Predictors of taxonomic splitting and its role in primate conservation

Maria J.A. Creighton[a], Alice Q. Luo[a], Simon M. Reader[a] & Arne Ø. Mooers[b]

[a]Department of Biology, McGill University, Montreal, Quebec, Canada

[b]Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada

*Author for correspondence: maria.creighton@mail.mcgill.ca; ph: +1-514 398 7020; McGill University, Stewart Biology Building 1205 ave Docteur Penfield, Montreal, QC Canada, H3A 1B1 ORCID: https://orcid.org/0000-0003-1337-1946

WORD COUNT: 4969

DECLARATIONS OF INTEREST: None

1   **ABSTRACT:**

2   Species are the main unit used to measure biodiversity, but different preferred diagnostic criteria

3   can lead to very different delineations. For instance, named primate species have more than

4   doubled in number since 1982. Such increases have been attributed to a shift away from the

5   'biological species concept' (BSC) in favour of less inclusive species criteria. Critics of recent

6   changes in primate taxonomy have suggested taxonomic splitting may be biased toward certain

7   clades and have unfavourable consequences for conservation. Here, we explore predictors of

8   taxonomic splitting across primate taxa since the initial shift away from the BSC nearly 40 years

9   ago. We do not find evidence that diversification rate, the rate of lineage formation over

10   evolutionary time, is significantly linked to splitting, contrary to expectations if new species

11   concepts and taxonomic methods identify incipient species. We also do not find evidence that

12   research effort in fields where work has been suggested to motivate splitting is associated with

13   increases in species numbers among genera. To test the suggestion that splitting groups is likely

14   to increase their perceived risk of extinction, we test whether genera that have undergone more

15   splitting have also observed a greater increase in their proportion of threatened species since the

16   initial shift away from traditional taxonomic methods. We find no cohesive signal of taxonomic

17   splitting leading to higher threat probabilities across primate genera. Overall, this analysis sends

18   a positive message: threat statuses of primate species are not being overwhelmingly affected by

19   splitting. Regardless, we echo warnings that it is unwise for conservation to be reliant on

20   taxonomic stability. Species (however defined) are not independent from one another, thus,

21   monitoring and managing them as such may not meet the overarching goal of conserving

22   biodiversity.

23

24   **KEYWORDS:** species; phylogenetic species concept; primates; taxonomic splitting;

25   conservation

26

27   **INTRODUCTION:**

28   "Species" are an integral unit of biodiversity used across many sub-disciplines of biology, yet

29   how scientists define species has been subject to change. Notably, in the last 40 years, the

30   emergence of new methods for identifying diagnostic differences between populations (e.g.,

31   molecular phylogenetic methods) and changes in preferred species criteria have led to large

32   increases in species numbers across many clades (Agapow *et al.*, 2004).

33

34   Groves (2014) provides a brief overview of popular species definitions employed by taxonomists

35   through the late nineteenth to twentieth century. One notable phenomenon is the considerable

36   decrease in diagnosed species that occurred following the rapid adoption of the polytypic species

37   concept beginning in the 1890's. The polytypic species concept emphasizes that species should

38   be inclusive and that one should delineate taxa that resemble one another as subspecies (Groves,

39   2014). The popularization of the polytypic species concept was eventually accompanied by the

40   widespread adoption of the 'biological species concept' (BSC) beginning in the early 1960's

41   (Groves, 2014). The BSC defines species as populations/meta-populations that do not interbreed

42   with other populations/meta-populations under natural conditions (Mayr, 1963; Groves, 2014).

43   While this definition has been subject to revisions, the central premise of the BSC is that

44   reproductive barriers are key to diagnosing species (Groves, 2014). The BSC was widely

45   accepted and layered onto the pre-existing polytypic species concept, creating a period of relative

46   taxonomic stability for vertebrates from the 1960's to 1980's (Isaac *et al.*, 2004). However,

47 various criticisms of the BSC did emerge, the most notable being the practical difficulty of

48 diagnosing species under the BSC because of the need for information on reproductive barriers

49 (Donoghue, 1985; Tattersall, 2007). Such criticisms suggested the need for new approaches to

50 delineating species that offered higher diagnostic power.

51

52 In the last 40 years the 'phylogenetic' or 'diagnostic' species concept (PSC) (Cracraft, 1983) has

53 been widely popularized in vertebrate taxonomy due to its diagnosable advantages over previous

54 species definitions (Isaac *et al.,* 2004; Cotton *et al.,* 2016). Under the PSC, a species is diagnosed

55 as the smallest population or meta-population that is distinct in heritable differences from other

56 populations or meta-populations (Cracraft, 1983; Groves & Grubb, 2011; Groves, 2014).

57 According to its proponents, the PSC's emphasis on diagnosable evidence gives it an advantage

58 over other species concepts because it allows users to rely on a range of data types, including

59 newly available molecular markers, to make distinctions (see, e.g., the variety of data types used

60 to describe the newest species of ape, *Pongo tapanuliensis*; Nater *et al.,* 2017). Together, new

61 species concepts and methodological advancements have characterized a shift away from

62 biological species and toward species that are delineated based on distinctive, diagnosable

63 differences.

64

65 Although using distinctive differences to delineate species offers advantages, this approach has

66 also been subject to criticism, notably, for its tendency to split species into a range of less-

67 inclusive units (Agapow *et al.*, 2004; Zachos *et al.*, 2013; Zachos & Lovari, 2013). Many

68 populations previously recognized as subspecies or morphological variants have been elevated to

69 the full species status, resulting in a large increase in the number of listed species. For instance,

70   181 species of primates were listed by Honacki *et al.* (1982), one year prior to Cracraft's (1983)

71   proposal of the PSC. Today the IUCN (International Union for Conservation of Nature) lists over

72   500 distinct primate species (Estrada *et al.*, 2017). Some families (e.g., Cheirogaleidae and

73   Indriidae) have more than tripled in species richness over the past ~40 years (see Figure 1).

74   While some new species have been added as a result of new field discoveries, a majority are

75   populations which were previously identified at lower taxonomic levels that have now been

76   elevated to species status following a shift in the accepted approach to diagnosing primate

77   species (Tattersall, 2007). Some attribute this taxonomic increase to the popularity of the PSC

78   (see, e.g., Isaac *et al.*, 2004), while others credit increased exploration and the development of

79   new techniques for evaluating diagnosability (see, e.g., Köhler *et al.,* 2005; Harris & Froufe,

80   2005; Padial & De la Riva, 2006; Sangster, 2009). The trend of increasing species numbers by

81   raising taxonomic statuses has been referred to as 'taxonomic inflation' by some (Isaac *et al.*,

82   2004; Rylands & Mittermeier, 2014) and has been criticized for being non-random, and biased

83   toward certain clades (Isaac *et al.,* 2004; Agapow *et al.,* 2004; Zachos *et al.,* 2013; Zachos &

84   Lovari, 2013; Rylands & Mittermeier, 2014).

85

86

87

88

89
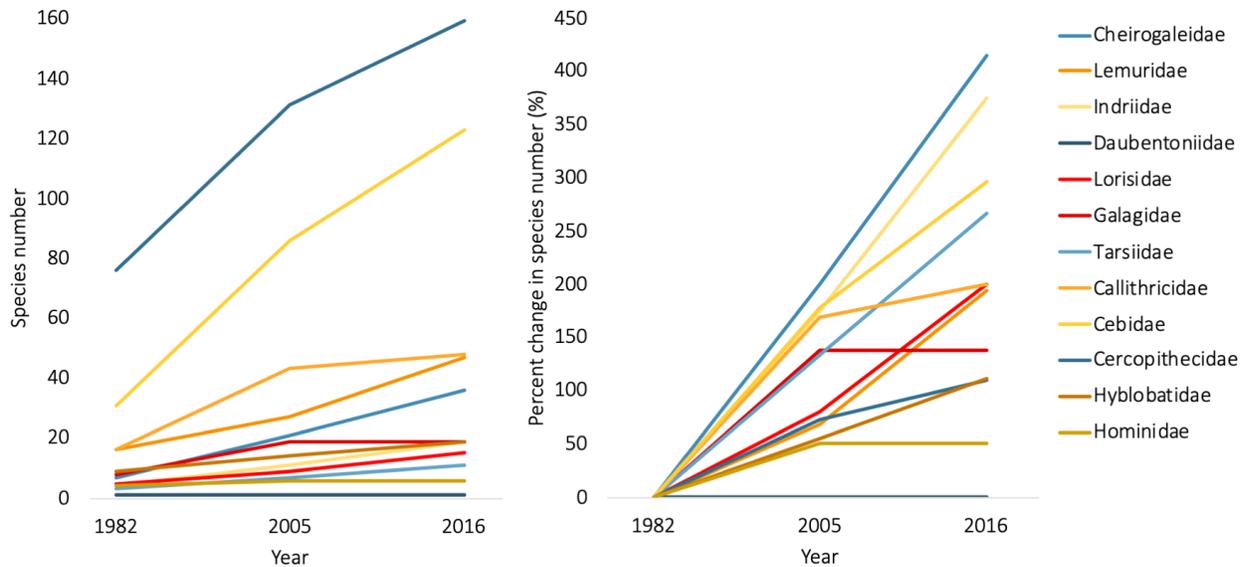
Figure 1: Species numbers and percentage change in species numbers for historic primate families recorded at three time points by Honacki *et al.* (1982), Wilson & Reeder (2005) and the IUCN species list from 2016 (data documented in Estrada *et al.*, 2017).

93

Disproportional splitting among taxonomic groups could have several causes. Variation in the number of new species described across taxa could be driven by variation in the rate at which lineages evolve, such that new species descriptions are tracking cryptic diversity or incipient species formation among rapidly evolving taxa. In this case, new species listings might point to situations where there is a discordance between patterns of genetic change and the evolution of gross morphological changes used by traditional taxonomists. This could be due to ecology – if some lineages are diversifying along ecological axes not captured in traditional taxonomic approaches – or due to demographics – if some lineages have ecologies and/or histories that lead to faster local genetic coalescent times and so diagnosability. Under either of these scenarios, variation in which clades observe the greatest amount of splitting under new taxonomic approaches would simply reflect underlying biological reality. However, it is also possible that

105   splitting is driven by other factors that may be prone to bias. Critics have argued that many

106   increases in species numbers are artificial, reflecting major shortcomings of new species criteria

107   and a reliance on insufficient data (Zachos *et al.*, 2013; Zachos & Lovari, 2013). Zachos *et al.*

108   (2013) provides evidence suggesting unwarranted splitting in select cases, advocating that

109   splitting has been taken to a "molecular extreme". Many different types of molecular data are

110   used to justify splitting (e.g., genetic data from mitochondrial DNA barcoding) and, by this logic,

111   groups may continue to be split as an increasing amount of molecular data become available for

112   them. If true, this would lead to the continual identification of increasingly exclusive diagnostic

113   features such that we are likely to find new species the more we look for them.

114

115   It is also possible that conservation interest in particular groups could motivate splitting. Limited

116   funding for conservation research is increasingly focused on 'biodiversity hotspots' and it has

117   been suggested that researchers could have a vested interest in declaring taxa in these regions as

118   endemic species (Karl & Bowen, 1999; Isaac *et al.*, 2004). There is evidence from some

119   charismatic groups that taxa receive more conservation attention and funding when comprised of

120   multiple, small, and taxonomically distinct populations (e.g., African apes; Stanford, 2001;

121   Oates, 2006; Gippoliti & Amori, 2007). Species lists are often used to determine which groups

122   should receive conservation attention (Mace, 2004) and so changing the way we define species

123   may also change which groups receive action.

124

125   Along with being potentially biased toward certain taxonomic groups and related to the point

126   made above, taxonomic splitting has also been suggested to result in individually more imperiled

127   populations which could distort conservation agendas (Agapow *et al.*, 2004; Isaac *et al.*, 2004;

128 Morrison *et al.*, 2009; Zachos, 2015; Robuchon *et al.*, 2019). One criterion used by the IUCN to

129 classify species as imperiled is population size: species may be designated "Vulnerable" if there

130 are fewer than 1000 mature individuals found in the wild and "Endangered" if there are fewer

131 than 250 (Agapow *et al.*, 2004; Frankham *et al.*, 2012). Thus, splitting one species into several

132 new species may result in one or more receiving a (more) imperiled status (Agapow *et al.*, 2004;

133 Isaac *et al.*, 2007). This could lead to seemingly rare but poorly defined species being prioritized

134 over well-defined and perhaps biologically more distinctive species (Pillon & Chase, 2007).

135 Recent evidence suggests species splitting is not a driver of threat status for birds (Simkins *et al.*,

136 2020), however, it is unknown whether these findings are generalizable across other taxonomic

137 groups (Garnett & Thomson, 2020).

138

139 Here, we set out to better understand the causes and consequences of taxonomic splitting in

140 primates. We test (i) predictors of taxonomic splitting, that is, whether recent taxonomic

141 increases are associated with the amount of research being done in fields suggested to motivate

142 splitting or with a lineage's underlying diversification rate, and (ii) impacts of taxonomic

143 splitting, that is, whether rates of splitting dictate which groups are most imperiled. Because

144 some newly described primate species have previously been described as subspecies or

145 subpopulations of more than one different biological species prior to being assigned to the full

146 species rank (which would make quantifying rates of splitting at the species-level difficult), we

147 ask these questions at the genus level.

148

149 To explore our first question regarding predictors of taxonomic splitting, we consider whether

150 biological factors or measures of human-induced bias explain increases in species numbers. To

151   test potential human-induced biases in splitting, we consider broad estimates of research effort

152   for each taxon and predict that more research done on a given taxon may be associated with

153   more splitting. We examine research effort in two fields: molecular genetics (since molecular

154   work could cause species to be split continuously as finer molecular distinctions are made) and

155   conservation (since splitting has been suggested to be motivated by conservation interests). To

156   explore possible biological explanations for trends in taxonomic splitting across clades, we test

157   whether recent taxonomic increases are explained by diversification rate. Clades with high recent

158   diversification rates are expected to contain more incipient or cryptic species than lineages

159   diversifying at a lower rate since they will contain more closely related lineages that resemble

160   one another. Therefore, if new approaches help to identify incipient or cryptic species,

161   diversification rate should be positively correlated with splitting.

162

163   To explore our second question linking splitting and risk, we test whether increases in species

164   numbers are associated with a change in the number of threatened species listed in genera

165   through time. Ideally we would look at changes in a weighted measure of threat score that

166   differentiates between threat categories of varying severity (e.g., using the Red List Index (RLI);

167   Butchart *et al.*, 2007; Bubb *et al.*, 2009); however, criteria for inclusion in Red List threat

168   categories have changed considerably over time, meaning weights assigned to categories for RLI

169   calculations do not match up with categories used in the past (see, e.g., statuses in IUCN

170   Conservation Monitoring Centre, 1986). Because of these changes in Red List criteria, we

171   instead ask whether clades with species that have been split more frequently have observed a

172   greater increase in their proportion of threatened species (defined below) through time in

173   comparison to those which have been split less frequently. We predict that if splitting is driving

174 an increase in threatened species, we should observe a positive association between taxonomic

175 increases caused by splitting and change in the proportion of threatened primate species.

176

177 **MATERIALS AND METHODS:**

178 ***Data***

179 To measure the number of primate species that were described before the introduction of the

180 PSC and new molecular techniques, we used the last pre-PSC taxonomy, that of Honacki *et al.*

181 (1982). This taxonomy contains 181 species and is considered a reliable estimate of the number

182 of species thought to exist during the popularity of the BSC (see Rylands & Mittermeier, 2014).

183 We then recorded if each species in this taxonomy was historically considered to be threatened

184 by referencing the most complete IUCN Red List published around the same time (IUCN

185 Conservation Monitoring Centre, 1986). Honacki *et al.* (1982) was contrasted with the IUCN

186 taxonomy from 2016 and attendant data documented in Estrada *et al.* (2017), which lists 503

187 species. For each of these 503 species we noted their taxonomic placement (genus and family),

188 whether or not they were considered threatened (VU=Vulnerable, EN=Endangered, or

189 CR=Critically Endangered), and their biogeographic region. We note that while IUCN

190 assessments have some shortcomings (see, e.g., Rueda-Cediel *et al.,* 2018), the IUCN provides

191 the largest global data on threat status and is influential in determining how most species are

192 managed. For each species described by the IUCN in 2016 that was not listed in Honacki *et al.*

193 (1982), we scored whether the species was a "*de novo*" species description (Burgin *et al.*, 2018):

194 there were 40 such cases where a new species had not been previously formally identified as a

195 subspecies or subpopulation of another species prior to splitting. These cases represented new

196   species descriptions where it was unclear whether a new species was a result of taxonomic

197   splitting or the discovery of an entirely new population.

198

199   We compiled all the species listed by the IUCN in 2016 into 12 families and 50 genera found in

200   Honacki *et al.* (1982). Family name "Callimiconidae" in Honacki *et al.* (1982) was not used as

201   this taxon has since been recognized as a genus of the larger family "Callitrichidae" (Wilson &

202   Reeder, 2005). We removed *Rungwecebus kipunji* from the IUCN species list from 2016 as it

203   represents a newly discovered genus that does not collapse into any of the genera provided by

204   Honacki *et al.* (1982).

205

206   Research effort in the fields of molecular genetics and conservation was estimated for each genus

207   through a literature review of publications in the Web of Science Core Collection published

208   between 1983 (when a general trend toward taxonomic splitting first began) and 2016. All search

209   terms for these literature reviews and further details on related methodology are documented in

210   the supplementary materials (Tables S1 and S2). In this study we used the genera listed in

211   Honacki *et al.* (1982) (n=50), many of which have since been further separated into multiple

212   genera. Thus, when appropriate we included new genus names in addition to those listed by

213   Honacki *et al.* (1982) in the literature searches (see Table S3). In total, our final sample was 688

214   publications on molecular genetics and 2222 on conservation.

215

216   Diversification rate was estimated with the method-of-moments approach described in Magallon

217   & Sanderson (2001) (i.e., ln(taxa richness)/stem age). Diversification rate estimates generated

218   using this method often rely on species numbers as their estimate of taxa richness, meaning that

219   diversification rate estimates are inherently biased by the splitting phenomenon we are studying

220   (i.e., frequently split genera will receive disproportionally high diversification rates). Therefore,

221   richness scores for our diversification rate calculations were determined as the counts of the

222   well-resolved "lineages" described in Creighton *et al.* (2021). These lineages were determined by

223   creating a time cut-off in the 10kTrees consensus primate phylogeny (Arnold *et al.*, 2010) in an

224   attempt to eliminate very young newly described species and obtain a consistent (unbiased)

225   estimate of diversity across clades. These lineages were assigned to each of the 50 genera

226   described in Honacki *et al.* (1982). Diversification rate was then estimated by taking the natural

227   log lineage richness for each genus and dividing by the stem age of that genus (Magallon &

228   Sanderson, 2001). Stem ages for each genus were extracted from the 10kTrees consensus

229   phylogeny (version 3) (Arnold *et al.*, 2010) trimmed to contain a single branch representing each

230   genus. During this process, there were several instances where genera described in Honacki *et al.*

231   (1982) were non-monophyletic within the more recent primate phylogeny we used (Arnold *et al.*,

232   2010), making it unclear how to assign a divergence date for these clades. We therefore removed

233   eight genera from the analyses where diversification rate was a variable of interest: *Presbytis*,

234   *Lemur*, *Galago*, *Cebuella*, *Cercocebus*, *Cercopithecus*, *Papio*, and *Pygathrix*.

235

Analysis

237   To test our questions about the predictors of species-splitting and its consequences for

238   conservation, we fit a series of linear effects and mixed effects models. We note that the response

239   variables used in these analyses (i.e., measures of taxonomic increase and extinction risk) are

240   likely to be phylogenetically clustered, and phylogenetic models could be used to account for

241   this influence of phylogeny; however, many genera listed by Honacki *et al.* (1982) are non-

242   monophyletic, making it unclear how to designate them a single branch in modern phylogenies

243   (see discussion above on diversification rate). Importantly, after accounting for regional

244   differences, family contributed little to no variance in any of our models, indicating that

245   phylogenetic relationships at that level were not confounding our results. We provide further

246   discussion on model choice in the supplementary materials.

247

248   Data were analysed using R version 4.0.5 (R Core Team, 2021).

249

250   *Predictors of Taxonomic Splitting*

251   To determine if measures of potential human bias (i.e., research effort) or diversification rate

252   explained discrepancies in splitting across taxa, we tested to see if these variables were

253   significantly associated with the number of species added to primate genera since 1982 while

254   controlling for the original number of described species (i.e., biological species) and region. We

255   fit three generalized linear mixed effects models with Poisson distributions using the lme4

256   package in R (Bates *et al.*, 2014), and obtained p-values using the lmerTest package (Kuznetsova

257   *et al.*, 2017). In these models the response variable was the number of species assigned to a given

258   genus by the IUCN in 2016 that had not been previously described by Honacki *et al.* (1982).

259   Each model had either conservation research effort, molecular genetics research effort or

260   diversification rate included as a fixed effect, as well as biogeographical region and the number

261   of species in the genus per Honacki *et al.* (1982) to control for their effects on splitting. The

262   natural logarithm (ln) of the number of species listed for each genus in Honacki *et al.* (1982) was

263   included as both a linear and a quadratic term following inspection of raw plots and plots against

264   scaled residuals from the simulation output. A square-root transformation was used on molecular

265   genetics research effort to decrease the impact of outliers on model fit. To assist with model

266   stability and convergence, we scaled all continuous variables in the model to have a mean of zero

267   and standard deviation of one (Becker *et al.*, 1988). Mainland Africa and Asia (hereafter

268   mainland Africa + Asia) were grouped together and served as the baseline region in our models

269   based on previous studies that have shown that the taxonomy of primates from these regions has

270   been relatively stable compared to Madagascar and the Neotropics (e.g., Isaac & Purvis, 2004;

271   Isaac *et al.*, 2004; Tattersall, 2007). We also chose to group Asia and Africa together because

272   one genus (*Macaca*) is found in both regions. Genus ID was included as a random effect in these

273   models to correct for overdispersion. Family (nested within region) was originally included as a

274   random effect but contributed very little to model fit and created issues with convergence due to

275   overfitting, and thus was dropped from the final models. We tested potential interaction terms

276   with all variables and region to test for regional effects, but none were significant and so these

277   terms were also dropped from the final models. Before running models with other predictors

278   included, we also ran a model including only the linear and quadratic terms for the number of

279   species listed for each genus in Honacki *et al.* (1982) as predictors to assess their association

280   with taxonomic increase in the simplest model. We ran all models a second time after removing

281   the 40 *de novo* species from our response (Tables S4, S5 and S6). We checked model

282   assumptions and fit by plotting residuals versus the fitted values and versus each covariate in the

283   model. Residual plots and analyses done with the Diagnostics for Hierarchical Regression

284   Models (DHARMa) R package (Hartig, 2017) indicated acceptable model fits.

285

286   *Taxonomic Splitting and Threat Score*

287   To test whether taxonomic splitting over time is associated with a change in the proportion of

288   threatened species within genera, we conducted a two-step (hierarchical) analysis on species'

289   threat probability between 1982 to 2016. We first fit a generalized linear mixed effects model

290   with a binomial distribution for the number of threatened and non-threatened species within a

291   genus using the lme4 package in R (Bates *et al.*, 2014). In this first model, the response

292   estimated the proportional counts of species at risk within genera (equivalent to the per-species

293   probability of threat) in the periods of 1982 and 2016. Predictors for this model included fixed

294   effects for the time period (the baseline of 1982 and the change to 2016), the region that

295   encompasses each genus' natural distribution (Madagascar, Neotropics, mainland Africa + Asia),

296   and an interaction between region and time period to account for geographic differences in threat

297   probabilities through time. Genus identity was included as a random intercept, to account for

298   repeated measures in 1982 and 2016, and as a random slope with time period to account for

299   differences in changes to threat probabilities among genera. As discussed above, we included the

300   taxonomic rank of family (nested within region) as a random effect but found that the variance in

301   threat probabilities among families was minimal and that including this term also created

302   convergence issues; given this, we subsequently omitted family from our models. Residual plots

303   and analyses with the DHARMa R package (Hartig, 2017) indicated acceptable model fits for the

304   final model. From this first model describing genus level changes in threat probabilities, we then

305   used the REextract function implemented in the merTools package in R (Knowles & Frederick,

306   2020) to extract the genus level random slopes for time period and their associated standard

307   errors, giving us an estimate of the varying effect of change in threat probability among genera

308   (while conditioning on regional trends) between 1982 and 2016. We then fit a linear model

309   where varying effect of change in threat probability from 1982 to 2016 was the response, and the

310   proportional change in species within each genus (i.e., the number of new species in the IUCN

311   species list from 2016 / original number of species in Honacki *et al.* (1982) – our measure of

312   taxonomic increase) – was the predictor. We weighted each estimate of genus level change in

313   threat probability by its standard error (w = 1/SE) to propagate the error of model estimated

314   random effects. We ran this model a second time after removing *de novo* species descriptions

315   from our response (Table S7). Raw data used for these tests are visualized in Figure 2.

316

317   Plots of the raw data indicated that patterns in taxonomic increases, the proportion of species at

318   risk in primate genera today, and the changes in the proportion of threatened species in primate

319   genera between 1982 and 2016 appeared to show regional differences (Figures S1 and S2). We

320   thus ran a subsequent set of models to test for regional variation in the effect of splitting. In this

321   analysis, we removed region from the first generalized linear mixed effects model and

322   subsequently included region interacting with taxonomic increase as a predictor of the varying

323   effect of change in genus threat probability in the second-order model. However, interaction

324   effects were not significant in this model and thus we only considered the results of the first set

325   of models reported above.

326

327   **<u>RESULTS:</u>**

328   *Predictors of Taxonomic Splitting*

329   None of our measures – conservation research effort (β = -0.220; p = 0.180; Table S4), molecular

330   genetics research effort (β= -0.188; p= 0.212; Table S5) or diversification rate (β= 0.132; p=

331   0.477; Table S6) – were significantly associated with increases in species numbers across

332   primate genera. Removing *de novo* species did not impact this pattern of results (Tables S4, S5

333   and S6). Notably, in addition to sharing a significant linear relationship with taxonomic increase

334   as expected (Tables S4, S5 and S6), the quadratic term added for the original number of species

335   in 1982 was significant in the model with diversification rate, indicating a downwardly concave

336   association with taxonomic increase regardless of whether de novo species were included (Table

337   S6); this term had a p-value less than 0.1 in models that included conservation or molecular

338   genetics research effort as predictors (Tables S4 and S5), indicating that the positive increase in

339   splitting by initial genus size decelerates as genera become very large, even after accounting for

340   other variables. We note that this quadratic term was significant without other predictors (i.e.,

341   research effort or diversification rate) present in the model, both with *de novo* species included in

342   the response (β= -2.268; p= 0.011) and without *de novo* species included in the response (β= -

343   2.086; p= 0.013). In all three models without *de novo* species, taxonomic increases were

344   significantly higher in the Neotropics compared to Africa and Asia (Tables S4 to S6).

345

346   *Taxonomic Splitting and Threat Score*

347   Visual inspection of the raw data on taxonomic increases and proportion of species at risk for all

348   genera revealed no clear indication that highly split genera have observed a disproportional

349   increase in perceived extinction risk (Figures 2 and S3). Results from our linear model confirmed

350   that taxonomic increases did not predict a change in threat score through time (β= 0.001; p=

351   0.975; Table S7; Figure 3). Removing *de novo* species did not impact this pattern of results

352   (Table S7; Figure S4).

353

354
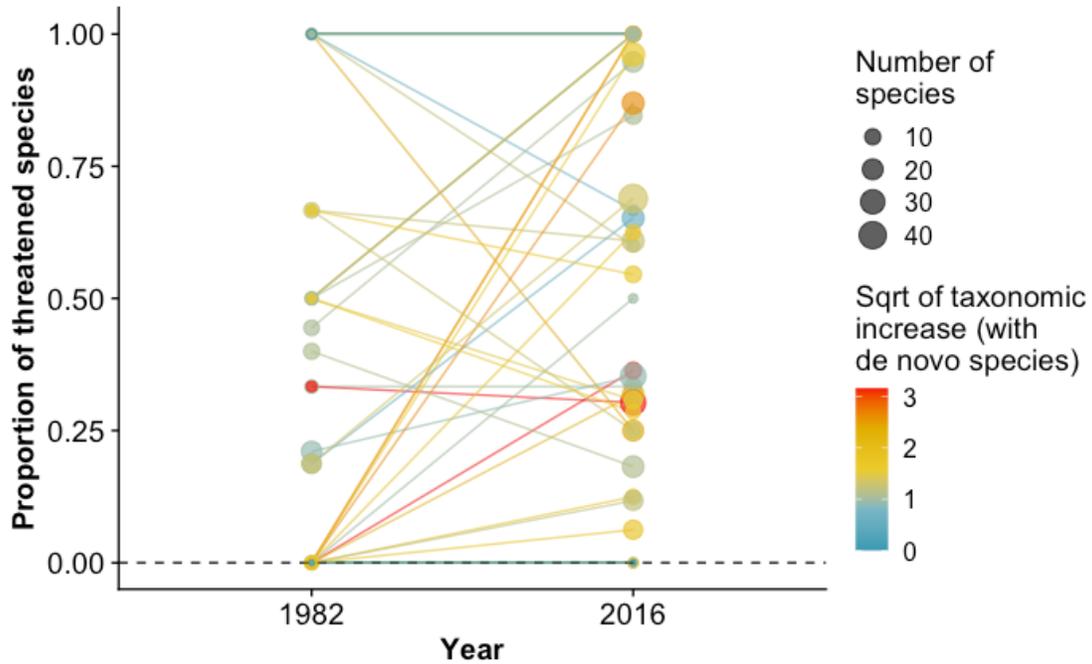
355

356   Figure 2: Scatterplots with trendlines showing the change in the proportion of species identified
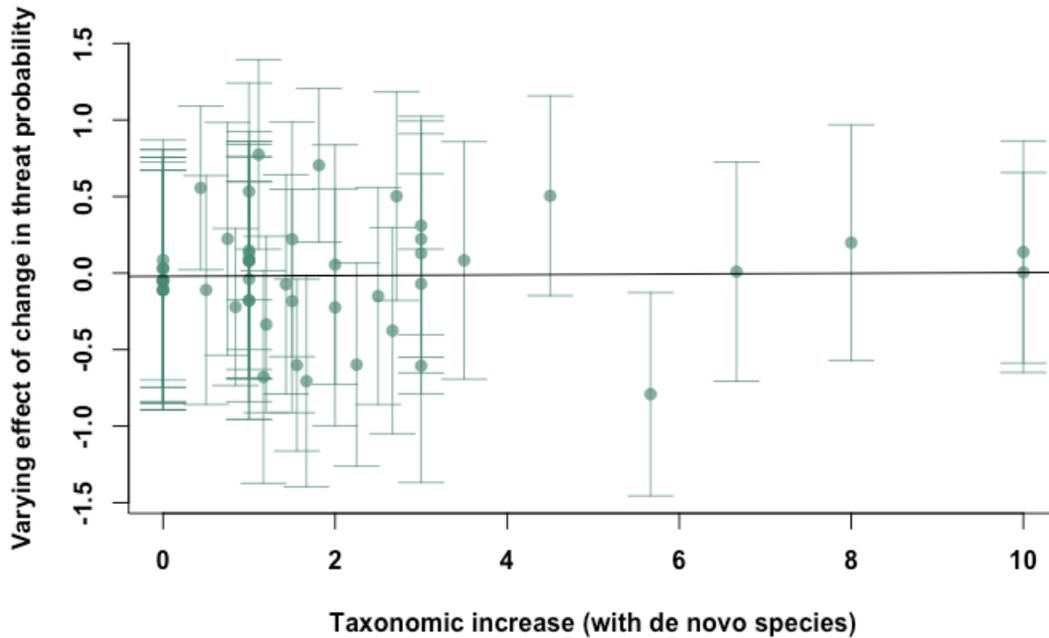
357   as being threatened in primate genera in 1982 and 2016 painted by the square root of taxonomic

358   increase (including *de novo* species). Total number of species in each genus is indicated by point

359   size.

360

361

Figure 3: Taxonomic increase including *de novo* species descriptions versus varying effect of change in threat probability (i.e., our measure of genus level change in threat probability between 1982 and 2016; see Methods for an explanation of how this was estimated) ($\beta= 0.001$; $p= 0.975$; Table S7). Vertical bars indicate standard errors on varying effect of change in threat probability values.

DISCUSSION:

Our results support neither a strong biological nor a strong interest-driven mechanism for splitting across primate genera. If the naming of new species under new taxonomic approaches captures incipient speciation, we predicted that diversification rate should predict taxonomic increases. However, diversification rate was not a significant predictor of taxonomic increases in our models. We also tested the prediction that splitting is motivated by conservation interest (Karl & Bowen, 1999; Isaac *et al.*, 2004) or by increasing molecular research within certain taxa (Zachos *et al.*, 2013; Zachos & Lovari, 2013). However, we found no evidence that research

377　effort in either of these areas was associated with the amount of taxonomic splitting observed

378　across genera. We did find that initial genus size predicts increased splitting, but this effect

379　decelerates as genera approach the largest sizes. This could indicate that splitting is being driven

380　by variables not captured by our models. For example, genera that already contained many

381　species in 1982 may have been more likely to have closely related populations described as

382　separate species under the BSC (e.g., if we knew more about their hybrid statuses compared to

383　other clades). In this case, we may have already discovered much diversity in these large clades

384　prior to applications of the PSC and new molecular techniques. Future studies could aim to

385　elucidate the origins of this decelerating association, however, our results indicate it is not

386　explained by the rate at which lineages diverge (which should capture the presence of cryptic

387　species) or research effort in the two fields studied here.

388

389　While inspection of the raw data suggested that genus level measures of splitting, threat

390　probability, and change in threat probability all varied by region, we found that there was no

391　cohesive signal of taxonomic splitting leading to higher threat probability across these genera

392　and regions. Therefore, while there are idiosyncratic regional differences in both splitting and

393　threat probability there was no evidence for causal links between the two. We do note that while

394　the act of splitting itself does not seem to be having an overwhelming effect on threat status, it is

395　still possible splitting may have uncaptured negative consequences downstream in conservation

396　management (e.g., loss of genetic variation in captive breeding programs; Zachos, 2013).

397　However, our analysis sends a positive message that splitting is not currently a significant

398　determinant of the relative conservation priority of primate clades, consistent with recent

399　findings in birds (Simkins *et al.*, 2020).

400

401 Importantly, we note that our analyses have some limitations. First, quantifying research effort of

402 any kind is difficult due to the abundance of work published in different media. As such, our

403 estimates of research effort in conservation and molecular genetics may overlook some types of

404 research. More work could be done to determine if additional estimates of research bias can

405 explain increases in species numbers. For instance, cumulative funding estimates from various

406 sources (e.g., the IUCN, non-governmental organizations and regional governments) per taxon

407 could provide an additional or alternative measure of conservation interest. Second, it was

408 necessary to remove eight non-monophyletic genera (some of which have undergone substantial

409 splitting) from analyses that included diversification rate, leading to a considerable reduction in

410 statistical power for those analyses. Third, due to changes in Red List criteria, species status, and

411 lack of information about which species had been assessed in the 1980s, we were not able to

412 consider differences in threat status severity (e.g., Vulnerable versus Endangered) when

413 considering associations between taxonomic splitting and changes in threat score through time.

414 Future studies could look at the association between splitting and changes in weighted measures

415 of threat score (e.g., RLI) over time using a more recent starting point (i.e., after 1982) where

416 Red List criteria become stable and there is available information on which non-threatened

417 species have been evaluated.

418

419 Overall, we do not find support for biological processes or research bias driving taxonomic

420 splitting across primate genera. We also find no cohesive signal of splitting leading to higher

421 threat probabilities. Generally speaking, relying on species as the central unit of conservation and

422 primary object of biological study behooves taxonomy to remain stable, while changing ideas

423 about the concept of species makes taxonomy inherently unstable (Mayr, 1996). We suggest that

424 areas of research requiring consistent estimates of diversity (e.g., conservation, macroecology, or

425 evolutionary biology) may benefit from (i) weighing evolutionary distinctiveness when

426 determining how species are listed/treated if attempting to capture true biological diversity (see,

427 e.g., Redding & Mooers, 2010; Redding *et al.*, 2015); or, for applied conservation specifically

428 (ii) shifting more resources toward regional management efforts that are less likely to be

429 influenced by changing species designations. It is well-known that closely related species are

430 more similar to one another than they are to more distantly related taxa. Thus, treating all species

431 independently and of equal weights in conservation listing may not lead to desired outcomes

432 (Redding & Mooers, 2010). As of 2016, approximately 60 percent of all primate species were

433 threatened with extinction according to the IUCN Red List of Threatened Species (Estrada *et al.*,

434 2017), making it imperative that conservation efforts are spent wisely to ensure optimal

435 conservation of primate biodiversity writ large.

436

445   Paul Sims and Dan Greenberg for statistical advice. We also thank Sandro Lovari, Frank Zachos,

446   and one anonymous reviewer for their useful comments on earlier drafts.

447

448   **<u>DECLARATIONS OF INTEREST:</u>** None.

449

450   Data   and   code   accompanying   the   manuscript   are   available   on   Zenodo:

451   https://doi.org/10.5281/zenodo.6614874

452

453

454   **<u>REFERENCES:</u>**

455   Agapow, P.M., Bininda-Emonds, O.R., Crandall, K.A., Gittleman, J.L., Mace, G.M., Marshall,

456   J.C. & Purvis, A. (2004). The impact of species concept on biodiversity studies. *Q. Rev.*

457   *Biol.* **79**, 161–179.

458   Arnold, C., Matthews, L. J. & Nunn, C.L. (2010). The 10kTrees website: a new online resource

459   for primate phylogeny. *Evol. Anthropol.* **19**, 114–118.

460   Bates, D., Maechler, M., Bolker, B. & Walker, S. (2014). *lme4: Linear mixed-effects models*

461   *using   Eigen   and   S4*.   R   package   version   1.1.21.   https://cran.r-

462   project.org/web/packages/lme4.

463   Becker, R.A., Chambers, J.M. & Wilks, A.R. (1988). *The new S language: A programming*

464   *environment for data analysis and graphics*. Wadsworth & Brooks/Cole, Pacific Grove.

465   Bubb, P.J., Butchart, S.H.M., Collen, B., Dublin, H., Kapos, V., Pollock, C., Stuart, S.N. & Vié,

466   J.–C. (2009). *IUCN Red List Index – guidance for national and regional use*. IUCN,

467   Gland.

468   Burgin, C.J., Colella, J.P., Kahn, P.L. & Upham, N.S. (2018). How many species of mammals

469   are there? *J. Mammal.* **99**, 1–14.

470   Butchart, S.H., Akçakaya, H.R., Chanson, J., Baillie, J.E., Collen, B., Quader, S., Turner, W.R.,

471   Amin, R., Stuart, S.N. & Hilton-Taylor, C. (2007). Improvements to the Red List Index.

472   *PLoS One* **2**, e140.

473   Cotton, A., Clark, F., Boubli, J.P. & Schwitzer, C. (2016). IUCN Red List of threatened primate

474   species. In S.A. Wich & A.J. Marshall (Eds.), *An introduction to primate conservation* (p.

475   31–38). Oxford University Press, Oxford.

476   Cracraft, J. (1983). Species concepts and speciation analysis. In R.F. Johnson (Ed.), *Current

477   ornithology* (p. 159–187). Plenum Press, New York.

478   Creighton, M.J.A., Greenberg, D.A., Reader, S.M. & Mooers, A.Ø. (2021). The role of

479   behavioural flexibility in primate diversification. *Anim. Behav.* **180**, 269–290

480   Donoghue, M.J. (1985). A critique of the biological species concept and recommendations for a

481   phylogenetic alternative. *Bryologist* **88**, 172–181.

482   Estrada, A., Garber, P.A., Rylands, A.B., Roos, C., Fernandez-Duque, E., Di Fiore, A., Nekaris,

483   K.A., Nijman, V., Heymann, E.W., Lambert, J.E., Rovero, F., Barelli, C., Setchell, J.M.,

484   Gillespie, T.R., Mittermeier, R.A., Arregoitia, L.V., de Guinea, M., Gouveia, S.,

485   Dobrovolski, R., Shanee, S., Shanee, N., Boyle, S.A., Fuentes, A., MacKinnon, K.C.,

486   Amato, K.R., Meyer, A.L.S., Wich, S., Sussman, R.W., Pan, R., Kone, I. & Li, B. (2017).

487   Impending extinction crisis of the world's primates: Why primates matter. *Sci. Adv.* **3**,

488   e1600946.

489   Frankham, R., Ballou, J.D., Dudash, M.R., Eldridge, M.D.B., Fenster, C.B., Lacy, R.C.,

490   Mendelson III, J.R., Porton, I.J., Ralls, K. & Ryder, O.A. (2012). Implications of

491   different species concepts for conserving biodiversity. *Biol. Conserv.* **153**, 25–31.

492   Garnett, S.T. & Thomson, S.A. (2020). Are the implications for conservation of a major

493   taxonomic revision of the world's birds' simply serendipity? *Anim. Conserv.* **23**, 355-

494   356.

495   Gippoliti, S. & Amori, G. (2007). The problem of subspecies and biased taxonomy in

496   conservation lists: the case of mammals. *Folia Zool.* **56**, 113–117.

497   Groves, C.P. (2014). Primate taxonomy: inflation or real? *Annu. Rev. Anthropol.* **43**, 27–36.

498   Groves, C. & Grubb, P. (2011). *Ungulate taxonomy*. JHU Press, Baltimore.

499   Hartig, F. (2017). *DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression*

500   *models*. R package version 0.3.0. https://cran.r-project.org/web/packages/DHARMa.

501   Harris, D.J., & Froufe, E. (2005). Taxonomic inflation: species concept or historical geopolitical

502   bias? *Trends Ecol. Evol.* **20**, 6–7.

503   Honacki, J.H., Kinman, K.E. & Koeppl, J.W. (1982). *Mammal species of the world; a taxonomic*

504   *and geographic reference*. Allen Press, Inc. and Association of Systematic Collections,

505   Lawrence.

506   Isaac, N.J., Mallet, J. & Mace, G.M. (2004). Taxonomic inflation: its influence on macroecology

507   and conservation. *Trends Ecol. Evol.* **19**, 464–469.

508   Isaac, N.J. & Purvis, A. (2004). The 'species problem' and testing macroevolutionary

509   hypotheses. *Divers. Distrib.* **10**, 275–281.

510   Isaac, N.J., Turvey, S.T., Collen, B., Waterman, C. & Baillie, J.E. (2007). Mammals on the

511   EDGE: conservation priorities based on threat and phylogeny. *PLoS One* **3**, e296.

512   IUCN Conservation Monitoring Centre. (1986). 1986 *IUCN Red List of threatened animals*.

513   IUCN, Gland and Cambridge.

514   Karl, S.A. & Bowen, B.W. (1999). Evolutionary significant units versus geopolitical taxonomy:

515   molecular systematics of an endangered sea turtle (genus *Chelonia*). *Conserv. Biol.* **13**,

516   990–999.

517   Köhler, J., Vieites, D.R., Bonett, R.M., García, F.H., Glaw, F., Steinke, D., & Vences, M.

518   (2005). New amphibians and global conservation: a boost in species discoveries in a

519   highly endangered vertebrate group. *BioScience* **55**, 693–696.

520   Knowles, J.E. & Frederick, C. (2020). *merTools: Tools for Analyzing Mixed Effect Regression*

521   *Models*. R package version 0.5.2. https://CRAN.R-project.org/package=merTools.

522   Kuznetsova, A., Brockhoff, B. & Christensen, H.B. (2017). lmerTest Package: Tests in Linear

523   Mixed Effects Models. *J. Stat. Softw.* **82**, 1–26.

524   Mace, G.M. (2004). The role of taxonomy in species conservation. *Philos. Trans. R. Soc. B:*

525   *Biol. Sci.* **359**, 711–719.

526   Magallon, S. & Sanderson, M.J. (2001). Absolute diversification rates in angiosperm clades.
527   *Evolution* **55**, 1762–1780.

528   Mayr, E. (1963). *Animal species and evolution*. Belknap Press, Cambridge.

529   Mayr, E. (1996). What is a species, and what is not? *Philos. Sci.* **63**, 262–277.

530   Morrison, W.R., Lohr, J.L., Duchen, P., Wilches, R., Trujillo, D., Mair, M. & Renner, S.S.
531   (2009). The impact of taxonomic change on conservation: Does it kill, can it save, or is it
532   just irrelevant? *Biol. Conserv.* **142**, 3201–3206.

533   Nater, A., Mattle-Greminger, M. P., Nurcahyo, A., Nowak, M. G., De Manuel, M., Desai, T., ...
534   & Krützen, M. (2017). Morphometric, behavioral, and genomic evidence for a new
535   orangutan species. *Curr. Biol.* **27**, 3487–3498.

536   Oates, J. F. (2006). Is the chimpanzee, *Pan troglodytes*, an endangered species? It depends on
537   what "endangered" means. *Primates* **47**, 102–112.

538   Padial, J.M., & De la Riva, I. (2006). Taxonomic inflation and the stability of species lists: the
539   perils of ostrich's behavior. *Syst. Biol.* **55**, 859–867.

540   Pillon, Y. & Chase, M.W. (2007). Taxonomic exaggeration and its effects on orchid
541   conservation. *Conserv. Biol.* **21**, 263–265.

542   R Core Team. (2021). *R: A Language and environment for statistical computing*. R Foundation
543   for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

544   Redding, D.W. & Mooers, A.Ø. (2010). Can systematists help decide the relative worth of bits of
545   biodiversity? *Systematist* **32**, 4–8.

546   Redding, D.W., Mooers, A.Ø., Şekercioğlu, Ç.H., & Collen, B. (2015). Global evolutionary

547   isolation measures can capture key local conservation species in Nearctic and Neotropical

548   bird communities. *Philos. Trans. R. Soc. B: Biol. Sci.* **370**, 20140013.

549   Robuchon, M., Faith, D.P., Julliard, R., Leroy, B., Pellens, R., Robert, A., Thévenin, C., Véron,

550   S. & Pavoine, S. (2019). Species splitting increases estimates of evolutionary history at

551   risk. *Biol. Conserv.* **235**, 27–35.

552   Rueda-Cediel, P., Anderson, K.E., Regan, T.J., & Regan, H.M. (2018). Effects of uncertainty

553   and variability on population declines and IUCN Red List classifications. *Conserv. Biol.*

554   **32**, 916-925.

555   Rylands, A.B. & Mittermeier, R.A. (2014). Primate taxonomy: species and conservation. *Evol.*

556   *Anthropol.* **23**, 8–10.

557   Sangster, G. (2009). Increasing numbers of bird species result from taxonomic progress, not

558   taxonomic inflation. *Proc. Royal Soc. B* **276**, 3185–3191.

559   Simkins, A.T., Buchanan, G.M., Davies, R.G. & Donald, P.F. (2020). The implications for

560   conservation of a major taxonomic revision of the world's birds. *Anim. Conserv.* **23**, 345–

561   352.

562   Stanford, C.B. (2001). The subspecies concept in primatology: the case of mountain gorillas.

563   *Primates* **42**, 309–318.

564   Tattersall, I. (2007). Madagascar's lemurs: cryptic diversity or taxonomic inflation? *Evol.*

565   *Anthropol.* **16**, 12–23.

566   Wilson, D.E. & Reeder, D.M. (2005). *Mammal species of the world: a taxonomic and*

567   *geographic reference (vol. 1)*. JHU Press, Baltimore.

568   Zachos, F.E. (2013). Species splitting puts conservation at risk. *Nature* **494**, 35–35.

569   Zachos, F.E. (2015). Taxonomic inflation, the phylogenetic species concept and lineages in the

570   tree of life–a cautionary comment on species splitting. *J. Zoolog. Syst. Evol.* **53**, 180–184.

571   Zachos, F.E., Apollonio, M., Bärmann, E.V., Festa-Bianchet, M., Göhlich, U., Habel, J.C.,

572   Haring, E., Kruckenhauser, L., Lovari, S., McDevitt, A.D., Pertoldi, C., Rössner, G.E.,

573   Sánchez-Villagra, M.R., Scandura, M. & Suchentrunk, F. (2013). Species inflation and

574   taxonomic artefacts – A critical comment on recent trends in mammalian

575   classification. *Mamm. Biol.* **78**, 1–6.

576   Zachos, F.E. & Lovari, S. (2013). Taxonomic inflation and the poverty of the Phylogenetic

577   Species Concept – a reply to Gippoliti and Groves. *Hystrix* **24**, 142–144.