

# A New Global Palaeobiogeographical Model for the late Mesozoic and early Tertiary

MARTÍN D. EZCURRA AND FEDERICO L. AGNOLÍN

## APPENDIX 3

### *Treatment of the “Widespread Taxa” Problem*

When a taxon is found in two or more of the employed geographic areas this is considered a “widespread taxon”. Such distributions would be consequence of several different reasons, including missing data and biotic dispersal. Cladistic biogeographers have coped with this problem following three different alternative assumptions (assumption 0, 1 or 2), which differ in the allowed cladogenetic relationships among the areas in which the taxon is distributed (Nelson and Platnick 1981; Zandee and Roos 1987; Sanmartin and Ronquist 2002). The use of these assumptions allows widespread taxa to be included in these analyses (Upchurch et al. 2002).

In the present cladistic palaeobiogeographical analysis some widespread taxa are present, including *Theriosuchus* (Europe and Asia), *Ceratosaurus* (North America and Europe), *Torvosaurus* (North America and Europe), *Allosaurus fragilis* (North America and Europe), and the Mononykinae + *Albertonykus* clade (North America and Asia) (see Section i). With the exception of the genus *Theriosuchus*, all the other taxa are a single widespread species or no sufficient information is currently available to separate specimens at an alfa-taxonomic level. Accordingly, these taxa were treated following the assumption 2 of Nelson and Platnick (1981), which allows a failure to vicariate, extinction, dispersal, or any combination of these events as an explanation for the origin of widespread distributions (van Veller et al. 1999). By contrast, *Theriosuchus* has five distinct recognized species in two of our geographic areas (i.e. *Theriosuchus pusillus*, *Theriosuchus ibericus*, *Theriosuchus*

*guimarotae*, and *Theriosuchus sympiestodon* from Europe; and *Theriosuchus grandinaris* from Asia; Owen 1879; Brinkman 1992; Schwarz and Salisbury 2005; Lauprasert et al. 2010; Martin et al. 2010). This genus can be treated following the assumption 0 of Zandee and Roos (1987), in which the widespread distribution of the taxon is result of a failure to speciate in response to vicariant events affecting other lineages and it is considered as a synapomorphy of the areas in which it occurs (Sanmartin and Ronquist 2002). However, no currently available analysis successfully reconstructed the phylogenetic relationships among the species of *Theriosuchus*, and since the programs used here to conduct the TRAs cannot deal with polytomies, the assumption 2 has been also used for the genus *Theriosuchus*.

#### *Optimal Area Cladograms (OACs) Search Parameters*

The OACs of each time-frame have been reconstructed with the program COMPONENT 2.0 (Page 1993) based on the information provided by the time-sliced cladograms derived from the archosaur semi-strict supertree. The geographic areas (hosts) without associates were interpreted as missing information. A heuristic search employing nearest-neighbor interchanges as branch swapping algorithm was conducted and “losses” was used as the optimal criterion instead of the “leaves added” optimization.

The use of “leaves added” as the optimal criterion during the search of the OACs maximizes the number of codivergences among the geographic areas. However, when some of the employed geographic areas have a poor taxonomic sample, a very likely problem when we are dealing with palaeontological data, the “leaves added” optimization will tend to recover these areas closer to the root of the tree or clade. This is result of the low number of taxa sampled from these areas, which result in a lower probability of codivergences with other areas than that which would be recovered between two well-sampled territories. By contrast, the “losses” optimization criteria is less sensitive to these poorly sampled geographic areas, and will be more effective to reconstruct their palaeobiogeographic affinities

(Ezcurra 2010). Accordingly, we have used the latter optimal criterion in order to perform the search of the OACs.

### *Randomisation Test Analyses*

Reconstructions of biogeographical events for each time-slice were conducted in TreeMap 1.0 (Page 1995) through heuristic searches. Regarding the statistical analyses, a randomisation test which generates random taxa cladograms and reconciles each of them with the phylogenetic cladogram was performed (Page 1991). This test has the goal of determining the probability that the observed biogeographical pattern could have occurred only by chance (Page 1994, 1995). 10,000 randomised replications were conducted for each time slice using the “proportional to distinguishable” algorithm. This randomisation test was conducted in TreeMap 1.0. The topology recovered in the OACs was considered as statistically significant if less than 5% of the replicates ( $p < 0.05$ ) reached the number of codivergences found in the original TRA.

In the following the histograms recovered from the randomisation test analyses are figured for each time slice (Figs. S7-12).

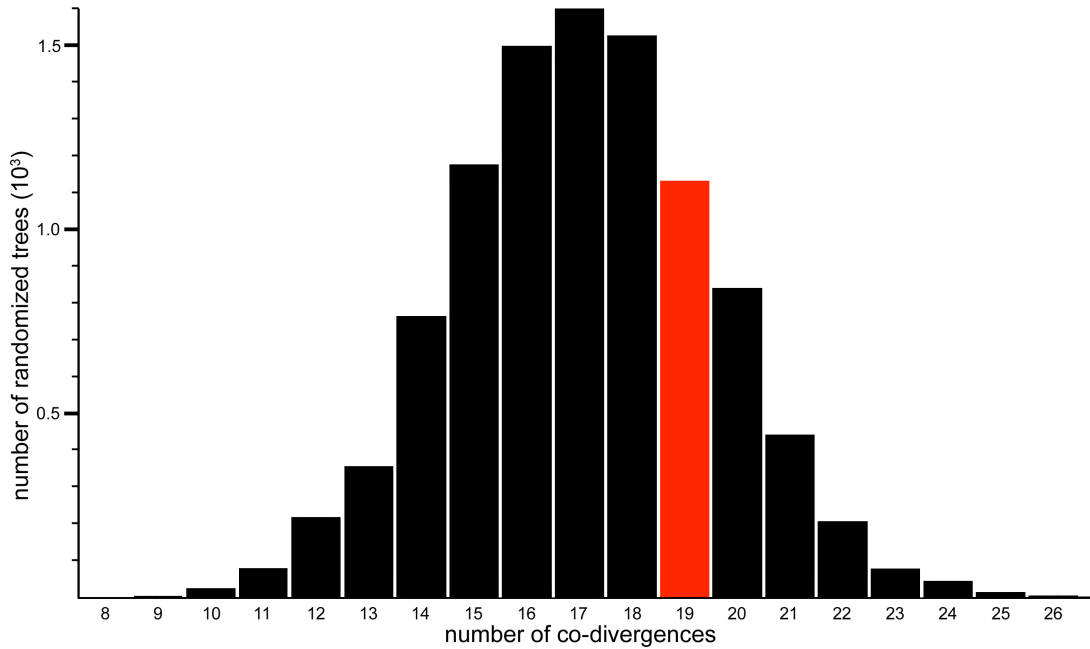


FIGURE S7. Histogram recovered from the randomization test of the Late Jurassic time slice. The values in the horizontal axis are the number of codivergences and in the vertical axis the number of randomized trees generated ( $10^3$ ). The red column depicts the number of codivergent events found in the original analysis ( $p = 0.1131$ ).

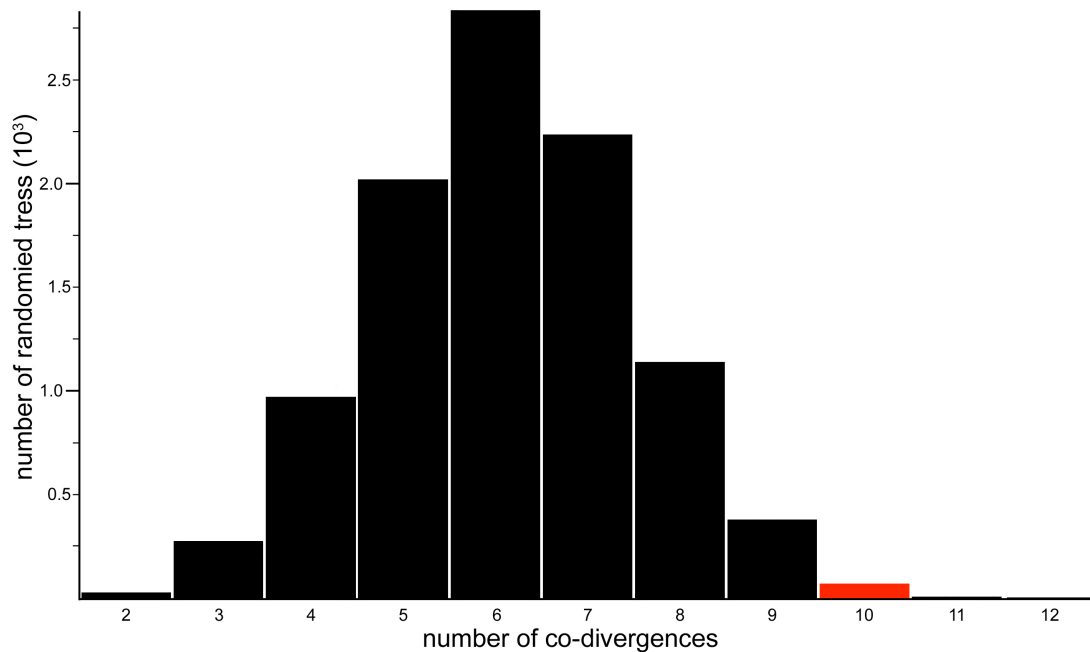


FIGURE S8. Histogram recovered from the randomization test of the Berriasian-Hauterivian time slice. The values in the horizontal axis are the number of codivergences and in the vertical axis the number of randomized trees generated ( $10^3$ ). The red column depicts the number of codivergent events found in the original analysis ( $p = 0.0070$ ).

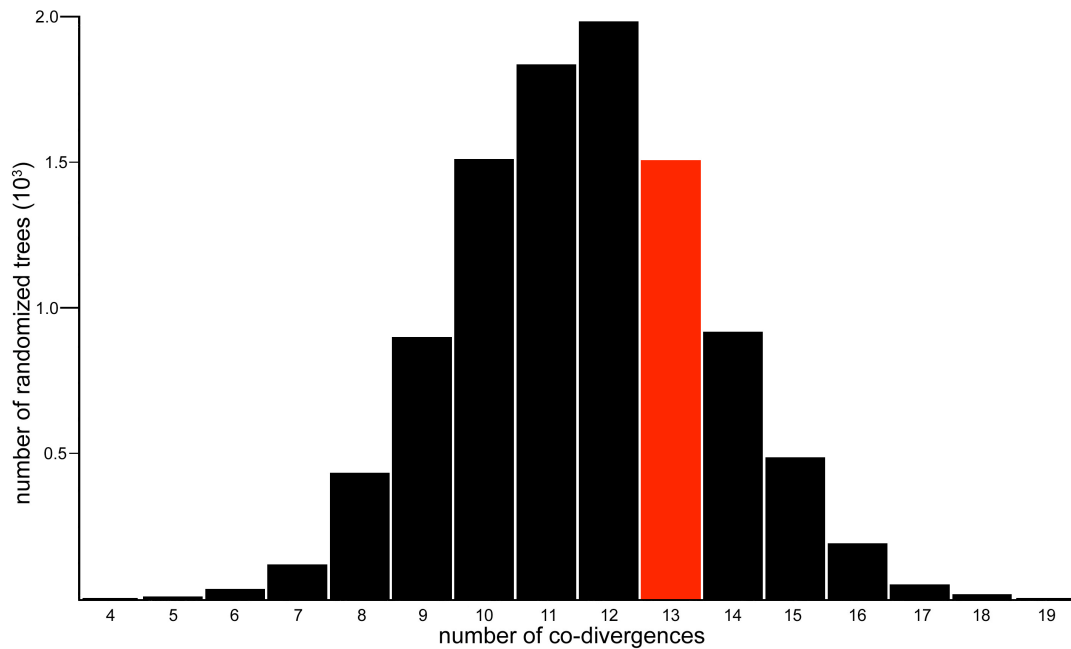


FIGURE S9. Histogram recovered from the randomization test of the Barremian time slice. The values in the horizontal axis are the number of codivergences and in the vertical axis the number of randomized trees generated ( $10^3$ ). The red column depicts the number of codivergent events found in the original analysis ( $p = 0.1508$ ).

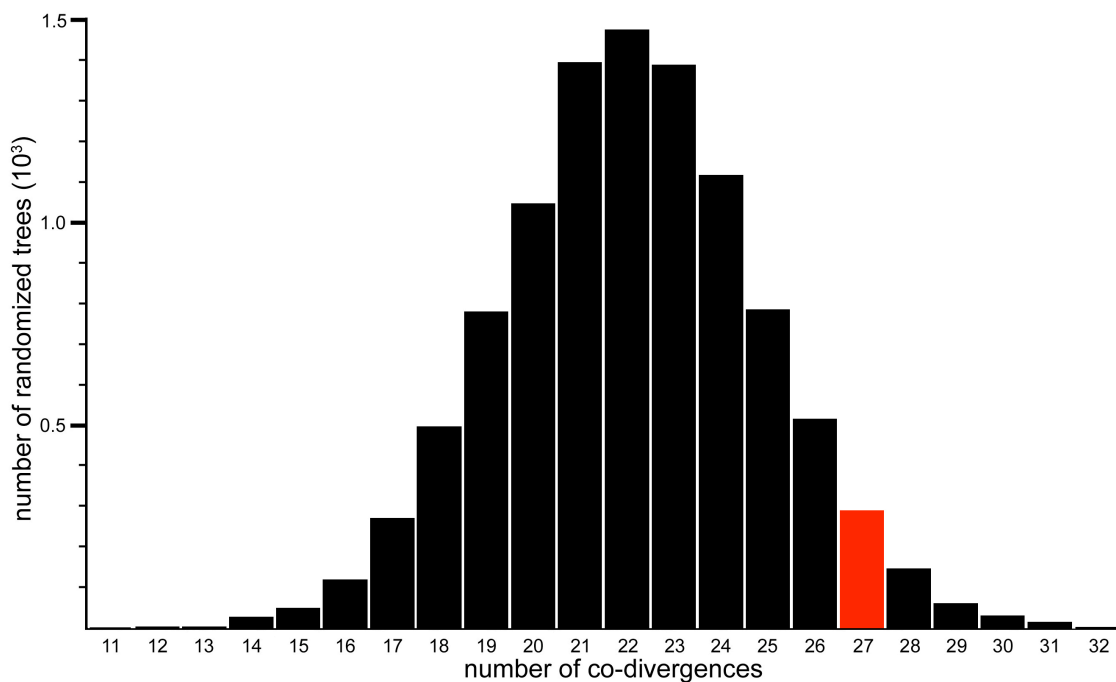


FIGURE S10. Histogram recovered from the randomization test of the Aptian-Albian time slice. The values in the horizontal axis are the number of codivergences and in the vertical axis the number of randomized trees generated ( $10^3$ ). The red column depicts the number of codivergent events found in the original analysis ( $p = 0.0289$ ).

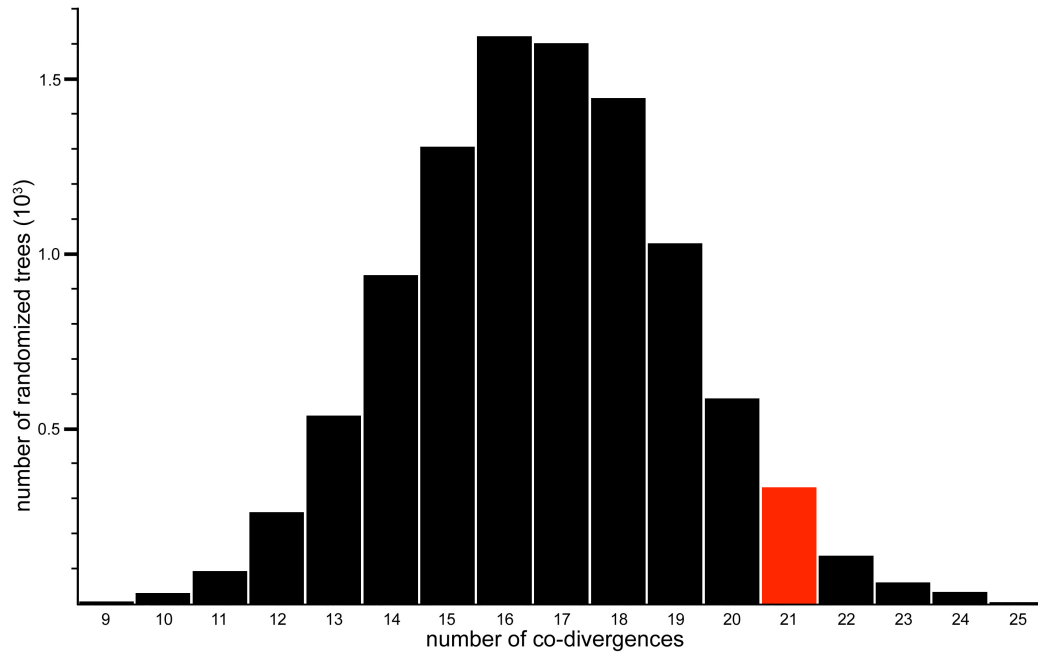


FIGURE S11. Histogram recovered from the randomization test of the Cenomanian-Santonian time slice. The values in the horizontal axis are the number of codivergences and in the vertical axis the number of randomized trees generated ( $10^3$ ). The red column depicts the number of codivergent events found in the original analysis ( $p = 0.0333$ ).

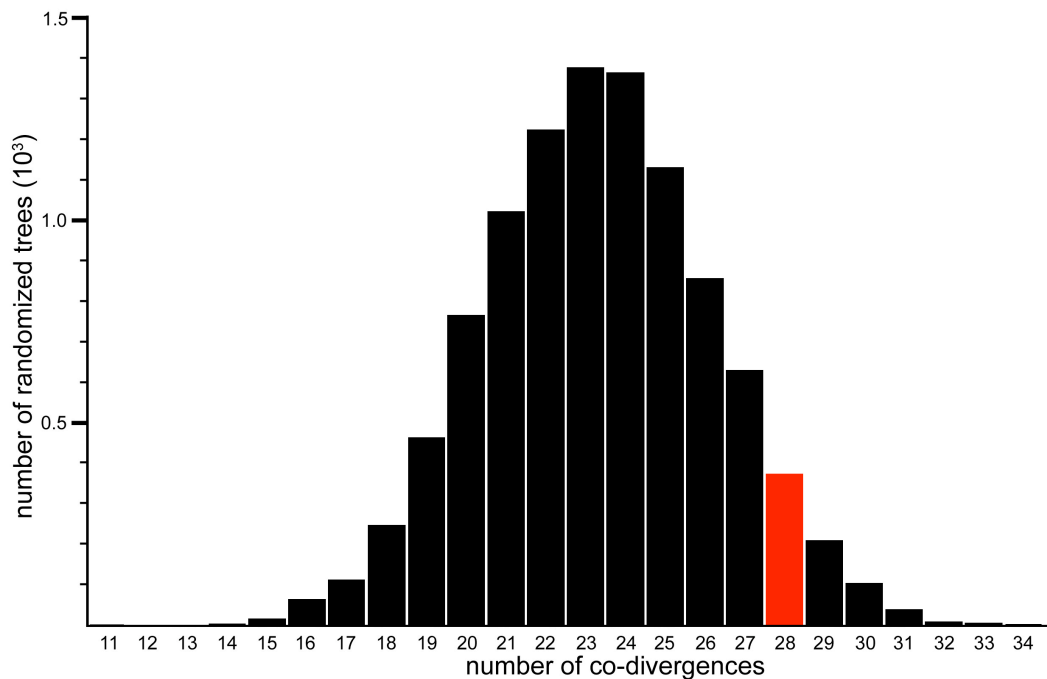


FIGURE S12. Histogram recovered from the randomization test of the Campanian-Maastrichtian time slice. The values in the horizontal axis are the number of codivergences and in the vertical axis the number of randomized trees generated ( $10^3$ ). The red column depicts the number of codivergent events found in the original analysis ( $p = 0.0373$ ).

*Results of the Test of Biogeographical Hypotheses on the Phylogeny of Archosaur Tetrapods*

Test of biogeographical hypothesis: null hypothesis (0); alternative hypothesis (1).

The biogeographical hypothesis consistency index was calculated after the following equation (1):

$$\text{BHCI} = \text{Bhm}/\text{BHs} \quad (1)$$

in which BHCI: biogeographical hypothesis consistency index; Bhm: number of states of the biogeographical hypothesis – 1 (i.e. minimum number of changes in any tree); and BHs: number of steps recovered in the optimization of the biogeographical hypothesis on the tree.

Bonferroni corrections were applied to these multiple analyses. Accordingly, the significance levels will be those described by the following equation (2):

$$\alpha^* = \alpha/n \quad (2)$$

in which  $\alpha$  is 0.05, and  $n$  is the number of test performed.

The bold p-values denote when a statistically significant result was found ( $\alpha^* < \alpha/n$ ).

Abbreviations: AS, Asia; EU, Europe; NA, North America.

Late Jurassic

Bonferroni correction:  $\alpha^* = \alpha/n = 0.05/8 = 0.00625$

Gondwana / Laurasia

p-value: 0.3419

Eurogondwana / Asiamerica

p-value: 0.0087

Eurogondwana / AS / NA

p-value: 0.0098

Euroasia / NA+Gondwana

**p-value: 0.0005**

Euroasia / NA / Gondwana

p-value: 0.0451

AS+Gondwana / NA+EU  
p-value: 0.2435

Asia / Gondwana / NA+EU  
p-value: 0.0166

Eurogondwana+AS / NA  
p-value: 0.0922

*Conclusion.*—The results of this test indicate that the null hypothesis of an adjustment of a dichotomy between North America-Gondwana versus Euroasia in the topology of the Late Jurassic time-sliced supertree was not refuted ( $p < 0.00625$ ). The latter is in agreement with the presence of several closely related taxa between North American and African Late Jurassic assemblages (e.g. *Brachiosaurus-Jiraffatitan*, *Barosaurus-Tornieria*, *Dryosaurus altus-Dryosaurus lettowvorbecki*). However, it must be considered that African Late Jurassic archosaur taxa are not numerous and the remaining Gondwanan landmasses only present a few Late Jurassic archosaur species. Europe shares with North America and Asia several closely related Late Jurassic taxa that suggest the presence of strong biogeographical affinities between these northern areas. However, the scanty African Late Jurassic archosaurs are mostly closely related to North American forms, a fact that lead to find significant the dichotomy between North America and Euroasia. The increment of the Late Jurassic Gondwanan fossil record will probably change this pattern.

Berriasian-Hauterivian

Bonferroni correction:  $\alpha^* = \alpha/n = 0.05/5 = 0.01$

Gondwana / Laurasia  
p-value: 0.2455

Eurogondwana / Asiamerica  
**p-value: 0.0069**

Eurogondwana / Asia / NA  
**p-value: 0.0096**

Euroasia / NA+Gondwana  
p-value: 0.0569



Euroasia / NA / Gondwana  
p-value: 0.0247

*Conclusion.*—The significant ( $p < 0.01$ ) results recovered from these multiple test were those that included Eurogondwana and Asiamerica as a single region or North America and Asia as separate areas. These results perfectly match with that recovered by the TRA and the Eurogondwanan model.

Barremian

Bonferroni correction:  $\alpha^* = \alpha/n = 0.05/2 = 0.025$

Gondwana / Laurasia  
**p-value: 0.0118**

Eurogondwana / Asiamerica  
**p-value: 0.0014**

*Conclusion.*—Equivocal results. The latter would be consequence of a mixed Gondwanan and Laurasian fauna in Europe.

Aptian-Albian

Bonferroni correction:  $\alpha^* = \alpha/n = 0.05/2 = 0.025$

Gondwana / Laurasia  
**p-value: 0.0001**

Eurogondwana / Asiamerica  
**p-value: 0.0001**

*Conclusion.*—Equivocal results. The latter would be consequence of a mixed Gondwanan and Laurasian fauna in Europe.

Cenomanian-Santonian

Bonferroni correction:  $\alpha^* = \alpha/n = 0.05/2 = 0.025$

Gondwana / Laurasia  
**p-value: 0.0001**

Eurogondwana / Asiamerica  
**p-value: 0.0001**

*Conclusion.*—Equivocal results. The latter would be consequence of a mixed Gondwanan and Laurasian fauna in Europe.

Campanian-Maastrichtian

Bonferroni correction:  $\alpha^* = /n = 0.05/2 = 0.025$

Gondwana / Laurasia  
**p-value: 0.0001**

Eurogondwana / Asiamerica  
**p-value: 0.0001**

*Conclusion.*—Equivocal results. The latter would be consequence of a mixed Gondwanan and Laurasian fauna in Europe.

*Linkage between Sea Level and Biogeographical Models*

The biotic connection between Africa and the European region during the Early Cretaceous seems to have been possible through a series of microplates and the Adriatic-Dinaric Carbonate Platform (Bossellini 2002). Accordingly, the presence of a Eurogondwanan fauna would be favoured during a low sea-level period.

We tested the relationships between the Eurogondwana or Gondwana/Laurasia models and the global sea-level through a plot of the CI of the “Eurogondwana and Gondwana/Laurasia characters” and the sea-level (following Miller et al. 2005) against time (Fig. S13). We plot the CI as  $CI^* = 1 - CI$  because in this form a lesser degree of homoplasy, and in consequence a better adjustment of the biogeographical model to the phylogeny, will be recovered when the value is closer to 0. In this regard, if we have peaks of strong biogeographical signal, they will coincide with peaks in the sea-level fall. In order to have in the plot the sea-level and the  $CI^*$  at the same scale, the  $CI^*$  was multiplied by 100. Finally, the lowest value of the sea-level and of the  $CI^* \times 100$  were standardized in the same value, in which the each value of the sea-level was sum by 87.114 (Table S1).

The resultant plot shows that the point of greater difference between the Eurogondwana and Gondwana-Laurasia hypotheses coincides with a drastic sea level fall during the Berriasian-Hauterivian time span (Fig. S13). In addition, this is the time-span in which we suggested the presence of a Eurogondwanan fauna. Accordingly, the presence of a Eurogondwana fauna during the Berriasian-Hauterivian time-span could be, at least in part, related with a drastic fall in the global sea-level. The latter would have allowed a larger aerial exposure of the Carbonate Platforms connecting Afro-Arabia (Bossellini 2002), the probable route of biotic interchange between the European region and Africa.

TABLE S1. Data employed to perform the plot in Fig. S12. In the plot the X-axis is the Ma midpoint and the Y-axis is  $100 \times CI^* \text{ Eurog}$ ,  $100 \times CI^* \text{ Gond/Lau}$ , and  $\text{Sea level} + 87.114$ , respectively. The Sea level is depicted taking into account a modern sea level of 0 meters. Abbreviations: Ap-Al, Aptian-Albian; Berr-Haut, Berriasian-Hauterivian; Cam-Maas, Campanian-Maastrichtian; Cen-San, Cenomanian-Santonian; CI, consistency index;  $CI^*$ ,  $1 - CI$ ; Eurog, Eurogondwana; Gond/Lau, Gondwana/Laurasia; Late Jura, Late Jurassic; m, meters above or below modern sea level; Ma, million of years.

Age	Ma midpoint	$CI^*$ Eurog	$CI^*$ Gond/Laur	$100 \times CI^*$ Eurog	$100 \times CI^*$ Gond/Lau	Sea level (m)	Sea level+87,114
Cam-Maas	72.55	0.933	0.909	93.3	90.9	31.5457729	118.6597729
Cen-San	90.04	0.929	0.917	92.9	91.7	27.588219	114.702219
Ap-Al	112.16	0.941	0.944	94.1	94.4	-	87.00594308
Barremian	127.51	0.909	0.833	90.9	83.3	-	86.07006077
Berr-Haut	138.12	0.667	0.833	66.7	83.3	-	66.70000000
Late Jura	153.29	0.947	0.889	94.7	88.9	-	79.20665281

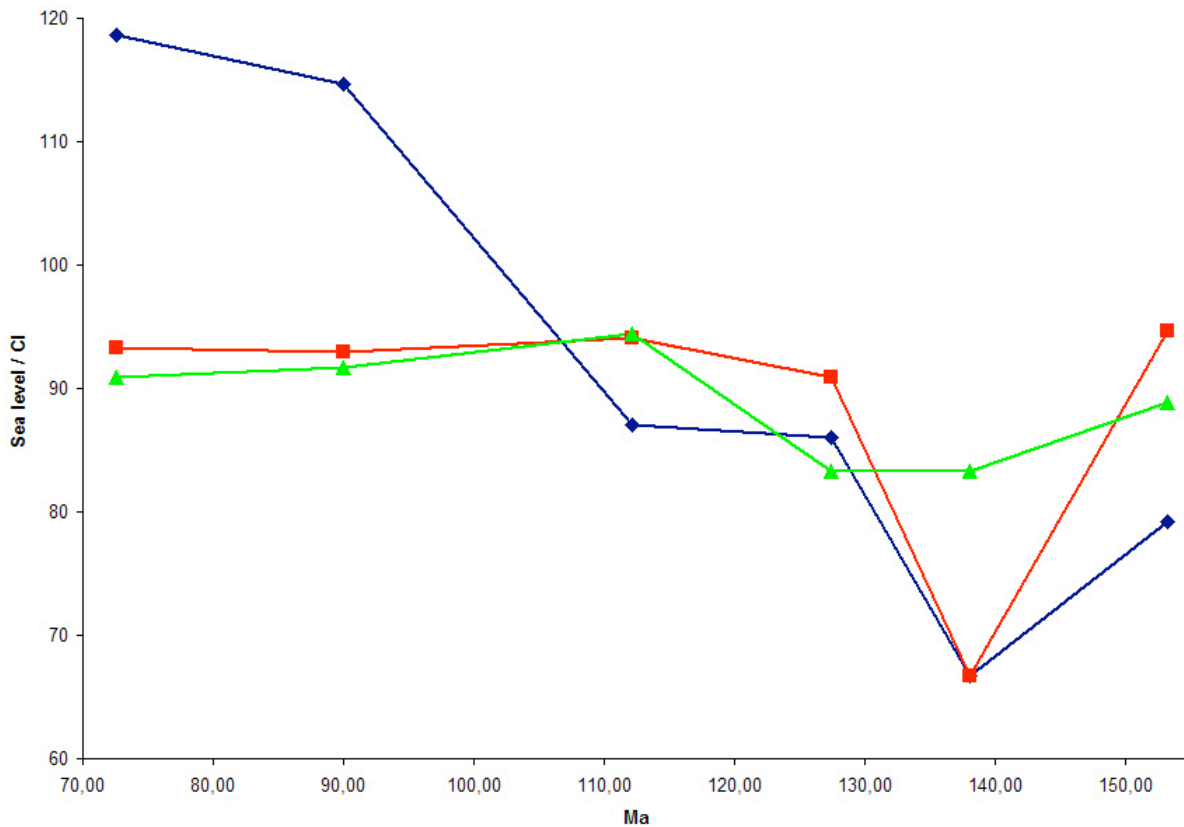


FIGURE S13. Sea level (blue, sea level+87.114) and  $100 \times CI^*$  of the “Eurogondwanan character” (red) and “Gondwana/Laurasia character” (green) against geological time during the Late Jurassic and Cretaceous. Note that the greatest difference between the “Eurogondwanan” and “Gondwana-Laurasia”

characters was found when the sea level reaches its lower-most record. The roughly similar 100xCI\* of both biogeographical characters during post-Barremian times would be consequence of a low degree of endemism between Gondwanan and Laurasian faunas. It must be noted that the 100xCI\* cannot be compared among different time-slices because of tree-shape biases.