The biogeographic patterns of the olive fly and its primary symbiont Candidatus Erwinia dacicola across the distribution area of the olive tree

scientific reports

Check for updates

OPEN

The biogeographic patterns of the olive fly and its primary symbiont *Candidatus* Erwinia dacicola across the distribution area of the olive tree

Isabel Martinez-Sañudo[1]✉, M. Alejandra Perotti[2], Ivana Carofano[1], Giacomo Santoiemma[1], Laura Marri[3] & Luca Mazzon[1]

The olive fly, *Bactrocera oleae* (Rossi, 1790), is the major insect pest of olives attacking both cultivated and wild olive. *Bactrocera oleae* carries a primary and vertically transmitted symbiont, the bacterium *Candidatus* Erwinia dacicola. As any primary symbiont, it plays an important role in the reproduction and lifespan of the fly. The genetic 16S rRNA diversity of the primary symbiont and the mitochondrial haplotype variation of the insect host were simultaneously examined in 54 olive fly populations. The aim was to unravel the biogeographic patterns of this economically relevant host-bacteria interaction across a wide distribution area. Three symbiont haplotypes were identified. The primary symbiont showed a lower haplotype diversity than that of its host, a characteristic indicative of a long-term interaction. A significant genetic and geographic association between host and primary symbiont was observed, with an East-West genetic differentiation pattern in the Mediterranean basin, coinciding with the historical genetic distribution of the olive tree. The study shows promise, informing and aiding the development of future tools for the control of the olive fly.

Keywords Primary symbiont, *Bactrocera oleae*, 16S rRNA, Tephritidae, Oesophageal bulb, *Olea europaea*

Primary or obligate symbionts can serve as genetic markers to investigate host biodiversity and distribution patterns[1–3]. Primary symbionts are vertically transmitted from mother to offspring and can accurately mirror the history of their hosts. They can act as reliable markers for past events, giving in some instances even higher resolution than their hosts [e.g., [3,4]].

The olive fly *Bactrocera oleae* (Rossi, 1790) (Diptera, Tephritidae), is considered the major insect pest of olive crops worldwide. The larvae feed exclusively on wild (*Olea europaea* L. var. *sylvestris* (Miller) Lehr. (oleaster) and cultivated olive fruits of the *Olea* genus[5], having a significant impact on the quantity and quality of olives. Annually, infested crops experience an average yield loss of up to 15%[6], and this loss could potentially reach to as much as 90% without proper control measures[7]. The olive fly distribution is primarily limited to regions where both cultivated and wild trees coexist[8]. At present, the fly is reported throughout the Mediterranean basin, the Canary Islands, Africa, and Middle East. It has also been found in Iran, China, India, Pakistan[9–11], and ultimately it was introduced in the American continent in 1998[12]. Australia seems to be the only olive-growing area worldwide still unaffected by the olive fruit fly[6]. While the fly is abundant and well-known in the Mediterranean basin, it is most likely associated with the wild olive origin. Studies on the original geographical distribution of the wild variety of *Olea europaea* indicate its presence from South Africa all the way to South Asia, including the Saharan mountains to the Mediterranean[13]. The exact location of the origin of domestication of the olive tree is highly debated. However, in the Sahara region climatic conditions changed, becoming more favorable between 11,800 and 4,900 years ago, perhaps suggesting divergence of subspecies[14] and coinciding with events of domestication. According to archaeological records, from the upper Paleolithic and during the

[1]Department of Agronomy, Food, Natural Resources, Animals and the Environment (DAFNAE), University of Padova, Viale dell'Università, 16, 35020 Legnaro, PD, Italy. [2]Ecology and Evolutionary Biology Section, School of Biological Sciences, University of Reading, Reading, GB, Great Britain. [3]Department of Life Sciences, University of Siena, Siena, Italy. ✉email: isabel.martinez@unipd.it

Holocene, human activity increased in eastern and western areas of the Mediterranean[15–17] and grafting and selection of olive cultivars seems to have originated in the Levant[17].

Based on the close relationship between the Mediterranean and the Saharan wild olive[18], alongside the decreasing genetic diversity observed from African to Mediterranean olive fly populations, some authors propose the African origin of *B. oleae* and its subsequent expansion into the Mediterranean basin[19,20].

Like other widely distributed insects, it is generally accepted that understanding structure and dynamics of *B. oleae* populations is fundamental for the design and implementation of effective monitoring and control strategies[21]. Several efforts have been made to identify the genetic structure of olive fly populations across the distribution area by using a variety of molecular markers such as microsatellites[19,22–24], mitochondrial DNA fragments [e.g., [10,11,21,25,26]] and the full mitochondrial *B. oleae* genome[27]. Despite the significant progresses that have been achieved through these studies, many questions concerning the structure of *B. oleae*'s populations across its range of distribution remain open.

The primary symbiont of the fly, *Candidatus* Erwinia dacicola, is essential for colonization of the harsh habitat of unripe olives. According to Ben-Yosef et al.[28] symbiosis in olive flies is exceptional within tephritids. The counteractive effect of the bacterium, by inhibiting toxic compounds like oleuropein makes this symbiosis a tripartite association functionally interdependent. Studies must consider the three partners altogether, the fly, the bacterium and the chemistry of olives.

The genetic structure of *Ca*. E. dacicola needs to be studied together with that of the insect host. In a previous work, we revealed the presence of just two bacterial lineages distributed among Italian olive fly populations[29]. Subsequently, these two symbiont haplotypes were also detected in four Mediterranean populations, and two additional haplotypes were recently found[26]. The values of nucleotide substitution rates can give an estimate of the age of the symbiotic association[30]. The key is to explore the genetic variability of the primary symbiont in synchrony with that of the host, taking into consideration that human selection of domesticated cultivars has likely favored *B. oleae* survival and spread.

This study aimed to explore and analyze the bacterial diversity of *Ca*. E. dacicola along with the genetic diversity of its insect host. A comprehensive sampling of the olive fly populations was conducted throughout its distribution area with focus on the Mediterranean basin. Additionally, sampling was extended to locations outside the Mediterranean, encompassing areas of America, Africa and Asia. The data gathered was used to provide new insights into the genetic structure and colonization history of the olive fly.

Results

Genetic diversity of the olive fly symbiont

To investigate the genetic variability of the primary symbiont *Ca.* E. dacicola, a total of 441 sequences of 780 bp were obtained for the bacterial 16S rRNA gene. Almost all sequences of all fly populations showed high identity (> 99% sequence similarity) with the olive fly symbiont *Ca.* E. dacicola (e.g., GenBank HQ667588). Other different bacteria species (mostly *Pseudomonas* spp.) were detected with frequencies lower than 5%. All sequence files are available from the NCBI-GenBank database (accession numbers PP278062-PP278102 for *B. oleae* and PP264486 for *Ca.* E. dacicola).

After alignment of the sequences obtained from this study, a total of three variable sites at positions 11, 282 and 657 from the 5' of the 16S rRNA, were identified. Three different haplotypes were revealed among the fly populations analyzed. The two haplotypes already described were named htA and htB according to the nomenclature proposed by Savio et al.[29]. The third haplotype was named htP because it was specific to Pakistani populations (Table S1). At nucleotide positions 11 and 282, htB and htP showed Thymine (T), while htA showed a transition to Cytosine (C); at nucleotide 657 htA and htP had Guanine (G), while htB had a transition to Adenine (A) (Table S2). All haplotype htP sequences were extended in the 3' direction up to 1280 bp using the primer fl2[29], and two further mutations were identified at nucleotide positions 943 and 945 where htP had A in both the point mutations, as htB, while htA had two transversions to C and T, respectively.

Some populations included individuals harboring htA and individuals showing htB. To test the co-presence of different bacterial haplotypes in a single olive fly, the content of the specialized oesophageal diverticulum harboring primary symbionts (i.e., oesophageal bulb) of seven individuals from some of these populations (i.e., one sample from locations 7, 40, 43, 46, 48 and 2 samples from location 36, Table S3) were cloned. Results revealed the presence of a unique *Ca*. E. dacicola haplotype in each olive fly oesophageal bulb.

Haplotypes htA, htB and htP were identified in the studied populations with the following proportions: 47.39% (209/441), 50.11% (221/441) and 2.49% (11/441), respectively (Fig. 1). All individuals of the Pakistan populations shared the haplotype htP, whereas all South-African olive flies and 90.47% of individuals of Western Mediterranean populations (Iberian Peninsula, Morocco, Algeria, and France, including Corsica and Sardinia) presented htA. Haplotype htB, instead, was exclusive of the Eastern Mediterranean populations (Egypt, Israel, Palestine, Cyprus and Turkey except Crete), as well as the Iranian and Californian populations. Individuals of populations located in the Central Mediterranean area (Italy, Croatia, Slovenia and Greece) predominantly harbored *Ca.* E. dacicola htB (82%; 116/142) while a low percentage of them showed htA (18%; 26/142) (Table S1).

SAMOVA was performed to identify genetic groups of *Ca.* E. dacicola populations among the 54 olive fly populations analyzed. The F_{CT} value reached a plateau for k = 3 (0.832; $P < 0.01$) while the highest F_{CT} value (0.838; $P < 0.01$) was obtained when fly populations were divided into 4 groups. For k > 4 the grouping started to disappear. Fixation indices (F) are reported for each k tested in Fig. S1. For k = 3, three genetically different groups were detected. Group 1 included locations 1–12, 17, 19, 21, 25, 28, 34, 36, 37, 41, 42, 48 corresponding with populations from South-Africa, Iberian Peninsula, Morocco, Algeria, France including Corsica, Sardinia (Italy), and Crete. The second group, clustered individuals from locations 13–16, 18, 20, 22–24, 26, 27, 29–33, 35, 38–40, 43–47, 50–54 corresponding with populations from Tunisia, Egypt, Italy, including Sicily, Slovenia, Croatia,

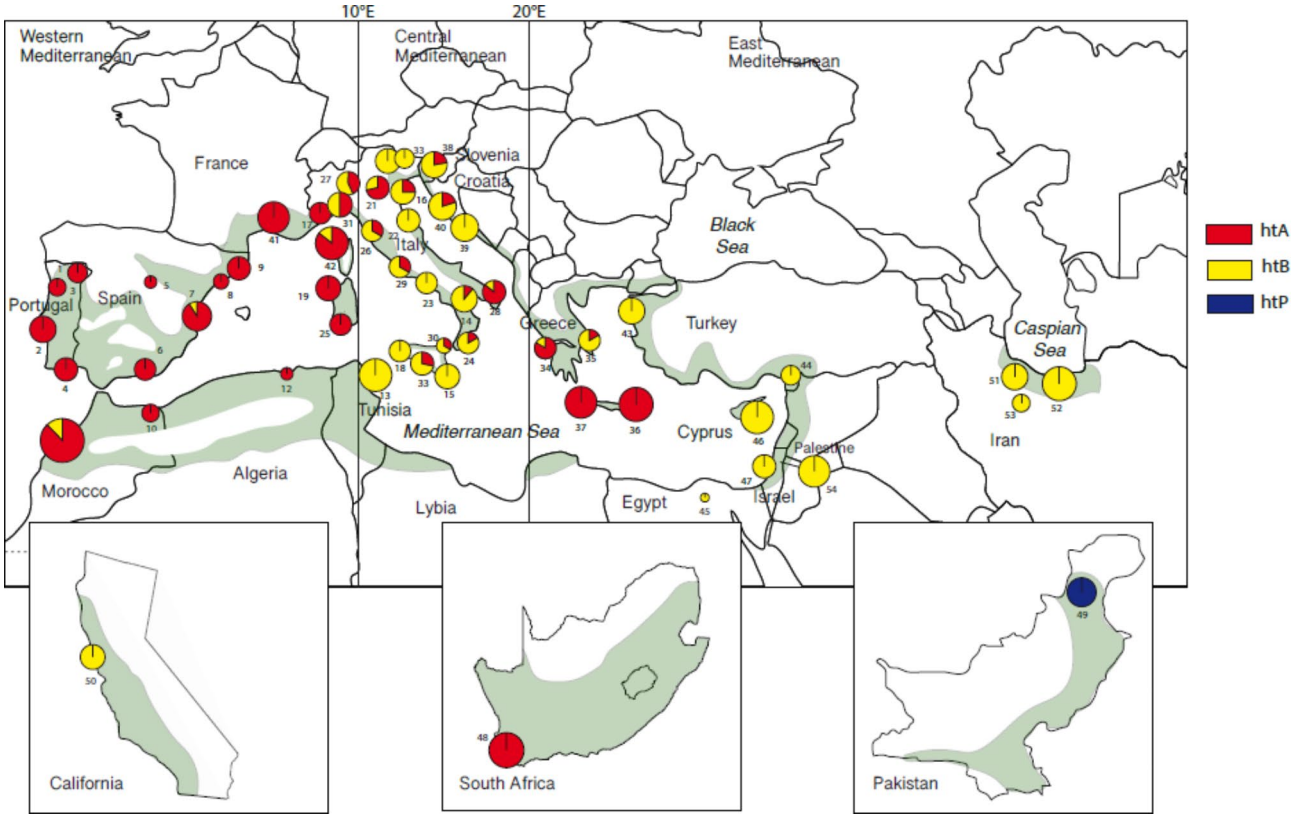

Fig. 1. Geographic distribution of the three '*Ca.* Erwinia dacicola' haplotypes (htA, htB and htP) among individuals of the 54 *Bactrocera oleae* populations (coded according to Table S3). The areas of the circles are proportional to the sampling. The shaded areas represent the distribution range of olive trees[31–33]. Maps source: https://d-maps.com.

Israel, Iran, Cyprus and Turkey while group 3, included location 49, corresponding to the Pakistani population. For k = 4, a new cluster emerged, comprising locations 27 and 31 from two distinct North Italian populations, Camogli and Campione, while the remaining three clusters remained consistent with those previously observed for k = 3. These results show an East-West distribution pattern of the *Ca.* E. dacicola haplotypes, with the exception of populations 36 (Heraklion) and 37 (Anopolis) (Fig. 1).

Mitochondrial haplotype analysis and population structure of the host

A fragment of 757 bp of the mtDNA (NADH dehydrogenase subunit 1, the leucine tRNA and the 16S rRNA) was obtained for all 441 specimens in which the symbiotic bacteria were amplified and sequenced. Twenty-seven (3,5%) polymorphic sites were identified. Calculation of population diversity indices was limited to those populations represented by more than five individuals (Table 1). Population 11 from Morocco revealed the highest number of haplotypes (N_{HT}=7) followed by population 37 (Anopolis, Greece) and 41 (Avignon, France) (N_{HT}=6). On the other hand, population 32 and 44 from Fanna (Italy) and Antakya (Turkey) respectively, exhibit only one haplotype. In general, most populations showed high levels of genetic variability. Haplotype diversity (h) ranged from 0 (Antakya, Turkey and Fanna, Italy) to 0.933 (Athens, Greece) while sequence divergence-nucleotide diversity (π) among the haplotypes changed from 0 (Antakya, Turkey and Fanna, Italy) to 0.0037 (Avignon, France).

A haplotype network, including a total of 472 sequences, was built by combining *B. oleae* sequences from GenBank (Table S4) and from our dataset. A total of 42 haplotypes were observed and designated from h1 to h42 (Fig. 2A).

Three haplotypes, h6, h2 and h1, were found at higher frequencies comprising 62.8% of all the individuals (296/471 individuals).

Among them h6, was the most common haplotype shared by 132 individuals from 37 populations, where only Central Mediterranean populations represent the 48.03%. The next two most widespread haplotypes, h2 and h1, were observed at frequencies of 22.5% (106/471 individuals) and 12.3% (58/471 individuals) respectively. In haplotype h2, 42% (60/143) individuals of the Western Mediterranean and 22.87% (35/153 individuals) of the Central Mediterranean samples were grouped together. Several populations from the eastern Mediterranean shared h1, at a frequency of 36.71% (29/79 individuals) (Figs. 2A and 3).

The network showed that some populations were exclusively represented by rare haplotypes. African individuals, for example, were represented by haplotypes h35 and h39 and Pakistan individuals were included in

Code†	Haplotypes distribution	*N*	N_{HT}	h	π	k
1	2h2, 1h4, 2h7	5	3	0.800	0.0013	1.000
2	4h2, 2h4, 2h6, 1h7	9	4	0.778	0.0015	1.166
4	6h2, 1h22	7	2	0.286	0.0003	0.285
6	4h2, 1h4, 1h5	6	3	0.600	0.0008	0.666
7	8h2, 3h7	11	2	0.436	0.0005	0.436
9	5h2, 2h6	7	2	0.476	0.0012	0.952
11	2h1, 16h2, 3h4, 1h6, 1h7, 1h31, 1h32	25	7	0.587	0.0009	0.753
13	1h2, 2h4, 7h6, 1h20, 3h26	14	5	0.725	0.0015	1.175
14	1h1, 2h2, 4h6, 1h14, 1h20	9	5	0.806	0.0022	1.666
15	2h1, 2h4, 3h25, 1h26	8	4	0.821	0.0023	1.750
16	1h2, 5h6, 1h9, 1h26	8	4	0.643	0.0015	1.178
17	5h6, 1h7	6	2	0.333	0.0013	1.000
18	1h1, 2h2, 1h4, 2h6	6	4	0.867	0.0019	1.466
19	1h2, 5h6, 1h7, 1h27	8	4	0.643	0.0017	1.357
20	2h1, 2h2, 3h6, 1h28	8	4	0.821	0.0024	1.821
21	2h1, 2h6, 3h10	7	3	0.762	0.0026	2.000
22	5h2, 2h6	7	2	0.476	0.0012	0.952
23	2h6, 1h9, 1h10, 2h26	6	4	0.867	0.0032	2.466
24	5h2, 1h6	6	2	0.333	0.0008	0.666
25	2h2, 3h6, 1h21	6	3	0.733	0.0018	1.400
26	1h1, 3h6, 2h10	6	3	0.733	0.0030	2.333
27	1h4, 3h6, 2h10, 1h22	7	4	0.810	0.0031	2.380
28	1h2, 1h4, 4h6, 1h24	7	4	0.714	0.0013	1.047
29	1h2, 5h6	6	2	0.333	0.0008	0.666
31	2h2, 5h6, 1h23	8	3	0.607	0.0020	1.571
32	5h20	5	1	0.000	0.0000	0.000
33	2h2, 5h6	7	2	0.476	0.0012	0.952
34	1h4, 4h7, 1h17	6	3	0.600	0.0020	1.533
35	1h1, 1h2, 2h4, 1h8, 1h14	6	5	0.933	0.0027	2.066
36	1h1, 4h2, 8h6, 2h16	15	4	0.667	0.0017	1.333
37	1h1, 2h2, 5h6, 1h8, 3h10, 1h15	13	6	0.821	0.0029	2.230
38	1h1, 3h2, 4h6, 1h7	9	4	0.750	0.0020	1.555
39	3h2, 4h6, 1h7, 1h18, 1h19	10	5	0.800	0.0025	1.911
40	1h2, 8h6, 1h8	10	3	0.378	0.0007	0.600
41	2h6, 5h8, 2h10, 1h11, 2h12, 1h13	13	6	0.833	0.0037	2.846
42	2h2, 1h4, 9h6, 1h8, 1h9	14	5	0.593	0.0012	0.912
43	2h1, 1h2, 5h6, 1h37	9	4	0.694	0.0020	1.555
44	5h1	5	1	0.000	0.0000	0.000
46	12h1, 1h2, 1h3	14	3	0.275	0.0003	0.285
47	5h1, 1h29, 1h30	7	3	0.524	0.0010	0.761
48	15h35, 1h39	16	2	0.125	0.0001	0.125
49	1h6, 4h33, 1h34, 1h35, 4h36	11	5	0.782	0.0020	1.563
50	4h1, 1h6, 3h38	8	3	0.679	0.0021	1.607
51	2h4, 3h6, 1h7, 3h28	9	4	0.806	0.0020	1.444
52	1h1, 1h4, 4h6, 2h7, 7h28	15	5	0.733	0.0022	1.676
54	9h1, 1h6, 1h30, 1h41, 1h42	13	5	0.538	0.0012	0.923

Table 1. Genetic diversity indices of *Bactrocera oleae* populations analyzed. *N*, number of individuals for each population; N_{HT}, number of haplotypes in each population; h, haplotype diversity; π, nucleotide diversity; k, mean number of pairwise differences per sequence. †Population codes correspond to those listed in (Table S3).

four unique haplotypes (h33, h34, h36 and h40) as well as in h6 even if for just one haplotype (Table 1). Haplotypes h41 and h42 were unique haplotypes exclusively represented by individuals of the Palestine population.

The haplotype network of *B. oleae* did not reveal a clear geographical structure. Nevertheless, certain populations such as those from Africa and Pakistan were well structured and differentiated from the Mediterranean and American haplotypes.

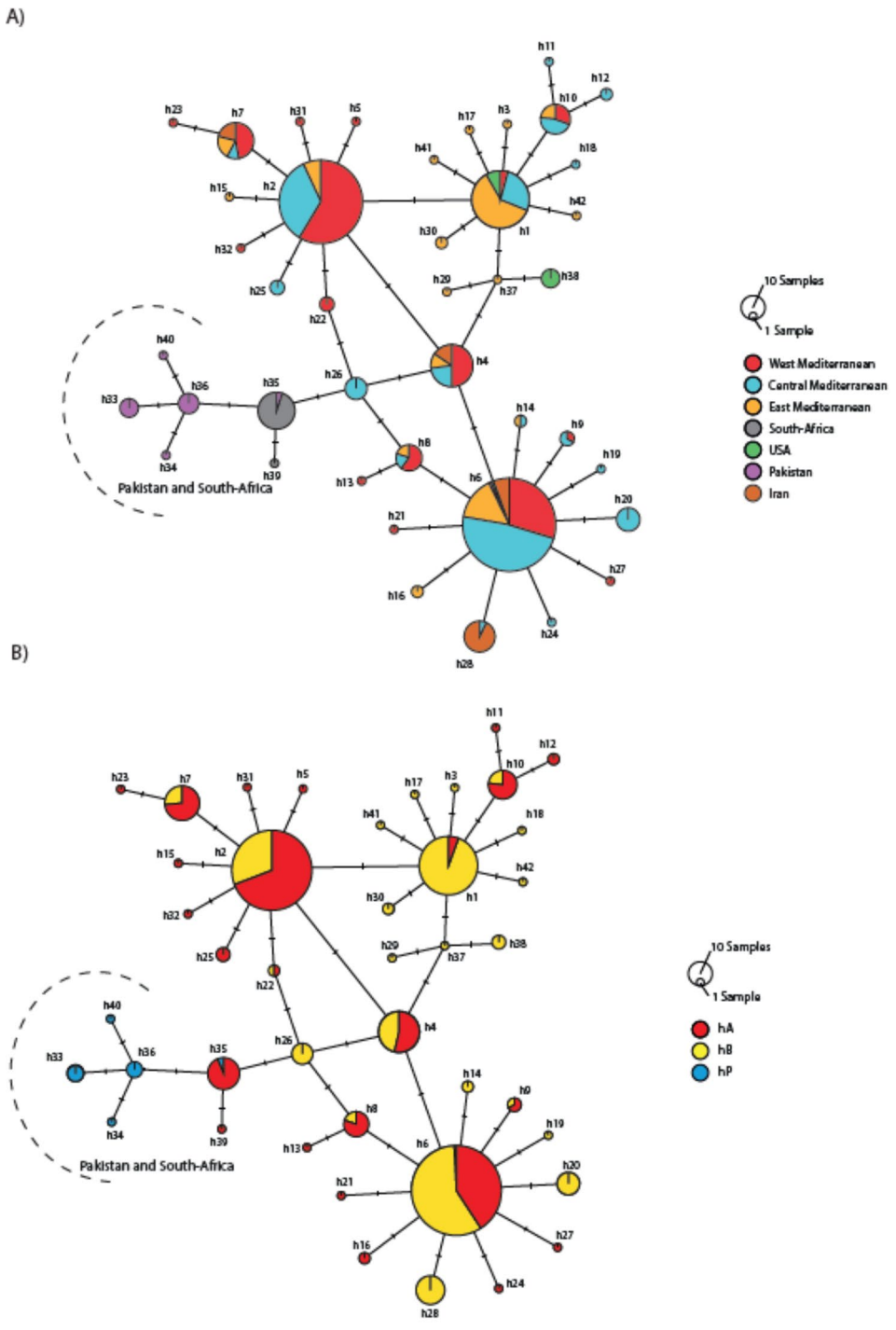

Fig. 2. (**A**) Haplotype network of *Bactrocera oleae* populations inferred by TCS v1.21[34]. The size of the circles is proportional to the haplotype frequency. Sampling region of each haplotype is color coded as in the legend. (**B**). Haplotype network of *Ca.* Erwinia dacicola among *B. oleae* populations inferred by TCS v1.21. The size of the circles is proportional to the haplotype frequency and color coded as in the legend.

Tests of homogeneity among populations were performed using AMOVA. For this purpose, population groupings were chosen according to the SAMOVA results obtained for *Ca.* E. dacicola analyses (k = 3 and k = 4) and to the six major geographical regions of the sampling area (West, Central and Eastern Mediterranean areas, South Africa, California, Iran and Pakistan) (Table S3). AMOVA results (Table 2), revealed significant

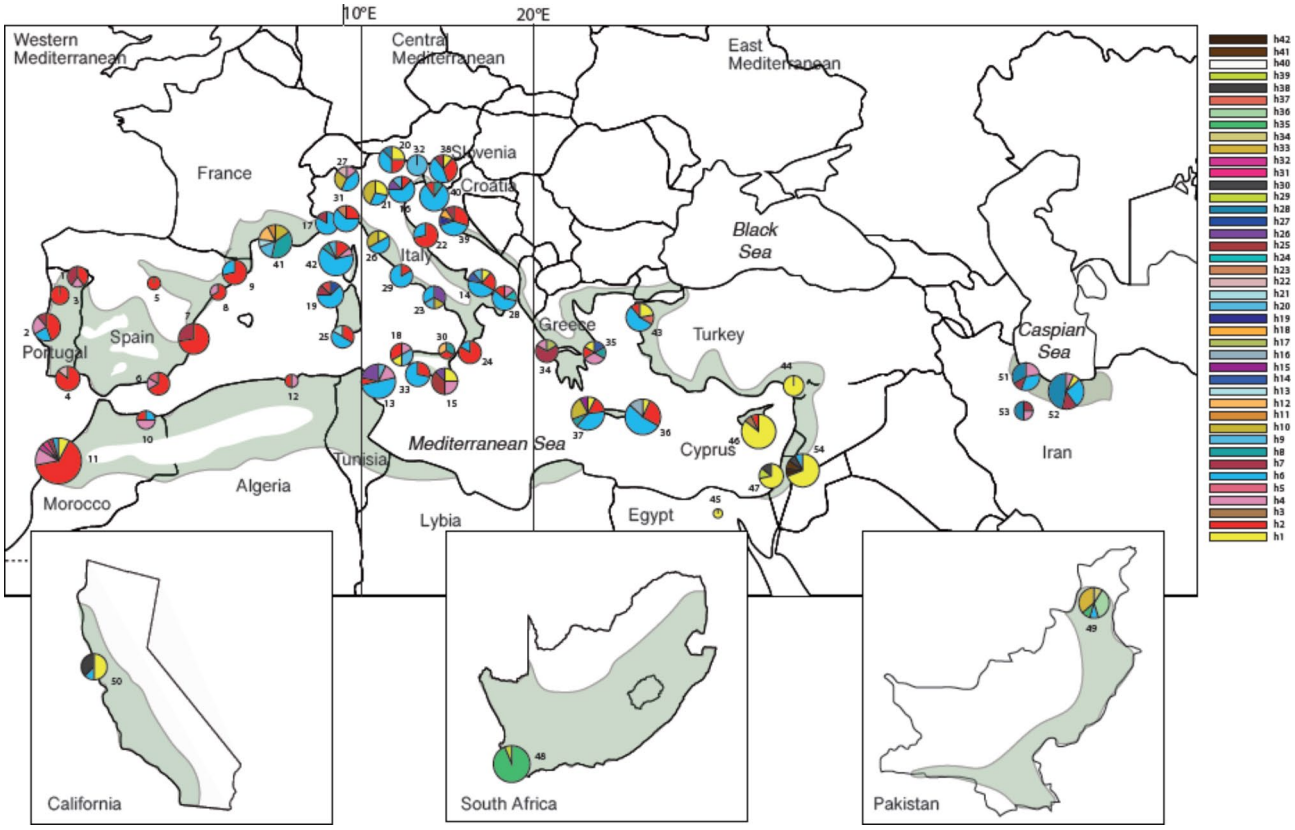

Fig. 3. Geographic distribution of mitochondrial haplotypes observed among individuals of Mediterranean, Californian, South African and Pakistani populations of *B. oleae* coded according to Table S3. The areas of the circles are proportional to the sampling The shaded areas represent the distribution range of olive trees according to[31–33]. Maps source: https://d-maps.com.

Structure	Source of variation	Variance (%)	Fixation indices	*P* value
Grouping by SAMOVA results (k = 3)	Among groups	7.856	F_{CT}=0.0743	0.03
	Among populations within groups	35.68	F_{SC}=0.3644	< 0.001
	Within populations	62.23	F_{ST}=0.4116	< 0.001
Grouping by SAMOVA results (k = 4)	Among groups	6.7	F_{CT}= 0.0589	0.078
	Among populations within groups	36.35	F_{SC}=0.3687	< 0.001
	Within populations	62.23	F_{ST}=0.4059	< 0.001
Grouping by geographical regions	Among groups	31.82	F_{CT}=0.2265	< 0.001
	Among populations within groups	20.43	F_{SC}=0.2500	< 0.001
	Within populations	62.23	F_{ST}=0.4238	< 0.001

Table 2. Analysis of molecular variance (AMOVA) for three different groups of *Bactrocera oleae* populations based on SAMOVA results of *Ca.* E. Dacicola and geographic subdivisions.

differences among groups of populations for the three SAMOVA hierarchical groups ($P = 0.02$) and when olive fly populations were divided based on geographical areas ($P < 0.001$). However, a significant population structure among groups was not detected for the SAMOVA groups k = 4 ($P = 0.078$).

Relationships between *Ca.* E. dacicola and *B. oleae* haplotypes

A bipartite interaction matrix between the haplotypes of the insect host *B. oleae* and the primary symbiont *Ca.* E. dacicola was built considering samples in which the sequence of both the symbiont and the host were obtained (Fig. 4). The network of interactions consisted of three symbiont haplotypes and 41 host haplotypes. The specialization index H2' showed significantly high network specialization (H2'=0.357, Z-score = 22.466, $P < 0.001$). A haplotype network was constructed based on the three symbiont lineages and the insect haplotypes observed in our dataset to define symbiont-host relationships (Fig. 2B). Both networks showed a non-random association of the host-symbiont haplotypes. Olive fly populations sharing haplotype h2, harbored at a frequency

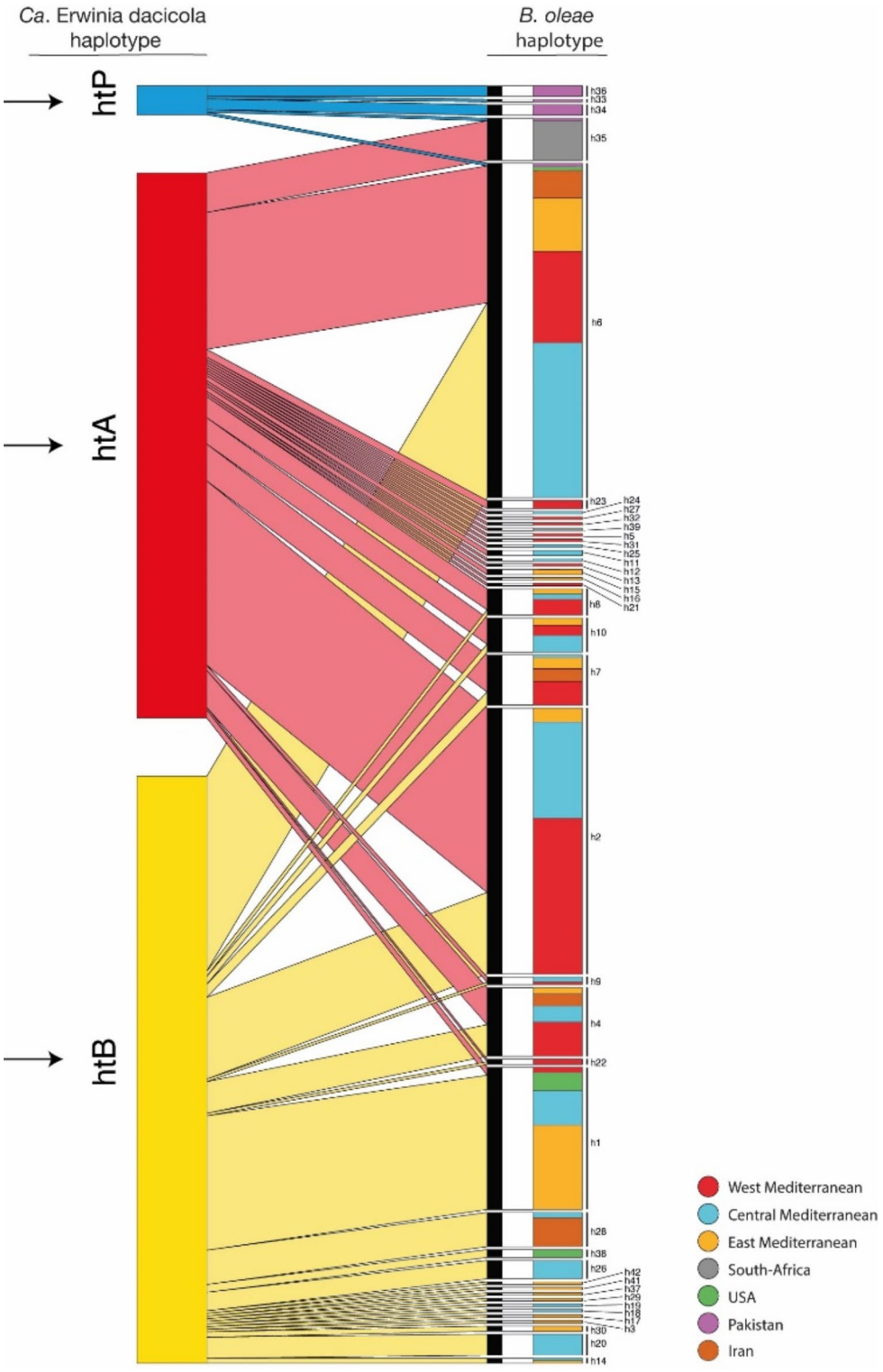

Fig. 4. Weighted bipartite networks of interactions between the symbiont *Ca.* Erwinia dacicola (left bars) and the host *Bactrocera oleae* (right bars). The width of the bars reflects the relative frequency of the symbiont and the host haplotypes, respectively. The width of the lines connecting bars indicates the frequency with which *Ca.* Erwinia dacicola haplotype was found in each *B. oleae* haplotype. The geographical origin of each sample is represented by colored horizontal rectangles.

of 69.3% (70/101) the bacterial haplotype htA. In contrast, the populations within haplotype h1 presented 94.4% (51/54) of the bacterial genetic variant htB. Populations with haplotype h6, mostly located in the Central Mediterranean region, were associated with both symbiont haplotypes htB and htA (58.2% and 41%) respectively (Table S5). Several host haplotypes were associated with only one symbiont haplotype. Haplotypes h33, h34 and h36 shared by Pakistani samples harbored exclusively the third haplotype htP. The South African population sharing the haplotype h35 and h39, harbored almost exclusively the bacterial haplotype htA. In addition, the Californian and the Iranian olive fly populations harbored only the variant htB.

Discussion

Three bacterial haplotypes of *Ca.* E. dacicola were identified in the olive fly populations across the studied geographical range. Haplotypes htA and htB, previously recovered in Italian[29], Portuguese, Spanish, French, Tunisian[26] and USA[35] populations, were present across the Mediterranean populations. Furthermore, htA was found in South Africa olive fly populations while htB was found in Iran and in California populations. A third and new haplotype, htP, was exclusively found in fly populations collected in Pakistan (Fig. 1). Despite the extensive sampling across a wide area of the olive fly distribution range, haplotypes htC and htD, identified by Nobre[26] in Tunisia were not detected in this study. As already detected by Savio et al.[29], the results confirmed the presence of a unique *Ca.* E. dacicola haplotype in each olive fly oesophageal bulb, suggesting that the different haplotypes did not coexist in the same olive fly.

The results revealed a significant genetic and geographic structure of the symbiont with an East-West genetic differentiation pattern in the Mediterranean basin. Western populations (Algeria, Morocco, Iberian Peninsula, France, Corsica, Sardinia, and Liguria Region) mainly harbored the bacterial symbiont htA while most of the olive flies collected from Eastern Mediterranean (Egypt, Israel, Palestine, Cyprus, Turkey), harbored the genetic lineage htB. In the Central Mediterranean area haplotypes htA and htB, were both present with the latter being predominant. This parallels with the selection of olives during domestication, as the genetics of olive cultivars demonstrate that one early lineage is restricted to the East, Levant (clorotype E1)[36,31]. Some archaeological findings go back to up to 7,000 years, submerged ruins on the coast of Israel[37]. Currently, and still debated, the main cultivated olive (E1) originated and spread from an area between nowadays Syria and Turkey[36,31]. It is suspected that during the domestication processes, the selected olive varieties, which correspond to lineage E1 (clorotype) correlated with *Ca.* E. dacicola htB. This is expected due to the selected traits of the olives either for oil or direct consumption, offering the favorable habitat to *B. oleae* carrying symbiont htB. The wild olive is considerably variable in terms of traits.

The Central Mediterranean area could be a confluence zone between Eastern and Western areas. These Mediterranean populations may result from admixture between Western individuals carrying the predominant symbiont haplotype htA and Eastern individuals harboring the symbiont haplotype htB. For olive cultivars, clorotype E1 is restricted to the East and clorotypes E2 and E3 are specific to West and Central areas[36]. Again, the results resemble the tripartite association, with the synchronic distribution of symbiont (able to pre-digest the olive for colonization), *B. oleae*, and its host plant. Settlements from the Bronze age, in SE Spain show remarkable changes in olive shape due to domestication[38], with an increase in the number of olive-stone remains on Late Bronze Age[16]. Interestingly, traits characteristic of Middle East olives are only seen later, towards the change in the millennium[16]. Migrations and introductions of olive cultivars from the East towards the Centre and West Mediterranean could explain the origin of mixture of symbiont strains (haplotypes).

The presence of both symbiont haplotypes in the same populations could be also attributed to different events of losses and new acquisitions of different *Ca.* E. dacicola haplotypes bypassing the model of strict vertical transmission. While vertical transmission is the primary and prevalent mechanism for ensuring symbiotic bacterial transmission, an exchange of *Ca.* E. dacicola haplotypes by horizontal transmission at specific stages of the *B. oleae* life cycle cannot be excluded. The olive fly spends most of its life cycle at the larva stage on, or within host plant tissues. Therefore, shared breeding sites and replacement of bacterial haplotypes could likely occur among larvae living in the same olive fruit. The occurrence in adult olive flies of horizontal transmission of the bacterial symbiont through cohabitation was reported under laboratory conditions[39]. Furthermore, the occurrence of accidental invasion and replacement of phylogenetically related symbiotic bacteria species has also been hypothesized between different species of non-frugivorous fruit flies (subfamily Tephritinae) visiting or developing in the same host plant[40].

Similar to the primary symbiont data, the mitochondrial haplotype analysis (AMOVA) of the host flies revealed a significant genetic and geographic structure supporting the subdivision of populations into different geographical subgroups. This genetic structure is in accordance with the findings of previous studies based on different mitochondrial and microsatellite markers which suggested the existence of three different clusters in the Mediterranean basin (Wester, Central and Eastern populations)[21,23,25,26] plus other clusters from Pakistan and South Africa[19,27]. Indeed, Nardi[19] estimated the split of African and Mediterranean fly populations around 4000–5000 years ago, which parallels the domestication of olive trees in the Middle East.

The whole dataset, considering the genetic diversity of both the primary symbiont and the host fly, allowed to observe some geographic and genetic patterns of the olive fly, *B. oleae*. The Pakistani populations showed significant symbiont-host haplotypes associations. The exclusive and (newly described) symbiont lineage htP was found with host haplotypes h36, h37, h38 differing from the rest of the populations. This outcome aligns with the findings of Nardi et al.[19,27], who suggested an ancient separation of the Pakistani fly population, and it is consistent with its description as a distinct taxonomic subspecies (e.g., var. asiatica: Silvestri, 1916). Fly haplotype h36 exclusive of the Pakistani populations, and perhaps symbiont htP corresponds to the earliest settlement of olives in the region.

The two large fly populations from Crete (Anopolis and Heraklion) harboring the western symbiont htA represent an exception to the predominance of htB in the eastern Mediterranean. The prevalence of htA in

the island populations could be explained assuming an effect of geographic isolation, which contributed to the absence of external gene flow. A study on the natural distribution of *Olea europaea* by the identification of wood-charcoal, reported the absence or late appearance of the olive tree in Crete during the Holocene, in contrast to the abundant presence of the botanical species in Cyprus and in the Levantine basin[15]. The introduction of olive to Crete occurred relatively late, sometime after the Middle Neolithic, as a result of overseas contacts[41]. Although the commercial trades with the neighbors have been extensive since the introduction of the olive cultivation, it is likely that local fly populations were largely contained in the island. This hypothesis is further supported by the presence in the island of exclusive mitochondrial haplotypes of the fly host *B. oleae*, h15 and h16.

Olive fly populations found in California were suggested to have their origins in the Eastern Mediterranean, particularly in Turkey as a result of human activities [e.g.,[19,23]]. The presence in the Californian populations of the eastern symbiont haplotype htB, associated with host haplotypes characteristic of the eastern Mediterranean populations, including Turkey, offers additional support for this hypothesis.

Bactrocera oleae was apparently absent in Iranian olive groves until 2004 when it changed from a quarantine pest to becoming the most important pest of olive trees in this region [e.g.,[11]]. In this study, all individuals within the Iranian populations harbored the symbiont haplotype htB and olive fly haplotypes shared with individuals from central and eastern Mediterranean populations. The high predominance of the symbiont haplotype htB in central populations combined with the presence of host haplotypes typical from the central Mediterranean (Fig. 2B) provides support to the hypothesis advanced by Ramezani et al.[11] of a Central European origin for the Iranian olive flies. Further genetic studies using deeper sequencing methods will significantly enhance our understanding of the genetic population structure and will add useful information regarding the functional properties of the different symbiont haplotypes.

The high genetic diversity of the host, *B. oleae*, displaying 42 haplotypes which interact with only three *Ca.* E. dacicola haplotypes, can be expected from a long-term host-symbiont association. A high genetic diversity of the insect host compared to that of the primary symbiont was also found in the brown marmorated stink bug, *Halyomorpha halys*, and its bacterial primary symbiont *Pantoea carbekii*, seven haplotypes of the primary symbiont were found against 25 haplotypes of the insect host[3].

A phylogenetic study of the primary symbionts of the genus *Ca*. Stammerula sp., harbored in flies of the subfamily Tephritinae, showed that the stable lineages are conserved and that the different fruit fly species host a specific and unique single bacterial symbiont lineage[40,42]. Moreover, *Ca*. E. dacicola is uncultivable as well as *Ca*. Stammerula[42,43]. This pattern is generally observed and expected in endosymbionts, or intracellular bacterial symbionts that have achieved a stable, long-standing host association living in the protected intracellular environment[30]. Despite their extracellular location, the primary symbionts described for the subfamily Tephritinae[40,42] and *B. oleae*[43] are harbored and protected within specialized structures. In particular, the adult olive fly hosts the primary symbiont within a specialized oesophageal diverticulum called the oesophageal bulb[44,45]. Similarly, in the subfamily Tephritinae, species-specific primary symbionts are hosted within epithelial evaginations or crypts of the fly midgut protected by the peritrophic tube[46,47]. The presence of primary symbiont host structures and their uncultivability accounts for the presence of well-established and co-evolved associations. These associations likely involve the selective uptake of a single symbiotic bacterial species, being the organ able to discriminate from non-symbiotic bacteria[48]. This is a characteristic of enduring and ancient host-symbiont associations.

Primary symbionts seem to be exposed to selective pressure over time, removing deleterious mutations with reduction in base substitutions; this feature is dominating in older or ancient associations, and it slows down the molecular evolutionary rates of the symbionts[30]. Nevertheless, as the *B. oleae* clade might have diverged from other species of the genus approximately 25 million years ago[49], and it separated from its sister species *B. biguttula* by only about 10 million years[50], the symbiotic association with *Ca*. E. dacicola might have been established concurrent with the emergence of the clade and thus be in line with younger associations for which a high substitution rate is attributed[30]. Nonetheless, the presence of strong selective pressure, would explain the low diversity found for *Ca*. E. dacicola. In contrast to other frugivorous and polyphagous fruit flies, *B. oleae* larvae feed exclusively on olive pulp. In unripe olives, the pulp is particularly rich in anti-nutritive and antimicrobial secondary metabolites, including oleuropein, a phenolic glycoside[28]. The persistent presence of the symbiont in the larvae's gut plays a crucial role in enabling *B. oleae* to colonize the harsh olive substrates. This fundamental function has led to a close co-evolutionary relationship between the symbiont and the olive fly, resulting in mutual dependence and selective pressure to maintain the symbiont in the complex system of domestication of the olive.

The results obtained from studies on the genetic variability of *Ca*. E. dacicola may stimulate some hypothetical assumptions regarding the long history of coevolution between host insect and symbiont as well as on the separation of the three identified lineages (A, B and P) during the phytogeographical events involving both insect and host plant.

It is interesting to note that the geographical distribution of the three haplotypes of *Ca*. E. dacicola described in this paper, align closely with *B. oleae* genetic macro-groups identified by Nardi et al.[27]. They classify the olive fly into three genetic groups: an African population, regarded as the species' center of differentiation; a Pakistani population, considered the relict of an ancient separation; and a Mediterranean population, which is further divided into three subgroups: eastern, central, and western. This correspondence invites speculation about the origin and the divergence among the three identified lines of *Ca*. E. dacicola. The current distribution of the three haplotypes could be the result of the ancient basal separation of the wild olive tree into the three lines (African, Asian and Mediterranean) that is traced back to the early Pliocene[18,51] followed by the repeated fragmentation of the Mediterranean range into different refuge stations that occurred during the Pleistocene[27]. These events forced *B. oleae* populations into long periods of geographical isolation. In the light of these results, it becomes

evident how the presence of a vertical transmission pattern for the *B. oleae* symbiont provides a valuable tool in understanding the evolutionary history and population genetics of its host insect.

Materials and methods

Sample collection and preparation

Adults of *B. oleae* and/or infested olives were collected from 54 olive growing sites from 19 states around the Mediterranean basin, the Middle East, South Africa, Iran, Pakistan, and USA. According to[52], the Mediterranean basin has been divided in three subregions as follows: western Mediterranean (WM) area, extending between 350°W–10°E and 30°N–47°N, central Mediterranean (CM) area, extending between 10–20°E and 30°N–47°N and eastern Mediterranean (EM), extending between 20°E–38°E and 30°N–47°N (Table S3).

Collected olive flies were kept in 96%, ethanol and shipped to the laboratory where they were stored at − 20 °C until being processed. In the case of the collected infested olives, they were kept in transparent plastic boxes (20 × 15 × 8 cm) topped with a net until larvae emerged and developed into adults. Emerged insects were then transferred into net rearing cages (10 × 10 × 10 cm) and fed with a sugar diet (50% w/v glucose solution) and water *ad libitum* according to Savio et al.[29]. In order to ensure a sufficient number of bacteria in the oesophageal bulbs, flies were processed after 2–3 days[45]. Before performing the molecular studies, flies were dissected to extract the esophageal bulb hosting symbionts, under a laminar flow hood[43]. Oesophageal bulbs were then gently transferred into Eppendorf tubes and kept at − 20 °C for further analysis.

The DNA extraction protocol described by[53] was used to extract both the insect and the bacterial DNA content of the oesophageal bulbs, with an average of eight individuals per population.

Symbiont analyses

The bacterial DNA was analyzed by amplifying a fragment of 16S rRNA gene, using the universal primers fD1 and rP1[54,55]and 63 F and 1389R[56]. Amplifications were performed in 20 µl reaction volumes containing 1X PCR GoTaq Flexi Buffer (Promega), 2.5 mM MgCl2, 0.1 mM dNTPs, 0.5 mM each primer, 1 U GoTaq Flexi DNA polymerase (Promega) and a 1 ml sample of a 1: 30 dilution of the DNA extract. The standard thermal profile for the amplification of the 16S rRNA genes included an initial denaturation at 95 °C for 2 min, 27 cycles at 95 °C 30 s, 56 °C for 30 s, 72 °C for 90 s and a final extension at 72 °C for 5 min.

In addition, to check the co-presence of different bacterial haplotypes in a single olive fly, PCR products obtained from seven oesophageal bulbs were cloned into *Escherichia coli* JM109 cells. The oesophageal bulbs of olive fly individuals from populations that exhibited both bacterial haplotypes were mainly chosen. The P-GEM-T Easy vector (Promega) was used for cloning, following the manufacturer's recommendations. Transformations were verified using PCR assays with the M13-T7 universal primer pair. PCR products were checked by electrophoresis on 1% agarose gel stained with SYBR® Safe (Invitrogen), then purified with a mix of Exonuclease and Antarctic Phosphatase (New England Biolabs) and sequenced at the BMR Genomics Service (Padua, Italy).

Bacterial sequence chromatograms were quality checked and aligned using MEGA 11[54]. Low-quality regions found at the beginning and end of each sequence were trimmed and low-quality sequences were not included in the analysis. A GenBank BLAST analysis of the sequences obtained was run through the NCBI website (www.ncbi.nlm.nih.gov) to assess the identity of the sequences.

Spatial analysis of molecular variance was performed using SAMOVA v12.02[57] to assess the presence of geographical patterns in the genetic structure. Different numbers of populations (K) in the range of K = 2–7 based on 100 simulated annealing processes were simulated to search for the optimal grouping option that maximizes the among-group component (F_{CT}) of the overall genetic variance.

Insect host analyses

A region of the mitochondrial DNA of the olive fly, corresponding to the NADH dehydrogenase subunit 1, the leucine tRNA and the 16S mitochondrial region, was amplified using the universal pair of primers: N1-J12261m and LRN13398[58,59]. PCR reactions were carried out using identical volumes as those employed for the symbiont DNA amplification. The cycling program included an initial denaturation at 96 °C for 5 min, followed by 35 cycles at 96 °C for 50 s, 56 °C for 50 s and 72 °C for 2 min, with a final extension step at 72 °C for 5 min. PCR amplified products were purified and sequenced as reported above.

Haplotype distributions and frequencies within the populations (N_{HT}), gene diversity (h), nucleotide diversity (π) and mean number of pairwise differences (k) were assessed using ARLEQUIN v3.5 software package[60]. Hierarchical analysis of molecular variance (AMOVA)[61] of *B. oleae* populations was performed with ARLEQUIN software to evaluate the distribution of molecular variance among groups, populations and individuals based on the groups defined by SAMOVA analysis on *Ca.* E. dacicola and by geographical areas, with 10,000 random permutations.

Sequences of the same mitochondrial marker available in the NCBI database were retrieved and added to this study dataset (Table S4). A haplotype parsimony network with a probability cut-off of 95% was reconstructed using the TCS 1.21 software[62] and PopART 1.7[34] and used for depicting the geographical relationships among haplotypes. Ambiguous connections (loops) were resolved using approaches from coalescent theory based on three criteria: frequency, network location and geography[63].

Symbiont-host haplotype analysis

Relationships between the symbiont and its insect host haplotypes were defined by a haplotype network constructed by TCS v1.21[62] and a bipartite network. The overall network specialization was determined using the specialization index (H2'). It ranges between 0 (no specialization between symbiont and host) and 1 (complete specialization between symbiont and host)[64]. The metric significance was checked by z-score and

calculated using 1,000 null models obtained with the Patefield algorithm[65]. Graph and metrics were generated by the bipartite package[66] in R 3.0.1 software (R Development Core Team—https://www.r-project.org).

Data availability

All data used in this paper is available through GenBank accession PP264486 for the new symbiont sequences and PP278062-PP278102 for the host sequences.

Received: 26 February 2024; Accepted: 12 September 2024
Published online: 28 September 2024

References

1. Wang, Y. et al. Phylogeography of *Diaphorina citri* (Hemiptera: Liviidae) and its primary endosymbiont, '*Candidatus* Carsonella ruddii': an evolutionary approach to host–endosymbiont interaction. *Pest. Manag. Sci.***74**, 2185–2194 (2018).
2. Otero-Bravo, A. & Sabree, Z. L. Comparing the utility of host and primary endosymbiont loci for predicting global invasive insect genetic structuring and migration patterns. *Biol. Control* **116**, 10–16 (2018).
3. Martinez-Sañudo, I. et al. Co-haplotyping symbiont and host to unravel invasion pathways of the exotic pest *Halyomorpha halys* in Italy. *Sci. Rep.***10**, 18441 (2020).
4. Funk, D. J., Helbling, L., Wernegreen, J. J. & Moran, N. A. Intraspecific phylogenetic congruence among multiple symbiont genomes. *Proc. R. Soc. B: Biol. Sci.* 267, 2517–2521 (2000).
5. Tzanakakis, M. E. *Insects and Mites Feeding on Olive: Distribution, Importance, Habits, Seasonal Development and Dormancy* (Brill Academic, 2006).
6. Malheiro, R., Casal, S., Baptista, P. & Pereira, J. A. A review of *Bactrocera oleae* (Rossi) impact in olive products: from the tree to the table. *Trends Food Sci. Technol.***44**, 226–242 (2015).
7. Ordano, M. et al. Olive fruit fly (*Bactrocera oleae*) population dynamics in the Eastern Mediterranean: influence of exogenous uncertainty on a monophagous frugivorous insect. *PloS ONE***10**, e0127798. https://doi.org/10.1371/journal.pone.0127798 (2015).
8. Daane, K. M. & Johnson, M. W. Olive Fruit fly: managing an ancient pest in Modern Times. *Annu. Rev. Entomol.***55**, 151–169 (2010).
9. Copeland, R. S., White, I. M., Okumu, M., Machera, P. & Wharton, R. A. Insects associated with fruits of the Oleaceae (Asteridae, Lamiales) in Kenya, with special reference to the Tephritidae (Diptera). *Bish. Mus. Bull. Entomol.***12**, 135–164 (2004).
10. Bon, M. C. et al. Populations of *Bactrocera oleae* (Diptera: Tephritidae) and its parasitoids in Himalayan Asia. *Ann. Entomol. Soc. Am.***109**, 81–91 (2016).
11. Ramezani, S., Blibech, I., Trindade Rei, F., Van Asch, B. & Da Costa, L. T. *Bactrocera oleae* (Diptera: Tephritidae) in anan: An invasion from the Middle West. *Eur. J. Entomol.***112**, 713–721 (2015).
12. Rice, R. E. Olive fruit fly, *Bactrocera* (Dacus) *oleae*. *UC Plant. Prot. Q***9**, 2 (1999).
13. Green, P. S. A revision of Olea L. (Oleaceae). *Kew Bull.***57**, 91–140 (2002).
14. Tierney, J. E., Pausata, F. S. & de Menocal, P. B. Rainfall regimes of the Green Sahara. *Sci. Adv.***3**, e1601503. https://doi.org/10.1126/sciadv.1601503 (2017).
15. Carrión, Y., Ntinou, M. & Badal, E. *Olea europaea* L. in the north Mediterranean Basin during the Pleniglacial anearly middle Mholocenelocene. *Quat. Sci. Rev.***29**, 952–968 (2010).
16. Newton, C., Lorre, C., Sauvage, C., Ivorra, S. & Terral, J. F. On the origins and spread of *Olea europaea* L.(olive) domestication: evidence for shape variation of olive stones at Ugarit, late bronze age, Syria - a window on the Mediterranean Basin and on the westward diffusion of olive varieties. *Veg. Hist. Archaeobot.***23**, 567–575 (2014).
17. Juniper, B. E. & Maberly, D. J. *The Story of the Apple* (Timber, 2006).
18. Besnard, G., Khadari, B., Baradat, P. & Bervillè, A. *Olea europea* (Oleaceae) phylogeography based on chloroplast DNA polymorphism. *Theor. Appl. Genet.***104**, 1353–1361 (2002).
19. Nardi, F., Carapelli, A., Dallai, R., Roderick, G. K. & Frati, F. Population structure and colonization history of the olive fly, *Bactrocera oleae* (Diptera. Tephritidae). *Mol. Ecol.***14**, 2729–2738 (2005).
20. White, I. M. Taxonomy of the Dacina (Diptera: Tephritidae) of Africa and the Middle East. *Afr. Entomol. Mem.***2**, 1–156 (2006).
21. van Asch, B., Pereira-Castro, I. & Rei, F. T. Costa, L. T. Marked genetic differentiation between Western Iberian and italic populations of the olive fly: Southern France as an intermediate area. *PLoS ONE***10**, e0126702. https://doi.org/10.1371/journal.pone.0126702 (2015).
22. Augustinos, A. A. et al. Microsatellite analysis of olive fly populations in the Mediterranean indicates a westward expansion of the species. *Genetica***125**, 231–241 (2005).
23. Zygouridis, N. E., Augustinos, A. A., Zalom, F. G. & Mathiopoulos, K. D. Analysis of olive fly invasion in California based on microsatellite markers. *Heredity***102**, 402–412 (2009).
24. Knap, T. & Bandelj, D. Microsatellite analysis revealed a different approach of control of olive fly population (*Bactrocera oleae*) in Slovenia. *J. Appl. Entomol.***141**, 256–265 (2017).
25. Matallanas, B., Lantero, E., M'Saad, M., Callejas, C. & Ochando, M. D. Genetic polymorphism at the cytochrome oxidase I gene in Mediterranean populations of *Batrocera oleae* (Diptera: Tephritidae). *J. Appl. Entomol.***137**, 624–630 (2013).
26. Nobre, T. Olive fruit fly and its obligate symbiont *Candidatus* Erwinia dacicola: two new symbiont haplotypes in the Mediterranean basin. *PloS ONE***16**, e0256284. https://doi.org/10.1371/journal.pone.0256284 (2021).
27. Nardi, F. et al. Domestication of olive fly through a multi-regional host shift to cultivated olives: comparative dating using complete mitochondrial genomes. *Mol. Phylogenet. Evol.***57**, 678–686 (2010).
28. Ben-Yosef, M., Pasternak, Z., Jurkevitch, E. & Yuval, B. Symbiotic bacteria enable olive fly larvae to overcome host defences. *R Soc. Open. Sci.***2**, 150170. https://doi.org/10.1098/rsos.150170 (2015).
29. Savio, C. et al. Evidence of two lineages of the symbiont '*Candidatus* Erwinia Dacicola'in Italian populations of *Bactrocera oleae* (Rossi) based on 16S rRNA gene sequences. *Int. J. Syst. Evol. Microbiol.***62**, 179–187 (2012).
30. Allen, J. M., Light, J. E., Perotti, M. A., Braig, H. R. & Reed, D. L. Mutational meltdown in primary endosymbionts: selection limits Muller's ratchet. *PLoS ONE***4**, 2 (2009).
31. Besnard, G., Terral, J. F. & Cornille, A. On the origins and domestication of the olive: a review and perspectives. *Ann. Bot.***121**, 385–403 (2018).
32. Langgut, D. & Garfinkel, Y. 7000-year-old evidence of fruit tree cultivation in the Jordan Valley, Israel. *Sci. Rep.***12**, 7463 (2022).
33. Ashraf, U. et al. Ecological niche model comparison under different climate cenarios: a case study of *Olea* Spp. *Asia Ecosphere*. **8**, e01825 (2017).
34. Leigh, J. W. & Bryant, D. P. O. P. A. R. T. Full-feature software for haplotype network construction. *Methods Ecol. Evol.***6**, 1110–1116 (2015).
35. Estes, A. M., Hearn, D. J., Bronstein, J. L. & Pierson, E. A. The olive fly endosymbiont, *Candidatus* Erwinia Dacicola, switches from an intracellular existence to an extracellular existence during host insect development. *Appl. Environ. Microbiol.***75**, 7097–7106 (2009).

36. Besnard, G. et al. Population genetics of Mediterranean and Saharan olives: geographic patterns of differentiation and evidence for early generations of admixture. *Ann. Bot.***112**, 1293–1302 (2013).
37. Kislev, M. E. Wild olive stones at submerged Chalcolithic Kfar Samir, Haifa, Israel. *J. Isr. Prehist Soc.***26**, 134–195 (1995).
38. Brun, J. P. *Archéologie du vin et de l'huile. De la Préhistoire à l'époque Hellénistique*Errance,. (2004).
39. Bigiotti, G., Pastorelli, R., Guidi, R., Belcari, A. & Sacchetti, P. Horizontal transfer and finalization of a reliable detection method for the olive fruit fly endosymbiont, *Candidatus* Erwinia dacicola. *BMC Biotechnol.***19**, 1–12 (2019b).
40. Mazzon, L. et al. Phylogenetic relationships between flies of the Tephritinae subfamily (Diptera, Tephritidae) and their symbiotic bacteria. *Mol. Phylogenet. Evol.***56**, 312–326 (2010).
41. Bottema, S. & Sarpaki, A. Environmental change in Crete: a 9000-year record of Holocene vegetation history and the effect of the Santorini eruption. *Holocene***13**, 733–749 (2003).
42. Mazzon, L. et al. Presence of specific symbiotic bacteria in flies of the subfamily Tephritinae (Diptera Tephritidae) and their phylogenetic relationships: proposal of '*Candidatus* Stammerula tephritidis'. *Int. J. Syst. Evol. Microbiol.***58**, 1277–1287 (2008).
43. Capuzzo, C., Firrao, G., Mazzon, L., Squartini, A. & Girolami, V. *Candidatus* ErDacicolacicola', a coevolved symbiotic bacterium of the olive fly *Bactrocera oleae* (Gmelin). *Int. J. Syst. Evol. Microbiol.***55**, 1641–1647 (2005).
44. Petri, L. Ricerche Sopra i Batteri Intestinali della Mosca Olearia. *Mem. Staz.Pat.Veg.* 1–129 (1909).
45. Girolami, V. Reperti morfo-istologici sulle batteriosimbiosi del *Dacus oleae* Gmelin E Di Altri Ditteri Tripetidi, in natura e negli allevamenti su substrati artificiali. *Redia***54**, 269–294 (1973).
46. Stammer, H. J. Die bakteriensymbiose Der Trypetiden (Diptera). *Zoomorphology***15**, 481–523 (1929).
47. Mazzon, L., Martinez-Sanudo, I., Savio, C., Simonato, M. & Squartini, A. Stammerula and other symbiotic bacteria within the fruit flies inhabiting Asteraceae flowerheads in Manipulative Tenants: bacteria Associated with Arthropods, (eds Zchori-Fein, E. & Bourtzis, K.) 89–111, (CRC, 2011).
48. Ohbayashi, T. et al. Insect's intestinal organ for symbiont sorting. *Proc. Natl. Acad. Sci. USA***112**, E5179–E5188 (2015).
49. Zhang, Y. et al. Mitochondrial phylogenomics reveals the evolutionary and biogeographical history of fruit flies (Diptera: Tephritidae). *Entomol. Gen.***43**, 359–368 (2022).
50. da Costa, L. T. et al. The complete mitochondrial genome of *Bactrocera biguttula* (Bezzi) (Diptera: Tephritidae) and phylogenetic relationships with other Dacini. *Int. J. Biol. Macromol.***126**, 130–140 (2019).
51. Besnard, G., Rubio de Casas, R. & Vargas, P. Plastid and nuclear DNA polymorphism reveals historical processes of isolation and reticulation in the olive tree complex (*Olea europaea*). *J. Biogeogr.***34**, 736–752 (2007).
52. Kouroutzoglou, J., Flocas, H. A., Hatzaki, M., Keay, K. & Simmonds, I. On the dynamics of Mediterranean explosive cyclogenesis in Advances in Meteorology, Climatology and Atmospheric Physics (eds Helmis, C. & Nastos, P.) 563–570 (Springer, (2013).
53. Palmano, S., Firrao, G. & Locci, R. Sequence analysis of domains III and IV of the 23S rRNA gene of verticillate streptomycetes. *Int. J. Syst. Evol. Microbiol.***50**, 1187–1191 (2000).
54. Tamura, K., Stecher, G. & Kumar, S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.***38**, 3022–3027 (2021).
55. Weisburg, W. G., Barns, S. M., Pelletier, D. A. & Lane, D. J. 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol.***173**, 697–703 (1991).
56. Osborn, A. M., Moore, E. R. & &Timmis, K. N. An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics. *Environ. Microbiol.***2**, 39–50 (2000).
57. Dupanloup, I., Schneider, S. & Excoffier, L. A simulated annealing approach to define the genetic structure of populations. *Mol. Ecol.***11**, 2571–2581 (2002).
58. Simon, C. et al. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.***87**, 651–701 (1994).
59. Simon, C., Buckley, T. R., Frati, F., Stewart, J. B. & Beckenbach, A. T. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annu. Rev. Ecol. Evol. Syst.***37**, 545–579 (2006).
60. Excoffier, L. & Schneider, S. LavalG. ARLEQIN (ver. 3.0): An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1, 47–50 (2007).
61. Excoffier, L., Smouse, P. E. & Quattro, J. M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics***131**, 479–491 (1992).
62. Clement, M., Posada, D. & Crandall, K. A. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.***9**, 1657–1659 (2000).
63. Crandall, K. A. & Templeton, A. R. Empirical tests of some predictions from coalescent theory with applications to intraspecific phytogeny reconstruction. *Genetics***134**, 959–969 (1993).
64. Dáttilo, W. Different tolerances of symbiotic and nonsymbiotic ant-plant networks to species extinctions. *Netw. Biol.***2**, 127 (2012).
65. Dormann, C. F. & Strauss, R. A method for detecting modules in quantitative bipartite networks. *Methods Ecol. Evol.***5**, 90–98 (2014).
66. Dormann, C. F., Gruber, B. & Fründ, J. Introducing the bipartite package: analysing ecological networks. *Interaction***1**, 8–11 (2008).

Acknowledgements

Special thanks to Professor Vincenzo Girolami for his intuition, crucial suggestions, and unwavering support during the research on the olive fruit fly symbiont. We would also like to thank C. Savio and E. Viale for assistance with laboratory analysis and H. Mollashahi for support with Iranian samples. Moreover, we thank all colleagues who helped us with the sampling collection: Y. Ben-Dov, M. Borin, (A) Bozzolo, V. Caleca, E. Cargnus, S. Deliperi, N. Mori, C. Duso, M.M. El Idrissi, I. Elmaghraby, M. Faccoli, P. Fontana, D. Fornasiero, G. Frigimelica, H. Genç, J. Giliomee, C. Manoli, M. Marzaro, (B) Massa, H. Mladi, A. Mojdehi, F. Mraichi, M.E. Noce, G. Pellizzari, E. Petrucco Toffolo, (C) Rapisarda, M.A. Rashid, C. Rempoulakis, P. Salvato, C. Sañudo, C. Sartori, V. Vacante, C. Villari, L. Vivan, Shahsavari and A. Arjomandinejad. This work was partially supported by the Italian national project 'Decoupling a long-term relationship between the olive fly and its symbiont for pest control (EndOFsymb)' PRIN2022DM104.23_01, funded by the Italian Ministry for Education, University and Research.

Author contributions

I.M.S. and Lu.M. planned the research and I.M.S. and I.C. performed the research. I.M.S., M.A.P., G.S. and Lu.M. analysed the data. I.M.S., M.A.P., La.M. and Lu.M. wrote the paper. All authors read and approved the final manuscript.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-73055-x.

Correspondence and requests for materials should be addressed to I.M.

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

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