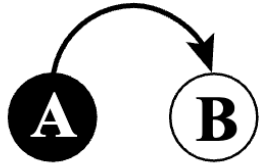
Test statistic value

Biologically what are we testing?

Dollo's law of irreversibility: An organism never returns exactly to a former state, even if it finds itself placed in conditions of existence identical to those which it has previously lived (Louis Dollo (1893)

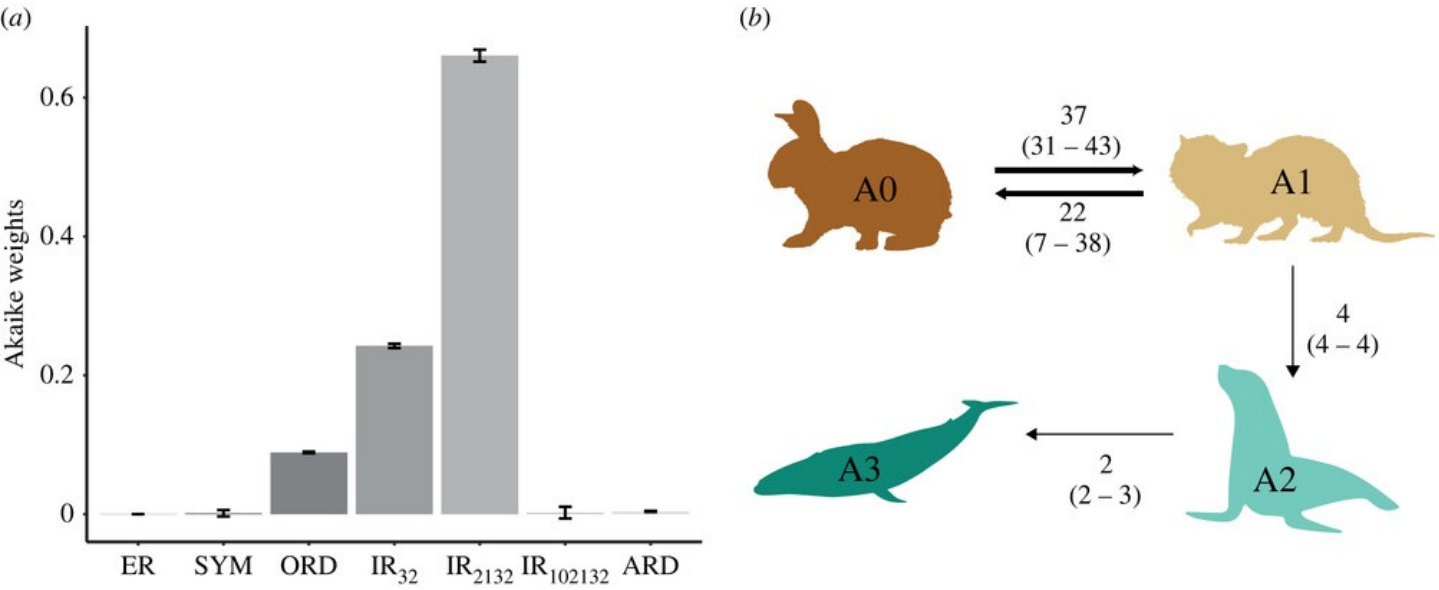


Goldberg and Igić. 2008

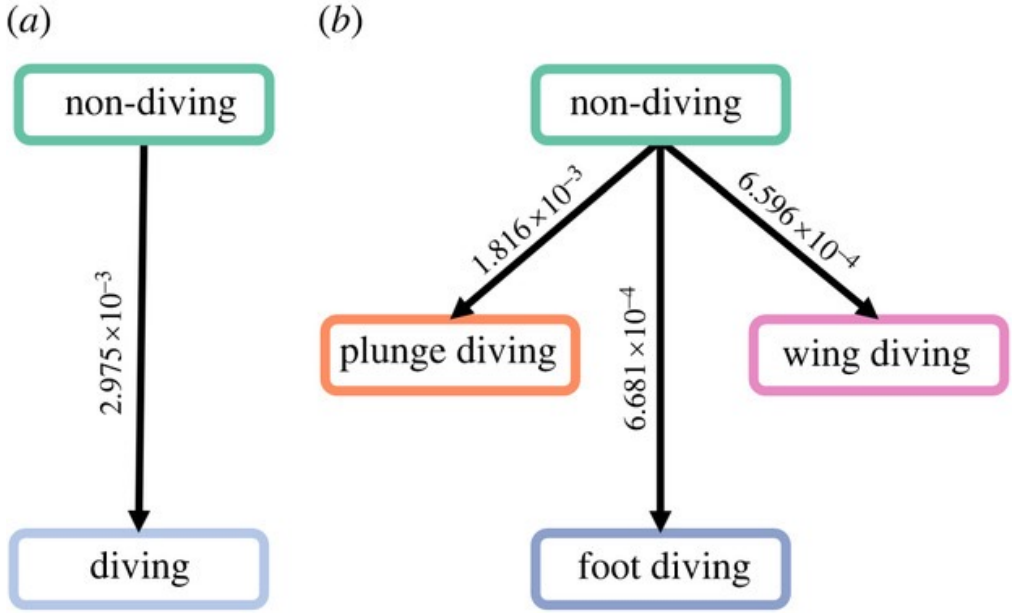


Biologically what are we testing?

Dollo's law of irreversibility: An organism never returns exactly to a former state, even if it finds itself placed in conditions of existence identical to those which it has previously lived (Louis Dollo (1893)



Secondary adaptation to aquatic environments (Farina and Silvestro, 2023)



Waterbirds diving
Tyler and Younger, 2022