

Plant fossils reveal major biomes occupied by the late Miocene Old-World Pikermian fauna

Thomas Denk^{1*}, Constantin M. Zohner², Guido W. Grimm⁴ and Susanne S. Renner^{3*}

Reconstruction of palaeobiomes, ancient communities that exhibit a physiognomic and functional structure controlled by their environment, depends on proxies from different disciplines. Based on terrestrial mammal fossils, the late Miocene vegetation from China to the eastern Mediterranean and East Africa has been reconstructed as a single cohesive biome with increasingly arid conditions, with modern African savannahs the surviving remnant. Here, we test this reconstruction using plant fossils spanning 14–4 million years ago from sites in Greece, Bulgaria, Turkey, the Tian Shan Mountains and Baode County in China, and East Africa. The western Eurasian sites had a continuous forest cover of deciduous or evergreen angiosperms and gymnosperms, with 15% of 1,602 fossil occurrences representing conifers, which were present at all but one of the sites. Raup–Crick analyses reveal high floristic similarity between coeval eastern Mediterranean and Chinese sites, and low similarity between Eurasian and African sites. The disagreement between plant-based reconstructions, which imply that late Miocene western Eurasia was covered by evergreen needleleaf forests and mixed forests, and mammal-based reconstructions, which imply a savannah biome, throws into doubt the approach of inferring Miocene precipitation and open savannah habitats solely from mammalian dental traits. Organismal communities are constantly changing in their species composition, and neither animal nor plant traits by themselves are sufficient to infer entire ancient biomes. The plant fossil record, however, unambiguously rejects the existence of a cohesive savannah biome from eastern Asia to northeast Africa.

Late Miocene Eurasia was inhabited by antelopes (including Hippotragini), giraffes, horses (extinct genus *Hipparion* and other Equinae), rhinoceroses, elephants and mastodons, pigs, hyenas, cats and many other mammals^{1–5}. Among the best-known localities with this type of fauna is a site near the village of Pikermi not far from Athens, Greece, which has led to the name ‘Pikermian’ palaeobiome for the late Miocene chronofauna that extended from the Balkans to Afghanistan^{2,4,6–8}. Especially important for the reconstruction of Pikermian biomes have been teeth, which are thought to offer a quasi-quantitative proxy for environmental aridity because height and enamel abrasion reflect food hardness and water content⁹. Hypsodonty has been quantified by mean hypsodonty values based on a site’s average height-to-length ratio of the second (upper or lower) molar of Artiodactyla, Perissodactyla, Proboscidea and Primates. Mean hypsodonty scores in combination with genus-level faunal comparisons between late Miocene faunas from a region spanning from Greece south to Lower Nawata in Kenya, Africa, and east to Baode in Shanxi Province (near Inner Mongolia), China, led to the reconstruction of a single cohesive Old-World savannah palaeobiome of which the modern African savannah fauna is the surviving branch¹⁰. Earlier palaeozoologists had instead reconstructed the Eurasian Pikermian biome as sclerophyllous Mediterranean-type woodland, with trees, chaparral and herbaceous undergrowth, but no continuous grass layer^{5–7,11}.

Here we use the plant fossil record to test the hypothesis of a single cohesive biome, that is, a predominant vegetation type controlled by the environment and climate that extended from Mongolia to East Africa and at its peak covered much of the Old World¹⁰. Regarding the relevant time range, we follow Kaya et al.¹⁰ who consider the Pikermian as the interval spanning the latest middle Miocene to the early Pliocene (12–5 Ma, million years ago) divided into the

eight European Neogene fossil mammal (MN) biochronologic units MN7 + 8, MN9, MN10, MN11, MN12, MN13 and MN14. For comparison, we included 4 Ma-old Pliocene occurrences from northeast Africa. Our study is based on fossil woods, fruits, leaves, spores and plant silica bodies, each taxon’s biogeographic affinity, and climate-related vegetation reconstruction based on the modern preferences of the lineage to which a fossil belongs. We apply Raup–Crick similarity analysis^{12,13} to facilitate direct comparison with genus-level faunistic studies that have used this approach^{4,10}.

Results and discussion

The plant fossil record of Pikermian sites. We analysed macrofossils, pollen and spores from 1,602 fossil occurrences from 45 sites in Greece, Bulgaria and Turkey, which provided us with 39 floras (Tables 1 and 2; Supplementary Tables 1 and 2) and 1,480 records for the Raup–Crick similarity analysis^{12,13}; 88 macrofossil records from six sites in northeast Africa; and 154 microfossil records from eight sites in China, covering European Neogene mammal units MN7 + 8 to MN14, Serravallian (middle Miocene) to mid-Zanclean (early Pliocene), 13.8–4.2 Ma^{14,15}. We also inferred palaeo-environments from the modern ecological requirements of the relevant genera and from pollen signatures expressed in abundances of particular taxa. A summary of the vegetation types inferred for western Eurasia (Greece to East Anatolia) is given in Fig. 1 and Tables 1 and 2 (also Supplementary Tables 1 and 2), with a cartoon illustrating abundant plant genera and formations in Supplementary Fig. 1. The similarity of these sites to each other and to coeval sites in East Africa and northeast Asia through time is quantified in Figs. 2 and 3, Supplementary Tables 2 and 3, and Supplementary Fig. 3. No more than five genera were shared between East African and Eurasian or Asian assemblages between MN7 + 8 to 14 (13.8–4.2 Ma), although

¹Department of Palaeobiology, Swedish Museum of Natural History, Stockholm, Sweden. ²Institute of Integrative Biology, ETH Zurich (Swiss Federal Institute of Technology), Zurich, Switzerland. ³Systematic Botany and Mycology, University of Munich (LMU), Munich, Germany.

⁴Unaffiliated: gwgrimm@gmx.de. *e-mail: thomas.denk@nrm.se; renner@lmu.de

Table 1 | Vegetation of the western Eurasian palynoflora sites

Site	Age	Palaeo-environmental characterization							
		MF, DF	ENF	ENF gen.	S	Poac.	Amaranth.	Artemisia	Asteraceae
GR01	MN7-8	MF	-30	3		n.i.	n.i.	n.i.	n.i.
GR01	MN7-8	MF	-40	5		0-10	-5 to <40	0 to <5	0 to <5
TR01	MN7-8	MF	-25	5		-5	<5	<5	0 (-20)
TR02	MN7-8	MF	-30	4		0	n.i.	0	n.i.
TR03	MN9	DF?	n.i.	1		n.i.	n.i.	n.i.	n.i.
TR04	MN9 +10	MF	90	4		0	0	0	<5
GR04	MN9-10	MF	20-30	5		-5 (10)	n.i.	n.i.	n.i.
GR05	MN10	MF	-20	6		0-10	0-15	<1	<5
TR05	MN10		-5	1	S	-5	<5	<1	-85
GR07	MN10-11	MF	-20	7	SF?	<5	5-15	<5	0
TR06	MN10-11		-30	2	S	0-10	20 to >40	10 to >20	10-30
BG01	MN10-12	MF	10-40	6		-5 (20)	<<5-10	<1	<1-5
GR08	MN11	MF	n.i.	7		n.i.	n.i.	n.i.	n.i.
TR07	MN11 +12	MF	90	4		-5	<15	0	<5
GR04	MN11-12	MF	10 to <40	4-5		-5 (35)	n.i.	n.i.	n.i.
TR08	MN11-12		2	1	S	0-2	15-20	0	-10
GR09	MN12	DF	5-20	7		>8	<2	0-1	<1-5
TR09	MN12	DF	-40	4	SF?	0-15	5-30	0-5	<<10
TR10	MN12		-35	3	S	n.i.	n.i.	<5	n.i.
TR11	MN12		<20	2	S	n.i.	n.i.	<5	n.i.
TR12	MN12-13		0 to <10	3	S	5-40	20 to >60	<1	<<5 to <15
GR11	MN13	MF	50-70	7		<10	<10	<1	0
GR12	MN13	MF	-20	6	SF?	<10	-15	<1	0
TR13	MN13		<5(-20)	4	GS	-35	<15	<1	<<10
TR14	MN13	DF	-55	6		-5	<10	<10	0
TR15	MN13		90	5	SF?	5-10	<5	<5	<10
TR16	MN13-14		30-40	5	S/SF	10 to <20	-20	<1	<<10
TR17	MN14	DF	>20 to >70	5	SF?	<5	<5-15	0	20-70
TR18	MN14		90	5	SF?	5-10	<5	<5	<10
TR19	MN14		5 to >10	4	?	20 (40)	<5 to <15	0	<5

MF, DF, mesic or dry forest; ENF, percentage of conifers in pollen diagrams representing the evergreen needleleaf forest biome of ref. ¹⁶; ENF gen., number of evergreen conifer genera in pollen assemblages; S, steppe elements—dominating or codominating the pollen diagrams (S, steppe; SF, steppe forest with Pinaceae; GS, grass steppe); Poac., percentage of Poaceae in pollen diagrams; Amaranth., percentage of Amaranthaceae in pollen diagrams; Artemisia, percentage of *Artemisia* in pollen diagrams; Asteraceae, percentage of Asteraceae in pollen diagrams; n.i., not indicated. Ranges of pollen percentages indicate fluctuations in a palynological section; percentage values in brackets indicate single peaks in abundance.

Africa became slightly more similar to the other regions by 6–4 Ma (Figs. 2 and 3; Supplementary Table 3).

We find that western Eurasia was continuously covered by the biomes ENF and mixed forests sensu Woodward et al.¹⁶ or temperate broadleaf and mixed forests (TBMF) sensu Olson et al.¹⁷ (Fig. 1, black colour in the pie diagrams: montane conifer forest and mesic forest). The importance of needle trees is underscored by 15% of the 1,602 fossil records belonging to 17 conifer genera endemic to the Northern Hemisphere and by subsets of these genera found at >90% of the western Eurasian sites. Similarly, 17% of the Chinese records belong to conifers, of which *Pinus* is found at all nine investigated East Asian sites, and *Picea* (seven sites) and *Abies* (five sites) at more than half of them covering all three main time slices (14–9 Ma, 9–5 Ma, 5–4 Ma; Supplementary Table 3).

Forests with (tropical) subtropical tree taxa (Taxodiaceae, Engelhardioideae, Lauraceae, Sapotaceae) and taxa extending into the temperate zone (*Cathaya*, *Cedrus*, Fagaceae) characterized the Greek islands Chios, Crete, Gavdos and Samos at circa 13–9 Ma,

and similar forests grew at the sites Muğla-Çatakbağyaka and Yeni Eskihsar (MN7 + 8, 13.8–11.2 Ma; Supplementary Tables 1 and 2) in Central Anatolia. Younger (MN9 and MN10, 11.2–8.7 Ma) floras in Central Anatolia are less diverse, but still contain abundant mesic forest elements. The pollen assemblage of Sivas-Karaözü is dominated by Asteraceae, with Fagaceae, Betulaceae and Pinaceae occurring more sparsely. Forest-dominated vegetation under a humid warm temperate climate (Köppen Cf/Cw climates; Table 2) has also been documented for more northern Pliocene sites (MN7 + 8 to 10, 13.8–8.7 Ma), including Bulgaria, the Czech Republic, France, Hungary, Italy, Poland, Romania and the Vienna Basin^{18–21}. As in older (MN6, 14.8–13.8 Ma) floras of the eastern Mediterranean (for example, ref. ²²), the herbs probably grew along streams, lake shores and rocky outcrops, or were part of the understorey of these forests. Coeval sites in the Ukrainian Plain (Volhynian and Bessarabian assemblages, MN7 + 8, 13.8–11.2 Ma) had mixed mesophytic forests and, north of the modern Black Sea, steppe vegetation dominated by herbaceous Amaranthaceae–Chenopodiaceae^{20,23}.

Table 2 | Environments and climates inferred from the plant fossil record for the Pikermian, Baodean and Nawatian chronofaunas, that is, the Old-World savannah palaeobiome of Kaya et al.¹⁰

Köppen climate type ^a	Terrestrial biomes ^b represented by fossil plant assemblages		
	Western Eurasia	Northeast Asia	East Africa
Cf, Cw, Df, Dw	Temperate broadleaf and mixed forest	Temperate broadleaf and mixed forest	
Cf, Df, Dw	Temperate conifer forest	Temperate conifer forest	
Cf, Df (western Eurasia), BSk (northeast Asia)	Temperate grasslands, savannahs and shrublands	Temperate grasslands, savannahs and shrublands	
A			Tropical and subtropical moist broadleaf forest
BSh			Tropical and subtropical grasslands, savannahs and shrublands
Cwb, Cfb			Montane grasslands and shrublands

Temperate forest biomes are well-represented in western Eurasia (see Fig. 1), while temperate grasslands/steppes predominate in northeast Asia. ^aRef. 69, ^bRef. 17.

Between 9 and 6 Ma (MN10+11 to MN13), there were marked environmental changes. The first *Artemisia* steppe appeared in Central Anatolia (MN10+11, circa 9 Ma; Cappadocia, Turkey), with abundant Amaranthaceae (Chenopodioideae) and Asteraceae, and rarer Poaceae (grassland biome of Woodward et al.¹⁶; temperate grasslands, savannahs and shrublands biome of Olson et al.¹⁷). Western Turkey was covered by a steppe forest with typical Northern Hemisphere woody genera, such as *Acer*, *Carpinus*, *Carya*, *Castanea*, *Cathaya*, *Juglans*, *Liquidambar*, *Parrotia*, *Pterocarya*, *Quercus*, *Tsuga*, *Ulmus* and *Zelkova*. We find no evidence of the Mediterranean-type (Köppen Cs climate; Table 2) sclerophyllous stressed in some Pikermian reconstructions^{6,24}. Instead, the sclerophyllous species of *Quercus* found in Samos, Crete and southwest Anatolia show the morphology of Himalayan members of sect. *Ilex* and indicate (summer-)humid climate (Köppen Cf/Cw climates^{25,26}). The vegetation of Crete, Samos, the Greek mainland and South Bulgaria continued to be dominated by mesic forests with Fagaceae, Pinaceae and Lauraceae, and this remained the case for mainland Greece and Italy into the Pliocene²⁷. The west–east differentiation recorded here roughly corresponds to the change from ‘forest faunas’ to ‘steppe faunas’ inferred by Kurtén²⁸.

The angiosperms and gymnosperms at the northeast Asian sites from 7.2–5.3 Ma (Messinian) are essentially the same as those in coeval Central Anatolian sites (see high Raup–Crick similarity in Fig. 3 and Supplementary Fig. 3), except that Asian sites also have high levels of *Artemisia*, Chenopodioideae and Tamaricaceae, pointing to a steppe grassland biome and an arid climate (Köppen BSk climate; Table 2), as also inferred by Fig. 5 in Kaya et al.¹⁰. Records of xeric grasslands are known only from the Russian Plain, beginning from the late–middle Miocene²⁹. The Mongolian steppe biome seems to have existed for at least 20 million years³⁰ and may have originated with the aridification caused by the westward Paratethys retreat, dated to between 47 and 37 Ma, which reduced the moisture load of westerlies blowing into the Asian continental interior³¹.

Western Eurasian sites contain little grass pollen, and grass pollen peaks >20% occur in only four palynofloras (Table 1). Isotope analyses of palaeosols from Samos, Pikermi, Rhodes and the lower Axios Valley in Macedonia and of fossil mammal teeth from Samos and from Paşalar in northwest Turkey also imply a restricted extent of C3 grasses and an absence of C4 grasses³². Nevertheless, C4 grasses, including Panicoideae, have been reported from at least one locality in Central Anatolia (7.1 ± 0.1 Ma³³), and sites in Bulgaria and Greece (MN12+13, 7.6–4.8 Ma^{34,35}). They are also documented from early–middle Miocene (20–10 Ma) localities in southwestern Europe^{36,37} and classic Pikermian localities from northern China (less than 30% of the vegetation consumed by gazelles and *Hipparion* grazers³⁸, but up to 76% in Inner Mongolian herbivore diets after 7.5 Ma³⁹).

On the Iranian Plateau, C4 grasslands were present by 8.2 Ma⁴⁰, in Indo–Pakistan by 8 Ma and in East Africa by 10 Ma^{40–43}.

In sum, the plant fossil record reveals forest biomes in Miocene sites relevant to the late Miocene fauna, namely TBMF (corresponding to deciduous broadleaf forests and mixed forests of Woodward et al.¹⁶); temperate coniferous forests (corresponding to ENF) and temperate grasslands, savannahs and shrublands (corresponding to woody savannahs and grassland biomes^{16,17}; see Methods for details of assignments), with the western and eastern Asian sites sharing around 75% of the genera among contemporaneous assemblages (Figs. 2 and 3).

Conclusions

The contrasting views obtained from zoological versus botanical proxies used to reconstruct Old-World biomes in the late Miocene highlight the difficulty of inferring past organismal communities. As per the original definition of the concept, a biome ‘is fundamentally controlled by the habitat and exhibits a corresponding development and structure’⁴⁴, p. 120. Biomes are initially characterized by plant life forms but include animals whose morphological, physiological and behavioural traits are adapted to the plant communities in which they live⁴⁵. Among the world’s most important biomes in terms of spatial extent are grasslands and savannahs, which together may comprise about 40% of the Earth’s unfrozen land surface⁴⁶. Savannahs are mixed tree–grass systems, with a more or less continuous grass understorey, and ‘wooded savannahs’ may have tree cover of up to 80%^{16,47}, depending on fire frequency and grazing, which can yield mosaics of savannah and forest in ecotone regions. Savannahs may be further categorized into temperate C3 grasslands (steppes, prairies) and subtropical/tropical savannahs with C4 grasses^{17,47–51}. Some savannah definitions include a low mean annual precipitation and a high mean annual temperature⁵². This variation in the definition of what constitutes a savannah makes it extremely difficult to infer past savannah biomes.

A further difficulty in inferring past organismal communities is that the flora and fauna of any area is the result of the fluctuating and fortuitous immigration of plants and animals in an equally fluctuating and variable environment^{45–55}. Spatial communities, such as biomes, are therefore epiphenomena, and their components are not all co-evolved or co-adapted to the extent that a few suffice to predict the remaining species’ traits or the conditions of the environment. For a reliable reconstruction, both animal and plant traits are required, along with soil types, isotope composition and other proxies. Mean hypsodonty as a measure of aridity has never been tested and does not reliably indicate the presence of grassland per se^{10,32,56,57}. Plant micro- and macrofossils, by contrast, readily differentiate extensive C3 or C4 grasslands, needleleaf forests, evergreen

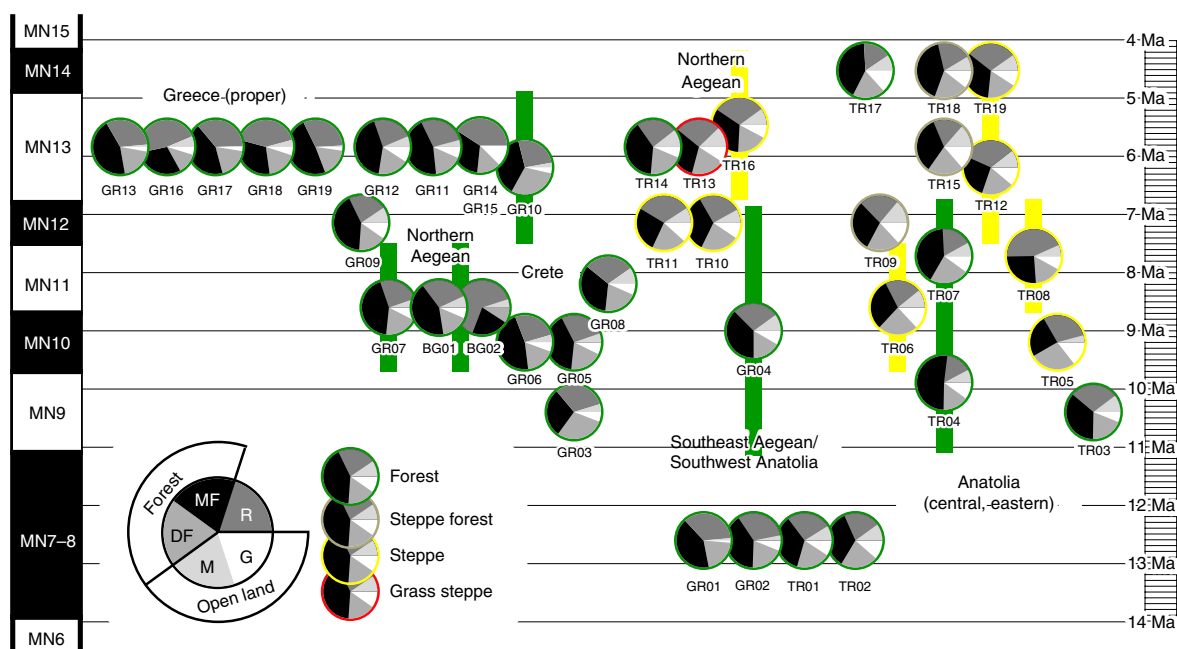


Fig. 1 | Vegetation and landscapes inhabited by the Pikermian fauna from Greece to Anatolia. More open vegetation evolved in Central Anatolia between MN10 and MN11. The vegetation units (as in Supplementary Fig. 1) are: MF, montane conifer forest and mesic forest (with or without *Fagus*); DF, edaphically and aspect-wise dry forest; R, swamp and riparian forests; M, intramontane meadows; and G, grassland and open patches. Localities GR01–GR19 refer to locations in Greece, BG01 and BG02 to locations in Bulgaria, TR01–TR19 to locations in Turkey (for coordinates see Supplementary Table 1). European Neogene Mammal units are on the left, Ma scale on the right. See Methods for further explanation.

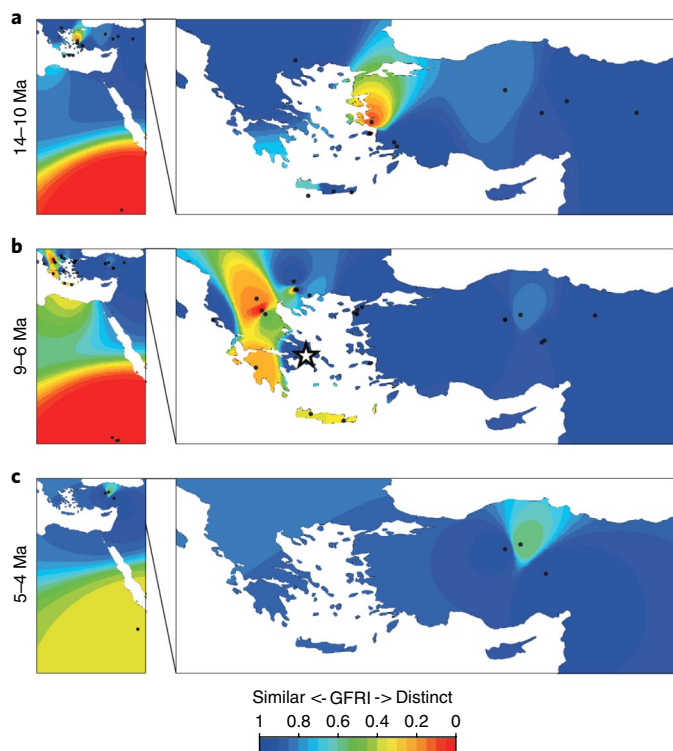


Fig. 2 | Raup–Crick genus-level floristic similarity^{12,13} to Pikermi for different time periods. **a**, 14–10 Ma; **b**, 9–6 Ma; **c**, 5–4 Ma; original data in Supplementary Table 3. Inverse distance-weighting was used to interpolate between sampling points. Left panels show Africa and southern Eurasia; right panels show East Europe/West Asia. Localities from each time slice are shown as black dots. A black star indicates the location of Pikermi at 7 Ma. GFRI, Global Floristic Resemblance Index.

broadleaf forests and deciduous forests, and therefore they can help directly discern past vegetation. The plant fossil record analysed here shows that the Pikermian chronofauna lived in different biomes, including mixed coniferous broadleaved forests, mesophytic forests and more herb-rich lakeside/riverine associations (in agreement with refs^{18,19}), Asian steppes and some C3- or C4-dominated grasslands. More research is needed to resolve the geographic extent of C4 grasslands relevant to Pikermian animals. Another open question is the role of interlocking vegetation mosaics that permitted co-occurring animals to feed on diverse sets of plants (more diverse than today). Clearly, however, plant fossil data do not support an interpretation of modern African savannahs as a surviving branch of a single cohesive Old-World savannah palaeobiome.

Methods

Inference of vegetation types and biomes. We investigated macrofossils, pollen and spores from 1,602 fossil occurrences from 45 sites in Greece, Bulgaria and Turkey; 88 macrofossil records from two sedimentary basins in East Africa; and 154 microfossil records from eight sites in Mongolia and China, covering European Neogene Mammal units MN7 + 8 to MN14, that is, 13.8–4.2 Ma (Serravallian to mid-Zanclean). Of our floristic assemblages, 15 western Eurasian localities correspond to the climax of the Pikermi fauna between 8 and 7 Ma. For the 9–6 Ma time slice when the Pikermian fauna stretched as far west as present-day Spain and France, north to the Ukraine, Kazakhstan and Mongolia, and east to China¹⁰ we have >25 western Eurasian localities. We compiled taxon lists based on macrofossils (foliage, fruits and seeds) and pollen and spores; see Supplementary Tables 1–3. The revised fossil-taxon records were assigned to one or more modern vegetation types based on the modern occurrence of taxa (clades) related to the fossil taxon. The following vegetation types were scored: R: riparian, swamp forest, including aquatic plants (Supplementary Fig. 1a,c). MF: mesic forests. Most taxa assigned to this vegetation unit are typical members of the TBMF biome according to Olson et al.¹⁷ and World Wide Fund for Nature (WWF)⁵⁸. The most typical trees of this biome in the Northern Hemisphere are Fagaceae, Betulaceae, *Acer* and the Pinaceae (Supplementary Fig. 1a). The rich conifer component represents the ENF biome of Woodward et al.¹⁶. DF: dry forests. Taxa assigned to this vegetation unit are members of the TBMF^{17,58} and (partly) the shrubland biome¹⁶. It also includes some members of genera typical of mesic forests that adapted to edaphically drier conditions (for example, certain types of *Acer* and *Quercus*, *Ulmus*, *Zelkova*, Supplementary Fig. 1b,c). Very few woody taxa are found exclusively in dry forests

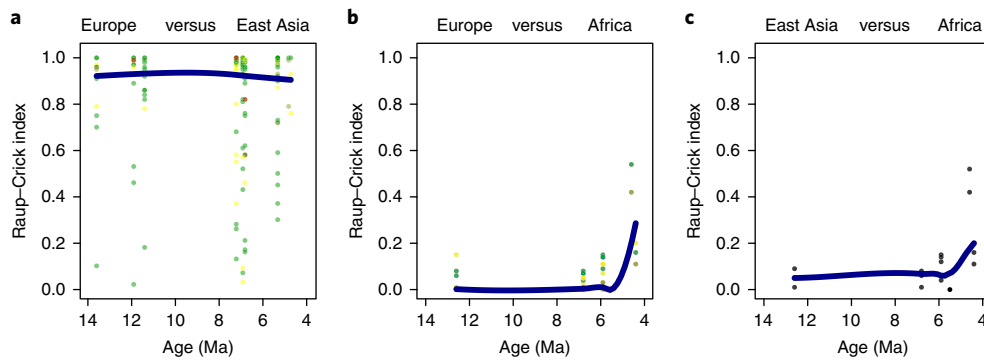


Fig. 3 | Raup-Crick similarities^{12,13} over time between Europe and East Asia, Europe and Africa, and East Asia and Africa computed on the fossil dataset, using reference sites from the same time periods (14–10 Ma, 9–6 Ma, 5–4 Ma). a–c, European (a,b) or Asian (c) sites were used as reference sites.

Each dot represents the comparison between two localities, with dot colours indicating the biome type of the respective reference site according to Fig. 1 (green, forest; olive, steppe forest; yellow, steppe; red, grass steppe). For interpolation, local polynomial regression fitting (loess) with the second-degree polynomial and 0.75 degree of smoothing was used (same parameters as in Kaya et al.¹⁰ to facilitate direct comparison with their Supplementary Fig. 3).

and shrublands (for example, *Cistus*, *Ephedra*). M: open areas not dominated by grasses ('meadows' in a very broad sense). This vegetation unit denotes open areas that are inhabited by both grasses and non-grass herbaceous plants. Included here are plants growing in bogs, wet meadows, on rocky substrates, *Artemisia* steppe, steppe dominated by *Amaranthaceae*, and disturbed areas among others (Supplementary Fig. 1b). G: grassland and woody grassland. This vegetation unit refers to open landscapes with a grass-dominated ground layer. This includes both temperate grasslands^{59,60} (Supplementary Fig. 1c) and tropical grasslands (savannah^{47,49}). M and G correspond to the temperate grasslands, savannahs and shrublands biome of Olson et al.¹⁷ and to the savannah and grassland biomes of Woodward et al.¹⁶ Taxa occurring in tropical African savannahs are marked with an asterisk (Supplementary Table 2). Taxa occurring in more than one vegetation type were scored 0.5, 0.333 etc. for each of two or more vegetation types.

Taphonomy. The plant fossil data used to infer palaeo-environments represent different organs of plants. Taken alone, most of these datasets can be biased towards particular environments, life forms and aspects of the vegetation. For example, the carpological record (fruit and seed remains) commonly overrepresents the vegetation surrounding a lake and underrepresents the vegetation from the hinterland (regional). In contrast, leaf floras commonly represent both the local and regional vegetation but strongly underrepresent the herbaceous component of the flora. Leaf size, frequency and leaf texture can be used to discern local vegetation from hinterland vegetation²⁰. Plant silica reflect the diversity of grasses but may underestimate the contribution of other plants to the vegetation⁶¹. Palynofloras (dispersed spores and pollen) represent both local and regional vegetation and cover woody and herbaceous plants. Wind-pollinated plants are overrepresented whereas insect-pollinated plants are underrepresented. Including information from modern pollen rain and abundances of different pollen types, palynofloras provide a highly accurate picture of past vegetation types (for example, ref. ⁶²). For example, modern pollen rain studies across Anatolia⁶³ showed that local peaks of *Asteraceae* pollen might represent the temporary exposure of a lakeshore, while peaks of *Artemisia* might indicate early forms of Anatolian steppe.

Pollen signatures. Pollen diagrams reflect local and regional vegetation and can be used to discriminate between TBMF, northern temperate steppe and steppe forest, and tropical African savannahs. Pollen signatures indicative of TBMF include moderate to high amounts of Pinaceae, *Acer*, Betulaceae, Fagaceae and Juglandaceae, which are among the most characteristic elements of TBMF according to WWF⁵⁸. In addition, taxa from the Northern Hemispheric ENF biome of Woodward et al.¹⁶ are strongly represented in TBMF (Table 1). Pollen diagrams indicating northern temperate steppes and steppe forest have high values of *Artemisia*, *Amaranthaceae*–*Chenopodioidae*, *Asteraceae* and Poaceae (temperate grasslands, savannahs and shrubland biome of Olson et al.¹⁷) and relatively low values of TBMF elements. Pine pollen can reach moderately high values. Pollen signatures of tropical savannahs, by contrast, include >40% to >80% Poaceae and Cyperaceae pollen, along with pollen of mainly (Afro)tropical genera in low amounts^{65,64} (Table 2). As a rule, they do not include any of the TBMF taxa mentioned above. Using these criteria, pollen floras are tagged as mainly reflecting MF, DF, SF, S and GS based on comparison to modern vegetation types or potential vegetation in the eastern Mediterranean region^{59,65,66}. Northeast Asian steppe has a similar pollen signature as western Eurasian steppe and steppe forest, but the TBMF component is much less prominent⁶⁷.

Phytoliths (plant silica bodies). Strömberg et al.⁶⁸ analysed plant silica bodies from late Miocene sites in Anatolia. They recognized the following four groups:

AQ (wetland plants and aquatics); FI TOT (forest indicator, FI, taxa) comprising palms and other FI (basal angiosperms and eudicots, conifers, ferns etc.); NDG (non-diagnostic phytoliths of 'grasses', other monocots and conifers (6–14% in samples EM15 to EM28)); NDO (undetermined phytoliths (30–50% in samples EM15 to EM28)). They further subdivided Poaceae phytoliths (grass short silica cells) into CH TOT (closed habitat grasses, that is, basal Poaceae grade and Bambusoideae, Erhartoideae); POOID-D (diagnostic Pooideae); POOID-ND (non-diagnostic Pooideae); PACCAD TOT (Panicoidae, Chloridoideae); PACCAD general (indetermined PACCAD); and OTHG (other grasses, unknown Poaceae). Combining their four groups and Poaceae subdivisions, Strömberg et al. devised seven categories: (1) palms, (2) other FI, (3) CH TOT, (4) POOID-D + POOID-ND (termed: Pooide open habitat grasses), (5) Chloridoide, (6) Panicoid + PACCAD general and (7) non-diagnostic phytoliths outside Poaceae (phytolith classes 'NDG', 'NDO'). In our study (Supplementary Fig. 2b), the AQ group was coded as WL (wetlands); FI TOT as NG (non-grass component), NDO and NDG as ?; CH TOT as BE (Bambusoideae, Erhartoideae clades); POOID-D as Po (unambiguous Pooideae); PACCAD TOT as PC (Panicoidae, Chloridoideae); and POOID-ND and OTHG as ?P (undetermined Poaceae).

Statistical similarity. We used the Raup-Crick similarity index^{12,13} to quantify genus-level floristic resemblance between localities, relying on an implementation of the index in the Community Ecology Package ('vegan') in R (<https://cran.r-project.org/package=vegan>). The index assumes that taxa are distributed randomly in space and time and that species lists are made up from the taxa that happen to fall in certain areas and in certain stratigraphic intervals. An average expected similarity ('k') and the expected variation in this number is predictable from the numbers of taxa, areas and stratigraphic intervals involved. The expected 'k' and its probable variation constitute the appropriate null hypothesis for assessing organismal (taxonomic) similarity. When dealing with assemblages from different continents, one would expect to be able to reject the null hypothesis most of the time; when dealing with assemblages from the same formation in a local area, one would expect not to reject the null hypothesis and to conclude that the compositional differences are the result of chance differences in sampling. The index has the advantage that distributional data are weighted on the basis of frequency so that widespread taxa do not have a disproportionate influence on similarity, and there is therefore no need to discard taxa on the a priori grounds that they are too widespread or too localized. Problems arise where one of the assemblages is very small. Two localities had only six taxa identified at least to the genus level, all others had more than six. To compute Raup-Crick similarities^{12,13} over time between continents, we did pair-wise comparisons between each African or Asian locality and every European locality (Fig. 3a,b) and between each African locality and every Asian locality (Fig. 3c) from the same time period (14–10 Ma, 9–6 Ma, 5–4 Ma). Thus, when comparing African to European sites between 4–5 Ma, all European sites between 4–5 Ma were chosen as reference; when comparing sites between 6–9 Ma, all sites between 6–9 Ma were the reference; when comparing sites between 10–14 Ma, all sites between 10–14 Ma were the reference. Each dot represents the comparison between two localities, with dot colours indicating the biome type of the respective reference site according to Fig. 1 (green, forest; olive, steppe forest; yellow, steppe; red, grass steppe). For temporal interpolation, local polynomial regression fitting (loess) with the second-degree polynomial and 0.75 degree of smoothing was used (same parameters as in Kaya et al.¹⁰ to facilitate direct comparison with their Supplementary Fig. 3). To create maps showing Raup-Crick genus-level floristic similarities, Pikermi at 7 Ma was used as reference site (Fig. 2). Inverse distance-weighting was used to interpolate between sampling points.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The authors declare that all data supporting the findings of this study are available in the supplementary tables.

Received: 11 April 2018; Accepted: 13 September 2018;

Published online: 29 October 2018

References

- Solounias, N. & Dawson-Saunders, B. Dietary adaptations and paleoecology of the late Miocene ruminants from Pikermi and Samos in Greece. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **65**, 149–172 (1988).
- Bernor, R. L., Solounias, N., Swisher, C. C. III & van Couvering, J. A. in *The Evolution of Western Eurasian Neogene Mammal Faunas* (eds Bernor, R. L., Fahlbusch, V. & Mittmann, H.-W.) 137–154 (Columbia Univ. Press, New York, 1996).
- Fortelius, M. et al. in *The Evolution of Western Eurasian Neogene Mammal Faunas* (eds Bernor, R. L., Fahlbusch, V. & Mittmann, H.-W.) 414–448 (Columbia Univ. Press, New York, 1996).
- Eronen, J. T. et al. Distribution history and climatic controls of the late Miocene Pikermian chronofauna. *Proc. Natl Acad. Sci. USA* **106**, 11867–11871 (2009).
- Solounias, N., Rivals, F. & Semperebon, G. M. Dietary interpretation and paleoecology of herbivores from Pikermi and Samos (late Miocene of Greece). *Paleobiology* **36**, 113–136 (2010).
- Bernor, R. L., Andrews, P. J., Solounias, N. & van Couvering, J. A. H. The evolution of ‘Pontian’ mammal faunas: some zoogeographic, paleoecologic and chronostratigraphic considerations. *Ann. Géol. Pays Hellén. Hors Sér.* **1**, 81–89 (1979).
- Solounias, N., Plavcan, J. M., Quade, J. & Witmer, L. in *The Evolution of Neogene Terrestrial Ecosystems in Europe* (eds Agustí, J., Rook, L. & Andrews, P.) 436–453 (Cambridge Univ. Press, Cambridge, 1999).
- Jernvall, J. & Fortelius, M. Common mammals drive the evolutionary increase of hypsodonty in the Neogene. *Nature* **417**, 538–540 (2002).
- Fortelius, M. et al. Fossil mammals resolve regional patterns of Eurasian climate change over 20 million years. *Evol. Ecol. Res.* **4**, 1005–1016 (2002).
- Kaya, F. et al. The rise and fall of the Old-World savannah fauna and the origins of the African savannah biome. *Nat. Ecol. Evol.* **2**, 241–246 (2018).
- Bernor, R. L. et al. Systematic, stratigraphic, and paleoenvironmental contexts of first-appearing *Hipparion* in the Vienna Basin, Austria. *J. Vertebr. Paleontol.* **8**, 427–452 (1988).
- Raup, D. & Crick, R. E. Measurement of faunal similarity in paleontology. *J. Paleontol.* **53**, 1213–1227 (1979).
- Chase, J. M., Kraft, N. J. B., Smith, K. G., Vellend, M. & Inouye, B. D. Using null models to disentangle variation in community dissimilarity from variation in α -diversity. *Ecosphere* **2**, art24 (2011).
- Harzhauser, M. & Piller, W. E. Benchmark data of a changing sea—palaeogeography, palaeobiogeography and events in the central Paratethys during the Miocene. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **253**, 8–31 (2007).
- Cohen, K. M., Finney, S. C., Gibbard, P. L. & Fan, J.-X. The ICS international chronostratigraphic chart. *Episodes* **36**, 199–204 (2013).
- Woodward, F. I., Lomas, M. R. & Kelly, C. K. Global climate and the distribution of plant biomes. *Phil. Trans. R. Soc. Lond. B* **359**, 1465–1476 (2004).
- Olson, D. M. et al. Terrestrial ecosystems of the world: a new map of life on Earth. *Bioscience* **51**, 933–938 (2001).
- Kovar-Eder, J. Pannonian (upper Miocene) vegetational character and climatic inferences in the central Paratethys area. *Ann. Naturhist. Mus. Wien* **88**, 117–129 (1987).
- Kovar-Eder, J., Kvaček, Z., Martinetto, E. & Roiron, P. Late Miocene to early Pliocene vegetation of southern Europe (7–4 Ma) as reflected in the megafossil plant record. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **238**, 321–339 (2006).
- Ivanov, D. et al. Miocene vegetation and climate dynamics in eastern and central Paratethys (southeastern Europe). *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **304**, 262–275 (2011).
- Draxner-Höck, G., Harzhauser, M. & Göhlich, U. B. Fossil record and dynamics of late Miocene small mammal faunas of the Vienna Basin and adjacent basins, Austria. *C. R. Palevol.* **15**, 855–862 (2016).
- Bouchal, J. M. et al. Miocene palynofloras of the Tinaz lignite mine, Muğla, southwest Anatolia: taxonomy, paleoecology and local vegetation change. *Rev. Palaeobot. Palyno.* **243**, 1–36 (2017).
- Syabryaj, S., Utescher, T., Molchanoff, S. & Bruch, A. A. Vegetation and palaeoclimate in the Miocene of Ukraine. *Palaeogeography, Palaeoclimatology, Palaeoecology* **253**, 153–168 (2007).
- Merceron, G., Novello, A. & Scott, R. S. Paleoenvironments inferred from phytoliths and dental microwear texture analyses of meso-herbivores. *Geobios* **49**, 135–146 (2016).
- Velitzelos, D., Bouchal, J. M. & Denk, T. Review of the Cenozoic floras and vegetation of Greece. *Rev. Palaeobot. Palyno.* **204**, 56–117 (2014).
- Denk, T., Güner, H. T. & Grimm, G. W. From mesic to arid: leaf epidermal features suggest preadaptation in Miocene dragon trees (*Dracaena*). *Rev. Palaeobot. Palyno.* **200**, 211–228 (2014).
- Bertini, A. & Martinetto, E. Messinian to Zanclean vegetation and climate of northern and central Italy. *B. Soc. Paleontol. Ital.* **47**, 105–121 (2008).
- Kurtén, B. The Chinese Hipparion fauna. A quantitative survey with comments on the ecology of the machairodonts and hyaenids and the taxonomy of the gazelles. *Comment. Biol.* **13**, 1–82 (1952).
- Kovar-Eder, J. & Kvaček, Z. The integrated plant record (IPR) to reconstruct Neogene vegetation: the IPR-vegetation analysis. *Palaios* **23**, 97–111 (2007).
- Jiang, H. & Ding, Z. A 20 Ma pollen record of East Asian summer monsoon evolution from Guyuan, Ningxia, China. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **265**, 30–38 (2008).
- Bosboom, R. et al. Linking Tarim Basin sea retreat (West China) and Asian aridification in the late Eocene. *Basin Res.* **26**, 621–640 (2014).
- Quade, J., Solounias, N. & Cerling, T. E. Stable isotopic evidence from paleosol carbonates and fossil teeth in Greece for forest or woodlands over the past 11 Ma. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **108**, 41–53 (1994).
- Kayseri-Özer, M. S. et al. Palaeoclimatic and palaeo-environmental interpretations of the late Oligocene, late Miocene–early Pliocene in the Çankiri-Çorum Basin. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **467**, 16–36 (2017).
- Spassov, N., Böhme, M., Geraads, D., Kötter, S. & van Baak, C. P. Pikermian mammal event, post-Pikermian mammal turnover and appearance of *Graecopithecus*. In *15th Congress of the Regional Committee on Mediterranean Neogene Stratigraphy, Book of Abstracts*, 29 (RCMNS, 2017); <http://www.rcmns2017.com/files/pdf/BOOK-OF-ABSTRACTS.pdf>
- Böhme, M. et al. Messinian age and savannah environment of the possible hominin *Graecopithecus* from Europe. *PLoS ONE* **12**, e0177347 (2017).
- Urban, M. A. et al. Isotopic evidence of C₄ grasses in southwestern Europe during the early Oligocene–middle Miocene. *Geology* **38**, 1091–1094 (2010).
- Urban, M. A., Nelson, D. M., Jiménez-Moreno, G. & Hu, F. S. Carbon isotope analyses reveal relatively high abundance of C₄ grasses during early–middle Miocene in southwestern Europe. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **443**, 10–17 (2016).
- Passey, B. H., Eronen, J. T., Fortelius, M. & Zhang, Z. Paleodiets and paleoenvironments of late Miocene gazelles from north China: evidence from stable carbon isotopes. *Vertebrat. Palasiatic.* **45**, 118–127 (2007).
- Zhang, C. et al. C₄ expansion in the central Inner Mongolia during the latest Miocene and early Pliocene. *Earth Planet. Sc. Lett.* **287**, 311–319 (2009).
- Bernor, R. L., Ataabadi, M. M., Meshida, K. & Wolf, D. The Maragheh hipparions, late Miocene of Azarbaijan, Iran. *Palaeobio. Palaeoenv.* **96**, 453–488 (2016).
- Barry, J. C. et al. Faunal and environmental change in the Late Miocene Siwaliks of northern Pakistan. *Paleobiology* **28**, 1–71 (2002).
- Feakins, S. J., Levin, N. E., Liddy, H. M., Sieracki, A., Eglinton, T. I. & Bonnefille, R. Northeast African vegetation change over 12 m.y. *Geology* **41**, 295–298 (2013).
- Flynn, L. J., Pilbeam, D., Barry, J. C. & Morgan, M. E. Siwalik synopsis: a long stratigraphic sequence for the later Cenozoic of South Asia. *C. R. Palevol.* **15**, 877–887 (2016).
- Clements, F. E. The development and structure of biotic communities. *J. Ecol.* **5**, 120–121 (1917).
- Egerton, F. N. History of ecological sciences, part 59: niches, biomes, ecosystems, and systems. *Bull. Ecol. Soc. Am.* **98**, 298–337 (2017).
- Bastin, J.-F. et al. The extent of forest in dryland biomes. *Science* **356**, 635–638 (2017).
- Parr, C. L., Lehmann, C. E. R., Bond, W. J., Hoffmann, W. A. & Andersen, A. N. Tropical grassy biomes: misunderstood, neglected, and under threat. *Trends Ecol. Evol.* **29**, 205–213 (2014).
- Ratnam, J. et al. When is a ‘forest’ a savanna, and why does it matter? *Glob. Ecol. Biogeogr.* **20**, 653–660 (2011).
- Ratnam, J., Tomlinson, K. W., Rasquinha, D. N. & Sankaran, M. Savannas of Asia: antiquity, biogeography, and an uncertain future. *Phil. Trans. R. Soc. B* **371**, 20150305 (2016).
- Lehmann, C. E., Archibald, S. A., Hoffmann, W. A. & Bond, W. J. Deciphering the distribution of the savanna biome. *New Phytol.* **191**, 197–209 (2011).
- Griffith, D. M. et al. Comment on ‘The extent of forest in dryland biomes’. *Science* **358**, eaao1309 (2017).
- Whittaker, R. H. *Communities and Ecosystems* 2nd edn (Macmillan, New York, 1975).
- Gleason, H. A. The individualistic concept of the plant association. *Torrey Botanical Club Bull.* **53**, 7–26 (1926).
- Ricklefs, R. E. History and diversity: explorations at the intersection of ecology and evolution. *Am. Nat.* **170**, S56–S70 (2007).

55. Ricklefs, R. E. Disintegration of the ecological community. *Am. Nat.* **172**, 741–750 (2008).
56. Hayek, L.-A. C., Bernor, R. L., Solounias, N. & Steigerwald, P. Preliminary studies of hipparionine horse diet as measured by tooth microwear. *Ann. Zool. Fenn.* **28**, 187–200 (1992).
57. Kaiser, T. M., Solounias, N., Fortelius, M., Bernor, R. & Schrenk, F. Tooth mesowear analysis on *Hippotherium primigenium* from the Vallesian Dinotheriensande (Germany)—a blind test study. *Carolinea* **58**, 103–114 (2000).
58. *Terrestrial Ecoregions* (WWF, 2018); <https://www.worldwildlife.org/biome-categories/terrestrial-ecoregions>
59. Walter, H. & Breckle, S.-W. *Ökologie der Erde. 3. Spezielle Ökologie der gemäßigten und Arktischen Zonen Euro-Nordasiens* (Gustav Fischer Verlag, Stuttgart, 1986).
60. Woldring, H. & Cappers, R. The origin of the ‘wild orchards’ of Central Anatolia. *Turk. J. Bot.* **25**, 1–9 (2001).
61. Bremond, L., Alexandre, A., Vêla, E. & Guiot, J. Advantages and disadvantages of phytolith analysis for the reconstruction of Mediterranean vegetation: an assessment based on modern phytolith, pollen and botanical data (Luberon, France). *Rev. Palaeobot. Palyno.* **129**, 213–228 (2004).
62. Lisitsyna, O. V., Giesecke, T. & Hicks, S. Exploring pollen percentage threshold values as an indicator for the regional presence of major European trees. *Rev. Palaeobot. Palyno.* **166**, 311–324 (2011).
63. Salzmann, U. Are modern savannas degraded forests? A Holocene pollen record from the Sudanian vegetation zone of NE Nigeria. *Veg. Hist. Archaeobot.* **9**, 1–15 (2000).
64. Zhou, Y. et al. Vascular flora of Kenya, based on the flora of tropical East Africa. *PhytoKeys* **90**, 113–126 (2017).
65. Van Zeist, W., Woldring, H. & Stapert, D. Late Quaternary vegetation and climate of southwestern Turkey. *Palaeohistoria* **17**, 53–143 (1975).
66. Sirenko, E. A. Microrhythms in the evolution of Pliocene and early Pleistocene vegetation in eastern Ukraine. *Paleontol. J.* **34**(suppl.1), S81–S86 (2000).
67. Qin, F. et al. Utility of surface pollen assemblages to delimit eastern Eurasian steppe types. *PLoS ONE* **10**, e0119412 (2015).
68. Strömberg, C. A. E., Werdelin, L., Friis, E. M. & Saraç, G. The spread of grass-dominated habitats in Turkey and surrounding areas during the Cenozoic: phytolith evidence. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **250**, 18–49 (2007).
69. Peel, M. C., Finlayson, B. L. & McMahon, T. A. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.* **11**, 1633–1644 (2007).

Acknowledgements

This work was supported by a Swedish Research Council grant to T.D. and an Austrian Science Fund Grant M 1751 (to G.W.G.). We thank F. Kaya for discussion of his Raup–Crick analyses and the shortcomings of different proxies for inferring past biomes.

Author contributions

T.D. designed the initial study and performed analyses. C.M.Z. performed Raup–Crick analyses. G.W.G. performed Köppen climate analyses and designed most figures. S.S.R. provided conceptual input and wrote the first draft. All authors co-wrote the final paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41559-018-0695-z>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to T.D. or S.S.R.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2018

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a | Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection

We compiled taxon lists based on macrofossils (foliage, fruits, and seeds) and pollen and spores; see Supplementary Tables 1–3. In all, we investigated macrofossils, pollen, and spores from 1602 fossil occurrences from 45 sites in Greece, Bulgaria and Turkey; 88 macrofossil records from two sedimentary basins in East Africa; and 154 macrofossil records from eight sites in Mongolia and China, covering European Neogene Mammal units MN7+8 to MN14, that is, 13.8–4.2 Ma (Serravallian to mid-Zanclean).

Data analysis

We used the Raup-Crick similarity index (Raup and Crick, 1979; Chase et al., 2011) to quantify genus-level floristic resemblance between localities, relying on an implementation of the index in the Community Ecology Package ('vegan') in R (<https://cran.r-project.org/package=vegan>).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The authors declare that all data supporting the findings of this study are available in the supplementary tables.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Based on terrestrial mammal fossils, late Miocene vegetation in Eurasia and East Africa has been reconstructed as a single cohesive biome of which modern African savannahs are the surviving remnant. Here, we test this reconstruction using plant fossils spanning 14–4 Ma from sites in Greece, Bulgaria, Turkey, the Tian Shan Mountains and Baode County in China, and East Africa. The Western Eurasian sites had a continuous forest cover of deciduous or evergreen angiosperms and gymnosperms, with 15% of 1602 fossil occurrences representing conifers, which are present at all but one of the sites. Raup-Crick analyses reveal high floristic similarity between coeval Western and Eastern Eurasian sites, and low similarity between Eurasian and African sites. At least four distinct biomes are represented.
Research sample	We compiled taxon lists based on macrofossils (foliage, fruits, and seeds) and pollen and spores; see Supplementary Tables 1–3. In all, we investigated macrofossils, pollen, and spores from 1602 fossil occurrences from 45 sites in Greece, Bulgaria and Turkey; 88 macrofossil records from two sedimentary basins in East Africa; and 154 macrofossil records from eight sites in Mongolia and China, covering European Neogene Mammal units MN7+8 to MN14, that is, 13.8–4.2 Ma (Serravallian to mid-Zanclean). We used the Raup-Crick similarity index (Raup and Crick, 1979; Chase et al., 2011) to quantify genus-level floristic resemblance between localities, relying on an implementation of the index in the Community Ecology Package ('vegan') in R (https://cran.r-project.org/package=vegan).
Sampling strategy	Please see under "Research sample"
Data collection	Please see under "Research sample"
Timing and spatial scale	1602 fossil occurrences from 45 sites in Greece, Bulgaria and Turkey; 88 macrofossil records from two sedimentary basins in East Africa; and 154 macrofossil records from eight sites in Mongolia and China, covering European Neogene Mammal units MN7+8 to MN14, that is, 13.8–4.2 Ma (Serravallian to mid-Zanclean).
Data exclusions	No data were excluded.
Reproducibility	Using the data in our Supplementary Tables 1-3 and the Raup-Crick statistical test for similarity, our results are reproducible.
Randomization	The Raup-Crick index assumes that taxa are distributed randomly in space and time and that species lists are made up from the taxa that happen to fall in certain areas and in certain stratigraphic intervals. The randomization is done automatically in the R package the implements the test.
Blinding	Not applicable
Did the study involve field work?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involvement	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Unique biological materials
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Eukaryotic cell lines
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Human research participants

Methods

n/a	Involvement	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/>	MRI-based neuroimaging

Unique biological materials

Policy information about [availability of materials](#)

Obtaining unique materials

Palaeontology

Specimen provenance

Specimen deposition

Dating methods

Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.