Journal of Pest Science (2025) 98:2383–2402
https://doi.org/10.1007/s10340-025-01945-8

# Assessment of *Trioza erytreae* microbiome and mitochondrial genome variability by integrated high-throughput sequencing approach

**Michela Chiumenti[1] · Vittorio Nicoloso[1] · Alberto Fereres[2] · José Alberto Pereira[3] · Hans J. Maree[4,5] · Rachelle Bester[4,5] · Bernard Reynaud[6,7] · Hélène Delatte[6] · Leandro Peña[8] · Vicente Pallás[8] · Pedro Serra[8] · Beatriz Navarro[1] · Francesco Di Serio[1]**

Received: 13 January 2025 / Revised: 16 June 2025 / Accepted: 9 July 2025 / Published online: 5 August 2025


## Abstract

Huanglongbing, a devastating citrus disease, is associated with '*Candidatus* Liberibacter asiaticus', '*Ca.* L. africanus' or '*Ca* L. americanus', bacteria transmitted by the psylloids *Diaphorina citri* and *Trioza erytreae*. Using a DNA-Seq and metabarcode sequencing integrated approach, the first catalogue of endosymbionts associated with *T. erytreae* from the Iberian Peninsula, South Africa and African Islands, was generated. The almost complete genome of two new bacteria, one facultative and one obligate, tentatively named Asaia-like endosymbiont of *T. erytreae* and Sodalis-like endosymbiont of *T. erytreae,* respectively, was assembled and annotated. The complete mitochondrial genomes of *T. erytreae* from the geographical areas studied were also assembled and phylogenetic analyses were performed, suggesting that *T. erytreae* populations currently present in the Iberian Peninsula and specimens analyzed from South Africa may have originated from a common ancestor. Similar results were obtained when the genetic distances between Sodalis-like endosymbiont of *T. erytreae* were taken into consideration, thus supporting the symbiont–host codivergence which suggests that this bacterium is approaching to an obligate status. Finally, a new genetic marker of *T. erytreae,* an insertion in the mitochondrial tRNA-Ser gene, was identified only in some European samples, showing for the first time the existence of two mixed subpopulations of *T. erytreae*. The integrated DNA-Seq and metabarcode sequencing approach used in this study, besides generating a catalogue of *T. erytreae* endosymbionts, provided novel data on the sequence variability of bacterial and insect mitochondrial genomes from different geographic areas, highlighting the possible original sources of currently spreading *T. erytreae* populations may be more complex than previously reported.

**Keywords** African citrus triozid · Endosymbiont · Phylogeography · DNA-Seq · 16S rRNAs · Haplotype

## Introduction

In arthropods, bacterial endosymbionts strongly interfere with the biology, ecology and evolution of their hosts (Moran et al. 2008). They can be vertically transmitted

Communicated by Antonio Biondi.

✉ Beatriz Navarro
beatriz.navarro@cnr.it

✉ Francesco Di Serio
francesco.diserio@cnr.it

[1] Institute for Sustainable Plant Protection, National Research Council, Via Amendola, 122/D, 70126 Bari, Italy

[2] Spanish National Research Council, Madrid, Spain

[3] CIMO, SusTEC, Instituto Politécnico de Bragança, Campus Sta Apolónia, 5300-253 Bragança, Portugal

[4] Department of Genetics, Stellenbosch University, Stellenbosch, South Africa

[5] Citrus Research International, Stellenbosch, South Africa

[6] Université de La Réunion, UMR PVBMT, 97410 Saint Pierre, La Réunion, France

[7] CIRAD, UMR PVBMT, 97410 St Pierre, La Réunion, France

[8] Institute of Molecular and Cellular Biology of Plants, Polytechnic University of Valencia-Spanish National Research Council (CSIC), Valencia, Spain

to the host progeny, being in this case defined as primary (P-) endosymbionts (Balmand et al. 2013; Baumann 2005; Rosell et al. 2009). P-endosymbionts may supplement the unbalanced diets of phloem-feeding insects, such as hemipterans, and may largely depend on the host for their own survival, thus forming obligate associations (Moran and Telang 1998). On the contrary, secondary (S-) endosymbionts are facultative, erratically distributed, not required for host reproduction and can be transmitted horizontally (Chiel et al. 2009; Moran et al. 2008).

While relationships between obligate endosymbionts and their hosts are considered to be ancient and stable, resulting from a strict co-diversification with their hosts, the association of facultative endosymbionts with their insect hosts are generally considered more recent and generally less stable, although they can largely affect the host fitness (Douglas 2007; Moran et al. 2008; Thao et al. 2000a). Endosymbionts may also modify the immune response of the insect or its interactions with parasitoids and pathogens, thus suppressing the insect vector competence (reviewed by Douglas 2016). Moreover, approaches based on the transfer of genetically manipulated endosymbionts into the insect of interest (paratransgenesis) have also been validated for the control of insect-vectored human, animal and plant diseases (Whitten et al. 2016; Ratcliffe et al. 2022). Therefore, interesting perspectives exist for the exploitation of knowledge on insect-associated microbiomes to develop novel control strategies against economically important insects and/or insect-transmitted diseases and to study their phylogeography.

Huanglongbing (HLB), the most devastating disease of citrus, is caused by '*Candidatus* (*Ca*.) Liberibacter asiaticus', '*Ca.* L. africanus' and '*Ca* L. americanus'. These bacteria are transmitted by citrus psylloids, including *Diaphorina citri* and *Trioza erytreae*. The African citrus triozid *T. erytreae* has been reported from Southern, Central and Northern-Eastern Africa, the Arabian Peninsula, Madeira, Canary Islands (EPPO 2024) and, since the last decade, the Iberian Peninsula (EPPO Reporting Service, 2015a; 2015b). The pest is currently spreading in both Spain and Portugal (Arenas-Arenas et al. 2018; 2019; Benhadi-Marín et al. 2022), thus posing a major threat to citrus production in Europe. Understanding the phylogenetic relationship of these *T. erytreae* populations is a key step to infer possible entry and spread in the EU, thus reducing the risk of introducing the associated HLB causal agents. The origin of *T. erytreae* outbreaks in European countries has been recently assessed using microsatellite markers and COI (cytochrome c oxidase subunit 1) gene (Ruíz-Rivero et al. 2021). However, phylogeographic studies based on the complete mitochondrial genome of *T. erytreae* determined by wide-genome sequencing have not yet been performed. This approach may allow deep analyses of intra-population sequence variability and, not being limited to a single mitochondrial gene, may provide additional genetic markers relevant to further dissect the phylogeography of *T. erytreae*. Moreover, whether the insect entered once or multiple times from different locations is still unclear.

Data on *T. erytreae* endosymbionts are currently limited to two studies, one on insect populations from Kenya and on one insect from Cameroon, both based on sequencing of PCR amplicons obtained using general eubacterial and species-specific 16S rRNA primers (Rasowo et al. 2021; Kwak et al. 2021). No information is available on endosymbionts associated with other *T. erytreae* populations in Africa and the Iberian Peninsula. Here, we identified and characterized the endosymbionts associated with *T. erytreae* from several geographical areas where the pest is currently emerging (Spain and Portugal), and from areas where the pest is already established (Madeira, Canary Islands, São Tomé and Príncipe, South Africa and Reunion Island). The study of endosymbionts associated with *T. erytreae* was performed using high-throughput sequencing (HTS) of total DNA (DNA-Seq), an approach that overcame the potential bias generated by primer design in PCR-based and metabarcode studies, allowed the obtention of the full-length sequence of the 16S rRNA and the housekeeping genes of endosymbionts and, in some cases, provided information on the complete genome of the bacteria, partially filling the gap of knowledge regarding endosymbionts of psylloids in this respect (Mauck et al. 2024). This DNA-Seq unbiased approach also unveiled a previously undetected sequence variability of the insect mitochondrial genome from different geographic areas. Taken together, these data provided interesting information on the phylogenetic relationship between *T. erytreae* populations currently present in the Iberian Peninsula and Africa.

## Materials and methods

### Sample collection, DNA preparations and preparations of libraries

A total of 23 locations in Portugal mainland (northern and central coastal regions), Spain mainland (coastal regions of Asturias, Cantabria, Galicia and Pais Vasco), South Africa (provinces of Eastern Cape, Gauteng, Mpumalanga, Western Cape), São Tomé and Príncipe, Reunion (France), Madeira (Portugal) and Tenerife (Spain) islands were sampled (Additional file 7), thus covering almost the complete areas where *T. erytreae* is emerging in Europe and the Iberian Peninsula and other geographic areas where *T. erytreae* is already established since long. Adults (males and females) of *T. erytreae* were collected in citrus orchards, washed twice in PBS, immediately immersed in DNAgard stabilization solution (Sigma-Aldrich), and stored at 4 °C until the DNA was

extracted. DNA preparations were obtained from a pool of five insects from each population under study using DNeasy Blood&Tissue kit (Qiagen), implemented with a pre-incubation for 30 min at 37 °C with lysozyme (12.5 mg/ml in 1 × PBS). Bacterial DNA was further enriched using NEBNext® Microbiome DNA Enrichment Kit (New England Biolabs). When specified, DNA preparations were prepared from individual insect samples following the same protocol.

DNA-sequencing libraries were prepared according to Illumina protocols using the TruSeq Nano DNA LT Library Prep Kit or DNA Nextera XT Library Prep kit. The libraries were then sequenced on an Illumina apparatus in a 2 × 150 bp paired-end format.

## Meta-assembly, binning and annotation

Raw reads were quality checked and processed to remove ambiguities (Qscore < 20), Illumina adapters and sequences having < 50 bp in length using cutadapt v 1.8 (Martin 2011). For each library, quality-filtered reads were assembled using IDBA-UD (Peng et al. 2012) with a multiple kmer approach (from 31 to 111 bp, by 10) obtaining a meta-assembly for each sampled location. The initial annotation of all contigs was obtained by a nucleotide search with a BLAST + v2.9.0 (Camacho et al. 2009) run vs the GenBank nucleotide collection (nt) database (downloaded on 3 January 2021). Contigs from each meta-assembly were used as reference sequences for mapping and reads from each library and generated alignments were converted to bam format using bowtie2 v2.3.5.1 (Langmead and Salzberg 2012). Mapped reads and meta-assemblies were used for an initial binning step using MetaBAT 2 (Kang et al. 2019). Completeness and contamination of bins were evaluated using CheckM v1.0.18 (Parks et al. 2015). Each bin was considered as a metagenomic assembled genome (MAG). Bins with 80% completeness and less than 2% contamination were retained for further analysis. A preliminary taxonomic classification of each MAG was obtained using MIGA (Rodriguez et al. 2018) and Ribosomal Database Project Classifier (RDP; Cole et al. 2014). MAGs were annotated using PROKKA (Seemann 2014). The coding density of MAGS76_1 (Sodalis-like endosymbiont sequenced from the Mafra library) was calculated using CheckM and the pseudogenes were annotated using pseudofinder v3 (Syberg-Olsen et al. 2022). A graphical representation of endosymbiont genomes was generated using CGView v2.0.2 (Stothard et al. 2019).

To identify complete bacterial 16S rRNA sequences from each sequenced library, the assembled contigs were searched using BLAST + against the SILVA LTPs 132 SSU database, which is a quality checked and regularly updated dataset of aligned 16S rRNAs (https://www.arb-silva.de/). To obtain a more accurate taxonomic assignment, such 16S rRNA sequences identified in the libraries were further searched for homologies using the online version of NCBI BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi; Altschul et al. 1990). Additionally, the retrieved 16S rRNA sequences were used as references for a bowtie2 mapping (Langmead and Salzberg 2012), with no mismatches or gaps allowed. These alignments were subsequently used for microbiome profiling using the phyloseq package in R (McMurdie and Holmes 2013).

## 16S rRNA metabarcoding sequencing and analysis

To perform bacterial metabarcoding, DNA extracts of three biological replicates (pools of 5 individuals each) from geographical populations under study were amplified using the primers designed by Klindworth et al. (2013) in the hypervariable V3-V4 region of 16S rRNA gene. Paired-end sequences were quality checked, trimmed, denoised, joined, dereplicated and chimeric sequences filtered-out using dada2 plugin in the QIIME2 software suite (Bolyen et al. 2019). Obtained amplicon sequence variants (ASVs) were taxonomically classified using the SILVA LTPs 132 SSU database. After SILVA taxonomy assignment, a further BLASTn search was performed against the NCBI non-redundant nucleotide database to confirm taxonomic assignments and to exclude the presence of heterologous sequences (e.g., chloroplast and mitochondria). Alpha- and beta-diversity metrics and statistics of the *T. erytreae* microbial population were also performed using QIIME2 tools.

## Endosymbiont intraspecific diversity, pathway analysis and comparative genomics

To study the intraspecific diversity of the Sodalis-like and the Asaia-like endosymbionts, the 16S rRNA and the housekeeping gene sequences (*atpA, gyrB, rpoC, rpoB, rpoD, secA*) were isolated from the Prokka annotation files. The 16S rRNAs and the concatenated housekeeping genes were aligned using the MUSCLE algorithm (Edgar 2004) and a maximum-likelihood phylogenetic tree was generated in MEGAX (Kumar et al. 2018). Metabolic pathway predictions were obtained using gapseq (Zimmermann et al. 2021) and pathway tools v24.5 tier1 (Karp et al. 2021).

A phylogenetic tree of the Asaia-like endosymbiont of *T. erytreae* (Asaia-TE) and its closest type-strain genomes was performed using the Type (Strain) Genome Server (TYGS) platform (https://tygs.dsmz.de). The tree was based on the 16S rRNA gene. The closest type-strain genomes were determined by TYGS using a dual procedure of selection based on the whole genome comparison through the MASH algorithm and the 16S rRNA gene. The tree was inferred with FastME 2.1.6.1 (Lefort et al. 2015) using the Genome BLAST Distance Phylogeny (GBDP) calculated from 16S rRNA gene sequences.

A phylogenetic tree of *Sodalis* sp. and Sodalis-like endosymbionts was generated using IQ-tree software (Nguyen et al. 2015). The 16S rRNA gene of the Sodalis-like endosymbionts of *T. erytreae* (Sodalis-TE) and those of *Sodalis* sp. and Sodalis-like endosymbionts available in public repositories were aligned by MAFFT (Katoh et al. 2002) and used for inferring a tree by maximum-likelihood method (ML) with the best-fit model TVMe + I + G4, obtained by ModelFinder (Kalyaanamoorthy et al. 2017). A total of 3000 bootstrap replicates were performed using ultrafast bootstrap (UFBoot2) (Hoang et al. 2018).

### *T. erytreae* mitochondrial genome characterization

The complete mitochondrial genome of *T. erytreae* was identified by BLASTn searches of the assembled contigs (see above). The intraspecific diversity of the mitochondrial genomes was studied by Bowtie2 alignments. Mitochondrial genomes (excluding the AU-rich putative non-coding control region) and COI sequences were aligned by Clustal Omega (Sievers and Higgins 2014) and phylogenetic trees were inferred by maximum-likelihood method (ML) with the best substitution model identified for each alignment using MEGAX (Kumar et al. 2018).

Single nucleotide polymorphisms (SNPs) in the mitochondrial genomic sequences were found using alignments of reads (Bowtie2) against the mitochondrial genome of *T. erytreae* from South Africa (accession number MT416550) and further investigated by HTS metabarcoding approach. DNA preparations from five pooled insects and from a single insect were extracted and used for amplicon sequencing (2 × 250 bp, paired-end on Illumina MiSeq sequencer) of the fragments generated with the primers Mito-TE-21F_HTS and Mito-TE-22R_HTS (Additional file 1) targeting the tRNA serine (tRNA-Ser) gene. Metabarcoding libraries were QC checked and trimmed using BBduk, overlapping pairs were reconstructed using BBmerge (BBtools; Bushnell 2023) and aligned against the tRNA-Ser sequence obtained from the DNA-Seq libraries.

## Results

### Characterization of the *T. erytreae* microbiome through DNA-Seq and MIGA analysis

To have a wide overview of the microbiome associated with *T. erytreae* populations from the Iberian Peninsula and other geographic areas where *T. erytreae* is already established, samples were collected from different geographical areas, including 13 municipalities in the Atlantic coast of Portugal and the Atlantic and Cantabrian coast of Spain mainland, six municipalities in South Africa, one in each of the Atlantic islands of Madeira (Portugal), Tenerife (Spain) and São Tomé and Príncipe, and one in Reunion, in the Indian Ocean (Fig. 1).

Total reads (Qscore ≥ 20, length ≥ 50) obtained by DNA-Seq of libraries generated from the collected samples ranged from about 29 M to 80 M, with a similar GC content (mostly 40–43%, Additional file 2).

Bins, representing bacterial metagenomic assembled genomes (MAGs), were reconstructed from the respective assembled contigs and their completeness and contamination evaluated. A total of 38 MAGs corresponded to bacterial genomes with more than 80% completeness and less than 2% contamination (Table 1). MAGs were tentatively classified by MIGA annotation and by the Ribosomal Database Project Classifier (RDP), providing a preliminary list of bacteria associated with *T. erytreae* belonging to several families of the order Enterobacteriales (Table 1). Low quality reconstructed genomes with no or unclear taxonomic assignment were discarded from further analysis at this stage (Additional file 3).

### Characterization of the *T. erytreae* microbiome through DNA-Seq and the analyses of complete prokaryotic 16S rRNAs

Since the MAG approach allows the identification only of almost complete bacterial genomes, we used the alternative strategy based on the analysis of the complete prokaryotic 16S rRNA genes to confirm and extend the preliminary microbial profile. To this end, the assembled contigs from each sample were searched against the SILVA database using BLAST and the identified complete prokaryotic 16S rRNAs were filtered-out to perform further analysis (see below). At this stage, however, we noted that neither MIGA nor SILVA annotations allowed the identification of *Ca.* Carsonella ruddii, which is a known vertically transmitted P-endosymbiont of all psylloids, with evidence supporting co-speciation events between the insect hosts and this P-endosymbiont (Hall et al. 2016; Kuechler et al. 2013; Thao et al. 2000b).

Therefore, a further BLAST search, using the contigs from MAGs shown as low-quality genomes by CheckM, identified six Carsonella-like contigs in MAG54_3 from Cobreces, Spain (Additional file 3). PROKKA annotations revealed that one of these contigs corresponded to a complete 16S gene sharing the highest similarity (sequence identity 94.00%, query coverage 100%) with the 16S gene of *Ca.* Carsonella ruddii from *Diaphorina* cf. *continua* (accession AP023214). Moreover, the *Ca.* Carsonella ruddii housekeeping gene RNA polymerase β' subunit (rpoC) was identified by PROKKA annotation in MAG54_3 and confirmed by PCR, cloning and sequencing in one *T. erytreae* sample using specific primers (Additional file 1 and 4). Finally,

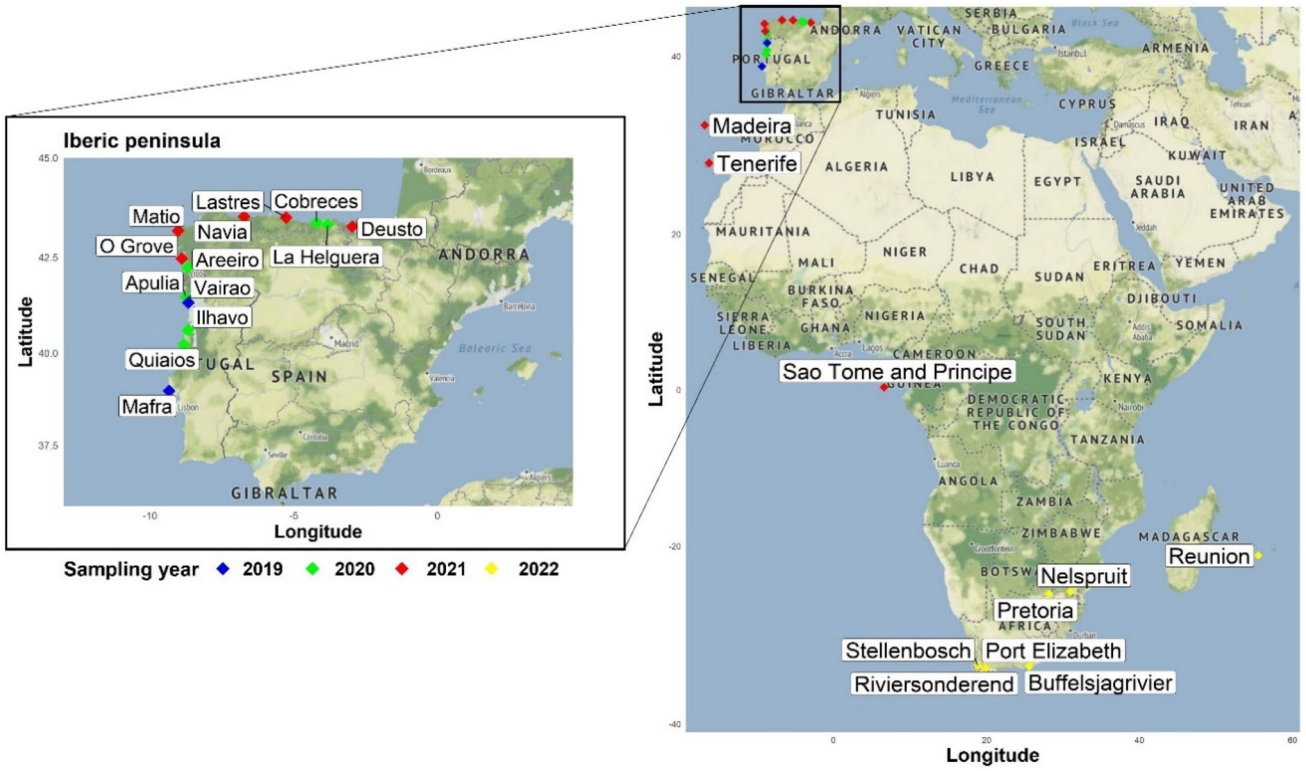



**Fig. 1** Sampling maps. Geographic origin and sampling years of *T. erytreae* samples analyzed in this study

metabarcoding sequence analysis (see below) confirmed the presence of *Ca.* Carsonella ruddii in all the samples from the tested locations. Altogether, these data support the presence of this primary endosymbiont in *T. erytreae*, as expected.

To assess the microbial diversity in each sequenced library, all the 16S rRNA genes identified above were used as reference sequences for performing a bowtie2 mapping followed by phyloseq analysis. This approach identified three to eight different bacterial families associated with *T. erytreae* samples coming from the different areas under investigation (Fig. 2). Interestingly, this 16S-based approach not only confirmed the microbial variability already identified by MIGA, but also highlighted the presence of additional endosymbionts in several samples. In particular, bacteria belonging to three families (Yiersinaceae, Pectobacteriaceae and the unknown family including *Ca.* C. ruddii) were found in all the sequenced samples, thus representing the *T. erytreae* potential core endosymbiont population; those classified in the family Acetobacteriaceae were identified in six samples, while the remaining ones were retrieved from only one or two samples. The highest diversity of taxonomic entities was observed in the samples from Ilhavo and Reunion (Fig. 2). Interestingly, the 16S rRNA gene of bacteria classified in the same family in the sequenced libraries showed identity higher than 99% with each other when compared by multiple sequence alignments, supporting that they belong to the same taxonomic entity regardless the considered library (see also below).

Additional Blastn searches in the NCBI nucleotide collection using the 16S rRNA consensus gene sequences of the potential core endosymbiont populations, showed the highest sequence identity with i) the secondary endosymbiont of *Ctenarytaina eucalypti* (CP003546, 93.92%—query coverage 100%, family Pectobacteriaceae), ii) the secondary endosymbiont of *Trioza magnoliae* (AF077607, sequence identity 88,37%—query coverage 100%, family Yiersinaceae), and iii) *Ca.* Carsonella ruddii from *Diaphorina,* already mentioned above. The first bacterium was classified as a species of the genus Sodalis by RDP (Supplementary Table 2). Considering that the sequence identities with the closest Blast hit reported above are always below 97%, which is the identity threshold between 16S rRNA genes generally accepted to identify new bacterial species, the potential core endosymbiont bacteria are proposed to consist of new species, hereafter named *Ca.* Carsonella ruddii of *T. erytreae* (C. ruddii-TE), Yersinaceae-like endosymbiont of *T. erytreae* (Yersinaceae-TE) and Sodalis-TE (Fig. 2).

In the case of endosymbionts not found in all samples, we focused on one endosymbiont of the family Acetobacteriaceae found in six samples, which shared 97–100% sequence identity of its 16S rRNA gene with *Asaia* spp. and Asaia-like bacteria. Therefore, the tentative name

**Table 1** List of metagenomics assembled genomes (MAGs) reconstructed from *T. erytreae* samples collected in different geographic areas. Only MAGs with completeness > 80% and contamination < 2% are listed. For each sample, origin (country and municipality), number of bins, MAG identity code (MAG id), MAG completeness and contamination (estimated by CheckM), number of contigs included in each bin and MIGA annotation are indicated. MAGs with lower quality (< 80% completeness, > 2% contamination) or with no taxonomic allocation were not listed here; together with the RDP annotation, they are reported in the **Additional file 3**

| Country | Location | #bins (Metabat) | MAG Id | MAG Completeness (CheckM) | MAG Contamination (CheckM) | #contigs in the MAG | MIGA annotation (p-value) |
|---|---|---|---|---|---|---|---|
| *Dem. Rep. of São Tomé and Príncipe* | São Tomé and Príncipe | 3 | MAGSTP_2 | 91.73 | 1.25 | 77 | f_Pectobactericeae (p-value 0.349*) |
| | | | MAGSTP_3 | 91.28 | 1.36 | 39 | c_Gammaproteobacteria (p-value 0.0689**) |
| *France* | Reunion | 5 | MAGReu_3 | 96.3 | 0.03 | 252 | f_Xanthomonadaceae (p-value 0.000995****)<br>g_qG maxbin 005 fasta contigs u10412 (p-value: 0.0063****) |
| | | | MAGReu_5 | 98.75 | 0.0 | 40 | f_Pectobacteriaceae (p-value 0.296*) |
| *Portugal* | Apulia | 4 | MAGApu_1 | 98.75 | 0.0 | 56 | f_Pectobactericeae (p-value 0.284*) |
| | | | MAGApu_3 | 99.15 | 0.25 | 65 | f_Acetobacterieceae (p-value 0.203*) |
| | Ilhavo | 6 | MAGIlv_5 | 97.5 | 0.0 | 2 | f_Pectobactericeae (p-value 0.284*) |
| | | | MAGIlv_6 | 97.34 | 0.25 | 98 | f_Acetobacterieceae (p-value 0.191*) |
| | Madeira | 2 | MAGMad_2 | 98.77 | 0.62 | 23 | f_Pectobactericeae (p-value 0.308*) |
| | Mafra | 2 | MAGS76_1 | 100.0 | 0.0 | 5 | f_Pectobactericeae (p-value 0.296*) |
| | Vairao | 2 | MAGS73_1 | 97.5 | 0.0 | 2 | f_Pectobactericeae (p-value 0.284*) |
| | Quiaios | 3 | MAGQui_1 | 98.75 | 0.0 | 66 | f_Pectobactericeae (p-value 0.284*) |
| | | | MAGQui_2 | 99.4 | 0.0 | 33 | f_Acetobacterieceae (p-value 0.203*) |
| *Spain* | Areeiro | 4 | MAGAre_3 | 100.0 | 0.0 | 111 | f_Pectobactericeae (p-value 0.284*) |
| | | | MAGAre_4 | 81.17 | 0.17 | 249 | f_Acetobacterieceae (p-value 0.182*) |
| | Cobreces | 3 | MAG54_1 | 97.5 | 0.0 | 2 | f_Pectobactericeae (p-value 0.284*) |
| | Deusto | 11 | MAGBil_1 | 98.0 | 0.43 | 165 | f_Enterobacteriaceae (p-value 0.000626****)<br>g_Kosakonia (p-value 0.00364****)<br>s_Kosakonia cowanii (p-value 0,0491***) |
| | | | MAGBil_3 | 99.43 | 0.0 | 50 | f_Moraxellaceae (p-value 0.000616****)<br>g_Acinetobacter (p-value 0.00426****)<br>s_Acinetobacter calcoaceticus (p-value 0.0574**) |

**Table 1** (continued)

| Country | Location | #bins (Metabat) | MAG Id | MAG Completeness (CheckM) | MAG Contamination (CheckM) | #contigs in the MAG | MIGA annotation (p-value) |
|---|---|---|---|---|---|---|---|
| | | | MAGBil_6 | 98.47 | 1.18 | 128 | f_Yersiniaceae (p-value 0.000369****)<br>g_Rahnella (p-value 0.00154****)<br>s_Rahnella victoriana (p-value 0.0181***) |
| | | | MAGBil_7 | 95.89 | 1.1 | 211 | f_ Erwiniaceae (p-value 0.00179****)<br>g_Pantoea (p-value 0.00974****) |
| | | | MAGBil_8 | 100.0 | 0.0 | 24 | f_Pectobacteriaceae (p-value 0.296*) |
| | | | MAGBil_10 | 94.28 | 1.84 | 409 | f_Yersiniaceae (p-value 0.000368****)<br>g_Rouxiella (p-value 0.00153****)<br>s_Rouxiella chamberiensis (p-value 0.0198***) |
| | LaHelguera | 2 | MAG51_1 | 97.5 | 0.0 | 27 | f_Pectobactericeae (p-value 0.296*) |
| | Lastres | 10 | MAGLas_1 | 96.09 | 0.74 | 330 | g_Pantoea (p-value 0.209*) |
| | | | MAGLas_2 | 99.4 | 0.0 | 15 | f_Acetobacteraceae (p-value 0.203*) |
| | | | MAGLas_4 | 98.38 | 0.54 | 81 | f_Morganellaceae (p-value 0.00102****)<br>g_Morganella (p-value 0.00633****) |
| | | | MAGLas_7 | 95 | 0 | 167 | f_Enterobacteriaceae (p-value 0.196*) |
| | | | MAGLas_9 | 99.02 | 0.79 | 74 | f_Hafniaceae (p-value 0.000371****)<br>g_Hafnia (p-value 0.00148****)<br>s_Hafnia alvei (p-value 0.0166***) |
| | Matio | 2 | MAGMat_1 | 98.75 | 0.0 | 78 | f_Pectobacteriaceae (p-value 0.284*) |
| | Navia | 2 | MAGAst_2 | 100 | 0.11 | 11 | f_Pectobactericeae (p-value 0.296*) |
| | O Grove | 2 | MAGGal_1 | 98.75 | 0.0 | 126 | f_Pectobacteriaceae (p-value 0.296*) |
| | Tenerife | 2 | MAGTen_2 | 100.0 | 0.0 | 39 | f_Pectobactericeae (p-value 0.296*) |
| *South Africa* | Buffelsjagrivier | 2 | MAGBJ1_1 | 100.0 | 0.07 | 14 | f_Pectobacteriaceae (p-value 0.296*) |
| | Port Elizabeth | 2 | MAGPE2_1 | 100.0 | 0.07 | 11 | f_Pectobacteriaceae (p-value 0.296*) |
| | Riviersonderend | 4 | MAGRO2_3 | 68.75 | 0.0 | 118 | f_Pectobacteriaceae (p-value 0.198*) |

**Table 1** (continued)

| Country | Location | #bins (Metabat) | MAG Id | MAG Completeness (CheckM) | MAG Contamination (CheckM) | #contigs in the MAG | MIGA annotation (p-value) |
|---|---|---|---|---|---|---|---|
| | Stellenbosch | 2 | MAGSTEL1_1 | 100.0 | 0.0 | 20 | f_Pectobacteriaceae (p-value 0.296*) |
| | Nelspruit | 2 | MAGTIM2_1 | 100.0 | 0.0 | 78 | f_Pectobacteriaceae (p-value 0.296*) |
| | Pretoria | 3 | MAGUP2_1 | 100.0 | 0.0 | 86 | f_Pectobacteriaceae (p-value 0.284*) |

Asaia-TE is proposed for this microorganism. Altogether, these findings show that *T. erytreae* is associated with endosymbionts not previously reported.

## Confirmation of DNA-Seq *T. erytreae* taxonomic profile by 16S metabarcoding sequencing

Nine out of the 23 sampled populations assessed using the DNA-Seq approach were further investigated by metagenomics sequencing of a partial 16S DNA fragment (about 400 bp) amplified using V3-V4 region primers (Klindworth et al. 2013). The taxonomic profile of bacteria from each sample generated by this last approach and subsequent QIIME2 analyses (Additional file 5) is comparable to that previously obtained by the DNA-Seq approach and yielded a total of 388 ASVs. Moreover, the mapping of amplicon sequences reconstructed by QIIME2 onto the contigs from DNA-Seq data allowed us to confirm the presence of 16S rRNA sequences sharing 100% identity with Carsonella ruddii-TE, Yersinaceae-TE, and Sodalis-TE in all libraries, and with the Asaia-TE in the six libraries containing the MAG or the complete 16S rRNA gene of this new tentative species. Therefore, data obtained by DNA-Seq were confirmed by 16S metabarcode sequencing.

Based on sequencing of PCR amplification products obtained with 16S universal eubacterial primers, Rasowo et al. (2021) reported *Rickettsia, Arsenophonus* and *Wolbachia* species as endosymbionts of *T. erytreae* populations from Kenya. The 16S metagenomics-based approach used in our study did not detect *Rickettsia* and *Arsenophonus*. However, a *Wolbachia*-related species was identified in one sample by metabarcode sequencing, although with a low number of reads and in only one of the three biological replicates (Additional file 5). To further assess this issue, the contigs generated from our 23 DNA-Seq libraries were searched by Blast either against sequences corresponding to the *Wolbachia, Rickettsia, Arsenophonus* amplification products reported by Rasowo et al. (2021) or against all the *Wolbachia* complete genomes available in GenBank. The results of these searches were always negative. These data suggest that *T. erytreae* from the locations considered in our survey were not associated with the endosymbionts reported previously, likely because of the different geographical origins of the tested populations.

## Assembly and annotation of the draft genome of Asaia-TE

Five samples originating from different areas of Portugal (Apulia, Ilhavo, Quiaios) and Spain (Areeiro and Lastres) showed one MAG with a completeness of more than 81% and a contamination level lower than 0.25% (Table 1). In three cases, the completeness of such a MAG was higher than 99% (MAGQui_2, MAGApu_3 and MAGLas_2). MAGApu_3 consisted of 65 contigs for a total sequence length of 2,126,524 bp encoding 1,942 proteins and is considered the representative draft genome of the Asaia-TE. Phylogenetic analyses using the complete 16S rRNA from the MAGApu_3 and those from the closest related strains identified by the TYGS, showed that the Asaia-TE clusters with the other members of the family Acetobacteraceae and could represent a new tentative species between *Neoasaia chiangmaiensis* and those included in the Asaia-like clade (Fig. 3). Considering the tree topology, the possibility that this new bacterial species is representative of a new genus should be taken into consideration.

Based on the PROKKA annotation of MAGApu_3, six housekeeping genes, DNA-directed RNA polymerase subunit beta (*rpo*B), DNA-directed RNA polymerase subunit beta' (*rpo*C), RNA polymerase sigma factor (*rpo*D), RNA gyrase subunit B (*gyr*B), ATP synthase subunit alpha (*atp*A) and protein translocase subunit SecA (*sec*A), were identified in the five Asaia-like MAGs. Multiple alignments of concatenated sequences of these genes (for a total of 16,106 bp) from each Asaia-like MAG showed sequence identity of 100% to each other, a result consistent with previous data on the complete 16S rRNA gene, thus further supporting the conclusion that the five *T. erytreae* populations under study contained the same endosymbiont.

To further verify whether this bacterium was absent from *T. erytreae* populations other than the five considered,

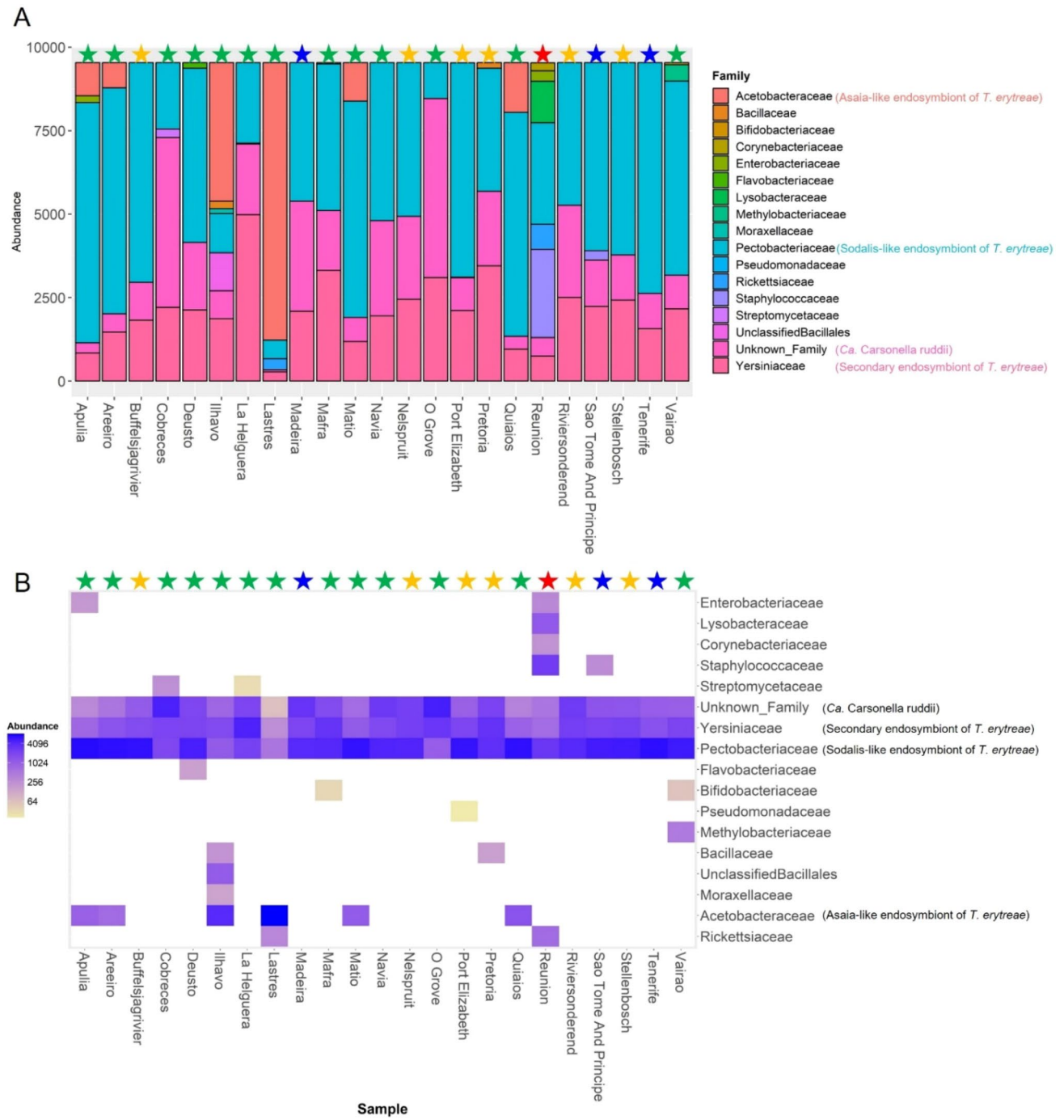



**Fig. 2** Graphical representation of the taxonomic diversity of the endosymbionts found in the DNA-Seq libraries of *T. erytreae* populations under study. Barplot with relative abundances **A** and heatmap **B** are based on the study the 16S rRNA full genes retrieved from the assembled libraries. For each identified microorganism, its taxonomic classification at family level has been reported. The closest BLAST hits of searches using the 16S rRNA full genes as a query are reported on the right for some families. Stars refer to populations sampled in South Africa (yellow), Reunion (red), Iberian Peninsula (green) and Atlantic islands (blue)

the DNA-Seq libraries of the remaining populations studied were further searched for sequences mapping to the Asaia-like housekeeping genes using Bowtie2. No mapping sequences were found, thus confirming the absence of an Asaia-like endosymbiont in these *T. erytreae* populations. Presence and absence of this endosymbiont in representative samples from Portugal and Spain was further confirmed by PCR and sequencing using specific primers targeting the

*rpoC* gene of Asaia-TE (Additional file 1 and 4). Altogether, these findings strongly suggest that this novel bacterial species might be a facultative endosymbiont.

## Assembly and annotation of the draft genome of Sodalis-TE

According to MIGA annotation, one MAG from each of the 23 analyzed samples corresponded to the genome of a member of the family Pectobacteriaceae (genus Sodalis) (Additional file 3). The predicted completeness of these genomes ranged from 91.73 to 100% (contamination levels 0–1.25%, Additional file 3) for all samples, excluding RO2 from South Africa whose genome was incomplete (genome completeness 68.75%, contamination 0%). To identify as complete as possible high-quality bacterial genomes of the Sodalis-like endosymbiont, we focused on MAGs showing a CheckM value corresponding to 100% completeness and no contamination. Among them, the MAG obtained from the sample collected in Mafra (MAGS76_1, Portugal), from which the longest sequence (1,301,002 bp) was assembled with the lowest number of contigs (5 contigs), was selected to be considered as the representative draft genome of Sodalis-TE (Fig. 4A, https://www.ncbi.nlm.nih.gov/genome/) and for further analysis.

The complete 16S rRNAs of Sodalis-TE from samples collected in the Iberian Peninsula (Spain and Portugal), in the Islands of Madeira and Tenerife, and South Africa, are identical to each other and slightly differ from São Tomé and Príncipe and Reunion samples. In a 16S rRNA-based maximum-likelihood (ML) phylogenetic tree including also other Sodalis-related bacteria, all the Sodalis-TE clustered in a single clade separated from all the other Sodalis species, suggesting that the Sodalis-TE could represent a novel endosymbiont species.

However, Sodalis-TE from São Tomé and Príncipe and Reunion islands segregated into two different subclades with respect to those from the other populations under study (Fig. 4B). These data suggest a closer phylogenetic relationship between endosymbionts from European (including Madeira and Tenerife) and South African populations than with insect populations from São Tomé and Príncipe and Reunion (Additional file 6A). The higher genetic distance between the samples from São Tomé and Príncipe and Reunion, and samples from the continental Europe, Madeira, Tenerife and South Africa, was also observed when an ML phylogenetic tree was generated using the complete sequences of six concatenated housekeeping genes (*atpA, gyrB, rpoC, rpoB, rpoD, secA*) of Sodalis-TE. In this case, however, a slightly larger variability was observed among samples from South Africa (Additional file 6B).

The presence of Sodalis-TE was further confirmed by PCR and sequencing of the *rpoC* gene amplification product (Additional file 4) obtained by specific primers (Additional file 1).

The genome of Sodalis-TE has an intermediate size (1,3 Mb), similar to the S-endosymbiont of *Heteropsylla cubana* (1,12 Mb; Sloan and Moran 2012) and the S-endosymbiont of *Ctenarytaina eucalypti* (1,44 Mb, Sloan and Moran 2012), and smaller than those of other endosymbionts species of genus *Sodalis*, such as *Ca*. Sodalis pierantonious str. SOPE (4,5 Mb; Oakeson et al. 2014) and *S. glossinidius* (4,1 Mbp; Toh et al. 2006), which are maternally transmitted endosymbionts. These Sodalis-related bacteria are considered to be in a middle-stage of reductive evolution toward the primary symbiotic status (Santos-Garcia et al. 2017; Wilson et al. 2010). Interestingly, according to Pathway Tools (v24.5 tier1) and gapseq analysis, the genome of Sodalis-TE showed typical signatures of obligate endosymbionts (Santos-Garcia et al. 2017), such as AT enrichment (38,7% GC content), low coding density (51.4%), a reduced number of coding genes (811), a general reduction in the number of metabolic pathways, enzymatic and transport reactions and encoded polypeptides and enzymes with respect to *Ca*. Sodalis pierantonious str. SOPE and *S. glossinidius* (Table 2). Notably, the number of pseudogenes identified in the genome of Sodalis-TE (53) is more than twice the number of pseudogenes found in the S-endosymbiont of *Ctenarytaina eucalypti* and the S-endosymbiont of *Heteropsylla cubana,* which have already been reported to be endosymbionts with genomes of similar size (Morrow et al. 2017).

Another signature of reductive evolution toward obligate endosymbiosis is the decreased number of rRNA gene copies within the prokaryotic genome (Oakeson et al. 2014; Toh et al. 2006). Interestingly, in contrast to *Sodalis*-related bacteria with a larger genome, a single copy of 23S, 16S and 5S rRNA genes was present in the genome of Sodalis-TE. Moreover, in the Sodalis-TE, the 5S rRNA gene is unlinked (about 42,8 kb far) from the 16S-23S operon (Fig. 4A), which is an additional signature of bacteria with a closer relationship with their host (Ahn et al. 2020).

Altogether, these molecular features and the confirmed presence of Sodalis-TE in all insect populations (see above), strongly suggest that it is a long-term endosymbiont (Mauck et al. 2024) moving toward an obligate status.

## Phylogenetic analysis of *T. erytreae* populations based on mitochondrial genome

Taking advantage of the DNA-Seq data generated in this study, the complete mitochondrial genome of the *T. erytreae* populations from the different geographic areas under consideration was assembled, thus phylogenetic studies were performed using both the mitochondrial gene Cytochrome c Oxidase I (COI) and the complete mitochondrial genome. Phylogenetic analysis based on the mitochondrial COI gene

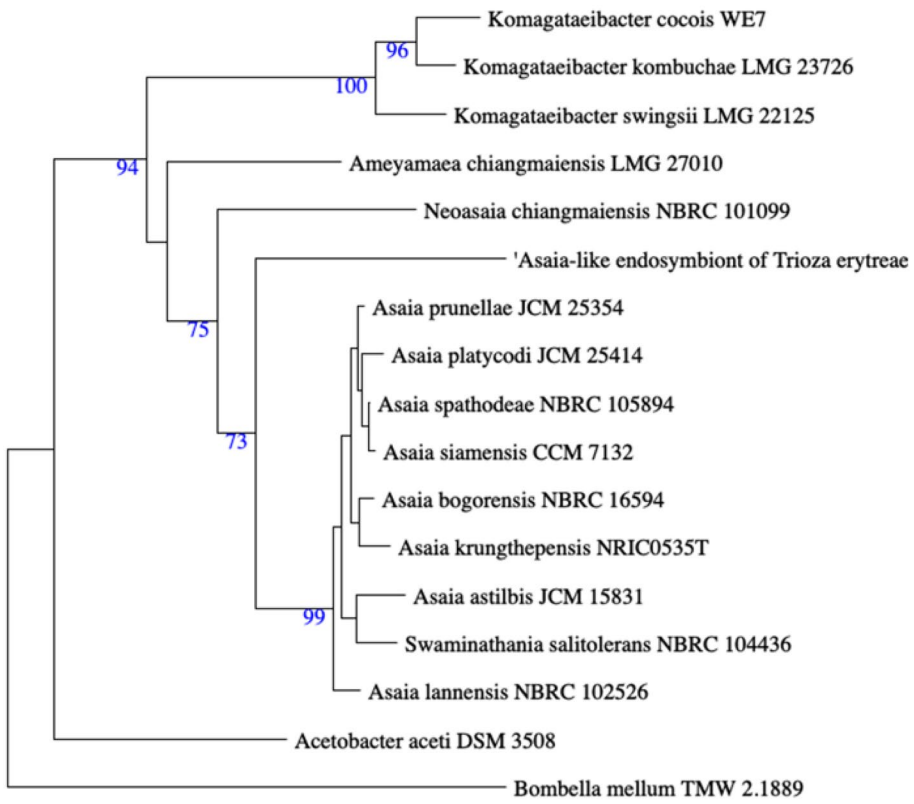



**Fig. 3** Phylogenetic tree based on the 16S rRNA gene of Asaia-like endosymbionts. The tree was inferred, with FastME 2.1.6.1 (Lefort et al. 2015) in the Type (Strain) Genome Server (TYGS) platform, using the Genome BLAST Distance Phylogeny (GBDP) of 16S rRNA gene sequences from the closest type-strain genomes. The branch lengths were scaled in terms of GBDP distance formula d5. The numbers above branches are GBDP pseudo-bootstrap support values > 60% from 100 replications, with an average branch support of 58.6%. The tree was rooted at the midpoint and visualized with PhyD3

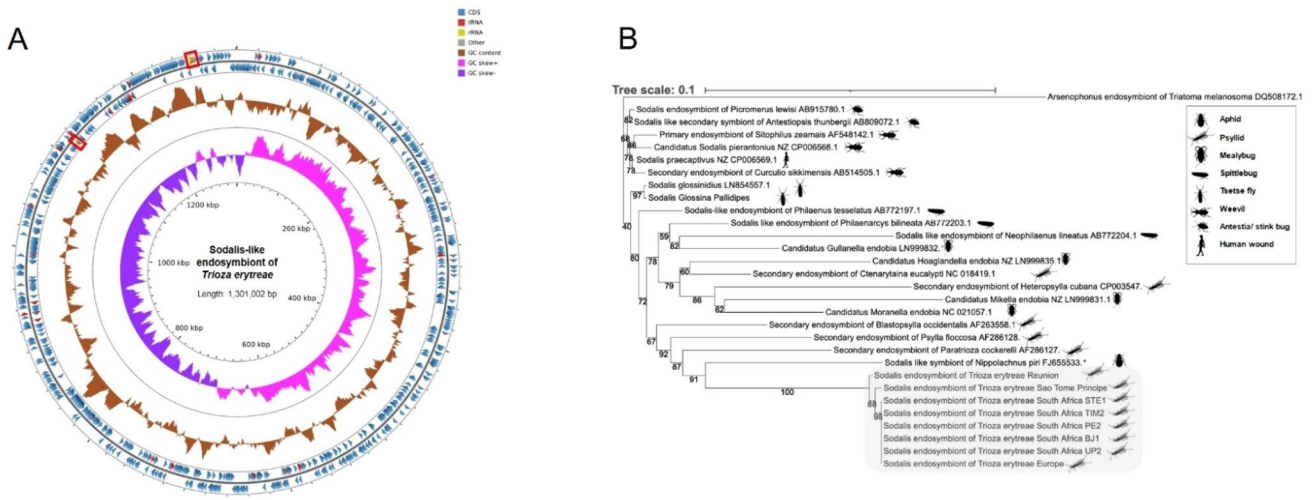



**Fig. 4** **A** Graphical representation of the draft genome of Sodalis-TE. The representation obtained by cgview contains all the genes annotated by PROKKA, including CDS, tRNAs, rRNAs (red boxes) and other genes. **B** Phylogenetic tree generated using the 16S rRNA gene of *Sodalis* sp. and Sodalis-like endosymbionts. The maximum-likelihood tree was inferred in IQ-Tree adopting the best-fit model TVMe + I + G4. Bootstrap probability values (3000 replicates) higher than 50% are shown at the branch nodes. The Sodalis-TE identified in this study clustered together and are denoted by a gray background. Name, accession number and host (shown by an icon) of each endosymbiont are indicated at each branch tip. *Arsenophonus* endosymbiont of *Triatoma melanosoma* was used as an outgroup

showed that all the samples from Iberian Peninsula and the islands of Tenerife (Spain) and Madeira (Portugal) and one from Stellenbosch (South Africa), clustered in a single clade, together with other *T. erytreae* populations characterized previously from Madeira, Tenerife and South Africa that were included in the haplotype 4 (Fig. 5) (Ajene et al. 2020). The other *T. erytreae* samples from South Africa reported in this study clustered within the previously established

**Table 2** Comparative data on the Sodalis-TE and closely related species. Underlined headers in the first column indicate general information on the genomes, simple italics headers introduce detailed data on pathways, reactions, polypeptides and other cellular components. Genome erosion and endosymbiosis were already reported for *Ca.* Sodalis pierantonius str. SOPE, *Sodalis glossinidius* str. 'morsitans', Secondary endosymbiont of *Ctenarytaina eucalypti,* Secondary endosymbiont of *Heteropsylla cubana*

| | *Candidatus* Sodalis pierantonius str. SOPE (NZ_CP006568) | Sodalis glossinidius str. 'morsitans' (NZ_LN854557) | Sodalis endosymbiont of *T. erytreae* | Secondary endosymbiont of *Ctenarytaina eucalypti* (NC018419) | Secondary endosymbiont of *Heteropsylla cubana* (NC018420) |
|---|---|---|---|---|---|
| *Genome size (bp)* | 4,513,140 | 4,171,874 | 1,301,002 | 1,441,139 | 1,121,596 |
| *Tot genes* | 4643 | 4836 | 847 | 779 | 619 |
| *Protein genes* | 3240 | 3593 | 811 | 713 | 552 |
| *RNA genes* | 61 | 77 | 36 | 45 | 44 |
| *Pseudogenes* | 1342 | 1166 | 53 | 21 | 23 |
| *Pathways* | 207 | 196 | 98 | 122 | 82 |
| *Enzymatic reactions* | 1111 | 1107 | 586 | 772 | 628 |
| *Transport reactions* | 11 | 14 | 2 | 11 | 8 |
| *Polypeptides* | 3240 | 3593 | 811 | 713 | 552 |
| *Enzymes* | 648 | 709 | 335 | 385 | 294 |
| *Transporters* | 19 | 19 | 9 | 17 | 5 |
| *Compounds* | 778 | 800 | 472 | 597 | 456 |
| *tRNAs* | 55 | 71 | 36 | 40 | 38 |

haplotypes 3 and 5 of *T. erytreae*, both of which contained sequences from South African specimens characterized previously (Fig. 5) (Ajene et al. 2020). Therefore, our data on the COI gene confirmed the *T. erytreae* variability reported previously in South Africa (Ajene et al. 2020) and agree with the variability recently reported by Ruíz-Rivero et al. (2021) in *T. erytreae* populations from mainland Spain and Portugal, and Madeira and the Canary Islands. In addition, in our analysis, the isolate from Reunion was assigned to haplotype 2, while the one from São Tomé and Príncipe did not cluster with any previously reported haplotype (Fig. 5), thus supporting a genetic divergence of these populations from the insects collected in Europe, South Africa and Madeira and Canary Islands.

When the complete consensus mitochondrial genomes of *T. erytreae* currently reported in literature and those determined in our study were considered to infer an ML phylogenetic tree, again, all the insect populations from Spain, Portugal, Madeira and Tenerife clustered together with those from South Africa and separately from all the other ones (Fig. 6), supporting a closer relationship between the European and South African populations. Interestingly, the consensus mitochondrial genome of the *T. erytreae* population from Stellenbosch showed only six single nucleotide polymorphic (SNP) positions with respect to samples from mainland Spain and Portugal, and from Madeira and Tenerife that were almost identical to each other. Remarkably, the topology of the phylogenetic tree based on the insect complete mitochondrial genome paralleled that observed in the ML phylogenetic tree generated using the six housekeeping genes of the Sodalis-TE populations (Fig. 6), thus strongly supporting co-speciation event between *T. erytreae* and the associated Sodalis-like endosymbionts.

## Intra-population sequence variability of the *T. erytreae* mitochondrial genome

The availability of DNA-Seq data also allowed further investigation of intra-population sequence variability in the *T. erytreae* mitochondrial genome. When the reads from each library were mapped to the respective complete mitochondrial genome, a variability consisting of the presence/absence of an AU indel in the T-loop of the secondary structure of the tRNA-Ser gene (Fig. 7A) was identified in most *T. erytreae* populations from the Iberian Peninsula and Madeira (Table 3). The sequence variants lacking the AU were identical to those previously reported in the consensus mitochondrial genome of the South African haplotype 5 (accession number MT416550); the other ones, containing the AU insertion in the T-loop of the tRNA-Ser, were never reported before (Fig. 7A). In the population from Tenerife used to generate the DNA-Seq library we found the AU insertion in 100% of the reads. However, since this population corresponds to a laboratory colony, it is not representative of the situation in nature. In fact, when DNA preparations from five *T. erytreae* specimens collected in a field of Tenerife were tested by PCR with tRNA-Ser specific primers, seven out of the nine cloned amplicons contained the AU insertion, indicating the existence of two different tRNA-Ser genes also in natural *T. erytreae* populations of Tenerife. Such a

variability was not observed in any of the *T. erytreae* populations from South Africa, where only tRNA-Ser identical to those of haplotype 5 were found. Similarly, no variability was also observed in *T. erytreae* populations from Reunion and São Tomé and Príncipe, which had a tRNA-Ser with a T-loop characterized by deletions with respect to the South African sequence (Fig. 7A). Moreover, the AU insertion in the T-loop observed in most *T. erytreae* from the Iberian Peninsula, Tenerife and Madeira was not reported in any other annotated mitochondrial genome sequence of *T. erytreae* (from Uganda and Ethiopia) so far (Fig. 7A).

Since our study was based on DNA preparations from five insects, at this stage we could not establish whether the detected sequence variability in tRNA-Ser derived from the pooling process of different specimens before the DNA extraction, or from the occurrence of heteroplasmy (simultaneous presence of two or more mitochondrial DNA haplotypes in a single individual) in *T. erytreae.* To clarify this point, PCR-metabarcoding using specific primers targeting the tRNA-Ser gene and DNA preparations from single specimens and from five pooled specimens, followed by HTS of the respective amplicons was performed. To this aim, populations from Madeira, Mafra, Quiaios, Vairao, plus a new population from Maceda (not studied previously) were used. These analyses showed that DNA preparations from single insects were always associated with reads corresponding to a single tRNA-Ser, with or without the AU insertion (Fig. 7B). In contrast, in the pooled samples, reads of tRNA-Ser with and without AU insertion were detected, confirming that the tRNA-Ser variability observed in the *T. erytreae* populations is due to the pooling process and not to a heteroplasmy phenomenon. This finding is consistent with the observation that *T. erytreae* populations currently present in Spain, Portugal and Madeira contain specimens with a mitochondrial genome differing at least in the sequence of the tRNA-Ser. To further explore whether the insertion of AU in the mitochondrial tRNA-Ser could be present in other South African populations of *T. erytreae*, 34 additional single insects from three different areas of this country (13 specimens from the Stellenbosch district, 15 from the Swellendam district, and 6 specimens from the Cape Town metropole) were collected to perform amplification with specific primers (Mito-tRNAser_TE_For/Mito-tRNAser_TE_Rev, Additional file 1) of the mitochondrial tRNA-Ser gene. Sanger sequencing of the PCR amplicon confirmed the absence of the AU insertion in the tRNA-Ser gene of all the South African *T. erytreae* specimens analyzed.

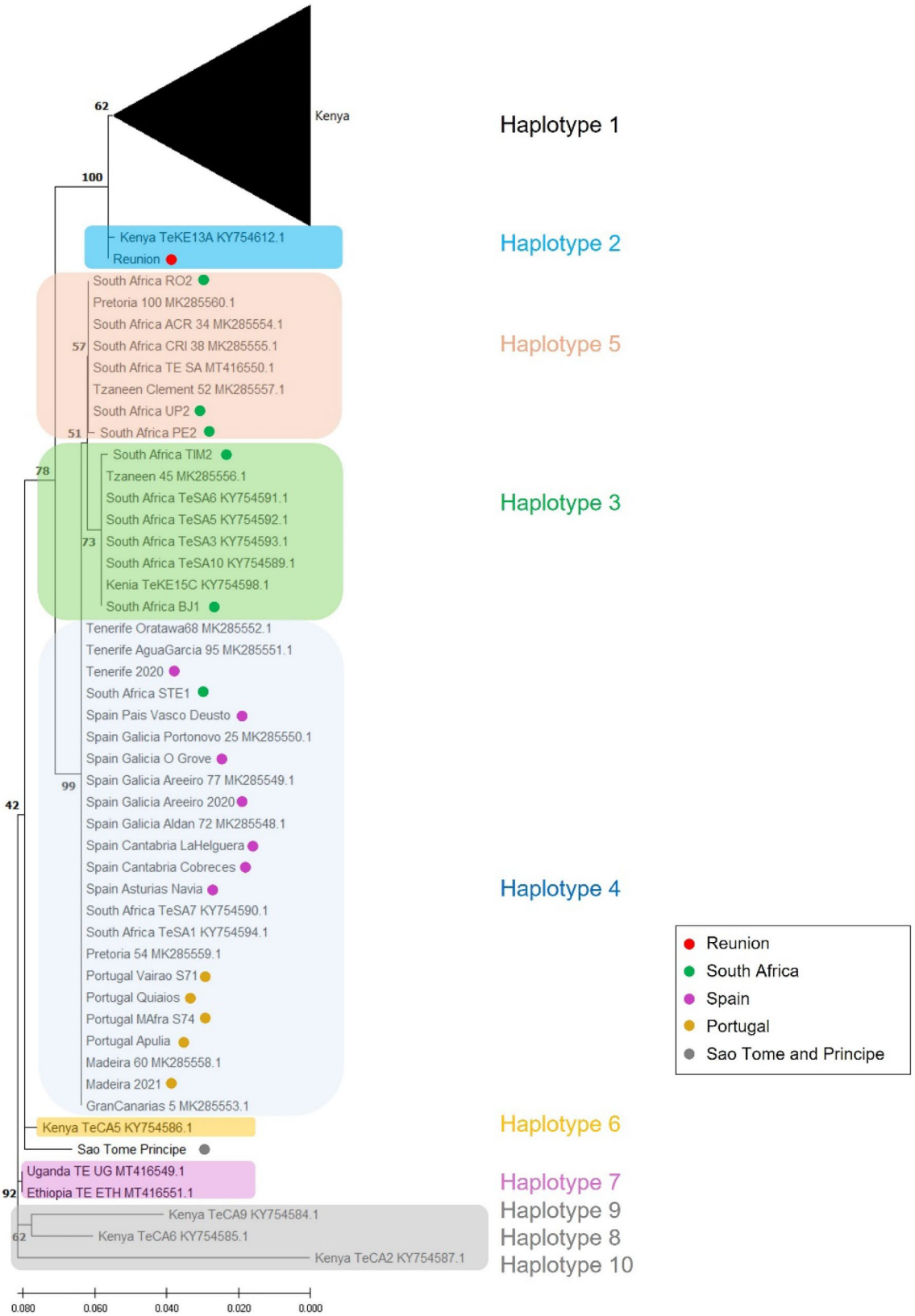
Taking this into account, while the possible original source of the European specimens without the AU insertion in the tRNA-Ser gene could have been the same as the specimens from South Africa, the origin of specimens showing the AU insertion in the T-loop of tRNA-Ser remains undetermined.

**Table 3** Assessment of sequence variability of mitochondrial tRNA-Ser gene in *T. erytreae*. Percentage of reads aligned to the South Africa mitochondrial genome (acc. nr. MT416550) showing the AU insertion in the tRNA-Ser gene for each sequenced library collected in the geographic areas under study. The value 0.0 (in bold) indicates absence of reads with AU insertion in the tRNA-Ser gene

| Geographic origin | Sample | % reads with AU insertion |
|---|---|---|
| *Europe* | Vairao | 54.7 |
| | Mafra | 63.4 |
| | Areeiro | 50.0 |
| | Apulia | 42.3 |
| | Ilhavo | 12.0 |
| | Quiaios | 32.5 |
| | O' Grove | **0.0** |
| | Matio | **0.0** |
| | Navia | **0.0** |
| | Lastres | 46.0 |
| | La Helguera | 31.1 |
| | Cobreces | 18.1 |
| | Deusto | 2.6 |
| | Tenerife (laboratory colony) | 100 |
| | Madeira | 64.7 |
| *South Africa* | Port Elizabeth | **0.0** |
| | Pretoria | **0.0** |
| | Buffelsjagrivier | **0.0** |
| | Nelspruit | **0.0** |
| | Riviersonderend | **0.0** |
| | Stellenbosch | **0.0** |
| | MT416550 (Nelspruit) | |
| *Dem. Rep. of São Tomé and Príncipe* | São Tomé and Príncipe | **0.0** |
| *Reunion* | Reunion | **0.0** |
| *Uganda* | MT416549 | |
| *Ethiopia* | MT416551 | |

## Discussion

*'Ca.* Liberibacter asiaticus', '*Ca* L. americanus' and '*Ca* L. africanus', the casual agents of the citrus devastating disease HLB, are bacteria transmitted by the citrus psylloids *D. citri* and *T. erytreae*. Since 2014, *T. erytreae* has been recorded in several geographical areas of Spain and Portugal mainland and is currently spreading in the Iberian Peninsula. The possible origin of these outbreaks has been proposed to be from South Africa based on microsatellites and mitochondrial DNA barcoding approaches in a previous study (Ruíz-Rivero et al. 2021). However, whether the insect entered once or multiple times in continental Europe is unknown. Moreover, only limited knowledge on endosymbionts possibly associated with *T. erytreae* is available (Rasowo et al. 2021; Kwak et al. 2021), with most information obtained by using


Kenya
Haplotype 1
Kenya TeKE13A KY754612.1
Reunion
Haplotype 2
South Africa RO2
Pretoria 100 MK285560.1
South Africa ACR 34 MK285554.1
South Africa CRI 38 MK285555.1
South Africa TE SA MT416550.1
Tzaneen Clement 52 MK285557.1
South Africa UP2
South Africa PE2
Haplotype 5
South Africa TIM2
Tzaneen 45 MK285556.1
South Africa TeSA6 KY754591.1
South Africa TeSA5 KY754592.1
South Africa TeSA3 KY754593.1
South Africa TeSA10 KY754589.1
Kenia TeKE15C KY754598.1
South Africa BJ1
Haplotype 3
Tenerife Oratawa68 MK285552.1
Tenerife AguaGarcia 95 MK285551.1
Tenerife 2020
South Africa STE1
Spain Pais Vasco Deusto
Spain Galicia Portonovo 25 MK285550.1
Spain Galicia O Grove
Spain Galicia Areeiro 77 MK285549.1
Spain Galicia Areeiro 2020
Spain Galicia Aldan 72 MK285548.1
Spain Cantabria LaHelguera
Spain Cantabria Cobreces
Spain Asturias Navia
South Africa TeSA7 KY754590.1
South Africa TeSA1 KY754594.1
Pretoria 54 MK285559.1
Portugal Vairao S71
Portugal Quiaios
Portugal MAfra S74
Portugal Apulia
Madeira 60 MK285558.1
Madeira 2021
GranCanarias 5 MK285553.1
Haplotype 4
Kenya TeCA5 KY754586.1
Haplotype 6
Sao Tome Principe
Uganda TE UG MT416549.1
Ethiopia TE ETH MT416551.1
Haplotype 7
Kenya TeCA9 KY754584.1
Kenya TeCA6 KY754585.1
Kenya TeCA2 KY754587.1
Haplotype 9
Haplotype 8
Haplotype 10
62
100
57
51
78
73
99
42
92
62
0.080
0.060
0.040
0.020
0.000
Reunion
South Africa
Spain
Portugal
Sao Tome and Principe

◂**Fig. 5** Phylogenetic analyses of *T. erytreae* haplotypes based on Cytochrome c Oxidase subunit 1 (COI). Maximum-likelihood (ML) tree generated using an alignment (T92 + G as the best substitution model) of COI sequence fragments (571 bp) of *T. erytreae* available in databases and obtained in this study. Nodal support was calculated using 1,000 bootstrap replicates. The length of the branches is proportional to the number of substitutions per site. The different haplotypes are numbered according to Ajene et al. (2020) and are denoted with a color background. The circles on the right indicate sequences from this study and are colored according to the country of origin of the *T. erytreae* population

amplification methods that may generate information biased by the primers used. Here, we reported the first characterization of the microbiome associated with *T. erytreae* by an unbiased DNA-Seq approach, with results implemented and confirmed by 16S rRNA metabarcoding data. This approach generated a catalogue of endosymbionts, based on full-length 16S rRNA gene, associated with *T. erytreae* populations from different geographical areas of Europe and Africa, which includes three core bacterial endosymbionts: Sodalis-TE, Yersinaceae-TE, and C. ruddii-TE. The latter bacterium has been detected for the first time in *T. erytreae* in this study, a finding in line with previous reports that identified *Ca* C. ruddi as an obligate endosymbiont with relevant biological roles in all psyllids (Thao et al. 2000b; Morrow et al. 2017; Nakabachi et al. 2022; Mauck et al. 2024). The absence or underdetection of *Ca.* C. ruddii in the previously published and our 16S metabarcoding studies performed in *T. erytreae* (Rasowo et al. 2021; Kwak et al. 2021) was likely due to known biases in 16S eubacterial primers that underamplify this P-endosymbiont. In the *T. erytreae* microbiome obtained by Kwak et al (2021), seven OTUs were identified, all belonging to the order Enterobacteriales, and one of them with the best blast matching to Sodalis endosymbiont of *Nezara antennata*, further supporting our results. Moreover, Morrow et al. (2017) identified ubiquitous Sodalis-like bacteria and other Enterobacteriaceae in other psyllid species.

The almost complete genomes of Sodalis-TE and Asaia-TE were assembled and annotated. Phylogenetic analysis based on their complete 16S rRNA and housekeeping genes supports their classification as two new species.

The identification of the Asaia-like bacterium in *T. erytreae* only in some of the tested samples supports its facultative behavior. Other *Asaia* species are part of the gut microbiota of several human and plant disease insect vectors (Favia et al. 2007; Damiani et al. 2010; Chen et al. 2020; De Freece et al. 2014; Li et al. 2019; Cappelli et al. 2019) and are in vitro cultured, transformable and easily horizontally transmitted (Hughes et al. 2014; Rossi et al. 2015; Cappelli et al. 2019). Moreover, natural acquisition of this endosymbiont by ingestion of contaminated nectar has been shown for *Anopheles* mosquitoes (Bassene et al. 2020). Therefore, they are potentially useful to achieve symbiont-based control of insect-vectored pathogens, as already shown by experimental data on *Anopheles* mosquitoes and *Aedes aegypti,* which are relevant vectors of human diseases (Chouaia et al. 2012; Mancini et al. 2020; Crotti et al. 2009; Souza et al. 2019). The finding of Asaia-TE calls for future studies focusing on their role in physiology, development, and vector competence of *T. erytreae*.

Regarding the Sodalis-TE*,* when compared with other Sodalis species, its genome has a low GC content and gene density, reduced numbers of coding genes and metabolic pathways, which are associated with a high number of pseudogenes. Moreover, in this genome, a single copy of 23S, 16S and 5S rRNA genes was identified, with the 5S rRNA gene unlinked (about 42,8 kb far) from the 16S-23S operon. Unlinked rDNA operons are another characteristic of symbiotic relationship of bacteria with their hosts (Ahn et al. 2020) and have been reported in the genome of several bacteria (Lui et al. 2021). However, while in most of them, including several Sodalis species (Toh et al. 2006; Oakeson et al. 2014; Morrow et al. 2017), the 16S rRNA gene is separated from the 23S-5S operon, in Sodalis-TE the 5S gene was found unlinked from 16S-23S rRNA operon, a situation reported before only for some Archea (Garrett and Klenk 2008) and mycoplasms (Stemke et al. 1994). All these features are consistent with reductive evolution toward obligate endosymbiosis (Santos-Garcia et al. 2017; Morrow et al. 2017) and, together with the confirmed presence of Sodalis-TE in all insect populations regardless the geographical origin, strongly support that it is a long-term endosymbiont (Mauck et al. 2024).

In addition to information regarding endosymbionts of *T. erytreae*, this study extended the number of complete mitochondrial sequences of *T. erytreae,* from the four previously available [from Ethiopia, Uganda and South Africa (Ajene et al. 2020) and an unknown location—NC038142], to twenty-three. Nineteen of these mitochondrial sequences were generated in the frame of this study from Europe and Africa. All together these data provide evidence of a common ancestor for the specimens analyzed from South Africa, Spain and Portugal. Notably, the topology of the phylogenetic tree based on the insect mitochondrial genome mirrored the tree generated using the six housekeeping genes of Sodalis-TE*,* supporting the possible use of the bacterial genetic markers to assess the phylogeography of its host insect. A similar situation was reported previously for *Diaphorina citri* and its obligate endosymbiont *Ca.* C. ruddii, sustaining a co-diversification during the migration of the insect from South to East and Southeast Asia (Wang et al. 2018).

The DNA-Seq data generated in our study also allowed deep analyses of the intra-population sequence variability of the *T. erytreae* mitochondrial genomes, revealing a polymorphism (an AU indel) in the T-loop of tRNA-Ser. Such

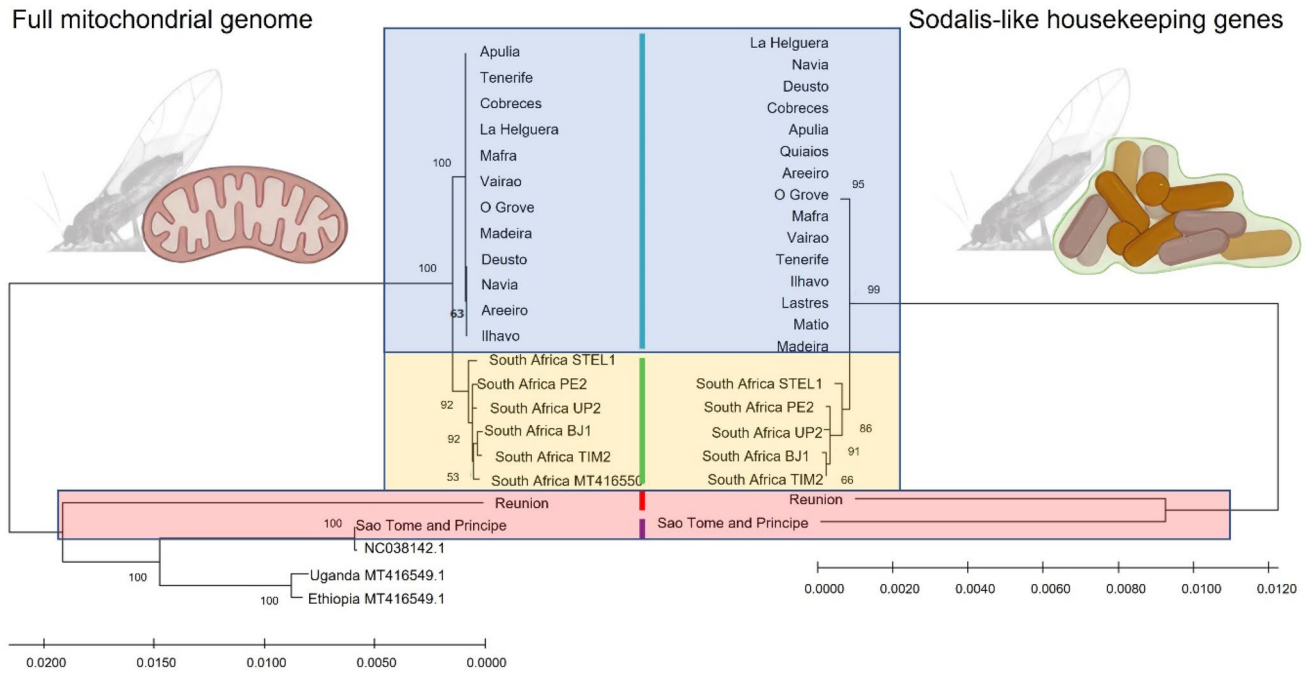



**Fig. 6** Comparative phylogenesis of *T. erytreae.* mitochondrial genome and housekeeping genes of an associated endosymbiont. Maximum-likelihood phylogenetic trees obtained (i) Using full mitochondrial genome of *T. erytreae* from this study and those deposited in databases (on the left, best substitution model HKY + G) and (ii) Using the concatemers of the house keeping genes of Sodalis-like endosymbiont of *T. erytreae* (on the right, best substitution model TN93 + G). The three main clusters identified in both trees are indicated by different colored background. Nodal support was calculated using 1,000 bootstrap replicates. The length of the branches is proportional to the number of substitutions per site. *T. erytreae* picture, source: S.P. van Vuuren, Citrus Research International, Bugwood.org. Mitochondrion and bacteria images created in BioRender.com

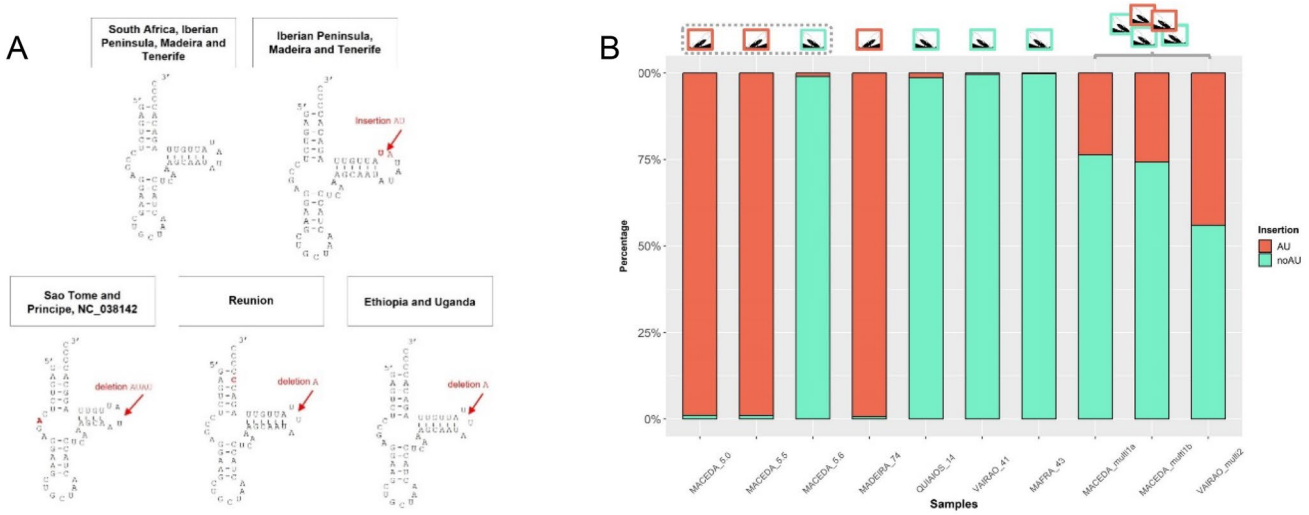



**Fig. 7** **A** Primary and secondary structure of mitochondrial tRNA-Ser of *T. erytreae* from different geographic areas. The changes respect to tRNA-Ser from South Africa (NCBI acc. number MT416550) are indicated in red. **B** Barplot showing the frequency of distribution of indel AU in the tRNA-Ser sequences. Icons on the top of the bars indicate if the amplicon sequencing has been done on single or pooled insects. Orange color indicates the frequency of AU insertion, green color the absence of the AU insertion. Gray dotted box surrounds single insects collected in the same location (Maceda, Portugal). From some locations (Maceda and Vairao) pooled insects were also tested

a polymorphism was present, with different prevalence, in all the studied *T. erytreae* populations from mainland Spain and Madeira. PCR-metabarcoding approach showed that the AU indel was not due to mitochondrial heteroplasmy (Zhang and Hewitt 1997), supporting instead the existence in mainland Spain and Portugal, Tenerife and Madeira, of *T. erytreae* populations composed of a mix of individuals bearing or not such polymorphism. Since the indel was not detected in any African or South African insect populations, the question of the origin of *T. erytreae* individuals bearing the AU insertion in the T-loop of the tRNA-Ser and currently spreading in Europe remains unanswered. The identification of the T-loop of tRNA-Ser rRNA as a possible genetic marker, provides a molecular tool for performing more extensive analyses. By including specimens from *T. erytreae* populations collected from particularly the African continent and the Arabian Peninsula can assist to solve this conundrum.

This study revealed the first catalogue of endosymbionts associated with *T. erytreae* from different geographical areas of Europe and Africa (including Iberian Peninsula, South Africa, Tenerife, Madeira, Reunion and São Tomé and Príncipe). An integrated approach, based on the analyses of DNA-Seq libraries (from *T. erytreae* populations from 23 different geographical areas), 16S rRNA-amplicon (from nine insect populations), and tRNA-Ser amplicon libraries (from 10 insects and/or pool of insects) allowed (i) The identification of a core of associated endosymbionts; (ii) The assembly and annotation of the almost complete genome of one likely long-term endosymbiont (Sodalis-TE) and a facultative (Asaia-TE) bacterial species, thus providing the basic knowledge to further investigate the insect-endosymbionts interactions and their relevance for HLB transmission; (iii) The assessment of the intra- and inter-species variability of the 16S rRNA and/or housekeeping genes of these bacteria; (iv) The assembly and annotation of the mitochondrial genome of *T. erytreae* populations from 19 different European and African geographical areas and v) the identification of a new genetic marker corresponding to a specific indel in the tRNA-Ser gene of the insect mitochondrial genome that highlighted the co-existence in mainland Spain and Portugal, Tenerife and Madeira of two subpopulations of *T. erytreae,* one of which showed a polymorphism not found elsewhere. So, while this study confirms that *T. erytreae* populations, present in South Africa and currently spreading in Spain and Portugal, originated from a common ancestor, the original source of the subpopulations of *T. erytreae* containing a specific indel in the tRNA-Ser remains unknown. This finding indicates that not all the possible original sources of *T. erytreae* populations currently spreading in Europe have been identified yet, asking for a more extensive assessment of *T. erytreae* populations in Africa and other countries where the insect has long been established.

## Author contributions

MC, FDS and BN conceived the study; AF, JAP, BR, HD, HM and RB collected samples and contributed to the experimental design; HM and RB tested samples; MC and VN performed bioinformatics analysis; BN and VN performed molecular and phylogenetic analysis; PS, VP, LP contributed to the experimental design and provided experimental guidance; FDS, MC and BN drafted the manuscript. All the authors approved the final manuscript.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s10340-025-01945-8.

**Acknowledgements** The authors are grateful to Miclay Carvalho (University of São Tomé and Principe), Glynnis Cook and Wayne Kirkman (Citrus Research International) for the help in collecting *T. erytreae* samples.

**Funding** Open access funding provided by Consiglio Nazionale Delle Ricerche (CNR) within the CRUI-CARE Agreement. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817526 (PRE-HLB: Preventing HLB epidemics for ensuring citrus survival in Europe. H2020-SFS-2018–2 Topic SFS-05–2018–2019–2020—new and emerging risks to plant health). This article reflects only the authors' views, and the Agency is not responsible for any use that may be made of the information it contains.

**Data availability** The raw sequence data and bacterial draft genomes reported in this paper have been deposited in the Sequence Read Archive in National Center for Biotechnology information (BioProject ID: PRJNA1099175) that are accessible in read-only format for review at https://dataview.ncbi.nlm.nih.gov/object/PRJNA1099175?reviewer=f1jpg3gc3fuudk45h7ii3qlg3n. The data will be released upon manuscript acceptance for publication. The metadata for all samples included in this study is presented in Additional file 7. Mitochondrial genomes have been submitted to GenBank with accession numbers PQ167384- PQ167402.

## Declarations

**Conflict of interest** The authors declare no competing interests.

## References


Ahn H, Seol D, Cho S, Kim H, Kwak W (2020) Enhanced symbiotic characteristics in bacterial genomes with the disruption of rRNA operon. Biology 9:440. https://doi.org/10.3390/biology9120440

Ajene I, Khamis FM, Pietersen G, van Asch B (2020) Mitochondrial genetic variation reveals phylogeographic structure and cryptic diversity in *Trioza erytreae*. Sci Rep 10:8893. https://doi.org/10.1038/s41598-020-65880-7

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Arenas-Arenas FJ, Duran-Vila N, Quinto J, Hervalejo Á (2018) Is the presence of *Trioza erytreae*, vector of huanglongbing disease, endangering the Mediterranean citrus industry? Survey of its population density and geographical spread over the last years. J Plant Pathol 100:567–574. https://doi.org/10.1007/s42161-018-0109-8

Arenas-Arenas FJ, Duran-Vila N, Quinto J, Hervalejo Á (2019) Geographic spread and inter-annual evolution of populations of *Trioza erytreae* in the Iberian Peninsula. J Plant Pathol 101:1151–1157. https://doi.org/10.1007/s42161-019-00301-x

Balmand S, Lohs C, Aksoy S, Heddi A (2013) Tissue distribution and transmission routes for the tsetse fly endosymbionts. J Invertebr Pathol 112:S116–S122. https://doi.org/10.1016/j.jip.2012.04.002

Bassene H, Niang EHA, Fenollar F, Doucoure S, Faye O, Raoult D, Mediannikov O (2020) Role of plants in the transmission of Asaia sp., which potentially inhibit the Plasmodium sporogenic cycle in Anopheles mosquitoes. Sci Rep 10:7144. https://doi.org/10.1038/s41598-020-64163-5

Baumann P (2005) Biology of bacteriocyte-associated endosymbionts of plant sap-sucking insects. Annu Rev Microbiol 59:155–189. https://doi.org/10.1146/annurev.micro.59.030804.121041

Benhadi-Marín J, Fereres A, Pereira JA (2022) Potential areas of spread of *Trioza erytreae* over mainland Portugal and Spain. J Pest Sci 95:67–78. https://doi.org/10.1007/s10340-021-01440-w

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Asnicar F et al (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37:852–857. https://doi.org/10.1038/s41587-019-0209-9

Bushnell B (2023) BBTools software package. 2014. https://sourceforge.net/projects/bbmap/.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421

Cappelli A, Damiani C, Mancini MV, Valzano M, Rossi P, Serrao A, Favia G (2019) Asaia activates immune genes in mosquito eliciting an anti-Plasmodium response: implications in malaria control. Front Genet 10:836. https://doi.org/10.3389/fgene.2019.00836

Chen S, Zhang D, Augustinos A, Doudoumis V, Bel Mokhtar N, Maiga H, Bourtzis K (2020) Multiple factors determine the structure of bacterial communities associated with Aedes albopictus under artificial rearing conditions. Front Microbiol 11:605. https://doi.org/10.3389/fmicb.2020.00605

Chiel E, Zchori-Fein E, Inbar M, Gottlieb Y, Adachi-Hagimori T, Kelly SE, Hunter MS (2009) Almost there: transmission routes of bacterial symbionts between trophic levels. PLoS ONE 4:e4767. https://doi.org/10.1371/journal.pone.0004767

Chouaia B, Rossi P, Epis S, Mosca M, Ricci I, Damiani C, Favia G (2012) Delayed larval development in Anopheles mosquitoes deprived of Asaia bacterial symbionts. BMC Microbiol 12:S2. https://doi.org/10.1186/1471-2180-12-S1-S2

Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Tiedje JM (2014) Ribosomal database project: data and tools for high throughput rRNA analysis. Nucleic Acids Res 42:D633–D642. https://doi.org/10.1093/nar/gkt1244

Crotti E, Damiani C, Pajoro M, Gonella E, Rizzi A, Ricci I et al (2009) Asaia, a versatile acetic acid bacterial symbiont, capable of cross-colonizing insects of phylogenetically distant genera and orders. Environ Microbiol. 11:3252–64. https://doi.org/10.1111/j.1462-2920.2009.02048.x

Damiani C, Ricci I, Crotti E, Rossi P, Rizzi A, Scuppa P, Genchi M (2010) Mosquito-bacteria symbiosis: the case of *Anopheles gambiae* and Asaia. Microb Ecol 60:644–654. https://doi.org/10.1007/s00248-010-9704-8

De Freece C, Damiani C, Valzano M, D'Amelio S, Cappelli A, Ricci I, Favia G (2014) Detection and isolation of the α-proteobacterium Asaia in Culex mosquitoes. Med Vet Entomol 28:438–442. https://doi.org/10.1111/mve.12045

Douglas AE (2007) Symbiotic microorganisms: untapped resources for insect pest control. Trends Biotechnol 25:338–342. https://doi.org/10.1016/j.tibtech.2007.06.003

Douglas AE (2016) How multi-partner endosymbioses function. Nat Rev Microbiol 14:731–743. https://doi.org/10.1038/nrmicro.2016.151

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340

EPPO Reporting Service (2015a) *Trioza erytreae* occurs in mainland Portugal. EPPO Reporting Service 11. https://gd.eppo.int/media/data/reporting/rs-2015-11-en.pdf

EPPO Reporting Service (2015b) Updated situation of *Trioza erytreae* in Spain. EPPO Reporting Service 11. https://gd.eppo.int/media/data/reporting/rs-2015-11-en.pdf

EPPO (2024) EPPO Global database. https://gd.eppo.int/

Favia G, Ricci I, Damiani C, Raddadi N, Crotti E, Marzorati M, Daffonchio D (2007) Bacteria of the genus Asaia stably associate with *Anopheles stephensi*, an Asian malarial mosquito vector. Proc Natl Acad Sci USA 104:9047–9051. https://doi.org/10.1073/pnas.0610451104

Garrett RA, Klenk HP (2008) Archaea: evolution, physiology, and molecular biology. John Wiley & Sons, Chichester

Hall AAG, Morrow JL, Fromont C, Steinbauer MJ, Taylor GS, Johnson SN, Riegler M (2016) Codivergence of the primary bacterial endosymbiont of psyllids versus host switches and replacement of their secondary bacterial endosymbionts. Environ Microbiol 18:2591–2603. https://doi.org/10.1111/1462-2920.13351

Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS (2018) UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol 35:518–522. https://doi.org/10.1093/molbev/msx281

Hughes GL, Dodson BL, Johnson RM, Murdock CC, Tsujimoto H, Suzuki Y, Rasgon JL (2014) Native microbiome impedes vertical transmission of Wolbachia in Anopheles mosquitoes. Proc Natl Acad Sci USA 111:12498–12503. https://doi.org/10.1073/pnas.1408888111

Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14:587–589. https://doi.org/10.1038/nmeth.4285

Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z (2019) MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 7:e7359. https://doi.org/10.7717/peerj.7359

Karp PD, Midford PE, Billington R, Kothari A, Krummenacker M, Latendresse M, Fulcher C (2021) Pathway Tools version 23.0 update: software for pathway/genome informatics and systems biology. Brief Bioinform 22:109–126. https://doi.org/10.1093/bib/bbz104

Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059–3066. https://doi.org/10.1093/nar/gkf436

Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res 41:e1. https://doi.org/10.1093/nar/gks808

Kuechler SM, Gibbs G, Burckhardt D, Dettner K, Hartung V (2013) Diversity of bacterial endosymbionts and bacteria–host co-evolution in Gondwanan relict moss bugs (Hemiptera: Coleorrhyncha: Peloridiidae). Environ Microbiol 15(7):2031–2042. https://doi.org/10.1111/1462-2920.12101

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/molbev/msy096

Kwak Y, Sun P, Meduri VR, Percy DM, Mauck KE, Hansen AK (2021) Uncovering symbionts across the psyllid tree of life and the discovery of a new Liberibacter species, "*Candidatus*" Liberibacter capsica. Front Microbiol 12:739763. https://doi.org/10.3389/fmicb.2021.739763

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923

Lefort V, Desper R, Gascuel O (2015) FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. Mol Biol Evol 32:2798–2800. https://doi.org/10.1093/molbev/msv150

Li F, Hua H, Ali A, Hou M (2019) Characterization of a bacterial symbiont Asaia sp. in the white-backed planthopper, *Sogatella furcifera*, and its effects on host fitness. Front Microbiol 10:2179. https://doi.org/10.3389/fmicb.2019.02179

Lui LM, Nielsen TN, Arkin AP (2021) A method for achieving complete microbial genomes and improving bins from metagenomics data. PLoS Comput Biol 17:e1008972. https://doi.org/10.1371/journal.pcbi.1008972

Mancini MV, Damiani C, Short SM, Cappelli A, Ulissi U, Capone A, Kalogris C (2020) Inhibition of Asaia in adult mosquitoes causes male-specific mortality and diverse transcriptome changes. Pathogens 9:380. https://doi.org/10.3390/pathogens9050380

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12. https://doi.org/10.14806/ej.17.1.200

Mauck KE, Gebiola M, Percy DM (2024) The hidden secrets of Psylloidea: biology, behavior, symbionts, and ecology. Annu Rev Entomol 69(1):277–302. https://doi.org/10.1146/annurev-ento-120120-114738

McMurdie PJ, Holmes S (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS ONE 8:e61217. https://doi.org/10.1371/journal.pone.0061217

Moran NA, Telang A (1998) Bacteriocyte-associated symbionts of insects. Bioscience 48:295–304. https://doi.org/10.2307/1313356

Moran NA, McCutcheon JP, Nakabachi A (2008) Genomics and evolution of heritable bacterial symbionts. Annu Rev Genet 42:165–190. https://doi.org/10.1146/annurev.genet.41.110306.130119

Morrow JL, Hall AAG, Riegler M (2017) Symbionts in waiting: the dynamics of incipient endosymbiont complementation and replacement in minimal bacterial communities of psyllids. Microbiome 5:58. https://doi.org/10.1186/s40168-017-0276-4

Nakabachi A, Inoue H, Hirose Y (2022) Microbiome analyses of 12 psyllid species of the family Psyllidae identified various bacteria including *Fukatsuia* and *Serratia symbiotica*, known as secondary symbionts of aphids. BMC Microbiol 22(1):15. https://doi.org/10.1186/s12866-021-02429-2

Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268–274. https://doi.org/10.1093/molbev/msu300

Oakeson KF, Gil R, Clayton AL, Dunn DM, von Niederhausern AC, Hamil C, Silva FJ (2014) Genome degeneration and adaptation in a nascent stage of symbiosis. Genome Biol Evol 6:76–93. https://doi.org/10.1093/gbe/evt210

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114

Peng Y, Leung HC, Yiu SM, Chin FY (2012) IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics 28:1420–1428. https://doi.org/10.1093/bioinformatics/bts174

Rasowo BA, Copeland RS, Khamis FM, Aidoo OF, Ajene IJ, Mohamed SA, Borgemeister C (2021) Diversity and phylogenetic analysis of endosymbionts from *Trioza erytreae* (Del Guercio) and its parasitoids in Kenya. J Appl Entomol 145:104–116. https://doi.org/10.1111/jen.12807

Ratcliffe NA, Furtado Pacheco JP, Dyson P, Castro HC, Gonzalez MS, Azambuja P, Mello CB (2022) Overview of paratransgenesis as a strategy to control pathogen transmission by insect vectors. Parasites Vectors 15:112. https://doi.org/10.1186/s13071-021-05132-3

Rodriguez-R LM, Gunturu S, Harvey WT, Rosselló-Mora R, Tiedje JM, Cole JR, Konstantinidis KT (2018) The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. Nucleic Acids Res 46:W282–W288. https://doi.org/10.1093/nar/gky467

Rosell RC, Blackmer JL, Czosnek H, Inbar M (2009) Mutualistic and Dependent Relationships with Other Organisms. In: Stansly PA, Naranjo SE (eds) Bemisia: Bionomics and Management of a Global Pest. Springer, Dordrecht, pp 161–183. https://doi.org/10.1007/978-90-481-2460-2_5

Rossi P, Ricci I, Cappelli A, Damiani C, Ulissi U, Mancini MV, Favia G (2015) Mutual exclusion of Asaia and Wolbachia in the reproductive organs of mosquito vectors. Parasites Vectors 8:278. https://doi.org/10.1186/s13071-015-0888-0

Ruíz-Rivero O, Garcia-Lor A, Rojas-Panadero B, Franco JC, Khamis FM, Kruger K, Urbaneja A (2021) Insights into the origin of the invasive populations of *Trioza erytreae* in Europe using microsatellite markers and mtDNA barcoding approaches. Sci Rep 11:18651. https://doi.org/10.1038/s41598-021-97824-0

Santos-Garcia D, Silva FJ, Morin S, Dettner K, Kuechler SM (2017) The all-rounder Sodalis: a new bacteriome-associated endosymbiont of the lygaeoid bug *Henestaris halophilus* (Heteroptera: Henestarinae) and a critical examination of its evolution. Genome Biol Evol 9:2893–2910. https://doi.org/10.1093/gbe/evx202

Seemann T (2014) Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Sievers F, Higgins DG (2014) Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences. In: Russell DJ (ed) Multiple Sequence Alignment Methods. Humana Press, Totowa, NJ, pp 105–116. https://doi.org/10.1007/978-1-62703-646-7_6

Sloan DB, Moran NA (2012) Genome reduction and co-evolution between the primary and secondary bacterial symbionts of psyllids. Mol Biol Evol 29(12):3781–3792. https://doi.org/10.1093/molbev/mss180

Souza RS, Virginio F, Riback TIS, Suesdek L, Barufi JB, Genta FA (2019) Microorganism-based larval diets affect mosquito development, size and nutritional reserves in the yellow fever mosquito Aedes aegypti (Diptera: Culicidae). Front Physiol 10:152. https://doi.org/10.3389/fphys.2019.00152

Stemke GW, Huang Y, Laigret F, Bove JM (1994) Cloning the ribosomal RNA operons of *Mycoplasma flocculare* and comparison with

those of *Mycoplasma hyopneumoniae*. Microbiology 140:857–860. https://doi.org/10.1099/00221287-140-4-857

Stothard P, Grant JR, Van Domselaar G (2019) Visualizing and comparing circular genomes using the CGView family of tools. Brief Bioinform 20:1576–1582. https://doi.org/10.1093/bib/bbx081

Syberg-Olsen MJ, Garber AI, Keeling PJ, McCutcheon JP, Husnik F (2022) Pseudofinder: detection of pseudogenes in prokaryotic genomes. Mol Biol Evol 39:msac153. https://doi.org/10.1093/molbev/msac153

Thao ML, Clark MA, Baumann L, Brennan EB, Moran NA, Baumann P (2000a) Secondary endosymbionts of psyllids have been acquired multiple times. Curr Microbiol 41:300–304. https://doi.org/10.1007/s002840010138

Thao ML, Moran NA, Abbot P, Brennan EB, Burckhardt DH, Baumann P (2000b) Cospeciation of psyllids and their primary prokaryotic endosymbionts. Appl Environ Microbiol 66:2898–2905. https://doi.org/10.1128/AEM.66.7.2898-2905.2000

Toh H, Weiss BL, Perkin SA, Yamashita A, Oshima K, Hattori M, Aksoy S (2006) Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. Genome Res 16:149–156. https://doi.org/10.1101/gr.4106106

Wang Y, Lu J, Beattie GA, Islam MR, Om N, Dao HT, Cen Y (2018) 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. https://doi.org/10.1002/ps.4917

Whitten MMA, Facey PD, Del Sol R, Fernández-Martínez LT, Evans MC, Mitchell JJ, Dyson PJ (2016) Symbiont-mediated RNA interference in insects. Proc Biol Sci 283:20160042. https://doi.org/10.1098/rspb.2016.0042

Wilson ACC, Ashton PD, Calevro F, Charles H, Colella S, Febvay G, Douglas AE (2010) Genomic insight into the amino acid relations of the pea aphid, *Acyrthosiphon pisum*, with its symbiotic bacterium *Buchnera aphidicola*. Insect Mol Biol 19(s2):249–258. https://doi.org/10.1111/j.1365-2583.2009.00942.x

Zhang DX, Hewitt GM (1997) Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies. Biochem Syst Ecol 25:99–120. https://doi.org/10.1016/S0305-1978(96)00042-7

Zimmermann J, Kaleta C, Waschina S (2021) gapseq: informed prediction of bacterial metabolic pathways and reconstruction of accurate metabolic models. Genome Biol 22:81. https://doi.org/10.1186/s13059-021-02295-1