



A coherent biogeographical framework for Old World Neogene and Pleistocene mammals

Corentin Gibert, Axelle Zacaï, Frédéric Fluteau, Gilles Ramstein, Olivier Chavasseau, Ghislain Thiery, Antoine Souron, William Banks, Franck Guy, Doris Barboni, et al.

► To cite this version:

Corentin Gibert, Axelle Zacaï, Frédéric Fluteau, Gilles Ramstein, Olivier Chavasseau, et al.. A coherent biogeographical framework for Old World Neogene and Pleistocene mammals. *Palaeontology*, 2022, 65 (2), 10.1111/pala.12594 . hal-03651752

HAL Id: hal-03651752

<https://hal.science/hal-03651752>

Submitted on 26 Apr 2022



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

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

Copyright



A coherent biogeographical framework for Old World Neogene and Pleistocene mammals

by CORENTIN GIBERT^{1,2} , AXELLE ZACAÏ¹, FRÉDÉRIC FLUTEAU³,
GILLES RAMSTEIN⁴, OLIVIER CHAVASSEAU¹, GHISLAIN THIERY¹,
ANTOINE SOURON², WILLIAM BANKS², FRANCK GUY¹, DORIS BARBONI⁵,
PIERRE SEPULCHRE⁴, CÉCILE BLONDEL¹, GILDAS MERCERON¹ and
OLGA OTERO¹ 

¹Laboratoire Paléontologie Evolution Paléoécosystèmes Paléoprimatologie (PALEVOPRIM), UMR-CNRS 7262, Université de Poitiers, 6 rue M. Brunet, 86073 Poitiers, France; corentingibert@gmail.com

²Laboratoire de la Préhistoire à l'Actuel: Culture, Environnement et Anthropologie (PACEA, UMR 5199 CNRS, INEE), University of Bordeaux, Bordeaux, France

³Université Paris Cité, Institut de physique du globe de Paris, CNRS, Université de Paris, Paris, France

⁴Laboratoire des Sciences du Climat et de l'Environnement (LSCE), Institut Pierre Simon Laplace (IPSL), CEA-CNRS-UVSQ, Université Paris-Saclay, 91191 Gif-sur-Yvette, France

⁵Centre Européen de Recherche et d'Enseignement des Géosciences de l'Environnement (CEREGE), Université Aix-Marseille, Europôle Méditerranéen de l'Arbois BP80, Aix-en-Provence cedex 4, 13545 France

Typescript received 4 May 2021; accepted in revised form 19 December 2021

Abstract: In order to understand mammalian evolution and compute a wide range of biodiversity indices, we commonly use the 'bioregion', a spatial division adapted to ecological and evolutionary constraints. While commonly conducted by neontologists, the establishment of bioregions in palaeontology is generally a secondary analysis, shaped on subjective time scales and areas specific to the investigated questions and groups. This heterogeneity, coupled with the scale-dependency of biodiversity indices, prevents the clear identification of macroecological and macroevolutionary trends for large taxonomic groups like extinct mammals. Here we tackle this issue by providing a coherent framework for Neogene and Pleistocene mammals of the Old World following two steps: (1) a temporal scale adapted to mammalian evolutionary history (i.e. evolutionary fauna) is defined by poly-cohort analysis; (2) bioregions are then computed for each evolutionary fauna by clustering, ordination and intermediate approaches at

multiple spatial scales (i.e. continental to regional) for Eurasia and Africa. Additionally, providing a coherent framework for a wide range of mammalian datasets, our results show: (1) the synchronous emergence and fall of five mammalian evolutionary faunas identified at chronological scales varying from the epoch to the geological stage; (2) a transition from a longitudinal to a latitudinal biogeographical structuring between the Miocene and Pliocene, especially in Europe; (3) the long-term affinity of southern Asian with African faunas, in sharp contrast with the modern Palaearctic bioregion extension; and (4) the establishment of a vast Mediterranean bioregion from fragmented areas in the Late Miocene to its full extent in the Pleistocene.

Key words: bioregion, macroecology, hierarchical clustering on principal components, Neogene mammal, Pleistocene mammal, Old World.

UNDERSTANDING the spatial and temporal distribution of biodiversity is one of the most important issues in ecological and evolutionary sciences. At the end of the nineteenth century, Alfred Russel Wallace (1876) and Adolf Engler (1879) conducted pioneering analyses of the spatial distribution of modern continental biotas, which separated the world's fauna and flora into six biogeographical 'realms', more commonly known as biogeographical regions (or bioregions). These areas are defined by the geographical limits of metacommunities and can be used as spatial units to study biodiversity in its various

dimensions (e.g. taxonomic, phylogenetic or functional), whether specific to an area (i.e. alpha diversity) or characterizing differences between areas (i.e. beta diversity). The study of bioregions has since been extended to the marine realm (e.g. Longhurst 2010), to modern as well as palaeontological data, and conducted at various spatial scales (Bernor 1978; Bonis *et al.* 1992a; Heikinheimo *et al.* 2007; Kostopoulos 2009; He *et al.* 2017). Cenozoic mammals and the definition of their bioregions have been the subject of intense research for more than 50 years (Tobien 1967; Bernor *et al.* 1979, 1996; Bernor 1983,

1984; Tedford 1987; Bonis *et al.* 1992b; Fortelius *et al.* 1996; Geraads 1998, 2010; Casanovas-Vilar *et al.* 2005; Maridet *et al.* 2007; Costeur & Legendre 2008; Bibi 2011; He *et al.* 2017, 2018). However, their study has been conducted at a wide range of spatial (i.e. from regional to continental) and temporal scales (i.e. from unique biozone to geological era), depending on the investigated question and often resulting from arbitrary choices. The resulting multiplicity of mammalian palaeontological bioregion shape, scale and size is mainly related to the nature of the studied datasets, often limited to a short period of time, and/or a single trophic guild or taxon. This heterogeneity prevents straightforward investigations of the spatio-temporal dynamics of mammalian metacommunities (Chave 2013), as well as comparisons of biodiversity values between studies, since these computed values are scale dependent (Brocklehurst & Fröbisch 2018).

Furthermore, bioregions established with palaeontological data may consist of geographically close (but temporally distant) fossil sites occupied/inhabited by a wide range of faunal elements that could have lived in drastically dissimilar environments. This phenomenon increases as a function of the chosen temporal subdivision: the larger the temporal scale, the more temporally distant fossil sites are gathered in a same bioregion. Mammalian bioregions have been defined and used at scales varying from the mammal biozone (Geraads 1998; Costeur *et al.* 2004; Maridet *et al.* 2007) through the geological epoch (Button *et al.* 2017; He *et al.* 2018) to the entire geological period (e.g. Neogene; Tedford 1987; Costeur & Legendre 2008). Short temporal scales are valuable, but they often imply taxonomically and geographically limited studies (Kostopoulos 2009), highly dependent on sampling variations (Costeur *et al.* 2004; Casanovas-Vilar *et al.* 2005; Maridet *et al.* 2007). Yet, grouping faunas at larger scales, such as the geological epoch, is insufficient to track regional abiotic changes and therefore produce exploitable biogeographical regions coherent with the tempo of biotic evolution (Escarguel *et al.* 2011). For example, throughout the Miocene (i.e. geological epoch scale), the Earth experienced major climate changes (Barnosky 2001; Zachos *et al.* 2001; Mosbrugger *et al.* 2005); the first half of the epoch is marked by a warming leading to the Middle Miocene Climatic Optimum (MMCO), followed by a long and continuous global cooling. These global changes had important consequences at the regional scale (Cano *et al.* 2014) by altering the distribution and intensity of precipitation and seasonal temperature variation, leading to major changes within all communities adapted to these regional conditions (Janis 1993).

We address this issue by establishing a coherent spatio-temporal framework that can be used to examine spatial dynamics of Neogene and Pleistocene Old World mammalian

metacommunities. This is accomplished by: (1) determining the temporal scales consistent with the evolutionary history of mammals; and (2) identifying their bioregions at multiple spatial scales, so that these can be used for a wide range of datasets.

In order to establish this coherent framework, one must acknowledge that: (1) a biogeographical region (in the context of a palaeontological study) is defined as a large area encompassing similar faunas and/or floras with a common evolutionary history, characterized by homogeneous environmental and climatic conditions (Kreft & Jetz 2010); and (2) the spatial and temporal dimensions of all ecological and evolutionary processes are fundamentally intertwined (Escarguel *et al.* 2011), a relationship often overlooked in palaeobiogeographical studies.

Consequently, two steps are required to produce a coherent framework of bioregions using the Neogene and Pleistocene fossil record of Old World mammals: first, time-dynamic bioregions (Costeur *et al.* 2004; Reygondeau *et al.* 2013) must be computed on temporal scales large enough to avoid strong variation in sampling effort (He *et al.* 2018), and short enough to track regional environmental changes that are driving ecological and evolutionary dynamics (e.g. dispersal, niche differentiation) that form the basis of bioregion building (Maridet *et al.* 2007). The identification of these appropriate temporal resolutions is performed using poly-cohort analyses (Escarguel & Legendre 2006). This approach can detect the succession of evolutionary faunas (i.e. temporally coherent sets of mammals appearing and disappearing synchronously; Figueirido *et al.* 2012) thus ensuring that geographical clusters of fossil sites are built on large temporal scales consistent with the evolution of mammalian faunas and their ecosystems (Herbert *et al.* 2016; Barbolini *et al.* 2020; Li *et al.* 2020). Secondly, bioregions must be established at multiple spatial scales, so that the operational framework proposed here can be used with as many mammal groups and biodiversity patterns as possible. This is crucial as many ecological patterns emerge only at particular scales, and are perceived as noise or constants, or even go undetected, at other scales (Levin 1992; Chave 2013). Bioregion analysis is conducted using two methods: a classical hierarchical clustering (Kreft & Jetz 2010) and a hierarchical clustering on principal components (HCPC; Husson *et al.* 2010). We use an incremental approach with both methods (Heikinheimo *et al.* 2007) to generate continental as well as finer regional bioregions, and create a variable number of clusters to identify large-scale patterns structuring the geographical distribution of mammals (minimum number of clusters) as well as more detailed distribution of Neogene and Pleistocene mammal diversity (maximum number of clusters). Finally, the nature of these subdivisions and their temporal and spatial dynamics can be observed and

interpreted in the light of the well-known changes in mammalian faunas and environments.

MATERIAL AND METHOD

Dataset

All occurrence data of Neogene and Pleistocene mammalian species from Eurasia and Africa were extracted from the open access New and Old Worlds fossil mammal database (NOW; <https://www.helsinki.fi/science/now>; downloaded 1 March 2020). This fossil record was supplemented by our own unpublished dataset, extending over the last five million years, which includes all occurrences of mammalian species identified since 2006 by palaeontologists of the Omo Group Research Expedition (OGRE) in the lower Omo valley, Ethiopia. Occurrences defined at higher taxonomic levels than the species, as well as uncertain identification using open nomenclature (i.e. '?', 'cf.') were excluded from this study. Finally, to limit the dataset to species with similar dispersal constraints, bats and marine mammals were also excluded (Heikinheimo *et al.* 2007). We chose to use the European land mammal ages (MN; Mein 1976; Agustí *et al.* 2001; Fig. S1) as temporal units to ensure the homogeneity of the dataset. Because these have been defined on European faunas, we assigned all African and Asian occurrences to ad hoc 'MN-equivalents' (MNEQ; Ataabadi *et al.* 2013, Madern & Van den Hoek Ostende 2015) based on their absolute and relative ages (extracted from the database). Only localities defined at the scale of the geological stage (or ~3 MNEQ) were selected to maximize temporal resolution and homogeneity. Palaeocoordinates of fossil localities were calculated by rotating their present-day coordinates using the software GPlates version 2.0.0 (Müller *et al.* 2018), and the rotation file supplied by Scotese (2016). To reduce the effect of spatial variability in sampling as well as duplicates between databases, assemblages from the same MNEQ were grouped together when located within a range of 0.1° of latitude and longitude following He *et al.* (2018). Furthermore, only localities with at least five different species were retained for poly-cohort and cluster analyses (Fig. S2A). Even though Old World Neogene and Pleistocene mammals are documented by one of the most complete terrestrial palaeontological datasets in the world, sampling effort is spatially and temporally highly variable (Peláez-Campomanes & Van der Meulen 2009; Madern & Van den Hoek Ostende 2015). In comparison with Europe and Asia, the African continent is almost free of rich localities (>5 species) before 5 Ma (Fig. S2B), except in eastern and southern Africa. Overall, the final dataset consists of 44 338 occurrences, distributed within 1995 localities, which corresponds

to 5147 species of non-flying terrestrial mammals that have inhabited the Old World during the last 23 million years.

Poly-cohort analysis and evolutionary faunas

To ensure that biogeographical regions are built on a temporal scale consistent with the evolution of mammalian ecosystems and faunas, the selected time scale is that of evolutionary fauna ('EF' herein; Sepkoski 1981; Figueirido *et al.* 2012) or chronofauna (Olson 1952; Eronen 2007). Computing EF is a simple way to build distinct sets of bioregions for each major evolutionary stage of the Old World mammals. Some evolutionary faunas, such as the Pikermian and the Baodean chronofaunas (respectively in Europe and Asia), are already well known and documented (Eronen *et al.* 2009; Kostopoulos 2009; Ataabadi *et al.* 2013). However, mammalian evolutionary faunas remain unknown for several regions of Eurasia and very large temporal intervals of the Neogene and Pleistocene. We thus computed evolutionary faunas for Europe, Asia and Africa for the entire Neogene–Pleistocene interval, using a poly-cohort analysis developed by Escarguel & Legendre (2006), based on the work of Simpson (1944). Distinct methods like clustering-based palaeocommunities (PCOMs) developed by Raia *et al.* (2005, 2006), or the Q-mode factor analysis used by Figueirido *et al.* (2012) are reliable and proven alternatives to produce evolutionary faunas at multiples temporal scale. However, for the sake of clarity, this study focus on poly-cohort and avoids using clustering/factor approach for both temporal and spatial analysis. The poly-cohort approach consists of observing all cohorts (whole sets of species occurring during a specific time span; i.e. one MNEQ) of a dataset to distinguish sets of cohorts originating and disappearing synchronously. This is accomplished by creating two curves: a pre-nascence curve that reflects the cohort origination dynamics through the previous time interval, and a survivorship curve that reflects the extinction of the cohort during the following time interval (Fig. S3). In turn, the joint analysis of all cohorts (here, corresponding to each MNEQ) reveals evolutionary breakpoints at which all pre-existing cohorts are impacted by high extinction rates and give way to a new set of cohorts that appear during or just before that breakpoint. Finally, breakpoints identified by the poly-cohort analysis are used to define the temporal limits of evolutionary faunas and regroup MNEQ units. As poly-cohort analysis is based on the percentage temporal trajectories of all cohorts, it shows low sensitivity to sampling heterogeneity. Nevertheless, strong temporal variations in sampling effort could generate arbitrary breakpoints. Attention should thus be directed toward the potential relationship

between the computed evolutionary breakpoints and the temporal variation in sampling heterogeneity (Peláez-Campomanes & Van der Meulen 2009; Madern & Van den Hoek Ostende 2015).

Establishment of biogeographical regions

Taxonomic dissimilarity was calculated between all pairs of localities using Raup & Crick (1979) and Simpson (1944) indices, as recommended for fossil datasets (Kreft & Jetz 2010; Ataabadi *et al.* 2013) to reduce the effect of sampling heterogeneity on dissimilarity values. Both indices are computed within each evolutionary fauna obtained with the poly-cohort analyses. The determination of biogeographical region was first made with hierarchical clustering (unweighted pair group method with arithmetic mean (UPGMA) using ‘ward.D2’ method) directly on the inter-locality dissimilarity values. Hierarchical clustering was also computed on new coordinates calculated using non-metric multidimensional scaling (nMDS) on the dissimilarity values. This second method is known as hierarchical clustering on principal components (HCPC, FactoMineR package; Lê *et al.* 2008), and represents an intermediate solution between ordination and clustering methods. To validate the clustering of localities into bioregions, their taxonomic dissimilarity was subjected to significance testing using analyses of similarity (ANOSIM) based on Raup & Crick and Simpson indices (Fig. S4A). Only clusters that are significantly dissimilar and geographically coherent (i.e. non-overlapping and contiguous) were retained (Brocklehurst & Fröbisch 2018). For example, when a few localities of a cluster A were found geographically isolated within a cluster B after HCPC or UPGMA clustering, ANOSIM was conducted on the clusters initially determined by UPGMA and HCPC, as well as on the clusters subsequently modified to account for both geographical and taxonomic dimensions (i.e. with the few isolated localities of A integrated within B; this procedure illustrated in Fig. S4B).

Limiting clustering analyses to hierarchical clustering (e.g. UPGMA) prevents the identification of geographical gradients of composition within faunas (Brayard *et al.* 2007). Given that the importance of gradients in the organization of biodiversity has long been demonstrated (McCoy & Connor 1980; Marcot *et al.* 2016), the use of ordination and/or HCPC methods to identify gradients is imperative. The distribution of fossil sites within the

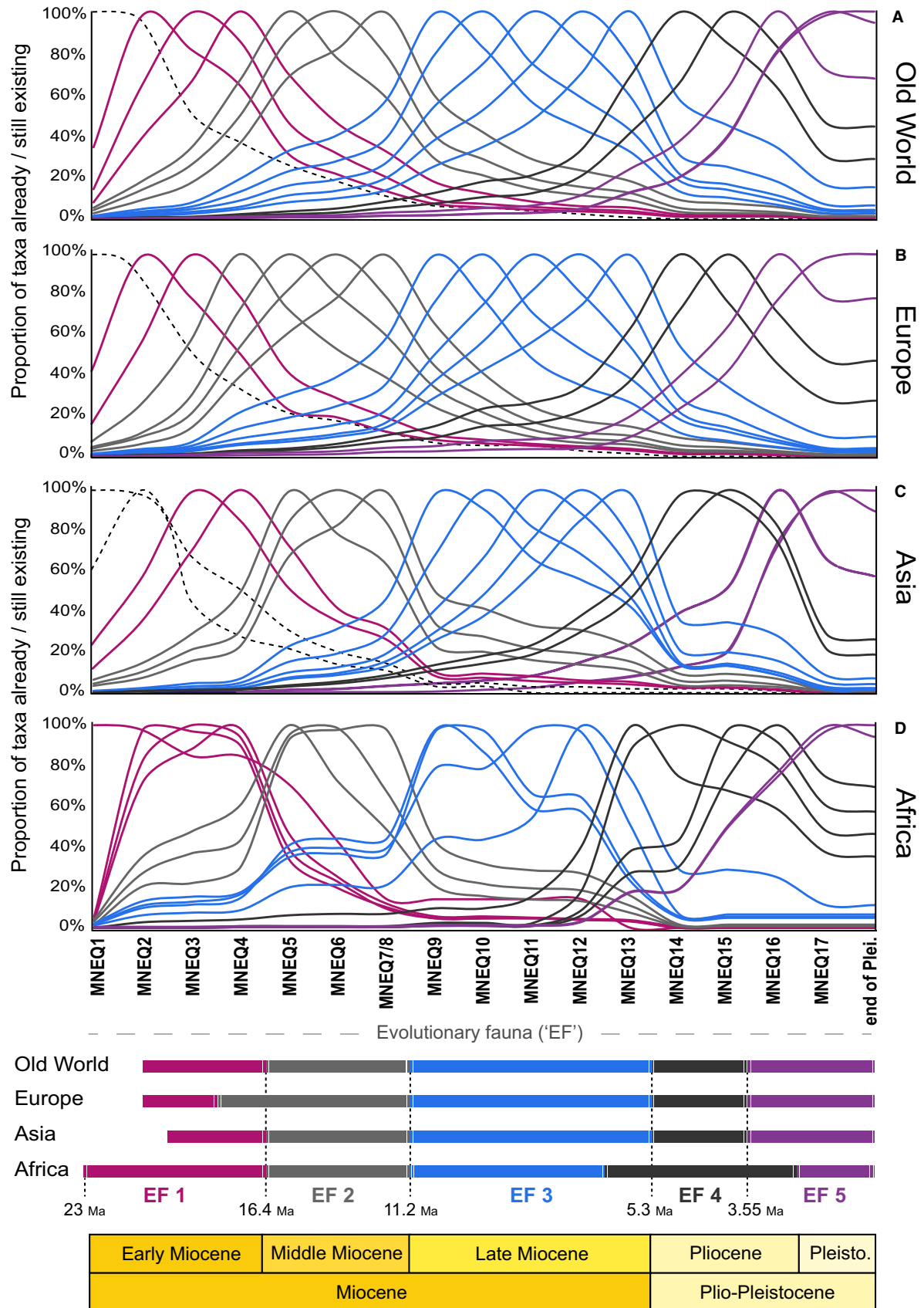
HCPC analysis is also illustrated on the UPGMA hierarchical clustering tree, directly calculated on the dissimilarity matrices to highlight the differences in results among methods (Dommergues *et al.* 2009; Appendix S2). HCPC and ordination methods should be more effective at detecting very large-scale spatial structure by generating two or three continental-scale bioregions of equivalent size, whereas UPGMA is more likely to isolate handfuls of localities inhabited by the rarest taxa from a homogeneous whole. As the number of clusters has to be specified for UPGMA and HCPC analyses, we started with the minimal number of clusters (i.e. two) to identify the major structure within the geographical distribution of species at the scale of the entire Old World (called hereafter ‘global scale’) and then at continental scales (i.e. Europe, Asia, Africa). The number of clusters was then gradually incremented (rooted increment for UPGMA, non-rooted for HCPC) in order to generate more detailed subdivisions of the spatial distribution of fossil mammals. Incrementing was stopped when subdivisions were concentrated within continents (for the global scale) or when geographically coherent bioregions would no longer be formed with both methods at continental scales. Bioregional maps shown in the main text are a selection of the most geographically consistent bioregions formed using HCPC and UPGMA analysis. All the maps produced with both methods are shown in Appendix S2 (Figs S5, S20). See Data Archiving Statement, below, for locations of raw data extracted from NOW database and all derived data (i.e. locality and taxon names from NOW and OGRE ordered by evolutionary faunas and bioregions). The Shungura occurrences extracted from OGRE unpublished dataset are grouped within a unique locality after thinning and included in the data files at the temporal scale of evolutionary faunas. An example Rmarkdown script with comments and figures can be found in Appendix S4, and in the GitHub and Dryad repositories, to help reproduce all analyses conducted in this study (Gibert Bret 2022).

RESULTS

Neogene and Pleistocene evolutionary faunas

Five evolutionary faunas (EF), corresponding to five large temporal units ranging from two to five MNEQs, were identified in Neogene and Pleistocene mammalian assemblages (Fig. 1). They are distinctly identified by the

FIG. 1. Poly-cohorts curves at an: A, Old World; B, European; C, Asian; D, African scale. Mammal zonation (here represented by MNEQ1–17; Agustí *et al.* 2001) are extracted from NOW database. Black dashed lines are cohorts excluded from the clustering analysis. The colour of each cohorts illustrates its respective evolutionary fauna; ‘end of Plei.’ refers to the period of the Pleistocene (previously MN18), not covered by MN17 (i.e. from 1.9 Ma to 0.011 Ma).



amplitude of extinctions and originations within the poly-cohort matrices, almost synchronously in all three continents (Table S1). While the mean background rates of origination and extinction of a given cohort from one MNEQ to another are 17% and 13%, respectively, four major turnover events led to the quasi-synchronous origination and extinction of a large number of taxa. On average, these turnovers depict 39% of new and 35% of extinct taxa from one MNEQ to the next one, greatly exceeding the background origination and extinction rates (Table S1).

At the Old World scale, a first EF (Fig. 1A; Early Miocene fauna) originated during MNEQ2 (21.7–19.5 Ma) and became extinct in a massive pulse at the end of MNEQ4 (16.4 Ma; Table S1A). The first cohort of the dataset (MNEQ1: 23–21.7 Ma) was not included in the first EF because it experienced an intensive extinction pulse in MNEQ2, independently of MNEQ2–4 cohorts (48% of MNEQ1 species were already extinct during MNEQ3). In addition, MNEQ2–4 cohorts were almost absent during MNEQ1 with respectively only 16% and 7% of taxa from MNEQ3 and MNEQ4 cohorts recorded. The second (MNEQ5–7/8: 16.4–11.2 Ma) and third (MNEQ9–13: 11.2–5.3 Ma) EFs match the division of MN in the Middle and Late Miocene (Agustí *et al.* 2001; Fortelius *et al.* 2014). The fourth EF combines cohorts from MNEQ14 and MNEQ15 and coincides with the Pliocene Zanclean stage. The last EF ('Plio-Pleistocene fauna') originated during the next MNEQ zone (MNEQ16: 3.55–2.5 Ma) and lasted until the end of our fossil record (part of Pleistocene after MNEQ17: 1.95–0.0117 Ma).

Similarly, five consecutive EFs are also identified at a continental scale by poly-cohort analysis (Fig. 1B–D). Their timing of origination and extinction are close but do not perfectly match those of the global scale poly-cohort analysis (Fig. 1A). For European mammals, the second evolutionary fauna originates during MN4 (17.2–16.4 Ma) rather than MNEQ5 (16.4–14.2 Ma) at the global scale (Fig. 1B; Table S1B), but the following faunas match the temporal dynamics of evolutionary faunas identified at the global scale. Fahlbusch (1989) proposed a similar grouping for European rodent faunas between MN4 and MN7/8 (12.85–11.2 Ma). Finally, a minor turnover event can be identified between MNEQ10 and 11, in particular among European mammalian faunas (Table S1).

At the Asian scale, the origination and extinction times of evolutionary faunas coincide with those described at the global scale, except for the first one consisting only of the MNEQ3 and MNEQ4 cohorts. Indeed, a large part (61%) of MNEQ2 taxa (Fig. 1C) were already present during MNEQ1 and disappeared (33%) before MNEQ3.

In Africa, the first EF included the MNEQ1 cohort. The fourth EF included MNEQ13 (Messinian) and MNEQ16 (Zanclean) cohorts that are respectively parts of EF3 and EF5 at all other scales (Fig. 1D). The last EF in

Africa is restricted to Pleistocene assemblages (from MNEQ17 to the end of the Pleistocene).

Distribution at the Old World scale

Following the establishment of EFs, bioregions were computed at the global and continental scales for each of the five evolutionary faunas. All bioregions illustrated in this paper were obtained using the Raup & Crick index analyses and a selection of HCPC and UPGMA bioregions was made according to their respective geographical coherence (Figs 2–5). Bioregions resulting from the Simpson index analyses, largely identical to the bioregions produced by Raup & Crick (except for one bioregion in Asia during the Plio-Pleistocene), are presented in Appendix S2, with all HCPC and UPGMA maps, and trees computed using both indices (Figs S5, S71). The detailed description of the iterative identification of bioregions at both global and continental scales is provided in Appendix S3, while only a short summary highlighting the major temporal and spatial changes in bioregional structure is provided here.

At the Old World scale, the first spatial structure is the division between faunas in northern Eurasia (Europe + northern Asia) and those in Africa and southern Asia (Fig. 2A, G, M). This pattern is illustrated by the first dimension of nMDS for EF1 (Fig. S21), EF3 (Fig. S27), EF5 (Fig. S33) and the first dichotomy of UPGMA and HCPC trees for EF2 (Figs S26, S25). Another major pattern at this scale is the spatial division of northern Eurasia. In the Early Miocene (EF1), northern Asia and northern Europe formed a large bioregion (Fig. 2C), latitudinally separated from taxa in south-eastern (central Europe to Anatolia) and south-western Europe (southern France, Iberian Basin). From the Middle Miocene to the Pliocene (EF2–4; Fig. 2F, I, K), a new spatial pattern emerges: the longitudinal separation of the northern Eurasian faunas, reaching a climax during Late Miocene (Fig. 2I; illustrated by the second dimension of nMDS: Fig. S27). Finally, at the end of the Pliocene and during the Pleistocene (EF5), the major structuring pattern of bioregions on a global scale is latitudinal. A large southern European region is formed along the Mediterranean coast starting in the Pliocene (Fig. 2L; EF4) and northern Eurasian faunas no longer split longitudinally into multiple bioregions but form a large bioregion reminiscent of the modern extension of the Palaearctic biogeographical realm, with the exception of its northern African part.

Distribution at the continental scale

Europe. In Europe, the main spatial structures of Miocene mammalian faunas are longitudinal (Fig. 3A, E, G). The

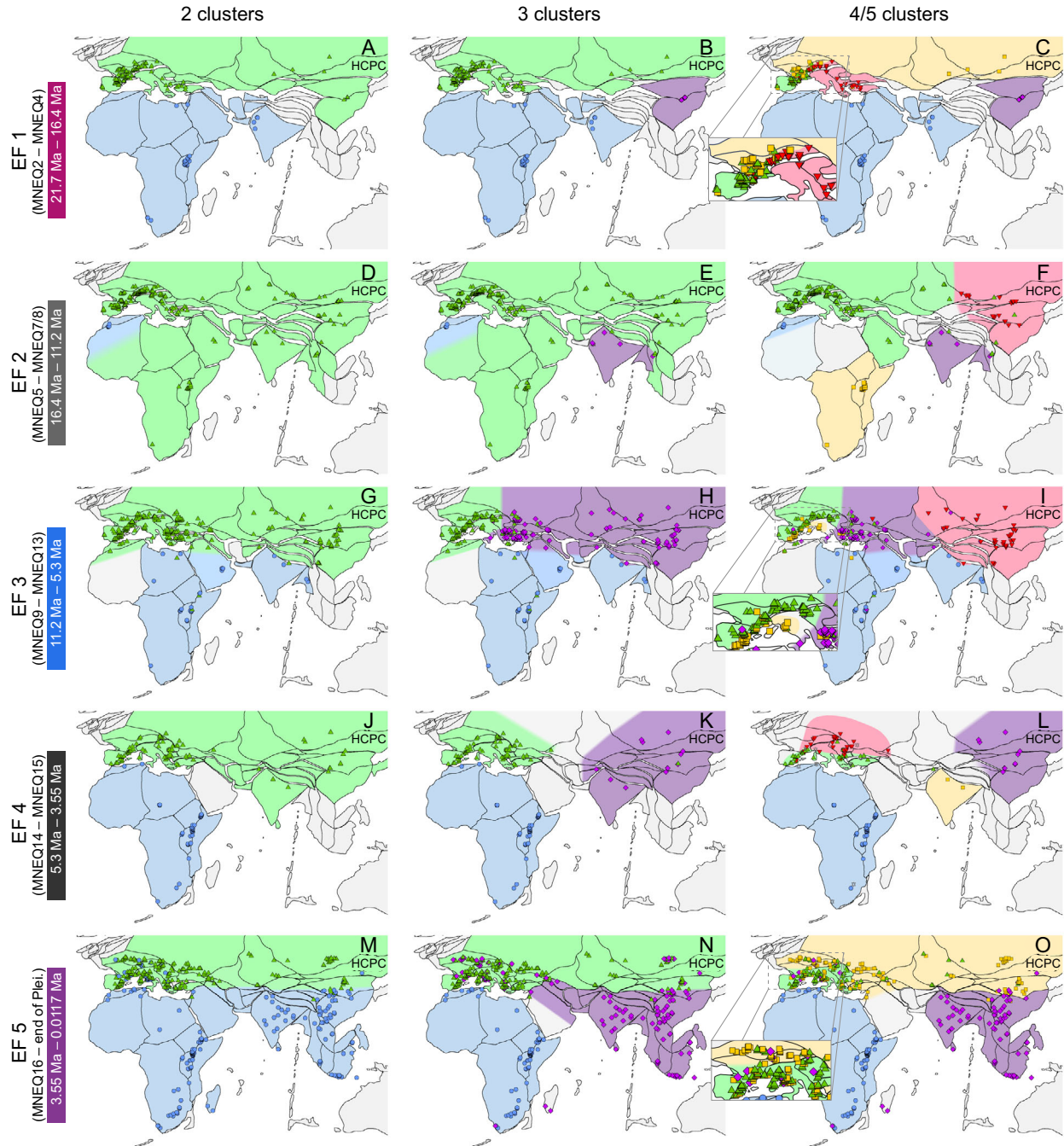


FIG. 2. Illustration of bioregions at the Old World scale computed with hierarchical clustering on principal components (HCPC) and Raup & Crick dissimilarity index. Iterative production of clusters from left (two clusters) to right (with a minimum of four clusters). Only continuous and significantly dissimilar clusters are pictured as bioregions (Table S3). Clusters are coloured according to their order of computation, from first to last: green, blue, purple, yellow, red and grey. Plate limits are extracted from GPlates (Müller *et al.* 2018).

first dimension of nMDS represents a longitudinal structure opposing western and eastern Europe for the Early, Middle and Late Miocene faunas (Figs S36, S39, S42). This pattern reaches a maximum during the Late Miocene (EF3) when the European faunas are divided into two

large bioregions of equal size. On the contrary, after the Mio-Pliocene transition (EF4 and EF5), the first axis of the nMDS depicts a latitudinal structure (Figs S45, S48), resulting in the emergence of a large Mediterranean biogeographical region separated from the north of the

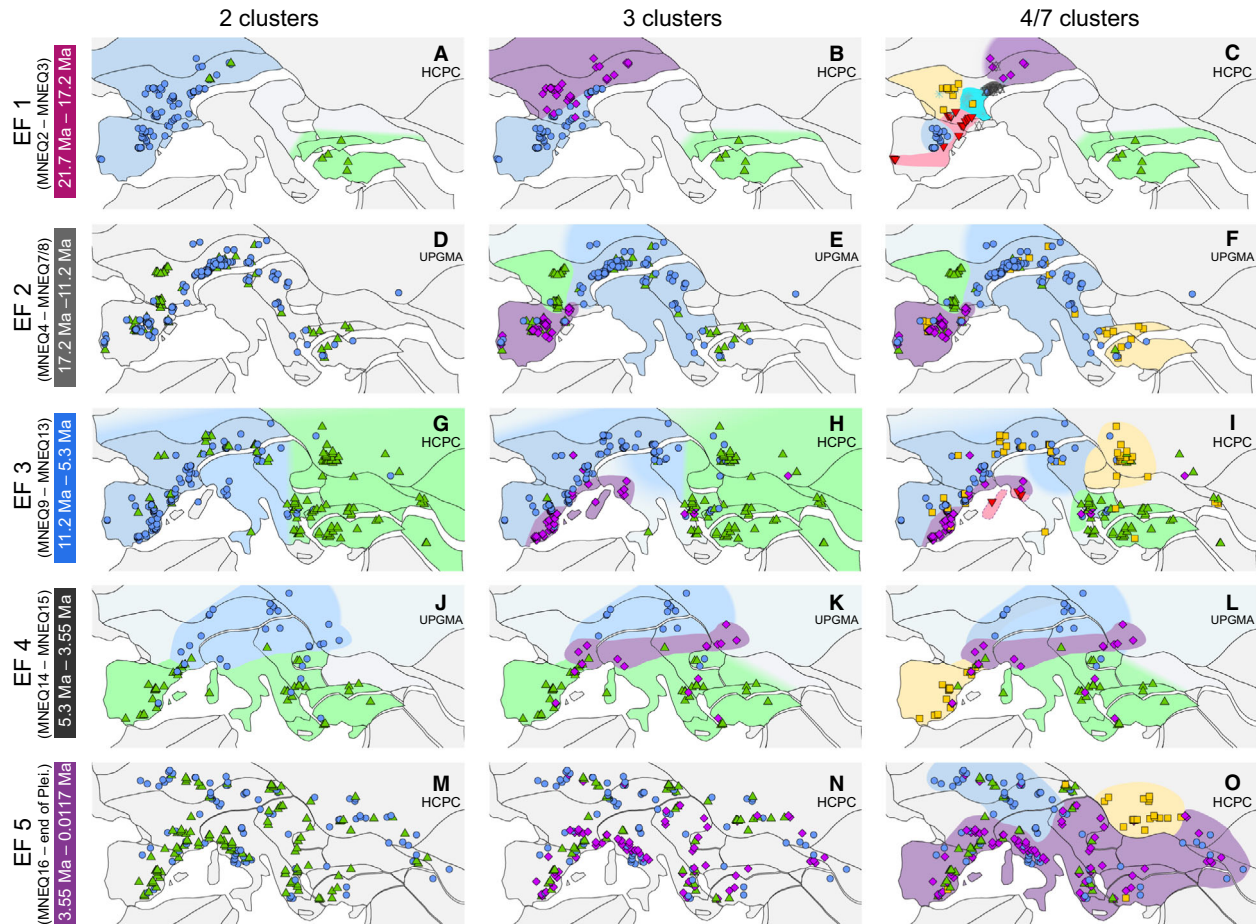


FIG. 3. Illustration of bioregions at the European scale computed with hierarchical clustering on principal components (HCPC) and hierarchical clustering on dissimilarity matrix (UPGMA) with Raup & Crick dissimilarity index. Iterative production of clusters from left (two clusters) to right (with a minimum of four clusters). Only continuous and significantly dissimilar clusters are pictured as bioregions (Table S3). Clusters are coloured according to their order of computation, from first to last: green, blue, purple, yellow, red, grey and light-blue. Plate limits are extracted from GPlates (Müller *et al.* 2018).

continent (Fig. 3J, O). This Mediterranean bioregion is emerging in the Late Miocene but remained restricted to western coasts (Fig. 3H). During the Pliocene (EF4), it extended to the eastern coast (Fig. 3) and progressed northwards continuously throughout the Pleistocene (Fig. 3O). Finally, whereas the first evolutionary fauna (EF1) can be spatially divided into numerous bioregions (Fig. 3C), the last one (EF5) struggles to form geographically coherent clusters (Fig. 3N).

Asia. In Asia, the main geographical pattern of mammalian assemblages is very stable over time: the assemblages surrounding the Indian sub-continent and south-eastern assemblages are consistently differentiated from northern assemblages in ordination and hierarchical clustering, with both dissimilarity indices for the five evolutionary faunas (Fig. 4; Figs S9, S10). The second major spatial

pattern is the longitudinal division of bioregions in northern Asia (Fig. 4B, F, I, K, O). In contrast to Europe, which experienced a change from a longitudinal to a latitudinal structure between the Miocene and Pliocene, the northern Asian fauna remained longitudinally structured from the Early Miocene to the Pleistocene. This longitudinal division reaches a maximum in the Middle and Late Miocene (EF2, EF3: Fig. 4F, I). Finally, during Pleistocene (EF5), the south-eastern mammalian fauna is distinct from the northern and Indian faunas (Fig. 4O); however, the temporal dynamic of this bioregion cannot be clearly identified as its sampling is insufficient during the preceding EFs.

Africa. On the African continent, the only EFs analysed with UPGMA and HCPC were the Early Pliocene (EF4) and Plio-Pleistocene (EF5) ones, as the record is too

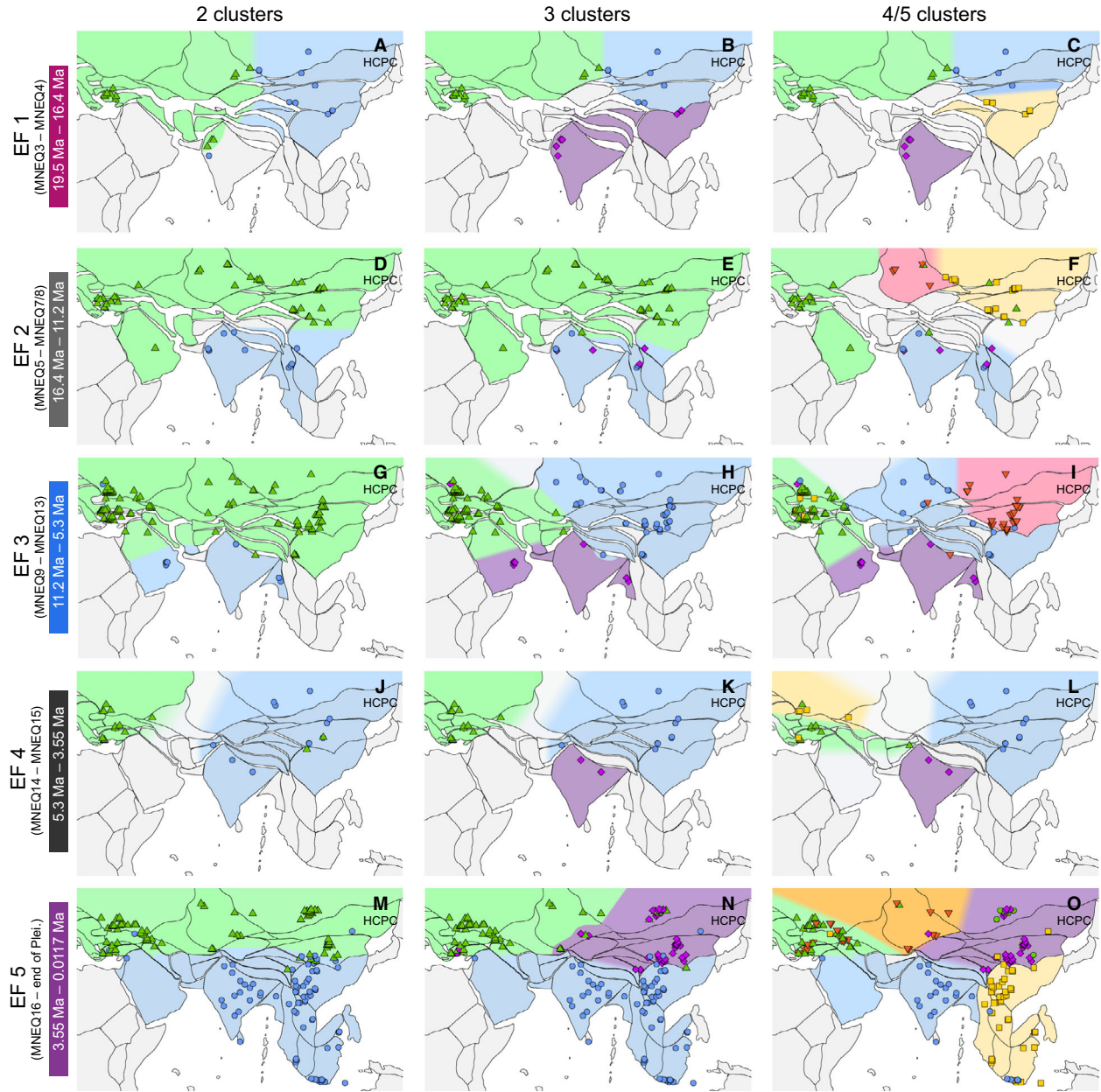


FIG. 4. Illustration of bioregions at the Asian scale computed with hierarchical clustering on principal components (HCPC) with Raup & Crick dissimilarity index. Iterative production of clusters from left (two clusters) to right (with a minimum of four clusters). Only continuous and significantly dissimilar clusters are pictured as bioregions (Table S3). Clusters are coloured according to their order of computation, from first to last: green, blue, purple, yellow and red. Plate limits are extracted from GPlates (Müller *et al.* 2018).

sparse and limited to eastern Africa throughout the Miocene. During the Early Pliocene (EF4), HCPC failed to produce two large scale coherent biogeographical regions (Fig. 5A). Increasing the number of clusters leads to four coherent bioregions (Fig. 5B, C): one in the east, one in the south, and two in the north made of assemblages from Morocco, Algeria and Tunisia on one hand, and

from Libya, Egypt and the Arabian Peninsula on the other hand.

The Plio-Pleistocene fauna (EF5) has a comparable spatial structure, first with eastern assemblages isolated from the rest of the continent (Fig. 5D, E), second with the emergence of a northern bioregion made of assemblages surrounding the south Mediterranean coast, and finally

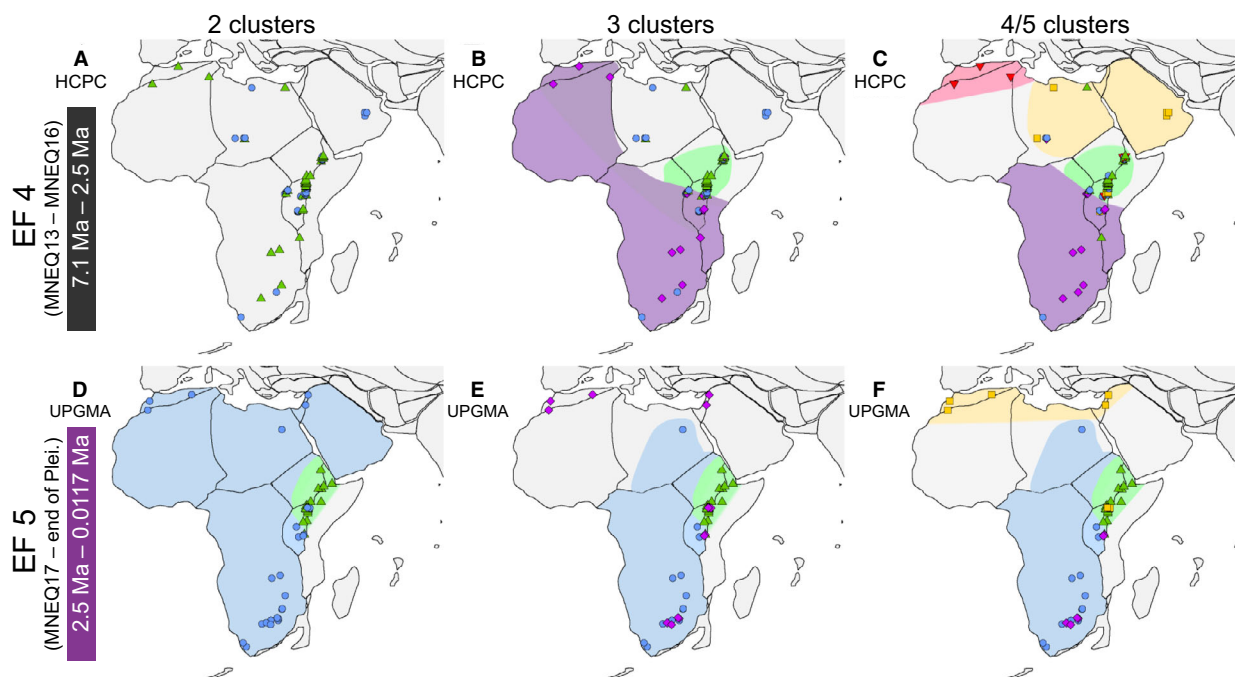


FIG. 5. Illustration of bioregions at the African scale computed with hierarchical clustering on principal components (HCPC) and hierarchical clustering on dissimilarity matrix (UPGMA) with Raup & Crick dissimilarity index. Iterative production of clusters from left (two clusters) to right (with a minimum of four clusters). Only continuous and significantly dissimilar clusters are pictured as bioregions (Table S3). Clusters are coloured according to their order of computation, from first to last: green, blue, purple, yellow and red. Plate limits are extracted from GPlates (Müller *et al.* 2018).

with one larger bioregion made of southern assemblages (Fig. 5F). For EF4 as well as EF5, eastern Africa encompasses many localities and is clearly separated from the other bioregions of the continent. Nevertheless, in the southern part of it, a transition zone with the southern bioregion is evidenced by the presence of closely spaced localities attributed to both bioregions, notably in Kenya and Uganda for EF4. In contrast, during EF5 these localities are attributed to the eastern bioregion and the southern bioregion begins in Tanzania. Between these two EFs, the separation between the eastern and southern bioregions became stronger (i.e. Fig. 5A, D) and the two bioregions previously on the same side of the clustering tree (Fig. S68) are now separated by the first dichotomy structuring African faunas at EF5 (Fig. S71).

DISCUSSION

Five evolutionary faunas

Our poly-cohort analysis reveals five major evolutionary faunas over the entire Neogene to Pleistocene interval, regardless of spatial scale (i.e. Old World or continents; Fig. 1). Their respective durations, as well as times of

emergence and decay, closely match the sub-epoch chronology for Miocene faunas and the stage scale for the Plio-Pleistocene faunas. Their periods of emergence and decay (i.e. breakpoints highlighted by colour changes in Fig. 1) are decoupled from known periods of strong sampling variations (e.g. in Europe: MN6, MN10, MN11 and MN12; Peláez-Campomanes & Van der Meulen 2009; Madern & Van den Hoek Ostende 2015). This non-subjective temporal division adapted to the evolution of Neogene and Pleistocene mammals, between the sub-epoch (e.g. EF2, EF3) and the stage scale (EF4), is consistent with the division into five faunas proposed by Fahlbusch (1989) for European rodents. With our poly-cohort analysis, this macroevolutionary pattern holds true not only at the global scale, but also for Asia and Africa. This synchronization between Old World, European, Asian and African EFs identified via poly-cohort analysis, reflects the global nature and scale of the forcings that gives rise to the breakpoints dividing EFs (Escarguel *et al.* 2011; Cantalapiedra *et al.* 2014, 2015), such as global variations in temperature and precipitation leading to major change of biomes or palaeogeographical evolutions, along with new dispersal routes.

One of these five faunas is well known and already identified at the Old World scale: the Late Miocene evolutionary fauna, the key taxon of which (the equid

Hipparion) dispersed from North America (Eronen *et al.* 2012). This Late Miocene evolutionary fauna (EF3) is already described at the scale of Europe and Asia, where it is respectively known as the Pikermian and Baodean chronofaunas (Eronen *et al.* 2009; Atabadi *et al.* 2013). The composition of this evolutionary fauna as well as its emergence and fading timing have been interpreted as the consequences of variations in aridity and temperature at the global scale (Bernor 1983; Agustí & Moyà-Solà 1990; Fortelius *et al.* 2003, 2014). At the end of the Middle Miocene, the global temperature decreased, aridity spread (Van der Meulen & Daams 1992), and steppe punctuated of open forests extended until it formed a nearly continuous open biome between the three continents of the Old World (Denk 2016; Kaya *et al.* 2018). A new fauna adapted to this open environment emerged at the expense of the Oligocene and Early Miocene taxa (e.g. *Tragulidae*; Barry *et al.* 2002), which were more adapted to the extensive forest cover that previously occupied a vast part of the Old World (Casanovas-Vilar *et al.* 2005; Costeur & Legendre 2008; Larsson *et al.* 2011). This Late Miocene fauna (EF3) declined and was replaced by EF4 during the Early Pliocene, when temperature increased again at global scale (Zachos *et al.* 2001) as well as habitat homogeneity increased in Europe (Maridet *et al.* 2007; Maridet & Costeur 2010; Sniderman *et al.* 2016).

Beyond these major climatic changes, evolutionary fauna succession is also the result of inter-continental dispersal and faunal mixing through the formation of land bridges via plate tectonics and eustatic variations. The Early Miocene evolutionary fauna (MNEQ2–4), for example, gradually disappears (Scherler *et al.* 2013) after two events of intense faunal mixing between Africa and Eurasia, known as the Proboscidean Datum Event (PDE) or *Gomphotherium* Landbridge (Harzhauser *et al.* 2007) and *Creodonta/Deinotherium* landbridge event (Tassy *et al.* 1989; Van der Made 1999; Sen 2013), when connections between Africa and Eurasia began to be frequent. The Arabia–Eurasia collision led to the narrowing and shallowing of the Neotethys seaway at the beginning of the Miocene (McQuarrie & Van Hinsberger 2013; Pirouz *et al.* 2017). A drastic reduction of water exchanges between the Mediterranean Sea and the Indian Ocean across the Neotethys seaway (Mesopotamian Trough) occurred during the early Burdigalian (*c.* 20 Ma) according to marine biogeographic divergences (Harzhauser *et al.* 2007) and neodymium (Nd) isotope records (Bialik *et al.* 2019) and ceased during the Langhian. This is in good agreement with the onset of faunal exchange between Africa and Eurasia during MN4 (Fig. S1: *c.* 16–17 Ma, end Burdigalian and Early Miocene; Rögl 1999). Intermittent Arabia–Eurasia continental connections induced the PDE during the late Early Miocene, coinciding with the end of EF1, and the permanent closure of

the Neotethys Seaway at the end of the Middle Miocene corresponds to the end of EF2 (Harzhauser *et al.* 2007). Furthermore, the fate of the Paratethys, an epicontinental sea in Eurasia, has been a crucial factor in the dynamics of mammalian faunas. During the Early Miocene, the Paratethys consisted in a series of connected basins extending over Eurasia from the Vienna basin in the west to the Kopet Dag in Turkmenistan in the east (Popov *et al.* 2006). The central Paratethys was isolated from the eastern Paratethys at 11.3 Ma (MN7/8; end of EF2) in response to eustatic level fall and/or to the Carpathian orogeny (Borghi *et al.* 2013). Extension of the Neotethys and Paratethys seaways affected not only the dispersal routes of species, but also had a profound effect on Eurasian and African climates and biomes (Zhang *et al.* 2014). Similarly, the transition period between EF3 and EF4 in Africa (Fig. 1D) coincides with the Messinian Salinity Crisis (MSC; *c.* 6.8 Ma to *c.* 5.3 Ma) at the African scale, when the Mediterranean Sea shrank during the Messinian (MN13) as a result of the restriction of the Atlantic inflow across the Rifian corridor (Warny *et al.* 2003). In three short phases, north-western Africa (Algeria, Morocco, Tunisia) and south-western Europe (Italy, Spain) faunas were connected by landbridges and mixed (Agustí *et al.* 2006; Gibert *et al.* 2013; García-Alix *et al.* 2016). The MN13 cohort is the first of EF4 in Africa (Fig. 1D) whereas it is the last cohort of EF3 for Europe, Asia and the Old World (Fig. 1A–C), illustrating the stronger impact on African faunas than on Eurasian faunas. This difference can be explained by three factors: (1) the dispersal of taxa from Africa to Europe via Anatolia was already occurring in the Vallesian and Turolian (MNEQ9–12; i.e. EF3; Azzaroli & Guazzone 1979), reducing the impact of the MSC on faunal composition; (2) the new taxa arising from the MSC disappeared rapidly (50% after 1 Ma; Jaeger *et al.* 1987); and (3) Late Miocene localities with diversified faunas are rare (>5 taxa; Fig. S2) in Africa (MNEQ9–12), amplifying the effect of MSC dispersals in comparison with Europe (Fig. 1B, D). Finally, for Africa and contrary to others poly-cohorts, the EF5 starts at the Plio-Pleistocene boundary (MNEQ17), when aridity increased significantly and caused large changes in faunal composition and abundance (deMenocal 2004), as documented in bovids (Bobe & Eck 2001).

The step-like shape and more pronounced turnover of both Asian and African poly-cohorts (Tables S1, S2) compared to those of Europe is an effect of the biostratigraphic zonation used here (i.e. MN), defined by the simultaneous occurrence of wide-range European species, and in which species are often defined at the scale of a single or a pair of MNs. This is thus a mechanical/biostratigraphical rather than a macroevolutionary effect. By using MNEQ, African and Asian taxa are more likely to cross MN boundaries than European ones and generate

higher origination/extinction rates. Furthermore, African and Asian localities are more often defined at a lower temporal resolution than the stage, and the greater temporal variability of their respective sampling intensity also contributes to differences in poly-cohort curve shapes.

Patterns of distribution

From longitudinal to latitudinal bioregions. At the global scale, and throughout the Neogene and Pleistocene, the major distribution pattern identified by the computation of a minimum of bioregions (Fig. 2) confirms the proximity of southern Asian faunas with African ones (Flynn *et al.* 1990; Bibi 2011). This similarity can be the consequence of iterative dispersal events from Africa and the Middle East to the west, such as those documented between 20–16 Ma and 9.5–7.4 Ma (Barry *et al.* 1985; Bibi 2011). The two EFs including these dispersal events (Early and Late Miocene) are marked by the stability of the African and southern Asian bioregion through iterative clustering (Fig. 2A–C, G–I) and its dissimilarity with northern Eurasia. Furthermore, during EF5, this pattern opposing northern Eurasia and southern Asia + Africa is reinforced by south-eastern Asian assemblages that are significantly better sampled in comparison to the Neogene (Fig. 2M–O). The biogeographical pattern that separates north and south Asian faunas along the Himalayan range is observed in all of our analyses (Fig. 4) and reveals the importance of an orographic barrier to the temporal stability of modern Palaearctic and Indomalayan biogeographical realm limits (Udvardy 1975). In fact, these two biogeographical realms are isolated from each other by relief and high plateaus uplifted during the Cenozoic along the Alpine–Himalayan orogen. The southern Tibetan plateau (Lhasa block) had reached high elevations (*c.* 3.1–4.7 km) prior to the Early Miocene (Ingalls *et al.* 2020) whereas the Himalayas attained present day elevations around the late Early Miocene (Gébelin *et al.* 2013; Ding *et al.* 2017). Further west, the Arabian and Eurasian continents collided during the late Oligocene leading to the uplift of the Zagros fold belt (Pirouz *et al.* 2017). These orographic barriers structure the nature of mammalian faunas to such an extent that their effect is also visible among groups rarely associated with the establishment of biogeographical regions, such as Miocene carnivores (Grohé *et al.* 2020). Interestingly, when African faunas are strongly associated with the faunas of southern Asia (*i.e.* in the Early and Late Miocene) the southern border of the Eurasian (*i.e.* the future Palaearctic realm) biogeographical realm diverges from modern boundaries and is located in Eurasia rather than in Africa (Fig. 2A, G).

The rest of the Old World is marked by a strong affinity between European and northern Asian faunas. In the

Early Miocene, as in the Plio-Pleistocene, the northern European and northern Asian assemblages have more similar compositions than northern and southern European assemblages (Fig. 2A–C, M–O; Wang *et al.* 2015). From the Middle Miocene to the Pliocene, the northern faunas of Eurasia can be longitudinally subdivided into bioregions, a partitioning that corresponds to important environmental and climatic zones (*e.g.* continentality) and changes. For example, the biome of warm-temperate evergreen broadleaf and mixed forest that dominated Europe and Asia during the Early and Middle Miocene contracted during the Late Miocene, leaving space for a new biome of temperate deciduous savannah (Strömberg 2011; Pound *et al.* 2012). The Central European Wet Zone (Utescher *et al.* 2011), identified with palaeobotanical evidence, structured the Early and Middle Miocene (EF1, EF2) of Europe. This vast wet zone in central Europe can be identified within the Middle Miocene (EF2) central European bioregions (Fig. 3E–F, blue circles) and began to be replaced in its eastern part by the Late Miocene favoured by the shrinkage of the Paratethys. This new savannah biome covered a vast region from eastern Europe (28° of longitude) to Iran, Kazakhstan and western China (Costeur *et al.* 2004; Kaya *et al.* 2018), and was inhabited by a new steppe-adapted fauna (the Pikermian or Baodean chronofaunas) comprising an increasing proportion of Bovidae, Equidae and Giraffidae (Fortelius *et al.* 2014). This fauna/biome ecosystem is identified by our HCPC analysis at the Old World scale (Fig. 2H, L, purple diamonds) and in the east of Europe at the European scale (Fig. 3G–I, green triangles). During the Messinian and the Mio-Pliocene transition, temperature and humidity increased for a short time (Griffin 2002; García-Alix *et al.* 2008), the steppe biome fragmented (Kaya *et al.* 2018), and the Pikermian and Baodean chronofaunas disappeared (Fortelius *et al.* 2006). The decline of this bioregion is less obvious on a global scale because the fossil record to the east of Europe is very scarce during the Early Pliocene (EF4) compared to the Late Miocene (EF3; Fig. 2G–I). However, like the great savannah biome, the longitudinal division inherited from the Eocene and Oligocene faunas (Legendre 1987), and dominating during the Miocene in Europe, fades away to be replaced by a latitudinal structuring during the Pliocene (*i.e.* at a global (Fig. 2I, L, O) and European (Fig. 3L, O) scale), as the latitudinal gradient of precipitation and temperature becomes steeper at the end of Miocene (Utescher *et al.* 2011).

Regarding the temporal evolution of Neogene mammals, the extension of Neotethys and Paratethys was an important factor of their spatial distribution in Eurasia and Africa. The broad continental extension of Paratethys at the beginning of the Miocene allowed the presence of the Central European Wet Zone (EF1, EF2), and its

retreat in combination with the closure of Neotethys seaway resulted in the aridification and increased seasonality of central Asia, eastern Europe and northern Africa, with the formation of Sahara Desert during the Tortonian (EF3) (Zhang *et al.* 2014; Tzanova *et al.* 2015). Combined with Paratethys shrinking, the global cooling following the MMCO (EF1/EF2) also increased seasonality and caused major changes in temperature and precipitation gradients (Ramstein *et al.* 1997; Fluteau *et al.* 1999). The latitudinal gradient of temperature, weaker at the beginning of Miocene, only increased slightly in intensity in the Late Miocene (EF3) to become more comparable to the modern gradient in the Pliocene and Pleistocene (EF4, EF5; Bruch *et al.* 2004). Consequently, in the western Asian and European bioregions, latitudinal structures emerged from the Pliocene (EF4) and persisted during the Pleistocene, particularly in Europe (EF5). Furthermore, in deep time, latitudinal gradients of diversity are only detected during icehouse periods and particularly when seasonality is strong (Mannion *et al.* 2014). A similar transition from an east–west structuring during the Miocene to a more north–south dynamic during the Pliocene has been shown for spatio-temporal patterns of hypsodonty (a proxy for palaeoprecipitation: Fortelius *et al.* 2006). This change in climatic and palaeogeographical patterns impacts the expression of evolutionary traits in mammals (e.g. hypsodonty) but also the taxonomic composition of assemblages, by selection and/or dispersal (Fig. 2H compared to Fig. 2K). Regarding flora, from the Pliocene onwards temperate deciduous and mixed forests became the dominant biome from northern Europe to Siberia, while continuing their southward progression (Mai 1995).

Emergence of a Mediterranean bioregion. At the scale of Europe, the latitudinal structuring identified by both HCPC and UPGMA led to the formation of a Mediterranean bioregion during the Pliocene. The onset of this spatial structuring has been identified for small mammals at the end of the Miocene (i.e. MN9–MN14; Maridet *et al.* 2007) but the pattern remained difficult to identify in the Pliocene because of the lack of rodent-rich localities in eastern Europe. In our dataset consisting of all non-flying terrestrial mammals, we can observe the persistence of this Mediterranean bioregion during the Pliocene and Pleistocene. This bioregion finds its roots in the Late Miocene when Spanish and French coastal assemblages became similar in composition (Fig. 3H) but not until the Pliocene did it extend further east to include the Balkan, Anatolian and Bulgarian assemblages (Fig. 3J). This first increase of similarity between the Spanish and French coastal faunas during the Late Miocene is also visible among large mammals alone (Costeur *et al.* 2004).

The large peri-Mediterranean region formed during the Pliocene and Pleistocene (EF4, EF5) is limited to southern

Europe or northern Africa only and does not extend around the whole Mediterranean Sea (Figs 2L, O, 5F). Despite the fact that numerous faunal exchanges have taken place since the Late Miocene, notably through Anatolia (Agustí *et al.* 2006; Fortelius *et al.* 2014), faunas on the two sides of the Mediterranean Sea remain significantly dissimilar (Table S3). However, during EF3 (Fig. 2G), faunas from north-western Africa were similar to European ones. This compositional similarity may be the consequence of direct faunal exchanges across the dried-up Mediterranean Sea (Jaeger *et al.* 1987; Gibert *et al.* 2013) during the MSC, since MN13 (i.e. Messinian) is integrated at the global scale in EF3. On the scale of the African continent, the Messinian event is included in EF4, and for this evolutionary fauna, the north-western part of Africa (i.e. the localities of Algeria, Morocco and Tunisia; García-Alix *et al.* 2016), where mammals dispersed from south-western Europe and vice versa (i.e. Italy, France, Spain; Azzaroli & Guazzone 1979) are separated from the rest of northern Africa (Fig. 5C, red cluster). During the Pliocene and Pleistocene, the Mediterranean Sea was flooded again and this separation disappears (Fig. 5F) in favour of a single Mediterranean region in Africa separated from the Mediterranean region in southern Europe.

The establishment of modern bioregions dates back to the Pleistocene for Chinese mammal families (He *et al.* 2018). Looking at the bioregional division in Europe, we also observe the emergence of a modern pattern (Heikinheimo *et al.* 2007) at the very end of the Pliocene and during the Pleistocene, with the appearance of a large Mediterranean bioregion and two (east and west) European bioregions further north (Fig. 3O). The division of European bioregions in the Pliocene and older periods (Fig. 3A–L) shows patterns that are significantly different from modern ones, therefore the timing of emergence of modern biogeographical patterns is comparable between Asia and Europe. At the global scale, the Plio-Pleistocene evolutionary fauna does not spatially split exactly like modern biogeographical realms. The Palaearctic ecozone extending from northern Asia through Europe to northern Africa, does not exist in our analyses (Fig. 2M–O). Northern African faunas remain closer to southern Asian assemblages than to the rest of the continent (Fig. 2M), and the ‘Palaearctic’ Plio-Pleistocene ecozone remains limited to northern Eurasia. In fact, the modern boundary of Palaearctic ecozone shifted southward to northern Africa at the very end of the Pleistocene (Tropic of Cancer; Geraads 2010; Bibi 2011), therefore on our time scale, we can only observe past configuration of this ecozone where African faunas are more linked with southern Asian faunas than with European faunas, especially for large mammals (Geraads 2010; Bibi 2011).

On the African continent, if fossils localities are concentrated in eastern Africa (almost exclusively during

the Miocene, and to a lesser extent during the Plio-Pleistocene), it is because mammalian sites have mostly been preserved in active sedimentary basins such as those resulting from the East African Rift System (EARS). Their formations are related to active rifting which began during the Middle Miocene (Chorowicz 2005) or as early as the latest Oligocene (Roberts *et al.* 2012). During the best-sampled evolutionary faunas, the eastern and southern bioregions experienced stronger dissimilarities in the Pleistocene (EF5) in comparison with the Pliocene (EF4), as evidenced by the disappearance of the transition zones between them (Fig. 5B, E), and their respective clustering trees (Figs S58, S71). In the Pliocene, this transition zone (e.g. localities in Malawi, Tanzania) is inhabited by taxa associated with both bioregions (Grubb *et al.* 1999; Fig. 5C), and disappears when the heterogeneity between these two bioregions increases (Patterson *et al.* 2014; Fig. 5F) in parallel with the increasing aridification between 2.8 and 2.3 Ma (i.e. MNEQ17; Bobe & Eck 2001; deMenocal 2004), because dispersal between these two bioregions depends on variations in aridity. The eastern and southern bioregions are separated by a large arid region and connected by a discontinuous route along the Kingdon line (Grubb *et al.* 1999). As humidity increases, this 'southern route' becomes continuous, as it did in the Pliocene, disappearing along with the transition zone between 3 and 2 Ma when aridity increased again.

Establishing bioregions for extinct mammals

Methods, indices and transition zones. The spatial distributions of all evolutionary faunas were analysed using both UPGMA and HCPC (Figs S5–S20). However in Figs 2–5, only 9 out of the 51 selected configurations (i.e. the most geographically coherent) were calculated using UPGMA. This significant difference shows the importance of using methods designed to identify taxonomic composition gradients such as ordination methods (here, HCPC), rather than only hierarchical clustering methods, when trying to compute bioregions within the fossil record (Brayard *et al.* 2007; Dommergues *et al.* 2009). Another major benefit of the HCPC method lies in its ability to objectify the decision of clustering localities from an ordination analysis such as the nMDS, for which resulting clusters additionally possess, here, the advantage of being more geographically coherent than clusters resulting from only a hierarchical method.

While the comparison of UPGMA (e.g. Fig. S5E–H) and HCPC (e.g. Fig. S6E–H) maps indicates a greater geographical coherence of clusters established by the HCPC method, several problems classically associated with clustering persist. For example, localities associated

with one bioregion can be found isolated within another, clusters without any geographical coherence can be computed (e.g. Fig. 2L), and the problem of transition zones remains pronounced (Brocklehurst & Fröbisch 2018). The Pyrenean area during EF4 (Fig. 3K) or central Europe during EF3 (Fig. 3G) are good examples of transition zones that are potentially problematic for defining a boundary between two bioregions. The rule applied here is continuity: if a locality is surrounded by localities from another cluster, it is not used to define the boundary of a bioregion.

Beyond these specific cases, where localities situated close to bioregion boundaries can be problematic, one must be aware that the boundaries illustrated in the figures here have no proper existence in clustering and ordination analyses. In fact, neither distance nor the presence of biogeographical barriers such as mountain ranges or seas are considered to divide localities into bioregions, but only their taxonomic compositions. For example, bioregions at the European scale during EF1 are clearly not clustered along natural barriers (Fig. 3C), nor are the longitudinally structured bioregions for EF2 and EF3 in the north of Eurasia at global scale (Figs 2F, I, 4F, I). On the other hand, even if those barriers are not part of the analysis, they stand-out in time-stable boundaries, such as the one between the northern and southern Eurasian faunas (e.g. Fig. 2A, E, H) or the separation of African and European faunas (e.g. Fig. 2A, J), respectively located along a mountain range (Himalayas) and a sea (Mediterranean). Finally, as our analyses output only bioregions and not boundaries, we tried to avoid allocating empty areas to bioregions (e.g. Figs 2K, 3B, 4F). However, in some cases large empty areas are assigned to a bioregion in order to connect very distant localities (e.g. central and western Asia in EF1; Fig. 4C).

When clustering fails. For some evolutionary faunas, the computation of only two or three geographically coherent clusters is impossible, as in Europe during the Middle Miocene (Fig. 3D), in Africa for EF4 (Fig. 5A), and particularly in Europe for EF5 (Fig. 3M–N). During the Middle Miocene of Europe, the establishment of only two bioregions is prevented by the computation of a vast central region (Fig. 3D, blue circles) surrounded by two small taxonomic clusters (Fig. 3D, green triangles) in contradiction of the continuity rule. Furthermore, for the Pliocene evolutionary fauna in Africa (Fig. 5A), or the Plio-Pleistocene fauna in Europe, both methods employed, and indices, completely fail to compute any geographically coherent clusters when the minimum number of cluster is generated (Fig. 3M). In Europe, a similar incremental production of bioregions for present-day mammals also failed to form only two geographically

coherent clusters by generating a large central bioregion, mirroring the results for Middle Miocene mammals (Fig. 3D; Heikinheimo *et al.* 2007). But when the number of modern bioregions is increased, the large Mediterranean bioregion also appears in present-day mammals and is opposed to faunas in western and eastern Europe, as seen in the Plio-Pleistocene mammalian evolutionary fauna (Fig. 3O; Heikinheimo *et al.* 2007). However, in contrast with present-day European mammalian faunas, we are unable to provide more than three coherent bioregions with the Plio-Pleistocene faunas (Fig. S7T); this problem of only being able to produce two (or at most three) geographically coherent bioregions persists when only Pleistocene taxa are analysed (end of Pleistocene, later than MN17: 1.95–0.0117 Ma; Fig. S72). In Europe, over the successive evolutionary faunas, the oldest one provides the largest number of bioregions (Fig. 3C). While it is difficult to establish more than four coherent and significantly dissimilar bioregions for each evolutionary fauna in Europe, the Early Miocene faunas can be divided into seven bioregions, and the Plio-Pleistocene faunas into three, counterintuitively illustrating that there is no link between the quality/age (number of localities/occurrence) of a fossil record and our ability to establish biogeographical regions at a small spatial scale.

If the age of a fauna is not a good predictor of bioregion reconstruction, it would appear that compositional heterogeneity is more reliable. The study of similarity indices for assemblages of small (Maridet *et al.* 2007) and large Neogene mammals (Costeur *et al.* 2004) converge towards maximum heterogeneity in the Late Miocene and Late Pliocene, and greater homogeneity in the Middle Miocene and the Early Pliocene. Compared to our bioregion results, we cannot establish a clear link between the mean level of heterogeneity of a fauna and its ability to produce multiple coherent and significantly different bioregions. However, periods of high compositional heterogeneity among European mammalian faunas are not linked to a better spatial division into bioregions (e.g. Late Miocene bioregions: Fig. 3I). On the contrary, Early Miocene assemblages are on average more homogeneous (especially for small mammals) and can nevertheless be spatially divided into seven coherent bioregions (Fig. 3C). In fact, the strong compositional heterogeneity described in the Late Miocene for both small and large mammals is usually linked to an increase in spatial and temporal environmental heterogeneity (Costeur & Legendre 2008; Koufos & Konidaris 2011) when European environments open up and evergreen forests become fragmented (Kaya *et al.* 2018), seasonality increases, and tree species change from evergreen to deciduous (Pound *et al.* 2012). Conversely, periods of low heterogeneity at the beginning of the Pliocene and Miocene periods are linked to an increase in humidity and a progression of the evergreen

forest (Fortelius *et al.* 2014). Therefore, in the European Neogene and Pleistocene mammal dataset, the greatest number of bioregions can be formed when heterogeneity is low but temporal stability and environmental homogeneity are high. Smith *et al.* (2020) showed similar results with modern faunas, demonstrating that tropical regions with high temperature and low seasonality form bioregions more easily than temperate regions with high seasonality. This biogeographical pattern is explained by the temporal stability of tropical regions (Condamine *et al.* 2012), as tropical faunas have a longer evolutionary history, allowing them to accumulate greater spatial diversity than faunas in temperate regions. This contrasts with regions with strong seasonality or temporal instability (e.g. higher latitudes) where taxa will disperse rather than adapt. Our inability to form many (or only two) bioregions with Plio-Pleistocene evolutionary faunas in Europe (Fig. 3O) is therefore a consequence of the strong Pleistocene climatic instability generated by glacial–interglacial cycles (Lisiecki & Raymo 2007; Kahlke *et al.* 2011). These cycles are known to have deeply impacted the spatial distribution of European mammalian species (Sommer & Nadachowski 2006; Banks *et al.* 2008; Lister & Stuart 2008), thus grouping all European localities over the last two million years into a single biozone prevents the computation of spatially coherent clusters outside of major bioregions like the Mediterranean one. This is especially true in high latitude regions, where climatic change leads to more drastic temperature variation and stronger seasonality change than at low latitudes (Roots 1989; Hampe & Petit 2005; Wang *et al.* 2012). This differential effect between low and high latitudes can be observed in Asia, where Plio-Pleistocene faunas in the northern part of the continent form geographically less coherent clusters than the tropical and intertropical faunas of southern Asia (Fig. 4N–O). Finally, while bioregions produced with the Asian dataset show a strong latitudinal structuring from the beginning of the Miocene to the Plio-Pleistocene, bioregions in Europe show a shift from a longitudinal to a latitudinal structuring at the end of the Miocene. The reason for these different temporal dynamics between continents is that Europe is limited to latitudes above the northern tropics (23°N), while in Asia strong latitudinal differences between tropical/intertropical ecosystems and the open and arid ecosystems of higher latitudes, leads to strong latitudinal difference in composition.

CONCLUSION

At both the Old World and the continental scale, the Neogene and Pleistocene mammals split into five consecutive evolutionary faunas. Ranging from the sub-epoch

(i.e. Early, Middle and Late Miocene for EF1, EF2, EF3) to the stage level (e.g. Zanclean for EF4), they emerge and fade near synchronously on the three continents, considered independently as well as at the Old World scale. This synchronicity highlights the global scale of the abiotic and/or biotic phenomena driving these faunistic shifts, such as the Middle Miocene climatic transition following the MMCO, the intense global cooling peaking during the Late Miocene, the convergence of African and Eurasian plates leading to the closure of the Neotethys seaway as well as the shrinkage of the Paratethys. The temporal scale of the evolutionary fauna is well suited for the study of mammalian evolution and allows us to provide a spatial framework at multiple scales for the computation of a variety of biodiversity indices (e.g. α and β -diversity, taxonomic, functional, phylogenetic and morphologic diversity) for very diverse datasets.

The focus on a minimum number of bioregions reveals the major distribution patterns of mammalian biodiversity, such as the persistent subdivision between northern Eurasian and southern Asian faunas. The latter rather displays a very strong affinity with African faunas, a long-term similarity in taxonomic composition that results from important dispersal episodes between Africa and the Himalayan foothills, notably in the Early and Late Miocene. These inter-continental exchanges reflect global-scale events like the convergence of continental plates leading to landbridges formation (Early Miocene), or the emergence of continental-scale biomes after the Middle Miocene climatic transition and the final closure of epicontinental seaways. In addition, producing only two bioregions at the Old World scale reveals the constant similarity between northern Asian and European faunas and their longitudinal subdivision from the Middle Miocene to the Pleistocene. Furthermore, in Europe, a significant shift of distribution pattern from a longitudinal to a latitudinal structure can be observed between the Miocene and the Pliocene. This reorganization of faunas, probably related to the gradual establishment of the strong modern latitudinal temperature gradient (LTG) at temperate latitudes from the very end of the Miocene, is consistent with the spatio-temporal dynamics of large-herbivore hypsodonty patterns observed in Europe. The appearance/establishment of this strong LTG led to the simultaneous emergence of a large Mediterranean bioregion, resulting from the convergence of Spanish, French and Italian coastal faunas. This bioregion then extended eastward during the Pliocene and further north to reach its modern extent during the Pleistocene.

The consistent framework applicable to a wide range of Neogene and Pleistocene mammalian datasets we provide in this study could benefit from modern ecological

sciences (Gao & Kupfer 2018). The integration of, for example, observed (Pound *et al.* 2012) or modelled (Salzmann *et al.* 2008) environmental information such as biome, palaeobotanical and/or palaeoaltitudinal data could notably improve the definition of bioregion boundaries or help to address transition zone problems, by focusing on parameters other than taxonomic similarity and geographical continuity. Nevertheless, although these environmental parameters are not considered in our analyses of the spatio-temporal evolution of Old World mammal bioregions, they frequently track the major climatic and tectonic changes known in the fossil record. This shows that the existence of bioregions always primarily reflects, albeit indirectly, environmental filtering and its consequences on the evolution and distribution of taxa.

Acknowledgements. We acknowledge Gilles Escarguel (University of Lyon 1) for his help in the choice of clustering methods. Jean-Renaud Boisserie (University of Poitiers) for OGRE dataset.

This study was supported by the French Agence Nationale de la Recherche (ANR) project HADoC under reference ANR-17-CE31-0010.

Two anonymous referees commented on an earlier draft of this manuscript.

Author contributions. CG is the main contributor of this paper and conceived the idea. CG, AZ and GT designed the methodology and R codes. CG and AZ led the writing of the manuscript. FF, GR, PS, DB provided palaeogeographical, palaeobotanical and palaeoclimatical interpretations. OC, OO, AS, FG, WB, CB and GM provided faunistic interpretations and ideas. All authors contributed critically to the drafts and gave final approval for publication.

DATA ARCHIVING STATEMENT

Raw data extracted from NOW database and all derived data are included within .RData files and can be downloaded from GitHub (<https://github.com/Corentin-Gibert-Paleontology/A-coherent-biogeographic-framework-for-Old-World-Neogene-and-Pleistocene-mammals>) and the Dryad Digital Repository (<https://doi.org/10.5061/dryad.18931zczb>).

Editor. Arnaud Brayard

SUPPORTING INFORMATION

Additional Supporting Information can be found online (<https://doi.org/10.1111/pala.12594>):

Appendix S1. (Figs S1–S4) Sampling distributions, temporal extent and European land mammal ages, schematic diagram for poly-cohorts, flowcharts of the complete analytical process.

Appendix S2. (Figs S5–S72) Maps with Simpson index, factor maps and clustering tree for HCPC, UPGMA computed with Simpson and Raup & Crick indexes.

Appendix S3. Supplementary results (more detailed explanation of bioregions computation at continental scale).

Appendix S4. R script computed with Rmarkdown to help reproduce the analysis.

Table S1. Percentage of new and extinct taxa from one MNEQ to another.

Table S2. Poly-cohorts for all terrestrial mammal species (excluding bats) at the following scales: A, Old World; B, European; C, Asian; D, African.

Table S3. ANOSIM *p*-value and R statistic for bioregions at the following scales: A, Old World, illustrated in Figure 2; B, European, illustrated in Figure 3; C, Asian, illustrated in coloured hull on Figure 4; D, African illustrated in Figure 5.

Table S4. ANOSIM *p*-value and R statistic for clusters computed by HSPC at the following scales: A, Old World; B, European; C, Asian; D, African.

REFERENCES

- AGUSTÍ, J. and MOYÀ-SOLÀ, S. 1990. Mammal extinctions in the Vallesian (Upper Miocene). 425–432. In KAUFFMAN, E. G. and WALLIER, O. H. (eds). *Extinction events in earth history*. Lecture Notes in Earth Sciences, **30**. Springer.
- AGUSTÍ, J., CABRERA, L., GARCÉS, M., KRIJGSMAN, W., OMS, O. and PARÉS, J. M. 2001. A calibrated mammal scale for the Neogene of Western Europe. State of the art. *Earth-Science Reviews*, **52** (4), 247–260.
- AGUSTÍ, J., GARCÉS, M. and KRIJGSMAN, W. 2006. Evidence for African–Iberian exchanges during the Messinian in the Spanish mammalian record. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **238** (1–4), 5–14.
- ATAABADI, M., LIU, L., ERONEN, J. T., BERNOR, R. L. and FORTÉLIUS, M. 2013. Continental scale patterns in Neogene mammal community evolution and biogeography: a Europe–Asia perspective. 629–655. In WANG, X.-M., FLYNN, L. J. and FORTÉLIUS, M. (eds). *Fossil mammals of Asia*. Columbia University Press.
- AZZAROLI, A. and GUAZZONE, G. 1979. Terrestrial mammals and land connections in the Mediterranean before and during the Messinian. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **29**, 155–167.
- BANKS, W. E., D’ERRICO, F., PETERSON, A. T., KAGEYAMA, M. and COLOMBEAU, G. 2008. Reconstructing ecological niches and geographic distributions of caribou (*Rangifer tarandus*) and red deer (*Cervus elaphus*) during the Last Glacial Maximum. *Quaternary Science Reviews*, **27** (27–28), 2568–2575.
- BARBOLINI, N., WOUTERSEN, A., DUPONT-NIVET, G., SILVESTRO, D., TARDIF, D., COSTER, P. M. C., MEIJER, N., CHANG, C., ZHANG, H.-X., LICHT, A., RYDIN, C., KOUTSODENDRIS, A., HAN, F., ROHRMANN, A., LIU, X.-J., ZHANG, Y., DONNADIEU, U., FLUTEAU, F., LADANT, J.-B., LE HIR, G. and HOORN, C. 2020. Cenozoic evolution of the steppe-desert biome in Central Asia. *Science Advances*, **6** (41), eabb8227.
- BARNOSKY, A. D. 2001. Distinguishing the effects of the Red Queen and Court Jester on Miocene mammal evolution in the northern Rocky Mountains. *Journal of Vertebrate Paleontology*, **21** (1), 172–185.
- BARRY, J. C., JOHNSON, N. M., RAZA, S. M. and JACOBS, L. L. 1985. Neogene mammalian faunal change in southern Asia: correlations with climatic, tectonic, and eustatic events. *Geology*, **13** (9), 637–640.
- BARRY, J. C., MORGAN, M. E., FLYNN, L. J., PILBEAM, D., BEHRENSMEYER, A. K., RAZA, S. M., KHAN, I. A., BADGLEY, C., HICKS, J. and KELLEY, J. 2002. Faunal and environmental change in the late Miocene Siwaliks of northern Pakistan. *Paleobiology*, **28**, 1–71.
- BERNOR, R. L. 1978. The mammalian systematics, biostratigraphy and biochronology of Maragheh and its importance for understanding Late Miocene Hominoid zoogeography and evolution. PhD thesis. University of California, Los Angeles.
- BERNOR, R. L. 1979. The evolution of ‘Pontian’ mammal faunas: some zoogeographic, paleoecologic and chronostratigraphic considerations. 81–89. In *7th International Congress on Mediterranean Neogene, Athens, 1979*. Annales Géologiques Pays Helléniques, hors série, **1**.
- BERNOR, R. L. 1983. Geochronology and zoogeographic relationships of Miocene Hominoidea. In CIOCHON, R. L. and CORRUCINI, R. S. (eds). *New interpretations of ape and human ancestry*. Advances in Primatology. Springer.
- BERNOR, R. L. 1984. A zoogeographic theater and biochronology play: the time/biofacies phenomena of Eurasian and African Miocene mammal provinces. *Paléobiologie continentale*, **14** (2), 121–142.
- BERNOR, R. L., KOUFOS, G. D., WOODBURN, M. O. and FORTÉLIUS, M. 1996. The evolutionary history and biochronology of European and Southwest Asian Late Miocene and Pliocene hipparionine horses. 307–338. In BERNOR, R. L., FAHLBUSCH, V. and MITTMAN, H.-W. (eds). *The evolution of western Eurasian Neogene mammal faunas*. Columbia University Press.
- BIALIK, O. M., FRANK, M., BETZLER, C., ZAMMIT, R. and WALDMANN, N. D. 2019. Two-step closure of the Miocene Indian Ocean Gateway to the Mediterranean. *Scientific reports*, **9** (1), 1–10.
- BIBI, F. 2011. Mio-Pliocene faunal exchanges and African biogeography: the record of fossil bovids. *PLoS One*, **6** (2), e16688.
- BOBE, R. and ECK, G. G. 2001. Responses of African bovids to Pliocene climatic change. *Paleobiology*, **27** (S2), 1–47.
- TER BORGH, M., VASILIEV, I., STOICA, M., KNEŽEVIĆ, S., MATENCO, L., KRIJGSMAN, W. and CLOETINGH, S. 2013. The isolation of the Pannonian basin (Central Paratethys): new constraints from magnetostratigraphy and biostratigraphy. *Global & Planetary Change*, **103**, 99–118.
- DE BONIS, L., BOUVRAIN, G., GERAADS, D. and KOUFOS, G. 1992a. Diversity and paleoecology of Greek late Miocene mammalian faunas. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **91** (1–2), 99–121.

- DE BONIS, L., BOUVRAIN, G., GERAADS, D. and KOUFOS, G. 1992b. Multivariate study of late Cenozoic mammalian faunal compositions and paleoecology. *Paleontologia i Evolució*, **24** (25), 93–101.
- BRAYARD, A., ESCARGUEL, G. and BUCHER, H. 2007. The biogeography of Early Triassic ammonoid faunas: clusters, gradients, and networks. *Geobios*, **40** (6), 749–765.
- BROCKLEHURST, N. and FRÖBISCH, J. 2018. The definition of bioregions in palaeontological studies of diversity and biogeography affects interpretations: Palaeozoic tetrapods as a case study. *Frontiers in Earth Science*, **6**, 200.
- BRUCH, A. A., UTESCHER, T., OLIVARES, C. A., DOLÁKOVÁ, N., IVANOV, D. and MOSBRUGGER, V. 2004. Middle and Late Miocene spatial temperature patterns and gradients in Europe – Preliminary results based on palaeobotanical climate reconstructions. *Courier-Forschungsinstitut Senckenberg*, 15–28.
- BUTTON, D. J., LLOYD, G. T., EZCURRA, M. D. and BUTLER, R. J. 2017. Mass extinctions drove increased global faunal cosmopolitanism on the supercontinent Pangaea. *Nature Communications*, **8** (1), 1–8.
- CANO, A. R. G., CANTALAPIEDRA, J. L., ALVAREZ-SIERRA, M. A. and HERNÁNDEZ FERNÁNDEZ, M. 2014. A macroecological glance at the structure of late Miocene rodent assemblages from Southwest Europe. *Scientific Reports*, **4**, 6557.
- CANTALAPIEDRA, J. L., HERNÁNDEZ FERNÁNDEZ, M. and MORALES, J. 2014. The biogeographic history of ruminant faunas determines the phylogenetic structure of their assemblages at different scales. *Ecography*, **37** (1), 1–9.
- CANTALAPIEDRA, J. L., HERNÁNDEZ FERNÁNDEZ, M., AZANZA, B. and MORALES, J. 2015. Congruent phylogenetic and fossil signatures of mammalian diversification dynamics driven by Tertiary abiotic change. *Evolution*, **69** (11), 2941–2953.
- CASANOVAS-VILAR, I., MOYÁ-SOLÁ, S., AGUSTÍ, J. and KÖHLER, M. 2005. The geography of a faunal turnover: tracking the Vallesian Crisis. 247–300. In ELEWA, A. M. T. (ed.) *Migration of organisms*. Springer.
- CHAVE, J. 2013. The problem of pattern and scale in ecology: what have we learned in 20 years? *Ecology Letters*, **16**, 4–16.
- CHOROWICZ, J. 2005. The east African rift system. *Journal of African Earth Sciences*, **43** (1–3), 379–410.
- CONDAMINE, F. L., SPERLING, F. A., WAHLBERG, N., RASPLUS, J. Y. and KERGOAT, G. J. 2012. What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. *Ecology Letters*, **15** (3), 267–277.
- COSTEUR, L. and LEGENDRE, S. 2008. Spatial and temporal variation in European Neogene large mammals diversity. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **261** (1–2), 127–144.
- COSTEUR, L., LEGENDRE, S. and ESCARGUEL, G. 2004. European large mammals palaeobiogeography and biodiversity during the Neogene. Palaeogeographic and climatic impacts. *Revue de Paléobiologie*, **9**, 99–109.
- DEMENOCAL, B. P. 2004. African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth & Planetary Science Letters*, **220** (1–2), 3–24.
- DENK, T. 2016. Palaeoecological interpretation of the late Miocene landscapes and vegetation of northern Greece: a comment to Merceron et al., 2016 (*Geobios* 49, 135–146). *Geobios*, **49** (6), 423–431.
- DING, L., SPICER, R. A., YANG, J., XU, Q., CAI, F., LI, S., LAI, Q., WANG, H., SPICER, T. E. V., YUE, Y., SHUKLA, A., SRIVASTAVA, G., KHAN, M. A., BERA, S. and MEHROTRA, R. 2017. Quantifying the rise of the Himalaya orogen and implications for the South Asian monsoon. *Geology*, **45** (3), 215–218.
- DOMMERGUES, J. L., FARA, E. and MEISTER, C. 2009. Ammonite diversity and its palaeobiogeographical structure during the early Pliensbachian (Jurassic) in the western Tethys and adjacent areas. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **280** (1–2), 64–77.
- ENGLER, A. 1879. *Versuch einer Entwicklungsgeschichte der Pflanzenwelt: insbesondere der Florengebiete seit der Tertiärperiode*. Vol. 1. W. Engelmann.
- ERONEN, J. T. 2007. Locality coverage, metacommunities and chronofauna: concepts that connect paleobiology to modern population biology. *Vertebrata Palasiatica*, **45** (2), 137–144.
- ERONEN, J. T., ATAABADI, M. M., MICHEELS, A., KARME, A., BERNOR, R. L. and FORTELIUS, M. 2009. Distribution history and climatic controls of the Late Miocene Pikermian chronofauna. *Proceedings of the National Academy of Sciences*, **106** (29), 11867–11871.
- ERONEN, J. T., FORTELIUS, M., MICHEELS, A., PORTMANN, F. T., PUOLAMÄKI, K. and JANIS, C. M. 2012. Neogene aridification of the Northern Hemisphere. *Geology*, **40** (9), 823–826.
- ESCARGUEL, G. and LEGENDRE, S. 2006. New methods for analysing deep-time meta-community dynamics and their application to the Paleogene mammals from the Quercy and Limagne area (Massif Central, France). *Strata*, **13**, 245–273.
- ESCARGUEL, G., FARA, E., BRAYARD, A. and LEGENDRE, S. 2011. Biodiversity is not (and never has been) a bed of roses! *Comptes Rendus Biologies*, **334** (5–6), 351–359.
- FAHLBUSCH, V. 1989. European Neogene rodent assemblages in response to evolutionary, biogeographic, and ecological factors. 129–139. In BLACK, C. C. and DAWSON, M. R. (eds). *Papers on fossil rodents in honor of Albert Elmer Wood*. Science Series, **33**. Natural History Museum of Los Angeles County.
- FIGUEIRIDO, B., JANIS, C. M., PÉREZ-CLAROS, J. A., DE RENZI, M. and PALMQVIST, P. 2012. Cenozoic climate change influences mammalian evolutionary dynamics. *Proceedings of the National Academy of Sciences*, **109** (3), 722–727.
- FLUTEAU, F., RAMSTEIN, G. and BESSE, J. 1999. Simulating the evolution of the Asian and African monsoons during the past 30 Myr using an atmospheric general circulation model. *Journal of Geophysical Research: Atmospheres*, **104** (D10), 11995–12018.
- FLYNN, L. J., PILBEAM, D., JACOBS, L. L., BARRY, J. C., BEHRENSMEYER, A. K. and KAPPELMAN, J. W.

1990. The Siwaliks of Pakistan: time and faunas in a Miocene terrestrial setting. *The Journal of Geology*, **98** (4), 589–604.
- FORTELIUS, M., WERDELIN, L., ANDREWS, P., BERNOR, R. L., GENTRY, A., HUMPHREY, L., MITTMANN, H.-W. and VIRANTA, S. 1996. Provinciality, diversity, turnover, and paleoecology in land mammal faunas of the later Miocene of western Eurasia. 414–418. In BERNOR, R. L., FAHLBUSCH, V. and MITTMANN, H.-V. (eds). *The evolution of western Eurasian Neogene mammal faunas*. Columbia University Press.
- FORTELIUS, M., ERONEN, J., LIU, L. P., PUSHKINA, D., TESAKOV, A., VISLOBOKOVA, I. and ZHANG, Z. Q. 2003. Continental-scale hypsodonty patterns, climatic paleobiogeography, and dispersal of Eurasian Neogene large mammal herbivores. *Deinsea*, **10** (1), 1–12.
- FORTELIUS, M., ERONEN, J., LIU, L., PUSHKINA, D., TESAKOV, A., VISLOBOKOVA, I. and ZHANG, Z. 2006. Late Miocene and Pliocene large land mammals and climatic changes in Eurasia. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **238** (1–4), 219–227.
- FORTELIUS, M., ERONEN, J. T., KAYA, F., TANG, H., RAIA, P. and PUOLAMÄKI, K. 2014. Evolution of Neogene mammals in Eurasia: environmental forcing and biotic interactions. *Annual Review of Earth & Planetary Sciences*, **42**, 579–604.
- GAO, P. and KUPFER, J. A. 2018. Capitalizing on a wealth of spatial information: improving biogeographic regionalization through the use of spatial clustering. *Applied Geography*, **99**, 98–108.
- GARCÍA-ALIX, A., MINWER-BARAKAT, R., MARTÍN SUÁREZ, E., FREUDENTHAL, M. and MARTÍN, J. M. 2008. Late Miocene–Early Pliocene climatic evolution of the Granada Basin (southern Spain) deduced from the paleoecology of the micromammal associations. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **265** (3–4), 214–225.
- GARCÍA-ALIX, A., MINWER-BARAKAT, R., MARTÍN SUÁREZ, E., FREUDENTHAL, M., AGUIRRE, J. and KAYA, F. 2016. Updating the Europe–Africa small mammal exchange during the late Messinian. *Journal of Biogeography*, **43** (7), 1336–1348.
- GÉBELIN, A., MULCH, A., TEYSSIER, C., JESSUP, M. J., LAW, R. D. and BRUNEL, M. 2013. The Miocene elevation of Mount Everest. *Geology*, **41** (7), 799–802.
- GERAADS, D. 1998. Biogeography of circum-Mediterranean Miocene–Pliocene rodents: a revision using factor analysis and parsimony analysis of endemism. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **137** (3–4), 273–288.
- GERAADS, D. 2010. Biogeographic relationships of Pliocene and Pleistocene North-western African mammals. *Quaternary International*, **212** (2), 159–168.
- GIBERT, L., SCOTT, G. R., MONTOYA, P., RUIZ-SÁNCHEZ, F. J., MORALES, J., LUQUE, L., ABELLA, J. and LERÍA, M. 2013. Evidence for an African-Iberian mammal dispersal during the pre-evaporitic Messinian. *Geology*, **41** (6), 691–694.
- GIBERT BRET, C. 2022. Supplementary materials, R scripts and RData files. Dryad Digital Repository. <https://doi.org/10.5061/dryad.18931zczb>
- GRIFFIN, D. L. 2002. Aridity and humidity: two aspects of the late Miocene climate of North Africa and the Mediterranean. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **182** (1–2), 65–91.
- GROHÉ, C., DE BONIS, L., CHAIMANEE, Y., CHAVASSEAU, O., RUGBUMRUNG, M., YAMEE, C., JAEGER, J. J. 2020. The late middle Miocene Mae Moh Basin of northern Thailand: the richest Neogene assemblage of Carnivora from Southeast Asia and a paleobiogeographic analysis of Miocene Asian carnivorans. *American Museum Novitates*, **2020** (3952), 1–57.
- GRUBB, P., SANDROCK, O., KULLMER, O., KAISER, T. M. and SCHRENK, F. 1999. Relationships between eastern and southern African mammal faunas. 253–267. In BROMAGE, G. T. and SCHRENK, F. (eds). *African biogeography, climate change, and human evolution*. Oxford University Press.
- HAMPE, A. and PETIT, R. J. 2005. Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, **8**, 461–467.
- HARZHAUSER, M., KROH, A., MANDIC, O., PILLER, W. E., GÖHLICH, U., REUTER, M. and BERNING, B. 2007. Biogeographic responses to geodynamics: a key study all around the Oligo-Miocene Tethyan Seaway. *Zoologischer Anzeiger*, **246** (4), 241–256.
- HE, J., KREFT, H., GAO, E., WANG, Z. and JIANG, H. 2017. Patterns and drivers of zoogeographical regions of terrestrial vertebrates in China. *Journal of Biogeography*, **44** (5), 1172–1184.
- HE, J., KREFT, H., LIN, S., XU, Y. and JIANG, H. 2018. Cenozoic evolution of beta diversity and a Pleistocene emergence for modern mammal faunas in China. *Global Ecology & Biogeography*, **27** (11), 1326–1338.
- HEIKINHEIMO, H., FORTELIUS, M., ERONEN, J. and MANNILA, H. 2007. Biogeography of European land mammals shows environmentally distinct and spatially coherent clusters. *Journal of Biogeography*, **34** (6), 1053–1064.
- HERBERT, T. D., LAWRENCE, K. T., TZANOVA, A., PETERSON, L. C., CABALLERO-GILL, R. and KELLY, C. S. 2016. Late Miocene global cooling and the rise of modern ecosystems. *Nature Geoscience*, **9** (11), 843–847.
- HUSSON, F., JOSSE, J., PAGÈS, J. 2010. Analyse de données avec R - Complémentarité des méthodes d'analyse factorielle et de classification. 42èmes Journées de Statistique, Marseille, France.
- INGALLS, M., ROWLEY, D. B., CURRIE, B. S. and COLMAN, A. S. 2020. Reconsidering the uplift history and peneplanation of the northern Lhasa terrane, Tibet. *American Journal of Science*, **320** (6), 479–532.
- JANIS, C. M. 1993. Tertiary mammal evolution in the context of changing climates, vegetation, and tectonic events. *Annual Review of Ecology & Systematics*, **24** (1), 467–500.
- JAEGER, J. J., COIFFAIT, B., TONG, H. and DENYS, C. 1987. Rodent extinctions following Messinian faunal exchanges between Western Europe and Northern Africa. *Mémoires de la Société Géologique de France*, **NS**, **150**, 153–158.
- KAHLKE, R.-D., GARCÍA, N., KOSTOPOULOS, D. S., LACOMBAT, F., LISTER, A. M., MAZZA, P. P. A., SPASSOV, N. and TITOV, V. V. 2011. Western Palaearctic palaeoenvironmental conditions during the Early and early

- Middle Pleistocene inferred from large mammal communities, and implications for hominin dispersal in Europe. *Quaternary Science Reviews*, **30** (11–12), 1368–1395.
- KAYA, F., BIBI, F., ŽLIOBAITĖ, I., ERONEN, J. T., HUI, T. and FORTELIUS, M. 2018. The rise and fall of the Old World savannah fauna and the origins of the African savannah biome. *Nature Ecology & Evolution*, **2** (2), 241–246.
- KOSTOPOULOS, D. S. 2009. The Pikermian Event: temporal and spatial resolution of the Turolian large mammal fauna in SE Europe. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **274** (1–2), 82–95.
- KOUFOS, G. D. and KONIDARIS, G. E. 2011. Late Miocene carnivores of the Greco-Iranian Province: composition, guild structure and palaeoecology. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **305** (1–4), 215–226.
- KREFT, H. and JETZ, W. 2010. A framework for delineating biogeographical regions based on species distributions. *Journal of Biogeography*, **37** (11), 2029–2053.
- LARSSON, L. M., DYBKJÆR, K., RASMUSSEN, E. S., PIASECKI, S., UTESCHER, T. and VAJDA, V. 2011. Miocene climate evolution of northern Europe: a palynological investigation from Denmark. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **309** (3–4), 161–175.
- LÊ, S., JOSSE, J. and HUSSON, F. 2008. FactoMineR: an R package for multivariate analysis. *Journal of Statistical Software*, **25** (1), 1–18.
- LEGENDRE, S. 1987. Les immigrations de la «Grande Coupure» sont-elles contemporaines en Europe occidentale? 141–148. In SCHMIDT-KITTLER, N. (ed.) *International Symposium on Mammalian Biostratigraphy and Paleocology of the European Paleogene* Mainz. F. Pfeil.
- LEVIN, S. A. 1992. The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. *Ecology*, **73** (6), 1943–1967.
- LI, P., ZHANG, C., KELLEY, J., DENG, C., JI, X., JABLONSKI, N. G., WU, H., FU, Y., GUO, Z. and ZHU, R. 2020. Late Miocene climate cooling contributed to the disappearance of hominoids in Yunnan region, southwestern China. *Geophysical Research Letters*, **47** (11), e2020GL087741.
- LISIECKI, L. E. and RAYMO, M. E. 2007. Plio-Pleistocene climate evolution: trends and transitions in glacial cycle dynamics. *Quaternary Science Reviews*, **26** (1–2), 56–69.
- LISTER, A. M. and STUART, A. J. 2008. The impact of climate change on large mammal distribution and extinction: evidence from the last glacial/interglacial transition. *Comptes Rendus Geoscience*, **340** (9–10), 615–620.
- LONGHURST, A. R. 2010. *Ecological geography of the sea*. Elsevier.
- MADERN, P. A. and VAN DEN HOEK OSTENDE, L. W. 2015. Going south: latitudinal change in mammalian biodiversity in Miocene Eurasia. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **424**, 123–131.
- MAI, D. H. 1995. *Tertiäre Vegetationsgeschichte Europas: Methoden und Ergebnisse*. Spektrum Akademischer Verlag.
- MANNION, P. D., UPCHURCH, P., BENSON, R. B. and GOSWAMI, A. 2014. The latitudinal biodiversity gradient through deep time. *Trends in Ecology & Evolution*, **29** (1), 42–50.
- MARCOT, J. D., FOX, D. L. and NIEBUHR, S. R. 2016. Late Cenozoic onset of the latitudinal diversity gradient of North American mammals. *Proceedings of the National Academy of Sciences*, **113** (26), 7189–7194.
- MARIDET, O., ESCARGUEL, G., COSTEUR, L., MEIN, P., HUGUENEY, M. and LEGENDRE, S. 2007. Small mammal (rodents and lagomorphs) European biogeography from the Late Oligocene to the mid Pliocene. *Global Ecology & Biogeography*, **16** (4), 529–544.
- MARIDET, O. and COSTEUR, L. 2010. Diversity trends in Neogene European ungulates and rodents: large-scale comparisons and perspectives. *Naturwissenschaften*, **97** (2), 161–172.
- MCCOY, E. D. and CONNOR, E. F. 1980. Latitudinal gradients in the species diversity of North American mammals. *Evolution*, **34** (1), 193–203.
- MCQUARRIE, N. and VAN HINSBERGEN, D. J. J. 2013. Retrodeforming the Arabia-Eurasia collision zone: age of collision versus magnitude of continental subduction. *Geology*, **41** (3), 315–318.
- MEIN, P. 1976. Biozonation du Néogène méditerranéen à partir des mammifères. 77–81. In *Proceedings of the VI Congress of the Regional Committee on Mediterranean Neogene Stratigraphy, Bratislava, September 4–7, 1975*.
- MOSBRUGGER, V., UTESCHER, T. and DILCHER, D. L. 2005. Cenozoic continental climatic evolution of Central Europe. *Proceedings of the National Academy of Sciences*, **102** (42), 14964–14969.
- MÜLLER, R. D., CANNON, J., QIN, X., WATSON, R. J., GURNIS, M., WILLIAMS, S., PFAFFELMOSE, T., SETON, M., RUSSELL, S. H. J. and ZAHIROVIC, S. 2018. GPlates: building a virtual Earth through deep time. *Geochemistry, Geophysics, Geosystems*, **19** (7), 2243–2261.
- OLSON, E. C. 1952. The evolution of a Permian vertebrate chronofauna. *Evolution*, 181–196.
- PATTERSON, D. B., FAITH, J. T., BOBE, R. and WOOD, B. 2014. Regional diversity patterns in African bovids, hyaenids, and felids during the past 3 million years: the role of taphonomic bias and implications for the evolution of *Paranthropus*. *Quaternary Science Reviews*, **96**, 9–22.
- PELÁEZ-CAMPOMANES, P. and VAN DER MEULEN, A. J. 2009. Diversity of mammals in the Neogene of Europe: comparing data quality of large and small mammals in the NOW database. *Hellenic Journal of Geosciences*, **44**, 105–115.
- POUND, M. J., HAYWOOD, A. M., SALZMANN, U. and RIDING, J. B. 2012. Global vegetation dynamics and latitudinal temperature gradients during the Mid to Late Miocene (15.97–5.33 Ma). *Earth-Science Reviews*, **112** (1–2), 1–22.
- PIROUZ, M., AVOUAC, J. P., HASSANZADEH, J., KIRSCHVINK, J. L. and BAHROUDI, A. 2017. Early Neogene foreland of the Zagros, implications for the initial closure of the Neo-Tethys and kinematics of crustal shortening. *Earth & Planetary Science Letters*, **477**, 168–182.
- POPOV, S. V., SHCHERBA, I. G., ILYINA, L. B., NEVESS-KAYA, L. A., PARAMONOVA, N. P., KHONDKARIAN, S. O. and MAGYAR, I. 2006. Late Miocene to Pliocene palaeogeography of the Paratethys and its relation to the Mediterranean. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **238** (1–4), 91–106.

- RAIA, P., PIRAS, P. and KOTSAKIS, T. 2005. Turnover pulse or Red Queen? Evidence from the large mammal communities during the Plio-Pleistocene of Italy. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **221** (3–4), 293–312.
- RAIA, P., PIRAS, P. and KOTSAKIS, T. 2006. Detection of Plio-Quaternary large mammal communities of Italy. An integration of fossil faunas biochronology and similarity. *Quaternary Science Reviews*, **25** (7–8), 846–854.
- RAMSTEIN, G., FLUTEAU, F., BESSE, J. and JOUS-
SAUME, S. 1997. Effect of orogeny, plate motion and land-sea distribution on Eurasian climate change over the past 30 million years. *Nature*, **386** (6627), 788–795.
- RAUP, D. M. and CRICK, R. E. 1979. Measurement of faunal similarity in paleontology. *Journal of Paleontology*, 1213–1227.
- REYDONDEAU, G., LONGHURST, A., MARTINEZ, E., BEAUGRAND, G., ANTOINE, D. and MAURY, O. 2013. Dynamic biogeochemical provinces in the global ocean. *Global Biogeochemical Cycles*, **27** (4), 1046–1058.
- ROOTS, E. F. 1989. Climate change: high-latitude regions. *Climatic Change*, **15** (1–2), 223–253.
- RÖGL, F. 1999. Mediterranean and Paratethys. Facts and hypotheses of an Oligocene to Miocene paleogeography (short overview). *Geologica Carpathica*, **50** (4), 339–349.
- ROBERTS, E. M., STEVENS, N. J., O'CONNOR, P. M., DIRKS, P. H. G. M., GOTTFRIED, M. D., CLYDE, W. C., ARMSTRONG, R. A., KEMP, A. I. S. and HEMMING, S. 2012. Initiation of the western branch of the East African Rift coeval with the eastern branch. *Nature Geoscience*, **5** (4), 289–294.
- SALZMANN, U., HAYWOOD, A. M., LUNT, D. J., VALDES, P. J. and HILL, D. J. 2008. A new global biome reconstruction and data-model comparison for the middle Pliocene. *Global Ecology & Biogeography*, **17** (3), 432–447.
- SCHERLER, L., MENNECART, B., HIARD, F. and BECKER, D. 2013. Evolutionary history of hoofed mammals during the Oligocene–Miocene transition in Western Europe. *Swiss Journal of Geosciences*, **106** (2), 349–369.
- SCOTese, C. R. 2016. Tutorial: PALEOMAP PaleoAtlas for GPlates and the PaleoData Plotter Program. <https://doi.org/10.13140/RG.2.2.34367.00166>
- SEN, S. 2013. Dispersal of African mammals in Eurasia during the Cenozoic: ways and whys. *Geobios*, **46** (1–2), 159–172.
- SEPKOSKI, J. J. Jr 1981. A factor analytic description of the Phanerozoic marine fossil record. *Paleobiology*, 36–53.
- SIMPSON, G. G. 1944. *Tempo and mode in evolution*. Columbia Biological Series, **15**. Columbia University Press.
- SNIDERMAN, J. K., WOODHEAD, J. D., HELLSTROM, J., JORDAN, G. J., DRYSDALE, R. N., TYLER, J. J. and PORCH, N. 2016. Pliocene reversal of late Neogene aridification. *Proceedings of the National Academy of Sciences*, **113** (8), 1999–2004.
- SMITH, J. R., HENDERSHOT, J. N., NOVA, N. and DAILY, G. C. 2020. The biogeography of ecoregions: descriptive power across regions and taxa. *Journal of Biogeography*, **47** (7), 1413–1426.
- SOMMER, R. S. and NADACHOWSKI, A. 2006. Glacial refugia of mammals in Europe: evidence from fossil records. *Mammal Review*, **36** (4), 251–265.
- STRÖMBERG, C. A. 2011. Evolution of grasses and grassland ecosystems. *Annual Review of Earth & Planetary Sciences*, **39**, 517–544.
- TASSY, P., SEN, S., JAEGER, J. J., MAZIN, J. M. and DALFES, N. 1989. Une sous-espèce nouvelle de *Choerolophodon pentelici* (Proboscidea, Mammalia) à Esme Akçaköy, Miocène supérieur d'Anatolie occidentale. *Comptes rendus de l'Académie des sciences, Paris, Série 2*, **309**, 2143–2146.
- TEDFORD, R. H. 1987. Faunal succession and biochronology of the Arikareean through Hemphillian interval (Late Oligocene through earliest Pliocene epochs) in North America. 153–210. In WOODBURN, M. O. (ed.) *Cenozoic mammals of North America*. University of California Press.
- TOBIEN, H. 1967. Subdivision of Pontian mammalian faunas. *Giornale di Geologia*, **35** (2), 1–5.
- TZANOVA, A., HERBERT, T. D. and PETERSON, L. 2015. Cooling Mediterranean sea surface temperatures during the Late Miocene provide a climate context for evolutionary transitions in Africa and Eurasia. *Earth & Planetary Science Letters*, **419**, 71–80.
- UDVARDY, M. D. 1975. *A classification of the biogeographical provinces of the world*, Vol. 8. International Union for Conservation of Nature and Natural Resources, Morges.
- UTESCHER, T., BRUCH, A. A., MICHEELS, A., MOSBRUGGER, V. and POPOVA, S. 2011. Cenozoic climate gradients in Eurasia—a palaeo-perspective on future climate change? *Palaeogeography, Palaeoclimatology, Palaeoecology*, **304** (3–4), 351–358.
- VAN DER MADE, J. 1999. Biogeography and stratigraphy of the Mio-Pleistocene mammals of Sardinia and the description of some fossils. *Deinsea*, **7** (1), 337–360.
- VAN DER MEULEN, A. J. and DAAMS, R. 1992. Evolution of Early-Middle Miocene rodent faunas in relation to long-term palaeoenvironmental changes. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **93** (3–4), 227–253.
- WALLACE, A. R. 1876. *The geographical distribution of animals*. Harper & Brothers Publishers, New York.
- WANG, S., LU, H., HAN, J., CHU, G., LIU, J. and NEGENDANK, J. F. 2012. Palaeovegetation and palaeoclimate in low-latitude southern China during the Last Glacial Maximum. *Quaternary International*, **248**, 79–85.
- WANG, X., WANG, Y., LI, Q., TSENG, Z. J., TAKEUCHI, G. T., DENG, T., XIE, G., CHANG, M.-M. AND WANG, N. 2015. Cenozoic vertebrate evolution and paleoenvironment in Tibetan Plateau: progress and prospects. *Gondwana Research*, **27** (4), 1335–1354.
- WARNY, S. A., BART, P. J. and SUC, J. P. 2003. Timing and progression of climatic, tectonic and glacioeustatic influences on the Messinian Salinity Crisis. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **202** (1–2), 59–66.
- ZACHOS, J., PAGANI, M., SLOAN, L., THOMAS, E. and BILLUPS, K. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, **292** (5517), 686–693.
- ZHANG, Z., RAMSTEIN, G., SCHUSTER, M., LI, C., CONTOUX, C. and YAN, Q. 2014. Aridification of the Sahara desert caused by Tethys Sea shrinkage during the Late Miocene. *Nature*, **513** (7518), 401–404.