

1 ***JNplots*: an R package to visualize outputs from the Johnson-Neyman**
2 **technique for categorical and continuous moderators, including options for**
3 **phylogenetic regressions.**

4
5 Ken S. Toyama

6
7 Department of Ecology and Evolutionary Biology, University of Toronto.
8

9

10 ABSTRACT

11

12 The analysis of two-way interactions in linear models is common in the fields of ecology and
13 evolution, being often present in allometric, macroevolutionary, and experimental studies, among
14 others. However, the interpretation of significant interactions can be incomplete when limited to
15 the examination of model coefficients and significance tests. The Johnson-Neyman technique
16 represents a step forward in the interpretation of significant two-way interactions, allowing the
17 user to examine how changes in the moderator variable, it being categorical or continuous, affect
18 the significance of the relationship between the dependent variable and the predictor. Despite its
19 implementation in several software since its initial development, the available options to perform
20 the method lack certain functionality aspects, including the visualization of regions of non-
21 significance when the moderator is categorical, the implementation of phylogenetic corrections,
22 and more intuitive graphical outputs. Here I present the R package *JNplots*, which aims to fill
23 gaps left by previous software regarding the calculation and visualization of regions of non-
24 significance when fitting two-way interaction models. *JNplots* includes two basic functions
25 which allow the user to investigate different types of interaction models, including cases where
26 the moderator variable is categorical or continuous. The user can also specify whether the model
27 to explore should be phylogenetically informed and choose a particular phylogenetic correlation
28 structure to be used. Finally, the functions of *JNplots* produce plots that are largely customizable
29 and allow a more intuitive interpretation of the interaction term. Here I provide a walkthrough on
30 the use of *JNplots* using three different examples based on empirical data, each representing a
31 different common scenario in which the package can be useful. Additionally, I present the
32 different customization options for the graphical outputs of *JNplots*.

33 INTRODUCTION

34

35 The analysis of two-way interactions in linear models (i.e., models of the form: *dependent*
36 *variable ~ predictor * moderator*) is common in the fields of ecology and evolution (Hilborn and
37 Stearns, 1982; Dochtermann and Jenkins, 2011; Spake et al., 2023). For example, one might be
38 interested in the interaction between sex and size when examining the ontogenetic allometry of a
39 trait in a population or species (*do males and females of a given species show different*
40 *allometries?*). Macroevolutionary studies also necessitate the analysis of interactions. For
41 instance, one could be interested in testing whether a set of species follows Bergmann's rule (i.e.,
42 a positive association between body size and latitude; Bergmann, 1847), but also whether this
43 hypothetical relationship changes depending on the degree of precipitation these species
44 experience (*do different levels of precipitation affect the relationship between latitude and size*
45 *across species?*).

46

47 These two examples illustrate four different characteristics of models in ecology and
48 evolutionary biology that may involve two-way interactions. First, the moderator (which is an
49 independent variable that modulates the effect the predictor has on the dependent variable) can
50 be (1) categorical (e.g., sex) or (2) continuous (e.g., precipitation). Next, an interaction analysis
51 can be (3) phylogenetically-independent (e.g., the comparison of the ontogenetic allometry of
52 males and females of the same species) or can be (4) phylogenetically-informed (e.g., the
53 examination of the effect of precipitation on the evolutionary relationship between temperature
54 and size across species). In the latter case, the influence of shared evolutionary history needs to
55 be accounted for through the modification of the variance-covariance matrix of the taxa involved
56 in the analysis (Revell, 2010; Symonds and Blomberg, 2014).

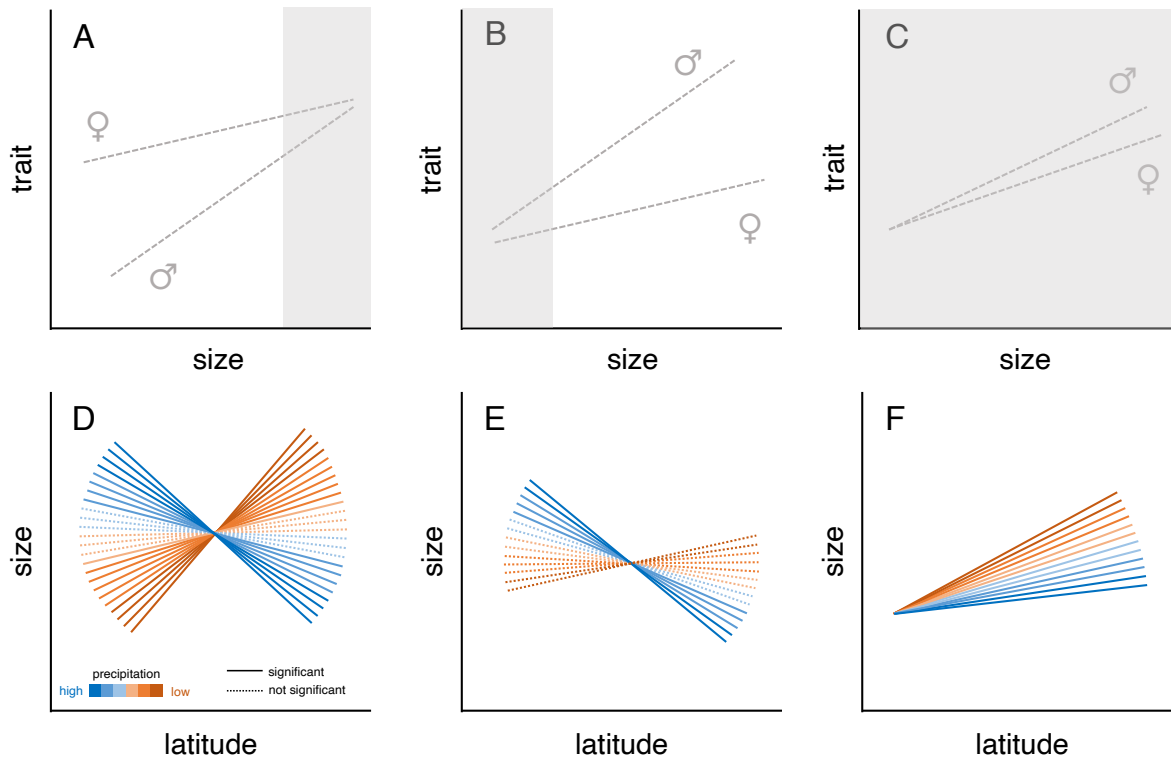
57

58 Regardless, a significant interaction term in either type of model provides evidence of an effect
59 of the moderator on the relationship between the dependent variable and the predictor. Once this
60 is confirmed, more information about the nature of the interaction can be obtained by looking at
61 the model coefficients. Let us consider the first example (*trait ~ size * sex*). If the interaction
62 term is significant and presents a positive coefficient a researcher would now know that the slope
63 of the *trait ~ size* relationship is significantly higher for one of the sexes. Similarly, if we obtain a
64 significant interaction term in the second example (*size ~ latitude * precipitation*) with a negative
65 coefficient, one could infer that the slope of the *size ~ latitude* relationship becomes more
66 negative (or less positive) as precipitation increases.

67

68 However, even with this information the significant effect of an interaction might have relevant
69 biological implications that are not immediately obvious. For example, males might have a
70 steeper allometric slope than females for a given trait based on our inference of a significant
71 interaction, but this does not eliminate the possibility that males and females might not be
72 different in shape at large (Figure 1A) or small sizes (Figure 1B), or even that they might not
73 significantly differ in shape at any size value that is biologically relevant (Figure 1C). Moreover,
74 in Figure 1A and 1B we cannot statistically conclude that, overall, one sex has relatively larger
75 trait values than the other, even though the visualization of the data suggests this is the case at
76 least for most of the size range. The reason is that the assumption of homogeneity of slopes,
77 necessary to compare groups when performing analyses of covariance (i.e., ANCOVA), is not
78 met when the interaction term is significant (Sokal and Rohlf, 2012). Similarly, precipitation

79 might have a significant and negative modulatory effect on the relationship between size and
 80 latitude. However, with this information alone one cannot know the precipitation values for
 81 which the size ~ latitude relationship is significant (Figure 1D, E). It is also possible that, despite
 82 precipitation affecting the slope of this association, the size ~ latitude relationship stays
 83 significant for all realistic precipitation values (Figure 1F). In either case, the examination of
 84 coefficients when obtaining a significant interaction term might be insufficient to interpret an
 85 interaction model and obtain useful statistical conclusions, evidencing the need for better
 86 analytical and visualization techniques.
 87



88
 89
 90 Figure 1. Hypothetical examples where the interactions between the moderator and the predictor are significant. A,
 91 B, and C depict cases where sex, a categorical moderator, influences the relationship between a given trait and size.
 92 In A and B, the trait value is larger in one sex than in the other, but because of the slope difference between sexes
 93 the difference in trait values might be non-significant at some unknown values of size (here depicted as grey
 94 regions). In C, the interaction between sex and size is also significant, but the different slopes do not result in
 95 significant differences between sexes for any biologically relevant value of size. In D, E, and F, precipitation, a
 96 continuous moderator, has a significant effect on the relationship between size and latitude. In D, the relationship
 97 between size and latitude is significant at high (blue) and low (brown), but not intermediate levels of precipitation
 98 (dotted lines). In E, the association between size and latitude is only significant and negative when precipitation is
 99 high. In F, the relationship between size and latitude stays positive and significant regardless of precipitation level,
 100 although it significantly influences the slope of the relationship. In D–F, solid and dotted lines represent significant
 101 and not significant size ~ latitude associations, respectively.

102
 103
 104
 105

107

108 The Johnson-Neyman technique (Johnson and Neyman, 1936; Johnson and Fay, 1950) is a
109 method that allows a more thorough examination of interaction effects. Originally developed to
110 account for the effect of a categorical moderator, it allows the identification of a range of
111 predictor values for which the interaction between predictor and moderator results in non-
112 significant differences in the dependent variable between categories (White, 2003; Huitema,
113 2011). For instance, it would allow one to identify size values for which trait differences between
114 males and females are not significant in figures 1A and 1B (indicated by grey areas). The
115 Johnson-Neyman technique has already been employed in empirical research in ecology and
116 evolution. For example, Hünicken et al (2022) utilized the method to identify regions of non-
117 significance in their allometric analysis of two species of *Corbicula* clams. They found the clam
118 species showed different height ~ length relationships (i.e., a significant *length* * *species*
119 interaction). However, despite this significant interaction, the Johnson-Neyman technique
120 indicated that the two species differed in height only at the extremes of the length distribution,
121 while differences in height were not significant for most length values (see Figure 4D in
122 Hünicken et al., 2022).

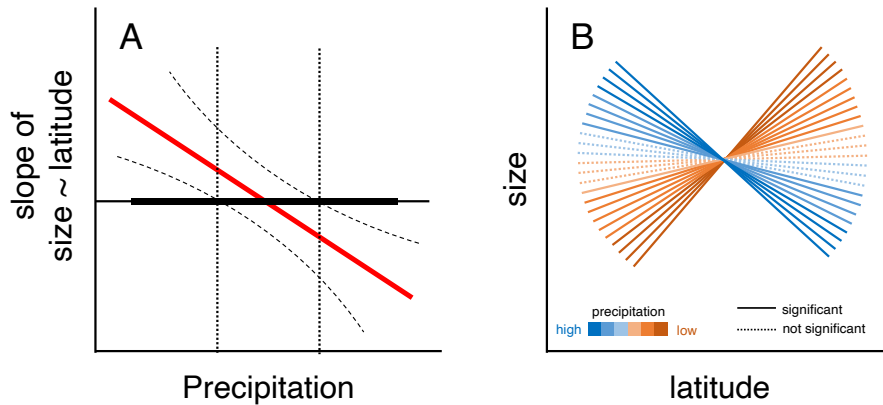
123

124 The Johnson-Neyman technique has also been expanded to account for continuous moderators
125 (Bauer and Curran, 2005). Unlike the case with categorical moderators like 'sex' or 'species', one
126 might be more interested in assessing how a gradual change in variables like precipitation or
127 temperature affect the relationship between the dependent variable and the predictor (Figure 1D–
128 F). For example, Jaime et al (2022) estimated the rates at which trees in 130 experimental plots
129 were attacked by bark beetles and how these rates were affected by the climatic distance between
130 a given plot and the niche optima of the host tree ($\text{distance}_{\text{host}}$) and that of the beetle species
131 ($\text{distance}_{\text{beetle}}$) (i.e., $\text{attack rate} \sim \text{distance}_{\text{host}} * \text{distance}_{\text{beetle}}$). The Johnson-Neyman technique
132 allowed the authors to conclude that, although attack rates decrease with $\text{distance}_{\text{host}}$, this
133 relationship weakens and even disappears as $\text{distance}_{\text{beetle}}$ values increase (see Figure S5 in Jaime
134 et al., 2022).

135

136 Despite being a relatively unknown method, a number of software have been developed to
137 perform the Johnson-Neyman technique and its expanded application for continuous moderators
138 (Preacher et al., 2006; Hayes and Matthes, 2009; Carden et al., 2017; Hayes and Montoya, 2017;
139 Montoya, 2019; Lin, 2020), the most complete being the R (R Core Team, 2021) package
140 *interactions* (Long, 2019), which includes all the functionality provided by other software and
141 overall includes a wide range of visualization and analysis options. Nonetheless, this and
142 previous software lack some functions that might prove useful for users exploring model
143 interactions. Regarding the issue of phylogenetic relatedness, previous methods do not provide
144 an option to directly incorporate phylogenetic information in the calculation of regions of non-
145 significance, limiting the use of the technique in macroevolutionary studies. Regarding
146 categorical moderators, other software do not provide an option to visualize regions of non-
147 significance (i.e., values of the predictor for which there are no significant differences between
148 categories, e.g., Figure 1A–C). Indeed, the uses of the Johnson-Neyman technique for categorical
149 moderators reported in the literature are usually based on custom-made programming scripts
150 (e.g., the study on *Corbicula* clams described above, Hünicken et al, 2022). Finally, regarding
151 the effect of continuous moderators, the function *johnson_neyman* of the R package *interactions*

152 provides a numerical output as well as a plot showing the association between the value of the
 153 moderator and the slope of the relationship between the dependent variable and the predictor.
 154 Although this type of plot (Figure 2A) resembles the one originally presented by Bauer and
 155 Curran (2005) and has been used to describe interaction effects in the literature (e.g., the study
 156 on bark beetles described above, Jaime et al., 2022), its interpretation is not straightforward
 157 because the relationship between the dependent variable and the predictor (e.g., as in Figure 2B)
 158 is not presented other than through the value of its slope.
 159



160 Figure 2. Hypothetical example of a significant interaction effect between latitude and precipitation and its effect on
 161 body size. Values of precipitation (moderator) that result in non-significant relationships between size (dependent
 162 variable) and latitude (predictor) can be obtained through an extension of the Johnson-Neyman technique. (A) The
 163 output of the method can be visualized as a plot showing the relationship between the slope of the size ~ latitude
 164 association and precipitation values (see Bauer and Curran, 2005 and the R package *interactions*, Long, 2019). Here
 165 the horizontal thin line represents a slope of zero, the horizontal thick line represents the range of precipitation data,
 166 the red line shows the negative relationship between the size ~ latitude slope and precipitation, the dashed lines
 167 represent 95% confidence intervals, and the vertical dotted lines represent the range of precipitation values that
 168 result in a non-significant size ~ latitude relationship. (B) Alternatively, the different slopes of this relationship
 169 could be illustrated in a size ~ latitude plot, showing how the relationship between size and latitude changes under
 170 the effect of different precipitation values. Both types of plots show the same information (higher precipitation
 171 decreases the value of the size ~ latitude slope, but intermediate values of precipitation result in non-significant
 172 relationships), but B is easier to interpret.
 173

174
 175 Here I present the R package *JNplots* as a solution to fill gaps left by previous software regarding
 176 the calculation and visualization of non-significance regions through the Johnson-Neyman
 177 technique. As will be explained next, *JNplots* allows the user to calculate Johnson-Neyman
 178 intervals when including categorical or continuous moderators in interaction models, and to
 179 produce graphical outputs that depict them in an intuitive way. It also allows the user to modify
 180 the correlation structure of the data, allowing the consideration of phylogenetic relationships
 181 when calculating Johnson-Neyman intervals.
 182

183 *JNplots*: IMPLEMENTATION AND EXAMPLES

184
 185 The *JNplots* R package can be used to analyse two-way interaction models that exhibit any of the
 186 four characteristics presented above (and their combinations) using the Johnson-Neyman
 187 technique and its variants. Its two basic functions, *jnt_cat* and *jnt_cont*, can be used to explore
 188 two-way interactions in which the moderator is categorical or continuous, respectively. Both

189 functions allow the analysis of phylogenetically-informed models through the use of the function
190 *gls* from the package *nlme* (Pinheiro et al., 2017). Both functions can be used to calculate and
191 visualize 'regions of significance' in an intuitive way. Finally, the functions allow plotting
192 flexibility as they include arguments that are passed on to the R base plot function. *JNplots* is
193 publicly available at <https://github.com/kenstoyama/JNplots>, and depends on the packages *ape*
194 (Paradis and Schliep, 2019), *nlme* (Pinheiro et al., 2017), and *scales* (Wickham and Seidel,
195 2022), which are downloaded from CRAN during the installation of *JNplots*. The package can be
196 installed from github using the following commands:

```
197  
198 devtools::install_github('kenstoyama/JNplots')  
199 library(JNplots)
```

201 Using the following empirical examples, I present possible scenarios in which the functions from
202 *JNplots* can be used and provide a detailed walkthrough of their implementation. All the data
203 needed to reproduce these examples are publicly available from their respective sources and are
204 also included in the installation of *JNplots*.

205 206 *Example 1: head length allometry in two lizard species*

207
208 Data from this example comes from the study of Toyama et al (2018). In the original study the
209 authors tested whether ontogenetic changes in the diet of a lizard (from insectivory to herbivory)
210 corresponded to changes in its morphology (from slender to robust heads). As part of their
211 analyses, the authors compared the head shape allometry of the semi-herbivorous species to other
212 congeners that showed mainly insectivorous habits throughout their life (see Figure 4 in Toyama
213 et al., 2018). Using their original data, I compared the relationship between head length and body
214 size in a pair of these species: *Microlophus thoracicus*, a semi-herbivore species, and *M.*
215 *peruvianus*, a species that rarely includes plant material in its diet.

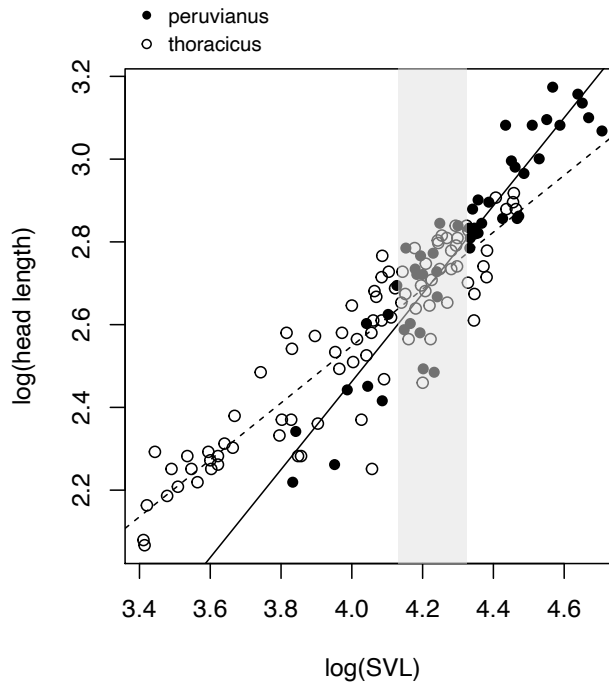
216
217 I prepared a subset of the original dataset (dataset 'microlophus', included in *JNplots*), which
218 included data on body size (measured as SVL (snout-vent-length) in millimetres), head length
219 (also in mm), and species. Measurements were log-transformed. Since the moderator in this case
220 is categorical (i.e., species), I proceeded to test for a possible two-way interaction between
221 species and size (i.e., head length ~ size * species) using the function *jnt_cat*. The only necessary
222 arguments in *jnt_cat* are the names of the predictor (X), the dependent variable (Y), and the
223 moderator (m). They are added to the function as character strings. The dataset also needs to be
224 specified:

```
225  
226 jnt_cat(X='svl', Y='hl', m='species', data=microlophus)
```

227
228 Notice that the character strings must coincide with the column names in the dataset
229 'microlophus'. These four arguments are the minimum needed for the function to work. The
230 output of the function consists in the summary table of the two-way interaction model (head
231 length ~ size * species), and the lower and upper limits of the region of non-significance (i.e.,
232 values of the predictor for which the difference between categories is not significant) (Table 1).
233 The function also produces a plot showing the association between the dependent variable (e.g.,
234 head length) and the predictor (e.g., size), with the two categories (e.g., species) represented by

235 different symbols and/or colors, and regression lines for each individual category based on the
236 output of the interaction model (Figure 3).

237
238 This re-analysis of the data using *jnt_cat* indicated that the interaction between sex and size was
239 significant ($t = -5.499, p < 0.001$), and evidenced the existence of a region of non-significance
240 along the examined size range (Figure 3). Specifically, the results indicated that the head lengths
241 of both species are not significantly different for individuals with $\log(\text{SVL})$ values between 4.13
242 and 4.33 (approximately between 62.33 and 75.63 mm).
243



244
245 Figure 3. Graphical output of a model relating head length to body size (SVL) and its interaction with species of
246 *Microlophus* lizards (model: head length ~ body size * species) obtained with the function *jnt_cat* from *JNplots*.
247 Solid and dashed lines represent head length ~ body size relationships for individuals of each of the two species
248 (also represented by closed and open circles), as shown in the legend. These relationships were obtained from the
249 output of the interaction model. Grey area represents the non-significance area calculated with the Johnson-Neyman
250 technique. Data obtained from Toyama et al (2018) and available to use with *JNplots*.
251

252 In this particular case, the calculation of regions of non-significance using *jnt_cat* provided
253 predictor (size) values that defined regions where differences between categories (and lack
254 thereof) are statistically supported, which provides more rigor when interpreting the results of an
255 interaction model. At this point is worth mentioning that regions of non-significance can exist
256 and be relevant for the data of study despite the interaction term being non-significant (Rogosa,
257 1980, 1981; Bauer and Curran, 2005), thus it might be worth performing the Johnson-Neyman
258 technique regardless of the significance of the interaction term.
259
260

261 Table 1. Two-way interaction fitted models obtained with *JNplots* for three empirical examples. Significant p-values
 262 are shown in bold. The limits of significance obtained using the Johnson-Neyman technique (*min JN value* and *max*
 263 *JN value*) and the minimum and maximum values found in the data (*min value data* and *max value data*) are shown
 264 at the bottom of each table. In the first example the moderator is categorical and these limits refer to values of the
 265 predictor (e.g., for which predictor values are the differences between moderator categories non-significant?), while
 266 in the second and third examples the moderator is continuous and these limits refer to values of the moderator (e.g.,
 267 for which moderator values is the relationship between the dependent variable and the predictor (non)significant?).
 268

***Microlophus* lizards**

	Coefficient	t-value	p-value
Intercept	-1.79	-6.83	<0.001
log(SVL)	1.06	17.57	<0.001
species	1.59	5.47	<0.001
log(SVL) x species	-0.38	-5.50	<0.001

min JN value	max JN value	min value data	max value data
4.133	4.326	—	—

Lizard home range

	Coefficient	t-value	p-value
Intercept	4.26	23.56	<0.001
overlap	1.77	9.11	<0.001
network	-0.07	-0.39	0.700
overlap x network	0.86	6.13	<0.001

min JN value	max JN value	min value data	max value data
-3.296	-1.360	-2.169	2.481

Bird coloration

	Coefficient	t-value	p-value
Intercept	0.32	18.40	<0.001
precipitation	-5.03E-05	-5.13	<0.001
temperature	-4.40E-06	-0.007	0.9947
precip. x temp.	1.30E-06	2.80	0.0056

min JN value	max JN value	min value data	max value data
31.106	84.853	1.7	27.5

269
 270
 271
 272
 273
 274

275 *Example 2: drivers of home range size in a lizard*

276

277 Data from this example comes from the study of Payne et al (2022a). In the original study, the
278 authors were interested in uncovering the factors influencing the home range size (i.e., the area in
279 which an individual interacts with the environment) of individuals of the lizard species *Tiliqua*
280 *rugosa*. One of their main results indicated that the size of the home range of an individual
281 increases with its degree of overlap with the home range of their neighbours. Additionally, this
282 relationship is stronger for individuals that interact with more neighbours (i.e., degree of social
283 network).

284

285 I prepared a subset of the data (dataset 'lizard_home_range', included in *JNplots*, see Payne et al.,
286 2022b for original dataset) that included information on the home range size of each individual
287 ('hrsize95'), degree of overlap ('PHR95_overlap_z'), and social network degree ('degree_z'). To
288 analyze the two-way interaction between overlap and degree of social network (i.e., home range
289 size ~ overlap * social network) I used the function *jnt_cont*, as the moderator (i.e., degree of
290 social network) is continuous. As with *jnt_cat*, the necessary arguments for the function are the
291 names of the predictor (X), the dependent variable (Y), and the moderator (m) as they appear in
292 the dataset, which also needs to be specified:

293

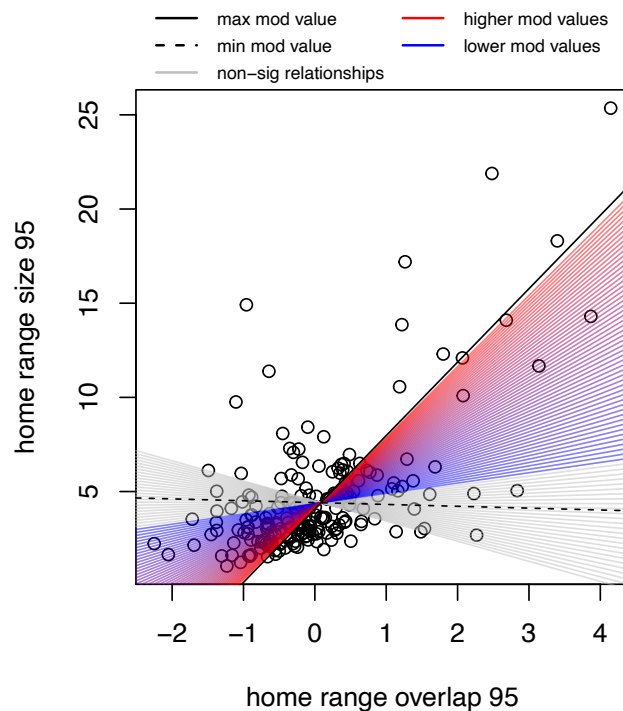
```
294 jnt_cont(X='PHR95_overlap_z', Y='hrsize95', m='degree_z',  
295         data=lizard_home_range)
```

296

297 As with *jnt_cat*, the output of the function consists in the summary table of the two-way
298 interaction model (home range size ~ overlap * social network), the values of the moderator that
299 represent the limits between significance and non-significance, and the minimum and maximum
300 moderator values in the data (Table 1). The function also produces a plot showing the association
301 between the dependent variable (e.g., home range size) and the predictor (e.g., overlap) (Figure
302 4). However, when the moderator is continuous the interpretation of the interaction effect differs
303 from the output of *jnt_cat*. In this example, and in agreement with the original study, home range
304 size increases with overlap. However, the degree of social interactions has a positive effect on
305 this relationship (i.e., the positive effect of overlap on home range size is stronger for lizards that
306 interact more with neighbours). This positive effect is represented by multiple regression lines
307 plotted in the output figure (Figure 4). The multiple grey regression lines that constitute the grey
308 'area' represent regressions fitted using moderator values that are outside the range of
309 significance, i.e., values of the moderator that make the relationship between the dependent
310 variable and the predictor not significant. (Figure 4). The 'area' in color consists of multiple
311 regression lines that represent models fitted using moderator values that fall within the
312 significance range (i.e., moderator values for which the relationship between the dependent
313 variable and the predictor is significant). The significant regression lines are colored in a blue-
314 red gradient that represent lower and higher moderator values, respectively, illustrating how
315 changes in the magnitude of the moderator (i.e., degree of social network) affect the relationship
316 between home range size and overlap (Figure 4). The plot also shows two additional lines. The
317 solid and dashed black lines represent fitted models that use the maximum and minimum values
318 of the moderator in the data, respectively. This aids in the interpretation of the plot because not
319 all moderator values might be relevant for the study system or the data at hand.

320

321 In this example, a higher degree of social interactions (moderator) increases the slope between
 322 home range size (dependent variable) and overlap (predictor) (Figure 4). However, a low degree
 323 of social interactions (specifically below a value of -1.360, Table 1) makes that relationship not
 324 significant, keeping home range size small regardless of the degree of overlap (grey area in
 325 Figure 4). Importantly, some moderator values that would result in non-significant relationships
 326 are found in the data, suggesting that this result is biologically relevant (see grey regression lines
 327 between solid and dashed black lines in Figure 4).
 328



329
 330 Figure 4. Graphical output of a model relating home range size to home range overlap and degree of social network
 331 in the lizard *Tiliqua rugosa* (model: home range size ~ overlap * social network) obtained with the function *jnt_cont*
 332 from *JNplots*. Colored lines represent significant linear models obtained using different 'degree of social network'
 333 values, the blue-red gradient represents different degrees of social network going from low to high, respectively.
 334 Grey lines represent non-significant linear models. Solid and dashed black lines represent the maximum and
 335 minimum precipitation values from the dataset, respectively. Data obtained from Payne et al (2022b).
 336

337 *Example 3: drivers of coloration in birds*

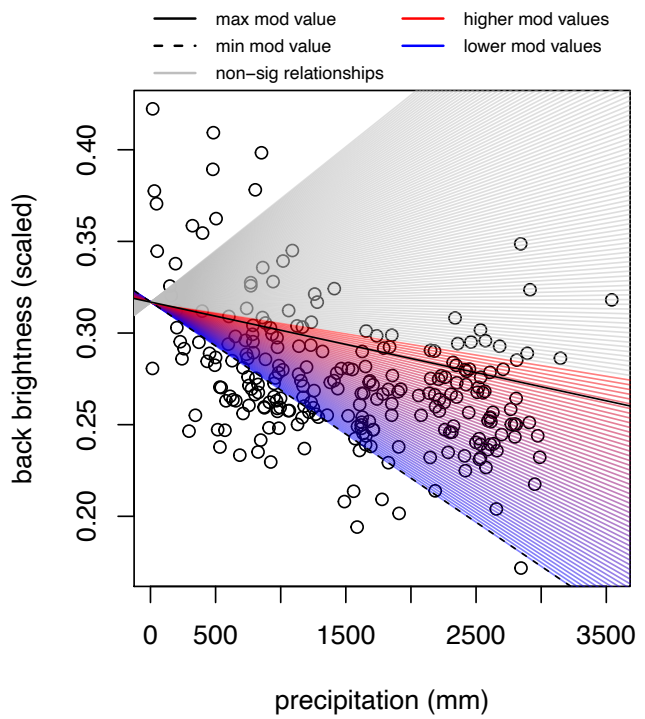
338
 339 Data from this example were originally described in a study by Marcondes and Brumfield (2019)
 340 and reanalysed in a follow-up study (Marcondes et al., 2021). In the latter study, the authors
 341 assessed how climatic variables and light environments influence the plumage coloration of bird
 342 species of the family Furnariidae. Among other findings, the authors found that the brightness (a
 343 proxy for overall melanin content, with less bright plumage patches having less melanin) of the
 344 back plumage was negatively related to precipitation. Furthermore, an interaction between
 345 temperature and precipitation was detected, indicating that the negative effect of precipitation on
 346 brightness is stronger when temperature is lower (see Figure 1A in Marcondes et al., 2021).
 347

348 I reanalysed a subset of the data used by Marcondes et al (2021) (dataset 'birds_colors', included
 349 in *JNplots*, original data by Marcondes and Brumfield, 2021 and Seeholzer et al., 2017) using the

350 *jnt_cont* function as in the previous example. The model of interest in this case was brightness ~
 351 precipitation * temperature. However, in contrast to the previous example, this analysis implies
 352 the non-independence of datapoints due to phylogenetic relationships. To account for this, I used
 353 the argument 'correlation' in *jnt_cont*. The argument 'correlation' specifies the correlation
 354 structure of the model (as one would do in the *gls* function of *nlme*). Phylogenetic correlation
 355 structures (e.g., 'corBrownian', 'corPagel', 'corBlomberg', etc) in turn need a phylogeny to be
 356 specified. Here I chose 'corPagel' as the correlation structure and used a phylogenetic tree of the
 357 Furnariidae ('tree_Furnariidae', also included in *JNplots*, Harvey et al., 2020), selecting '1' as the
 358 initial value of lambda (see Paradis and Schlieb, 2019, for details on using different correlation
 359 structures):

```
360
361 jnt_cont(X='bio12', Y='back_bright', m='bio1', data=bird_colors,
362         correlation=corPagel(1, tree_Furnariidae))
363
```

364 The output of *jnt_cont* showed that, in agreement with the original study, plumage brightness
 365 decreased with precipitation and the interaction between temperature and precipitation was
 366 significant (Table 1). Specifically, the effect of precipitation on brightness was stronger at lower
 367 temperatures. The limits of significance represented in the plot confirmed this pattern and also
 368 showed that the statement is generalizable for the entire range of temperature values experienced
 369 by species in the data, as it completely overlaps with the region of significance (Figure 5).
 370



371
 372 Figure 5. Graphical output of a model relating back plumage brightness to precipitation and temperature in
 373 Furnariidae bird species (model: brightness ~ precipitation * temperature) obtained with the function *jnt_cont* from
 374 *JNplots*. Colored lines represent significant linear models obtained using different temperature values, the blue-red
 375 gradient represents different temperatures going from low to high, respectively. Grey lines represent non-significant
 376 linear models. Solid and dashed black lines represent the maximum and minimum precipitation values from the
 377 dataset, respectively. Data obtained from Marcondes and Brumfield (2021).

378 CUSTOMIZATION OF GRAPHICAL OUTPUTS IN *JNplots*

379

380 One of the main aims of *JNplots* is to provide graphical outcomes that allow the user to interpret
381 interaction models in an intuitive way. To aid in this objective, the graphical outputs of its
382 functions allow for some aesthetic flexibility.

383

384 In the case of *jnt_cat* the regions of non-significance might not overlap the predictor values in
385 the data. This would result in the region of non-significance not appearing or only partially
386 appearing in the graphical output. Specifying the option 'plot.full = T' ('plot.full' defaults to F)
387 will result in the plot always showing the entire region of non-significance regardless of its
388 overlap with the predictor values of the data (compare Figure 6A and 6B). Other basic aspects of
389 the plot that can be modified are the symbols representing both categories (default: pch =
390 c(16,1)), colors (default: cols = c('black', 'black')), line types (default: lty = c(1,2)), line widths
391 (default: lwd = c(1,1)), and line colors (default: line.col = c('black', 'black')). As an example,
392 compare Figure 6A, which uses only default settings, and Figure 6C.

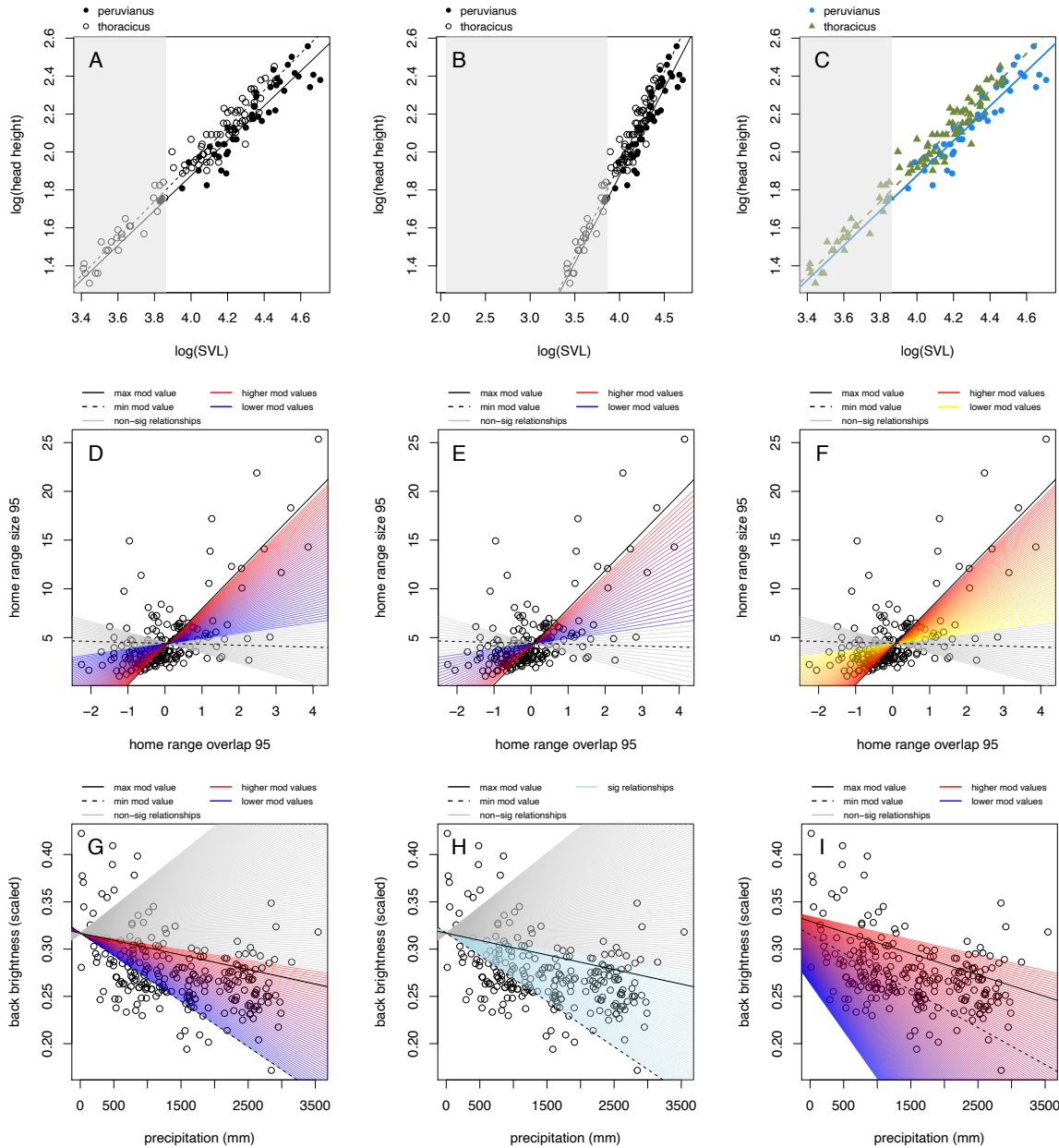
393

394 Plotting characteristics can also be specified in *jnt_cont*. The user can control the relative number
395 of regression lines to be plotted with the argument 'res', which defaults to 100. The exact number
396 of lines to be plotted is equal to the value of 'res' - 1, meaning that the number of plotted
397 regressions increases with the value specified in 'res' (Compare Figure 6D and 6E, which have
398 'res' values of 80 and 40, respectively). The gradient of colors shown by the significant
399 regression lines can also be modified. The arguments 'max_col_grad' and 'min_col_grad' define
400 the colors of the regression lines when using the maximum and minimum moderator values that
401 result in significant relationships, respectively. The colors of the regression lines in-between will
402 form a gradient between the extreme colors ('max_col_grad' and 'min_col_grad' default to 'red'
403 and 'blue', respectively). For example, compare Figure 6D and Figure 6F. If a color gradient
404 indicating different moderator values is not desired then 'col.gradient = F' (defaults to T), and all
405 the lines representing significant fitted regressions will take the color specified in the argument
406 'sig_color', which defaults to 'lightblue' (Compare Figure 6G and 6H). The color of the non-
407 significant regression lines can also be specified in the argument 'nonsig_color' (defaults to
408 'grey').

409

410 Finally, as previously mentioned, the correlation structure of the data can be modified in both
411 *jnt_cat* and *jnt_cont*. Although this is not an aesthetic specification, changing the correlation
412 structure will most likely change the aspect of the graphical outcome of either function. For
413 example, compare Figure 6G and 6I, which use Pagel's lambda and Brownian motion correlation
414 models, respectively.

415



416
 417 Figure 6. Graphical flexibility of *JNplots*. (A) *jnt_cat* was applied to a head height ~ size * species model. (B) Here
 418 the argument 'plot.full' was changed to 'T', which allows to see the entirety of the non-significance region regardless
 419 of the range of the predictor values. (C) Here plot.full = F, but other arguments were modified to change the
 420 aesthetics of the plot (pch = c(16,17), cols = c('dodgerblue2', 'darkolivegreen4'), lwd = c(2,2), line.col =
 421 c('dodgerblue2', 'darkolivegreen4')). (D) Same as figure 4, *jnt_cont* was applied to the model home range size ~
 422 overlap * network. (E) Here the argument 'res' was specified to be 40 (default = 100). Notice the lower number of
 423 regression lines and the larger space between them. (F). Here res = 80, and gradient colors are changed specifying
 424 min_col_grad = 'yellow' and max_col_grad = 'red'. (G) Same as figure 5, *jnt_cont* was applied to the model
 425 brightness ~ precipitation * temperature, res=150, correlation = corPagel(1, tree_Furnariidae). (H) Argument
 426 col.gradient = F and sig_color = 'lightblue'. The argument sig_color is only considered when col.gradient = F and
 427 defines a single color to be used for all significant regression lines. The argument nonsig_color works similarly for
 428 non-significant regression lines. (I) In this case the correlation structure is based on a Brownian motion model of
 429 evolution (correlation = corBrownian(1, tree_Furnariidae).

430 CONCLUSIONS

431

432 Multiple model testing is common in ecological and evolutionary studies, and understanding
433 how variables included in such models interact is indispensable for their interpretation (Hilborn
434 and Stearns, 1982; Dochtermann and Jenkins, 2011; Spake et al., 2023). Although the Johnson-
435 Neyman technique was initially developed in the context of educational and psychological
436 studies (Johnson and Neyman, 1936; Johnson and Fay, 1950), its application to other fields is
437 evident (e.g., White, 2003), as was the need to expand its application beyond categorical
438 moderators and two-way interactions (e.g., Bauer and Curran, 2005). In the same vein, *JNplots*
439 aims to be a tool that facilitates the application of the method in ecological and evolutionary
440 studies through the direct implementation of phylogenetic corrections and the possibility to
441 analyze categorical and continuous moderators, thus going beyond what is possible with existing
442 software. Equally important, *JNplots* aims to aid in the interpretation of two-way interactions
443 through more intuitive graphical outputs.

444

445 Although its main functions are readily available, *JNplots* still has room for expansion. For
446 example, the Johnson-Neyman technique can be applied to three-way or higher-level interactions
447 (Pothoff, 1964; Hunka, 1995; Hunka and Leighton, 1997; Curran et al., 2004; Bauer and Curran,
448 2005). Other types of regressions, like type II or reduced major axis regressions, and even non-
449 linear models also represent alternatives to traditional linear models not yet included as
450 analytical options in this package. These and other variations in the analysis of interactions
451 remain to be implemented in *JNplots* (or any other software). Before then, users interested in
452 such variations are free to copy and modify the functions from *JNplots*
453 (<https://github.com/kenstoyama/JNplots>) and adapt them to their needs.

454

455 Together with the release of this package, I provided a quick start guide online
456 (<https://kenstoyama.wordpress.com/2023/04/28/jnplots-quick-guide/>) for users that are more
457 familiar with the Johnson-Neyman technique and are specifically interested in the numerical and
458 graphical outputs of *JNplots*. The same page can be used to report issues with the use of the
459 package.

460

461 ACKNOWLEDGEMENTS

462

463 I want to thank Miriam Ahmad-Gawel, Rafael Marcondes, and Liam Revell for their helpful
464 comments and suggestions in earlier versions of this manuscript.

465 REFERENCES

466

467 Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression:
468 Inferential and graphical techniques. *Multivariate behavioral research*, 40(3), 373-400.

469

470 Bergmann C. (1847). Ueber die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse.
471 Gottinger Studien 3: 595–708.

472

473 Carden, S. W., Holtzman, N. S., & Strube, M. J. (2017). CAHOST: An excel workbook for
474 facilitating the Johnson-Neyman technique for two-way interactions in multiple
475 regression. *Frontiers in psychology*, 8, 1293.

476

477 Curran, P. J., Bauer, D. J., & Willoughby, M. T. (2004). Testing main effects and interactions in
478 latent curve analysis. *Psychological Methods*, 9(2), 220.

479

480 Dochtermann, N. A., & Jenkins, S. H. (2011). Developing multiple hypotheses in behavioral
481 ecology. *Behavioral Ecology and Sociobiology*, 65, 37-45.

482

483 Harvey, M. G., G. A. Bravo, S. Claramunt, A. M. Cuervo, G. E. Derryberry, J. Battilana, G. F.
484 Seeholzer, et al. 2020. The Evolution of a Tropical Biodiversity Hotspot. *Science*, 370, 1343–
485 1348.

486

487 Hayes, A. F., & Matthes, J. (2009). Computational procedures for probing interactions in OLS
488 and logistic regression: SPSS and SAS implementations. *Behavior research methods*, 41(3), 924-
489 936.

490

491 Hayes, A. F., & Montoya, A. K. (2017). A tutorial on testing, visualizing, and probing an
492 interaction involving a multicategorical variable in linear regression analysis. *Communication*
493 *Methods and Measures*, 11(1), 1-30.

494

495 Hilborn, R., & Stearns, S. C. (1982). On inference in ecology and evolutionary biology: the
496 problem of multiple causes. *Acta Biotheoretica*, 31(3), 145-164.

497

498 Huitema, B. (2011). *The analysis of covariance and alternatives: Statistical methods for*
499 *experiments, quasi-experiments, and single-case studies*. John Wiley & Sons.

500

501 Hünicken, L. A., Sylvester, F., & Paolucci, E. M. (2022). Physiological and morphological
502 assessments suggest opposite structural allocation strategies between closely related invasive
503 clams. *Hydrobiologia*, 849(12), 2859-2875.

504

505 Hunka, S. (1995). Identifying regions of significance in ANCOVA problems having non-
506 homogeneous regressions. *British journal of mathematical and statistical psychology*, 48(1),
507 161-188.

508

509 Hunka, S., & Leighton, J. (1997). Defining Johnson-Neyman regions of significance in the three-
510 covariate ANCOVA using Mathematica. *Journal of Educational and Behavioral*
511 *Statistics*, 22(4), 361-387.
512

513 Jaime, L., Batllori, E., Ferretti, M., & Lloret, F. (2022). Climatic and stand drivers of forest
514 resistance to recent bark beetle disturbance in European coniferous forests. *Global Change*
515 *Biology*, 28(8), 2830-2841.
516

517 Johnson, P. O., & Fay, L. C. (1950). The Johnson-Neyman technique, its theory and
518 application. *Psychometrika*, 15(4), 349-367.
519

520 Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their application to
521 some educational problems. *Statistical research memoirs*, 1, 57-93.
522

523 Lin, H. (2020). Probing two-way moderation effects: A review of software to easily plot
524 Johnson-Neyman figures. *Structural Equation Modeling: A Multidisciplinary Journal*, 27(3),
525 494-502.
526

527 Long, J. A. (2019). Interactions: Comprehensive, user-friendly toolkit for probing interactions. R
528 package version 1.1. 0. *The Comprehensive R Archive Network*.
529

530 Marcondes, R. S., & Brumfield, R. T. (2019). Fifty shades of brown: Macroevolution of plumage
531 brightness in the Furnariida, a large clade of drab Neotropical passerines. *Evolution*, 73(4), 704-
532 719.
533

534 Marcondes, R. S., & Brumfield, R. T. (2021). Data from: Fifty shades of brown: macroevolution
535 of plumage brightness in the Furnariida, a large clade of drab Neotropical passerines. *Evolution*
536 73:704–719, Dryad Digital Repository, <https://doi.org/10.5061/dryad.s86434s>.
537

538 Marcondes, R. S., Nations, J. A., Seeholzer, G. F., & Brumfield, R. T. (2021). Rethinking
539 Gloger's rule: climate, light environments, and color in a large family of tropical birds
540 (Furnariidae). *The American Naturalist*, 197(5), 592-606.
541

542 Montoya, A. K. (2019). Moderation analysis in two-instance repeated measures designs: Probing
543 methods and multiple moderator models. *Behavior research methods*, 51, 61-82.
544

545 Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and
546 evolutionary analyses in R. *Bioinformatics*, 35(3), 526-528.
547

548 Payne, E., Spiegel, O., Sinn, D. L., Leu, S. T., Gardner, M. G., Godfrey, S. S., Wohlfeil, C., &
549 Sih, A. (2022a). Intrinsic traits, social context, and local environment shape home range size and
550 fidelity of sleepy lizards. *Ecological Monographs*, 92(3), e1519.
551

552 Payne, E., Spiegel, O., Sinn, D. L., Leu, S. T., Gardner, M. G., Godfrey, S. S., Wohlfeil, C., &
553 Sih, A. (2022b). Intrinsic traits, social context, local environment, and home range size and

554 fidelity data from a field study of sleepy lizards, 2009 - 2017, Dryad Digital Repository,
555 <https://doi.org/10.25338/B8D922>
556

557 Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Heisterkamp, S., Van Willigen, B., & Maintainer,
558 R. (2017). Package ‘nlme’. *Linear and nonlinear mixed effects models, version, 3*(1), 274.
559

560 Potthoff, R. F. (1964). On the Johnson-Neyman technique and some extensions
561 thereof. *Psychometrika, 29*(3), 241-256.
562

563 Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions
564 in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of*
565 *educational and behavioral statistics, 31*(4), 437-448.
566

567 R Core Team (2021). R: A language and environment for statistical computing. R Foundation for
568 Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
569

570 Revell, L. J. (2010). Phylogenetic signal and linear regression on species data. *Methods in*
571 *Ecology and Evolution, 1*(4), 319-329.
572

573 Rogosa, D. (1980). Comparing nonparallel regression lines. *Psychological bulletin, 88*(2), 307.
574

575 Rogosa, D. (1981). On the relationship between the Johnson-Neyman region of significance and
576 statistical tests of parallel within-group regressions. *Educational and Psychological*
577 *Measurement, 41*(1), 73-84.
578

579 Seeholzer, G. F., Claramunt, S., & Brumfield, R. T. (2017). Niche evolution and diversification
580 in a Neotropical radiation of birds (Aves: Furnariidae). *Evolution, 71*(3), 702-715.
581

582 Sokal, R. R., Rohlf, F. J. 2012. Biometry, 4th edn. New York, NY, USA: Freeman and Co.
583

584 Spake, R., Bowler, D., Callaghan, C. T., Blowes, S. A., Doncaster, C. P., Antao, L. H., ... &
585 Chase, J. M. (2023). Understanding ‘it depends’ in ecology: a guide to hypothesising, visualising
586 and interpreting statistical interactions. *Biological Reviews*.
587

588 Symonds, M. R., & Blomberg, S. P. (2014). A primer on phylogenetic generalised least squares.
589 Modern phylogenetic comparative methods and their application in evolutionary biology:
590 concepts and practice, 105-130.
591

592 Toyama, K. S., Junes, K., Ruiz, J., Mendoza, A., & Pérez, J. M. (2018). Ontogenetic changes in
593 the diet and head morphology of an omnivorous tropidurid lizard (*Microlophus thoracicus*).
594 *Zoology, 129*, 45-53.
595

596 Wickham, H. & Seidel D. (2022). *scales*: Scale Functions for visualization. R package version
597 1.2.1. <https://CRAN.R-project.org/package=scales>