# Phylogenetic Independent Contrasts in R

Casey Dunn

April 8, 2013

## Contents

## 1 Introduction

This document introduces you to a variety of tools and approaches to analyzing phylogenetic data with R.

There is detailed information on phylogenetic analysis with R at both the R phylo wiki and the CRAN task page for phylogenetics. Also consult the excellent book by Paradis (2012).

Make sure that you have the most recent version of R installed. For the analyses below, you will need to install the following libraries in R:

- ape

- geiger

- phylobase

Update them if they are not the latest version.

## 2 Simulations

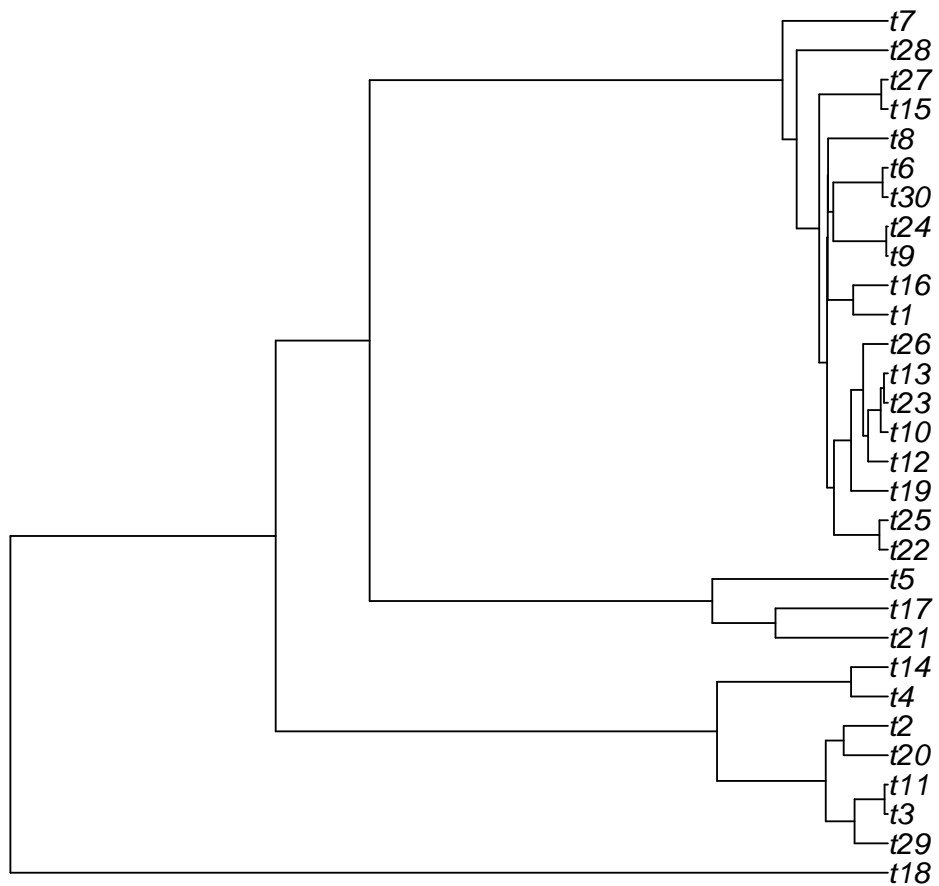### 2.1 Simulating trees

First, load ape:

```
library(ape)
```

The simplest way to generate a random tree is with the rcoal() function:

```
t <- rcoal(30)
plot(t)
```



The only required argument is the number of taxa. rcoal() generates ultrametric trees, the related function rtree() can be used to generate trees whose tips are not contemporaneous.

## 2.2   Simulating phenotypic data on trees

The sim.char() function from the geiger library is a convenient tool for simulating character evolution.

We'll simulate three characters, each with a variance of 1. The fist two characters will have a covariance of 0.7 with each other. The third characters will have a covariance of 0 with the other characters. First, make a variance-covariance matrix describes this relationship:

```
vcv <- diag(3)
vcv[1, 2] <- 0.7
vcv[2, 1] <- 0.7
vcv

#      [,1] [,2] [,3]
# [1,]  1.0  0.7    0
# [2,]  0.7  1.0    0
# [3,]  0.0  0.0    1
```

Now, generate multiple simulated datasets on tree t according to vcv:

```
library(geiger)
D <- sim.char(t, vcv, nsims = 100, model = "brownian", root.state = 0)
D[, , 1]

#           [,1]     [,2]     [,3]
# t18 -0.49165 -0.12393 -0.07618
# t29 -0.73760 -0.40257  0.40294
# t3  -0.42216  0.10705  0.14649
# t11 -0.48251  0.09523  0.18946
# t20 -0.50808 -0.11292  0.25305
# t2  -0.58485 -0.12972  0.07370
# t4  -0.40454  0.84516  0.18182
# t14 -0.65656  0.43805  0.50867
# t21  0.47845 -0.15693 -0.12136
# t17 -0.41837 -0.52663  0.51227
# t5   0.49333  0.31534  0.97985
# t22 -0.58436 -0.85311  0.60058
# t25 -0.49611 -0.63954  0.54087
# t19  0.08629 -0.48944  0.60645
# t12 -0.24220 -0.71394  0.64994
# t10 -0.98530 -1.22016  0.75746
# t23 -0.71701 -0.71615  0.78872
# t13 -0.65692 -0.74032  0.88594
# t26 -0.49009 -0.64183  0.64411
# t1  -0.40792 -1.25974  0.67816
# t16  0.05793 -0.79269  0.88397
```

Independent contrasts

```
# t9  -0.50197 -0.86244  0.57600
# t24 -0.50057 -0.94100  0.58466
# t30 -0.32149 -0.80582  0.39914
# t6  -0.25654 -0.69368  0.31728
# t8  -0.07820 -1.06217  0.69155
# t15  0.17422 -0.39603  0.83282
# t27  0.01495 -0.35081  0.81576
# t28 -0.04190 -0.85619  0.70131
# t7   0.12225 -0.85276  0.56476
```

This shows just the first simulation.

You can plot the data right onto the tree. The adephylo library has some nice tools for plotting miltivariate data, but we first need to combine the tree and data into a phylo4d object:

```
library(adephylo)
D4 <- phylo4d(t, D[, , 1])
table.phylo4d(D4, box = FALSE)
```

## 3  Independent contrasts

Now construct independent contrasts for each of the three variables:

```
ic1 <- pic(D[, 1, 1], t)
ic2 <- pic(D[, 2, 1], t)
ic3 <- pic(D[, 3, 1], t)

cov(cbind(ic1, ic2, ic3))
```

Independent contrasts

```
#          ic1      ic2      ic3
# ic1  0.92763  0.62642 -0.06047
# ic2  0.62642  1.04380 -0.08868
# ic3 -0.06047 -0.08868  0.28707
```

This matrix is based on a single simulation, and will differ by chance from the original vcv matrix that we defined above.

# 4   Phylogenetic signal

Many studies use $K$ (Blomberg et al., 2003) to assess the "phylogenetic signal" of a character. Under a Brownian motion model on a phylogenetic tree, a $K$ of 1 is expected. For $K < 1$, relatives resemble each other less than expected. For $K > 1$, relatives resemble each other more than expected.

This measure has been implemented in the picante library.

```
library(picante)
Kcalc(D[, 1, 1], t)
```

```
#         [,1]
# [1,] 0.1467
```

Now take a look at $K$ for a variable that has been simulated without consideration of the tree:

```
ntips <- length(t$tip.label)
x <- rnorm(ntips, mean = 0, sd = 1)
names(x) <- names(D[, 1, 1])
Kcalc(x, t)
```

```
#         [,1]
# [1,] 0.02887
```

Neither of these values are exactly 1, but it is not clear if the difference is significant. This can be addressed via randomization tests as implemented by the phylosignal() function

```
phylosignal(D[, 1, 1], t)
```

```
#       K PIC.variance.obs PIC.variance.rnd.mean PIC.variance.P
# 1 0.1467           0.9416                  5.74          0.001
#   PIC.variance.Z
# 1         -1.782
```

```
phylosignal(x, t)
```

```
#        K PIC.variance.obs PIC.variance.rnd.mean PIC.variance.P
# 1 0.02887            41.43                 39.87          0.631
#   PIC.variance.Z
# 1        0.08261
```

## 5 How this document was made

This document is a computable data report compiled directly from the data. To recreate this file from the data, you will need to install:

- R (`http://www.r-project.org`). This document was generated with version 2.15.2.

- The R package knitr (`http://yihui.name/knitr/`), which can be installed from within R. This document was generated with version 0.9.

- pdflatex, which comes with LaTeX distributions (`http://www.latex-project.org/ftp.html`). This document was generated with version 3.1415926-2.4-1.40.13.

From within the knitr directory, launch R and run:

```
library(knitr)
knit("comparative_r.Rnw")
quit()
```

This will generate a new tex file. To compile this tex file into a pdf file, run the following at the shell command line:

```
pdflatex comparative_r.tex
```

In addition to recreating this document as-is, you can directly edit and add to the analyses in the .Rnw source file. You can also copy the R source code from the .Rnw file.

## References

Blomberg, S. P., T. Garland, and A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57:717–745.

Paradis, E. 2012. Analysis of Phylogenetics and Evolution With R. Springer Science+Business Media.