1   **Title: Phylogeny, ancestors and anagenesis in the hominin fossil record**

2   **Authors:** Caroline Parins-Fukuchi[1]*, Elliot Greiner[2], Laura M. MacLatchy[2,3], Daniel C.
3   Fisher[4,3,1]

4   **Affiliations:**

5   [1]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor,
6   Michigan, USA

7   [2]Department of Anthropology, University of Michigan, Ann Arbor, Michigan, USA

8   [3]Museum of Paleontology, University of Michigan, Ann Arbor, Michigan, USA

9   [4]Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor,
10   Michigan, USA

11   *Correspondence to: cfukuchi@umich.edu

12

15

16   **Author contributions:** CPF designed and implemented the ML stratocladistic procedure,

17   performed analyses, interpreted results, and drafted the manuscript. DCF made conceptual

18   contributions to the stratocladistic approach developed here, contributed to the manuscript, and

19   helped interpret results. EG gathered temporal occurrence ranges, helped interpret results, and

20   edited the manuscript. LMM helped focus the scope of the hominin analysis, helped interpret

21   results, and contributed to the manuscript.

22

23   **Supplementary data:** Software used in this study is implemented in the *mandos* package and is

24   available on GitHub (github.com/carolinetomo/mandos). All trees tested using the stratocladistic

25   likelihood procedure and their associated AIC scores are available on GitHub (WEB

26   ADDRESS—will upload prior to submission). This repository also contains the morphological

27   matrix and stratigraphic range files and all scripts needed to reproduce results. Stratigraphic

28   occurrences and their associated citations are provided as supplementary information. The

29   supplement also includes an online appendix that describes and tests the morphological

30   likelihood calculation on trees containing anagenetic relationships.

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49   **Abstract**: Probabilistic approaches to phylogenetic inference have recently gained traction in

50   paleontological studies. Because they directly model processes of evolutionary change,

51   probabilistic methods facilitate a deeper assessment of variability in evolutionary pattern by

52   weighing evidence for competing models. Although phylogenetic methods used in

53   paleontological studies have generally assumed that evolution proceeds by splitting cladogenesis,

54   extensions to previous models help explore the potential for morphological and temporal data to

55   provide differential support for contrasting modes of evolutionary divergence. Recent

56   methodological developments have integrated ancestral relationships into probabilistic

57   phylogenetic methods. These new approaches rely on parameter-rich models and sophisticated

58   inferential methods, potentially obscuring the respective contributions of data and models. In this

59   study, we describe a simple likelihoodist approach that combines probabilistic models of

60   morphological evolution and fossil preservation to reconstruct both cladogenetic and anagenetic

61   relationships. By applying this approach to a dataset of fossil hominins, we demonstrate the

62   capability of existing models to unveil evidence for anagenesis presented by morphological and

63   temporal data. This evidence was previously recognized by qualitative assessments, but largely

64   ignored by quantitative phylogenetic analyses. For example, we find support for directly

65   ancestral relationships in multiple lineages: *Sahelanthropus* is ancestral to later hominins;

66   *Australopithecus anamensis* is ancestral to *Au. afarensis*; *Au. garhi* is ancestral to *Homo*; *H.*

67   *antecessor* is ancestral to *H. heidelbergensis*, which in turn is ancestral to both *H. sapiens* and *H.*

68   *neanderthalensis*. These results show a benefit of accommodating direct ancestry in

69   phylogenetics. By so doing, quantitative results align more closely with previous qualitative

70   expectations.

71   **Keywords:** anagenesis, phylogenetics, morphology, paleontology, hominin

72 **Introduction:** Phylogenetic methods that rely upon probabilistic evolutionary models have

73 recently begun to emerge as important tools for addressing paleobiological issues at

74 macroevolutionary timescales (Wright and Hillis 2014; Puttick et al. 2017). Probabilistic

75 approaches appear promising in their performance relative to earlier cladistic methods and offer

76 several important advantages. These include their tendency to increase the clarity with which

77 models are specified and their ability to weigh evidence under competing evolutionary models

78 using modern inferential machinery. While probabilistic approaches have long been dominant in

79 molecular phylogenetics, their application to paleontological datasets is largely untapped. Yet

80 the extension of approaches developed for extant taxa to fossil taxa will allow for the testing of a

81 broader range of evolutionary hypotheses that are designed to reflect differences in the

82 morphological patterns underlying divergence between fossil taxa.

83 Most previous approaches to phylogenetic reconstruction depict evolutionary relatedness

84 using patterns of 'hypothetical' ancestry in which more closely related taxa are represented as

85 possessing more recent common (hypothetical) ancestors. However, when fossil taxa are

86 included in phylogenetic analyses, we might reasonably consider whether direct ancestry

87 between temporally disjunct taxa or samples can be detected (Foote 1996). Although there is a

88 spectrum of possible patterns of temporal distribution reflecting directly ancestral relationships,

89 these may be dichotomized into cladogenesis and anagenesis.

90 Characterization of the relative prevalence of cladogenesis and anagenesis in the fossil

91 record and their bearing on the study of evolutionary mode has long been a fundamental question

92 in paleobiology (Simpson 1944; Stanley 1998; Levinton 2001). Although the significance of

93 these contrasting models of evolutionary change in interpreting patterns in the fossil record has

94 been fiercely debated (Gould and Eldredge 1977; Gould 1980; Levinton and Chris 1980;

4

95   Gingerich 1985; Futuyma 1987), anagenetic change has been widely observed in the fossil

96   record and can impact inferences of evolutionary processes (MacLeod 1991). This may be

97   particularly important when constructing and testing hypotheses of phylogenetic relationships,

98   where accommodation of anagenetic change may improve accuracy and yield insights otherwise

99   unattainable (Gingerich 1979; Fox et al. 1999; Aze et al. 2011; Strotz and Allen 2013; Aze et al.

100   2013).

101   Quantitative phylogenetic approaches have usually assumed the universality of

102   cladogenesis. Until recently, this fundamental limitation has remained largely intact, despite the

103   development of increasingly sophisticated probabilistic approaches to phylogenetic inference.

104   Since most of the major developments in probabilistic approaches to phylogenetics were

105   developed for molecular data (but see Lewis 2001), their use has been largely restricted to

106   neontological applications. As a result, there has been relatively little impetus in the areas most

107   engaged in phylogenetic research to consider phyletic evolution, leaving the role of anagenesis in

108   phylogenetic reconstruction under-explored.

109   The conceptual issues discussed above are particularly relevant in taxa such as the

110   hominin clade (humans and all other taxa more closely related to humans than to genus *Pan*),

111   where hypotheses of direct ancestry are often proposed. Discerning human evolutionary patterns

112   is a focus traceable to Darwin (1871), and ancestry is almost certainly more frequently proposed

113   (even if informally) for clusters hominin fossils than for those from other taxonomic groups.

114   Nonetheless, there are few formal routes for recognizing this status based on quantitative

115   evaluations of traits.

116   Although quantitative approaches to phylogenetics have typically ignored anagenesis,

117   stratocladistic methods were developed to leverage temporal occurrence data to help reconstruct

5

118   phylogenetic trees that have the potential to express direct ancestor-descendant relationships

119   (Fisher 2008). Stratocladistics uses the criterion of maximum parsimony (MP) to minimize both

120   the number of homoplasious character changes and unsampled stratigraphic intervals implied by

121   topologies that include both bifurcating and serially linked segments of phyletic lineages. Several

122   authors have expressed objections both to the use of temporal data in phylogenetic inference and

123   the capability of available data to adequately test ancestral and collateral relationships between

124   species (Smith 2000). Nevertheless, direct ancestors are expected to occur in the fossil record

125   (Foote 1996), and the integration of temporal data has been shown to improve reconstruction

126   accuracy over morphological analyses alone (Fox et al. 1999). In addition, those earlier

127   criticisms, which have been frequently raised by proponents of cladistic methodologies, are also

128   less relevant when placed in the context of the modern probabilistic approaches. While cladistics

129   operates at the level of cladograms, both stratocladistics and recent probabilistic approaches

130   reconstruct phylogenetic trees. This renders earlier criticism of temporal data largely irrelevant

131   given the current landscape of phylogenetic methodology. As such, it stands to reason that

132   probabilistic approaches are remiss when they fail to accommodate the possibility of ancestor-

133   descendant relationships between temporally heterogeneous taxa.

134   There have been several attempts to explicitly extend the intent and logic of

135   stratocladistics into probabilistic approaches using maximum-likelihood (ML) (Huelsenbeck and

136   Rannala 1997; Wagner 1998). However, these have not found widespread use. Aspects of these

137   have been appropriated more recently by methods that use some combination of morphological,

138   fossil preservation, and lineage diversification models to infer relationships and lineage

139   divergence times in a Bayesian context (Pyron 2011). This framework has been extended to

140   explicitly accommodate ancestor-descendant relationships by modelling lineage diversification

141   and fossil preservation processes (Stadler 2010; Bapst and Hopkins 2016; Stadler et al. 2017), in

142   particular through the use of the 'Fossilized Birth-Death' (FBD) prior (Heath et al. 2014; Zhang

143   et al. 2016). However, the assumptions and ramifications of doing so have not been well

144   explored.

145   One outstanding question regarding these approaches is the extent to which

146   morphological data themselves are able to illuminate questions of direct ancestor-descendant

147   relationships without explicitly modelling the process of lineage diversification. This may be

148   especially important in clades with gappy fossil records, such as many lineages of terrestrial

149   vertebrates. In these cases, there may be less information available to reconstruct speciation and

150   extinction parameters, and so a characterization of the identifiability of hypotheses of direct

151   ancestry from morphology alone can benefit future studies in these taxa. In addition,

152   methodological extensions of Bayesian tip-dating methods that accommodate serial sampling of

153   lineages have been largely developed for use in epidemiological systems, where they are used to

154   model molecular sequence evolution along single lineages. Although these systems are useful

155   models of patterns in the fossil record, their sampling is often incomplete, and so they may

156   sometimes call for different approaches in practice. Non-Bayesian *a posteriori* time-scaling

157   (APT) approaches, such as *cal3*, that accommodate ancestor-descendant relationships also exist

158   (Bapst 2013). These apply divergence times to unscaled cladograms that have been inferred from

159   character data alone. Although both Bayesian and APT approaches have been shown effective

160   when applied to fossil taxa (Bapst 2013; Bapst and Hopkins 2016; Gavryushkina et al. 2017;

161   Stadler et al. 2017), there has been limited discussion of the statistical properties, limitations, and

162   implications of using existing paleontological data and probabilistic models to identify

163   anagenesis. Nevertheless, anagenesis is an evolutionary question of critical importance, and so

7

164   the development of a probabilistic approach using likelihoodist criteria may be broadly beneficial

165   by clearly quantifying the evidence for anagenesis present in paleontological data.

166   Although cladistic methods have been applied to improve understanding of certain

167   aspects of hominin evolution (Delson et al. 1977; Chamberlain and Wood 1987; Strait et al.

168   1997; Irish et al. 2013), their exclusive focus on bifurcating relationships has precluded the study

169   of direct ancestry except in cases where character polarity and specimen sampling have been

170   carefully considered (Kimbel et al. 2006). More recently, Bayesian approaches have been

171   suggested to provide a clearer view of hominin evolutionary relationships (Dembo et al. 2015;

172   2016), but their results still contain ambiguities. For example, like earlier cladistic approaches

173   (Strait et al. 1997; Strait 1999), results from Bayesian analyses conflict with qualitative

174   interpretations of hominin relationships in several key instances. This divide is exemplified by

175   considering the taxa *Homo heidelbergensis* and *Homo antecessor,* for whom hypotheses of

176   anagenesis are often considered. For example, previous studies have alternatively suggested that

177   either *Homo heidelbergensis* or *Homo antecessor* may be directly ancestral to *Homo sapiens* and

178   *Homo neanderthalensis* (Mounier et al. 2009), while some have debated the suggestion that

179   *Homo antecessor* is ancestral to *Homo heidelbergensis* (Stringer 2012). Others have suggested

180   alternately that *Homo heidelbergensis* is either a chronospecies directly ancestral to *H.*

181   *neanderthalensis*, or directly ancestral to both *H. sapiens* and *H. neanderthalensis* (Rightmire

182   1998; Rosas and Bermúdez De Castro 1998; Stringer 2012). This uncertainty is underscored by

183   the suggestion that cladistic methods and data are unreliable in their ability to describe hominin

184   evolutionary patterns and history (Collard and Wood 2000).

185   In this study, we describe an approach to phylogenetic inference that combines models of

186   stratigraphic preservation and morphological evolution to reconstruct time-scaled phylogenies

187   and distinguish between anagenesis and cladogenesis using ML and the Akaike Information

188   criterion (AIC). This approach seeks to simplify existing methods, such as the APT and FBD

189   approaches described above, to clearly identify the signal for directly ancestral relationships

190   presented by morphological data alone. Unlike more complex approaches, ours seeks only to

191   identify information in morphological and temporal data that establishes differential support for

192   ancestor-descendant and bifurcating relationships. As a result, our approach explicitly relies on

193   models of morphological evolution and stratigraphic preservation rather than on models of

194   lineage diversification.

195   We apply our new approach in hominins using a morphological supermatrix borrowed

196   from the literature (Dembo et al. 2015; 2016). Hominins are a particularly compelling case study

197   for our questions surrounding the identifiability of phylogenetic models that incorporate

198   anagenesis. Although cladistic and phylogenetic methods have been applied to hominins

199   (Chamberlain and Wood 1987; Strait et al. 1997; Strait 1999; Irish et al. 2013; Dembo et al.

200   2015; 2016), previous methods' exclusive focus on bifurcating relationships has led to their

201   inability to fully address outstanding hypotheses of hominin evolution, which are often

202   predicated upon the occurrence of anagenesis. Our analysis identified several areas of anagenetic

203   change in hominins, demonstrating cases where temporal data both corroborate and refute results

204   achieved using morphology alone. We recognize that increasingly thorough compilation of

205   morphological data and new fossil discoveries is likely to continue to revise and refine current

206   understanding of hominin evolution. Nevertheless, our approach sheds light on the capabilities of

207   existing models and data to accommodate anagenetic change without becoming entangled in a

208   complex web of Bayesian assumptions. Moving forward, application of our method in hominins

209   provides general demonstration of the importance of accommodating anagenetic change in

9

210   phylogenetic methods to generate deeper understanding of evolutionary patterns in the fossil

211   record.

212

213   **Materials and methods:**

214   *Inference of ML topology.* Our approach evaluates the likelihood of candidate topologies using

215   probabilistic models of fossil preservation and morphological evolution (Huelsenbeck and

216   Rannala 1997; Lewis 2001). We perform a semi-automated tree search by calculating the

217   likelihoods of these models on a set of candidate topologies. This approach tests hypotheses of

218   direct ancestorship by combining branches with non-overlapping ranges and comparing

219   cladogenetic and anagenetic models using the AIC. All code developed for these analyses is

220   publicly available and implemented in the *mandos* package

221   (www.github.com/carolinetomo/mandos).

222   *Identification of a fully-bifurcating ML tree.* Our approach has yet to be implemented with a fully

223   automated tree searching algorithm, so we combine semi-automated rearrangements with manual

224   perturbations to search for the ML topology. This is done by exploring tree-space surrounding a

225   starting tree estimated from morphological characters alone. In this study, we obtained a ML

226   starting tree using RAxML, version 8.2.11, using the Mk model of morphological evolution

227   (Lewis 2001). The morphological data were separated into partitions according to the number of

228   possible states (i.e., binary, trinary, etc.) and analyzed under separate models. This partitioning

229   scheme was maintained for all subsequent morphological likelihood calculations, including those

230   used below in the AIC comparisons. We then performed a series of nearest neighbor interchange

231   (NNI) and subtree pruning and regrafting (SPR) operations. These yielded a set of 1700

232   candidate topologies from which we identified the fully-bifurcating topology best supported by

10

233   both morphologic and stratigraphic data. This tree provided a starting point drawn from both

234   morphologic and stratigraphic lines of evidence. From here, we explored ancestor-descendant

235   relationships using the model-testing approach described below.

236   *Modelling stratigraphic preservation*. Stratigraphic likelihoods were calculated under a

237   homogeneous Poisson process of geologic preservation. When applied to phylogenetics,

238   inference under this model has been shown to accurately recover simulated phylogenetic

239   relationships (Huelsenbeck and Rannala 1997). The likelihood function is derived from a

240   Poisson process as the probability of observing the first occurrence (o_f), last occurrence (o_l) and

241   number (n_o) of occurrences in the stratigraphic record given some origination and extinction time

242   (t_f and t_l, respectively) and preservation rate (λ). These likelihoods are calculated independently

243   for each (i) of b total lineages and multiplied to yield the overall tree likelihood (see

244   (Huelsenbeck and Rannala 1997) for full derivation):

$$\hat{L} = \prod_{i=1}^{b} \frac{\left(o_f^i - o_f^i\right)^{n_o^i-2} \lambda^{n_o^i} e^{-\lambda\left(t_f^i - t_l^i\right)}}{\left(n_o^i - 2\right)!}$$

(Eq. 1)

246   This equation reaches its maximum as branching and extinction times approach the first and last

247   occurrences in the fossil record, and so the likelihood is maximized across the tree when the total

248   amount of unsampled time implied by the topology is minimized. Although this approach differs

249   from stratocladistic parsimony in its treatment of occurrence data as continuous rather than

250   discrete, this property causes the preservation model to behave as a statistical formalization of

251   the stratigraphic parsimony debt calculations undertaken in parsimony-based stratocladistic

252   analyses. We estimate lineage origination and termination times along with preservation rate

253   using multivariate numerical optimization routines implemented in SciPy (Jones et al. n.d.).

11

254   When combined with morphological data and models of character evolution, this approach

255   represents a comprehensive extension of traditional stratocladistics based entirely upon

256   probabilistic models.

257   *Identification of anagenesis*. Using the bifurcating topology with the highest likelihood under

258   both stratigraphic and morphologic models, we manually identified a set of potential ancestor-

259   descendant relationships. To explore a more comprehensive range of both anagenetic and

260   cladogenetic arrangements, we also extensively perturbed results manually and compared

261   likelihoods. Although a fully-automated tree searching approach will ultimately be desirable in

262   future versions of our method, our approach to tree-searching is similar to those used in previous

263   stratocladistic studies. Starting with the fully bifurcating tree, we identified putative ancestor-

264   descendant arrangements by collapsing each branch with a temporal range beginning earlier than

265   the range represented by its sister lineage. We isolated each putative episode of anagenesis and

266   compared the morphologic and stratigraphic likelihoods of anagenetic and cladogenetic

267   arrangements using AIC scores. This was required because cladogenetic nodes assume one more

268   parameter than anagenetic nodes (i.e., the branch length of the new lineage), and so bifurcating

269   trees contain more parameters than anagenetic trees. Comparison of AIC scores enables a

270   comparison of the relative quality of models with different numbers of parameters. AIC score is

271   calculated from the number of model parameters (k), and the log-likelihood (L):

272
$$AIC = 2k - 2\ln L$$

(Eq. 2)

273   Since models with larger numbers of parameters are more prone to overfitting, they are biased

274   toward possessing higher likelihoods than less complex models. As a result, likelihood scores of

12

275   models that differ in dimension are not directly comparable. AIC scores represent the amount of

276   information lost by a model when representing data, with lower scores indicative of models that

277   preserve a greater amount of information. AIC accommodates for differences in parameter count

278   between models by penalizing the addition of additional parameters, seeking to optimize the

279   trade-off between improvements in model fit associated with added parameters and the loss of

280   statistical power that results from overparameterization. In our use, AIC facilitates comparison of

281   phylogenetic models of differing dimensions by penalizing the addition of branches that are

282   better explained through an anagenetic pattern. The parameter count, k, for each phylogeny is

283   calculated by summing the number of estimated branch lengths for each tree with the number of

284   parameters used in the partitioned Mk substitution model.

285   Under the Poisson preservation model, stratigraphic likelihood predictably improves

286   when unsampled time implied by the phylogeny is reduced, and so the acceptance of anagenetic

287   arrangements also requires the support of morphology. This requires a novel calculation of

288   morphological likelihood where the probability of transitioning from an observed, rather than an

289   uncertain, parental character state to a single or multiple descendant character states is calculated

290   under the Mk model (Fig. 1). This calculation differs from that used on multi-furcating nodes.

291   Since the sequences at internal nodes representing unobserved taxa are unknown, the likelihood

292   of character data at the tips is typically calculated by summing over all possible states at each

293   unobserved internal node (Felsenstein 1981). However, when dealing with anagenesis, the sets of

294   character states at some internal nodes are known. In these cases, the likelihood is calculated as

295   the conditional probability of observing the set of traits at the tips given the set of traits

296   possessed by the putative ancestor. Anagenetic arrangements are only accepted when there has

297   been a sufficiently small amount of character change. This procedure resembles model testing

13

298   procedures used to reconstruct ancestral DNA sequences, which compare conditional likelihoods

299   of different permutations of character states at ancestral nodes (Yang et al. 1995). Like

300   stratocladistics, this use of probabilistic models of character change enables morphology to

301   occupy a central role in identifying anagenesis. As a result, this approach can in principle be used

302   to test anagenesis even in the absence of a temporal model. A demonstration and test of this

303   method using simulated data is provided in the supplement.

304   *Morphological matrix.* We performed our analysis on a supermatrix of 391 discrete craniodental

305   characters compiled by Dembo and colleagues (Dembo et al. 2015; 2016). We removed all

306   ambiguous character states, as researchers did not identify whether these were truly ambiguous

307   or polymorphic. While ambiguous character codings are unlikely to provide significant

308   phylogenetic information, existing Markov models of discrete character evolution do not

309   accommodate polymorphism. We excluded the taxa *Kenyanthropus platyops* and *Homo naledi*

310   from the present analysis. The features that are diagnostic of *K. platyops* have been suggested to

311   result from taphonomic distortion resulting from matrix expansion, rather than from true

312   biological differences (White 2003). Thus, we omitted this taxon in hopes of shedding greater

313   light on the remaining, more widely accepted hominin taxa. *Homo naledi* was omitted because

314   the data provided in the original study yielded an ML topology placing *H. naledi* as sister to *H.*

315   *sapiens*. Although the phylogenetic affinity of *H. naledi* is a major outstanding question in

316   paleoanthropology, the confusing signal presented by the *H. naledi* data, which are relatively

317   recently acquired and therefore represent less well-studied fossils overall, reduced our

318   confidence in the ability of this dataset to resolve its placement. Therefore, to avoid any

319   confounding effects from the potential unreliability of the *H. naledi* data, we performed our

14

320   analyses on the remaining subset of the data after *H. naledi* was removed. This enabled us to

321   explore the phylogenetic relationships between better-known hominins.

322   *Geologic occurrence times.* We surveyed the literature to obtain the observed temporal range of

323   each taxon in continuous time. Reported radiometric dates for the oldest and youngest fossils

324   were taken as the first and last observations. Some specimens are ambiguous in their taxonomic

325   assignment; these were excluded from the analysis. We also gathered the number of total

326   occurrences as the number of localities where each taxon has been identified as listed in

327   MacLatchy et al. (2010), and supplemented these with additional localities identified in the

328   literature. Cases where multiple specimens belonging to the same taxon have been identified at a

329   single locality were treated as single occurrences. Although we recognize the potential ambiguity

330   in delineating between sites, localities, and occurrences, we attempted to coarsely characterize

331   the total number of occurrences using the number of sites at which each taxon occurs. This

332   approach is more likely to underestimate the number of occurrences than overestimate them,

333   which we expect to yield more conservative statistical support for competing topologies under

334   the preservation model. A comprehensive list of the sites used to define temporal ranges for all

335   taxa is provided in the supplement.

336

337   **Results and Discussion:**

338   *Anagenesis in the hominin fossil record:* Our analysis yielded evidence for several instances of

339   anagenesis in the hominin fossil record (Fig. 2). Our analysis reconstructed *Australopithecus*

340   *anamensis* as directly ancestral to *Au. afarensis*. This result agrees with broad acceptance of *Au.*

341   *anamensis* and *Au. afarensis* as phyletically linked chronospecies (Leakey et al. 1995; Ward et

342   al. 2001; Kimbel et al. 2006). Although our analysis recovered *Ar. ramidus* as sister to the

15

343   anagenetic *Au. anamesis-Au. afarensis* branch, the morphological data did not support the

344   collapse of *Ar. ramidus*. Nevertheless, *Ar. ramidus* possessed poor character sampling in the

345   matrix, and so its placement should be regarded as tentative. *Sahelanthropus tchadensis* is

346   recovered as a direct ancestor to the rest of the hominin clade. This result should also be treated

347   cautiously due to the small number of characters recovered for analysis of this portion of the

348   phylogeny, but it is in line with with *Sahelanthropus'* status as the oldest recognized hominin

349   (Brunet et al. 2002; Guy et al. 2005; Zollikofer et al. 2005).

350   Consistent with an early appraisal (Asfaw et al. 1999), our final analysis inferred *Au.*

351   *garhi* to be directly ancestral to the *Homo* clade (Fig.2a). This conflicts with cladistic analyses

352   that placed *Au. garhi* as outgroup to *Au. africanus*, *Paranthropus*, and *Homo* (Strait 1999).

353   However, when anagenesis is not considered and phylogeny is inferred from morphology alone,

354   we recover the same placement for *Au. garhi* as the cladistic result (Fig. 2b). Like the example

355   above, this may reflect the constraint that strictly bifurcating methods impose on phylogenetic

356   reconstructions among fossil taxa. However, preference for *Au. garhi* as ancestral to *Homo* is

357   weak, with an only slightly improved AIC score over the *Au. garhi* outgroup hypothesis (Table

358   1).

359   Our results also better reconcile quantitative and qualitative interpretations of the

360   evolutionary mode and relationships among the *Homo* species leading to modern humans and

361   Neanderthals. Paleoanthropologists have variously interpreted *Homo heidelbergensis* as either:

362   1) a direct ancestor to both Neanderthals and modern humans (Rightmire 1998; Mounier et al.

363   2009; Stringer 2012; Buck and Stringer 2014), or 2) a chronospecies leading to Neanderthals

364   (Rosas and Bermúdez De Castro 1998). However, quantitative phylogenetic analysis has placed

365   *H. heidelbergensis* as sister to Neanderthals (Dembo et al. 2015). Our ML result based on

16

366   morphological characters alone places *H. heidelbergensis* as sister to *H. sapiens*. However, when

367   temporal data are incorporated and anagenesis is accommodated, AIC support improves

368   substantially*, and H. heidelbergensis* is collapsed to represent a direct ancestor of modern

369   humans and Neanderthals (Fig. 3). We also uncover one other instance of anagenesis in this

370   clade. Opinions are divided as to whether *H. antecessor* represents a direct ancestor of later

371   hominin species or is an evolutionary dead end (Bermúdez de Castro et al. 1997; Stringer 2012;

372   Dembo et al. 2015), but our results provide support for combining *H. antecessor* and *H.*

373   *heidelbergensis* into a single lineage. This suggests a long episode of anagenetic evolution

374   immediately prior to the divergence between modern humans and Neanderthals.

375   Overall, our reconstruction of relationships among later species of *Homo* immediately

376   preceding and encompassing modern humans and Neanderthals supports the hypothesis that *H.*

377   *heidelbergensis* is directly ancestral to both modern humans and Neanderthals. This result

378   differs from the hypothesis supported by the analysis performed by Dembo and colleagues

379   (Dembo et al. 2015), and is instead more consistent with an earlier exploratory statistical analysis

380   (Mounier et al. 2009), and with the position frequently suggested by paleoanthropologists

381   (Rightmire 1998; Stringer 2012; Buck and Stringer 2014). While previous phylogenetic analyses

382   have yielded results that equivocate or disagree with the common interpretation of *H.*

383   *heidelbergensis* as the last common ancestor of modern humans and Neanderthals, our analysis

384   shows that the consideration of direct ancestry can generate statistical support for phylogenetic

385   results that conform more closely to positions generated through researchers' subjective

386   interpretations and exploratory data analyses. This finding supports a general argument against

387   the use of cladistic and phylogenetic methods that are restricted to bifurcating relationships in

388   fossil taxa, where the possibility of variability in evolutionary mode (i.e., occurrence of both

17

389   anagenesis and cladogenesis) is at odds with an assumption that evolution proceeds by

390   cladogenesis alone (Fig. 2).

391   Our results differ markedly from previous phylogenetic studies seeking to reconstruct

392   hominin phylogeny using probabilistic and cladistic methods. In key regions of the tree, results

393   achieved under our method reveal support for hypotheses more consistent with many qualitative

394   interpretations of hominin relationships, demonstrating the importance of explicitly

395   accommodating anagenesis in the phylogenetic reconstruction of fossil taxa. This may explain

396   some of the historical difficulty in reconciling paleontological interpretations of hominin

397   relationships with cladistic results. For instance, Dembo and colleagues' results are inconsistent

398   with earlier suggestions that *Au. anamensis* and *Au. afarensis* are chronospecies differentiated

399   through anagenesis. However, by considering ancestor-descendant relationships and

400   incorporating temporal data, our analysis reveals that a linked *Au. anamensis-Au. afarensis*

401   lineage is the arrangement most strongly supported by the data. Generally, this suggests that

402   cladograms and strictly bifurcating phylogenies may be inadequate when describing evolutionary

403   relationships between fossil taxa, and so ancestor-descendant relationships should be considered

404   during topological inference. Further, we argue that explicit testing of ancestor-descendant

405   relationships is important even in cases where bifurcating trees are not wholly misleading, as

406   their omission precludes us from considering the full range of possible evolutionary scenarios.

407   Although previous studies addressing hominin phylogeny using probabilistic methods represent a

408   significant step forward in weighing alternative evolutionary hypotheses, we suggest that their

409   phylogenetic reconstructions have suffered from methodological limitations that were not

410   generally perceived. We suggest that overcoming these limitations can provide a substantial step

411   forward in closing the gap between the paleobiologists' interpretations and previous cladistic and

18

412   phylogenetic results. In particular, we show through our analyses that the apparent discordance

413   between quantitative and qualitative assessments of evolutionary relationships can be reconciled

414   by extending phylogenetic models to explicitly accommodate anagenesis.

415

416   *Ancestors, anagenesis, and evolutionary processes.* Our method does not seek to distinguish

417   between speciation modes at a mechanistic level. As noted by Fisher (2008), overlap between the

418   extinction and origination times of taxon pairs does not necessarily preclude an ancestor-

419   descendant relationship. Instead, it is possible for ancestral and descendant lineages to coexist.

420   This reality may complicate the identification and evolutionary interpretation of ancestor-

421   descendant relationships from temporal data alone. Doing so requires diversification and

422   preservation models that contain additional parameters that quantify completeness of the fossil

423   record. Such models are currently implemented in the *cal3* time-scaling method, which seeks to

424   distinguish between splitting, budding, and anagenesis from a cladogram with a fixed topology

425   (Bapst and Hopkins 2016). Our approach is distinct both theoretically and operationally from

426   *cal3*, instead relying most heavily on morphological evidence to weigh the likelihood of

427   ancestor-descendant relationships without considering the completeness of fossil sampling. At

428   the population level, one might expect some temporal overlap between ancestral and descendant

429   taxa undergoing anagenetic change, especially if the participants are widely distributed

430   geographically. As a result, even if the temporal ranges corresponding to taxa identified as

431   ancestor-descendant pairs are discovered to slightly underestimate the time of coexistence, the

432   inference may simply highlight the fuzziness in the taxonomic placement of fossils belonging to

433   lineages undergoing continuous transformation and in discerning between anagenesis and

434   evolutionary budding given incomplete sampling (Fig. 4). This interpretation is consistent with

19

435   previous authors' treatment of temporal ranges when identifying anagenesis between taxa, which

436   have allowed a period of overlap between putative ancestor-descendant pairs (Aze et al. 2011;

437   Strotz and Allen 2013).

438   We use the term 'ancestor' modelled loosely after Gingerich (1979). Ancestors identified

439   through our method represent collections of samples possessing a suite of morphological

440   character states that is not sufficiently differentiated from a single or a set of subsequently

441   occurring samples to warrant assignment to distinct lineages. In our usage, we consider

442   anagenesis as any evolutionary change occurring along these serially linked phyletic lineage

443   segments. Thus, the taxonomic and phylogenetic units of analysis are fundamentally important to

444   the formulation and interpretation of our results. Hominin taxa are often represented by only one

445   or a small number of samples, and the number of species may be overestimated (White 2003).

446   This may complicate the ability to meaningfully characterize phylogenetic relationships from a

447   precise mechanistic view. Nevertheless, morphological and temporal data can be combined to

448   summarize the relative support for anagenetic and cladogenetic patterns in the inheritance of

449   observed character states (i.e. even in cases where distinctiveness may not be at the species

450   level).

451   Under our method, ancestor-descendant relationships might be interpreted either as true

452   anagenesis (i.e., a single population undergoing gradual transformation), or as some form of

453   budding cladogenesis. Previous researchers have argued that true anagenesis is rare compared to

454   budding when analyzing the fossil records of densely-sampled marine invertebrate lineages using

455   more complex preservation models (Bapst and Hopkins 2016). Nevertheless, we suggest that

456   distinction between these two modes may often be impossible in terrestrial vertebrate lineages

457   with large sampling gaps. For example, our results among early hominin species include multiple

458   inferred direct ancestors, but the large gaps in stratigraphic sampling throughout this region of

459   the tree hamper the ability to determine whether these relationships represent true anagenesis or

460   budding that has been obfuscated by poor sampling.

461   Instead of relying upon hope that more complex preservation models are robust to wide

462   sampling gaps, our method disregards the completeness of sampling to evaluate the support for

463   ancestor-descendant pairs using morphology. Although temporal data occupy an important place

464   in our approach, they are largely used as a guide to constrain the set of possible ancestors and

465   descendants and to provide additional insight when morphological data are equivocal. Coarsely

466   speaking, our approach focuses on patterns in morphological differentiation and lineage

467   *disparification*, while approaches such as *cal3* model lineage *diversification*. Further empirical

468   and simulation-based work is needed to determine the relative merits of methods of testing

469   ancestor-descendant relationships, such as *cal3*, that explicitly focus on diversification patterns,

470   and ours, which weighs morphological patterns more heavily. We speculate that their relative

471   accuracy may depend largely on the completeness of sampling in the rock record and the

472   correlation strength between morphological change and lineage diversification, although other

473   factors may also be important.

474   The scales at which phylogenetic data are sampled may further complicate mechanistic

475   evolutionary interpretations. Morphological character matrices often lack samples across the

476   entire stratigraphic range of each taxon, and so generally assume morphological stasis within

477   lineages. It is therefore often impossible to observe gradual morphological change within and

478   between taxa. These considerations might cause anagenetic relationships identified here to

479   represent either true anagenesis or some form of 'pseudo-anagenesis', where stratigraphic and

480   morphological data appear consistent with anagenesis but the persistence of the ancestor has not

21

481   been sampled. The ancestor-descendant relationships identified by our method may be

482   interpreted in several ways. As sampled, these results may be roughly conceived as anagenesis in

483   the sense that the mode of evolutionary change between taxa is indistinguishable from evolution

484   occurring along a single lineage, depending upon the completeness of sampling and the degree to

485   which morphological disparity correlates with true biological species diversity. This

486   interpretation is consistent with historical usage by paleobiologists (Gingerich 1979; Smith 1994;

487   Levinton 2001). Thus, our approach seeks to reveal the extent to which existing cladistic and

488   temporal data can provide evidence for non-branching evolutionary modes and does not seek to

489   resolve conceptual issues that may stem from incomplete sampling, lineage diversification, or

490   population-level evolutionary change. Regardless of the fine-scale evolutionary interpretation,

491   failure to accommodate phyletic change and ancestor-descendant relationships when inferring

492   phylogenetic relationships can generate views of evolutionary history that are positively

493   misleading in the sense that incorrect results cannot improve with the addition of new data.

494   *Some practical methodological considerations.* Concerns regarding the accuracy of probabilistic

495   approaches have been raised, stemming from the reliance of these methods on the overly

496   simplistic Lewis Mk model of morphological evolution (Goloboff et al. 2017). These critics

497   advocate the use of cladistic methods, arguing that Markov models inadequately capture the

498   complexities of morphological evolution. Although we agree that existing substitution models

499   oversimplify these processes, our results suggest that the accommodation of ancestor-descendant

500   hypotheses in probabilistic methods can improve the fidelity of phylogenetic reconstructions,

501   even when Lewis Mk is used as the underlying model of morphological change. As a result,

502   concerns regarding the adequacy of existing morphological substitution models may be partially

503   alleviated by considering hypotheses of direct ancestorship. This is supported by simulation

22

504   work showing that stratocladisics outperforms cladistics in topology reconstruction (Fox et al.

505   1999). Further exploration is needed to demonstrate more thoroughly the limitations of our new

506   approach, which builds upon stratocladistics by incorporating the benefits of probabilistic

507   analyses, including 1) more explicit statements of the assumptions involved, and 2) the ability to

508   weigh competing models using modern inferential criteria.

509   The method we describe seeks to enhance understanding of the fossil record by explicitly

510   testing support for existing hypotheses of direct ancestorship while attempting to make more

511   explicit assumptions than stratocladistics or recently developed Bayesian methods. Although it

512   shares features with recent methods that use mechanistic evolutionary models in a Bayesian

513   context to infer ancestral relationships (Zhang et al. 2016), our approach is intended as a

514   foundational, minimally complex framework for exploring the behavior of probabilistic models

515   to evaluate support for anagenesis in temporal and morphological data. As such, our method

516   should be viewed as a complement to, rather than a simplification of, existing Bayesian

517   approaches. That is, our method encourages examination of the informativeness of the data

518   without the increased complexity of assessing prior probabilities. Our method further differs

519   from both existing Bayesian and parametric APT approaches by explicitly omitting

520   diversification parameters and instead placing morphological data in a central role when

521   evaluating hypotheses of direct ancestry. In doing so, temporal data help to delineate the set of

522   possible ancestors and play an important role in measuring the fit of candidate trees to the

523   observed stratigraphic record. Our method does not seek to reconstruct diversification processes,

524   and instead focuses on identifying hypotheses that best describe only the information contained

525   within morphological and temporal datasets. Assessment of information contained within

526   datasets and tests of hypotheses can also be achieved using Bayesian approaches (Lewis et al.

527   2016), but likelihoodist approaches such as ours streamline these procedures by reducing

528   complications presented by prior probabilities and Bayesian Markov-chain Monte Carlo

529   sampling.

530   Although Bayesian methods can be beneficial in certain circumstances, our method

531   simplifies identification of anagenetic hypotheses using evolutionary and stratigraphic models.

532   We observe that the likelihood surface surrounding certain nodes may possess low peaks (Table

533   1), which likely results from sparse sampling and relatively low information in the

534   morphological characters. Since Bayesian approaches often average results across this surface, it

535   is possible that they may fail to capture those relationships best supported by the data by

536   including information from weakly supported hypotheses. This is of greater concern in

537   paleontological than neontological data because the increased abundance of molecular data is

538   often likely to result in more clearly defined peaks in the likelihood surface. In cases where

539   information is sparse, likelihoodist approaches such as ours offer the benefit of filtering through

540   noisy and equivocal signal to reveal the hypothesis best supported by the data. Although these

541   benefits are also achievable through Bayesian approaches, additional caution must be taken to

542   select priors that do not dominate weakly informative data. In addition, careful thought should be

543   given when summarizing the posterior/likelihood surface. Averaging across a relatively flat

544   surface might yield poor results (Yang and Zhu 2018), while the comparison of individual point

545   estimates, as is done here, may more clearly shed light on models best supported by the data

546   while clearly capturing their relative support.

547   *How can we proceed?* Anagenesis and splitting cladogenesis were most pertinent to our analysis

548   of hominin evolution, and the straightforward dichotomy between these simplified the

549   assumptions and interpretation of our tests. Our approach may need to explicitly accommodate

24

550   budding before being applied to groups with very dense fossil records. However, the assumptions

551   required, which may often include morphological stasis within lineages, may make application to

552   some phylogenetic datasets impractical. This is especially true in cases such as hominins, for

553   which the fossil record implies large sampling gaps, and characters representative across

554   stratigraphic ranges are often sampled from only a single individual. Expansions of our approach

555   through implementation of new models will further test the implications of existing

556   paleontological datasets for reconstructing complex evolutionary and geologic processes over

557   deep timescales. We hope that the example provided here will encourage integration of more

558   diverse evolutionary modes into phylogenetic methods yielding better explanations of temporal

559   patterns in critical parts of life's history.

560   As we emphasize above, the approach described here should be viewed as an attempt to

561   explore the capability of phylogenetic methods to identify the signal of anagenesis using existing

562   models to interrogate morphological and stratigraphic data. In doing so, we acknowledge that

563   there are many complicated biological and geological factors that could be incorporated into this

564   framework. For instance, previous researchers have accommodated heterogeneity in fossil

565   preservation rates across time and among lineages (Foote 1997; 2001; Gavryushkina et al. 2014).

566   There have also been several concerns raised in the literature regarding the adequacy of existing

567   models of discrete trait evolution to inform complex evolutionary scenarios (Goloboff et al.

568   2017; Brown et al. 2017). Alternative models that use continuous characters may help to improve

569   some of these issues (Felsenstein 1988; Goloboff et al. 2006; Parins-Fukuchi 2017). Moving

570   forward, elaborations making use of new data sources and models will only continue to improve

571   resolution of evolutionary patterns in the fossil record.

572   Finally, we acknowledge that our empirical results beg qualification. In particular, we

573   expect that future studies will generate a more comprehensive and authoritative view of hominin

574   evolution as improved data continue to become available. For instance, although we are currently

575   cautious about making strong statements concerning the ancestral position of *Sahelanthropus*

576   using this dataset, additional information may resolve this issue. This may be the case for several

577   other areas of the hominin tree, which may be better resolved as temporal and taxonomic gaps in

578   sampling are better filled by new discoveries. In addition, we concede the possibility that more

579   comprehensive automated tree searching routines may reveal support for hypotheses that we

580   failed to consider under our semi-automated approach. Therefore, instead of providing an

581   authoritative view of hominin evolution, our study provides a springboard for future studies by

582   showing that the accommodation of anagenesis can improve our view of the processes and

583   relationships underpinning the evolution of fossil taxa. Nevertheless, due to the improved

584   support and congruence of hypotheses that explicitly consider ancestor-descendant relationships,

585   we recommend that future phylogenetic studies in hominins avoid methods that only consider

586   bifurcating relationships. Future studies that build upon existing work in other taxa will also be

587   important to better characterize the extent to which this suggestion can be generalized across the

588   tree of life. Although the accommodation of directly ancestral relationships is especially relevant

589   in hominin taxa, for which hypotheses of anagenesis have been long entertained through

590   qualitative anatomical assessment, these results may also be important in other taxa. Further

591   empirical work will be needed to develop a better understanding of the extent to which the

592   consideration of direct ancestors can improve resolution of evolutionary patterns throughout the

593   fossil record.

594

References

Asfaw, B., T. White, O. Lovejoy, B. Latimer, S. Simpson, and G. Suwa. 1999. Australopithecus garhi: a new species of early hominid from Ethiopia. Science 284:629–635. American Association for the Advancement of Science.

Aze, T., T. H. G. Ezard, A. Purvis, H. K. Coxall, D. R. M. Stewart, B. S. Wade, and P. N. Pearson. 2011. A phylogeny of Cenozoic macroperforate planktonic foraminifera from fossil data. Biol. Rev. 86:900–927.

Aze, T., T. H. G. Ezard, A. Purvis, H. K. Coxall, D. R. M. Stewart, B. S. Wade, and P. N. Pearson. 2013. Identifying anagenesis and cladogenesis in the fossil record. Proc. Natl. Acad. Sci. U.S.A. 110:E2946–E2946.

Bapst, D. W. 2013. A stochastic rate-calibrated method for time-scaling phylogenies of fossil taxa. Methods in Ecology and Evolution 4:724–733.

Bapst, D. W., and M. J. Hopkins. 2016. Comparing cal3 and other a posteriori time-scaling approaches in a case study with the pterocephaliid trilobites. Paleobiology 43:49–67. Cambridge University Press.

Bermúdez de Castro, J. M., J. L. Arsuaga, E. Carbonell, A. Rosas, I. Martínez, and M. Mosquera. 1997. A hominid from the lower Pleistocene of Atapuerca, Spain: possible ancestor to Neandertals and modern humans. Science 276:1392–1395.

Brown, J. W., C. Parins-Fukuchi, G. W. Stull, O. M. Vargas, and S. A. Smith. 2017. Bayesian and likelihood phylogenetic reconstructions of morphological traits are not discordant when taking uncertainty into consideration: a comment on Puttick et al. Proc. Roy. Soc. B 284:20170986. The Royal Society.

Brunet, M., F. Guy, D. Pilbeam, H. T. Mackaye, A. Likius, D. Ahounta, A. Beauvilain, C. Blondel, H. Bocherens, J.-R. Boisserie, L. De Bonis, Y. Coppens, J. Dejax, C. Denys, P. Duringer, V. Eisenmann, G. Fanone, P. Fronty, D. Geraads, T. Lehmann, F. Lihoreau, A. Louchart, A. Mahamat, G. Merceron, G. Mouchelin, O. Otero, P. P. Campomanes, M. P. De Leon, J.-C. Rage, M. Sapanet, M. Schuster, J. Sudre, P. Tassy, X. Valentin, P. Vignaud, L. Viriot, A. Zazzo, and C. Zollikofer. 2002. A new hominid from the Upper Miocene of Chad, Central Africa. Nature 418:145–151. Nature Publishing Group.

Buck, L. T., and C. B. Stringer. 2014. Homo heidelbergensis. Current Biology 24:R214–R215.

Chamberlain, A., and B. Wood. 1987. Early hominid phylogeny. J. Hum. Evol. 16:119–133.

Collard, M., and B. Wood. 2000. How reliable are human phylogenetic hypotheses? Proc. Natl. Acad. Sci. U.S.A. 97:5003–5006. National Acad Sciences.

Darwin, C. 1871. *The Descent of Man, and Selection in Relation to Sex*. 1st ed. John Murray, London, UK.

630   Delson, E., N. Eldredge, and I. Tattersall. 1977. Reconstruction of hominid phylogeny: A
631   testable framework based on cladistic analysis. J. Hum. Evol. 6:263–278.

632   Dembo, M., D. Radovčić, H. M. Garvin, M. F. Laird, L. Schroeder, J. E. Scott, J. Brophy, R. R.
633   Ackermann, C. M. Musiba, D. J. de Ruiter, A. Ø. Mooers, and M. Collard. 2016. The
634   evolutionary relationships and age of Homo naledi: An assessment using dated Bayesian
635   phylogenetic methods. J. Hum. Evol. 97:17–26.

636   Dembo, M., N. J. Matzke, A. Ø. Mooers, and M. Collard. 2015. Bayesian analysis of a
637   morphological supermatrix sheds light on controversial fossil hominin relationships. Proc. Roy.
638   Soc. B 282:20150943.

639   Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach.
640   J. Mol. Evol. 17:368–376.

641   Felsenstein, J. 1988. Phylogenies and Quantitative Characters. Annual Review of Ecology and
642   Systematics 19:445–471.

643   Fisher, D. C. 2008. Stratocladistics: Integrating Temporal Data and Character Data in
644   Phylogenetic Inference. Annual Review of Ecology, Evolution, and Systematics 39:365–385.

645   Foote, M. 1997. Estimating taxonomic durations and preservation probability. Paleobiology
646   23:278–300.

647   Foote, M. 2001. Inferring temporal patterns of preservation, origination, and extinction from
648   taxonomic survivorship analysis. Paleobiology 27:602–630.

649   Foote, M. 1996. On the probability of ancestors in the fossil record. Paleobiology 22:141–151.

650   Fox, D. L., D. C. Fisher, and L. R. Leighton. 1999. Reconstructing Phylogeny with and without
651   Temporal Data. Science 284:1816–1819. American Association for the Advancement of Science.

652   Futuyma, D. J. 1987. On the Role of Species in Anagenesis. Am. Nat. 130:465–473.

653   Gavryushkina, A., D. Welch, T. Stadler, and A. J. Drummond. 2014. Bayesian inference of
654   sampled ancestor trees for epidemiology and fossil calibration. 10:e1003919.

655   Gavryushkina, A., T. A. Heath, D. T. Ksepka, T. Stadler, D. Welch, and A. J. Drummond. 2017.
656   Bayesian Total-Evidence Dating Reveals the Recent Crown Radiation of Penguins. Syst. Biol.
657   66:57–73.

658   Gingerich, P. D. 1985. Species in the fossil record: concepts, trends, and transitions.
659   Paleobiology 11:27–41.

660   Gingerich, P. D. 1979. Stratophenetic approach to phylogeny reconstruction. *in* J. Cracraft and
661   N. Eldredge, eds. Phylogenetic Analysis and Paleontology. Columbia University Press, New
662   York, NY.

28

663 Goloboff, P. A., A. Torres, and J. S. Arias. 2017. Weighted parsimony outperforms other
664 methods of phylogenetic inference under models appropriate for morphology. Cladistics
665 281:20141278.

666 Goloboff, P. A., C. I. Mattoni, and A. S. Quinteros. 2006. Continuous characters analyzed as
667 such. Cladistics 22:589–601. Blackwell Publishing Ltd.

668 Gould, S. J. 1980. Is a new and general theory of evolution emerging? Paleobiology 6:119–130.

669 Gould, S. J., and N. Eldredge. 1977. Punctuated equilibria: the tempo and mode of evolution
670 reconsidered. Paleobiology 3:115–151.

671 Guy, F., D. E. Lieberman, D. Pilbeam, M. P. de Leon, A. Likius, H. T. Mackaye, P. Vignaud, C.
672 Zollikofer, and M. Brunet. 2005. Morphological affinities of the Sahelanthropus tchadensis (Late
673 Miocene hominid from Chad) cranium. Proceedings of the National Academy of Sciences
674 102:18836–18841.

675 Heath, T. A., J. P. Huelsenbeck, and T. Stadler. 2014. The fossilized birth-death process for
676 coherent calibration of divergence-time estimates. Proc. Natl. Acad. Sci. U.S.A. 111:E2957–66.

677 Huelsenbeck, J. P., and B. Rannala. 1997. Maximum Likelihood Estimation of Phylogeny Using
678 Stratigraphic Data. Paleobiology 23:174–180. Paleontological Society.

679 Irish, J. D., D. Guatelli-Steinberg, S. S. Legge, D. J. de Ruiter, and L. R. Berger. 2013. Dental
680 morphology and the phylogenetic "place" of Australopithecus sediba. Science 340:1233062–
681 1233062.

682 Jones, E., T. Oliphant, and P. Peterson. n.d. SciPy: Open Source Scientific Tools for Python.

683 Kimbel, W. H., C. A. Lockwood, C. V. Ward, M. G. Leakey, Y. Rak, and D. C. Johanson. 2006.
684 Was Australopithecus anamensis ancestral to A. afarensis? A case of anagenesis in the hominin
685 fossil record. J. Hum. Evol. 51:134–152.

686 Leakey, M. G., C. S. Feibel, I. McDougall, and A. Walker. 1995. New four-million-year-old
687 hominid species from Kanapoi and Allia Bay, Kenya. Nature 376:565–571. Nature Publishing
688 Group.

689 Levinton, J. S. 2001. Genetics, Paleontology, and Macroevolution. Cambridge University Press.

690 Levinton, J. S., and M. S. Chris. 1980. A Critique of the Punctuated Equilibria Model and
691 Implications for the Detection of Speciation in the Fossil Record. Syst. Biol. 29:130–142.

692 Lewis, P. O. 2001. A Likelihood Approach to Estimating Phylogeny from Discrete
693 Morphological Character Data. Syst. Biol. 50:913–925.

694 Lewis, P. O., M.-H. Chen, L. Kuo, L. A. Lewis, K. Fučíková, S. Neupane, Y.-B. Wang, and D.
695 Shi. 2016. Estimating Bayesian Phylogenetic Information Content. Syst. Biol. 65:1009–1023.

696   MacLatchy, L. M., J. Desilva, W. J. Sanders, and B. Wood. 2010. Hominini. *in* W. J. Sanders
697   and L. Werdelin, eds. *Cenozoic mammals of Africa*. University of California Press, Oakland, CA.

698   MacLeod, N. 1991. Punctuated anagenesis and the importance of stratigraphy to paleobiology.
699   Paleobiology 17:167–188.

700   Mounier, A., F. Marchal, and S. Condemi. 2009. Is Homo heidelbergensis a distinct species?
701   New insight on the Mauer mandible. J. Hum. Evol.

702   Parins-Fukuchi, C. 2017. Use of Continuous Traits Can Improve Morphological Phylogenetics.
703   Syst. Biol., doi: 10.1093/sysbio/syx072.

704   Puttick, M. N., J. E. O'Reilly, A. R. Tanner, J. F. Fleming, J. Clark, L. Holloway, J. Lozano-
705   Fernandez, L. A. Parry, J. E. Tarver, D. Pisani, and P. C. J. Donoghue. 2017. Uncertain-tree:
706   discriminating among competing approaches to the phylogenetic analysis of phenotype data.
707   Proc. Roy. Soc. B 284:20162290.

708   Pyron, R. A. 2011. Divergence Time Estimation Using Fossils as Terminal Taxa and the Origins
709   of Lissamphibia. 60:466–481. Oxford University Press.

710   Rightmire, G. P. 1998. Human evolution in the Middle Pleistocene: the role of Homo
711   heidelbergensis. Evolutionary Anthropology.

712   Rosas, A., and J. M. Bermúdez De Castro. 1998. The Mauer mandible and the evolutionary
713   significance of Homo heidelbergensis. Geobios 31:687–697.

714   Simpson, G. G. 1944. Tempo and Mode in Evolution. Columbia University Press, New York.

715   Smith, A. B. 2000. Stratigraphy in phylogeny reconstruction. Journal of Paleontology 74:763–
716   766.

717   Smith, A. B. 1994. Systematics and the Fossil Record. Blackwell Science Ltd, Oxford, UK.

718   Stadler, T. 2010. Sampling-through-time in birth–death trees. Journal of Theoretical Biology
719   267:396–404.

720   Stadler, T., A. Gavryushkina, R. C. M. Warnock, A. J. Drummond, and T. A. Heath. 2017. The
721   fossilized birth-death model for the analysis of stratigraphic range data under different speciation
722   concepts.

723   Stanley, S. M. 1998. Macroevolution: Pattern and Process. 2nd ed. Johns Hopkins University
724   Press, Baltimore, MD.

725   Strait, D. S. 1999. Cladistics and Early Hominid Phylogeny. Science 285:1209c–1209.

726   Strait, D. S., F. E. Grine, and M. A. Moniz. 1997. A reappraisal of early hominid phylogeny. J.
727   Hum. Evol. 32:17–82.

728   Stringer, C. 2012. The status of Homo heidelbergensis (Schoetensack 1908). Evolutionary
729   Anthropology: Issues, News, and Reviews 21:101–107. Wiley-Blackwell.

730   Strotz, L. C., and A. P. Allen. 2013. Assessing the role of cladogenesis in macroevolution by
731   integrating fossil and molecular evidence. Proc. Natl. Acad. Sci. U.S.A. 110:2904–2909.

732   Wagner, P. J. 1998. A likelihood approach for evaluating estimates of phylogenetic relationships
733   among fossil taxa. Paleobiology 24:430–449. Cambridge University Press.

734   Ward, C. V., M. G. Leakey, and A. Walker. 2001. Morphology of Australopithecus anamensis
735   from Kanapoi and Allia Bay, Kenya. J. Hum. Evol. 41:255–368.

736   White, T. 2003. Paleoanthropology. Early hominids--diversity or distortion? Science 299:1994–
737   1997.

738   Wright, A. M., and D. M. Hillis. 2014. Bayesian analysis using a simple likelihood model
739   outperforms parsimony for estimation of phylogeny from discrete morphological data.
740   9:e109210. Public Library of Science.

741   Yang, Z., and T. Zhu. 2018. Bayesian selection of misspecified models is overconfident and may
742   cause spurious posterior probabilities for phylogenetic trees. Proc. Natl. Acad. Sci. U.S.A.
743   115:1854–1859. National Academy of Sciences.

744   Yang, Z., S. Kumar, and M. Nei. 1995. A new method of inference of ancestral nucleotide and
745   amino acid sequences. Genetics 141:1641–1650.

746   Zhang, C., T. Stadler, S. Klopfstein, T. A. Heath, and F. Ronquist. 2016. Total-Evidence Dating
747   under the Fossilized Birth-Death Process. Syst. Biol. 65:228–249.

748   Zollikofer, C. P. E., M. S. Ponce de León, D. E. Lieberman, F. Guy, D. Pilbeam, A. Likius, H. T.
749   Mackaye, P. Vignaud, and M. Brunet. 2005. Virtual cranial reconstruction of Sahelanthropus
750   tchadensis. Nature 434:755–759. Nature Publishing Group.
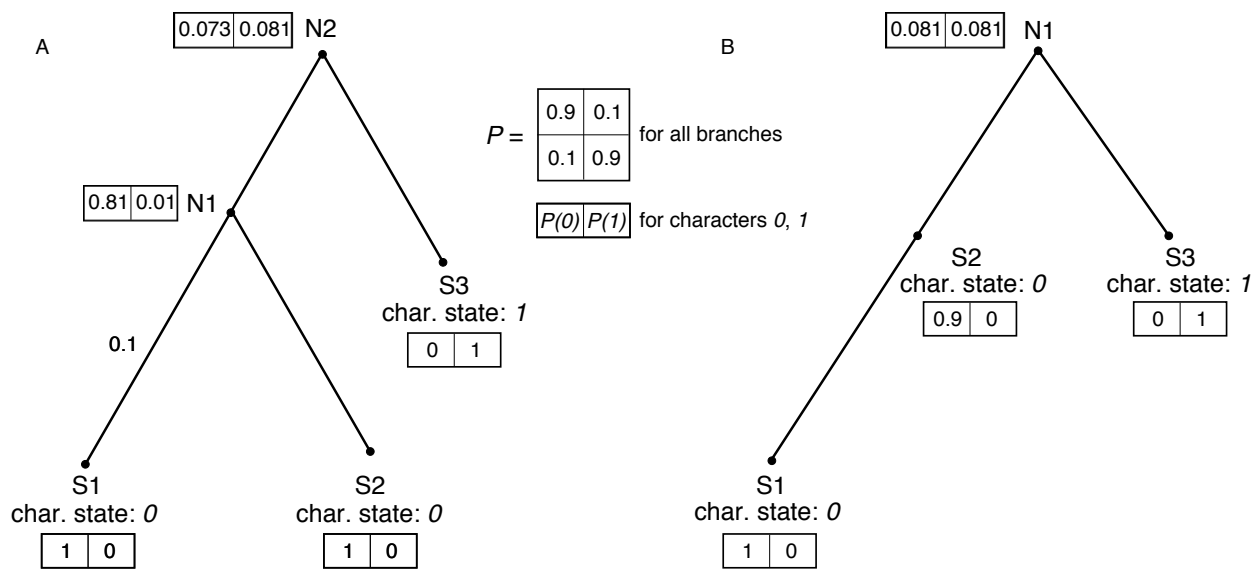
751

752

753

754

755

756

757

758   **Tables:**

Hypothesis	AIC score
Au. garhi outgroup to *Paranthropus* and *Homo*	9271.11
Au. garhi directly ancestral to *Homo*	9264.30

759 **Table 1.** Support for alternative hypotheses regarding the placement of *Au. garhi*.
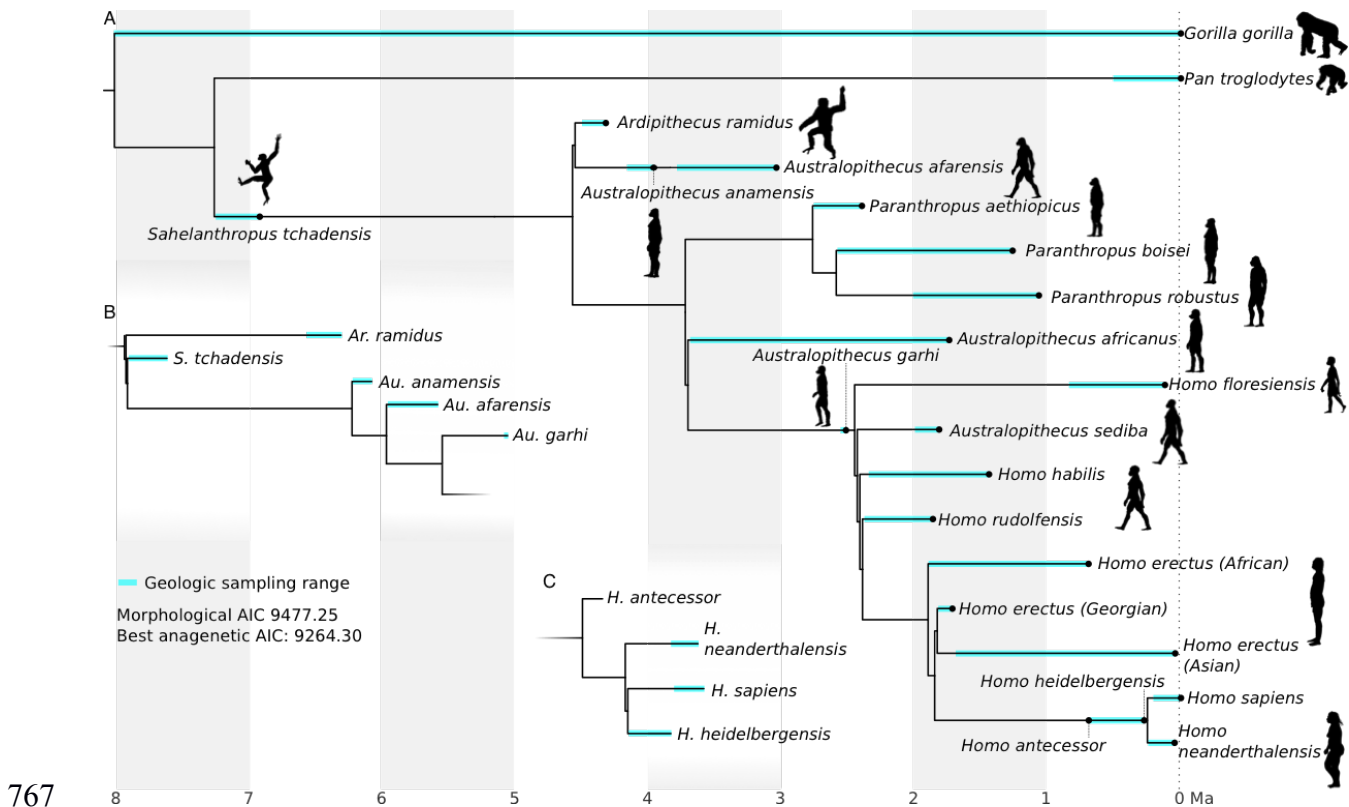
760

761 **Figures:**

762

763

764 **Fig. 1**. Comparison between anagenetic and cladogenetic trees. A) Likelihoods are calculated on

765 bifurcating arrangements using the standard 'pruning algorithm'. B) Anagenetic likelihoods are

766 calculated using a novel approach, where ancestral states may be known.

Fig. 2. Phylogenetic relationships between hominin species. A) Reconstruction with the best AIC score when temporal data are considered alongside morphological data and anagenesis is accommodated. B, C) Areas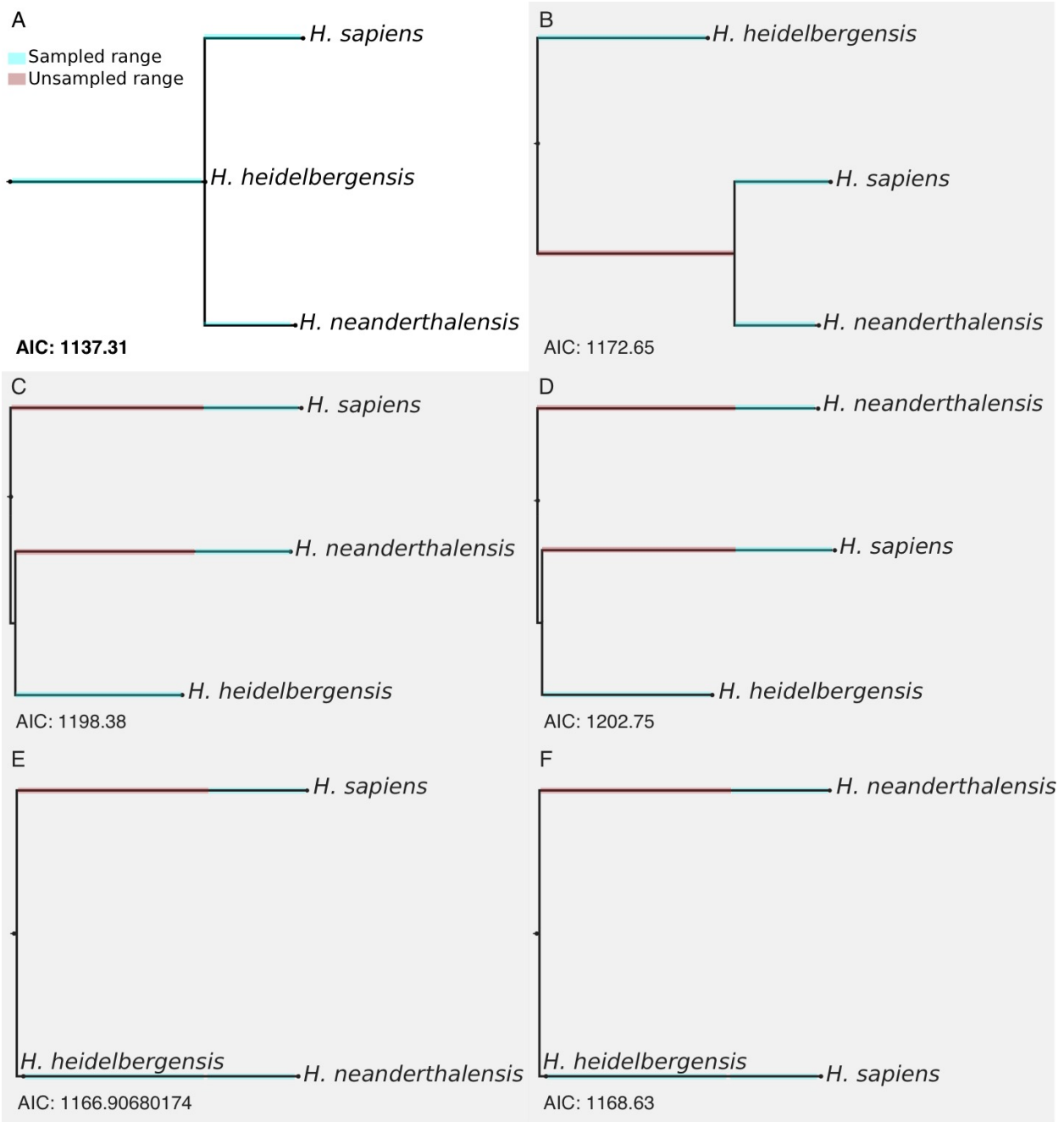 of the tree inferred from morphology alone that differ from panel A. Both regions display misleading results when anagenesis is not considered. This is reflected in the improvement in AIC observed in the preferred anagenetic topology. Silhouettes obtained from phylopic.org.
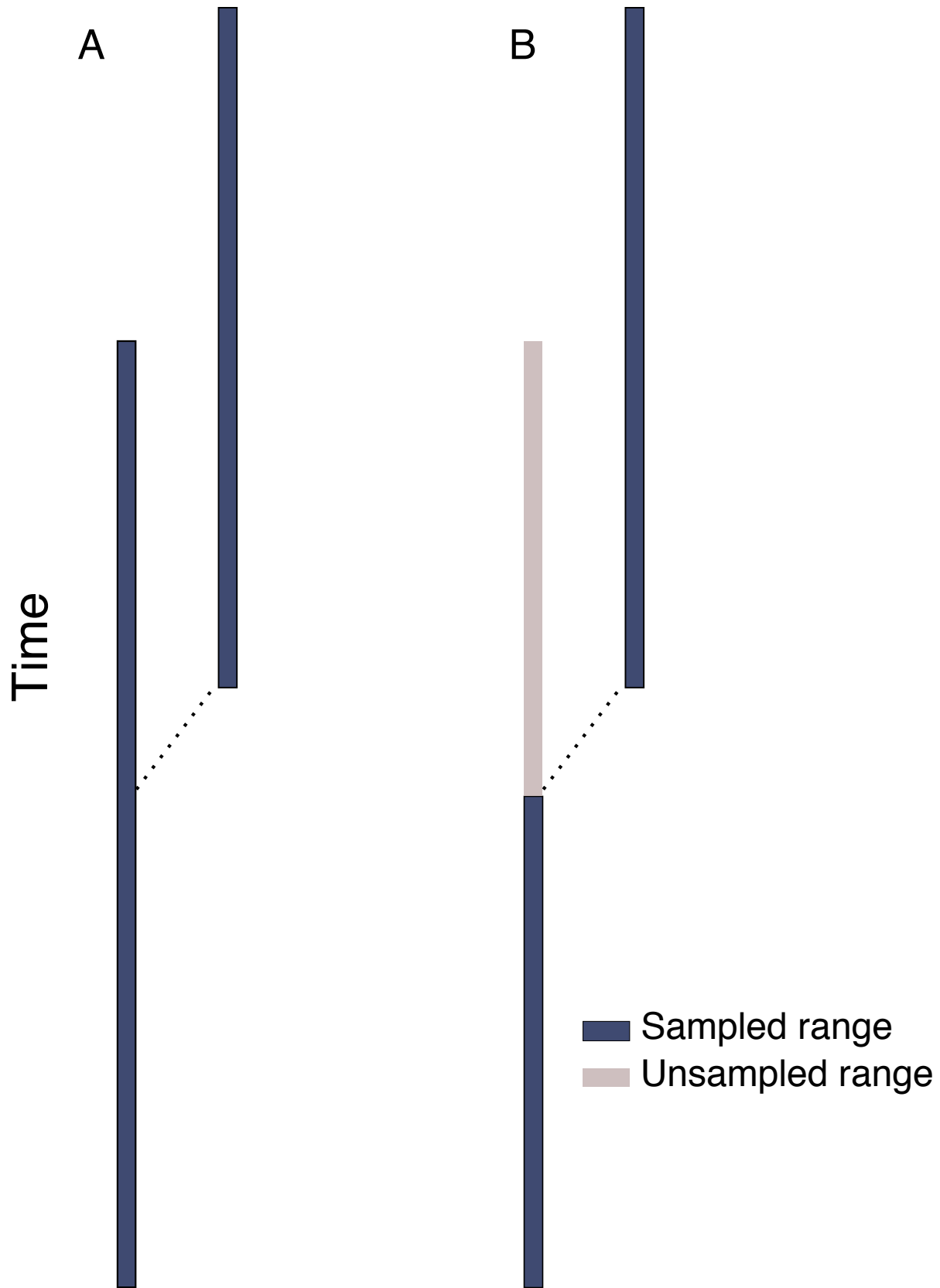
774

775

Fig. 3. AIC scores calculated for each possible arrangement between *H. sapiens*, *H. neanderthalensis,* and *H. heidelbergensis*.

779
780 **Figure 4**. A) Speciation mode interpreted as budding when sampling is complete. B) Incomplete

781   stratigraphic sampling may create an inability to distinguish between anagenesis and

782   budding when sampling is sparse.