

# The Comparative Method

# Multivariate Brownian Motion

Univariate:

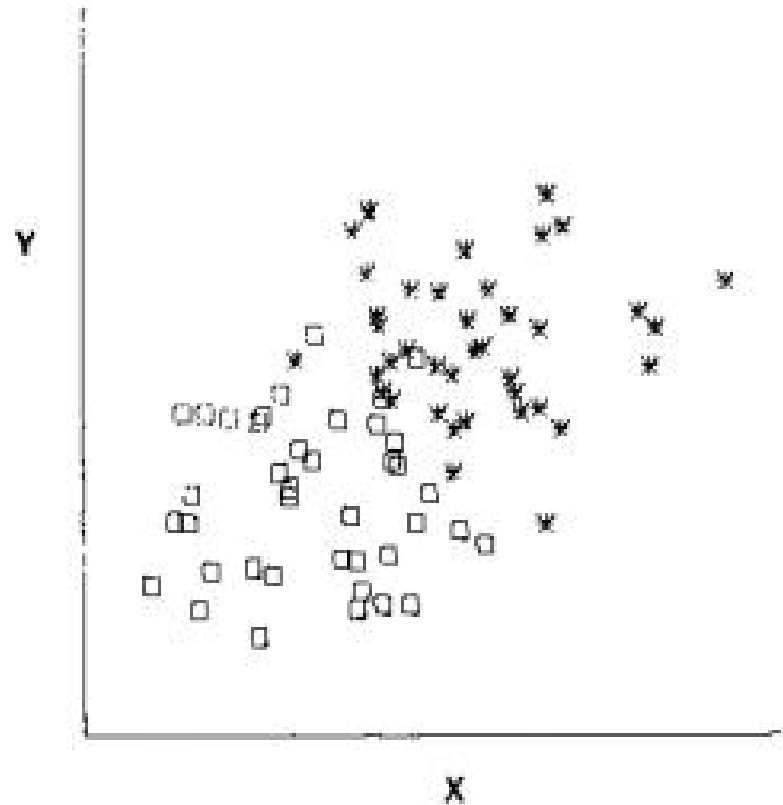
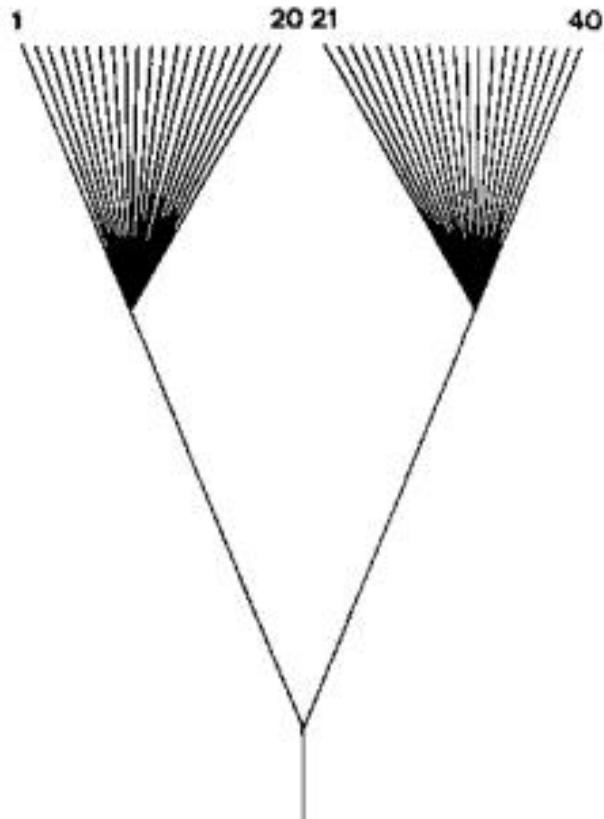
$$\text{root} = \mu$$

$$\text{step variance} = \sigma^2$$

Multivariate (N traits):

$$\text{root} = c(\mu_1, \dots, \mu_N)$$

$$\text{step variance} = \mathbf{R} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}$$



**Felsenstein 1985**

Brownian motion is a random process modeling the wanderings of a molecule affected by thermal noise. If we measure the position of the molecule along one axis, its successive displacements are independent. This has the effect that the displacement after time  $\nu$  has elapsed is the sum of a large number of small displacements, each of which is equally likely to be either positive or negative. The result is that the total displacement is drawn from a normal distribution with mean zero and a variance proportional to  $\nu$ . In the present model the different characters undergo Brownian motion at different rates, so that after one unit of time the change in  $X$  has variance  $s_X^2$  and the (possibly correlated) change in  $Y$  has variance  $s_Y^2$ . After  $\nu$  units of time their variances are, respectively,  $s_X^2\nu$  and  $s_Y^2\nu$ .

There is one case in which the problem does not arise. That is when the characters respond essentially instantaneously to natural selection in the current environment, so that phylogenetic inertia is essentially absent. In that case we could correlate a phenotype with the environment. We could also correlate two characters with each other, provided that we realized that their correlation might simply reflect response to a common environmental factor. It may be doubted how often phylogenetic inertia is effectively absent. In any case the presumption of the absence of phylogenetic inertia should be acknowledged whenever it is proposed to do comparative studies without taking account of the phylogeny.

**What is meant by this passage? Why does Felsenstein make this argument?**

There is no reason to believe that the normal distribution is particularly plausible as the distribution from which changes in individual branches of the phylogeny are drawn, except insofar as the net change in a branch is the resultant of a series of bursts of change and thus might be approximately normal.

The matter of the model is an obvious point for future development and (to the extent that this is possible) empirical study. One rather serious problem that confronts comparative studies is that the relationship under study may change through time. Harvey and Mace (1982) have discussed the problem of change of the slope of the relationship between two variables with taxonomic level, which appears to be quite common. It should be possible to use the current model to study statistically whether there is any connection between the variance of a contrast and the slope of the regression of one variable on another.

### *What if We Do Not Take the Phylogeny into Consideration?*

Some reviewers of this paper felt that the message was “rather nihilistic,” and suggested that it would be much improved if I could present a simple and robust method that obviated the need to have an accurate knowledge of the phylogeny. I entirely sympathize, but do not have a method that solves the problem. The best we can do is perhaps to use pairs of close relatives as suggested above, although this discards at least half of the data. Comparative biologists may understandably feel frustrated upon being told that they need to know the phylogenies of their groups in great detail, when this is not something they had much interest in knowing. Nevertheless, efforts to cope with the effects of the phylogeny will have to be made. Phylogenies are fundamental to comparative biology; there is no doing it without taking them into account.

# Phylogenetically Independent Contrasts Regression

N-1 contrasts for N tips

Regress  $Y \sim X$  (or  $X \sim Y$ ) with *intercept fit through the origin*

Assumes multivariate Brownian Motion of *both X and Y*

Can incorporate measurement error in both X and Y (Felsenstein 2008)



# Phylogenetic Generalized Least Squares Regression

Downloaded from <http://rstb.royalsocietypublishing.org/> on October 17, 2018

*Phil. Trans. R. Soc. Lond. B* 326, 119–157 (1989) [ 119 ]

*Printed in Great Britain*

## THE PHYLOGENETIC REGRESSION

By A. GRAFEN†

*Animal Behaviour Research Group, Department of Zoology, University of Oxford, South Parks Road,  
Oxford OX1 3PS, U.K.*

*(Communicated by W. D. Hamilton, F.R.S. – Received 13 February 1989)*

# Turns out...

*Syst. Biol.* 61(3):382–391, 2012

© The Author(s) 2012. Published by Oxford University Press, on behalf of the Society of Systematic Biologists. All rights reserved.

For Permissions, please email: journals.permissions@oup.com

DOI:10.1093/sysbio/syr118

Advance Access publication on January 3, 2012

## Independent Contrasts and PGLS Regression Estimators Are Equivalent

SIMON P. BLOMBERG<sup>1,\*</sup>, JAMES G. LEFEVRE<sup>2</sup>, JESSIE A. WELLS<sup>1</sup>, AND MARY WATERHOUSE<sup>3</sup>

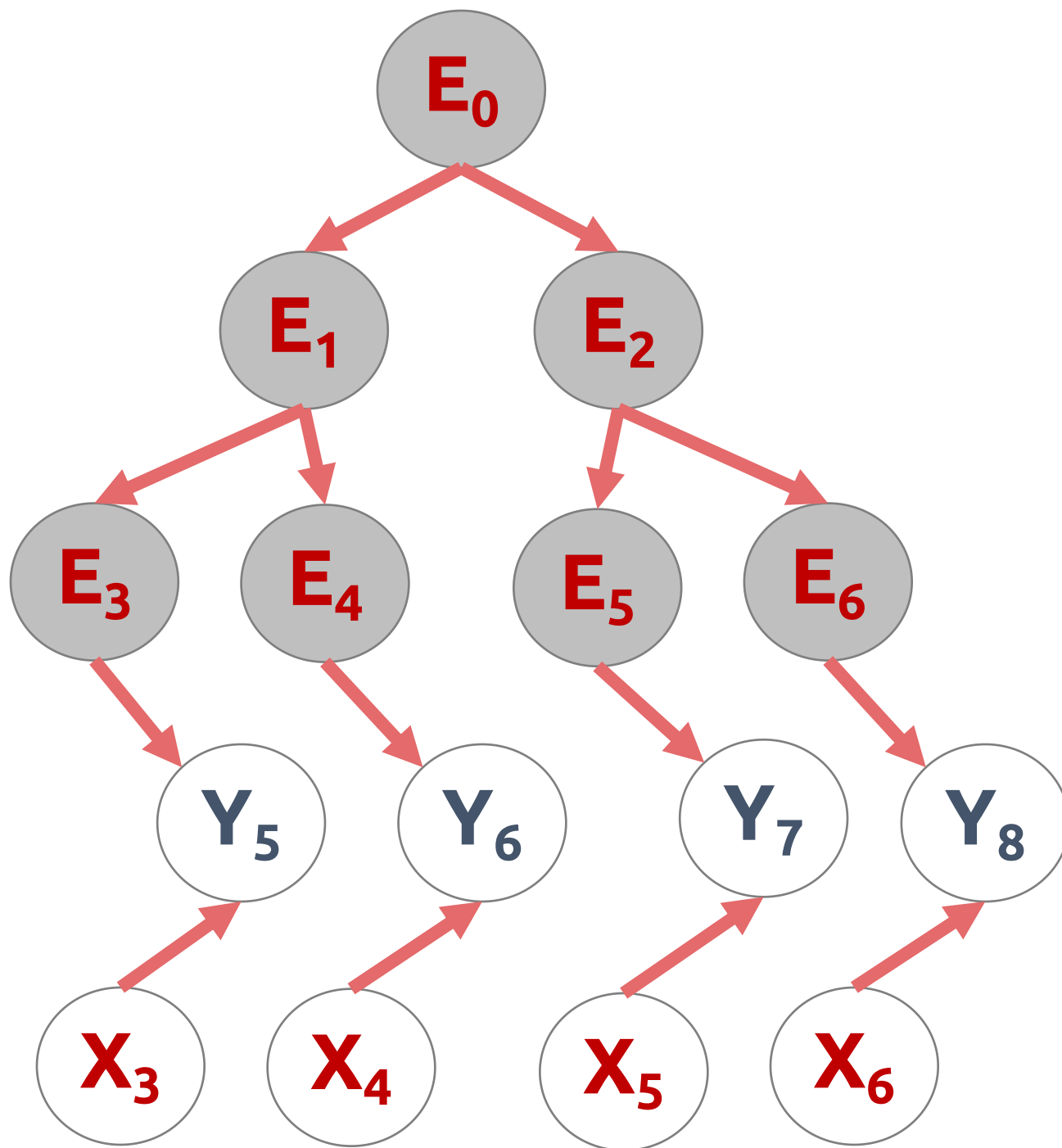
<sup>1</sup>*School of Biological Sciences, University of Queensland, St. Lucia, Queensland 4072, Australia;* <sup>2</sup>*Centre for Discrete Mathematics and Computing, University of Queensland, St. Lucia, Queensland 4072, Australia;* and <sup>3</sup>*Discipline of Mathematical Sciences, Faculty of Science and Technology, School of Mathematical Sciences, Queensland University of Technology, Brisbane, Queensland 4001, Australia;*

*\*Correspondence to be sent to: School of Biological Sciences, University of Queensland, St. Lucia, Queensland 4072, Australia;  
E-mail: s.blomberg1@uq.edu.au.*

*Received 13 October 2010; reviews returned 11 January 2011; accepted 15 September 2011*

*Associate Editor: Luke Harmon*

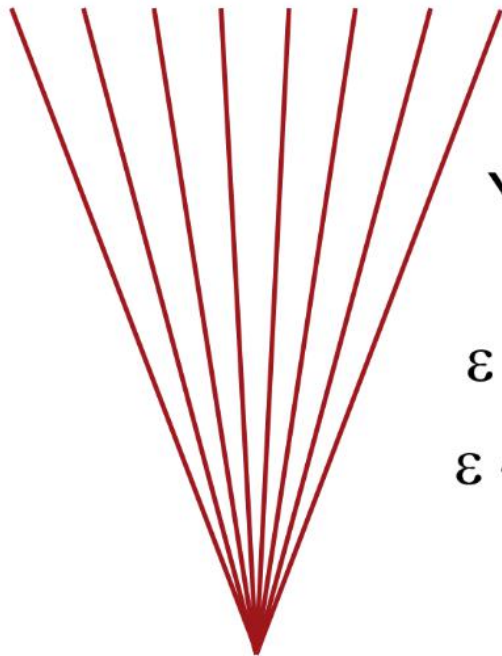
$$Y \sim X$$



# Just like before...multivariate Normal distribution

To define a mvNormal distribution, just need a mean (vector) and a variance (matrix).

A B C D E F G H

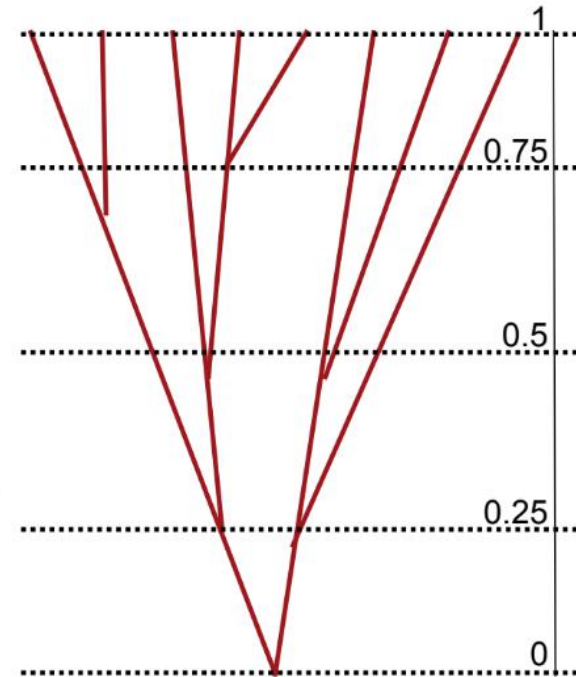


$$Y = \beta X + \varepsilon$$

$$\varepsilon \sim N(0, \sigma^2 \mathbf{I})$$

$$\varepsilon \sim N(0, \sigma^2 \mathbf{V})$$

A B C D E F G H



# Extensions of PGLS

## Tree transformations

$\lambda$  - “non phylogenetic slop” (heritability)

= 1 BM

= 0 No tree structure

$\kappa$  - Speciation

= 1 BM

= 0 Only speciation

$\delta$  - Node heights raised/lowered to power of delta

= 1 BM

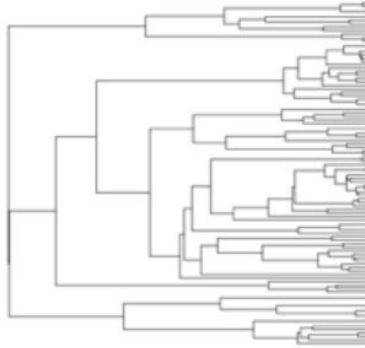
> 1 Accelerating

<1 Decelerating

Ornstein-Uhlenbeck - We will cover a lot more later!

# Tree transformations

Starting tree

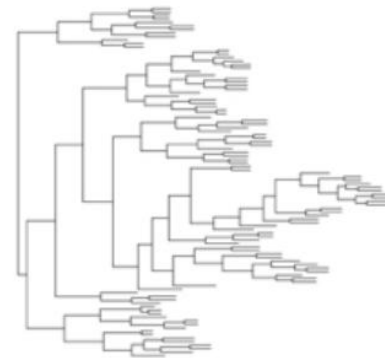
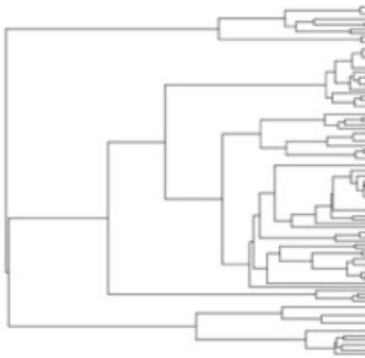
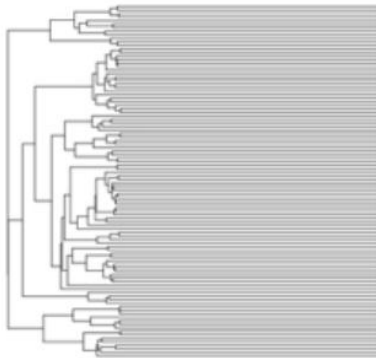


$$\mathbf{C}_\lambda = \begin{bmatrix} \sigma_1^2 & \lambda \cdot \sigma_{12} & \dots & \lambda \cdot \sigma_{1r} \\ \lambda \cdot \sigma_{21} & \sigma_2^2 & \dots & \lambda \cdot \sigma_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda \cdot \sigma_{r1} & \lambda \cdot \sigma_{r2} & \dots & \sigma_r^2 \end{bmatrix}$$

Tree Transformations

$\delta = 0.3$

$\kappa = 0.3$



$$\mathbf{C}_\delta = \begin{bmatrix} (\sigma_1^2)^\delta & (\sigma_{12})^\delta & \dots & (\sigma_{1r})^\delta \\ (\sigma_{21})^\delta & (\sigma_2^2)^\delta & \dots & (\sigma_{2r})^\delta \\ \vdots & \vdots & \ddots & \vdots \\ (\sigma_{r1})^\delta & (\sigma_{r2})^\delta & \dots & (\sigma_r^2)^\delta \end{bmatrix}$$

# Model adequacy



VOL. 186, NO. 2 THE AMERICAN NATURALIST AUGUST 2015

E-ARTICLE

## Model Adequacy and the Macroevolution of Angiosperm Functional Traits

Matthew W. Pennell,<sup>1,\*</sup> Richard G. FitzJohn,<sup>2</sup> William K. Cornwell,<sup>3</sup> and Luke J. Harmon<sup>1</sup>

1. Department of Biological Sciences and Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, Idaho 83844;  
2. Department of Biological Sciences, Macquarie University, Sydney, New South Wales 2109, Australia; 3. School of Biological, Earth,  
and Environmental Sciences, University of New South Wales, Sydney, New South Wales 2052, Australia

*Submitted April 7, 2014; Accepted March 31, 2015; Electronically published June 12, 2015*

*Online enhancement:* supplemental PDF.

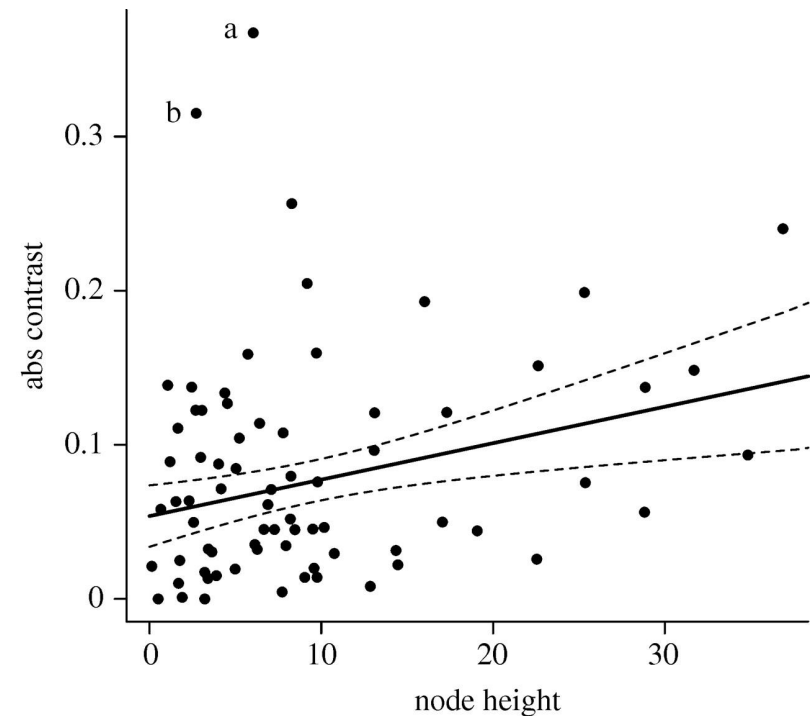
# Node-Height test

Slope = 0: BM prediction

Positive - OU-like

Negative - Early burst like

Outliers - Not BM!





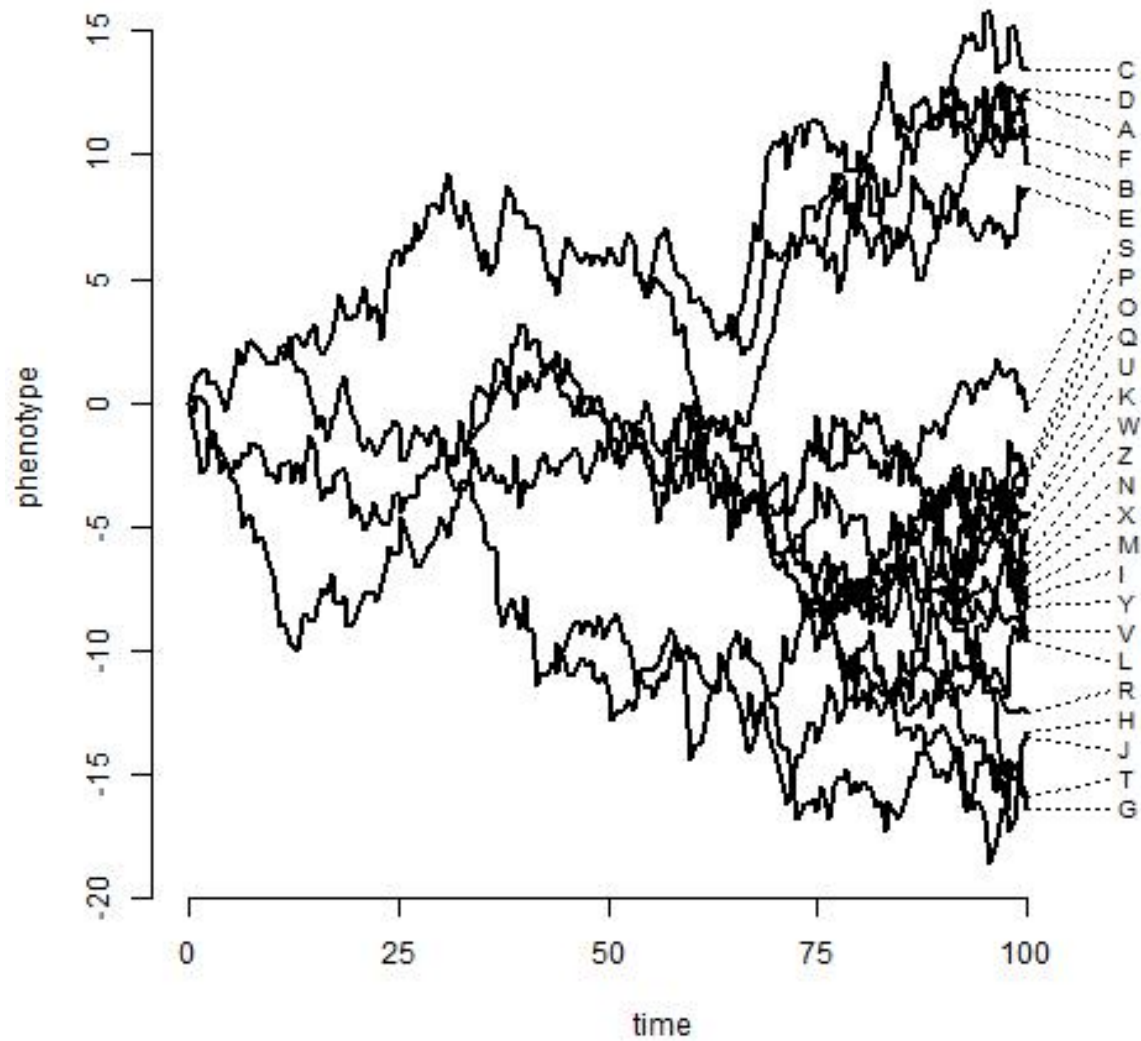
# What will be the effects if:

You don't have a fully resolved tree?

You have a tree with error?

You have Y values with error?

You have X values with error?



Give me your best estimate of the evolutionary rate (back of the napkin calculation)