Patterns of Competitive Exclusion
in the Mammalian Fossil Record

Patterns of Competitive Exclusion [running headline]

AUTHORS:

Esther Galbrun (corresponding author)
esther.galbrun@uef.fi
School of Computing, University of Eastern Finland
Technopolis, Microkatu 1, FI-70210 Kuopio, Finland.

Jo Skeie Hermansen
Centre for Ecological and Evolutionary Synthesis, University of Oslo
P.O. Box 1066, University of Oslo, NO-0316 Oslo, Norway.

Indrė Žliobaitė
Department of Computer Science, University of Helsinki,
P.O. Box 64, FI-00014 University of Helsinki, Finland.

Abstract

Due to recent common ancestry, species belonging to the same genus are expected to be more similar with respect to their phenotype, and hence exhibit less niche divergence than species belonging to different genera. Consequently, congeneric species are expected to compete intensely for resources, and therefore to be segregated in space. Yet, despite the longstanding history of this hypothesis of congeneric competitive exclusion, empirical evidence in support of it is at best limited. Here, we analyze co-occurrence patterns of species that belong to the same genera in the mammalian fossil record kept in the NOW database, considering separately Europe during the Neogene, and North America during the Oligocene–Neogene. We assess co-occurrence patterns in comparison to baselines where competitive exclusion is obfuscated through randomization. We find that congeneric species occur together notably less than would be expected at random, with large herbivores being more segregated than large carnivorans and small mammals.

1 Introduction

Understanding the processes that govern how species assemble into communities is a key challenge for ecology (Diamond 1975; Weiher et al. 2011; HilleRisLambers et al. 2012). Such processes include positive and negative species interactions, environmental filtering due to shared or distinct habitat tolerances, and limited or concordant dispersal (Lyons et al. 2016). Among the best studied of these processes are negative species interactions in the form of interspecific competition for limited resources, famously summarized in the principle of competitive exclusion and paraphrased in Hardin's maxim that "complete competitors cannot coexist" (Hardin 1960). Closely related to the principle of competitive exclusion is the competition-relatedness hypothesis of community assembly (Darwin 1859), stating that more closely related species should compete more intensely due to greater niche overlap, and consequently co-occur less in space (Cahill et al. 2008). Congeneric species descend from a recent common ancestor and are therefore expected to be more similar with respect to their phenotype on average, and hence exhibit less niche divergence than species belonging to different genera.

Empirical support for competition-related segregation in extant communities is not univocal. Meta-analyses have found an excess of segregated species pairs relative to null expectations (Gotelli and McCabe 2002; Ulrich and Gotelli 2010), but interpretations are not settled yet. There is little evidence in support of this excess resulting from intensified competition due to a recent common ancestry (Sfenthourakis et al. 2006).

Even fewer empirical analyses exist for species segregation in the past. Klompmaker and Finnegan (2018) analyzed a range of Phanerozoic marine benthic communities and found no support for increased segregation in more closely related species pairs. Tóth et al. (2019) analyzed the persistence of co-occurrence patterns over time in the Pleistocene and Holocene and identified structural changes attributable to the end-Pleistocene megafaunal extinction. Similarly, Lyons et al. (2016) analyzed the co-occurrence structure in the Holocene relative to the present-day community structure. Neither of these studies of the fossil record explicitly looked at congeneric species and, to the best of our knowledge, ours is the first attempt to investigate this for fossil mammals.

We investigate the competition-relatedness hypothesis in the fossil record by assessing patterns of co-occurrence among congeneric species relative to those of species belonging to different genera in Cenozoic Era mammal communities of the Northern Hemisphere. To obtain a baseline for comparison, we randomize the affiliations of species to genera within time bins. This way, whether species that occur at the same localities are congeneric or not is determined by chance, simulating the situation where competitive exclusion does not act. We do this separately for three faunal groups: large herbivores, small mammals and large carnivorans, over two continents: Europe during the Neogene and North America during the Oligocene–Neogene. We analyze obtained patterns in the context of the mean ordinated hypsodonty, which relates to environmental aridity and net primary productivity, and of the number of species present, which relates to overall observed diversity.
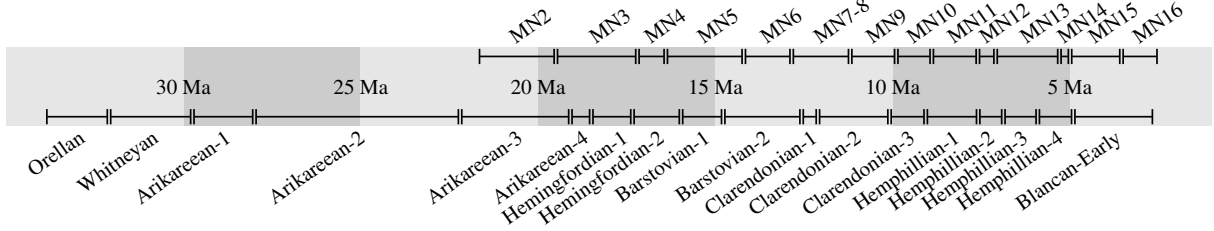
Figure 1: Timeline of the time units used in our study for Europe (top) and North America (bottom).

2 Data and computational methods

2.1 Data

We analyze records from the NOW database of fossil mammals (The NOW Community 2021). We use a dump of the public database downloaded on 25 November 2020, containing a total of 67 547 records. Each record represents the occurrence of an extinct species at a locality and contains taxonomic and ecological information about the species as well as details about the locality where it was found, including its geographic coordinates, estimated age boundaries and palaeo-environmental context.

Our study covers two continents: Europe (in a broad sense) during the Miocene, and North America during the Oligocene and the Neogene. Accordingly, we first filter records by geography. The subset covering Europe contains records with latitudes ranging from N 14° to N 82° and longitudes ranging from W 24° to E 75°, further restricted to a subset of present-day countries.[1] The subset covering North America contains records with latitudes ranging from N 19° to N 84° and longitudes ranging from W 140° to W 70°.

We further filter and split the records based on time. We select time spanning from 21.7 Ma to 2.5 Ma for Europe and from 33.9 Ma to 2.63 Ma for North America, time spans that are relatively well represented in the NOW database. As time bins for our analysis, we use the Mammalian Neogene zonation (MN, *europeanmammalzones*) in Europe and American Land Mammal Ages (NALMA) in North America, with their corresponding age boundaries as currently reported. Our resulting time bins are depicted in Fig. 1 while the precise age boundaries are listed in the Appendix. Any locality for which more precise time information than a mammalian time unit is available is assigned to the containing time bin. We leave out from our analysis localities that span across several time bins, unless the overlap with adjacent time bins is very small (less than 10% of the locality's time span).

Within each continent, we separately analyze three faunal groups consisting of the following orders:

---

[1] Armenia, Austria, Azerbaijan, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, France, Georgia, Germany, Greece, Hungary, Iran, Iraq, Italy, Kazakhstan, Moldova, North Macedonia, Poland, Portugal, Romania, Russia, Serbia, Serbia and Montenegro, Slovakia, Spain, Switzerland, Syria, Turkey and Ukraine.

large herbivores: *Artiodactyla*, *Hyracoidea* (absent from North America), *Perissodactyla*, *Primates* and *Proboscidea*.

large carnivorans: *Carnivora*, *Creodonta* and *Tubulidentata* (absent from North America).

small mammals: *Cimolesta* (absent from Europe), *Didelphimorphia*, *Eulipotyphla*, *Lagomorpha*, *Leptictida* (absent from Europe) and *Rodentia*.

Although not all the species within those orders belong to the same dietary group, the primary goal of using this grouping is to justify the assumption of competition between species within each group.

The database contains occurrences that have incomplete or uncertain taxonomic information. For the purpose of our analysis, we make an operational distinction between (yet) unnamed species (annotated in NOW by "gen. sp.") and indeterminate species for which affiliation to the species level is uncertain (annotated by "indet."). For the purpose of this analysis, we ignore all the other uncertainty qualifiers (e.g. containing qualifiers such as "cf") and treat those occurrences as identified. We do not discard incomplete or uncertain occurrences from the master dataset used for analyses, but we consider incompleteness and uncertainty separately when computing co-occurrence statistics. Summary statistics of the master dataset ready for analysis for each continent, time bin and faunal group as well as plots of extant genera and species across time bins can be found in the Appendix.

The counts of species and of localities in each time bin for the two continents and three faunal groups studied are shown in Fig. 2.

For each continent and time bin the sets of localities vary across faunal groups since some localities may contain only large herbivores, others only large carnivorans, or combinations of the three groups.

In histograms in the bottom of each panel of Fig. 2, localities are colored according to mean ordinated hypsodonty. The mean ordinated hypsodonty (Fortelius et al. 2002) at a locality is computed only over large mammals occurring at that locality. Localities for which the value is not available, e.g. because no large mammal occurs there, are drawn in gray. The macro-average mean ordinated hypsodonty for each time bin is indicated under the corresponding bar.

In histograms in the bottom of each panel of Fig. 2, species are colored and grouped based on the order to which they belong. Indeterminate and unnamed species are excluded. Species occurring at just one locality in a time bin are counted separately from species occurring in several localities and are depicted with lighter shades. We can see that the overall distribution of single-occurrence species is relatively stable over time and across the three faunal groups. Here, they typically make around one-third or less of the species pool within each faunal group in Europe and half or more of the species pool within each faunal group in North America.

Looking at the height of the bars across time in Fig. 2, we note that the number of extant species in a time bin partly correlates with the number of localities belonging to that time bin. Both the total number of species and the total number of localities

represented vary over time, rather consistently across the three faunal groups, which may suggest that diversity variations are indeed real, rather than spurious artifacts of potential undersampling. To ensure the interpretability of analysis, we do not make any adjustments for potentially unequal sampling when estimating the degree of segregation of congeneric species, but treat occurrence data as is and account for potential sampling effects via randomization experiments, instead. For the same reason, we deliberately choose not to manually amend the dataset extracted from the NOW database. Rather, we aim for our analysis tools and tests to be as robust as possible to potential noise in the data, which is always present one way or the other.

2.2 Methods

Our goal is to study patterns of co-occurrence among species, in general, and the extent to which congeneric species co-occur as compared to what would be expected at random, in particular. For each continent, faunal group and time bin, the collection of fossil occurrences can be seen as a localities-by-species binary occurrence matrix, with localities as rows and species as columns. If a species has been found at a locality, the entry in the corresponding matrix cell will be one, otherwise it will be zero. Our methods are best understood by considering how such occurrence matrices are analyzed.

2.2.1 Statistics of species co-occurrences

We consider in turn each triple of continent, faunal group and time bin (e.g. North America–large herbivores–Whitneyan), and process the corresponding data subset, i.e. the occurrence matrix, separately. Given such a data subset, for a species i, we let S_i and s_i denote the subset and number of localities where the species occurs, respectively. That is, S_i contains the identifiers of the rows where positive entries are found in column i and s_i is the marginal sum for that column, also the size of S_i. We consider three different co-occurrence statistics, which can be divided into two types.

Segregation of species pairs. The first type of co-occurrence statistics can be computed for any pair of species, regardless of whether they belong to the same genus or not. For a pair of species i and j, let s_{ij} denote the number of localities where both species occur, i.e. $s_{ij} = |S_i \cap S_j|$. The *C-score* (Stone and Roberts 1990) is then defined as

$$C(s_{ij}, s_i, s_j) = (s_i - s_{ij}) \cdot (s_j - s_{ij}),$$

so that a higher C-score indicates that the species in the pair are more segregated. The C-score takes positive values; the maximum attainable value is the product of the number of occurrences of the involved species, i.e. it is not normalized.

The *mid-P variant of Fisher's exact test* (Berry and Armitage 1995; Kallio et al. 2011) is the probability for the overlap between the two species to have a size smaller than observed, plus half of the probability for the overlap to be of the observed size, if the occurrences were drawn independently at random from either species. This is computed using the hypergeometric distribution. Formally, given the total number of localities n,
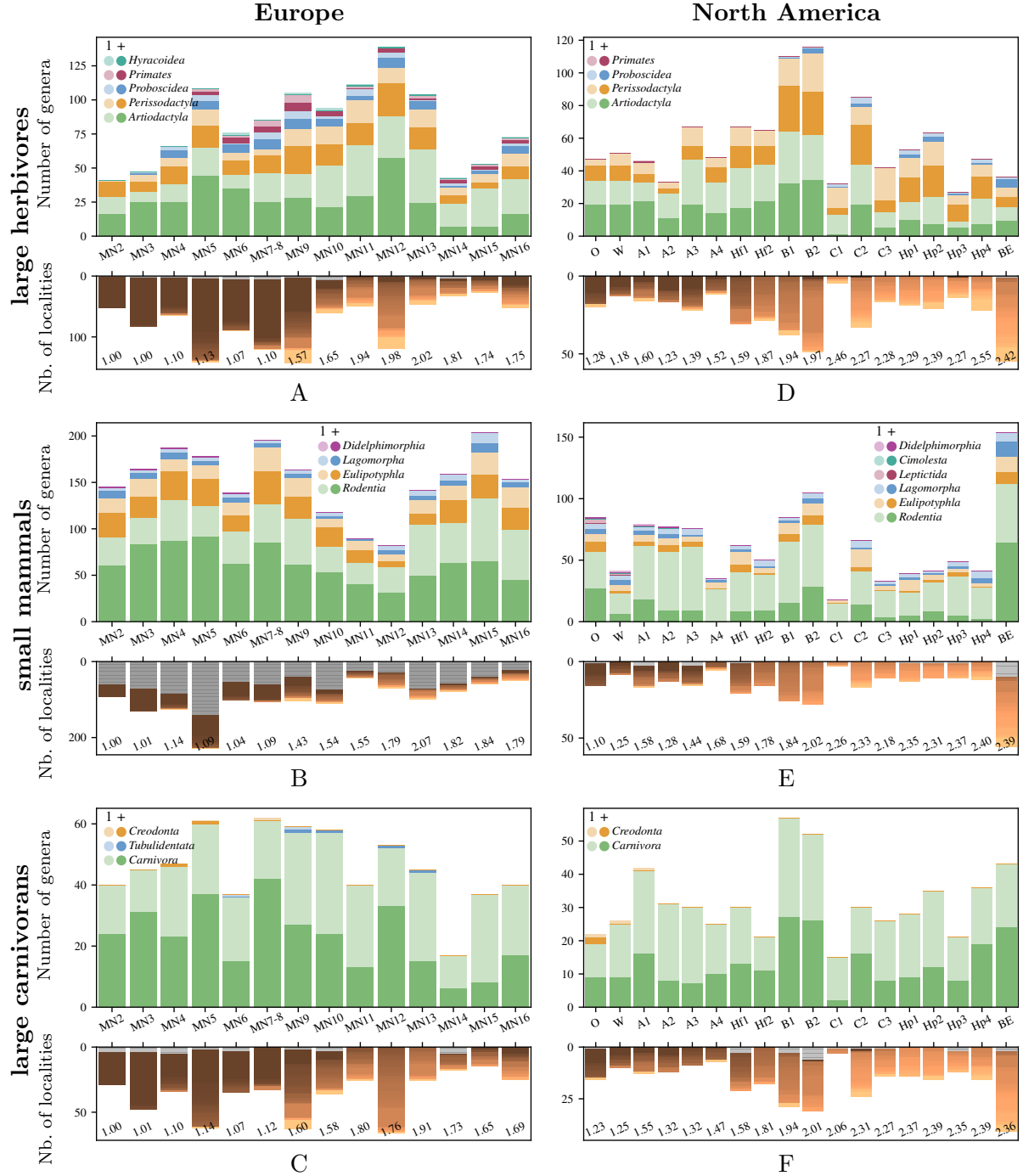
Figure 2: Number of species and of localities in each time bin for the two continents (columns) and three faunal groups (rows). Species are colored and grouped based on the order to which they belong, with single-occurrence (1) and multi-occurrence (+) species depicted with lighter and standard shades, respectively. Indeterminate and unnamed species are excluded. Localities are colored according to the mean ordinated hypsodonty of large mammals, with lighter shades of brown correspond to higher mean ordinated hypsodonty. Localities for which the value is not available are drawn in gray. The macro-average mean ordinated hypsodonty for each time bin is indicated under the corresponding bar.

the number of localities where each of the two considered species occurs, s_i and s_j, and where both occur together, s_{ij}:

$$p_F(s_{ij}, s_i, s_j, n) = \sum_{k=0}^{s_{ij}} P(X = k \mid n, s_i, s_j) - \frac{1}{2} P(X = s_{ij} \mid n, s_i, s_j)$$

where

$$P(X = k \mid n, s_i, s_j) = \frac{\binom{s_i}{k} \binom{n-s_i}{s_j-k}}{\binom{n}{s_j}} \ .$$

The mid-P value p_F can be interpreted as a probability and takes value within the unit interval, with species that co-occur less than expected, i.e. segregated pairs, corresponding to values close to zero.

These pair-based co-occurrence statistics do not take indeterminate and unnamed species into account. Since the complete species information is not available for them, we simply ignore the corresponding columns.

Distribution of species within genera. In order to account for multiple (more than two) species within a genus, we consider a second type of co-occurrence statistics and look, for a given genus, at occurrence localities with multiple species. Specifically, we consider the *fraction of multi-species occurrence localities*, where the numerator is the number of localities where at least two species of the considered genus co-occur, while the denominator is the number of localities where at least one species of the genus occurs. Note that, for a genus with a single extant species within a time bin, which we refer to as *singleton genus*, this ratio is always zero by definition. Clearly, such a genus does not contribute congeneric species.

Furthermore, we also simply consider the *number of genera with co-occurring species*, i.e. the number of genera for which the above fraction is strictly greater than zero.

Note that, according to the NOW database conventions, if the genus of a record at a given locality is determined and named but the species is undetermined or unnamed, the record is assumed to be a different species than any other record of the same genus at that locality. Therefore, such a record increases the count of distinct species occurring at the locality.

2.2.2   Null-models for assessing the strength of co-occurrences

In order to assess the extent to which congeneric species are segregated from each other, we need to compare co-occurrence statistics to baselines where competitive exclusion potentially does not act. We obtain such null models through randomizations over species occurrence. This way we generate variants of the dataset where some properties, such as sample sizes, are maintained but species co-occurrence is randomly distributed. Repeating a randomization multiple times allows to obtain a distribution of the statistics of interest, and then compare the value observed in the original data to that distribution to argue about the strength of co-occurrence (or segregation).

6

We use several types of randomizations, as follows. Two of the most common null models are the fixed–fixed (FF) model (Connor and Simberloff 1979) and the proportional–proportional (PP) model (Gotelli 2000). Both approaches generate randomized variants of an occurrence matrix that have the same number of columns, of rows, and of positive entries. The fixed–fixed (FF) model also maintains the marginal distribution exactly, whereas the proportional–proportional (PP) model imposes proportional constraints on the marginals, i.e. the sums of rows and columns in the randomized versions are only required to match the original data on average. Efficient algorithms have been proposed to generate null models, including, in particular, the four-step proportional–proportional algorithm by Ulrich and Gotelli (2012), later referred to as `UG`, and the `Curveball` fixed–fixed algorithm by Strona et al. (2014). Here, we use the implementations by the authors, as support material in (Strona et al. 2018), with small modifications to `UG` to better handle large proportions of singletons.

In addition, we consider a further simple method to which we refer as *taxonomic shuffling* (`shuffle`). Specifically, the taxonomic labels associated to the columns of the occurrence matrix, that indicate which genus and species a column represents, are reassigned at random. As a result, each genus contains the same number of distinct species, but the species occurrences are different.

Randomization is done within each time bin separately while excluding undetermined and unnamed species.

3   Results

Results for analysis are obtained in three major steps: *(i)* computing co-occurrence statistics from the original data, *(ii)* comparing the values from the original data to null models, *(iii)* examining trends in the environmental context. Such three-step computational experiment is carried out separately for each faunal group over each continent. Furthermore, we consider three different co-occurrence statistics and three different null models. Hence, we performed a total of $2 \times 3 \times 3 \times 3 = 54$ three-steps computational experiments, one for each of the combinations outlined below.

continent		faunal group		co-occurrence statistic		null model

Europe
North America
	$\times$	large herbivores
		small mammals
		large carnivorans
			$\times$	C-score
				mid-P value
				nb. of genera with
				co-occurring species
					$\times$	`UG`
						`Curveball`
						`shuffle`

An exhaustive report of our computational experiments can be found in the Appendix. Here in the main text, we focus on one co-occurrence statistic, namely the number of genera with co-occurring species, and on one null model, namely the fixed–fixed model generated by the `Curveball` randomization algorithm. First, we walk through the three steps of the experiments for the large herbivores, in order to explain our computational experiment workflow. Then, we analyze the results from the final step of the experiments with this co-occurrence statistic and null model (number of genera with co-occurring species, `Curveball`) for all three faunal groups and both continents, shown in Fig. 5.
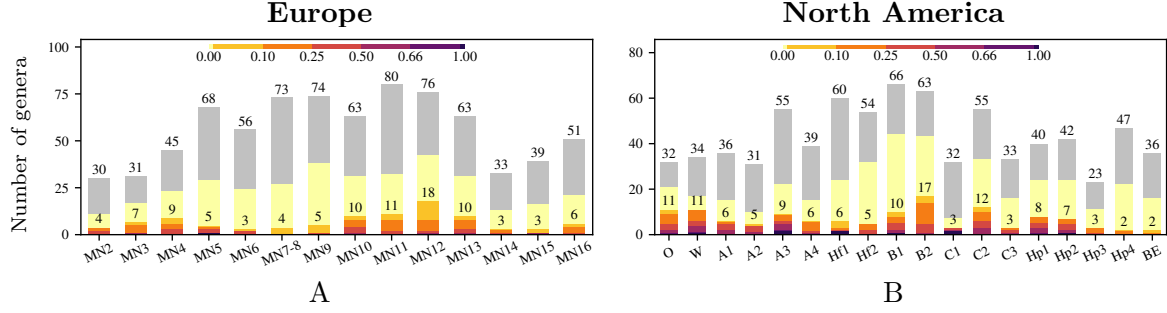
7

Figure 3: Counts of genera by fraction of multi-species occurrence localities for each time bin, large herbivores. The total number of genera in a time bin is indicated above each bar and the number of these genera having co-occurring species is indicated above the portion of the bar that represents them. Singleton genera are drawn in gray.
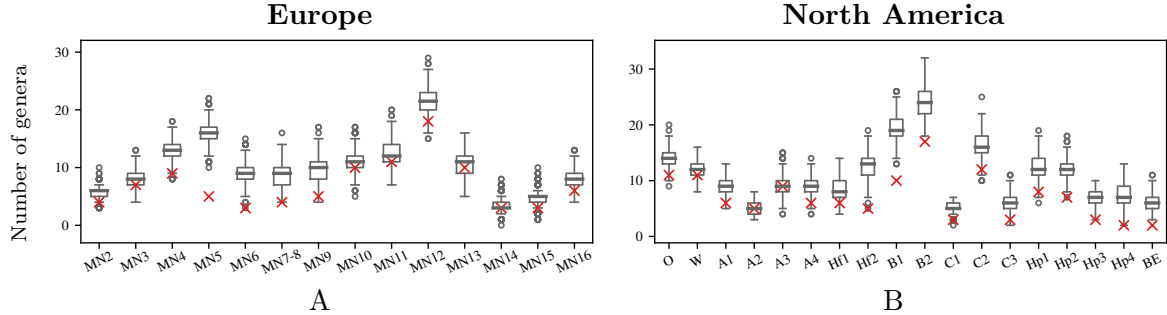

Figure 4: Number of genera with co-occurring species for each time bin, large herbivores. Comparing the value computed for the original data (red crosses) to its distribution across one thousand `Curveball` null models (gray boxplots).

3.1 Workflow for computational experiments

First, we use large herbivores in both Europe and North America as example and go step by step through the workflow of these two experiments.

Computing co-occurrence statistics from the original data. The distribution of genera according to their fraction of multi-species occurrence localities is shown in Fig. 3. That is, genera are grouped depending on the proportion of localities where multiple species of the genus co-occur, among all localities where any species of the genus occurs. A *singleton genus*, here defined as a genus with a single extant species within a time bin, does not contribute congeneric species. By definition, the fraction of multi-species occurrence localities of a singleton genus is zero, and the number of genera with co-occurring species in a time bin is at most the number of non-singleton genera. Singleton genera are drawn in gray in Fig. 3.

While there are overall more genera in Europe than in North America, a larger fraction contains multiple occurring species and there is a greater overlap of congeneric species in the latter.
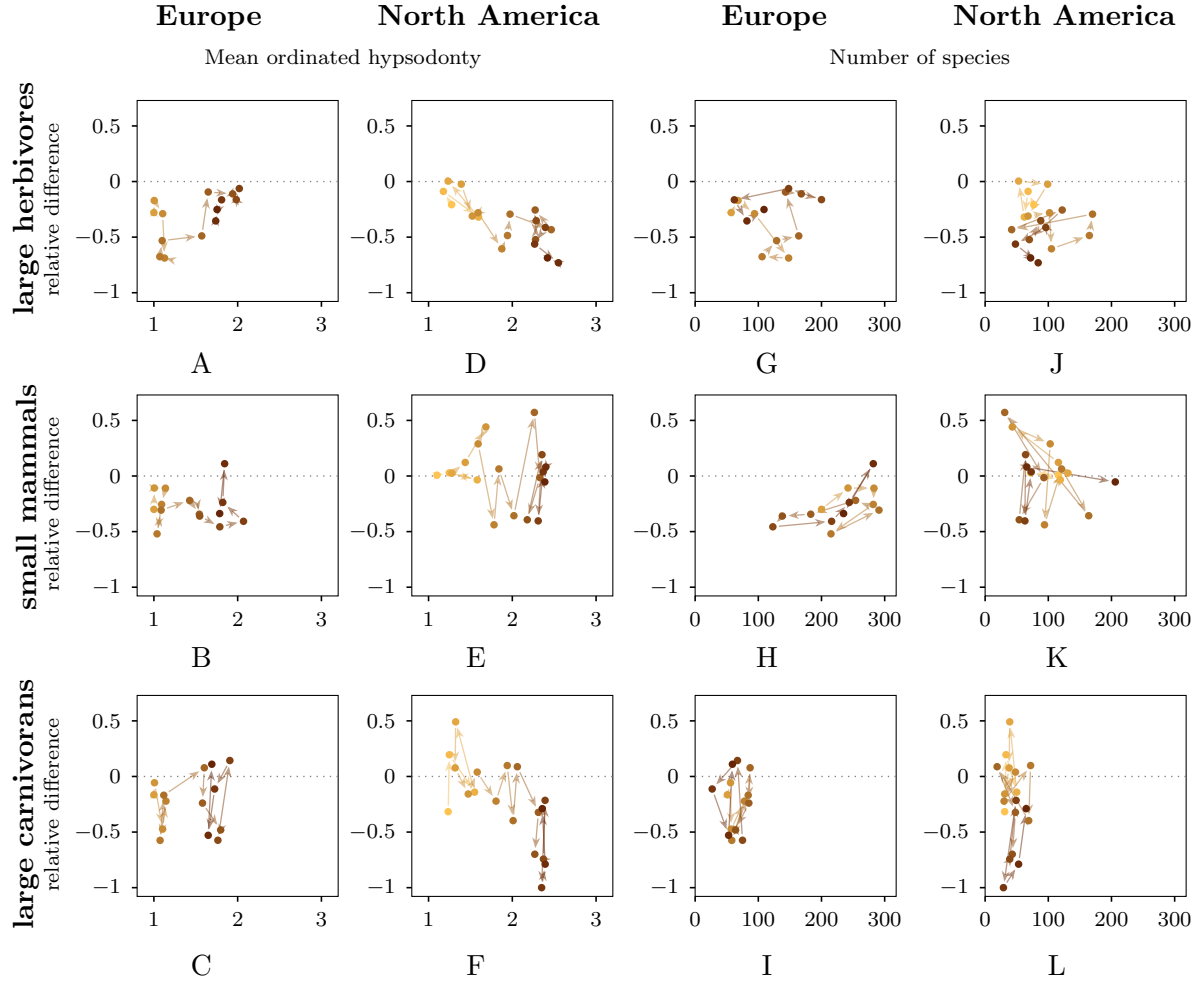
Figure 5: Relative difference between observed and expected number of genera with co-occurring species versus environmental context variables, respectively mean ordinated hypsodonty (panels A–F) and number of species (panels G–L), over time. Each dot represents a time bin, with more recent times depicted in darker shades of brown and arrows going from the oldest to most recent time bins. The vertical axis represents the relative difference between the number of genera with co-occurring species computed from the original data and the `Curveball` null model expected value. The dashed horizontal line in each panel corresponds to equality between the observed and expected values, indicating no segregation pattern. Larger negative values indicate stronger segregation of congeneric species. Time bins span from 21.7 to 2.5 MA for Europe and from 33.9 to 2.6 Ma for North America.

However, it is difficult to tell whether the observed values are specific to a considered time bin and related to some underlying process, or a straight-forward consequence of more general properties of the data. In addition, because different time bins contain different numbers of localities and of species, it is difficult to meaningfully compare values across time bins and across continents.

Comparing the values from the original data to null models. In order to alleviate these issues, we resort to a comparison to null models. Values observed across several randomized variants of the data provide a background distribution against which we can contrast the value observed in the original data. However, we cannot apply this process while considering the collection of values across the different genera for each dataset. Instead, the values must be aggregated to obtain a single numerical summary for each dataset.

In Fig. 4, the number of genera with co-occurring species in the original data (red crosses) against the baseline provided by the distribution of this value measured across one thousand `Curveball` null models (gray boxplots). Intuitively, red crosses correspond to the lower count reported for each bar in Fig. 3, whereas the gray boxplots show the distribution of this value computed for randomized copies of the data.

Examining trends in the environmental context. Finally, we examine segregation trends in the environmental context, as captured by mean ordinated hypsodonty and number of species, in particular. To do so, we look at the direction and magnitude of the deviation between the value observed in the original data and the null models. Formally, denoting the value from the original data as x and the mean of the null model distribution as μ, we compute the *relative difference* as $\delta = (x - \mu)/\mu$. Intuitively, this difference captures how far the red crosses in Fig. 4 are from the associated boxplots. A red cross far below the boxplot, i.e. a case where the value observed in the original data is much lower than the expected value in the null model, corresponds to a large negative difference.

We then plot the relative differences computed in this way against the two contextual variables, in Fig. 5. The two columns on the left-hand side show the relative difference against the mean ordinated hypsodonty, for Europe and North America respectively, whereas the two columns on the right-hand side show the relative difference against the number of species, for Europe and North America respectively. The three rows correspond to the three faunal groups of large herbivores, small mammals and large carnivorans, respectively, from top to bottom. In a given panel, each dot represents a time bin, with more recent times depicted in darker shades of brown and arrows going from the oldest to most recent time bins.

3.2 Analysis and interpretations

Now we turn to the analysis of the results from the experiments with the number of genera with co-occurring species and the `Curveball` null model, for all three faunal groups and both continents.

Are congeneric species more segregated than expected at random? Most dots in all plots of Fig. 5 lie below the dashed line indicating equality of the observed value to the expected value. This pattern, showing that the number of genera with co-occurring species is smaller in the original data than in the null models in most cases, suggests that congeneric species are generally more segregated than expected at random. Indeed, points lower along the vertical axis correspond to time bins with more segregated congeneric species.

The distribution of points along the vertical axis suggests that among the three faunal groups, congeneric large herbivores species are the most segregated, followed by large herbivores in Europe. However, small mammals appear to be much less segregated in North America. This pattern may or may not relate to the overall exceptional distribution of records of small mammals in North America. As observed in Fig. 2, the North American record of small mammals contains many more single-occurrence species than any other group in this analysis. Congeneric species of large carnivorans appear to be the least segregated among the three faunal groups in both continents.

How do patterns of segregation relate to changing environments? Panels A–F in Fig. 5 show distinct patterns for Europe and North America in relation to mean ordinated hypsodonty. Mean ordinated hypsodonty is a proxy for the harshness of the environment, with higher values related to harsher environments (Fortelius et al. 2002), i.e. increased aridity and decreased net primary productivity. Each dot represents a time bin, with darker shades of brown representing more recent times. We can observe a general trend of darkening shades when moving left-to-right along the horizontal axis, suggesting that the environment is getting harsher over time in both continents, with the absolute harshness eventually reaching higher levels in North America than in Europe. This is a well-known pattern (Fortelius et al. 2014).

Our analysis brings forward a mild but interesting nuance: overall segregation of congeneric species diminishing over time in Europe, but increasing over time in North America, with the exception of small mammals in North America, for which no clear trend over time emerges. The pattern for large herbivores is especially intriguing —at the milder end of environmental conditions, the level of segregation is similar in Europe and North America, but when environmental conditions become harsher, the segregation of congeneric species appears to go in opposite directions in the two continents, decreasing in Europe and increasing in North America.

Panels G–L in Fig. 5 show the relative differences between observed and expected number of genera with co-occurring species against the total number of species diversity over time. No notable pattern emerges, suggesting that segregation of congeneric species operates regardless of whether species diversity is low or high.

4   Discussion

Assessing the segregation of extinct species is a challenging task. Ours was, to the best of our knowledge, the first attempt to investigate competitive exclusion from the mammalian

fossil record through computational experiments. How to design or select suitable randomization tests and how to understand their implications have been recurrent questions during several months of active work on this manuscript.

Among many challenges, a major one was obtaining summary statistics over a region or a faunal group, while different types of competitive relationships might potentially have been in operation within them. And, if some species might have been segregated whereas others might have co-occurred, due to their cooperative interactions or, more likely, to their affinities to similar environmental conditions, those opposing relationships would cancel out when computing aggregated co-occurrence statistics, leading to inconclusive tests. The results that we obtained when looking at the distribution of species within genera, and the number of genera with co-occurring species in particular, suggest that segregation was at play, especially among large mammals. Experiments with pairs of species, on the other hand turned out not to reveal strong consistent trends and did not yield conclusive evidence either way. The number of pairs grows quadratically with the number of species. Congeneric species pairs represent a small and diminishing fraction of this number. Using null models as comparison baselines can help tackling this imbalance problem but did not allow us to overcome it.

The fact that fossil localities are not at all equal to communities of living specimens constitutes another important challenge. Due to time averaging, fossil localities may contain contemporaneous specimens, but they may also contain specimens that lived during successive time periods. If that is the case, congeneric species that lived in succession will appear as coexisting, although competitive exclusion might have been the very reason for the replacement of one species by the other. Intuitively, this should weaken the test outcomes. Thus, if even under such circumstances we obtain a signal suggesting segregation taking place, then we might infer that either successor species are rarely congeneric, or that the underlying segregation mechanism is sufficiently strong to not get diluted through such time averaging.

Yet another important challenge is taxonomic. The decision to assign a species to a genus is made by taxonomic experts and, for the fossil record, is primarily based on morphological differences between species. These differences leave more room for interpretation in some groups than others. Thus, affiliation to the same genus is only an approximation of the distance to the last common ancestor. It would be interesting to perform a similar analysis taking into account phylogenetic information, when reliable supertrees become available for the mammalian record at large. Randomization over trees would certainly require more intricate methods than when approximating close common ancestry through affiliation to the same genus, as we did in this analysis.

5   Conclusion

We tested whether mammalian species that belong to the same genus are less likely to occur together than would be expected at random. We found a rather weak but consistent signal over time and across continents that indeed, congeneric species tend to be segregated in space. The strongest segregation seems to be at play among large herbivores, followed by small mammals. Congeneric species of large carnivorans appear

to be the least segregated among the three faunal groups in both continents.

Our analysis brings forward a mild but interesting nuance: overall segregation of congeneric species diminishing over time in Europe, but increasing over time in North America, with the exception of small mammals in North America, for which no clear trend over time emerges. The pattern for large herbivores is especially intriguing —at the milder end of environmental conditions, the level of segregation is similar in Europe and North America, but when environmental conditions become harsher, the segregation of congeneric species appears to go in opposite directions in the two continents, decreasing in Europe and increasing in North America. Interpretation of the latter pattern remains open.

References

Berry, G. & Armitage, P. (1995). Mid-p confidence intervals: A brief review. *Journal of the Royal Statistical Society: Series D (The Statistician)*. 44(4):417–423

Cahill, J. F., Kembel, S. W., Lamb, E. G., & Keddy, P. A. (2008). Does phylogenetic relatedness influence the strength of competition among vascular plants? *Perspectives in Plant Ecology, Evolution and Systematics*. 10(1):41–50

Connor, E. F. & Simberloff, D. (1979). The assembly of species communities: Chance or competition? *Ecology*. 60(6):1132–1140

Darwin, C. R. (1859). *On the Origin of Species by Means of natural Selection, or The Preservation of favoured Races in the Struggle for Life*. John Murray, London

Diamond, J. M. (1975). Assembly of species communities. In: Cody, M. L. & Diamond, J. M. (eds) *Ecology and evolution of communities*, Harvard University Press, Cambridge, pp 342–444

Fortelius, M., Eronen, J., Jernvall, J., Liu, L., Pushkina, D., Rinne, J. et al. (2002). Fossil mammals resolve regional patterns of Eurasian climate change over 20 million years. *Evolutionary Ecology Research*. 4(7):1005–1016

Fortelius, M., Eronen, J. T., Kaya, F., Tang, H., Raia, P., & Puolamäki, K. (2014). Evolution of neogene mammals in eurasia: Environmental forcing and biotic interactions. *Annual Review of Earth and Planetary Sciences*. 42:579–604

Gotelli, N. J. (2000). Null model analysis of species co-occurrence patterns. *Ecology*. 81(9):2606–2621

Gotelli, N. J. & McCabe, D. J. (2002). Species co-occurrence: A meta-analysis of J. M. Diamond's assembly rules model. *Ecology*. 83(8):2091–2096

Hardin, G. (1960). The competitive exclusion principle. *Science*. 131(3409):1292–1297

HilleRisLambers, J., Adler, P., Harpole, W., Levine, J., & Mayfield, M. (2012). Rethinking community assembly through the lens of coexistence theory. *Annual Review of Ecology, Evolution, and Systematics.* 43(1):227–248

Kallio, A., Puolamäki, K., Fortelius, M., & Mannila, H. (2011). Correlations and co-occurrences of taxa: the role of temporal, geographic, and taxonomic restrictions. *Palaeontologia Electronica.* 14(1)

Klompmaker, A. A. & Finnegan, S. (2018). Extreme rarity of competitive exclusion in modern and fossil marine benthic ecosystems. *Geology.* 46(8):723–726

Lyons, K. S., Amatangelo, K. L., Behrensmeyer, A. K., Bercovici, A., Blois, J. L., Davis, M. et al. (2016). Holocene shifts in the assembly of plant and animal communities implicate human impacts. *Nature.* 529(7584):80–83

Sfenthourakis, S., Tzanatos, E., & Giokas, S. (2006). Species co-occurrence: the case of congeneric species and a causal approach to patterns of species association. *Global Ecology and Biogeography.* 15(1):39–49

Stone, L. & Roberts, A. (1990). The checkerboard score and species distributions. *Oecologia.* 85(1):74–79

Strona, G., Nappo, D., Boccacci, F., Fattorini, S., & San-Miguel-Ayanz, J. (2014). A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals. *Nature Communications.* 5(1):4114

Strona, G., Ulrich, W., & Gotelli, N. J. (2018). Bi-dimensional null model analysis of presence-absence binary matrices. *Ecology.* 99(1):103–115

The NOW Community (2021). New and old worlds database of fossil mammals (NOW). `https://nowdatabase.org/`. Accessed 25 November 2020.

Tóth, A. B., Lyons, S. K., Barr, W. A., Behrensmeyer, A. K., Blois, J. L., Bobe, R. et al. (2019). Reorganization of surviving mammal communities after the end-pleistocene megafaunal extinction. *Science.* 365(6459):1305–1308

Ulrich, W. & Gotelli, N. J. (2010). Null model analysis of species associations using abundance data. *Ecology.* 91(11):3384–3397

Ulrich, W. & Gotelli, N. J. (2012). A null model algorithm for presence–absence matrices based on proportional resampling. *Ecological Modelling.* 244:20–27

Weiher, E., Freund, D., Bunton, T., Stefanski, A., Lee, T., & Bentivenga, S. (2011). Advances, challenges and a developing synthesis of ecological community assembly theory. *Philosophical Transactions of the Royal Society B: Biological Sciences.* 366(1576):2403–2413