1   Special Paper

2 **A null biogeographic test for assessing ecological niche evolution**

3 Laura A. Nunes[1,2*] and Richard G. Pearson[1]

4   **1.** Centre for Biodiversity and Environment Research, Department of Genetics, Evolution and

5   Environment, University College London, London WC1E 6BT, UK

6   **2.** Vertebrate Division, Department of Life Sciences, Natural History Museum, London, SW7

7   5BD, UK

8 *Correspondence: Laura A. Nunes, Centre for Biodiversity and Environment Research, Department of Genetics,

9 Evolution and Environment, University College London, London WC1E 6BT, UK

10 Email: laura.nunes.13@ucl.ac.uk

11 **Abstract:**

12 **Aims**:

13 Quantification of the degree to which ecological niches change over evolutionary timescales

14 is important for deepening our understanding of evolutionary and ecological processes.

15 Phylogenetic niche conservatism (PNC) is when closely related species differ less ecologically

16 than expected by chance, whereas Phylogenetic Niche Divergence (PND) is when closely

17 related species differ more ecologically than expected by chance. We present a new null

18 model to test for PNC and PND (the RTR significance test), which we combine with a novel

19 metric for quantifying niche overlap.

20 **Location**:

21 Europe, North America and Madagascar

Wordcount:8154   Pages for Figures and Tables: 5

22   **Methods**:

23   The RTR null model comprises many thousands of replicates generated by randomly

24   translocating and rotating the set of occurrence records for two populations (e.g., sister

25   species) while maintaining the spatial configuration between all occurrences within each

26   replicate. For each replicate we calculate niche overlap as the proportion of the combined

27   niche breadth that is shared by the two species, averaged over n environmental dimensions.

28   This approach enables us to test whether the observed niche overlap is more or less than

29   expected by chance given the environmental conditions present in the study area. We test

30   the performance of our approach in comparison to other methods using both simulated and

31   real case scenarios, including crested newts in Europe, pocket gophers in North America,

32   and lemurs in Madagascar.

33   **Results**:

34   We find that our measure of niche overlap performs better than other metrics in an artificial

35   simulation scenario, and we find evidence for both PNC and PND in our case studies for

36   Europe, North America and Madagascar. Our results demonstrate that both the RTR

37   significance test and the novel metric of niche overlap are consistent with evolutionary

38   theory and are suitable methods to test for PNC and PND.

39   **Main Conclusions:**

40   We make available scripts to implement the RTR test and metric of niche overlap, and

41   expect that the methods will prove useful for addressing a broad set of questions relating to

42   ecological niche evolution and speciation, particularly for restricted-range species for which

43   few known occurrence records are available.

44   **Keywords: Ecological niche, niche conservatism, niche overlap, phylogeography, null**

45   **biogeographic model, speciation**

46

47

48

49

50

51

52

53

54

55

56

57

58 **Introduction**

59 Knowledge of how species respond to environmental conditions over evolutionary

60 timescales is important for understanding the causes of biodiversity proliferation, change

61 and persistence (Barraclough, 1998; Wiens & Donoghue, 2004). In particular, there is

62 interest in whether tolerance to environmental stressors tends to be conserved across a

63 phylogeny (Wiens *et al.*, 2004) or divergent from species to species (Losos *et al.*, 2008;

64 Ogburn & Edwards, 2015). Phylogenetic niche conservatism (PNC) is the tendency for closely

65 related species to differ less ecologically than expected by chance, and phylogenetic niche

66 divergence (PND) is the tendency for closely related species to differ more ecologically than

67 expected by chance (Pyron *et al.*, 2015). These concepts are of particular importance for

68 understanding the biogeography of speciation; for instance, allopatric speciation has been

69 inferred in cases of PNC (Peterson *et al.*, 1999; Wiens, 2004) whereas PND is expected in

70 cases of parapatric speciation (whereby ecological divergence along an environmental

71 gradient results in species with distinct ecological niches). Studies to-date have revealed

72 evidence both for PNC and PND but methodological difficulties have made the search for

73 general patterns difficult (Wiens & Graham, 2005; Peterson, 2011; Warren *et al.*, 2008;

74 Graham *et al.*, 2004).

75   A combination of newly available phylogenies, growing databases of species

76 occurrence records, new fine-resolution environmental variables derived from remote

77 sensing, and recently developed GIS-based statistical and machine-learning tools (e.g.,

78 ecological niche models, ENMs; also termed Species Distribution Models) provide

79 opportunities to substantially advance understanding of PND and PNC (McCormack *et al.*,

80 2010; Soberón, 2007). A general methodology for testing for PNC and PND is to: (i) collect

81   georeferenced occurrence records for populations with known phylogenetic relationships

82   (e.g., sister species); (ii) couple the occurrence records with a set of georeferenced

83   environmental variables (e.g., temperature, precipitation, vegetation cover); and (iii)

84   statistically compare the environments occupied by different populations. We note that this

85   approach tests for differences in the niches currently occupied by the populations (the

86   'occupied niche' , *sensu* Peterson *et al.*, 2011) rather than the niches within which the

87   populations could exist (either the 'full fundamental niche' or 'existing fundamental niche'

88   *sensu* Peterson *et al.*, 2011).

89   PNC and PND can be quantified in terms of niche overlap; that is, the proportion of an

90   environmental niche that is shared between two species (Colwell & Futuyma, 1971). These

91   measures of niche overlap range from 1 (niche equivalency) to 0 (niche divergence), with

92   varying degrees of niche similarity in between (Warren *et al.*, 2008). An important advance

93   proposed by Warren *et al.* (2008) has been the use of null model tests to assess observed

94   niche differences (or similarities) in the context of the environmental conditions available in

95   the study area. Generating a suitable null model allows us to ask whether the observed

96   similarity or difference between the niches of two populations is statistically meaningful

97   given the available environments. Warren *et al.* (2008) proposed two null tests, and several

98   other authors have proposed methods for testing niche similarity (see Appendix S3 in

99   Supporting Information for review of methods and their limitations).

100   Here we present a new method that takes an alternative approach to testing for PNC

101   and PND and addresses some of the problems with other approaches. Specifically, we have

102   devised a method to test whether two populations are currently distributed in such a way

103   that niche overlap is higher (PNC) or lower (PND) than would be expected by chance. We

104   introduce a new null model (the RTR significance test) and a novel metric for quantifying

105   niche overlap. The new methods are suitable for the study of range-restricted species with

106   few known occurrence records, and were designed to aid our understanding of the

107   landscape-scale ecological processes involved in speciation. We highlight that the current

108   RTR approach is designed to test for PNC among closely related species and is not well

109   suited to applications outside of phylogeography; for example, studies of niche

110   differentiation among invasive species require tests that cover two geographic areas (native

111   and invaded ranges) but the RTR test is limited to a single study area. We make available R

112   scripts to implement the methods (see Appendix S1).

113

114   **Materials and Methods**

115   We propose a methodology that involves four main steps: 1) Collection of georeferenced

116   occurrence data for two populations (e.g., sister species) and environmental variables for

117   the region of interest; 2) Measurement of observed niche overlap; 3) Production of a null

118   reference frequency distribution of niche overlap values; and 4) Comparison of the observed

119   niche overlap value to the null reference distribution to make the decision of rejecting or

120   accepting the null hypothesis. The null hypothesis is that the observed niche overlap is no

121   different to that expected at random, as defined by our null model (see below). Rejection of

122   the null hypothesis would suggest that environmental conditions have played an active role

123   in defining distributions (i.e., populations are adapted to particular ecological niches).

124   Failure to reject the null hypothesis would suggest that spatial rather than ecological

125   processes have been dominant in defining present day distributions.

126   The two methods that we present below – the new metric and the RTR null model –

127   can be used together (as we do here) but may also be used separately with existing metrics

128   of niche overlap and significance tests (e.g., the RTR approach could form the basis for

129   running ENMs and calculating the distribution of niche overlap using Schoener's *D* or

130   Hellinger's *I* under the null hypothesis). All the analyses were conducted in R (R Core Team,

131   2014)

A metric for measuring multidimensional niche overlap

133   We quantify the overlap (*x*) along a given environmental axis (*e*) between two species (*i and i'*) as:

$$x_e(i, i') = \frac{\min\left(y_{e,i}^{max}, y_{e,i'}^{max}\right) - \max\left(y_{e,i}^{min}, y_{e,i'}^{min}\right)}{\max\left(y_{e,i}^{max}, y_{e,i'}^{max}\right) - \min\left(y_{e,i}^{min}, y_{e,i'}^{min}\right)} \qquad x_e(i, i') \in [0,1] ; \quad i \neq i'$$

134   where $y_{e,}^{max}$ and $y_{e,}^{min}$ are the maximum and minimum values of *e* for species *i* or *i'*

135   respectively and the overlap ($x_e(i, i')$) is a real-value ranging between 0 and 1. In

136   instances where the environmental breadths of the species do not overlap, our equation

137   will return a negative value which we replace by a value of 0 to indicate no overlap

138   between the axes.

139   We next average all the axes overlap values across all of the dimensions used to

140   define the niches of the two species using a metric we term *MO* (for Multidimensional

141   Overlap):

142    $$MO_{i,i'} = \frac{1}{N} \sum_{e=1}^{e=N} x_e(i, i') \qquad\qquad MO_{i,i'} \in [0,1] ; \quad i \neq i'$$

143   such that the overall overlap ($MO_{i,i'}$) is a real-value between 0 and 1 (see Appendix

144   S4 for further details on the *MO* metric) .

145   The *MO* metric is a presence-only approach that is has similarities with the BIOCLIM

146   method (Busby, 1991) in that it constructs simple climate envelopes around the occurrence

147   records (Booth *et al.,* 2014). One limitation of such methods is that they are sensitive to

148   occurrence records that are ecological outliers, which can lead to overestimation of the

149   ecological niche (Farber & Kadmon, 2003). Overestimation due to outliers can be

150   addressed by selecting a 'core niche region', such as the 5-95% percentile of the niche

151   (Carpenter *et al.*, 2003). We therefore implemented a method to undertake an optional

152   'trimming' function to remove ecological outliers (see Appendix S1) and we have tested

153   sensitivity of results to removal of outliers (see Appendix S4). A second limitation is that

154   very limited sampling of the species' distribution (e.g., fewer than five occurrence records)

155   is likely to lead to underestimation of the species' niche. However, all methods are limited

156   by the availability of empirical data and our approach has the advantage of avoiding

157   making unfounded extrapolations beyond the range of the available data. A third limitation

158   of BIOCLIM-like methods is that all environmental variables are treated as equally

159   important (there is no weighting of variable importance, unlike in methods such as MAXENT;

160   Phillips *et al.* 2006). It is therefore important to apply a priori ecological knowledge (e.g.,

161   Blair *et al.* 2013) and/or statistical assessment of variable importance (e.g., by jackknifing,

162   Wielstra *et al.*, 2012; Soto-Centeno *et al.*, 2013) to select relevant environmental variables

163   for the taxa under consideration. In our three case studies (see below) we used the same

164   variables that were used in the original studies, each of which applied a priori statistical

165   analysis or expert knowledge to identify important variables. A fourth limitation is that our

166   approach examines only elements of niche evolution that affect the minimum and

167   maximum values along each niche dimension. Unlike methods that fit response curves in

168   SDMs, our method does not attempt to identify more subtle differences in niches that are

169   reflected in the shape of the functional response to the environment. The benefit of this is

170   that we avoid the many assumptions that go into fitting response curves to limited data

171   and uncertainty over what form the curves should take (Elith *et al.,* 2009).

172   The RTR test is suitable for populations (e.g., of sister species) that have restricted

173   ranges and few occurrence records, and has several advantages over alternative tests,

174   including: (i) there is no need for reciprocal tests because an individual 'background' region

175   does not need to be defined for each species; (ii) it maintains the spatial autocorrelation of

176   the point data; (iii) the test is based on overlap in N-dimensional niche space; (iv) there is

177   no limit to the number of ecological dimensions or occurrence records that can be included;

178   and (v) the test returns a p-value to assess statistical significance rather than two p-values

179   from a reciprocal test due to the need for individual background regions.

180

181   **RTR null biogeographic model**

182   We have developed a novel test that uses what we term the Random Translocation and

183   Rotation (RTR) null model. The RTR null model comprises many thousands of replicates that

184   are generated by randomly translocating and rotating the pooled set of occurrence records

185   for two populations while maintaining the spatial configuration between all occurrences

186   (i.e., of both compared species together) within each replicate. Niche overlap is calculated

187   for each replicate and the observed niche overlap is then compared to the distribution of

188   overlap values from the null model (Fig. 1). If the observed niche overlap falls outside a

189   critical boundary, we reject the null hypothesis and infer that the niches are conserved (e.g.,

190   above 95% percentile of null distribution, PNC) or divergent (e.g., below 5% percentile of

191   null distribution, PND). The significance threshold of this null model approach is not

192   restricted to upper and lower 5% boundaries, thus the investigator is able to select the

193   critical threshold of the model, as well as choose between a two-tailed or one-tailed test for

194   PNC or PND (see Appendix S1).

195   The RTR null model maintains the spatial configuration, and thus spatial

196   autocorrelation, between all occurrences within each replicate (note somewhat comparable

197   methods by Beale *et al.,* 2008 and Chapman, 2010). The method enables us to test whether

198   the observed distributions are likely to be driven by environmental factors: we ask whether

199   the niches occupied by two species are more or less similar than would be expected at

200   random if the spatial configuration of the set of occurrences is maintained within a given

201   landscape (background region). In effect, we keep the spatial configuration constant so we

202   can ask whether there is something ecologically 'special' about the way that the two species

203   are currently located on the landscape. Notice that here we are not assessing the present

204   day spatial configuration of the two species (e.g., are the ranges adjacent?) but rather we

205   are assessing whether the ecological niches currently occupied are more similar or different

206   than expected by chance within the landscape.

207   As with other null models (e.g., Warren *et al.*, 2008), results from the RTR approach

208   are impacted by the extent of the landscape over which the replicates are run. We address

209   sensitivity to selection of the study region below (see *Testing sensitivity to extent of study*

210   *region*). However, an important difference to current methods is that we do not define

211   separate background regions for each species; instead, we randomly translocate and rotate

212   within a single region the set of occurrence records for the two species combined, thus

213   maintaining the spatial configuration between species. We therefore make no assumptions

214   about geographic constraints that might separate the distributions of the two populations.

215   In some instances it may be appropriate to identify likely geographic constraints (e.g., the

216   Isthmus of Tehuantepec in southern Mexico; Peterson *et al.,* 1999, Warren *et al.,* 2008);

217   however, in most cases geographic constraints are less clear and we do not want to impose

218   a priori any geographic separation when we are testing whether a single population might

219   have split into separate populations (e.g., sister species) due to ecological divergence.

220   **Comparing niche overlap metrics using simulated species**

221   To assess the performance of different measures of niche overlap, we compared our *MO*

222   metric against three alternatives: (i) Warren *et al.*'s (2008) *I* statistic, which is based on a

223   comparison of ENMs and has been widely used; (ii) Broennimann *et al.*'s (2012) PCA-env

224   framework with the *D* statistic, a more recent methodology that has been applied to

225   invasive species; and (iii) Blonder *et al.*'s (2014) *n*-dimensional technique, in which niches

226   are built as multidimensional hypervolumes.

227   We compared the metrics using simulated (artificial) species. Simulated species

228   were preferable here to real case studies from nature because the degree of PND and PNC

229   could be precisely defined, providing a 'known truth' against which the different metrics

230   could be compared. We simulated the environmental niche overlap of two simulated

231   species in a two-dimensional environmental domain of 100 x 100 grid cells (following

232   Broennimann *et al.,* 2012, and Colwell *et al.*, 2009). We generated two opposing gradients

233   across the artificial landscape to represent two uncorrelated environmental variables. We

234   represented each species' distribution as a square of 30x30 grid cells (see Appendix S5 for

235   further information on methods and expected outcomes). One species' distribution was

236   kept static in the bottom left corner of the environmental domain, while the other was

237   initially placed in the bottom left corner but then moved one grid cell at a time in either

238   direction (up or right) or diagonally away from the other species. The movement of one

239   simulated species away from the other across the environmental gradient meant that

240   there was increasing ecological divergence with increasing geographic separation (Fig. S5 in

241   Appendix S5).

242   To calculate the *I* statistic, we needed to construct ENMs. We used the R package

243   'DISMO' 1.0-5 (Hijmans *et al.*, 2011) to build ENMs and the package 'SDMTOOLS' 1.1-221

244   (VanDerWal, *et al.*, 2012) to calculate *I*. We opted for ENMs constructed using MAXENT

245   3.3.3K (Phillips *et al.*, 2006) as this method performs well in comparison with other ENM

246   approaches (Elith *et al.*, 2006) and was used in Warren *et al.*'s original paper (2008). We

247   followed Warren *et al.* (2008) in maintaining default values for all program settings,

248   including regularization and feature selection. To calculate *D* we used the R script for

249   uncalibrated PCA-env functions provided by Broennimann *et al.* (2012). To build and

250   measure the intersection of two hypervolumes following Blonder *et al.* (2014), we used

251   their package 'HYPERVOLUME' 1.4.1 with 1,000 random numbers, a bandwidth of 0.1, a

252   quantile of 0, and a reduction factor of 0.5, as recommended by Blonder *et al.* (2014).

253   **Testing the null model using three case studies**

254   We also tested our methodology (RTR null model combined with the new *MO* metric of

255   niche overlap) using three real-world case studies. We selected case studies based on the

256   following criteria: (i) a previous study has been published that includes estimates or

257   hypotheses regarding PNC or PND; and (ii) occurrence records are available for the included

258   sister species. Following each of the original studies, we used georeferenced environmental

259   layers from Worldclim at 30 arc-seconds resolution (Hijmans *et al.*, 2005), in each case

260   selecting the same set of variables as the original study (see Appendix S6). We selected two

261   case studies with a continental setting (the Balkan Peninsula and North America) and one

262   island (Madagascar) since islands have more obvious natural boundaries and therefore raise

263   different issues concerning the selection of a suitable study region:

264   i. Crested newts in the Balkans. The phylogeography of the crested newts has been

265   extensively studied (Arntzen *et al.,* 2007, Wielstra *et al.*, 2010, Wielstra *et al.*, 2012), which

266   allows us to make some predictions about the potential for niche conservatism. There is

267   evidence for vicariance among two sister pairs: (i) *Triturus karelinii* and *T. ivanbureschi*,

268   which is thought to have split due to the uplift of the Armenian Plateau; and (ii) *T. carnifex*

269   and *T. macedonicus*, which is thought to have split due to the formation of the Adriatic Sea

270   (Wielstra *et al.*, 2010). Wielstra *et al.* (2012) used ordination methods to assess niche

271   evolution and found evidence for PNC. This finding is in line with the theoretical expectation

272   that adaptation to new niches tends to be slower than extinction rates, leading to PNC in

273   cases of allopatric speciation (Peterson *et al.* 1999). We aimed to test this using the RTR

274   method.

275   Occurrence records for the crested newt species were obtained from Wielstra *et al.*

276   (2012). We set the western and eastern boundaries of Europe (Fig. 2a) based on the overall

277   distribution of the *Triturus* clade in Europe (Wielstra *et al.*, 2012). The environmental

278   layers used (see Appendix S6) were clipped to the same extent using the crop function

279   from the 'RASTER' 2.5-2 package (Hijmans, 2015), and all layers were converted to a Lambert

280   Conformal Conic projection (at central meridian 18°E, standard parallels 42°N and 46°N),

281   which represents low spatial distortion for the Balkan region (Zagmajster *et al.*, 2008).

282   ii. Pocket gophers in North America. Though studies of ecological divergence are

283   commonly conducted at the species level, there is also interest in looking at a population

284   level in order to understand biogeographic patterns of within-species divergence (Graham

285   *et al.*, 2004; Glor & Warren, 2011). *Geomys pinetis* is the only species of pocket gopher

286   found in south-eastern USA and within this species there is a geographic and genetic

287   subdivision between eastern (*G. pinetis*) and western populations (hereafter *G. mobilensis*)

288   on either side of the Apalachicola River (Soto-Centeno *et al.*, 2013). Soto-Centeno *et al.*

289   (2013) used the 'identity' test and the 'blob' range-breaking test of *ENMTools* (Glor &

290   Warren, 2011) and could not reject a hypothesis of niche similarity between the two

291   populations across the river barrier. Here we aimed to test whether a signal for PNC or

292   PND could be detected using our RTR method.

293   Occurrence records for *G. pinetis* and *G. mobilensis* were obtained from MaNIS

294   (manisnet.org), following Soto-Centeno *et al.* (2013). We ran the RTR test for the south-

295   eastern region of the country (Fig. 2b) which is equivalent to the extent used in Soto-

296   Centeno *et al.* (2013). The Lambert conformal conic projection (central meridian 96°E,

297   standard parallels 20°N and 60°N) was used to transform the environmental layers (see

298   Appendix S6) because this has low spatial distortion for North America (Les *et al.*, 2013).

299   iii. Lemurs in Madagascar. Blair *et al.* (2013) found that two sister pairs of *Eulemur*

300   lemurs in Madagascar (*E. collaris-E. cinereisceps* and *E. rufus - E. rufifrons*) have clear

301   riverine barriers; for one of these pairs (*E. rufus - E. rufifrons*) they found evidence of PNC*,*

302   and for the other pair (*E. collaris - E. cinereisceps*) they found no significant signal, based on

303   the null background tests of Warren *et al.* (2008). By contrast, two other sister pairs (*E.*

304   *albifrons - E. sanfordi* and *E. flavifrons - E. macaco*) were found to have less well defined

305   geographic barriers; for one of these pairs they found support for significant PND for one

306   pair (*E. flavifrons -E. macaco*) and for the other pair (*E. albifrons - E. sanfordi*) they found

307   no significant signal (Blair *et al.*, 2013).

308   We obtained the *Eulemur* occurrence records from Blair *et al.* (2013) and restricted

309   the analysis to the island of Madagascar (Fig. 2c). The environmental layers used (see

310 Appendix S6) were transformed to an oblique Mercator projection for Madagascar

311 (following Pearson *et al.* 2007).

312 **Testing sensitivity to extent of study region**

313 The RTR test requires selection of an overall study region within which the RTR replicates

314 are generated. The selection of this overall background region will affect the model output

315 and the interpretation of the findings (e.g, the observed niche overlap being unique in

316 relation to a particular region but not to another). Selecting an extent that is too wide might

317 include environments that are too different from what the species or clade are likely to

318 experience thus biasing the null library to unrealistic measurements of niche overlap.

319 Selecting a small extent will lead to a reduction in the available environmental

320 heterogeneity being sampled which could result in the exclusion of such unrealistic

321 environments, thus reducing irrelevancy in null library. However, a too narrow extent could

322 mean that there is more similarity between the environments occupied by the observed

323 distribution and the environments sampled within the study region due to higher likelihood

324 of partial overlap within and between the simulated distributions and the observed

325 distributions. Smaller study regions are therefore expected to return lower type I error rates

326 (i.e., false rejection of the null hypothesis) than tests performed across larger areas. In

327 general, a study region should be selected that bounds the landscape, and hence the set of

328 environmental conditions, that the species could reasonably be expected to have had the

329 opportunity to occur in. Thus, factors such as the dispersal capacity of the species,

330 topographic features in the landscape (e.g., barriers to dispersal), and the distribution of

331 major clades to which the species belong might be considered when selecting the study

332 region.

333   We tested sensitivity of the RTR significance test to the extent of the study region by

334   undertaking *Eulemur* analysis for the whole island of Madagascar and also for alternative

335   smaller extents. We did this for the two most range-restricted sister pairs: *E. collaris - E.*

336   *cinereiceps* and *E. falvifrons - E. macaco*. We divided Madagascar latitudinally,

337   approximately cutting the island in half, to generate sub-island extents that take into

338   account the island's general north-south zonation (Blair *et al.*, 2013). We also generated

339   further restricted study regions by approximating areas of lemur endemism proposed by

340   Pastorini *et al.* (2003). Thus, *E. collaris - E. cinereiceps* was tested with a southern extent and

341   a more restricted south-eastern extent (corresponding to region E2 in Pastorini *et al.* (2003);

342   Fig. 2c) and *E. falvifrons - E. macaco* was tested with a northern extent and a smaller north-

343   western extent (corresponding to region X in Pastorini *et al.*, 2003; Fig. 2c).

344

345   **Results**

346   **Performances of different niche overlap metrics in a common simulated scenario**

347   We found contrasting performance between niche overlap metrics when tested using

348   simulated species, with some metrics having a tendency to overestimate niche overlap

349   while others tend to underestimate niche overlap in relation to the 'known truth' scenario

350   (Fig. 3b).

351   The novel metric presented in this study, *MO*, captures the expected results more

352   closely than the other three metrics (Fig. 3), with the expected ranges for each region being

353   matched precisely with the outcomes observed in the new metric (Fig. 3c).

354   Warren's *I* metric over-predicted niche overlap in instances where no overlap should

355   be observed in region D (Fig. 3d). This metric also did not capture the expected gradient in

356   niche overlap in region C (Fig. 3d). However, it captured the expected range of range of

357   niche overlap in region B (Fig. 3d). Moreover, niche identity (complete niche overlap) was

358   correctly predicted to occur only in Region A (Fig. 3d).

359   Broennimann's *D* statistic also tended to over-predict niche overlap in region D (Fig.

360   3e); however, this was to a lesser extent than Warren's *I* statistic, as it was able to identify

361   some cases of complete niche divergence in that region. The expected gradient from 0 to

362   0.5 was observed in region C (Fig. 3e). The metric was the only one to overestimate niche

363   overlap (ranging from 0.2-1) in region B (Fig. 3e). Broennimann *et al.*'s method also

364   overpredicted the occurrence of identical niches, which was predicted beyond region A (Fig.

365   3e). Nevertheless, overall it was able to capture the expected range of 0 to 1 (Fig. 3e).

366   Blonder *et al.*'s (2014) hypervolume approach also captured the 0 to 1 range, though

367   niche overlap was under-predicted in regions C and D, where only niche divergence was

368   observed (Fig. 3f). However, the approach accurately captured the expected range from 0 to

369   1 in region B, and complete niche overlap was correctly detected only in Region A.

370   **Performance of the RTR null biogeographic model in real case scenarios**

371   The RTR test supports a finding of PNC for both sister pairs of crested newts (Table 1). This is

372   in agreement with previous findings (Wielstra *et al.*, 2012).

373   For the populations of pocket gophers, we found no significant signal for either PNC

374   or PND, which is consistent with the results of Soto-Centeno *et al.* (2013). However, the

375   observed niche overlap was close to significant for PNC (observed *MO* = 0.59 and 95%

376   threshold for significance = 0.61), indicating that further studies may be appropriate to test

377   this finding (Table 1).

378   Results for lemurs show contrasting findings between the RTR null model and the

379   null model of Warren *et al.* (2008), with agreement in the findings for only one out of four

380   sister pairs. The RTR test finds support for PNC for the pair *E. collaris - E. cinereiceps* (Table

381   1). This is in contrast to the findings of Blair *et al.* (2013), who found no significant signal

382   using the tests of Warren *et al.* (2008). For the pair *E. rufus - E.rufifrons*, we find evidence for

383   PND. This is again in contrast to Blair *et al.* (2013), who found some support for PNC. For the

384   other two *Eulemur* pairs, we find no significant signal using the RTR test. This is in

385   agreement with Blair *et al.* (2013) for the *E. albifrons -E. sanfordi* pair, but in conflict for *E.*

386   *flavifrons -E. macaco* (Blair *et al.* (2013) found evidence for PND).

387   Changing the extent of the study region for *Eulemur* pairs showed that the RTR test is

388   sensitive to selection of the area over which the test is run (Table 2). For one sister pair (*E.*

389   *collaris -E. cinereiceps*) we observed loss of significant signal as the extent of the study

390   region was reduced; however this loss of signal was only observed when the background

391   area was very close in extent to that of the two species. Thus, use of a small study extent

392   produced a result consistent with the findings from Blair *et al.* (2013), who also used a small

393   extent ('background' regions were selected by constructing minimum convex polygons

394   around occurrence records). For the other sister pair, no significant signal was found

395   regardless of the extent of background area used (Table 2).

396   **Discussion**

397   **Performances of different niche overlap metrics in a common simulated scenario**

398   Our analysis showed contrasting outcomes across niche overlap metrics for the same

399   simulated scenario. These findings highlight the importance of choosing an appropriate

400   metric for answering specific ecological and/or evolutionary questions. We found that the

401   statistic introduced in this paper, the *MO* metric, performed better than other metrics by

402   producing results closer to those expected from the 'known truth' scenario.

403   We found poorer performance of Warren's *I* statistic and Broenniman's PCA-env

404   procedure with the *D* statistic for quantifying niche overlap when compared to other

405   metrics. The two statistics tend to be correlated (Warren *et al.*, 2008) and we have

406   demonstrated a common tendency to overestimate the simulated niche overlap when no

407   niche overlap is expected to be found and underestimated when one environmental

408   variable overlaps but not the other. This tendency to overestimate the niche overlap has

409   been reported previously in simulated environments (Broennimann *et al.*, 2012) and is likely

410   because the approaches measured the intersection of predictions from two ENMs, which

411   are designed to estimate suitability and therefore tend to overestimate species'

412   distributions (because some suitable habitats will be unoccupied; Peterson *et al.,* 2011).

413   We found a tendency for the *n*-dimensional hypervolume approach to underestimate

414   niche overlap in our simulation. This is particularly noticeable in instances where the niches

415   do not overlap on at least one axis. There are multiple aspects that could account for

416   underestimation of niche overlap using the hypervolume approach (Blonder *et al.,* 2014). In

417   our simulations, the most likely explanation for underestimation is that the hypervolume

418   method measures niche overlap as the intersection of two volumes such that the volumes

419   will not intersect at all (niche overlap is measured as 0) if they have one or more non-

420   overlapping variables. None of the other metrics we assessed make this strict assessment of

421   niche overlap, either because they calculate niche overlap by doing a cell-by-cell comparison

422   (e.g., Warren's *I* and Broenniman's *D*) or they calculate an averaged proportion of the niche

423   breadth shared across all axis (our *MO* metric). However, we note that in cases when the

424   volumes did overlap on both axes in our simulation, the performance of the hypervolume

425   approach was comparable to that of the *MO* metric (which closely followed the 'known

426   truth').

427   In comparison with the other metrics tested, we found that our new *MO* metric better

428   captured niche overlap in the simple simulated scenario that we tested. There is, however,

429   scope to refine this metric in light of limitations (see Methods). One limitation is the

430   potential sensitivity to ecological outliers (see Appendix S4). We have implemented a

431   function to remove ecological outliers (see Appendix 1) but removal of outliers is only

432   advised when it is expected that some points may be erroneous (e.g., misidentifications) or

433   there are likely sink populations that do not represent the niche. Removing true ecological

434   extremes will lead to a misrepresentation of the species occupied niche. Although there are

435   limitations, we have demonstrated in our comparisons that the conceptually simple *MO*

436   metric is a useful approach for quantifying niche overlap.

437   **Performance of the RTR null biogeographic model in real case studies**

438   By applying the RTR test to real case studies, we have shown that the new test can

439   provide comparable outcomes to expected and previously observed patterns from the

440   literature. We have also shown differences in the outputs of the RTR method and Warren *et*

441   *al.*'s (2008) background test, with agreement for only one out of four pairs, but congruency

442   between the RTR method and ordination techniques and the 'blob' range-breaking test for

443   the remaining pairs. Application of the *MO* metric and the RTR significance test supports a

444   hypothesis of PNC for two allopatric sister pairs of crested newts in the Balkans (*T. carnifex -*

445   *T. macedonicus* and *T. karelinii -T. ivanbureschi*). This finding is in line with previous work

446   that has used ordination techniques to identify PNC in this study system (Wielstra *et al.*,

447   2012) and offers a way of further testing the theoretical expectation that PNC is a key

448   pattern emerging from allopatric speciation (Cooper *et al.*, 2010; Losos, 2011; Crisp & Cook,

449   2011). For pocket gophers in south-eastern North America, we found no significant signal

450   for either PNC or PND, which implies that observed ecological differences between the

451   populations are not likely due to selection for a particular set of conditions. Rather, our

452   results suggest that niche differences are a coincidental result of different environments

453   available on each side of a geographic barrier (the Apalachicola river).

454   The endemic lemurs of Madagascar have been assessed previously in tests of which, if

455   any, mode of speciation may have been most important in driving local endemism and

456   speciation across the island (Pearson & Raxworthy, 2009; Blair *et al.*, 2013). Our

457   measurements of niche overlap using the *MO* metric were congruent with Warren *et al.*'s *I*

458   and *D* statistics, with our results consistently falling within the range of values presented by

459   Blair *et al.*, (2013). However, the RTR test identified significant statistical support for PNC in

460   the pair *E. collaris- E. cinereiceps*, which is not congruent with the results using Warren *et*

461   *al.*'s background test (Blair *et al.*, 2013). Our finding, combined with strong genetic support

462   for the Mananara river acting as a barrier to gene flow (Wyner *et al.*, 2002), suggests an

463   allopatric mode of speciation for this pair, with niches failing to evolve on either side of a

464   geographic divide.

465   Our analyses find no significant support for either PNC or PND in two *Eulemur* sister

466   pairs (*E. flavifrons-E. macaco* and *E. albifrons-E. sanfordi*). These results are in contrast to

467   those of Blair *et al.* (2013), who found support for PND for one of the pairs (*E. flavifrons -E.*

468   *macaco*). Closer analysis of the differences between the null libraries produced by Blair *et*

469   *al.* (2013) using Warren's background test and our RTR significance test shows that the RTR

470   test captured a broader null distribution for *E. flavifrons -E. macaco*. This explains the

471   difference in statistical inference between the two null models for these sister pairs, despite

472   comparable observed niche overlap values between the two studies. However, it is

473   important to highlight that both *E. flavifrons-E. macaco* and *E. albifrons –E. sanfordi* have no

474   complete geographical barriers and have potential hybrid zones (see Blair *et al.*, 2013 and

475   references therein). This suggests that other factors besides geographic or environmental

476   separation may be responsible for speciation within these pairs (e.g., microhabitat selection,

477   Rakotondranary & Ganzhorn, 2011) which would account for the lack of signal found in our

478   analysis.

479   We found support for PND for *E. rufus-E. rufifrons*, which opposes the hypothesis of

480   PNC of Blair *et al.* (2013), who found weak support for PNC. However, there is limited spatial

481   overlap between the potential distributions of the pairs (Blair *et al.,* 2013), which suggests

482   that the species have different ecological preferences and is consistent with our finding of

483   PND.

484   One explanation for differences in results between the RTR test and Warren *et al.*'s

485   (2008) null models is that the RTR approach focuses on the ecological dimension of niche

486   evolution between species while Warren *et al.* (2008) focuses on the geographical

487   dimension (i.e., our RTR method measures overlap in niche space, whereas Warren *et al.*

488   (2008) measures the spatial overlap of ENMs). This difference likely accounts for different

489   biological inferences from the alternative methods. We contend here that our approach of

490   measuring overlap in niche space is most appropriate for robustly testing for PNC and PND,

491   but further research is needed to decipher the different facets of niche evolution that are

492   picked up by each of the approaches. Consequently, we expect that the new methods

493   presented here will complement, rather than replace, existing methods.

494   We highlight that the RTR approach is best suited to study range-restricted species

495   where many unique RTR replicates can be generated within a background region. Instances

496   where a species is wide-ranging relative to the study region, or where the 'shapes' of the

497   species' distribution and study region are such that RTR replicates can be located in only a

498   limited number of ways, will result in few replicates and potential spatial biases. We

499   illustrate that spatial biases are case specific in Appendix S7 by showing the locations of

500   10,000 RTR replicates for a pair of newts and a pair of lemurs. Spatial bias in the RTR null

501   model acts to curtail the background region, and therefore the range of environments

502   considered (note that this does not negate the meaning of the significance test for PNC or

503   PND, but it does mean that the range of conditions over which the calculation is performed

504   is only a subset of those in the selected background region). As with selection of the

505   background region, spatial bias within the RTR null model will be an important area for

506   future research. To facilitate this we provide R code in Appendix S1 for replicating the

507   analyses we have done in Appendix S7.

508   As with other tests for PNC and PND, the RTR test found instances with no significant

509   signal. In addition to the actual absence of either PNC or PND, there are a number of

510   methodological factors that can explain non-significant results, including: (i) the choice of

511   predictor variables (some key variables may be excluded from the analysis); (ii) the coarse

512   resolution of analysis (niche differentiation within the 1km^2 cells used will not be picked up);

513   (iii) incomplete knowledge of occurrences (modern-day loss and fragmentation of habitats

514   means we have an incomplete picture of the niches of the species), and (iv) inappropriate

515   extent of the study region (explored in more detail below).

516   **Testing sensitivity to extent of study region**

517

518   We have shown that the RTR test is sensitive to the extent of the study region. Our tests

519   for Madagascar suggest that sensitivity to the selection of study region extent is low, unless

520   the extent becomes very small (e.g., of similar size to the extent of the two species'

521   distributions). Further exploration of this sensitivity, and of different strategies for

522   selecting the study region (e.g., based on dispersal capacity, following Anderson & Raza,

523   2010), is warranted.

524   Discordance between our results and those of Blair *et al.* (2013) may be due in part

525   to the different sensitivities of the RTR test and Warren *et al.*'s (2008) test to the extent of

526   the study region. Blair *et al.* (2013) defined the background area of each species based on a

527   minimum convex polygon bound by the occurrence records of each species. A serious

528   drawback of using minimum convex polygons around occurrence records for each species

529   is the assumption that all locations within the polygon are suitable habitat for the species.

530   This assumption is unlikely to hold in many cases (e.g., consider a species that occupies

531   warm lowlands around the base of a mountain: drawing a minimum convex polygon

532   around the species' occurrence records will encompass both the lowlands and also the

533   cooler mountain top).

534   A result that is non-significant is not expected to become significant if the extent is

535   made smaller. This was observed in our analysis of *E. flavifrons - E. macaco*. We thus find

536   that a tight study region, such as those defined by Blair *et al.* (2013), is more prone to

537   result in falsely accepting the null hypothesis. A wider extent, by contrast, is more likely to

538   result in rejection of the null hypothesis.

539

540   **Conclusions**

541   Overall, we find that the new metric of niche overlap, the *MO* metric, and the new RTR

542   significance test are suitable methods for testing for PNC and PND, particularly when

543   applied to range-restricted species with few occurrence records. Given its novel approach to

544   the study of niche dynamics between populations, the RTR method holds great promise for

545   testing for PNC and PND across large phylogenies with many sister pairs, and thus shedding

546   new light on evolutionary processes, in particular speciation.

547

552   **References**

553

554   Anderson, R. P. & Raza, A. (2010) The effect of the extent of the study region on GIS

555   models of species geographic distributions and estimates of niche evolution: Preliminary

556   tests with montane rodents (genus *Nephelomys*) in Venezuela. *Journal of Biogeography*,

557   **37**,1378–1393.

558   Anderson, R.P. (2013) A framework for using niche models to estimate impacts of climate

559   change on species distributions. *Annals of the New York Academy of Sciences* ,**1297**, 8-28.

560   Araújo, M.B., Thuiller, W. & Yoccoz, N.G. (2009) Reopening the climate envelope reveals

561   macroscale associations with climate in European birds. *Proceedings of the National*

562   *Academy of Sciences USA*, **106**, E45–E46.

563   Arntzen J.W., Themudo G. Espregueira & Wielstra B. (2007) The phylogeny of crested newts

564   (*Triturus cristatus* superspecies): nuclear and mitochondrial genetic characters suggest a

565   hard polytomy, in line with the paleogeography of the centre of origin. *Contributions to*

566   *Zoology*, **76**, 261-278.

567   Barraclough, T. G. (1998) Revealing the factors that promote speciation. *Philosophical*

568   *Transactions of the Royal Society B: Biological Sciences*, **353**,241–249.

569   Beale CM, Lennon JJ, Gimona A. (2008) Opening the climate envelope reveals no

570   macroscale associations with climate in European birds. *Proceedings of the National*

571   *Academy of Sciences USA :* **105**,14908–12.

572   Blair, M. E., Sterling, E. J., Dusch, M., Raxworthy, C. J., & Pearson, R. G. (2013) Ecological

573   divergence and speciation between lemur (*Eulemur*) sister species in Madagascar.

574   *Journal of Evolutionary Biology*, **26**, 1790–801.

575   Blonder, B., Lamanna, C., Violle, C., & Enquist, B. J. (2014) The n-dimensional

576   hypervolume. *Global Ecology and Biogeography*, **23**, 595–609.

577   Booth, T. H., Nix, H. A., Busby, J. R. & Hutchinson, M. F. (2014) BIOCLIM: the first species

578   distribution modelling package, its early applications and relevance to most current MAXENT

579   studies. *Diversity and Distributions*, **20**, 1–9.

580   Broennimann, O., Fitzpatrick, M. C., Pearman, P. B., Petitpierre, B., Pellissier, L., Yoccoz, N.

581   G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N. E., Graham, C. H., & Guisan,

582   A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental

583   data. *Global Ecology and Biogeography*, **21**,481–497.

584   Busby, J.R. (1991) BIOCLIM – a bioclimate analysis and prediction system*. Plant Protection*

585   *Quarterly,* **6***,* 8–9*.*

586   Carpenter, G., Gillison, A.N. & Winter, J. (1993) DOMAIN: a flexible modelling procedure for

587   mapping potential distributions of plants and animals. *Biodiversity and Conservation*, **2**,

588   667–680.

589   Chapman, D. S. (2010), Weak climatic associations among British plant distributions. *Global*

590   *Ecology and Biogeography*, **19**, 831–841.

591   Colwell, R. K. & Futuyma, D. J. (1971) On the measurement of niche b readth and

592   o verlap. *Ecology*, **52**,567–576.

593   Colwell, R. K., Gotelli, N. J., Rahbek, C., Entsminger, G. L., Farrell, C., & Graves, G. R. (2009)

594   Peaks, plateaus , canyons , and craters : the complex geometry of simple mid-domain effect

595   models. *Evolutionary Ecology Research*, **11**,355–370.

596   Cooper, N., Jetz, W. & Freckleton, R. P. (2010), Phylogenetic comparative approaches for

597   studying niche conservatism. *Journal of Evolutionary Biology*, **23**,2529–2539.

598 Crisp, M. D. & Cook, L. G. (2012), Phylogenetic niche conservatism: what are the underlying

599 evolutionary and ecological causes? *New Phytologist*, **196**, 681–694.

600 Elith, J. *et al.* (2006) Novel methods improve prediction of species' distributions from

601 occurrence data. *Ecography*, **29**, 129–151

602 Elith, J. & Graham, C. H. (2009) Do they? How do they? Why do they differ? On finding

603 reasons for differing performances of species distribution models. *Ecography*, **32**, 66–77.

604 Elith, J., Kearney, M. & Phillips, S. (2010), The art of modelling range-shifting species.

605 *Methods in Ecology and Evolution*, **1**, 330–342.

606 Estrada-Peña, A., Estrada-Sánchez, A., & de la Fuente, J. (2014). A global set of Fourier-

607 transformed remotely sensed covariates for the description of abiotic niche in

608 epidemiological studies of tick vector species. *Parasites & Vectors*, **7**,302.

609 Farber, O. & Kadmon, R. (2003) Assessment of alternative approaches for bioclimatic

610 modelling with special emphasis on the Mahalanobis distance. *Ecological*

611 *Modelling*, **160**, 115–130.

612 Glor, R. E. & Warren, D. (2011) Testing ecological explanations for biogeographic

613 boundaries. *Evolution*, **65**,673–683.

614 Graham, C. H., Ron, S. R., Santos, J. C., Schneider, C. J., & Moritz, C. (2004) Integrating

615 phylogenetics and environmental niche models to explore speciation mechanisms in

616 dendrobatid frogs. *Evolution*, **58**,1781–1793.

617 Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005) Very high

618 resolution interpolated climate surfaces for global land areas. *International Journal of*

619 *Climatology*, **25**,1965–1978.

620 Hijmans, R.J., Phillips, S., Leathwick, J. & Elith, J. (2011) Package 'DISMO'. Available online

621 at: http://cran.r-project.org/web/packages/dismo/index.html.

622 Hijmans, R.J., (2015). RASTER: Geographic data analysis and modeling. R package version 2.3

623 24. http://CRAN.R-project.org/package=raster

624 Hutchinson, G. E. (1957) Concluding remarks. *Cold Spring Harbour Symposia on*

625 *Quantitative Biology*, **22**,415-427.

626 Les, D.H., Peredo, E.L.., Benoit, L.K., Tippery, N.P., King, U.M. & Sheldon, S.P. (2013)

627 Phytogeography of *Najas gracillima* (Hydrocharitaceae) in North America and its cryptic

628 introduction to California. *American Journal of Botany*, **100**, 1905–1915.

629 Losos, J.B. (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship

630 between phylogenetic relatedness and ecological similarity among species. *Ecology Letters*,

631 **11**, 995–1003.

632 Losos, J.B. (2011). Seeing the forest for the trees: the limitations of phylogenies in

633 comparative biology.*American Naturalist,* **177**, 709–727.

634 McCormack, J. E., Zellmer, A. J., & Knowles, L. L. (2010) Does niche divergence

635 accompany allopatric divergence in *Aphelocoma* jays as predicted under ecological

636 speciation? Insights from tests with niche models. *Evolution*, **64**, 1231–1244.

637 Ogburn, R.M. & Edwards, E.J. (2015) Life history lability underlies rapid climate niche

638 evolution in the angiosperm clade *Montiaceae. Molecular Phylogenetics and Evolution*, **92**,

639 181–192.

640   Owen, J.G. (1989). Patterns of herpetofaunal species richness: Relation to temperature,

641   precipitation, and variance in Elevation. *Journal of Biogeography*, **16**, 141–150.

642   Pastorini, J., Thalmann, U., & Martin, R. D. (2003) A molecular approach to comparative

643   phylogeography of extant Malagasy lemurs. *Proceedings of the National Academy of*

644   *Sciences USA*, **100**, 5879–5884.

645   Pearson, R. G. & Raxworthy, C. J. (2009) The evolution of local endemism in Madagascar:

646   watershed versus climatic gradient hypotheses evaluated by null biogeographic models.

647   *Evolution*, **63**,959–967.

648   Pearson, R. G., Raxworthy, C. J., Nakamura, M. & Townsend Peterson, A. (2007) Predicting

649   species distributions from small numbers of occurrence records: a test case using cryptic

650   geckos in Madagascar. *Journal of Biogeography*, **34**,102–117.

651   Pearson, R.G., Stanton. J.C., Shoemaker, K.T., Aiello-Lammens, M.E., Ersts, P.J., Horning, N.,

652   Fordham, D.A., Raxworthy, C.J., Ryu, H.Y., McNees, J., & Akçakaya, H.R.((2014). Life history

653   and spatial traits predict extinction risk due to climate change. *Nature Climate Change*,

654   **4**,217-221.

655   Peterson, A. T. (2011) Ecological niche conservatism: a time-structured review of

656   evidence. *Journal of Biogeography*, **38**, 817–827.

657   Peterson, A. T., Soberón, J. & Sánchez-Cordero, V. (1999) Conservatism of ecological

658   niches in evolutionary time. *Science,* **285**, 265-267.

659   Peterson, A.T., Barve, N., Bini, L.M., Diniz-Filho, J.A., Jiménez-Valverde, A., Lira-Noriegaa,

660   A., Lobo, J., Maher, S., de Marco, P.,Martínez-Meyer, E., Nakazawa, Y. & Soberón,

661   J. (2009) The climate envelope may not be empty. *Proceedings of the National Academy of*

662   *Sciences USA*, **106**, E47.

663   Peterson, A.T., Soberón, J., Anderson, R.P., Pearson, R.G., Martínez-Meyer,E., Nakamura,M.

664   & Araújo, M.B. (2011) *Ecological Niches and Geographic Distributions: A Modeling*

665   *Perspective*, Princeton University Press, Princeton, NJ.

666   Phillips, S. J., Anderson, R. P. & Schapire, R. E. (2006) Maximum entropy modelling of

667   species geographic distributions. *Ecological Modelling*, **190**,231–259.

668   Pyron, R.A., Costa, G.C., Patten, M.A. & Burbrink, F.T. (2015) Phylogenetic niche

669   conservatism and the evolutionary basis of ecological speciation. *Biological Reviews*, **90**,

670   1248-1262.

671   R Core Team (2014) R: A language and environment for statistical computing. R Foundation

672   for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

673   Rakotondranary, S. J. & Ganzhorn, J. U. (2011) Habitat separation of sympatric *Microcebus*

674   spp. in the dry spiny forest of south-eastern Madagascar. *Folia Primatologica*, **82**, 212–223.

675   Rödder, D. & Engler, J. O. (2011) Quantitative metrics of overlaps in Grinnellian niches:

676   advances and possible drawbacks. *Global Ecology and Biogeography*, **20**,915–927.

677   Soberón, J. (2007). Grinnellian and Eltonian niches and geographic distributions of species.

678   *Ecology Letters*, **10**,1115–1123.

679   Soto-Centeno, J. A., Barrow, L. N., Allen, J. M. & Reed, D. L. (2013) Re-evaluation of a

680   classic phylogeographic barrier: new techniques reveal the influence of microgeographic

681   climate variation on population divergence. *Ecology and Evolution*, **3**,1603–1613.

682   Vanderwal J., Falconi L.,Januchowski S., Shoo L. & Storlie C. (2012) Package 'SDMTOOLS'.

683   Species Distribution Modelling Tools: Tools for processing data associated with species

684   distribution modelling exercises. R-package version 1.1.12, URL: http://cran.r-project.org/.

685   Warren, D. L., Glor, R. E. & Turelli, M. (2008) Environmental niche equivalency versus

686   conservatism: quantitative approaches to niche evolution. *Evolution*, **62**,2868–2883.

687   Wielstra, B., Beukema, W., Arntzen, J. W., Skidmore, A. K., Toxopeus, A. G. & Raes, N.

688   (2012) Corresponding mitochondrial DNA and niche divergence for crested newt

689   candidate species. *PLoS ONE*, **7**,e46671.

690   Wielstra B., Espregueira Themudo G., Guclu O., Olgun K., Poyarkov N.A. & Arntzen J.W.

691   (2010) Cryptic crested newt diversity at the Eurasian transition: The mitochondrial DNA

692   phylogeography of Near Eastern *Triturus* newts. *Molecular Phylogenetics and Evolution*, **56**,

693   888-896.

694   Wiens, J. J. (2004) Speciation and ecology revisited: Phylogenetic niche conservatism

695   and the origin of species. *Evolution*, **58**,193–197.

696   Wiens, J.J. & Graham, C.H. (2005) Niche conservatism: integrating evolution, ecology, and

697   conservation biology. *Annual Review of Ecology, Evolution and Systematics*, **36**, 519–539.

698   Wiens, J. J., Ackerly, D. D., Allen, A. P., Anacker, B. L., Buckley, L. B., Cornell, H. V.,

699   Damschen, E. I., Jonathan Davies, T., Grytnes, J.A., Harrison, S. P., Hawkins, B. A., Holt, R.

700   D., McCain, C. M. & Stephens, P. R. (2010) Niche conservatism as an emerging principle

701   in ecology and conservation biology. *Ecology Letters*, **13**, 1310–1324.

702 Wiens, J. J. & Donoghue, M. J. (2004) Historical biogeography, ecology and species richness.

703 *Trends in Ecology and Evolution*, **19**, 639–644.

704 Wollenberg, K.C., Vieites, D.R., Glaw, F. & Vences, M. (2011) Speciation in little: the role of

705 range and body size in the diversification of Malagasy mantellid frogs. *BMC Evolutionary*

706 *Biology.* **11**,217.

707 Wyner, Y., Johnson, S.E., Stumpf, R. & DeSalle, R. (2002) Genetic assessment of a white

708 collared X red-fronted lemur hybrid zone at Andringitra, Madagascar. *American Journal of*

709 *Primatology*, **57**, 51–66.

710 Zagmajster, M., Culver, D.C. & Sket, B. (2008) Species richness patterns of obligate

711 subterranean beetles in a global biodiversity hotspot – effect of scale and sampling

712 intensity. *Diversity and Distributions*, **14**, 95–10.

713 **Supporting Information**
714 Additional Supporting Information may be found in the online version of this article:

715 **Appendix S1** R codes.
716 **Appendix S2** Methods for R code of the RTR test.

717 **Appendix S3** Review of existing methods.
718 **Appendix S4** *MO* metric.
719 **Appendix S5** Virtual scenarios for testing niche overlap metrics.
720 **Appendix S6** Environmental variables for case studies.
721 **Appendix S7** Assessing potential spatial bias.

722

723 **Biosketch**

724 **Laura Nunes** is currently doing a Ph.D. on ecological niche evolution and the biogeography

725 of speciation, particularly among chameleons in Madagascar. Richard Pearson is interested

726 in the ecological and evolutionary processes that determine species distributions.

727   Author contributions: R.G.P. and L.A.N. conceived the ideas and designed the analyses.

728   L.A.N. wrote the code and performed the analyses. L.A.N. wrote the paper with input from

729   R.G.P.

730   Editor: Daniel Chapman

731   **Figures and Tables**

732

733   **Table 1.** Observed niche overlap measured with the *MO* metric and outputs from the RTR

734   null biogeographic model as a significance test for the observed niche overlap value against

735   a null library for two sister pairs of crested newts in the Balkans (*T. karelinii – T. ivanbureschi*

736   and *T. carnifex-T. macedonicus*), a sister pair of pocket gophers (*G.pinetis –G.mobilensis*) in

737   North America and four sister pairs of lemurs from in Madagascar (*E. collaris – E.cinereiceps*;

738   *E.flavifrons-E.macaco*; *E.albifrons –E. sanfordi* and *E.rufus-E.rufifrons*) . PNC refers to

739   Phylogenetic Niche Conservatism when the observed niche overlap is higher than expected

740   by chance, and PND refers to Phylogenetic Niche Divergence for cases where the observed

741   niche overlap is lower than expected by chance. No significant signal refers to cases where

742   the observed niche overlap does not occur less often than expected by random chance.

743   Results from the RTR test are compared against expectations based on published papers.

Sister pair	Original study	Observed Niche Overlap (*MO* statistic)	5%, mean and 95% tails for Niche Overlap based on a null RTR distribution	Hypothesis from original study	RTR test result
T. karelinii - T. ivanbureschi	Wielstra *et al.*,2012	**0.61**	0.26;0.37;0.51	**PNC**	**PNC**
T. carnifex - T. macedonicus	Wielstra *et al.*,2012	**0.69**	0.23;0.42;0.65	**PNC**	**PNC**

G. pinetis - *G. mobilensis*	Soto-Centeno *et al.,* 2013	**0.59**	0.31; 0.47 ;0.61	**No significant** **signal**	**No significant** **signal**
E. collaris - *E. cinereiceps*	Blair *et al.,* 2013	**0.56**	0.05;0.24;0.51	**No significant** **signal**	**PNC**
E. flavifrons - *E. macaco*	Blair *et al.,* 2013	**0.34**	0.09;0.31;0.57	**PND**	**No significant** **signal**
E. albifrons - *E. sanfordi*	Blair *et al.,* 2013	**0.24**	0.09;0.25;0.45	**No significant** **signal**	**No significant** **signal**
E.rufus – *E. rufifrons*	Blair *et al.,* 2013	**0.19**	0.27;.0.48;0.67	**PNC/No** **significant** **signal**	**PND**

744

745 **Table 2.** Observed niche overlap measured with the *MO* metric and outputs from the RTR

746 null biogeographic model as a significance test for the observed niche overlap value against

747 a null library for two *Eulemur* sister pairs in Madagascar using alternative study region

748 extents. Regions E2 and X refer to proposed areas of endemism for lemurs according to

749 Pastorini *et al.* (2003) which coincide with the range-restricted sister pairs (see main text).

750 The results from the RTR null model are compared against hypotheses based on published

751 papers. Warren *et al.*'s (2008) background similarity test is based on results from Blair *et al.*

752 (2013). PNC refers to Phylogenetic Niche Conservatism when the observed niche overlap is

753 higher than expected by chance, and PND refers to Phylogenetic Niche Divergence for cases

754 where the observed niche overlap is lower than expected by chance. No significant signal

755 refers to cases where the observed niche overlap does not occur less often than expected

756   by random chance.

| Sister pair	Extent of study region	Niche Overlap (*MO* metric)	5%, mean and 95% tails for Niche Overlap (*MO*) based on a null RTR distribution	P-value (obtained from RTR test)	Background similarity test	RTR test result
E. collaris - E. cinereiceps	**Region E2**	**0.56**	0.10 ; 0.40 ;0.70	**0.26**	**No significant signal**	**No significant signal**
	Southern Region		0.06 ; 0.23 ;0.48	**0.02**		**PNC**
	Whole island		0.05 ; 0.24 ;0.52	**0.03**		**PNC**
E. flavifrons - E. macaco	**Region X**	**0.36**	0.20;0.37;0.56	**0.41**	**PND**	**No significant signal**
	Northern region		0.10;0.32;0.60	**0.42**		**No significant signal**
	Whole island		0.10;0.31; 0.57	**0.40**		**No significant signal**

757

758

759

760

761

763   **Figure 1.** Random translation and rotation (RTR) null model test for phylogenetic niche

764   divergence (PND) and phylogenetic niche conservatism (PNC), illustrated for a pair of

765   sister lemur species in Madagascar (see Blair *et al.*, 2013). (a) Occurrence records for the

766   sister species, visualized on an example environmental layer (multiple environmental

767   layers (e.g., temperature, precipitation) are used in the test to characterize the n-

768   dimensional ecological niche space). (b) Thousands of null replicates are generated by

769   randomly rotating and translating the set of occurrence records for the two species. (c)

770   For the observed distributions and for all null replicates, the niche overlap is calculated

771   and the observed overlap is compared against the null model. In this instance the

772   observed niche overlap is in the highest 5% of the null distribution, so we infer PNC.
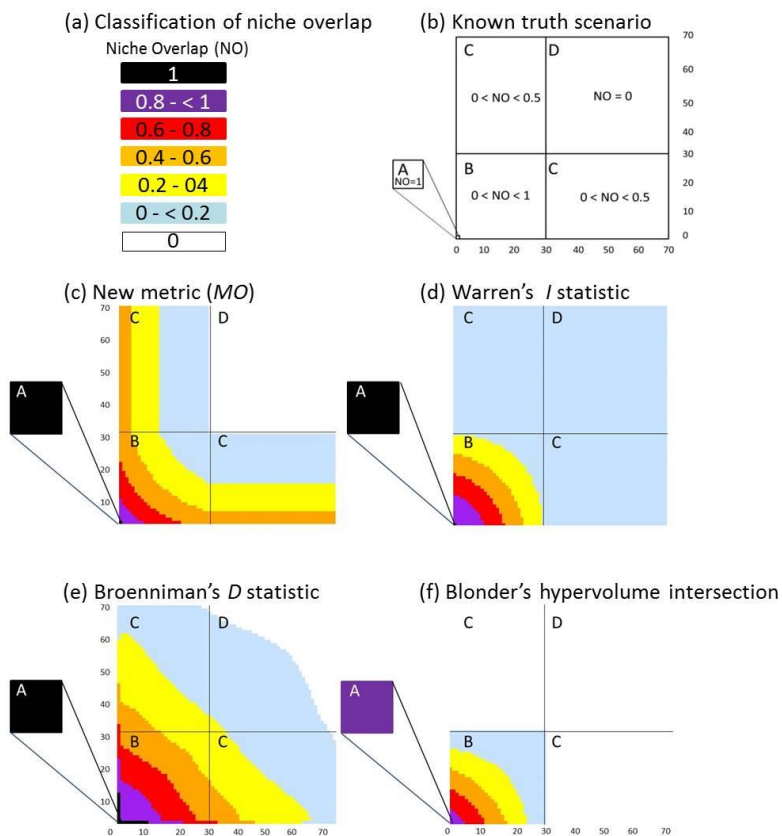

773

774
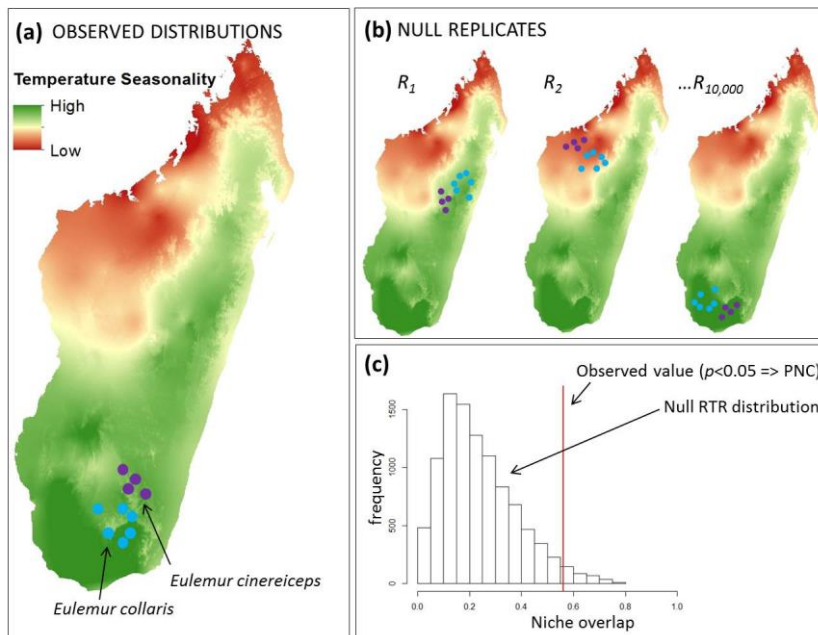
Figure 2. Species' occurrence records and extents of the study regions used in the three case studies. (a) Crested newts in the (*T. karelinii – T. ivanbureschi* and *T. carnifex-T. macedonicus)*; (b) Populations of pocket gophers in south-eastern USA (*G.pinetis – G.mobilensis*); (c) Lemurs in Madagascar (*E. collaris – E.cinereiceps; E.flavifrons-E.macaco; E.albifrons –E. sanfordi* and *E.rufus-E.rufifrons*). Regions E2 and X refer to proposed areas of endemism for lemurs according to Pastorini *et al.* (2003) which coincide with the range-restricted sister pairs (see main text). The outlines (dashed lines in (c)) represent the alternative extents used to test sensitivity of RTR null model to differences in the extent of the study region.
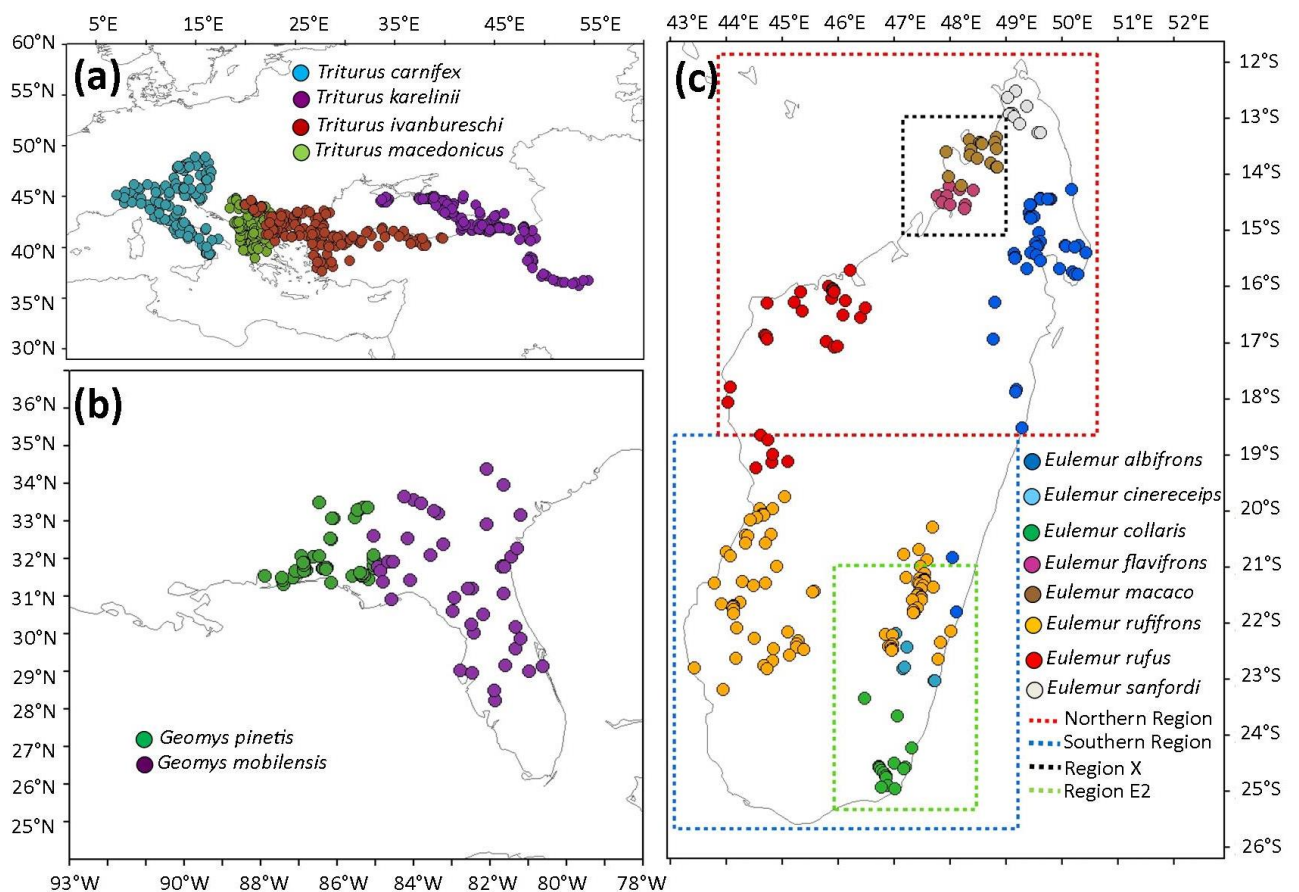
786   **Figure 3.** Niche overlap values generated by four alternative metrics for a simulated

787   scenario. (a) We used the classification of niche overlap established by Rödder & Engler

788   (2011) but also distinguished between completely divergent niches (niche overlap [NO]=0)

789   and identical niches (NO=1). (b) The 'known truth' represents the expected ranges of niche

790   overlap in each section of a 2-dimensional 70x70 grid (see Appendix S4 for more details).

791   (c) The new niche overlap statistic (*MO*) introduced in this paper. (d) Warren *et al.*'s (2008)

792   background similarity test with the *I* statistic. (e) Broenniman *et al.*'s (2012) PCA

793   uncalibrated technique with the *D* statistic. (f) Blonder *et al.*'s (2014) intersection of

794   hypervolumes between two species.

795

796