



**HAL**  
open science

## Reconstructing the post-glacial spread of the sand fly *Phlebotomus mascittii* Grassi, 1908 (Diptera: Psychodidae) in Europe

Edwin Kniha, Vít Dvořák, Stephan Koblmüller, Jorian Prudhomme, Vladimir Ivović, Ina Hoxha, Sandra Oerther, Anna Heitmann, Renke Lühken, Anne-Laure Bañuls, et al.

### ► To cite this version:

Edwin Kniha, Vít Dvořák, Stephan Koblmüller, Jorian Prudhomme, Vladimir Ivović, et al.. Reconstructing the post-glacial spread of the sand fly *Phlebotomus mascittii* Grassi, 1908 (Diptera: Psychodidae) in Europe. *Communications Biology*, 2023, 6 (1), pp.1244. 10.1038/s42003-023-05616-1 . hal-04510287

**HAL Id: hal-04510287**








**<https://hal.science/hal-04510287v1>**

Submitted on 29 Mar 2024

**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.

## Reconstructing the post-glacial spread of the sand fly *Phlebotomus mascittii* Grassi, 1908 (Diptera: Psychodidae) in Europe

Edwin Kniha<sup>1</sup>, Vít Dvořák<sup>2</sup>, Stephan Koblmüller <sup>3</sup>, Jorian Prudhomme<sup>4,5</sup>, Vladimir Ivović<sup>6</sup>, Ina Hoxha<sup>1</sup>, Sandra Oerther<sup>7,8,9</sup>, Anna Heitmann<sup>10</sup>, Renke Lühken<sup>10,11</sup>, Anne-Laure Bañuls<sup>4</sup>, Denis Sereno <sup>4,12</sup>, Alice Michelutti<sup>13</sup>, Federica Toniolo<sup>13</sup>, Pedro M. Alarcón-Elbal<sup>14,15</sup>, Daniel Bravo-Barriga <sup>16</sup>, Mikel A. González<sup>17,18</sup>, Javier Lucientes<sup>19</sup>, Vito Colella<sup>20</sup>, Domenico Otranto<sup>21,22</sup>, Marcos Antônio Bezerra-Santos <sup>21</sup>, Gernot Kunz<sup>3</sup>, Adelheid G. Obwaller<sup>23</sup>, Jerome Depaquit<sup>24</sup>, Amer Alić<sup>25</sup>, Ozge Erisoz Kasap<sup>26</sup>, Bulent Alten<sup>26</sup>, Jasmin Omeragic<sup>27</sup>, Petr Volf <sup>2</sup>, Julia Walochnik <sup>1</sup>, Viktor Sebestyén<sup>28</sup> & Attila J. Trájer <sup>28</sup>✉

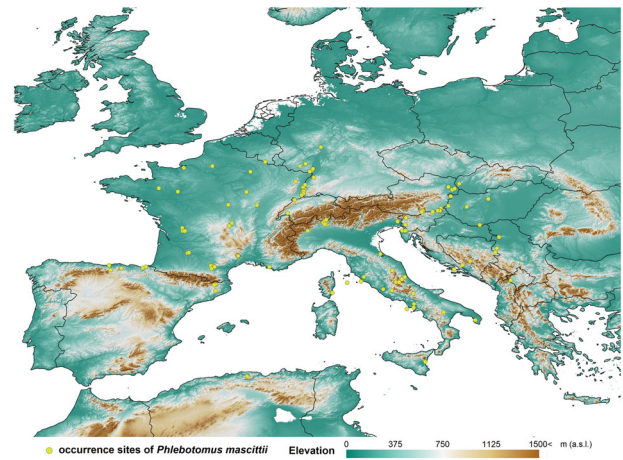
Phlebotomine sand flies (Diptera: Phlebotominae) are the principal vectors of *Leishmania* spp. (Kinetoplastida: Trypanosomatidae). In Central Europe, *Phlebotomus mascittii* is the predominant species, but largely understudied. To better understand factors driving its current distribution, we infer patterns of genetic diversity by testing for signals of population expansion based on two mitochondrial genes and model current and past climate and habitat suitability for seven post-glacial maximum periods, taking 19 climatic variables into account. Consequently, we elucidate their connections by environmental-geographical network analysis. Most analyzed populations share a main haplotype tracing back to a single glacial maximum refuge area on the Mediterranean coasts of South France, which is supported by network analysis. The rapid range expansion of *Ph. mascittii* likely started in the early mid-Holocene epoch until today and its spread possibly followed two routes. The first one was through northern France to Germany and then Belgium, and the second across the Ligurian coast through present-day Slovenia to Austria, toward the northern Balkans. Here we present a combined approach to reveal glacial refugia and post-glacial spread of *Ph. mascittii* and observed discrepancies between the modelled and the current known distribution might reveal yet overlooked populations and potential further spread.

**P**hlebotomine sand flies (Diptera: Psychodidae: Phlebotominae) are small hematophagous insects inhabiting tropical, subtropical, and temperate regions. They are of significant medical and veterinary relevance as vectors of *Leishmania* spp. protozoans, bacteria, and several arboviruses capable of infecting humans in various regions of the Old and New Worlds<sup>1</sup>. Human leishmaniasis are among the top ten neglected tropical diseases globally, with a burden of approximately 50,000 deaths per year and one billion people at risk of infection<sup>2</sup>. Leishmaniasis caused by *Leishmania infantum* occurs predominantly in the tropics and subtropics, including the Mediterranean area<sup>3</sup>, and its incidence is expected to be underestimated in most countries in the WHO European Region<sup>4</sup>. This region, as well as Israel, Turkey, Turkmenistan, and Uzbekistan, are the most affected countries and account for almost 80% of the total number of cases reported in the region. Moreover, the Mediterranean North African Maghreb region is highly endemic for cutaneous leishmaniasis (CL), with Algeria being the second most affected country worldwide. Moreover, leishmaniasis also occurs in the Balkans as well as the southern Caucasus and Central Asia<sup>2,5</sup>.

In Afro-Eurasia, various *Leishmania* species are transmitted to mammals by *Phlebotomus* sand flies. While the occurrence of sand flies and the endemicity of leishmaniasis in Mediterranean Europe have been known for decades, the occurrence of sand flies in the more continental territories (particularly Central Europe) has been less understood; however, recent findings record their emergence or sporadic and potentially overlooked presence<sup>6</sup>. In Central Europe, several sand fly species focus the attention of epidemiologists, namely *Phlebotomus (Larroussius) perniciosus* Newstead, 1911, *Ph. (L.) neglectus* Tonnoir, 1921, *Ph. (L.) perfiliwei* Parrot, 1930, *Ph. (L.) tobbsi* Adler & Theodor, 1930, *Ph. (L.) ariasi* Tonnoir, 1921, and *Ph. (Phlebotomus) papatasi* (Scopoli, 1786) as vectors for *Leishmania* spp. and/or for phleboviruses<sup>7–11</sup>. Additionally, *Ph. (Transphlebotomus) mascittii* Grassi, 1908, and *Ph. (Adlerius) simici* Nitzulescu, 1931, are suspected but unproved vector species<sup>12–15</sup>.

*Phlebotomus mascittii* has the widest and northernmost distribution in Europe (Fig. 1). It has been reported from many Mediterranean countries of Europe, including the western Mediterranean countries of Spain<sup>16,17</sup>, France<sup>18,19</sup> including Corsica<sup>20</sup> and Italy<sup>21–23</sup>, several Balkan countries of ex-Yugoslavia<sup>24–28</sup>, and historically also from eastern Mediterranean countries of Greece and Turkey, even though these records shall be reviewed with respect to later descriptions of new species of the subgenus *Transphlebotomus* from this region<sup>29</sup>. The presence, however, is not restricted to the Mediterranean region and reaches into higher latitudes; the northernmost record is known from Germany<sup>30</sup>. Despite some recent records from Austria<sup>31–33</sup>, Germany<sup>34</sup>, Slovakia<sup>35</sup>, and Hungary<sup>36</sup>, its exact distribution in Central Europe is still largely unknown. The species distribution outside Europe is limited to a single finding in northern Algeria that represents its southernmost record<sup>37</sup>.

Data about the capacity of sand fly species to adapt to environmental changes allow us to predict changes in sand fly distribution<sup>38</sup>. For instance, previous studies allowed to increase scientific knowledge about the paleoogeographical history of *Larroussius*<sup>39</sup>, *Paraphlebotomus*<sup>40–42</sup> and *Phlebotomus*<sup>43,44</sup> subgenera. Despite being one of the most widespread sand fly species in Europe, an accurate appraisal of the distribution and dispersal patterns of *Ph. mascittii* has yet to be conducted. It is known that glacial-interglacial climatic changes significantly impacted sand fly distributional ranges. During the last glacial period, which ended about 10,000 to 12,000 years ago, vast parts of Central and Eastern Europe were covered by permafrost and



**Fig. 1** Described records of *Phlebotomus mascittii* used for modeling analyses.

Occurrence data in Sardinia and North Macedonia were not included in the analyses and were georeferenced to the center of the country due to a lack of point data. Molecularly analyzed specimens from this study are included in the map; details are given in Supplementary Data 1. Country border data is freely available at TM WORLD BORDERS 0.3 (URL: <https://koordinates.com/layer/7354-tm-world-borders-03/>) and elevation data at ETOPO Global Relief Model (URL: <https://www.ncei.noaa.gov/products/etopo-global-relief-model>).

tundra, resulting in a considerable decrease in pre-glacial biodiversity and probably in the total disappearance of sand fly taxa from the continental regions of Europe<sup>45</sup>. Thus, being typical Mediterranean insects, sand flies either became extinct or survived in (extra-) Mediterranean refugia<sup>46,47</sup>. The coldest periods drove sand flies and their pathogens to the southern Mediterranean coastlines<sup>45</sup>. Due to the climate-ameliorating effect of the Mediterranean Sea, some sand fly populations also survived on the Mediterranean islands, which were colonized during the Messinian salinity crisis in the latest Miocene epoch by sand flies<sup>29,39,41,48</sup>. For example, genetic evidence suggests that some species, such as *Ph. ariasi* Tonnoir, 1921, had a refugium north of the Pyrenees during the Last Glacial Period<sup>49</sup>.

Intensive surveys over the last decades have revealed that sand fly species occur in several, sometimes geographically distant regions in Europe, e.g., southern Belgium and northern Germany<sup>18,30</sup>. After the last glacial period, excluding the short but significant cooling of the Younger Dryas stadial<sup>50</sup>, the onset of the Holocene was marked by a fast temperature rise and rapid transgression<sup>51</sup>. Two warm periods around 6500 and 4500 years ago (the Holocene optima) were characterized by mean temperatures similar to current climatic conditions. During these times, numerous animal and plant species from southern refugia colonized more northern regions, including Central Europe<sup>47,52–54</sup>. Also, sand flies (re-)colonized northern areas, established, and survived in small stable populations in microclimatically favorable areas until today<sup>45,47,55</sup>. Species distribution models suggest that sand fly populations of different Mediterranean species could have occurred as far north as the British Isles during the Holocene optima<sup>45</sup>. The migration routes of sand flies to central parts of Europe are, however, complex and still poorly understood.

The present study aimed to elucidate the effect of climatic and sea level changes on the expansion of *Ph. mascittii* from its glacial refugia within the last 17 kiloyears (kys). Here, we used an integrated approach linking phylogeographic data, climate modeling, and network analysis to unlock new insights into the process of post-glacial sand fly colonization of Central Europe.

## Results

**Patterns of genetic diversity.** Altogether, 92 COI and 95 Cytb sequences of *Ph. mascittii* specimens were included in the genetic analyses (Supplementary Data 1). COI sequences without gaps and stop codons (Supplementary Data 2) with a length of 626 bp clustered into four haplotypes with three variable sites. The haplotype diversity (Hd) was 0.4190, and the nucleotide diversity ( $\pi$ ) was 0.00074. One major haplotype (COI\_2) was shared among 69 of 92 sequences from all included countries/territories except Corsica Island, with a distinct and unique haplotype (COI\_4). While haplotype 1 (COI\_1) consisted of specimens from Austria and Serbia, haplotype 3 (COI\_3) only included sequences originating from Austrian representatives (Fig. 2a). The Cytb sequences with a total length of 641 bp also showed no gaps or stop codons (Supplementary Data 3). In total, nine haplotypes were identified, having eight variable sites, of which one was parsimony informative. The haplotype diversity (Hd) was 0.2708, and the nucleotide diversity ( $\pi$ ) was 0.0006. Like in the COI haplotype network, one major haplotype was observed for Cytb, i.e., Cytb\_1, including 81 of 95 the analyzed sequences, which originated from all included countries. Again, sequences of Corsican specimens clustered within a distinct haplotype (Cytb\_3), except for one sequence that belonged to Cytb\_1. Unlike the COI network, several singletons (Cytb\_2, Cytb\_4–Cytb\_8) were observed in the Cytb network (Fig. 2b).

For COI haplotypes, a mean pairwise distance of 0.24% (range: 0.16–0.32%) was observed thereby (Supplementary Table 1), and Cytb haplotypes showed an overall mean pairwise distance of 0.3% (range: 0.16–0.49%) (Supplementary Table 2).

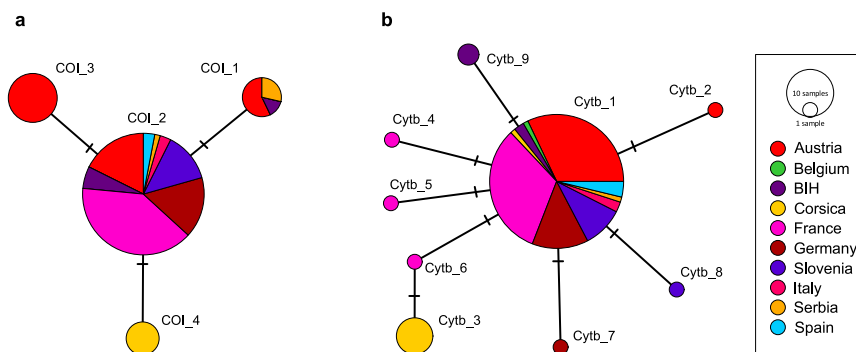
The star-shaped haplotype networks (Fig. 2), together with a significantly negative Tajima's D ( $D = -1.655$ ;  $p = 0.018$ ) and a unimodal mismatch distribution (Fig. 3a) with non-significant raggedness index ( $rg = 0.097$ ;  $p = 0.711$ ) but a marginally non-significant sum of squared differences ( $SSD = 0.004$ ;  $p = 0.064$ ), indicate a recent population expansion. Also, the Bayesian Skyline plot provided clear evidence for strong recent post-glacial population expansion (Fig. 3b). The time to the most recent common ancestor was estimated at 42.97 KY (95% highest posterior density interval 8.02–86.35 KY) to 107.42 KY (95% highest posterior density interval 20.05–215.87 KY), depending on the assumed substitution rate. No evidence for isolation by distance was present in our data (Supplementary Fig. 1, Supplementary Data 4).

**Modeled alterations of climatic suitability patterns and former ranges.** The model shows that areas with high climatic suitability in Heinrich Stadial 1 existed along the currently partly submerged coasts of the Gulf of Lyon, as well as on the narrow coastline of the Ligurian Sea, as well as in the south-central

regions of the Apennine Peninsula. Compared to the previous period, the high-suitability regions of the Bølling–Allerød Interstadial show a larger extension in present-day France, including territories adjacent to the Bay of Biscay and the Gulf of Lyon. However, for the Younger Dryas Stadial, the suitability patterns could be like those in the Heinrich Stadial 1 period. Although in the Greenlandian (early Holocene) period, the climatically suitable areas exhibit a more minor increase compared to the late Pleistocene conditions. From the Northgrippian (mid-Holocene) period onwards, the extension of the climatically suitable regions began to expand rapidly, covering large areas of continental western Europe and the Apennine Peninsula. This rapid expansion in regions with high climatic suitability was continuous between the Northgrippian and the Anthropocene and led to the colonization of Central Europe (Fig. 4).

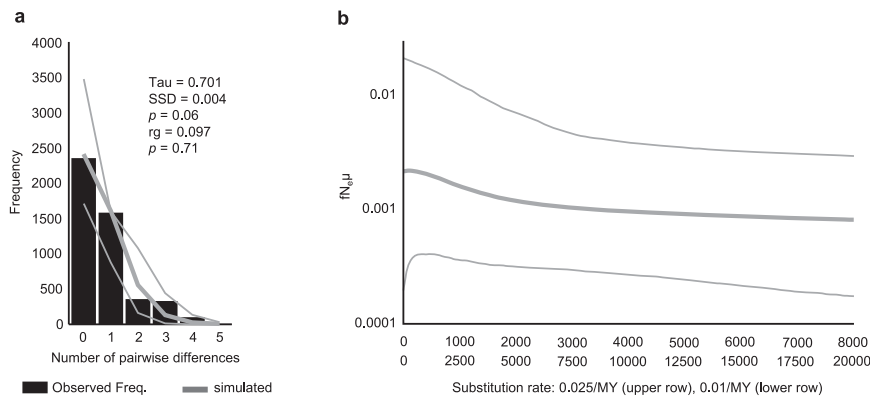
The model results indicate that the occurrence of *Ph. mascittii* in the late Pleistocene era could have been restricted to relatively small regions of the Mediterranean coasts of Southwest Europe and the Apennine Peninsula. Between 17 and 11.7 ka, the patterns of the potential range of the species increased, but the significant expansion of the species only started after the Greenlandian period, in the early mid-Holocene era. Parallel to this areal expansion, the species could have lost a notable area of their glacial refugia due to the post-glacial transgression. Based on the modeled ranges, a relatively dispersed Southwest European along the Mediterranean coasts and an Apennine Peninsula glacial refugia can be hypothesized. Among the sampling sites of this study, only the sites in Southwest France can be found within the Southwest European glacial refugium of the species (Fig. 5).

**Network analysis results.** Considering the percental frequency (%) of the number of periods with climatic suitability for *Ph. mascittii*, when the local climatic suitability overwhelmed 90% in a site, the Mediterranean coasts of South France could have been a stable refugia over the past 17 kys. The geographical trend of the climatic suitability values suggests that the post-glacial migration of *Ph. mascittii* had an eastward character. The relative centrality values show that the Gulf of Lyon and the Swiss Middle Mountain region could have played a key role in the eastward migration of the species in the post-glacial era (Fig. 6). The network analysis results indicate that the peripheral region of this refugium likely played the most significant role as a source of the spread of the species, while Southwest France represents the climatically most stable glacial refugial core region (Fig. 7). The result of network similarity hierarchical clustering denotes a strong connection between the two genes ( $r^2 = 0.31$ ) and between climatic and spatial factors ( $r^2 = 0.21$ ). A weak correlation was observed between genetic and climatic as well as spatial factors, and the



**Fig. 2** Statistical parsimony networks of *Phlebotomus mascittii* from eight countries. Analysis based on COI (a) and Cytb (b) sequences. BIH Bosnia and Herzegovina.





**Fig. 3 Signatures of population expansion in *Phlebotomus mascittii*.** **a** Mismatch distribution for *Ph. mascittii* in Europe, based on combined COI and Cytb sequences. Black columns represent the observed frequency of pairwise differences. Gray lines refer to the expected distribution based on parameter estimates and their 95% confidence limits simulated under a model of population growth. The sum of squared differences (SSD) and raggedness index (rg) and their respective  $P$  values are given to describe the fit of the observed mismatch distribution to the expectation based on growth parameter estimates. **b** Bayesian Skyline Plot (BSP), assuming a minimum and maximum substitution rate for mitochondrial protein-coding genes of insects of 1% and 2.5% per. The thick line denotes the median estimate; thin lines indicate the 95% highest posterior density (HPD) interval.  $fN_e\mu$  = female effective population size scaled by substitution rate. Raw data of the analyses are given in Supplementary Data 2.

phylogeographic patterns were not supported by the sites' general climatic suitability or the points' distance (Supplementary Fig. 2).

## Discussion

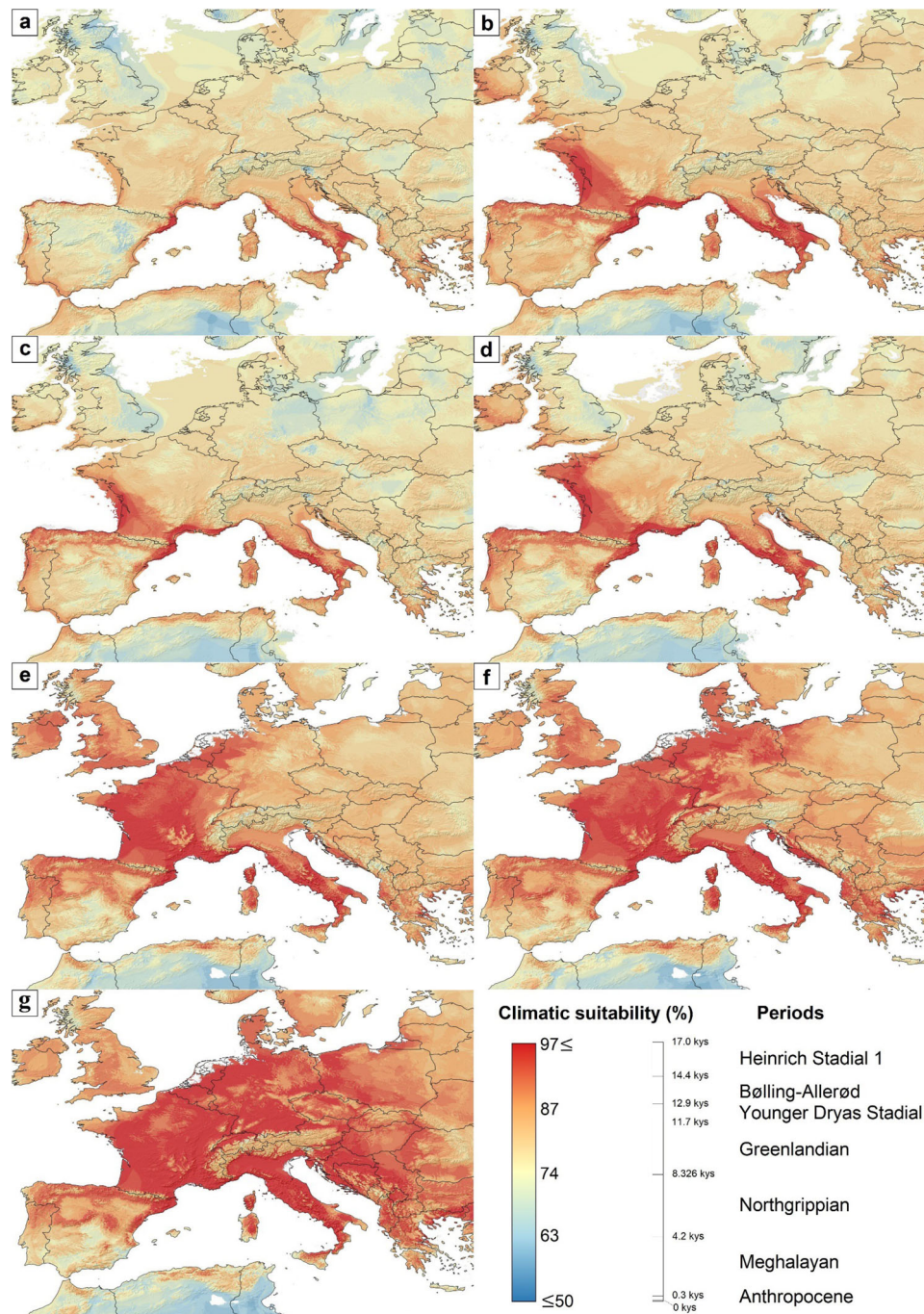
Using an integrated approach that encompasses phylogeographic data, climate modeling, and network analysis, we unlocked insights into the post-glacial expansion of *Ph. mascittii* in Europe and investigated its refugia and colonization routes. The molecular results suggest that *Ph. mascittii* populations experienced a notable genetic bottleneck before their post-glacial expansion from a southwestern European refugium. A major glacial refugium was identified in southern France by comparing the genetic data and modeled former climatic patterns. Another potential Apennines refugial area might be assumed but needs further clarification with currently unavailable data. The rapid post-glacial expansion of *Ph. mascittii* started in the early mid-Holocene era, and it might have reached its current range during the Meghalayan stage. The Southwest European region north of the Pyrenees likely played a key role in the start of the post-glacial migration of the species; however, the survival of *Ph. mascittii* during cold periods was due to the Southern France refugium. A partly similar region is assumed to have served as the most important glacial refugium for *Ph. ariasi*<sup>49</sup>.

Notably, the currently submerged epicontinental shelf areas, which were drylands during the Last Glacial Maximum, could have played an important role in the survival of sand fly species. This fact is most visible, e.g., in the case of the Gulf of Venice in the North Adriatic region. Despite little knowledge of the glacial paleoenvironmental conditions of these regions, some authors suggest that *Quercus* species, which formed deciduous forests, returned to West and Central Europe in the postglacial era<sup>56</sup>. However, recent investigations showed that temperate seasonal forest habitats were widely present in the peri-Mediterranean area even in the late glacial era<sup>57</sup>. Currently, *Ph. mascittii* is known to occur in deciduous forests in Central Europe, e.g., in the Kaposcs Valley, Hungary<sup>58</sup>, in close vicinity to a deciduous forest dominated by *Quercus cerris* L. Trájer<sup>57</sup> showed that the southern regions of France during the Bølling-Allerød period could have supported temperate seasonal forests, which would have been a suitable habitat for *Ph. mascittii*. This is noteworthy, considering that the (now submerged) coastal plains could have provided a tempered and relatively humid environment for *Ph. mascittii* populations—environmental conditions in which this sand fly

species thrives<sup>59</sup>. We, therefore, hypothesize that *Ph. mascittii* populations could have survived the Last Glacial Maximum cold in their South European refugia in deciduous forest habitats, and their postglacial expansion was linked to the northward spread of the temperate seasonal forests.

The recent post-glacial expansion of *Ph. mascittii* is well supported by the observed star-shaped haplotype networks and the unimodal mismatch distribution. Similar processes have been hypothesized for *Ph. papatasi*<sup>60</sup> and *Ph. perniciosus*<sup>61</sup>. Also, the inferred Skyline Plot shows a recent population expansion, particularly in the last 10,000 to 5000 years, which aligns well with our climate modeling results. Similar to the post-glacial migration patterns of *Ph. perniciosus*<sup>62</sup>, we did not observe evidence of genetic isolation by distance due to the short analyzed timespan. The weak correlation between the two genetic, climatic, and spatial factors might be best explained by the lack of recombination in the mitochondrial genome and the linkage between the two analyzed loci. Hence, the low correlation between the two loci in this dataset has a stochastic effect, namely a very recently analyzed time scale in which only a few substitutions have occurred. In other possible cases, e.g., with the multiple previous refugia, at least the spatial factors would show a moderate or more substantial correlation with the genetic traits. We assume that small populations of sand flies migrated to the east after the post-glacial era from the Southwest Ice Age refugium in several waves during the last 6 kys, as also demonstrated for other insect taxa, e.g., grasshoppers (Orthoptera)<sup>63,64</sup>.

Based on our combined phylogeographic and modeling approaches, we hypothesize two potential, possibly parallel, dispersal routes from a southwestern European refugium: firstly, through the Côte d'Azur, the Po Valley, and the Slovenian mountain passes; and/or secondly, a northern route via Northeast France and South Germany. Geographical and climatic conditions must be considered regarding potential post-glacial migration routes, starting in the Pyrenees region. After the mid-Holocene warming, the main barriers to the west-east spread were the Alps in Central Europe, the Apennines in the Mediterranean area, and the Dinaric mountains. According to the model results, *Ph. mascittii* populations could have avoided the Alps by the northern route and reached the Carpathian Basin. Considering the southern way, crossing the northern Dinaric areas could have occurred via passes at lower altitudes. However, anthropogenic factors must be considered. Human farmer populations reached mainland Europe



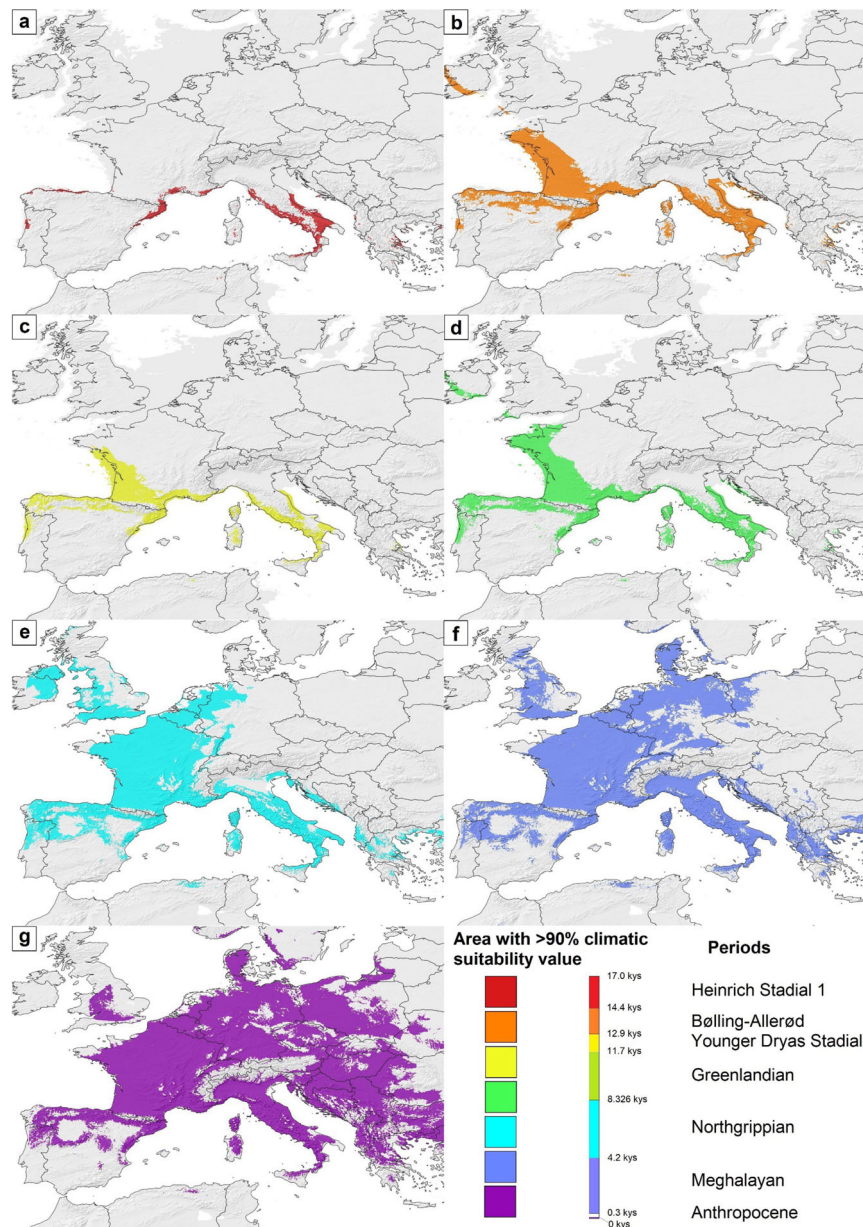
**Fig. 4** The changes in climatic suitability for *Phlebotomus mascittii* over the last 19 kys. **a** Heinrich Stadial 1 (17.0–14.7 kys), **b** Bølling-Allerød (14.7–12.9 kys), **c** Younger Dryas Stadial (12.9–11.7 kys), **d** Greenlandian (11.7–8.326 kys), **e** Northgrippian (8.326–4.2 kys), **f** Meghalayan (4.2–0.3 kys), and **g** Anthropocene (1979–2013). Modeling results were georeferenced in paleoclimatic maps, openly available at PaleoClim.org (URL: <http://www.paleoclim.org/>) and WorldClim (URL: <https://www.worldclim.org/>).

roughly in 6700 Before Christian Era (BCE), the broader environment of the Pyrenees in 5800 BCE, and from 5200 BCE onwards, both the northern and southern forelands of the Alps and the Dinaric mountains were populated by Neolithic societies<sup>65</sup>. Overall, the abovementioned factors may have contributed to the colonization and spread of *Ph. mascittii* in Europe. Noteworthy, the currently observed distribution shows marked discrepancies from the current modeled climatic suitability of this species. This might be explained by the fact that *Ph. mascittii* has been reported to favor urban and semi-urban breeding sites such

as animal farms, old barns with natural floors or even cemeteries and is, thus, mostly restricted to these sites<sup>16,33</sup>.

To uncover additional potential migration routes, it will be necessary to carefully study currently unavailable specimens from Italy and the northern Balkans to understand the role of potential refugia in the Apennine Peninsula and the Balkans. The hypothesized southwest-to-northeast or south-to-north direction of post-glacial migration trends is in agreement with observed diversity patterns in other cold-sensitive animal groups like reptiles and amphibians<sup>66</sup>.



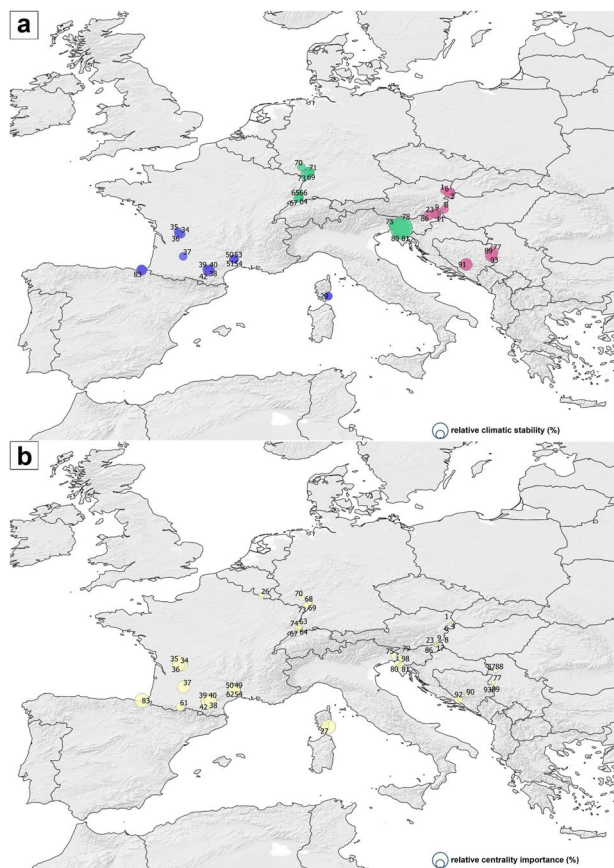


**Fig. 5** The changing expansiveness of areas with more than 90% climatic suitability values for *Phlebotomus mascittii* in the last 17 kys. **a** Heinrich Stadial 1 (17.0–14.7 kys), **b** Bølling-Allerød (14.7–12.9 kys), **c** Younger Dryas Stadial (12.9–11.7 kys), **d** Greenlandian (11.7–8.326 kys), **e** Northgrippian (8.326–4.2 kys), **f** Meghalayan (4.2–0.3 kys), and **(g)** Anthropocene (1979–2013). Modeling results were georeferenced in paleoclimatic maps, openly available at PaleoClim.org (URL: <http://www.paleoclim.org/>) and WorldClim (URL: <https://www.worldclim.org/>).

The results herein presented also indicate a post-glacial dispersal of *Ph. mascittii* to Corsica. The present-day distance between the nearest dryland points of Corsica and the Apennine Peninsula is about 81 km, and between the islands of Corsica and Elba is approximately 50 km. Although these distances were shorter in the Last Glacial Maximum due to the low sea level, this still means a considerable distance for sand flies due to their frequently publicized weak flying capability. For example, the maximal travel distances recorded for *Ph. papatasi* females and males were 1.91 and 1.51 km, respectively<sup>67</sup>. However, Tonelli et al.<sup>68</sup> showed that different sand fly species may show different flight abilities and might reach greater distances than previously reported. Thus, this mode of natural dispersal, particularly involving wind, should at least be considered, although never observed for sand flies and controversially discussed. Sand flies are tiny insects, very sensitive to low air humidity (and the

consequent desiccation), large temperature variations, and intensive solar radiation<sup>69</sup>. Thus, since these conditions prevail during transport by wind over the sea, it remains uncertain if these insects could have survived such a dispersal mode. However, the dispersal from the mainland to Corsica could also have been facilitated by human activities. Humans settled in Corsica around 6000 BCE<sup>70</sup> and the seafaring population brought sheep, goats, and pigs to the island<sup>71</sup>, which implicates the transport of litter as well as bulks of sand or soil below deck. This mode of dispersal has also been discussed for the island of Santorini, Greece<sup>72</sup>, and for *Sergentomyia clydei* (Sinton, 1928) to the Seychelles, and *Sergentomyia babu* (Annandale, 1910) to Mauritius<sup>73</sup>.

This study provides to the best of our knowledge the first detailed insight into the most likely post-glacial dispersal of *Ph. mascittii*, the currently most widely distributed sand fly species in Europe. Our approach, combining phylogeographic, climatic, and



**Fig. 6** Relative climatic and relative centrality measures of the sampled *Phlebotomus mascittii* sites. The diameter of the circles represents the relative climatic suitability values (%) (a) and centrality importance (%) (b). The numbers are in accordance with the site numbers sorted in Supplementary Data 1. Network results were georeferenced in paleoclimatic maps, openly available at PaleoClim.org (URL: <http://www.paleoclim.org/>) and WorldClim (URL: <https://www.worldclim.org/>).

networking data, revealed a crucial glacial refugia in southwestern France. Based on the results of this study, we hypothesize a recent post-glacial dispersal of *Ph. mascittii* by two different routes. Elucidating post-glacial sand fly migration patterns to Central Europe might help to predict future sand fly dispersal and that of pathogens potentially transmitted by them.

## Material and methods

**Sampling and morphological sand fly identification.** For this study, 38 complete specimens and 48 DNA samples of *Ph. mascittii* obtained during various entomological surveys, including published and yet unpublished material, were used. In addition, the dataset was complemented with 14 previously published *Ph. mascittii* sequences available in the GenBank database. The head and terminal segments of the abdomen of all unidentified specimens were dissected and slide-mounted on a glass slide in CMCP-10 high-viscosity mountant (Polysciences Europe GmbH, Hirschberg an der Bergstraße, Germany). Identification was based on the morphological parameters of the male genitalia, female spermatheca, and pharyngeal armature<sup>74</sup>.

**DNA extraction, PCR amplification and sequencing.** The remaining body parts of slide-mounted specimens were transferred into individual vials (1.5 mL), 180  $\mu$ L ATL buffer and 20  $\mu$ L Proteinase K were added, and homogenization was performed with a hand homogenizer (Thomas Scientific, Svedesboro, NJ, USA).

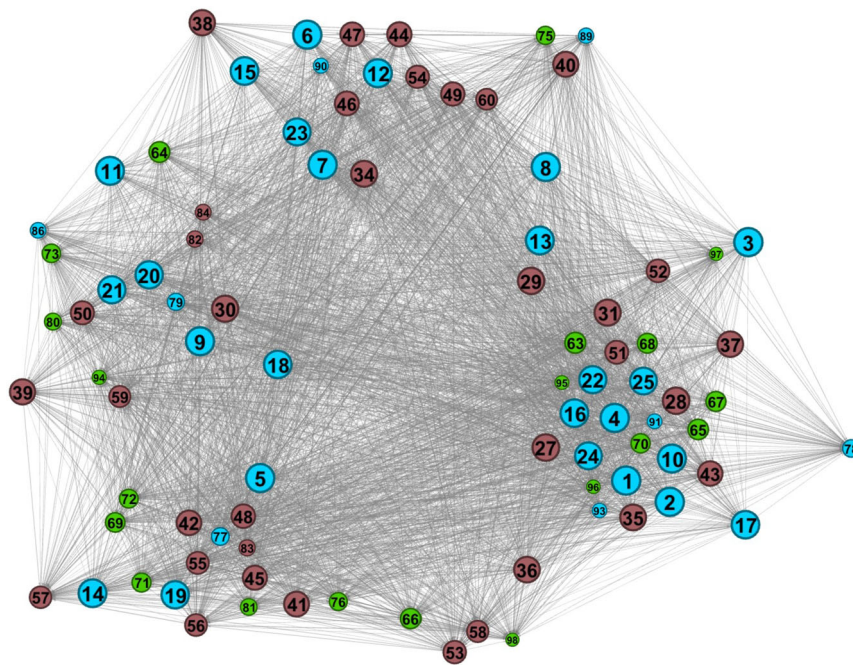
DNA isolation was performed with a QIAamp® DNA Mini Kit 250 (Qiagen, Hilden, Germany), following the manufacturer's protocol. DNA was eluted in a final volume of 100–200  $\mu$ L of elution buffer. Sand flies from Germany were homogenized in 180  $\mu$ L of Dulbecco's Modified Eagle Medium (Sigma Aldrich, St. Louis, Missouri, United States (US)), 100  $\mu$ L homogenate was used for DNA extraction using the 5x MagMax Pathogen RNA/DNA Kit (Thermo Fisher Scientific, Waltham, Massachusetts, US), DNA was eluted in 80  $\mu$ L. Partial sequences of two mitochondrial genes were amplified by PCR, namely cytochrome c oxidase subunit 1 (COI) and cytochrome b (Cytb). For the COI, a 658-bp fragment was amplified using the primers LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA-3')<sup>75</sup> and for Cytb, a 737-bp fragment was amplified using the primers CB1-SE (5'-TATGTACTACCTGAGGACAAATATC-3') and CB-R06 (5'-TATCTAATGGTTTCAAAACAATTGC-3')<sup>76</sup>. All PCR amplifications were run with the following PCR conditions: 95 °C for 15 min, followed by 35 cycles of 95 °C for 1 min (denaturation), 52 °C for 1:30 min (annealing), and 72 °C for 2 min (elongation), followed by a final extension of 72 °C for 10 min. PCR amplification was performed in reaction volumes of 50  $\mu$ L containing 10 $\times$  reaction buffer B, 2.5 mM MgCl<sub>2</sub>, 1.6 mM dNTPs, 1  $\mu$ M primers, 1.25 units DNA polymerase, and 1–5  $\mu$ L DNA; sterile H<sub>2</sub>O was added to the final volume of 50  $\mu$ L. Thermal cycling was performed with an Eppendorf Mastercycler modular PCR system (Eppendorf AG, TM, Hamburg, Germany).

Bands were visualized and analyzed with a Gel Doc™ XR+ Imager (Bio-Rad Laboratories, Inc., Hercules, CA, USA), cut out of the gel, and purified with an Illustra™ GFX™ PCR DNA and Gel Purification kit (GE Healthcare, Buckinghamshire, UK). Sanger sequencing was performed using a BigDye® Terminator v.1.1 Cycle Sequencing kit (Thermo Fisher Scientific Inc, Waltham, MA, USA) and run on a SeqStudio® Genetic Analyzer (Thermo Fisher Scientific Inc., Waltham, MA, USA).

**Sequence editing and alignment.** Overall, 86 new COI sequences with a length of 658 bp and 87 new Cytb sequences with a length of 748 bp of individual *Ph. mascittii* specimens were obtained from both strands, and aligned and consensus sequences were generated using the DNA sequence analysis tool GeneDoc 2.7.0<sup>77</sup>. DNA chromatogram files were checked for double signals, if necessary. Translation to amino acid sequences of coding regions (COI and Cytb) showed intact reading frames for all included sequences, and no internal stop codons were observed (Supplementary Data 3, 4). Sequence identities were checked by comparing the obtained sequences to available sequences in GenBank using BLAST. Sequences were submitted to GenBank and are available under the following accession numbers: OQ064321.1–OQ064395.1 and OR573677.1–OR573687.1 for COI and OQ067398.1–OQ067475.1 and OR574850.1–OR574858.1 for Cytb (Supplementary Data 1). In addition, six COI sequences and eight Cytb sequences of *Ph. mascittii* were downloaded from GenBank and included in the analysis (Supplementary Data 1). For each locus, sequences were aligned using ClustalX<sup>78</sup>.

**Genetic diversity through space and time.** Unique haplotypes, haplotype (Hd), and nucleotide ( $\pi$ ) diversity per locus were calculated with DnaSP v.5<sup>79</sup>, and uncorrected pairwise distances between all haplotypes per locus were calculated in MEGAX<sup>80</sup>. Phylogenetic relationships among haplotypes were visualized as a statistical parsimony network<sup>81</sup>, as inferred in PopART<sup>82</sup>. To test for signals of population expansion, we calculated Tajima's D<sup>83</sup> (1000 simulated samples) and a mismatch distribution (1000 bootstrap replicates) in Arlequin v3.11<sup>84</sup> based on the





**Fig. 7 Results of the network analysis of the sampled sites.** Analysis based on the geographical, cytochrome c oxidase subunit I (COI) and cytochrome b (Cytb) sequence distances and the climatic suitability differences of the *Phlebotomus mascittii* sites. The size of the nodes represents their relative importance in the network based on the analyzed aspects of the sites. The numbers follow the site numbers given in Supplementary Data 1.

concatenated dataset, including both loci. The fit between the observed mismatch distribution and the expectations based on growth parameter estimates was evaluated by the sum of squared differences (SSD) and the raggedness index (rg). Furthermore, past population size trajectories and time to the most recent common ancestor (tMRCA) were inferred using a Bayesian coalescent approach in BEAST 2.7.1<sup>85</sup>. Due to the very low levels of variation observed within each locus, we refrained from partitioning the data per gene or codon position. We employed the best-fitting model of molecular evolution<sup>86</sup> selected by the Bayesian Information Criterion in ModelFinder<sup>87</sup>, a strict molecular clock with a minimum and maximum substitution rate of 0.01 and 0.025 substitutions per site per million years (MY)<sup>8,88</sup>, which covers the range of substitution rates typically inferred from / employed for mitochondrial protein coding genes in insects<sup>60,89–91</sup>, and the Bayesian Skyline tree prior. Four independent MCMC runs of ten million generations each were done, sampling every 1000th step and a burn-in of the first 10% of sampled generations. LogCombiner (part of the BEAST package) was used to combine the individual log and tree files. Assessment of run convergence, verification of effective sample sizes [ESS > 200 for all parameters, indicating that the parameter log file accurately reflects the posterior distribution<sup>92</sup>], and visualization of past population size changes were done in Tracer 1.7<sup>93</sup>. The Mantel test, as implemented in Alleles In Space 1.0<sup>94</sup>, was used to test for a correlation between the genetic and geographical distances of the sampled individuals.

**Climatic and topographic data sources.** Four Holocene and three late Pleistocene models were applied in the modeling process. As a reference (current Anthropocene: 1979–2013) period, the model of Karger et al.<sup>95</sup> was used. Three other Holocene periods, the Greenlandian, Northgrippian, and Meghalayan periods, were also used. The Heinrich Stadial 1, Bølling-Allerød interstadial, and Younger Dryas Stadial periods' models represented the late Pleistocene (Table 1).

**Table 1 The sources of the used climatic models and data. The values of the Anthropocene climate model were used for the derivation of the range-limiting climatic extrema of sand flies.**

Model period	Age	Reference
Current period (Anthropocene)	1979–2013	Karger et al. <sup>95</sup>
late-Holocene, Meghalayan	4.2–0.3 ka	Fordham et al. <sup>105</sup>
mid-Holocene, Northgrippian	8.326–4.2 ka	Fordham et al. <sup>105</sup>
early-Holocene, Greenlandian	11.7–8.326 ka	Fordham et al. <sup>105</sup>
Younger Dryas Stadial	12.9–11.7 ka	Fordham et al. <sup>105</sup>
Bølling-Allerød	14.7–12.9 ka	Fordham et al. <sup>105</sup>
Heinrich Stadial 1	17.0–14.7 ka	Fordham et al. <sup>105</sup>

**Geographical data and mapping.** A literature search for records of *Ph. mascittii* species was performed (Google Scholar and PubMed), and coordinates or locations were extracted. Published as well as yet unpublished *Ph. mascittii* trapping sites were georeferenced into a distribution map using Quantum GIS 3.4.11<sup>96</sup> (Fig. 1). For most of the occurrences, the exact coordinates were known (coordinates were known for all molecularly analyzed specimens). In the remaining cases, the location of the original trappings has been defined as accurately as possible. In some cases, the exact coordinates of the catch sites could be reconstructed. By identifying the geographical position of the streets and individual houses within the small settlements, the coordinate extraction was possible within a resolution accuracy of ten meters. In the case of other small settlements where a schematic map was not provided, the center coordinates of the villages were considered in georeferencing. Sometimes, the resolution was between  $1 \times 1 <$  to  $\leq 10 \times 10$  km. Coordinate accuracy was ensured by a general resolution that ranged from  $1 \times 1 <$  to  $\leq 10 \times 10$  km. Of all occurrence data points, 102 were exact point coordinates, 44 were at a scale of  $1 \times 1$  km or lower, and eleven were at  $10 \times 10$  km or lower. From the point of view of distribution modeling, even the accuracy of  $\leq 10 \times 10$  km resolution

is generally acceptable because these model experiments aim to predict large-scale occurrence patterns and not local habitats. Furthermore, a cut of 2–2 percentiles from the sampled climatic data was utilized to filter the possibly existing non-relevant climatic values. Finally, it is worth mentioning that the resolution of the used paleoclimate models is 2.5 arcminutes (in the mid-latitudes, it is equal to ~5 km) which means that the accuracy of the site coordinates fits the resolution of the climatic models. Supplementary Data 1 shows the essential data relevant to the studied sites.

**Modeling of climatic suitability patterns.** For the acquisition of the range-limiting extrema of *Ph. mascittii*, the 2.5 arcmin resolution model of Karger et al. (2021) related to the reference period of 1979–2013 was used. A total of 19 bioclimatic variables<sup>97</sup> were utilized among which 11 (bio1–11) have temperature and 8 (bio12–19) have precipitation nature (Supplementary Table 3). To gain data on distribution-limiting climatic extrema, 2–2 percentiles were cut from the absolute maximum and minimum values of the factors to avoid the involvement of unrealistic climatic contracts according to the generally applied considerations of environmental modeling<sup>98</sup>. The climatic values used in modeling are available in Supplementary Table 3.

The modeling of the former distribution areas followed the logic of Boolean algebra<sup>99</sup>. This technique is based on climate envelope or multidimensional modeling<sup>100</sup>, which considers independently the effects of environmental factors on the occurrence probability of species. Then, it estimates distribution areas or species-related climatic suitability values regarding its recorded occurrences. Using the climatic hypervolume contained by the minimum and maximum values of climatic variables within the native range of the researched species, a Boolean map of risk zones is produced in this procedure. The theoretical basis of this modeling technique is fundamentally based on the barrel analogy of Liebig's law of the minimum<sup>101</sup>.

To generate the results, the raster calculator function of Quantum GIS 3.4.11<sup>96</sup> was used. Supplementary Data 5 shows the script of the utilized equation.

The area of all range areas was modeled according to the following general equations:

$$1(v_n) = \begin{cases} 0 & \text{if } v_{n\_limit\_min} > v_n \text{ and } v_n > v_{n\_limit\_max} \\ 1 & \text{if } v_{n\_limit\_min} \leq v_n \text{ and } v_n \leq v_{n\_limit\_max} \end{cases}$$

Where  $v_n$  represents the  $n$ th climatic constraint of the distribution area of a species,  $v_{n\_limit\_min}$  and  $v_{n\_limit\_max}$  are the lower and upper distribution-limiting values related to the climatic constraint (Supplementary Table 4).

The potential area-based suitability patterns were determined according to the following mathematical formalism:

$$A(v_1; v_2; \dots; v_n) = 1(v_1) - v_1 \cap 1(v_2) - v_2 \cap \dots \cap 1(v_n) - v_n$$

where  $A(v_1; v_2; \dots; v_n)$  shows the potential distribution area of the given species, which contains the remaining areas after considering the factor-related limitations.

Then, the modeled values were transformed into percentage (%) values and colorized.

**Network-based evaluation of the spatial similarities of modeled climatic suitability and genetic patterns.** The sampling points are considered a network, where  $N = 88$  represents the nodes connected in an  $A = N \times N$  adjacency matrix. The  $A$  adjacency matrix is symmetric due to the undirectedness of the defined network. The edge from node  $j$  to node  $i$  can be expressed as  $e_{ij}$ . Networks are defined for each examined variable, so the following

multilayer  $M$  network can be written<sup>102</sup>:

$$M = \left( COI, Cytb, d, \sum_{a=1}^7 S_a \right)$$

where  $M$  represents the multilayer network, COI and Cytb stand for the identified genetic haplotypes,  $d$  denotes the coordinates, and  $S_a$  refers to the modeled climatic suitability values in the given  $a$  model period. Points with missing genetic information ( $n = 10$ ) were excluded from the network analysis (Supplementary Data 1). Network analysis was performed in Gephi 0.10.1<sup>103</sup>.

The genetic distance values are given for all ( $e_{ij}$ ) edges, and the geographic distance ( $d$ ) can be calculated from the coordinates. The climatic suitability values can be obtained from the difference in the cumulated  $S_a$  value of the two nodes, so it is used to express a degree of climatic similarity. Finally, integrating the effects of the variables, an edge with multiple weights is introduced as follows:

$$w_{e_{ij}} = dCOI_{e_{ij}} + dCytb_{e_{ij}} + d_{e_{ij}} + S_{a_{e_{ij}}}$$

where the raw edge weights are normalized and their reciprocal value is used. Therefore, an edge with a high weight represents genetic similarity, relatively close nodes, and similar climatic conditions.

Based on the multi-weighted edges, the centrality measures of the identified network nodes are used to compare genetic, spatial, and climatic factors and to express the similarity of the patterns of the edges based on the Jaccard index<sup>104</sup>.

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

#### Data availability

All the data supporting the findings of this study are available within the article and its Supplementary Information files. All DNA sequence data was uploaded to GenBank under the accession numbers OQ064321.1–OQ064395.1, OR573677.1–OR573687.1, OQ067398.1–OQ067475.1, and OR574850.1–OR574858.1, and accession numbers are additionally given in Supplementary Data 1. Source data for charts and graphs are provided in Supplementary Data 2.

Received: 5 March 2023; Accepted: 21 November 2023;

Published online: 08 December 2023

#### References

1. Ready, P. D. Biology of phlebotomine sand flies as vectors of disease agents. *Annu. Rev. Entomol.* **58**, 227–250 (2013).
2. WHO. Fact sheets: Leishmaniasis. <https://www.who.int/news-room/fact-sheets/detail/leishmaniasis> (accessed on 19 July 2023) (2021).
3. Inceboz, T. Epidemiology and Ecology of Leishmaniasis. in *Current Topics in Neglected Tropical Diseases* (IntechOpen, 2019). <https://doi.org/10.5772/intechopen.86359>.
4. Ejov, M. & Dagne, D. Strategic framework for leishmaniasis control in the WHO European Region 2014–2020, World Health Organization. Regional Office for Europe, Copenhagen, Denmark. <https://iris.who.int/handle/10665/329477> (2014).
5. Benallal, K. E. et al. Phlebotomine sand flies (Diptera: Psychodidae) of the Maghreb region: a systematic review of distribution, morphology, and role in the transmission of the pathogens. *PLoS Negl. Trop. Dis.* **16**, e0009952 (2022).
6. Kniha, E., Aspöck, H., Auer, H. & Walochnik, J. Leishmania infections and Leishmania species in central Europe. *Wiener Tierärztliche Monatsschrift Vet Med. Austria* **110**, Doc1 (2023).
7. Antoniou, M., Gramiccia, M., Molina, R., Dvořák, V. & Volf, P. The role of indigenous phlebotomine sandflies and mammals in the spreading of leishmaniasis agents in the Mediterranean region. *Eurosurveillance* **18**, 1–9 (2013).

8. Depaquit, J., Grandadam, M., Fouque, F., Andry, P. & Peyrefitte, C. Arthropod-borne viruses transmitted by Phlebotomine sandflies in Europe: a review. *Eurosurveillance* **15**, 19507 (2010).
9. Ready, P. D. Leishmaniasis emergence in Europe. *Eurosurveillance* **15**, 19505 (2010).
10. Dvořák, V., Shaw, J. & Volf, P. *The Leishmaniasis: Old Neglected Tropical Diseases. The Leishmaniasis: Old Neglected Tropical Diseases* (Springer International Publishing, 2018). <https://doi.org/10.1007/978-3-319-72386-0>.
11. Ayhan, N., Prudhomme, J., Laroche, L., Bañuls, A. L. & Charrel, R. N. Broader geographical distribution of Toscana virus in the Mediterranean region suggests the existence of larger varieties of sand fly vectors. *Microorganisms* **8**, 114 (2020).
12. Artemiev, M. M. A revision of sandflies of the subgenus *Adlerius* (Diptera, Phlebotominae, Phlebotomus). *Zool. Zh.* **59**, 1177–1192 (1980).
13. Chaskopoulou, A., Giantsis, I. A., Demir, S. & Bon, M. C. Species composition, activity patterns and blood meal analysis of sand fly populations (Diptera: Psychodidae) in the metropolitan region of Thessaloniki, an endemic focus of canine leishmaniasis. *Acta Trop.* **158**, 170–176 (2016).
14. Zanet, S. et al. Epidemiology of *Leishmania infantum*, *Toxoplasma gondii*, and *Neospora caninum* in *Rattus rattus* in absence of domestic reservoir and definitive hosts. *Vet. Parasitol.* **199**, 247–249 (2014).
15. Obwaller, A. G. et al. Could *Phlebotomus mascittii* play a role as a natural vector for *Leishmania infantum*? New data. *Parasit. Vectors* **9**, 458 (2016).
16. Alarcón-Elbal, P. M. et al. First findings and molecular data of *phlebotomus mascittii* (Diptera: Psychodidae) in the Cantabrian Cornice (Northern Spain.). *J. Med. Entomol.* **58**, 2499–2503 (2021).
17. Bravo-Barriga, D., Ruiz-Arroondo, I., Peña, R. E., Lucientes, J. & Delacour-Estrella, S. Phlebotomine sand flies (Diptera, Psychodidae) from Spain: an updated checklist and extended distributions. *Zookeys* **1106**, 81–99 (2022).
18. Depaquit, J., Naucke, T. J., Schmitt, C., Ferté, H. & Léger, N. A molecular analysis of the subgenus *Transphlebotomus* *Artemiev*, 1984 (Phlebotomus, Diptera, Psychodidae) inferred from ND4 mtDNA with new northern records of *Phlebotomus mascittii* Grassi, 1908. *Parasitol. Res.* **95**, 113–116 (2005).
19. Prudhomme, J. et al. Ecology and spatiotemporal dynamics of sandflies in the Mediterranean Languedoc region (Roquedur area, Gard, France). *Parasit. Vectors* **8**, 1–14 (2015).
20. Naucke, T. J., Menn, B., Massberg, D. & Lorentz, S. Winter activity of *Phlebotomus* (*Transphlebotomus*) *mascittii*, Grassi 1908 (Diptera: Psychodidae) on the island of Corsica. *Parasitol. Res.* **103**, 477–479 (2008).
21. Dantas-Torres, F. et al. Ecology of phlebotomine sand flies and *Leishmania infantum* infection in a rural area of Southern Italy. *Acta Trop.* **137**, 67–73 (2014).
22. Michelutti, A. et al. Occurrence of phlebotomine sand flies (Diptera: Psychodidae) in the Northeastern plain of Italy. *Parasit. Vectors* **14**, 164 (2021).
23. Gradoni, L. et al. Monitoring and detection of new endemic foci of canine leishmaniasis in northern continental Italy: an update from a study involving five regions (2018–2019). *Vet. Parasitol. Reg. Stud. Rep.* **27**, 100676 (2022).
24. Dvorak, V. et al. Sand flies (Diptera: Psychodidae) in eight Balkan countries: historical review and region-wide entomological survey. *Parasit. Vectors* **13**, 573 (2020).
25. Vaselek, S. et al. Sand fly and *Leishmania* spp. survey in Vojvodina (Serbia): first detection of *Leishmania infantum* DNA in sand flies and the first record of *Phlebotomus* (*Transphlebotomus*) *mascittii* Grassi, 1908. *Parasit. Vectors* **10**, 444 (2017).
26. Vaselek, S. et al. A survey of sand flies (Diptera, Phlebotominae) along recurrent transit routes in Serbia. *Acta Trop.* **197**, 105063 (2019).
27. Ivočić, V., Kalan, K., Zupan, S. & Bužan, E. Illegal waste sites as a potential micro foci of Mediterranean Leishmaniasis: first records of phlebotomine sand flies (Diptera: Psychodidae) from Slovenia. *Acta Vet. Brno.* **65**, 348–357 (2015).
28. Praprotnik, E., Zupan, S. & Ivočić, V. Morphological and molecular identification of *phlebotomus mascittii* Grassi, 1908 populations from Slovenia. *J. Med. Entomol.* **56**, 565–568 (2019).
29. Kasap, O. E. et al. Phylogeography of the subgenus *Transphlebotomus* *Artemiev* with description of two new species, *Phlebotomus anatolicus* n. sp. and *Phlebotomus killicki* n. sp. *Infect. Genet. Evol.* **34**, 467–479 (2015).
30. Melaun, C., Krüger, A., Werblow, A. & Klimpel, S. New record of the suspected leishmaniasis vector *Phlebotomus* (*Transphlebotomus*) *mascittii* Grassi, 1908 (Diptera: Psychodidae: Phlebotominae) — the northernmost phlebotomine sandfly occurrence in the Palearctic region. *Parasitol. Res.* **113**, 2295–2301 (2014).
31. Naucke, T. J., Lorentz, S., Rauchenwald, F. & Aspöck, H. *Phlebotomus* (*Transphlebotomus*) *mascittii* Grassi, 1908, in Carinthia: first record of the occurrence of sandflies in Austria (Diptera: Psychodidae: Phlebotominae). *Parasitol. Res.* **109**, 1161–1164 (2011).
32. Poepl, W. et al. Emergence of sandflies (Phlebotominae) in Austria, a Central European country. *Parasitol. Res.* **112**, 4231–4237 (2013).
33. Kniha, E. et al. Integrative approach to *phlebotomus mascittii* Grassi, 1908: first record in Vienna with new morphological and molecular insights. *Pathogens* **9**, 1032 (2020).
34. Oerther, S. et al. Phlebotomine sand flies in Southwest Germany: an update with records in new locations. *Parasit. Vectors* **13**, 173 (2020).
35. Dvořák, V., Hlavackova, K., Kocisova, A. & Volf, P. First record of *Phlebotomus* (*Transphlebotomus*) *mascittii* in Slovakia. *Parasite* **23**, 48 (2016).
36. Trájer, A. J. Checklist, distribution maps, bibliography of the Hungarian *Phlebotomus* (Diptera: Psychodidae) fauna complementing with the climate profile of the recent sandfly distribution areas in Hungary. *Folia Faun. Slov.* **22**, 7–12 (2017).
37. Berdjane-Brouk, Z., Charrel, R. N., Bitam, I., Hamrioui, B. & Izri, A. Record of *Phlebotomus* (*Transphlebotomus*) *mascittii* Grassi, 1908 and *phlebotomus* (*Laroussius*) *chadli* Rioux, Juminer & Gibily, 1966 female in Algeria. *Parasite* **18**, 337–339 (2011).
38. Fouque, F. & Reeder, J. C. Impact of past and on-going changes on climate and weather on vector-borne diseases transmission: a look at the evidence. *Infect. Dis. Poverty* **8**, 51 (2019).
39. Essegir, S. & Ready, P. D. Speciation of *Phlebotomus* sandflies of the subgenus *Laroussius* coincided with the late Miocene-Pliocene aridification of the Mediterranean subregion. *Biol. J. Linn. Soc.* **70**, 189–219 (2000).
40. Depaquit, J. et al. ITS 2 sequences heterogeneity in *Phlebotomus sergenti* and *Phlebotomus similis* (Diptera, Psychodidae): Possible consequences in their ability to transmit *Leishmania tropica*. *Int. J. Parasitol.* **32**, 1123–1131 (2002).
41. Trájer, A. J., Sebestyén, V. & Padišák, J. The impacts of the Messinian Salinity Crisis on the biogeography of three Mediterranean sandfly (Diptera: Psychodidae) species. *Geobios* **65**, 51–66 (2021).
42. Cruaud, A., Lehrter, V., Genson, G., Rasplus, J.-Y. & Depaquit, J. Evolution, systematics and historical biogeography of sand flies of the subgenus *Paraphlebotomus* (Diptera, Psychodidae, Phlebotomus) inferred using restriction-site associated DNA markers. *PLoS Negl. Trop. Dis.* **15**, e0009479 (2021).
43. Depaquit, J. et al. Molecular homogeneity in diverse geographical populations of *Phlebotomus papatasi* (Diptera, Psychodidae) inferred from ND4 mtDNA and ITS2 rDNA. Epidemiological consequences. *Infect. Genet. Evol.* **8**, 159–170 (2008).
44. Pavlou, C. et al. A molecular phylogeny and phylogeography of Greek Aegean Island sand flies of the genus *Phlebotomus* (Diptera: Psychodidae). *Arthropod. Syst. Phylogeny* **80**, 137–154 (2022).
45. Trájer, A. J. & Sebestyén, V. The changing distribution of *Leishmania infantum* Nicolle, 1908 and its Mediterranean sandfly vectors in the last 140 kys. *Sci. Rep.* **9**, 11820 (2019).
46. Aspöck, H., Gersdorfer, T., Formayer, H. & Walochnik, J. Sandflies and sandfly-borne infections of humans in Central Europe in the light of climate change. *Wien. Klin. Wochenschr.* **120**, 24–29 (2008).
47. Aspöck, H. & Walochnik, J. When sandflies move north. *Public Heal. J.* **20**, 24–31 (2009).
48. Trájer, A. J. Palaeoclimatic models—predicted changes in the potential Neogene distribution patterns of *Phlebotomus similis* and *Phlebotomus sergenti* (Insecta: Diptera: Psychodidae). *Palaeoenvir.* **102**, 149–172 (2022).
49. Mahamdallie, S. S., Pesson, B. & Ready, P. D. Multiple genetic divergences and population expansions of a Mediterranean sandfly, *Phlebotomus ariasi*, in Europe during the Pleistocene glacial cycles. *Heredity (Edinb.)* **106**, 714–726 (2011).
50. Golledge, N. R. Glaciation of Scotland during the Younger Dryas stadial: a review. *J. Quat. Sci.* **25**, 550–566 (2009).
51. Smith, D. E., Harrison, S., Firth, C. R. & Jordan, J. T. The early Holocene sea level rise. *Quat. Sci. Rev.* **30**, 1846–1860 (2011).
52. Aspöck, H. Postglacial formation and fluctuations of the biodiversity of Central Europe in the light of climate change. *Parasitol. Res.* **103**, 7–10 (2008).
53. Schmitt, T. *6 Eurasien. Molekulare Biogeographie: Gene in Raum und Zeit*. 1st edn. Haupt Verlag, Bern (2020).
54. Schmitt, T. & Varga, Z. Extra-Mediterranean refugia: the rule and not the exception? *Front. Zool.* **9**, 22 (2012).
55. Aspöck, H. Fluctuations of biodiversity in Europe in light of climate change. *Nov. Acta Leopoldina* **111**, 35–44 (2010).
56. Birks, H. J. & Tinner, W. Past forests of Europe. [https://boris.unibe.ch/80787/1/Past\\_forests\\_of\\_Europe.pdf](https://boris.unibe.ch/80787/1/Past_forests_of_Europe.pdf) (2016).
57. Trájer, A. J. Placing of the second oldest red ochre mine in mainland Europe, Lovas, Hungary, to human, paleoenvironmental and paleobiogeographic context. *Quat. Sci. Rev.* **292**, 107670 (2022).
58. Tánzos, B. et al. First Record of Autochthonous Canine Leishmaniasis in Hungary. *Vector Borne Zoonotic Dis.* **12**, 588–594 (2012).
59. Kniha, E. et al. Ecology, seasonality and host preferences of Austrian *Phlebotomus* (*Transphlebotomus*) *mascittii* Grassi, 1908, populations. *Parasit. Vectors* **14**, 291 (2021).
60. Essegir, S., Ready, P. D., Killick-Kendrick, R. & Ben-Ismael, R. Mitochondrial haplotypes and phylogeography of *Phlebotomus* vectors of *Leishmania major*. *Insect Mol. Biol.* **6**, 211–225 (1997).
61. Aransay, A. M., Ready, P. D. & Morillas-Marquez, F. Population differentiation of *Phlebotomus perniciosus* in Spain following postglacial dispersal. *Heredity (Edinb.)* **90**, 316–325 (2003).



62. Perrotey, S. et al. Postglacial dispersal of *Phlebotomus perniciosus* into France. *Parasite* **12**, 283–291 (2005).
63. Hewitt, G. M. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* **58**, 247–276 (1996).
64. Hewitt, G. M. Post-glacial re-colonization of European biota. *Biol. J. Linn. Soc.* **68**, 87–112 (1999).
65. Haak, W. et al. Ancient DNA from the First European farmers in 7500-year-old neolithic sites. *Science (80-)* **310**, 1016–1018 (2005).
66. Sillero, N. et al. Updated distribution and biogeography of amphibians and reptiles of Europe. *Amphib. Reptil.* **35**, 1–31 (2014).
67. Orshan, L. et al. Distribution and Dispersal of *Phlebotomus papatasi* (Diptera: Psychodidae) in a Zoonotic Cutaneous Leishmaniasis Focus, the Northern Negev, Israel. *PLoS Negl. Trop. Dis.* **10**, e0004819 (2016).
68. Tonelli, G. B., Binder, C., Margonari, C. & Andrade Filho, J. D. Sand fly behavior: much more than weak-flying. *Mem. Inst. Oswaldo Cruz* **116**, e210230 (2021).
69. Trájer, A. J., Hammer, T. & Padišák, J. Reflection of the Neogene–Quaternary phylogeography in the recent distribution limiting climatic factors of eight Mediterranean *Phlebotomus* species (Diptera: Psychodidae). *J. Nat. Hist.* **52**, 1763–1784 (2018).
70. Costa, L. J. Préhistoire de la Corse. [The prehistory of Corsica]. *Kyrnos Publ. pour l'archéologie. Arch. from Orig. 2004-11-07* [In French] (2005).
71. Patton, M. *Islands in Time: Island Sociography and Mediterranean prehistory* (Routledge, 1996).
72. Trájer, A. J. The changing malaria risk patterns in East-Central Europe and the North Balkans in the last 27 000 years. *J. Quat. Sci.* **36**, 1234–1253 (2021).
73. Depaquit, J. et al. Molecular and morphological systematics of the sandfly *Sergentomyia* (*Sintonius*) *clydei* Sinton, 1928 and questions about its record in the Seychelles. *Infect. Genet. Evol.* **21**, 41–53 (2014).
74. Lewis, D. J. A taxonomic review of the genus *Phlebotomus* (Diptera: Psychodidae). *Bull. Br. Museum* **45**, 121–209 (1982).
75. Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* **3**, 294–299 (1994).
76. Parvizi, P. & Ready, P. D. Molecular investigation of the population differentiation of *Phlebotomus papatasi*, important vector of *L. major*, in different habitats and regions of Iran. *Iran. Biomed. J.* **10**, 69–77 (2006).
77. Nicholas, K. B. Genedoc: a tool for editing and annotating multiple sequence alignments. <http://www.psc.edu/biomed/genedoc> (1997).
78. Larkin, M. A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
79. Librado, P. & Rozas, J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452 (2009).
80. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
81. Templeton, A. R., Crandall, K. A. & Sing, C. F. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**, 619–633 (1992).
82. Leigh, J. W. & Bryant, D. POPART: full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**, 1110–1116 (2015).
83. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
84. Excoffier, L., Laval, G. & Schneider, S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* **1**, 47–50 (2007).
85. Bouckaert, R. et al. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* **10**, e1003537 (2014).
86. Felsenstein, J. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**, 368–376 (1981).
87. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermini, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
88. Brower, A. V. Z. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc. Natl Acad. Sci.* **91**, 6491–6495 (1994).
89. Farrell, B. D. Evolutionary assembly of the milkweed Fauna: cytochrome oxidase I and the age of tetrapesbeetles. *Mol. Phylogenet. Evol.* **18**, 467–478 (2001).
90. Quek, S.-P., Davies, S. J., Itino, T. & Pierce, N. E. Codiversification in an ant-plant mutualism: stem texture and the evolution of host use in *Crematogaster* (Formicidae: Myrmicinae) inhabitants of *Macaranga* (Euphorbiaceae). *Evolution (N.Y.)* **58**, 554–570 (2004).
91. Papadopoulou, A., Anastasiou, I. & Vogler, A. P. Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. *Mol. Biol. Evol.* **27**, 1659–1672 (2010).
92. Kuhner, M. K. Coalescent genealogy samplers: windows into population history. *Trends Ecol. Evol.* **24**, 86–93 (2009).
93. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior summarization in Bayesian phylogenetics using tracer 1.7. *Syst. Biol.* **67**, 901–904 (2018).
94. Miller, M. P. Alleles In Space (AIS): computer software for the joint analysis of interindividual spatial and genetic information. *J. Hered.* **96**, 722–724 (2005).
95. Karger, D. N. et al. CHELSA-TraCE21k v1.0. Downscaled transient temperature and precipitation data since the last glacial maximum. *Clim. Past Discuss.* **19**, 439–456 (2021).
96. QGIS Development Team. QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org> (2019).
97. Nix, H. A. A biogeographic analysis of Australian snakes. in *Atlas of elapid snakes of Australia: Australia flora and fauna series 7* (ed Longmore, R.), 4–15 (Australian Government Publishing Service, Canberra, 1986).
98. Trájer, A. J. et al. The effect of climate change on the potential distribution of the European *Phlebotomus* species. *Appl. Ecol. Appl. Ecol. Environ. Res.* **11**, 189–208 (2013).
99. Halmos, P. & Givant, S. *Introduction to Boolean algebras*. 8–13 (Springer, 2009).
100. Hijmans, R. J. & Graham, C. H. The ability of climate envelope models to predict the effect of climate change on species distributions. *Glob. Chang. Biol.* **12**, 2272–2281 (2006).
101. Austin, M. Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecol. Modell.* **200**, 1–19 (2007).
102. Hammoud, Z. & Kramer, F. Multilayer networks: aspects, implementations, and application in biomedicine. *Big Data Anal.* **5**, 2 (2020).
103. Bastian, M., Heymann, S. & Jacomy, M. Gephi: an open source software for exploring and manipulating networks. in *Proceedings of the Third International ICWSM Conference* 361–362 (2009).
104. Besta, M. et al. Communication-Efficient Jaccard similarity for High-Performance Distributed Genome Comparisons. in *2020 IEEE International Parallel and Distributed Processing Symposium (IPDPS)* 1122–1132 (IEEE, 2020). <https://doi.org/10.1109/IPDPS47924.2020.00118>.
105. Fordham, D. A. et al. PaleoView: a tool for generating continuous climate projections spanning the last 21,000 years at regional and global scales. *Ecography* **40**, 1348–1358 (2017).

## Acknowledgements

The research presented in this article was Prepared in the “Climate Change Multi-disciplinary National Laboratory” project (RRF-2.3.1-21-2022-00014), within the framework of the Széchenyi Plan Plusz program, with the support of the Recovery and Resilience Instrument of the European Union. Data collection and analysis was supported by the Austrian Agency for Education and Internationalization (OeAD) [grant number WTZ CZ02/2020], the VectorNet project, a European network for sharing data on the geographic distribution of arthropod vectors, transmitting human and animal disease agents (Contract OC/EFSA/AHAW/2013/02-FWC1) funded by the European Food Safety Authority (EFSA) and the European Center for Disease Prevention and Control (ECDC), and by the Austrian defense research program FORTE of the Federal Ministry of Finance (BMF) [grant number 886318]. R.L. is funded by the Federal Ministry of Education and Research of Germany (grant number 01K12022). The authors are also deeply thankful to Iveta Häfeli for technical support and to Univ.-Prof. Dr. Horst Aspöck for substantial discussions, both members of the Institute of Specific Prophylaxis and Tropical Medicine, Medical University of Vienna.

## Author contributions

Design of the study: E.K., V.D., J.W., V.S. and A.J.T. Entomological fieldwork: E.K., V.D., J.P., V.I., S.O., A.L.B., D.S., A.M., F.T., P.M.A.E., D.B.B., M.A.G., J.L., V.C., D.O., M.B.S., G.K., A.G.O., P.V., A.A., O.E.K., B.A. and J.O. Laboratory work: E.K., V.D., V.I., I.H., A.H., R.L., A.M., F.T. and O.E.K. Data analyses: E.K., V.D., S.K., J.D., J.W., V.S. and A.J.T. Drafting of manuscript: E.K., J.W., V.S. and A.J.T. Critical review of manuscript: all authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-023-05616-1>.

**Correspondence** and requests for materials should be addressed to Attila J. Trájer.

**Peer review information** *Communications Biology* thanks Ammara Fareen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: David Favero. A peer review file is available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

<sup>1</sup>Institute of Specific Prophylaxis and Tropical Medicine, Center for Pathophysiology, Infectiology and Immunology, Medical University of Vienna, Vienna, Austria. <sup>2</sup>Department of Parasitology, Faculty of Science, Charles University, Prague, Czech Republic. <sup>3</sup>Institute of Biology, University of Graz, Graz, Austria. <sup>4</sup>UMR MIVEGEC (Université de Montpellier—IRD—CNRS), Institute of Research for Development, Montpellier, France. <sup>5</sup>INTHERES, Université de Toulouse, INRAE, ENVT, Toulouse, France. <sup>6</sup>Department of Biodiversity, FAMNIT, University of Primorska, Koper-Capodistria, Slovenia. <sup>7</sup>Institute of Global Health, Heidelberg University, Heidelberg, Germany. <sup>8</sup>German Mosquito Control Association (KABS), Speyer, Germany. <sup>9</sup>Institute for Dipterology (IfD), Speyer, Germany. <sup>10</sup>Department of Arbovirology, Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany. <sup>11</sup>Research Group Vector Control, Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany. <sup>12</sup>Institut de Recherche pour le Développement, Université de Montpellier, UMR INTERTRYP, Parasite Infectiology and Public Health Research group. IRD, CIRAD, Montpellier, France. <sup>13</sup>Laboratory of Parasitology, Micology and Medical Entomology, Istituto Zooprofilattico Sperimentale delle Venezie, Legnaro Padova, Italy. <sup>14</sup>Departamento de Producción y Sanidad Animal, Salud Pública Veterinaria y Ciencia y Tecnología de los Alimentos (PASAPTA), Facultad de Veterinaria, Universidad CEU Cardenal Herrera, Valencia, Spain. <sup>15</sup>Laboratorio de investigación de Entomología, Departamento de Zoología, Facultad de Ciencias Biológicas, Bloque B, Universidad de Valencia, Valencia, Spain. <sup>16</sup>Department of Animal Health, Animal Health and Zoonosis Research Group (GISAZ), UIC Zoonosis and Emerging Diseases (ENZOEM), University of Cordoba, Cordoba, Spain. <sup>17</sup>Department of Animal Production and Health, Veterinary Public Health and Food Science and Technology (PASAPTA), Facultad de Veterinaria, Universidad Cardenal Herrera-CEU, CEU Universities, Valencia, Spain. <sup>18</sup>Applied Zoology and Animal Conservation Group, University of the Balearic Islands (UIB), Palma de Mallorca, Spain. <sup>19</sup>Animal Health Department, The AgriFood Institute of Aragon (IA2), School of Veterinary Medicine, University of Zaragoza, Zaragoza, Spain. <sup>20</sup>Faculty of Science, The University of Melbourne, Parkville, Australia. <sup>21</sup>Department of Veterinary Medicine, University of Bari, Bari, Italy. <sup>22</sup>Faculty of Veterinary Sciences, Bu-Ali Sina University, Hamedan, Iran. <sup>23</sup>Division of Science, Research and Development, Federal Ministry of Defence, Vienna, Austria. <sup>24</sup>Université de Reims Champagne Ardenne, ESCAPE EA7510, USC ANSES VECPAR, SFR Cap Santé, UFR de Pharmacie, Reims, France. <sup>25</sup>Department of Clinical Sciences of Veterinary Medicine, Faculty of Veterinary Medicine, University of Sarajevo, Sarajevo, Bosnia and Herzegovina. <sup>26</sup>Department of Biology, Ecology Section, Faculty of Science, VERG Laboratories, Hacettepe University, Ankara, Turkey. <sup>27</sup>Department of Pathobiology and Epidemiology, Veterinary Faculty, University of Sarajevo, Sarajevo, Bosnia and Herzegovina. <sup>28</sup>University of Pannonia, Sustainability Solutions Research Lab, Veszprém, Hungary.

✉email: [attilatrajaj@gmail.com](mailto:attilatrajaj@gmail.com); [trajaj.attila@mk.uni-pannon.hu](mailto:trajaj.attila@mk.uni-pannon.hu)