



**HAL**  
open science

# Evolution, systematics and historical biogeography of sand flies of the subgenus *Paraphlebotomus* (Diptera, Psychodidae, Phlebotomus) inferred using restriction-site associated DNA markers

Astrid Cruaud, Véronique Lehrter, Guenaëlle Genson, Jean-Yves Rasplus, Jérôme Depaquit

## ► To cite this version:

Astrid Cruaud, Véronique Lehrter, Guenaëlle Genson, Jean-Yves Rasplus, Jérôme Depaquit. 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*, 2021, 15 (7), pp.e0009479. 10.1371/journal.pntd.0009479 . hal-03296081

**HAL Id: hal-03296081**

**<https://hal.science/hal-03296081>**

Submitted on 5 Aug 2021

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

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



Distributed under a Creative Commons Attribution 4.0 International License

## RESEARCH ARTICLE

# Evolution, systematics and historical biogeography of sand flies of the subgenus *Paraphlebotomus* (Diptera, Psychodidae, *Phlebotomus*) inferred using restriction-site associated DNA markers

Astrid Cruaud<sup>1</sup> , Véronique Lehrter<sup>2</sup> , Guenaëlle Genson<sup>1</sup>, Jean-Yves Rasplus<sup>1</sup> , Jérôme Depaquit<sup>1</sup> \*

**1** CBGP, INRAE, CIRAD, IRD, Montpellier SupAgro, Montpellier, Univ Montpellier, France, **2** Université de Reims Champagne Ardenne, ESCAPE EA7510, USC ANSES VECPAR, SFR Cap Santé, UFR de Pharmacie, Reims, France

 These authors contributed equally to this work.

‡ These authors are joint senior authors on this work.

\* [jerome.depaquit@univ-reims.fr](mailto:jerome.depaquit@univ-reims.fr)



## OPEN ACCESS

**Citation:** 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 Negl Trop Dis* 15(7): e0009479. <https://doi.org/10.1371/journal.pntd.0009479>

**Editor:** Remi N. Charrel, Aix-Marseille Université, FRANCE

**Received:** February 25, 2021

**Accepted:** May 15, 2021

**Published:** July 19, 2021

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pntd.0009479>

**Copyright:** © 2021 Cruaud et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Fastq reads for analysed samples are available as a NCBI

## Abstract

Phlebotomine sand flies are the main natural vectors of *Leishmania*, which cause visceral and tegumentary tropical diseases worldwide. However, their taxonomy and evolutionary history remain poorly studied. Indeed, as for many human disease vectors, their small size is a challenge for morphological and molecular works. Here, we successfully amplified unbiased copies of whole genome to sequence thousands of restriction-site associated DNA (RAD) markers from single specimens of phlebotomines. RAD markers were used to infer a fully resolved phylogeny of the subgenus *Paraphlebotomus* (11 species + 5 outgroups, 32 specimens). The subgenus was not recovered as monophyletic and we describe a new subgenus *Artemievus* subg. nov. Depaquit for *Phlebotomus alexandri*. We also confirm the validity of *Ph. riouxi* which is reinstated as valid species. Our analyses suggest that *Paraphlebotomus sensu nov.* originated ca 12.9–8.5 Ma and was possibly largely distributed from peri-Mediterranean to Irano-Turranian regions. Its biogeographical history can be summarized into three phases: i) a first split between *Ph. riouxi* + *Ph. chabaudi* and other species that may have resulted from the rise of the Saharan belt ca 8.5 Ma; ii) a Messinian vicariant event (7.3–5.3 Ma) during which the prolonged drought could have resulted in the divergence of main lineages; iii) a recent radiation event (3–2 Ma) that correspond to cycles of wet and dry periods in the Middle East and the East African subregions during the Pleistocene. Interestingly these cycles are also hypothetical drivers of the diversification of rodents, in the burrows of which *Paraphlebotomus* larvae develop. By meeting the challenge of sequencing pangenomics markers from single, minute phlebotomines, this work opens new avenues for improving our understanding of the epidemiology of leishmaniasis and possibly other human diseases transmitted by arthropod vectors.

Sequence Read Archive (PRJNA725143). COI sequences were uploaded in Genbank IDs #MZ049642-MZ049672. Data sets and trees have been uploaded on Zenodo (<https://doi.org/10.5281/zenodo.4721254>).

**Funding:** The authors received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist

## Author summary

Phlebotomine sand flies are tiny insects transmitting unicellular parasites called *Leishmania* worldwide. They are bad flyers and are known to be a group dispersing according to historical events. No invasive species are known within Phlebotomine sandflies. Their taxonomy includes about 40 genera. In the Old World, the subgenus *Paraphlebotomus* includes the main vectors of *Leishmania tropica*, the agent of a cutaneous leishmaniasis mostly in North Africa visceral and the Middle West. The goal of this study was to explore the phylogeny and the biogeographical history of this subgenus using a new generation sequencing technique. We successfully amplified DNA copies from the selected DNA extracts, some of them being more than 20 years old. We fully resolved the phylogeny of the subgenus *Paraphlebotomus* who was not recovered as monophyletic, meaning it does not include all the ascendants of last common ancestor. Consequently, the taxonomy was revised by the creation of a new subgenus and suggest that *Paraphlebotomus* originated about 10 million years ago. After a first split isolation North African species, the Messinian crisis caused by the aridification of the Tethys (the ancestor of the Mediterranean Sea) occurring 7 million years ago allowed the divergence of the main lineages including that of *Phlebotomus sergenti* the most important vector of the group. Most recent radiation events occurred during the Pleistocene. The historical biogeography of these Phlebotomine sandflies are closely related to that of the *Leishmania* they transmit.

## Introduction

Phlebotomine sand flies (Diptera, Psychodidae) are the main natural vectors of *Leishmania*, a genus of protozoan parasites which cause visceral and tegumentary neglected tropical diseases worldwide. Of the approximately 1,000 species of sand flies, about 70 are vectors of *Leishmania* [1,2]. Leishmaniasis are endemic in nearly 100 countries, mainly in Africa, Asia and Latin America [3]. The estimated annual incidence of cutaneous leishmaniasis in humans is 0.7 to 1 million cases and 50,000 to 90,000 cases of visceral leishmaniasis were reported in 2017 [4]. However, most leishmaniasis are zoonotic diseases affecting wild or domestic animals. Canine leishmaniasis is for instance a serious veterinary problem in many countries [14].

As for any vector-borne disease, an accurate systematics of the vectors is critical for a better understanding of the epidemiology of diseases. However, although phlebotomine sand flies are the subject of many research programs, their taxonomy and evolutionary history remain poorly studied [5], which limits our progress to reduce the incidence of leishmaniasis.

Phlebotomine are minute flies (2–3 mm), which makes their identification to species difficult, especially for non-specialists. Diagnostic characters are almost exclusively located on the head and genitalia of both sexes and their observation require that specimens are slide mounted, which further complicates taxonomic work. The small size of the phlebotomine sand flies is also a technical challenge for molecular work, which may explain why only few molecular studies were conducted so far. Because species complexes may exist, the sequencing of a pool of specimens is hazardous and DNA extracted from single specimens should be used for molecular work. In addition, as dissection of the head and genitalia is required for identification, DNA extraction must be performed only on remains of the body. As a result, DNA yield is low, which drastically limits the number of genes that can be amplified and sequenced from a single specimen using PCR and Sanger approaches. Thus, the amount of genetic

information that can be analysed to infer robust relationships and highlight possible species complexes is low.

In the Old World, most vectors of *Leishmania* belong to the genus *Phlebotomus* [5], and some are included in the genus *Sergentomyia* [6]. Eleven subgenera have been described in the genus *Phlebotomus* [7] among which, four are involved in the transmission of *Leishmania*, namely *Larroussius*, *Phlebotomus*, *Euphlebotomus* and *Paraphlebotomus*, and two others are suspected or local vectors (*Adlerius* and *Synphlebotomus*) (Table 1).

Based on morphological characters [8], *Paraphlebotomus*, *Phlebotomus* and *Synphlebotomus* are closely related (presence of a basal lobe on gonocoxite, short aedeagal or spermathecal ducts, short parameral sheath, annealed spermathecae without neck), while *Larroussius* (lack of basal lobe on gonocoxite, long aedeagal or spermathecal ducts, long parameral sheath, annealed spermathecae with neck) appears to be part of another evolutionary lineage also including *Adlerius*, *Euphlebotomus*, *Anaphlebotomus* or *Transphlebotomus*. Two studies based on morphological characters consider these two groups of subgenera as separate clades [9,10]. Studies based on ribosomal DNA, highlighted that the subgenera *Paraphlebotomus* and *Phlebotomus* belong to the same clade [11,12] but no molecular study has been published so far that included members of these related subgenera.

Here, we focus on the subgenus *Paraphlebotomus* that currently includes 14 species (Table 1), four of which being demonstrated vectors of *Leishmania* spp. The most important vector is undoubtedly *Ph. sergenti*, which is the main vector of *L. tropica*, the etiological agent of an anthroponotic cutaneous leishmaniasis occurring in a large area from Pakistan to Morocco [13,14]. However, *L. tropica* can be locally transmitted by other vectors belonging to

Table 1. Subgenera of *Phlebotomus* sand flies involved in *Leishmania* transmission.

Subgenus	Species	Main <i>Leishmania</i> species transmitted (suspected or proven)	References	Included in the study
<i>Adlerius</i>	<i>Ph. simici</i>	<i>L. infantum</i>	[1]	x
	<i>Ph. arabicus</i>	<i>L. tropica</i>	[17]	
<i>Euphlebotomus</i>	<i>Ph. argentipes</i>	<b><i>L. donovani</i></b>	[1,94]	/
<i>Larroussius</i>	<i>Ph. neglectus</i>	<b><i>L. infantum</i></b>	[95]	x
<i>Paraphlebotomus</i> <i>sensu</i> Lewis or Depaquit [7, 54]	<i>Ph. alexandri</i>	<i>L. donovani</i>	[1,96]	x
		<b><i>L. infantum</i></b>	[97]	
	<i>Ph. andrejevi</i>	<i>L. major</i>	[98]	/
	<i>Ph. caucasicus</i>	<b><i>L. major</i></b>	[99]	x
	<i>Ph. chabaudi</i>	/	/	x
	<i>Ph. gemetchi</i>	/	/	/
	<i>Ph. jacusieli</i>	/	/	x
	<i>Ph. kazeruni</i>	<i>L. major</i>	[1,100]	x
	<i>Ph. mireillae</i>	/	/	x
	<i>Ph. mongolensis</i>	/	/	x
	<i>Ph. nuri</i>	/	/	/
	<i>Ph. riouxi</i>	/	/	x
	<i>Ph. saevus</i>	<b><i>L. tropica</i></b>	[101]	x
	<i>Ph. sergenti</i>		<b><i>L. tropica</i></b>	[13,14,101]
		<b><i>L. aethiopica</i></b>	[101]	
	<i>Ph. similis</i>	/	/	x
<i>Phlebotomus</i>	<i>Ph. duboscqi</i>	<b><i>L. major</i></b>	[102]	x
	<i>Ph. papatasi</i>	<b><i>L. major</i></b>	[103]	x
<i>Synphlebotomus</i>	<i>Ph. ansarii</i>	<i>L. major</i>	[104]	x

<https://doi.org/10.1371/journal.pntd.0009479.t001>

subgenera *Larroussius* in Kenya [15,16] and *Adlerius* in Israel/Palestine [17]. So far, only three molecular studies have included enough species of *Paraphlebotomus* to provide preliminary hypotheses about its systematics. Depaquit et al. [18] have conducted the only molecular study that attempted to resolve phylogenetic relationships within *Paraphlebotomus* using nine ingroup and two outgroup species. However, a single marker (*Internal transcribed spacer 2*, ITS2) was used and the resolution was low. Besides, we cannot exclude that possible paralogs of ITS2 may have biased results. A second study [19], whose goal was to describe the diversity of sand flies in military camps in Afghanistan using mitochondrial Cytochrome b gene (*Cytb*), generated data for five species of *Paraphlebotomus*. In this study, *Paraphlebotomus* appeared as a polyphyletic assemblage. *Ph. alexandri* grouped with species from the subgenus *Phlebotomus* (*Ph. papatasi* and *Ph. bergeroti*) with poor support. Finally, Grace-Lema et al. [20] used a patchy matrix of 7 genes (65% missing data, ca 45% of the species with more than 85% missing data) to infer a phylogeny of phlebotomine sand flies (113 terminals). However, as for the rest of the tree, relationships between the four species of *Paraphlebotomus* were poorly resolved. Thus, the systematics of the subgenus is still unclear.

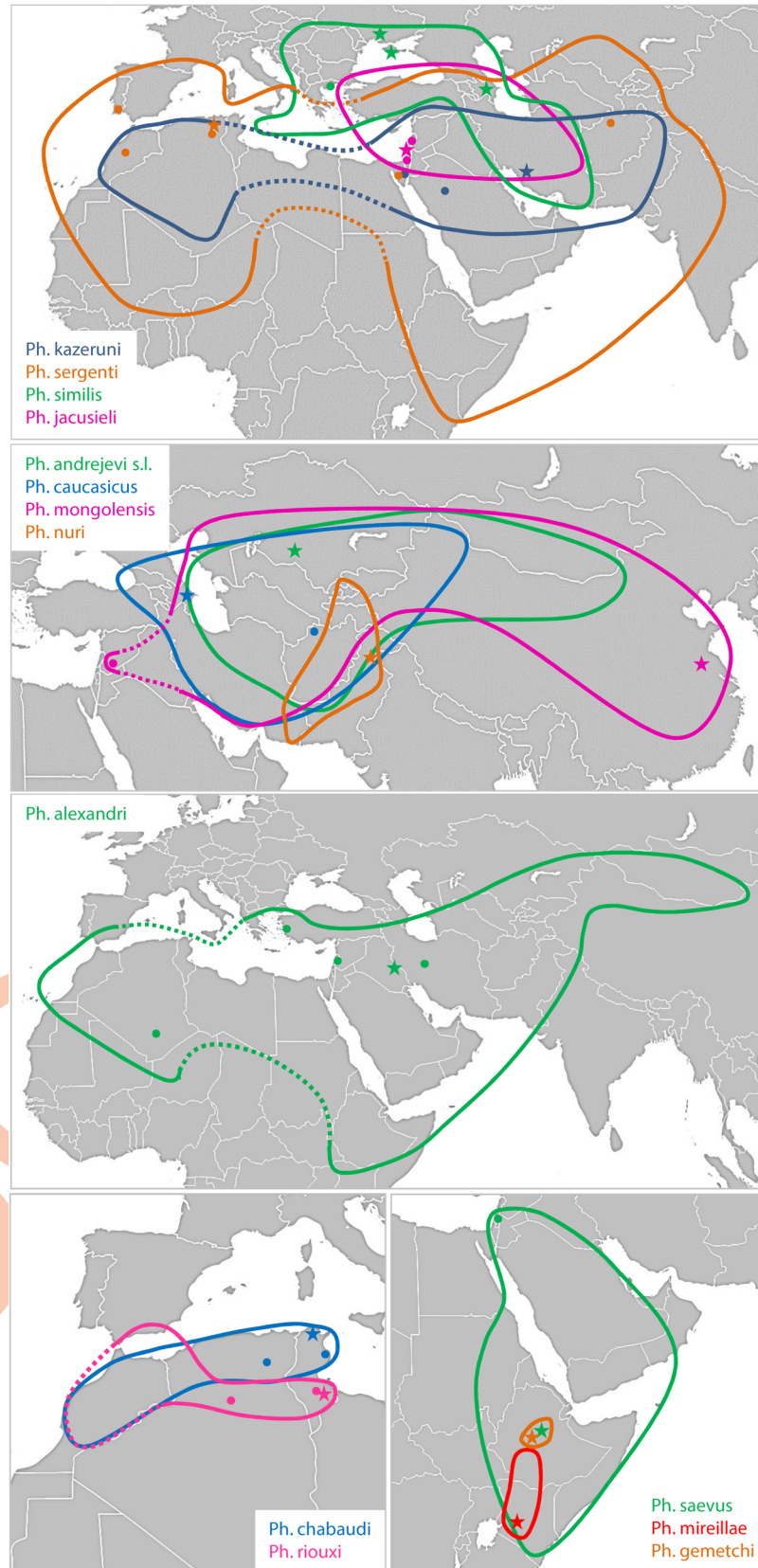
Most species of *Paraphlebotomus* have restricted distribution (Fig 1; see also [21]). *Ph. mir-eillae* is found in Kenya and southern Ethiopia [22]; *Ph. saevus* occurs in Arabian Peninsula and in the Horn of Africa; *Ph. nuri* is a Middle-East endemic and *Ph. gemetchi* is reported from Ethiopia [23]. *Ph. chabaudi* and *Ph. riouxi* are Ibero-Maghrebian and *Ph. similis* occurs in the north-east of the Mediterranean basin, with recent records from central Turkey [24], Iran [25] and Ukraine [26]. *Ph. jacusieli* is distributed from eastern Mediterranean to eastern Iran [27]. *Ph. caucasicus*, *Ph. andrejevi* and *Ph. mongolensis* are located in central Asia or Middle-east, but the range of *Ph. mongolensis* extends to East China sea (Suizhong, Liaoning province) [28]. To the opposite, a few species have extremely large distribution. *Ph. sergenti* and *Ph. alexandri* occur from Europe and North Africa to Middle and Far East. *Ph. kazeruni* is almost always a rare species, mainly recorded from Iran, but recent records in Turkey [29], southern Algeria [30] and Pakistan [31,32] have widened its known range. As suggested by Depaquit et al. [18], the ancestor of the subgenus *Paraphlebotomus* may have originated in Eastern Africa or in the Arabic peninsula, however this needs to be formally tested.

With technical advances, it is now possible to generate unbiased copies of whole genomes (Whole Genome Amplification, WGA) to circumvent low DNA amount issues [33–35]. Although this technic is still rarely used for such purpose, it opens new avenues for the sequencing of pangenomic markers to resolve relationships among tiny species. Here we used WGA and RAD (Restriction site Associated DNA) sequencing [36] of single specimens of phlebotomine sand flies to infer a robust phylogeny of the subgenus *Paraphlebotomus* and five outgroups. We also performed divergence time estimates and probabilistic inference of historical biogeography of the subgenus.

## Materials and methods

### Taxonomic sampling

Twenty-three ingroup specimens representing 11 species (78% of the known diversity) were selected as ingroup. In average, one or two specimens per species were analysed. We failed to collect two rare species (*Ph. nuri* and *Ph. gemetchi*) and to process old DNA extracts of *Ph. andrejevi*. For demonstrated vectors, four to five specimens collected in several locations were included. Nine outgroup specimens (five species) were chosen in the subgenera *Phlebotomus*, *Synphlebotomus*, *Larroussius* and *Adlerius*. Sampling was carried out during the last twenty years, from a large area extending from Senegal to India (west-east axis) and from Italy to Kenya (north-south axis) (Fig 1 and S1 Table). Samples were collected using CDC trap, sticky



**Fig 1. Distribution maps and sampling sites for the different species of *Paraphlebotomus sensu Lewis* [7] and Depaquit [54].** Ranges are delimited by color lines. Lines are dotted when the presence is expected but has never been confirmed (lack of sampling, doubtful taxonomic status). Sampling sites and type localities are respectively represented by dots and stars. The background map was in the public domain and has been downloaded from <https://commons.wikimedia.org/wiki/File:BlankMap-World.svg>.

<https://doi.org/10.1371/journal.pntd.0009479.g001>

trap, or came from colonies reared at lab. Specimens were preserved in 95% ethanol at room temperature or at  $-20^{\circ}\text{C}$  until dissection. Some samples were dried freeze at  $-80^{\circ}\text{C}$  or in liquid nitrogen for longer conservation. After thawing, specimens were dissected in 95% ethanol. For morphological identification under stereomicroscope, head and genitalia were mounted in chloral gum or in Euparal between slide and cover slide after clarification in boiling Marc André solution, washing and dehydration. The rest of the body was dried freeze at  $-20^{\circ}\text{C}$  for molecular biology analyses. Identification to species was performed by JD and vouchers (slides + body remains after DNA extractions) have been deposited in the University of Reims Champagne-Ardenne, collection of J. Depaquit.

### DNA extraction and whole genome amplification

DNA extractions were performed with either CTAB/Chloroform-Isoamyl Alcohol or extraction kits. For the CTAB/Chloroform-Isoamyl Alcohol extraction, samples were crushed and incubated at  $60^{\circ}\text{C}$  in a solution containing CTAB buffer [37]. Proteins were eliminated by the addition of chloroform-Isoamyl Alcohol. RNase (0.5 units) was added to the aqueous phase, which was then incubated at  $37^{\circ}\text{C}$  for 30 min to remove RNA. Total genomic DNA was precipitated by the addition of two-thirds of the total volume of isopropanol and stored at  $4^{\circ}\text{C}$  overnight. After centrifugation, the pellet was washed, dried and suspended in 50  $\mu\text{L}$  of ultra-pure water. Genomic DNA was also extracted using the QIAamp DNA Mini Kit (Qiagen) following the manufacturer's instructions. DNA was eluted using 100  $\mu\text{L}$  to 160  $\mu\text{L}$  of elution buffer.

DNA quantity was measured using the Qubit double-stranded DNA High Sensitivity Assay kit (ThermoFisher Scientific) following the manufacturer's instructions. DNA extracts were then subjected to Whole Genome Amplification using the GenomiPhi V2 DNA Amplification kit (GE Healthcare). Ethanol precipitation of DNA was performed prior to WGA to fit with the recommendations of the GenomiPhi protocol (1  $\mu\text{L}$  DNA input at 10 ng/ $\mu\text{L}$ ). 1/10 volume of sodium acetate 3 M pH 5.2 was added to DNA. Then, 2 volumes of cooled absolute ethanol were added to the mix. The mix was incubated at  $-20^{\circ}\text{C}$  overnight. The pellet of DNA extract obtained by centrifugation (30 min, 13 000 rpm,  $4^{\circ}\text{C}$ ) was washed by 500  $\mu\text{L}$  of cooled 70% ethanol. After a new centrifugation (15 min, 13 000 rpm,  $4^{\circ}\text{C}$ ), the pellet was dried at room temperature, and suspended in 4  $\mu\text{L}$  of sterile molecular biology ultrapure water. Whole Genome Amplification (WGA) was then performed on 1 to 2  $\mu\text{L}$  of concentrate DNA following the manufacturer protocol.

### DNA barcoding

The standard *COI* barcode fragment was sequenced for all specimens to obtain calibrations for divergence time estimates (see below). For each sample, the cocktail of M13 tailed primers and the PCR conditions described in Germain et al. [38] were used on 4  $\mu\text{L}$  of DNA obtained with WGA. Unpurified PCR products were sent to Eurofins Genomics for Sanger sequencing. Forward and reverse strands were assembled using Geneious 10.2.2 [39]. Consensuses were aligned with MAFFT-7.313 (linsi option) [40] and translated to amino acids using MEGA 7 [41] to detect frame-shift mutations and premature stop codons, which may indicate the

presence of pseudogenes. Sequences were merged with all barcodes available in BOLD [42] for the genus *Phlebotomus* (last access December 4, 2019) and IQ-TREE v1.6.7 [43] was used to infer a maximum likelihood phylogenetic tree. The best-fit evolutionary model was selected with ModelFinder as implemented in IQ-TREE [44] and branch supports were assessed with ultrafast bootstrap [45] and SH-aLRT test [46] (1000 replicates).

### Preparation and sequencing of RAD library

In silico digestion of the genome of *Ph. papatasi* available on GenBank (AJVK00000000.1) was performed using the recognition site of the *PstI* enzyme to determine the number of expected cut-sites and calculate the maximum number of samples to be sequenced on a single lane of an Illumina HiSeq 2500. 51,712 cut sites (103,424 RAD markers) were inferred. Considering that ca 150,000,000 passing filter reads would be usable to construct RAD loci according to our experience (after removal of reads with mutation in barcodes /restriction sites and PCR duplicates), 48 samples were included in the library to get a sequencing depth per marker of ca 30X (only 32 samples were analysed for this study, the 16 others were sequenced for another purpose). Library construction followed by Baird et al. [36] and Etter et al. [47] with modifications detailed in Cruaud et al. [48]. The quantity of P1 adapters to be added to saturate restriction sites (result = 2  $\mu$ L) as well as the optimal time for DNA sonication in a Covaris S220 ultrasonicator (result = 70s) that are both specific to the studied group were evaluated in a preliminary experiment. After tagging with barcoded P1 adapters and prior to sonication, specimens were pooled eight by eight (six pools were thus obtained and tagged with barcoded P2 adapters). 2\*125nt paired-end sequencing of the library was performed at MGX-Montpellier GenomiX on one lane of an Illumina HiSeq 2500 flow cell.

### RAD data cleaning and phylogenetic inference

The quality of the raw data was assessed using FastQC (Babraham Institute, Bioinformatics Group <http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>). Raw data were processed using the Perl wrapper RADIS [49] that relies on Stacks [50,51] for demultiplexing of data, removing PCR duplicates and building individual and catalog loci. For data cleaning (RADIS\_step1\_data\_cleaning.pl), parameters in the RADIS.cfg files were set as follows (radis\_readlen = 125; radis\_ntrim\_read1\_5p = 5 to remove the complementary overhang of the restriction site; radis\_ntrim\_read1\_3p = 0; radis\_ntrim\_read2\_5p = 0, radis\_ntrim\_read2\_3p = 0).

Individual RAD loci were built using m (minimum depth of coverage to create a stack) = 3 and, M (maximum distance allowed between stacks) = 2; with removal (r) and deleveraging (d) algorithms enabled. Six data sets were then built to test a possible impact of analytical parameters on our phylogenetic analyses. Catalogs of loci were built with n (number of mismatches allowed between sample tags when generating the catalog) set to 6, 8 or 10. Loci / samples selection was performed as follows: only samples with more than 100 RAD loci in the catalog were kept (radis\_nloci\_min), only locus for which at least either 50% or 75% of the samples had sequences were kept (radis\_nsamples\_min) and loci for which at least one sample had 3 sequences or more were removed to discard possible paralogs (radis\_npbloci\_cutoff). Phylogenetic inferences were performed using RAXML 8.2.9 (multi-threaded version with AVX extensions) [52] and a rapid bootstrap search with 100 replicates, followed by a thorough ML search (GTR GAMMA), without partitioning of the data set. Trees were visualized and annotated with TreeGraph 2.13.0-748beta [53]. Percentage of missing data and percentage of identity between sequence pairs were calculated using Geneious 10.2.2 [39].



## Species distribution, divergence time estimation and ancestral range estimation

Species distribution for subgenus *Paraphlebotomus* were estimated from maps of two precedent studies [7,54] and a review of recent papers (see [results](#) for an exhaustive list). There is no fossil record of *Phlebotomus* and paleogeographic events that could be used as calibrations are unknown. Consequently, we used calibration priors based on mitochondrial DNA. We estimated minimum and maximum pairwise sequence divergences among *COI* sequences of samples belonging to the two species pairs recovered both in the RAD and *COI* trees (i.e. *Ph. chabaudi-Ph. riouxi* and *Ph. papatasi-Ph. duboscqi*) using MEGA7 (evolutionary model = K2P). Divergences were then divided by the *COI* clock rate for insects estimated by Papadopoulou et al. [55] (i.e.  $3.54 \pm 0.38\%$  divergence per Myr) to obtain minimum and maximum ages for the split between the two species of each pair following Godefroid et al. [56]. A timetree was built with the RelTime method [57] as implemented in MEGA X [58] using the RelTime-ML option, a GTR+G evolutionary model (with 4 discrete Gamma categories as for the RAxML analysis) and all nucleotide sites (including gaps and missing data). The RAxML RAD tree obtained with the *cstacks*  $n = 10$ , *radis\_nsample\_min* = 50% of the samples (Table 2; ie the tree obtained with the largest data set) was used as input and *Ph. (Adlerius) simici* and *Ph. (Larrousius) neglectus* were considered as outgroups. Uniform distributions were used as calibration densities.

We inferred the biogeographic history of the subgenus *Paraphlebotomus* using the R package BioGeoBEARS 1.1.1 [59]. The chronogram built with RelTime was used as input but only one specimen per species was included in the analysis and outgroups (including *Ph. alexandri*) were removed to avoid artefacts (taxa were pruned from the chronogram with the R package *ape* [60]). We compared the fit of the models DEC, BAYAREALIKE, and DIVALIKE with AICc but did not consider the jump parameter for founder events (+) following recent criticisms [61]. Geographic ranges of the species were coded using six geographical subregions of the West Palearctic: A: West Mediterranean; B: Saharo-Arabian; C: Sudanian; D: Somalian; E: East Mediterranean; F: Irano-Turanian. The maximum number of areas species may occupy was set to 6. To consider the main geologic events that occurred during the last 15 Myr in the peri-mediterranean area, we defined two time periods with different dispersal rate scalars: 1) from the Mid Miocene climatic optimum (ca 15 Ma) to the end of the Messinian crisis (5.3 Ma); 2) from the Zanclean flood (5.3 Ma) to present. The matrix of geographical ranges for the

**Table 2. Properties of the RAD data sets obtained with different parameters combinations.** *n* = number of mismatches allowed when merging individual loci in *cstacks*, *radis\_nsample\_min* = minimum number of samples that should have sequences for a RAD locus to be included in the analysis. In all data sets, loci for which at least one sample had three or more sequences (i.e. possible paralogs) were removed (*radis\_npbloci\_cutoff* = 3).

Combination of parameters	# RAD markers	# bases in alignment	% missing data	% variable sites	% parsimony informative sites	% GC	RAxML tree
<i>n</i> = 6, <i>radis_nsample_min</i> = 50% of the samples	5,050	580,750	34.6	12.0	8.40	45.4	<a href="#">S2A Fig</a>
<i>n</i> = 6, <i>radis_nsample_min</i> = 75% of the samples	1,423	163,645	16.7	11.5	8.20	45.9	<a href="#">S2B Fig</a>
<i>n</i> = 8, <i>radis_nsample_min</i> = 50% of the samples	7,194	827,310	34.1	14.7	10.6	45.8	<a href="#">S2C Fig</a>
<i>n</i> = 8, <i>radis_nsample_min</i> = 75% of the samples	2,086	239,890	16.6	14.4	10.6	46.4	<a href="#">S2D Fig</a>
<i>n</i> = 10, <i>radis_nsample_min</i> = 50% of the samples	8,906	1,024,190	33.1	16.9	12.4	46.1	<a href="#">S2E Fig</a>
<i>n</i> = 10, <i>radis_nsample_min</i> = 75% of the samples	2,808	322,920	16.7	16.9	12.7	46.7	<a href="#">S2F Fig</a>

<https://doi.org/10.1371/journal.pntd.0009479.t002>

different species as well as the dispersal multipliers for each time period are available as [S1 Text](#).

### Computational resources

Analyses were performed on the Genotoul Cluster (INRA, France, Toulouse).

## Results

### DNA barcoding and calibrations

*COI* sequences were successfully obtained from all samples but one (*Ph. saevus* for which all DNA was included in the RAD library). The maximum likelihood tree confirmed that *COI* is not informative enough to solve relationships among species of phlebotomine sand flies ([S1 Fig](#)). To avoid issues due to mitochondrial introgression and putative paralogs (heteroplasms, coding pseudogenes) that could be contained in the *COI* data set, only splits between highly supported species pairs that were recovered both in the RAD ([S2 Fig](#)) and the *COI* trees were used to calibrate our dating analysis, namely *Ph. chabaudi-Ph. riouxi* and *Ph. papatasi-Ph. duboscqi*. Minimum and maximum interspecific K2P distances resulted in minimum and maximum ages of 2.42–3.64 Myr for the *Ph. chabaudi-Ph. riouxi* split (rounded to 2.0–4.0 Myr for the RelTime analysis) and 2.83–4.27 Myr for the *Ph. papatasi-Ph. duboscqi* split (rounded to 2.0–5.0 Myr for the RelTime analysis).

### RAD data sets and phylogenetic inferences

In average, 97,054 RAD loci were obtained per specimen (ustacks, [S1 Table](#)). Depending on the value of the parameters *n* (number of mismatches allowed when merging individual loci) and *radis\_nsample\_min* (minimum number of samples that should have sequences for a RAD locus to be included in the analysis) data sets included 1,423 to 8,906 RAD loci of 115 bp each and contained only 16.6–34.6% missing data ([Table 2](#)).

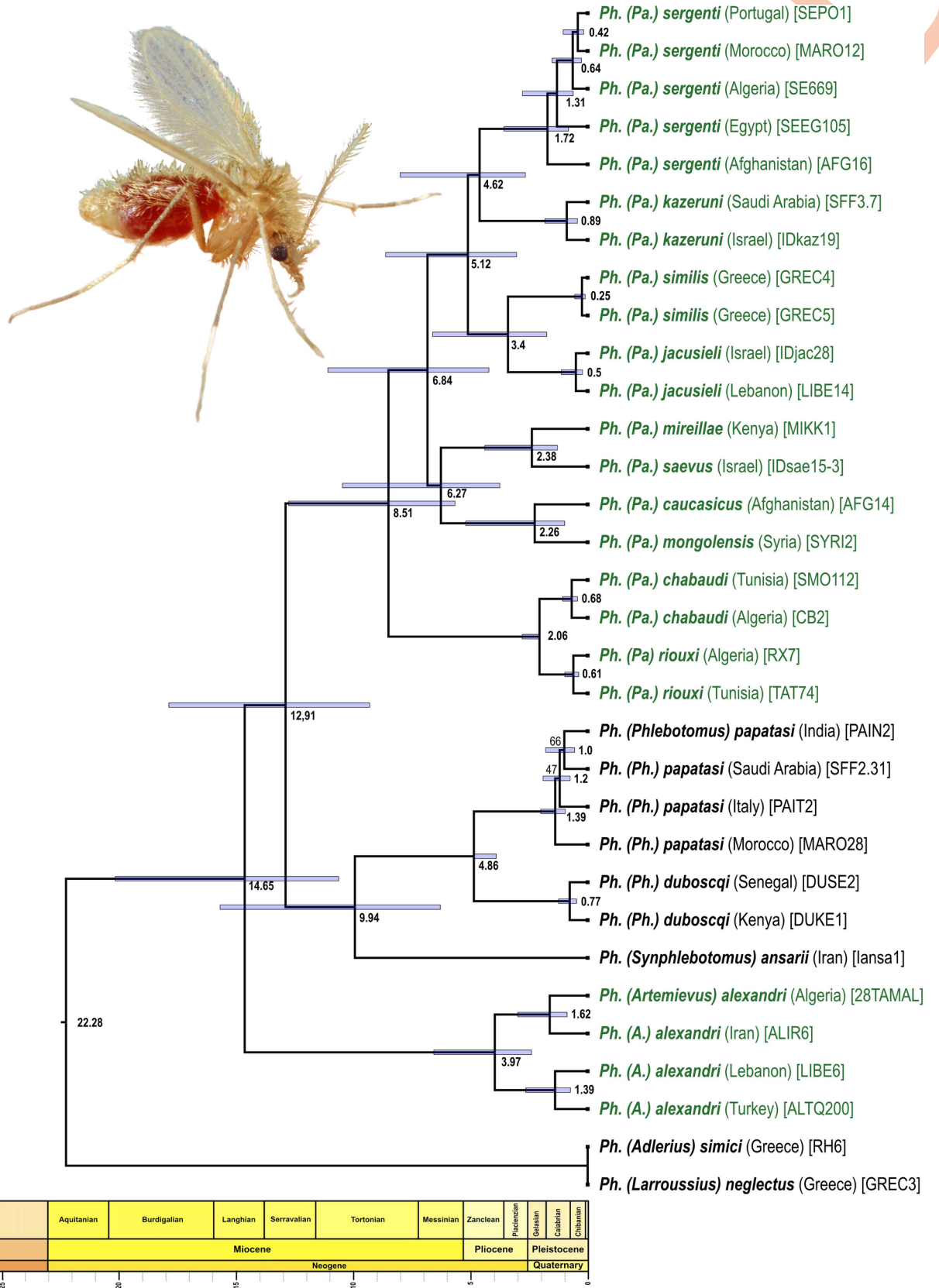
All RAxML analyses yielded the same fully resolved tree topology (Figs 2 and [S2](#)). The subgenus *Paraphlebotomus* was paraphyletic. *Ph. alexandri* was recovered sister to the rest of *Paraphlebotomus* species and a clade consisting of members of the *Phlebotomus* (*Ph. papatasi* and *Ph. duboscqi*) and *Synphlebotomus* (*Ph. ansarii*) subgenera. The remaining *Paraphlebotomus* species clustered into a strongly supported clade in which relationships between species were fully resolved.

### Species validity and intraspecific variability

When multiple conspecific specimens were analysed, they formed well-supported clades that were deeply separated from their nearest relative, suggesting that morphology alone mostly enables accurate identification of species. This was the case for the analysed specimens belonging to *Ph. chabaudi* and *Ph. riouxi* with an average pairwise percentage dissimilarity between specimens of the two clades of 1.7% for the *n* = 10, *radis\_nsample\_min* = 50% data set (i.e. on 1,024,190 nt). Significant genetic divergence was observed between individuals of species with large distribution ranges (*Ph. alexandri*—up to 1.6% dissimilarity and *Ph. sergenti*—up to 1.2%).

### Divergence-time estimates

RelTime analysis inferred that the split between *Ph. alexandri* and all other lineages has occurred about 15 Ma (14.7 Ma, confidence interval—CI = 10.6–20.2) and that the subgenus *Paraphlebotomus* (excluding *Ph. alexandri*) originated in the Middle—late Miocene between 12.9 (CI = 9.3–17.9; most ancient ancestor) and 8.5 Ma (CI = 5.7–12.8; most recent common



**Fig 2. RelTime timetree.** Bootstrap supports lower than 100 are reported at nodes (100 replicates, RAD data set used: cstacks n = 10, radis\_nsample\_min = 50% see Table 2 and S2 Fig). Species belonging to subgenus *Paraphlebotomus sensu* Seccombe et coll., 1993 are shown in green but terminals are annotated following the new classification proposed in this study with *Ph. alexandri* belonging to *Artemievus* subg. nov. Depaquit. Divergence times are not displayed in the outgroup (*Ph. (Adlerius) simici* and *Ph. (Larrousius) neglectus*) because RelTime uses evolutionary rates from the ingroup to calculate divergence times and does not assume that evolutionary rates in the ingroup clade apply to the outgroups. The image shows *Ph. papatasi* and has been modified from the image #10275 of the Centers for Disease Control and Prevention's Public Health Image Library (PHIL: <https://phil.cdc.gov/>; Content provider: CDC/ Frank Collins. Photo credit: James Gathany; last access January 22, 2021) that is in the public domain and thus free of any copyright restrictions.

<https://doi.org/10.1371/journal.pntd.0009479.g002>

ancestor MRCA) (Fig 2). Within *Paraphlebotomus*, the “*Ph. chabaudi* / *Ph. riouxi*” lineage split from all other *Paraphlebotomus* sand flies about 8.5 Ma by the end of Miocene. All deep splits between geographical lineages occurred between 6.8 and 5.1 Ma, although with relatively large confidence intervals. Finally, most splits between sister species occurred during the Pliocene between 4.6 Ma (*Ph. sergenti* / *Ph. kazeruni*) and 2.3 Ma (*Ph. caucasicus* / *Ph. mongolensis*). The main split within *Ph. alexandri* also occurred during this time period (ca 4.0 Ma).

## Biogeography

AICc values show that the DEC and DIVALIKE models have an equally significant fit (S3 Text). The ancestral range of *Paraphlebotomus* is equivocal with Irano-Turanian + West Mediterranean subregions as the most likely range (Fig 3 and S4 Text). Each of the five species pairs (namely *riouxi-chabaudi*; *mongolensis-caucasicus*; *saevus-mireillae*; *jacusieli-similis* and *kazeruni-sergenti*) originated in a different subregion of the West-Palaearctic (respectively West-Mediterranean, Irano-Turanian, Somalian, East-Mediterranean and Saharo-Arabian), suggesting an origin by vicariant fragmentation. The only difference between the DEC and DIVALIKE models is the most likely range for the MRCA of the clade *mongolensis-caucasicus*; *saevus-mireillae*; *jacusieli-similis* and *kazeruni-sergenti*, which is Somalian + Irano-Turanian subregions for DEC and Irano-Turanian only for DIVALIKE (S4 Text).

## Discussion

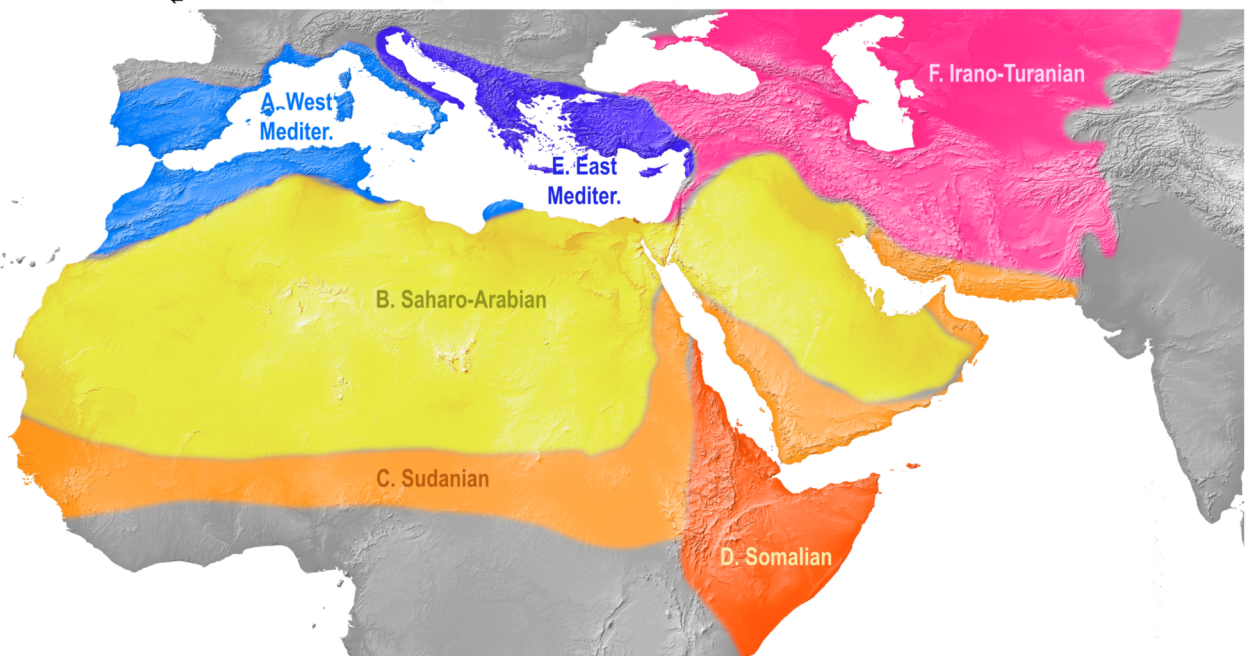
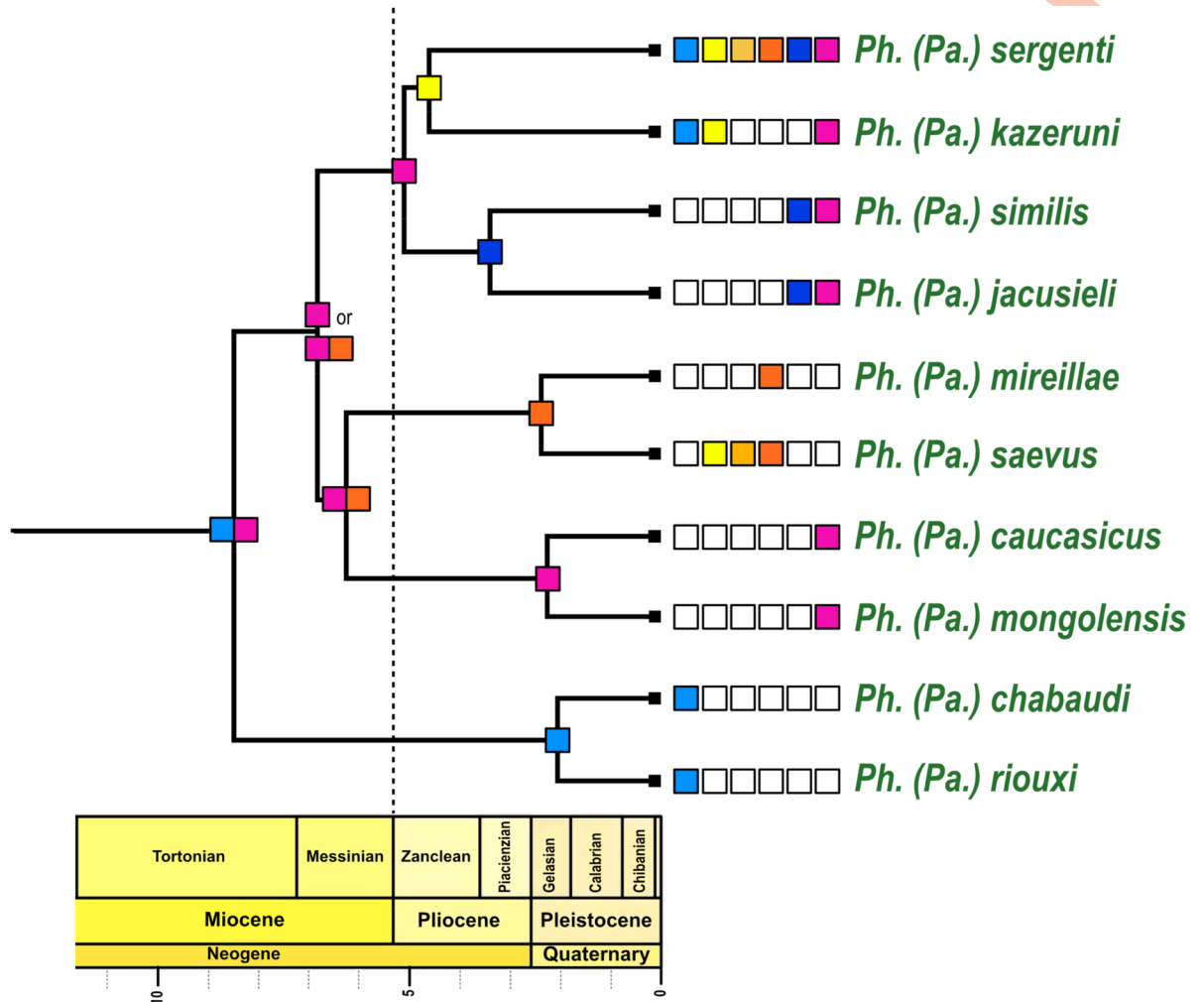
### Description of a new subgenus of Phlebotomus Rondani & Berté

*Ph. alexandri* did not cluster with other species of *Paraphlebotomus*, which questions its current classification. An early divergence of *Ph. alexandri* [18] as well as the paraphyly of the subgenus *Paraphlebotomus* have already been acknowledged [19] and is confirmed here with high statistical support. Based on morphology, *Ph. alexandri* greatly differs from other species of *Paraphlebotomus* [54]. The first flagellomere (f1 = antennal segment AIII) is extremely short and the basal lobe of the coxite is wider than long. The pharyngeal armature is poorly extended and composed by a network of conspicuous scaly teeth. These apomorphies lead us to propose a new subgenus for *Ph. alexandri*.

*Artemievus* subg. nov. *Depaquit*. Type-species: *Phlebotomus alexandri* Sinton, 1928.

*Artemievus* subg. nov. is defined by i) a short (<170µm) flagellomere 1 (= A III), shorter than labrum and shorter than the combined length of flagellomeres 2 and 3 (= AIV + AV) in both genders, ii) in females, a rectangular pharyngeal armature occupying less than a quarter of the pharynx, spermathecae with seven to 10 rings, ascoids reaching the next articulation, iii) in males, basal lobe of gonocoxite not incurved, style with four spines, the distal one carried by a long process, paramere simple and spoon-like and a narrow sperm pump (Figs 4 and 5).

This subgenus includes only *Ph. (A.) alexandri*, a widespread and polymorphic species extending from Morocco to Mongolia and from Ethiopia to southern Europe (Fig 1).



**Fig 3. Ancestral range estimation for the subgenus *Paraphlebotomus sensu nov*.** Geographic areas as delineated on the map are represented with colored boxes. A: West Mediterranean region; B: Saharo-Arabian region; C: Sudanian region; D: Somalian region; E: East Mediterranean region; F: Irano-Turanian. Most probable ancestral states (areas or combination of areas) as inferred by DEC and DIVALIKE are indicated at each node of the phylogeny. When states are different between the two models, the most probable state as inferred by DIVALIKE is indicated first. Input data (matrix of geographic ranges for species included in the analysis, time periods and dispersal multipliers); ancestral range estimations for all tested models (DEC, DIVALIKE, BAYAREALIKE), and results of model testing are available in S1–S4 Text. The background map was in the public domain and has been downloaded from Natural Earth (<http://www.natureearthdata.com>).

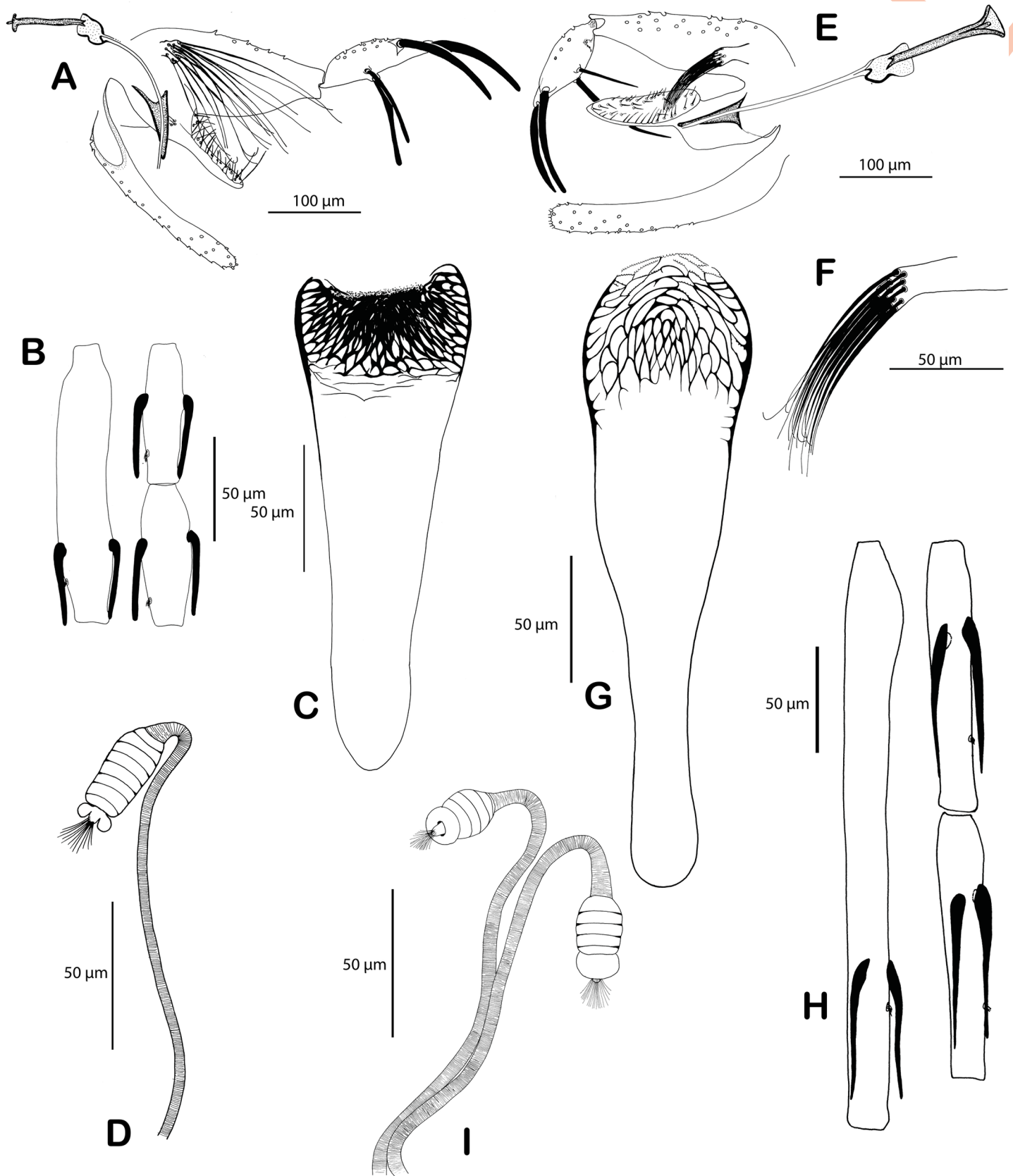
<https://doi.org/10.1371/journal.pntd.0009479.g003>

### Systematics of the subgenus *Paraphlebotomus sensu nov*

Our phylogenomic analysis (Figs 2 and S2) recovers fully resolved relationships within subgenus *Paraphlebotomus sensu nov*, that corresponds well to morphological affinities. The first clade formed by *Ph. chabaudi* and *Ph. riouxi* is supported by several morphological characters. Females of these North African species exhibit a typical spermathecae with a bellflower-like distal ring and males differ from all other species by their parameral sheaths that are apically pointed. *Ph. chabaudi* and *Ph. riouxi* can be differentiated by the size of the basal lobes of their gonocoxites and the number of setae they harbor. Based on dubious molecular evidences [and despite the tree presented in Fig 4 of their 2014 paper shows that specimens of the two species belong to two different clades], Tabbabi et al. [62,63] considered these species as synonyms. This result was acknowledged by Grace-Lema et al. [20], who wrongly interpreted a highly supported sister-taxa relationship between *Ph. chabaudi* and *Ph. riouxi* as a support for their synonymy, although they noticed the long branch between the two species. The high genetic distance is mis-interpreted by the author as a consequence of missing data, which is an inconsistent statement as other analysed species show the same level of divergence and were not considered as synonyms. The present study confirms the validity of *Ph. riouxi*, that was already assessed with Sanger data and morphological evidence by Lehrter et al. [64]. Therefore *Ph. riouxi* stat. rev. is reinstated as a valid species, from its status of junior synonym of *Ph. chabaudi*.

The clade grouping all other species of *Paraphlebotomus* is subdivided into four pairs of sister taxa that exhibit strong morphological affinities:

1. Females of the Afrotropical *Ph. saevus* and *Ph. mireillae* mainly differ by the presence of spines on spermathecal ducts. Males, which are extremely close, differ by the distal part of the parameral sheath (thicker in *Ph. saevus* than in *Ph. mireillae*), and the hairs of the gonocoxite basal lobe (longer in *Ph. mireillae* than in *Ph. saevus*). These two species are morphologically closely related to *Ph. gemetchi* and *Ph. nuri*, two rare species occurring respectively in Ethiopia and Pakistan, that we could not include in the present study.
2. Males of *Ph. caucasicus* and *Ph. mongolensis* are closely related but recognizable by the shape of gonostyle or of basal lobe of the gonocoxite. To the contrary, females are undistinguishable and share mitochondria [65], a putative introgression that supports the close relationships observed in our topologies. *Ph. andrejevi*, another species that could not be included in our analysis certainly belongs to this clade as females of *Ph. andrejevi* are impossible to discriminate from *Ph. caucasicus* and *Ph. mongolensis*.
3. Males of *Ph. similis* and *Ph. jacusieli* share many characters (shape of the parameral sheath or the flagellomere lengths) but differ by the longer process carrying the most distal spine of the gonostyle in *Ph. jacusieli* and the thick pharyngeal teeth of the female (occupying the whole pharynx in *Ph. similis* versus restricted to the central and anterior part of the armature in *Ph. jacusieli*).



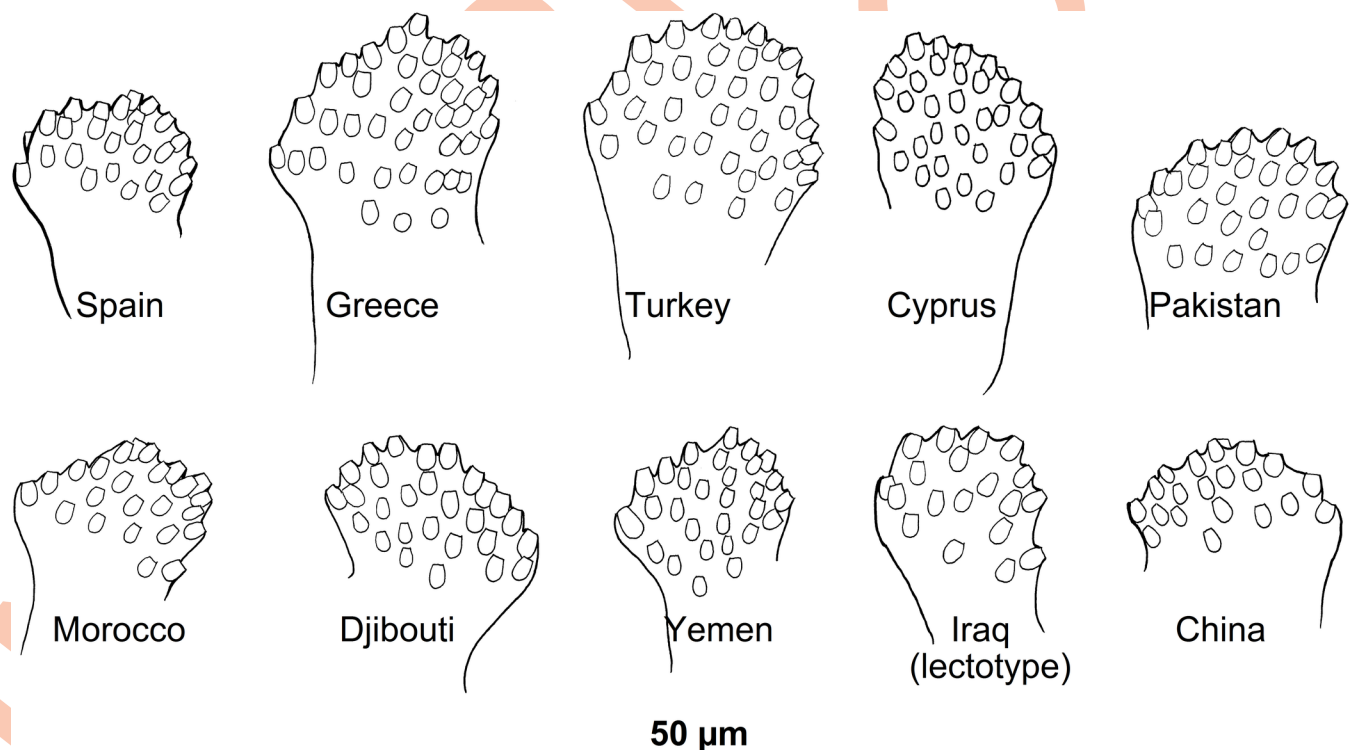
**Fig 4. Comparison of *Ph. alexandri*, the type-species of *Artemievus* subg. nov. and *Ph. sergenti*, the type-species of the subgenus *Paraphlebotomus*.** *Ph. alexandri*: genitalia of the lectotype from Amara, Iran (A), female from Syria: flagellomeres 1, 2, 3 (B), pharynx (C) and spermatheca (D). *Ph. sergenti* syntypes from Mac Mahon, Algeria: male genitalia (E) and detail of the basal lobe of the coxite (F); female pharynx (G), flagellomeres 1, 2, 3 (H) and spermathecae (I). Drawings by J. Depaquit.

<https://doi.org/10.1371/journal.pntd.0009479.g004>

4. Males of *Ph. sergenti* and *Ph. kazeruni* share a short and thin curved basal lobe of the coxite and similar gonostyle, whereas females share a similar pharyngeal armature. Females of *Ph. kazeruni* differs from those of *Ph. sergenti* by the lowest number of spermathecal rings (only one or two). *Ph. sergenti* + *Ph. kazeruni* cluster with *Ph. similis* + *Ph. jacusieli* in a strongly supported clade, in agreement with their morphological affinities (shape of the parameral sheath, flagellomeres lengths, size of the basal lobe of the gonocoxite).

### Evolutionary history and biogeography of *Paraphlebotomus sensu nov*

From the Mid-Miocene Climate Optimum (16.4–14.7 Ma) to present time, the Mediterranean region and the adjacent Middle Eastern areas experienced multiple geologic and climatic events that have strongly affected the distribution of organisms. Our analyses indicate that the most ancient ancestor of present *Paraphlebotomus* species originated ca 12.9 Ma at a time of global cooling and strong environmental changes (Fig 2). DEC and DIVALIKE models indicate that the range of the MRCA is ambiguous and suggest that the most likely range is composed of two disjoint areas: West-Mediterranean + Irano-Turania, which seems implausible (Fig 3 and S1 Text). This result can be partly explained by the fact that i) some lineages were lost; ii) distribution data are still patchy; iii) biogeographic studies for the subgenera *Phlebotomus* and *Synphlebotomus* are lacking and it was not possible to enable estimation of ancestral



**Fig 5. Morphological variability of the basal lobe of the gonocoxite of *Ph. alexandri* specimens collected in different localities.** Drawings by J. Depaquit.

<https://doi.org/10.1371/journal.pntd.0009479.g005>



range for these two subgenera in this study because very few species were included. Two different hypotheses could be considered for the MRCA of *Paraphlebotomus*: a restricted distribution or a large distribution (like *Ph. sergenti* or other phlebotomine sandflies). DEC and DIVALIKE appear to favor a largely distributed MRCA which is compatible with paleogeographical and paleoclimatic events.

During the Early and Middle Miocene, closure of the gateway of Proto-Mediterranean Sea [66] followed by the Messinian Salinity Crisis [67] and coupled with the uplift of the Atlas and Zagros Mountains [68], have strongly affected climatic conditions in the area under consideration. These mountain ranges, as well as the uplift of the Qinghai-Tibetan Plateau on the East, have acted as topographic barriers to moisture and have largely contributed to the aridification of the area. As a consequence, the vegetation of the area dominated by forests in the Early Miocene, underwent a transition to open grassland vegetation towards the late Miocene. Grasslands may have favored the expansion of the most ancient common ancestor of present *Paraphlebotomus*, enabling it to range from the Maghreb mountains to central Asia.

The expansion of the Sahara as an arid belt has long been dated back to the Quaternary Ice age (ca 2.5 Ma) [69] or to the Pliocene (5.33–2.58 Ma) [70,71]. Recent studies [72,73] have demonstrated that arid desert conditions existed repetitively since 7 Ma (Late Miocene). Zhang et al [74] used climate model simulations and confirmed that the Late Miocene (ca 7–11 Ma) has been the pivotal period for triggering north African aridity and creating the Sahara Desert. An early origin of the Sahara Desert is also suggested by the work of Böhme et al [75] who dated a saline aeolian dust deposit from Sahara in Greece back to 7.37–7.11 Ma. Therefore, the basal split between the Maghreb clade (*Ph. riouxi* + *Ph. chabaudi*) and other *Paraphlebotomus* species, ca 8.5 Myr ago, may have resulted from the rise of the Saharan belt and the aridification of the region located between the Atlas Mountains and the Turanian subregion.

During the Messinian, the Mediterranean region and the Middle East underwent rapid and dramatic climatic changes. Multiple factors probably operated simultaneously, the expansion of Antarctic ice-sheet [76], the rise of Sahara [77], the closure of the connection between the Atlantic Ocean and the Mediterranean sea that led to the nearly complete desiccation of the Mediterranean sea (Messinian salinity crisis) [67]. Consequently, steppic vegetation [78] and desert areas [79] expanded in the sub-desert belt south of Mediterranean region and extending to Middle and Central Asia. Therefore, the disjunct occurrence of the ancestors of the four lineages present in the East Mediterranean (*Ph. similis* + *Ph. jacusieli*), Saharo-Arabian (*Ph. sergenti* + *Ph. kazeruni*), Somalian (*Ph. mireillae* + *Ph. saevus*, possibly also *Ph. gemetchi*) and Turanian subregions (*Ph. caucasicus* + *Ph. mongolensis*, possibly also *Ph. nuri* and *Ph. andrejevi*) is probably the result of a Messinian vicariant event in which the extreme prolonged drought has been instrumental.

Our biogeographical analyses dated the last episode of speciation within the subgenus *Paraphlebotomus* during the Pleistocene (between 3 and 2 Ma). At that time, the Middle East areas and the East African subregion have experienced cycles of wet periods followed by period of strong drought [80]. Interestingly, these cycles have also favored the diversification of the rodents, in the burrows of which *Paraphlebotomus* larvae develop [81–83]. In arid natural areas, the presence of sand fly species is strongly limited by the low moisture of the soil [84]. *Paraphlebotomus* species therefore occur close to sites where larvae can develop. Rodent burrows maintain greater humidity and *Paraphlebotomus* larvae can develop upon organic matter (feces) that accumulate in the moist atmosphere of the deepest burrows [84]. Burrows also provide good shelter to adults during the warmer hours of the day. This coupled to the relatively low dispersal abilities of *Phlebotomus* species (less than 2 kilometers, [85]) and the relative scarceness of suitable larval habitats in the dryer and warmest areas, make sand fly population

scanty and prone to local extinction in desert and sub-desert habitats. Therefore, cycles of dry-wet periods may have resulted in reduction and separation of sand fly populations and finally by the formation of the current species. Subsequently, a few species have experienced large expansion in the last million years and have evolved toward anthropophily (i.e. *Ph. sergenti*).

### Putative species complexes that required further studies

The widely distributed species *Ph. alexandri* and *Ph. sergenti* exhibit high intra-specific divergences, which suggests that these species may represent complexes of closely related entities. Two subclades are recovered in *Ph. alexandri* (Figs 2 and S2). The genetic distance between these subclades (1.6% on 1,024,190 nt) is similar to the distance observed between other pairs of sister species present in our topology (e.g. 1.6% between *Ph. similis* and *Ph. jacusieli* and 1.7% between *Ph. chabaudi* and *Ph. riouxi*). *Ph. alexandri* exhibits a high variability in the number of setae on the basal lobe of the gonocoxite and this variability appears to be linked to the geographical origin of the studied specimens (Fig 5). In the same manner, the genetic distance observed between individuals of *Ph. sergenti* from the West Mediterranean area (Portugal, Algeria, Morocco) and those from Egypt and Afghanistan are respectively of 1.1% and 1.2%. The divergence observed within *Ph. sergenti* is slightly lower than the one observed in *Ph. alexandri* but however remains substantial. Such a high intraspecific divergence within these two species has already been observed [86–88] and deserves further studies to clarify the exact status of the subclades observed within *Ph. alexandri* and *Ph. sergenti*. This is of prime importance as it may have strong consequences on the epidemiology of *Leishmania infantum*, *L. major* and to a lesser extent of *L. donovani*. Importantly, our results highlight the power of nuclear pangenomic markers over mitochondrial genes to analyse genetic differentiation and phylogeography of phlebotomine sand flies. Indeed, several species-complexes of *Phlebotomus* may have experienced mitochondrial introgression and share haplotypes between species, even when nuclear gene flow is restricted [65,89,90]. In these complex, frequent mito-nuclear discordance due to introgression, incomplete lineage sorting, endosymbionts or mis-identification renders species delimitation difficult. This is probably where reliable markers for identification are mostly wanted but this is also where barcoding approaches (based on *COI* or *Cytb* genes) may be deficient and have to be used cautiously.

### RADseq to improve our understanding of the epidemiology of arthropod borne diseases

RAD sequencing is now widely used for population genetics and phylogenetics. Notably, the method has been shown effective to decipher species complexes or to detect introgression [91]. However, it is still rarely used to study minute insects due to the limiting DNA amount that can be extracted. Here, we confirm that multiple displacement amplification method can be used to generate unbiased copies of genomes to make possible the sequencing of RAD markers from tiny specimens [34]. Coupled with continuous progress in high-throughput sequencing technology, ever larger numbers of samples can be processed at the same time which make possible large-scale studies. For instance, with the *PstI* enzyme used here, about 400 individuals of phlebotomines could be sequenced on a S1 NovaSeq flow cell.

Several studies have shown that members of phlebotomine species complexes can display contrasted differences in terms of biology and vector capacity [92]. Therefore, deciphering species-complexes with such a high-throughput and powerful method may contribute to better identify biological entities of importance for the transmission of *Leishmania*. Further, RADseq characterization of species complexes may help to i) better delimit the geographical distribution of the different species; ii) identify phenotypic traits relevant to the disease epidemiology

and iii) analyse intraspecific population structure within entities and genetic exchanges, if any, between them.

Once thoroughly delimited molecularly, the specific level of these entities could be validated by crossbreeding experiments and/or other analyses focusing on chemical and behavioral characters (partner courtship, attraction by sex pheromones, etc). Subsequently, species need to be characterized morphologically, if possible, and named. The accurate species delimitation done using WGA and RAD markers, coupled with non-destructive extraction (see [93]), may facilitate the identification of diagnostic morphological characters. Finally, “barcoding-like” markers that enable accurate molecular identification of the species within analysed complexes could then be developed to enable fast and massive identification to survey large areas. The WGA+RAD approach developed here, that can be used on single specimen, may therefore represent a powerful tool to improve our understanding of the epidemiology of leishmaniasis and possibly also other arthropod borne human diseases.

## Supporting information

### **S1 Text. Input data and results of the ancestral range estimation using BioGeoBears.**

Matrix of geographic ranges and map of geographic regions  
(DOCX)

### **S2 Text. Input data and results of the ancestral range estimation using BioGeoBears.** Dispersal multipliers matrices and Time Periods

(DOCX)

### **S3 Text. Input data and results of the ancestral range estimation using BioGeoBears.**

Results of model testing  
(DOCX)

### **S4 Text. Input data and results of the ancestral range estimation using BioGeoBears.**

Ancestral range estimations for all tested models as inferred by BioGeoBears  
(DOCX)

### **S1 Fig. Phylogenetic tree obtained for the COI data set**

(PNG)

### **S2 Fig. Phylogenetic tree obtained for the RAD data sets.** cstacks n = 6; radis\_nsample\_min = 50%; 5,050 loci; RAxML tree (unpartitioned data set); Bootstrap values at nodes (100 replicates); %missing = % of missing RAD tags. Data sets are described in [Table 2](#).

(BMP)

### **S3 Fig. Phylogenetic tree obtained for the RAD data sets.** cstacks n = 6; radis\_nsample\_min = 75%; 1,423 loci; RAxML tree (unpartitioned data set); Bootstrap values at nodes (100 replicates); %missing = % of missing RAD tags. Data sets are described in [Table 2](#).

(BMP)

### **S4 Fig. Phylogenetic tree obtained for the RAD data sets.** cstacks n = 8; radis\_nsample\_min = 50%; 7,194 loci; RAxML tree (unpartitioned data set); Bootstrap values at nodes (100 replicates); %missing = % of missing RAD tags. Data sets are described in [Table 2](#).

(BMP)

### **S5 Fig. Phylogenetic tree obtained for the RAD data sets.** cstacks n = 8; radis\_nsample\_min = 75%; 2,086 loci; RAxML tree (unpartitioned data set); Bootstrap values at nodes (100 replicates); %missing = % of missing RAD tags. Data sets are described in [Table 2](#).

(BMP)

**S6 Fig. Phylogenetic tree obtained for the RAD data sets.** cstacks n = 10; radis\_nsamples\_min = 50%; 8,906 loci; RAxML tree (unpartitioned data set); Bootstrap values at nodes (100 replicates); %missing = % of missing RAD tags Data sets are described in [Table 2](#). (BMP)

**S7 Fig. Phylogenetic tree obtained for the RAD data sets.** cstacks n = 10; radis\_nsamples\_min = 75%; 2,808 loci; RAxML tree (unpartitioned data set); Bootstrap values at nodes (100 replicates); %missing = % of missing RAD tags Data sets are described in [Table 2](#). (BMP)

**S1 Table. Information about samples used in this study and results of the RAD-seq experiment.** (XLSX)

## Acknowledgments

The authors thank Carlos Alves-Pires, Azzedine Bounamous, Vit Dvořák, Nabil Haddad, Hanafi Hanafi, Najoua Haouas, Vladimir Ivovic, Kaouther Jaouadi, Mireille and Bob<sup>(†)</sup> Killick-Kendrick, Andreas Krüger, Nicole Léger, Michele Maroli, Laor Oshan, Yusuf Özbel, Jean-Antoine Rioux<sup>(†)</sup>, and Petr Volf for providing specimens and/or for fruitful discussions about the distribution of *Ph. similis* and *Ph. sergenti* in Turkey and in the Balkans. We thank Zoe Adams (NHM, London) and Zoubir Harrat (Pasteur Institute of Algeria) who arranged the loans of type-specimens of *Ph. alexandri* and *Ph. sergenti*, respectively. We also are grateful to Morad Belbagra for the technical assistance in data processing, to Laure Sauné for advices during library preparation, to Montpellier GenomiX (MGX, France) for sequencing and to the genotoul bioinformatics platform Toulouse Occitanie (Bioinfo Genotoul, doi: [10.15454/1.5572369328961167E12](https://doi.org/10.15454/1.5572369328961167E12)) for providing computing resources.

## Author Contributions

**Conceptualization:** Jean-Yves Rasplus, Jérôme Depaquit.

**Data curation:** Astrid Cruaud, Véronique Lehrter, Guenaëlle Genson.

**Formal analysis:** Astrid Cruaud, Véronique Lehrter, Guenaëlle Genson, Jean-Yves Rasplus, Jérôme Depaquit.

**Funding acquisition:** Jean-Yves Rasplus, Jérôme Depaquit.

**Investigation:** Astrid Cruaud, Véronique Lehrter, Guenaëlle Genson, Jean-Yves Rasplus.

**Methodology:** Astrid Cruaud, Véronique Lehrter, Guenaëlle Genson, Jean-Yves Rasplus, Jérôme Depaquit.

**Resources:** Jérôme Depaquit.

**Supervision:** Astrid Cruaud, Véronique Lehrter, Jean-Yves Rasplus, Jérôme Depaquit.

**Validation:** Astrid Cruaud, Véronique Lehrter, Guenaëlle Genson, Jean-Yves Rasplus, Jérôme Depaquit.

**Visualization:** Astrid Cruaud, Véronique Lehrter, Guenaëlle Genson, Jean-Yves Rasplus, Jérôme Depaquit.

**Writing – original draft:** Astrid Cruaud, Jean-Yves Rasplus.

**Writing – review & editing:** Astrid Cruaud, Véronique Lehrter, Guenaëlle Genson, Jérôme Depaquit.

## References

- Killick-Kendrick R. Phlebotomine vectors of the leishmaniasis: a review. *Med Vet Entomol.* 1990; 4(1):1–24. <https://doi.org/10.1111/j.1365-2915.1990.tb00255.x> PMID: 2132963
- Ready PD. Biology of Phlebotomine sand flies as vectors of disease agents. In: Berenbaum MR, editor. *Annual Review of Entomology*, Vol 58. Annual Review of Entomology. 58:2013. p. 227–50.
- Alvar J, Velez ID, Bern C, Herrero M, Desjeux P, Cano J, et al. Leishmaniasis worldwide and global estimates of its incidence. *PLoS One.* 2012; 7(5). <https://doi.org/10.1371/journal.pone.0035671> WOS:000305338500009. PMID: 22693548
- Burza S, Croft SL, Boelaert M. Leishmaniasis. *Lancet.* 2018; 392(10151):951–70. [https://doi.org/10.1016/S0140-6736\(18\)31204-2](https://doi.org/10.1016/S0140-6736(18)31204-2) WOS:000444475700028. PMID: 30126638
- Bates PA, Depaquit J, Galati EAB, Kamhawi S, Maroli M, McDowell MA, et al. Recent advances in phlebotomine sand fly research related to leishmaniasis control. *Parasites Vectors.* 2015; 8. <https://doi.org/10.1186/s13071-015-0712-x> WOS:000351040200001. PMID: 25885217
- Maia C, Depaquit J. Can *Sergentomyia* (Diptera, Psychodidae) play a role in the transmission of mammal-infecting *Leishmania*? *Parasite.* 2016; 23. <https://doi.org/10.1051/parasite/2016062> WOS:000390430800001. PMID: 27921993
- Lewis DJ. A taxonomic review of the genus *Phlebotomus* (Diptera: Psychodidae). *Bulletin of the British Museum (Natural History) Entomology.* 1982; 45:121–209.
- Galati EAB, Galvis-Ovallos F, Lawyer P, Leger N, Depaquit J. An illustrated guide for characters and terminology used in descriptions of Phlebotominae (Diptera, Psychodidae). *Parasite.* 2017; 24. <https://doi.org/10.1051/parasite/2017027> WOS:000406464900001. PMID: 28730992
- Rispail P, Léger N. Numerical taxonomy of Old World Phlebotominae (Diptera: Psychodidae). 1. Considerations of morphological characters in the genus *Phlebotomus* Rondani & Berte 1840. *Mem Inst Oswaldo Cruz.* 1998; 93(6):773–85. <https://doi.org/10.1590/s0074-02761998000600015> WOS:000076958300015. PMID: 9921302
- Rispail P, Léger N. Numerical taxonomy of Old World Phlebotominae (Diptera: Psychodidae). 2. Restatement of classification upon subgeneric morphological characters. *Mem Inst Oswaldo Cruz.* 1998; 93(6):787–93. <https://doi.org/10.1590/s0074-02761998000600016> WOS:000076958300016. PMID: 9921303
- Aransay AM, Scoulica E, Tselentis Y, Ready PD. Phylogenetic relationships of phlebotomine sandflies inferred from small subunit nuclear ribosomal DNA. *Insect Mol Biol.* 2000; 9(2):157–68. <https://doi.org/10.1046/j.1365-2583.2000.00168.x> WOS:000086700200005. PMID: 10762423
- Depaquit J, Perrotey S, Lecointre G, Tillier A, Tillier S, Ferte H, et al. Molecular systematics of Phlebotominae: a pilot study. Paraphyly of the genus *Phlebotomus*. *Comptes Rendus De L Academie Des Sciences Serie Iii-Sciences De La Vie-Life Sciences.* 1998; 321(10):849–55. ISI:000077073900006. [https://doi.org/10.1016/s0764-4469\(99\)80025-0](https://doi.org/10.1016/s0764-4469(99)80025-0) PMID: 9835021
- Al-Zahrani M, Peters W, Evans D, Chin C, Smith V, Lane R. *Phlebotomus sergenti*, a vector of *Leishmania tropica* in Saudi Arabia. *Trans Roy Soc Trop Med Hyg.* 1988; 83(3):416. [https://doi.org/10.1016/0035-9203\(88\)90142-3](https://doi.org/10.1016/0035-9203(88)90142-3) PMID: 3232175
- Guilvard E, Rioux JA, Gallego M, Pralong F, Mahjour J, Martinezortega E, et al. *Leishmania tropica* in Morocco. III. Identification of 89 isolates from the vector *Phlebotomus sergenti*. *Annales de Parasitologie Humaine et Comparée.* 1991; 66(3):96–9. <https://doi.org/10.1051/parasite/199166396> WOS:A1991GT41500002. PMID: 1776784
- Johnson RN, Ngumbi PM, Robert LL, Anjili CO, Killickkendrick R, Meredith SEO. Phlebotomine sand flies of Kenya (Diptera, Psychodidae). 2. *Phlebotomus aculeatus* as a probable vector of *Leishmania tropica* s.l. *Ann Trop Med Parasitol.* 1993; 87(5):541–3. <https://doi.org/10.1080/00034983.1993.11812807> WOS:A1993MM18100015. PMID: 8311581
- Lawyer P, Mbrahtu Y, Ngumbi PM, Mwanyumba P, Mbugua J, Kiilu G, et al. *Phlebotomus guggisbergi* (Diptera: Psychodidae), a vector of *Leishmania tropica* in Kenya. *Am J Trop Med Hyg.* 1991; 44:290–8. <https://doi.org/10.4269/ajtmh.1991.44.290> PMID: 2035750
- Svobodova M, Votypka J, Peckova J, Dvorak V, Nasereddin A, Baneth G, et al. Distinct transmission cycles of *Leishmania tropica* in 2 adjacent foci, northern Israel. *Emerg Infect Dis.* 2006; 12(12):1860–8. WOS:000242301900009. <https://doi.org/10.3201/eid1212.060497> PMID: 17326936
- Depaquit J, Ferte H, Leger N, Killick-Kendrick R, Rioux JA, Killick-Kendrick M, et al. Molecular systematics of the Phlebotomine sandflies of the subgenus *Paraphlebotomus* (Diptera, Psychodidae),

- Phlebotomus*) based on ITS2 rDNA sequences. Hypotheses of dispersion and speciation. *Insect Mol Biol.* 2000; 9(3):293–300. <https://doi.org/10.1046/j.1365-2583.2000.00179.x> WOS:000088167600009. PMID: 10886413
19. Kruger A, Struven L, Post RJ, Faulde M. The sandflies (Diptera: Psychodidae, Phlebotominae) in military camps in northern Afghanistan (2007–2009), as identified by morphology and DNA 'barcoding'. *Annals of Tropical Medicine and Parasitology.* 2011; 105(2):163–76. <https://doi.org/10.1179/136485911X12899838683241> WOS:000288593200007. PMID: 21396252
  20. Grace-Lema DM, Yared S, Quitadamo A, Janies DA, Wheeler WC, Balkew M, et al. A phylogeny of sand flies (Diptera: Psychodidae: Phlebotominae), using recent Ethiopian collections and a broad selection of publicly available DNA sequence data. *Systematic Entomology.* 2015; 40:733–44. <https://doi.org/10.1111/syen.12135>
  21. Cunze S, Kochmann J, Koch LK, Hasselmann KJQ, Klimpel S. Leishmaniasis in Eurasia and Africa: geographical distribution of vector species and pathogens. *Royal Society Open Science.* 2019; 6(5). <https://doi.org/10.1098/rsos.190334> WOS:000470127300055. PMID: 31218068
  22. Gebre-Michael T, Balkew M. On the occurrence of *Phlebotomus* (*Paraphlebotomus*) *mireillae* (Diptera: Psychodidae) in Ethiopia. *Ann Trop Med Parasitol.* 2002; 96(4):421–5. <https://doi.org/10.1179/000349802125001177> WOS:000176685300011. PMID: 12171625
  23. Gebre-Michael T, Balkew M. *Phlebotomus* (*Paraphlebotomus*) *gemetchi* (Diptera: Psychodidae), a new sand fly species from Ethiopia. *J Med Entomol.* 2003; 40(2):141–5. <https://doi.org/10.1603/0022-2585-40.2.141> WOS:000181992300003. PMID: 12693840
  24. Kavrur H, Arikan H, Ozbel Y. *Phlebotomus halepensis* (Diptera: Psychodidae) vectorial capacity in Afyon and Nigde Province, Turkey. *J Med Entomol.* 2018; 55(2):317–22. <https://doi.org/10.1093/jme/tjx210> WOS:000429320000008. PMID: 29202202
  25. Moin-Vaziri V, Depaquit J, Yaghoobi-Ershadi MR, Oshaghi MA, Derakhshandeh-Peykar P, Ferte H, et al. Intraspecific variation within *Phlebotomus sergenti* Parrot (1917) (Diptera: Psychodidae) based on mtDNA sequences in Islamic Republic of Iran. *Acta Trop.* 2007; 102(1):29–37. ISI:000247774400004. <https://doi.org/10.1016/j.actatropica.2007.02.012> PMID: 17451632
  26. Koch LK, Kochmann J, Klimpel S, Cunze S. Modeling the climatic suitability of leishmaniasis vector species in Europe. *Sci Rep.* 2017; 7. <https://doi.org/10.1038/s41598-017-13822-1> WOS:000413084800022. PMID: 29042642
  27. Akhoundi M, Baghaei A, Depaquit J, Parvizi P. Molecular characterization of *Leishmania* infection from naturally infected sand flies Caught in a focus of cutaneous leishmaniasis (Eastern Iran). *J Arthropod-Borne Dis.* 2013; 7(2):122–31. WOS:000326789700003. PMID: 24409437
  28. Guan LR, Zhou ZB, Jin CF, Fu Q, Chai JJ. Phlebotomine sand flies (Diptera: Psychodidae) transmitting visceral leishmaniasis and their geographical distribution in China: a review. *Infectious Diseases of Poverty.* 2016; 5. <https://doi.org/10.1186/s40249-016-0107-z> WOS:000370780300001. PMID: 26906187
  29. Toprak S, Ozer N. Sand fly species of Sanliurfa province in Turkey. *Med Vet Entomol.* 2005; 19(1):107–10. <https://doi.org/10.1111/j.0269-283X.2005.00545.x> WOS:000227390100015. PMID: 15752185
  30. Benallal K, Gassen B, Bouiba L, Depaquit J, Harrat Z. Entomological investigation following the resurgence of human visceral leishmaniasis in southern Algeria. *Acta Trop.* 2013; 128(3):518–21. <https://doi.org/10.1016/j.actatropica.2013.07.017> WOS:000327824800010. PMID: 23920502
  31. Ali N, Ullah A, Wahid S, Khisroon M, Rasheed SB. Distribution, species composition and relative abundances of sandflies in North Waziristan Agency, Pakistan. *Med Vet Entomol.* 2016; 30(1):89–94. <https://doi.org/10.1111/mve.12141> WOS:000369874600012. PMID: 26582763
  32. Kato H, Uezato H, Sato H, Bhutto AM, Soomro FR, Baloch JH, et al. Natural infection of the sand fly *Phlebotomus kazeruni* by *Trypanosoma* species in Pakistan. *Parasites & Vectors.* 2010; 3. <https://doi.org/10.1186/1756-3305-3-10> WOS:000276213000001. PMID: 20184773
  33. Blair C, Campbell CR, Yoder AD. Assessing the utility of whole genome amplified DNA for next-generation molecular ecology. *Mol Ecol Resour.* 2015; 15(5):1079–90. <https://doi.org/10.1111/1755-0998.12376> WOS:000359631600007. PMID: 25619406
  34. Cruaud A, Groussier G, Genson G, Polaszek A, Sauné L, Rasplus J-Y. Pushing the limits of whole genome amplification: successful sequencing of RADseq library from a single microhymenopteran (Chalcidoidea, Trichogramma). *PeerJ.* 2018; 5640. <https://doi.org/10.7717/peerj.5640> PMID: 30356952
  35. Jasmine F, Ahsan H, Andrusis IL, John EM, Chang-Claude J, Kibriya MG. Whole-Genome Amplification enables accurate genotyping for microarray-based high-density single nucleotide polymorphism array. *Cancer Epidemiol Biomarkers Prev.* 2008; 17(12):3499–508. <https://doi.org/10.1158/1055-9965.EPI-08-0482> WOS:000261724000030. PMID: 19064567

36. Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, et al. Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE*. 2008; 3(10). ARTN e3376 <https://doi.org/10.1371/journal.pone.0003376> WOS:000265121600002. PMID: 18852878
37. Winnepeninckx B, Backeljau T, Dewachter R. Extraction of high-molecular-weight dna from mollusks. *Trends Genet*. 1993; 9(12):407–. WOS:A1993MK28000003. [https://doi.org/10.1016/0168-9525\(93\)90102-n](https://doi.org/10.1016/0168-9525(93)90102-n) PMID: 8122306
38. Germain JF, Chatot C, Meusnier I, Rasplus J-Y, Cruaud A. Combining DNA barcoding and morphological analysis to identify *Epitrix* potato flea beetles (Coleoptera: Chrysomelidae). *Bulletin of Entomological Research*. 2013; 103:354–62. <https://doi.org/10.1017/S000748531200079X> PMID: 23448201
39. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 2012; 28:1647–9. <https://doi.org/10.1093/bioinformatics/bts199> PMID: 22543367
40. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*. 2013; 30:772–80. <https://doi.org/10.1093/molbev/mst010> PMID: 23329690
41. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets *Molecular Biology and Evolution*. 2016; 33:1870–4. <https://doi.org/10.1093/molbev/msw054> PMID: 27004904
42. Ratnasingham S, Hebert PDN. BOLD: The Barcode of Life Data System ([www.barcodinglife.org](http://www.barcodinglife.org)). *Mol Ecol Notes*. 2007; 7(3):355–64. ISI:000246244300001. <https://doi.org/10.1111/j.1471-8286.2007.01678.x> PMID: 18784790
43. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Molecular Biology and Evolution*. 2015; 32:268–74. <https://doi.org/10.1093/molbev/msu300> PMID: 25371430
44. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*. 2017; 14:587–9. <https://doi.org/10.1038/nmeth.4285> PMID: 28481363
45. Minh BQ, Nguyen MAT, von Haeseler A. Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution*. 2013; 30:1188–95. <https://doi.org/10.1093/molbev/mst024> PMID: 23418397
46. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*. 2010; 59:307–21. <https://doi.org/10.1093/sysbio/syq010> PMID: 20525638
47. Etter PD, Bassham S, Hohenlohe PA, Johnson EA, Cresko WA. SNP discovery and genotyping for evolutionary genetics using RAD sequencing. In: Orgogozo V, Rockman MV, editors. *Molecular methods for evolutionary genetics*. 772. Humana Press, New York.; 2011. p. 157–78.
48. Cruaud A, Gautier M, Galan M, Foucaud J, Sauné L, Genson G, et al. Empirical assessment of RAD sequencing for interspecific phylogeny. *Molecular Biology and Evolution*. 2014; 31:1272–4. <https://doi.org/10.1093/molbev/msu063> PMID: 24497030
49. Cruaud A, Gautier M, Rossi J-P, Rasplus J-Y, Gouzy J. RADis: Analysis of RAD-seq data for InterSpecific phylogeny. *Bioinformatics*. 2016; 32(19):3027–8. <https://doi.org/10.1093/bioinformatics/btw352> PMID: 27312412
50. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. Stacks: an analysis tool set for population genomics. *Mol Ecol*. 2013; 22(11):3124–40. <https://doi.org/10.1111/mec.12354> WOS:000319675300022. PMID: 23701397
51. Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. Stacks: Building and genotyping loci de novo from short-read sequences. *G3 (Bethesda)*. 2011; 1(3):171–82. <https://doi.org/10.1534/g3.111.000240> WOS:000312406900001. PMID: 22384329
52. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014; 30:1312–3. <https://doi.org/10.1093/bioinformatics/btu033> PMID: 24451623
53. Stöver BC, Müller KF. TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics*. 2010; 11:7. <https://doi.org/10.1186/1471-2105-11-7> PMID: 20051126
54. Depaquit J. Révision du sous-genre *Paraphlebotomus* (*Phlebotomus*—Phlebotominae—Psychodidae—Diptera). Approches morphologique et moléculaire. (Thèse d'Université). Reims1997.
55. Papadopoulou A, Anastasiou I, Vogler AP. Revisiting the insect mitochondrial molecular clock: The Mid-Aegean trench calibration. *Molecular Biology and Evolution*. 2010; 27:1659–72. <https://doi.org/10.1093/molbev/msq051> PMID: 20167609
56. Godefroid M, Meseguer A, Sauné L, Genson G, Streito J-C, Rossi J-P, et al. Restriction-site associated DNA markers provide new insights into the evolutionary history of the bark beetle genus

- Dendroctonus* Mol Phylogenet Evol. 2019; 139:106528. <https://doi.org/10.1016/j.ympev.2019.106528> PMID: 31176966
57. Tamura K, Tao Q, Kumar S, Russo C. Theoretical foundation of the RelTime method for estimating divergence times from variable evolutionary rates. *Molecular Biology and Evolution*. 2018; 35:1770–82. <https://doi.org/10.1093/molbev/msy044> PMID: 29893954
  58. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution*. 2018; 35:1547–9. <https://doi.org/10.1093/molbev/msy096> PMID: 29722887
  59. Matzke NJ. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography*. 2013; 5(4):242–8.
  60. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*. 2018; 35:526–8.
  61. Ree RH, Sanmartin I. Conceptual and statistical problems with the DEC+J model of founder-event speciation and its comparison with DEC via model selection. *Journal of Biogeography* 2018; 45:741–9.
  62. Tabbabi A, Ghrab J, Aoun K, Ready PD, Bouratbine A. Habitats of the sandfly vectors of *Leishmania tropica* and *L. major* in a mixed focus of cutaneous leishmaniasis in southeast Tunisia. *Acta Trop*. 2011; 119(2–3):131–7. <https://doi.org/10.1016/j.actatropica.2011.05.002> WOS:000293717200011. PMID: 21605538
  63. Tabbabi A, Rhim A, Ghrab J, Martin O, Aoun K, Bouratbine A, et al. *Phlebotomus (Paraphlebotomus) riouxii*: a synonym of *Phlebotomus chabaudi* without any proven vectorial role in Tunisia and Algeria. *Med Vet Entomol*. 2014; 28:51–9. <https://doi.org/10.1111/mve.12067> WOS:000341626200007. PMID: 25171607
  64. Lehrter V, Banuls AL, Léger N, Rioux JA, Depaquit J. *Phlebotomus (Paraphlebotomus) chabaudi* and *Phlebotomus riouxii*: closely related species or synonyms? *Parasite*. 2017; 24. <https://doi.org/10.1051/parasite/2017050> WOS:000416861000001. PMID: 29194032
  65. Parvizi P, Taherkhani H, Ready PD. *Phlebotomus caucasicus* and *Phlebotomus mongolensis* (Diptera: Psychodidae): indistinguishable by the mitochondrial cytochrome b gene in Iran. *Bull Entomol Res*. 2010; 100(4):415–20. <https://doi.org/10.1017/S0007485309990423> ISI:000280020700005. PMID: 19939318
  66. Torfstein A, Steinberg J. The Oligo–Miocene closure of the Tethys Ocean and evolution of the proto-Mediterranean Sea. *Sci Rep*. 2020; 10(1):13817. <https://doi.org/10.1038/s41598-020-70652-4> PMID: 32796882
  67. Krijgsman W, Hilgen FJ, Raffi I, Sierró FJ, Wilson DS. Chronology, causes and progression of the Messinian salinity crisis. *Nature*. 1999; 400(6745):652–5. <https://doi.org/10.1038/23231>
  68. Mouthereau F. Timing of uplift in the Zagros belt/Iranian plateau and accommodation of late Cenozoic Arabia–Eurasia convergence. *Geological Magazine*. 2011; 148(5–6):726–38. Epub 2011/04/18. <https://doi.org/10.1017/S0016756811000306>
  69. Kroepelin S. Revisiting the age of the Sahara desert. *Science*. 2006; 312:1138–9. <https://doi.org/10.1126/science.312.5777.1138b> PMID: 16728618
  70. Swezey CS. Revisiting the age of the Sahara Desert. *Science*. 2006; 312:1138e9. PMID: 16739253
  71. Swezey CS. Cenozoic stratigraphy of the Sahara, Northern Africa. *Journal of African Earth Sciences*. 2009; 53:89e121.
  72. Schuster M, Durringer P, Ghiennie J-F, Vignaud P, Mackaye HT, Likius A, et al. The age of the Sahara Desert. *Science*. 2006; 311:821. <https://doi.org/10.1126/science.1120161> PMID: 16469920
  73. Schuster M, Durringer P, Ghiennie J-F, Vignaud P, Mackaye HT, Likius A, et al. Revisiting the age of the Sahara Desert. *Response*. *Science*. 2006; 312:1139. <https://doi.org/10.1126/science.1125301> PMID: 16728619
  74. Zhang Z, Ramstein G, Schuster M, Li C, Contoux C, Yan Q. Aridification of the Sahara desert caused by Tethys Sea shrinkage during the Late Miocene. *Nature*. 2014; 513:401e4. <https://doi.org/10.1038/nature13705> PMID: 25230661
  75. Böhme M, Spassov N, Ebner M, Geraads D, Hristova L, Kirscher U, et al. Messinian age and savannah environment of the possible hominin *Graecopithecus* from Europe. *PLoS One*. 2017; 12:e0177347. <https://doi.org/10.1371/journal.pone.0177347> PMID: 28531204
  76. Ohneiser C, Florindo F, Stocchi P, Roberts AP, DeConto RM, Pollard D. Antarctic glacio-eustatic contributions to late Miocene Mediterranean desiccation and reflooding. *Nature Communications*. 2015; 6(1):8765. <https://doi.org/10.1038/ncomms9765> PMID: 26556503



77. Micheels A, Eronen J, Mosbrugger V. The Late Miocene climate response to a modern Sahara desert. *Global and Planetary Change*. 2009; 67(3):193–204. <https://doi.org/10.1016/j.gloplacha.2009.02.005>.
78. Popescu S-M, Suc J-P, Loutre M-F. Early Pliocene vegetation changes forced by eccentricity precession: Example from Southwestern Romania. *Palaeogeography, Palaeoclimatology, Palaeoecology*. 2006; 238:340–8.
79. Barbolini N, Woutersen A, Dupont-Nivet G, Silvestro D, Tardif D, Coster PMC, et al. Cenozoic evolution of the steppe-desert biome in Central Asia. *Science Advances*. 2020; 6(41):eabb8227. <https://doi.org/10.1126/sciadv.abb8227> PMID: 33036969
80. Rose C, Polissar PJ, Tierney JE, Filley T, deMenocal PB. Changes in northeast African hydrology and vegetation associated with Pliocene–Pleistocene sapropel cycles. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2016; 371(1698):20150243. <https://doi.org/10.1098/rstb.2015.0243> PMID: 27298473
81. Bryja J, Colangelo P, Lavrenchenko LA, Meheretu Y, Sumbera R, Bryjova A, et al. Diversity and evolution of African Grass Rats (Muridae: *Arvicanthis*). From radiation in East Africa to repeated colonization of northwestern and southeastern savannas. *Journal of Zoological Systematics and Evolutionary Research*. 2019; 57(4):970–88. <https://doi.org/10.1111/jzs.12290> WOS:000498850700018.
82. Aghova T, Palupcikova K, Sumbera R, Frynta D, Lavrenchenko LA, Meheretu Y, et al. Multiple radiations of spiny mice (Rodentia: *Acomys*) in dry open habitats of Afro-Arabia: evidence from a multi-locus phylogeny. *Bmc Evolutionary Biology*. 2019; 19. <https://doi.org/10.1186/s12862-019-1380-9> WOS:000460511700002. PMID: 30832573
83. Ndiaye A, Chevret P, Dobigny G, Granjon L. Evolutionary systematics and biogeography of the arid habitat-adapted rodent genus *Gerbillus* (Rodentia, Muridae): a mostly Plio–Pleistocene African history. *Journal of Zoological Systematics and Evolutionary Research*. 2016; 54(4):299–317. <https://doi.org/10.1111/jzs.12143> WOS:000386032500006.
84. Schlein Y, Warburg A, Schnur LF, Le Blancq SM, Gunders AE. Leishmaniasis in Israel: reservoir hosts, sandfly vectors and leishmanial strains in the Negev, Central Arava and along the Dead Sea. *Trans Roy Soc Trop Med Hyg*. 1984; 78:480–4. [https://doi.org/10.1016/0035-9203\(84\)90067-1](https://doi.org/10.1016/0035-9203(84)90067-1) PMID: 6385358
85. Orshan L, Elbaz S, Ben-Ari Y, Akad F, Afik O, Ben-Avi I, et al. Distribution and dispersal of *Phlebotomus papatasi* (Diptera: Psychodidae) in a zoonotic cutaneous Leishmaniasis focus, the Northern Negev, Israel. *Plos Neglected Tropical Diseases*. 2016; 10(7). <https://doi.org/10.1371/journal.pntd.0004819> WOS:000381017800027. PMID: 27427959
86. Ebrahimi S, Bordbar A, Parvizi P. Genetic dynamics in the sand fly (Diptera: Psychodidae) nuclear and mitochondrial genotypes: evidence for vector adaptation at the border of Iran with Iraq. *Parasites & Vectors*. 2016; 9. <https://doi.org/10.1186/s13071-016-1603-5> WOS:000377338300003. PMID: 27260204
87. Dvorak V, Votykka J, Aytakin AM, Alten B, Volf P. Intraspecific variability of natural populations of *Phlebotomus sergenti*, the main vector of *Leishmania tropica*. *J Vector Ecol*. 2011; 36:S49–S57. <https://doi.org/10.1111/j.1948-7134.2011.00111.x> WOS:000287927300006. PMID: 21366780
88. Baron S, Martin-Sanchez J, Gallego M, Morales-Yuste M, Boussaa S, Morillas-Marquez F. Intraspecific variability (rDNA ITS and mtDNA Cyt b) of *Phlebotomus sergenti* in Spain and Morocco. *Acta Trop*. 2008; 107(3):259–67. <https://doi.org/10.1016/j.actatropica.2008.07.003> WOS:000260707800008. PMID: 18703008
89. Pesson B, Ready JS, Benabdennbi I, Martin-Sanchez J, Esseghir S, Cadi-Soussi M, et al. Sandflies of the *Phlebotomus perniciosus* complex: mitochondrial introgression and a new sibling species of *P. longicuspis* in the Moroccan Rif. *Med Vet Entomol*. 2004; 18(1):25–37. WOS:000220087800005. <https://doi.org/10.1111/j.0269-283x.2004.0471.x> PMID: 15009443
90. dos Reis YV, Alevi KCC. Hybridization in Phlebotominae (Diptera: Psychodidae): A mini-review. *Infect Genet Evol*. 2020; 86. <https://doi.org/10.1016/j.meegid.2020.104593> WOS:000599880700011. PMID: 33053414
91. Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*. 2016; 17 81–92. <https://doi.org/10.1038/nrg.2015.28> PMID: 26729255
92. de Sousa-Paula LC, Pessoa FAC, Otrando D, Dantas-Forres F. Beyond taxonomy: species complexes in New World phlebotomine sand flies. *Med Vet Entomol*. 2021. <https://doi.org/10.1111/mve.12510> PMID: 33480064
93. Cruaud A, Nidelet S, Arnal P, Weber A, Fusu L, Gumovsky A, et al. Optimised DNA extraction and library preparation for minute arthropods: application to target enrichment in chalcid wasps used for biocontrol. *Mol Ecol Resour*. 2019; 19(3):702–10. <https://doi.org/10.1111/1755-0998.13006> PMID: 30758892

94. Swaminath CS, Shortt HE, Anderson LAP. Transmission of Indian kala-azar to man by the bites of *Phlebotomus argentipes*, Ann. and Brun. Indian J Med Res. 2006; 123(3):A473–A7. WOS:000238243500028.
95. Léger N, Gramiccia M, Gradoni L, Madulo-Leblond G, Pesson B, Ferté H, et al. Isolation and typing of *Leishmania infantum* from *Phlebotomus neglectus* on the island of Corfu, Greece. Trans Roy Soc Trop Med Hyg. 1988; 82:419–20. [https://doi.org/10.1016/0035-9203\(88\)90145-9](https://doi.org/10.1016/0035-9203(88)90145-9) PMID: 3232178
96. Guan L-R, Xu Y-X, Li B-S, Dong J. The role of *Phlebotomus alexandri* Sinton, 1928 in the transmission of kala-azar. Bulletin of the World Health Organization. 1986; 64:107–12. PMID: 3488133
97. Guan L-R. Current status of kala-azar and vector control in China. Bulletin of the World Health Organization. 1991; 69:595–601. PMID: 1959161
98. Neronov VM, Strelkova MV, Shurkhal AA, Lushekina AA, Artemyev MM. Natural focality of zoonotic cutaneous leishmaniasis in the Mongolian People's Republic; results and objectives of integrated research. Folia Parasitologica. 1987; 34:1–9. PMID: 3583127
99. Yaghoobi-Ershadi MR, Javadian E, Tahvildar-Ebidrini GH. The isolation of *Leishmania major* from *Phlebotomus (Paraphlebotomus) caucasicus*, in Isfahan province, Islamic Republic of Iran. Trans Roy Soc Trop Med Hyg. 1994; 88(5):518–9. [https://doi.org/10.1016/0035-9203\(94\)90142-2](https://doi.org/10.1016/0035-9203(94)90142-2) WOS: A1994PQ48700006. PMID: 7992325
100. Haouas N, Amer O, Alshammri FF, Al-Shammari S, Remadi L, Ashankyty I. Cutaneous leishmaniasis in northwestern Saudi Arabia: identification of sand fly fauna and parasites. Parasites Vectors. 2017; 10. <https://doi.org/10.1186/s13071-017-2497-6> WOS:000414331300003. PMID: 29096693
101. Gebre-Michael T, Balkew M, Ali A, Ludovisi A, Gramiccia M. The isolation of *Leishmania tropica* and *L. aethiopia* from *Phlebotomus (Paraphlebotomus)* species (Diptera: Psychodidae) in the Awash Valley, northeastern Ethiopia. Trans Roy Soc Trop Med Hyg. 2004; 98(1):64–70. [https://doi.org/10.1016/s0035-9203\(03\)00008-7](https://doi.org/10.1016/s0035-9203(03)00008-7) WOS:000187423600007. PMID: 14702839
102. Dedet JP, Desjeux P, Derouin F. Écologie d'un foyer de leishmaniose cutanée dans la région de Thiès (Sénégal, Afrique de l'Ouest) 4. Infestation spontanée et biologie de *Phlebotomus duboscqi* Neveu-Lemaire, 1906. Bulletin de la Société de Pathologie Exotique. 1980; 73:266–76.
103. Killick-Kendrick R, Leaney A, Peters W, Rioux J, Bray R. Zoonotic cutaneous leishmaniasis in Saudi Arabia: the incrimination of *Phlebotomus*. Trans Roy Soc Trop Med Hyg. 1985; 79(2):252–5. [https://doi.org/10.1016/0035-9203\(85\)90350-5](https://doi.org/10.1016/0035-9203(85)90350-5) PMID: 4002297
104. Yaghoobi-Ershadi MR, Javadian E. Seasonal variation of *Leishmania major* infection rates in sandflies from rodent burrows in Isfahan province, Iran. Med Vet Entomol. 1996; 10(2):181–4. <https://doi.org/10.1111/j.1365-2915.1996.tb00726.x> WOS:A1996UL34300011. PMID: 8744712