

RESEARCH

Open Access



Evolutionary genomics and divergence of *Cacopsylla* species with a special focus on the apple proliferation vectors *Cacopsylla melanoneura* and *Cacopsylla picta*

Lapo Ragionieri^{1*†}, Liliya Štarhová Serbina^{2,3†}, Erika Corretto¹, James M. Howie^{4,5}, Fernando Cruz^{6,7}, Tyler S. Alioto^{6,7}, Nicola Zadra⁸, Tobias Weil⁹, Gianfranco Anfora⁸, Christian Stauffer⁴, Lino Ometto^{10*}, Omar Rota-Stabelli^{8,9*} and Hannes Schuler^{1,11}

Abstract

Background The psyllid genus *Cacopsylla* includes several species that act as vectors for phytoplasma-associated diseases affecting plantations across Europe. Among them, *Cacopsylla melanoneura* and *Cacopsylla picta* are the primary vectors of 'Candidatus Phytoplasma mali', the phloem-restricted bacterium responsible for Apple Proliferation disease in Europe. To explore whether vector competence in these species reflects shared ancestry or independent evolution, we assembled mitochondrial and draft nuclear genomes of Italian populations of *C. melanoneura* and *C. picta* and reconstructed time-calibrated phylogenies using 13 mitochondrial protein-coding genes from 12 *Cacopsylla* species.

Results Phylogenetic analyses revealed two major *Cacopsylla* clades (Clade I and II) whose divergence times range from the Early Miocene (18.4 MYA; 95% HPD: 10.8–27.5) to the Middle Miocene (12.7 MYA; 95% HPD: 9.7–16.0). Both *C. melanoneura* and *C. picta* are within Clade I, which is predominantly composed of univoltine species that overwinter on conifers. Within this clade, *Cacopsylla melanoneura* is more closely related to the plum psyllid *Cacopsylla pruni* than to the apple-associated *Cacopsylla picta* and *Cacopsylla mali*, the latter belonging to Clade II. Draft nuclear genomes revealed significant differences in size (438 Mb in *C. melanoneura* vs. 631 Mb in *C. picta*), largely attributed to repetitive elements. Comparative analyses of repetitive elements across *Cacopsylla* species revealed a recent expansion of transposable elements, particularly LINE elements, which were slightly more abundant in Clade I and contributed to the larger genome size observed in *C. picta*.

[†]Lapo Ragionieri and Liliya Štarhová Serbina contributed equally.

*Correspondence:

Lapo Ragionieri
lapo.ragionieri@unibz.it
Lino Ometto
lino.ometto@unipv.it
Omar Rota-Stabelli
omar.rotastabelli@unitn.it

Full list of author information is available at the end of the article



Conclusions Collectively, our findings provide the first genomic resources for *C. melanoneura*, *C. picta*, and several other phytoplasma-vectoring *Cacopsylla* species. We established a robust mitogenomic phylogeny with divergence estimated for this genus showing the presence of two clades with the representatives predominantly associated with different overwintering strategies. Our results further indicate that vectorial capacity in *Cacopsylla* reflects an independent evolutionary trajectory rather than a shared ancestral origin. This evolutionary framework advances our understanding of the biology and origin of vector competence in this agriculturally important group.

Keywords Psylloidea, Genome assembly, Mitochondrial genes, Time-calibrated phylogeny, Transposable elements, Apple proliferations, Overwintering behaviour

Background

Psyllids (Hemiptera: Psylloidea) are a group of small phytophagous insects comprising ~ 4,000 described species belonging to seven families [1]. They are members of Sternorrhyncha, a suborder which includes several other well studied plant pathogen vectors such as aphids (Aphidoidea), whiteflies (Aleyrodoidea) and scale insects (Coccoidea) [1, 2]. Like many phloem feeding hemipterans, various psyllid species can vector phytoplasmas. Phytoplasmas are obligate, cell wall-less, prokaryotic plant pathogens (Mollicutes) inhabiting the phloem sieve elements of their host plant, and are responsible for a variety of economically important tree and crop diseases [3–5] including European Stone Fruit Yellows, Pear Decline, and Apple Proliferation [5, 6]. Phytoplasmas alter the plant's cellular processes by interfering with its ability to transport nutrients and water. This results in weakened and malformed structures, as well as in significant yield reductions [4–6].

Apple Proliferation (AP) is an important disease affecting apple trees, associated with the phytoplasma '*Candidatus Phytoplasma mali*' (hereafter, '*Ca. P. mali*'; for a review see Janik et al. [6]). Distinctive symptoms include characteristic shoots with a "witches' broom" appearance and fruits of reduced size and quality [7]. Infected trees cannot be cured, and they must be eradicated, leading to significant yield losses [6, 8, 9]. The transmission of phytoplasma occurs when the vector insects feed on the phloem of infected plants, acquiring phytoplasma which must then pass through the insect's digestive system, cross the gut barrier, replicate, and invade the salivary glands. When the infected psyllid later feeds on the phloem of a healthy apple tree, the pathogen can be transmitted thus establishing a new infection [10]. The genus *Cacopsylla* (Psyllidae) includes all known vectors for fruit tree phytoplasmas belonging to the phytoplasma group 16SrX [11, 12]. The main vectors of AP are *C. melanoneura* and *C. picta*, while other hemipteran species like the leafhopper *Fieberiella florii* (Stål, 1864), are playing only a minor role [13]. *Cacopsylla picta* (Föster, 1848) acts as the main vector of the AP in most European countries [14–16] while *C. melanoneura* (Föster, 1848) is the main vector in the areas of Northwestern Italy where *C. picta* is absent [14, 17, 18]. Due to variation in vector

capabilities of these two psyllids, various environmental and genetic factors are expected to influence the rate of acquisition and transmission of '*Ca. P. mali*'. These factors include the infectious status of the host tree, the genetics of the psyllid vector [16, 19], the insect's endosymbionts [20] and the phytoplasma itself [19].

Cacopsylla, like most other psyllid species, are host specific, with closely related psyllids tending to associate with closely allied host plant lineages [21, 22]. In temperate regions, many psyllid species migrate from their deciduous host plants to conifers, where they overwinter as adults or eggs [23]. In contrast, psyllids in tropical and Southern temperate regions are often associated with evergreen hosts and tend to remain on the same plant throughout the year, completing several generations [23]. Intriguingly, some *Cacopsylla* species can develop on multiple related host plants (i.e. Rosaceae), forming genetically distinct groups with differing fitness levels. This is exemplified by *C. melanoneura*, which appears to have at least two genetically distinct populations, associated either with apple trees or hawthorn [24].

The timing of *Cacopsylla* diversification has not yet been inferred using time-calibrated phylogenies, leaving the genus's evolutionary history, particularly its historical biogeography and host-plant associations, poorly understood. To date, a Neighbor-Joining analysis based on partial *cytochrome oxidase I* (COI) gene sequences provided weak support for most deep clades, including those containing the AP vectors *C. melanoneura* and *C. picta*, or other *Cacopsylla* species like *C. mali* and *C. pruni* [25]. A more recent phylogeny, based on whole mitochondrial genomes that included several *Cacopsylla* species, was not time-calibrated and omitted both *C. melanoneura* and *C. picta*, limiting its utility for resolving the evolutionary relationships among key vectors of '*Ca. P. mali*' [26]. Genomic information for these two species, and psyllids more broadly, remains limited, with existing studies focusing mainly on mitochondrial genomes [27]. These studies suggest the existence of two distinct clades, which distinguish species with differences in overwintering behaviour [28], although the diversification, including at the genetic and genomic level, is not yet established.

A comprehensive understanding of a species evolutionary history requires investigating the evolution of the

nuclear genome, which is shaped by various molecular mechanisms such as different classes of mutations, gene duplications, recombination, and horizontal gene transfer. In many species one major component involved in genome evolution and adaptation is represented by transposable elements [29]. Transposable elements (TEs) are mobile DNA sequences that can significantly influence the genome structure and its functions. In insects, the proportion of the genome occupied by TEs varies widely, reaching up to 60% in some species such as the migratory locust [30]. Although TE activity can have potentially deleterious effects on gene regulation, increasing evidence suggests that TEs also drive genomic innovation and can confer selective advantages to the host [31]. Comparative studies have revealed both similarities and differences in TE family diversity and abundance among insect genomes. Notably, significant variation in TE activity has been observed even among closely related species [32]. For example, within the order Hemiptera, contrasting TE profiles have been reported. In *Homalodisca vitripennis* (Hemiptera, Cicadellidae), LINES and DNA transposons show a distinct divergence distribution, while SINES and LTR elements are poorly represented. This pattern contrasts sharply with that of the pea aphid *Acyrtosiphon pisum* (Hemiptera, Aphididae), where SINES dominate the TE content, and LINES and LTRs are virtually absent.

To overcome the limited genomic resources available for *Cacopsylla* species, we present newly assembled mitochondrial genomes and draft nuclear genome assemblies of *C. melanoneura* and *C. picta*. The mitochondrial genomes, together with publicly available data from other *Cacopsylla* species, were used to reconstruct a robust, time-calibrated phylogeny of the genus, including several vector species with different host plant preferences. In parallel, we provide the first nuclear genome assemblies of *C. melanoneura* and *C. picta* and use them to investigate the role of repetitive elements in psyllid genome evolution. Given the established involvement of repetitive elements in adaptive evolution and gene regulation, we also performed a comparative analysis across multiple *Cacopsylla* species, with a particular focus on the AP vector species *C. melanoneura* and *C. picta*, to identify lineage-specific traits potentially associated with the accumulation of recently active TEs.

Methods

Psyllid collection, DNA extraction, library preparation and sequencing

A total of 20 individuals of each species, *C. melanoneura* and *C. picta*, were collected in 2015 using the beating tray method in apple orchards in Northeastern Italy, specifically from Laimburg (South Tyrol) and Bosentino (Trentino). Samples were preserved at -20°C in absolute ethanol. Total DNA was extracted from each individual,

and those with higher quality and yield were selected for subsequent library preparation and sequencing. For *C. melanoneura*, DNA was extracted using the NucleoSpin Tissue Kit (Macherey-Nagel), and three paired-end libraries were constructed from two individual specimens and one pool of six individuals, with different fragment sizes (PE290, PE310, and PE700; Table S1). For *C. picta*, an enhanced DNA extraction protocol was used, involving overnight lysis with DTT and extraction with the KingFisher Cell and Tissue DNA Kit (Thermo Fisher Scientific). Two paired-end libraries were prepared, each from a single individual and with different fragment sizes (PE510 and PE750; Table S1).

All five libraries were sequenced on the Illumina HiSeq 2000 platform with a 251 bp insert size (PRJNA1312285). This yielded 57.9, 59.4, and 75.3 Gb of raw data for *C. melanoneura*, corresponding to genome coverages of 131 \times , 135 \times , and 171.74 \times , respectively. For *C. picta*, sequencing produced 121.5 and 74.9 Gb of raw data, with corresponding coverages of 192 \times and 119 \times .

Nuclear genome size estimation in *Cacopsylla* species

Prior to the *de novo* assembly, the genome sizes of *C. melanoneura* (Table S2) and *C. picta* (Table S2) were estimated using a k-mer analysis of the raw reads of library PE700 and PE510, respectively, using *Jellyfish* v2.2.6 [33], *GenomeScope* v1.0 ($k = 17$) [34] and *SGA preqc* ($k = 31$) [35]. Additionally, using this approach we also estimated the genome size of *C. burckhardti* (PRJNA927338) and *C. jukyungi* (PRJNA927338), for which raw data with sufficient coverage are publicly available.

Assembly of the nuclear genomes of *C. melanoneura* and *C. picta*

Adapter sequences were removed and reads were trimmed in the *Trim Galore* wrapper (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) using *Cutadapt* [36]. Overlapping reads from the shorter fragment libraries were merged using *FLASH* [37], after which reads were filtered in *GEM-mapper* [38] (with up to 2% mismatches) against a contamination database including phiX, Univec sequences, *E. coli*, *Cacopsylla coccinea* mitochondrion complete genome (NC_027087.1) and 17 contaminant species' genomes (detected independently for each *Cacopsylla* species, using with *Kraken* v0.10.5 [39] as present in $\geq 0.03\%$ of the reads, see Table S3).

The *C. melanoneura* genome was assembled *de novo* using overlapping read-pairs from the PE700 library (detected with *FLASH*) using *DISCOVAR de novo* (experimental release v.51885) with default options [40, 41], after discarding all read-pairs shorter than 200 bp. Circular lines were discarded, and the *DISCOVAR* contigs were scaffolded with *BESST* v2.2.5 [42] using processed reads

from all three PE libraries (PE290, PE310 and PE700) with the option “*--separate_repeats*” and discarding all contigs classified as repeats. Mis-assemblies were removed as in [43, 44], using reads from all PE libraries, after introducing breaks at genomic intervals calculated to have negative consistency scores. This resulted in shorter but more reliable scaffolds, which were then broken again into contigs. Then contig assembly was “unchaffed” by filtering out those shorter than the read length (250 bp), thus only retaining sequences long enough to contain at least a partial exon (251–500 bp). Finally, the filtered contigs were re-scaffolded in *BESST* v2.2.5 yielding a final draft assembly (Table S4).

The *C. picta* genome was assembled with a modified version of the *C. melanoneura* pipeline. Overlapping read-pairs from library PE510 were assembled with *DISCOVAR de novo* [40, 41], discarding both read-pairs shorter than 200 bp and circular lines. The *DISCOVAR* contigs were collapsed with *Redundans* [45] to reduce heterozygous haplotypes and repeats, thus accounting for high fragmentation. *Redundans* contigs were scaffolded with *BESST* v2.2.5 [42] using both PE libraries with the “*--separate_repeats*” option, yielding the final assembly draft (Table S5).

Estimates of genomic completeness were obtained for each genome using two methods. *KAT* [46] compared k-mers in the pre-processed reads (libraries PE290, PE310, PE700 for *C. melanoneura*, PE510 for *C. picta*) with those in the final genomes to estimate sequence completeness. To assess genome assembly quality, we used *BUSCO* v5.8.3 [47], which estimates the completeness of genomic data based on the presence of universal single-copy orthologs (OrthoDB database odb12).

Repetitive element annotation and comparisons

The expected sample heterozygosity and genomic repetitiveness were estimated in *GenomeScope* v1.0 [34] (Table S2). We then conducted a comparative analysis of abundances, sequence length and genome percentage for multiple repetitive element classes, including retroelements (LINEs, SINEs, LTRs), DNA transposons, and interspersed repeats (sRNAs, DNA satellites, simple repeats) using *RepeatMasker* [48], by mapping masked reads to a *RepeatMasker* Combined Database (Dfam_Consensus-20181026, RepBase-20181026) with *RMblastn* v2.6.0+ (command: *RepeatMasker scaffold.fa -species arthropoda -pa 3*). In addition, *RepeatModeler2* [49] with the option “*-LTRStruct*” was used to identify dispersed transposable elements. Subsequently, *RepeatMasker* was used to identify and mask interspersed repeats in the nuclear draft assemblies of *C. melanoneura* and *C. picta*, using the *RepeatModeler2* library option. Additionally, we identified satellites and putative satellites in the raw reads and included them in *RepeatModeler2* families for

use in a *RepeatExplorer2* analysis. Briefly, the pipeline was run to characterize the repeatome [50, 51], including the *TAREAN* analysis within the *Galaxy* portal (<https://repeatexplorer-elixir.cerit-sc.cz>). For this, we randomly selected a total of 2,000,000 paired reads. Default options were selected, except that we set computing time to “long”, applied a filtration of the most abundant repeats, and set the threshold of the analysis to 0.001%.

Besides *C. melanoneura* and *C. picta* presented here, no nuclear genome assemblies are available for other *Cacopsylla* species. For some other species, sequencing reads are publicly available (Tables S6). To analyse the genomic content and compare nuclear repeats across *Cacopsylla* species, we used *dnaPipeTE* (<https://github.com/clemgoub/dnaPipeTE>), a tool optimized for repeatome analysis from low-coverage sequencing data [52] that uses raw reads as input. Before analysis, we removed reads matching to mitochondrial DNA, known symbionts and contaminants. Repeat annotation was performed using the Dfam-RepeatMasker.lib database for Arthropoda, which included 6,828 entries as of May 2025. Repeat landscapes were generated using the *dnaPTlandscape.sh* script from the *dnaPT_utils* package (https://github.com/clemgoub/dnaPT_utils), with the *-S* and *-U* options enabled to visualize only repeat superfamilies and to exclude unclassified elements.

Assembly of the mitochondrial genomes of *C. melanoneura* and *C. picta*

Mitochondrial genomes of *C. melanoneura* (PX243765) and *C. picta* (PX243766) were assembled and circularised in the *MitoZ* pipeline [53] operated in a *Singularity* container [54] using Illumina raw read libraries PE290 and PE750, respectively. Mitogenomes were annotated in *MITOS* [55] and *MITOS2* [56], a consensus annotation manually compiled, and the genomes visualised with annotations, GC content and coverages in *MitoZ* [53].

Mitochondrial phylogeny and divergence time estimation

To test the phylogenetic placement of *C. melanoneura* and *C. picta* within the *Cacopsylla* genus, we conducted phylogenetic analyses using Bayesian inference (BI) in *BEAST* v2.7.5 [57, 58] and maximum likelihood (ML) in *IQ-TREE* v1.6.12 [59]. The best-fit partitioning schemes and substitution models for subsequent phylogenetic analyses were predicted with *ModelFinder* [59] implemented in *IQ-TREE* release 2.1.4b [59]. We repeated the analyses using three datasets: amino acid sequences, the corresponding nucleotide (codon) sequences, and nucleotide sequences with the third codon position excluded. The mitochondrial sequences of 13 protein coding genes (PCGs) from 12 *Cacopsylla* species, eight psyllids from other families, two aphids and two white flies (Table S6) were aligned using the MAFFT-L-INS-I algorithm [60]

and concatenated using *catsequence* 1.3 (<https://zenodo.org/record/4409153#.YmJYT35Byot>).

The optimal substitution model for each partition of the BI analyses was determined using the model-averaging approach implemented in OBAMA [57] with all substitution models enabled for the amino acid data, and the bModelTest packages [61] for the nucleotide data. The BI analyses of 100 million MCMC generations, with trees sampled every 1,000 generations, were run on the CIPRES platform [62] using BEAST [57]. Random starting trees were assigned to each partition. Divergence time estimation was conducted using the Most Recent Common Ancestor (MRCA) approach, incorporating various calibration points using *BEAUTi* [57] under an Optimized Relaxed Clock model. A time constraint derived from the posterior estimate of Li et al. [63] was applied to the split between Psylloidea and its sister clade (Aleyrodoidea + Aphidoidea) set at 269 MYA and modeled as a normal prior distribution (mean = 269, S = 5.9). Outgroup taxa from Aleyrodoidea and Aphidoidea were represented by *Aleurodicus dugesii* + *Bemisia tabaci* and *Aphis gossypii* + *Myzus persicae*, respectively. To calibrate the Psylloidea crown node, we used the Eocene fossil *Eogyropsylla paveloctogenarius* (Aphalaridae; 41.3–47.8 MYA), the oldest known representative of Psylloidea s.str [64]. The youngest age (41.3 MYA) was applied as a soft minimum using a log-normal prior (M = 3.67, S = 0.32, offset = 41.3), with *Lanthanaphalara mira*, an extant member of the same family, used to define the calibrated node. An additional calibration was applied to a tribal-level node within the family Liviidae (Paurocephalini), using the Early Miocene fossil *Melanastera casca* (20.4–13.8 MYA), a congener of the extant *M. paucipunctata* [65]. The youngest age (13.8 MYA) was treated as a soft minimum, implemented as a log-normal prior (M = 3.3, S = 0.39, offset = 13.8).

To select the tree prior (Yule vs. birth–death models), we performed a marginal likelihood estimation (MLE) using *Tracer* [66, 67]. The marginal likelihood of both priors was calculated with 1,000 path steps, running chains for 1 million generations, and sampling log likelihood every 1,000 cycles. This analysis was conducted using *BEAST* v1.10.4 [68]. We calculated the Bayes factor (BF) as twice the natural logarithm (2lnBF) with respect to the highest-likelihood model, considering BF values > 10 as significantly favouring one model over another. The remaining prior parameters for both Yule and birth–death models were kept unchanged. The results were evaluated in *Tracer* v1.7.1 [66], using the effective sampling size (ESS > 200) criterion, and the first 30% of samples were discarded as burn-in. The tree files were then combined using *LogCombiner* v2.4.5, and the parameter values were annotated to the Maximum Clade Credibility (MCC) tree

using *TreeAnnotator* v2.4.5. Nodal support was assessed by posterior probabilities (PP).

ML and BI trees were visualized using *FigTree* v1.4.4 (<https://github.com/rambaut/figtree>). All figures were edited in *GIMP* v. 2.10.38.

Results

Nuclear genomes of *C. melanoneura* and *C. picta*

Based on raw sequencing reads, *C. melanoneura* has an estimated genome size of 438.67 Mb, while *C. picta* has a larger genome with 632.08 Mb (Table S2). The other species with sufficient and public available raw reads data showed intermediate genome sizes: 528.10 Mb for *C. jukyungi* and 549.12 Mb for *C. burckhardti*.

The final assembly of the *C. melanoneura* genome yielded a 688 Mb sequence, in a total of 678,097 contigs and 483,574 scaffolds with N50 values of 1.53 kb and 3.92 kb, respectively (Table S4). The assembled genome size was thus inflated relative to a priori estimates of 438.7–473.8 Mb (Table S2). The BUSCO completeness analysis revealed 32.9% of genes recovered as complete (of which 4.6% were duplicates), 31.4% as fragmented and 35.7% missed (Fig. S1). As expected by using data from a pool of individuals, this genome appears to be highly heterozygous (2.51%) and repetitive (estimated at 57.28%) a considerable number of *k*-mers were missed in the heterozygous and homozygous regions, and several artificial duplications and repeats emerged due to the non-perfect collapsing of haplotypes.

The assembled *C. picta* genome was larger, with 739.04 Mb and shows contig and scaffold N50s of 2.22 Kb and 6.3 Kb respectively (Table S5). The BUSCO completeness analysis indicated 30.1% of genes complete (of which 0.7% were duplicates), 33.1% fragmented (i.e. partial matches) and 36.8% missed. Because the genome was heterozygous (1.37%) and repetitive (estimated at 67.1%), there were some difficulties collapsing alleles in heterozygous regions and an excess of duplicated sequences in homozygous regions, although the effect on genome size inflation was weaker than in *C. melanoneura*. Although the genome size and repetitiveness seem to be properties inherent to both genomes, the heterozygosity and efficiency in collapsing haplotypes can be explained by the larger number of individuals sequenced and simultaneously assembled in *C. melanoneura*.

Mitochondrial genomes of *C. melanoneura* and *C. picta*

The final assemblies of the *C. melanoneura* and *C. picta* mitogenomes yielded 14,882 bp and 14,829 bp circular sequences, respectively, with 13 PCGs, 22 tRNAs and 2 rRNAs, as well as some control region and origin of L-strand replication elements (recovered in *MITOS2*) (Fig. S2). The mitochondrial genome size of *C. melanoneura*

Table 1 Biological traits, vector status, host associations and geographic distribution of 12 *Cacopsylla* species included in the phylogenetic and transposable elements analyses

Clade	<i>Cacopsylla</i> species	Volinism	Overwintering strategy	Dimorphism	Vector status	Host plant species	Host plant family	Distribution	References
<i>Cacopsylla</i> Clade I	<i>C. picta</i>	univoltine	adults, confers	absent	<i>Cz.</i> Phytoplasma mali	<i>Malus domestica</i> , <i>M. sylvestris</i>	Rosaceae	Western Palaearctic	Hodkinson, 2009; Ounvard, 2025
	<i>C. burckhardtii</i>	univoltine	adults, confers	absent	unknown	<i>Pyrus calleryana</i> , <i>P. communis</i> , <i>P. pyrifolia</i> var. <i>culta</i> , <i>P. ussuriensis</i>	Rosaceae	Asia	Cho et al., 2017
	<i>C. pyrisuga</i>	univoltine	adults, confers	absent	<i>Cz.</i> Phytoplasma pyri	<i>Pyrus amygdaliformis</i> , <i>P. communis</i> , <i>P. paschica</i> , <i>P. pyrifolia</i> var. <i>culta</i> , <i>P. salicifolia</i> , <i>P. ussuriensis</i>	Rosaceae	Palaearctic	Hodkinson, 2009; Ounvard, 2025
<i>Cacopsylla</i> Clade II	<i>C. melanoneura</i>	univoltine	adults, confers	present	<i>Cz.</i> Phytoplasma mali	<i>Crataegus maximowiczii</i> , <i>C. monogyna</i> , <i>C. oxyacanthifolia</i> , <i>Malus communis</i> , <i>Mespilus germanica</i>	Rosaceae	Palaearctic	Hodkinson, 2009; Ounvard, 2025; Sejjak, 2006
	<i>C. pruni</i>	univoltine	adults, confers	absent	<i>Cz.</i> Phytoplasma prunorum	<i>Prunus armeniaca</i> , <i>P. avium</i> , <i>P. cerasifera</i> , <i>P. domestica</i> , <i>P. instituta</i> , <i>P. padus</i> , <i>P. persica</i> , <i>P. salicina</i> , <i>P. spinosa</i>	Rosaceae	Palaearctic	Hodkinson, 2009; Ounvard, 2025
	<i>C. fuscescilla</i>	unknown	unknown	unknown	unknown	<i>Eriobotrya japonica</i>	Rosaceae	Asia	Zhao et al., 2023
<i>Cacopsylla</i> Clade III	<i>C. citrisuga</i>	uncertain	adults, host plants	present	<i>Cz.</i> Liberibacter asiaticus	<i>Citrus grandis</i> , <i>C. limon</i> , <i>C. meadica</i> , <i>C. trifoliata</i>	Rutaceae	Asia	Ounvard, 2025
	<i>C. mali</i>	univoltine	eggs, host plants	absent	unknown	<i>Malus asiatica</i> , <i>M. domestica</i> , <i>M. pumila</i> , <i>M. sylvestris</i>	Rosaceae	Palaearctic, North America, Asia	Hodkinson, 2009; Ounvard, 2025
	<i>C. coecinea</i>	polyvoltine	adults	present	unknown	<i>Akshiba quinata</i> , <i>A. trifoliata</i> , <i>Stantonia hexaphylla</i>	Lardizabaleaceae	Asia	Kwon, 1983; Ounvard, 2025
	<i>C. jukyangi</i>	polyvoltine	adults, host plants	dimorphic	unknown	<i>Pyrus communis</i> , <i>P. pyrifolia</i> var. <i>culta</i> , <i>P. ussuriensis</i>	Rosaceae	Asia	Cho et al., 2017; Ounvard, 2025
	<i>C. pyri</i>	polyvoltine	adults, host plants	dimorphic	<i>Cz.</i> Phytoplasma pyri	<i>Pyrus communis</i> , <i>P. elaeagnifolia</i>	Rosaceae	Western Palaearctic	Hodkinson, 2009; Ounvard, 2025
	<i>C. pyriscola</i>	polyvoltine	adults, host plants	dimorphic	<i>Cz.</i> Phytoplasma pyri	<i>Pyrus communis</i> , <i>P. pyraeaster</i> , <i>P. ussuriensis</i>	Rosaceae	Palaearctic, introduced in North America	Hodkinson, 2009; Ounvard, 2025

Volinism, the number of generations per year; Overwintering strategy, the developmental stage and plant on which psyllid overwinters; Dimorphism, the presence of seasonally distinct morphological forms; Vector status denotes whether the species is a confirmed vector of plant pathogens (*Liberibacter*, *Phytoplasma*); Host plant species/family, the plant taxa on which psyllid completes its life cycle; Distribution, the known geographic range of each species

presented here was almost identical to the one of specimens from Czech Republic (14,779 – 14,881 bp) [27], while the *C. picta* specimen of this study had a slightly larger mitochondrial genome compared to Czech individuals (14,801 and 14,802 bp) [27]. In general, variations in genome length among the investigated *Cacopsylla* species are consistently found in the control regions. The gene order and GC usage in both *C. melanoneura* and *C. picta* was identical to that observed across all known *Cacopsylla* species [26, 69–71] as well as most other psyllid species [26, 27].

Mitochondrial phylogeny and divergence time estimation of *Cacopsylla* species

We investigated the phylogenetic relationships of the newly sequenced samples of *C. melanoneura* and *C. picta* within the genus *Cacopsylla*. Our dataset comprised 13 PCGs obtained from whole-mitochondrial sequences across 12 *Cacopsylla* species and 12 outgroup taxa (Table 1 and Table S6). To assess the robustness of our phylogenetic inference, we employed two analytical frameworks, Bayesian inference (BI) and maximum likelihood (ML), and analysed three data types: amino acid sequences (3,404 bp), nucleotide sequences including all codon positions (10,211 bp), and nucleotide sequences with the third codon position excluded, hereafter referred to as the nucleotide tree (6,809 bp).

For the time-calibrated BI phylogenies, we assessed the fit of different tree priors by comparing marginal likelihood estimates (MLE) derived via stepping-stone sampling. The MLE values for the Yule and birth–death models were –106,096.04 and –106,088.62, respectively. The resulting Bayes Factor (BF = -7.42) favoured the birth–death model, which we therefore used in the final analyses.

In general, the BI trees based on the amino acid dataset and the nucleotide dataset excluding the third codon position, provided stronger support for both deep and internal branches suggesting that these values provide more reliable approximate estimates compared to the all-codon tree (Fig. 1; S3, S4, S5). A notable difference in the all-codon tree, however, was the recovery of the outgroup family Liviidae as monophyletic, including *Euphyllura phillyreae* (Euphyllurinae) and three representatives of the subfamily Liviinae (*Livia junci*, *Melanastera paucipunctata*, *Paurocephala sauteri*). In addition, the all-codon tree recovered a monophyletic grouping of *C. coccinea* and *C. mali* (Fig. S3).

The divergence time estimates from the amino acid tree were in general older compared to the nucleotide tree. In both trees, the crown diversification of Psylloidea was estimated to have occurred in the Paleocene–Eocene, at approximately 52.1 MYA (95% HPD: 45.7–70.2 MYA) based on the nucleotide dataset, or 62.3 MYA (95% HPD:

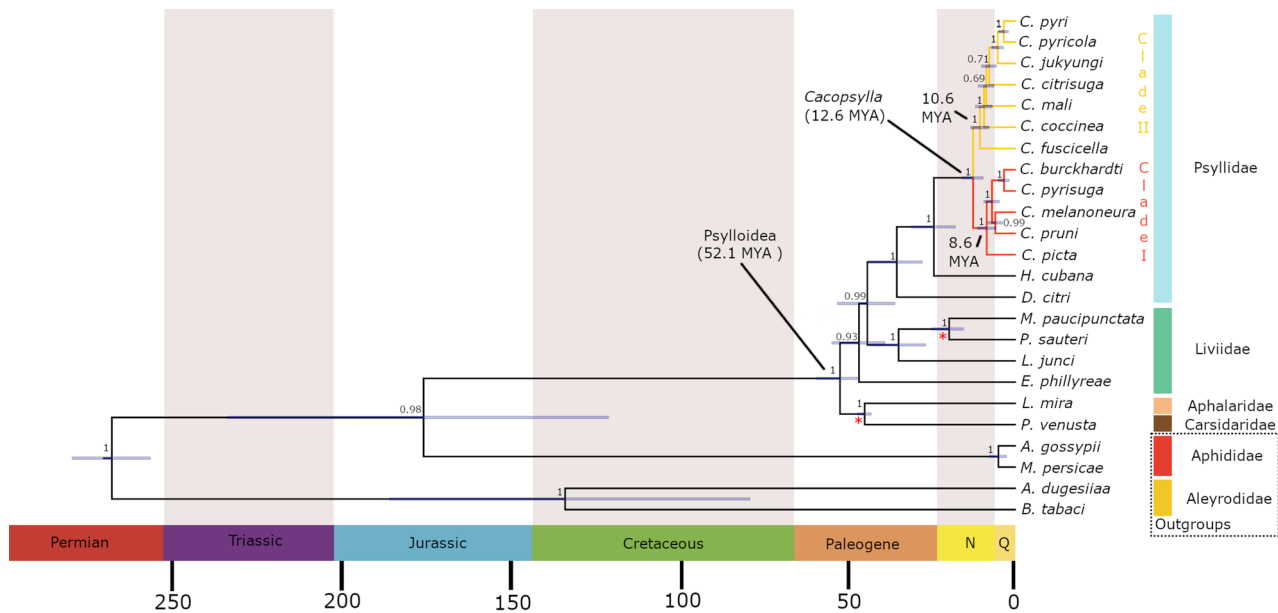


Fig. 1 Time-calibrated Bayesian phylogeny of 12 *Cacopsylla* species, based on 13 mitochondrial PCGs. Red branches correspond to Clade I and yellow branches to Clade II. The * refers to selected fossil records used for calibrations. Number at nodes correspond to posterior probabilities derived from the amino acid dataset. Node divergence age estimates \pm 95% HPD are shown to be overlapping based on nucleotide third codon removed phylogeny

46.3–82.7 MYA) based on the amino acid dataset (Fig. 1). These estimates are substantially older than the fossil prior minimum of 41.3 MYA. In the amino acid tree, the crown age of the genus *Cacopsylla* was estimated to be in the Early Miocene, at approximately (95% HPD: 10.8–27.5 MYA). In contrast, the nucleotide tree suggested a later divergence, placing the crown age of *Cacopsylla* in the Middle Miocene, at approximately 12.7 MYA (95% HPD: 9.7–16.0 MYA). Both the amino acid and nucleotide trees strongly supported (PP=1.0) the presence of two major *Cacopsylla* clades: Clade I and Clade II (Figs. 1 and 2). All species within *Cacopsylla* Clade I (red branches) are associated with host plants from the family Rosaceae (Table 1; Fig. 2). In all trees, *C. melanoneura* forms a clade with the plum psyllid *C. pruni*, with strong support in the amino acid tree (PP=0.96) and in the nucleotide tree (PP=1.0). The divergence between these two species was estimated at approximately 6.1–6.9 MYA (amino acid tree: 95% HPD: 2.4–11.8 MYA; nucleotide tree: 95% HPD: 3.9–8.3 MYA). This pair consistently clusters as a sister group to the pear psyllids *C. burckhardti* and *C. pyrisuga*, with strong support across both datasets (amino acid tree PP=1.0; nucleotide tree PP=0.99). In turn, *C. picta* forms a sister lineage to the above-mentioned clades in both trees (PP=1.0), with the split between *C. picta* and the most recent common ancestor (MRCA) of these species estimated at approximately 8.6–11.5 MYA, in the Late Miocene (amino acid tree: 95% HPD: 5.5–18.7 MYA; nucleotide tree: 95% HPD: 6.0–11.4 MYA). These results suggest independent

evolutionary trajectories for the two main AP vector species in Europe, *C. melanoneura* and *C. picta*.

The *Cacopsylla* species in Clade II (yellow branches) form a separate group, also including psyllids primarily associated with Rosaceae, along with single representatives from the families Lardizabalaceae and Rutaceae (Table 1). The two European species associated with *Pyrus*, *C. pyri* and *C. pyricola*, cluster together with the Asian *C. jukyungi*, also associated with *Pyrus*, as sister species (Fig. 2). These nodes exhibit strong support across all phylogenetic inference methods. In contrast, the phylogenetic relationships among *C. citrisuga*, *C. coccinea*, and *C. mali* remain unresolved and vary considerably among different phylogenetic reconstructions, likely due to incomplete taxon sampling. Interestingly, *C. citrisuga* and *C. coccinea* are Asian species that primarily reproduce on citrus and *Akebia*, respectively, while *C. mali*, found in Europe and North America, reproduces on apple trees. Notably, although it is the only species in this group associated with apple, *C. mali* is phylogenetically distant from *C. melanoneura* and *C. picta* and has never been recovered as their sister taxon.

Another notable finding is that the divergence of the four pear-associated *Cacopsylla* species from the two clades, *C. pyri* / *C. pyricola* (Clade II) and *C. burckhardti* / *C. pyrisuga* (Clade I), represents the two most recent speciation events among the analysed taxa, estimated at around 4.7 MYA (amino acid tree: 95% HPD: 3.6–6.1 MYA; nucleotide tree: 95% HPD: 4.6–6.1 MYA). Similar to the results of the BI all-codon tree, the ML all-codon

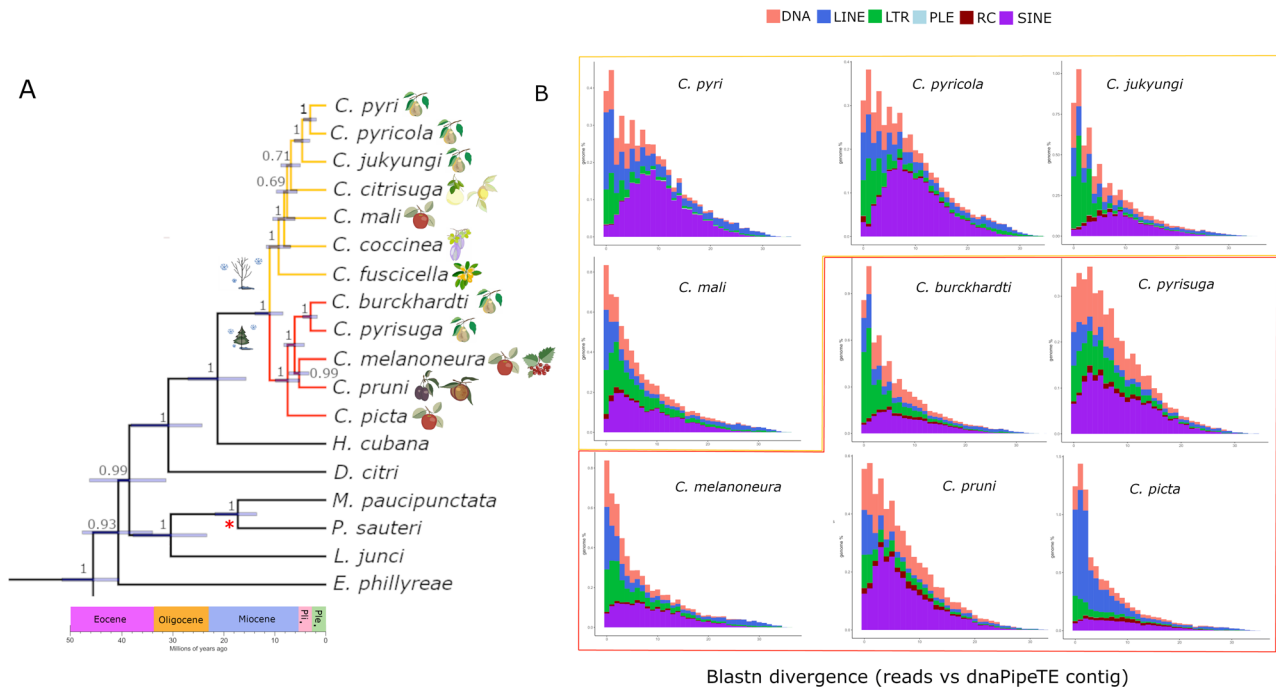


Fig. 2 Subtree with dated phylogeny of *Cacopsylla* species with repeat landscape plots. **A** Subtree from Fig. 1. Icons of the leafless host plant tree and the conifer indicate differences in overwintering behaviour (migrating - in red (Clade I); non-migrating - in yellow (Clade II), whereas icons of different fruits represent the most agriculturally important host plants for each species. **B** Landscapes divergence in TEs, the X-axis of each plot reflects how much a TE has diverged from its consensus sequence, with low divergence suggesting younger TEs (peaks on the left), while high divergence indicates ancient insertions (broad right-shifted distributions). The Y-axis represents the genomic proportion of the genome occupied by TEs at each divergence level. The large fraction of unclassified elements was omitted for these plots to focus only on annotated elements

tree also recovered the family Liviidae as monophyletic, albeit with weak support (Fig. S6). In contrast, both the BI and ML trees based on amino acid and third codon-excluded nucleotide datasets did not support the monophyly of Liviidae, while both major *Cacopsylla* clades (Fig. S7 and S8) were strongly supported in all ML analyses ($BS > 0.98$), with internal branch support across *Cacopsylla* being generally lower than in the BI inferences.

Repetitive element landscape diversity in *Cacopsylla* species

TE composition was investigated using landscape plots that report the frequency and variability of the different TE families. In this analysis, sequence divergence of TE copies from their consensus sequence is shown on the X-axis, and their genomic proportion on the Y-axis. Peaks toward the left (low divergence) indicate more recent TE insertions, while those toward the right (high divergence) reflect older insertions. In *C. mali* and *C. picta*, prominent recent bursts of DNA and LINE elements support relatively recent TE bursts. In contrast, species like *C. pruni* and *C. pyricola* show older TEs expansions (Fig. 2). The dominant TE classes vary across samples, indicating that different TE families drive genome changes in different species (Fig. 2). Specifically, in *C. pyri*, there is high recent TE activity, particularly in DNA and LINE

elements, while in its sister species, *C. pyricola*, recent TE activity is lower and the TE landscape suggests older insertions. *Cacopsylla jukyungi* shows DNA and LTR elements dominant in recent bins, suggesting a likely recent expansion. In *C. mali*, there is a very pronounced peak of low divergent TE copies, supporting strong recent activity of DNA, LINE and RC elements. This is likely one of the highest recent TE activity profiles. In *C. burckhardtii*, there is moderate to high recent TE activity, mostly driven by LTR, DNA and LINE elements. The sister species, *C. pyrisuga*, has a flatter and more right-shifted TE distribution, which suggests that TE activity has slowed down. *Cacopsylla melanoneura* shows very active recent TE insertions, especially in DNA, LINE and LTR elements. The prominent peak at 0–2% divergence suggests a recent expansion event. In contrast, the sister species *C. pruni* shows low recent TE activity, with most insertions being older. The TE peak is broad and right shifted, indicating a more stable TE landscape. Finally, *C. picta* has a sharp recent peak, especially in LINE and DNA elements, which is comparable to *C. mali* and *C. melanoneura*, making it one of the most dynamic genomes in this dataset. In summary, the species that show strong peaks at very low divergence values (0–2%), indicating ongoing or recent TE bursts, are *C. mali* from Clade II and all the species from Clade I, except *C. pruni*. The remaining

species from Clade II and *C. pruni* show a broader peak indicating relatively low and moderate divergence values (3–10%).

TE content contributes to genome size in *C. melanoneura* and *C. picta*

Given the substantial difference in estimated genome size between *C. melanoneura* and *C. picta*, we examined in more detail the contribution of repetitive elements to their genome size. Repetitive elements accounted for a moderate fraction of the assembled genome of both *C. melanoneura*, at 6.7%, and *C. picta*, at 12.2%. In both species, simple repeats were the most common element, followed by retroelements (LINEs, SINEs and LTRs), and DNA transposons (Fig. S9). This enrichment was consistent when considering total element count, length, and percentage of genome occupied. Focusing on comparative differences between species, simple repeats were more numerous and accounted for a larger proportion of the genome in *C. picta* compared to *C. melanoneura* (Fig. S9). A similar pattern was observed for retroelements, which was driven by the relative abundances of LINE elements (Fig. S9). Unclassified repetitive elements were also more common in *C. picta*. No notable differences in counts, length or percentage of genome occupied were found for SINE or LTR retroelements, sRNA or satellites, while DNA transposons were slightly more numerous in *C. picta*, although the percentage of genome occupied by these TEs was identical in the two species. Similar results were obtained using *RepeatMasker* with species-specific repeat libraries generated by *RepeatModeler2*, where the main difference observed was in the proportion of masked bases, 52.95% in *C. melanoneura* and 60.86% in *C. picta*. Discrepancies between the two approaches were primarily due to differences in the proportion of unclassified repetitive elements in both species (Fig. S9).

Retroelements and DNA transposons in the draft genome of *C. melanoneura* and *C. picta*

We further explored the genomic abundance in the draft genome assembly of both the retroelement and DNA transposon sub-classes. LINEs were the most abundant subclass in both *C. melanoneura* and *C. picta*, followed by SINEs and LTRs (Fig. S9; S10). However, LINEs were more abundant in *C. picta*, while SINEs and LTRs were more common in *C. melanoneura*. Most of the difference in LINE content between the two species is associated to the much higher number of *RTE/Bov-B* elements in *C. picta*. On the other hand, the higher abundance of LTRs in *C. melanoneura* is mostly attributed to *BEL/Pao* and *Gypsy/DIRS1* being more abundant in this species (Fig. S10). We also note that *Tcl-IS630-Pogo* and *hobo-activator* were the most common DNA transposons in both *C. melanoneura* and *C. picta*, the latter hosting in particular

almost twice as much *hobo-activator* than its related species (Fig. S10).

Discussion

Psyllids from the genus *Cacopsylla* are phytophagous insects with a wide range of host plants and include several vectors of plant pathogens. Although several *Cacopsylla* species are among the most important pests of various fruit trees, information on their evolutionary history, including their genetic and genomic makeup and evolution, remains scarce. We therefore conducted comparative genomic and phylogenetic analyses of *Cacopsylla* species, with a special focus on the two vectors of apple phytoplasma, ‘*Ca. P. mali*’: *C. melanoneura* and *C. picta*. One specific aim was to assess whether their ability to transmit ‘*Ca. P. mali*’ reflects a shared evolutionary origin or independent acquisition. For this purpose, we sequenced and assembled the nuclear and mitochondrial genomes of these species and, leveraging on publicly available data from other *Cacopsylla* and other psyllid species, reconstructed their phylogeny to trace their relationships with other vectors within the genus *Cacopsylla*.

The phylogenetic trees based on amino acid and nucleotide sequences of 12 psyllid species showed largely similar topologies, although divergence times were consistently estimated to be older when inferred from the amino acid data. In both, the amino acid and nucleotide phylogenies, *Euphyllura phillyreae* from the subfamily Euphyllurinae was recovered as paraphyletic with respect to the other analysed Liviinae representatives. In contrast, the family Liviidae was recovered as monophyletic in the all-codon trees. A similar pattern was reported by Percy et al. [26], who found that Liviidae is not monophyletic in either their conserved codon or combined datasets. Nevertheless, our amino acid and nucleotide phylogenies recovered stronger support for both deep and internal clades compared to the all-codon trees, suggesting that codon composition and substitution saturation may affect the inferred topology. This contrasts with findings from Wang et al. [72], where excluding the third codon position from nuclear nucleotide datasets did not resolve compositional heterogeneity or correct the inferred topology. We estimated that Psylloidea started diversifying during the Paleocene–Eocene, approximately at 52.1 or 62.3 million years ago (MYA), with confidence intervals ranging from 46.3 to 82.7 MYA. These estimates are substantially younger than those reported in the large-scale Hemiptera phylogenies of Li et al. [63] that placed the crown diversification of Psylloidea at ~ 120 MYA during the Early Cretaceous, while Johnson et al. [73] estimated it at ~ 220 MYA in the Late Triassic. This discrepancy may be explained by our inclusion of psyllid fossil calibrations, which constrained internal psyllid nodes, as well as our broader taxon sampling within

Psylloidea. Our results are also consistent with paleontological evidence presented by Ouvrard [64] and Drohojowska et al. [65, 74], which suggested that, based on fossil evidence, extant psyllids could not have diversified earlier than the mid-Eocene Lutetian stage. Although our point estimates predate the Lutetian, the lower bound of the 95% HPD interval extends into this stage, making our divergence estimates more consistent with both paleontological and molecular evidence than previous studies.

The crown divergence of the genus *Cacopsylla* was estimated to have occurred in the Early Miocene at 18.4 MYA (amino acid phylogeny) or in the Middle Miocene at 12.7 MYA (nucleotide phylogeny). These estimates are also in agreement with the fossil evidence of *Cacopsylla* dated to the Miocene [75]. Because the upper and lower bounds of the amino acid and nucleotide-based estimates also overlap, the approximate dating is reliable. Nevertheless, the amino acid or nucleotide-based estimate is probably the closest estimate of the true time of divergence given its better alignment with the available fossil evidence. This time estimation, based on 13 mitochondrial protein coding genes and psyllid fossil calibrations, contrasts with the results of a recent study by Darnis et al. [76], who estimated the crown divergence of *Cacopsylla* at approximately 85 MYA. Their analysis was based on a 621 bp fragment of the Cytochrome c oxidase subunit I-tRNA-COII region, combined with limited fossil evidence and external calibration derived from the phylogeny of Rosaceae. The discrepancy between the two studies may be attributed to the use of a single gene in the latter study as well as differences in calibration methods.

Our topologies across all trees consistently support a close relationship between *C. melanoneura* and *C. pruni*, the latter being the major vector of ‘*Ca. Phytoplasma prunorum*’, the causal agent of European Stone Fruit Yellows [77]. Interestingly, both *C. melanoneura* and *C. pruni* may represent cryptic species complexes, comprising genetically distinct lineages that are morphologically nearly identical and lack diagnostic morphological characters to distinguish them. These lineages are generally associated with different host preferences and show considerable molecular divergence within each nominal species [24, 28, 76, 78, 79].

Crucially, across nearly all analyses, *C. picta* was consistently placed as a sister taxon to a clade comprising four other Rosaceae-associated species, *C. melanoneura* and *C. pruni*, as well as two pear psyllids *C. burckhardti* and *C. pyrisuga*. This pattern suggests that the two primary vectors of AP, *C. melanoneura* and *C. picta*, likely evolved their vector competence independently rather than inheriting it from a common ancestor. Furthermore, both AP vector species are even more distantly related to another apple psyllid *C. mali*, which has not been confirmed as a vector of ‘*Ca. P. mali*’. Similarly, pear psyllids

C. pyri and *C. pyricola* (Clade II), as well as *C. burckhardti* and *C. pyrisuga* (Clade I), are distributed across distinct phylogenetic clades, suggesting independent evolutionary origins of pear-associated lineages. While *C. pyri* and *C. pyricola* may have inherited vector competence from a common ancestor, the more distantly related *C. pyrisuga* likely acquired this trait through a separate evolutionary pathway.

While *Cacopsylla* species do not cluster based on their capability to transmit phytoplasma or host preference, the two clades recovered in our study appear to differ in key biological traits. *Cacopsylla* Clade I consists of species which are strictly associated with Rosaceae. All species in this clade ((*C. burckhardti* + *C. pyrisuga*), (*C. melanoneura* + *C. pruni*), *C. picta*) are univoltine and migrate to conifers for overwintering (Table 1). This migratory behaviour may represent an adaptation to harsh environmental conditions and the seasonal unavailability of suitable phloem sources on deciduous hosts, serving as an overwintering strategy of many psyllid species in Northern temperate regions [23]. In contrast, most species assigned to Clade II (*C. pyri*, *C. pyricola*, *C. jukyungi*, *C. citrisuga*, *C. mali*) overwinter on their host plants without migrating to conifers. Although the genetic basis and evolution of the migratory behaviour in psyllids remain unclear, our data suggest a potential correlation between phylogenetic structure and ecological traits. Moreover, all species belonging to Clade I are restricted to the family Rosaceae, whereas several species from Clade II are associated with other plant families, i.e. *C. citrisuga* develops on Rutaceae and *C. coccinea* on Lardizabala-ceae. Additionally, Clade I contains only univoltine species, while various species belonging to Clade II are polyvoltine. This pattern is consistent with ecological strategies observed in psyllids from tropical and Southern temperate regions, where the psyllid species richness is the highest and continuous host availability facilitates multiple generations per year [1]. Taken together, these observations suggest that life history traits, such as voltinism and overwintering strategy, might be linked to the phylogenetic structure within *Cacopsylla*, and may help trace the evolutionary transitions in host association and overwintering behaviour. However, broader taxon sampling and mitogenome sequencing across the genus are needed to confirm this hypothesis.

We also investigated the repeatome to better understand the genomic processes shaping evolution in the *Cacopsylla* genus. We found that TEs dynamics vary widely across species, with lineage-specific patterns of accumulation and turnover likely contributing to genome evolution and adaptation. Species such as *C. mali*, *C. melanoneura*, *C. picta* and *C. pyri* show evidence of recent TE bursts, particularly of DNA transposons and LINE elements, suggesting ongoing genome re-modelling. In

contrast, *C. pruni*, *C. pyricola*, and *C. pyrisuga* exhibit older and likely more stable TE landscapes. According to our phylogenetic analyses, Clade I and Clade II are similar in age, although Clade I may be slightly younger. Moreover, the two clades differ in life history traits that may influence TE activity (Table 1). Clade I species inhabit less stable environments and are univoltine and migratory, with seasonal movement to conifers for overwintering. In contrast, most Clade II species are mostly polyvoltine, with several of them occupying more stable environments (without seasonal changes), which may reduce the stressful conditions that promote TE activity [80–82]. Environment aside, polyvoltine species may accumulate TEs faster due to shorter generation times [83], while univoltine species may experience slower TE accumulation and stronger selection against harmful insertions and more stable epigenetic silencing [84]. Additionally, the presence of seasonal migration in some species may also lead to periods of genomic “silence” potentially affecting TE activity compared to species that undergo continuous generations on their host plants. For example, studies have shown that in some insect groups, TEs are differentially expressed during diapause versus active developmental stages. In the bee species *Megachile rotundata* and *Osmia lignaria*, TE expression varied significantly during diapause and in response to temperature stress, suggesting that environmental conditions and life stage transitions can modulate TE activity [85]. Similarly, in *Drosophila montana*, a cold-adapted fly species, specific TEs were found to be upregulated during diapause, implying a role in regulating genes associated with diapause and cold tolerance [86].

We further compared the TEs in the shallow genome assemblies of *C. melanoneura* and *C. picta*. The primary finding was that simple repeats, LINE-class retrotransposons, and unclassified repeats were approximately twice as abundant in the genome of *C. picta* compared to *C. melanoneura*. The analysis of retroelements and DNA transposons revealed that these classes significantly contributed to the genome size differences between the two species. Retroelements such as RTE/Bov-B subclass were especially overrepresented in *C. picta*. In contrast, *C. melanoneura* had a greater proportion of non-assigned LINEs, although this did not affect the overall retroelement composition. Overall, while the composition of element types was broadly conserved, *C. picta* consistently exhibited a higher abundance of most repeat types.

Transposable elements may contribute to genomic variability and regulatory plasticity that may facilitate resistance to chemical contaminants such as pesticides, primarily through gene upregulation, loss of function, and alternative splicing. Among TE types, LINEs have been shown in *Drosophila melanogaster* to alter the splicing and regulation of genes involved in insecticide

resistance, conferring resistance to DDT and other insecticides [87, 88]. Similarly, in *Helicoverpa armigera*, several TE families, including LINEs, have been identified near detoxification genes, supporting the association between TE activity and gene regulation [89, 90]. Similarly, in the whitefly *Bemisia tabaci*, multiple TE insertions were reported, including six linked to known insecticide resistance genes, suggesting a role in both resistance and genome evolution [91]. In the aphid *Myzus persicae*, TEs contribute to resistance by disrupting a dominant susceptibility allele, enabling the expression of a recessive resistance allele [92]. These studies demonstrate how TEs can drive adaptation under environmental and genetic stressors highlighting a potential role of TEs in the adaptation of agriculturally important pest species such as *C. picta* and *C. melanoneura* on pesticides.

Conclusions

Cacopsylla is a diverse genus of phytophagous insects which includes several important vectors of phytoplasma, such as Apple Proliferation (AP), a serious threat to agriculture. Using a time-calibrated phylogenetic approach, we estimated the crown divergence of this genus between the Early Miocene (~18.4 MYA) and the Middle Miocene (~12.7 MYA). Within *Cacopsylla*, two major clades were identified, largely differentiated by their overwintering strategies and voltinism, ecological traits that might reflect adaptations to environmental changes over the past 10 million years. Interestingly, although *C. melanoneura* and *C. picta* are the only *Cacopsylla* species known to be vector of ‘*Ca. P. mali*’ and both develop on apple trees, we found them to be phylogenetically distantly related. The divergence of *C. picta* from the most recent common ancestor of *C. melanoneura* was estimated to have occurred between 8.6 and 11.5 MYA, during the Late Miocene. These findings reinforce the hypothesis that these two species vectoring AP phytoplasma followed independent evolutionary trajectories and likely acquired the ability to transmit ‘*Ca. P. mali*’ independently. Furthermore, *C. mali*, another apple-associated psyllid that is not a confirmed vector of ‘*Ca. P. mali*’, was found to belong to a distinct clade and was never recovered as a sister taxon to either *C. melanoneura* or *C. picta*. Based on the current results, it is not possible to clearly identify the ancestral clade, although Clade I appears relatively younger than Clade II. Finally, our analyses of transposable elements suggest that life history traits, such as host preference, overwintering strategy and voltinism, may influence recent genome rearrangements. These patterns are more pronounced in the representatives from the generally univoltine and overwintering Clade I, supporting the idea that some repetitive elements may play functional roles in seasonal migration, ecological adaptation, pathogen transmission

and potentially in the evolution of pesticide resistance. Future studies are needed to clarify how transposable element dynamics, and the expansion of repetitive sequences shape genomic plasticity and resilience in these agriculturally important pests.

Abbreviations

AP	Apple Proliferation
BF	Bayes Factor
BI	Bayesian Inference
BS	Bootstrap
COI	Cytochrome c Oxidase Subunit I
ESS	Effective Sampling Size
HPD	Highest Posterior Density
LINEs	Long Interspersed Nuclear Elements
LTRs	Long Terminal Repeats
MCC	Maximum Clade Credibility
ML	Maximum Likelihood
MLE	Marginal Likelihood Estimation
MRCA	Most Recent Common Ancestor
MYA	Million Years Ago
PCGs	Protein-Coding Genes
PP	Posterior Probabilities
RC	Rolling Circle Transposon
SINEs	Short Interspersed Nuclear Elements
TEs	Transposable Elements

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-026-12622-0>.

Supplementary Material 1.

Acknowledgements

We are very grateful to Daniel Burckhardt for his critical comments and suggestions. The authors thank the Department of Innovation, Research and University of the Autonomous Province of Bozen/Bolzano for covering the open access publication charges.

Authors' contributions

LR, LSS, EC, JMH, GA, CS, LO, ORS, and HS conceived and designed the study. HS, CS, GA and ORS secured the funding. EC, TW and LO coordinated the sequencing. FC and TA assembled the nuclear genomes. JMH assembled the mitochondrial genomes with input from NZ. LSS performed the phylogenetic analyses with input from ORS and JMH. LR and LO performed the transposable elements analyses. LR, LSS, EC and JMH wrote the paper with contribution from all authors. All authors discussed the results and approved the article.

Funding

This work was supported by a joint project funded by the province of Bolzano-Bozen and the Austrian Science Fund FWF (project I4639-B) to Hannes Schuler and Christian Stauffer, by a joint project of the province of Bolzano-Bozen and the Luxembourg National Research Fund FNR to Hannes Schuler and by a cofounded Fondazione Edmund Mach/APOT (project SCOPAZZI) to Omar Rota Stabelli, Guanfranco Anfora and Lino Ometto.

Data availability

NCBI BIO PROJECT: PRJNA1312285.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Competence Centre for Plant Health, Free University of Bozen- Bolzano, Bozen-Bolzano, Italy

²Center for Integrative Biodiversity Discovery, Leibniz Institute for Evolution and Biodiversity Science, Museum für Naturkunde, Berlin, Germany

³Museum of Nature Hamburg, Leibniz Institute for the Analysis of Biodiversity Change, Hamburg, Germany

⁴Department of Ecosystem Management, Climate and Biodiversity, BOKU University, Vienna, Austria

⁵Center for Cancer Research, Medical University of Vienna, Vienna, Austria

⁶Centro Nacional de Análisis Genómico (CNAG), Baldiri Reixac 4, Barcelona 08028, Spain

⁷Universitat de Barcelona (UB), Barcelona, Spain

⁸University of Trento, Center Agriculture Food Environment (C3A), Trento, Italy

⁹Centro Ricerca ed Innovazione, Fondazione Edmund Mach, San Michele all' Adige, Italy

¹⁰Department of Biology and Biotechnology, University of Pavia, Pavia, Italy

¹¹Faculty of Agricultural, Environmental and Food Sciences, Free University of Bozen-Bolzano, Bozen 39100, Italy

Received: 2 September 2025 / Accepted: 2 February 2026

Published online: 19 March 2026

References

- Burckhardt D, Ouvrard D, Percy DM. An updated classification of the jumping plant-lice (Hemiptera: Psylloidea) integrating molecular and morphological evidence. *Eur J Taxon.* 2021;736:137–82. <https://doi.org/10.5852/ejt.2021.736.1257>.
- Li D, Zhang C, Tong Z, Su D, Zhang G, Zhang S, et al. Transcriptome response comparison between vector and non-vector aphids after feeding on virus-infected wheat plants. *BMC Genomics.* 2020;21:638. <https://doi.org/10.1186/s12864-020-07057-0>.
- Jarausch B, Jarausch W. Psyllid vectors and their control. In: Weintraub PG, Jones P, editors. *Phytoplasmas: Genomes, plant hosts and vectors.* Wallingford, UK: CAB International; 2010.
- Rao GP, Bertaccini A, Fiore N, Lieferting LW, Phytoplasmas. *Plant Pathogenic Bacteria - I: Characterisation and Epidemiology of Phytoplasma - Associated Diseases.* 2018. <https://doi.org/10.1007/978-981-13-0119-3>.
- Seemüller E, Schneider B. *Candidatus* Phytoplasma mali, *Candidatus* Phytoplasma pyri and *Candidatus* Phytoplasma prunorum, the causal agents of Apple proliferation, Pear decline and European stone fruit yellows, respectively. *Int J Syst Evol Microbiol.* 2004;54:1217–26. <https://doi.org/10.1099/ijs.0.02823-0>.
- Janik K, Barthel D, Oppedisano T, Anfora G. Apple Proliferation: A Joint Review. 2020. ISBN 978-88-784-3054-9.
- Seemüller E, Carraro L, Jarausch W, Schneider B. CHAPTER 14: Apple proliferation Phytoplasma. In: Hadidi A, Barba M, Candresse T, Jelkmann W, editors. *Virus and Virus-Like diseases of pome and stone fruits.* APS; 2011. pp. 67–73. <https://doi.org/10.1094/9780890545010.014>.
- Tedeschi R, Jarausch B, Delic D, Weintraub PG. Actual distribution of fruit tree and grapevine Phytoplasma diseases and their vectors in Europe and neighboring regions. *Phytopathogenic Mollicutes.* 2013;3:3–4.
- Bertaccini A. COST Action FA0807 Phytoplasmas and phytoplasma disease management: how to reduce their economic impact. 2014. ISBN 978-88-909922-0-9.
- Hogenhout SA, Oshima K, Ammar E, Kakizawa S, Kingdom HN, Namba S. Phytoplasmas: bacteria that manipulate plants and insects. *Mol Plant Pathol.* 2008;9:403–23. <https://doi.org/10.1111/j.1364-3703.2008.00472.x>.
- Weintraub PG, Beanland L. Insect vectors of phytoplasmas. *Annu Rev Entomol.* 2006;51:91–111. <https://doi.org/10.1146/annurev.ento.51.110104.151039>.
- Burckhardt D, Ouvrard D. A revised classification of the jumping plant-lice (Hemiptera: Psylloidea). *Zootaxa.* 2012;3509. <https://doi.org/10.11646/zootaxa.3509.1.1>.

13. Tedeschi R, Alma A. *Fieberiella florii* (Homoptera: Auchenorrhyncha) as a vector of *Candidatus* Phytoplasma mali. Plant Dis. 2006;90:284–90. <https://doi.org/10.1094/PD-90-0284>.
14. Fischnaller S, Parth M, Messner M. Occurrence of different *Cacopsylla* species in Apple orchards in South Tyrol (Italy) and detection of Apple proliferation Phytoplasma in *Cacopsylla melanoneura* and *Cacopsylla picta*. Cicadina. 2017;17:37–51.
15. Jarausch B, Fuchs A, Schwind N, Krzczal G, Jarausch W. *Cacopsylla picta* as most important vector for *Candidatus* Phytoplasma mali in Germany and neighbouring regions. Bull Insectology. 2007;60:189–90.
16. Oppedisano T, Panassiti B, Pedrazzoli F, Mittelberger C, Bianchedi PL, Angeli G et al. Importance of psyllids' life stage in the epidemiology of apple proliferation phytoplasma. J Pest Sci. 2004;2020(93)49–61. <https://doi.org/10.1007/s10340-019-01130-8>.
17. Tedeschi R, Baldessari M, Mazzoni V, Trona F, Angeli G. Population dynamics of *Cacopsylla melanoneura* (Hemiptera: Psyllidae) in Northeast Italy and its role in the Apple proliferation epidemiology in Apple orchards. J Econ Entomol. 2012;105:322–8. <https://doi.org/10.1603/ec11237>.
18. Candian V, Monti M, Tedeschi R. Temporal dynamics of *Ca.* Phytoplasma mali load in the insect vector *Cacopsylla melanoneura*. Insects. 2020;11. <https://doi.org/10.3390/insects111090592>
19. Corretto E, Trenti M, Starhova Serbina L, Howie JM, Dittmer J, Kerschbamer C et al. Multiple factors driving the acquisition efficiency of apple proliferation phytoplasma in *Cacopsylla melanoneura*. J Pest Sci (2004). 2024;97:1299–314. <https://doi.org/10.1007/s10340-023-01699-1>.
20. Schuler H, Dittmer J, Borruso L, Galli J, Fischnaller S, Anfora G, et al. Investigating the microbial community of *Cacopsylla* spp. As potential factor in vector competence of Phytoplasma. Environ Microbiol. 2022;24:4771–86. <https://doi.org/10.1111/1462-2920.16138>.
21. Ouvrard D, Chalise P, Percy DM. Host-plant leaps versus host-plant shuffle: a global survey reveals contrasting patterns in an oligophagous insect group (Hemiptera, Psylloidea). Syst Biodivers. 2015;13:434–54. <https://doi.org/10.1008/014772000.2015.1046969>.
22. Burckhardt D, Ouvrard D, Queiroz D, Percy D. Psyllid Host-Plants (Hemiptera: Psylloidea): resolving a semantic problem. Fla Entomol. 2014;97:242–6. <https://doi.org/10.1653/024.097.0132>.
23. Hodkinson ID. Life cycle variation and adaptation in jumping plant lice (Insecta: hemiptera: Psylloidea): a global synthesis. J Nat Hist. 2009;43:65–179. <https://doi.org/10.1080/00222930802354167>.
24. Malagnini V, Pedrazzoli F, Papetti C, Cainelli C, Zasso R, Gualandri V, et al. Ecological and genetic differences between *Cacopsylla melanoneura* (Hemiptera, Psyllidae) populations reveal species host plant preference. PLoS ONE. 2013;8:e69663–69663. <https://doi.org/10.1371/journal.pone.0069663>.
25. Oetti S, Schlink K. Molecular identification of two vector species, *Cacopsylla melanoneura* and *Cacopsylla picta* (Hemiptera: Psyllidae), of Apple proliferation disease and further common psyllids of Northern Italy. J Econ Entomol. 2015;108:2174–83. <https://doi.org/10.1093/jee/tov204>.
26. Percy DM, Crampton-Platt A, Sveinsson S, Lemmon AR, Lemmon EM, Ouvrard D, et al. Resolving the psyllid tree of life: phylogenomic analyses of the superfamily psylloidea (Hemiptera). Syst Entomol. 2018;43:762–76. <https://doi.org/10.1111/syen.12302>.
27. Šafářová D, Zrníková E, Holušová K, Ouředníčková J, Starý M, Navrátil M. Molecular characterization of mitogenome of *Cacopsylla picta* and *Cacopsylla melanoneura*, two vector species of *Candidatus* Phytoplasma mali. Agronomy. 2023;13:2210. <https://doi.org/10.3390/agronomy13092210>.
28. Corretto E, Štarhová Serbina L, Dittmer J, et al. Evolutionary dynamics of obligate endosymbiosis in the psyllid genus *Cacopsylla*. Commun Biol. 2025;8:1540. <https://doi.org/10.1038/s42003-025-08904-0>.
29. Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, Hammel M, et al. Ten things you should know about transposable elements. Genome Biol. 2018;19:199. <https://doi.org/10.1186/s13059-018-1577-z>.
30. Wang X, Fang X, Yang P, Jiang X, Jiang F, Zhao D, et al. The locust genome provides insight into swarm formation and long-distance flight. Nat Commun. 2014;5:2957. <https://doi.org/10.1038/ncomms3957>.
31. Makalowski W, Gotea V, Pande A, Makalowska I. Transposable Elements: Classification, Identification, and Their Use As a Tool For Comparative Genomics. 2019. p.177–207. https://doi.org/10.1007/978-1-4939-9074-0_6.
32. Petersen M, Armisén D, Gibbs RA, Hering L, Khila A, Mayer G, et al. Diversity and evolution of the transposable element repertoire in arthropods with particular reference to insects. BMC Ecol Evol. 2019;19:11. <https://doi.org/10.1186/s12862-018-1324-9>.
33. Marçais G, Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics. 2011;27:764–70. <https://doi.org/10.1093/bioinformatics/btr011>.
34. Vurture GW, Sedlazeck FJ, Nattestad M, Underwood CJ, Fang H, Gurtowski J, et al. GenomeScope: fast reference-free genome profiling from short reads. Bioinformatics. 2017;33:2202–4. <https://doi.org/10.1093/bioinformatics/btx153>.
35. Simpson JT. Exploring genome characteristics and sequence quality without a reference. Bioinformatics. 2014;30:1228–35. <https://doi.org/10.1093/bioinformatics/btu023>.
36. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 2011;17. <https://doi.org/10.14806/ej.17.1.200>.
37. Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics. 2011;27:2957–63. <https://doi.org/10.1093/bioinformatics/btr507>.
38. Marco-Sola S, Sammeth M, Guigo R, Ribeca P. The GEM mapper: fast, accurate and versatile alignment by filtration. Nat Methods. 2012. <https://doi.org/10.1038/nmeth.2221>.
39. Wood D, Salzberg S. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 2014;15:R46–46.
40. Weisenfeld NI, Kumar V, Shah P, Church DM, Jaffe DB. Direct determination of diploid genome sequences. Genome Res. 2017;27:757–67. <https://doi.org/10.1101/gr.214874.116>.
41. Weisenfeld NI, Yin S, Sharpe T, Lau B, Hegarty R, Holmes L, et al. Comprehensive variation discovery in single human genomes. Nat Genet. 2014;46:1350–5. <https://doi.org/10.1038/ng.3121>.
42. Sahlin K, Vezzi F, Nystedt B, Lundeberg J, Arvestad L. BESST - Efficient scaffolding of large fragmented assemblies. BMC Bioinformatics. 2014;15:281. <https://doi.org/10.1186/1471-2105-15-281>.
43. Cruz F, Julca I, Gomez-Garrido J, Loska D, Marcet-Houben M, Cano E, et al. Genome sequence of the Olive tree, *Olea Europaea*. Gigascience. 2016;5:29. <https://doi.org/10.1186/s13742-016-0134-5>.
44. Puerma E, Orengo DJ, Cruz F, Gomez-Garrido J, Librado P, Salguero D, et al. The high-quality genome sequence of the oceanic Island endemic species *Drosophila guanche* reveals signals of adaptive evolution in genes related to flight and genome stability. Genome Biol Evol. 2018;10:1956–69. <https://doi.org/10.1093/gbe/evy135>.
45. Przytycki LP, Gabaldon T. Redundans: an assembly pipeline for highly heterozygous genomes. Nucleic Acids Res. 2016;44. <https://doi.org/10.1093/nar/gkw294>.
46. Mapleson D, Garcia Accinelli G, Kettleborough G, Wright J, Clavijo BJ. KAT: a K-mer analysis toolkit to quality control NGS datasets and genome assemblies. Bioinformatics. 2017;33:574–6. <https://doi.org/10.1093/bioinformatics/btw663>.
47. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31. <https://doi.org/10.1093/bioinformatics/bt v351>.
48. Smit AFA, Hubley R, Green P. RepeatMasker Open. <http://www.repeatmasker.org>
49. Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, et al. RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci. 2020;117:9451–7. <https://doi.org/10.1073/pnas.1210461117>.
50. Novák P, Neumann P, Macas J. Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. BMC Bioinformatics. 2010;11:378. <https://doi.org/10.1186/1471-2105-11-378>.
51. Novák P, Neumann P, Pech J, Steinhaisl J, Macas J. RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. Bioinformatics. 2013;29:792–3. <https://doi.org/10.1093/bioinformatics/btt054>.
52. Goubert C. Assembly-Free detection and quantification of transposable elements with DnaPipeTE. 2023. p.25–43. https://doi.org/10.1007/978-1-0716-2883-6_2.
53. Meng G, Li Y, Yang C, Liu S. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. Nucleic Acids Res. 2019;47:e63–63. <https://doi.org/10.1093/nar/gkz173>.
54. Kurtzer GM, Sochat V, Bauer MW, Singularity. Scientific containers for mobility of compute. PLoS ONE. 2017;12:e0177459–0177459. <https://doi.org/10.1371/journal.pone.0177459>.

55. Bernt M, Donath A, Juhling F, Externbrink F, Florentz C, Fritzsche G, et al. MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol.* 2013;69:313–9. <https://doi.org/10.1016/j.ympev.2012.08.023>.
56. Donath A, Juhling F, Al-Arab M, Bernhart SH, Reinhardt F, Stadler PF, et al. Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. *Nucleic Acids Res.* 2019;47:10543–52. <https://doi.org/10.1093/nar/gkz833>.
57. Bouckaert R, Vaughan TG, Barido-Sottani J, Duchene S, Fourment M, Gavryushkina A, et al. BEAST 2.5: an advanced software platform for bayesian evolutionary analysis. *PLoS Comput Biol.* 2019;15:e1006650–1006650. <https://doi.org/10.1371/journal.pcbi.1006650>.
58. Drummond AJ, Rambaut A. BEAST: bayesian evolutionary analysis by sampling trees. *BMC Evol Biol.* 2007;7:214. <https://doi.org/10.1186/1471-2148-7-214>.
59. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol.* 2020;37:1530–4. <https://doi.org/10.1093/molbev/msaa015>.
60. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30:772–80. <https://doi.org/10.1093/molbev/mst010>.
61. Bouckaert RR, Drummond AJ, bModelTest. Bayesian phylogenetic site model averaging and model comparison. *BMC Evol Biol.* 2017;17:42. <https://doi.org/10.1186/s12862-017-0890-6>.
62. Miller MA, Schwartz T, Hoover P, Yoshimoto K, Sivagnanam S, Majumdar A. The CIPRES workbench. In: Proceedings of the 2015 XSEDE Conference on Scientific Advancements Enabled by Enhanced Cyberinfrastructure - XSEDE '15. New York, New York, USA: ACM Press; 2015. p.1–8. <https://doi.org/10.1145/2792745.2792784>.
63. Li H, Leavengood JM Jr, Chapman EG, Burkhardt D, Song F, Jiang P, et al. Mitochondrial phylogenomics of hemiptera reveals adaptive innovations driving the diversification of true bugs. *Proc Biol Sci.* 2017;284. <https://doi.org/10.1098/rspb.2017.1223>.
64. Ouvrard. The oldest jumping plant-louse (Insecta: Hemiptera: Sternorrhyncha) with comments on the classification and nomenclature of the palaeogene psylloidea. *Acta Musei Moraviae Sci Biologicae (Brno).* 2013;98:21–33.
65. Burckhardt D, Drohojowska J, Štarhová Šerbina L, Malenovský I. First record of jumping plant lice of the family Liviidae (Hemiptera, Sternorrhyncha, Psylloidea) from Dominican amber. *Neues Jahrbuch für Geologie und Paläontologie - Abhandlungen Band.* 2024;215–27.
66. Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol Biol Evol.* 2012;29:2157–67. <https://doi.org/10.1093/molbev/mss084>.
67. Xie W, Lewis PO, Fan Y, Kuo L, Chen M-H. Improving marginal likelihood Estimation for bayesian phylogenetic model selection. *Syst Biol.* 2011;60:150–60. <https://doi.org/10.1093/sysbio/syq085>.
68. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* 2018;4. <https://doi.org/10.1093/ve/vey016>.
69. Jo E, Cho G. The complete mitochondrial genome of *Cacopsylla burckhardti* (Hemiptera, Psylloidea, Psyllidae). *Biodivers Data J.* 2022;10:e85094–85094. <https://doi.org/10.3897/BDJ.10.e85094>.
70. Que S, Yu L, Xin T, Zou Z, Hu L, Xia B. Complete mitochondrial genome of *Cacopsylla coccinae* (Hemiptera: Psyllidae). *Mitochondrial DNA DNA Mapp Seq Anal.* 2016;27:3169–70. <https://doi.org/10.3109/19401736.2015.1007319>.
71. Wang Y, Cen Y, He Y, Wu Y, Huang S, Lu J. The first complete mitochondrial genome sequence of *Cacopsylla citrisuga* (Yang & Li), a new insect vector of Huanglongbing in Yunnan Province, China. *Mitochondrial DNA B Resour.* 2021;6:575–7. <https://doi.org/10.1080/23802359.2021.1875908>.
72. Wang W, Dong Z, Du Z, Wu P. Genome-scale approach to reconstructing the phylogenetic tree of psyllids (superfamily Psylloidea) with account of systematic bias. *Mol Phylogenet Evol.* 2023;189:107924. <https://doi.org/10.1016/j.ympev.2023.107924>.
73. Johnson KP, Dietrich CH, Friedrich F, Beutel RG, Wipfler B, Peters RS et al. Phylogenomics and the evolution of hemipteroid insects. 2018;115. <https://doi.org/10.5061/dryad.t4f4g85>.
74. Drohojowska J, Węgierek P, Solórzano Kraemer MM. First psylloidea (Hemiptera: Sternorrhyncha) in miocene Mexican amber. *Palaentol Z.* 2016;90:185–8. <https://doi.org/10.1007/s12542-016-0290-z>.
75. Zhang T, Luo X, Yao Y, Ren D. A new fossil psyllid, *Cacopsylla trigona* sp. nov. (Hemiptera: psylloidea: Psyllidae), from the miocene of China. *Palaeoentomology.* 2019;2:413–7. <https://doi.org/10.11646/palaeoentomology.2.5.2>.
76. Darnis M, Ravigné V, Sauvion N. Origin. Divergence and migration routes of psyllids of the *Cacopsylla Pruni* complex (Hemiptera: Psylloidea) inferred by approximate bayesian computation methods. *J Biogeogr.* 2025;52. <https://doi.org/10.1111/jbi.15177>.
77. Steffek R, Follak S, Sauvion N, Labonne G, MacLeod A. Distribution of 'Candidatus Phytoplasma prunorum' and its vector *Cacopsylla Pruni* in European fruit-growing areas: a review. *EPPO Bull.* 2012;42:191–202. <https://doi.org/10.1111/epp.2567>.
78. Peccoud J, Labonne G, Sauvion N. Molecular test to assign individuals within the *Cacopsylla Pruni* complex. *PLoS ONE.* 2013;8:e72454. <https://doi.org/10.1371/journal.pone.0072454>.
79. Sauvion N, Peccoud J, Meynard C, Ouvrard D. Occurrence data for the two cryptic species of *Cacopsylla Pruni* (Hemiptera: Psylloidea). *Biodivers Data J.* 2021;9. <https://doi.org/10.3897/BDJ.9.e68860>.
80. Milyaeva PA, Kukushkina IV, Kim AI, Nefedova LN. Stress induced activation of LTR retrotransposons in the *Drosophila melanogaster* genome. *Life.* 2023;13:2272. <https://doi.org/10.3390/life13122272>.
81. de Oliveira DS, Rosa MT, Vieira C, Loreto ELS. Oxidative and radiation stress induces transposable element transcription in *Drosophila melanogaster*. *J Evol Biol.* 2021;34:628–38. <https://doi.org/10.1111/jeb.13762>.
82. Kneuss E, Munafò M, Eastwood EL, Deumer U-S, Preall JB, Hannon GJ, et al. Specialization of the *Drosophila* nuclear export family protein Nxf3 for PiRNA precursor export. *Genes Dev.* 2019;33:1208–20. <https://doi.org/10.1101/gad.328690.119>.
83. Lawlor MA, Ellison CE. Evolutionary dynamics between transposable elements and their host genomes: mechanisms of suppression and escape. *Curr Opin Genet Dev.* 2023;82:102092. <https://doi.org/10.1016/j.gde.2023.102092>.
84. Slotkin RK, Martienssen R. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet.* 2007;8:272–85. <https://doi.org/10.1038/nrg2072>.
85. Signor S, Yocum G, Bowsher J. Life stage and the environment as effectors of transposable element activity in two bee species. *J Insect Physiol.* 2022;137:104361. <https://doi.org/10.1016/j.jinsphys.2022.104361>.
86. Tahami MS, Vargas-Chavez C, Poikela N, Coronado-Zamora M, González J, Kankare M. Transposable elements in *Drosophila Montana* from harsh cold environments. *Mob DNA.* 2024;15:18. <https://doi.org/10.1186/s13100-024-00328-7>.
87. Schmidt JM, Good RT, Appleton B, Sherrard J, Raymont GC, Bogwitz MR, et al. Copy number variation and transposable elements feature in Recent, ongoing adaptation at the Cyp6g1 locus. *PLoS Genet.* 2010;6:e1000998. <https://doi.org/10.1371/journal.pgen.1000998>.
88. Daborn PJ, Yen JL, Bogwitz MR, Le Goff G, Feil E, Jeffers S et al. A Single P450 Allele Associated with Insecticide Resistance in *Drosophila*. *Science* (1979). 2002;297:2253–6. <https://doi.org/10.1126/science.1074170>.
89. Klai K, Chénais B, Zidi M, Djebbi S, Caruso A, Denis F, et al. Screening of *Helicoverpa armigera* mobilome revealed transposable element insertions in insecticide resistance genes. *Insects.* 2020;11:879. <https://doi.org/10.3390/insects11120879>.
90. Rahman M-M, Omoto C, Kim J. Genome-Wide exploration of long Non-Coding RNAs of *Helicoverpa armigera* in response to pyrethroid insecticide resistance. *Insects.* 2024;15:146. <https://doi.org/10.3390/insects15030146>.
91. Zidi M, Klai K, Confais J, Chénais B, Caruso A, Denis F, et al. Genome-Wide screening of transposable elements in the Whitefly, *Bemisia tabaci* (Hemiptera: Aleyrodidae), revealed insertions with potential insecticide resistance implications. *Insects.* 2022;13:396. <https://doi.org/10.3390/insects13050396>.
92. Panini M, Chiesa O, Troczka BJ, Mallott M, Manicardi GC, Cassanelli S et al. Transposon-mediated insertional mutagenesis unmasks recessive insecticide resistance in the aphid *Myzus persicae*. *Proceedings of the National Academy of Sciences.* 2021;118. <https://doi.org/10.1073/pnas.2100559118>.

Publisher's Note

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