Computational Biology

Lecturers:
Tanja Stadler, Carsten Magnus & Tim Vaughan

Teaching Assistants:
Jūlija Pečerska, Jérémie Sciré,
Sarah Nadeau & Marc Manceau

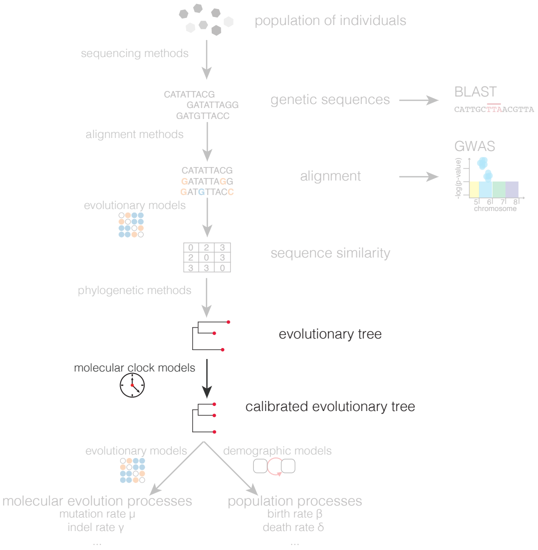
Computational Evolution
Department of Biosystems Science and Engineering

HS 2019

Statistical testing: Questions

- Is there a way to test how to best root a maximum likelihood tree?
- Can you use the bootstrapping ideas for assessing confidence in a UPGMA tree?
- What is required to infer the direction of transmission from a phylogeny?

Overview


population of individuals
sequencing methods
CATATTACG
GATATTAGG
GATGTTACC
genetic sequences
BLAST
CATTGCTTAACGTTA
alignment methods
GWAS
-log(p-value)
5 6 7 8
chromosome
CATATTACG
GATATTAGG
GATGTTACC
alignment
evolutionary models
0 2 3
2 0 3
3 3 0
sequence similarity
phylogenetic methods
evolutionary tree
molecular clock models
calibrated evolutionary tree
evolutionary models
demographic models
molecular evolution processes
mutation rate μ
indel rate γ
...
population processes
birth rate β
death rate δ
...

Outline of lecture 08

We will answer the following question:

> How can we compare **phenotypic traits/characters** between individuals/species that **evolved on a phylogeny**?

Outline of lecture 08

We will answer the following question:

How can we compare **phenotypic traits/characters** between individuals/species that **evolved on a phylogeny**?

such characters/traits could be:

- spike numbers of HIV virions
- number of legs in arthropods
- fur patterns in rodents

- height
- surface to weight ratio
- virulence of influenza
- shape of dinosaur jaws

Outline of lecture 08

We will answer the following question:

How can we compare **phenotypic traits/characters** between individuals/species that **evolved on a phylogeny**?

such characters/traits could be:

- spike numbers of HIV virions
- number of legs in arthropods
- fur patterns in rodents

discrete characters

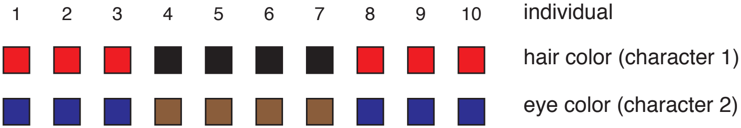
- height
- surface to weight ratio
- virulence of influenza
- shape of dinosaur jaws

continuous characters

Comparing discrete characters.

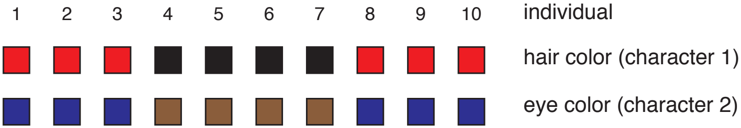
The problem

Example: We want to know whether eye color is correlated with hair color. We examine 10 individuals:



The problem

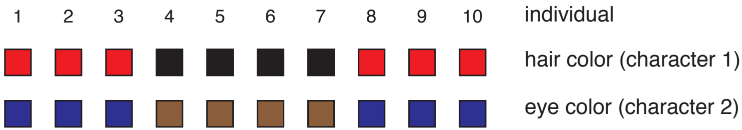
Example: We want to know whether eye color is correlated with hair color. We examine 10 individuals:



To test whether there is a true correlation we need to perform a statistical test.

The problem

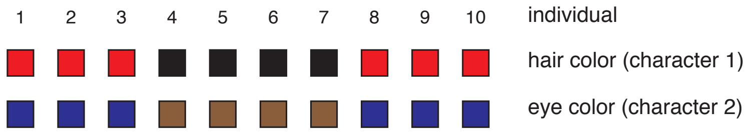
Example: We want to know whether eye color is correlated with hair color. We examine 10 individuals:



To test whether there is a true correlation we need to perform a statistical test. In this situation we can apply **Fisher's exact test** with a significance level of 0.05:

$\mathcal{H}_0$: Having brown eyes is equally likely among red- and black-haired individuals.

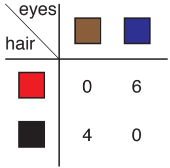
Fisher's exact test I

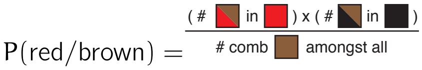


$\mathcal{H}_0$: Having brown eyes is equally likely among red- and black-haired individuals.

With Fisher's exact test, we test whether the observed result happened due to chance alone:

contingency table:



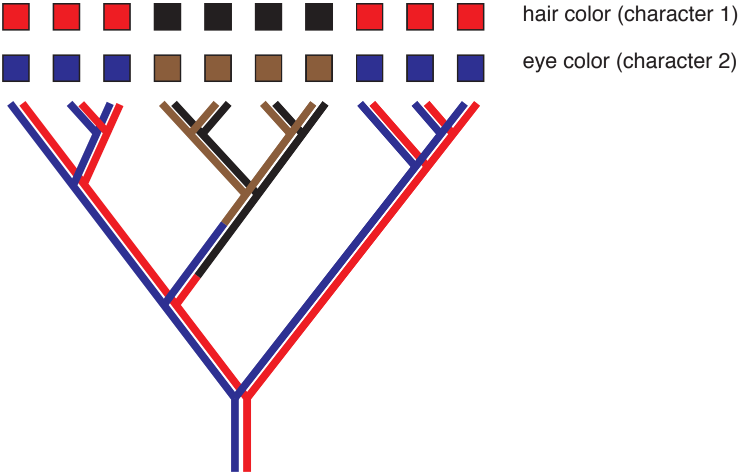


$$= \frac{\binom{6}{0}\binom{4}{4}}{\binom{10}{4}} = 0.0048 < 0.05$$

⇒ We reject the hypothesis of independent character evolution, i.e. we can see a correlation.

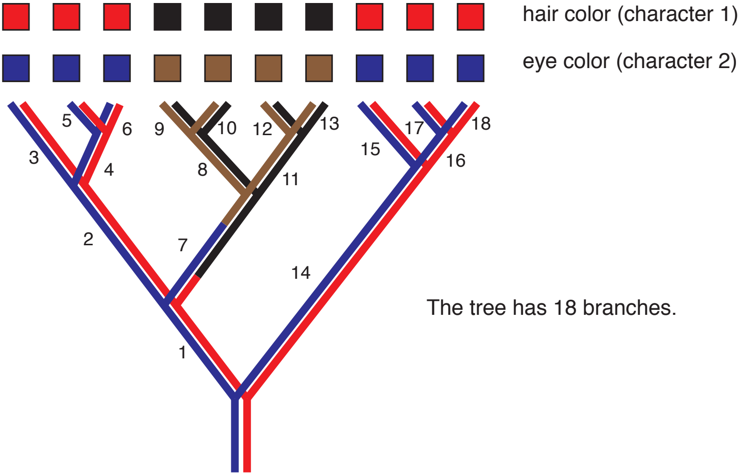
The problem: a phylogenetic approach

However, the analysis could be biased due to relatedness of the individuals:



The problem: a phylogenetic approach

However, the analysis could be biased due to relatedness of the individuals:



The tree has 18 branches.

Correct way to look at the problem: Is the change of characters on the branches correlated?

Fisher's exact test II

$\mathcal{H}_0$: The character changes are equally likely on every branch.

Fisher's exact test II

$\mathcal{H}_0$: The character changes are equally likely on every branch.

contingency table:

hair \ eyes	yes	no
yes	1	0
no	0	17

$$\begin{aligned}\mathrm{P}(2 \text{ changes on 1 branch}) &= \frac{\binom{1}{1}\binom{17}{0}}{\binom{18}{1}} \\ &= 0.05555 > 0.05\end{aligned}$$

$\Rightarrow \mathrm{p}\text{-value} > 0.05$

Fisher's exact test II

$\mathcal{H}_0$: The character changes are equally likely on every branch.

contingency table:

hair \ eyes	yes	no
yes	1	0
no	0	17

$$
\begin{aligned}
\mathrm{P}(2 \text{ changes on } 1 \text{ branch}) &= \frac{\binom{1}{1}\binom{17}{0}}{\binom{18}{1}} \\
&= 0.05555 > 0.05
\end{aligned}
$$

$\Rightarrow \mathrm{p}\text{-value} > 0.05$

$\Rightarrow$ We cannot reject the hypothesis that character change is equally likely, i.e. we cannot say that there is a correlation between the characters.

Fisher's exact test II

$\mathcal{H}_0$: The character changes are equally likely on every branch.

contingency table:

hair \ eyes	yes	no
yes	1	0
no	0	17

$$P(2 \text{ changes on 1 branch}) = \frac{\binom{1}{1}\binom{17}{0}}{\binom{18}{1}} = 0.05555 > 0.05$$

$$\Rightarrow p\text{-value} > 0.05$$

$\Rightarrow$ We cannot reject the hypothesis that character change is equally likely, i.e. we cannot say that there is a correlation between the characters.

To summarize: Neglecting the phylogenetic background can lead to false conclusions on correlations between characters. This is mainly the case because of the non-independence of species data points as a result of shared ancestry.

Methods to find correlations on discrete characters

- In our approach we do not consider differences in branch lengths. Changes are more likely to happen on longer branches than on short branches.
- Several methods correct for varying branch lengths in the phylogenies.
- [Ridley, 1983] used parsimony, [Pagel, 1994] used likelihood methods. For an overview of further methods see [Felsenstein, 2004].

Comparing continuous characters.

Continuous characters

So far, we have mostly looked at evolution on a **discrete space**:

- nucleotide substitution models in lectures 3 and 4
- codon and amino acid substitution models in lecture 4
- space of tree topologies (not considering branch lengths) in phylogenetic reconstruction based on nucleotide sequences in lectures 5-7
- correlation between discrete phenotypic characters in this lecture

Continuous characters

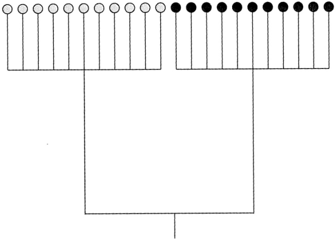
So far, we have mostly looked at evolution on a **discrete space**:

- nucleotide substitution models in lectures 3 and 4
- codon and amino acid substitution models in lecture 4
- space of tree topologies (not considering branch lengths) in phylogenetic reconstruction based on nucleotide sequences in lectures 5-7
- correlation between discrete phenotypic characters in this lecture

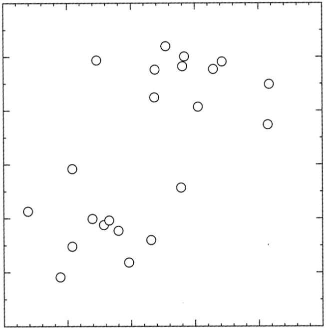
For the rest of the lecture we want to learn how evolution of **continuous phenotypic characters** (e.g. height, weight, virulence) can be modelled and how we can test for correlations amongst continuous traits.

Why linear regression cannot be used to compare two characters evolved on a phylogeny

Imagine a species tree



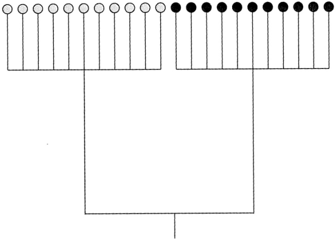
and two characters evolved on this tree:



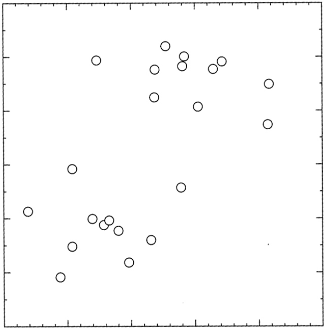
[Felsenstein, 2004]

Why linear regression cannot be used to compare two characters evolved on a phylogeny

Imagine a species tree



and two characters evolved on this tree:

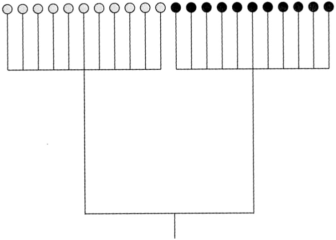


[Felsenstein, 2004]

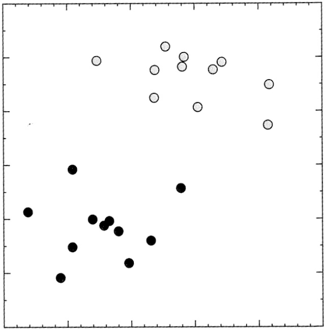
Is there a correlation?

Why linear regression cannot be used to compare two characters evolved on a phylogeny

Imagine a species tree



and two characters evolved on this tree:



[Felsenstein, 2004]

Is there a correlation?
No correlation. Only clade effects.

Toolbox: Brownian motion

One commonly used model to describe evolution of continuous traits on a phylogeny is the Brownian motion model.

- named after Robert Brown (1773 - 1858) who observed crop seeds on water under the microscope
- Albert Einstein (1879 -1955): crop seed movement due to hits from water molecules (described in a paper from 1905), seen as proof for molecules and atoms
- football stadium analogy

visit: `https://en.wikipedia.org/wiki/Brownian_motion`

Toolbox: Brownian motion – mathematical description

The theory behind Brownian motion is quite advanced as it is a time-continuous stochastic process on a continuous state space.

In short: Brownian motion is described as a Wiener* process, $(W_t)_{t \in T}$, which fulfils the following four conditions:

1. $W_0 = 0$
2. W_t is almost surely continuous
3. W_t has independent increments (implies **memorylessness**)
4. for $0 \leqslant s \leqslant t$, the $W_t - W_s \sim \mathcal{N}(0, \sigma^2(t - s))$

* named after Norbert Wiener (1894-1964), American mathematician

For further (hard-core math) reading: [Bertoin, 1994]

Toolbox: Brownian motion – mathematical description

The theory behind Brownian motion is quite advanced as it is a time-continuous stochastic process on a continuous state space.

In short: Brownian motion is described as a Wiener* process, $(W_t)_{t \in T}$, which fulfils the following four conditions:

1. $W_0 = 0$
   - the process starts in 0
2. W_t is almost surely continuous
3. W_t has independent increments (implies **memorylessness**)
4. for $0 \leqslant s \leqslant t$, the $W_t - W_s \sim \mathcal{N}(0, \sigma^2(t - s))$

* named after Norbert Wiener (1894-1964), American mathematician

For further (hard-core math) reading: [Bertoin, 1994]

Toolbox: Brownian motion – mathematical description

The theory behind Brownian motion is quite advanced as it is a time-continuous stochastic process on a continuous state space.

In short: Brownian motion is described as a Wiener* process, $(W_t)_{t \in T}$, which fulfils the following four conditions:

1. $W_0 = 0$
   - the process starts in 0
2. W_t is almost surely continuous
   - $P(W_t \text{ continuous}) = 1$
3. W_t has independent increments (implies **memorylessness**)

4. for $0 \leqslant s \leqslant t$, the $W_t - W_s \sim \mathcal{N}(0, \sigma^2(t - s))$

* named after Norbert Wiener (1894-1964), American mathematician

For further (hard-core math) reading: [Bertoin, 1994]

Toolbox: Brownian motion – mathematical description

The theory behind Brownian motion is quite advanced as it is a time-continuous stochastic process on a continuous state space.

In short: Brownian motion is described as a Wiener* process, $(W_t)_{t \in T}$, which fulfils the following four conditions:

1. $W_0 = 0$
   - the process starts in 0
2. W_t is almost surely continuous
   - $P(W_t \text{ continuous}) = 1$
3. W_t has independent increments (implies **memorylessness**)
   - for $0 \leqslant s_1 \leqslant t_1 < s_2 \leqslant t_2$, $(W_{t_1} - W_{s_1})$ and $(W_{t_2} - W_{s_2})$ are independent
4. for $0 \leqslant s \leqslant t$, the $W_t - W_s \sim \mathcal{N}(0, \sigma^2(t-s))$

* named after Norbert Wiener (1894-1964), American mathematician

For further (hard-core math) reading: [Bertoin, 1994]

Toolbox: Brownian motion – mathematical description

The theory behind Brownian motion is quite advanced as it is a time-continuous stochastic process on a continuous state space.

In short: Brownian motion is described as a Wiener* process, $(W_t)_{t \in T}$, which fulfils the following four conditions:

1. $W_0 = 0$
   - the process starts in 0
2. W_t is almost surely continuous
   - $P(W_t \text{ continuous}) = 1$
3. W_t has independent increments (implies **memorylessness**)
   - for $0 \leqslant s_1 \leqslant t_1 < s_2 \leqslant t_2$, $(W_{t_1} - W_{s_1})$ and $(W_{t_2} - W_{s_2})$ are independent
4. for $0 \leqslant s \leqslant t$, the $W_t - W_s \sim \mathcal{N}(0, \sigma^2(t-s))$
   - $\mathcal{N}(\mu, \sigma^2)$: normal distribution with density function:

     $f_{\mu,\sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

* named after Norbert Wiener (1894-1964), American mathematician

For further (hard-core math) reading: [Bertoin, 1994]

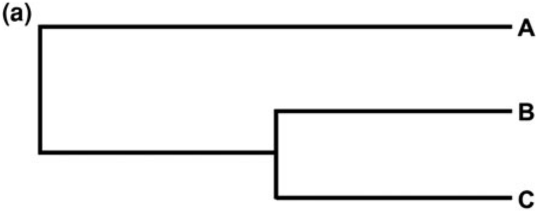
Toolbox: Brownian motion – A model for continuous character evolution

Analogies between models for evolution on discrete and continuous character space:

discrete	continuous
probability to visit any state	probability density on state space
memorylessness due to Markov Chain model	memorylessness due to Brownian motion
transition probabilities scale with time	variance scales with branch length

Toolbox: Brownian motion – On a phylogeny

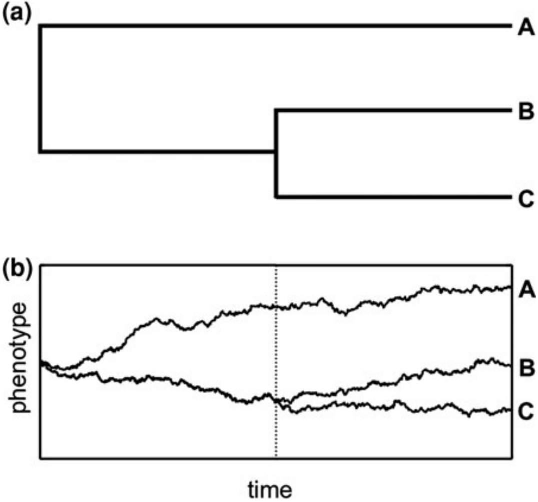
Given a phylogeny, we can apply a Brownian motion model on this phylogeny to evolve a continuous character (e.g. height):



picture from: [Symonds and Bomberg, 2014]

Toolbox: Brownian motion – On a phylogeny

Given a phylogeny, we can apply a Brownian motion model on this phylogeny to evolve a continuous character (e.g. height):



picture from: [Symonds and Bomberg, 2014]

Toolbox: Linear regression

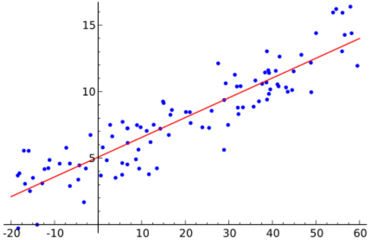
Mathematical method to determine the dependency of a variable Y on another variable X. We measure X and Y for n independent realisations and fit a regression model to the data. The observations $(x_1, y_1), \ldots, (x_n, y_n)$ need to be

- independent
- with the same (normally) distributed errors

Toolbox: Linear regression

Mathematical method to determine the dependency of a variable Y on another variable X. We measure X and Y for n independent realisations and fit a regression model to the data. The observations $(x_1, y_1), \ldots, (x_n, y_n)$ need to be

- independent
- with the same (normally) distributed errors



en.wikipedia.org/wiki/Linear_regression

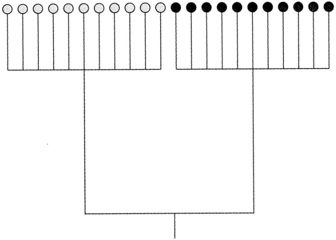
Model:
$y_i = \beta x_i + b + \epsilon$
where $\epsilon \sim \mathcal{N}(0, \sigma^2)$

Fitting:
Least squares method

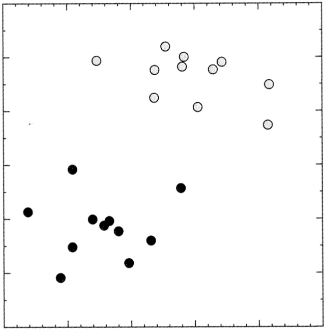
Goodness of fit:
R^2: perfect fit if close to 1; no dependency if close to 0

Why linear regression cannot be used to compare two characters evolved on a phylogeny

Imagine a species tree



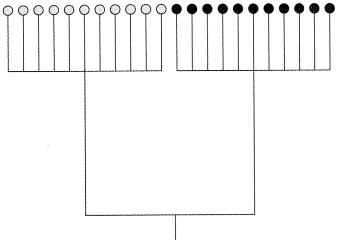
and two characters evolved on this tree:



[Felsenstein, 2004]

Is there a correlation?
No correlation. Only clade effects.

Why linear regression cannot be used to compare two characters evolved on a phylogeny



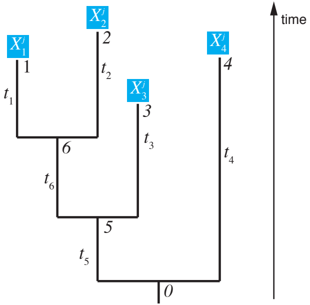
When two characters evolve on a tree

- they share common evolutionary history (**not independent realisations!!!**)
- the "error" (variance added by Brownian motion) is not equally distributed

Constructing independent variables

One method to overcome interdependencies of the evolutionary trait is the contrast method.

Suppose a phylogeny of 4 species:

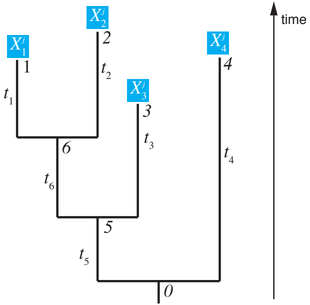


traits X_1^j and X_2^j are not independent as they share the evolutionary lineages t_6, t_7

Constructing independent variables

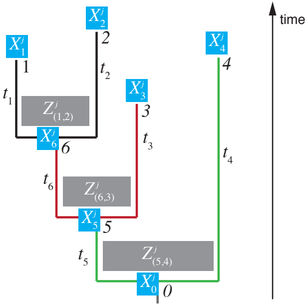
One method to overcome interdependencies of the evolutionary trait is the contrast method.

Suppose a phylogeny of 4 species:



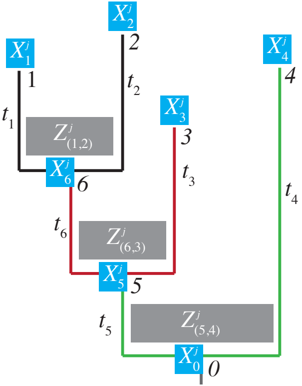
traits X_1^j and X_2^j are not independent as they share the evolutionary lineages t_6, t_7

Instead of characters, we look at their contrasts:



Independent contrasts

We need to calculate/estimate the **values of the contrasts** and their **variances** in order to perform a linear regression on the contrasts:



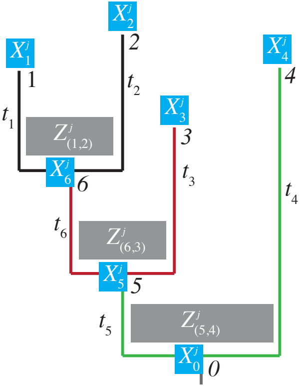
$$Z^j_{(1,2)} = X^j_1 - X^j_2$$

$$Z^j_{(6,3)} = X^j_6 - X^j_3$$

$$Z^j_{(5,4)} = X^j_5 - X^j_4$$

Independent contrasts

We need to calculate/estimate the **values of the contrasts** and their **variances** in order to perform a linear regression on the contrasts:



$$Z^j_{(1,2)} = X^j_1 - X^j_2$$

$$Z^j_{(6,3)} = X^j_6 - X^j_3$$

$$Z^j_{(5,4)} = X^j_5 - X^j_4$$

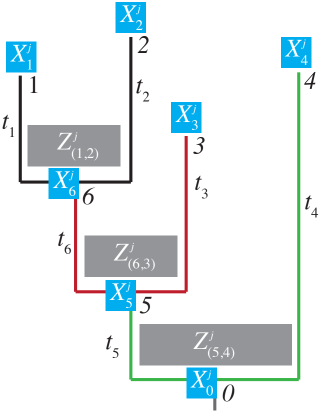
- We assume character evolution according to Brownian motion.
- We observed the tip values, but we have to estimate the values at internal nodes.

Variance of sum of random variables

To calculate the variance, we apply the following formula:

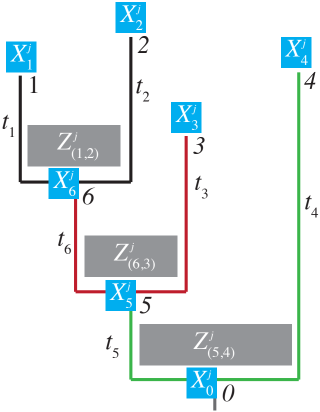
$$Var[\alpha X + \beta Y] = \alpha^2 Var[X] + \beta^2 Var[Y] + 2\alpha\beta Cov[X, Y]$$

Contrasts at cherries



$$Z^j_{(1,2)} = X^j_1 - X^j_2$$

Contrasts at cherries



$$Z_{(1,2)}^j = X_1^j - X_2^j$$

$$\mathrm{Var}[X_1^j] = \sigma^2(t_1 + t_6 + t_5)$$
$$\mathrm{Var}[X_2^j] = \sigma^2(t_2 + t_6 + t_5)$$

$$\begin{aligned} \mathrm{Var}[Z_{(1,2)}^j] &= \mathrm{Var}[X_1^j - X_2^j] \\ &= \\ \mathrm{Var}[X_1^j] + \mathrm{Var}[X_2^j] &- 2\mathrm{Cov}[X_1^j, X_2^j] \\ &= \sigma^2\big(t_1 + t_6 + t_5 + t_2 + t_6 + t_5 \\ &\qquad -2(t_6 + t_5)\big) \\ &= \sigma^2(t_1 + t_2) \end{aligned}$$

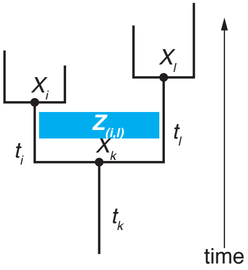
☞ The value for contrasts at cherries can easily be calculated. The variance is proportional to the branch lengths between the two external nodes.

Contrasts further down in the tree

For the sake of simpler notation: superscripts j are omitted here!

Wanted:

$Z_{(i,l)} = X_i - X_l$ and
$Var[Z_{(i,l)}]$

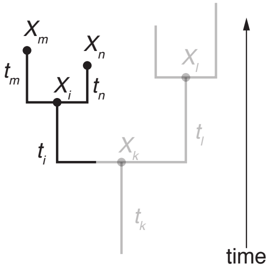


Contrasts further down in the tree

For the sake of simpler notation: superscripts j are omitted here!

Wanted:

$Z_{(i,l)} = X_i - X_l$ and $Var[Z_{(i,l)}]$



We have to calculate the values at internal nodes:

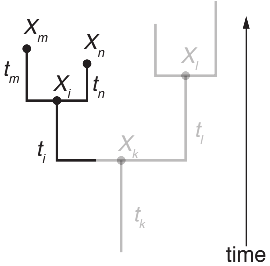
$X_i =$

Contrasts further down in the tree

For the sake of simpler notation: superscripts j are omitted here!

Wanted:

$Z_{(i,l)} = X_i - X_l$ and $Var[Z_{(i,l)}]$



We have to calculate the values at internal nodes:

$$X_i = \frac{t_n}{t_m+t_n} X_m + \frac{t_m}{t_m+t_n} X_n$$

and their variances:

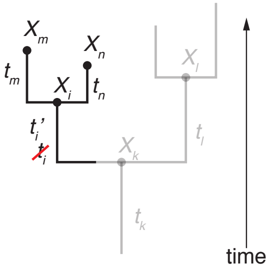
$$Var[X_i] = Var[\frac{t_n}{t_m+t_n} X_m + \frac{t_m}{t_m+t_n} X_n]$$

Contrasts further down in the tree

For the sake of simpler notation: superscripts j are omitted here!

Wanted:

$Z_{(i,l)} = X_i - X_l$ and $Var[Z_{(i,l)}]$



We have to calculate the values at internal nodes:

$$X_i = \frac{t_n}{t_m+t_n} X_m + \frac{t_m}{t_m+t_n} X_n$$

and their variances:

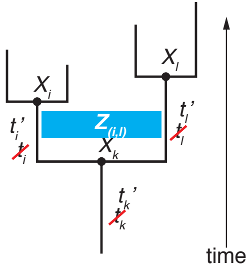
$$Var[X_i] = Var[\frac{t_n}{t_m+t_n} X_m + \frac{t_m}{t_m+t_n} X_n]$$
$$= \sigma^2 \left(\underbrace{\frac{t_m t_n}{t_m + t_n} + t_i}_{t_i'} + t_k + \dots \right)$$

Contrasts further down in the tree

For the sake of simpler notation: superscripts j are omitted here!

Wanted:

$Z_{(i,l)} = X_i - X_l$ and $Var[Z_{(i,l)}]$



We have to calculate the values at internal nodes:

$$X_i = \frac{t_n}{t_m+t_n}X_m + \frac{t_m}{t_m+t_n}X_n$$

and their variances:

$$Var[X_i] = Var[\frac{t_n}{t_m+t_n}X_m + \frac{t_m}{t_m+t_n}X_n]$$
$$= \sigma^2 \left(\underbrace{\frac{t_m t_n}{t_m + t_n} + t_i}_{t_i'} + t_k + \dots \right)$$

calculate values at internal nodes, and the corrected branch lengths, then calculate the contrasts and their variances:

$$Var[Z_{(i,l)}] = \sigma^2 \left(t_i' + t_l' \right)$$

Normalisation of contrasts

To be able to compare the independent contrast, all contrasts need to have the same variance. Thus, we need to normalise the contrasts in a last step.

Normalisation of contrasts

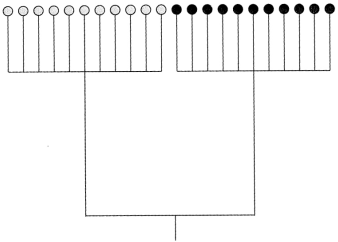
To be able to compare the independent contrast, all contrasts need to have the same variance. Thus, we need to normalise the contrasts in a last step.

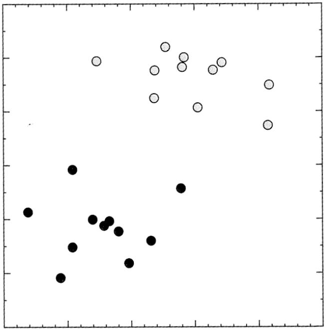
Given the contrast $Z^j_{(i,l)}$ with variance $Var[Z^j_{(i,l)}] = \sigma^2 c_{Z^j_{(i,l)}}$
We know that $Var(\alpha X) = \alpha^2 Var(X)$. Thus, we can replace the contrasts by

$$Z^j_{(i,l)} = Z^j_{(i,l)} / \sqrt{c_{Z^j_{(i,l)}}}$$

Therefore all $Z^j_{(i,l)} \sim \mathcal{N}(0, \sigma^2)$ and are ready for a linear regression. Remember that the superscript j indicates the j-th trait.

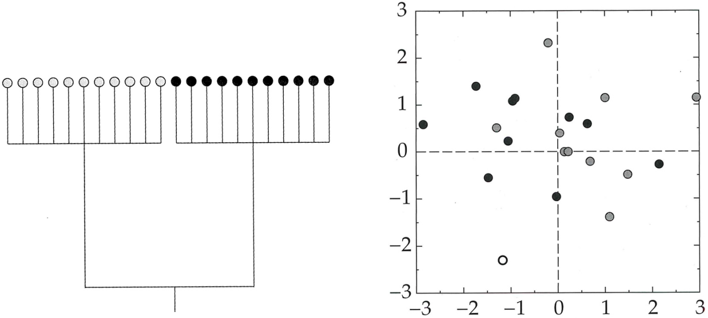
Independent contrasts in the first example





[Felsenstein, 2004]

Independent contrasts in the first example



[Felsenstein, 2004]

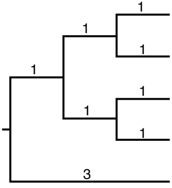
The independent contrast method supports our early suspicion that no correlation between the two characters evolved on this particular phylogeny can be found.

Example: Fiddler crabs

male fiddler crab
[Fiddlercrab, 2015]

Is there a correlation between carapace breadth and propodus length in five *Uca* fiddler crab species?

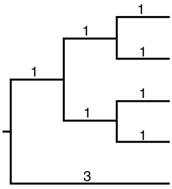
Carapace breadth and propodus length



	c	p
chloroph-talmus	1.02	1.38
crassipes	1.06	1.41
inversa	0.96	1.36
sindensis	0.92	1.22
argilicola	0.89	1.13

$\mathfrak{c}$ carapace breadth
$\mathfrak{p}$ propodus length
[Symonds and Bomberg, 2014]

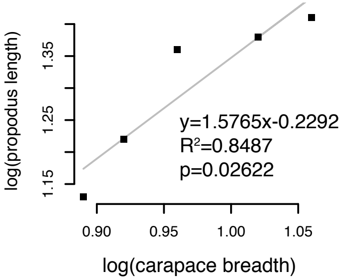
Carapace breadth and propodus length



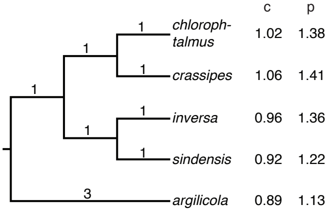
	c	p
chloroph-talmus	1.02	1.38
crassipes	1.06	1.41
inversa	0.96	1.36
sindensis	0.92	1.22
argilicola	0.89	1.13

$\mathfrak{c}$ carapace breadth
$\mathfrak{p}$ propodus length
[Symonds and Bomberg, 2014]

Linear regression on the data gives:

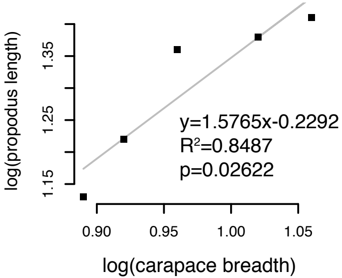


Carapace breadth and propodus length



c carapace breadth
p propodus length
[Symonds and Bomberg, 2014]

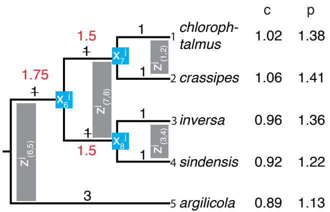
Linear regression on the data gives:



At first sight, carapace breadth seems to be a predictor for the propodus length.

Independent contrasts

We can now derive the independent contrasts by first calculating the corrected branch lengths and the character values in the nodes:



Character values at internal nodes:

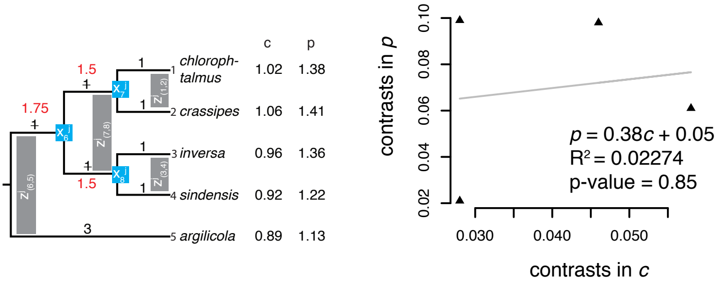
$j=$	c	p
x_7^j	1.04	1.395
x_6^j	0.94	1.29
x_8^j	0.99	1.3425

Contrasts:

	c_{raw}	p_{raw}	SD	c_{st}	p_{st}
$z^j_{(1,2)}$	0.04	0.03	$\sqrt{2}$	0.028	0.021
$z^j_{(3,4)}$	0.04	0.14	$\sqrt{2}$	0.028	0.099
$z^j_{(7,8)}$	0.1	0.105	$\sqrt{3}$	0.058	0.061
$z^j_{(6,5)}$	0.1	0.2125	$\sqrt{4.75}$	0.046	0.098

Regression on independent contrasts

Regression on the standardized contrasts leads to:

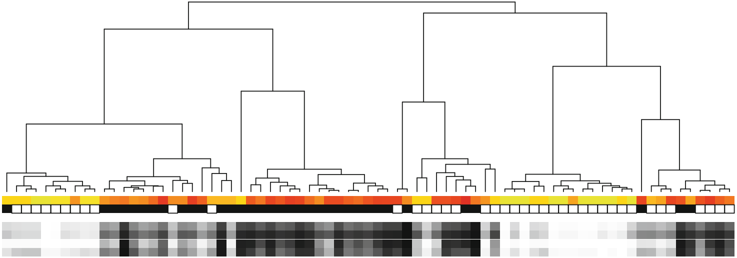


No signal for dependency between carapace breadth and propodus length.

`Comparison of a discrete and a continuous character.`

Antelope group sizes as predictors for antipredator behavior

[Ives and Garland, 2014] studied the dependency between a **continuous** variable ($\log_{10}$ group size in yellow-red) and the **discrete** antipredator behaviour (hide = white, flee/fight = black) in 75 antelope species:



The four bottom rows show predictions of the antipredator behaviour using different models.

Summary

- Continuous character evolution on a phylogeny can be modelled by Brownian motion.
  - More advanced models for character evolution can be used, e.g. Ornstein-Uhlenbeck-process in HIV virulence evolution
- Discrete and continuous character evolution cannot be analysed directly with a regression if species/individuals have a common evolutionary history.
- If species/individuals evolved on a phylogeny, one has to correct for the common evolutionary trajectories before comparing characters (e.g. using independent contrasts).

Continuous traits and comparative methods: Questions.

- In a Fisher's exact test, how would you calculate which values for one of the cells in the contingency table would lead to a rejection of the null hypothesis, given that row and column sums remain the same?
- Is the Brownian motion model a good model for all continuous traits? Could you imagine situations where this is not the case and which assumption in this model could be violated?
- Do you think it is a good strategy to first determine the species tree and then look at character evolution, or would a co-estimation of characters and the phylogeny make more sense?

References I

- Bertoin, J. (1994). *Lévy Processes*. Cambridge Tracts in Mathematics (Volume 121).
- Felsenstein, J. (2004). *Inferring phylogenies*. Sinauer associates Sunderland.
- Fiddlercrab (2015). Information on fiddler crabs: http://www.fiddlercrab.info .
- Ives, A. and Garland, T. J. (2014). *Phylogenetic Regression for Binary Dependent Variables. Book chapter in: Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology*. Springer.
- Pagel, M. (1994). Detecting correlated evolution on phylogenies: A general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society of London, Series B*, 255:37–45.
- Ridley, M. (1983). *The Explanation of Organic Diversity: The Comparative Method and Adaptations for Mating*. Oxford University Press.
- Symonds, M. and Bomberg, S. (2014). *A Primer on Phylogenetic Generalised Least Squares. Book chapter in: Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology*. Springer.