Catchments catch all in South African coastal lowlands: topography and palaeoclimate restricted gene flow in *Nymania capensis* (Meliaceae) – a multilocus phylogeographic and distribution modelling approach

Alastair J Potts [Corresp. 1, 2]

1 Botany Department, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa

2 Department of Biological Sciences, University of Cape Town, Cape Town, South Africa

Corresponding Author: Alastair J Potts
Email address: potts.a@gmail.com

Background. This study investigates orbitally-forced range dynamics at a regional scale by exploring the evolutionary history of *Nymania capensis* (Meliaceae) across the deeply incised landscapes of the subescarpment coastal lowlands of South Africa; a region home to three biodiversity hotspots (Succulent Karoo, Fynbos, and Maputaland-Pondoland-Albany hotspots).

Methods. A range of methods are used including: multilocus phylogeography (chloroplast and high- and low-copy nuclear DNA), molecular dating and species distribution modelling (SDM).

Results. The results support an 'evolutionarily distinct catchment' hypothesis where: (1 different catchments contain genetically distinct lineages, 2) limited genetic structuring was detected within basins whilst high structuring was detected between basins, and 3) within primary catchment populations display a high degree of genealogical lineage sorting. In addition, the results support a glacial refugia hypothesis as: a) the timing of chloroplast lineage diversification is restricted to the Pleistocene in a landscape that has been relatively unchanged since the late Pliocene, and b) the projected LGM distribution of suitable climate for N. capensis suggest fragmentation into refugia that correspond to the current phylogeographic populations.

Discussion. This study highlights the interaction of topography and subtle Pleistocene climate variations as drivers limiting both seed and pollen flow along these lowlands. This lends support to the region's large-scale conservation planning efforts, which used catchments as foundational units for conservation as these are likely to be evolutionary significant units.

1   **TITLE:**

2   **Catchments catch all in South African coastal lowlands: topography and palaeoclimate**

3   **restricted gene flow in *Nymania capensis* (Meliaceae) – a multilocus phylogeographic**

4   **and distribution modelling approach**

6   **SHORT HEADER: Catchments catch all in a lowland plant species**

8   **Alastair J. Potts**

9   Department Biological Sciences, University of Cape Town, Private Bag X3, Rondebosch, Cape

10   Town, South Africa, 7701.

11   Botany Department, Nelson Mandela Metropolitan University, Summerstrand Campus (South),

12   Port Elizabeth, South Africa, 6000.

14   **Abstract**

15   **Background.** This study investigates orbitally-forced range dynamics at a regional scale by

16   exploring the evolutionary history of *Nymania capensis* (Meliaceae) across the deeply incised

17   landscapes of the subescarpment coastal lowlands of South Africa; a region home to three

18   biodiversity hotspots (Succulent Karoo, Fynbos, and Maputaland-Pondoland-Albany hotspots).

19   **Methods.** A range of methods are used including: multilocus phylogeography (chloroplast and

20   high- and low-copy nuclear DNA), molecular dating and species distribution modelling (SDM).

21 **Results.** The results support an 'evolutionarily distinct catchment' hypothesis where: 1)

22 different catchments contain genetically distinct lineages, 2) limited genetic structuring was

23 detected within basins whilst high structuring was detected between basins, and 3) within

24 primary catchment populations display a high degree of genealogical lineage sorting. In

25 addition, the results support a glacial refugia hypothesis as: a) the timing of chloroplast lineage

26 diversification is restricted to the Pleistocene in a landscape that has been relatively unchanged

27 since the late Pliocene, and b) the projected LGM distribution of suitable climate for *N. capensis*

28 suggest fragmentation into refugia that correspond to the current phylogeographic

29 populations.

30 **Discussion.** This study highlights the interaction of topography and subtle Pleistocene climate

31 variations as drivers limiting both seed and pollen flow along these lowlands. This lends support

32 to the region's large-scale conservation planning efforts, which used catchments as

33 foundational units for conservation as these are likely to be evolutionary significant units.

35

36 **Introduction**

37 Climate changes driven by variations in the Earth's orbit, i.e. Milankovitch oscillations, have

38 influenced the distribution of species and clades across the globe. Jansson & Dynesius (2002)

39 termed this "orbitally-forced range dynamics" (ORD) and highlight that this process drives

40 extinction, splitting, and merging of gene pools and clades. The magnitude of ORD varies

41 geographically, primarily latitudinally with greater ORD in the poles and a decline towards the

42 equator (Dynesius & Jansson 2000). Gene pools should persist without going extinct or merging

43 in areas with low ORD, leading to long term splitting and diverging of clades; Jansson &

44 Dynesius (2002) call this β-clade formation where evolutionary change persists for more than

45 100 kyr. Here I explore the role of ORD on a regional scale, specifically in terms of the

46 interaction between glacial–inter-glacial climate changes and topography (i.e. watershed

47 configuration) on a tree species, *Nymania capensis* (Meliaceae), across a section of the

48 subescarpment coastal plains in South Africa.

49

50 *Nymania capensis* is associated with the subtropical thicket vegetation in South Africa, locally

51 termed the Albany Subtropical Thicket. This region forms the western part of the Maputaland-

52 Pondoland-Albany biodiversity hotspot (Steenkamp et al. 2004). Thicket vegetation is

53 characterised as dense, woody, semi-succulent and thorny, with an average height of 2-3 m,

54 and is relatively impenetrable in a pristine condition (Vlok et al. 2003). It is largely restricted to

55 the subescarpment coastal plains along the south coast of South Africa where it spans a

56 number of deeply incised primary catchments. Although thicket has a continuous or near-

57   continuous distribution across catchments, each catchment has been identified as a discrete

58   biogeographic unit (Vlok et al. 2003) suggesting that the watersheds have been dispersal

59   barriers to thicket plant species. This pattern, termed here as the "evolutionarily discrete

60   catchment" (EDC) hypothesis, has formed the foundation of large-scale conservation planning

61   that aims to ensure the persistence of evolutionary processes for thicket biota (Rouget et al.

62   2006). This hypothesis can be seen as a specific case of low ORD where the topographic

63   heterogeneity of the deeply incised subescarpment lowlands coupled with climate stability has

64   enabled β-clade formation, i.e. clades that have survived and remained genetically separate for

65   longer than 100 kyr. Under the EDC hypothesis, a number of predictions can be generated if

66   catchments and their associated watersheds are important landscape features responsible for

67   structuring genetic diversity (sensu Potts et al. 2013c; Price et al. 2010):

68   (1) Different catchments should contain genetically distinct lineages,

69   (2) There should be limited genetic structuring within catchments and a high degree of

70   genetic structuring among basins, and

71   (3) If watersheds have been long term barriers to gene flow, then isolated catchment

72   populations should contain evidence of genealogical sorting.

73

74   Southern Africa did not experience glaciation during the Pleistocene climate cycles (Partridge

75   1997). Nonetheless, these watershed barriers are suggested to have been greatly strengthened

76   during the Pleistocene as thicket is predicted to have severely contracted into, and fragmented

77   across, catchments (Cowling et al. 2005; Potts et al. 2013c): a glacial multiple refugia (GMR)

78 hypothesis (a low ORD scenario). If thicket species were adversely affected by glacial climates

79 through the Pleistocene, in interaction with the topography, then the following is predicted:

80   (4) The timing of population subdivision between the catchments should coincide with the

81   onset of glacial-interglacial cycles as these catchments are largely fixed and stable

82   landscape features through this period (Cowling et al. 2009), but also likely form β-

83   clades that have persisted for longer than 100 kyr (Jansson & Dynesius 2002), and

84   (5) The species' distribution should experience contraction and fragmentation that is

85   consistent with phylogeographic patterns.

86

87 There is growing support that these EDC-GMR hypotheses apply to thicket vegetation (Duker et

88 al. 2015a; Potts et al. 2013c). Plastid lineages from thicket taxa *Nymania capensis* and *Pappea*

89 *capensis* have previously been shown to be isolated to primary catchments (Potts et al. 2013b).

90 However, being strongly geographically constrained, plastid markers may give a biased

91 representation of a species' genetic coherence (Premoli et al. 2012) or evolutionary history, as

92 the smaller effective population size of uniparentally-inherited markers render them more

93 susceptible to stochastic processes (Edwards & Beerli 2000). Therefore, this study provides an

94 in-depth phylogeographic examination of *Nymania capensis* by combining the results from

95 Potts et al. (2013b) with data from high and low copy nuclear regions. These data are used to

96 test the EDC hypothesis by determining whether populations within populations are

97 evolutionarily significant units (sensu Moritz 1994). In addition, molecular dating of the

98 chloroplast dataset and hindcasting of species distribution models are used to explore the GMR

99   hypothesis. The results are contrasted with the predictions generated from these two

100   hypotheses.

101

Materials and Methods

Study system

104   Thicket is restricted to the year-round rainfall zone along the coastal lowlands of South Africa. A

105   winter rainfall zone is found to the west and a summer rainfall zone to the east. The

106   subescarpment coastal lowlands form a series of short but deeply incised catchments separated

107   from an unusually elevated interior plateau by the Great Escarpment. The coastal lowland

108   landscape has been topographically stable and relatively unchanged since the end of the

109   Pliocene (~2.6 Ma; Cowling et al. 2009) and possibly even longer (Scharf et al. 2013).

110

111   The distribution of *N. capensis* in the Albany Subtropical Thicket spans three primary

112   catchments along the coastal lowlands, specifically the Gouritz, Gamtoos and Sundays (named

113   after the major rivers within each catchment). However, the Gouritz catchment has additional

114   topographic complexity as two parallel mountain ranges associated with the Cape Fold Belt run

115   across it, creating an intermontane basin known as the Little Karoo. An inselberg, the Rooiberg

116   Mountain, splits this intermontane basin along a west-east orientation (see Fig. 1a). This

117   inselberg has been found to be a barrier to gene flow in another terrestrial plant species

118   endemic to the Little Karoo (Potts et al. 2013a), so the secondary catchments west and east of

119   the Rooiberg (Groot and Olifants, respectively) were sampled as extensively as the primary

120   catchments.

121

122   ***Study species***

123   *Nymania* is a monotypic, woody genus of Meliaceae restricted to southern Africa (see Appendix

124   S1 for images of the growth habit, flowers and inflated seed capsules). Solitary flowers are

125   borne on leaf axils and are generally pollinated by insects (J.H.J. Vlok and A.J. Potts, pers. obs.).

126   The fruit is an inflated and deeply lobed capsule with papery thin membranes, and contains

127   numerous small seeds. The seeds are carried away from the parent plant within the light

128   inflated capsules that are blown along the ground by wind or carried along by water. The

129   inflated capsule is peculiar in this species as it is not described anywhere else in the Meliaceae

130   (Pennington & Styles 1975).

131

132   This species occurs across a wide range of environmental conditions and is split into two

133   disjunct distributions, one in the southern semi-arid region (~300–600 mm annual rainfall) of

134   the Albany Subtropical Thicket and the other in the northern arid region (~100–300 mm annual

135   rainfall) of South Africa, which is a mosaic of Subtropical Thicket and Nama-Karoo vegetation.

136   Preliminary morphological (A.E. van Wyk, pers. com.) and genetic data (two samples used in

137   this study) suggest that the northern distribution that extends from the Nama-Karoo into

138   Namibia and the southern distribution, which is restricted to the Albany Thicket, are different

139   species. The northern and southern distributions of *N. capensis* are separated by a distance of

140   over 400 km spanning the Great Escarpment mountain range and the interior elevated plateau

141 of South Africa. This species was selected because (i) it is a common component of the Albany

142 Thicket, (ii) for its ease of location and identification in the landscape, and (iii) human activities,

143 such as livestock farming, do not appear to have affected its local distribution in the landscape

144 (personal observations).

145

146 ***Samples and DNA data***

147 Sampling consisted of one individual from each of 78 localities across the thicket. Sampling

148 permits were obtained from the Eastern Cape Parks and Tourism Agency and CapeNature

149 (0028-AAA008-00015). The sampling strategy was focussed on maximising the number of sites

150 sampled in order to explore the broad scale regional patterns rather than intra-population

151 differences. This scattered sampling strategy is also not affected by local and rapid coalescence

152 events (Städler et al. 2009) and thus gives an unbiased view of population structuring and

153 demographic history. Thus, each catchment was represented by 14 to 26 individuals, with

154 nearly all individuals selected from localities that were over 10 km apart. Two samples of *N.*

155 *capensis* from the northern distribution stored in the Bolus Herbarium were used to generate

156 outgroup accessions (BOL48535 and BOL60966).

157

158 Three DNA sequence markers were analysed: chloroplast DNA and nuclear DNA from high and

159 low copy regions. Chloroplast DNA accessions of the trnQ-5′–rps16 and atpI–atpH regions (from

160 Potts et al. 2013b) are treated as linked regions as they are inherited in tandem. Nuclear DNA

161 sequence data were obtained from two loci, the high-copy ITS region of the 35S rDNA cistron

162 (the 5.8S rDNA and the flanking internal transcribed spacers, ITS-1 and ITS-2) (from Potts et al.

163   2014, complemented with additional accessions) and the low-copy chloroplast-expressed

164   glutamine synthetase gene (ncpGS; new data). Extraction and PCR protocols are reported in

165   Appendix S2.

166

167   Details of sequence alignment and, in the case of nuclear DNA, the detection of intra-individual

168   site polymorphisms (2ISPs) are given in Potts et al. (2014). In brief, multiple variants of a nuclear

169   region may be present within an individual which give rise to 2ISPs. These can be observed as

170   double (and even triple) peaks in trace files and were coded using the extended IUPAC

171   nucleotide codes (e.g. Y, R); these codes are commonly referred to as 'ambiguity' codes, but it is

172   important to note that 2ISPs are not the result of uncertain coding due to scrappy trace files but

173   rather the clear presence of two or more bases. Numerous 2ISPs were observed in direct-PCR

174   sequences from both nuclear regions, thus indicating the presence of more than one DNA

175   variant present in the genome. This has been observed in many other plant species (Bailey et al.

176   2003), and is unsurprising given the complexity of the nuclear regions, especially ITS.

177

178   How accurate is direct-PCR sequencing for detecting 2ISPs? This question has not been

179   investigated in-depth in this study or elsewhere. However, using a small subset of *N. capensis*

180   samples, I found a high degree of congruence between direct-PCR and cloned samples with a

181   low false detection rate; two other studies show similar findings (see Appendix S2 for further

182   details). Both ITS and ncpGS cloned samples demonstrated that these regions contained

183   multiple copies of each region per sample (see supplementary data files). This rules out

184   statistically inferring phased haplotypes from polymorphic sequences as these methods assume

185   a maximum of two copies per individual (e.g. PHASE, Stephens et al. 2001) and cloning all

186   samples was prohibitively expensive.

187   All datasets aligned readily and reliable gaps corresponding to insertion or deletion events were

188   included as informative characters; indels that were not associated with homopolymer repeats

189   were coded as binary characters and used in all analyses (multiple base indels were treated as

190   single characters). Where 2ISPs comprised an indel and a base or bases (observed in the ITS

191   dataset), then these were given a unique code for inclusion into subsequent analyses as no

192   IUPAC nucleotide code exists for this combination.

193

194   ***Phylogenetic and phylogeographic analyses***

195   The phylogenetic relationships between chloroplast, ncpGS haplotypes and all ITS ribotypes

196   were inferred using Statistical Parsimony (SP) and NeighbourNet (NN) networks in TCS version

197   1.21 (Clement et al. 2000) and SplitsTree version 4.8 (Huson & Bryant 2006), respectively. In

198   order to accommodate the 2ISPs present in the ncpGS and ITS datasets, polymorphism *p*-

199   distances (Potts et al. 2014) were used to calculate both SP and NN networks; this measure

200   specifically incorporates 2ISPs when calculating genetic distances (these sites are usually

201   removed or averaged) and was obtained using the PHANGORN library (Schliep 2011) in R3.0.1 (R

202   Core Team 2015). Consensus networks generated from 1000 bootstraps using Neighbour

203   Joining, Maximum Parsimony or Maximum Likelihood were used to assess edge support in the

204   SP and NN networks; details of software used and settings are given in Potts et al. (2014) but, in

205   brief, PAUP* version 4b10 (Swofford 2002) and RAxML version 7.2.6 (Stamatakis 2006;

206   Stamatakis et al. 2008) were used under the 2ISP-informative approach. The standard

207   treatment of 2ISPs by phylogenetic software is that the actual nucleotide is uncertain (or

208   'ambiguous') and this usually leads to an erosion of phylogenetic support. The 2ISP-informative

209   approach avoids this by forcing the phylogenetic algorithms to treat 2ISPs as additional

210   information (Potts et al. 2014).

211

212   Bayesian software such as MrBayes (Ronquist & Huelsenbeck 2003) or BEAST (Drummond &

213   Rambaut 2007) cannot incorporate 2ISPs as informative characters (Potts et al. 2014) and

214   return unresolved topologies for the ITS and ncpGS alignments (results not shown).

215

216   The level of genealogical divergence was assessed using the genealogical sorting index (gsi,

217   Cummings et al. 2008). The *gsi* statistic is a standardised measure of the extent to which

218   predefined groups in a gene tree exhibit exclusive ancestry (fall within a distinct clade); the *gsi*

219   statistic ranges from 0 (a complete lack of genealogical divergence with respect to other

220   groups) to 1 (monophyly: all elements of a predefined group fall within the same, exclusive

221   subtree). Further details are provided in Appendix S2.

222

223   Pairwise genetic and geographic distances among sampling locations within and between

224   catchments were used to test patterns of isolation by distance using a Mantel test (Mantel

225   1967). Genetic distance for cpDNA was calculated using uncorrected *p*-distances, whereas ITS

226   and ncpGS genetic distances were calculated using polymorphism *p*-distances (Potts et al.,

227   2014). Standard genetic and geographic distances were calculated using the APE library version

228   3.0-9 in R (Paradis et al. 2004). The probability and significance of the correlation coefficient

229   (Spearman's R) was estimated after 10,000 permutations.

230

231   ***Molecular dating***

232   Given the dominance of 2ISPs in the nuclear datasets, only the cpDNA could be used for

233   molecular dating; a molecular clock approach using such individual consensus sequences is not

234   yet feasible and the alternative 2ISP resampling procedure suggested by Lischer et al. (2013)

235   generates exceptionally wide and uninformative date ranges due to the high proportion of

236   2ISPs in the ITS dataset (results not shown). Assigning a timescale to phylogenies typically

237   involves ingroup fossils as age constraints or applying a rate of molecular evolution. No fossil or

238   pollen data specific to *N. capensis* are available, so the hypothesis of Pleistocene lineage

239   evolution was tested using a highly conservative approach by using the 'extremes' of

240   substitution rates found in the literature for non-coding chloroplast DNA, specifically 1.0×10^{-9}

241   (Richardson et al. 2001) and 31×10^{-9} substitutions per site per year (Fu & Allaby 2010). The

242   dating of lineage divergence was carried out using BEAST version 1.4.8 (Drummond & Rambaut

243   2007), which estimates the tree topology and the date (height) of nodes simultaneously using a

244   Bayesian approach. The BEAST analysis was performed using all cpDNA samples as this sampling

245   is important for coalescent estimation of lineage divergence. Further details are provided in

246   Appendix S2.

247

248 ***Distribution modelling***

249 The details of the species distribution modelling, including hindcasting the models to Last

250 Glacial Maximum climate conditions are given in Potts et al. (2013b). In brief, an ensemble

251 species distribution modelling approach (sensu Araújo & New 2007) was used where the

252 uncertainty generated by locality selection, parameter selection of current climate and

253 distribution modelling algorithm could be assessed; this resulted in an ensemble of 216 models.

254 These models were then projected onto the statistically downscaled CCSM global climate

255 model that is available from http://www.worldclim.org. All projections were converted to

256 binary presence-absence maps using the equal sensitivity plus specificity threshold criterion.

257 These models were further interrogated, using the SDMTOOLS library version 1.1-13 (VanDerWal

258 et al. 2011) in R3.0.1, by calculating the area and patch cohesion index within each catchment

259 for each model combination within the ensemble for both current and past climate.

260

261

262 **Results**

263 ***Genetic data characteristics***

264 The final chloroplast dataset is comprised of 1948 bp (trnQ-5'-rps16: 791 bp; atpI-atpH: 1157

265 bp). The ITS and ncpGS datasets were comprised of 666 bp and 1089 bp, respectively. Sequence

266 characteristics for each dataset are summarised in Table 1. Haplotype/ribotype and nucleotide

267 diversity were high in the ITS dataset, moderate in the chloroplast dataset, and low in the

268   ncpGS dataset. Summaries of variable sites across the chloroplast and ncpGS haplotypes, as

269   well as a subset of ITS samples, are shown in Appendix S3, respectively.

270

271   Two individuals situated in the Gamtoos had chloroplast haplotypes and ITS genotypes

272   associated with the Sundays catchment lineage, however these were very close (≤5 km) to the

273   watershed suggesting incorrect watershed delineation or recent migration rather than

274   incomplete lineage sorting. Thus, these samples were grouped with the Sundays basin samples

275   for all analyses.

276   ***Phylogeographic analyses***

277   Under the evolutionarily discrete catchment (EDC) hypothesis, different catchments should

278   contain genetically distinct lineages (Prediction 1). The catchment association was strong with

279   chloroplast haplotypes, with all but one haplotype restricted to one of the primary or secondary

280   catchments (Fig. 1b). A strong association was also evident between primary catchments and

281   ITS clusters (Fig. 1d) with two exceptions (indicated with arrows, discussed further below). Only

282   rare ncpGS haplotypes were restricted to catchments, whereas two widespread haplotypes

283   were shared either among both subregions of the Gouritz or among both Gamtoos and Sundays

284   catchments (Fig. 1c). Two samples contained anomalous ncpGS haplotypes; these haplotypes

285   are found in an intermediate position between the western and eastern haplotypes. They may

286   represent inherited ancestral copy diversity or recent gene flow (subsequent intragenomic

287   recombination) between samples from the western and eastern basins. However, given the

288   slow rate of mutation in this region, evident by the low genetic diversity, and that one of these

289   intermediate haplotypes is also shared with an outgroup sample from the northern distribution

290   (BOL48535; Supplementary files), it is likely that these haplotypes represent ancestral and

291   unsorted ncpGS copies.

292

293   Samples AJP0537 and AJP0810 (indicated with arrows in Figure 1b-d) lie within the Gamtoos

294   catchment, but are close to sampled localities with a Sundays catchment genetic signature

295   found on the watershed boundary between the catchments. These samples have Gamtoos

296   chloroplast haplotypes, and for ncpGS, a rare Gamtoos haplotype, as well as a common and

297   widespread haplotype. However, the first sample has an ITS signal that nests it within the

298   Sundays ITS cluster 3, while the second displays a recombinant signal that lies between the

299   Gamtoos cluster 2 and cluster 3 (Fig. 1d). The contrasting basin associations between these two

300   samples suggest that they represent a contact zone between the Sundays and Gamtoos

301   lineages, possibly caused by pollen flow between plants in this zone. As these samples are

302   anomalous to the overall patterns of association, and are geographically restricted, they were

303   removed for all subsequent analyses.

304

305   Isolation by distance using the Mantel Test (MT) was used to determine if there was genetic

306   structuring within and between catchments (Prediction 2). In general, non-significant MT_R

307   values were observed within the primary or secondary catchments (Table 1) across the

308   different DNA regions; however, there were a few instances where significant but low MT_R (ITS

309   in Sundays; ncpGS in Olifants) and significant but high MT_R values (in comparison to the overall

310   value; ITS in Gamtoos) were observed. There were significant and high MT_R values between the

311   Olifants and Gamtoos basins across all DNA regions. Significant MT_R values between the Groot

312   and Olifants and the Gamtoos and Sundays were only observed in the cpDNA and ITS; however,

313   the MT_R values were low for the former pair of basins and high for the latter pair.

314

315   If watersheds have been evolutionarily long term barriers to gene flow then topographic-driven

316   isolation should be evident, culminating in sorting, and ultimately monophyly, of lineages

317   within catchments (Prediction 3). Although monophyletic lineages were detected with

318   moderate to high support (60%-98%) in the chloroplast bootstrap phylogenies, only samples

319   from the Gamtoos catchment displayed catchment-based monophyly (Fig. 1b). In contrast, the

320   ITS produces high support for Gouritz, Gamtoos and Sundays lineages (Fig. 1d). No

321   monophyletic lineages were found to be restricted to any catchments in the ncpGS data, but

322   this is due to two widespread haplotypes and low diversification rates. Despite lack of

323   supported monophyly in the cpDNA and ncpGS analyses, the *gsi* results do suggest a high level

324   of lineage sorting (>0.600) in the western Gouritz basin for both datasets, and in the Gamtoos

325   and Sundays for cpDNA (Table 1).

326

327   A BEAST analysis was used to date the diversification of chloroplast lineages in order to

328   determine whether the Pleistocene climate cycles have also played a role in population

329   isolation in catchments (Prediction 4). The timing of all of the thicket lineages fall firmly within

330   the Pleistocene whether a 'fast' or 'slow' rate of chloroplast mutation is used (Fig. 2). Species

331   distribution modelling suggests that *N. capensis* experienced major range contractions and

332   fragmentation during the Last Glacial Maximum (Fig. 3, Prediction 5), finding refugia within

333   each of the catchments and current distribution.

334

335 **Discussion**

336

337 In a broad sense, this study makes use of both genetic and geospatial data in order to explore

338 β-clade formation driven by "orbitally-forced range dynamics" (ORD; Jansson & Dynesius 2002)

339 in a region that remained free of glaciation during the Pleistocene. More specifically, it tests the

340 predictions of two ORD-related hypotheses regarding the effects of topography and Pleistocene

341 climate fluctuations on the genetic diversity of *Nymania capensis* within the Albany Subtropical

342 Thicket: the evolutionarily distinct catchment (EDC) hypothesis and the glacial multiple refugia

343 (GMR) hypothesis. The former hypothesis was generated from a phytosociological study of

344 thicket (Vlok et al. 2003) where catchments were identified as discrete biogeographic units,

345 whereas the latter was generated from biome-level distribution modelling (Potts et al. 2013c)

346 and observations of physiological sensitivity to declining temperatures of thicket species

347 (Cowling et al. 2005; Duker et al. 2015a). From these hypotheses, five predictions were

348 generated and tested using *Nymania capensis*. The results suggest that catchments have been

349 *in situ* refugia (Gavin et al. 2014; Tzedakis 1993) for *N. capensis* during glacial-interglacial cycles

350 with glacial refugia largely located within the interglacial distribution.

351

352 ***Evolutionarily discrete catchments: watersheds and catchments as drivers of diversification***

353

354   Investigating the role that catchment topography plays in reducing gene flow and driving

355   population divergence has largely been restricted to either obligate freshwater animal species

356   (e.g. redfins, Swartz et al. 2009) or freshwater animal species capable of terrestrial movement

357   (e.g. freshwater crayfish, crabs, salamanders, and spotted frogs; Cook et al. 2008; Funk et al.

358   2005; Giordano et al. 2007; Ponniah & Hughes 2006). Only recently have catchments and

359   watersheds been explored as drivers of diversification in terrestrial species that do not rely on

360   the riparian system, and these have mostly focussed on invertebrates (e.g. springtails and

361   cicadas; Garrick et al. 2007; Garrick et al. 2004; Price et al. 2010). The EDC hypothesis for the

362   catchments in this study area is supported by two other studies, one on fish (Swartz et al. 2009)

363   and the other on cicadas (Price et al. 2010).

364

365   The phylogeographic patterns of *N. capensis* are largely consistent with the three predictions

366   deduced from the EDC hypothesis. Firstly, the predominant pattern is one of genetically distinct

367   lineages restricted to single catchments (Fig. 1), although a few lineages span neighbouring

368   catchments. The number of genetically distinct lineages varies between the different markers;

369   this is expected given their different rates of mutation and effective population sizes. Secondly,

370   there is limited genetic structuring within catchments as most analyses did not detect a within-

371   basin signal of isolation by distance. In contrast, a large and significant isolation by distance

372   effect is evident across all catchments (Table 1). This suggests that gene flow is hindered

373   between basins, while this is not the case within basins. Lastly, although catchments do not

374   contain genetically homogenised and fully isolated lineages, there is strong evidence of

375   genealogical sorting within each catchment (Table 1).

376

377   The strong phylogeographic break between the western Gouritz and the eastern Gamtoos and

378   Sundays catchments detected in this study corresponds to previously documented breaks

379   observed within *Meleuphorbia* (Euphorbiaceae, Ritz et al. 2003) and *Platypleura plumosa*

380   (Hemiptera: Cicadidae, Price et al. 2010). The Rooiberg inselberg that separates the Groot and

381   Olifants basin has been observed as a phylogeographic break in *Platypleura karooensis*

382   (Hemiptera: Cicadidae, Price et al. 2010) and *Berkheya cuneata* (Asteraceae, Potts et al. 2013a).

383   However, although only one chloroplast haplotype is shared between these subcatchments,

384   which is suggestive of restricted seed flow, there is very weak support for isolation by distance

385   or genealogical sorting across all DNA regions.

386

387   Thus, the results of this study provide additional support for the EDC hypothesis as the marked

388   genetic structuring revealed in *N. capensis* is consistent with patterns predicted on the basis of

389   primary catchment divisions. However, although the landscape has been stable since the Late

390   Pliocene (Cowling et al. 2009), dramatic cycling between glacial and inter-glacial climates has

391   occurred during the Pleistocene. These shifts may have also affected the distribution,

392   fragmentation and divergence of thicket species such as *Nymania capensis* (Cowling et al. 2005;

393   Dynesius & Jansson 2000).

394

395   ***Glacial refugia: Pleistocene climatic cycles as drivers of diversification***

396   If the Pleistocene climate cycling between glacial and interglacial periods was responsible for

397   isolating populations in this landscape – which has been stable since the late Pliocene (~2.6 Ma;

398   Cowling et al. 2009) – then it would be expected that lineage diversification would coincide or

399   occur after the onset of these cycles. There are many potential problems with the molecular

400   clock approach (Graur & Martin 2004; Ho 2007), and selecting an accurate rate of substitution

401   for target taxa is of primary concern when fossil calibration is not possible. Using a wide range

402   of published substitution rates for non-coding chloroplast DNA should circumvent the lack of an

403   accurate rate for *Nymania capensis*. Using both slow and fast substitution rates, the divergence

404   of all thicket lineages falls well within the Pleistocene (Fig. 2), suggesting that this species has

405   experienced fragmentation and isolated diversification during this period. Under the fast

406   substitution rate, many of the lineages diverge after the Last Glacial Maximum. This is,

407   however, an unlikely scenario given that this is an exceptionally fast rate that has been derived

408   from a genus of annual herbs (*Linum*) and the fast turnover in generations is likely to have

409   greatly increased the substitution rate (Kay et al. 2006). *Nymania capensis* is a perennial plant

410   that requires more than five years of ideal environmental conditions before flowers and seeds

411   are produced (Jan Vlok, pers. comm.). It is thus highly unlikely that its substitution rate would

412   coincide with that of an annual herb, leading to the conclusion that lineage diversification fell

413   within the Pleistocene. The slower substitution rate also falls within the generally accepted

414   range of 1.0 to 3.0×10^{-9} (Wolfe et al. 1987).

415

416   Climatic changes during the Pleistocene glacial cycles have induced distributional shifts in

417   species, often resulting in fragmentation and divergence of populations (Hewitt 2004; Fig. 3;

418   Jansson & Dynesius 2002). Identifying refugial areas during glacial periods through the

419   Pleistocene has been a strong focus within many phylogeographic studies (Avise 2000), with the

420   majority of studies focussed on the previously glaciated northern hemisphere regions (Abbott

421   et al. 2000; Soltis et al. 1997; Taberlet et al. 1998). Southern Africa did not experience glaciation

422   during the Pleistocene climate cycles (Partridge 1997) and determining refugia for plant species

423   in areas that did not experience glaciation is a challenging process (e.g. Byrne 2007).

424   Nevertheless, in the Cape Floristic Region, which neighbours the thicket, climatic fluctuations

425   during the Pleistocene have been suggested to be the main driver of fragmentation and shifts in

426   faunal species distributions, culminating in allopatric diversification (e.g. rock agama, dwarf

427   chameleons, and cicadas; Price et al. 2007; Swart et al. 2009; Tolley et al. 2006). The lower

428   temperatures, both globally (Zachos et al. 2001) and regionally (Holmgren et al. 2003; Talma &

429   Vogel 1992), coupled with the frost sensitivity observed in many of the thicket's component

430   species (Duker et al. 2015a; Duker et al. 2015b) is suggested to have driven thicket vegetation

431   into fragmented refugia (Cowling et al. 2005), as reflected by *Nymania capensis*.

432

433   Climatic changes will drive shifts in geographic distributions of species (Jansson & Dynesius

434   2002) if they cannot adapt to new ecological conditions, at least over moderate periods of time.

435   Evidence suggests that niche conservatism may be general and pervasive across most species

436   over moderate periods of time, despite profound changes in climate and environmental

437   conditions (Martínez-Meyer & Peterson 2006; Martínez-Meyer et al. 2004; Peterson et al.

438   1999). Here I assume that the niche of *N. capensis* has been largely conserved from the LGM to

439   the present.

440   The species distribution modelling results of *N. capensis* suggest this species' range contracted

441   and fragmented into and within the primary catchments (Fig. 3). This is consistent with the

442   phylogeographic evidence that suggests isolation into at least three refugia which correspond

443   to the delimitation of primary basins (Fig. 1). Thus, the retraction into catchments during glacial

444   periods through the Pleistocene would have strengthened the effects of watershed barriers to

445   gene flow. This retraction is consistent with the limited palaeodata (reviewed in Cowling et al.

446   2005) and community distribution modelling of thicket subtypes (Potts et al. 2013c) that also

447   suggests that thicket suffered significant range constrictions during the most recent Pleistocene

448   glacial period. Also, the absence of vertebrates endemic to the thicket, which would be

449   expected given the present-day area of the biome, is suggestive of historical reduction and

450   fragmentation of the biome's distribution (Hoare et al. 2006).

451

452   The EDC hypothesis has been used as the cornerstone for conservation planning in the Albany

453   Subtropical Thicket in order to conserve both the biodiversity patterns and evolutionary

454   processes of this vegetation (Rouget et al. 2006). Specifically, a number of conservation

455   corridors have been identified to create a mega-conservancy network; these corridors are

456   predominantly focussed on conserving major environmental gradients primarily within

457   catchments. These results demonstrate that the patterns observed in the chloroplast

458   phylogeography (Potts et al. 2013b) are mirrored in the nuclear genome, suggesting that the

459   catchments house evolutionarily significant units (sensu Crandall et al. 2000; Moritz 1994).

460

461   In conclusion, the results from these phylogeographic analyses and ensemble niche modelling

462   suggest that the genetic structuring of *N. capensis* has been determined by landscape topology

463   coupled with the effects of fluctuating Pleistocene climates. Populations have been restricted

464   to three primary catchments with no extensive present-day or historical gene flow during the

465   previous glacial period. This is largely consistent with the predictions derived from a scenario

466   where landscape topography and climatic fluctuations were responsible for structuring

467   populations across the southern African lowlands. From their glacial refugia they (re-)colonised

468   their modern range, a process that is apparently accompanied by secondary contact and

469   introgression ('anomalous' Gamtoos samples). The findings validate and further highlight the

470   importance of including catchments in conservation planning as discrete and evolutionarily

471   significant units important for maintaining genetic resources.

472

473   **Acknowledgements**

474   The author gratefully acknowledges Richard M. Cowling, Terry A. Hedderson, Guido W. Grimm,

475   and three anonymous reviewers for comments that greatly improved this manuscript. In

476   addition, Jan Vlok is thanked for sharing his knowledge regarding his extensive natural history

477   observation of this species and the thicket vegetation.

478

479   **References**

480   Abbott RJ, Smith LC, Milne RI, Crawford RMM, Wolff K, and Balfour J. 2000. Molecular
481   analysis of plant migration and refugia in the Arctic. *Science* 289:1343-1346.
482   Araújo MB, and New M. 2007. Ensemble forecasting of species distributions. *Trends in Ecology*
483   *& Evolution* 22:42-47.
484   Avise JC. 2000. *Phylogeography: the History and Formation of Species*. Cambridge,
485   Massachusetts: Harvard University Press.
486   Bailey CD, Carr TG, Harris SA, and Hughes CE. 2003. Characterization of angiosperm nrDNA
487   polymorphism, paralogy, and pseudogenes. *Molecular Phylogenetics and Evolution*
488   29:435-455.

489   Byrne M. 2007. Phylogeography provides an evolutionary context for the conservation of a
490   diverse and ancient flora. *Australian Journal of Botany* 55:316-325.
491   Clement M, Posada D, and Crandall K. 2000. TCS: a computer program to estimate gene
492   genealogies. *Molecular Ecology* 9:1657-1660.
493   Cook BD, Pringle CM, and Hughes JM. 2008. Phylogeography of an island endemic, the Puerto
494   Rican freshwater crab (*Epilobocera sinuatifrons*). *Journal of Heredity* 99:157-164.
495   Cowling RM, Proches S, and Partridge TC. 2009. Explaining the uniqueness of the Cape flora:
496   incorporating geomorphic evolution as a factor for explaining its diversification.
497   *Molecular Phylogenetics and Evolution* 51:64-74.
498   Cowling RM, Proches S, and Vlok JHJ. 2005. On the origin of southern African subtropical
499   thicket vegetation. *South African Journal of Botany* 71:1-23.
500   Crandall KA, Bininda-Emonds ORP, Mace GM, and Wayne RK. 2000. Considering
501   evolutionary processes in conservation biology. *Trends in Ecology & Evolution* 15:290-
502   295.
503   Cummings MP, Neel MC, Shaw KL, and Otto S. 2008. A genealogical approach to quantifying
504   lineage divergence. *Evolution* 62:2411-2422.
505   Drummond AJ, and Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling
506   trees. *BMC Evolutionary Biology* 7:214.
507   Duker R, Cowling RM, du Preez DR, van der Vyver ML, Weatherall-Thomas CR, and Potts AJ.
508   2015a. Community-level assessment of freezing tolerance: frost dictates the biome
509   boundary between Albany Subtropical Thicket and Nama-Karoo in South Africa. *Journal*
510   *of Biogeography* 42:167–178.
511   Duker R, Cowling RM, Preez DRd, and Potts AJ. 2015b. Frost, *Portulacaria afra* Jacq., and the
512   boundary between the Albany Subtropical Thicket and Nama-Karoo biomes. *South*
513   *African Journal of Botany* 101:112-119.
514   Dynesius M, and Jansson R. 2000. Evolutionary consequences of changes in species'
515   geographical distributions driven by Milankovitch climate oscillations. *Proceedings of*
516   *the National Academy of Sciences, USA* 97:9115-9120.
517   Edwards SV, and Beerli P. 2000. Perspective: Gene divergence, population divergence, and the
518   variance in coalescence time in phylogeographic studies. *Evolution* 54:1839-1854.
519   Fu Y-B, and Allaby R. 2010. Phylogenetic network of *Linum* species as revealed by non-coding
520   chloroplast DNA sequences. *Genetic Resources and Crop Evolution* 57:667-677.
521   Funk WC, Blouin MS, Corn PS, Maxell BA, Pilliod DS, Amish S, and Allendorf FW. 2005.
522   Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by
523   the landscape. *Molecular Ecology* 14:483-496.
524   Garrick RC, Sands CJ, Rowell DM, Hillis DM, and Sunnucks P. 2007. Catchments catch all:
525   long-term population history of a giant springtail from the southeast Australian highlands
526   - a multigene approach. *Molecular Ecology* 16:1865-1882.
527   Garrick RC, Sands CJ, Rowell DM, Tait NN, Greenslade P, and Sunnucks P. 2004.
528   Phylogeography recapitulates topography: very fine-scale local endemism of a saproxylic
529   'giant' springtail at Tallaganda in the Great Dividing Range of south-east Australia.
530   *Molecular Ecology* 13:3329-3344.
531   Gavin DG, Fitzpatrick MC, Gugger PF, Heath KD, Rodríguez-Sánchez F, Dobrowski SZ,
532   Hampe A, Hu FS, Ashcroft MB, Bartlein PJ, Blois JL, Carstens BC, Davis EB, de
533   Lafontaine G, Edwards ME, Fernandez M, Henne PD, Herring EM, Holden ZA, Kong
534   W-s, Liu J, Magri D, Matzke NJ, McGlone MS, Saltré F, Stigall AL, Tsai Y-HE, and

535   Williams JW. 2014. Climate refugia: joint inference from fossil records, species
536   distribution models and phylogeography. *New Phytologist* 204:37-54.
537 Giordano AR, Ridenhour BJ, and Storfer A. 2007. The influence of altitude and topography on
538   genetic structure in the long-toed salamander (*Ambystoma macrodactulym*). *Molecular*
539   *Ecology* 16:1625-1637.
540 Graur D, and Martin W. 2004. Reading the entrails of chickens: molecular timescales of
541   evolution and the illusion of precision. *Trends in Genetics* 20:80-86.
542 Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary.
543   *Philosophical Transactions of the Royal Society of London Series B* 359:183-195.
544 Ho SYW. 2007. Calibrating molecular estimates of substitution rates and divergence times in
545   birds. *Journal of Avian Biology* 38:409-414.
546 Hoare DB, Mucina L, Rutherford MC, Vlok JHJ, Euston-Brown DIW, Palmer AR, Powrie LW,
547   Lechmere-Oertel RG, Proches S, and Dold AP. 2006. Albany Thicket biome. In: Mucina
548   L, and Rutherford MC, eds. *The vegetation of South Africa, Lesotho and Swaziland*. Cape
549   Town: South African National Biodiversity Institute.
550 Holmgren K, Lee-Thorp JA, Cooper GRJ, Lundblad K, Partridge TC, Scott L, Sithaldeen R,
551   Talma AS, and Tyson PD. 2003. Persistent millennial-scale climatic variability over the
552   past 25,000 years in Southern Africa. *Quaternary Science Reviews* 22:2311-2326.
553 Huson DH, and Bryant D. 2006. Application of phylogenetic networks in evolutionary studies.
554   *Molecular Biology and Evolution* 23:254-267.
555 Jansson R, and Dynesius M. 2002. The fate of clades in a world of recurrent climatic change:
556   Milankovitch oscillations and evolution. *Annual Review of Ecology and Systematics*
557   33:741-777.
558 Kay KM, Whittall JB, and Hodges SA. 2006. A survey of nuclear ribosomal internal transcribed
559   spacer substitution rates across angiosperms: an approximate molecular clock with life
560   history effects. *BMC Evolutionary Biology* 6:36.
561 Lischer HEL, Excoffier L, and Heckel G. 2013. Ignoring heterozygous sites biases
562   phylogenomic estimates of divergence times: implications for the evolutionary history of
563   Microtus voles. *Molecular Biology and Evolution*.
564 Mantel N. 1967. The detection of disease clustering and a generalized regression approach.
565   *Cancer Research* 27:209-220.
566 Martínez-Meyer E, and Peterson AT. 2006. Conservatism of ecological niche characteristics in
567   North American plant species over the Pleistocene-to-Recent transition. *Journal of*
568   *Biogeography* 33:1779-1789.
569 Martínez-Meyer E, Peterson AT, and Hargrove WW. 2004. Ecological niches as stable
570   distributional constraints on mammal species, with implications for Pleistocene
571   extinctions and climate change projections for biodiversity. *Global Ecology and*
572   *Biogeography* 13:305-314.
573 Moritz C. 1994. Defining 'evolutionarily significant units' for conservation. *Trends in Evolution*
574   *and Ecology* 9:373-375.
575 Paradis E, Claude J, and Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R
576   language. *Bioinformatics* 20:289-290.
577 Partridge TC. 1997. Cainozoic environmental change in southern Africa, with special emphasis
578   on the last 200 000 years. *Progress in Physical Geography* 21:3-22.
579 Pennington TD, and Styles BT. 1975. A generic monograph of the Meliaceae. *Blumea* 22:419-
580   540.

581 Peterson AT, Soberon J, and Sanchez-Cordero V. 1999. Conservatism of ecological niches in
582   evolutionary time. *Science* 285:1265-1267.
583 Ponniah M, and Hughes JM. 2006. The evolution of Queensland spiny mountain crayfish of the
584   genus Euastacus. II. Investigating simultaneous vicariance with intraspecific genetic data.
585   *Marine and Freshwater Research* 57:349-362.
586 Potts A, Hedderson T, Vlok JHJ, and Cowling RM. 2013a. Pleistocene range dynamics in the
587   eastern Greater Cape Floristic Region: a case study of the Little Karoo endemic *Berkheya*
588   *cuneata* (Asteraceae). *South African Journal of Botany* 88:401-413.
589 Potts AJ, Hedderson TA, and Cowling RM. 2013b. Testing large-scale conservation corridors
590   designed for patterns and processes: comparative phylogeography of three tree species.
591   *Diversity and Distributions* 19:1418-1428.
592 Potts AJ, Hedderson TA, Franklin J, and Cowling RMC. 2013c. The Last Glacial Maximum
593   distribution of Albany Subtropical Thicket inferred from community distribution
594   modelling. *Journal of Biogeography* 40:310-322.
595 Potts AJ, Hedderson TA, and Grimm GW. 2014. Constructing phylogenies in the presence of
596   *i*ntra-*i*ndividual *s*ite *p*olymorphisms (2ISPs) with a focus on the nuclear ribosomal
597   cistron. *Systematic Biology* 63:1-16.
598 Premoli AC, Mathiasen P, Cristina Acosta M, and Ramos VA. 2012. Phylogeographically
599   concordant chloroplast DNA divergence in sympatric Nothofagus *s.s.* How deep can it
600   be? *New Phytologist* 193:261-275.
601 Price BW, Barker NP, and Villet MH. 2007. Patterns and processes underlying evolutionary
602   significant units in the *Platypleura stridula* L. species complex (Hemiptera : Cicadidae)
603   in the Cape Floristic Region, South Africa. *Molecular Ecology* 16:2574-2588.
604 Price BW, Barker NP, and Villet MH. 2010. A watershed study on genetic diversity:
605   phylogenetic analysis of the *Platypleura plumosa* (Hemiptera: Cicadidae) complex
606   reveals catchment-specific lineages. *Molecular Phylogenetics and Evolution* 54:617-626.
607 R Core Team. 2015. *R: A language and environment for statistical computing*. Vienna, Austria:
608   R Foundation for Statistical Computing
609 Richardson JE, Pennington RT, Pennington TD, and Hollingsworth PM. 2001. Rapid
610   diversification of a species-rich genus of neotropical rain forest trees. *Science* 293:2242-
611   2245.
612 Ritz CM, Zimmermann NFA, and Hellwig FH. 2003. Phylogeny of subsect. *Meleuphorbia* (A.
613   Berger) Pax & Hoffm. (*Euphorbia* L.) reflects the climatic regime in South Africa. *Plant*
614   *Systematics and Evolution* 241:245-259.
615 Ronquist F, and Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under
616   mixed models. *Bioinformatics* 19:1572-1574.
617 Rouget M, Cowling RM, Lombard AT, Knight AT, and Kerley GIH. 2006. Designing large-
618   scale conservation corridors for pattern and process. *Conservation Biology* 20:549-561.
619 Scharf TE, Codilean AT, de Wit M, Jansen JD, and Kubik PW. 2013. Strong rocks sustain
620   ancient postorogenic topography in southern Africa. *Geology*.
621 Schliep KP. 2011. phangorn: phylogenetic analysis in R. *Bioinformatics* 27:592-593.
622 Soltis DE, Gitzendanner MA, Strenge DD, and Soltis PS. 1997. Chloroplast DNA intraspecific
623   phylogeography of plants from the Pacific Northwest of North America. *Plant*
624   *Systematics and Evolution* 206:353-373.

625 Städler T, Haubold B, Merino C, Stephan W, and Pfaffelhuber P. 2009. The impact of sampling
626   schemes on the site frequency spectrum in non-equilibrium subdivided populations.
627   *Genetics* 182:205-216.
628 Stamatakis A. 2006. RAxML-VI-HPC: Maximum Likelihood-based phylogenetic analyses with
629   thousands of taxa and mixed models. *Bioinformatics* 22:2688-2690.
630 Stamatakis A, Hoover P, and Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML
631   web servers. *Systematic Biology* 57:758-771.
632 Steenkamp Y, Van Wyk B, Victor J, Hoare D, Smith G, Dold T, and Cowling R. 2004.
633   Maputaland-Pondoland-Albany. In: Mittermeier RA, Robles Gil P, Hoffman M, Pilgrim
634   JD, Brooks TM, Mittermeier CG, Lamoreux J, and da Fonseca GAB, eds. *Hotspots
635   revisited: Earth's biologically richest and most endangered terrestrial ecoregions*.
636   Mexico City: CEMEX, 219-228.
637 Stephens M, Smith N, and Donnelly P. 2001. A new statistical method for haplotype
638   reconstruction from population data. *American Journal of Human Genetics* 68:978-989.
639 Swart BL, Tolley KA, and Matthee CA. 2009. Climate change drives speciation in the southern
640   rock agama (Agama atra) in the Cape Floristic Region, South Africa. *Journal of
641   Biogeography* 36:78-87.
642 Swartz ER, Skelton PH, and Bloomer P. 2009. Phylogeny and biogeography of the genus
643   *Pseudobarbus* (Cyprinidae): shedding light on the drainage history of rivers associated
644   with the Cape Floristic Region. *Molecular Phylogenetics and Evolution* 51:75-84.
645 Swofford DL. 2002. *PAUP*: Phylogenetic analysis using parsimony (*and other methods).
646   Version 4.10b.* Sunderland, Massachusetts: Sinauer Associates.
647 Taberlet P, Fumagalli L, Wust-Saucy AG, and Cosson JF. 1998. Comparative phylogeography
648   and postglacial colonization routes in Europe. *Molecular Ecology* 7:453-464.
649 Talma AS, and Vogel JC. 1992. Late Quaternary paleotemperatures derived from a speleothem
650   from Cango Caves, Cape Province, South Africa. *Quaternary Research* 37:203-213.
651 Tolley KA, Burger M, Turner AA, and Matthee CA. 2006. Biogeographic patterns and
652   phylogeography of dwarf chameleons (*Bradypodion*) in an African biodiversity hotspot.
653   *Molecular Ecology* 15:781-793.
654 Tzedakis P. 1993. Long-term tree populations in northwest Greece through multiple Quaternary
655   climatic cycles. *Nature* 364:437-440.
656 VanDerWal J, Falconi L, Januchowski S, Shoo L, and Storlie C. 2011. SDMTools: Species
657   distribution modelling tools: Tools for processing data associated with species
658   distribution modelling exercises. *R package version 11-13* http://CRAN.R-
659   project.org/package=SDMTools.
660 Vlok JHJ, Euston-Brown DIW, and Cowling RM. 2003. Acocks' Valley Bushveld 50 years on:
661   new perspectives on the delimitation, characterisation and origin of subtropical thicket
662   vegetation. *South African Journal of Botany* 69:27-51.
663 Wolfe KH, Li W-H, and Sharp PM. 1987. Rates of nucleotide substitution vary greatly among
664   plant mitochondrial, chloroplast and nuclear DNAs. *Proceedings of the National
665   Academy of Sciences, USA* 84:9054-9058.
666 Zachos J, Pagani M, Sloan L, Thomas E, and Billups K. 2001. Trends, rhythms, and aberrations
667   in global climate 65 Ma to present. *Science* 292:686-693.

668 **Data accessibility**

669 *DNA sequences*
670 Genbank rps16-trnQ intergenic spacer: KF180321-98
671 Genbank atpH-atpI intergenic spacer: KF180399-476
672 Genbank internal transcribe spacer 1, 5.8S ribosomal RNA gene, internal transcribed spacer 2:
673 KF443002-33, KY095713-57.
674 Genbank ncpGS gene: KY095758-827.
675
676
677

678

679 **Table 1** DNA summary statistics.

Genome	Catchment	n	h	s	π	MT$_R$	gsi$_T$
cpDNA	Overall	78	20	28	0.0022	0.4510 ***	
	Groot[1]	14	4	7	0.0015	-0.0517	0.290 (0.081) *
	Olifants[1]	26	8	11	0.0016	0.0638	0.508 (0.092) **
	Gamtoos	18	5	11	0.0008	-0.0250	**0.991** (0.067) **
	Sundays	18	7	8	0.0013	0.1077	**0.683** (0.140) **
	Gro+Oli[1]	40	10	12	0.0017	0.0885 *	**0.947** (0.155) **
	Oli+Gam					0.3930 ***	0.353 (0.097) **
	Sun+Gam	38	11	16	0.0017	0.3460 ***	**0.783** (0.119) **
ITS	Overall	75	68	54	0.0258	0.6234 ***	
	Groot[1]	14	14	20	0.0128	0.1889	0.289 (0.025) **
	Olifants[1]	24	22	28	0.0133	0.1078	0.474 (0.039) **
	Gamtoos	15	13	23	0.0127	0.3433 *	**0.996** (0.018) **
	Sundays	22	21	21	0.008	0.1619 *	**0.994** (0.034) **
	Gro+Oli[1]	38	34	29	0.0133	0.0829 *	**0.945** (0.007) **
	Oli+Gam					0.6448***	0.456 (0.083) **
	Sun+Gam	37	34	35	0.0174	0.5622 ***	**0.990** (0.082) **
ncpGS	Overall	69	13	11	0.0025	0.5717 ***	
	Groot[1]	13	3	4	0.0006	-0.1778	0.251 (0.098) **
	Olifants[1]	22	5	3	0.001	-0.1435 *	0.393 (0.132) **
	Gamtoos	14	5	5	0.0016	0.2020	0.187 (0.060) **
	Sundays	20	2	1	0.0001	-0.1236	0.546 (0.174) **
	Gro+Oli[1]	35	7	7	0.0009	0.0372	**0.842** (0.284) **
	Oli+Gam					0.5454 ***	0.000 (0.000) **
	Sun+Gam	34	6	6	0.0008	0.0875	**0.848** (0.267) **

680 n: number of samples; h: number of haplotypes; s: number of segregating sites; π: nucleotide
681 diversity; MT$_R$: Mantel tests r value (* p<0.05; ** p<0.01; *** p<0.001); gsi$_T$: mean and standard
682 deviation (in brackets) of the ensemble genealogical sorting index, and values above a
683 threshold of 0.6 are in bold.
684 [1] The Groot and Olifants are sub-catchments within the primary Gouritz catchment that is referred to in the text.

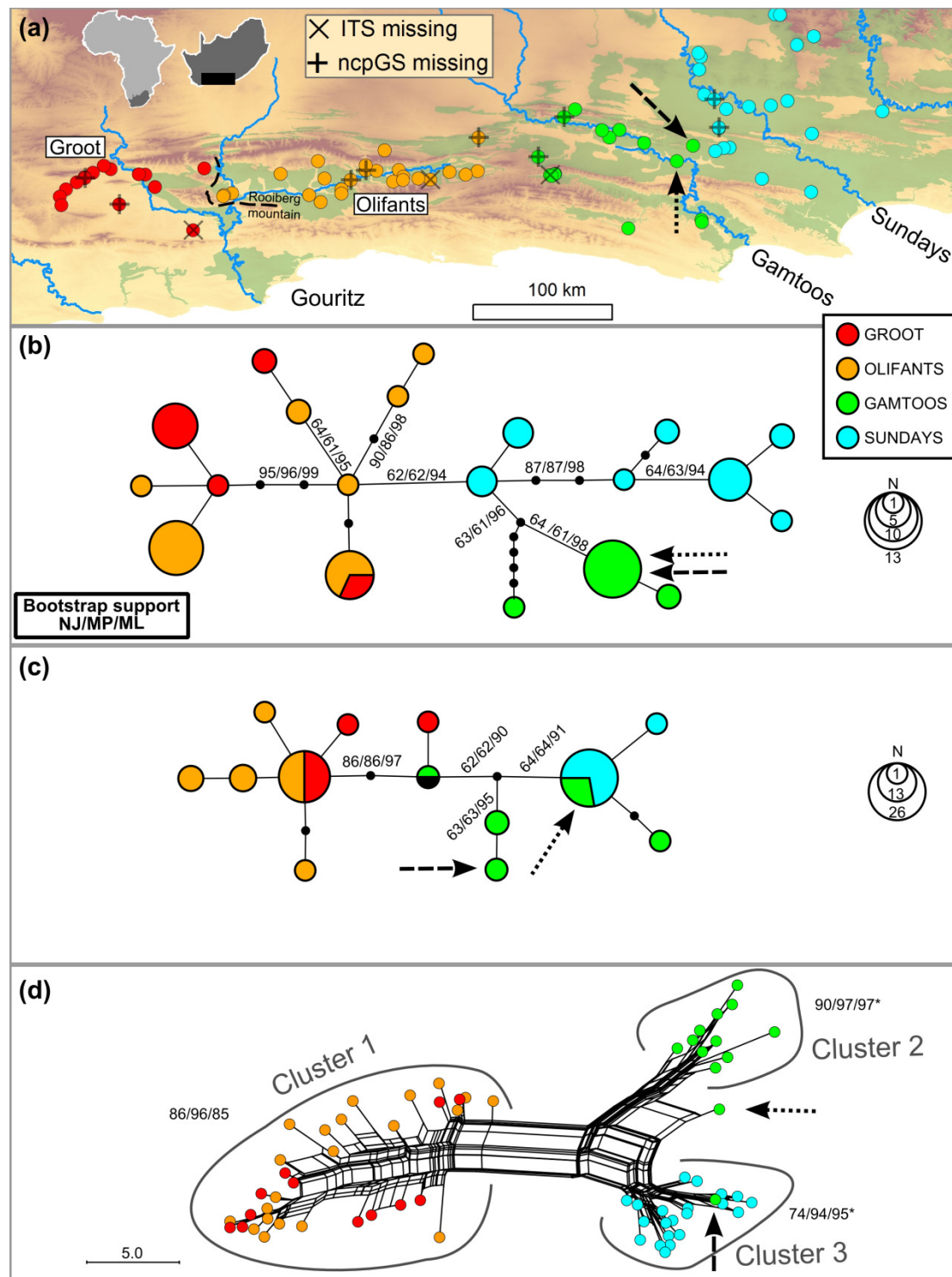



Figure 1 Distribution of *Nymania capensis* sampling localities within the Albany Subtropical

Thicket (light green) (a), and evolutionary relationships among accessions based on three

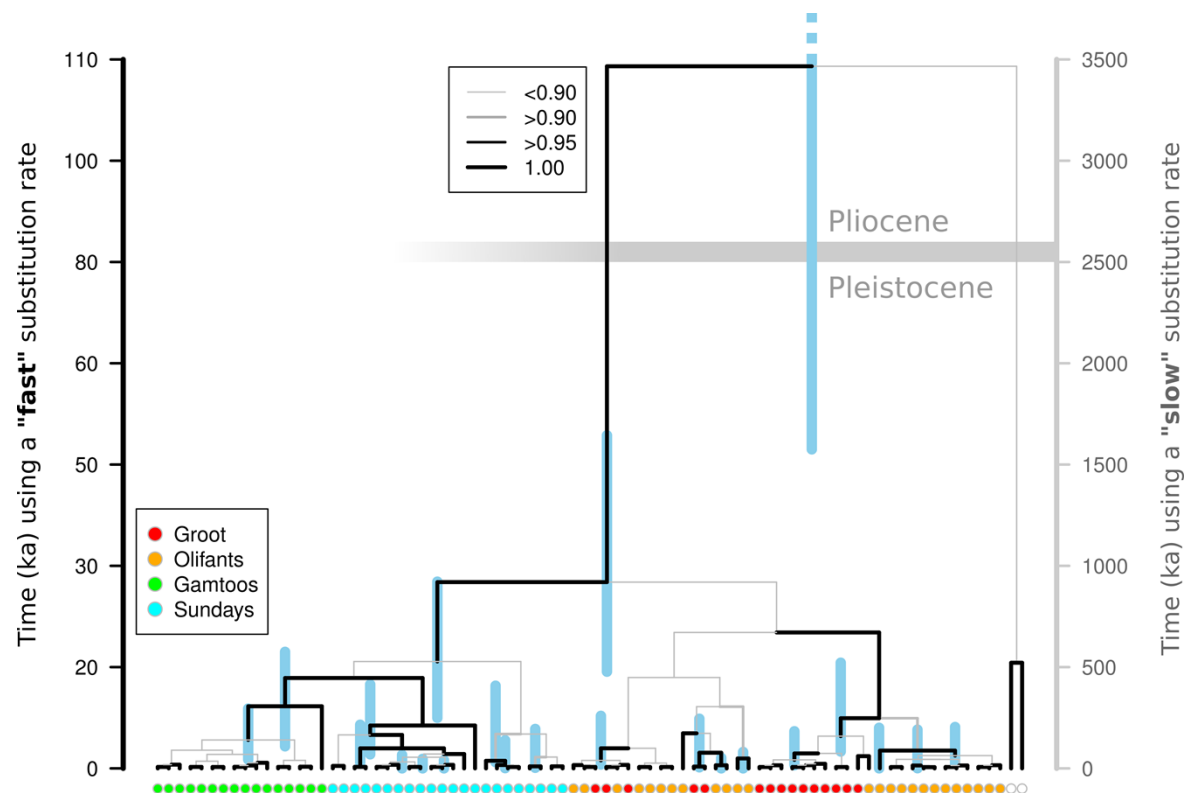
different molecular regions: cpDNA (b), low-copy nDNA (ncpGS) (c), and high-copy nDNA (ITS)

689   (d). Arrows represent samples AJP0537 and AJP0810, respectively, which are discussed in the

690   text. Outgroup samples were pruned from these networks.

691

692

693

694 **Figure 2** Molecular dating of *Nymania capensis* cpDNA sequences using majority rule Bayesian

695 chronogram generated in Bᴇᴀsᴛ. The posterior probabilities of branches are shown using a

696 combination of branch width and colour. Nodes are centred on the median time to the most

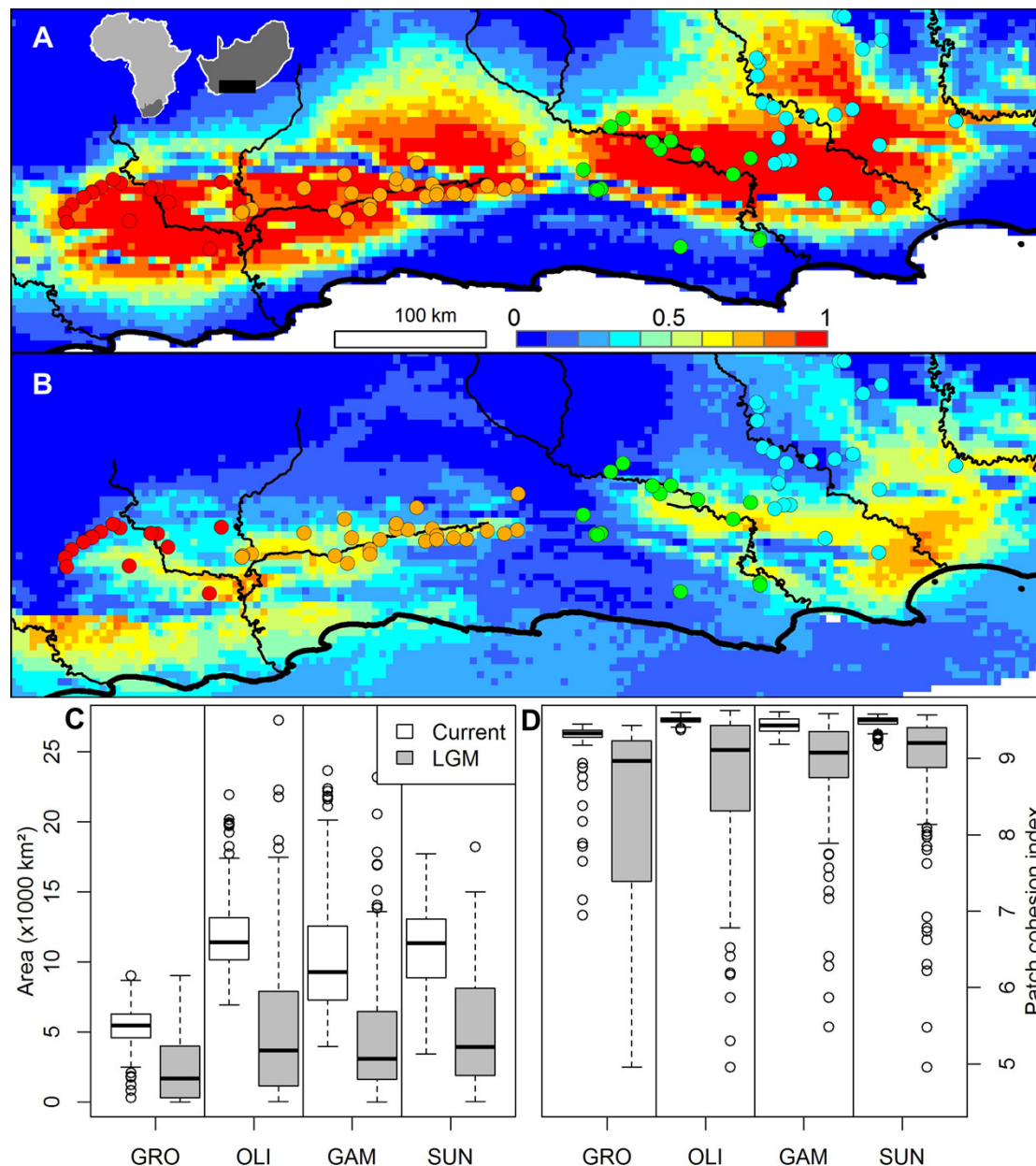
697 common recent ancestor (TMRCA) with blue shaded bars indicating the distribution of the 95%

698 HPD for each estimate. The timing of divergence estimates are show using a "slow" and "fast"

699 substitution rate (see text for details). The horizontal grey bar indicates the Pliocene-

700 Pleistocene boundary (~2.6 Ma) under the "slow" substitution rate.

701

703

704   **Figure 3** The ensemble of 216 species distribution models of *Nymania capensis* projected onto

705   (a) present and (b) Last Glacial Maximum (CCSM) climates, as well as the changes in climatically

706   suitable (c) area, and (d) patch cohesion index, within each basin. In the distribution maps, red

707   indicates greater certainty of presence, blue indicates greater certainty of absence, and green

708   (~0.5) indicates areas of greatest model uncertainty.