



# A molecular phylogeny and phylogeography of Greek Aegean Island sand flies of the genus *Phlebotomus* (Diptera: Psychodidae)

Christoforos Pavlou<sup>1</sup>, Emmanouil Dokianakis<sup>2</sup>, Nikolaos Tsirigotakis<sup>2</sup>, Vasiliki Christodoulou<sup>2</sup>, Yusuf Özbel<sup>3</sup>, Maria Antoniou<sup>2</sup>, Nikos Poulakakis<sup>1,4</sup>

1 *Biology Department, School of Sciences and Engineering, University of Crete, Voutes University Campus, GR-70013, P.O. Box 2208, Heraklion, Crete, Greece*

2 *Laboratory of Clinical Bacteriology, Parasitology, Zoonoses and Geographical Medicine, Faculty of Medicine, University of Crete, Voutes University Campus, GR-70013, P.O. Box 2208, Heraklion, Crete, Greece*

3 *Department of Parasitology, Medical School, Ege University, Bornova 35100, Izmir, Turkey*

4 *Natural History Museum of Crete, School of Sciences and Engineering, University of Crete, Knossos Av., GR-71409, P.O. Box 2208, Heraklion, Crete, Greece*

<http://zoobank.org/2E923688-4766-424F-ABBB-4FFFE29018C6>

Corresponding authors: Christoforos Pavlou (christoforos\_pm@hotmail.com)

**Received** 22 November 2021

**Accepted** 22 February 2022

**Published** 20 May 2022

**Academic Editors** Bradley Sinclair, Martin Wiemers

**Citation:** Pavlou C, Dokianakis E, Tsirigotakis N, Christodoulou V, Özbel Y, Antoniou M, Poulakakis N. A molecular phylogeny and phylogeography of Greek Aegean Island sand flies of the genus *Phlebotomus* (Diptera: Psychodidae). *Arthropod Systematics & Phylogeny* 80: 137–154. <https://doi.org/10.3897/asp.80.e78315>

## Abstract

The genus *Phlebotomus* (Diptera: Psychodidae: Phlebotominae) comprises a group of small winged insect species of medical importance. To date, ten species of *Phlebotomus* are known to be present in Greece; yet their evolutionary history is poorly studied due to the lack of comprehensive phylogenetic and phylogeographic studies. Herein, we aim to clarify the phylogenetic relationships amongst the local species collected from 12 Aegean Islands, Cyprus and Turkey; and to identify which of the palaeogeographic events may have influenced their biogeographic history. Our analyses revealed for the first time the presence of *P. cf. major* and *P. sergenti* in the Aegean Islands. All studied local species were retrieved as monophyletic and the mtDNA and nDNA phylogenetic trees indicated a plausible mitochondrial introgression between the closely related species of the *P. major* complex. From a palaeogeographic viewpoint, the major driving force that shaped the biogeographic history of the studied *Phlebotomus* species seems to be the dispersal that started in the Oligocene epoch, followed by several speciation events that occurred at the end of Miocene and the Plio-Pleistocene, including multiple dispersal events of Asiatic origin. The Messinian Salinity Crisis, the bimodal Mediterranean climate, and the glacial and interglacial periods were identified as key drivers for the diversification of the local species of *Phlebotomus*.

## Key words

Biogeography, Greece, molecular systematics, Phlebotominae, species delimitation

## 1. Introduction

The genus *Phlebotomus* Rondani & Berte belongs to the subfamily of Phlebotominae (sand flies). They are haematophagous insects and vectors of the protozoan parasites of the genus *Leishmania* Ross (Killick-Kendrick 1990) as well as the bacterium *Bartonella bacilliformis* Strong, Tyzzer, Brues, Sellards & Gastiaburu (Chamberlin et al. 2002) and arthropod-borne viruses (phleboviruses and vesiculoviruses) (Bichaud et al. 2011, 2014; Alwassouf et al. 2016). According to the database Systema Dipterorum, approximately 576 sand fly species are found in the Old World belonging to the genera of *Phlebotomus*, *Sergentomyia* Franca & Parrot and *Chinius* Leng (Akhoundi et al. 2016; Evenhuis and Pape 2022). Thirteen sand fly species are known from Greece, ten belonging to the genus *Phlebotomus* and three to the genus *Sergentomyia* (Léger et al. 1986; Ivočić et al. 2007; Xanthopoulos et al. 2011; Christodoulou et al. 2012; Tsirigotakis et al. 2018). Most of the *Phlebotomus* species present in Greece have wide geographical distribution including the Middle East, parts of Asia and the Mediterranean basin. *Phlebotomus (Adlerius) creticus* Antoniou, Depaquit & Dvořák and *Phlebotomus (Transphlebotomus) killicki* Dvořák, Votýpka & Volf on the other hand, are restricted to the Aegean and eastern Mediterranean, respectively. Five of them transmit or are suspected of transmitting *Leishmania* parasites: *P. (Larroussius) neglectus* Tonnoir, *P. (Larroussius) tobbi* Adler, Theodor & Lourie, *P. (Larroussius) perfiliewi* Parrot, *P. (Paraphlebotomus) similis* Perfiliev and *P. (Phlebotomus) papatasi* Scopoli.

*Phlebotomus* is considered a monophyletic group (Secombe et al. 1993; Aransay et al. 2000) and comprises 14 subgenera (Akhoundi et al. 2016; Cruaud et al. 2021), which are widely distributed in the Old World. Six of them are present in Greece; however, they consist of several species that are not found in the country (except *Artemievus* Depaquit). Nevertheless, comprehensive phylogenetic analyses are limited, and most research is focused on a limited number of loci and species. Due to its medical importance, subgenus *Larroussius* Nitzulescu is the most researched group in the genus, with multiple studies focusing on different species (Esseghir et al. 2000; Pesson et al. 2004; Latrofa et al. 2011; Depaquit et al. 2013). Phylogeographic studies on the genus are also limited and focus on specific species or subgenera (Esseghir et al. 2000; Depaquit et al. 2002; Kasap et al. 2015; Cruaud et al. 2021).

The Aegean archipelago is located in the eastern Mediterranean at the crossroads of three continents (Europe, Asia and Africa). It consists of more than 7,500 islands and islets with various island sizes and geological histories (Sfenthourakis and Triantis 2017). The Greek Aegean Islands provide an interesting area for phylogenetic and phylogeographic studies because of their topography (nearby Greek mainland or Anatolia) and geological history. This region was the epicentre of many studies, which either focused on the speciation that occurred within the Aegean region or on the dispersal of faunas from the

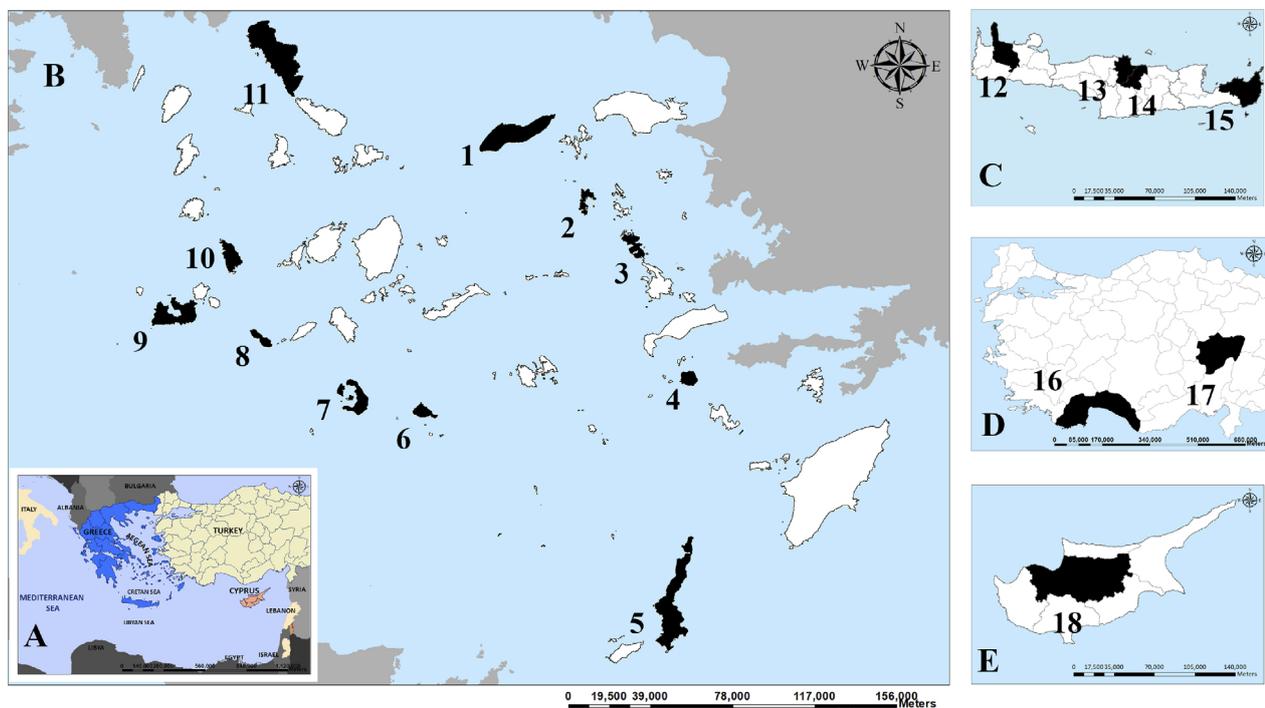
neighbouring regions (e.g. Turkey or Greek Mainland) to the Aegean. Over the last two and a half decades, approximately 75 animal genera have been studied in the Aegean (Poulakakis et al. 2015; Trichas et al. 2020). Such studies included, for example, the green lizard (*Lacerta* Linnaeus) (Kornilios et al. 2019), land snails (*Albinaria* Vest) (Douris et al. 2007), the centipede species *Scolopendra cingulata* Latreille (Simaiakis et al. 2012), tenebrionid beetles (Papadopoulou et al. 2009, 2010), the cave cricket genus *Dolichopoda* Boliva (Borissov et al. 2020) and the woodlouse genus *Trachelipus* Budde-Lund (Parmakelis et al. 2008). The Aegean region was subject to different geological and climatic events that contributed to the existing landscape and influenced different faunas. Such events were, for example, the formation of the Aegean Barrier [AB; a sea barrier often termed the mid-Aegean Trench (Poulakakis et al. 2015; Kornilios et al. 2019)], the Messinian Salinity Crisis (MSC) and the glacial and interglacial periods in the Pleistocene (Sfenthourakis and Triantis 2017). These events may have played an essential role in the diversification or the dispersal of sand flies from neighbouring regions.

Esseghir et al. (2000) and Kasap et al. (2015) suggested that the palaeogeographic events in the Mediterranean and Aegean Sea played important roles in the history of the *Phlebotomus* subgenera *Larroussius* and *Transphlebotomus* Artemiev & Neronov. It was proposed that *Larroussius* lineages were derived from speciation events associated with habitat shifts by aridification during the Pliocene (Esseghir et al. 2000). The *Transphlebotomus* lineages were associated with the palaeogeographic events of the eastern Mediterranean, which may have been the main drivers of their diversification (Kasap et al. 2015). Depaquit et al. (2002) suggested that *P. sergenti* Parrot and *P. similis* resulted from two migration routes during Miocene, one north of the Paratethys and the other south of the Tethys (Depaquit et al. 2002). Nevertheless, Trájer et al. (2021) suggested that *P. similis*, *P. sergenti* and *P. (Paraphlebotomus) jacusieli* Theodor could have separated after the tectonic subsidence of the Hellenic Orogenic Belt in late Miocene (Trájer 2021). In 2021, Cruaud et al. examined the systematics and the phylogeography of the subgenus *Paraphlebotomus* Theodor. The subgenus was not recovered as monophyletic, and thus, they described a new subgenus (*Artemievus*) for *P. (Artemievus) alexandri* Sinton. They also suggested that *Paraphlebotomus* originated between 8.5 and 12.9 mya and was distributed from the peri-Mediterranean to the Irano-Turanian regions (Cruaud et al. 2021).

The main objectives of our study include the investigation of the phylogenetic relationships amongst the local sand fly species found in 12 Greek Aegean Islands (Fig. 1), the estimation of the divergence times and the reconstruction of the ancestral distributions. Also, we aim to establish which palaeogeographic and palaeoclimatic events may have contributed to the evolutionary history of these species and associate them with divergence times. To accomplish these, two mitochondrial and four nuclear loci were analysed using multilocus phylogenetic and coalescence-based methods. The sample-set includ-

**Table 1.** Sand fly species collected from all study areas (Tsirigotakis et al. 2018; Dvořák et al. 2020; Omondi et al. 2020). The corresponding numbers for each study area in Fig. 1 are given in parenthesis.

Species/Region	Crete (C: 12–15)	Cyclades (A: 6–11)	Dodecanese (A: 2–5)	North Aegean (A: 1)	Cyprus (E: 18)	Turkey (D: 16–17)
<i>P. neglectus</i>	√	√	√	√		√
<i>P. perfiliewi</i>		√	√		√	
<i>P. tobbi</i>	√	√	√	√	√	√
<i>P. halepensis</i>						√
<i>P. simici</i>	√	√	√	√		√
<i>P. creticus</i>	√					
<i>P. (Adlerius) sp.</i>		√				
<i>P. papatasi</i>	√	√	√		√	
<i>P. sergenti</i>					√	
<i>P. similis</i>	√	√	√	√		
<i>P. alexandri</i>	√	√				√
<i>P. killicki</i>	√					



**Figure 1.** Geographical map indicating all study areas. **A** Key map; **B** Aegean Islands (1: Ikaria, 2: Patmos, 3: Leros, 4: Nisyros, 5: Karpathos, 6: Anafi, 7: Santorini, 8: Folegandros, 9: Milos, 10: Sifnos, 11: Andros); **C** Crete (12: Botanical Garden-Chania, 13: Fodele-Heraklion, 14: Foinikia-Heraklion, 15: Xerokampos-Lasithi); **D** Turkey (16: Antalya, 17: Kayseri); **E** Cyprus (18: Agioi Trimithias-Nicosia).

ed multiple specimens from all collected species in the Aegean Islands, with additional specimens from Cyprus and Turkey.

## 2. Methods

### 2.1. Sampling

Sand fly trapping in the Greek Aegean Islands was conducted during previous field work (2016–2019) (Tsirigotakis et al. 2018; Dvořák et al. 2020) using CDC min-

ature light traps (John W. Hock Co., Gainesville, FL, USA). The studied Aegean Islands included Crete, Cyclades (Andros, Sifnos, Milos, Folegandros, Santorini & Anafi), Dodecanese (Patmos, Leros, Nisyros & Karpathos), and North Aegean (Ikaria) (Fig. 1) (Tsirigotakis et al. 2018; Dvořák et al. 2020). Due to the presence of specimens with morphological uncertainties (e.g. specimens of *P. similis*), additional sampling took place during 2018 and 2019 in Cyprus (Agioi Trimithias) and Turkey (Kayseri & Antalya) (Omondi et al. 2020), in order to collect closely related species (e.g. *P. sergenti*) and enable us to resolve these uncertainties. Overall, 300 traps were used, with each trap being set for one night. Sand fly head, genitalia and wing of all collected specimens were mounted

**Table 2.** Primers and PCR conditions.

Locus	Primers	Sequence (5'-3')	Size (bp)	PCR conditions
COI	LCO1490	GGTCAACAAATCATAAAGATATTGG	~710	3mM MgCl, 0.4μM primers, 0.2mM dNTP's, 1U Taq DNA polymerase (94°C/1min, 42-50°C/1min, 72°C/1min) 35 cycles
	HCO2198	TAAACTTCAGGGTGACCAAAAAATCA		
	C1J1632	TGATCAAATTTATAAT	~560	3mM MgCl, 0.4μM primers, 0.2mM dNTP's, 1U Taq DNA polymerase (94°C/1min, 42°C/1min, 72°C/1min) 35 cycles
	C1N2191	GGTAAAATTTAAAATATAAACTTC		
CytB	CB3-PDR	CAYATTC AACWGAATGATA	~550	3mM MgCl, 0.6μM primers, 0.2mM dNTP's, 1U Taq DNA polymerase (94°C/1min, 43-46°C/1min, 72°C/1min) 35 cycles
	N1N-PDR	GGCAYWTTGCCTCGAWTTCGWTATGA		
ITS2	C1a	CCTGGTTAGTTTCTTTCCCTCCGCT	~530	2.5mM MgCl, 0.6μM primers, 0.2mM dNTP's, 1.5U Taq DNA polymerase (94°C/1min, 54-64°C/1min, 72°C/1min) 35 cycles
	JTS3	CGCAGCTAACTGTGTGAAATC		
28S	rc28H	CTACTATCCAGCGAAACC	~680	2mM MgCl, 0.6μM primers, 0.2mM dNTP's, 1.5U Taq DNA polymerase Touchdown PCR (94°C/1min, 60+58°C/1min, 72°C/1min) 5+32 cycles
	28K	GAAGAGCCGACATCGAAG		
EF1-α	EF-F05	CCTGGACATCGTGATTCAT	~500	2.5mM MgCl, 0.6μM primers, 0.2mM dNTP's, 1.5U Taq DNA polymerase (94°C/0.5min, 44-48°C/0.5min, 72°C/0.5min) 35 cycles
	EF-R08	CCACCAATCTTGTAGACATCCTG		
TPI	TPI111Fb	GGNAAYTGGAARATGAAYGG	~460	3mM MgCl, 0.6μM primers, 0.2mM dNTP's, 1.5U Taq DNA polymerase Touchdown PCR (94°C/0.5min, 54+50+45°C/0.5min, 72°C/1.5min) 5+5+30 cycles
	TPI275R	GCCANACNGGYTCRTANGC		

in CMCP-10 mounting medium (Polysciences, Inc., Warrington, PA, USA). At the same time, the rest of the body was stored in 70% ethanol at -20°C. Species identification was made using the morphological characters of the head, wing and genitalia (Lewis 1982; Killick-Kendrick et al. 1991; Kasap et al. 2015; Dvořák et al. 2020).

All morphologically identified *Phlebotomus* spp. are shown in Table 1. The collected *Phlebotomus* fauna from our study areas included 12 species, with the most common being *P. neglectus*, *P. tobbi*, *P. (Adlerius) simici* Nitzulescu and *P. similis*. *Phlebotomus killicki* and *P. creticus* were found only in Crete, *P. sergenti* only in Cyprus and *P. (Adlerius) halepensis* Theodor only in Turkey. Also, a male specimen ("C147B") of the subgenus *Adlerius* Nitzulescu collected from Andros could not be identified; nevertheless, it had morphological similarities with *P. creticus*.

## 2.2. DNA extraction, PCR amplification and sequencing

The sand fly genomic DNA was extracted from the thorax, legs and the anterior part of the abdomen of each chosen specimen using the Qiagen QIAamp DNA micro kit (Qiagen, Hilden, Germany). Double-stranded PCR was performed to amplify partial sequences of two mitochondrial loci (mtDNA) and four nuclear loci (nDNA). These loci were cytochrome c oxidase subunit 1 (COI), cytochrome b (CytB), internal subscribed spacer 2 (ITS2), domain 9 and 10 of 28S ribosomal RNA (28S), elongation factor 1 alpha (EF1-α) and triose-phosphate isomerase (TPI). All PCR reactions had 25μl volume and were performed in

T100 thermal cycler (Bio-Rad Laboratories, California, USA). Primers and PCR conditions are described in Table 2. Single-stranded Sanger sequencing was performed in CeMIA SA (Larisa, Greece) using the same primers as in PCR.

## 2.3. Sequence editing, alignment and model selection

Sequence chromatograms were viewed and edited using CodonCode Aligner v.9.0.1 (CodonCode Corporation, Centerville, USA). Multiple sequence alignments for each locus were performed using MAFFT v.7.475 (Kato and Standley 2013) with the default settings and the E-INS-i alignment strategy. Our alignments consisted of our samples and some additional sequences retrieved from the GenBank database, belonging to species that are closely related to our samples. The NCBI accession numbers of these sequences are shown in Table S1 of Supporting information. Protein coding sequences were translated into amino acids using Sequence Manipulation Suite (Stothard 2000) to check for the presence of stop codons.

The optimal nucleotide substitution model for each locus was identified using PartitionFinder v.2.1.1 (PF) (Lanfear et al. 2016). For each protein-coding locus, the alignment was partitioned into three blocks containing the 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> codon positions. We ran PF three times for each locus using the greedy search algorithm with linked branch lengths in calculations of likelihood scores under the Bayesian information criterion (BIC). The difference in those runs was the restriction of the candidate

**Table 3.** Conserved, variable and informative sites at each locus (without the outgroup).

Locus	Total sites	Conserved sites	Variable sites	Parsimony informative
COI	619	398	221	211
CytB	461	273	188	174
EF1- $\alpha$	423	294	129	126
TPI	373	281	92	90
28S	654	543	111	110
ITS2	526	272	224	204

models to those that were available in MrBayes v.3.2.6 (Ronquist et al. 2012), IQ-TREE v.1.6.12 (Nguyen et al. 2015), and BEAST2 v.2.6.2 (Bouckaert et al. 2019). The models which included both gamma distribution and invariable sites were ignored because of the high correlation between the proportion of the invariable sites and the gamma shape, making it extremely difficult to estimate both parameters reliably (Yang 2006).

All retrieved sequences were submitted to GenBank and their NCBI accession numbers are given in Table S1 of Supporting information. The aligned mtDNA, nDNA and concatenated datasets consisted of 1,080, 1,976 and 3,056 base pairs (bp), respectively. Information on conserved, variable and informative sites is given in Table 3. PartitionFinder indicated the best-fit nucleotide substitution models for each locus and each analysis, which are summarised in Table S2 of Supporting information.

## 2.4. Genetic distances and phylogenetic analyses

Genetic distances were calculated using the Tamura-Nei model (Tamura and Nei 1993) as implemented in MEGA v.7 (Kumar et al. 2016) to better compared with other studies. Phylogenetic trees were constructed using Bayesian Inference (BI) and Maximum Likelihood (ML) on the mtDNA, nDNA gene trees and the concatenated alignments. The BI analysis was performed in MrBayes with four runs and eight chains per run for 10,000,000 generations, with a sampling frequency of 100 using the substitution models selected by PF. To check for convergence and stationarity, we used several convergence diagnostics, such as the average standard deviation of split frequencies (ASDSF-values below 0.01 indicate convergence), the log-likelihood values, the Effective Sample Size (ESS-values below 100 may indicate insufficient sampling) and the average Potential Scale Reduction Factor (PSRF-values should be close to 1 for all parameters). The  $-\ln L$  stabilised after approximately  $10^6$  generations, and the first 25% of the trees were discarded as 'burn-in' phase. A majority-rule consensus tree relied on the remaining trees, and posterior probabilities were calculated as the percentage of samples recovering any particular clade (Huelsenbeck and Ronquist 2001). The ML analysis was performed in IQ-TREE with bio-neighbor joining starting tree under the suggested models selected in PF. Two independent runs were performed with bootstrap (Felsenstein 1985) and ultrafast bootstrap (Minh

et al. 2013) values, which were estimated by 1,000 and 10,000 replicates, respectively. *Sergentomyia minuta* Rondani represented the outgroup in these analyses. Fig-Tree v.1.4.4 (Rambaut 2018) was used to display the phylogenetic trees and Tracer v.1.7.1 (Rambaut et al. 2018) was used to assess the convergence statistics.

## 2.5. Coalescent species tree and divergence time estimation

We performed a molecular clock test in MEGA before the divergence times estimation analysis was applied to the studied loci. The null hypothesis of equal evolutionary rates throughout the tree was rejected at a 5% significance level ( $p < 0.05$ ), thus we applied an uncorrelated lognormal relaxed clock for the time estimation (Drummond et al. 2006).

The coalescent species tree and the estimation of divergence times were calculated using Starbeast2 (Ogilvie et al. 2017) as included in the BEAST2 software package (Bouckaert et al. 2019). We used BEAUti v.2.6.2, as implemented in the BEAST2 software package, to create the input file. We used the partitions and substitution models selected by PF (Lanfear et al. 2016). Also, the Yule Model was used for the speciation and the Uncorrelated Lognormal Models for the relaxed molecular clock. The Markov chain Monte Carlo (MCMC) analysis run for  $4 \times 10^8$  generations and saved the result every 5,000<sup>th</sup> generation. Tracer was used to verify that the analysis had converged and that the satisfactory Effective Sample Size had been obtained. The  $-\ln L$  stabilised after approximately  $4 \times 10^7$  generations, and the first 10% of the trees were discarded. *Sergentomyia minuta* represented the outgroup, and FigTree was used to display the species tree. To estimate divergence times, we used the available gene-specific substitution rates on CytB from previous studies. Specifically, the substitution rate for CytB was 2.1% (1.57–2.64%), and the clock rate was set at 0.0105 (Esseghir et al. 2000).

## 2.6. Biogeographic analysis

Biogeographic analysis was constructed in Reconstruct Ancestral State in Phylogenies (RASP) v.4.2 (Yu et al. 2015) and the ancestral distributions were reconstructed using the Statistical Dispersal-Vicariance Analysis (S-DIVA) (Yu et al. 2010), allowing a maximum of

two areas per node. The analysis was conducted using the trees and the dated species tree generated from the BEAST2 software package (Bouckaert et al. 2019). Five geographical regions were defined: A, Middle East; B, Aegean islands; C, East Europe; D, West Europe; E, North Africa.

## 2.7. Species delimitation

For conducting species delimitation, we used BP&P v.4.3 (Flouri et al. 2018), which uses the multispecies coalescent model to compare different models of species delimitation and phylogeny in the Bayesian framework. This method accounts for incomplete lineage sorting due to polymorphism and differences between gene trees and species tree (Yang and Rannala 2010). We conducted the A11 analysis for the joint species delimitation and species tree inference of unguided species delimitation (Yang and Rannala 2014), which attempts to group different populations into one species and explore different phylogenetic relationships among the delimited species (Yang and Rannala 2014). Our dataset comprised of five different loci (mtDNA, EF1- $\alpha$ , TPI, 28S & ITS2). The samples were divided into 17 potential distinct species that corresponded to the major clades of the phylogenetic concatenated tree and the morphological identification. The rjMCMC analysis was conducted for 100,000 generations, with a sampling interval of three, ‘burn-in’ phase of 2,500 and each species delimitation model had equal prior probability. We tested three combinations of priors (Leaché and Fujita 2010), where the prior distribution on the ancestral population size ( $\theta$ ) and the prior for the root age ( $\tau_0$ ) were set either to 0.1 [IG(3, 0.2)] or 0.001 [IG(3, 0.002)]. The first combination assumed relatively large ancestral population sizes and large divergences [ $\theta \sim$ IG(3, 0.2) &  $\tau_0 \sim$ IG(3, 0.2)]; the second combination assumed low ancestral population sizes and low divergences [ $\theta \sim$ IG(3, 0.002) &  $\tau_0 \sim$ IG(3, 0.002)] and the third combination assumed relatively large ancestral population sizes and low divergences among species [ $\theta \sim$ IG(3, 0.2) &  $\tau_0 \sim$ IG(3, 0.002)]. Each analysis was run twice with different seed numbers (positive and negative integer) to confirm the consistency between the runs. The starting tree had the topology of the coalescent species tree.

## 3. Results

### 3.1. Genetic distances

The genetic distances for mtDNA between the local *Phlebotomus* spp. ranged from 2.78% to 28.35% and for nDNA ranged from 0.4% to 23.22%. Genetic distances among and within species for mtDNA and nDNA (except ITS2) are given in Table 4. Sequence divergence for the COI ranged from 4.32% to 22.04%, for CytB from 3.79%

to 32.24%, for EF1- $\alpha$  from 0% to 13.99%, for TPI from 0.93% to 21.81%, for 28S from 0.73% to 9.92% and for ITS2 from 1.94% to 32.8%. The genetic distances for all loci separately are presented in Tables S3–S5 of Supporting information.

### 3.2. Combined phylogenetic tree and mtDNA/nDNA gene trees

Both BI and ML analyses on the concatenated dataset resulted into similar and well-supported phylogenetic trees (BI: lnL = -15,458.60; ML: Bootstrap lnL = -15,314.19 and Ultrafast Bootstrap lnL = -15,314.12) (Fig. 2). According to our results, all local species were monophyletic with very good statistical support. Also, each subgenus representative formed distinct monophyletic clades, with *Artemievus* being more closely related with *Phlebotomus* rather than *Paraphlebotomus*. Moreover, subgenus *Transphlebotomus* is more closely related to *Adlerius* than *Larroussius*. The *Adlerius* specimen (“C147B”) from Andros Island, which was morphologically similar to *P. creticus*, was placed in the same clade with *P. creticus* from Crete; thus, we assigned it to *P. creticus*, whose closest relative appears to be *P. (Adlerius) balcanicus* Theodor. As for the species of the *P. major* complex, *P. (Larroussius) syriacus* Adler & Theodor was the earliest branching lineage and *P. major* Annadale was the sister species to *P. neglectus*. The male specimens “M292” and “M293” (*P. cf. major*), which were morphologically different from the males of *P. neglectus*, were grouped with *P. major* sequences.

The mtDNA and the nDNA datasets produced trees with lnL = -7,489.09 and -7,848.33 for BI, respectively, and; lnL = -7,312.67 and -7,679.85 for Bootstrap ML, respectively and lnL = -7,312.06 and -7,679.21 for Ultrafast Bootstrap ML, respectively. The phylogenetic analyses produced less resolved trees but with similar topologies to that of the concatenated gene tree (Fig. S1 in Supporting information). The most important differences were the positions of the *P. major* complex, *P. sergenti* and *P. similis*. More specifically, according to the nDNA phylogenetic tree, *P. sergenti* and *P. similis* form a single well-supported clade. In contrast, the topology of these two species in the mtDNA gene tree was identical to the concatenated phylogenetic tree. Regarding the *P. major* complex, *P. neglectus* was resolved as paraphyletic in the mtDNA alignment tree since its lineage contained also the specimens of *P. major* and *P. cf. major* (Fig. S1 in Supporting information).

### 3.3. Coalescent species tree and divergence times

The coalescent species tree analysis resulted in good ESS values (lnL = -16,111.83) and the species tree presented very good posterior probabilities (Fig. 3). *Phlebotomus alexandri* is more closely related to *P. papatasi* rather than *P. sergenti* and *P. similis*, which are closely related.

*Phlebotomus creticus* is more closely related to *P. balcanicus* rather than *P. halepensis* or *P. simi-ci*. *Phlebotomus perniciosus*, which is distributed in the west Mediterranean (Léger and Depaquit 2002), appears to be closely related to *P. tobbi* (east Mediterranean). The species of the *P. major* complex appear to be closely related between them, with *P. syriacus* being the root lineage of the complex.

According to divergence time estimation (Fig. 3), the studied local species of the genus *Phlebotomus* separated from *S. minuta* at 34.37 mya (95% HPD 21.79–49.27) and started to diverge in the Oligocene at 29.92 mya (95% HPD 18.89–42.36), with the separation of *Artemievus*, *Phlebotomus* and *Paraphlebotomus* from all the other subgenera. The closely related species *P. similis* and *P. sergenti* diverged in the Pliocene at 3.62 mya (95% HPD 1.86–5.55). The estimated age for the split between the representatives of the subgenus *Larroussius* and those of the subgenera *Transphlebotomus* and *Adlerius* was 21.87 mya (95% HPD 14.72–29.73) (early Miocene). *Phlebotomus major* complex started diverging in the early Pleistocene at 2.45 mya (95% HPD 0.87–4.47), and *P. creticus* diverged from its closest relative (*P. balcanicus*) also in the early Pleistocene at 2.11 mya (95% HPD 0.7–3.73).

### 3.4. Biogeographic analysis

The results of the biogeographic analysis are shown in Fig. 3 and Fig. S2 of Supporting Information. The S-DIVA analysis proposed that the distributional history for the studied local species of *Phlebotomus* was dominated by dispersal. More specifically, the results indicated that 33 total dispersal events and only two vicariant events (nodes 5 & 8 in Fig. 3) were necessary to explain the current species distribution. Most dispersal events occurred from A (Middle East) to B (Aegean Islands) (11 events) and from A to C (East Europe) (eight events), while A was also the area with the most speciation events (14). According to the analysis, Middle East was the main contributing geographical area for most *Phlebotomus* species in this study. The majority of the geographical regions for the ancestral nodes were Middle East.

### 3.5. Species delimitation (BPP)

The results of the species delimitation analyses are given in Table 5. Assuming large ancestral population sizes and large divergences among species, the posterior probabilities of the species in the analysis ranged between 0.89 to 1, with the least supported species being *P. major* (0.89), *P.*

**Table 4.** Genetic distance (%) under the Tamura-Nei model between species. Genetic distance between species for mtDNA (COI & CytB) is given below diagonal and for nDNA (EF1- $\alpha$ , TPI & 28S) above diagonal. Diagonal values represent the genetic distance within species (in parenthesis the nDNA distances). n.a.: not available.

Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1 <i>S. minuta</i>	n.a. (n.a.)	18.99	12.28	9.77	10.72	11.27	15.25	16.17	n.a.	15.15	16.88	17.22	16.03	n.a.	20.02	16.74	13.93
2 <i>P. (Phl.) papatasi</i>	21.84	<b>0.42 (0.29)</b>	12.89	6.81	10.06	9.57	19.70	20.32	n.a.	17.07	18.09	23.22	19.11	n.a.	21.72	19.21	18.29
3 <i>P. (Art.) alexandri</i>	20.17	15.32	<b>0.10 (0.10)</b>	6.12	6.42	10.06	13.46	14.48	n.a.	12.72	13.55	14.45	13.50	n.a.	14.77	15.02	14.61
4 <i>P. (Par.) similis</i>	24.56	20.68	22.89	<b>2.58 (0.01)</b>	0.40	9.07	11.50	11.29	n.a.	11.41	10.45	10.50	10.66	n.a.	11.24	12.00	13.44
5 <i>P. (Par.) sergenti</i>	24.31	19.18	20.29	13.97	<b>1.46 (0.47)</b>	9.48	13.30	13.20	n.a.	12.91	11.87	12.02	11.63	n.a.	13.86	13.95	14.59
6 <i>P. (Transphl.) killicki</i>	19.78	15.27	16.48	21.43	21.48	<b>n.a. (n.a.)</b>	3.11	1.42	n.a.	2.27	6.50	n.a.	7.17	n.a.	n.a.	8.31	7.33
7 <i>P. (Adl.) simici</i>	19.60	17.22	15.82	22.53	21.39	13.09	<b>1.77 (0.03)</b>	5.89	n.a.	4.80	10.88	12.22	10.62	n.a.	12.84	10.63	9.74
8 <i>P. (Adl.) halepensis</i>	21.73	17.01	16.91	22.24	21.86	13.42	14.20	<b>0.10 (0)</b>	n.a.	1.59	12.07	12.18	9.86	n.a.	11.11	10.86	9.34
9 <i>P. (Adl.) balcanicus</i>	20.35	17.05	17.04	21.36	23.17	11.87	12.45	8.23	<b>n.a. (n.a.)</b>	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
10 <i>P. (Adl.) creticus</i>	20.34	16.00	16.48	21.87	21.27	11.73	13.15	10.46	4.44	<b>1.20 (0)</b>	10.61	10.19	9.03	n.a.	10.49	9.80	8.68
11 <i>P. (Larr.) peffilewi</i>	20.59	17.17	17.18	22.32	21.50	14.42	16.27	17.46	15.17	14.34	<b>0.95 (0.03)</b>	3.68	3.41	n.a.	6.69	6.65	6.40
12 <i>P. (Larr.) perniciosus</i>	21.56	15.42	15.39	21.46	21.29	11.94	14.63	15.82	14.97	14.45	7.82	<b>n.a. (n.a.)</b>	2.66	n.a.	8.97	8.15	9.31
13 <i>P. (Larr.) tobbi</i>	19.92	16.35	17.34	21.16	19.91	12.22	15.74	15.82	15.03	14.32	8.55	7.26	<b>0.71 (0.09)</b>	n.a.	8.50	8.11	7.37
14 <i>P. (Larr.) syriacus</i>	22.84	16.97	17.78	22.31	22.87	16.22	17.48	16.45	15.44	16.26	14.83	14.15	13.71	<b>n.a. (n.a.)</b>	n.a.	n.a.	n.a.
15 <i>P. (Larr.) major</i>	27.83	21.08	18.76	26.81	28.35	15.14	17.20	17.87	16.91	16.34	15.30	13.18	13.59	3.65	<b>n.a. (0.78)</b>	2.04	0.94
16 <i>P. (Larr.) neglectus</i>	20.78	17.74	16.77	22.41	21.32	16.06	16.19	16.55	15.83	15.46	14.17	13.08	13.32	5.59	2.78	<b>1.81 (0.10)</b>	1.51
17 <i>P. (Larr.) cf. major</i>	22.43	18.29	17.92	22.21	21.04	15.98	17.15	17.94	16.16	15.32	14.29	14.11	14.16	5.86	2.98	3.58	<b>0 (0)</b>

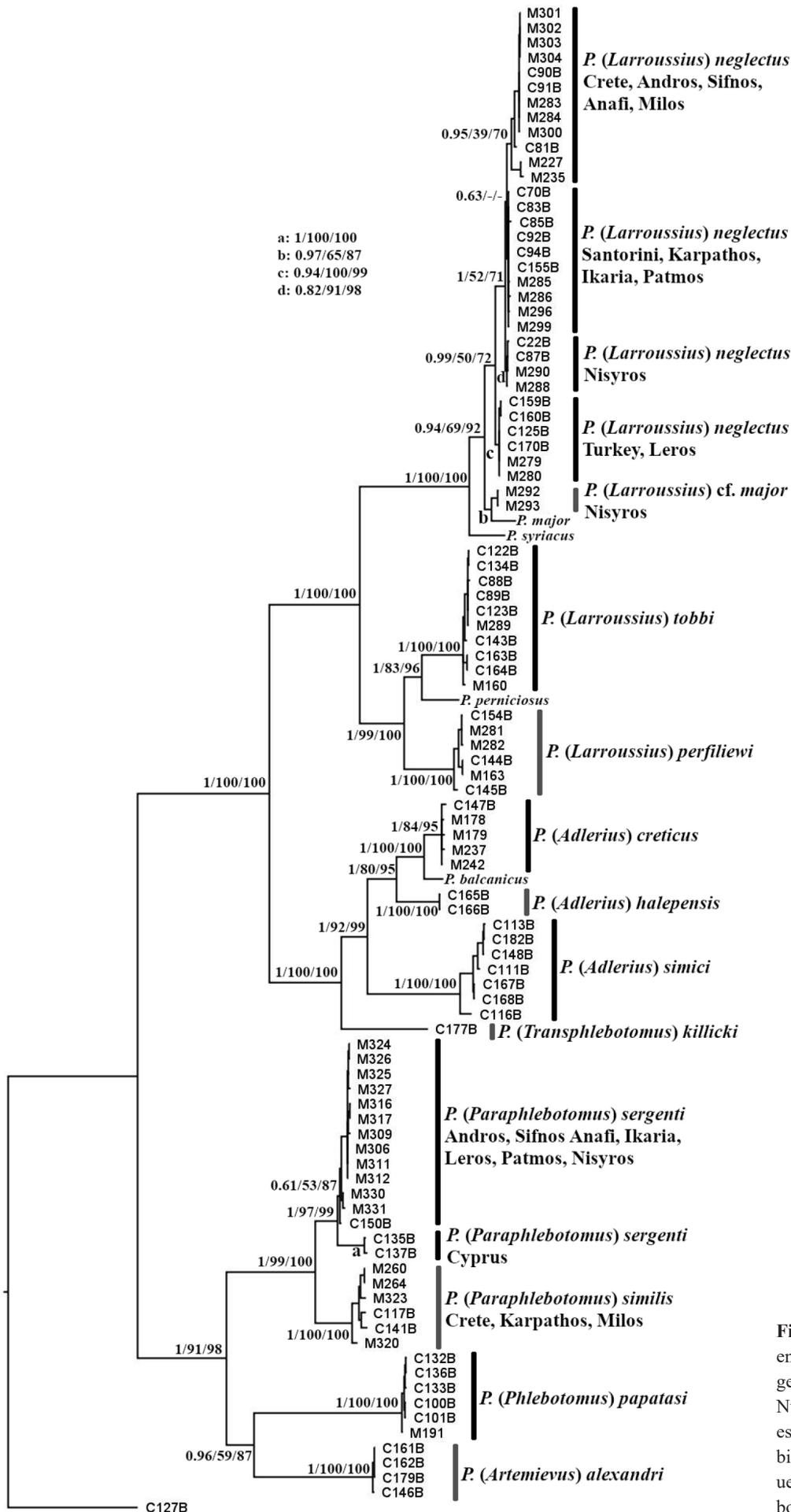
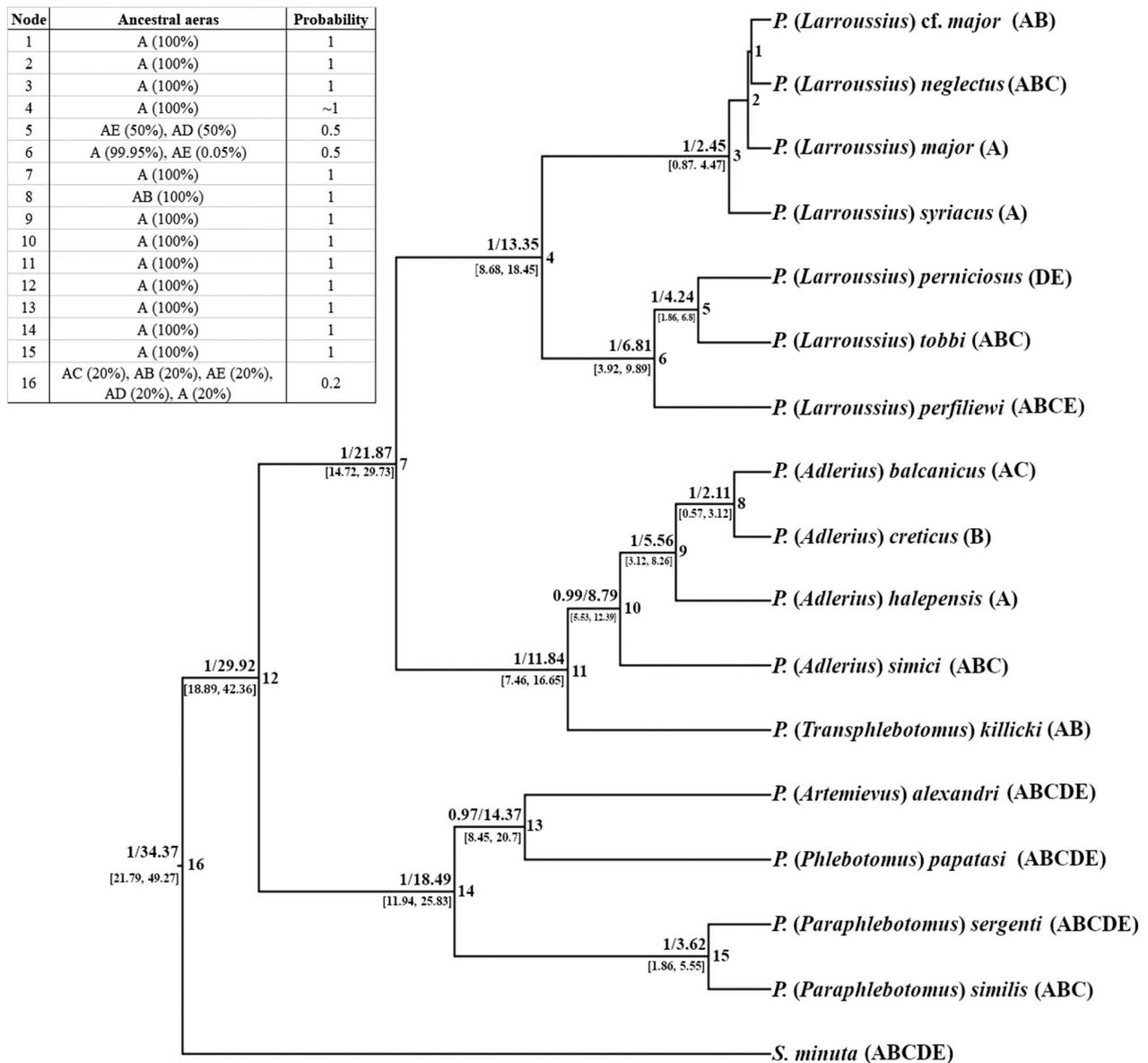


Figure 2. Bayesian inference concatenated phylogenetic tree of the six loci. Numbers next to the branches represent posterior probabilities (left), bootstrap values (middle) and ultrafast bootstrap values (right).



**Figure 3.** Dated coalescent species tree with posterior probabilities and mean divergence times (above the nodes). The 95% HPD are displayed below the nodes; ancestral geographical distribution is displayed in the up-left corner; letters next to the species correspond to their geographical distribution for the S-DIVA analysis (A: Middle East; B: Aegean Islands; C: East Europe; D: West Europe; E: North Africa).

*cf. major* (0.89) and *P. balcanicus* (0.90). The best-supported scheme included all 17 species groups and had a posterior probability equal to 0.76. The next proposed scheme included 16 species groups, with the fusion of *P. major*-*P. cf. major*. Assuming low ancestral population sizes and low divergences among species, species delimitation analysis supported a scheme of all 17 species-level clusters with a posterior probability equal to 0.81. The posterior probabilities of the species groups ranged from 0.85 to 1, with the least supported species groups being *P. major* (0.85) and *P. cf. major* (0.86). The posterior probabilities of the species for the third combination of priors (large ancestral population sizes and low divergences among species), ranged from 0.84 to 1. The least supported species were *P. balcanicus* (0.84) and *P. major* (0.90).

### 4. Discussion

This study is the first multilocus phylogenetic and phylogeographic approach of the *Phlebotomus* genus in the Greek Aegean Islands. The Aegean area is of great importance regarding biodiversity, biogeography and epidemiology due to its geographic location and history. Our analyses report for the first time the presence of *P. sergenti* and *P. cf. major* in the Aegean Islands, which were morphologically identified as *P. similis* and *P. neglectus*, respectively by Tsirigotakis et al. (2018). *Phlebotomus sergenti* was found in Nisyros, Leros, Patmos, Sifnos, Ikaria, Andros, Anafi, and *P. cf. major* in Nisyros. *Phlebotomus sergenti* is a proven vector of *L. tropica* Wright, and the species of the *P. major* complex are proven vec-

**Table 5.** Species delimitation results assuming 17 species and using different prior schemes. Posterior probabilities are the average of the two runs with different seed number.

Prior scheme	$\theta \sim \text{IG}(3, 0.2)$ & $\tau_0 \sim \text{IG}(3, 0.2)$		$\theta \sim \text{IG}(3, 0.002)$ & $\tau_0 \sim \text{IG}(3, 0.002)$		$\theta \sim \text{IG}(3, 0.2)$ & $\tau_0 \sim \text{IG}(3, 0.002)$	
Candidate species	Posterior probability		Posterior probability		Posterior probability	
<i>S. minuta</i>	1		1		1	
<i>P. similis</i>	1		1		1	
<i>P. alexandri</i>	1		1		1	
<i>P. tobbi</i>	1		1		1	
<i>P. neglectus</i>	1		1		1	
<i>P. perfiliewi</i>	1		1		1	
<i>P. simici</i>	1		1		1	
<i>P. sergenti</i>	1		1		1	
<i>P. papatasi</i>	1		1		0.99	
<i>P. creticus</i>	0.98		0.98		0.99	
<i>P. perniciosus</i>	0.96		1		0.92	
<i>P. halepensis</i>	0.96		0.97		0.96	
<i>P. killicki</i>	0.96		1		0.91	
<i>P. syriacus</i>	0.95		0.99		0.90	
<i>P. balcanicus</i>	0.90		0.95		0.84	
<i>P. cf. major</i>	0.89		0.86		0.92	
<i>P. major</i>	0.89		0.85		0.90	
<i>P. major</i> & <i>P. cf. major</i>	0.11		0.14		0.08	
<i>P. balcanicus</i> & <i>P. halepensis</i>	0.03		0.03		0.02	
<i>P. creticus</i> & <i>P. balcanicus</i>	0.03		0.02		0.02	
<i>P. major</i> & <i>P. syriacus</i>	0.01				0.03	
<i>P. cf. major</i> & <i>P. syriacus</i>	0.01					
Number of possible species	Posterior probability	Prior probability	Posterior probability	Prior probability	Posterior probability	Prior probability
15	0.03	0.06	0.01	0.06	0.05	0.06
16	0.22	0.03	0.19	0.03	0.26	0.03
17	0.76	0.01	0.81	0.01	0.68	0.01

tors of *L. infantum* Nicolle (Akhoundi et al. 2016). These findings can aid in predicting the geographical distribution of leishmaniasis in the region, by defining the geographical distribution of the various sand fly species.

#### 4.1. Phylogenetic relationships

The interspecific genetic distances of the studied local species of *Phlebotomus* were similar to those found by Esseghir et al. (2000). Our phylogenetic analyses indicated that all studied local species were monophyletic and each subgenus representative formed distinct monophyletic clades. Nevertheless, species complexes such as *P. major*, *P. perniciosus* and *P. perfiliewi*, need further investigation by sampling the remaining taxa, in order to resolve their taxonomic status. Subgenus *Artemievus* (which is comprised of only *P. alexandri*) appears to be more closely related with *Phlebotomus* rather than *Paraphlebotomus* (where *P. alexandri* was previously classified). This is in agreement with previously published work, which showed that *P. alexandri* is more closely related to *P. papatasi*, *P. (Phlebotomus) dubosqi* Neveu-Lemarie and *P. (Phlebotomus) bergeroti* Parrot rather than the species of the subgenus *Paraphlebotomus* (Krüger et al. 2011;

Cruaud et al. 2021). The two species of *Paraphlebotomus* in our study are closely related to each other, since they formed a single well-supported clade in the nDNA gene tree, while they formed two distinct, well-supported clades in the mtDNA gene tree (Figure S1 of Supporting information).

Subgenus *Larrousius* is closely related to *Transphlebotomus* and *Adlerius*, which agrees with the classification of morphological characters by Rispaill and Léger (1998). *Larrousius* was divided into two major clades, with the first one comprising *P. perfiliewi*, *P. perniciosus* Newstead and *P. tobbi*. *Phlebotomus perfiliewi* was the earliest branching lineage, and this is congruent with the phylogenetic analysis made by Esseghir et al. (2000). The second clade included the species of the *P. major* complex with *P. syriacus* as the earliest branching lineage. *Phlebotomus cf. major* from Nisyros was placed in the same subclade with *P. major*, and *P. neglectus* appears as their sister species. Kasap et al. (2013) observed similar phylogenetic relationships in this species complex. The *P. major* complex consists of six species with *P. neglectus*, *P. major* and *P. syriacus* present in Turkey (Kasap et al. 2013, 2019). Kasap et al. (2013) examined the distribution of these species in Turkey and found three major lineages: *P. syriacus* (distributed in southeast Turkey); *P.*

*neglectus* s.str. (sympatric with *P. syriacus* and the third lineage); and the third lineage composed of *P. major* sequences (distributed in central and southwest Turkey). These *P. major* sequences, which match those in our study, originated from specimens from Iran. Badakhshan et al. (2011) reported that most specimens from Iran were not representatives of *P. major* s.str. Kasap et al. (2013) suggested that these specimens may belong to a different cryptic species of the *P. major* complex. We assume that the specimens (seven males) found in Nisyros may belong to the same latter species.

The first lineage of subgenus *Adlerius* to branch off was *P. simici*, while the other three species appear to be closely related. According to our analyses, *P. creticus* formed a single well-supported clade with *P. balcanicus* as its closest relative. These results were also observed by Dvořák et al. (2020). The *Adlerius* specimen (“C147B”) from Andros was grouped with the specimens of *P. creticus* from Crete, and also, the genetic distance from *P. creticus* was 1.63% for mtDNA and 0% for nDNA. Consequently, we assume it represents *P. creticus*, which was found in cave entrances across all of Crete, while the Andros specimen was found in an abandoned stone building.

Another important finding of our study is the probable mitochondrial introgression between the species of the *P. major* complex. In the nDNA gene tree (Figure S1 of Supporting information), three distinct lineages were present, corresponding to the three taxa, and with *P. syriacus* as the earliest branching lineage. In the mtDNA gene tree (Figure S1 of Supporting information), all taxa were included in a single clade without separating the different species. Pesson et al. (2004) had suggested mtDNA introgression (most likely by hybridization) between *P. perniciosus* and *P. (Larrousius) longicuspis* Nitzulescu (Pesson et al. 2004), which are sympatric in southwest Mediterranean (Léger and Depaquit 2002). Esseghir et al. (2000) also suggested introgressive hybridizations between species whose ranges overlap [*P. (Larrousius) orientalis* Parrot & *P. (Larrousius) langeroni* Nitzulescu].

According to Bayesian species delimitation analyses, the least supported species are *P. balcanicus*, *P. cf. major* and *P. major*. The taxonomic status of *P. balcanicus* is unresolved probably because of the only two mitochondrial loci (CytB & COI) that are available in the GenBank database. Further molecular data (nuclear data) of *P. balcanicus* are needed to resolve its status. The relationship between *P. cf. major* and *P. major* is unresolved probably because they may belong to the same species or due to the few molecular data (CytB and EF1- $\alpha$ ) available for *P. major*.

## 4.2. Phylogeography of Greek Aegean Island sand flies

According to our dating analysis, the studied local species of *Phlebotomus* were separated from *S. minuta* at 34.37 mya, which is congruent with the suggestion of Tuon et al. (2008) and Akhoundi et al. (2016) that the genus

*Phlebotomus* emerged in the Eocene epoch (33.9 to 55.8 mya). The emergence of *Larrousius* occurred at 21.87 mya (early Miocene) and started diverging at 13.35 mya (Serravalian age). Léger and Pesson (1987) associated the speciation of Mediterranean *Larrousius* with vicariance due to tectonic activities during early Miocene, while Esseghir et al. (2000) concluded that the speciation was not dependent on tectonic vicariance, but was affected by palaeoecological changes during late Miocene and that the faunal turnover and dispersion has determined the current species distributions (Léger and Pesson 1987; Esseghir et al. 2000). To extract safer assumptions on the factors that played key roles to the divergence of the subgenus *Larrousius*, additional sampling for more taxa is needed.

*Phlebotomus perfiliewi* diverged from *P. tobbi* and *P. perniciosus* during the pre-evaporitic stage of the Messinian age (7.25–5.96 mya) (Prista et al. 2015) through dispersal events from an Asiatic origin. However, this hypothesis contradicts the conclusion made by Trájer et al. (2021), who emphasized that during Messinian, the active dispersal of sandflies was very difficult or impossible due to the thermal conditions (Trájer et al. 2021). The only vicariant event in *Larrousius* lineages occurred at 4.24 mya (Zanclean age), diverging the lineages of *P. tobbi* and *P. perniciosus*. This lineage split occurred after the refilling of the Mediterranean Sea and the re-establishment of the marine conditions (Prista et al. 2015). *Phlebotomus tobbi* is distributed in eastern Mediterranean and Middle East (Léger and Depaquit 2002), while *P. perniciosus* is distributed in western Mediterranean, from Morocco to Libya and from Portugal to Croatia (Pesson et al. 2004). Therefore, we hypothesize that the ancestor of these two species may have occupied the Mediterranean region sometime before the refilling of the Mediterranean Sea (probably during MSC) and the latter may have acted as the vicariant event that split their lineages.

The speciation within the *P. major* complex appears to have coincided with the glacial and interglacial periods during the early Pleistocene (2.46 to 2.11 mya) (Poulakakis et al. 2015; Suc et al. 2019), and ancestral reconstruction showed that the complex has an Asiatic origin. During these periods, cycles of alternating hot, dry and cold, wet seasons were present in the Aegean area, with glacial periods outlasting the interglacial (Poulakakis et al. 2015). Consequently, we assume that the glacial and interglacial periods were the key factors for the speciation within the complex and the current distribution of each species was achieved through dispersal. Esseghir et al. (2000) also suggested an Asiatic origin; however, they suggested that the speciation occurred during or shortly after the MSC, which acted as a vicariant event.

The subgenus *Transphlebotomus* is restricted to the Mediterranean basin (Kasap et al. 2015) and the subgenus *Adlerius* is distributed mostly in Asia and eastern Mediterranean. Our subgenera representatives of *Transphlebotomus* and *Adlerius*, appear to have an Asiatic origin, and their split occurred at 11.84 mya (late Miocene) when the AB began its formation in the Aegean area (Poulakakis et al. 2015). This dating is congruent with that of Kasap et al. (2015). The studied local species of *Adlerius*

have an Asiatic origin and the speciation amongst them was dominated by dispersal. *Phlebotomus simici* separated from the other species of *Adlerius* at 8.79 mya (after the formation of the AB). Therefore, we assume that the ancestor of *P. simici* was able to disperse from the Middle East to the Aegean and East Europe despite the presence of the AB. Due to the apparent crossing of the AB from the ancestor of *P. simici*, we propose that the AB had a minor influence on this diversification. Species of the woodlouse genus *Dolichopoda* and the tenebrionid beetle *Dichomma dardanum* are similar cases where the crossing of the AB was observed (Allegrucci et al. 2009, 2011; Papadopoulou et al. 2009).

The separation of *P. halepensis* from *P. balcanicus* and *P. creticus* coincided with the MSC (5.96–5.33 mya) (Krijgsman et al. 1999), which may have provided the opportunity for the dispersal of the ancestor of *P. balcanicus* and *P. creticus* from the Middle East to the Aegean through the formation of steppes and saline deserts. A similar case has been suggested for the separation of bush cricket lineages of the *Poecilimon jonicus* group (Orthoptera), which they may have taken the opportunity for dispersal during MSC, and were isolated during the refilling of the Mediterranean (Borissov et al. 2020). The ancestor of *P. creticus* and *P. balcanicus* had occupied the Middle East and the Aegean, and its diversification was due to a vicariant event during the Pleistocene. This event may have been the alternating glacial and interglacial periods. The current distribution of *P. balcanicus* consists of the Balkans, Turkey, Armenia and Georgia (Léger and Depaquit 2002; Cazan et al. 2019). Consequently, we assume that these periods may have acted as the vicariant event that separated these lineages, while the ancestor of *P. balcanicus* was able to disperse in the Balkans.

Finally, the separation of the subgenera *Paraphlebotomus*, *Phlebotomus* and *Artemievus* from each other appears to coincide with the Mid-Miocene Climatic Optimum (~17–15 mya), which represents a geologically warming event (Böhme 2003; Song et al. 2018). As for the closely related *P. sergenti* and *P. similis*, they appear to have separated before the establishment of the Mediterranean climate regime as we know it today (3.2 to 2.8 mya) (Blondel et al. 2010) in late Pliocene through dispersal from an Asiatic origin. According to Trájer et al. (2018), the ancestor of these species adapted to the hot and dry summers of the Mediterranean during late Neogene and the lowest minimum tolerable temperature of *P. sergenti* and *P. similis* is  $-7.2^{\circ}\text{C}$  and  $-4^{\circ}\text{C}$ , respectively (Trájer et al. 2018). Subsequently, we assume that this climate regime was the main factor of their divergence through climate adaptation. This assumption is incongruent with the hypothesis made by Depaquit et al. (1998), who suggested that during the Miocene, both species followed different migration routes, one north of the Paratethys and the other south of the Tethys (Depaquit et al. 1998). Our assumption is also incongruent with the alternative proposed hypothesis made by Trájer et al. (2021), in which *P. similis*, *P. sergenti* and *P. (Paraphlebotomus) jacusieli* could have separated after the tectonic

subsidence of the Hellenic Orogenic Belt in late Miocene (Trájer 2021). Furthermore, Cruaud et al. (2021) suggested that the separation of the *Paraphlebotomus* lineages was probably the result of a Messinian vicariant event in which extreme prolonged drought has been instrumental (Cruaud et al. 2021).

## 5. Conclusions

Under the framework of our regional sampling, all studied taxa were recovered as mutually monophyletic. *Phlebotomus sergenti* and *P. cf. major* were recorded for the first time in the Greek Aegean Islands. Furthermore, our results indicated a probable mitochondrial introgression between the species of the *P. major* complex, while their genetic diversification appears to be low. *Phlebotomus creticus* was indicated as the sister species of *P. balcanicus*, with their diversification being the most recent one amongst all studied species. According to our phylogeographic analyses, the palaeoecological events in the Mediterranean region such as the MSC, the establishment of the Mediterranean climate, and glacial and interglacial periods were identified as the major drivers for the diversification of the studied species. Dispersal was the major driving force that shaped the biogeographic history and the current geographical species distributions since most of the species' diversification was due to dispersal events from Middle East.

Further molecular and morphological research studies are needed for resolving the *P. major* complex and the relationships between the species comprising it. Likewise, additional molecular data (nuclear DNA) are needed for a more comprehensive study of the relationship and the status between *P. balcanicus* and *P. creticus*. Due to the close relationship of these two species, additional studies are required to determine the vector capability of *P. creticus*. Additional sampling in the Aegean Islands is necessary to locate more specimens of *P. creticus* and determine its geographical range. This study highlighted the importance of the Aegean Islands and the need for more studies. They may host important sand fly species that play a crucial role in the biodiversity and the epidemiology of leishmaniasis of the area. Finally, further exhaustive sampling in other areas is crucial, in order to collect as many taxa as possible, to clarify the taxonomic status of species complexes and to carry out a more comprehensive phylogenetic study on the genus.

## 6. Availability of data and materials

All generated sequences were submitted to GenBank and their NCBI accession numbers are given in Table S1 of Supporting information. All remaining DNA samples and sand fly specimens are deposited in Natural History Museum of Crete (University of Crete).

## 7. Authors' contributions

Specimen collection and morphological identification: CP, NT, VC, YO & MA; molecular procedures: CP & ED; map design: VC; phylogenetic and phylogeographic analyses: CP; data interpretation: CP, MA & NP; project planning: CP, MA & NP; manuscript preparation: CP, MA & NP. All authors have read and approved the manuscript.

## 8. Acknowledgements

We would like to express our gratitude towards the people who helped us with field and laboratory work. Especially, we are thankful to the personnel of the Molecular Systematics and Arthropods laboratories of the Natural History Museum of Crete for their help and assistance with the molecular procedures.

## 9. References

- Akhoundi M, Kuhls K, Cannet A, Votýpka J, Marty P, Delaunay P, Sereño D (2016) A historical overview of the classification, evolution, and dispersion of *Leishmania* parasites and sandflies. *PLOS Neglected Tropical Diseases* 10(3): e0004349. <https://doi.org/10.1371/journal.pntd.0004349>
- Allegrucci G, Rampini M, Gratton P, Todisco V, Sbordoni V (2009) Testing phylogenetic hypotheses for reconstructing the evolutionary history of *Dolichopoda* cave crickets in the eastern Mediterranean. *Journal of Biogeography* 36(9): 1785–1797. <https://doi.org/10.1111/j.1365-2699.2009.02130.x>
- Allegrucci G, Trucchi E, Sbordoni V (2011) Tempo and mode of species diversification in *Dolichopoda* cave crickets (Orthoptera, Rhaphidophoridae). *Molecular Phylogenetics and Evolution* 60(1): 108–121. <https://doi.org/10.1016/j.ympev.2011.04.002>
- Alwassouf S, Christodoulou V, Bichaud L, Ntais P, Mazeris A, Antoniou M, Charrel RN (2016) Seroprevalence of sandfly-borne phleboviruses belonging to three serocomplexes (Sandfly fever Naples, Sandfly fever Sicilian and Salehabad) in dogs from Greece and Cyprus using neutralization test. *PLOS Neglected Tropical Diseases* 10(10): e0005063. <https://doi.org/10.1371/journal.pntd.0005063>
- Aransay AM, Scoulcia E, Tselentis Y, Ready PD (2000) Phylogenetic relationships of phlebotomine sandflies inferred from small subunit nuclear ribosomal DNA. *Insect Molecular Biology* 9(2): 157–168. <https://doi.org/10.1046/j.1365-2583.2000.00168.x>
- Badakhshan M, Sadraei J, Moin-Vaziri V (2011) Morphometric and morphological variation between two different populations of *Phlebotomus major* s.l. from endemic and non-endemic foci of visceral leishmaniasis in Iran. *Journal of Vector Ecology* 36(1): 153–158. <https://doi.org/10.1111/j.1948-7134.2011.00152.x>
- Bichaud L, Izri A, de Lamballerie X, Moureau G, Charrel RN (2014) First detection of Toscana virus in Corsica, France. *Clinical Microbiology and Infection* 20(2): O101–O104. <https://doi.org/10.1111/1469-0691.12347>
- Bichaud L, Souris M, Mary C, Ninove L, Thirion L, Piarroux RP, Piarroux R, De Lamballerie X, Charrel RN (2011) Epidemiologic relationship between Toscana virus infection and *Leishmania infantum* due to common exposure to *Phlebotomus perniciosus* sandfly vector. *PLOS Neglected Tropical Diseases* 5(9): e1328. <https://doi.org/10.1371/journal.pntd.0001328>
- Blondel J, Aronson J, Bodiou J-Y, Boeuf G (2010) *The Mediterranean region: biological diversity in space and time*. New York, NY: Oxford University Press Inc.
- Böhme M (2003) The Miocene Climatic Optimum: evidence from ectothermic vertebrates of Central Europe. *Palaeogeography, Palaeoclimatology, Palaeoecology* 195(3–4): 389–401. [https://doi.org/10.1016/S0031-0182\(03\)00367-5](https://doi.org/10.1016/S0031-0182(03)00367-5)
- Borissov SB, Bobeva A, Çıplak B, Chobanov D (2020) Evolution of *Poecilimon jonicus* group (Orthoptera: Tettigoniidae): a history linked to the Aegean Neogene paleogeography. *Organisms Diversity & Evolution* 20(4): 803–819. <https://doi.org/10.1007/s13127-020-00466-9>
- Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, De Maio N, Matschiner M, Mendes FK, Müller NF, Ogilvie HA, du Plessis L, Poppinga A, Rambaut A, Rasmussen D, Siveroni I, Suchard MA, Wu C-H, Xie D, Zhang C, Stadler T, Drummond AJ, Pterea M (2019) BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology* 15(4): e1006650. <https://doi.org/10.1371/journal.pcbi.1006650>
- Cazan CD, Păstrav IR, Ionică AM, Oguz G, Erisoz Kasap O, Dvořák V, Halada P, Dumitrache MO, Volf P, Alten B, Mihalca AD (2019) Updates on the distribution and diversity of sand flies (Diptera: Psychodidae) in Romania. *Parasites & Vectors* 12: 247. <https://doi.org/10.1186/s13071-019-3507-7>
- Chamberlin J, Laughlin LW, Romero S, Solórzano N, Gordon S, Andre RG, Pachas P, Friedman H, Ponce C, Watts D (2002) Epidemiology of endemic *Bartonella bacilliformis*: a prospective cohort study in a Peruvian mountain valley community. *The Journal of Infectious Diseases* 186(7): 983–990. <https://doi.org/10.1086/344054>
- Christodoulou V, Antoniou M, Ntais P, Messaritakis I, Ivočić V, Dedet J-P, Pralong F, Dvořák V, Tselentis Y (2012) Re-emergence of visceral and cutaneous Leishmaniasis in the Greek island of Crete. *Vector-Borne and Zoonotic Diseases* 12 (3): 214–222. <https://doi.org/10.1089/vbz.2011.0004>
- Creutzburg N (1963) Paleogeographic evolution of Crete from Miocene till our days. *Cretan Annals* 15(16): 336–342.
- Cruaud A, Lehrter V, Genson G, Rasplus J-Y, Depaquit J (2021) Evolution, systematics and historical biogeography of sand flies of the subgenus *Paraphlebotomus* (Diptera, Psychodidae, *Phlebotomus*) inferred using restriction-site associated DNA markers. *PLOS Neglected Tropical Diseases* 15(7): e0009479. <https://doi.org/10.1371/journal.pntd.0009479>
- Depaquit J, Bounamous A, Akhoundi M, Augot D, Sauvage F, Dvořák V, Chaibullinova A, Pesson B, Volf P, Léger N (2013) A taxonomic study of *Phlebotomus (Larrousius) perfliewi* s. l. *Infection, Genetics and Evolution* 20: 500–508. <https://doi.org/10.1016/j.mee- gid.2013.10.006>
- Depaquit J, Ferté H, Léger N, Lefranc F, Alves-Pires C, Hanafi H, Maroli M, Morillas-Marquez F, Rioux J-A, Svobodova M, Volf P (2002) ITS 2 sequences heterogeneity in *Phlebotomus sergenti* and *Phlebotomus similis* (Diptera, Psychodidae): possible consequences in their ability to transmit *Leishmania tropica*. *International Journal for Parasitology* 32(9): 1123–1131. [https://doi.org/10.1016/s0020-7519\(02\)00088-7](https://doi.org/10.1016/s0020-7519(02)00088-7)
- Depaquit J, Léger N, Ferté H (1998) Le statut taxinomique de *Phlebotomus sergenti* Parrot, 1917, vecteur de *Leishmania tropica* (Wright, 1903) et *Phlebotomus similis* Perfliev, 1963 (Diptera-Psychodidae). *Approches morphologique et morphométrique*. Corollaires biogé-

- graphiques et épidémiologiques. *Bulletin de la Société de Pathologie Exotique* 91: 346–352.
- Douris V, Giokas S, Thomaz D, Lecanidou R, Rodakis GC (2007) Inference of evolutionary patterns of the land snail *Albinaria* in the Aegean archipelago: Is vicariance enough? *Molecular Phylogenetics and Evolution* 44(3): 1224–1236. <https://doi.org/10.1016/j.ympev.2007.01.004>
- Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed Phylogenetics and Dating with Confidence. *PLoS Biology* 4(5): e88. <https://doi.org/10.1371/journal.pbio.0040088>
- Dvořák V, Tsigotakis N, Pavlou C, Dokianakis E, Akhoundi M, Halada P, Volf P, Depaquit J, Antoniou M (2020) Sand fly fauna of Crete and the description of *Phlebotomus (Adlerius) creticus* n. sp. (Diptera: Psychodidae). *Parasites & Vectors* 13: 547. <https://doi.org/10.1186/s13071-020-04358-x>
- Esseghir S, Ready PD, Ben-Ismaïl R (2000) Speciation of *Phlebotomus* sandflies of the subgenus *Larroussius* coincided with the late Miocene-Pliocene aridification of the Mediterranean subregion. *Biological Journal of the Linnean Society* 70(2): 189–219. <https://doi.org/10.1111/j.1095-8312.2000.tb00207.x>
- Evenhuis NL, Pape T (2022) *Systema Dipteriorum* version 3.6 <http://diptera.org> [Accessed on 18.02.2022]
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39(4): 783–791. <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x>
- Flouri T, Jiao X, Rannala B, Yang Z (2018) Species tree inference with BPP using genomic sequences and the multispecies coalescent. *Molecular Biology and Evolution* 35: 2585–2593.
- Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17(8): 754–755. <https://doi.org/10.1093/bioinformatics/17.8.754>
- Ivočić V, Patakakis M, Tselentis Y, Chaniotis B (2007) Faunistic study of sandflies in Greece. *Medical and Veterinary Entomology* 21(1): 121–124. <https://doi.org/10.1111/j.1365-2915.2006.00649.x>
- Kasap OE, Dvořák V, Depaquit J, Alten B, Votýpka J, Volf P (2015) Phylogeography of the subgenus *Transphlebotomus* Artemiev with description of two new species, *Phlebotomus anatolicus* n. sp. and *Phlebotomus killicki* n. sp. *Infection, Genetics and Evolution* 34: 467–479. <https://doi.org/10.1016/j.meegid.2015.05.025>
- Kasap OE, Linton Y-M, Karakus M, Ozbel Y, Alten B (2019) Revision of the species composition and distribution of Turkish sand flies using DNA barcodes. *Parasites & Vectors* 12: 410. <https://doi.org/10.1186/s13071-019-3669-3>
- Kasap OE, Votýpka J, Alten B (2013) The distribution of the *Phlebotomus major* complex (Diptera: Psychodidae) in Turkey. *Acta Tropica* 127(3): 204–211. <https://doi.org/10.1016/j.actatropica.2013.05.001>
- Katoh K, Standley DM (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* 30(4): 772–780. <https://doi.org/10.1093/molbev/mst010>
- Killick-Kendrick R (1990) Phlebotomine vectors of the leishmaniasis: a review. *Medical and Veterinary Entomology* 4(1): 1–24. <https://doi.org/10.1111/j.1365-2915.1990.tb00255.x>
- Killick-Kendrick R, Tang Y, Killick-Kendrick M, Sang DK, Sirdar MK, Ke L, Ashford RW, Schorscher J, Johnson RH (1991) The identification of female sandflies of the subgenus *Larroussius* by the morphology of the spermathecal ducts. *Parassitologia* 33(Suppl): 335–347.
- Kornilios P, Thanou E, Lymberakis P, Ilgaz Ç, Kumlutaş Y, Leaché A (2019) Genome-wide markers untangle the green-lizard radiation in the Aegean Sea and support a rare biogeographical pattern. *Journal of Biogeography* 46(3): 552–567. <https://doi.org/10.1111/jbi.13524>
- Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS (1999) Chronology, causes and progression of the Messinian salinity crisis. *Nature* 400(6745): 652–655. <https://doi.org/10.1038/23231>
- Krüger A, Strüven L, Post RJ, Faulde M (2011) The sandflies (Diptera: Psychodidae, Phlebotominae) in military camps in northern Afghanistan (2007–2009), as identified by morphology and DNA ‘barcoding’. *Annals of Tropical Medicine & Parasitology* 105(2): 163–176. <https://doi.org/10.1179/136485911X12899838683241>
- Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33(7): 1870–1874. <https://doi.org/10.1093/molbev/msw054>
- Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B (2016) PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution* 34(3): 772–773. <https://doi.org/10.1093/molbev/msw260>
- Latrofa MS, Dantas-Torres F, Weigl S, Tarallo VD, Parisi A, Traversa D, Otranto D (2011) Multilocus molecular and phylogenetic analysis of phlebotomine sand flies (Diptera: Psychodidae) from southern Italy. *Acta Tropica* 119(2–3): 91–98. <https://doi.org/10.1016/j.actatropica.2011.04.013>
- Leaché AD, Fujita MK (2010) Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). *Proceedings of the Royal Society B: Biological Sciences* 277(1697): 3071–3077. <https://doi.org/10.1098/rspb.2010.0662>
- Léger N, Depaquit J (2002) Systématique et biogéographie des Phlébotomes (Diptera: Psychodidae). *Annales de la Société entomologique de France* 38(1–2): 163–175.
- Léger N, Pesson B (1987) Sur la taxonomie et la répartition géographique de *Phlebotomus (Adlerius) chinensis* sl et *P. (Larroussius) major* sl (Psychodidae-Diptera): statut des espèces présentes en Grèce. *Bulletin de la Société de pathologie exotique* 80(2): 252–260.
- Léger N, Pesson B, Madulo-Leblond G (1986) Les phlébotomes de Grèce. *Biologia gallo-hellenica* 11(2): 165–191.
- Lewis DJ (1982) A taxonomic review of the genus *Phlebotomus* (Diptera: Psychodidae). *Bulletin of the British Museum (Natural History)* 45(2): 121–209.
- Minh BQ, Nguyen MAT, von Haeseler A (2013) Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution* 30(5): 1188–1195. <https://doi.org/10.1093/molbev/mst024>
- Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32(1): 268–274. <https://doi.org/10.1093/molbev/msu300>
- Ogilvie HA, Bouckaert RR, Drummond AJ (2017) StarBEAST2 brings faster species tree inference and accurate estimates of substitution rates. *Molecular Biology and Evolution* 34(8): 2101–2114. <https://doi.org/10.1093/molbev/msx126>
- Omondi ZN, Demir S, Arserim SK (2020) Entomological survey of the sand fly fauna of Kayseri Province: focus on visceral and cutaneous Leishmaniasis in Central Anatolia, Turkey. *Turkish Journal of Parasitology* 44(3): 158–163. <https://doi.org/10.4274/tpd.galenos.2020.6751>
- Papadopoulou A, Anastasiou I, Keskin B, Vogler AP (2009) Comparative phylogeography of tenebrionid beetles in the Aegean archipelago: the effect of dispersal ability and habitat preference.

- Molecular Ecology 18(11): 2503–2517. <https://doi.org/10.1111/j.1365-294X.2009.04207.x>
- Papadopoulou A, Anastasiou I, Vogler AP (2010) Revisiting the insect mitochondrial molecular clock: the Mid-Aegean Trench calibration. *Molecular Biology and Evolution* 27(7): 1659–1672. <https://doi.org/10.1093/molbev/msq051>
- Parmakelis A, Klossa-Kilia E, Kiliadis G, Triantis KA, Sfenthourakis S (2008) Increased molecular divergence of two endemic *Trachelipus* (Isopoda, Oniscidea) species from Greece reveals patterns not congruent with current taxonomy. *Biological Journal of the Linnean Society* 95(2): 361–370. <https://doi.org/10.1111/j.1095-8312.2008.01054.x>
- Pesson B, Ready JS, Benabdennbi I, Martin-Sanchez J, Esseghir S, Cadi-Soussi M, Morillas-Marquez F, Ready PD (2004) Sandflies of the *Phlebotomus perniciosus* complex: mitochondrial introgression and a new sibling species of *P. longicuspis* in the Moroccan Rif. *Medical and Veterinary Entomology* 18(1): 25–37. <https://doi.org/10.1111/j.0269-283x.2004.0471.x>
- Poulakakis N, Kapli P, Lymberakis P, Trichas A, Vardinoyiannis K, Sfenthourakis S, Mylonas M (2015) A review of phylogeographic analyses of animal taxa from the Aegean and surrounding regions. *Journal of Zoological Systematics and Evolutionary Research* 53(1): 18–32. <https://doi.org/10.1111/jzs.12071>
- Prista GA, Agostinho RJ, Cachão MA (2015) Observing the past to better understand the future: a synthesis of the Neogene climate in Europe and its perspectives on present climate change. *Open Geosciences* 7(1): 65–83. <https://doi.org/10.1515/geo-2015-0007>
- Rambaut A (2018) FigTree v1.4.4, a graphical viewer of phylogenetic trees. University of Edinburgh: Institute of Evolutionary Biology.
- Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology* 67(5): 901–904. <https://doi.org/10.1093/sysbio/syy032>
- Rispail P, Léger N (1998) Numerical taxonomy of Old World Phlebotominae (Diptera: Psychodidae): 2. Restatement of classification upon subgeneric morphological characters. *Memórias do Instituto Oswaldo Cruz* 93(6): 787–793. <https://doi.org/10.1590/S0074-02761998000600016>
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61(3): 539–542. <https://doi.org/10.1093/sysbio/sys029>
- Secombe AK, Ready PD, Huddleston LM (1993) A catalogue of old world phlebotomine sandflies (Diptera: Psychodidae, Phlebotominae). *Occasional Papers on Systematic Entomology* 8: 1–57.
- Sfenthourakis S, Triantis KA (2017) The Aegean archipelago: a natural laboratory of evolution, ecology and civilisations. *Journal of Biological Research-Thessaloniki* 24: 4. <https://doi.org/10.1186/s407-09-017-0061-3>
- Simaiakis SM, Dimopoulou A, Mitrakos A, Mylonas M, Parmakelis A (2012) The evolutionary history of the Mediterranean centipede *Scolopendra cingulata* (Latreille, 1829) (Chilopoda: Scolopendridae) across the Aegean archipelago. *Biological Journal of the Linnean Society* 105(3): 507–521. <https://doi.org/10.1111/j.1095-8312.2011.01813.x>
- Song Y, Wang Q, An Z, Qiang X, Dong J, Chang H, Zhang M, Guo X (2018) Mid-Miocene climatic optimum: Clay mineral evidence from the red clay succession, Longzhong Basin, Northern China. *Palaeogeography, Palaeoclimatology, Palaeoecology* 512: 46–55. <https://doi.org/10.1016/j.palaeo.2017.10.001>
- Stothard P (2000) The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *BioTechniques* 28(6): 1102–1104. <https://doi.org/10.2144/00286ir01>
- Suc J-P, Popescu S-M, Fauquette S, Bessedik M, Jiménez Moreno G, Bachiri Taoufiq N, Zheng Z, Médail F, Klotz S (2019) Reconstruction of Mediterranean flora, vegetation and climate for the last 23 million years based on an extensive pollen dataset. *Ecologia Mediterranea* 44(2): 53–85. <https://doi.org/10.3406/ecmed.2018.2044>
- Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* 10(3): 512–526. <https://doi.org/10.1093/oxfordjournals.molbev.a040023>
- Trájer AJ (2021) Palaeoclimatic models – predicted changes in the potential Neogene distribution patterns of *Phlebotomus similis* and *Phlebotomus sergenti* (Insecta: Diptera: Psychodidae). *Palaeobiodiversity and Palaeoenvironments* 102(1): 149–172. <https://doi.org/10.1007/s12549-021-00483-2>
- Trájer AJ, Hammer T, Padišák J (2018) Reflection of the Neogene–Quaternary phylogeography in the recent distribution limiting climatic factors of eight Mediterranean *Phlebotomus* species (Diptera: Psychodidae). *Journal of Natural History* 52(27–28): 1763–1784. <https://doi.org/10.1080/00222933.2018.1485981>
- Trájer AJ, Sebestyén V, Padišák J (2021) The impacts of the Messinian Salinity Crisis on the biogeography of three Mediterranean sandfly (Diptera: Psychodidae) species. *Geobios* 65: 51–66. <https://doi.org/10.1016/j.geobios.2021.02.003>
- Trichas A, Smirli M, Papadopoulou A, Anastasiou I, Keskin B, Poulakakis N (2020) Dispersal versus vicariance in the Aegean: combining molecular and morphological phylogenies of eastern Mediterranean *Dendarus* (Coleoptera: Tenebrionidae) sheds new light on the phylogeography of the Aegean area. *Zoological Journal of the Linnean Society* 190(3): 824–843. <https://doi.org/10.1093/zoolinnean/zlaa-022>
- Tsirigotakis N, Pavlou C, Christodoulou V, Dokianakis E, Kourouniotis C, Alten B, Antoniou M (2018) Phlebotomine sand flies (Diptera: Psychodidae) in the Greek Aegean Islands: ecological approaches. *Parasites & Vectors* 11(97): 14. <https://doi.org/10.1186/s13071-018-2680-4>
- Tuon FF, Amato Neto V, Sabbaga Amato V (2008) *Leishmania*: origin, evolution and future since the Precambrian. *FEMS Immunology & Medical Microbiology* 54(2): 158–166. <https://doi.org/10.1111/j.1574-695X.2008.00455.x>
- Xanthopoulou K, Anagnostou V, Ivočić V, Djurkovic-Djakovic O, Rogozi E, Sotiraki S, Papa A (2011) Distribution of sandflies (Diptera, Psychodidae) in two Ionian islands and Northern Greece. *Vector-Borne and Zoonotic Diseases* 11(12): 1591–1594. <https://doi.org/10.1089/vbz.2011.0750>
- Yang Z (2006) *Computational molecular evolution*. Oxford: University Press. 376 pp.
- Yang Z, Rannala B (2010) Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences* 107(20): 9264–9269. <https://doi.org/10.1073/pnas.0913022107>
- Yang Z, Rannala B (2014) Unguided species delimitation using DNA sequence data from multiple loci. *Molecular Biology and Evolution* 31(12): 3125–3135. <https://doi.org/10.1093/molbev/msu279>
- Yu Y, Harris AJ, Blair C, He X (2015) RASP (Reconstruct Ancestral State in Phylogenies): A tool for historical biogeography. *Molecular*

Phylogenetics and Evolution 87: 46–49. <https://doi.org/10.1016/j.ympev.2015.03.008>

Yu Y, Harris AJ, He X (2010) S-DIVA (Statistical Dispersal-Vicariance Analysis): A tool for inferring biogeographic histories. *Molecular Phylogenetics and Evolution* 56(2): 848–850. <https://doi.org/10.1016/j.ympev.2010.04.011>

*Molecular Phylogenetics and Evolution* 56(2): 848–850. <https://doi.org/10.1016/j.ympev.2010.04.011>

## Supplementary material 1

### Table S1

**Authors:** Pavlou C, Dokianakis E, Tsirigotakis N, Christodoulou V, Özbel Y, Antoniou M, Poulakakis N (2022)

**Data type:** .docx

**Explanation note:** Information on studied specimens with NCBI accession numbers (n/a: not available).

**Copyright notice:** This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

**Link:** <https://doi.org/10.3897/asp.80.e78315.suppl1>

## Supplementary material 2

### Table S2

**Authors:** Pavlou C, Dokianakis E, Tsirigotakis N, Christodoulou V, Özbel Y, Antoniou M, Poulakakis N (2022)

**Data type:** .docx

**Explanation note:** The best-fit nucleotide substitution models for each locus/partition selected from PF under the BIC criterion.

**Copyright notice:** This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

**Link:** <https://doi.org/10.3897/asp.80.e78315.suppl2>

## Supplementary material 3

### Table S3

**Authors:** Pavlou C, Dokianakis E, Tsirigotakis N, Christodoulou V, Özbel Y, Antoniou M, Poulakakis N (2022)

**Data type:** .docx

**Explanation note:** Genetic distance (%) under the Tamura-Nei model between species for COI (below diagonal) and for CytB (above diagonal). Diagonal values represent the genetic distance within species (in parenthesis the CytB distances). n.a.: not available.

**Copyright notice:** This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

**Link:** <https://doi.org/10.3897/asp.80.e78315.suppl3>

## Supplementary material 4

### Table S4

**Authors:** Pavlou C, Dokianakis E, Tsirigotakis N, Christodoulou V, Özbel Y, Antoniou M, Poulakakis N (2022)

**Data type:** .docx

**Explanation note:** Genetic distance (%) under the Tamura-Nei model between species for EF1- $\alpha$  (below diagonal) and for TPI (above diagonal). Diagonal values represent the genetic distance within species (in parenthesis the TPI distances). n.a.: not available.

**Copyright notice:** This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

**Link:** <https://doi.org/10.3897/asp.80.e78315.suppl4>

## Supplementary material 5

### Table S5

**Authors:** Pavlou C, Dokianakis E, Tsirigotakis N, Christodoulou V, Özbel Y, Antoniou M, Poulakakis N (2022)

**Data type:** .docx

**Explanation note:** Genetic distance (%) under the Tamura-Nei model between species for 28S (below diagonal) and for ITS2 (above diagonal). Diagonal values represent the genetic distance within species (in parenthesis the ITS2 distances). n.a.: not available.

**Copyright notice:** This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

**Link:** <https://doi.org/10.3897/asp.80.e78315.suppl5>

## Supplementary material 6

### Figure S1

**Authors:** Pavlou C, Dokianakis E, Tsirigotakis N, Christodoulou V, Özbel Y, Antoniou M, Poulakakis N (2022)

**Data type:** .docx

**Explanation note:** Bayesian inference phylogenetic gene trees (mtDNA and nDNA). Numbers next to the branches represent posterior probabilities (left), bootstrap values (middle) and ultrafast bootstrap values (right).

**Copyright notice:** This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

**Link:** <https://doi.org/10.3897/asp.80.e78315.suppl6>

## Supplementary material 7

### Figure S2

**Authors:** Pavlou C, Dokianakis E, Tsirigotakis N, Christodoulou V, Özbel Y, Antoniou M, Poulakakis N (2022)

**Data type:** .docx

**Explanation note:** Geographical map indicating the information on dispersal and speciation events between and within areas, as calculated by S-DIVA analysis [A: Middle East (and Cyprus); B: Aegean islands; C: East Europe; D: West Europe; E: North Africa]. \*created with mapchart.net.

**Copyright notice:** This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

**Link:** <https://doi.org/10.3897/asp.80.e78315.suppl7>