



## OPEN Combining phylogeography and climate models to track the diversification and spread of *Phlebotomus simici*

Edwin Kniha<sup>1</sup>, Stephan Koblmüller<sup>2</sup>, Katharina Platzgummer<sup>1</sup>, Oscar Kirstein<sup>3</sup>, Debora Diaz<sup>3</sup>, Vít Dvořák<sup>4</sup>, Ozge Erisoz Kasap<sup>5</sup>, Betim Xhekaj<sup>6</sup>, Kurtesh Sherifi<sup>6</sup>, Julia Walochnik<sup>1</sup> & Attila J. Trájer<sup>7</sup>✉

Phlebotomine sand flies (Diptera: Psychodidae: Phlebotominae) are the principal vectors of *Leishmania* spp. (Kinetoplastida: Trypanosomatidae) worldwide. The subgenus *Adlerius* is taxonomically challenging and currently comprises about 20 species with a wide geographic distribution from eastern Asia to southeastern Europe. Some species are confirmed or suspected vectors of *Leishmania donovani/infantum*, *L. major*, and *L. tropica*, and are thus of high medical and veterinary relevance. A single record of *Phlebotomus (Adlerius) simici* in Austria from 2018 marks its sporadic northernmost and westernmost occurrence, with the origin of its appearance remaining unclear. To better understand *Adlerius* diversification and particularly post-glacial spread of *Ph. simici* to northern parts of Europe, we combined phylogenetic analyses with climatic suitability modelling. Divergence time estimates well supported the currently observed geographic distribution of the studied species and revealed several taxonomic challenges in the subgenus. We clearly delineated three distinct genetic and geographic *Ph. simici* lineages and phylogeographically assessed diversification that were well supported by climatic models. This study provides a comprehensive phylogenetic analysis of the subgenus *Adlerius*, enhancing our understanding of the diversification in relation to changing climate of this understudied group, and we present new insights into the post-glacial spread of *Ph. simici*, a suspected vector of *L. infantum*.

**Keywords** Central Europe, Phylogeography, *Adlerius*, Divergence time, COI

Phlebotomine sand flies (Diptera: Psychodidae: Phlebotominae) are small hematophagous insects and the principal vectors of the protozoan parasites *Leishmania*, the causative agents of leishmaniasis. They inhabit tropical, subtropical, arid, and temperate regions worldwide<sup>1</sup>. Species of the subgenus *Adlerius* Nitzulescu, 1931, show a wide geographic distribution range, being endemic in China, Middle Eastern countries (e.g. Iran, Iraq, Israel, Lebanon and Syria), Türkiye and Greece including Aegean islands, and many Balkan countries as far north as Serbia<sup>2–5</sup>.

*Adlerius* is a subgenus little studied, which presents a taxonomically interesting and challenging group. Its morphological characterization is inconclusive for some valid taxa, as females of some species are morphologically indistinguishable, for example, *Phlebotomus chinensis* and *Phlebotomus sichuanensis* in China<sup>6</sup>, or *Phlebotomus brevis* and *Phlebotomus simici* that have a sympatric distribution in Lebanon<sup>7</sup>. The last systematic revision of the subgenus, now outdated by over four decades, was based solely on the assessment of morphological characters and regarded some of them inadequate for conclusive species identification<sup>8</sup>. A few later studies using also molecular techniques enabled the description of a new species, *Phlebotomus creticus* on the island of Crete<sup>9</sup>, suggested the presence of previously undescribed *Adlerius* species from Türkiye and Armenia<sup>4,10</sup>, and another

<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>Institute of Biology, University of Graz, Graz, Austria. <sup>3</sup>Laboratory of Entomology, Ministry of Health, Jerusalem, Israel. <sup>4</sup>Department of Parasitology, Faculty of Science, Charles University, Prague, Czech Republic. <sup>5</sup>Department of Biology, Ecology Section, Faculty of Science, VERG Laboratories, Hacettepe University, Ankara, Turkey. <sup>6</sup>Faculty of Agriculture and Veterinary, University of Prishtina, Prishtina, Kosovo. <sup>7</sup>Sustainability Solutions Research Lab, University of Pannonia, Veszprém, Hungary. ✉email: trajer.attila@mk.uni-pannon.hu

from Moldova<sup>11</sup> and proposed resurrection of previously synonymized *Ph. davidi* as a valid species<sup>12</sup>. Albeit the use of molecular taxonomy approaches has been only sporadic in *Adlerius* species, the data obtained clearly highlights the need of a careful revision of the subgenus, based on a complex assessment of a combination of morphological as well as molecular markers. In addition to its taxonomical peculiarities, the subgenus *Adlerius* is also of medical and veterinary relevance as some species are proven competent vectors for several *Leishmania* species, such as *Ph. chinensis* for *L. infantum/donovani* and *Ph. arabicus* for *L. tropica*<sup>13,14</sup>, or suspected vectors such as *Ph. halepensis* for *L. major* and *L. tropica*, or *Ph. simici* and *Ph. balcanicus* for *L. infantum*<sup>15–17</sup>. Their proven or expected involvement in various transmission cycles of *Leishmania* spp. suggests that other species of the subgenus, whose vectorial competence has never been tested, may also be relevant for understanding the epidemiology of sand fly-borne diseases in the regions of their presence.

In Europe, *Phlebotomus mascittii* is the most widespread species and in some countries north of the usual sand fly Mediterranean distribution, it is the only endemic species<sup>18</sup>. Unexpectedly, a single female *Ph. simici* specimen was recorded in Austria in 2018, marking the western- and northernmost occurrence of an *Adlerius* species<sup>10</sup>. Despite many trapping efforts since then, it has never been collected again in Austria, and neither in any bordering countries. Kniha et al.<sup>10</sup> identified three distinct lineages of *Ph. simici* from Israel, Türkiye and the Balkans, respectively, suggesting a natural or anthropogenic post-glacial dispersal from a Balkan country to Austria.

Several previous studies showed that the western branch of the Paratethys and the Aegean-Asia Minor region could have played an important role in the diversification of Eastern Mediterranean sand fly taxa<sup>19,20</sup>. The evolutionary history of phlebotomine sand flies, particularly within the subgenera *Adlerius*, *Paraphlebotomus*, and *Transphlebotomus*, reflects a complex interplay of climatic, geographic, and ecological transformations across the Oligocene, Miocene, and Quaternary periods. Key geological and climatic shifts in the peri-Mediterranean region shaped both the distribution and diversification of sand fly species, influencing their speciation, range expansion, and ecological adaptation. During the Oligocene and early Miocene epochs, the humid, subtropical climate of the Mediterranean basin posed limited suitability for many sand fly ancestors, particularly for those with Mediterranean affinities<sup>21,22</sup>. However, with the gradual transition to a drier subtropical climate by the middle Miocene, habitats became more conducive to the diversification of certain sand fly groups, including *Larrousius* and *Paraphlebotomus*. Modelling studies of climate suitability from the Rupelian to the Tortonian stages in the Central Paratethys suggest that the ancestor of the *Larrousius* group could have adapted well to these evolving conditions, while the Central Paratethys may have served as a geographic barrier to eastward-migrating ancestors of *Paraphlebotomus* and *Phlebotomus* species<sup>22</sup>. As the Miocene progressed, these barriers began to fragment with the tectonic subsidence of landmasses like the Hellene Orogenic Belt, facilitating niche expansion and the gradual migration of *Phlebotomus* species, which adapted to the more extreme Mediterranean climate<sup>23</sup>.

The diversification of various sand fly species is linked to significant paleogeographic and paleoclimatic events, particularly the desiccation of the Mediterranean Basin during the Messinian Salinity Crisis<sup>24,25</sup>. This event led to the isolation and subsequent divergence of sand fly populations, driving the initial stages of speciation<sup>20</sup>. Following the Zanclean re-flood of the Mediterranean Basin, *Phlebotomus* species began to reoccupy new habitats, with many adapting to the distinct climatic regimes of the Iberian, Apennine, and Balkan refugia<sup>26</sup>. These refugia offered stable microclimates through glacial maxima, fostering the differentiation of sand flies into “Trans-Mediterranean,” “East Mediterranean,” and “West Mediterranean” climatic groups<sup>27</sup>.

The *Adlerius* group encapsulates the effects of both ancient and recent climatic forces, and thus can provide understanding of Mediterranean sand fly evolution<sup>19,20</sup>. While phylogeographic reconstructions of the region highlight the role of the Mediterranean Sea and adjacent landforms in shaping the evolutionary pathways of sand flies, the genetic, ecological, and bioclimatic adaptations of *Adlerius* underscore the adaptability of sand flies to shifting ecological landscapes. As vectors of *Leishmania*, their evolutionary history offers crucial insights into the biogeographical and ecological processes that continue to influence disease transmission dynamics in these regions, shedding light on both ancient speciation events and contemporary distributions.

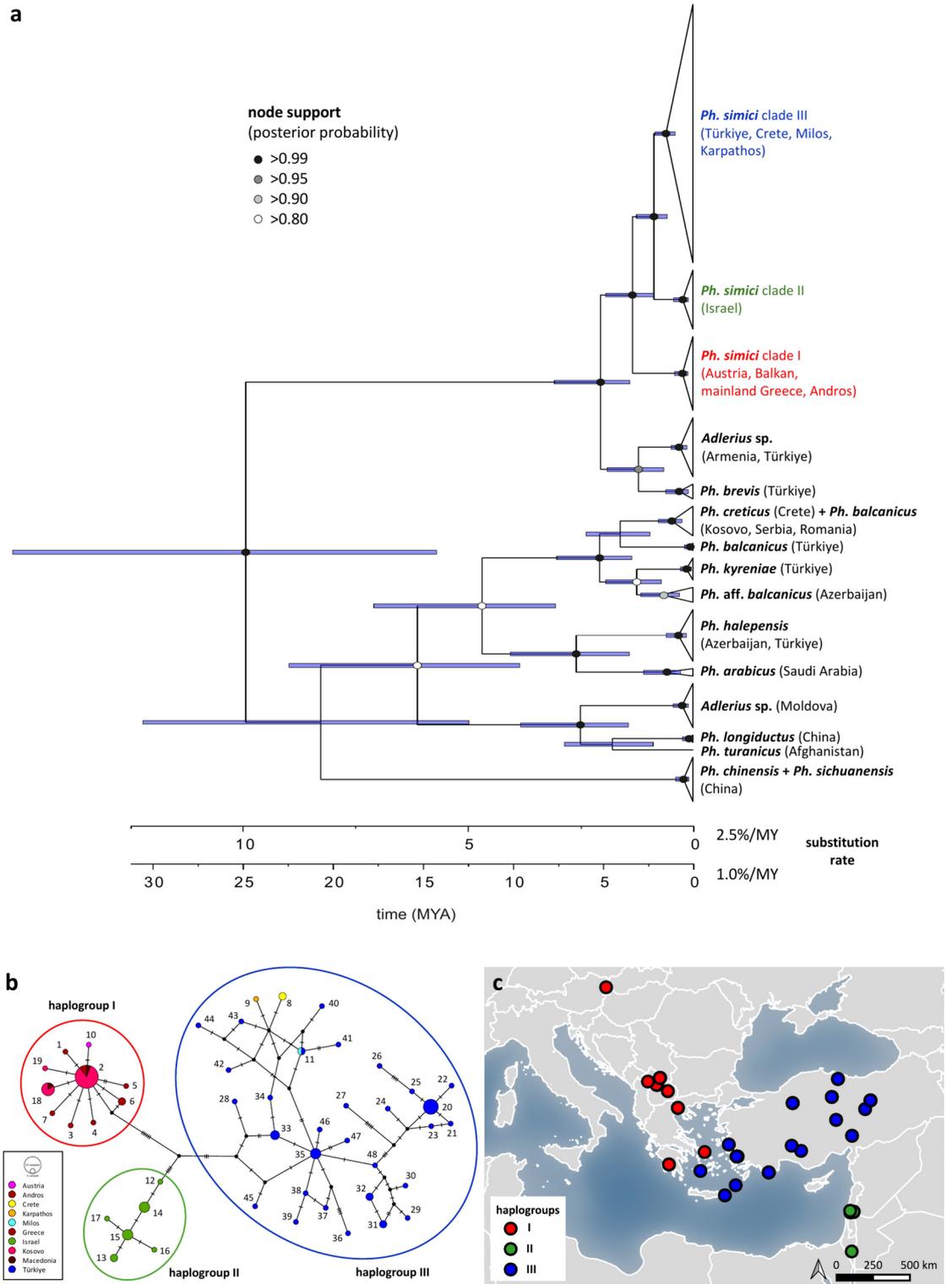
The present study aimed to assess phylogenetic relationships of *Adlerius* species and reconstruct precise temporal patterns of diversification of *Ph. simici* lineages by combining phylogeographic data and climate modelling. Results should further elucidate and support our understanding of the post-glacial dispersal and the currently known distribution.

## Results

### Phylogenetic relationships of *Adlerius*

Phylogenetic analyses based on COI sequences revealed two major clades: one comprising *Ph. simici*, *Ph. brevis* and an unidentified *Adlerius* species from Armenia and Türkiye (Fig. 1a), the second, major clade comprising all other included *Adlerius* species (Fig. 1a). In the first major clade, *Ph. simici* was observed to form three distinct lineages. *Phlebotomus simici* lineage I comprised specimens from Austria, Balkan, Greece and Andros, lineage II all specimens from Israel and lineage III all specimens from Türkiye, Crete, Milos and Karpathos. Lineages II and III were resolved as sister groups, together forming the sister group of *Ph. simici* lineage I. *Phlebotomus brevis* and *Adlerius* sp. resulted as sister species, together forming the sister group of *Ph. simici* (Fig. 1a).

In the second major clade, *Phlebotomus balcanicus* (incl. *Ph. aff. balcanicus*) comprised three different groups. *Phlebotomus balcanicus* originating from Kosovo, Serbia and Romania constituted the first clade together with the closely related species *Ph. creticus*, forming the sister group of *Ph. balcanicus* from Türkiye (Fig. 1a). *Phlebotomus aff. balcanicus* from Azerbaijan formed the sister group of *Ph. kyreniae* from Türkiye and together represented the sister group of *Ph. balcanicus* (Balkan)/*Ph. creticus* + *Ph. balcanicus* (Türkiye). Further, *Ph. halepensis* and *Ph. arabicus* formed sister species, and *Ph. longiductus* + *Ph. turanicus* were resolved as the sister



**Fig. 1.** Dated phylogeny of *Adlerius* based on COI sequences and corresponding *Ph. simici* haplotypes. Timetree with divergence times calculated assuming a maximum and minimum substitution rate of 2.5% and 1.0% per million years (my) for COI, respectively (a). The tree shows the median divergence ages, with the node bars representing the 95% highest posterior density (HPD) intervals for node ages. Node support is only shown for those nodes with posterior probabilities > 0.8. Statistical parsimony network of *Phlebotomus simici* from six countries (including four Greek islands) based on COI sequences (b). Colored circles represent major haplogroups identified. Geographical origin of identified *Ph. simici* haplogroups (I–III) among analyzed samples (c). Macedonia = North Macedonia.

group of an unidentified *Adlerius* species from Moldova. *Phlebotomus chinensis* and *Ph. sichuanensis*, which cannot be discriminated by COI sequences, resulted as a rather distinct lineage (Fig. 1a).

### Divergence times estimates

Median divergence time estimates of all included *Adlerius* species ranged from 25 million years ago (mya) (based on 1.0% divergence rate/my) and 10 mya (based on 2.5% divergence rate/my) to 2.5 mya and 1 mya. The majority of major splits were estimated to be between 6.5 mya (1.0% divergence/my) and 1 mya (2.5% divergence/my). Within major clade 1, the split of *Ph. simici* from *Ph. brevis* + *Adlerius* sp. were estimated to have happened between 5 mya and 2.3 mya. The first split between *Ph. simici* lineage I and lineage II + lineage III appeared between 3.5 mya and 1.25 mya, and the split between *Ph. simici* lineage II and lineage III was estimated between 2.5 mya and 1 mya (Fig. 1a).

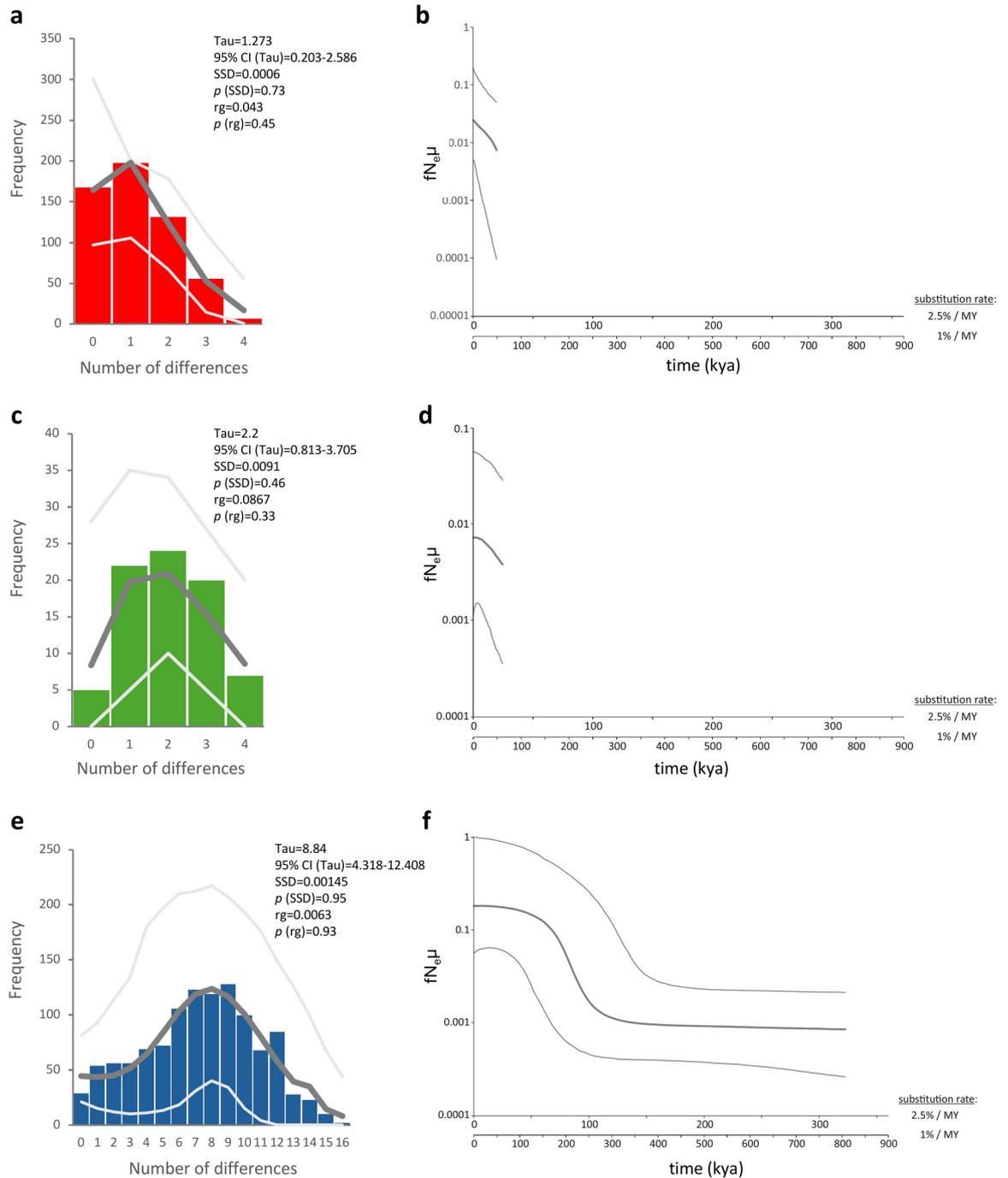
Divergence times estimates involving *Ph. balcanicus* in the second major clade were similar to divergence times observed in the first major clade. The first split between *Ph. creticus*/*Ph. balcanicus* (Balkan) + *Ph. balcanicus* (Türkiye) and *Ph. kyreniae* + *Ph. aff. balcanicus* appeared between 5 mya and 2.3 mya, the split between *Ph. creticus*/*Ph. balcanicus* (Balkan) and *Ph. balcanicus* (Türkiye) was estimated between 4 and 2 mya. Finally, the split between *Ph. kyreniae* and *Ph. aff. balcanicus* was estimated between 3 mya and 1.1 mya (Fig. 1a). Also, divergence time estimates between 6.5 mya and 2.5 mya were obtained for *Ph. halepensis* and *Ph. arabicus* as well as *Ph. longiductus* + *Ph. turanicus* and the unidentified *Adlerius* species from Moldova (Fig. 1a).

### Phylogeographic patterns of *Ph. simici*

Overall, 48 COI haplotypes of *Ph. simici* were identified, comprising 49 variable sites (29 parsimony informative) with a haplotype diversity (Hd) of 0.946, a nucleotide diversity ( $\pi$ ) of 0.0195, and a mean group distance of 1.9% (S.E. = 0.4%). Three major haplogroups (I–III) corresponding to the three *Ph. simici* lineages of the phylogenetic tree were observed (Fig. 1b), revealing a distinct geographic origin (Fig. 1c). Haplogroup I (pd = 0.2%, S.E. = 0.07%) comprised specimens from Austria, Kosovo, North Macedonia, mainland Greece and the Aegean Greek island Andros. Haplogroup II (pd = 0.3%, S.E. = 0.1%) comprised all specimens originating from Israel, and haplogroup III (pd = 1.3%, S.E. = 0.3%) included all specimens from Türkiye as well as the Aegean Greek islands of Crete, Milos and Karpathos (Fig. 1b). The mean group distance between haplogroup I and II was 2.6% (S.E. = 0.7%), between haplogroup I and III was 2.6% (S.E. = 0.7%), and 2.3% (S.E. = 0.6%) between haplogroup II and III (Supplementary Table 1).

For the star-shaped haplogroup I, a significantly negative Tajima's D ( $D = -2.235$ ;  $p = 0.0002$ ), and Fu's Fs ( $F_s = -28.866$ ;  $p < 0.001$ ), and a unimodal mismatch distribution (Fig. 2a) with non-significant raggedness index ( $rg = 0.043$ ;  $p = 0.447$ ) and non-significant sum of squared differences (SSD = 0.001;  $p = 0.726$ ) 0.0867;  $p = 0.33$ ), indicated a recent population expansion. The Bayesian Skyline plot clearly demonstrated evidence of a strong, recent, post-glacial population expansion (Fig. 2b). The time to the most recent common ancestor was estimated at 54.0 (95% HPD interval 20.7–101.1) kya to 135.1 (95% HPD interval 51.8–252.9) kya, depending on the assumed substitution rates. For haplogroup II, a negative, but non-significant, Tajima's D ( $D = -0.387$ ;  $p = 0.395$ ), a significantly negative Fu's Fs ( $F_s = -16.072$ ;  $p < 0.001$ ), and a unimodal mismatch distribution (Fig. 2c) with non-significant raggedness index ( $rg = 0.087$ ;  $p = 0.328$ ) and sum of squared differences (SSD = 0.009;  $p = 0.461$ ), indicated a recent population expansion. Also, the Bayesian Skyline plot provided clear evidence for strong recent post-glacial population expansion (Fig. 2d). The time to the most recent common ancestor was estimated at 81.9 (95% HPD interval 24.6–153.3) kya to 204.7 (95% HPD interval 61.4–383.3) kya, depending on the assumed substitution rate. For haplogroup III, a negative, but non-significant, Tajima's D ( $D = -0.794$ ;  $p = 0.237$ ), a significantly negative Fu's Fs ( $F_s = -25.012$ ;  $p < 0.001$ ), and a skewed unimodal mismatch distribution (Fig. 2e) with non-significant raggedness index ( $rg = 0.006$ ;  $p = 0.928$ ) and non-significant sum of squared differences (SSD = 0.001;  $p = 0.945$ ), indicated past population expansion. The Bayesian Skyline plot provided clear evidence for population expansion, but earlier than the two other haplogroups (Fig. 2f). The time to the most recent common ancestor was estimated at 561.6 (95% HPD interval 325.1–807.2) kya to 1403.9 (95% HPD interval 812.8–2018.1) kya, depending on the assumed substitution rate.

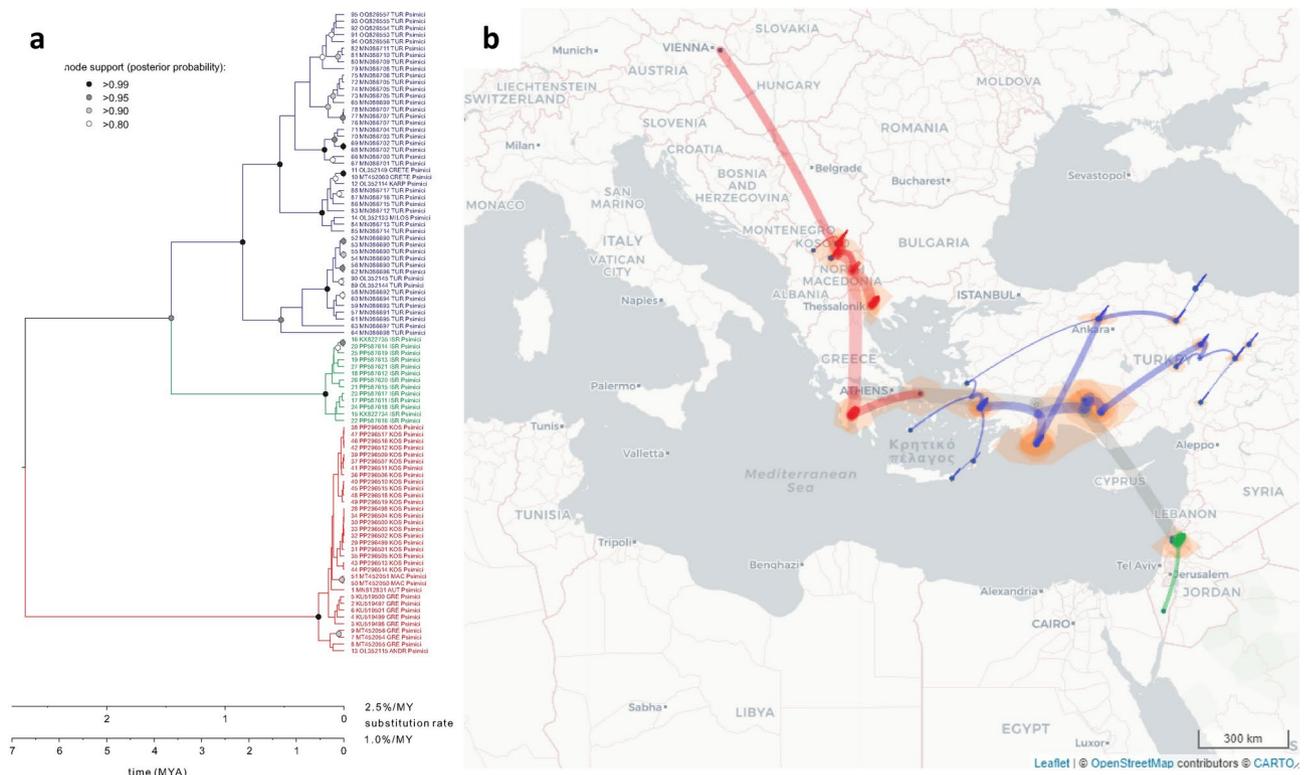
Our spatial diffusion analysis suggested that the most recent common ancestor (MRCA) of all extant *Ph. simici* (or more specifically, their mtDNA) occurred in southwestern Türkiye. From this region, the species expanded its range westward to Greece and further into the Balkans, reaching as far north as Austria (lineage I), and eastward, where it subsequently colonized the Near East (Israel; clade II) and spread further within Türkiye, including the southern Aegean islands (lineage III). In clade I (MRCA ~ 200–500 kya), the initial divergence occurred between southern Greece and the rest of the Balkans. Notably, the divergence of the Austrian sample from the central Balkan samples was estimated to have occurred 60–150 kya, depending on the substitution rate used. Lineage II was exclusively found in Israel, with levels of intra-clade divergence comparable to those observed in clade I. The MRCA of lineage III was inferred to have occurred in central southern Türkiye, from where it expanded through Türkiye via three distinct routes. The first lineage, which colonized central Türkiye, branched off approximately 0.8–2.0 mya. Initially present in southern central Türkiye near the Mediterranean coast (with the MRCA dating to ~ 0.5–1.25 mya), it spread toward the north and east more recently. The split between the other two major lineages within clade III occurred approximately 0.5–1.25 mya in southwestern Türkiye. One of these lineages, with an MRCA dating to ~ 300–750 kya, colonized the southwestern Turkish coast and progressively expanded northward, eventually reaching the central northern part of Türkiye near the Black Sea coast and also spreading to the western coast of Türkiye. The second lineage, with an MRCA of approximately 200–500 kya, colonized the western Turkish Mediterranean coast and subsequently spread to the southern Greek islands (Fig. 3).



**Fig. 2.** Signatures of population expansion in *Phlebotomus simici*. Mismatch distributions for *Ph. simici* haplogroup I (a), haplogroup II (c), and haplogroup III (e), based on COI 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. Estimated Bayesian Skyline Plots (BSPs) for all three haplogroups (b, d, f), assuming a minimum and maximum substitution rate for mitochondrial protein-coding genes of insects of 1% and 2.5% per my. 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.

**Elevation and climatic characterization of *Phlebotomus simici* sites**

The studied *Ph. simici*-related sites were found between sea level and 1656 m a.s.l. with a mean of 387 m (SD = 388 m); 72.3% of the studied *Ph. simici* sites were found in hot-summer Mediterranean (Köppen–Geiger Csa) climate areas, followed by cold semi-arid areas (Köppen–Geiger BSk) (14.3%). An additional 8.2% of the



**Fig. 3.** Phylogeographic scenario of the spread of *Ph. simici* based on COI sequences. **(a)** Intraspecific time-calibrated MCC tree of *Ph. simici*, assuming a minimum and maximum substitution rate of 1% and 2.5% per my for the COI gene in insects. The tree shows the median divergence ages; node support is only shown for nodes with posterior probabilities > 0.8. Different colors indicate to the distinct *Ph. simici* main clades. **(b)** Map showing the transition paths of the spread of *Ph. simici*. Transition paths are colored according to clades in **(a)**; width and opacity of transition paths reflect their depth of the transition in the phylogenetic tree (older paths are wider and more transparent). Orange areas indicate 80% HPD intervals for ancestral locations. The figure was created with *EvoLaps2* v.2.42 (<https://www.evolaps.org>).

sites were found in Mediterranean-influenced and non-Mediterranean-influenced continental (Köppen–Geiger Dsa, DSb, Dfb) climate regions, and 4.8% in temperate oceanic and humid subtropical (Köppen–Geiger Cfa, Cfb) climate regions. A single site (0.4%) occurred in an arid (Köppen–Geiger BWh) climate area (Supplementary Fig. 1).

Conclusively, *Ph. simici* is a sand fly species of predominantly hot summer Mediterranean and cold semi-arid (continental steppe and forest-steppe) climate regions. The dominance of the Mediterranean climate-influence is even more conspicuous when adding hot summer Mediterranean and Mediterranean-influenced humid continental climate locations, which together represent 74.6% of *Ph. simici*-inhabited sites.

## Climatic suitability patterns

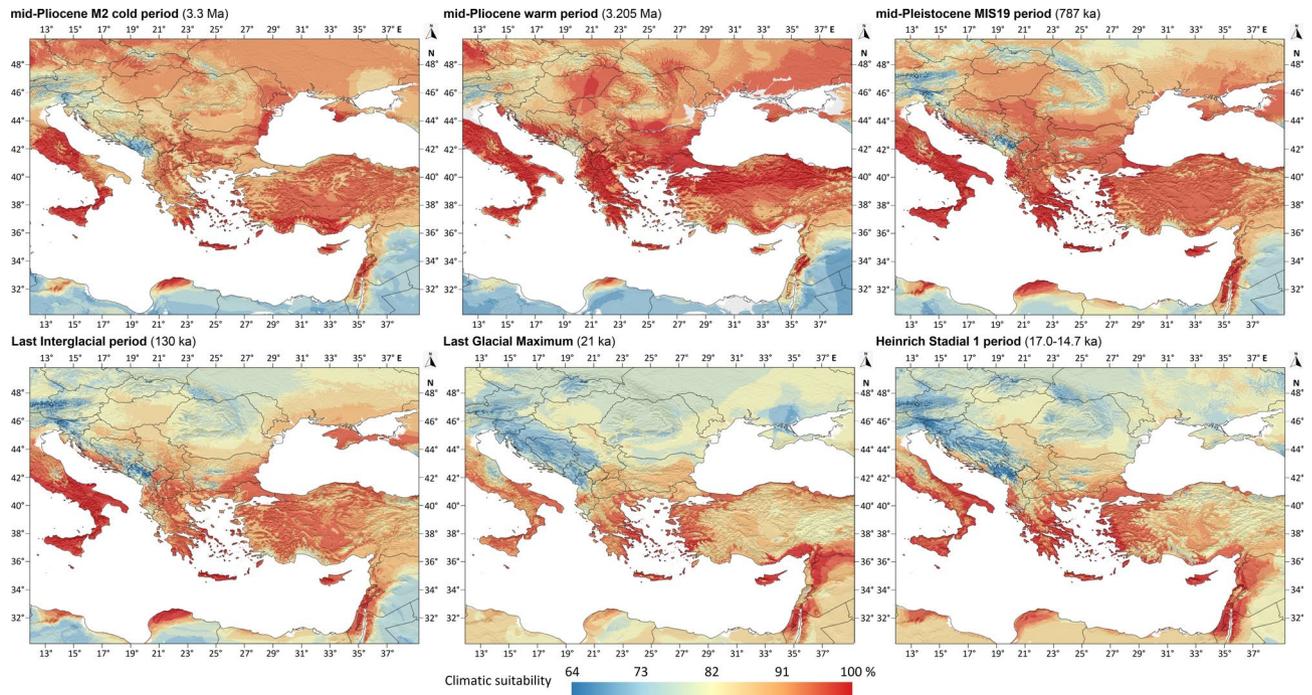
### Early late Miocene conditions

The Tortonian (11.608–7.246 mya) climatic suitability patterns of a *Ph. simici*-ancestor species with climatic limits similar to the present-day *Ph. simici* showed that climate of the large regions of Southeast Europe and the eastern part of the Mediterranean Basin was climatically suitable. These suitable areas included the southeastern part of central Europe, long coastal areas of the Pannonian Lake, the Balkan Peninsula except for the Dinarides and certain Aegean Lowland regions, Crete, the Bulgarian Lowland and ranges, large regions of Asia Minor, certain Levantine regions of the Middle East and the present-day Northern Cyrenaica region of Libya in North Africa (Supplementary Fig. 2).

### Middle-Pliocene conditions

In the M2 cold period of the Piacenzian (3.3 mya), the potential high climatic suitability areas of a *Ph. simici* could have been restricted to the southern regions of Asia Minor, Cyprus, the Aegean Archipelago, the Levant, the Crimean Peninsula and to the present-day western Romania in Southeast Europe and the Eastern Mediterranean Basin (Fig. 4a).

In contrast, in the warm period of the Piacenzian (3.205 mya), *Ph. simici* could have inhabited the large regions of the East Mediterranean basin including the area of the present-day continental Greece, the Aegean Archipelago, Crete, Cyprus, the North-central regions of Asia Minor, and certain regions of the Levant (Fig. 4b).



**Fig. 4.** The climatic suitability patterns of *Ph. simici* in Southeast Europe and the Eastern Part of the Mediterranean Basin from the Piacenzian to the Heinrich Stadial 1. Ancient shorelines were not depicted on the maps.

#### Quaternary conditions

In the MIS19 (787 kya) period, the modelled climatic suitability of *Ph. simici* was highest in the coastal regions of the Southern Balkan, including the whole Peloponnese Peninsula, in the Aegean Islands, including Crete, in the western coastal region of Asia Minor, in the Bosphorus Region, in Cyprus, the Levantine region of the Middle east and in Northern Cyrenaica in North Africa (Fig. 4c). Compared to the MIS19 interglacial, in the Last Interglacial (130 kya), the high suitability regions were notably narrower but already covered southern parts of the Balkan and Asia Minor. The maximum suitability regions were restricted to the Aegean lands, Cyprus, and the Levant (Fig. 4d). In the Last Interglacial, the climatic suitability area was only restricted to the coastal regions of the Aegean Sea, including Euboea, Boeotia and Attica in the mainland Greece, Crete, the valley of river Büyük Menderes and the Adana Lowland in Asia Minor, as well Adana Lowland in Asia Minor, and the Levant. These regions together form the potential glacial refugia of *Ph. simici* (Fig. 4e). For the Heinrich Stadial 1, the extension of the high suitability regions increased in the South Balkan, the Aegean region, and the western part of Asia Minor, but somewhat decreased in the northern part of the Levant (Fig. 4f).

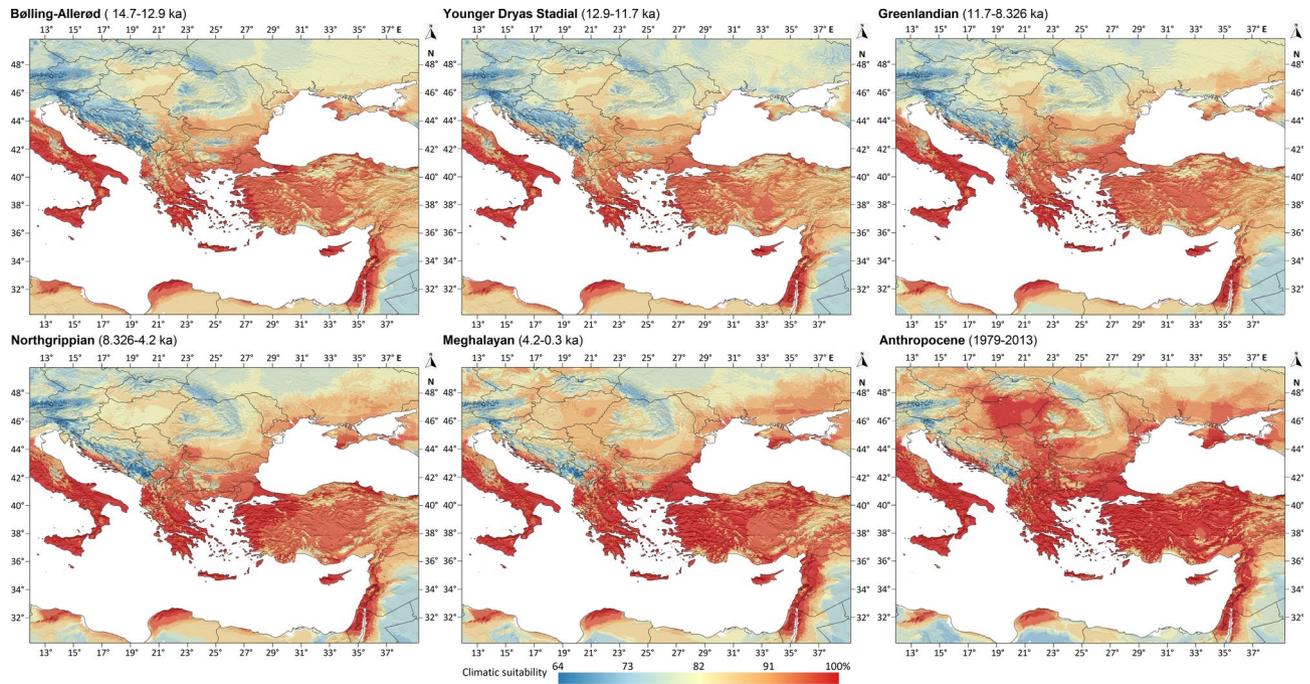
Between the Bølling-Allerød (14.7–12.9 kya) and the Greenlandian stages (11.7–8.326 kya) (Fig. 5a–c), the potential climatic suitability values showed no notable range expansion compared to the above-presented Heinrich Stadial 1. The spread of *Ph. simici* could have started in the Northgrippian (8.326–4.2 kya) (Fig. 5d), and the rate of the territorial expansion of the suitable regions notably increased between the Meghalayan (4.2–0.3 kya) and the Anthropocene stages (Fig. 5e,f). In the studied region, currently the high climatic suitability regions of *Ph. simici* cover almost the entire Balkans, Asia Minor, the Levant, and show increasing habitability tendency also towards the Carpathian Basin.

The 95% climatic suitability areas show that around the Last Glacial maximum's peak, the climatic suitability of *Ph. simici* were stable and mainly restricted to the per-Aegean region and the Levant. However, the potential spread of the species could have been slow until the end of the Younger Dryas Stadial. In the Holocene, the most rapid enlargement of the suitable climate habitats can be observed only in the last some hundred years (Fig. 6).

#### Discussion

Despite excluding some *Adlerius* species such as *Ph. kabulensis*, *Ph. comatus*, or *Ph. fengi* due to a lack of available sequence data, this study provides the most comprehensive phylogenetic and divergence time analyses of the subgenus *Adlerius* based on the frequently used and readily available marker gene cytochrome oxidase subunit I (COI), which should support our understanding of diversification of this understudied subgenus. Additionally, we phylogeographically assessed the emergence of three geographically distinct lineages of *Phlebotomus simici*, a suspected vector for *L. infantum*, combined with climate suitability modeling to elucidate its post-glacial spread and potential current distribution.

Inferred divergence times were based on substitution rates of 1.0% and 2.5%/my, a minimum and maximum of a range that covers the rates typically inferred from/employed for mitochondrial protein coding genes in insects<sup>28–30</sup>. Phylogenetic relationships resembled the presently known geographic distribution of the analyzed



**Fig. 5.** The climatic suitability patterns of *Ph. simici* in Southeast Europe and the Eastern Part of the Mediterranean Basin from the Bølling-Allerød to the Anthropocene. Ancient shorelines were not depicted on the maps.

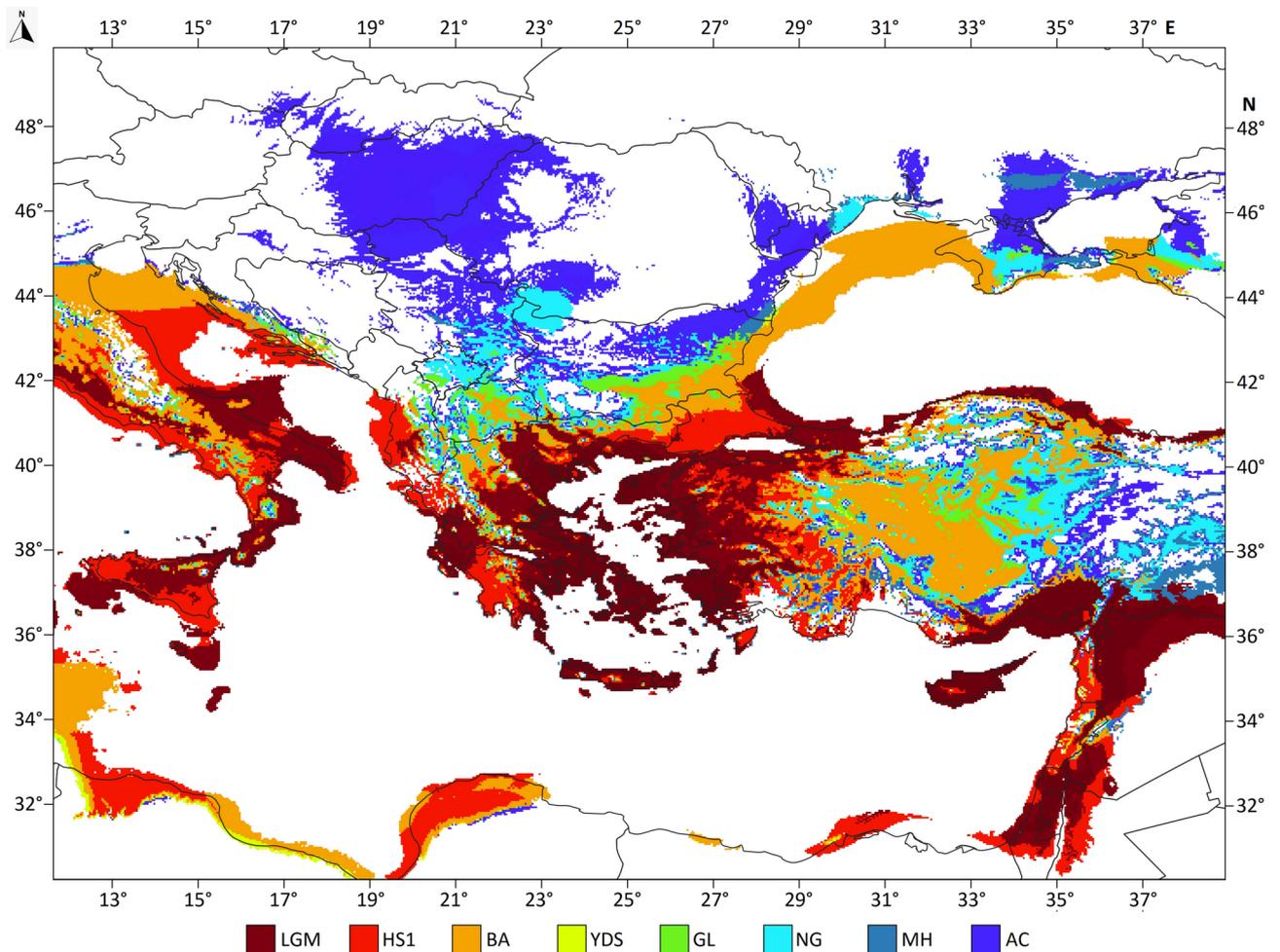
species and provided insights into their divergence. Interestingly, the first major clade, which includes the easternmost species *Ph. simici* along with *Ph. brevis* and an unidentified *Adlerius* species from Türkiye and Armenia, diverged early (10 mya based on 2.5% and 25 mya based on 1.0% substitution rate) from the second major clade. This was also observed by Pavlou et al.<sup>20</sup>, who calculated a Cytb-based split of 8.79 mya between *Ph. simici* and all other analyzed *Adlerius* species.

The diversification of species comprising major clade II likely followed a western direction of dispersal, with the Chinese sand fly species *Ph. chinensis* and *Ph. sichuanensis* representing the geographically easternmost clade. The ancestor of *Ph. longiductus* and *Ph. turanicus* potentially lived on the southeastern coasts of the Paratethys Sea. However, it is currently difficult to find a climatically and geographically analogous area in relation to the late Miocene Central Paratethys. As it was shown, an ancestral species with similar climatic needs to present-day *Ph. simici* could have occupied the coasts and peninsulas of the Pannonian Lake. Another study reported similar findings for other Western Eurasian sand fly taxa of three other subgenera, namely *Larrousius*, *Paraphlebotomus* and *Phlebotomus*<sup>22</sup>. The joint analysis of weathering-produced sediments and paleoclimatic reconstructions indicates that the coastal regions of the Pannonian Lake had subtropical climatic conditions similar to the subtropical regions of the eastern United States or Southeast China<sup>31</sup>. Currently, large parts of the Japanese archipelago have similar climatic and geographical features, but, possibly due to the relative isolation of the Japanese archipelago from the mainland, it is home to only one sand fly species, *Sergentomyia squamirostris*<sup>32</sup>.

The geographical split between those two clades is very well visible between *Ph. chinensis* inhabiting central and eastern parts of China and *Ph. longiductus* being endemic in the north-eastern part of the country<sup>13</sup>. *Phlebotomus halepensis* and *Ph. arabicus* existed in the middle eastern side of the Paratethys Sea and split around 6.25 to 2.5 mya based on calculated divergence times from the clade comprising *Ph. aff. balcanicus* from Azerbaijan, *Ph. kyreniae* and *Ph. balcanicus* from Türkiye as well as *Ph. balcanicus* from the Balkans, altogether resembling the geographically westernmost clade of major clade II. The projected splits of these clades within major clade II might be connected to the formation of the Eurasian Mountain belt, and the climatic and geographical changes in the former Paratethys area which extended from the Northern foreland of the Alps to the Aral Sea in the Oligocene–Miocene epoch, around 33.9–5 mya. Interestingly, the Neogene climatic changes and the Paratethys provided special Eurasian geographical conditions that deeply impacted the phylogeography and climatic needs of current *Transphlebotomus* species<sup>33</sup>.

Interestingly, almost parallel splits can be observed between different lineages, namely *Ph. turanicus* from *Ph. longiductus* and undescribed *Adlerius* sp. from Moldova, *Ph. arabicus* from *Ph. halepensis*, splits within the clade comprising *Ph. aff. balcanicus*, *Ph. balcanicus* and *Ph. kyreniae* as well as *Ph. brevis* + *Adlerius* sp. from *Ph. simici*, which coincide with the Messinian Salinity Crisis 5.96–5.33 mya and the concurrent reflooding, which was also shown to be an important event for the diversification of *Transphlebotomus* species in the Aegean region<sup>34</sup>.

In major clade II, some species urge for taxonomic attention and clarification. Firstly, the yet undescribed *Adlerius* species from Moldova<sup>11</sup> was observed to be the sister species of *Ph. longiductus* + *Ph. turanicus* and potentially split around 6.25 to 2.5 mya. Considering the western Asian and Middle eastern distribution of *Ph. turanicus*<sup>35,36</sup> and the presence of *Ph. longiductus* in e.g. Kazakhstan<sup>37</sup> and as far west as Ukraine<sup>38</sup>, the species from



**Fig. 6.** Changes of the 95% climatic suitable areas for *Ph. simici* between the Last Glacial maximum and the Anthropocene. The overlap of the colorized areas follows the Last Glacial Maximum (uppermost layer) to Anthropocene order. LGM: Last Glacial Maximum (lowermost layer), HS1: Heinrich Stadial 1 (17.0–14.7 kya), BA: Bølling-Allerød (14.7–12.9 kya), YDS: Younger Dryas Stadial (12.9–11.7 kya), GL: Greenlandian (11.7–8.326 kya), NG: Northgrippian (8.326–4.2 kya), MH: Meghalayan (4.2–0.3kya), AC: Anthropocene (1979–2013). The map was created by QGIS 3.34.11 with GRASS GIS v.7.6.1 (<https://grass.osgeo.org>) using country borders from Natural Earth (<https://www.naturalearthdata.com>).

Moldova likely followed a dispersal north of the Black Sea. Secondly, the clade including *Ph. (aff.) balcanicus*, *Ph. creticus* and *Ph. kyreniae* is of high taxonomic importance, as *Ph. balcanicus* is a suspected vector of *L. infantum* causing visceral leishmaniasis (VL) in Georgia<sup>17</sup>. In our analysis, we discriminated three distinct lineages that include *Ph. balcanicus* and *Ph. aff. balcanicus*. The species affinis (*Ph. aff. balcanicus*) found in Azerbaijan was shown to be the sister species of *Ph. kyreniae* and should thus be regarded as a distinct species. Also, *Ph. balcanicus* collected from Türkiye<sup>4</sup> forms a distinct lineage as the sister species of *Ph. balcanicus* including only specimens from Balkan countries<sup>5</sup> + *Ph. creticus*, which has so far been recorded only on the Greek Aegean island of Crete<sup>9</sup>. Similar in this respect to *Ph. simici*, *Ph. balcanicus* is a species with two genetically and geographically distinct lineages that cannot be discriminated morphologically. Based on geographic distribution, *Ph. balcanicus* in Georgia and Türkiye might represent the same lineage, however, no sequence data from Georgia is currently available to test this hypothesis.

Similarly to clade II, species of major clade I likely followed a westward dispersal. This is highlighted by the present-day distribution range of *Ph. simici*, *Ph. brevis* and undescribed *Adlerius* sp. from Middle Eastern countries to South-East Europe and supported by our Tortonian (11.608–7.246 mya) climatic suitability model that showed hypothetical climatic suitability for a *Ph. simici*-like ancient sand fly species as far as the southeastern part of central Europe, long coastal areas of the Pannonian Lake, the Balkan Peninsula except for the Dinarides and certain Aegean Lowland regions, Crete, the Bulgarian Lowland and ranges, large regions of Asia Minor, certain Levantine regions of the Middle East and the present-day Northern Cyrenaica region of Libya in North Africa.

Finally, our analysis provides strong support for the diversification of the three distinct *Ph. simici* lineages. The first and older split, 1.3 to 3.3 mya, between lineage I (Balkan) and the other two clades, lineage II (Israel) + lineage III (Türkiye), well correlates with the end-Neogene climatic changes including aridification

and the increasing fluctuation of the climatic conditions in the Northern Hemisphere. Here, the climatic change between the mid-Pliocene warm (3.3 mya) and cold (3.205 mya) periods potentially supported the diversification since the glacial-interglacial alterations caused severe habitat contractions during the glacial phases and notable geographical extensions during the interglacial periods. Later, in the late Pleistocene, climatic changes caused similar habitat losses and gains for sand fly species in the Mediterranean<sup>39</sup>. Our modelled climatic suitability for *Ph. simici* in the study region changed considerably between these two study periods, however, climatic suitability was stable in coastal areas of south-western Türkiye. Particularly, in the warm period parts of northern Türkiye, mainland Greece and southern Balkan became climatically suitable and thus promoted the further western spread of *Ph. simici*. This coincides with our spatial diffusion analysis, which suggested the most recent common ancestor of *Ph. simici* in southwestern Türkiye and a consequent westward spread of lineage I. It should also be noted that the effect of mid-Pliocene warming could not be uniformly favorable for sand fly species in Europe, since models show that others, such as *Phlebotomus neglectus* or *Phlebotomus papatasi*-like ancient sand fly species, could suffer notable habitat losses due to the more humid climate during certain warm periods<sup>40</sup>. The second and younger split between lineage II and lineage III, 0.8 to 2.0 mya, and likely the diversification of different haplotypes in Türkiye as projected by our spatial diffusion analysis can be correlated to the start of the first glacial events in the early Pleistocene. This abrupt climatic change might coincide with the increasing extremes of climate, which resulted in, among others, the formation of mega-yardangs in central Europe and the appearance of such steppe-dwelling taxa like camelid and struthionid bird species in regions that previously had humid subtropical climate<sup>41</sup>.

Based on our phylogeographic analyses, we observed similar low levels of intra-clade divergence in lineage I (Balkan) and lineage II (Israel) with very recent population expansion, which is supported by the climatic models that showed high climatic suitability during the Last Glacial Maximum (LGM) only in southernmost part of the Balkans (e.g. now southern North Macedonia), coastal Greece including Aegean islands, and small regions of southern Israel. After the last Glacial period, a considerable change in climatic suitability was only noticed from the Northgrippian (8.326–4.2 kya) and after, supporting likely recent post-glacial expansion of *Ph. simici*. Despite having another distribution area, a pattern similarly projected for the post-glacial expansion of *Phlebotomus mascittii*<sup>18</sup>. Interestingly, expansion of the Turkish lineage (III) started well before the LGM between 200 and 700 kya, which is also very well supported by large areas of high climatic suitability during the Chibanian/MIS19 (787 kya) and last interglacial (130 kya) periods in Türkiye. Contrary to the phylogeographic data indicating little impact of LGM on this lineage, our climate models suggest limited climatic suitability in Türkiye during the LGM, with only parts of western Türkiye as well as small regions of coastal northern and southern parts of the country, however, several more microclimatic areas might have been suitable during LGM. This apparent contradiction can be resolved by noting that, although generally relatively small coastal areas remained suitable habitats for *Ph. simici* during the Last Glacial Maximum in the East Mediterranean Region, Figure 6 shows that areas of the eastern Aegean Archipelago and western Asia Minor that are now largely submerged under the sea could have also served as glacial refugia for the species during the Last Glacial Maximum.

Noteworthy, since the first and only record of *Ph. simici* in Austria in 2018<sup>10</sup>, this species was despite multiple trapping efforts never caught again at the same trapping site or any other locality surveyed in Austria for presence of sand flies. Likewise, the species has never been found in any of the bordering countries, albeit it is to note that except for sustained surveillance in Italy and Slovenia, only sporadic and localized surveys were performed there. Our phylogeographic analyses suggested a divergence of the Austrian sample from Balkan samples 60–150 kya. This does not necessarily mean that Austrian *Ph. simici* diverged from their Balkan conspecifics at that time but more likely indicates that we have not yet sampled all haplotypes present in the Balkan region or that the haplotype recorded in Austria is extinct there. However, the (rare) presence of *Ph. simici* in Austria leaves open questions about its dispersal. While we modelled current potential but patchy climatic suitability as far north as small areas of eastern Austria (including the trapping location), southern Czech Republic, eastern Slovakia, and large parts of Hungary, climatic suitability was only shown after the late-Meghalayan for the recent Anthropocene reference period. Considering the absence of *Ph. simici* in Hungary<sup>42</sup> and western Croatia<sup>5</sup> as well as its potentially slow active dispersal ability<sup>43</sup>, anthropogenic dispersal should be taken into account. For example, Vaselek et al.<sup>44</sup> investigated sand fly diversity along clusters of main transit and migration routes in Serbia and found significant changes of the sand fly fauna in Serbia during the end of the 20<sup>th</sup> and beginning of the 21<sup>st</sup> Century. However, the role of anthropogenic dispersal for sand flies remains unclear and calls for further elucidation.

## Conclusion

Our study provides valuable phylogenetic insights into the diversification and geographic appearance of the taxonomically complex and yet understudied subgenus *Adlerius*. We highlight that particularly suspected vector species forming intra-species lineages such as *Ph. balcanicus* and *Ph. simici* are of high medical and veterinary relevance and the emergence and involvement of various lineages in transmission should be carefully assessed.

The combination of phylogeographic analyses and climate modelling revealed a detailed understanding of diversification of three distinct *Ph. simici* lineages and their spread. Particularly, for the lineage comprising the Austrian specimen, the southern Balkan and coastal Greece formed an important glacial refugial area. The combination of the currently known distribution of *Ph. simici* and the modelled climatic suitability in Central Europe urges for further studies on the involvement of anthropogenic and natural factors in sand fly dispersal and spread. Our data clearly highlights the impact of changing climatic conditions to shifts of sand fly distribution and might help to better understand future dispersal connected to climate change, particularly in combination with, if feasible, deep population genetic analyses.

## Material and methods

### Sequence dataset, alignment and editing

For this study, available sequences obtained from GenBank as well as previously unpublished *Ph. simici* sequences were used (Supplementary Table 2). Amplification and sequencing of new specimens followed the protocol previously published by Studentsky et al.<sup>45</sup>. All sequences were aligned with ClustalX<sup>46</sup>, edited in GeneDoc 2.7<sup>47</sup> and translated to amino acid sequences to assure intact reading frames and the absence of internal stop codons. The final alignment consisted of 168 individual *Adlerius* sequences with a length of 655 base pairs (bp), missing nucleotides were replaced with “N”.

### Phylogenetic relationships of *Ph. simici* and temporal patterns of diversification of the subgenus *Adlerius*

Phylogenetic relationships were inferred employing Maximum Likelihood (ML) and Bayesian (BI) approaches in IQ-TREE<sup>48</sup> and MrBayes 3.2.6<sup>49</sup> plugins of the PhyloSuite 1.2.2 package<sup>50</sup>. The best fitting model of evolution to be used in phylogenetic tree inference was selected based on the Bayesian information criterion (BIC) in Modelfinder<sup>51</sup>, as implemented in PhyloSuite. A standard ML tree search was conducted in IQ-TREE and node support was assessed with 1000 standard bootstrap replicates. For Bayesian phylogenetic inference with MrBayes, posterior probabilities were obtained from Metropolis-coupled Markov chain Monte Carlo (MCMCMC) simulations (2 independent runs, 6 chains, 10 million generation, sampling frequency 1000, 25% burn-in). Chain stationarity and parameter convergence (effective sample sizes (ESS) > 200, indicating that the parameter log file accurately reflects the posterior distribution<sup>52</sup>; were checked in Tracer 1.7<sup>53</sup> and via monitoring the average standard deviation of split frequencies (<0.01). Post burn-in trees were summarized in a 50% majority rule consensus tree. Visualization of ML and BI trees was done in FigTree v.1.4.2 (available at <http://tree.bio.ed.ac.uk/software/figtree>).

In addition, we inferred a chronogram in BEAST2.7.3<sup>54</sup>. Prior to this analysis, the originally 168 sequences were collapsed into 109 haplotypes using FaBox 1.61<sup>55</sup>. We employed the best fitting model of evolution as suggested by Modelfinder, an optimized relaxed-clock model, a substitution rate of 0.01 and 0.025 substitutions per site per million years (my)<sup>28,56</sup>, which covers the range of substitution rates typically inferred from/employed for mitochondrial protein coding genes in insects<sup>29,30,57,58</sup>, and a birth–death tree prior. We ran three independent MCMC runs for 150 million generation each and sampled parameters every 1000 generation. LogCombiner (part of the BEAST2 package) was then used to combine the three runs after discarding the first 10% of generations as burn-in. Chain stationarity and parameter convergence were assessed in Tracer. Pooled post-burn-in ESS was > 200 for all parameters. A maximum clade credibility tree was computed from the post-burn-in trees using TreeAnnotator (part of the BEAST2 package) and visualized in FigTree.

### Genetic diversity of *Ph. simici* through space and time

DnaSP v.5<sup>59</sup> was used to calculate unique haplotypes, haplotype (Hd), and nucleotide ( $\pi$ ) diversity. Uncorrected pairwise distances between all species and all *Ph. simici* haplotypes were calculated in MEGAX<sup>60</sup>. Phylogenetic relationships among *Ph. simici* haplotypes were visualized as a statistical parsimony network<sup>61</sup> inferred in PopART<sup>62</sup>. To test for signatures of recent population expansion in major clades of *Ph. simici*, we calculated mismatch distributions (1,000 bootstrap replicates) and the neutrality test statistics Tajima’s D<sup>98</sup> (10,000 simulated samples) and Fu’s Fs<sup>99</sup> (10,000 simulated samples) in Arlequin 3.5.2.2<sup>63</sup>. The fit between the observed mismatch distribution and expectations based on growth parameter estimates was evaluated by the sum of squared differences (SSD) and the raggedness index (rg). In addition, we inferred past population size changes by means of Bayesian skyline plots (BSPs) in BEAST 2.7.3. We employed the best-fitting models of molecular evolution selected by the BIC in Modelfinder, and a strict molecular clock with again a minimum and maximum substitution rate of 0.01 and 0.025 substitutions per site per my, and the Bayesian Skyline tree prior. Four independent MCMC runs of three to twelve million generations (depending on dataset size; adjusted after preliminary runs), each were done, sampling every 1000th step and a burn-in of the first 10% of sampled generations. LogCombiner was used to combine the individual log and tree files. Assessment of run convergence (ESS > 200 for all parameters), and visualization of past population size changes were done in Tracer.

Finally, the species’ spatial history was assessed employing a cauchy relaxed random walk (RRW) diffusion model implemented in BEAST 1.10.4<sup>64</sup>. We used a normally distributed diffusion rate, Bayesian Skyline tree prior, the best-fitting modes of molecular evolution selected by the BIC in Modelfinder, a strict molecular clock with again minimum and maximum substitution rates of 0.01 and 0.025 substitutions per site per my, and a Bayesian Skyline coalescent tree prior. The jitter option was set to 0.01 to create noise in identical coordinates. Runs were run for 300 million generations, sampled every 1000th step and with a burn-in of the first 10% of the sampled generations. Assessment of run performance (ESS > 200 for all parameters) was done in Tracer. A maximum clade credibility obtained in TreeAnnotator (part of the BEAST1 package) was used as input for EvoLaps2 v. 2.42<sup>65</sup> (<https://www.evolaps.org>) to visualize the spatiotemporal pattern.

### Geographical data and mapping

A literature search for records of *Ph. simici* was performed (Google Scholar and PubMed), and coordinates or locations were extracted<sup>14,5,9,10,16,44,66–82</sup>. In addition, a dataset acquired in the frame of the VectorNet project was exploited for presence/absence data<sup>83</sup>. Coordinates of trapping sites were georeferenced into a distribution map using Quantum GIS 3.4.11<sup>84</sup> (Supplementary Fig. 3).

### Climatic and topographic data sources

The climatic classification of the *Ph. simici* sites was determined using the Köppen–Geiger system. For this purpose, the present-day Köppen–Geiger map produced by Beck et al.<sup>85</sup> was used. A total of one climatic

reference period and twelve paleoclimatic data sources were used, covering a period spanning from the early Late Miocene to the Anthropocene. By epochs, one data source provided the Tortonian paleoclimatic data, two models represent the mid-Pliocene epoch, one model covers the mid-Pleistocene (the MIS19 interstadial), five models belong to the “Late” Pleistocene, three models represent the Holocene, and one the Anthropocene, which forms the reference period in this study (Supplementary Table 3). Topographic data were sourced from the ETOPO Global Relief Model<sup>86</sup>.

The Tortonian climatic model was produced by the paleoclimatic reconstruction of Bruch et al.<sup>87</sup>. It contains paleoclimatic reconstructions by fossil sites based on the coexistence analysis of fossil plant assemblages. This approach is used for the quantitative reconstructions of the Tertiary terrestrial paleoclimates. It is based on the idea that tertiary plant taxa require similar climates to those of their closest extant relatives<sup>88</sup>. It is a frequently used method to produce paleoclimatic reconstructions for the Cenozoic period, including even the Paleogene period<sup>89</sup>. To create the Tortonian paleoclimatic model for Europe and the neighboring regions (see details in Supplementary Table 4), the following steps were conducted:

- (1) The Tortonian plant fossil sites were georeferenced with the corresponding paleoclimatic data in the attribute table using QGIS version 3.8.3 with GRASS GIS version 7.6.1.
- (2) After that the differences between the Tortonian paleoclimatic<sup>87</sup> and the reference period's<sup>90</sup> climatic values were calculated.
- (3) In the next step, the calculated difference values were georeferenced by climatic factors, and the point-like difference values were IDW-interpolated.
- (4) In the final step, the original reference period's values were modified by the IDW-interpolated values.

### IDW interpolation of paleoclimatic values

As mentioned before, to produce paleoclimatic maps for the Tortonian, an Inverse Distance Weighting (IDW) method was used in QGIS<sup>91</sup>, which is a spatial interpolation method used to estimate values at unsampled locations based on values from nearby sampled points. The core idea of IDW is that points closer to the estimation location have more influence on the predicted value than points further away. This method assumes that the influence of the interpolated variable decreases with distance, hence the term “inverse distance” (Supplementary Fig. 4). IDW calculates the value at an unsampled location as a weighted average of the values of nearby points. The weights are inversely proportional to the distance between the known points and the point being estimated, giving more influence on closer points. The influence of the distance is controlled by a power parameter  $p$ . A higher  $p$  value increases the influence of nearer points and reduces the influence of farther points more sharply. The method typically uses a set number of nearest neighbors or a fixed radius to include surrounding points in the interpolation calculation.

The value  $V(x)$  at an unsampled location  $x$  is calculated using the following equation:

$$V(x) = \frac{\sum_{i=1}^n \frac{V_i}{d_i^p}}{\sum_{i=1}^n \frac{1}{d_i^p}} \quad (1)$$

where  $V(x)$  is the interpolated value at location  $x$ ;  $V_i$  is the known value at the  $i$ -th sampled point;  $d_i$  is the distance between the unsampled location  $x$  and the  $i$ -th sampled point;  $p$  is the power parameter, typically a positive value that determines the weight of the distance (common values are between 1 and 3);  $n$  is the number of sampled points used for interpolation.

Because the original paleoclimatic reconstruction of Bruch et al.<sup>87</sup> contains only four factors (the mean annual temperature, the mean temperature of the coldest months, the mean temperature of the warmest months, and the annual precipitation (sum), two additional factors were created based on the existing ones to enhance the precision of climatic suitability models:

$$MATR = Tm_{07} - Tm_{01} \quad (2)$$

where  $MATR$  is the mean annual temperature range,  $Tm_{07}$  and  $Tm_{01}$  are the mean July and January temperature. The unit is °C.

This derived factor characterizes the fluctuations of the thermal conditions during the year, which can be an important factor for sand fly seasonality and occurrence<sup>92</sup>.

In addition, the Thornthwaite Agrometeorological Index<sup>93</sup> was introduced to characterize the general humidity conditions which is also an important factor for sand fly activity and occurrence<sup>94</sup>. The applied equation was as follows based on the modified equation of<sup>95</sup>:

$$TAI = 1.65 \times \frac{\frac{bio12}{12}^{\frac{10}{9}}}{bio1 + 12.2} \quad (3)$$

where  $TAI$  is the Thornthwaite Agrometeorological Index,  $bio1$  is the mean annual temperature,  $bio12$  is the annual precipitation sum. The unit is  $\text{mm } ^\circ\text{C}^{-1}$ .

### Determination of lower and upper extrema

Climate values at *Ph. simici* occurrence points were sampled across 16 bioclimatic layers (2 monthly and 14 bioclimatic maps) using the QGIS sampling tool. To refine the data, the highest and lowest 1% of values for

each bioclimatic variable were excluded before determining the upper and lower extremes. This process ensured sampling based on the known occurrences of *Ph. simici* and improved data quality.

Formally, let  $X$  be the set of climate values sampled at occurrence points across  $n$  climatic variables  $\{X_1, X_2, \dots, X_n\}$ . For each variable  $X_i$ , let  $X_i^{(sorted)}$  represent the ordered values of  $X_i$ .

To determine the purified data for each variable  $X_i$ :

$$X_i^{cleaned} = X_i^{(sorted)} [k], X_i^{(sorted)} [k + 1], \dots, X_i^{(sorted)} [m - k] \quad (4)$$

where  $k = \lceil 0.01 \times m \rceil$  and  $m$  is the total number of samples in  $X_i$ .

### Model identification: modelling of climatic suitability patterns

Climate Envelope Method (CEM) forms the basic concept of the creation of climatic suitability methods. Climate envelope modelling is an approach used to predict the potential distribution of a species or phenomenon based on the climatic conditions that are currently associated with its presence<sup>96</sup>. This method assumes that the current distribution of a species is largely determined by a set of climatic variables (e.g. temperature, precipitation) that fall within specific upper and lower thresholds suitable for the species. By defining these climatic “envelopes,” or ranges, climate envelope models can be used to predict where the species could potentially exist under different climate scenarios.

The mathematical basis of CEM is the Boolean algebra<sup>97</sup>. In Boolean climate envelope modelling, the presence or absence of a species in a given area is determined by whether the climatic conditions fall within the specified limits for each variable. The presence-absence maps for each climatic factor are combined using Boolean algebra to generate an overall suitability map. The general Boolean equation for the climate envelope model can be expressed as:

$$S(x) = \bigwedge_{i=1}^n (L_i \leq V_i(x) \leq U_i) \quad (5)$$

where  $S_x$  is the overall suitability (1 for presence, 0 for absence) at location,  $n$  is the number of climatic variables which is 6 in the case of the Tortonian and 14 in the case of the Pliocene, Pleistocene and Holocene models (see Supplementary Table 3);  $V_i(x)$  is the value of the  $i$ -th climatic variable at location  $x$ ;  $L_i$  and  $U_i$  are the lower and upper limits for the  $i$ -th climatic variable;  $\bigwedge$  represents the logical AND operation.

Each variable's presence-absence condition is defined as 1 (suitable) if  $L_i \leq V_i(x) \leq U_i$ , otherwise 0 (unsuitable). The overall suitability map is obtained by the intersection (logical operator AND,  $\bigwedge$ ) of all these individual presence-absence maps. There were three kinds of climatic models: temperature like (T), precipitation-like (P), and humidity-like (H) ones.

Two model environments were built: one for the Tortonian and another one for the other climatic suitability models. The model equation for the Tortonian is as follows:

$$S(x) = \left( \bigwedge_{i=1}^4 (L_{Ti} \leq T_i(x) \leq U_{Ti}) \right) \bigwedge (L_{P1} \leq P_1(x) \leq U_{P1}) \bigwedge (L_{H1} \leq H_1(x) \leq U_{H1}) \quad (6)$$

where  $S(x)$  is the overall climatic suitability at location  $x$ ;  $T_i(x)$  is the value of the  $i$ -th temperature factor at location  $x$ ;  $P_j(x)$  is the value of the precipitation-like factor at location  $x$ ;  $H_j(x)$  is the value of the humidity-like factor at location  $x$ ;  $L_{Ti}$  and  $U_{Ti}$  are the lower and upper limits for the  $i$ -th temperature-like factor;  $L_{Pj}$  and  $U_{Pj}$  are the lower and upper limits for the precipitation-like factor;  $L_{Hj}$  and  $U_{Hj}$  are the lower and upper limits for the humidity-like factor.  $\bigwedge$  represents the logical AND operation.

This equation ensures that the suitability at a location  $x$  is determined only if all four temperature-like factors, the precipitation-like factor, and the humidity-like factor fall within their respective suitable ranges. The model equation for the Pliocene, Quaternary, and Holocene periods is as follows:

$$S(x) = \left( \bigwedge_{i=1}^6 (L_{Ti} \leq T_i(x) \leq U_{Ti}) \right) \bigwedge \left( \bigwedge_{j=1}^8 (L_{Pj} \leq P_j(x) \leq U_{Pj}) \right) \quad (7)$$

where  $S(x)$  is the overall climatic suitability at location  $x$ ;  $T_i(x)$  is the value of the  $i$ -th temperature factor at location  $x$ ;  $P_j(x)$  is the value of the  $j$ -th precipitation-like factor at location  $x$ ;  $L_{Ti}$  and  $U_{Ti}$  are the lower and upper limits for the  $i$ -th temperature factor;  $L_{Pj}$  and  $U_{Pj}$  are the lower and upper limits for the  $j$ -th precipitation-like factor.

This above-presented equation combines the suitability conditions for all temperature and precipitation-like factors to produce a final climatic suitability map. The species or phenomenon is considered present at a location  $x$  only if all the climatic variables fall within their specified suitable ranges simultaneously. Supplementary Table 5 shows the used climatic factors according to the two model environments.

### Data availability

All data are included in the article and supplementary material.

Received: 26 November 2024; Accepted: 14 March 2025

Published online: 25 March 2025

## References

- Cecílio, P., Cordeiro-da-Silva, A. & Oliveira, F. Sand flies: Basic information on the vectors of leishmaniasis and their interactions with *Leishmania* parasites. *Commun. Biol.* **5**, 305 (2022).
- Jia, L. Y. & Min, Z. L. Chinese phlebotomine sandflies of subgenus *Adlerius* NITZULESCU, 1931, (Diptera: Psychodidae) and the identity of *Phlebotomus sichuanensis* LENG & Yin, 1983, Part I—Taxonomical study and geographical distribution. *Parasite* **8**, 3–9 (2001).
- Jacobson, R. L. Leishmaniasis in an era of conflict in the Middle East. *Vector-Borne Zoon. Dis.* **11**, 247–258 (2011).
- Kasap, O. E., Linton, Y.-M., Karakus, M., Ozbel, Y. & Alten, B. Revision of the species composition and distribution of Turkish sand flies using DNA barcodes. *Parasit. Vectors* **12**, 410 (2019).
- Dvorak, V. et al. Sand flies (Diptera: Psychodidae) in eight Balkan countries: Historical review and region-wide entomological survey. *Parasit. Vectors* **13**, 573 (2020).
- Zhang, L., Ma, Y. & Xu, J. Genetic differentiation between sandfly populations of *Phlebotomus chinensis* and *Phlebotomus sichuanensis* (Diptera: Psychodidae) in China inferred by microsatellites. *Parasit. Vectors* **6**, 115 (2013).
- Perrotey, S., Benabdennbi, I., Haddad, N., Pesson, B. & Léger, N. Electrophoretic and morphological differentiation between two sympatric species of *Adlerius*: *Phlebotomus brevis* and *Phlebotomus simici* (Diptera: Psychodidae). *J. Med. Entomol.* **37**, 289–294 (2009).
- Artemiev, M. M. A revision of sandflies of the subgenus *Adlerius* (Diptera, Phlebotominae, Phlebotomus). *Zool. Zhurnal* **59**, 1177–1192 (1980).
- Dvořák, V. et al. Sand fly fauna of Crete and the description of *Phlebotomus (Adlerius) creticus* n. sp. (Diptera: Psychodidae). *Parasit. Vectors* **13**, 1–17 (2020).
- Kniha, E. et al. *Phlebotomus (Adlerius) simici* NITZULESCU, 1931: First record in Austria and phylogenetic relationship with other *Adlerius* species. *Parasit. Vectors* **14**, 20 (2021).
- Șuleșco, T. et al. Phlebotomine sand fly survey in the Republic of Moldova: Species composition, distribution and host preferences. *Parasit. Vectors* **14**, 371 (2021).
- Pareyn, M. et al. An integrative approach to identify sand fly vectors of leishmaniasis in Ethiopia by morphological and molecular techniques. *Parasit. Vectors* **13**, 580 (2020).
- Guan, L.-R., Zhou, Z.-B., Jin, C.-F., Fu, Q. & Chai, J.-J. Phlebotomine sand flies (Diptera: Psychodidae) transmitting visceral leishmaniasis and their geographical distribution in China: A review. *Infect. Dis. Poverty* **5**, 15 (2016).
- Svobodová, M., Volf, P. & Votýpka, J. Experimental transmission of *Leishmania tropica* to hyraxes (*Procapra capensis*) by the bite of *Phlebotomus arabicus*. *Microbes Infect.* **8**, 1691–1694 (2006).
- Sádllová, J., Hajmová, M. & Volf, P. *Phlebotomus (Adlerius) halepensis* vector competence for *Leishmania major* and *Le. tropica*. *Med. Vet. Entomol.* **17**, 244–250 (2003).
- 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).
- Giorgobiani, E. et al. Incrimination of *Phlebotomus kandelakii* and *Phlebotomus balcanicus* as vectors of *Leishmania infantum* in Tbilisi, Georgia. *PLoS Negl. Trop. Dis.* **6** (2012).
- Kniha, E. et al. Reconstructing the post-glacial spread of the sand fly *Phlebotomus mascittii* Grassi, 1908 (Diptera: Psychodidae) in Europe. *Commun. Biol.* **6**, 1244 (2023).
- 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).
- 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).
- Alten, B. Speciation and dispersion hypotheses of phlebotomine sandflies of the subgenus *Paraphlebotomus* (Diptera: Psychodidae): The case in Turkey. *Hacettepe J. Biol. Chem.* **38**, 229–246 (2010).
- Trájer, A. J. Investigation of the possible role of the Central Paratethys as a migration route and speciation area of the ancestors of Mediterranean *Larrousius*, *Paraphlebotomus* and *Phlebotomus* species. *Palaeobiodiversity Palaeoenvironments* **103**, 165–192 (2023).
- Trájer, A. J. Palaeoclimatic models - predicted changes in the potential Neogene distribution patterns of *Phlebotomus similis* and *Phlebotomus sergenti* (Insecta: Diptera: Psychodidae). *Palaeobiodiversity Palaeoenvironments* **102**, 149–172 (2022).
- Essegir, S. & Ready, P. D. Speciation of *Phlebotomus* sandflies of the subgenus *Larrousius* coincided with the late Miocene-Pliocene aridification of the Mediterranean subregion. *Biol. J. Linn. Soc.* **70**, 189–219 (2000).
- García-Castellanos, D. et al. Catastrophic flood of the Mediterranean after the Messinian salinity crisis. *Nature* **462**, 778–781 (2009).
- Trájer, A. J., Sebestyén, V. & Padisák, J. The impacts of the Messinian Salinity Crisis on the biogeography of three Mediterranean sandfly (Diptera: Psychodidae) species. *Geobios* **65**, 51–66 (2021).
- Trájer, A. J., Hammer, T. & Padisák, 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).
- 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).
- Farrell, B. D. Evolutionary assembly of the milkweed fauna: Cytochrome oxidase I and the age of tetraopes beetles. *Mol. Phylogenet. Evol.* **18**, 467–478 (2001).
- 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*. **58**, 554–570 (2004).
- Trájer, A. J. Characterization of the palaeoenvironmental conditions in the Pannonian Basin during the last 34 mys related to the formation of haematitic and kaolinitic sedimentary rocks. *Int. J. Earth Sci.* **112**, 1361–1387 (2023).
- Sanjoba, C., Omachi, S., Sato, K. & Matsumoto, Y. Additional distribution record of *Sergentomyia (Neophlebotomus) squamirostris* (Newstead) (Diptera: Psychodidae) from Tokyo, Japan. *Med. Entomol. Zool.* **68**, 45–48 (2017).
- Trájer, A. J., Walochnik, J. & Kniha, E. The possible region of the Late Miocene split of the sandfly subgenus *Transphlebotomus* Artemiev and the early late Neogene to late Quaternary dispersal of the ancestor of *Phlebotomus mascittii* Grassi. *Palaeobiodiversity Palaeoenvironments* <https://doi.org/10.1007/s12549-022-00570-y> (2023).
- 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).
- Usarov, G. X. et al. Phlebotomine sand fly (Diptera: Phlebotominae) diversity in the foci of cutaneous leishmaniasis in the Surxondaryo Region of Uzbekistan: 50 years on. *Parasitol. Res.* **123**, 170 (2024).
- Akhoundi, M., Parvizi, P., Baghaei, A. & Depaquit, J. The subgenus *Adlerius* Nitzulescu (Diptera, Psychodidae, *Phlebotomus*) in Iran. *Acta Trop.* **122**, 7–15 (2012).

37. Dergacheva, T. I. & Strelkova, M. V. Epidemiological role of sandflies *Phlebotomus smirnovi* Perfiliev, 1941 and *P. longiductus* Parrot, 1928 in visceral leishmaniasis foci in the Kazakh SSR. *Trans. R. Soc. Trop. Med. Hyg.* **79**, 34–36 (1985).
38. Ježek, J., Chvojka, P., Manko, P. & Oboňa, J. Faunistic and bibliographical inventory of moth flies from Ukraine (Diptera, Psychodidae). *Zookeys* **693**, 109–128 (2017).
39. Trájer, A. J. & Sebestyén, V. The changing distribution of *Leishmania infantum* Nicolle, 1908 and its Mediterranean sandfly vectors in the last 140 yrs. *Sci. Rep.* **9**, 11820 (2019).
40. Trájer, A. J. Paradox negative effects of the mid-pliocene warming on the climatic suitability of six Mediterranean sandfly species in Europe. *Biosis: Biol. Syst.* **1**, 141–156 (2020).
41. Trájer, A. J. Age and environmental conditions for the formation of the Pannonian mega-yardang system. *Int. J. Earth Sci.* **113**, 875–901 (2024).
42. 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).
43. 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).
44. Vaselek, S. et al. A survey of sand flies (Diptera, Phlebotominae) along recurrent transit routes in Serbia. *Acta Trop.* **197**, 105063 (2019).
45. Studentsky, L. et al. *Leishmania donovani* transmission cycle associated with human infection, *Phlebotomus alexandri* sand flies, and hare blood meals, Israel. *Emerg. Infect. Dis.* **29**, 945–955 (2023).
46. Larkin, M. A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
47. Nicholas, K. B. Genedoc: A tool for editing and annotating multiple sequence alignments. <http://www.pscedu/biomed/genedoc>. (1997).
48. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
49. Ronquist, F. et al. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
50. Zhang, D. et al. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **20**, 348–355 (2020).
51. 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).
52. Kuhner, M. K. Coalescent genealogy samplers: Windows into population history. *Trends Ecol. Evol.* **24**, 86–93 (2009).
53. 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).
54. Bouckaert, R. et al. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* **15**, e1006650 (2019).
55. Villesen, P. FaBox: An online toolbox for FASTA sequences. *Mol. Ecol. Notes* **7**, 965–968 (2007).
56. 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).
57. Essegir, S., Ready, P. D., Killick-Kendrick, R. & Ben-Ismaïl, R. Mitochondrial haplotypes and phylogeography of *Phlebotomus* vectors of *Leishmania major*. *Insect Mol. Biol.* **6**, 211–225 (1997).
58. López-López, A., Abdul Aziz, A. & Galián, J. Molecular phylogeny and divergence time estimation of *Cosmodela* (Coleoptera: Carabidae: Cicindelinae) tiger beetle species from Southeast Asia. *Zool. Scr.* **44**, 437–445 (2015).
59. Librado, P. & Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452 (2009).
60. Kumar, S., Stecher, G., Li, M., Nnyaz, C. & Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
61. 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).
62. Leigh, J. W. & Bryant, D. POPART: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**, 1110–1116 (2015).
63. 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).
64. Suchard, M. A. et al. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, 1–5 (2018).
65. Chevenet, F., Fargette, D., Bastide, P., Vitré, T. & Guindon, S. EvoLaps 2: Advanced phylogeographic visualization. *Virus Evol.* **10** (2024).
66. Xhekaj, B. et al. A cross-sectional study on phlebotomine sand flies in relation to disease transmission in the Republic of Kosovo. *Med. Vet. Entomol.* **38**, 573–585 (2024).
67. Boutsini, S. et al. Phlebotomine sandflies and factors associated with their abundance in the leishmaniasis endemic area of Attiki, Greece. *Parasitol. Res.* **117**, 107–113 (2018).
68. Aransay, A. M., Scoulica, E. & Tselentis, Y. Detection and identification of *Leishmania* DNA within naturally infected sand flies by seminested PCR on minicircle kinetoplastic DNA. *Appl. Environ. Microbiol.* **66**, 1933–1938 (2000).
69. Xanthopoulou, K. et al. Distribution of sandflies (Diptera, Psychodidae) in two Ionian Islands and Northern Greece. *Vector-Borne Zoon. Dis.* **11**, 1591–1594 (2011).
70. Fotakis, E. A., Giantsis, I. A., Demir, S., Vontas, J. G. & Chaskopoulou, A. Detection of pyrethroid resistance mutations in the major Leishmaniasis vector *Phlebotomus papatasi*. *J. Med. Entomol.* **55**, 1225–1230 (2018).
71. Tsigotakis, N. et al. Phlebotomine sand flies (Diptera: Psychodidae) in the Greek Aegean Islands: Ecological approaches. *Parasit. Vectors* **11**, 1–14 (2018).
72. Karanis, P. et al. Sandfly (Diptera: Psychodidae) distribution in Northern Greece. *Entomol. Hell.* **13**, 13 (2017).
73. Svobodova, M. et al. Distinct transmission cycles of *Leishmania tropica* in 2 Adjacent Foci, Northern Israel. *Emerg. Infect. Dis.* **12**, 1860–1868 (2006).
74. Vaselek, S. et al. Sandfly surveillance and investigation of *Leishmania* spp. DNA in sandflies in Kosovo. *Med. Vet. Entomol.* **34**, 394–401 (2020).
75. Yaman, M. & Özbel, Y. The sandflies (Diptera: Psychodidae) in the Turkish province of Hatay: Some possible vectors of the parasites causing human cutaneous leishmaniasis. *Ann. Trop. Med. Parasitol.* **98**, 741–750 (2004).
76. Kavur, H., Arikan, H. & Özbel, Y. *Phlebotomus halepensis* (Diptera: Psychodidae) vectorial capacity in Afyon and Niğde Province, Turkey. *J. Med. Entomol.* **55**, 317–322 (2018).
77. Omond, Z. N., Demir, S. & Arserim, S. K. Entomological survey of the sand fly fauna of Kayseri Province: Focus on visceral and cutaneous leishmaniasis in Central Anatolia, Turkey. *Turkiye parazitolojii Derg.* **44**, 158–163 (2020).
78. Ergunay, K. et al. Molecular evidence indicates that *Phlebotomus major* sensu lato (Diptera: Psychodidae) is the vector species of the recently-identified sandfly fever sicilian virus variant: Sandfly fever Turkey virus. *Vector-Borne Zoon. Dis.* **12**, 690–698 (2012).
79. Léger, N., Pesson, B. & Madulo-Leblond, G. Les phlébotomes de Grèce. *Bull. Soc. Pathol. Exot.* **79**, 386–397 (1986).
80. Léger, N. & Pesson, B. Sur la taxonomie et al répartition géographique de *Phlebotomus* (Adlerius) chinensis s. l. et de P. (Larrousius) major s. l. (Psychodidae-Diptera). *Bull. Soc. Pathol. Exot.* **80**, 252–260 (1987).

81. Leger, N. et al. Les phlébotomes de Crète. *Biol. Gall.* **20**, 135–143 (1993).
82. Depaquit, J., Ferte, H. & Leger, N. Les phlébotomes (Diptera - Psychodidae) de l'île de Rhodes (Grèce). *Bull. la Société Française Parasitol.* **14**, 28–38 (1996).
83. Wint, G. R. W. et al. VectorNet: Collaborative mapping of arthropod disease vectors in Europe and surrounding areas since 2010. *Eurosurveillance* **28** (2023).
84. QGIS Development Team. QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org> (2019).
85. Beck, H. E. et al. Present and future Köppen–Geiger climate classification maps at 1-km resolution. *Sci. Data* **5**, 180214 (2018).
86. MacFerrin, M. et al. ETOPO 2022: An updated NOAA global relief model. *AGUFM* **2021**, NH25B-0560 (2021).
87. Bruch, A. A., Uhl, D. & Mosbrugger, V. Miocene climate in Europe—Patterns and evolution. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **253**, 1–7 (2007).
88. Mosbrugger, V. & Utescher, T. The coexistence approach—a method for quantitative reconstructions of tertiary terrestrial palaeoclimate data using plant fossils. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **134**, 61–86 (1997).
89. Trájer, A. J. The correspondence between the physiological cold tolerance and the distribution border of Mediterranean sandflies in Southeast Europe. *J. Vector Borne Dis.* **61**, 376–388 (2024).
90. Karger, D. N. et al. CHELSA-TraCE21k v1.0. Downscaled transient temperature and precipitation data since the last glacial maximum. *Clim. Past Discuss.*, 1–27 (2021).
91. le Roux, R., Henrico, S., Bezuidenhout, J. & Henrico, I. Inverse distance weighting as an alternative interpolation method to create radiometric maps of natural radionuclide concentrations using QGIS. *Proc. ICA* **5**, 1–7 (2023).
92. Boussaa, S., Guernaoui, S., Pesson, B. & Boumezzough, A. Seasonal fluctuations of phlebotomine sand fly populations (Diptera: Psychodidae) in the urban area of Marrakech, Morocco. *Acta Trop.* **95**, 86–91 (2005).
93. Thornthwaite, C. W. An approach toward a rational classification of climate. *Geogr. Rev.* **38**, 55 (1948).
94. Nayakarathna, N. M. N. G., Ganehiarachchi, G. A. S. M., Rajapakse, R. P. V. J. & Jayanetti, S. R. Influence of humidity and temperature variation of natural breeding sites on abundance of Leishmaniasis vector *Phlebotomus argentipes* population in Anuradhapura District. In *Proceedings of the International Research Symposium on Pure and Applied Sciences 12* (Faculty of Science, University of Kelaniya, 2016).
95. Kemp, D. *Global Environmental Issues* (Routledge, 2002). <https://doi.org/10.4324/9780203425305>
96. Hijmans, R. J. & Graham, C. H. The ability of climate envelope models to predict the effect of climate change on species distributions. *Glob. Change Biol.* **12**, 2272–2281 (2006).
97. Whitesitt, J. E. *Boolean Algebra and Its Applications* (Courier Corporation, 2010).
98. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
99. Fu, Y.-X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915–925 (1997).

## Acknowledgements

This study is co-funded by European Commission grant 101057690 and UKRI grants 10038150 and 10039289, and is catalogued by the CLIMOS Scientific Committee as CLIMOS number 021 (<http://www.climos-project.eu>). The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission, the Health and Digital Executive Agency, or UKRI. Neither the European Union nor granting authority nor UKRI can be held responsible. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. For the purposes of Open Access, the authors have applied a CC BY [option: CC BY-ND] public copyright license to any Author Accepted Manuscript version arising from this submission. The six Horizon Europe projects, BlueAdapt, CATALYSE, CLIMOS, HIGH Horizons, IDAlert, and TRIGGER, form the Climate Change and Health Cluster. This paper was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (bo/00896/24/5). This work has been also supported by the National Multidisciplinary Laboratory for Climate Change (RRF-2.3.1-21-2022-00014) project within the framework of Hungary's National Recovery and Resilience Plan supported by the Recovery and Resilience Facility of the European Union. Sand fly samples used in this work were collected during bio-surveillance activities funded by the Scientific and Technological Research Council of Türkiye (TBAG 105T205 and SBAG: 114S999), Hacettepe University Scientific Researches Unit (09D01601002 and 01001601001), and the Armed Forces Health Surveillance Board, Global Emerging Infections Surveillance and Response System (AFHSB-GEIS), United States of America; FY18 award P0034\_18\_WR under US Army subcontract W911QY-16-C-0160 to Hacettepe University). The analyzed presence/absence sand fly data was kindly provided by the European Centre for Disease Prevention and Control (ECDC) and European Food Safety Authority (EFSA)<sup>83</sup>.

## Author contributions

E.K.: Conceptualization, Methodology, Validation, Formal analysis, Writing—Original Draft, Writing—Review and Editing, Visualization, Funding acquisition; S.K.: Conceptualization, Methodology, Validation, Formal analysis, Writing—Original Draft, Writing—Review and Editing, Visualization; K.P.: Data curation, Formal analysis, Writing—Review and Editing; O.K.: Methodology, Validation, Writing—Review and Editing; D.D.: Methodology, Validation, Writing—Review and Editing; V.D.: Methodology, Validation, Writing—Review and Editing; O.K.: Methodology, Validation, Writing—Review and Editing, Funding acquisition; B.Xh.: Resources, Validation, Writing—Review and Editing, K.S.: Resources, Validation, Writing—Review and Editing, J.W.: Validation, Supervision, Writing—Review and Editing, A.T.: Conceptualization, Methodology, Validation, Formal analysis, Writing—Original Draft, Writing—Review and Editing, Visualization, Funding acquisition.

## Funding

Open access funding provided by University of Pannonia.

## Declarations

## 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/s41598-025-94601-1>.

**Correspondence** and requests for materials should be addressed to A.J.T.

**Reprints and permissions information** is available at [www.nature.com/reprints](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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025