

1 Supplementary Information
2 *MonoPhy*: A simple R package to find
3 and visualize monophyly issues -
4 Commands for Examples

5
6 Orlando Schwery*¹ and Brian C. O'Meara¹

7
8 ¹Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville,
9 TN 37996-1610, USA

10
11
12 Below you find the commands used to run the two example cases from the paper.
13 We included comments and outputs, except where outputs are too large to be
14 displayed in this word file, in which case the separate file will be referred to. For a
15 more general treatment of usage and applications of the package, one should refer
16 to the tutorial vignette (`vignette("MonoPhyVignette")`), which contains stepwise
17 instructions for the different functions and their options. For further package details
18 consult the help files (`help("MonoPhy")`).

19 **Example 1 – Ericaceae**

20 For the first demonstration, we use the example files contained in the package. The
21 phylogeny is of the plant family Ericaceae, as published in Schwery *et al.* (2015)
22 pruned down to 77 taxa (original data see Schwery *et al.* (2014). Two taxon files
23 assign tribes or tribes and subfamilies to the tips (in both files, errors in the taxon
24 assignment have been introduced for the sake of demonstrating the functionality of
25 the package).

26 First we load the package:

27 `> library(MonoPhy)`

28

29 Next, we load the example files with the tree and the taxonomy table:

30 `> data(Ericactree)`

31 `> data(Ericactribes)`

* Correspondence author. Email: oschwery@vols.utk.edu

32 Now we can run the main analysis command. In order to access the output later,
33 we should save it to an object. We will run one analysis without taxonomy table,
34 assessing the monophyly of the genera (`solutionG`), and one with the table
35 assessing monophyly of the tribes (`solutionT`):

```
36 > solutionG <- AssessMonophyly(Ericactree)  
37 > solutionT <- AssessMonophyly(Ericactree, Ericactribes)  
38
```

39 Running these steps took around 0.045 and 0.093 seconds respectively on a
40 MacBook Pro with 2.4 GHz Intel Core i5 and 8GB Ram. Now we can get a first
41 impression of the monophyly in our tree. Note that 'Taxa' refers to higher order
42 taxa:

```
43 > GetSummaryMonophyly(solutionG)
```

```
44 $Taxlevel_1
```

| | Taxa Tips | |
|------------------|-----------|----|
| Total | 32 | 77 |
| Monophyletic | 9 | 25 |
| Non-Monophyletic | 6 | 35 |
| Monotypic | 17 | 17 |
| Intruder | 7 | 11 |
| Outlier | 1 | 2 |

```
52
```

```
53 > GetSummaryMonophyly(solutionT)
```

```
54 $Taxlevel_1
```

| | Taxa Tips | |
|------------------|-----------|----|
| Total | 10 | 77 |
| Monophyletic | 7 | 44 |
| Non-Monophyletic | 2 | 32 |
| Monotypic | 1 | 1 |
| Intruder | 1 | 2 |
| Outlier | 1 | 1 |

```
62
```

63 We can see that on both taxonomic levels monophyly issues exist to different
64 extents. For more detailed information, we can now look at the results tables for
65 genera (**Table 1**) and tribes (**Table 2**).

```
66 > GetResultMonophyly(solutionG)
```

```
67 > GetResultMonophyly(solutionT)
```

```
68
```

69 We can see that the genera *Dimorphanthera*, *Erica*, *Paphia*, *Phyllodoce*, *Pyrola* and
70 *Vaccinium* are not monophyletic as well as the tribes Vaccinieae and Andromedeae.

71 If we want to know which other taxa particularly cause the above to be non-
72 monophyletic, we can list them:

```

73 > GetIntruderTaxa(solutionG)
74 $Taxlevel_1
75 $Taxlevel_1$Dimorphanthaera
76 [1] "Paphia"
77
78 $Taxlevel_1$Erica
79 [1] "Bruckenthalia"
80
81 $Taxlevel_1$Paphia
82 [1] "Dimorphanthaera"
83
84 $Taxlevel_1$Phyllodoce
85 [1] "Kalmiopsis"
86
87 $Taxlevel_1$Pyrola
88 [1] "Orthilia"
89
90 $Taxlevel_1$Vaccinium
91 [1] "Gaylussacia" "Agapetes"
92
93 > GetIntruderTaxa(solutionT)
94 $Taxlevel_1
95 $Taxlevel_1$Vaccinieae
96 [1] "Andromedeae"
97

```

98 If we want to figure out which species exactly are the intruders, we can extract
99 this information as well, *e.g.* for the genus *Vaccinium* alone and for both of the non-
100 monophyletic tribes:

```

101 > GetIntruderTips(solutionG, taxa="Vaccinium")
102 $Taxlevel_1
103 $Taxlevel_1$Vaccinium
104 [1] "Gaylussacia_dumosa" "Agapetes_hosseana" "Agapetes_serpens"
105 "Agapetes_buxifolia"
106
107 > GetIntruderTips(solutionT)
108 $Taxlevel_1
109 $Taxlevel_1$Vaccinieae
110 [1] "Andromeda_polifolia" "Vaccinium_vitis_idaea"
111

```

112 A look at the intentional errors in the taxonomy table would make clear why
113 these two species are intruders of the tribe Vaccinieae: the labels of *Vaccinium vitis-*
114 *idaea* (a member of Vaccinieae) and *Zenobia pulverulenta* (an actual Andromedeae)

115 were swapped, which makes it appear like *V. vitis-ideae* and the remaining actual
116 Andromedeae – *Andromeda polifolia* – would intrude the Vaccinieae clade.

117 Under the default cutoff (0.5), the genus *Vaccinium* has outliers. Analogous to the
118 intruders, we can display the involved taxa (which we already know in this case of
119 course) and the specific species:

```
120 > GetOutlierTaxa(solutionG)
121 $Taxlevel_1
122 [1] "Vaccinium"
123
124 > GetOutlierTips(solutionG)
125 $Taxlevel_1
126 $Taxlevel_1$Vaccinium
127 [1] "Vaccinium_poasanum" "Vaccinium_uliginosum"
128
```

129 Note that in this case, without a cutoff, all tips ‘in between’ these two outliers and
130 the remaining *Vaccinium* species would have been considered ‘intruders’ of
131 *Vaccinium*.

132 For the purpose of orientation on the tree, plotting or any other further
133 operations, we can display the MRCA nodes of any taxon desired, *e.g.* *Vaccinium*
134 again:

```
135 > GetAncNodes(solutionG, taxa="Vaccinium")
136 $Taxlevel_1
137 MRCA
138 Vaccinium 93
139
```

140 Finally, we can visualize the results using the plot function. The following
141 command will produce the basic monophyly plot, which is displayed in **Fig. S1**.

```
142 > PlotMonophyly(solutionG, Ericactree, plot.type="monophyly")
143
```

144 **Example 2 – Embriophyta**

145 For the second example, we demonstrate the package’s performance on a tree with
146 several thousand tips. We therefore run the analysis on a tree of 31,749 species of
147 Embriophyta, as published by Zanne *et al.* (2014) (data see Zanne *et al.* (2013)), this
148 times with a higher outlier-cutoff:

```
149 > library(MonoPhy)
150 > tree <- read.tree(file=~/.~/Vascular_Plants_rooted.dated.tre")
151 > solution <- AssessMonophyly(tree, outlierlevel=0.9)
152
```

153 On the same computer as in the previous example, this analysis completed after
154 1.78 hours. Again, we get a first impression by looking at the summary and results
155 tables (due to its size, the result table can be found in **Tab. S1**, see separate file).

```
156 > GetSummaryMonophyly(solution)  
157 $Taxlevel_1
```

| | Taxa | Tips |
|----------------------|------|-------|
| 158 Total | 8179 | 31749 |
| 159 Monophyletic | 1954 | 10841 |
| 160 Non-Monophyletic | 1813 | 16496 |
| 161 Monotypic | 4412 | 4412 |
| 162 Intruder | 335 | 498 |
| 163 Outlier | 1703 | 5275 |

```
164  
165  
166 > GetResultMonophyly(solution)  
167
```

168 Just as before, we could extract more detailed information from the output. Due
169 to the size of the phylogeny, it would not be convenient to inspect the plot in an R
170 plot window. Therefore, we choose to save the plot directly as a PDF, which will
171 make it easier to view. Additionally, we can limit the complexity of the plot, by
172 collapsing all clades for monophyletic taxa to one remaining branch. We create the
173 standard monophyly plot with and without collapsing the monophyletic groups
174 (**Fig. S2** and **Fig. S3** respectively), and the intruder plot without collapsing (**Fig. S4**),
175 colour-coding the intruders and outliers according to the taxa that they are assigned
176 to.

```
177 > PlotMonophyly(solution, tree, plot.type='monophyly', monocoll=TRUE,  
178 label.offset=1, type="fan", cex=0.05, edge.width=1.5, PDF=TRUE,  
179 PDF_filename='MEE_Fig2.pdf')  
180 > PlotMonophyly(solution, tree, plot.type='monophyly',  
181 monocoll=FALSE, label.offset=0.75, type="fan", cex=0.05,  
182 edge.width=1.5, PDF=TRUE, PDF_filename='SuppFig1.pdf')  
183 > PlotMonophyly(solution, tree, plot.type='intruders',  
184 monocoll=FALSE, label.offset=0.75, type="fan", cex=0.05,  
185 edge.width=1.5, PDF=TRUE, PDF_filename='SuppFig2.pdf')
```

186 Note that plotting such big trees can take as long as or longer than running the
187 analysis itself. Here, **Fig. S2** took 1.22 hours, **Fig. S3** 2.39 hours and **Fig. S4** 4.53
188 hours.

189 **References**

- 190 Schwery, O., Onstein, R.E., Bouchenak-Khelladi, Y., Xing, Y., Carter, R.J. & Linder, H.P.
191 (2014) Data from: As old as the mountains: the radiations of the Ericaceae.
192 Dryad Data Repository.
- 193 Schwery, O., Onstein, R.E., Bouchenak-Khelladi, Y., Xing, Y., Carter, R.J. & Linder, H.P.
194 (2015) As old as the mountains: the radiations of the Ericaceae. *New*
195 *Phytologist*, **207**, 355-367.
- 196 Zanne, A.E., Tank, D.C., Cornwell, W.K., Eastman, J.M., Smith, S.A., FitzJohn, R.G.,
197 McGlinn, D.J., O'Meara, B.C., Moles, A.T., Reich, P.B., Royer, D.L., Soltis, D.E.,
198 Stevens, P.F., Westoby, M., Wright, I.J., Aarssen, L., Bertin, R.I., Calaminus, A.,
199 Govaerts, R., Hemmings, F., Leishman, M.R., Oleksyn, J., Soltis, P.S., Swenson,
200 N.G., Warman, L. & Beaulieu, J.M. (2014) Three keys to the radiation of
201 angiosperms into freezing environments. *Nature*, **506**, 89-+.
- 202 Zanne, A.E., Tank, D.C., Cornwell, W.K., Eastman, J.M., Smith, S.A., FitzJohn, R.G.,
203 McGlinn, D.J., O'Meara, B.C., Moles, A.T., Reich, P.B., Royer, D.L., Soltis, D.E.,
204 Stevens, P.F., Westoby, M., Wright, I.J., Aarssen, L., Bertin, R.I., Calaminus, A.,
205 Govaerts, R., Hemmings, F., Leishman, M.R., Oleksyn, J., Soltis, P.S., Swenson,
206 N.G., Warman, L., Beaulieu, J.M. & Ordonez, A. (2013) Data from: Three keys
207 to the radiation of angiosperms into freezing environments. Dryad Data
208 Repository.
209

210 **Tables**

211 **Table 1: *MonoPhy* result table for the genera of Ericaceae.**

| | Monophyly | MRCA | #Tips | Delta-Tips | #Intruders | Intruders | #Outliers | Outliers |
|----------------|-----------|------|-------|------------|------------|--------------------------|-----------|---------------------------|
| Agapetes | Yes | 114 | 3 | 0 | 0 | | | NA |
| Allotropa | Monotypic | NA | 1 | NA | NA | | | NA |
| Andromeda | Monotypic | NA | 1 | NA | NA | | | NA |
| Bruckenthalia | Monotypic | NA | 1 | NA | NA | | | NA |
| Bryanthus | Monotypic | NA | 1 | NA | NA | | | NA |
| Calluna | Monotypic | NA | 1 | NA | NA | | | NA |
| Cassiope | Yes | 146 | 5 | 0 | 0 | | | NA |
| Chimaphila | Yes | 88 | 2 | 0 | 0 | | | NA |
| Costera | Monotypic | NA | 1 | NA | NA | | | NA |
| Daboecia | Yes | 126 | 2 | 0 | 0 | | | NA |
| Dimorphanthera | No | 96 | 3 | 2 | 1 | Paphia | | 0 |
| Enkianthus | Yes | 150 | 5 | 0 | 0 | | | NA |
| Epigaea | Monotypic | NA | 1 | NA | NA | | | NA |
| Erica | No | 128 | 9 | 1 | 1 | Bruckenthalia | | 0 |
| Gaylussacia | Monotypic | NA | 1 | NA | NA | | | NA |
| Gonocalyx | Yes | 120 | 2 | 0 | 0 | | | NA |
| Kalmiopsis | Yes | 142 | 2 | 0 | 0 | | | NA |
| Ledothamnus | Yes | 138 | 2 | 0 | 0 | | | NA |
| Moneses | Monotypic | NA | 1 | NA | NA | | | NA |
| Monotropa | Monotypic | NA | 1 | NA | NA | | | NA |
| Notopora | Monotypic | NA | 1 | NA | NA | | | NA |
| Orthaea | Yes | 101 | 2 | 0 | 0 | | | NA |
| Orthilia | Monotypic | NA | 1 | NA | NA | | | NA |
| Oxydendrum | Monotypic | NA | 1 | NA | NA | | | NA |
| Paphia | No | 97 | 2 | 1 | 1 | Dimorphanthera | | 0 |
| Phyllodoce | No | 140 | 4 | 2 | 1 | Kalmiopsis | | 0 |
| Pyrola | No | 83 | 4 | 1 | 1 | Orthilia | | 0 |
| Rhodothamnus | Monotypic | NA | 1 | NA | NA | | | NA |
| Symphysia | Monotypic | NA | 1 | NA | NA | | | NA |
| Utleya | Monotypic | NA | 1 | NA | NA | | | NA |
| Vaccinium | No | 93 | 13 | 17 | 2 | Gaylussacia, Agapetes | | 2 Vaccinium_uliginosum |
| Zenobia | Monotypic | NA | 1 | NA | NA | | | NA |

212

213 **Table 2: *MonoPhy* result table for the tribes of Ericaceae.**

| | Monophyly | MRCA | #Tips | Delta-Tips | #Intruders | Intruders | #Outliers | Outliers |
|----------------|-----------|------|-------|------------|------------|-------------|-----------|-----------------------|
| Vaccinieae | No | 91 | 30 | 2 | 1 | Andromedeae | 0 | |
| Monotropeae | Yes | 81 | 2 | 0 | 0 | | NA | |
| Andromedeae | No | 91 | 2 | 30 | 0 | | 1 | Vaccinium_vitis_idaea |
| Ericaceae | Yes | 125 | 13 | 0 | 0 | | NA | |
| Bryanthaeae | Yes | 137 | 3 | 0 | 0 | | NA | |
| Cassiopoideae | Yes | 146 | 5 | 0 | 0 | | NA | |
| Pyroleae | Yes | 82 | 8 | 0 | 0 | | NA | |
| Enkianthoideae | Yes | 150 | 5 | 0 | 0 | | NA | |
| Phyllodoceae | Yes | 139 | 8 | 0 | 0 | | NA | |
| Oxydendreae | Monotypic | NA | 1 | NA | NA | | NA | |

214