Errata for *Analysis of Phylogenetics and Evolution with R—Second Edition*

Page 6: In Table 1.1, distory is also discussed in Chapter 3.

Page 15: Using out-of-range, negative indices does not result in an error with recent versions of R, so the second bloc of code should be:

```
> z[5]
[1] NA
> z[-5]
[1] "order"   "family"  "genus"   "species"
```

Page 33: for a a set ↯ for a set (in footnote)

The output of `designSplits` (in phangorn) has been simplified, so the bloc at the end of the page is now:

```
> designSplits(4)
     [,1] [,2] [,3] [,4] [,5] [,6] [,7]
[1,]    1    1    0    0    1    1    0
[2,]    1    0    1    1    0    1    0
[3,]    1    0    1    0    1    0    1
[4,]    0    1    1    1    1    0    0
[5,]    0    1    1    0    0    1    1
[6,]    0    0    0    1    1    1    1
```

Page 35: In Table 3.1, Alignement ↯ Alignment

Page 40: The function `read.PDB` has been renamed `read_protdb`. The package Rpdb (on CRAN) offers an alternative to handle PDB files.

Page 44: The trees of the Pfam database have been moved to another web site. The example given on this page should now be:

```
> a <- "http://pfam.xfam.org/family/"
> b <- "PF01607/tree/download"
> ref <- paste0(a, b)
> tr <- read.tree(ref)
> tr

Phylogenetic tree with 209 tips and 207 internal nodes.

Tip labels:
        Q9VRL7_DROME/658-711, Q9VTR4_DROME/207-261, Q9VRL7_DROME/482-536,
        Q9VTR4_DROME/367-421, Q9VTR4_DROME/293-347, Q8SXL5_DROME/104-158, ...
Node labels:
        , 0.600, 0.450, 0.930, 0.830, 0.860, ...
```

```
Unrooted; includes branch lengths.
```

Page 44: Last sentence of Section 3.2: Exercices ↝ Exercises

Page 61: There is now a method to convert from the class `"dendrogram"` to the class `"hclust"` (in the package stats). So, the sentence right before Table 3.3 should be deleted.

Page 66: `data(woodmouse)` should be added before `x <- woodmouse[, 1:50]`.

Page 70: The two lines of command:

```
> save(sylvia.clu, taxa.sylvia, sylvia.eco,
+      file = "sylvia.RData")
```

should be (the 's' at the end of the first object name is missing):

```
> save(sylvia.clus, taxa.sylvia, sylvia.eco,
+      file = "sylvia.RData")
```

Page 77: sequence lenght ↝ sequence length

Page 102: can plotted ↝ can be plotted

Page 127: The conversion from the numeric matrix `X` into a data frame with factors is now (i.e., with a recent version of R) more complicated. The command (in the middle of the page):

```
> Y <- as.data.frame(apply(X, 2, factor))
```

returns a data frame with character vectors instead of factors. It should be replaced by:

```
> Y <- as.data.frame(X)
> for (i in seq_along(Y)) Y[[i]] <- factor(Y[[i]])
```

Page 128: alignement ↝ alignment; (on the next line:) detaild ↝ detailed

Page 131: addivity ↝ additivity

Page 132: this sentence about ultrametric distances is wrong:

> Consider the case where there are only three observations: it is clear that the above condition is met only if all distances are equal.

It should be replaced by:

> Consider the case where there are only three observations: it is clear that the above condition is met only if, either all distances are equal, or the two largest distances are equal.

The next sentence may be slightly modified by changing its first word: "In the former case, if we represent these three observations with a tree, . . . "

Page 149: This part of the R code and output:

```
> pml(tr, x, rate = .05/.25)

 loglikelihood: -5.288237
```

must be changed to:

```
> pml(tr, x, rate = 3 * 0.05)

 loglikelihood: -5.367834
```

This is because the rate of change is three times α (= 0.05), and there is no need to multiply this value by the base frequencies (π). Thus, the next four lines (top of p. 150) should be removed and replaced by: "If we use (5.10) above we find -5.367834 which is very close.". (*Many thanks to Don Klinkenberg for reporting this error.*)

Page 150: The call to `pml` now produces a different message:

```
> o1 <- pml(tw, x)
negative edges length changed to 0!
```

Page 165: lot work ↝ lot of work (fifth line)

Page 166: The command:

```
> pars.tr <- optim.parsimony(tr, x)
```

should be (`tr` ↝ `tw`):

```
> pars.tr <- optim.parsimony(tw, x)
```

Also, `optim.parsimony()` does not return branch lengths anymore, so the following command now returns `NULL`. In such a case, branch lengths can be calculated with the function `acctran` in `phangorn`:

```
> tw.acc <- acctran(tw, x)
> tw.acc$edge.length
 [1] 5.0000000 9.0000000 1.0000000 1.0000000 1.0000000
 [6] 0.0000000 3.0000000 1.0000000 1.0000000 6.0000000
[11] 2.0000000 3.0000000 2.0000000 3.0000000 3.0000000
[16] 0.0000000 2.0000000 3.0000000 4.5000000 4.5000000
```

```
[21] 2.0000000 1.0000000 2.0000000 5.0000000 1.3333333
[26] 1.3333333 0.3333333
> sum(tw.acc$edge.length)
[1] 68
```

Page 174: Printing objects of class `"prop.part"` has been improved, so the output at the bottom of the page should be:

```
> prop.part(tr)
1: a
2: b
3: c
4: d

==> 1 time(s):[1] 1 2 3 4
==> 1 time(s):[1] 1 2 3
==> 1 time(s):[1] 2 3
```

Page 175: The function `prop.clades` has now the option `rooted = FALSE`, so the output in the middle of the page depends on the value of this option (since `tr` is a rooted tree):

```
> prop.clades(tr, tr, rooted = TRUE)
[1] 1 1 1
> prop.clades(tr, tr)
[1]  1 NA  1
```

Page 175: So there are 108 different splits ↯ So there are 90 different splits

This number is actually random: 100 replications of the same bootstrap command resulted in the following summaries of the number of splits:

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  82.00   90.00   95.00   95.03  100.00  115.00
```

Page 178: `dist.topo` now checks whether the trees are rooted or unrooted and issues a warning in the former case. So, the output of the commands in the middle of this page is now:

```
> tr <- read.tree(text = "((a,b),(c,d));")
> tb <- read.tree(text = "((a,d),(c,b));")
> dist.topo(tr, tb)
      tree1
tree2     4
Warning message:
In dist.topo(tr, tb) :
```

```
  Some trees were rooted: topological distances may be spurious.
> dist.topo(tr, tr)
      tree1
tree2     0
Warning message:
In dist.topo(tr, tr) :
  Some trees were rooted: topological distances may be spurious.
```

An example using unrooted trees could be:

```
> tr <- read.tree(text = "(a,b,(c,d));")
> tb <- read.tree(text = "(a,d,(c,b));")
> dist.topo(tr, tb)
      tree1
tree2     2
> dist.topo(tr, tr)
      tree1
tree2     0
```

Page 183: mean path lenghts ↝ mean path lengths

Page 190: The object `sylvia.seq.ali` has not be created yet; thus, the command:

```
sylvia.seq.ali <- sylvia.clus
```

should be added before calculating the distances.

Page 198: brances ↝ branches

Page 292: the equation of $\hat{D}_S$ should be (the $\sum$ was forgotten):

$$\hat{D}_S = \frac{1}{N} \sum_{k=1}^{N} \hat{D}_k$$

Page 324: branch lenght ↝ branch length

Page 360: CA or CT ↝ CA or TA

Page 363: Francois O. ↝ François O. (ref. 30)

Pages 366–367: References 81, 87, and 110 have been badly formatted; they should be:

[81] Felsenstein J. 2005. Using the quantitative genetic threshold model for inferences between and within species. *Philosophical Transactions of the Royal Society of London. Series B. Biological Sciences* **360:** 1427–1434.

[87] Ford D. J. 2010. Encodings of cladograms and labeled trees. *Electronic Journal of Combinatorics* **17:** R54.

[110] Goldberg E. A., Lancaster L. T. & Ree R. H. 2011. Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology* doi:10.1093/sysbio/syr046.

Page 376: Mooers A. O. ↝ Mooers A. Ø. (ref. 263)

June 27, 2023

Thanks to Daniel Hoffmann for reporting some typos, and to Haruo Suzuki for reporting some updates.