

1 **Trends:**

2 1. Current definitions of cryptic species are inconsistent and can lead to biased estimates of
3 species diversity.

4 2. Cryptic species are often implied to represent taxa displaying low phenotypic disparity in
5 relation to divergence time, but this relationship is usually not formally quantified.

6 3. Here we propose a quantitative framework, which provides a formal characterization of
7 the intuitive concept of cryptic species.

8 4. The proposed framework facilitates understanding of evolutionary processes leading to
9 and resulting from cryptic species and provides a basis for estimates and modelling of
10 occurrences of cryptic species across taxa and environments.

11 5. The framework fosters a shift from pattern- to process-driven research concerning
12 cryptic species.

13

1 **Title: Finding evolutionary processes hidden in cryptic species**

2

3 **Authors:**

4 Torsten H. Struck¹, Jeffrey L. Feder², Mika Bendiksy^{1,3}, Siri Birkeland¹, José Cerca¹,
5 Vladimir I. Gusarov¹, Sonja Kistenich¹, Karl-Henrik Larsson¹, Lee Hsiang Liow^{1,4}, Michael
6 D. Nowak¹, Brita Stedje¹, Lutz Bachmann¹, Dimitar Dimitrov^{1,5}

7 (except for the two first and the two senior authors, the authors are in alphabetical order)

8

9 ¹Natural History Museum, University of Oslo, 0318 Oslo, Norway

10 ²Dept. Biological Sciences, University of Notre Dame, Notre Dame, IN 46556, USA

11 ³NTNU University Museum, Norwegian University of Science and Technology, 7491
12 Trondheim, Norway

13 ⁴Centre for Ecological & Evolutionary Synthesis (CEES), Department of Biosciences,
14 University of Oslo, 0316 Oslo, Norway

15 ⁵Current address: Center for Macroecology, Evolution and Climate, Natural History Museum
16 of Denmark, University of Copenhagen, Copenhagen, Denmark

17

18 **Corresponding author:**

19 Torsten H. Struck

20 Natural History Museum, Department of Research and Science, University of Oslo

21 PO Box 1172 Blindern

22 N-0318 Oslo

23 Norway

24 e-mail: t.h.struck@nhm.uio.no

25 phone: +4722851740

26 **Keywords:**

27 biodiversity; convergence; genetic divergence; parallelism; phenotypic stasis; taxonomy

28

29 **Abstract:**

30 Cryptic species could represent a substantial fraction of biodiversity. However, inconsistent
31 definitions and taxonomic treatment of cryptic species prevent informed estimates of their
32 contribution to biodiversity and impede our understanding of their evolutionary and
33 ecological significance. We propose a conceptual framework that recognizes cryptic species
34 based on their low levels of phenotypic (morphological) disparity relative to their degree of
35 genetic differentiation and divergence times as compared to non-cryptic species. We discuss
36 how application of a more rigorous definition of cryptic species in taxonomic practice will
37 lead to more accurate estimates of their prevalence in nature, better understanding of their
38 distribution patterns on the tree of life, and increased abilities to resolve the processes
39 underlying their evolution.

40

41 **Main text:**

42 **Cryptic species - taxonomic oddities or biologically relevant entities?**

43 'Cryptic species' is a common and increasingly used term that refers to taxa that cannot
44 readily be distinguished morphologically, yet evidence indicates they are on different
45 evolutionary trajectories (see Box 1). While researchers may not be able to visually recognize
46 cryptic species as different species, the organisms can. Cryptic species are found on all major
47 branches of the tree of life and probably represent a significant portion of undiscovered
48 biodiversity [1-4]. As such, cryptic species might significantly add to our understanding of
49 biodiversity, calling for increased conservation efforts [2, 4-9]. Cryptic species are also
50 important because they serve as an intellectual bridge connecting the study of taxonomy and
51 phylogenetic pattern with ecosystems functioning, evolutionary processes, and
52 macroevolutionary trends, including speciation, parallelism, convergence, and stasis.
53 However, problems with the definition, among others the linkage to the species' taxonomic
54 nomenclature history, and inconsistencies in the use of the term 'cryptic species' make it
55 difficult to draw firm conclusions about their prevalence in nature and their implications for
56 ecology and evolution.

57

58 Here, we discuss the general problem of defining cryptic species based on a literature survey
59 that revealed the wide latitude in what researchers call cryptic species. Some authors have
60 even suggested considering cryptic species as a temporary formalization problem of species
61 delineation, rather than as a natural phenomenon [10]. To help mitigate the problem, we
62 propose a more rigorous, multi-dimensional, and inter-disciplinary approach for cryptic
63 species. The approach focuses on better quantifying the extent of phenotypic disparity of taxa
64 (see Glossary Box) compared to the degree to which they have genetically diverged and
65 exchanged genes (have evolved reproductive isolation). Standardizing the delineation of

66 cryptic species will facilitate investigations into several outstanding questions concerning
67 their biological significance (see Outstanding Questions Box). It will also lead to a better
68 characterization and understanding of the different types of cryptic species, from recently
69 diverged to phylogenetically distant taxa. In doing so, conclusions concerning 1) evolutionary
70 parallelism, convergence, and stasis; 2) the role that cryptic species play in ecosystem
71 functioning; and 3) factors initiating and contributing to speciation can be more confidently
72 accepted.

73

74 **The problem of definition in theory and practice**

75 Cryptic species have generated both taxonomic and evolutionary ambiguity. A frequently
76 cited definition of cryptic species [5] describes them as two or more distinct species that were
77 earlier classified as one. Hence, cryptic species are defined based only on their taxonomic
78 nomenclature history. However, this is unsatisfactory because various biological factors or
79 taxonomic artefacts might result in erroneous species lumping. Additionally, it offers no
80 guidance for how morphologically similar or by how many characters species should differ to
81 be considered as cryptic. Moreover, one of the longest and most contentious debates in
82 evolution concerns what constitutes a species. If biologists cannot even agree on what to
83 consider different species, then how can we reach consensus on what represents cryptic
84 species?

85

86 Our literature survey of 606 studies indicates that the lack of philosophical clarity translates
87 into a serious empirical problem in the operational designation of cryptic species (see Box 2,
88 Supplementary Material and Tables S1-4). For example, 47% of them, even though claiming
89 cryptic species status for taxa, presented no phenotypic data, while 25.3% reported at least
90 one trait differing between cryptic species. Thus, morphological similarity is subjectively

91 evaluated and rarely quantified to address how similar cryptic species are [11-13]. Moreover,
92 non-morphological phenotypes, such as behaviour, were seldom considered (see Box 2). In
93 this regard, cryptic species designation was often pattern-driven with a focus on
94 morphological characters discriminating taxa and little else. When several phenotypic traits
95 were assessed, analyses seldom extended to species beyond the focal cryptic species. This is
96 relevant because rates of morphological evolution for cryptic 'ingroup' taxa should be
97 substantially (statistically) reduced compared to non-cryptic taxa to be considered cryptic.

98

99 The genetic data provided in the surveyed studies were also of limited utility in cryptic
100 species delineation. Of the 606 studies, 35.5% based cryptic species designation on only a
101 single molecular marker, most often from the plastid or mitochondrion, and lacked
102 information on phenotypic disparity. Only 15.4% of the surveyed studies combined different
103 types of molecular markers with morphological and/or other phenotypic data, and compared
104 genetic divergence of the cryptic taxa to other congeneric non-cryptic species. The results
105 show that there is remarkable inconsistency in the operational designation of 'cryptic species'
106 [5, 14, 15] and huge variation in the applied analytical rigor [11, 16-21]. Taxonomic practice
107 for identifying cryptic species thus requires attention if the term is to be useful for
108 comparative studies.

109

110 With recent advances in high-throughput DNA sequencing, visualization/microscopy, and
111 statistical analytical tools, there are no technological or methodological impediments
112 restricting higher standards in the empirical investigation of cryptic species [22, 23]. This is
113 important as informed estimates of species diversity and speciation rates are crucial for
114 understanding evolutionary processes and ecosystem functioning, and for developing
115 effective conservation strategies and sustainable usage of ecosystem services [2, 4-9]. Cryptic

116 species are one component of these estimates. Estimates of cryptic biodiversity based on
117 vague definitions are of little help and, like undiscovered species or lack of species lists, will
118 be counterproductive. For example, in ecology and conservation research, cryptic species are
119 usually taken at face value based on the original reports. In particular, studies investigating
120 patterns of cryptic species distribution across habitats, taxonomic groups, or life history
121 strategies, are often based on meta-analyses [5, 24-28]. Given the shaky foundation in which
122 cryptic species appear to be subjectively defined, it is difficult to place much confidence in
123 the conclusions drawn from such meta-analyses. Sympatric cryptic species might, for
124 example, contradict the ecological paradigm of competitive exclusion [29, 30], but based on
125 the current state it remains difficult to decide whether this is specifically or generally true.
126 Similar considerations apply to studies of parallelism, convergence, and stasis. Without better
127 standardization of the designation of cryptic species including details about phenotypic
128 variation, levels of genomic differentiation, and divergence times, it remains difficult to make
129 proper inference about evolutionary processes. Such standardizations as suggested herein will
130 substantially improve comparability across lineages, as taxonomic nomenclature traditions
131 are replaced with studies quantifying variation in a similar manner within and across groups.

132

133 **The conceptual framework**

134 Accurate pattern and process-driven research on cryptic species is possible. However, to
135 accomplish this, a sound and consistent foundation for defining cryptic species is needed. We
136 do not pretend to solve the cryptic species problem completely here, but offer a conceptual
137 framework to alleviate the problem by combining phenotypic disparity and genetic
138 divergence. The latter serving as a proxy for reduced gene flow and an estimate of the time
139 since divergence from the most recent common ancestor (MRCA). By doing so, we
140 emphasize the importance of reduced gene flow between taxa and the establishment of

141 reproductive isolation between sexually reproducing populations relative to the extent to
142 which they have changed in morphological and other phenotypic characters. As we explain
143 below, this approach facilitates studies of parallelism, convergence, and speciation. The
144 proposed framework provides a yardstick for the standardization of cryptic species
145 descriptions without getting too entangled in the issue of species concepts. We concentrate on
146 sexually reproducing organisms, for which a metric of gene flow and divergence time versus
147 phenotypic disparity are key considerations.

148

149 Our conceptual framework highlights two important elements for defining cryptic species
150 (Fig. 1). First, species have to be distinguishable, for example, as statistically separable and
151 diverged genotypic clusters of individuals (reflecting reproductive isolation) that do not form
152 diagnostic morphological clusters. Although estimates of reproductive isolation in nature are
153 only truly possible for taxa that geographically overlap, data from laboratory crosses, when
154 technically feasible, and other information can be used to help gauge the level of gene flow
155 and reproductive isolation. One major consideration is the time point when diverging
156 populations are considered as being genetically and reproductively distinguishable species
157 [e.g., 31, 32-35], as this will affect conclusions about recently diverged species.

158 Consequently, cases, where populations exhibit sufficient gene flow to not cluster
159 distinctively using methods like STRUCTURE or genetic network analyses, should be
160 considered, if at all, as races or ecotypes [34, 36], rather than cryptic species [37].

161

162 Second, the temporal dimension of cryptic species should be recognized by their showing of
163 statistically lower degrees of phenotypic (or more specifically morphological) disparity than
164 non-cryptic relatives given similar divergence time estimates from their MRCA (Fig. 1). By
165 placing morphological disparity directly in relation to time (genetic divergence), recognition

166 of cryptic species can become divorced from taxonomic nomenclature traditions based on the
167 numbers of previously recognized species (e.g., lumpers vs. splitters), and debates about
168 levels of 'crypticity' [10] more nuanced.

169

170 Although these two components of defining cryptic species seem self-evident, they are
171 seldom adequately performed to allow for quantitative comparisons. For example, the
172 temporal dimension is frequently ignored [38-40] and, of the 606 studies in our survey, only
173 3.3% and 4.5% of the reported divergence events could confidently be regarded as young or
174 old, respectively. For accurately determining genetic divergence, genome-wide sequence data
175 are highly preferred for any group of taxa. However, very few studies applied genome-scale
176 data (see Box 2) [18, 23]. Uniparentally inherited markers, such as the mitochondrial
177 cytochrome oxidase subunit I gene (COI) - the target marker for DNA barcoding in animals -
178 do not provide a comprehensive assessment of gene flow and reproductive isolation. There
179 are several examples of high genetic divergence in COI that reflect deep population structure
180 rather than species differences [18].

181

182 More importantly, to identify and quantify species that are cryptic from those that are not,
183 detailed information about phenotypic disparity has to be related to genetic divergence, levels
184 of gene flow, and reproductive isolation. Therefore, population to species level
185 morphological variation needs to be explicitly quantified in order to measure morphological
186 disparity among cryptic species and their relatives as for example done in [41-43]. Available
187 species descriptions can provide a good starting point for such morphological comparisons,
188 providing information on both discrete and continuous characters. Depending on the data,
189 appropriate methods for the quantification of morphological variation are available, including
190 geometric morphometrics [44], landmark-free approaches such as the generalized procrustes

191 surface analysis [45], and multivariate analysis like non-metric multidimensional scaling
192 [46]. These methods and clustering, principal component, and discriminant function analyses
193 should be employed to assess whether populations can be statistically distinguished from
194 another or not. Also statistical tools like disparity through time plots [47] allow for testing if
195 morphological disparity between hypothesized cryptic taxa is significantly lower than
196 expected given a null random walk expectation of drift. Tests of rate variation (e.g., variance
197 ratio test) among hypothesized cryptic and non-cryptic lineages can also indicate whether
198 morphological and other phenotypic traits (e.g., those related to behaviour, life history, and
199 physiology) deviate significantly from neutral expectation to statistically support cryptic
200 species status for taxa. Note that hybridization has the potential to complicate analyses by
201 reducing phenotypic disparity below levels seen for allopatric or completely reproductively
202 isolated populations. However, it can generally be expected that proportional reductions in
203 the level of genomic divergence would compensate for this and help to maintain the
204 standardization of cryptic species delineation.

205

206 Currently there are no studies that adhere completely to the proposed framework. There are
207 several examples, however, where most of the requirements are fulfilled, e.g. in studies of
208 unicellular eukaryotes [48], cnidarians [41], annelids [42], molluscs [43], vertebrates [46],
209 and plants [49]. However, the primary focus of these studies has been to find diagnostic
210 characters. Phenotypic disparity was usually not cast in relation to other non-cryptic taxa
211 and/or genetic divergence. One reason for this is that detailed examination of phenotypic and
212 genetic variation in a comparative context, as proposed here, is time-consuming and not
213 practical for projects whose primary focus is not the delineation of cryptic species (but then
214 they should also refrain from assigning them). However, accurate rather than quick science is
215 what should be aimed for, and when conducted properly, the proposed framework will

216 provide the rigor to move beyond suggestive evidence to full and more standardized
217 recognition of cryptic species.

218

219 **Evolutionary processes and cryptic species**

220 Given a standardized and more accurate characterization of cryptic species, it is possible to
221 examine their ecological and evolutionary implications in greater depth and with more
222 confidence. For example, one question of interest is the extent to which cryptic species
223 represent recently diverged versus more distantly related taxa. Other questions concerning
224 evolutionary processes like parallelism, stasis, and convergence that are often considered
225 primarily with respect to single traits [50-52] could also be extended to investigate whole
226 phenotypes by more robust analysis of cryptic species. In this regard, underlying selective
227 regimes might be expected to be more pronounced or generally constrained to impact the
228 entire (or nearly entire) suite of phenotypic traits [53], to which the term 'cryptic speciation'
229 has been misleadingly applied in recent years (Fig. 1A). We examine these questions below.

230

231 *Recent divergence (Fig. 2A)* – In this case, hypothesized cryptic species are sister taxa or
232 members of a species complex with short divergence times, which are too recent for
233 substantial morphological differences to accumulate [37, 54, 55]. In many of these instances,
234 the rate of accumulation of morphological disparity might actually not differ significantly
235 from older non-cryptic species (Fig. 2A). In speciation research it is commonly assumed that
236 in the early stages of speciation selection acts largely on physiological, immunological,
237 reproductive or behavioural traits rather than on morphology [16, 17, 19]. Hence, for very
238 young species, similarity in morphology might not be unexpected and it could take additional
239 time to visually observe differences between taxa [10]. However, recently diverged taxa
240 showing significantly lower rates of morphological disparity might be constrained by

241 stabilizing selection and represent early stages of stasis.

242

243 *Parallelism (Fig. 2B)* – Cryptic species that evolved by parallelism are not sister taxa, but are
244 phylogenetically separated from each other to such a degree that their similarity can no
245 longer be considered symplesiomorphic, but rather independently evolved from
246 morphologically similar ancestors. In comparison to more closely related and younger non-
247 cryptic species, morphological disparity changes less as the cryptic species evolve from one
248 similar morphotype to another similar one (Fig. 2B). However, if the evolution of the new
249 morphotype in one lineage precedes the other lineage in time, morphological disparity will
250 first increase and then decrease again (similar to the plot in Fig. 2C). Regardless, ancestral
251 character state reconstructions are important in order to distinguish between recent
252 divergence, convergence or parallelism, and to assess and test rates of morphological change.
253 Swift et al. [41], for example, showed that similar morphologies for lake species evolved by
254 parallelism in closely related scyphozoan species. Confirmation of parallelism begs the
255 question of whether similar morphotypes evolved due to intrinsic (e.g., developmental or
256 genetic constraints) or extrinsic factors (e.g., deterministic environmental pressures)
257 confining the available morphospace to only one selectively advantageous solution.

258

259 *Convergence (Fig. 2C)* – In this case, cryptic species are not closely related and their
260 morphological similarity results from independent evolution of morphologically dissimilar
261 ancestors. At early stages of divergence, cryptic and non-cryptic species pairs are expected to
262 show similar rates of morphological differentiation. However, at some point in time the
263 cryptic species pairs would begin to converge morphologically (Fig. 2C). Convergence as a
264 mechanism for cryptic species is rare, but has been reported in the deep sea [56]. In contrast
265 to parallelism, intrinsic factors are expected to be less important for convergence than

266 extrinsic ones, as convergent evolution is assumed to have started from different genetic and
267 developmental backgrounds.

268

269 *Stasis (Fig. 2D)* – Under stasis, cryptic species are sister taxa or members of a complex that
270 retain a high degree of morphological similarity over extended periods of time (Fig. 2D).
271 Hence, symplesiomorphies prevail for millions of years, and significantly longer than
272 expected by random drift. For example, one cryptic complex of annelid worms has been
273 shown to display little morphological variation over tens of millions of years [42]. The lack
274 of morphological diversification could result from low standing genetic variation and/or
275 developmental constraints on the morphospace [5, 57]. It is also possible that the ecology of
276 taxa showing stasis has remained relatively constant through time and strong stabilizing
277 selection has retained a common, shared morphology. This raises the question whether
278 cryptic species tend to be ecological generalists versus specialists, the answer to which might
279 hinge on how common adaptation to different environments underlies speciation and depends
280 on morphological change.

281

282 **Concluding remarks**

283 Current research practices regarding cryptic species require change. There is much insight to
284 be gained by standardizing and increasing the rigor in the way that cryptic species are defined
285 and studied. Current practices, however, do not allow firm conclusions to be made
286 concerning the number and significance of cryptic species in nature or the evolutionary
287 processes associated with them. Indeed, given the results of our literature survey it is likely
288 that many reported cryptic species should not be considered as such. Consequently, there is a
289 need for careful re-analyses of many proposed cryptic species complexes with more rigorous
290 criteria to better assess their true prevalence in nature. We propose an interdisciplinary

291 approach that involves combining comprehensive data on genomic and phenotypic traits to
292 statistically test for significant differences in rates of phenotypic disparity (e.g.,
293 morphological disparity) between cryptic versus non-cryptic species. This approach will
294 standardize the designation of cryptic species in the literature for taxonomic and comparative
295 purposes, eliminate the history of taxonomic nomenclature as a consideration and enable
296 meta-analyses based on comparisons involving taxa categorized as displaying similar versus
297 differing levels of disparity, time periods of divergence, and degree of reproductive isolation.
298 Adopting the approaches we advocate will provide a more sound basis for policy making in
299 conservation biology and make it possible to address a number of questions involving
300 evolutionary parallelism, convergence, and stasis associated with cryptic species (see
301 Outstanding Questions box), helping to reveal the biological meaning hidden in cryptic
302 species. Conducted across lineages, general principles and accurate predictions, e.g., to what
303 extent cryptic species prevail in certain groups or are affected by climate change can be
304 deduced.
305

References

- 306
307
308 1. Jörger, K. and Schrödl, M. (2013) How to describe a cryptic species? Practical challenges
309 of molecular taxonomy. *Front Zool* 10 (1), 59.
- 310 2. Pante, E. et al. (2015) Species are hypotheses: avoid connectivity assessments based on
311 pillars of sand. *Mol Ecol* 24 (3), 525-544.
- 312 3. Loxdale, H.D. et al. (2016) Known knowns and unknowns in biology. *Biol J Linn Soc* 117
313 (2), 386-398.
- 314 4. Nygren, A. (2013) Cryptic polychaete diversity: a review. *Zool Scr* 43, 172–183.
- 315 5. Bickford, D. et al. (2007) Cryptic species as a window on diversity and conservation.
316 *Trends Ecol Evol* 22 (3), 148-155.
- 317 6. Alizon, S. et al. (2008) Competition between cryptic species explains variations in rates of
318 lineage evolution. *Proc Natl Acad Sci* 105 (34), 12382-12386.
- 319 7. Nadler, S.A. and Perez-Ponce de Leon, G. (2011) Integrating molecular and morphological
320 approaches for characterizing parasite cryptic species: implications for parasitology.
321 *Parasitology* 138, 1688-1709.
- 322 8. Boykin, L.M. et al. (2012) Species delimitation and global biosecurity. *Evol Bioinform* 8,
323 1-37.
- 324 9. Krug, P.J. et al. (2013) Integrative species delimitation in photosynthetic sea slugs reveals
325 twenty candidate species in three nominal taxa studied for drug discovery, plastid symbiosis
326 or biological control. *Mol Phylogenet Evol* 69, 1101-1119.
- 327 10. Korshunova, T. et al. (2017) External diversity is restrained by internal conservatism:
328 New nudibranch mollusc contributes to the cryptic species problem. *Zool Scr* 46, 683–692.
- 329 11. Wu, Z.-Z. et al. (2014) Sequence analysis of mitochondrial ND1 gene can reveal the
330 genetic structure and origin of *Bactrocera dorsalis* s.s. *BMC Evol Biol* 14 (1), 55.

- 331 12. Sanchez, G. et al. (2016) Evaluation of the 5' end of the 16S rRNA gene as a DNA
332 barcode marker for the Cephalopoda. *Fish Sci* 82 (2), 279-288.
- 333 13. Schmidt, R.C. et al. (2016) High levels of endemism in suckermouth catfishes
334 (Mochokidae: Chiloglanis) from the Upper Guinean forests of West Africa. *Mol Phylogenet*
335 *Evol* 100, 199-205.
- 336 14. Wang, Y. et al. (2014) Morphology, molecular genetics, and bioacoustics support two
337 new sympatric *Xenophrys* toads (Amphibia: Anura: Megophryidae) in Southeast China.
338 *PLoS ONE* 9 (4), e93075.
- 339 15. Van Campenhout, J. et al. (2016) Transcription, signaling receptor activity, oxidative
340 phosphorylation, and fatty acid metabolism mediate the presence of closely related species in
341 distinct intertidal and cold-seep habitats. *Genome Biol Evol* 8 (1), 51-69.
- 342 16. Bensch, S. et al. (2004) Linkage between nuclear and mitochondrial DNA sequences in
343 avian Malaria parasites: multiple cases of cryptic speciation? *Evolution* 58 (7), 1617-1621.
- 344 17. Damm, S. et al. (2010) An integrative approach to species discovery in odonates: from
345 character-based DNA barcoding to ecology. *Mol Ecol* 19, 3881–3893.
- 346 18. Giska, I. et al. (2015) Deeply divergent sympatric mitochondrial lineages of the
347 earthworm *Lumbricus rubellus* are not reproductively isolated. *BMC Evol Biol* 15 (1), 217.
- 348 19. Derycke, S. et al. (2016) Coexisting cryptic species of the *Litoditis marina* complex
349 (Nematoda) show differential resource use and have distinct microbiomes with high
350 intraspecific variability. *Mol Ecol* 25, 2093–2110
- 351 20. Karanovic, T. et al. (2016) Cryptic species or inadequate taxonomy? Implementation of
352 2D geometric morphometrics based on integumental organs as landmarks for delimitation
353 and description of copepod taxa. *Syst Biol* 65 (2), 304–327.
- 354 21. Razkin, O. et al. (2017) Species delimitation for cryptic species complexes: case study of
355 *Pyramidula* (Gastropoda, Pulmonata). *Zool Scr* 46 (1), 55-72.

- 356 22. Richards, S. (2015) It's more than stamp collecting: how genome sequencing can unify
357 biological research. *Trends Genet* 31 (7), 411-21.
- 358 23. Janzen, D.H. et al. (2017) Nuclear genomes distinguish cryptic species suggested by their
359 DNA barcodes and ecology. *Proc Natl Acad Sci* 114 (31), 8313-8318.
- 360 24. Pfenninger, M. and Schwenk, K. (2007) Cryptic animal species are homogeneously
361 distributed among taxa and biogeographical regions. *BMC Evol Biol* 7 (1), 121.
- 362 25. Poulin, R. and Pérez-Ponce de León, G. (2017) Global analysis reveals that cryptic
363 diversity is linked with habitat but not mode of life. *J Evol Biol* 30 (3), 641-649.
- 364 26. Perez-Ponce de Leon, G. and Poulin, R. (2016) Taxonomic distribution of cryptic
365 diversity among metazoans: not so homogeneous after all. *Biol Lett* 12, 20160371.
- 366 27. Adams, M. et al. (2014) Global biodiversity assessment and hyper-cryptic species
367 complexes: More than one species of elephant in the room? *Syst Biol* 63 (4), 518-533.
- 368 28. Skoracka, A. et al. (2015) Cryptic speciation in the Acari: a function of species lifestyles
369 or our ability to separate species? *Exp Appl Acarol* 67, 165-182.
- 370 29. Chesson, P. (1991) A need for niches. *Trends Ecol Evol* 6, 26–28.
- 371 30. Gause, G.F. (1934) *The struggle for existence*, Hafner Pub. Co.
- 372 31. Norris, R.D. and Hull, P.M. (2012) The temporal dimension of marine speciation. *Evol*
373 *Ecol* 26 (2), 393-415.
- 374 32. The Marie Curie SPECIATION Network (2012) What do we need to know about
375 speciation? *Trends Ecol Evol* 27 (1), 27-39.
- 376 33. De Queiroz, K. (2007) Species concepts and species delimitation. *Syst Biol* 56 (6), 879-
377 86.
- 378 34. Roux, C. et al. (2016) Shedding light on the grey zone of speciation along a continuum of
379 genomic divergence. *PLoS Biol* 14 (12), e2000234.

- 380 35. Bernardo, J. (2011) A critical appraisal of the meaning and diagnosability of cryptic
381 evolutionary diversity, and its implications for conservation in the face of climate change. In
382 Climate Change, Ecology and Systematics. (Hodkinson, T.R. et al. eds), pp. 380-438,
383 Cambridge University Press.
- 384 36. Sukumaran, J. and Knowles, L.L. (2017) Multispecies coalescent delimits structure, not
385 species. Proc Natl Acad Sci 114 (7), 1607-1612.
- 386 37. Reidenbach, K.R. et al. (2012) Patterns of genomic differentiation between ecologically
387 differentiated M and S forms of *Anopheles gambiae* in West and Central Africa. Genome
388 Biol Evol 4 (12), 1202-1212.
- 389 38. Harder, A.M. et al. (2016) Diversity and distribution within the sea spider genus
390 *Pallenopsis* (Chelicerata: Pycnogonida) in the Western Antarctic as revealed by
391 mitochondrial DNA. Polar Biol 39 (4), 677-688.
- 392 39. Sauvage, T. et al. (2016) A metabarcoding framework for facilitated survey of endolithic
393 phototrophs with tufA. BMC Ecol 16 (1), 8.
- 394 40. Williams, J.T. and Viviani, J. (2016) *Pseudogramma polyacantha* complex (Serranidae,
395 tribe Grammistini): DNA barcoding results lead to the discovery of three cryptic species,
396 including two new species from French Polynesia. Zootaxa 4111 (3), 15.
- 397 41. Swift, H.F. et al. (2016) Three routes to crypsis: Stasis, convergence, and parallelism in
398 the *Mastigias* species complex (Scyphozoa, Rhizostomeae). Mol Phylogenet Evol 99, 103-
399 115.
- 400 42. Struck, T.H. et al. (2017) Two new species in the annelid genus *Stygocapitella*
401 (Orbiniida, Parergodrilidae) with comments on their biogeography. Zootaxa 4286 (3), 301-
402 332.
- 403 43. Wada, S. et al. (2013) Long-term stasis and short-term divergence in the phenotypes of
404 microsnailes on oceanic islands. Mol Ecol 22 (18), 4801-4810.

- 405 44. Bookstein, F.L. (1991) Morphometric tools for landmark data. Geometry and biology,
406 Cambridge University Press.
- 407 45. Pomidor, B.J. et al. (2016) A landmark-free method for three-dimensional shape analysis.
408 PLoS ONE 11 (3), e0150368.
- 409 46. Shirley, M.H. et al. (2014) Rigorous approaches to species delimitation have significant
410 implications for African crocodylian systematics and conservation. Proc R Soc B: Biol Sci
411 281 (1776).
- 412 47. Harmon, L.J. et al. (2003) Tempo and mode of evolutionary radiation in iguanian lizards.
413 Science 301, 961–964.
- 414 48. Krenek, S. et al. (2015) New *Paramecium* (Ciliophora, Oligohymenophorea) congeners
415 shape our view on its biodiversity. Org Divers Evol 15 (2), 215-233.
- 416 49. Vigalondo, B. et al. (2015) Unmasking cryptic species: morphometric and phylogenetic
417 analyses of the Ibero-North African *Linaria incarnata* complex. Bot J Linn Soc 177 (3), 395-
418 417.
- 419 50. Huang, S. et al. (2015) Convergence, divergence, and parallelism in marine biodiversity
420 trends: Integrating present-day and fossil data. Proc Natl Acad Sci 112 (16), 4903-4908.
- 421 51. Hunt, G. et al. (2015) Simple versus complex models of trait evolution and stasis as a
422 response to environmental change. Proc Natl Acad Sci 112 (16), 4885-4890.
- 423 52. Ralph, P.L. and Coop, G. (2015) Convergent evolution during local adaptation to patchy
424 landscapes. PLoS Genet 11 (11), e1005630.
- 425 53. Futuyma, D.J. (2010) Evolutionary constraint and ecological consequences. Evolution 64
426 (7), 1865-1884.
- 427 54. Knowlton, N. (1993) Sibling species in the sea. Annu Rev Ecol Syst 24 (1), 189-216.
- 428 55. Gustafsson, A.L.S. et al. (2014) Genetics of cryptic speciation within an arctic mustard,
429 *Draba nivalis*. PLoS ONE 9 (4), e93834.

- 430 56. Vrijenhoek, R.C. (2009) Cryptic species, phenotypic plasticity, and complex life
431 histories: Assessing deep-sea faunal diversity with molecular markers. *Deep Sea Res Part 2*
432 *Top Stud Oceanogr* 56 (19), 1713-1723.
- 433 57. Appeltans, W. et al. (2012) The magnitude of global marine species diversity. *Curr Biol*
434 22 (23), 2189-2202.
- 435 58. Winker, K. (2005) Sibling species were first recognized by William Derham (1718). *Auk*
436 122, 706-707.
- 437 59. Mayr, E. (1970) *Populations, species, and evolution: an abridgment of Animal species*
438 *and evolution*, Belknap Press of Harvard University Press.
- 439 60. Wills, M.A. (2001) *Morphological disparity: A primer*. In *Fossils, phylogeny, and form:*
440 *An analytical approach*. (Adrain, J.M. et al. eds), pp. 55-144, Kluwer Academic/Plenum
441 Publishers.
- 442
443

444 **Figures:**

445 **Figure 1.** Our conceptual framework for cryptic species. The x-axis represents the time of
446 divergence between taxa since their most recent common ancestor (MRCA) approximated by
447 genetic divergence. The y-axis represents phenotypic (morphological) disparity. Intraspecific
448 variation (polymorphism) within a taxon is depicted by the dark green area in the lower left
449 corner of the figure. The null hypothesis is that morphological disparity between taxa relative
450 to sister species should increase proportionately with divergence time (light green area).
451 However, morphological disparity could increase at a significantly higher rate than the null
452 expectation due to, for example, a recent adaptive radiation (orange area in the upper left
453 corner of the figure). Alternatively, morphological disparity could also be substantially lower
454 than expected over time (blue area in the lower right corner), the hallmark of cryptic species.
455

456 **Figure 2.** Expected signatures of four evolutionary processes that can lead to cryptic species,
457 with the colours of lines in phylogenies and graphs corresponding to the different areas in
458 Figure 1 and species with similar (identical) morphotypes denoted with 'sim.'. Panels on the
459 left denote the phylogenetic relationships among taxa, while the panels to the right depict the
460 evolution of morphological disparity through time for pairs of cryptic and non-cryptic species
461 (e.g., A1/A2 vs. A1/A3). (A) Recent divergence: cryptic species are very closely related and
462 only recently diverged from each other. However, the rate of morphological disparity is not
463 necessarily substantially different from that for non-cryptic species and, as such, these taxa
464 may not actually represent cryptic species. The supposed cryptic species might indeed be on a
465 trajectory, which with time might lead from the borders of the dark green area to the light
466 green area in Figure 1. (B) Parallelism: the cryptic species are not very closely related to each
467 other and the rate of morphological disparity for non-cryptic species is much greater than that
468 for cryptic species. While disparity between non-cryptic species evolved from the dark to the

469 light green area, disparity between the cryptic species progressed into the dark blue area of
470 Figure 1. (C) Convergence: the cryptic species are also not closely related to each other.
471 Initially, morphological disparity for cryptic species can change in a manner similar to that
472 for the non-cryptic species pair. However, at some point, morphological disparity decreases
473 for the cryptic species, while continuing to increase between non-cryptic taxa. Hence, in their
474 past the level of disparity of the cryptic species was first within the light green area of Figure
475 1, but then evolved towards the dark blue area associated with the low level of disparity of
476 cryptic species. (D) Stasis: the cryptic species are closely related to each other or are part of a
477 species complex and diverged a long time ago. In comparison to non-cryptic species, the rate
478 of morphological change is substantially reduced, as cryptic species evolved from the dark
479 green to the dark blue area of Figure 1.

480

481 **Figure I** (as part of Text box 1). Scientific publications on the subject of cryptic species since
482 1940. (A) The number of papers found with the search term ‘cryptic speci*’ (dark green line)
483 and ‘cryptic speciation’ (light green). Of note is the marked increase in publications since
484 1990. (B) The number of papers included in the literature survey (see Box 2) that included
485 molecular data in the study (orange line) is also increasing similar to the overall numbers in
486 A. Dark blue bars indicate the percentages of molecular papers that analysed more than one
487 genetic marker and light blue bars studies those based on genomic data. Note that these
488 percentages are not increasing through time.

489

490 **Text boxes:**

491 **Text box 1: Cryptic species - history and definitions**

492 The English clergyman William Derham reported cryptic species in the avian genus
493 *Phylloscopus* as early as 1718 [58]. Cryptic species have thus been recognized for several
494 hundred years. In the last few decades the number of publications referring to cryptic species
495 has increased dramatically (Fig. IA), likely due to more researchers in the field and the
496 increased use of genetic methods to distinguish taxa (Fig. IB and, for example, [5, 10]).
497 However, criteria used in the literature to designate taxa as cryptic have often been vague and
498 non-uniform. In the few cases where an explicit definition has been stated, the wording is
499 often similar to that of Bickford et al. [5]: Cryptic species are “two or more distinct species
500 that are erroneously classified (and hidden) under one species name”. This taxonomy-based
501 definition is often elaborated upon to highlight that cryptic species are morphologically
502 indistinguishable [5, 35]. Others have included an additional requirement of genetic
503 divergence or distinctiveness between cryptic species ([15]; see Supplementary Table S4 for
504 a list of definitions). How genetically diverged populations must be to be considered cryptic
505 species is usually not specified, but one can assume that this will be of the same magnitude as
506 for non-cryptic species (e.g., a certain barcode gap) [5]. On the other hand, several definitions
507 seem to mostly follow trends and concepts related to the research topic of the paper or field
508 of the researcher. For example, in speciation research, definitions tend to highlight
509 reproductive isolation and the biological species concept [37]. Mayr [59], for instance,
510 defined cryptic species as “morphologically similar or identical natural populations that are
511 reproductively isolated”. Other terms like 'semi-cryptic', 'pseudo-cryptic', 'sibling', and
512 'hypercryptic' indicating different degrees of 'crypticity' have also been proposed [10],
513 complicating the debate of the biological relevance of cryptic species. Regardless, our
514 literature survey (see Box 2) revealed that many cryptic species have been defined based on

515 molecular data and taxonomic history, with little regard for actually quantifying

516 morphological disparity.

517

518 **Text Box 2: Characteristics of published studies of cryptic species**

519 Our literature survey was based on the ISI web of science '*Life Sciences*' database, using the
520 search term '*cryptic speci**' for '*Topic*' on June 17, 2016. The initial search returned 6,002
521 entries (see Supplementary Table S1), from which ~15% were discarded as they were either
522 not research papers, did not use our search term in a taxonomic context, or were not written
523 in English. From the remaining publications, 606 were randomly chosen (see Supplementary
524 Table S2) and assessed according to (I) how cryptic species were defined; (II) whether and
525 which types of genetic markers were scored; (III) the analyses conducted; and (IV) the
526 conclusions that could be drawn (see Supplementary Material and Table S3 for additional
527 details). For these 606 papers, 72.4% involved animals, 7.5% plants, 10.1% fungi, and 6.4%
528 other groups, including protozoans. Only 14.0% of the studies explicitly referred to a specific
529 definition of the term 'cryptic species', indicating the degree of subjectivity in the field.
530 Moreover, according to the Code species, including cryptic ones, are only valid when
531 accompanied by a formal description. However, only 19.3% of the studies provided such
532 formal descriptions. This low number can be indicative of uncertainties of the species status,
533 ignorance of taxonomic practice or that the species were formally described elsewhere.

534

535 The majority of studies (84.2%) provided molecular data, but many (35.5%) used only one
536 locus. In comparison, only 42.7% of the studies included explicit analyses of morphological
537 data and 23.9% of other phenotypic traits. Overall, 56.6% of the studies targeted
538 mitochondrial loci and 52.6% nuclear markers. Of the studies using nuclear data, 48.3%
539 contained results for multiple loci. Very few studies included genome-scale data (3.1%). The
540 relative numbers of studies with more than one marker or genomic data have not increased in
541 recent years (Fig. IB in Box 1). Most studies (73.9%) provided an estimate of genetic
542 divergence of some form (e.g. distance estimates or phylograms) and included congeneric

543 species in the comparison (61.4%). However, only 16.0% of the studies applied genetic
544 dating methods to estimate the time to the most recent common ancestor (MRCA) and only
545 4.3% used fossil calibrations.

546

547 **Glossary box:**

548 **Convergence:** Independent evolution of a derived character state between taxa from different
549 ancestral traits [41].

550 **Disparity:** The morphological or phenotypic difference between taxa [60].

551 **Most recent common ancestor (MRCA):** The last ancestor genetically shared by a group of
552 individuals.

553 **Parallelism:** Independent evolution of a character state in different taxa from a similar and
554 shared ancestral trait [41].

555 **Pattern-driven research:** Research focusing on the detection of biological patterns in
556 empirical data.

557 **Process-driven research:** Research focusing on the underlying processes generating
558 observed patterns.

559 **Stasis:** Retention of the same ancestral character state over an extended period of time [41].

560 **Symplesiomorphy:** Character state of the MRCA present in descendant taxa.

561

562 **Acknowledgments:**

563 We would like to thank the Natural History Museum of the University of Oslo for funding
564 and Sarah Samadi and Gene Hunt for discussion on the subject. We also appreciated the
565 comments of three anonymous reviewers. This is NHM Evolutionary Genomics Lab
566 contribution No. ??.

1 **Outstanding question box:**

- 2 1. What is the general relationship between phenotypic disparity and reproductive isolation
3 and genetic divergence through time?
- 4 2. Do thresholds of phenotypic disparity indicating the presence of cryptic species exist or
5 is the relationship a continuum, with taxa lying in the tail of the distribution warranting
6 cryptic species status?
- 7 3. Which methods for assessing phenotypic disparity and their significance are most
8 universally applicable and most powerful with regards to discerning cryptic species?
- 9 4. Is it possible to establish an *a priori* best-practice strategy for defining cryptic species
10 across a broad range of diverse taxonomic groups?
- 11 5. Are there more cryptic species in certain branches of the tree of life, among taxa with
12 certain life histories (e.g., generalists vs. specialists), or in certain habitats?
- 13 6. Which cryptic species are the results of recent speciation, parallelism, convergence or
14 stasis, and how common are they?
- 15 7. What are the relevant intrinsic and extrinsic factors affecting morphological evolution
16 and to what degree do they affect the phenotypic landscape of cryptic species?
- 17

Figure 1

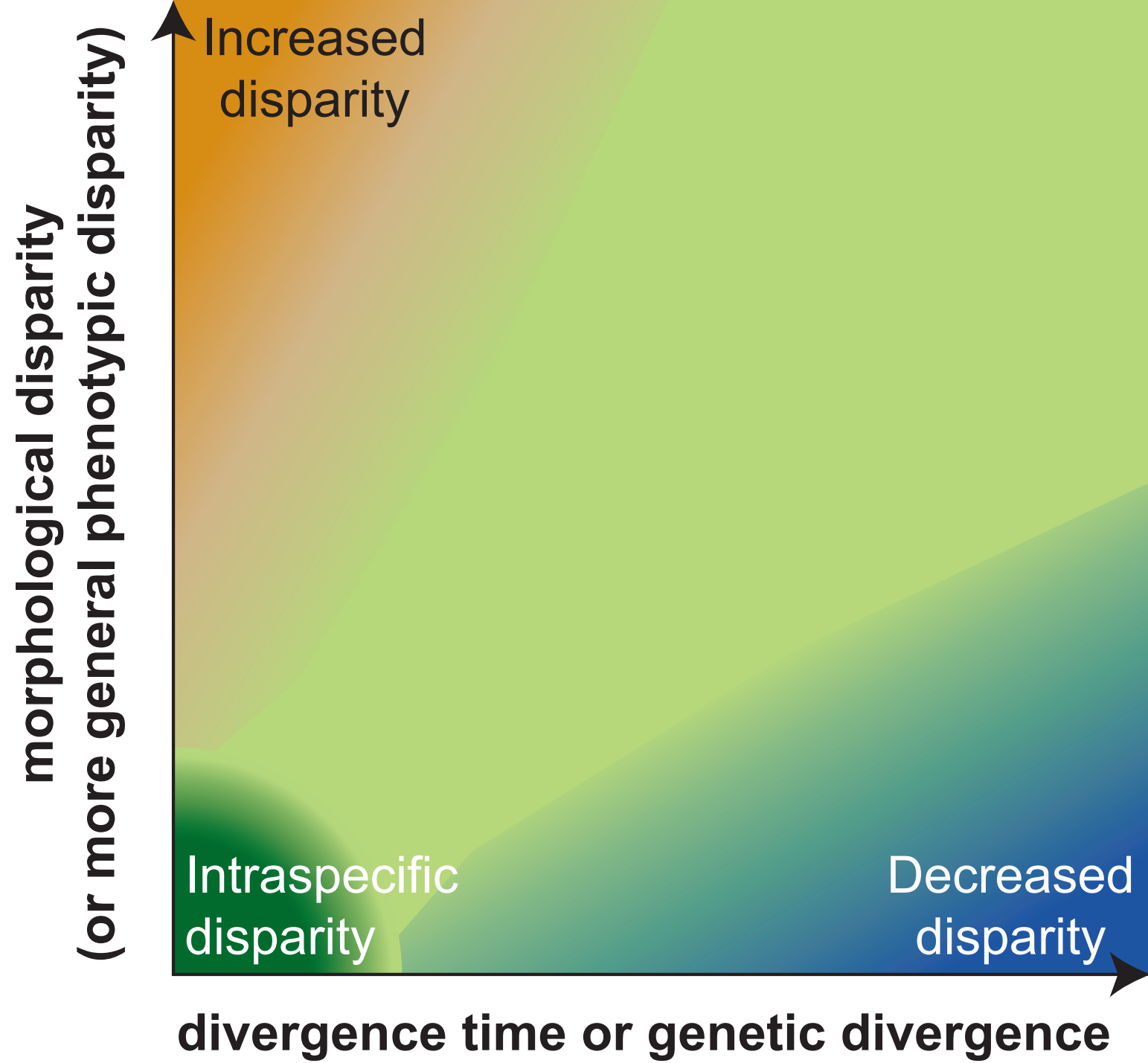


Figure 2

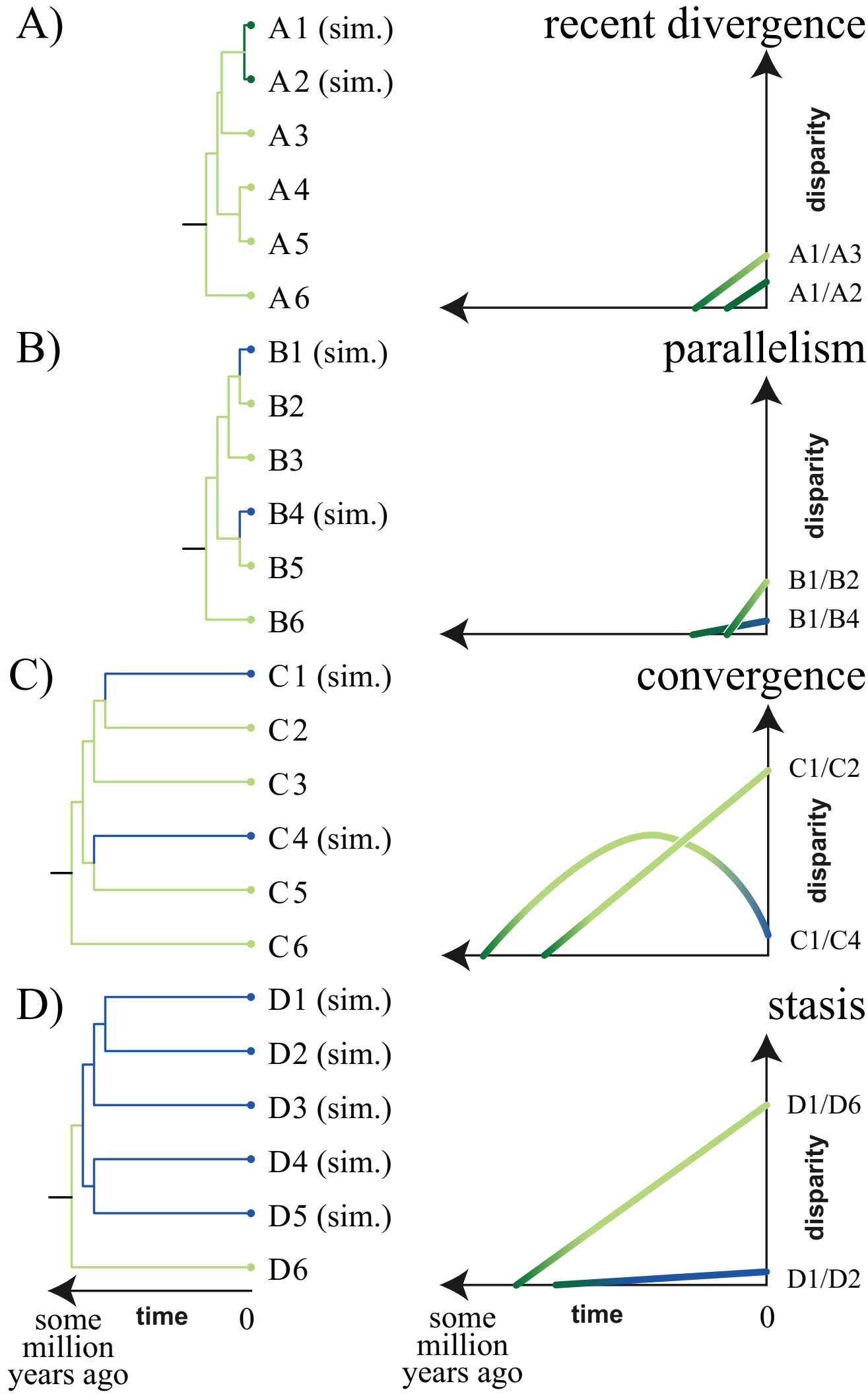


Figure I

