



HAL
open science

Bayesian Morphological Clock versus Parsimony: An Insight into the Relationships and Dispersal Events of Postvacuum Cricetidae (Rodentia, Mammalia)

Raquel López-Antoñanzas, Pablo Peláez-Campomanes

► **To cite this version:**

Raquel López-Antoñanzas, Pablo Peláez-Campomanes. Bayesian Morphological Clock versus Parsimony: An Insight into the Relationships and Dispersal Events of Postvacuum Cricetidae (Rodentia, Mammalia). *Systematic Biology*, 2022, 71 (3), pp.512-525. 10.1093/sysbio/syab059 . hal-03374784

HAL Id: hal-03374784

<https://hal.science/hal-03374784>

Submitted on 12 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Bayesian Morphological Clock versus Parsimony: An insight into**
2 **the relationships and dispersal events of postvacuum Cricetidae**
3 **(Rodentia, Mammalia)**

4 **Raquel López-Antoñanzas^{1,2*} and Pablo Peláez-Campomanes²**

5
6 ¹Laboratoire de Paléontologie, Institut des Sciences de l'Évolution (ISE-M, UMR 5554,
7 CNRS/UM/IRD/EPHE), Université de Montpellier, Montpellier, France

8 ²Departamento de Paleobiología, Museo Nacional de Ciencias Naturales-CSIC, Madrid, Spain

9 *Corresponding author email: raquel.lopez-antonanzas@umontpellier.fr

10
11 **ABSTRACT**

12 Establishing an evolutionary timeline is fundamental for tackling a great variety of topics
13 in evolutionary biology, including the reconstruction of patterns of historical biogeography,
14 coevolution and diversification. However, the tree of life is pruned by extinction and
15 molecular data cannot be gathered for extinct lineages. Until recently methodological
16 challenges have prevented the application of tip-dating Bayesian approaches in morphology-
17 based fossil-only datasets. Herein, we present a morphological dataset for a group of cricetid
18 rodents to which we apply an array of methods fairly new in palaeontology that can be used
19 by palaeontologists for the analysis of entirely extinct clades. We compare the tree topologies
20 obtained by traditional parsimony, time-calibrated and non-calibrated Bayesian inference
21 phylogenetic approaches and calculate stratigraphic congruence indices for each. Bayesian
22 tip-dated clock methods outperform parsimony in the case of our dataset, which includes

23 highly homoplastic morphological characters. Regardless, all three topologies support the
24 monophyly of Megacricetodontinae, Democricetodontinae and Cricetodontinae. Dispersal
25 and speciation events inferred through Bayesian Binary Markov chain Monte Carlo and
26 biodiversity analyses provide evidence for a correlation between biogeographic events,
27 climatic changes and diversification in cricetids.

28

29 Keywords: Cricetidae, Miocene, Bayesian tip-dating, morphological clock, parsimony,
30 STRAP, palaeobiogeography, palaeoecology, palaeobiodiversity

ANALYSING ENTIRELY EXTINCT CLADES

31 The fossil record is our only resource for providing a deep-time perspective of ecosystem
32 processes and, therefore, for understanding the dynamics that have shaped our current biota. It
33 is increasingly clear that fossil data are fundamental to infer species diversification, particularly
34 extinction rates (Mitchell et al., 2019). Despite the fact that morphological phylogenetics have
35 been increasingly marginalised in the last decades, morphological data remain the only
36 available information to reconstruct evolutionary scenarios and reconcile the fossil record with
37 molecular trees (Wright 2017; Lee and Palci 2015). The importance of morphological
38 phylogenetics for dating rigorously the tree of life is now widely recognized and has been
39 bolstered by recent methodological developments (Hunt and Slater 2016; Lee and Palci 2015;
40 Wright 2017). Recent developments applying Bayesian methods using fossil taxa as tips, what
41 is called the morphological clock, have been revitalizing the use of morphological data to
42 elucidate the dynamics of evolution over time and across the tree of life (Varela et al. 2019;
43 Simões et al. 2018, 2020a). While these analyses are usually carried out with data from both
44 fossils and extant taxa, a very interesting application of this methodology is that it can be
45 employed with data from extinct clades only, which adds to the palaeontologist's toolbox
46 another method of reconstructing the evolution. This makes it now possible to compare
47 phylogenies of extinct taxa obtained by means of evolutionary models with those resulting
48 from maximum parsimony, which is the most widely applied method for analysing
49 morphological data.

50 Time provides palaeontologists with a unique perspective on phylogeny. A few methods to
51 integrate stratigraphic data with parsimony analyses were already available in the 90's
52 (Wagner, 1995; Fisher, 1994). However, palaeontologists had to devote considerable time and
53 effort in the process of calibrate them manually because of the lack of user-friendly software
54 facilitating time-calibration of cladograms. Moreover, depending on the number of taxa
55 included in the dataset palaeontologists have to infer the distribution of morphological

ANALYSING ENTIRELY EXTINCT CLADES

56 characters without the inclusion of temporal data, with the subsequent loss of information.
57 Parsimony was until recently the only way for palaeontologists to analyse their morphological
58 fossil datasets. However, since the introduction of Bayesian tip-dated phylogenetic methods,
59 which were first applied with uniform tree prior (Ronquist et al. 2012) and then with fossilized
60 birth-death (FBD) tree priors (Stadler 2010; Heath et al. 2014), the inclusion of stratigraphic
61 data into phylogenetic analyses boomed. The development of tip-dating with FBD tree priors
62 that allow fossil species to be included as terminal tips has been particularly useful for
63 palaeontologists. Recent numerical methods such as PaleoTree (Bapst, 2012, 2014) or STRAP
64 (Stratigraphic Tree Analysis for Palaeontology) (Bell and Lloyd 2014) allow phylogenies
65 resulting from both parsimony and Bayesian analyses of fossil taxa to be dated. The package
66 STRAP allows besides to assess their stratigraphic congruence (Bell and Lloyd 2014;
67 O'Connor and Wills 2016; King and Beck 2019; King 2021). So, the development of all these
68 comparative methods have promoted the revival of phylogenetic analysis incorporating
69 stratigraphic data and the testing of different techniques of phylogenetic reconstruction using
70 morphological data (Bell and Lloyd 2014; Sansom et al., 2018; King 2021). This provides
71 palaeontologists with a golden opportunity to expand considerably their research toolkit.

72 Over the years, the development of numerical, parsimony and Bayesian methods to
73 determine trees has resulted in an intense study of certain clades (Benton 2015). However, the
74 most striking aspect of rodent palaeophylogenetics is the low ratio of phylogenetic hypotheses
75 to species number. In fact, few comprehensive morphological phylogenetic analyses have been
76 carried out for these mammals and none has been performed applying morphological clock
77 methods. Nonetheless, rodents have an excellent fossil record and their dentition provides an
78 ideal dataset to characterize phenotypic variation and calibrating divergence time analyses. The
79 present work rests on cricetids, which are an important group of rodents because they include
80 representatives involved in the first radiation of modern rodents in Eurasia during the Miocene.

ANALYSING ENTIRELY EXTINCT CLADES

81 They probably gave rise to other important lineages, such as arvicolines (voles), cricetines
82 (hamsters) and even murines (mice). The Miocene comprises the first phase in the development
83 of modern terrestrial ecosystems. This is a time interval that witnessed enormous geographical
84 and environmental changes and during which the mammal record experienced a major turnover
85 and successive dispersal events. Cricetids possibly appeared in China in Middle Eocene times
86 (Tong 1992). After an initial radiation in Central Asia, the group expanded in Western Europe
87 just after the “Grande Coupure” (Stehlin, 1910), when the Turgai Strait dried out and a land
88 bridge between Europe and Asia was established. They prospered in Europe during the
89 Oligocene but primitive cricetids disappeared before the arrival of the more derived, typical
90 Miocene ones. The time interval devoid of cricetids has been called the “cricetid vacuum” by
91 Daams and Freudenthal (1989). Post-vacuum cricetids are known worldwide and under various
92 morphologies since Miocene times, but their evolutionary history is poorly known. The
93 biogeographic relationships and phylogenetic affinities of cricetid species with one another
94 need to be studied in a time-scaled phylogenetic framework. This work aims to clear up the
95 phylogenetic relationships and estimate divergence times of postvacuum cricetids to elucidate
96 their origin and the dispersal events they underwent by applying parsimony and Bayesian
97 methods to an array of important extinct lineages of Miocene cricetids from Europe, the Middle
98 East and Asia. Moreover, the different approaches we use are assessed on the basis of the
99 stratigraphic coherence of the resulting topologies in order to arrive at an informed opinion on
100 which method is most likely to yield the most accurate results in the phylogenetic analyses of
101 palaeontological (morphological) dataset.

102 MATERIAL AND METHODS

103 *Material*

104 The systematic study presented below is based on the examination of original specimens
105 listed in Supplementary Tab. S1 provided in the Supplementary Information available in the

ANALYSING ENTIRELY EXTINCT CLADES

106 Dryad data repository [https://datadryad.org/stash/share/Dmda4GGW64PE8-](https://datadryad.org/stash/share/Dmda4GGW64PE8-c4fZiwP2VpbnTITZpWm4PRSm5L20Q)
107 [c4fZiwP2VpbnTITZpWm4PRSm5L20Q](https://datadryad.org/stash/share/Dmda4GGW64PE8-c4fZiwP2VpbnTITZpWm4PRSm5L20Q).

108 First, second and third lower molars are designated as m1, m2 and m3, respectively and first,
109 second and third upper molars as M1, M2 and M3, respectively.

110 *Maximum Parsimony Analyses*

111 *Implied weights maximum parsimony analysis*—A total of 82 phylogenetically informative
112 dental characters from 74 taxa have been coded. The data matrix (Supplementary file S1
113 available on Dryad) has been built using Mesquite 3.04 (Maddison & Maddison 2009) and
114 the analysis ran in TNT v.1.5 (Goloboff et al. 2003) using the new technology search
115 algorithms and the implied weighting algorithm (Supplementary file S2 available on Dryad)
116 (Goloboff et al. 2008). New technology search algorithms are recommended for large
117 datasets because they allow the sampling of trees from a broader spectrum of local optima
118 (Goloboff et al. 1999). Tree searches have been carried out using 1,000 initial trees by
119 random addition sequences with 100 iterations or rounds for each of the four NTS algorithms:
120 sectorial search, ratchet, drift and tree fusing. The analysis has been performed with $K = 10$
121 and collapsing all branches with support = 0. K values larger than the default (3.0) are more
122 accurate to perform analyses for large datasets (Goloboff et al. 2017). The number of
123 suboptimal trees to be retained was set at 10 and the relative fit difference at 0.1. The final
124 output trees (22 MPTs+suboptimal trees) have been filtered for all the most parsimonious
125 trees (MPTs). A total of 3 MPT with a length of 487 steps, a Consistency Index (CI) of 0.261
126 and a Retention Index (RI) of 0.765 have been obtained and have been used to calculate the
127 strict consensus tree (488 steps) (Supplementary Figure S1 available on Dryad). Character 1,
128 length of the M1, has been treated as additive. Branch support have been estimated through
129 two complementary indices: Bremer Support (Bremer 1994) and Relative Bremer Support
130 (Goloboff and Farris 2001).

ANALYSING ENTIRELY EXTINCT CLADES

131

132 *Equal weights maximum parsimony analysis*—The analysis has been also carried out with
133 TNT using equal weights. Tree searches have been conducted as performed for the implied
134 weighting parsimony analysis. The final output trees (440 MPTs+suboptimal trees) have been
135 filtered for all the most parsimonious trees (MPTs). A total of 375 MPT with 479 steps, a
136 Consistency Index (CI) of 0.265 and a Retention Index (RI) of 0.767 have been obtained and
137 have been used to calculate the strict (530 steps) and majority (488 steps) consensus trees
138 (Supplementary Figures S2 and S3 available on Dryad).

139 *Bayesian inference analyses*

140 Analyses have been carried out using Mr. Bayes v.3.2.6 (Ronquist et al. 2012) and the
141 BEAST2 package (Bouckaert R., et al. (2014) using the CIPRES Science Gateway v.3.3
142 (Miller et al., 2010).

143 *Non-clock Bayesian analysis*—The analysis has been performed with Mr. Bayes v.3.2.6
144 (Ronquist et al. 2012). The morphological dataset (Supplementary file S1 available on Dryad)
145 has been analysed with the MkV model (Lewis, 2001) under the γ model, 30 million
146 generations and four independent runs. Convergence of independent runs is assessed by an
147 average standard deviation of split frequencies of 0.0148, by an average potential scale
148 reduction factors (PSRF) of 1 for all parameters and an effective sample size (ESS) greater than
149 200 for each parameter.

150 *Time-calibrated relaxed-clock Bayesian inference analysis*—The analysis has been
151 performed in BEAST2 package (Bouckaert R. et al. 2014) (Supplementary file S3 available on
152 Dryad). The MkV model (Lewis 2001) has been used, with a gamma distribution with four rate
153 categories to account for rate variation across sites. The prior on the gamma shape parameter
154 is an exponential distribution. Following Simões et al. (2020a), the morphological dataset has

ANALYSING ENTIRELY EXTINCT CLADES

155 been analysed using the fossilized birth-death tree model with sampled ancestors (FBD-SA)
156 (Stadler 2010), under relaxed-clock models, 30 million generations and four independent runs
157 with the initial 30% of samples removed as “burn-in”. The relaxed-clock model is the
158 independent γ rate relaxed-clock model, which is a continuous uncorrelated relaxed-clock
159 model using a gamma distribution to assess clock rate variation across lineages (Simões et al
160 2018). It is compatible with the fossilized birth–death tree model. The base clock rate has been
161 given an informative prior that derived from the undated Bayesian inference analysis: the
162 median value for tree height (TL) in substitutions from the entire posterior trees sample
163 (TL=9.9825) divided by the age of the tree (37.6 Ma), which is based on the median of the
164 distribution for the root prior. So, the prior is: $9.9825/37.6=0.265$. In contrast to Mr. Bayes,
165 BEAST 2 does not allow different modelling strategies for how extant taxa are sampled and
166 assumes a random sampling strategy (Simões et al, 2020b), which is unimportant in studies
167 that do not include extant species. Our calibrations have been based on tip dating, which
168 accounts for the uncertainty in the placement of fossil taxa and avoids the issue of bound
169 estimates for node-based age calibrations (Ronquist et al. 2016, Simões et al. 2018). It has been
170 substantiated that sampling fossil ages instead of fixing them to the midpoint within their
171 stratigraphic age range avoids biases in divergence time estimations (Barido-Sotani et al.,
172 2019). Thus, the fossil ages used in this work for tip dating correspond to the uniform prior
173 distributions on the age range of the stratigraphic occurrence of the fossils. The age of the root
174 has been set with a soft lower bound. The minimum age of the root corresponds to the oldest
175 age for the oldest fossil belonging to *Eucricetodon* and the maximum root has been set at 41.2
176 Ma (Middle/Late Eocene boundary), which is the maximum soft age for the clade
177 Eucricetodontinae. Conditioning on rho has not been possible because the data set of the
178 analysis only contains extinct taxa so we have conditioned on the root and on sampling
179 (Supplementary file S3 available on Dryad). An effective sample size (ESS) greater than 200
180 for the main parameters assesses the convergence of independent runs.

ANALYSING ENTIRELY EXTINCT CLADES

181 *Stratigraphic congruence*

182 To assess the stratigraphic congruence of the parsimony and Bayesian inference-derived
183 phylogenies resulting from our analyses we used the package STRAP for R with the default
184 number of permutations (1000) (Bell and Lloyd, 2014). We have obtained the following
185 stratigraphic fit indices: MIG (minimum implied gap (Norell and Novacek, 1992; Norell et al.
186 1992)), SCI (Stratigraphy Consistency Index (Huelsenbeck 1994)), RCI (Relative
187 Completeness Index (Benton and Storrs 1994)), GER (Gap Excess Ratio (Wills 1999)),
188 MSM* (modified Manhattan Stratigraphic Measure (Pol, 2001)) and Wills' modifications of
189 GER (GERT and GER*, Wills et al. (2008)). MIG provides the sum of the branch lengths
190 excluding tip durations (the sum of the ghost ranges), SCI provides the proportion of nodes
191 that are stratigraphically congruent in a tree, RCI provides the measure of the extent of
192 observed ranges of taxa and the sum of the ghost ranges, MSM* corresponds to MIG for the
193 maximally stratigraphically consistent possible tree divided by the actual MIG and GER to
194 MIG minus the best possible stratigraphic fit, scaled by the contrast between the best and
195 worst fit values (Lloyd and Bell, 2014; Wright and Lloyd, 2020)). SCI, GER, GERT, and
196 MSM* scale between 0.0 (least congruent) and 1.0 (most congruent) (O'Connor and Wills
197 2016; Wright and Lloyd, 2020). Additionally, the program provides the significance test for
198 those measurements (p.SCI, p.RCI, p.GER, and p.MSM*) as well as a combined fit and
199 significance test measure (GER* and GERT) (Bell and Lloyd 2014). The p-values indicate the
200 probability of the null hypothesis, which indicates random tree topology. So, very small p-
201 values will indicate a significantly good fit to stratigraphy (Bell and Lloyd 2014).

202 *Biogeographic Inference using Reconstruct Ancestral State in Phylogenies*

203 In order to infer historical biogeography we have applied RASP 4.2 (Yu et al. 2015), a
204 method to reconstruct ancestral geographical distributions using a combination of phylogenetic
205 and geographical information. Bayesian Binary Markov chain Monte Carlo (BBM) method for

ANALYSING ENTIRELY EXTINCT CLADES

206 ancestral state has been performed on the Maximum Clade Credibility Tree (MCCT) in RASP
 207 4.2 (Yu et al. 2015). BBM calculates the probabilities of ancestral ranges using the probabilities
 208 of each unit area generated by MrBayes (Yu et al. 2015). The Markov chain Monte Carlo
 209 (MCMC) chains were run for five million generations and ten independent runs. The state was
 210 sampled every 100 generations. Fixed JC + G (Jukes-Cantor + Gamma) were used for BBM
 211 analysis. The distribution range of all species of cricetids included in this work has been divided
 212 into eight geographic areas, which have been carefully chosen according to dissimilarities in
 213 their faunal composition that provided evidence for palaeobiogeographic or
 214 palaeoenvironmental differences between them. They correspond to: A (southwestern Europe),
 215 B (Central Europe), C (Greece), D (Anatolia-Caucasus), E (west central Asia), F (Arabian
 216 Peninsula), G (southern Asia) and H (east central Asia).

217 *Biodiversity rates*

218

219 The combination of cladistic and biodiversity analyses highlights the phenomena of
 220 speciation, extinction, and diversity changes in a given group over time (Stigall 2010, López-
 221 Antoñanzas et al. 2015). This allows establishing the timing of biodiversity crises and, thus,
 222 deducing possible causes.

223 Per-capita rates for speciation (\hat{q}), extinction (\hat{p}), and diversity change (d) have been
 224 calculated (Supplementary Tab. S2 available on Dryad) following the equations given by
 225 Foote (2000) according to which:

$$226 \hat{p} = -\ln(N_{bt}/(N_{bt} + N_{ft})) / \Delta t$$

$$227 \hat{q} = -\ln(N_{bt}/(N_{bt} + N_{bL})) / \Delta t$$

$$228 d = \hat{p} - \hat{q}$$

229 where N_{bt} indicates the number of species that cross both the upper and lower interval
 230 boundaries, N_{ft} the number of species that originate within the interval and cross over the

ANALYSING ENTIRELY EXTINCT CLADES

231 upper interval boundary and N_bL the number of species that cross the lower interval boundary,
232 but become extinct during the interval and Δt the duration of the interval t_1-t_0 .

233 Rates of biodiversity change (R), speciation (S), and extinction (E) have been calculated
234 (Supplementary Tab. S2 available on Dryad) with the following equations (Stigall 2010):

$$235 \quad R = (\ln N_1 - \ln N_0) / \Delta t$$

$$236 \quad S = (\ln(N_0 + o_0) - \ln N_0) / \Delta t$$

$$237 \quad E = (\ln(N_0 + o_0) - \ln N_1) / \Delta t$$

238 where N_0 is the initial number of species in a clade at time t_0 , N_1 the number of species in a
239 clade at time t_1 , o_0 the number of speciation events during the interval t_1-t_0 , and Δt the duration
240 of the interval t_1-t_0 .

241 All rates have been calculated each 0.5 Ma using the phylogenetically corrected species
242 ranges obtained from the MCCT. Values calculated from the first and last intervals have been
243 excluded from the analysis to remove edge effects, following the criterion of Stigall (2010).

244 RESULTS AND DISCUSSION

245 *Tree Topology*

246 Maximum Parsimony and tip-dated and undated Bayesian trees (Figs. 1, 2, Supplementary
247 Figs.S2-S4 available on Dryad) show three large clades that correspond to the rodent
248 subfamilies Megacricetodontinae, Democricetodontinae and Cricetodontinae. Both in, the
249 majority consensus tree resulting from equal weights maximum parsimony and undated
250 Bayesian analysis (Supplementary Figures S3 and S4 available on Dryad) show
251 Megacricetodontinae and Cricetodontinae more closely related to one another than either one
252 is to Democricetodontinae. However, some studies have evidenced that downweighting
253 characters according to their homoplasy (applying implied weighting) improves
254 morphological data sets, particularly those that are highly homoplastic, and produces more
255 resolved and accurate trees than standard equal weights (Goloboff et al., 2008; Smith, 2019).

ANALYSING ENTIRELY EXTINCT CLADES

256 As a well-recognized problem when working on fossil rodents is the supposedly high
257 homoplasy of dental characters, discussion on parsimony results will be based on the
258 topology of the tree resulting from the implied weighting analysis.
259 Both implied weighting parsimony and tip-dated Bayesian analyses (Figs. 1 and 2) supported
260 Cricetodontinae as the sister group to the clade consisting of Megacricetodontinae plus
261 Democricetodontinae, which are sister clades of each other. These results place
262 Democricetodontinae and Megacricetodontinae closer to one another than either is to the
263 Cricetodontinae, which is in line with the phylogenetic hypothesis proposed by Flynn (2009).

264

265 *Implied weights maximum parsimony*—The topology of the tree shows three main clades:
266 Megacricetodontinae, Democricetodontinae and Cricetodontinae.

267 The calibration of the strict consensus tree (Fig.1) shows that Democricetodontinae (node
268 83) originated circa 23.7 Ma. They share two non-exclusive synapomorphies (Supplementary
269 Tab. S3, Supplementary Figure S1 (for numbers of each node in the tree) available on Dryad)),
270 which are the presence of a double protoloph on M2 (29(0→1)) and the metacone absent or
271 included in crest on M3 (50(0→2)). All the taxa included in this clade also share the exclusive
272 synapomorphy of having an m1 with anteroconid in lingual position (60(0→1)).

273 Megacricetodontinae (node 138) originated approximately 23.7 Ma. They share three
274 exclusive synapomorphies. The first one consists in having a divided anterocone with the
275 lingual anterocone smaller than the labial one on M1 (5(0→2)). The second one rests in having
276 m3 quite reduced (with an Lm1/Lm3 ratio between 1.4-1.6 mm) (67(0→2)) and the third one
277 in having the entoconid small but still distinct on m3. The taxa included in this clade (except
278 for *Megacricetodon andrewsi*) also share the non-exclusive synapomorphy of having M3 with
279 usually incomplete central atoll (45(0→1)). Megacricetodontinae comprise two clades. The

ANALYSING ENTIRELY EXTINCT CLADES

280 first one regroups the species originating from node 137 (*Shamalina tuberculata* (*Punjabemys*
281 *downsi*, *Sindemys shewanensis*)). It is supported by two non-exclusive synapomorphies: the
282 presence of a small enterostyle isolated from the protocone and the lack of a well-developed
283 anterolophid on m3 ((18(0→1); 73(0→1)). The second clade consists in all the species that
284 originate from node 145 (*Aktaumys dzhungaricus* and more derived megacricetodontines).
285 They share the non-exclusive synapomorphies of having on M1 the lingual anteroloph weak or
286 absent (10(0→1)), a long mesoloph (19(1→0)) and the posteroloph that continues beyond the
287 point where it meets the metalophule (23(0→1)), and in having the protoloph slightly anteriorly
288 directed on M3 (41(0→1)).

289 Cricetodontinae are the species originating from node 91. They share the two exclusive
290 synapomorphies of having the anterocone divided in two parts of similar size on M1 (5(0→1))
291 and a distinct anterior ectoloph (12(0→1)). They are also sustained by five non-exclusive
292 synapomorphies: a long posteroloph that continues beyond the point where it meets the
293 metalophule (23(0→1)), a large single anteroconid on m1 (54(0→1)) and the lingual
294 anterolophid weak or absent on the lower molars (59(0→2), 68(0→1), 73(0→1)).

295 The split of Cricetodontinae from the sister clades Democricetodontinae plus
296 Megacricetodontinae is set up circa 23.8 Ma. The basalmost taxa within Cricetodontinae are
297 *Cricetodon versteegi* and, one node up, *C. goklerensis*, both from the Early Miocene of Turkey.
298 The remaining taxa are split into two very asymmetrical branches. One is represented by
299 plesiomorphic species of *Cricetodon* from the Early Miocene of Turkey (*C. fikreti* + *C.*
300 *trallesensis*), whereas the other form a much larger group that includes all the remaining species
301 of the subfamily. At the base of this clade a polytomy involving *Cricetodon fengi*,
302 *Mixocricetodon dehmi* plus *Cricetodon sonidensis*, and a clade comprising all the more derived
303 species. The latter splits into two lineages. The smaller one includes the species originating
304 from node 103 with *C. wanhei* as most basal and *C. aliveriensis* plus *C. kasapligili* as most

ANALYSING ENTIRELY EXTINCT CLADES

305 derived and *C. tobieni* in an intermediate position. This clade is supported by the non-exclusive
 306 and unambiguous synapomorphy of having a complete central atoll on M3 (a reversal). This
 307 character is only shared by the most plesiomorphic species of *Cricetodon* (*C. verstegi*) and
 308 most of *Democricetodon* spp. The larger branch is constituted by a succession of lineages, the
 309 most basal of which is that of the most plesiomorphic true European species, *C. meini* plus *C.*
 310 *aureus*. One node up *C. orientalis* from the Middle Miocene of China branches and then *C.*
 311 *soriae* from the Middle Miocene of Europe. The more derived species divide first in a clade of
 312 Middle Miocene European species of *Cricetodon* (*C. jotae*, *C. jumaensis*, *C. albanensis*, *C.*
 313 *bolliegeri*, *C. nievei*, *C. engesseri*). They share the two non-exclusive synapomorphies of
 314 having the lingual anteroloph of M2 poorly developed (26(0→1)) and M2 elongated
 315 (28(0→1)). The remaining species, originating from node 93, consist in a succession of
 316 European (*C. sansaniensis*, *C. hungaricus*, *C. caucasicus*) and Turkish (*C. pasalarensis*) taxa
 317 that splits (node 114) to give rise to *Byzantinia* and *Hispanomys* during the Middle Miocene.
 318 The bifurcation in which *Hispanomys* originated (node 113) includes the late Middle Miocene
 319 species of ‘*Cricetodon*’ from Central Europe (‘*C.*’ *klariankae*, and ‘*C.*’ *venczeli*), which should
 320 be considered as belonging to *Hispanomys*. Similarly, the bifurcation from which *Byzantinia*
 321 evolved (node 116) contains the Middle Miocene ‘*Cricetodon*’ *fandli*, ‘*C.*’ *candirensis* and ‘*C.*’
 322 *cariensis* basally, which may be reinterpreted, according to this analysis, as plesiomorphic
 323 representatives of *Byzantinia*.

324 *Tip-dated Bayesian analysis*—The divergence between Cricetodontinae and
 325 Democricetodontinae + Megacricetodon is set up at approximately 28.2 Ma, and that between
 326 the sister clades Democricetodontinae and Megacricetodontinae circa 25.9 Ma.

327 Democricetodontinae: The topology of the tree (Fig. 2) shows two sister clades. The first
 328 includes two sister species: *Democricetodon sui* from the Early Miocene of China (21.9-21.16
 329 Ma), which is the oldest record of *Democricetodon* in Asia (Maridet et al., 2011), and *Primus*

ANALYSING ENTIRELY EXTINCT CLADES

330 *microps* from the Early Miocene of Pakistan (circa 21-23 Ma) (Bruijn et al. 1981). The second
331 one shows at the base the two plesiomorphic species of *Democricetodon* (*D. anatolicus* and *D.*
332 *doukasi*) from the Early Miocene of Turkey, which insert sequentially on the stem, leading to
333 more derived *Democricetodon* spp. *Democricetodon franconicus*, from the Early Miocene of
334 Aliveri (Greece), is the most basal species of this clade. It is followed successively by *D.*
335 *mutilus* and *D. gracilis*, which reveal the first entrance of Democricetodontinae in the central-
336 western European bioprovince (circa 18.27 Ma). *D. gracilis* is sister species to a clade including
337 mostly Asian species, providing evidence for an early Middle Miocene migration of
338 *Democricetodon* from Europe towards China at approximately 16.1 Ma.

339 Megacricetodontinae: The topology of the tree (Fig. 2) shows two main clades within this
340 subfamily of rodents. The first one includes an array of Early Miocene species of disputed
341 origin (*Vallaris zappai*, (*Sindemys shewanensis*, *Punjabemys downsi*)). *Punjabemys downsi*
342 from the Early Miocene of Pakistan was originally considered a member of
343 Megacricetodontinae (Lindsay, 1988) and then as belonging to the Myocricetodontinae by
344 Wessels (2009) together with *Vallaris zappai* and *Sindemys shewanensis* from the Early
345 Miocene of Turkey and Pakistan, respectively. According to the topology of the tree, these taxa
346 are either the basalmost representatives of Megacricetodontinae or could as well be considered
347 as their sister group. Be that as it may, they are sister group to the clade that includes the other
348 megacricetodontines, the basalmost taxon of which is *Shamalina tuberculata* from the Early
349 Miocene of Saudi Arabia. This result agrees with Lindsay (1994), who thought that *Shamalina*
350 could have given rise to *Megacricetodon*, but oppose the suggestion of Wessels (2009), who
351 included this taxon within the Myocricetodontinae. *Shamalina* may have led to two main
352 lineages. One consists of the sister species *Aktaumys dzhungaricus* and *Megacricetodon*
353 *beijiangensis* and points to a dispersal event from Saudi Arabia towards Kazakhstan and China
354 circa 21.28 Ma. The other lineage, which is more speciose, shows at its base *Megacricetodon*

ANALYSING ENTIRELY EXTINCT CLADES

355 *hellenicus* from the Early Miocene of Aliveri, Greece (Oliver and Peláez-Campomanes, 2014),
356 which provides evidence for the first entrance of *Megacricetodon* into the Aegean-Anatolian
357 region at approximately 19.5 Ma ago. One node further up two clades substantiates two
358 independent migration events. One involved *Megacricetodon* aff. *collongensis* (circa 18 Ma),
359 the most primitive species of the typical central European “*Megacricetodon bavaricus* group”
360 proposed by Oliver and Peláez-Campomanes (2013), and the other *M. andrewsi* and *M.*
361 *primitivus* at approximately 17 Ma. These results agree with Oliver and Peláez-Campomanes
362 (2016), who inferred at least three migration events in Europe for early *Megacricetodon* forms
363 on the basis of important morphological differences.

364 Cricetodontinae originated in Turkey at approximately 25 Ma (Fig. 2). The basalmost taxon
365 is *C. versteegi* and *C. goklerensis* is one node up in the clade. Both taxa come from the lower
366 Miocene of Turkey (de Bruijn et al. 1993; Joniak et al. 2017). *Cricetodon goklerensis* is sister
367 species to a large group that includes all the remaining species of the subfamily, the basalmost
368 clade of which includes the most plesiomorphic species of *Cricetodon*, mostly from Turkey
369 and China, and documents the first entrance of *Cricetodon* from Turkey into China at *circa*
370 19.4 Ma. The split of this basal clade and the clade comprising more derived species of
371 Cricetodontinae is set up at 22.3 Ma. Within the latter, the most basal taxa are represented by
372 the Turkish Early Miocene species *C. fikreti* and *C. trallesensis*. *Cricetodon trallesensis*
373 diverges from the main clade approximately 20 Ma ago and led to two clades. The first one
374 includes practically all European Middle Miocene species of *Cricetodon* and reveals the first
375 entrance of the Cricetodontinae into Europe circa 17.77 Ma. The presence inside this group of
376 the Asiatic species *C. orientalis* provides evidence for a second dispersal of Cricetodontinae
377 into China, this time from Europe, which took place approximately 16.8 Ma. The second clade
378 comprises the most derived species of Cricetodontinae. A sequence of the three most basal taxa
379 from the Middle Miocene of Central Europe and Turkey (*C. hungaricus*, *C. caucasicus*, *C.*

ANALYSING ENTIRELY EXTINCT CLADES

380 *pasalarensis*) insert sequentially on the stem, leading to two important lineages. One leads to
381 *Hispanomys* and includes the late Middle Miocene species of “*Cricetodon*” from Central
382 Europe (*C. fandli*, *C. klariankae* and *C. venzecli*), which should be considered to belong to
383 *Hispanomys*, according to this analysis. Similarly, the bifurcation from which *Byzantinia*
384 evolved contains, in a basal position, the Middle Miocene *Cricetodon cariensis* and *C.*
385 *candirensis*, which would be plesiomorphic representatives of *Byzantinia*.

386 *Bayesian analyses versus Parsimony*

387 Tree topologies of undated Bayesian and equal weights maximum parsimony searches are
388 more similar to each other than either is to the implied weighting parsimony and tip-dating
389 Bayesian trees. In the same way, the topologies of implied weighting parsimony and tip-dating
390 Bayesian trees are more similar to each other than either to the two other topologies. Despite
391 the fact that the topologies of implied weight parsimony and tip-dated Bayesian trees (Figs. 1,
392 2) are similar as far as the relationships between large clades are concerned, they differ in the
393 placement of some taxa amongst the lineages inside these clades. This issue holds particularly
394 true for the clades that are not comprehensively sampled (e.g., Democricetodontinae and
395 Megacricetodontinae vs Cricetodontinae) and those that are weakly supported. Because of the
396 differences recovered between the topologies resulting from different methods, stratigraphic
397 congruence indices have been calculated to assess how well each of these phylogenies fits with
398 the chronostratigraphy. Our results show high levels of stratigraphic congruence for all metrics
399 for parsimony (equally weights and implied weights) and both Bayesian (undated and tip-
400 dating) methods (Supplementary Tab. S4 available on Dryad). As expected, the maximum
401 credibility tree recovered by tip-dating has significantly higher stratigraphic congruence than
402 the majority consensus trees resulting from undated Bayesian and parsimony analyses
403 (Supplementary Tab. S4 available on Dryad). The majority consensus tree resulting from
404 implied weighting parsimony (similar to the strict consensus tree) has higher stratigraphic

ANALYSING ENTIRELY EXTINCT CLADES

405 congruence than the majority consensus trees recovered from equal weights parsimony and
406 undated Bayesian searches. Recent studies provided evidence that the inclusion of stratigraphic
407 age data in tip-dating impacts the topology of the tree, particularly when the datasets include
408 incompletely scored taxa or taxa having weak character support (King 2021). The differences
409 we found between the topologies of the tip-dated Bayesian tree and the undated Bayesian and
410 parsimony trees are probably due to the presence of highly homoplastic characters in our
411 dataset. Taken into account that the topology of the undated phylogenies shows some
412 anomalous results (e.g. derived position of the oldest democricetodontines) and that time-
413 scaling these topologies (Fig. 1) results in improbable long ghost lineages for a large number
414 taxa, we consider that the topology obtained via the tip-dating Bayesian analyses is most likely
415 to be more accurate with our kind of morphological dataset (fossil rodent teeth). Therefore, the
416 discussion on palaeogeography and biodiversity will be based on the Maximum Clade
417 Credibility Tree (MCCT) resulting from the tip-dating Bayesian analysis (Fig. 2). Interestingly,
418 Lee and Yates (2018) also found that tip-dating approaches were better able to elucidate the
419 relationship of some crocodylian groups that showed extensive convergent adaptations.

420

421 *Palaeobiogeographic framework*

422 During the Miocene, Democricetodontinae, Megacricetodontinae and Cricetodontinae were
423 successful and exhibited a wide distribution, ranging geographically from western Europe to
424 eastern Asia. Their evolution is complex and involves indigenous speciation phenomena and
425 numerous intercontinental dispersal events. According to our results most immigration events
426 involved Turkey (Fig. 3). Democricetodontinae may have originated in Turkey approximately
427 24.8 Ma and entered Asia about 23.9 Ma. The results of this work also provides evidence of a
428 probable Turkish origin for European Democricetodontinae with a first arrival in central and
429 western Europe circa 18.3 Ma. The delayed entrance of this group into Europe was probably

ANALYSING ENTIRELY EXTINCT CLADES

430 due to the geographical isolation of the Dinarian-Anatolian Island (Neubauer et al. 2015). A
431 Middle Miocene dispersal of *Democricetodon* from Europe to China has been evidenced in this
432 work and set up around 16.1 Ma.

433 A more comprehensive phylogenetic analysis that includes the Myocricetodontinae is
434 needed to elucidate if *Vallaris zappai* and more derived taxa belong to the Megacricetodontinae
435 or are, actually, closer to the Myocricetodontinae. Without this information, two scenarios are
436 plausible. The first one points to a Turkish origin of Megacricetodontinae circa 23.3 Ma,
437 whereas in the second one they would originate from either Turkey or Saudi Arabia around
438 22.6 Ma. The eastwards dispersal to Kazakhstan and western China is set circa 21.3 Ma and
439 that to the Aegean-Anatolian area (where they are recorded in Aliveri and K ese koy) circa 19.5
440 Ma. Two independent dispersal events took place then from this area towards the West first
441 and then toward eastern China, and Central Europe at approximately 18 Ma. A subsequent
442 dispersal from western Asia to the Aegean-Anatolian area occurred approximately one million
443 years later (circa 17 Ma) (Fig. 3).

444 Cricetodontinae are supposed to have originated from Turkey at 25 Ma, where several
445 lineages differentiated, and then passed into Greece at the earliest at 20 Ma and to eastern Asia
446 around 19.4 Ma (Fig. 3). Greek cricetodontines from the Early Miocene of Aliveri are not
447 closely related to the European Middle Miocene cricetodontines but to the Turkish Early
448 Miocene ones. In fact, at this time, Aliveri may not have belonged to the European region but
449 to the Aegean-Anatolian one, which was only intermittently in contact with Europe. Most of
450 the European cricetodontines originated from an independent Turkish lineage that dispersed
451 into the south-western part of the continent circa 18.5 Ma. Two dispersal events towards central
452 Europe occurred at 16.4 Ma and 15.5 Ma and two additional ones towards the Aegean-
453 Anatolian area and western China are inferred at 16.8 Ma and 16.2 Ma. On the basis of our
454 analysis, it turns out that the Late Miocene cricetodontines are not closely related to the Middle

ANALYSING ENTIRELY EXTINCT CLADES

455 Miocene European ones. In fact, one important dispersal event at about 16.4 Ma from Turkey
456 towards Europe together with an independent evolution in Turkey led to a later development
457 of two dynamic centres of speciation in southwestern and central Europe (*Hispanomys*) and
458 Anatolia (*Byzantinia*).

459 *Palaeoclimatic context*

460 During the Oligocene and the Early Miocene, the East Antarctic Ice Sheet was relatively
461 unstable with periods of growth and others of decline (Miller et al., 2020). At the
462 Eocene/Oligocene Transition and the beginning of the Oligocene (33.9-32 Ma), a large
463 deglaciation triggered successive sea level rises up to 50 m. On the contrary, sea-level drops
464 were recorded in early Late Oligocene times, at approximately 28 Ma (O1), and at the
465 beginning of the Miocene, circa 23 Ma (Mi1) (Miller et al., 2020). Interestingly, this sea-level
466 fall at the Oligocene/Miocene boundary coincides with a severe restriction of the marine
467 connections between the Indian Ocean and the Mediterranean Sea (Hüsing et al., 2009). This
468 could have opened dispersal land routes that could have allowed the migration of primitive
469 megacricetodontines, such as *Shamalina tuberculata* from Saudi Arabia towards Kazakhstan
470 and western China (~23.5 Ma). The dispersal of *Democricetodon* from Turkey to Pakistan and
471 China (~23.9 Ma) might have been elicited by the same phenomenon. Until the beginning of
472 the Miocene Climatic Optimum (17-15 Ma), eustatic sea level changes indicates the presence
473 of a moderate to large ice sheet in Antarctica (Miller et al. 2020). Low sea level that allowed
474 dispersal of faunas are indicators of cool intervals, whereas higher sea levels represents warmer
475 periods. Climate variations are expected to have played a major role in shaping the diversity of
476 small mammals in general and rodents in particular. So, confronting the timing of shifts of
477 diversification and climatic perturbations can provide lines of evidence regarding the influence
478 of the latter on the former. The analysis of the shifts of diversification across the phylogeny of
479 Cricetodontinae, the most comprehensive subclade in this work, is shown in Fig. 4.

ANALYSING ENTIRELY EXTINCT CLADES

480 Regrettably, the poor fossil record of Cricetodontinae in Turkey between 23 Ma and 19 Ma,
481 where and when this group originated and started diversifying, does not allow a robust
482 interpretation of our results during this interval of time. The MCO (~17-15 Ma) was a warm
483 period that is mostly characterized by high sea level and reduced ice volume. Since a little
484 before the Miocene Climatic Optimum (18.5 Ma) and nearly throughout all of it, the rodent
485 community of cricetodontines experienced a period of stability and success, during which the
486 biodiversity was slightly but constantly increasing (Fig. 4). This increase in biodiversity
487 resulted from a continuous decrease in the extinction rate of these rodents together with high
488 speciation rates that remained nearly constant. At approximately 16.0 Ma, an interruption of
489 the Miocene Climatic Optimum is evinced by a sea level fall of ~40 m (Mi2) (Miller et al.
490 2020), which would have allowed the second dispersal of *Democricetodon* from westwards to
491 eastwards. This cooling event (Mi2) seems to have promoted a drop in the speciation of Middle
492 Miocene European cricetodontines, which was fairly high up to that point (Fig. 4), while their
493 extinction rate continued to decrease. After the MCO and during the Middle Miocene Climatic
494 Transition (~15-13 Ma), three events of cooling and sea-level fall at 14.8 (Mi2a), 13.8 (Mi3)
495 and 12.8 Ma (Mi4) took place that culminated with the establishment of the “permanent East
496 Antarctic Ice Sheet” (Miller et al. 2020). After the Mi2 event, speciation and extinction rates
497 tended to decrease continuously. However, our results show high extinction and speciation
498 rates that are coupled with the Miocene cooling events Mi3 and Mi4 (Fig. 4). The most
499 prominent turnover is detected during the Mi3 event (circa 13.8 Ma). Interestingly, this event
500 seems to have been particularly impactful in southern Europe, where most of the
501 cricetodontines lived at this time. Actually, a sudden and important drop in temperatures was
502 evidenced in Spain after analysing oxygen isotopes of mammal teeth enamel (Domingo et al.
503 2009, 2012). The evolutionary radiation of various endemic lineages of *Byzantinia*, in
504 Anatolia, and *Hispanomys*, in Europe, took place during the Mi3 event (Fig. 4). In these areas,
505 parallel lineages showed similar trends towards an increase of size, a reduction of length of the

ANALYSING ENTIRELY EXTINCT CLADES

506 third molars and the development of a complete backwards paracone spur (Rummel 1998).
507 This suggests a specialisation toward food that includes more abrasive and fibrous plants,
508 indicating in turn increasingly open environments. When *Byzantinia* and *Hispanomys*
509 flourished, their potential competitors such as the Middle Miocene cricetodontines were driven
510 to extinction. At the beginning of the Late Miocene, at about 11.5 Ma (Mi5), another cooling
511 event associated with an increase in the ice volume in Antarctica took place (Cook et al., 2008).
512 This event is linked to a faunal turnover within the Cricetodontinae in which a drop in
513 biodiversity caused by elevated extinction rates (oldest taxa belonging to *Byzantinia*) was
514 followed by a second radiation of the most derived species of *Hispanomys* and *Byzantinia*.

515 It is evident that climatic events impacted the evolutionary history of the Cricetodontines.
516 However, given that some of the oldest taxa belonging to *Byzantinia* are not dated precisely,
517 the high extinction and speciation rates that we found in Mi5 may be exaggerated, whereas
518 those corresponding to Mi 4 might be minored.

519

520 CONCLUSION

521 Our analyses evidence that time-calibrated Bayesian searches yield trees that have higher
522 stratigraphic congruence compared with trees from undated Bayesian and parsimony
523 searches. Majority and strict consensus trees resulting from implied weights parsimony show
524 the second best stratigraphic concordance values. They are followed by equal character
525 weighting parsimony and non-clock Bayesian searches. So, according to our results, if
526 stratigraphic congruence is taken as a proxy for phylogenetic accuracy, then time-calibrated
527 Bayesian inference analysis (morphological clock) is probably the more accurate method for
528 analysing morphological characters with a high degree of homoplasy (such as those of rodent
529 teeth). However, more studies comparing morphological parsimony and time-calibrated and
530 not time-calibrated Bayesian methods are needed to conclude which of these approaches deal

ANALYSING ENTIRELY EXTINCT CLADES

531 better with the morphological characters provided by fossil rodents, which are mainly based
532 on the dentition and considered highly convergent.

533 The Miocene is marked by obvious manifestations of climatic changes, which turn out to
534 have been very influential in the evolution of animal life. Cricetids arrived at different times
535 in different areas around the Mediterranean and became extinct asynchronously as well. Our
536 results suggest that the Miocene cooling events, particularly Mi2, Mi3, Mi4 and Mi5 that
537 took place at 16 Ma, 13.8 Ma, 12.8 Ma and 11.5, respectively, impacted cricetid evolution by
538 promoting dispersal and triggering important origination/extinction events.

539

540 DATA AVAILABILITY

541 The data underlying this article are available in the article and from the Dryad Digital
542 Repository <https://datadryad.org/stash/share/Dmda4GGW64PE8-c4fZiwP2VpbnTITZpWm4PRSm5L20Q>

543

544 ACKNOWLEDGMENTS

545 We sincerely thank the instructors T. Simões (Harvard University, Cambridge, USA) and
546 O. Vernygora (University of Kentucky, Lexington, USA) of the Transmitting Science Course
547 “Morphological Phylogenetics” for kindly answering our inquiries. Further thanks for helpful
548 comments and constructive criticism are given to O. Vernygora, who also reviewed this
549 article together with two anonymous reviewers as well as the associate editor A. Wright
550 (Southeastern Louisiana University, Hammond, USA).

551

552 FUNDING

553 This research received support from the research project PGC2018-094122-B-100
554 (MICU/AEI/FEDER,EU) and the SYNTHESYS+ project (<http://www.synthesys.info>), which
555 is financed by European Commission.

556

ANALYSING ENTIRELY EXTINCT CLADES

557 REFERENCES

- 558 Bapst D.W. 2012. paleotree: an R package for paleontological and phylogenetic analyses of
559 evolution. *Methods Ecol. Evol.* 3:803–807.
- 560 Bapst D.W. 2014. Assessing the effect of time-scaling methods on phylogeny-based analyses
561 in the fossil record. *Paleobiology*.40:331–351.
- 562 Barido-Sottani J., Aguirre-Fernández G., Hopkins M.J., Stadler T., Warnock R. 2019.
563 Ignoring stratigraphic age uncertainty leads to erroneous estimates of species divergence
564 times under the fossilized birth–death process. *Proc. R. Soc. B.* 286:2019068520190685
- 565 Bell M.A. and Lloyd G.T. 2014 strap: An R package for plotting phylogenies against
566 stratigraphy and assessing their stratigraphic congruence. *Palaeontology.* 58:379–389.
- 567 Benton M.J., Storrs G.W. 1994. Testing the quality of the fossil record: paleontological
568 knowledge is improving. *Geology.* 22:111–114.
- 569 Benton MJ. 2015 Exploring macroevolution using modern and fossil data. *Proc. R. Soc. B.*
570 282:20150569.
- 571 Bouckaert R., Heled J., Kühnert D., Vaughan T., Wu C.H., Xie D., Suchard M.A., Rambaut
572 A., Drummond A.J. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary
573 Analysis. *PLoS Computational Biology.* 10:e1003537.
- 574 Bremer K. 1994. Branch support and tree stability. *Cladistics.* 10:295–304.
- 575 Casanovas-Vilar I, Garcés M. Van Dam J., García-Paredes I., Robles J.M., Alba D.M. 2016.
576 An updated biostratigraphy for the late Aragonian and Vallesian of the Vallès-Penedès
577 Basin (Catalonia, Spain). *Geol. Acta* 14:195-217.
- 578 Cooke P.J., Nelson C.S., Crundwell M.P. 2008. Miocene isotope zones, paleotemperatures,
579 and carbon maxima events at intermediate water-depth, Site 593, Southwest Pacific, New
580 Zealand. *J. Geol. Geophys.* 51:1–22.
- 581 Daams R., Freudenthal M. 1989. Cricetidae (Rodentia) from the type-Aragonian; the genus
582 *Megacricetodon*. *Scr. Geol.* 1:39–132.

ANALYSING ENTIRELY EXTINCT CLADES

- 583 De Bruijn, H. 1976. Vallesian and Turolian rodents from Biotia, Attica and Rhodes (Greece).
584 Proc. Kon. Nederl. Akad. Wetensch. B. 79:361–384.
- 585 De Bruijn H., Hussain S.T., Leinders, J.M. 1981. Fossil rodents from the Murree formation
586 near Banda Daud Shah, Kohat, Pakistan. Proc. Kon. Nederl. Akad. Wetensch. B. 84:71–
587 99.
- 588 De Bruijn H., Fahlbusch V., Saraç G., Ünay E. 1993. Early Miocene rodent faunas from the
589 eastern Mediterranean areas. III: the genera *Deperetomys* and *Cricetodon* with a
590 discussion of the evolutionary history of the Cricetodontini. Proc. Kon. Nederl. Akad.
591 Wetensch. B. 96:151–216.
- 592 Domingo L., Cuevas González J., Stephen T.G., Hernández Fernández M., López-Martínez
593 N., 2009. Multiproxy reconstruction of the palaeoclimate and palaeoenvironment of the
594 Middle Miocene Somosaguas site (Madrid, Spain) using herbivore dental enamel.
595 Palaeogeogr. Palaeoclimatol. Palaeoecol. 272:53–68.
- 596 Domingo L., Koch P.L., Grimes S.T., Morales J., López-Martínez N., 2012. Isotopic
597 paleoecology of mammals and the Middle Miocene Cooling event in the Madrid Basin
598 (Spain). Palaeogeogr. Palaeoclimatol. Palaeoecol. 339:98–113.
- 599 Fisher D.C. 1994. Stratocladistics: morphological and temporal patterns and their relation to
600 phylogenetic process. In: Grand L., Rieppel O., editors. Interpreting the hierarchy of
601 nature. San Diego: Academic Press, p. 133–171.
- 602 Flynn L.J. 2009. The antiquity of *Rhizomys* and independent acquisition of fossorial traits in
603 subterranean muroids. Bull. Am. Mus. Nat. Hist. 331:128–156.
- 604 Flynn L.J., Pilbeam D., Barry J.C., Morgan M.E., Raza S.M. 2016. Siwalik synopsis: A long
605 stratigraphic sequence for the Later Cenozoic of South Asia. C. R. Palevol. 15:877–887.
- 606 Foot M. 2000. Origination and extinction components of taxonomic diversity: general
607 problems. Paleobiol. 26: 74–102.

ANALYSING ENTIRELY EXTINCT CLADES

- 608 Goloboff P.A., Farris J.S. 2001. Methods for quick consensus estimation. *Cladistics*. 17:26–
609 34.
- 610 Goloboff P.A. 1999. Analyzing large data sets in reasonable times: solutions for composite
611 optima. *Cladistics*. 15:415–428.
- 612 Goloboff P.A., Farris J.S., Nixon K.C. 2003. TNT: Tree Analysis Using New Technology.
613 Program and documentation, available from the authors at www.lillo.org.ar/phylogeny/tnt/
- 614 Goloboff P.A., Carpenter J.M., Arias J.S., Esquivel D.R.M. 2008. Weighting against
615 homoplasy improves phylogenetic analysis of morphological data sets. *Cladistics*. 24:758–
616 773.
- 617 Goloboff P.A., Farris J.S., Nixon K.C. 2008. TNT, a free program for phylogenetic analysis.
618 *Cladistics*. 24: 774–786.
- 619 Goloboff P.A., Torres A., Arias J.S. 2017. Weighted parsimony outperforms other methods of
620 phylogenetic inference under models appropriate for morphology. *Cladistics*. 34:407–437.
- 621 Heath T.A., Huelsenbeck J.P., Stadler T. 2014. Fossilized birth–death process. *PNAS*.
622 111:E2957–E2966.
- 623 Hír J., Venczel M., Codrea V., Angelone C., van den Hoek Ostende L.W., Kirscher U., Prieto
624 J. 2016. Badenian and Sarmatian s.str. from the Carpathian area: Overview and ongoing
625 research on Hungarian and Romanian small vertebrate evolution. *C. R. Palevol*. 15:863–
626 875.
- 627 Huelsenbeck J.P. 1994 Comparing the stratigraphic record to estimates of phylogeny.
628 *Paleobiology*. 20:470–483.
- 629 Hüsing S.K., Zachariasse W.J., van Hinsbergen D.J.J., Murat Inceöz, W. K., Harzhauser M.,
630 Mandic O., Krohet A. 2009. Oligocene–Miocene basin evolution in SE Anatolia, Turkey:
631 constraints on the closure of the eastern Tethys gateway. *Geol. Soc. London. Spec. Publ.*
632 311:107–132.

ANALYSING ENTIRELY EXTINCT CLADES

- 633 Hunt G., Slater G. 2016. Integrating Paleontological and Phylogenetic Approaches to
634 Macroevolution. *Annu. Rev. Ecol. Evol. Syst.* 47:189–213.
- 635 Joniak P., De Bruijn H. 2014. Rodents from the Upper Miocene Tuğlu Formation (Çankiri
636 Basin, Central Anatolia, Turkey). *Palaontol Z.* 89:1039–1056.
- 637 Joniak P., Peláez-Campomanes P., van den Hoek Ostende L.W., Rojay B. 2017. Early
638 Miocene rodents of Gökler (Kazan Basin, Central Anatolia, Turkey), *Hist. Biol.* 31:982–
639 1007.
- 640 King B. 2021. Bayesian tip-dated phylogenetics in paleontology: topological effects and
641 stratigraphic fit. *Syst. Biol.* 70:283–294.
- 642 King B., Beck R. 2019. Bayesian Tip-dated Phylogenetics: Topological Effects, Stratigraphic
643 Fit and the Early Evolution of Mammals. *bioRxiv* 533885.
- 644 Lee M.S., Palci A. 2015. Morphological Phylogenetics in the Genomic Age. *Curr. Biol.*
645 25:R922–9.
- 646 Lee M.S.Y., Yates A.M. 2018. Tip-dating and homoplasy: reconciling the shallow molecular
647 divergences of modern gharials with their long fossil record. *Proc. R. Soc. B* 285:
648 20181071.
- 649 Lewis P.O. 2001. A likelihood approach to estimating phylogeny from discrete
650 morphological character data. *Syst. Biol.* 50:913–925.
- 651 Lindsay E.H. 1988. Cricetid rodents from Siwalik deposits near Chinji Village, part 1:
652 Megacricetodontinae, Myocricetontinae and Dendromurinae. *Paleovertebrata* 18:95–154.
- 653 Lindsay E.H. 1994. Fossil record of Asian Cricetidae with emphasis on Siwalik cricetids. In:
654 Tomida Y., Li C.K., Setoguchi T., editors. *Rodent and Lagomorph Families of Asian*
655 *Origins and Diversification*. National Science Museum Monographs, 8. p. 131–147.
- 656 Lindsay E.H. 2017. *Democricetodon fejfari* sp. nov. and replacement of Cricetidae by
657 Muridae in Siwalik deposits of Pakistan. *Foss. Impr.* 73:445–453.

ANALYSING ENTIRELY EXTINCT CLADES

- 658 López-Antoñanzas R., Knoll F., Wan S., Flynn L.J. 2015. Causal evidence between monsoon
659 and evolution of rhizomyine rodents. *Sci. Rep.* 5:9008.
- 660 Maddison W.P., Maddison D.R., 2009. *Mesquite: A Modular System for Evolutionary*
661 *Analysis*, ver. 2.6. Mesquite Project, Vancouver.
- 662 Maridet O., Wu W.Y., Ye J., Bi S.D., Ni X., Meng J. 2011. Early Miocene cricetids
663 (Rodentia) from the Junggar basin (Xinjiang, China) and their biochronological
664 implications. *Geobios.* 44: 445–459.
- 665 Miller K.G., Browning J.V., Schmelz W.J., Kopp R.E., Mountain G.S., Wright J.D. 2020.
666 Cenozoic sea-level and cryospheric evolution from deep-sea geochemical and continental
667 margin records. *Sci. Adv.* 6:eaz1346.
- 668 Miller M.A., Pfeiffer W., Schwartz T. 2010. Creating the CIPRES Science Gateway for
669 inference of large phylogenetic trees. In: 2010 gateway computing environments
670 workshop. p. 1–8.
- 671 Mitchell J., Etienne R.S., Rabosky D.L. 2019. Inferring diversification rate variation from
672 phylogenies with fossils. *Syst. Biol.* 68:1–18.
- 673 Norell M.A., Novacek M.J. 1992. Congruence between superpositional and phylogenetic
674 patterns: comparing cladistic patterns with fossil records. *Cladistics.* 8:319–337.
- 675 Norell M.A., Novacek M.J., Wheeler Q.D. 1992. Taxic origin and temporal diversity: the
676 effect of phylogeny. In: Novacek M.J., Wheeler Q.D., editors). *Extinction and phylogeny.*
677 Columbia University Press, p. 89–118.
- 678 O'Connor A. and Wills M.A. 2016. Measuring stratigraphic congruence across Trees, higher
679 taxa, and time. *Syst Biol.* 65:792–811.
- 680 Oliver A., Peláez-Campomanes P. 2014. Evolutionary patterns of early and middle
681 Aragonian (Miocene) of *Megacricetodon* (Rodentia, Mammalia) from Spain.
682 *Palaeontographica. A.* 303:85–135.
- 683 Pol D. 2001 Comments on the Manhattan stratigraphic measure. *Cladistics* 17, 285–289.

ANALYSING ENTIRELY EXTINCT CLADES

- 684 problems. *Paleobiology*. 26:74–102.
- 685 Prieto J., Rummel M. 2016. Some consideration on small mammal evolution in
686 SouthernGermany, with emphasis on Late Burdigalian-Earliest Tortonian (Miocene)
687 cricetid rodents. *C. R. Palevol*. 15:837–854.
- 688 Qiu Z., Li Q. 2016. Neogene rodents from central Nei Mongol, China. *Palaeontol. Sinica N.*
689 *Ser. C*. 198:1–684.
- 690 Reichenbacher B., Krijgsman W., Lataster Y., Pippèrr M., Van Baak C.G.C., Chang, L.,
691 Kálin, D., Jost J., Doppler, G., Jung, D., Prieto, J., Aziz, H.A., Böhme, M. Garnish,
692 J.,Kirscher U, Bachtadse, V. 2013. A new magnetostratigraphic framework for the Lower
693 Miocene (Burdigalian/Ottnangian, Karpatian) in the North Alpine Foreland Basin. *Swiss.*
694 *J. Geosci*. 106:309–334.
- 695 Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L,
696 Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic
697 inference and model choice across a large model space. *Syst. Biol*. 61:539–542.
- 698 Ronquist F., Lartillot N., Phillips M.J. 2016. Closing the gap between rocks and clocks using
699 total-evidence dating. *Phil. Trans. R. Soc. B*. 371:20150136.
- 700 Rummel M. 1998. Die Cricetiden aus dem Mittel- und Obermiozän der Türkei. *Doc. Naturae*.
701 123:1–300.
- 702 Sansom R.S., Choate P.G., Keating J.N., Randle E. 2018. Parsimony, not Bayesian analysis,
703 recovers more stratigraphically congruent phylogenetic trees. *Biol. Lett*. 14:20180263.
- 704 Simões T.R., Caldwell M.W., Tañanda M., Bernardi M., Palci A., Vernygora O., Bernardini
705 F., Mancini L., Nydam, R.L. 2018. The origin of squamates revealed by a Middle Triassic
706 lizard from the Italian Alps. *Nature*. 557:706–709.
- 707 Simões T.R., Vernygora O., Caldwell M.W., Pierce S.E. 2020a. Megaevolutionary dynamics
708 and the timing of evolutionary innovation in reptiles. *Nat. Commun*. 11:3322.
- 709 Simões T.R., Caldwell M.W., Pierce S.E. 2020b. Sphenodontian phylogeny and the impact

ANALYSING ENTIRELY EXTINCT CLADES

- 710 of model choice in Bayesian morphological clock estimates of divergence times and
711 evolutionary rates. *BMC Biology*.18:191.
- 712 Stadler T. 2010. Sampling-through-time in birth–death trees. *J. Theor. Biol.* 267:396–404.
- 713 Stehlin H.G. 1910. Remarques sur les faunules de Mammifères des couches éocènes et
714 oligocènes du Bassin de Paris. *Bull. Soc. Géol. France.* 4:488–520.
- 715 Stigall-Rode A.L., Lieberman B.S. 2005. Using environmental niche modeling to study the
716 Late Devonian biodiversity crisis. In: Over D.J., Morrow J.R., Wignall P.B., editors.
717 *Understanding Late Devonian and Permian-Triassic Biotic and Climatic Events: Towards*
718 *an Integrated Approach.* New York, NY: Elsevier. p. 93–179.
- 719 Stigall A.L. 2010. Invasive species and biodiversity crises: Testing the link in the Late
720 Devonian. *PLoS One.* 5:e15584.
- 721 Tong Y.S. 1992. *Pappocricetodon*, a pre-Oligocene cricetid genus (Rodentia) from Central
722 China. *Vertebrata Palasiatica* 30:1–16.
- 723 Ünay E., de Bruijn H., Saraç, G. 2003. A preliminary zonation of the continental Neogene of
724 Anatolia based on rodents. *Deinsea*, 10:539–547.
- 725 Van der Meulen A., García-Paredes I., Alvarez-Sierra M.A, van den Hoek Ostende L.W.,
726 Hordijk K., Oliver A., Pelaez-Campomanes P. 2012. Updated Aragonian biostratigraphy:
727 Small Mammal distribution and its implications for the Miocene European Chronology.
728 *Geol. Acta.* 10:159–179.
- 729 Varela L., Tambusso, P.S., McDonald, H.G., Fariña, R.A. 2019. Phylogeny,
730 macroevolutionary trends and historical biogeography of sloths: insights from a Bayesian
731 morphological clock analysis. *Syst. Biol.* 68:204–218.
- 732 Wagner P.J. 1995. Stratigraphic tests of cladistic hypotheses. *Paleobiol.* 21:153–178.
- 733 Wang X., Flynn L.J., Fortelius, M., editors. 2013. *Fossil Mammals of Asia: Neogene*
734 *Biostratigraphy and Chronology.* New York, Columbia University Press.

ANALYSING ENTIRELY EXTINCT CLADES

- 735 Wessels W. 2009. Miocene rodent evolution and migration: Muroidea from Pakistan, Turkey
736 and Northern Africa. *Geol. Ultraiectina*. 307:1–290.
- 737 Whybrow P.J., Collison M. E., Daams R., Gentry A.W., McClure H.A. 1982. Geology, fauna
738 (Bovidae, Rodentia) and flora from the early Miocene of eastern Saudi Arabia. *Ter.*
739 *Research*. 4:105–120.
- 740 Wills M.A. 1999 Congruence between phylogeny and stratigraphy: randomization tests and
741 the gap excess ratio. *Syst. Biol.* 48:559–580.
- 742 Wills M.A., Barret P.M, Heathcote J.F. 2008. The modified Gap Excess Ratio (GER*) and
743 the stratigraphic congruence of dinosaur phylogenies. *Syst. Biol.* 57:891–904.
- 744 Wright D.F. 2017. Bayesian estimation of fossil phylogenies and the evolution of early to
745 middle Paleozoic crinoids (Echinodermata). *J. Paleontol.* 91:799–814.
- 746 Wright A.M., Lloyd, G.T. 2020. Bayesian analyses in phylogenetic palaeontology:
747 interpreting the posterior sample. *Palaeontology* 63:997-1006.
- 748 Yu Y., Harris A.J., Blair C. and Xingjin H. 2015. RASP (Reconstruct Ancestral State in
749 Phylogenies): A tool for historical biogeography. *Mol. Phylogenetics Evol.* 87:46–49.

ANALYSING ENTIRELY EXTINCT CLADES

750

751 **Figure 1.** Calibrated strict consensus tree of cricetid rodents and their recorded temporal ranges
752 (black). Dashed bars indicate age uncertainty. Grey bars represent missing ranges and missing
753 ancestral lineages. Biochronological data from Casanovas et al. 2016; de Bruijn 1976; Flynn
754 2016; Hir et al. 2016; Joniak and de Bruijn 2014; Joniak et al. 2017; Lindsay 2017; Prieto and
755 Rummel 2016; Reichenbacher et al. 2013; Qiu and Li 2016; Ünay et al. 2003; Van der Meulen
756 et al. 2012; Wang et al., 2013; Wessels 2009; Whybrow et al. 1982.

757

758 **Figure 2.** Time-calibrated relaxed-clock Bayesian inference analysis with morphological tip-
759 dating using the fossilized birth–death tree model. Summary of the MCCT depicting the
760 median divergence time estimates for different cricetid lineages against a geological time
761 scale. Numbers at nodes indicate median estimates for the divergence times, and node bars
762 indicate the 95% highest posterior density for divergence times.

763

764 **Figure 3.** Ancestral state distributions at each node of the MCCT of cricetid obtained by BBM
765 analysis implemented in RASP against a geological time scale. Pie charts indicate probabilities
766 of alternative ancestral ranges. The colour indicates possible ancestral ranges at different nodes.
767 The schematic map shows the biogeographical areas used in this work and colours correspond
768 to: ocher (southwestern Europe), purple (central Europe), dark red (Greece), orange (Anatolia-
769 Caucasus), green (west central Asia), light blue (Arabian Peninsula), dark blue (southern Asia)
770 and yellow (east-central Asia).

771

772 **Figure 4.** Instantaneous per-capita (top) and deterministic (bottom) rates for Cricetodontinae
773 speciation (\hat{p} , S), extinction (\hat{q} , E), and biodiversity change (d, R) calculated using

ANALYSING ENTIRELY EXTINCT CLADES

774 phylogenetically constrained species ranges. Miocene cooling events against geological time
775 are indicated as Mi.
776

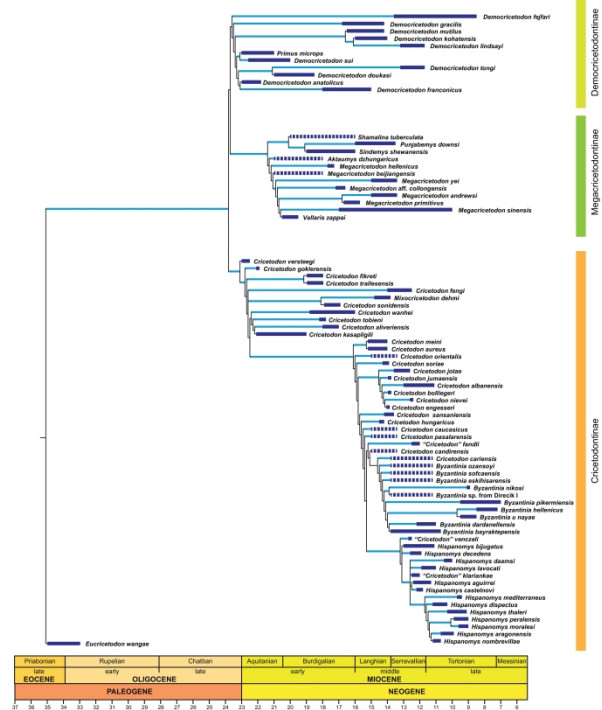


Figure 1. Calibrated strict consensus tree of cricetid rodents and their recorded temporal ranges (black). Dashed bars indicate age uncertainty. Grey bars represent missing ranges and missing ancestral lineages. Biochronological data from Casanovas et al. 2016; de Bruijn 1976; Flynn 2016; Hir et al. 2016; Joniak and de Bruijn 2014; Joniak et al. 2017; Lindsay 2017; Prieto and Rummel 2016; Reichenbacher et al. 2013; Qiu and Li 2016; Ünay et al. 2003; Van der Meulen et al. 2012; Wang et al., 2013; Wessels 2009; Whybrow et al. 1982.

432x307mm (300 x 300 DPI)

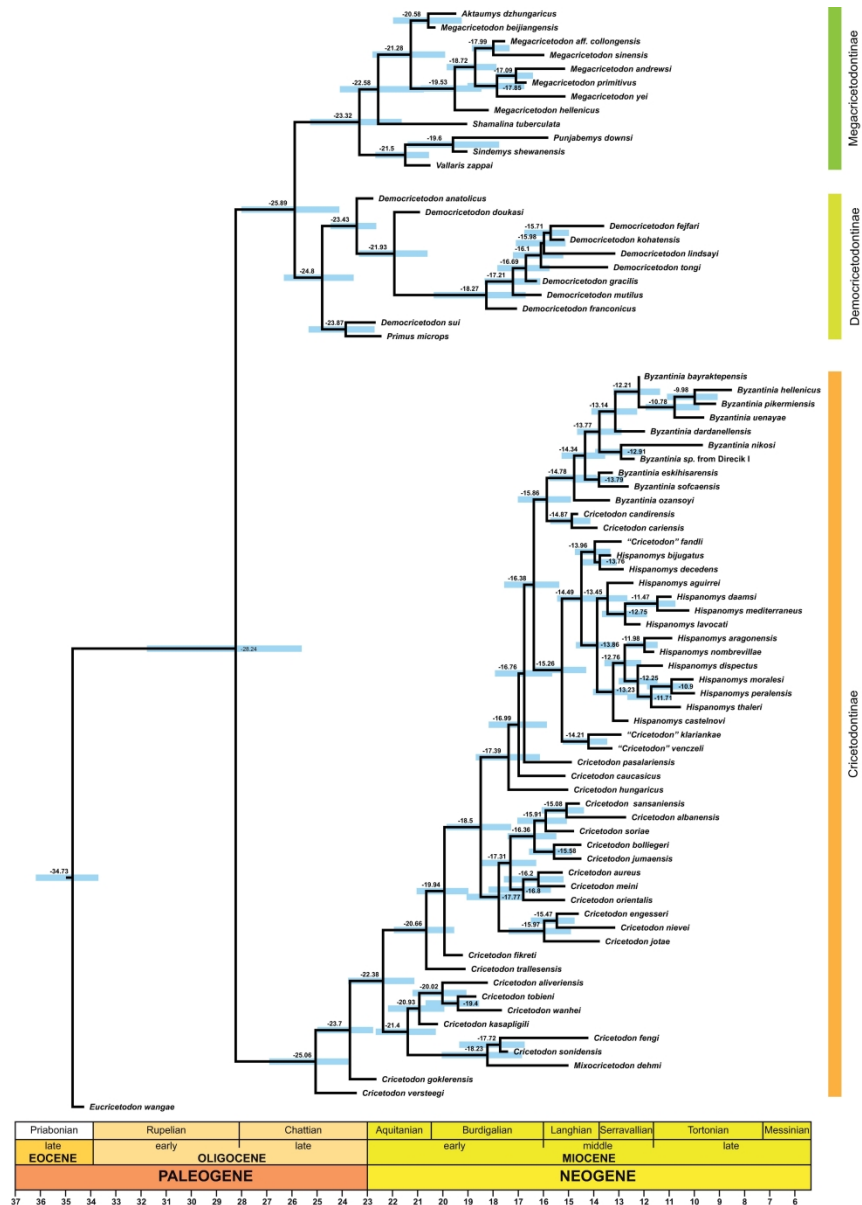


Figure 2. Time-calibrated relaxed-clock Bayesian inference analysis with morphological tip-dating using the fossilized birth–death tree model. Summary of the MCCT depicting the median divergence time estimates for different cricetid lineages against a geological time scale. Numbers at nodes indicate median estimates for the divergence times, and node bars indicate the 95% highest posterior density for divergence times.

274x385mm (300 x 300 DPI)



Figure 3. Ancestral state distributions at each node of the MCCT of cricetid obtained by BBM analysis implemented in RASP against a geological time scale. Pie charts indicate probabilities of alternative ancestral ranges. The colour indicates possible ancestral ranges at different nodes. The schematic map shows the biogeographical areas used in this work and colours correspond to: ochre (southwestern Europe), purple (central Europe), dark red (Greece), orange (Anatolia-Caucasus), green (west central Asia), light blue (Arabian Peninsula), dark blue (southern Asia) and yellow (east-central Asia).

168x409mm (300 x 300 DPI)

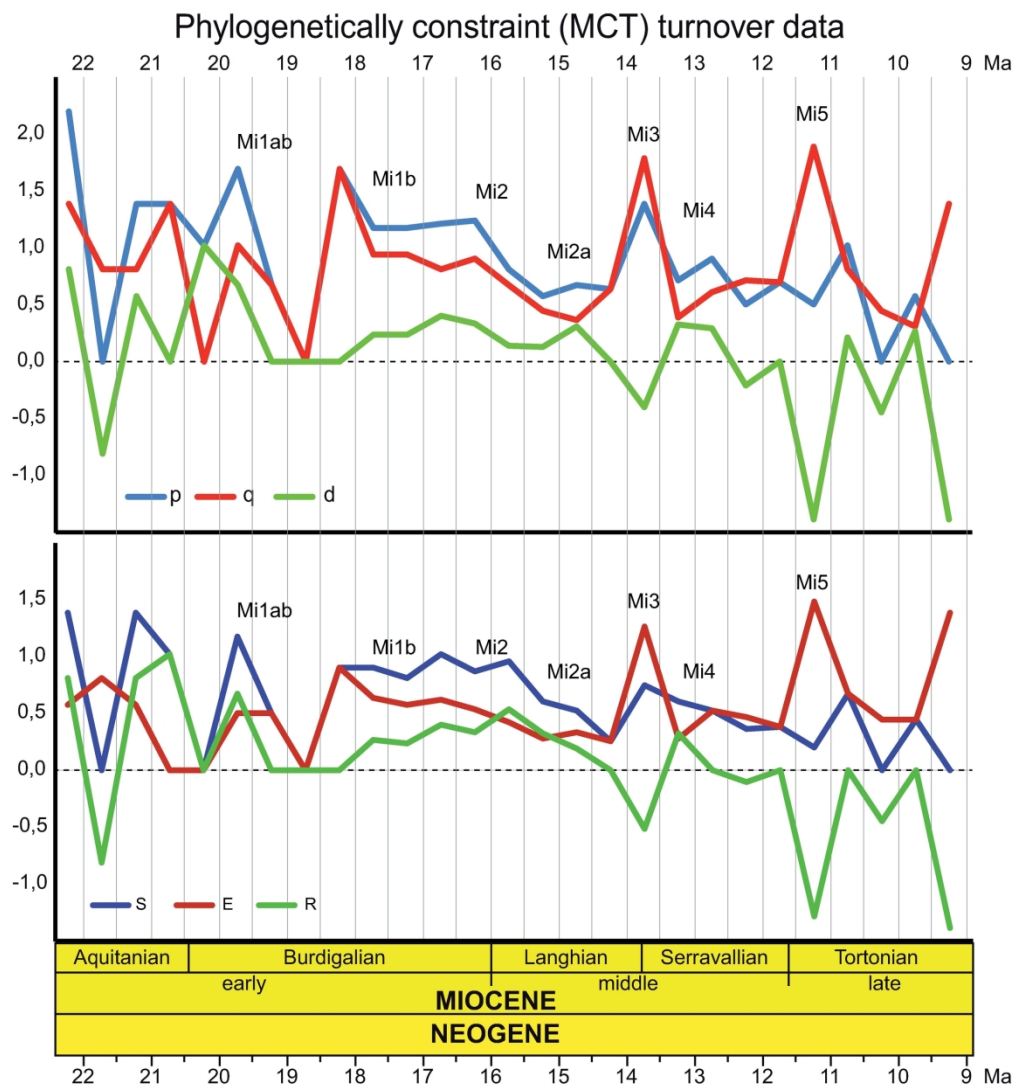


Figure 4. Instantaneous per-capita (top) and deterministic (bottom) rates for Cricetodontinae speciation (p, S), extinction (q, E), and biodiversity change (d, R) calculated using phylogenetically constrained species ranges. Miocene cooling events against geological time are indicated as Mi.

189x205mm (300 x 300 DPI)