



Supplementary Materials for

**Punctuated ecological equilibrium in mammal communities over evolutionary
timescales**

Fernando Blanco*, Joaquín Calatayud, David M. Martín-Perea, M. Soledad Domingo,
Iris Menéndez, Johannes Müller, Manuel Hernández Fernández, Juan L. Cantalapiedra

Correspondence to: fblancosegovia@gmail.com / fernando.blanco@mfn.berlin

This PDF file includes:

Materials and Methods
Figs. S1 to S10
Tables S1 to S12
Caption for Data S1

Other Supplementary Materials for this manuscript include the following:

Data S1:
Sites ages
Functional traits
Functional network MPO sampling
Functional network 75P sampling
Functional network bins
Taxonomic network MPO sampling
Taxonomic network 75P sampling
Taxonomic network bins

Materials and Methods

Database

We collected data for all large mammal species (Proboscidea, Hyracoidea, Primates, Pholidota, Creodonta, Carnivora, Perissodactyla, Artiodactyla) present in the Iberian Peninsula in the last 21 myr. The Neogene-Quaternary mammalian fossil record of the Iberian Peninsula is one of the most continuous vertebrate sequences in the world, with a median temporal gap between localities of 0.1025 myr (calculated with unique age values in order to remove zero age differences). This temporal resolution allowed us to study the assembly processes of deep-time communities with high temporal resolution.

Our dataset includes 167 fossil localities and 2 modern mammalian assemblages, which are representative of the current biogeographic differentiation in the Iberian Peninsula. These 167 fossil localities have been precisely dated using different methods, including, among others, Maximum Likelihood Appearance Event Ordination (ML-AEO), paleomagnetism, electron spin resonance and thermoluminescence (17) (see Dataset S1). When such age estimates were not available (30% of the cases), fossil localities were assigned to local mammal age units (MN units) or local zones (subdivisions of MN units). Yet, in these cases mean temporal precision of occurrences is high (0.8 myr). Our 169 localities provide presence/absence information for 396 mammalian taxa identified at the species level. This 396 x 169 presence/absence matrix represents the starting point for subsequent analyses (13).

Our aim was to track the tempo and mode of shifts in the ecological assembly of Iberian mammalian communities over the last 21 myr. To do so, we condensed the functional role of Iberian mammal species by focusing on three ecomorphological traits (functional traits) that are extensively used in paleobiological research due their ecological relevance: body size, diet and locomotion (Table S1). These traits reflect several facets of the functional role of species in ecosystems, such as habitat use, trophic level, range size, energetic requirements, and resource use (27, 28). We compiled information for these three traits through a review of the primary literature along with surveys of ecomorphological data available in the NOW database (29) and the Paleobiology Database (PBDB) (30). We classified species in eight body size categories modified from (31, 32): B (<1kg), C (1-10kg), D (10-45kg), E (45-90kg), F (90-180kg), G (180-360kg), H (360-1000kg) and I (>1000kg). We excluded the original category A (0-0.1kg), because we did not have species in this body size range in our database. Dietary strategies were grouped as follows: grazer, mixed feeder, browser, frugivore, omnivore, carnivore invertebrates, carnivore, meat bone, and hypercarnivore. Locomotion strategies were organized into 5 categories: aquatic, cursorial, ambulatory, scansorial and arboreal (Table S1).

Based on their body size, diet and locomotion, we assigned species to functional entities (FE), which are unique combinations of these traits (14). A total of 84 FEs were found in our dataset. The assignment of species to functional entities allowed us to tackle the

study of the functional structure of past ecosystems using a taxon-free method (1) (Table S1 and Dataset S1). Note that a certain functional entity may be represented by more than one species in a given site. See the sensibility analyses section below for an assessment of the impact of classification error on our analyses.

Since our focus is on community structure, we need to select fossil communities with a biological signal that represents their original faunal composition. In order to select such localities, we first applied a filter that discarded localities with less than 5 species. Then we applied a ‘minimum presence of orders’ (MPO) selection criterium, which only keeps sites with at least one species in the order Carnivora, and at least 2 species among the orders Perissodactyla, Proboscidea and Artiodactyla in such a way that at least two of these herbivore orders are represented. The philosophy behind the MPO filter is that well-preserved fossil sites need to have representatives of these groups in order to truly reflect the biodiversity and structure of the original ecosystem. Localities in the last 0.005 myr are not expected to fulfill the MPO condition, despite being well-sampled (they lack proboscideans due to the megafauna extinctions of the late Pleistocene; an absence of wild perissodactyls since the Pleistocene in current Iberian Peninsula ecosystems). After the MPO filter, our database includes 83 functional entities in 106 localities, with a resolution of one fossil site every 0.2 myr. Alternative data treatments provide very consistent results (see Sensitivity analysis to localities selection).

Functional assembly vs. taxonomical assembly

In order to compare the tempo of community assembly processes at functional and taxonomic levels, we built the taxonomical version of the described dataset for functional entities. By doing so, we obtained the same fossil sites as for the functional datasets. These datasets included 382 species in 106 localities for the MPO filter.

Network analysis

Input data. To explore the temporal organization of functional and taxonomical assemblages, we applied a community detection procedure borrowed from network theory (33-35). This analysis finds groups of localities sharing similar functional entities (or taxa) regardless of the temporal sequence, allowing us to delve into the functional evolution of communities. We first transformed the fossil occurrence data into bipartite networks. Bipartite networks consist of two different sets of nodes that are not mutually connected. In our case, these two sets of nodes are localities and functional entities (or taxa). A link depicts the presence of a functional entity (or taxa) in a locality. Moreover, for the functional network we weighted the links as the number of species belonging to a functional entity and present in a fossil site. By doing so, we quantified the relative importance of each functional entity in a locality. We did not weight the links in the taxonomic network as we do not have information about the number of specimens per species (Fig. S1 to S3).

Community detection algorithm. Once we generated the networks, we ran the *Infomap* community detection algorithm (36, 37), to find network communities or modules. As explained above, these modules represent groups of localities sharing functional entities (or taxa), which are, in turn, mostly distributed within the localities. That is to say, *Infomap* provides a simultaneous classification of localities and functional entities (or taxa) (38). *Infomap* capitalizes on the minimum description length principle of information theory, which equates finding regularities and compression: The model that finds most regularities in a given set of data can compress the data the most (39). In our case, modules of highly interconnected localities and functional entities (or taxa) form the regularities, and describing the network with an optimal set of communities corresponds to minimizing the description length (36, 38). *Infomap* minimizes the description length using a heuristic algorithm where nodes are stochastically placed into modules and movements are accepted if the description length reduces. We used this algorithm because it performs better than other community detection methods (38, 40). We ran *Infomap* 10,000 times, selecting the best partition as measured by the description or code length (36). This number of runs provided a complete solution landscape. Moreover, modules were consistent across runs and alternative community detection algorithms (see Sensitivity analysis to community detection).

Node characterization. We also characterized the importance of the nodes (i.e. functional entities/taxa and localities) in the definition of particular modules. To do so, we used the IndVal index (38, 41) that considers the affinity and fidelity of a node to the module where it was classified (Data S1). The affinity of a node, A_i , is calculated as the number of links to nodes classified in its same module, X_i , relative to the total number of nodes in the module, Z ,

$$A_i = X_i/Z \text{ (Eq. 1)}$$

For instance, a functional entity (or taxa) occurring in a large proportion of localities within the same module will show a high affinity for it. On the other hand, the fidelity, F_i , of a node is defined by the number of links inside its module, N_i , compared to its total number of links, L ,

$$F_i = N_i/L \text{ (Eq. 2)}$$

For instance, a species only occurring in localities of its own module will have a high-fidelity value to it. IndVal index is calculated as $IndVal = A_i \times F_i$. Thus, a node that links to most of the nodes of its own module (high affinity) and is not connected to nodes in other modules (high fidelity) will show a high IndVal.

To visualize our results, we plotted each locality IndVal score against its age (Fig. 1).

Sensitivity analysis

Sensitivity analysis to localities selection. We selected fossil communities representative of their original faunal composition using the MPO criteria. To explore whether our results were robust to other selection criteria we used a more restrictive selection method. We kept localities with species richness above the 75th percentile of their respective Neogene-Quaternary geological stage (11). We used as time bins the Neogene-Quaternary geological stages and one for the Recent (15 in total). After applying the ‘75th percentile’ (75P) filter, we obtained 79 functional entities in 53 localities, with a resolution of one fossil site every 0.4 myr on average. Moreover, to investigate whether assembly trends at the community level match patterns at the metacommunity level (regional scale), we constructed a new version of our data that resulted from aggregating the original presence absence matrix by 0.5 myr time bins instead of by localities. We obtained 84 FEs grouped in 38 time bins.

We repeated the above procedure for the taxonomic dataset obtaining 331 species in 52 localities in the 75P filter, and 396 species in 38 time bins, respectively. Moreover, to explore the effects of taxonomic resolution on our results, we grouped species by genera, obtaining 229 genera in 106 localities for the MPO filter, 209 genera in 52 localities in the 75P filter, and 234 genera in 38 time bins.

For each of these datasets, we derived a bipartite network and ran *Infomap* as explained above. The modules found were largely consistent across the different datasets both for the functional (Fig. S1) and taxonomic networks (Figs. S2 and S3).

Sensitivity analysis to trait assignment. The ecology of Iberian Neogene and Quaternary large mammals have been extensively studied applying a wide array of methodologies (isotopes, biomechanics, micro- and mesowear, etc) (17, 24, 42-44). Yet, we assessed the impact of a reasonable degree of miscategorization on the assignation of species to functional entities and, therefore, of localities to modules (functional faunas). Our body size categories are broad and conservative, ensuring that our categorization accounts for a potential error in the size estimation (Table S1 and Dataset S1). Thus, we focused on potential errors in the categorization of species’ diet and locomotion mode. To carry out such sensitivity analysis, we first defined the assignation errors of categories following biological rules (Fig. S5A). For instance, a species identified as *hypercarnivore* could randomly change to a *carnivore* or *meat-bone*, but never to *grazer* (Fig. S5A). We randomly applied the predefined changes in categories according to different probability errors (from 0% to 5%, with increments of 1%), for diet and locomotion mode separately. This results in a total of 36 combinations that could yield an approximate maximum of 10% of the species (around 40 species) being affected (Fig. S5). For each combination of error probability, we generated 100 new datasets, making a total of 3600 new datasets. For each of them, we ran the *Infomap* 100 times. Then, we compared the assignation of localities into modules of the original

network with the new network partitions using the similarity index proposed in (45) (Fig. S5B).

Similarity scores with the original MPO network are high (over 0.7) when categorization error affects 3% or less of the species (12 species or less). Given the broad dietary and locomotion categorization used here, and the large body of literature used to inform our trait scoring system, we consider it unlikely that the potential error would affect a higher proportion of species in our dataset. Still, even after introducing unrealistic changes in 10% of the species (around 40 species), similarity never drops below 0.5 (Fig. S5).

Sensitivity analysis to the effect of aggregation. We tested whether the emerging pattern of long-lasting functional faunas could be just a byproduct of merging species into functional entities. To do so, we compared the modularity of observed functional networks against that of null networks where the species were randomly classified into functional entities, keeping constant the number of species belonging to each functional entity. To calculate the network's modularity using the *Infomap* framework, we computed a relative code length by dividing the code length of the network with modules by the code length of the network without modules. When the network lacks a modular structure the index equals 1, whereas it approaches 0 for highly modular networks. To make the index increase with network modularity, we computed the complement of the relative code length,

$$M = 1 - CLm/CL \text{ (Eq. 3)}$$

where CLm represents the code length of the network with modules and CL the code length of the network without modules. Finally, we calculated the p-value of the observed network being more modular than random expectations as the proportion of 99 null networks plus the observed one being more modular than or equal to the observed network (Fig.S4). We used R software (46) for all the analyses, and the scripts are available at (12).

Sensitivity analysis to community detection. We identified two principal sources of uncertainty regarding the detection of network modules. Firstly, several community detection algorithms (*Infomap* included) use stochastic searches to find optimal network partitions. The most common procedure is to run the algorithm an elevated number of times and pick the best quality partition. This, however, brings questions on the number of runs needed to avoid local optima (i.e. obtaining a complete solution landscape) and the consistency of modules across runs of similar quality (45). To approach the first issue, we used a 10-fold cross validation exploring the probability that a test partition has a similarity higher than 0.75 to one of the training partitions ((45), implemented in <https://github.com/mapequation/solution-landscape>). We found that 10000 *Infomap* runs provided a complete solution landscape with ~1 probability for a test partition being more similar than 0.75 to training partitions, both for the functional and taxonomic

networks. To handle the second issue, we explored whether the modules of the best network partition were consistent across the remaining 9999 partitions. For a given module, we computed the probability of finding a module with similarity higher than a given threshold in the remaining partitions (45). We used a similarity threshold of 0.5 (which implies a one-to-one correspondence) and a more conservative threshold of 0.75. Regardless of the threshold, the modules observed in the best partitions were highly consistent across different solutions; in general, they showed probabilities > 0.85 of being observed in different solutions (Fig. 1, Tables S2 and S3).

Finally, we explored whether our results were robust to the choice of alternative community detection algorithms. To do so, we compared *Infomap* results with those of modularity maximization. For the MPO-based functional and taxonomic networks we optimized a modularity index for bipartite networks (47) using the algorithm proposed by (48), as implemented in the *bipartite* R package (49). For each network, we run the algorithm 100 times, selecting the partition with the highest modularity. We used the similarity index described in (45), which ranges between 0 and 1, to compare these partitions with the ones produced by *Infomap*. Our results revealed that both approaches yield similar network partitions with 0.79 and 0.75 similarities for the functional and taxonomic networks, respectively.

Diversification analysis

In order to provide a context of taxonomic turnover to our network community model, we computed speciation and extinction rates through the analysis interval (Fig. S7). To do so, we used the software *PyRate* (50, 51), a Bayesian method that uses occurrence data to simultaneously estimate extinction and speciation rates, interval-specific sampling rate, and speciation and extinction times of species (50). When conducting diversification analyses at regional scale, speciation and extinction rates also include immigration and local, temporal extirpations. To account for this phenomenon, and prior to our analyses, we considered that species that were absent during at least 2 myr were temporally extirpated. The quality of the dataset is high enough to fulfill this assumption (see below). The 2 myr gap was assessed based on midpoint ages from occurrence temporal ranges. The implementation of this condition was achieved in *PyRate* by renaming the species occurrences. For example, if species A is recorded at 16.0, 15.7 and 15.5, and later at 13.2, 13.0, 12.9 and 12.0 Ma, the first three occurrences were assigned to species A_1, and the latter four occurrences to species A_2. We ran *PyRate* for 23 million iterations, sampled every 2,500 iterations to obtain the posterior estimates, and discarded the first 400 iterations (equivalent to the first million) as burn-in. Shifts in speciation and extinction rates were allowed to take place every 0.5 myr (using the argument `-min_dt 0.5`). The sampling rate was allowed to change in 18 temporal bins based on well-established Iberian geological stages defined by these ages (in Ma): 22.4, 20.0, 16.9, 13.7, 12.75, 11.1, 9.7, 8.7, 7.75, 7.0, 4.9, 4.2, 3.2, 2.5, 2.0, 1.15, 0.5, and 0.01. In total, we ran 20 parallel *PyRate* runs on 20 resampled versions of the occurrence ages, retaining 1000 random samples from each run as final results.

Estimates of sampling rates and probability supports our assumption that 2 myr gaps very likely represent true temporal extirpations. The average probability of sampling a species that lasts 1 myr is 80%, emphasizing the exceptional quality of our fossil sites series. Also, and importantly, the probability of not sampling a species during 2 myr is 6% on average, which strongly suggests that 2 myr gaps in the species stratigraphic record in the Iberian Peninsula are actually due to transitory local extirpation rather than to imperfections of the fossil record. Every 1 myr, average speciation and extinction rates translate into a cumulative probability of speciation or immigration (per species) of 0.74, and a probability of extirpation (per species) of 0.69.

Trait-dependent Extinction

A novel multiple trait-dependent extinction (MTE) model was used to test for potential associations between ecological traits and extinction (18). MTE uses a Bayesian framework to assess how extinction rates change across species as a function of one or several traits (18). We analyzed the effect of the three ecological traits and their combinations as functional entities (FE) in 2-myrs bins. We found that the probability of any of the traits—and their combinations—having an effect on extinction (P_μ) is typically below 0.06 across all the 2-myrs bins examined, with just one exception: locomotion shows $P_\mu = 0.19$ in the bin 2-0 Ma, above the threshold of 0.125 which is equivalent to a Bayes factor of 2 given the default prior (0.05) (18).

We also used this approach to test whether species in long-lasting functional faunas had lower extinction risk. We considered a species to belong to one of the three main functional faunas if it was recorded in more than one locality, and at least 60% of them were assigned to that FF by our network analyses. This assignment becomes a trait (with three categories, i.e. the three main functional faunas) that can be incorporated in the analysis. FF-dependent extinction analyses were conducted by 1 myr bins (since the trait has just three states and our statistical power increases as more species enter each category), irrespectively of which categories the species present in a given time bin show. In general, only transitional bins will have species belonging to more than one FF, but some exceptions may occur. For example, we found species in localities of the FF1 that were assigned to a functional entity belonging to FF3 (they had 60% of their records in FF3).

Three bins showed significant P_μ values: 15-13 Ma ($P_\mu = 0.71$), 10-8 Ma ($P_\mu = 0.99$), and 9-7 Ma ($P_\mu = 0.72$). Significance here is guaranteed, because in the three cases P_μ is above the 0.514 threshold representing a Bayes factor of 6 given the default prior (0.05), suggesting a strong extinction selectivity (Fig. S9). A test conducted for the entire analysis interval did not show extinction selectivity ($P = 0.01$), suggesting that there is no association between the species extinction risk and the duration of functional faunas.

Linear models of community functional structure

The joint classification of functional entities and localities into modules allowed us to explore the influence of different local functional diversity measures on the duration of functional faunas. We did so by comparing diversity patterns produced by each specific functional fauna into the localities grouped in their same modules. The expectation is that longer-lasting functional faunas would generate localities with greater functional diversity (52). As a measure of functional diversity, we first use the Shannon's diversity index, which considers both the richness of functional entities and even distribution of species into functional entities:

$$H = - \sum_{i=1}^R p_i \ln(p_i) \text{ (Eq. 4)}$$

where R , is the functional entities richness, and p_i is the proportion of species belonging to the i functional entity. Then, we also considered each constituent part (i.e. richness and evenness) independently, using the Pielou's index to measure evenness. Calculated as,

$$J = H / \ln(R) \text{ (Eq. 5)}$$

where H' comes from Shannon's diversity index.

As above explained, we computed these indices for each locality based only on the functional entities belonging to the same module as the locality. To compare diversity values between modules, we fitted generalized linear models of each diversity index as a function of the module as a categorical variable. We also included the age of the localities, either as an independent explanatory variable or interacting with the module, to discover potential temporal trends beyond modules and changes in temporal trends between modules, respectively. Finally, we generated a model where only age was included and a null model without explanatory variables. We assumed a Gaussian error distribution in the case of Shannon's and Pielou's indices, and a Poisson error distribution with a log-link in the case of functional richness. We used the Akaike's information criterion corrected for small sample size (AICc) to select for the best model. We assumed a difference of two AICc units to consider that two models are significantly different. If two models were equivalent based on AICc we chose the one with fewer parameters as a rule of parsimony. To avoid sampling effects on our assessments, linear models were fit to the richest communities in our dataset (sites above the 75th percentile of sites with similar age). We used R software (46) for all the analyses (12).

Punctuated vs Gradual functional transitions

We explored whether the replacements of functional faunas were abrupt or, on the contrary, followed steady transitions. Under a gradual transition we should expect a steady decrease of the functional entities belonging to the declining functional fauna and a steady increase of the functional entities belonging to the incoming fauna (Fig. S10A). This should translate into transitional periods around the limit separating faunas

with a low proportion of functional entities belonging to each fauna (herein, characteristic functional entities; see (45) and Fig. S10A). On the contrary, an abrupt change should be characterized by the sudden replacement of functional entities, without a detectable transitional period with a reduced proportion of characteristic functional entities (Fig. S10B). Following this rationale, for each of the localities classified in two consecutive functional faunas, we first measured the proportion of characteristic functional entities of each of the two faunas. Then, we explored whether transitional periods of predefined lengths have localities with a low proportion of characteristic functional entities. We generated a transitional categorical variable by classifying localities as transitional or non-transitional if they fall in or out, respectively, of a predefined period around the limit between two faunas (i.e. a transitional bin). Then, we conducted generalized linear models of the proportion of characteristic functional entities as a function of the transitional variable.

If a gradual transition exists, transitional localities should show a lower proportion of characteristic entities and, thus, we expected the transitional variable to capture some variability of this proportion. Contrarily, if shifts of functional faunas were abrupt, the transitional variable should capture little or no variability. Because a hypothetical gradual transitional period may vary in length, we generated 8 versions of the analysis with transitional bins of different duration (from 0.5 to 4 my, with 0.5 my increases). We used these time windows to classify localities and generate the transitional variables, which were subsequently used as explanatory categorical variables in independent models. We repeated this procedure independently for the shifts from FF1 to FF2 (setting the center of the transitional periods at 14.5 Mya) and from FF2 to FF3 (with transitional periods centered at 9 Mya). In all cases, we used a binomial family error and a logit link function, as the dependent variable was a proportion. We also included the module as a covariate to account for potential variations between modules. To assess the importance of the transitional variable, we compared corresponding models with and without this variable by means of Akaike Information Criterion corrected for sample size (AICc). Finally, we computed a delta AICc as AICc of the transitional model minus the AICc of the model without the transitional variable. In this way, $\Delta \text{AICc} < 2$ would indicate a significant effect of the transitional variable, whereas $\Delta \text{AICc} > 2$ would indicate otherwise. Supporting the “abrupt shifts” idea, we found that, regardless of the shift or the length of the transitional period, the models including the transitional period were always significantly worse than the model without this variable (Fig. S6). We used R software (46) for all the analyses.

References and Notes

27. M. Chen, C. A. Strömberg, G. P. Wilson, *PNAS* **116**, 9931-9940 (2019).
28. B. Figueirido, C. M. Janis, J. A. Pérez-Claros, M. De Renzi, P. Palmqvist, *PNAS* **109**, 722-727 (2012).
29. The NOW Community, New and Old Worlds Database of Fossil Mammals (NOW). Licensed under CC BY 4.0. Release [September 2017], retrieved 5-2-2019 from (2019).
30. The Paleobiology database, PBDB. (<https://paleobiodb.org>, 2019).
31. P. Andrews, J. Lord, E. M. N. Evans, *Biol. J. Linn. Soc.* **11**, 177-205 (1979).
32. M. Hernández Fernández, *Quatern. Res.* **65**, 308-323 (2006).
33. A. D. Muscente *et al.*, *Nat. Comm.* **10**, 911 (2019).
34. A. D. Muscente *et al.*, *PNAS* **115**, 5217-5222 (2018).
35. H. K. Nenzen, D. Montoya, S. Varela, *PLoS ONE* **9**, (2014).
36. M. Rosvall, C. T. Bergstrom, *PNAS* **105**, 1118-1123 (2008).
37. D. Edler, M. Rosvall. (See <http://www.mapequation.org>, 2015).
38. R. Bernardo-Madrid *et al.*, *Ecol. Lett.* **22**, 1297-1305 (2019).
39. J. Rissanen, *Automatica* **14**, 465-471 (1978).
40. A. Lancichinetti, S. Fortunato, *Phys Rev E* **80**, 056117 (2009).
41. M. Dufrêne, P. Legendre, *Ecol. Monogr.* **67**, 345-366 (1997).
42. L. Domingo, J. Cuevas-González, S. T. Grimes, M. Hernández Fernández, N. López-Martínez, *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **272**, 53-68 (2009).
43. M. J. Salesa, M. Anton, A. Turner, J. Morales, *Zool. J. Linn. Soc.* **144**, 363-377 (2005).
44. D. DeMiguel, B. Azanza, J. Morales, *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **302**, 452-463 (2011).
45. J. Calatayud, R. Bernardo-Madrid, M. Neuman, A. Rojas, M. Rosvall, *Phys Rev E* **100**, 052308 (2019).
46. R Development Core team, *R Foundation for Statistical Computing*. Vienna, Austria. Internet: <http://www.R-project.org>, (2013).
47. M. J. Barber, *Phys Rev E* **76**, 066102 (2007).
48. S. J. Beckett, *R. Soc. Open Sci.* **3**, 140536 (2016).
49. C. F. Dormann, J. Fründ, N. Blüthgen, B. Gruber, *Open Eco. J.* **2**, (2009).
50. D. Silvestro, N. Salamin, A. Antonelli, X. Meyer, *Paleobiology* **45**, 546-570 (2019).
51. D. Silvestro, N. Salamin, J. Schnitzler, *Methods Eco. Evo.* **5**, 1126-1131 (2014).
52. C. S. Holling, *Annu. Rev. Ecol. Syst.* **4**, 1-23 (1973).

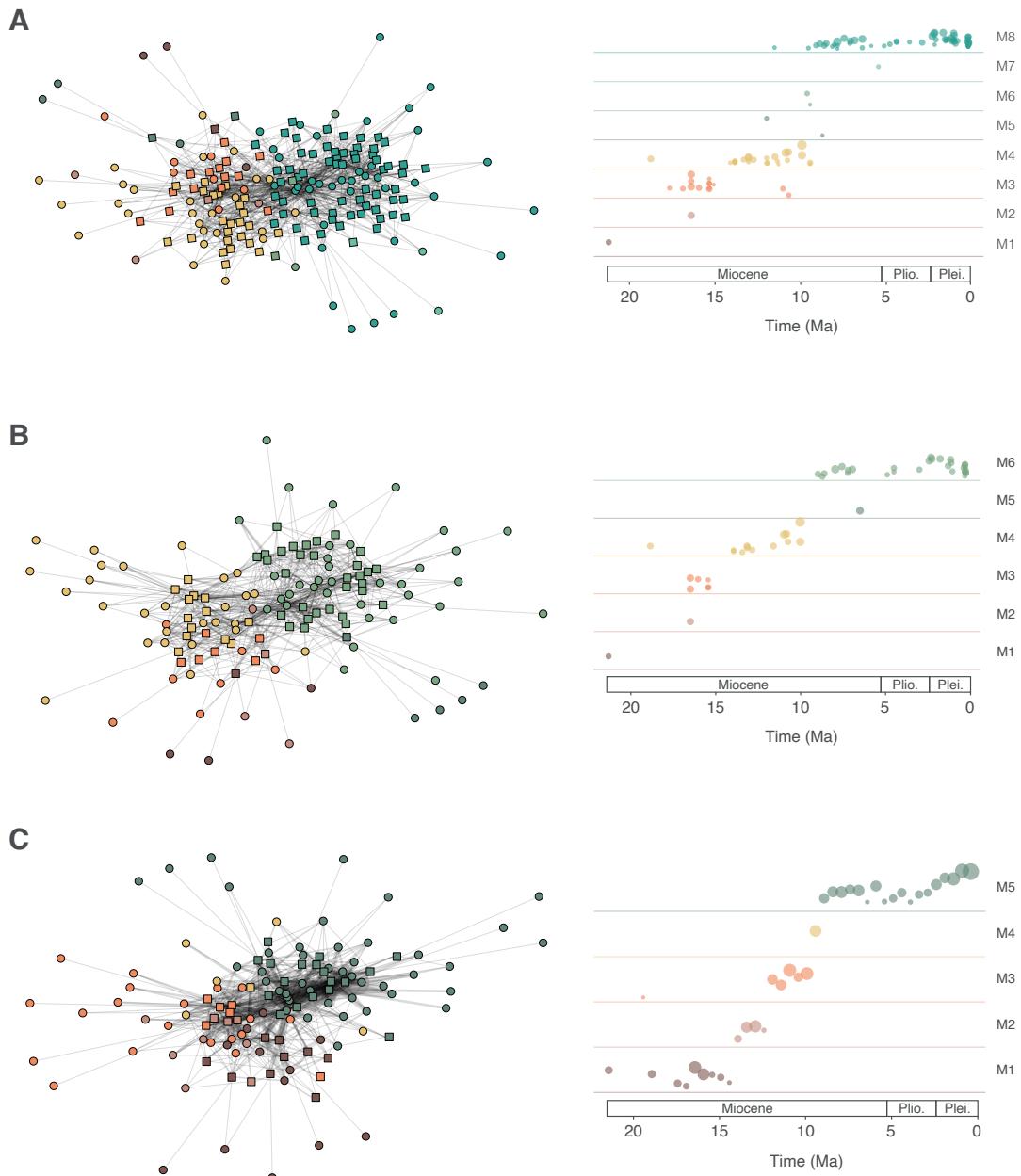


Fig. S1 Functional networks of Iberian mammal communities over the last 21 Ma. Networks (left) along with the plot of the identified Modules (right), which are represented as colored groups of localities (dots). The community level schemes are: minimum presence of orders (MPO, A), species richness above the 75th percentile (75P, B), and time bins (metacommunity, C) over time (Ma=Million years). In the module graphs (right), species richness of each locality is represented by the size of the dot, and its height corresponds to the importance of the locality in the formation of its module measured by the IndVal index (see Supplementary material). In the network graphs, squares represent the different functional entities. Plio.= Pliocene, Plei.= Pleistocene.

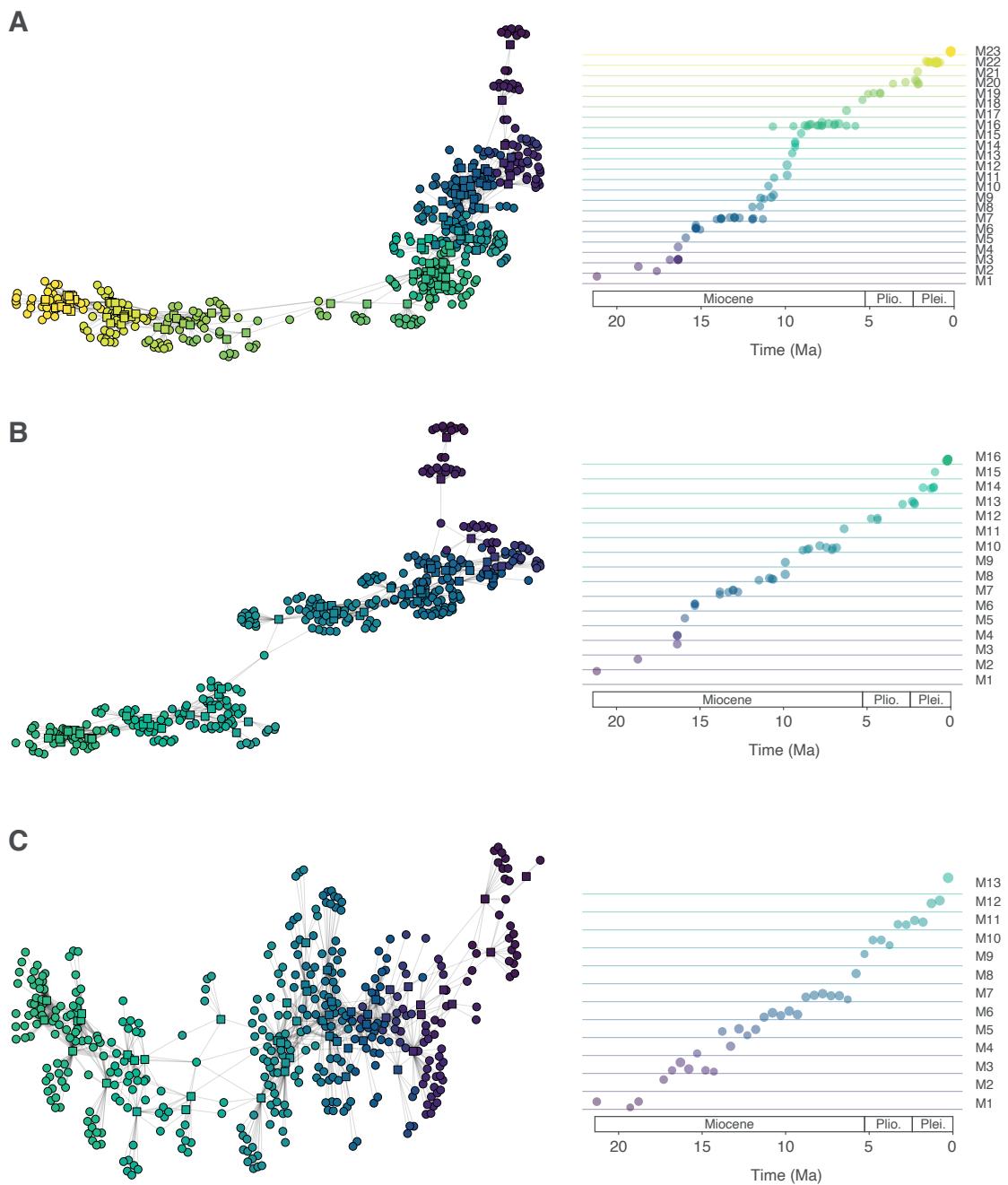


Fig. S2 Species networks of Iberian mammal communities over the last 21 Ma. Networks (left) along with the plot of the identified Modules (right), which are represented as colored groups of localities (dots). The community level schemes are: minimum presence of orders (MPO, A), species richness above the 75th percentile (75P, B), and time bins (metacommunity, C) over time (Ma=Million years). In the module graphs (right), species richness of each locality is represented by the size of the dot, and its height corresponds to the importance of the locality in the formation of its module measured by the IndVal index (see Supplementary material). In the network graphs, squares represent the different functional entities. Plio.= Pliocene, Plei.= Pleistocene.

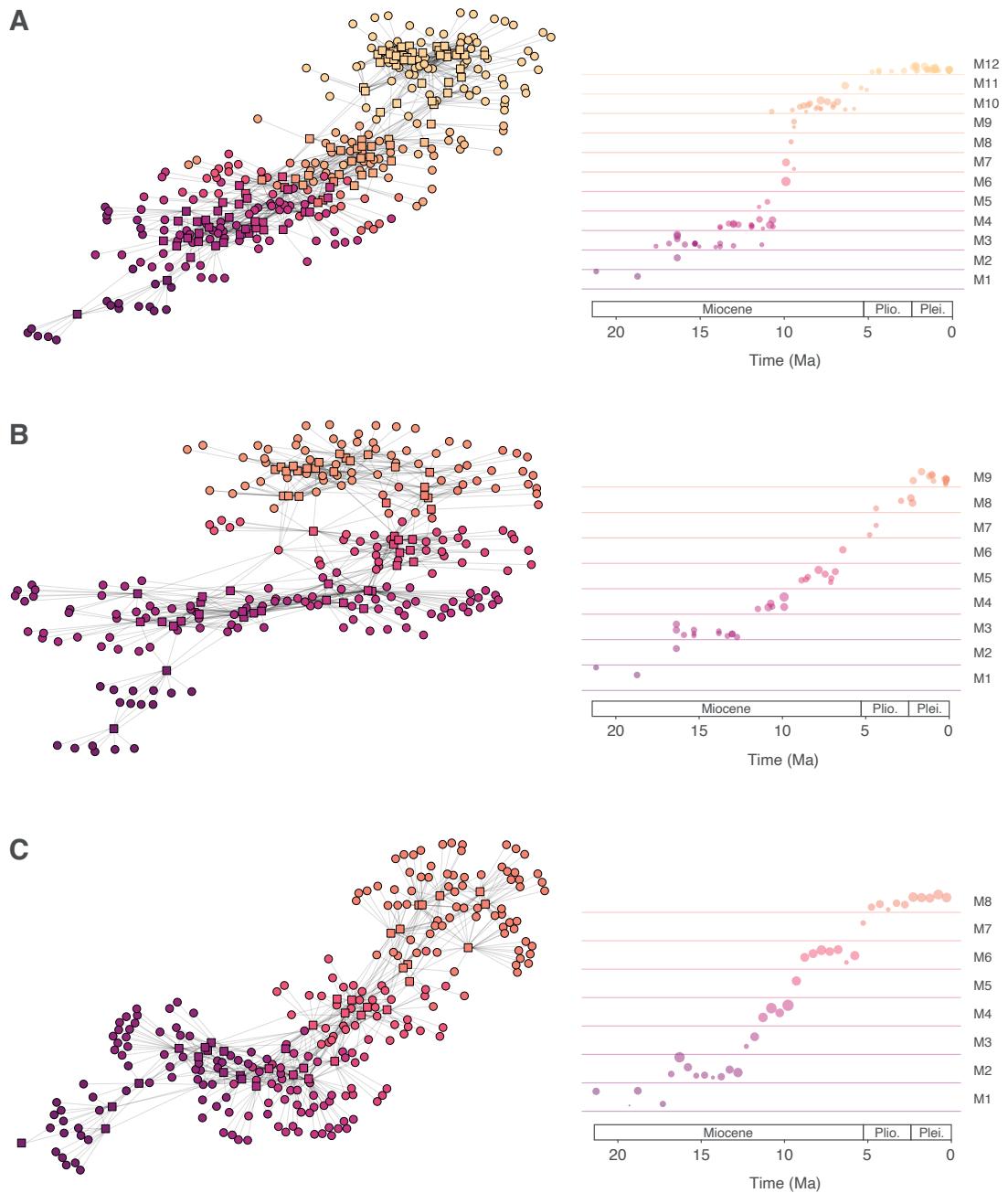


Fig. S3 Genus-level networks of Iberian mammal communities over the last 21 Ma. Networks (left) along with the plot of the identified Modules (right), which are represented as colored groups of localities (dots). The community level schemes are: minimum presence of orders (MPO, A), species richness above the 75th percentile (75P, B), and time bins (metacommunity, C) over time (Ma=Million years). In the module graphs (right), species richness of each locality is represented by the size of the dot, and its height corresponds to the importance of the locality in the formation of its module measured by the IndVal index (see Supplementary material). In the network graphs, squares represent the different functional entities. Plio.= Pliocene, Plei.= Pleistocene.

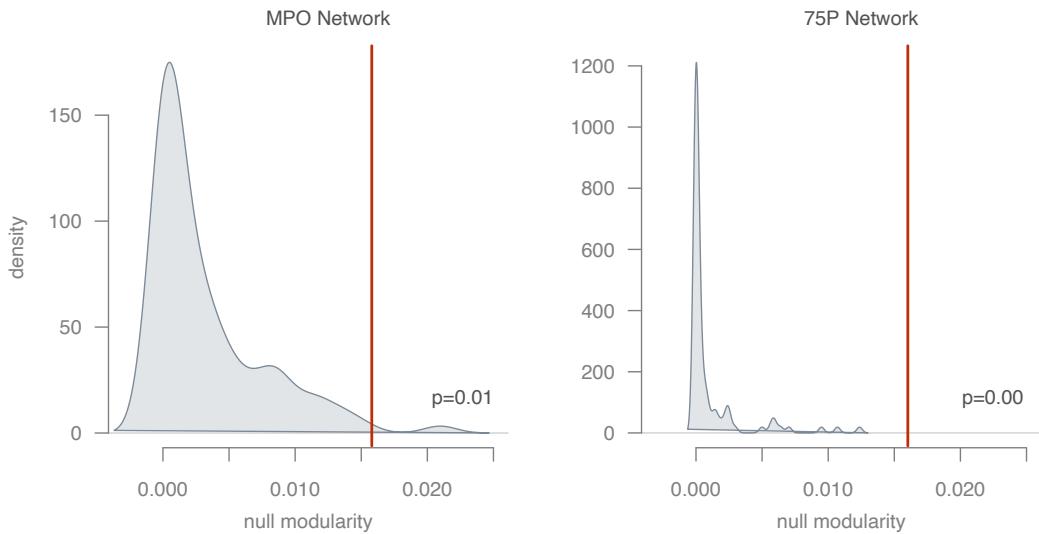


Fig. S4 Observed and simulated modularity of our functional-entities-based networks. Gray density diagrams show the modularity scores of 100 random networks compared with the observed network (red vertical line) for the two community level schemes: minimum presence of orders (left) and species richness above the 75th percentile (right).

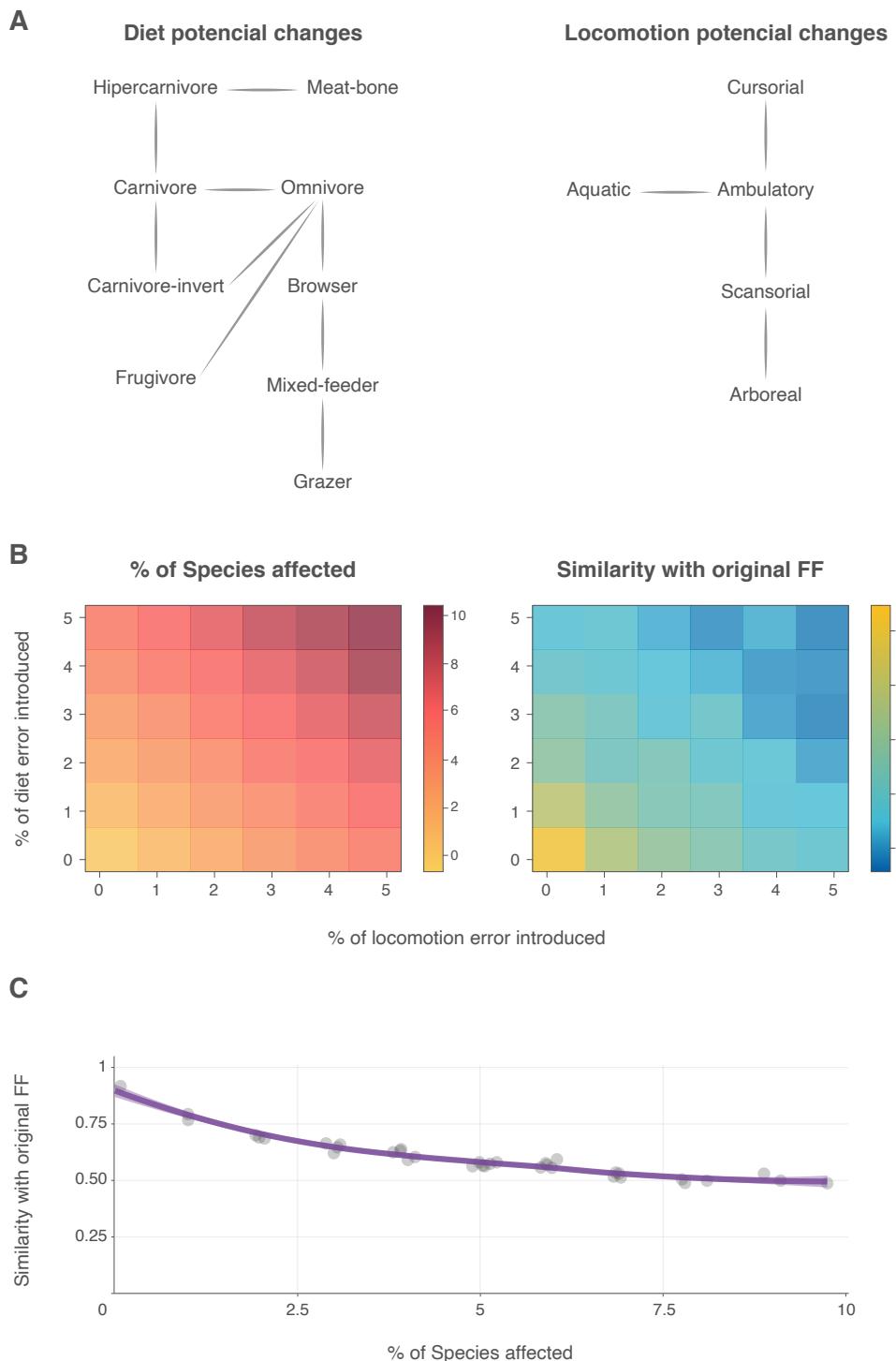


Fig. S5. Sensitivity Analysis of locomotion and diet categorization. Potential changes between trait categories following a biological meaning (A). Proportion of species affected by functional assignation with systematic error introduced for diet and locomotion (B, left). Similarity changes in module's localities composition compared to the original MPO functional faunas with different levels of systematic error introduced (B, right). Similarity of functional faunas generated by the introduction of different degrees of error in functional assignation with the original MPO functional network (C). We show averaged values over 100 replicates of each parameter combination.

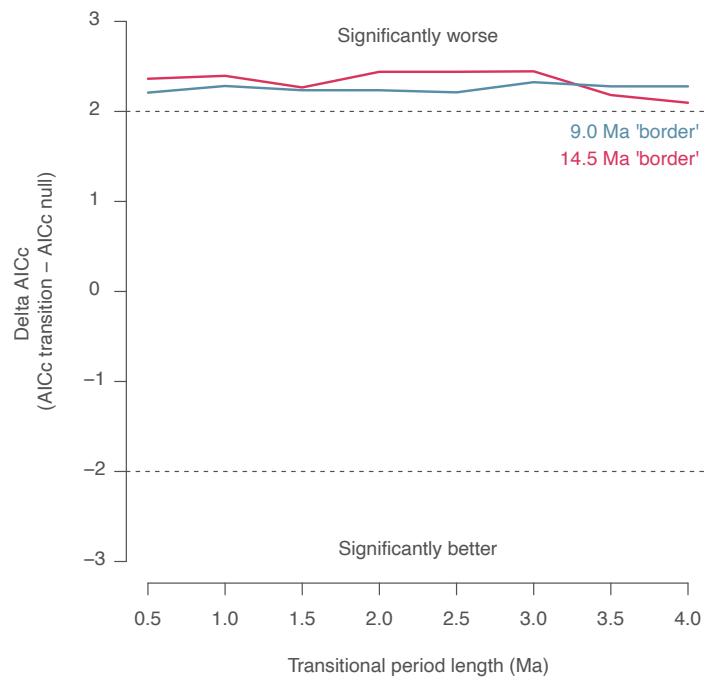


Fig. S6. Delta AIC_c of models with transitional variables compared to those without this variable for transitions between FF1 and FF2 (red), and FF2 to FF3 (blue). Delta AIC_c values are shown for different lengths of the transitional window (x-axis).

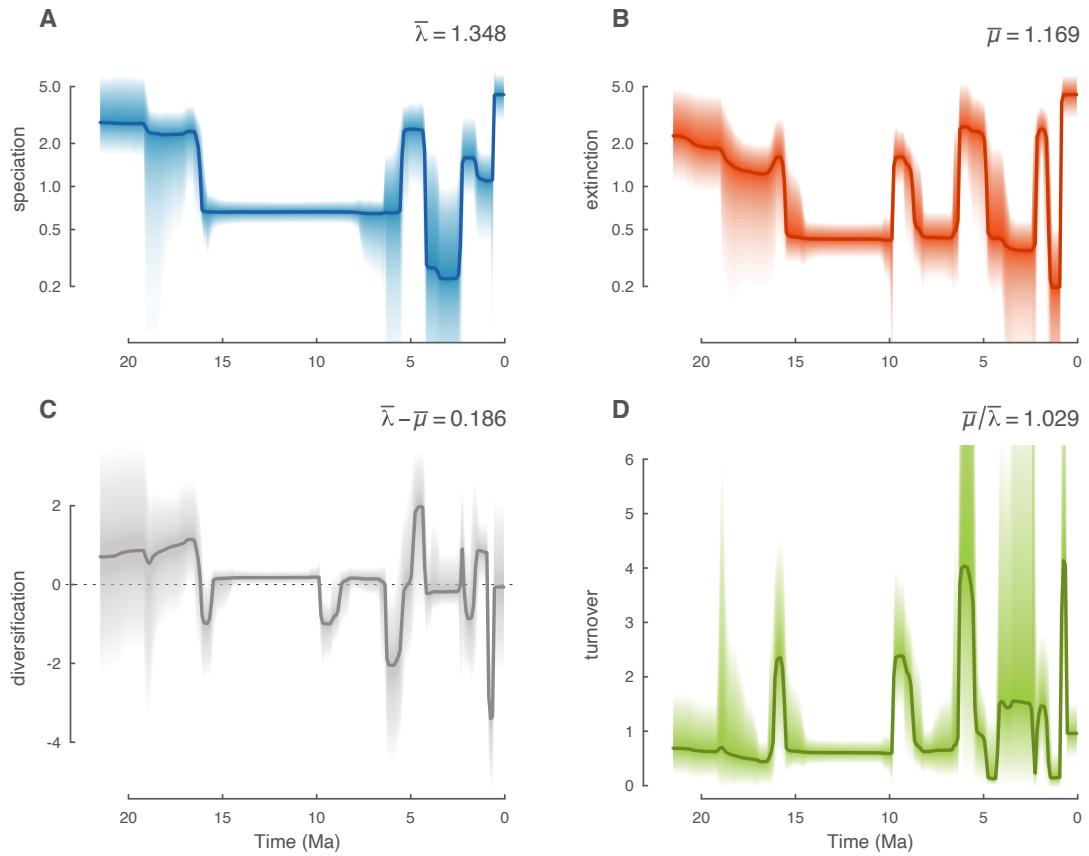


Fig. S7 Taxonomic evolution of Iberian faunas. Evolution of speciation (A), extinction (B), diversification (C) and turnover (D) rates over time. Shaded regions are the 95% confidence interval. Mean speciation (λ) and extinction rate (μ) are shown.

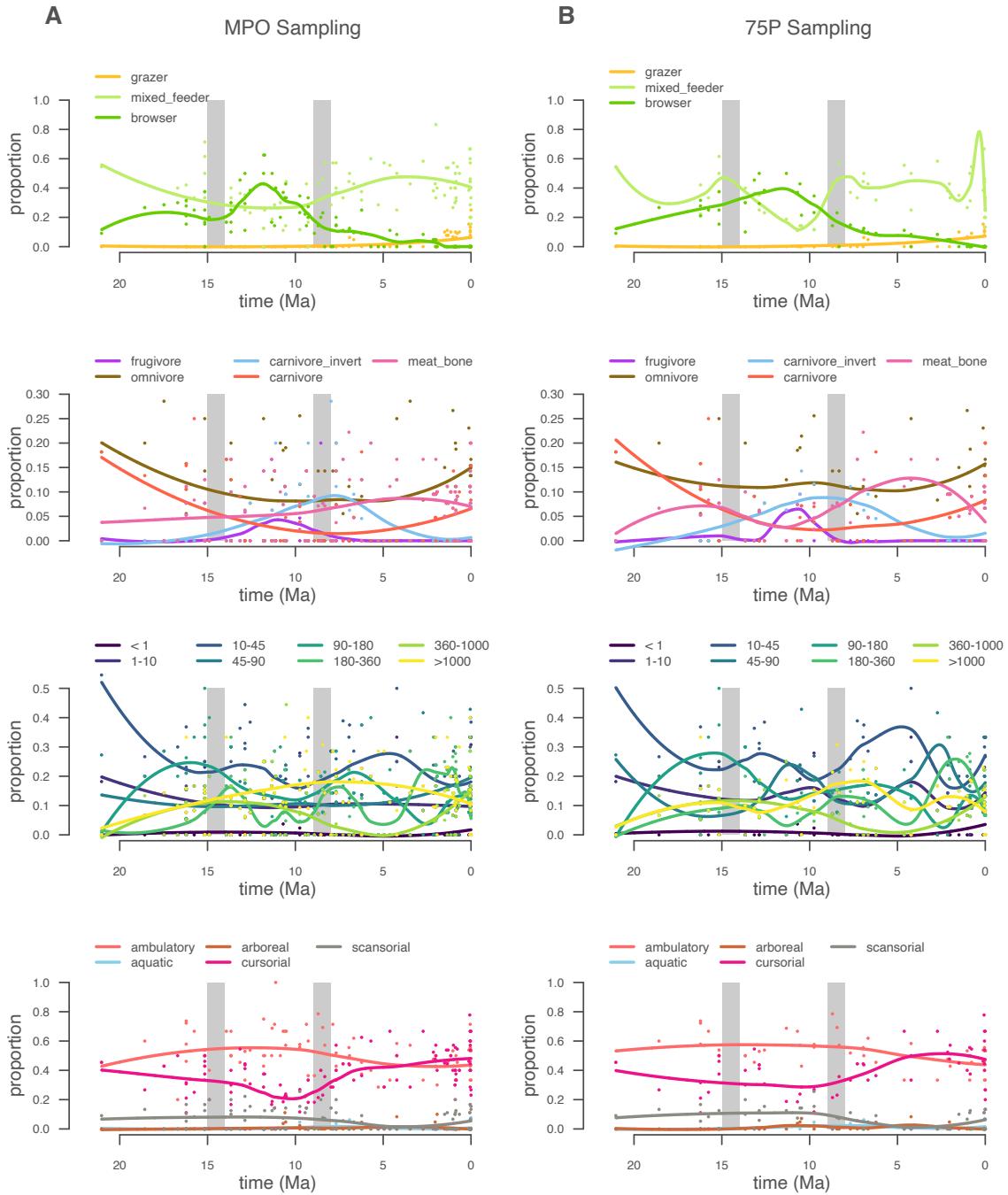


Fig. S8 Evolution of individual traits proportion in the localities (dots) through time, using MPO sampling (A) and 75P sampling (B). Vertical gray bars show the transitions between FF1-FF2 and FF2-FF3 (as in Fig. 1).

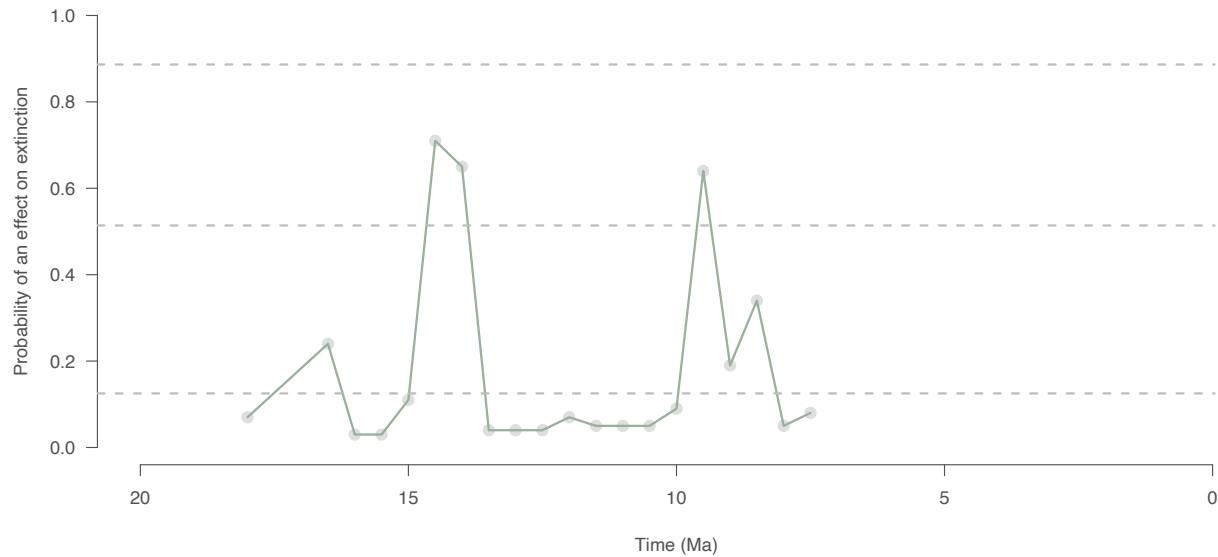


Fig. S9 Probability of an effect of module association on extinction probability (P_{μ} , effect of the trait on extinction risk) across the analysis interval. Dots represent time bins where species associated with at least two modules are found. Dotted lines show the P thresholds for Bayes factors 2, 6 and 10.

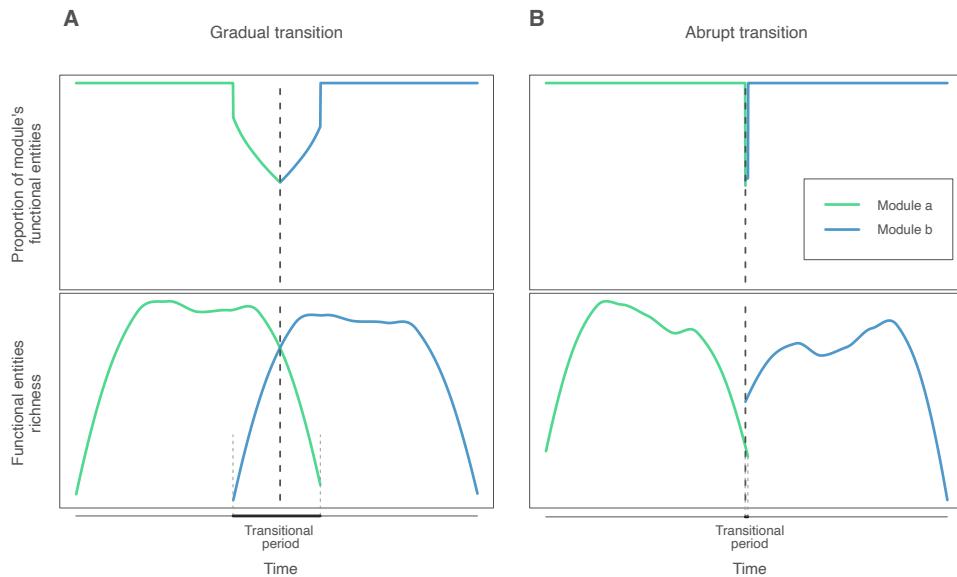


Fig. S10. Hypothetical scenarios expected under a gradual (A) or abrupt (B) transition between functional faunas. In the top row: the proportion of functional entities associated with a module (or functional fauna). In the bottom row: the richness of functional entities belonging to each module. The thick dotted line depicts the limit between both functional faunas.

Table S1. Explanation and categorization of functional traits used in the analysis.

Functional traits	
Body Size	Explanation
B	<1Kg
C	1-10Kg
D	10-45Kg
E	45-90Kg
F	90-180Kg
G	180-360Kg
H	360-1,000Kg
I	>1,000Kg
Diet	Explanation
Grazer	Consumes graze as their primary food
Mixed feeder	Consumes a mix of graze and shrubs
Browser	Consumes shrubs and small trees as their primary food
Frugivore	Consumes fruits as their primary food
Omnivore	Consumes a mix of animals and plants
Carnivore invertebrates	Consumes invertebrate animals as their primary food
Carnivore	Consumes animals as their primary food
Meat bone	Consumes bones as their primary food
Hypercarnivore	Consumes a large amount of animals as their primary food
Locomotion	Explanation
Ambulatory	Slow to medium range movement. Not capable of keeping a constant speed for a long distance
Cursorial	Capable of keeping a constant speed for a long distance
Arboreal	Spends most of the time in trees
Scansorial	Capable of climbing. Spends time both on trees and ground
Aquatic	Spends most of the time in the water

Table S2. Probability of obtaining the same module partition with similarity >0.5 and >0.75 (Robustness) for the functional network.

MPO functional	Robustness	
Module	0.5	0.75
1	0.84	0.79
2	0.4	0.39
3	0.96	0.88
4	0.91	0.85
5	0.98	0.52
6	0.95	0.88
7	0.43	0.42
8	1	0.96

Table S3. Probability of obtaining the same module partition with similarity >0.5 and >0.75 (Robustness) for the taxonomic network.

MPO taxonomic	Robustness	
Module	0.5	0.75
1	0.94	0.44
2	0.86	0.58
3	0.7	0.69
4	0.85	0.2
5	1	0.99
6	1	0.85
7	1	1
8	0.82	0.82
9	0.63	0.62
10	1	1
11	1	1
12	1	1
13	0.7	0.62
14	1	1
15	1	1
16	1	1
17	1	1
18	0.54	0.54
19	1	1
20	0.49	0.48
21	0.5	0.24
22	0.98	0.85
23	0.93	0.89

Table S4. Linear model comparison for functional richness using the 75P sampling. M= Modules, AICc= Akaike information criterion, R2= R square value, K= number of estimated parameters in the model.

M names	K	AICc	Delta_AICc	R2	Intercept	M 4	M 6	Age	M 4*Age	M 6*Age	M 3 - M 4	M 3 - M 6	M 4 - M 6	
M * Age	6	251.934	0	0.49	-3.309	7.078	5.679	0.3	-0.421	-0.32	-4.195	-3.492	0.703	
M + Age	4	252.039	0.105	0.42	1.938	0.804	0.473	-0.035	-	-	-	0.804'	-0.473	0.331
M	3	254.048	2.114	0.36	1.386	0.924	0.929	-	-	-	0.924'	0.929'	-0.005	
Age	2	264.929	12.995	0.18	2.423	-	-	-0.029	-	-	-	-	-	
Null	1	275.724	23.789	0	2.239	-	-	-	-	-	-	-	-	

" : p-value < 0.01

Table S5. Linear model comparison for functional richness using the MPO sampling. M= Modules, AICc= Akaike information criterion, R2= R square value, K= number of estimated parameters in the model.

M names	K	AICc	Delta_AICc	R2	Intercept	M 4	M 8	Age	M 4*Age	M 8*Age	M 3 - M 4	M 3 - M 8	M 4 - M 8
M * Age	6	491.088	0.592	0.3	-0.452	3.149	2.64	0.115	-0.178	-0.145	-1.906	-1.628	0.279
M + Age	4	490.496	0	0.27	1.735	0.56	0.451	-0.029	-	-	-0.56"	-0.451	0.109
M	3	493.462	2.966	0.24	1.293	0.647	0.805	-	-	-	0.647"	0.805"	-0.158
Age	2	498.955	8.459	0.19	2.211	-	-	-0.037	-	-	-	-	-
Null	1	526.84	36.344	0	1.977	-	-	-	-	-	-	-	-

" : p-value < 0.01

Table S6. Linear model comparison for functional richness using bins. M= Modules, AICc= Akaike information criterion, R2= R square value, K= number of estimated parameters in the model.

M names	K	AICc	Delta_AICc	R2	Intercept	M 3	M 5	Age	M 3*Age	M 5*Age	M 1 - M 3	M 1 - M 5	M 3 - M 5
M * Age	6	193.479	0	0.6	0.126	5.235	2.796	0.085	-0.359	-0.172	-1.922"	-1.211	0.711
M + Age	4	199.219	5.74	0.51	3.06	0.274	-0.123	-0.091	-	-	-0.274	0.123	0.398
M	3	214.074	20.595	0.37	1.564	0.705	0.992	-	-	-	0.705"	0.992"	-0.288
Age	2	198.688	5.209	0.47	2.905	-	-	-0.074	-	-	-	-	-
Null	1	253.649	60.17	0	2.312	-	-	-	-	-	-	-	-

" : p-value < 0.01

Table S7. Linear model comparison for functional diversity using the 75P sampling.
M= Modules, AICc= Akaike information criterion, R2= R square value, K= number of estimated parameters in the model.

M names	K	AICc	Delta_AICc	R2	Intercept	M 4	M 6	Age	M 4*Age	M 6*Age	M 3 - M 4	M 3 - M 6	M 4 - M 6
M * Age	7	28.5	1.649	0.52	-2.174	5.381	4.461	0.224	-0.307	-0.248	-3.28	-2.762	0.518
M + Age	5	26.851	0	0.51	1.858	0.748	0.457	-0.035	-	-	-	0.748"	-0.457
M	4	29.606	2.755	0.46	1.318	0.862	0.898	-	-	-	-	0.862"	0.898"
Age	3	44.946	18.094	0.25	2.349	-	-	-0.037	-	-	-	-	-
Null	2	58.138	31.287	0	2.099	-	-	-	-	-	-	-	-
": p-value < 0.01													

Table S8. Linear model comparison for functional diversity using the MPO sampling.
M= Modules, AICc= Akaike information criterion, R2= R square value, K= number of estimated parameters in the model.

M names	K	AICc	Delta_AICc	R2	Intercept	M 4	M 8	Age	M 4*Age	M 8*Age	M 3 - M 4	M 3 - M 8	M 4 - M 8
M * Age	7	102.582	0	0.35	-0.883	2.894	2.968	0.138	-0.156	-0.174	-	1.799"	1.754"
M + Age	5	107.679	5.096	0.3	1.532	0.527	0.511	-0.023	-	-	-	0.527"	-0.511
M	4	108.043	5.461	0.29	1.193	0.593	0.78	-	-	-	-	0.593"	-0.78"
Age	3	116.715	14.133	0.22	2.095	-	-	-0.04	-	-	-	-	-
Null	2	139.757	37.175	0	1.817	-	-	-	-	-	-	-	-
": p-value < 0.01													

Table S9. Linear model comparison for functional diversity using bins. M= Modules, AICc= Akaike information criterion, R2= R square value, K= number of estimated parameters in the model.

M names	K	AICc	Delta_AICc	R2	Intercept	M 3	M 5	Age	M 3*Age	M 5*Age	M 1 - M 3	M 1 - M 5	M 3 - M 5
M * Age	7	50.952	0	0.59	-0.388	5.397	2.963	0.106	-0.367	-0.181	-	2.014"	-1.291
M + Age	5	61.054	10.101	0.38	2.993	0.016	-0.323	-0.097	-	-	-0.016	0.323	0.338
M	4	65.475	14.522	0.25	1.38	0.453	0.855	-	-	-	-0.453	-	-0.402
Age	3	56.562	5.61	0.4	2.603	-	-	-0.073	-	-	-	-	-
Null	2	72.07	21.118	0	1.929	-	-	-	-	-	-	-	-
": p-value < 0.01													

Table S10. Linear model comparison for evenness using the 75P sampling. M= Modules, AICc= Akaike information criterion, R2= R square value, K= number of estimated parameters in the model.

M names	K	AICc	Delta_AICc	R2	Intercept	M 4	M 6	Age	M 4*Age	M 6*Age	M 3 - M 4	M 3 - M 6	M 4 - M 6
M * Age	7	-261.575	0	0.15	1.489	-0.542	-0.509	-0.033	0.035	0.031	0.301	0.3	-0.001
M + Age	5	-256.55	5.025	0	1	-0.005	-0.022	-0.002	-	-	0.005	0.022	0.018
M	4	-255.997	5.578	-0.04	0.976	0	-0.003	-	-	-	0	0.003	0.003
Age	3	-258.096	3.479	-0.02	0.975	-	-	0	-	-	-	-	-
Null	2	-260.28	1.295	0	0.975	-	-	-	-	-	-	-	-

" : p-value < 0.01

Table S11. Linear model comparison for evenness using the MPO sampling. M= Modules, AICc= Akaike information criterion, R2= R square value, K= number of estimated parameters in the model.

M names	K	AICc	Delta_AICc	R2	Intercept	M 4	M 8	Age	M 4*Age	M 8*Age	M 3 - M 4	M 3 - M 8	M 4 - M 8
M * Age	7	-484.862	6.334	0	0.942	-0.005	0.037	0.002	0.001	-0.003	-0.002	-0.015	-0.013
M + Age	5	-484.893	6.303	-0.03	0.981	-0.001	-0.004	0	-	-	0.001	0.004	0.003
M	4	-487.057	4.139	-0.02	0.978	0	-0.002	-	-	-	0	0.002	0.002
Age	3	-489.119	2.077	-0.01	0.976	-	-	0	-	-	-	-	-
Null	2	-491.196	0	0	0.977	-	-	-	-	-	-	-	-

" : p-value < 0.01

Table S12. Linear model comparison for evenness using the bins. M= Modules, AICc= Akaike information criterion, R2= R square value, K= number of estimated parameters in the model.

M names	K	AICc	Delta_AICc	R2	Intercept	M 3	M 5	Age	M 3*Age	M 5*Age	M 1 - M 3	M 1 - M 5	M 3 - M 5
M * Age	7	-108.677	6.414	0.07	0.862	0.466	0.071	0.005	-0.043	-0.007	-0.079	-0.005	0.074
M + Age	5	-110.671	4.42	0.01	0.965	-0.038	-0.038	-0.001	-	-	0.038	0.038	0
M	4	-113.331	1.759	0.04	0.946	-0.031	-0.024	-	-	-	0.031	0.024	-0.007
Age	3	-114.03	1.061	0.01	0.916	-	-	0.001	-	-	-	-	-
Null	2	-115.091	0	0	0.928	-	-	-	-	-	-	-	-

" : p-value < 0.01

Data S1. The dataset includes information of site ages, functional trait for fossil and extant species, and the results from the CDA network analyses for both the functional and taxonomic entities for different sampling schemes: sites with minimum presence of orders (MPO), the sites in the 75th percentile of species richness (75P), and sites aggregated in time bins.