METRICA
METRology Institutional CAtalog

ISTITUTO NAZIONALE DI RICERCA METROLOGICA
Repository Istituzionale

Distribution of β-casein variants and effects on milk composition in Podolian cows reared in Gargano Promontory (Southern Italy)

Original
Distribution of β-casein variants and effects on milk composition in Podolian cows reared in Gargano Promontory (Southern Italy) / Santillo, A., D'Angelo, F., Lamberti, C., Giuffrida, M.G., Romaniello, F., Albenzio, M.. - In: JOURNAL OF DAIRY SCIENCE. - ISSN 1525-3198. - 108:6(2025), pp. 5570-5579. [10.3168/jds.2025-26317]

Availability:
This version is available at: 11696/88159 since: 2026-02-25T08:33:32Z

Publisher:
Elsevier Inc.

Published
DOI:10.3168/jds.2025-26317

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

15 November 2026

scientific reports

Check for updates

OPEN

Rewiring the proteome of the *Euscelidius variegatus* holobiont in response to Flavescence dorée phytoplasma

Simona Abbà[1]✉, Marta Vallino[1], Simona Cirrincione[2], Cristina Lamberti[2], Beatrice Aiuto[2], Francesco Romaniello[3], Luciana Galetto[1], Cristina Marzachì[1], Domenico Bosco[4] & Marika Rossi[1]

The leafhopper *Euscelidius variegatus* is a laboratory vector of the phytoplasma associated to Flavescence dorée, a severe grapevine disease that threatens viticulture in Europe. Transcriptomic studies have already provided valuable insights into the mechanisms of insect-phytoplasma interactions, but proteomics can offer immediate insights into the cellular functions and metabolic adaptations of the insect and its microbiome to the presence of this plant bacterium. Here, the generation of new genomic data of the *E. variegatus* holobiont was instrumental in elaborating the first comprehensive proteomic profile of its response to Flavescence dorée phytoplasma (FDp). Both data-dependent acquisition and data-independent acquisition mass spectrometry were used to explore the complex molecular interactions between the insect host, its microbial community, and the phytoplasma. Results indicated a critical role of the insect mitochondria as a shared interface exploited by phytoplasmas for survival and propagation. Additionally, it appeared that the presence of FDp had a detrimental impact on the reciprocal metabolic support between the insect host and its two primary endosymbionts, predominantly resulting in a perturbation in amino acid synthesis and exchange. Proteins upregulated in response to FDp may represent promising targets for disrupting phytoplasma acquisition and transmission, either through rationally designed agrochemicals or gene silencing approaches.

Keywords Flavescence dorée, Genome assembly, Insect vector, DDA, DIA, Microbiome

Insect pests pose a significant threat to global agriculture, due to direct damage as well as to their ability to transmit a broad range of plant pathogens. Plant bacteria and viruses can have a profoundly deleterious effect on crops, resulting in substantial economic losses through decreased yields, diminished product quality, and elevated costs for disease control and prevention. The Hemiptera order includes several of the most important insect families that act as vectors of plant pathogens, including phytoplasmas. These are wall-less, obligate bacterial plant pathogens that infect a wide range of economically important crops worldwide[1]. A notable example is Flavescence dorée (FD), a quarantine plant disease[2] associated with the Flavescence dorée phytoplasma (FDp), which leads to severe vine decline and yield reduction across Europe. While the leafhopper *Scaphoideus titanus* Ball (Cicadellidae: Deltocephalinae) is recognized as the primary natural vector of FDp in vineyards[3], the phylogenetically close leafhopper *Euscelidius variegatus* Kirschbaum is a key model vector to study the mechanisms of pathogen acquisition and transmission under laboratory conditions[4]. Such mechanisms, as well as feeding behaviors, immune responses and reproductive success, can be strongly modulated by the composition of the insect microbiome[5,6]. As is the case with a multitude of insect species, *E. variegatus* harbors a diverse microbial community, including both obligate and facultative symbionts. Its primary endosymbionts, "*Candidatus* Nasuia deltocephalinicola" (*Pseudomonadota,* hereafter *Nasuia*) and "*Ca.* Karesulcia muelleri" (*Bacteroidota,* hereafter *Sulcia*), play a critical role in its survival[7]. These obligate endosymbionts are located within specialized cells (bacteriocytes) in the insect body and supply the insect host with essential nutrients

[1]Istituto per la Protezione Sostenibile delle Piante, Consiglio Nazionale delle Ricerche, IPSP-CNR, 10135 Turin, Italy. [2]Istituto di Scienze delle Produzioni Alimentari, Consiglio Nazionale delle Ricerche, ISPA-CNR, 10095 Grugliasco, TO, Italy. [3]Istituto Nazionale di Ricerca Metrologica, INRIM, 10135 Turin, Italy. [4]Dipartimento di Scienze Agrarie, Forestali ed Alimentari, Università degli Studi di Torino, 10095 Grugliasco, TO, Italy. ✉email: simona.abba@cnr.it

(e.g., amino acids, vitamins) that are typically absent from the insect phloem-sap diet[8–11]. In addition to these essential primary endosymbionts, *E. variegatus* may harbor a secondary, facultative, endosymbiont, initially known as 'BEV' and now classified as "*Ca.* Symbiopectobacterium sp" (*Pseudomonadota,* hereafter BEV)[12]. Unlike the obligate primary symbionts, BEV can be cultured in vitro, but its potential functional role in the biology and vector competence of *E. variegatus* is still under investigation. The genome of another symbiotic bacterium related to the *Asaia* species (*Pseudomonadota*) has been recently isolated from a laboratory rearing of *E. variegatus* and named *Sorlinia euscelidii*[13]. Furthermore, the iflavirus EVV1 has been identified as a stable component of the microbiome of the *E. variegatus* population reared in our laboratory[14,15].

Metagenomic and transcriptomic approaches have previously been applied to *E. variegatus* and other insect vectors of phytoplasmas, providing valuable insights into their biology and interactions with these plant bacteria[16,17]. However, genome-wide proteomics analyses remain a largely overlooked technique not only in the study of phytoplasma vectors, but more broadly across insect vectors. To date, most proteomic studies on vector-pathogen interactions have focused on blood-feeding insects transmitting human pathogens[18,19] or plant-infecting viruses[20–23], with limited attention given to vectors of bacterial plant pathogens[24]. Proteomics provides a robust methodology to comprehensively analyze the functional state of an organism and its molecular responses to environmental stimuli or pathogen challenges. Unlike genomics or transcriptomics, proteomics offers a direct measurement of the effector molecules of biological processes, thereby providing immediate insights into cellular functions and metabolic adaptations.

In this study, we used an integrative approach that combines newly generated genomic data of the *E. variegatus* holobiont (i.e. the insect vector and its associated microbiome) with existing transcriptomic resources to provide the first comprehensive proteomic profile of the holobiont response to FDp acquisition. Specifically, we compared two Liquid Chromatography-Tandem Mass (LC–MS/MS) spectrometry acquisition methods: Data-Dependent Acquisition (DDA) and Data-Independent Acquisition (DIA). DIA was ultimately selected due to its enhanced sensitivity and improved detection of low-abundance peptides, such as those that may originate from the insect microbiome[25].

Results

Construction of a reference proteomic data set for the *E. variegatus* holobiont

As a non-model organism, *Euscelidius variegatus* lacked sufficient genomic data to support the proteomic coverage required by software tools designed for analyzing large mass-spectrometric data sets. To address this limitation, the first step of our study was the construction of a comprehensive reference proteomic database for the *E. variegatus* holobiont. Following the holobiont genome assembly, we successfully reconstructed the complete genomes of its two primary symbionts *Sulcia* and *Nasuia* and the strain of *Sorlinia euscelidii* associated to our lab-reared *E. variegatus* population (Table 1). In addition, we unexpectedly obtained the complete genome of *Pantoea ananatis* (*Pseudomonadota,* hereafter *Pantoea*) (Table 1), a bacterium previously reported to be a component of the microbiome of other insects[26–29].

As regards the genomes of *E. variegatus* and the facultative symbiont BEV, the total length of the assembled scaffolds exceeded the estimated size of their respective reference genomes. However, the assembly remained fragmented and did not reach a chromosome-level resolution (Supplementary Information 1). In particular, BUSCO analysis estimated the completeness of the *E. variegatus* genome at 84.5% with respect to the hemipteran core proteome (Table 2). To enhance this completeness, we implemented a stepwise integration of *E. variegatus* transcriptomic data and *Macrosteles quadrilineatus* proteome (NCBI RefSeq assembly: GCF_028750875.1) into our reference database for mass spectrometry analysis, despite potential protein redundancy (Table 2). *Macrosteles quadrilineatus* is the closest phylogenetic species of *E. variegatus* with an available chromosome-level genome assembly. Successive integrations improved the completeness of the core hemipteran gene set, culminating in an almost 100% representation. Proteins with identical amino acid sequences were removed from the dataset.

To achieve a comprehensive representation of the *E. variegatus* holobiont, the reference database used for proteomic profiling (ProteomeXchange dataset identifier: PXD066715) ultimately included: the insect proteome,

	Candidatus Karelsulcia muelleri	*Candidatus* Nasuia deltocephalinicola	*Pantoea ananatis*	*Sorlinia euscelidii*
Genome length (bp)	188,541	107,039	4,666,554	2,331,141
Genes (total)	226	165	4543	2125
CDSs (total)	191	132	4420	2061
Genes (coding)	189	132	4341	2055
CDSs (with protein)	189	132	4341	2055
Genes (RNA)	35	33	123	64
rRNAs	1, 1, 1 (5S, 16S, 23S)	1, 1 (16S, 23S)	8, 7, 7 (5S, 16S, 23S)	3, 3, 3 (5S, 16S, 23S)
Complete rRNAs	1, 1, 1 (5S, 16S, 23S)	1, 1 (16S, 23S)	8, 7, 7 (5S, 16S, 23S)	3, 3, 3 (5S, 16S, 23S)
tRNAs	30	30	79	51
ncRNAs	2	1	22	4
Pseudo genes (total)	2	0	79	6

Table 1. Assembly statistics of the genome of the two primary and two facultative symbionts of *Euscelidius variegatus*. bp, base-pair; CDSs, coding sequences; ncRNA, non-coding RNA.

Total number of core genes used as reference set (Hemiptera)	2510		
	E. variegatus genome	+ *E. variegatus* transcriptome	+ *Macrosteles quadrilineatus* proteome
Number (percentage) of complete core genes detected	2121 (84.50%)	2462 (98.09%)	2508 (99.92%)
Number (percentage) of complete and partial core genes detected	2289 (91.20%)	2289 (91.20%)	2509 (99.96%)
Number (percentage) of missing core genes	221 (8.80%)	26 (1.04%)	1 (0.04%)

Table 2. Number of core proteins present in our reference proteomic database as assessed by BUSCO using the Hemiptera-specific reference set. BUSCO statistics were calculated under three conditions: (column 1) using the *Euscelidius variegatus* genome alone; (column 2) after adding *E. variegatus* transcriptomic data; and (column 3) after additionally incorporating the *Macrosteles quadrilineatus* proteome. Each column represents a cumulative addition to the reference data used in the previous step, highlighting how successive integration improves completeness of the core hemipteran gene set, culminating in an almost 100% representation.

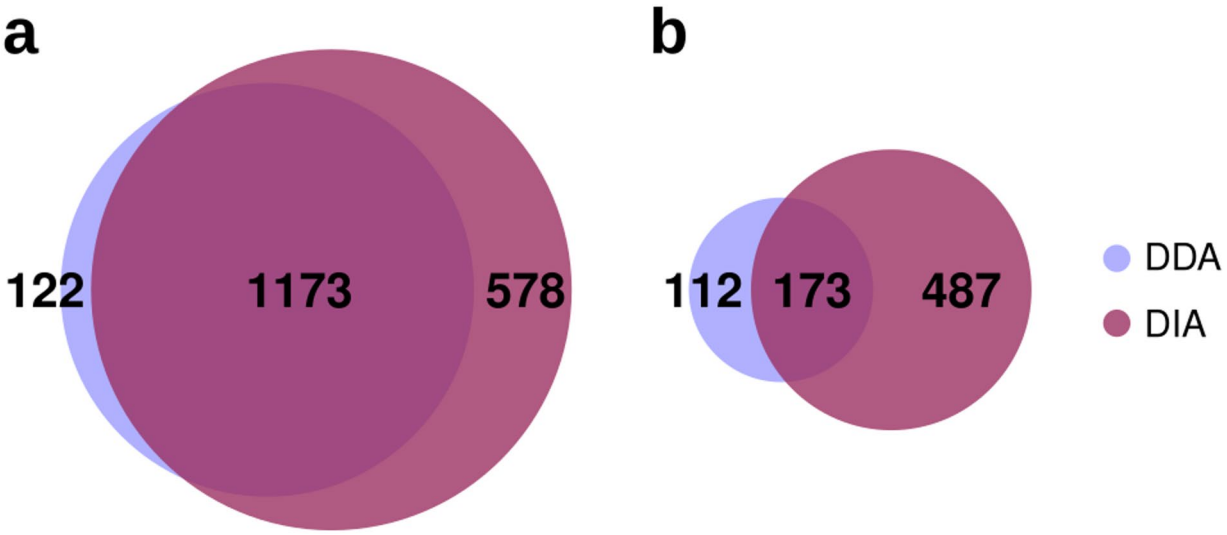


Fig. 1. Venn diagrams showing the results of the proteomic analyses conducted using DDA and DIA methods. (**A**) The total number of proteins identified by DDA only, by DIA only, and by both methods. (**B**) The total number of differentially expressed proteins (calculated as the sum of insect- and microbiome-derived proteins) identified by DDA only, by DIA only, and by both methods. 112: 108 insect proteins + 4 microbiome-derived proteins; 173: 168 insect proteins + 5 microbiome-derived proteins; 478: 461 insect protein + 17 microbiome-derived proteins.

the predicted proteins from the five newly sequenced bacterial genomes, the polyprotein of Euscelidius variegatus iflavirus 1 (EVV1, GenBank accession: NC_032087), and the predicted proteome of FDp (GenBank accession: CP097583).

Evaluating the performance of DDA and DIA in proteomic profiling

To compare protein identification across methods, we focused on the number of protein groups, since our aim was to evaluate overall expression rather than specific protein isoforms or post-translational modifications. Additionally, since the insect proteome used in this study comprises predicted proteins from the *E. variegatus* genome, the *E. variegatus* transcriptome and the *M. quadrilineatus* proteome, there is likely to be a high number of protein isoforms. This may limit the identification of unique peptides and consequently affect the reliability of individual protein quantification[30].

Overall, combining all the proteins identified in the eight samples, we identified 1295 distinct protein groups using the DDA approach compared to the 1751 protein groups identified with the DIA approach, representing an increase of 35%. While at least one member of 1173 protein groups was commonly identified by both methods (Fig. 1A), the higher number of protein groups detected exclusively by DIA highlights its superior efficiency and sensitivity in protein identification (ProteomeXchange dataset identifier: PXD066715). The integration of the *Macrosteles* proteome contributed approximately 4% and 16% of the total proteins identified by DDA and DIA, respectively.

Upon further analysis, the DDA approach enabled the identification of an average of 1154 insect proteins in non-exposed samples (H) and 1134 proteins in FDp-exposed samples (FD). The DIA approach facilitated the identification of a higher average number of proteins, with 1579 detected in non-exposed samples (+37%) and 1272 in FDp-exposed samples (+12%). For microbiome-related proteins, the DDA approach identified an average of 34 proteins in non-exposed samples and 31 proteins in FDp-exposed samples, whereas the DIA

approach detected 39 (+15%) and 36 (+16%) proteins in the respective groups. Notably, in FDp-exposed insects, two phytoplasma membrane proteins (the immunodominant membrane protein—imp and the Variable Membrane Protein A—VmpA) were detected by the DIA approach. Overall, very few proteins were detected from *Sorlinia, Nasuia* and *Pantoea*, whereas BEV and *Sulcia* appeared to be the most metabolically active symbionts, as indicated by the higher number of proteins identified.

Identification of differentially expressed insect proteins between FDp-exposed and non-exposed insects

DDA analysis identified 276 differentially expressed (DE) insect proteins between FDp-exposed and non-exposed insects, with 78 upregulated and 198 downregulated (false discovery rate [FDR] < 0.05, fold change ≥ ± 0.5; Supplementary Information 2, Fig. 2A). In comparison, the DIA approach identified 638 DE protein, with 64 upregulated and 574 downregulated (false discovery FDR < 0.05, fold change ≥ ± 0.5; Supplementary Information 2, Fig. 2B) representing a 31% increase in the number of DE proteins detected relative to DDA. Of these, 168 DE proteins were commonly identified by both methods (Fig. 1B). Notably, the DIA dataset revealed a considerable number of proteins (137) that were entirely absent in FDp-exposed samples but consistently detected in at least three out of four replicates from non-exposed insects. This threshold (i.e. presence in at least three replicates) was selected based on the expression pattern of two FDp-derived proteins, which were consistently detected in at least three replicates of FDp-exposed samples and absent in all non-exposed replicates. This criterion ensures robust identification of condition-specific expression while minimizing false positives.

To gain functional insights, we analyzed the pathways associated with the DE proteins using the KEGG database (http://www.kegg.jp/kegg/), focusing on BRITE functional hierarchies assigned to proteins upregulated and downregulated in response to FDp infection (Supplementary Information 3). A total of 17 and 29 BRITE functional categories were identified for upregulated and downregulated proteins, respectively, excluding the unclassified proteins. While this classification provided only a broad functional overview of the DE proteins, the results of the GO enrichment analysis (Fig. 3A–D and Supplementary Information 4) highlighted the significantly enriched molecular functions (MF), biological processes (BP), and cellular components (CC) associated to DE proteins. Regarding DE proteins identified by DIA, significantly overrepresented terms were detected in all three GO ontologies for both upregulated and downregulated proteins, with the only exception of the MF ontology among the upregulated proteins, where no significant enrichment was observed (Supplementary Information 4). Although the number of upregulated proteins was approximately nine times lower than that of downregulated proteins, GO enrichment analysis of the biological process ontology identified 27 significantly enriched terms among the upregulated proteins, compared to only three among the downregulated ones. Therefore, it appears that the upregulated proteins were more functionally clustered, leading to their statistical over-representation. In contrast, the downregulated proteins appeared to be spread across a broader range of biological processes, with only a few reaching statistical significance.

DIA results revealed that phytoplasma infection has a pronounced detrimental impact on mitochondrial functions, particularly affecting the respiratory chain complex (Fig. 3D). In contrast, phytoplasma infection led to a statistically significant increment of proteins associated to the extracellular space, i.e. proteins within the cellular matrix or associated to the outer membranes (Fig. 3C). In this regard, it is important to note that some DE predicted proteins that did not show any significant hits (E-value > 0.0001) in the NCBI nr protein database (as indicated by an asterisk in column "best_hit_blastp", Supplementary Information 2) were predicted to contain a signal peptide (as reported in column "Notes", Supplementary Information 2).

Both the DDA and DIA approaches consistently demonstrated that pathways associated with neuronal activity were particularly upregulated in FD-exposed insects (Fig. 3A, B). In addition, DIA data specifically revealed an overrepresentation of proteins linked to fatty acid biosynthesis and responses to external stimuli and sucrose (Fig. 3B).

Identification of differentially expressed microbiome proteins between FDp-exposed and non-exposed insects

Regarding microbiome proteins differentially expressed in response to FDp, the DDA approach identified 9 DE proteins (4 upregulated, 5 downregulated), while the DIA method identified 22 (7 upregulated, 15 downregulated) (Fig. 1B), representing a 44% increase. Of these, five proteins were detected by both methods. Importantly, FDp proteins were excluded from the count of upregulated proteins in FDp-exposed insects to focus solely on microorganisms common to both conditions. Although DIA revealed a higher number of DE microbiome proteins than DDA, the total number remained too low to yield statistically significant results in GO enrichment analysis. Notably, the 15 downregulated proteins largely originated from the two primary endosymbionts of *E. variegatus*, with nine derived from *Sulcia* and four from *Nasuia*. These proteins included three chaperonins and several key enzymes involved in amino acid biosynthesis, specifically within the valine, leucine, and isoleucine biosynthesis pathway, as well as the phenylalanine, tyrosine, and tryptophan biosynthesis pathway. In contrast, all upregulated microbiome-derived proteins were associated with the secondary endosymbiont BEV. Notably, among the upregulated proteins was a major capsid protein (GenBank Accession: QVJ82353.1), previously identified as part of a BEV temperate phage[31]. Additionally, three membrane proteins were identified: OmpA porin, involved in cell envelope stability, adhesion, invasion, biofilm formation, and phage recognition; KdgM, an oligogalacturonate-specific porin and a subunit of the ribose ABC transporter.

Discussion

Recent advances in instrumentation and strategies for data acquisition in mass spectrometry have significantly enhanced the levels of accuracy and sensitivity of proteomics so that it can now provide a more detailed and dynamic description of biological systems compared to transcriptomics, particularly by capturing the post-

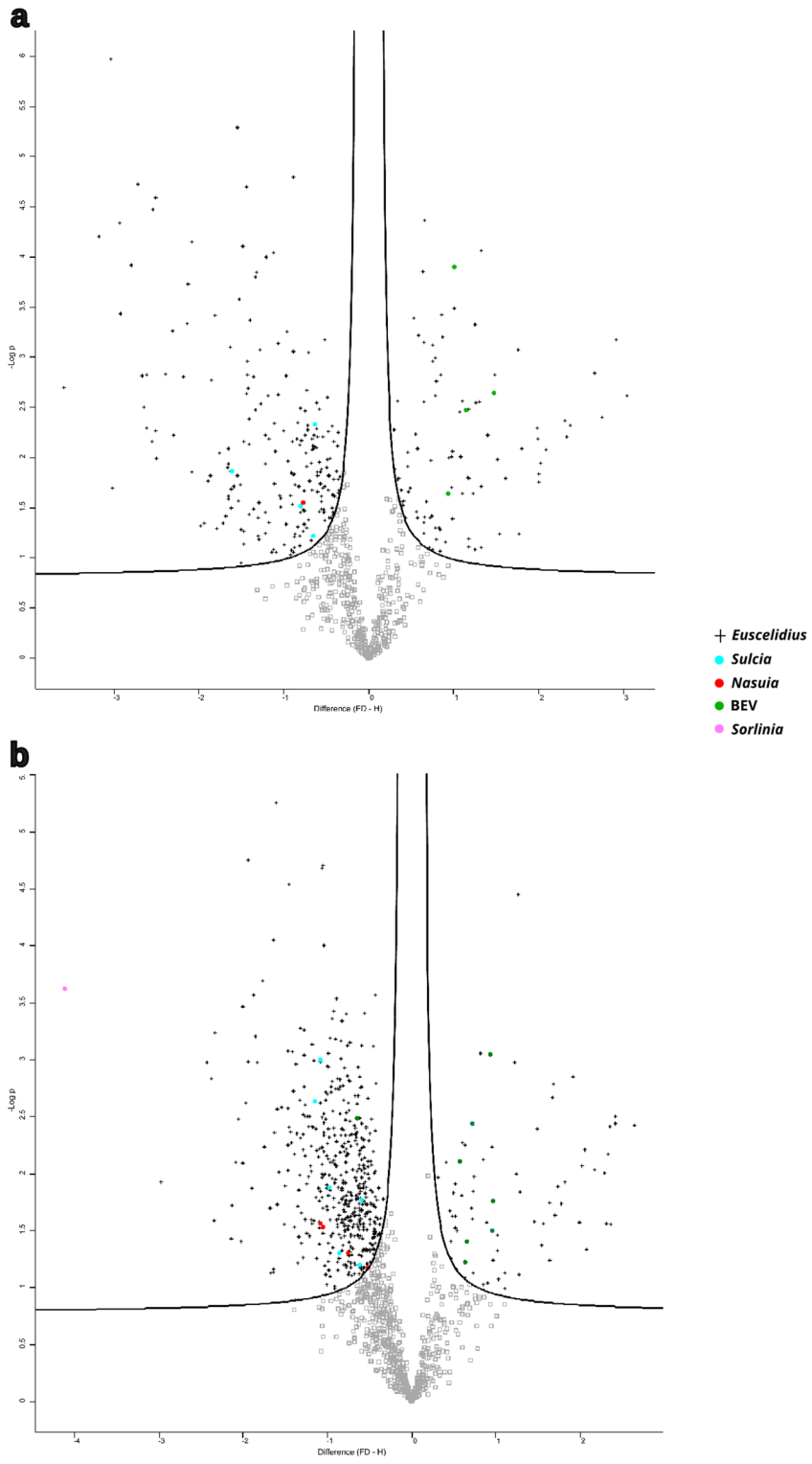


Fig. 2. Volcano plots of differentially expressed proteins identified using Perseus. (**A**) Volcano plot of the differentially expressed proteins as identified by Perseus following Data-Dependent Acquisition method. (**B**) Volcano plot of the differentially expressed proteins as identified by Perseus following Data-Independent Acquisition method. Proteins with $|\log_2 FC| > 0.5$ and $-\log_{10}(\text{p-value}) > 1.3$ are represented by black crosses (insect), light blue dots (*Sulcia*), greeen dots (BEV), red dots (*Nasuia*) and pink dots (*Sorlinia*). Positive $\log_2 FC$ indicates upregulation in FD-exposed samples compared to the control, while a negative $\log_2 FC$ indicates downregulation. The x-axis represents the Log_2 fold change ($\log_2 FC$). The y-axis represents the $-\log_{10}(\text{p-value})$ obtained using a two-sample t-test with a permutation-based false discovery rate threshold of 5%

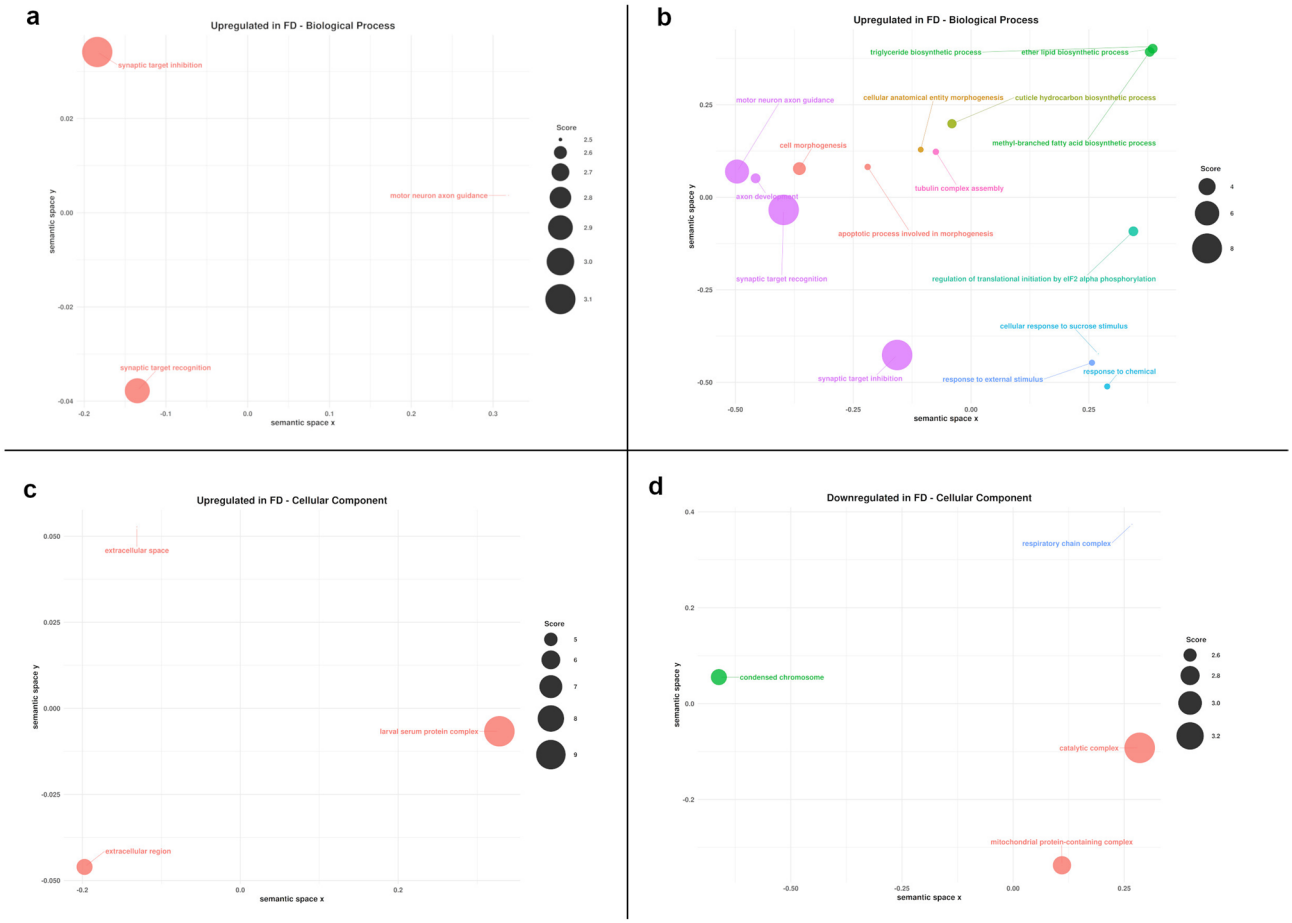


Fig. 3. Scatterplots showing semantic similarity of enriched GO terms among differentially expressed insect proteins in response to FDp. (**A**) Enriched Gene Ontology (GO) terms in the Biological Process (BP) ontology for upregulated proteins (DDA); (**B**) Enriched GO terms in the BP ontology for upregulated proteins (DIA); (**C**) Enriched GO terms in the Cellular Component (CC) ontology for upregulated proteins (DIA); (**D**) Enriched GO terms in the CC ontology for downregulated proteins (DIA). GO terms significantly enriched among differentially expressed proteins in FDp-exposed insects (see Supplementary Information 4) were analyzed using the rrvgo R package. Results are divided by GO ontologies: BP, CC and Molecular Function (MF). Redundant GO terms were reduced based on semantic similarity (threshold = 0.7), and the remaining representative terms were visualized in a two-dimensional semantic space. Therefore, the number of GO terms shown in the graph is lower than the one listed in Supplementary Information 4. Point size reflects the score, which is calculated as $-\log_{10}$(q-value). The q-value indicates statistical significance while controlling the False Discovery Rate. GO terms with higher scores are more statistically significant. Only GO terms with a score greater than 2 (i.e. a q-value < 0.01) are shown. Graphs were generated only when more than three enriched terms were identified within a GO category.

transcriptional and translational regulations of gene expression. However, a fundamental constraint of proteomic analyses is their reliance on comprehensive reference databases, a limitation that is particularly pronounced when working with non-model organisms. The initial step of our study was the creation of the most extensive possible database by sequencing the genome of the *E. variegatus* holobiont and predicting all its putative protein-coding genes.

The cumulative length of the obtained genomic scaffolds for *E. variegatus* should cover approximately the expected genome size of around 1.3 Gb, as estimated based on the chromosome-level genome assembly of the closest phylogenetically related insect, *M. quadrilineatus*[32]. Despite that, BUSCO analysis demonstrated that *E. variegatus* genome only encompasses 84.5% of the expected core proteome of hemipterans. Consequently, some scaffolds are likely to be redundant. This redundancy likely persisted because bioinformatic algorithms were unable to merge them, probably due to the large quantity of repetitive elements (e.g., tandem repeats and interspersed repeats) that have been previously observed in other leafhopper genomes[32–34].

Due to the current incompleteness of the *E. variegatus* genome, integrating genome-derived proteins with *Euscelidius* transcriptomic-derived proteins and proteomic information from *M. quadrilineatus* was essential for improving downstream protein identification. This approach is consistent with previous studies that have emphasized the difficulties associated with proteomics in non-model organisms with limited

genomic resources[35–37]. Although BUSCO analysis indicated that the *Macrosteles* proteome contributed only around 2% to the completeness of *E. variegatus* core proteome, it accounted for between 4 and 16% of the total identified protein groups, depending on the acquisition method used. This discrepancy likely reflects the higher number and accuracy of protein predictions derived from the chromosome-level genome assembly of *Macrosteles*, which includes not only core proteins, but also *Cicadellidae*/*Deltocephalinae*-specific proteins and isoforms. Consequently, certain conserved proteins, along with some missing *Euscelidius* proteins, are likely to be more accurately predicted and identified in the *Macrosteles* proteomic database than in the *E. variegatus* one. Additionally, the DIA approach resulted in the identification of a greater number of differentially expressed *Macrosteles* proteins, as well as more proteins in general, when compared to the DDA approach. While effective for identifying abundant proteins, DDA is frequently subject to limitations such as reduced reproducibility and suboptimal coverage of lower-abundance proteins. Conversely, DIA generates a comprehensive, high-resolution dataset that includes virtually all detectable peptides. It facilitates the identification of differentially expressed proteins, including the less abundant ones, such as those derived from the insect microbiome. Such particular achievement was supported by the reconstruction of complete, high-quality genomes of the two primary endosymbionts (*Sulcia* and *Nasuia*), two associated bacteria (*S. euscelidii* and *P. ananatis*), and a partial genome of BEV.

In view of the aforementioned advantages, DIA-based results were considered more comprehensive than those derived from DDA to provide a proteomic foundation for exploring the response of the insect and its microbiome to the challenge posed by the phytoplasma.

Notably, the presence of the plant pathogen in FDp-exposed insects was confirmed by the identification of two FDp surface-exposed proteins: VmpA and imp. VmpA facilitates the phytoplasma adhesion to specific molecules on the surface of insect midgut cells[38], while Imp is involved in interactions between FDp and insect gut proteins[39,40].

Overall, the presence of phytoplasma resulted in a greater number of proteins being downregulated than upregulated, indicating a general suppression of protein expression compared to control conditions. This observation aligns with previous findings reporting that FDp has detrimental effects on the survival and fecundity of its natural insect vector, *S. titanus*[41]. Similarly, infection with certain viruses, such as baculoviruses, can lead to a shutdown of protein synthesis in insect cells[42]. This strategy is employed by viruses to suppress the host's immune response and create an environment that is more favorable for their own replication. In particular, according to the GO enrichment analysis conducted on DIA data, *E. variegatus* responded to FDp with a downregulation of proteins associated with the mitochondrial compartment, including key enzymes involved in the respiratory electron transport chain. This pattern suggests a potential impairment of mitochondrial energy metabolism in response to infection. Indeed, previous RNAi experiments have shown that RNAi-mediated silencing of ATP synthase β in *E. variegatus* leads to a significant reduction in phytoplasma multiplication, likely due to ATP limitation[43]. It is plausible that the downregulation of respiratory chain components and mitochondrial functions observed in phytoplasma-infected insects may represent an insect defense mechanism aimed at limiting phytoplasma proliferation. Additionally, in *Euscelidius variegatus*, the Major Antigenic Membrane Protein (Amp) of *"Candidatus* Phytoplasma asteris" has been shown to selectively interact with the beta subunit of the mitochondrial ATP synthase[44], suggesting that phytoplasma may exploit host cell energy metabolism during colonization of the insect vector. No data is present in literature about the interaction of phytoplasma and insect mitochondria; however, in plants this association has been observed occasionally. In *Arabidopsis thaliana* infected with FDp, phytoplasma cells were found near mitochondria, particularly in root tissues[45]; similarly, in the phloem of tomato plants infected with the Columbia Basin PPT phytoplasma, bacterial cells were present at the contact site between the endoplasmic reticulum and mitochondria[46]. Although this association has not been studied across other phytoplasma–plant-vector interactions, it raises the possibility that phytoplasmas may similarly rely on host mitochondria as an external source of ATP. This is consistent with genomic evidence indicating that phytoplasmas lack key components of the ATP synthesis machinery[47]. The potential recruitment or manipulation of mitochondria by phytoplasmas may therefore serve a dual function: not only to access host-derived energy, but also as a strategy to interfere with mitochondrial-mediated defense signaling. Together, these observations point to a conserved and potentially critical role of host mitochondria in phytoplasma-host interactions across both plant and insect systems, warranting further investigation into mitochondria as a shared interface exploited by phytoplasmas for survival and propagation.

GO enrichment analysis also showed that axogenesis, lipid biosynthesis and response to external stimuli were overrepresented among the upregulated proteins in *E. variegatus* following exposure to FDp. We can speculate that the presence of phytoplasma induces changes in insect behaviour, such as increased movement or altered feeding habits, as demonstrated in other insect vectors[48,49]. The enrichment of axogenesis-related proteins suggests possible remodeling or development of neural circuits, which could affect sensory perception or motor function in response to external stimuli. Additionally, the upregulation of lipid biosynthesis pathways, particularly ether lipids and triacylglycerols (TAGs), could impact various physiological functions, including changes in membrane composition that may be required to accommodate the phytoplasma. In addition to their structural and energetic roles, lipids play crucial functions in antioxidant defense, cell signaling, immune regulation, and synaptic activity within the nervous system[50].

Arylphorins (a major class of hexamerins) and other hexamerins form the larval serum protein complex (GO: Cellular Component), a large multimeric storage complex that functions as an energy and amino-acid reserve in insects[51]. The increased abundance of these proteins upon FDp exposure may arise from two processes: (i) a compensatory response to phytoplasma-driven depletion of nutrient reserves normally stored in hexamerins, and/or (ii) their involvement in the humoral immune response to control the phytoplasma, similarly to what has been described for another leafhopper, *Circulifer haematoceps*, in response to the plant-pathogenic bacterium *Spiroplasma citri*[52].

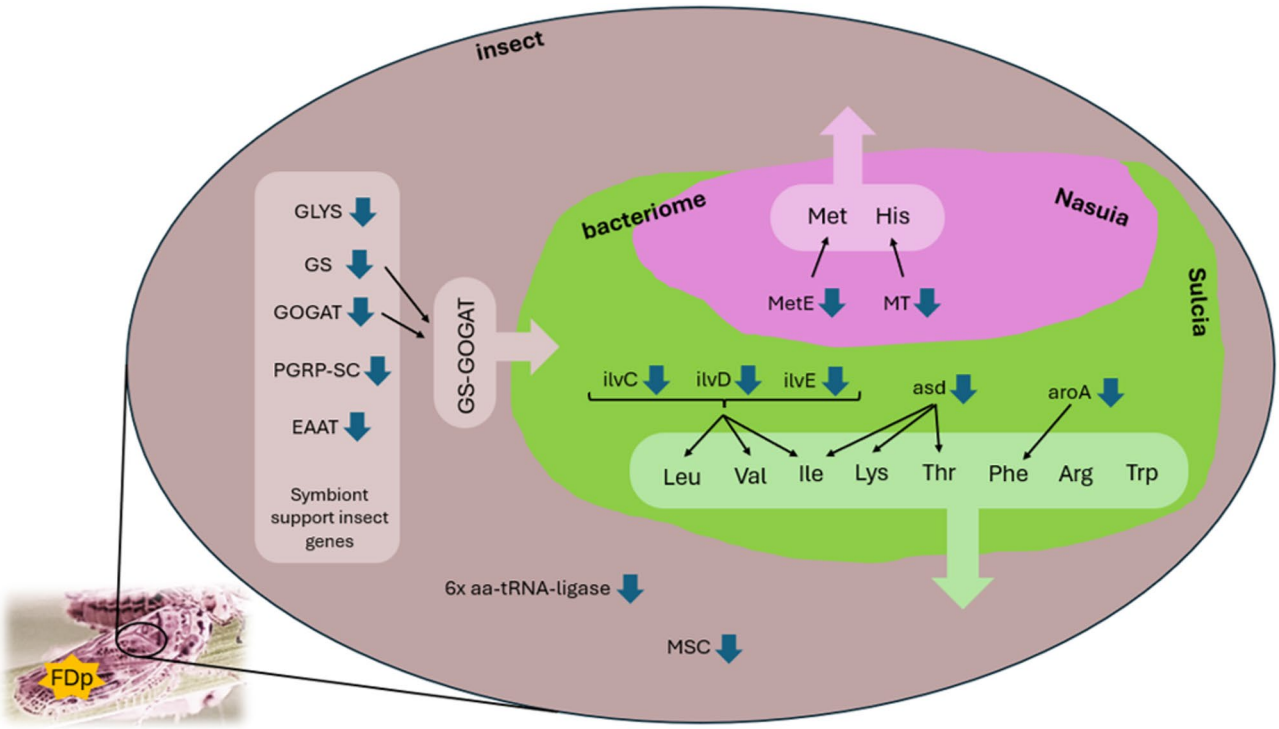


Fig. 4. Effect of *Euscelidius variegatus* exposure to FDp on protein involved in metabolic cooperation between the insect host and its primary endosymbionts. Brown area represents insect tissues, pink/green area represents a bacteriome with bacteriocytes hosting *Nasuia* (pink sector) and *Sulcia* (green sector) (Adapted from Fig. 2 in Bennet and Moran, 2013). Blue arrows indicate downregulation; black arrows indicate involvement in biosynthesis. Insect down-regulated proteins belonging to the "symbiont support genes set" as defined by Vasquez et al. 2024[32]: glycine-tRNA ligase (GLYS), glutamine synthetase 2 cytoplasmic (GS), glutamate synthase large chain (GOGAT), peptidoglycan recognition protein 3-like (PGRP-SC) and excitatory amino acid transporter-like (EAAT); aa-tRNA-ligase: insect down-regulated tRNA ligases; MSC: insect down-regulated multi-tRNA synthetase complex. *Nasuia* down-regulated proteins involved in aminoacid biosynthesis: 5-methyltetrahydropteroyltriglutamate-homocysteine (MetE), S-methyltransferase HisA/HisF-related TIM barrel protein (MT). *Sulcia* down-regulated proteins involved in aminoacid biosynthesis: ketol-acid reductoisomerase (ilvC), dihydroxy-acid dehydratase (ilvD), branched-chain amino acid aminotransferase 2 (ilvE), aspartate-semialdehyde dehydrogenase (asd), 3-deoxy-7-phosphoheptulonate synthase/chorismate mutase (aroA). GS-GOGAT: glutamine synthetase/glutamate synthase cycle.

Interestingly, five *E. variegatus* homologues of the 20 genes identified by Vasquez et al.[32] in *M. quadrilineatus* as putative symbiont-support genes were found to be downregulated in response to FDp (see column "Notes" in Supplementary Information 2). *Macrosteles quadrilineatus* harbors the same two obligate endosymbionts as *E. variegatus* and these 20 insect genes, predicted to be under positive selection, were identified among those likely to compensate for the incomplete metabolic functions of the highly degraded symbiont genomes. The five *E. variegatus* downregulated proteins are glycine-tRNA ligase, glutamine synthetase 2 cytoplasmic, glutamate synthase large chain (i.e. Uncharacterized protein LOC128983935), peptidoglycan recognition protein 3-like and excitatory amino acid transporter-like. In particular, it has been found in other hemipterans that both symbionts have lost the ability to produce and transport glutamine and glutamate, which are required to synthesize essential amino acids for their insect hosts[53]. Therefore, the activity of the insect GS/GOGAT (glutamine synthetase/glutamate synthase) pathway is essential for the survival of the entire holobiont. In addition to glycine-tRNA ligase, six other tRNA ligases and a component of the multi-tRNA synthetase complex were downregulated upon FDp challenge, indicating once more a decrease in the overall rate of protein production (Fig. 4).

From the perspective of the two primary symbionts, all differentially expressed proteins encoded by *Sulcia* and *Nasuia* were found to be downregulated in response to FDp infection. Such reduction of protein expression could result either from true protein downregulation or from a decrease in the number of bacterial cells present within the insect host. A global decline in bacterial cell number would be expected to reduce all identified bacterial proteins; however, only nine out of the 18 *Sulcia* proteins identified by DIA showed decreased levels. This pattern, therefore, suggests that specific protein downregulation, rather than a general reduction in bacterial titer, is the more likely explanation.

KEGG pathway analysis revealed that three out of the nine down-regulated *Sulcia* proteins (ketol-acid reductoisomerase, dihydroxy-acid dehydratase and branched-chain amino acid aminotransferase 2) are involved in the biosynthesis of the branched-chain amino acids valine, leucine, and isoleucine. Notably, isoleucine, leucine and valine are among the eight essential amino acids provided by *Sulcia* to the insect host, alongside arginine,

phenylalanine, tryptophan, lysine, and threonine. Two further *Sulcia* downregulated proteins are involved in the synthesis of amino acid precursors: the bifunctional enzyme 3-deoxy-7-phosphoheptulonate synthase/chorismate mutase (DAHPS/CM) type II, which contributes to phenylalanine biosynthesis, and the aspartate-semialdehyde dehydrogenase, which is involved in the synthesis of lysine, threonine, and isoleucine (Fig. 4).

Nasuia, which complements *Sulcia* by providing the insect with methionine and histidine, showed four downregulated proteins, which included precisely two enzymes involved in the biosynthesis of methionine and histidine: 5-methyltetrahydropteroyltriglutamate–homocysteine and S-methyltransferase HisA/HisF-related TIM barrel protein, respectively (Fig. 4).

Overall, these data suggest that FDp has a negative impact on the reciprocal metabolic support between the host insect and its primary endosymbionts, consequently resulting in an imbalance in amino acid synthesis and exchange. In this scenario, we observed an increase in the expression of some BEV membrane proteins involved in nutrient uptake, possibly reflecting an opportunistic adaptation of BEV to the phytoplasma-induced changes in host physiology. When it was first characterized in 1987, BEV showed, in fact, a certain degree of pathogenicity towards *E. variegatus*[54]

Conclusions

Recent advances in microbial ecology have highlighted the profound influence of insect-associated microbiomes, including bacteria, fungi, and viruses on host physiology, immunity, vector competence, and overall host fitness[6]. The identification of proteins from the primary symbionts, both localized in bacteriomes, as well as from BEV, predominantly found in gut and haemocoel, demonstrates that protein extraction from whole insect bodies allows also the detection of microbial proteins without the need for tissue dissection. However, targeted dissection of specific organs could increase the number and diversity of microbial proteins detected, thereby providing a more detailed understanding of the metabolic exchanges between the insect and its microbiome.

Proteins that are upregulated in response to FDp might represent promising targets for rational drug design (e.g., small molecules, peptides, or antibodies) or for gene silencing via dsRNA-mediated RNA interference. Their upregulation might be either a consequence of the phytoplasma manipulation of the insect metabolism or an active defense mechanism developed by the insect against the phytoplasma multiplication. For example, the silencing of natterin-1 in E. variegatus, which was identified as overexpressed in this study upon FDp exposure (Supplementary Information 2), was previously shown to impair phytoplasma multiplication[40].

While examples of targeted drug development exist for insect vectors of other pathogens, such as the use of mycophenolic acid to inhibit Dengue and Zika virus infections by targeting the mosquito inosine monophosphate dehydrogenase[55,56], no such strategies have yet been developed against vectors of phytoplasmas.

The remarkable potential of genome-wide proteomic analysis in insect pests does lie in the possibility to uncover new species-specific molecules or pathways that may be targeted to disrupt essential biological processes with high precision and minimal off-target effects, including those involved in symbiotic relationships with their microbiome.

Methods

Plants, insects, and phytoplasma strain

Oat (*Avena sativa*) and broad bean (*Vicia faba* var. "Agua-dulce Supersimonia") plants were grown from seed in pots in a greenhouse and maintained at 24 ± 2 °C. Two weeks after sowing, *A. sativa* plants were used to rear phytoplasma-free colonies of the leafhopper *Euscelidius variegatus* and *V. faba* plants were used as host plants to maintain isolates of the FDp. *Euscelidius variegatus* was originally collected in the Piedmont region (Italy) and has been continuously reared in plastic and nylon cages inside growth chambers set at 20–25 °C with a 16:8 light–dark photoperiod. The FDp strain FD-D[57], also originally isolated in Piedmont, is routinely maintained on *V. faba* through sequential transmission by *E. variegatus*[16]. For phytoplasma acquisition, fifth-instar nymphs from phytoplasma-free colonies were fed on FDp-infected broad beans for an acquisition access period (AAP) of 7 days and then transferred to oats for a 28-day latency period (LP). At 35 days post-acquisition, the surviving insects were analyzed for phytoplasma infection (as described in Rossi et al. 2023[58]) and used in experiments.

DNA extraction and sequencing

High molecular weight genomic DNA was extracted from a pool of 30 ethanol surface-sterilized *E. variegatus* individuals using the cetyltrimethylammonium bromide (CTAB) protocol[59]. A total of 6 µg of DNA was obtained, and its quality and quantity were assessed using a Qubit fluorometer (Thermo Fisher Scientific, USA) and agarose gel electrophoresis.

DNA was submitted to Novogene Co., Ltd. (Beijing, China) for library preparation and sequencing. A SMRTbell library was constructed and sequenced on the PacBio Sequel IIe platform (Pacific Biosciences, CA, USA), generating approximately 100 Gb of raw long-read data. The same DNA sample was also used to prepare a short-read paired-end library (insert size ~ 350 bp), which was sequenced on the Illumina NovaSeq 6000 platform, producing 140 Gb of raw data with 2 × 250 bp reads.

Genome assembly

To limit memory usage, random subsets corresponding to approximately one-third of the total PacBio long-read data (~ 31 Gb) were generated using seqtk v. 1.4-r122 (https://github.com/lh3/seqtk). The draft assemblies were produced using Flye v. 2.9.1-b1780[60] and polished using Illumina short-read data. Illumina reads were first subjected to quality control using the BBTools suite (BBMap—Bushnell B.—sourceforge.net/projects/bbmap/) to remove adapter sequences, low-quality bases, and potential contaminants. Subsequently, reads were aligned

to the Flye assemblies using BWA-MEM v. 0.7.15-r1140 (Liet al. 2009), and polishing was performed with Pilon v. 1.24[61] through three iterative rounds to correct residual base errors and indels.

As the sequenced sample was a mixture of the host and bacterial symbionts, the initial assembled contigs were identified using BLASTN searches against the NCBI nt database, separated into bacterial and insect sequences, and, within each group, assemblies were merged using Quickmerge (https://github.com/mahulchak/quickmerge) to improve contiguity. This separation was guided by the existing knowledge that *E. variegatus* harbors two primary bacterial symbionts, *Sulcia* and *Nasuia*, as well as two facultative symbionts, BEV and *S. euscelidii,* whose genome has been recently sequenced (BioProject: PRJNA1026792). Assembly quality was assessed using QUAST v. 5.3.0[62]. *Euscelidius variegatus* genome completeness was assessed using BUSCO v5 with the *hemiptera_v10* ortholog set.

Gene prediction

Prediction of insect coding sequencing and proteins was performed using the MAKER annotation pipeline v. 2.31.11[63] for eukaryotic organisms. The draft assembled genome and transcript evidence from a previous *E. variegatus* RNA-seq project (Bioproject PRJNA393620) were used as input. Repeat masking was performed within the MAKER pipeline using RepeatMasker, which was configured to use the Dfam database (bundled with MAKER) as the primary source for transposable element profiles. MAKER was used in two iterative rounds. In the first round, insect transcripts were aligned to the genome to generate initial gene predictions. In the second round, these predictions were refined using AUGUSTUS v. 3.4.0 with *Acyrthosiphon pisum* gene models and a high-confidence subset of gene models derived from the first round. Final gene models were extracted using gffread and used to train a species-specific AUGUSTUS ab initio predictor. This trained model was then employed in a final MAKER run to generate the definitive set of coding sequences and predicted proteins.

Scaffolds identified as bacterial were annotated for protein-coding genes using PROKKA v. 1.14.6[64].

Functional annotation of the assembled genomes

Functional annotation of all predicted protein-coding sequences was performed using blastp analysis against the NCBI nr database (threshold: Evalue < 0.0001) and eggNOG-mapper v2 (http://eggnog-mapper.embl.de/), based on orthology assignments from the eggNOG 5.0 database. In addition to orthologous group assignments, eggNOG-mapper provides functional annotations from multiple sources, including Gene Ontology (GO) terms and the BRITE hierarchy, which were used for downstream analyses. GO annotations characterized predicted gene products based on their putative biological roles, molecular functions, and subcellular localizations. BRITE codes, derived from the KEGG database, were used to assign high-level functional categories to the identified proteins. Functional pathway assignments were specifically derived from the KEGG reference pathways of the hemipteran insect *Halyomorpha halys* (hhal).

Proteins without a significant hit in the NCBI nr database were further analyzed with SignalP 5.0[65] to provide additional functional insights regarding their potential secretion (probability threshold > 0.5).

Protein extraction and quantification

Total proteins were extracted from insects either exposed (FD) or non-exposed (H) to FDp. Four biological replicates were analyzed for each condition, resulting in a total of 200 insects per treatment. For each replicate, 50 insects were homogenized in a mortar using liquid nitrogen. Once the insects were ground into a fine powder, 1 mL of RX buffer (0.1% Triton X-100, 100 mM KCl, 3 mM NaCl, 3.5 mM MgCl2, 1.25 mM EGTA, and 10 mM Hepes, pH 7.3)[47] was added. The samples were then sonicated for 1 min at room temperature and centrifuged at 13,000 rpm for 5 min. The supernatant was collected and precipitated with methanol/chloroform/water, as described by Wessel and Flügge (1984). The samples were quantified by means of the 2-D Quant Kit (Cytiva) with an adapted procedure for the 96-well microplate reader Tecan (Infinite M Nano). An aliquot of 100 µg of proteins was resuspended in 100 µL of 25 mM ammonium bicarbonate. The samples were reduced with dithiothreitol (DTT) to the final concentration of 20 mM, kept for 1 h at 37 °C, and alkylated with iodoacetamide (IAA) at the final concentration of 55 mM, for 1 h in the dark at room temperature (RT). The residual IAA was quenched in a second incubation step with DTT at the final concentration of 30 mM for 15 min at RT. The protein mixtures were digested O/N at 37 °C under shaking with modified porcine proteomic grade trypsin (1 µg/µL) at a 1:20 trypsin-protein ratio. An aliquot of 5 µL of pure formic acid (FA) was added to stop the digestion. The tryptic peptide mixtures were desalted using Strata C18-E SPE tubes. Briefly, the tubes were conditioned with 3 mL of methanol, then equilibrated with 3 mL of MilliQ water. The samples, previously diluted with 2 mL of 0.1% FA, were loaded in the Strata-X tubes. The tubes were then washed with 1 mL of 5% methanol, and the peptides were eluted with 3 mL of methanol. After desalting, the samples were dried in a 5301 Eppendorf Concentrator (Eppendorf, Hamburg, Germany) and suspended in 50 µL of 0.1% (v/v) FA/2% acetonitrile (ACN) before the analysis.

UHPLC-HRMS data acquisition: DDA and DIA methods

The high-resolution mass spectrometry (HRMS) analyses were performed using an Orbitrap Q Exactive Plus, coupled to an ultra-high-performance liquid chromatography (UHPLC) binary pump system (Vanquish Thermo Fisher Scientific, Waltham, Massachusetts, USA) with a HESI probe.

The stationary phase was a Luna C18 HPLC Column (150 × 1 mm, 3 µm; Phenomenex, Castelmaggiore, Italy). The oven temperature was set at 55 °C, and the autosampler was set at 4 °C. The mobile phases were 0.1% (v/v) FA in MilliQ water (A) and 0.1% (v/v) FA in 30% ACN 70% MetOH (B), and they were eluted at a flowrate of 70.0 µL/min, with an increasing concentration of solvent B, from 3 to 50%, over 80 min and with 70% in 5 min, and 100% in 1 min, for 5 min, and after reconditioned at 3% of B for 15 min. The injection volume was 4.0 µL.

The DDA method was in positive polarity mode, spray voltage of 3500 V, Sheat Gas 35, Aux Gas 10, probe heater 330 °C and S-Lens RF Level 60. Mass spectra were acquired in Full MS-ddMS2 mode. The instrument was set up so that Full MS spectra were acquired in an m/z scan range of 250–1500, the resolution was set at 70,000, the maximum IT was 200 ms, the AGC target was 3×10^6. Up to 6 of the most intense ions in MS1 were selected for fragmentation in the MS/MS mode. The fragmentation spectra resolution was set at 17,500 for the MS/MS spectra, with a dynamic exclusion of 10 s and an isolation window of 2.0 m/z. The normalized collision energy was set at 28, the maximum IT at 110 ms and the AGC target at 1×10^5.

The DIA method was composed of 20 consecutive MS2 windows acquired at 17,500 resolution, with an AGC target of 1×10^6 and a maximum injection time of 60 ms. In details the DIA method enclosed 4 windows with an isolation window of 31 m/z, 13 windows with an isolation of 21 m/z (AGC target of 2×10^5), and 3 windows with an isolation window of 51 m/z; the overlap for each window was equal to 0.5 m/z. The resulting m/z range was from 370 to 900.

Analysis parameters for DDA and DIA proteomic data

The Data-Dependent Acquisition (DDA) raw files were analyzed using MaxQuant (version 2.4.2.0), with peptide identification performed against a custom-made predicted proteome of the holobiont *E. variegatus*. The search was performed with trypsin specificity, allowing up to two missed cleavages. Carbamidomethylation of cysteine was set as a fixed modification, while oxidation of methionine and N-terminal acetylation was set as variable modifications. A false discovery rate (FDR) of 1% was applied at both the peptide-spectrum match (PSM) and protein levels. Label-free quantification (LFQ) was enabled, with a minimum ratio count of 2, and the 'match between runs' feature was activated.

The Data-Independent Acquisition (DIA) raw files were analyzed using DIA-NN (version 1.9.1) in library-free mode. In silico spectral libraries were generated from the same reference proteomes used in the MaxQuant analysis. An FDR of 1% was applied at both the precursor and the protein group levels. Protein quantification was performed with the following parameters enabled: trypsin was specified as the digestion enzyme, allowing for up to two missed cleavage site; charge range of 1–4 for peptides with 7–30 amino acids was set; peptidoform-level quantification, match-between-runs (MBR), heuristic protein inference, no shared spectra, single-pass mode, Robust LC (high precision), and retention time (RT)-dependent quantification mode were selected. The digestion settings and the allowed protein modifications in DIA-NN were configured to match those used in the MaxQuant analysis.

Statistical analysis of differentially expressed proteins

Proteomic quantification results from MaxQuant and DIA-NN were analyzed using Perseus (version 1.6.15.0). Label-free quantification (LFQ) intensities from MaxQuant and protein intensity values from DIA-NN were $\log_2$-transformed and normalized by median subtraction across columns. For differential expression analysis, only proteins quantified in at least three biological replicates per condition were retained. Differential expression analysis was performed using a two-sample *t*-test with a permutation-based false discovery rate (FDR) threshold of 5%. Proteins that were not detected in any replicate of one condition but were consistently detected in at least three replicates of the other condition were not included in the statistical analysis due to the absence of variance estimates. However, they were retained and reported as condition-specific proteins and considered overexpressed in the condition where they were detected.

Gene Ontology enrichment analysis was conducted using the GOEnrichment tool (https://github.com/DanFaria/GOEnrichment) with the GO file go-basic.obo. For each proteomic approach (DIA and DDA), the combined set of all proteins identified in non-exposed samples was used as the annotation (population) set, and the corresponding lists of differentially expressed proteins after the FDp challenge served as the study set. Multiple testing correction was applied using the Benjamini–Hochberg method, with a significance threshold set at $p < 0.01$. To reduce redundancy and improve visualization, the results were further refined using the rrvgo R package v.1.20.0. The terms with semantic similarity below 0.7 were grouped, and the org.Dm.eg.db annotation package (*Drosophila melanogaster* library) was used for functional annotation and graphical representation. REVIGO summarizes and visualizes GO terms by identifying a representative subset based on semantic similarity, when at least three terms are available within the same GO category.

Data availability

All the assembled genomes are available at the NCBI GenBank database under the BioProject accession number PRJNA1057536. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD066715 and [https://doi.org/10.6019/PXD066715] (https:/doi.org/https://doi.org/10.6019/PXD066715) Token: WByXJicLDy5A. This dataset comprises the reference proteomic database that was used for all the proteomic analyses and the raw output files generated by MaxQuant and DIA-NN. To maintain consistency between the protein database and the deposited output files of DDA and DIA analyses, in those files *Sorlinia* proteins keep the original prefix 'asaia', because *Asaia* sp. was the closest phylogenetic match to the bacterium identified in the *E. variegatus* microbiome before the releasing of *S. euscelidii* genome. The other prefixes are the same as the ones mentioned before.

Received: 3 September 2025; Accepted: 27 November 2025
Published online: 04 December 2025

References

1. Lee, I.-M., Davis, R. E. & Gundersen-Rindal, D. E. Phytoplasma: Phytopathogenic mollicutes. *Annu. Rev. Microbiol.* **54**, 221–255 (2000).
2. Nougadère, A. et al. Grapevine flavescence dorée phytoplasma Pest Report to support the ranking of EU candidate priority pests. *EFSA Support. Public.* **22**, 9567E (2025).
3. Malembic-Maher, S. et al. When a Palearctic bacterium meets a Nearctic insect vector: Genetic and ecological insights into the emergence of the grapevine Flavescence dorée epidemics in Europe. *PLoS Pathog.* **16**, e1007967 (2020).
4. Caudwell, A., Kuszala, C., Larrue, J. & Bachelier, J. Transmission de la Flavescence dorée de la fève à la fève par des cicadelles des genres Euscelis et Euscelidius. *Ann. Phytopathol.* **1572**, 181–189 (1972).
5. Gurung, K., Wertheim, B. & Falcao Salles, J. The microbiome of pest insects: it is not just bacteria. *Entomol. Exp. Appl.* **167**, 156–170 (2019).
6. Gupta, A. & Nair, S. Dynamics of insect-microbiome interaction influence host and microbial symbiont. *Front. Microbiol.* **11**, 1357 (2020).
7. Cooper, W. R. et al. Bacterial endosymbionts identified from leafhopper (Hemiptera: Cicadellidae) vectors of phytoplasmas. *Environ. Entomol.* **52**, 243–253 (2023).
8. Wilkinson, T. L. & Ishikawa, H. On the functional significance of symbiotic microorganisms in the Homoptera: A comparative study of *Acyrthosiphon pisum* and *Nilaparvata lugens*. *Physiol. Entomol.* **26**, 86–93 (2001).
9. Baumann, P. Biology of bacteriocyte-associated endosymbionts of plant sap-sucking insects. *Annu. Rev. Microbiol.* **59**, 155–189 (2005).
10. Kikuchi, Y. Endosymbiotic bacteria in insects: Their diversity and culturability. *Microbes Environ.* **24**, 195–204 (2009).
11. Douglas, A. E. The microbial dimension in insect nutritional ecology. *Funct. Ecol.* **23**, 38–47 (2009).
12. Martinson, V. G. et al. Multiple origins of obligate nematode and insect symbionts by a clade of bacteria closely related to plant pathogens. *Proc. Natl. Acad. Sci.* **117**, 31979–31986 (2020).
13. Marasco, R. et al. *Sorlinia euscelidii* gen. nov., sp. nov., a novel acetic acid bacterium isolated from the leafhopper *Euscelidius variegatus* (Hemiptera: Cicadellidae). *Int J Syst Evol Microbiol* **74**, 006544 (2024).
14. Abbà, S. et al. Genome sequence, prevalence and quantification of the first iflavirus identified in a phytoplasma insect vector. *Arch. Virol.* **162**, 799 (2017).
15. Ottati, S. et al. Biological characterization of *Euscelidius variegatus* iflavirus 1. *J. Invertebr. Pathol.* **173**, 107370 (2020).
16. Galetto, L. et al. Two phytoplasmas elicit different responses in the insect vector *Euscelidius variegatus* Kirschbaum. *Infect. Immun.* **86**, 10–128 (2018).
17. Weil, T. et al. Linking omics and ecology to dissect interactions between the apple proliferation phytoplasma and its psyllid vector *Cacopsylla melanoneura*. *Insect Biochem. Mol. Biol.* **127**, 103474 (2020).
18. Vijay, S., Rawat, M. & Sharma, A. Mass spectrometry based proteomic analysis of salivary glands of urban malaria vector *Anopheles stephensi*. *Biomed. Res. Int.* **2014**, 1–12 (2014).
19. Ramos, L. F. C. et al. Interspecies isobaric labeling-based quantitative proteomics reveals protein changes in the ovary of *Aedes aegypti* coinfected With ZIKV and Wolbachia. *Front. Cell Infect. Microbiol.* **12**, 900608 (2022).
20. Liu, B., Qin, F., Liu, W. & Wang, X. Differential proteomics profiling of the ova between healthy and Rice stripe virus-infected female insects of *Laodelphax striatellus*. *Sci. Rep.* **6**, 27216 (2016).
21. Zhao, J. et al. Comparative proteomic analysis provides new insight into differential transmission of two begomoviruses by a whitefly. *Virol. J.* **16**, 32 (2019).
22. Badillo-Vargas, I. E. et al. Proteomic analysis of *Frankliniella occidentalis* and differentially expressed proteins in response to *Tomato Spotted Wilt Virus* Infection. *J. Virol.* **86**, 8793–8809 (2012).
23. Tamborindeguy, C. et al. Genomic and proteomic analysis of *Schizaphis graminum* reveals cyclophilin proteins are involved in the transmission of cereal yellow dwarf virus. *PLoS ONE* **8**, e71620 (2013).
24. Kruse, A. et al. Combining 'omics and microscopy to visualize interactions between the Asian citrus psyllid vector and the Huanglongbing pathogen Candidatus *Liberibacter asiaticus* in the insect gut. *PLoS ONE* **12**, e0179531 (2017).
25. Zhao, J. et al. Data-independent acquisition boosts quantitative metaproteomics for deep characterization of gut microbiota. *NPJ. Biofilms. Microbiomes* **9**, 4 (2023).
26. Watanabe, K. & Sato, M. Gut colonization by an ice nucleation active bacterium, Erwinia(Pantoea)ananasReduces the cold hardiness of mulberry pyralid larvae. *Cryobiology* **38**, 281–289 (1999).
27. Gitaitis, R. D., Walcott, R. R., Wells, M. L., Perez, J. C. D. & Sanders, F. H. Transmission of *Pantoea ananatis*, causal agent of center rot of onion, by tobacco thrips *Frankliniella fusca*. *Plant Dis.* **87**, 675–678 (2003).
28. Krawczyk, K., Foryś, J., Nakonieczny, M., Tarnawska, M. & Bereś, P. K. Transmission of *Pantoea ananatis*, the causal agent of leaf spot disease of maize (*Zea mays*), by western corn rootworm (*Diabrotica virgifera* virgifera LeConte). *Crop Prot.* **141**, 105431 (2021).
29. Bing, X.-L. et al. Characterization of *Pantoea ananatis* from rice planthoppers reveals a clade of rice-associated *P. ananatis* undergoing genome reduction. *Microb. Genom.* **8**, 000907 (2022).
30. Blakeley-Ruiz, J. A. & Kleiner, M. Considerations for constructing a protein sequence database for metaproteomics. *Comput. Struct. Biotechnol. J.* **20**, 937–952 (2022).
31. Vallino, M. et al. Bacteriophage-host association in the phytoplasma insect vector *Euscelidius variegatus*. *Pathogens* **10**, 612 (2021).
32. Vasquez, Y. M., Li, Z., Xue, A. Z. & Bennett, G. M. Chromosome-level genome assembly of the aster leafhopper (*Macrosteles quadrilineatus*) reveals the role of environment and microbial symbiosis in shaping pest insect genome evolution. *Mol. Ecol. Resour.* **24**, 13919 (2024).
33. Kwak, Y., Argandona, J. A., Degnan, P. H. & Hansen, A. K. Chromosomal-level assembly of *Bactericera cockerelli* reveals rampant gene family expansions impacting genome structure, function and insect-microbe-plant-interactions. *Mol. Ecol. Resour.* **23**, 233–252 (2023).
34. Li, Z. et al. The genomic basis of evolutionary novelties in a leafhopper. *Mol. Biol. Evol.* **39**, msac184 (2022).
35. Heck, M. & Neely, B. A. Proteomics in non-model organisms: A new analytical frontier. *J. Proteome Res.* **19**, 3595–3606 (2020).
36. Armengaud, J. et al. Non-model organisms, a species endangered by proteogenomics. *J. Proteomics* **105**, 5–18 (2014).
37. Miura, N., Tabata, T., Ishihama, Y. & Okuda, S. Phylogenetic tree-based amino acid sequence generation for proteomics data analysis of unknown species. *Comput. Struct. Biotechnol. J.* **27**, 2313–2322 (2025).
38. Canuto, F. et al. A knockdown gene approach identifies an insect vector membrane protein with leucin-rich repeats as one of the receptors for the VmpA adhesin of flavescence dorée phytoplasma. *Front. Cell Infect. Microbiol.* **13**, 1289100 (2023).
39. Trivellone, V. et al. Evidence suggesting interactions between immunodominant membrane protein Imp of Flavescence dorée phytoplasma and protein extracts from distantly related insect species. *J. Appl. Microbiol.* **127**, 1801–1813 (2019).
40. Galetto, L. et al. Natterin-like and legumain insect gut proteins promote the multiplication of a vector-borne bacterial plant pathogen. *Microbiol. Res.* **291**, 127984 (2025).
41. Bressan, A., Girolami, V. & Boudon-Padieu, E. Reduced fitness of the leafhopper vector *Scaphoideus titanus* exposed to Flavescence dorée phytoplasma. *Entomol Exp Appl* **115**, 283–290 (2005).
42. Du, X. & Thiem, S. M. Responses of insect cells to baculovirus infection: Protein synthesis shutdown and apoptosis. *J. Virol.* **71**, 7866–7872 (1997).

43. Galetto, L. et al. Silencing of ATP synthase β reduces phytoplasma multiplication in a leafhopper vector. *J. Insect. Physiol.* **128**, 104176 (2021).
44. Galetto, L. et al. The major antigenic membrane protein of "*Candidatus phytoplasma* asteris" selectively interacts with ATP synthase and actin of leafhopper vectors. *PLoS ONE* **6**, e22571 (2011).
45. Rossi, M. et al. Application of laser microdissection to study phytoplasma site-specific gene expression in the model plant Arabidopsis thaliana. *Microbiol. Res.* **217**, 60–68 (2018).
46. Inaba, J., Kim, B. M., Zhao, Y., Jansen, A. M. & Wei, W. The endoplasmic reticulum is a key battleground between phytoplasma aggression and host plant defense. *Cells* **12**, 2110 (2023).
47. Suzuki, S. et al. Interaction between the membrane protein of a pathogen and insect microfilament complex determines insect-vector specificity. *Proc. Natl. Acad. Sci.* **103**, 4252–4257 (2006).
48. Martini, X., Hoffmann, M., Coy, M. R., Stelinski, L. L. & Pelz-Stelinski, K. S. Infection of an insect vector with a bacterial plant pathogen increases its propensity for dispersal. *PLoS ONE* **10**, e0129373 (2015).
49. Roosien, B. K. et al. Conditional vector preference aids the spread of plant pathogens: Results from a model. *Environ. Entomol.* **42**, 1299–1308 (2013).
50. Prasad, S. S. et al. Correction: Prasad et al. Patterns of variation in the usage of fatty acid chains among classes of ester and ether neutral lipids and phospholipids in the Queensland Fruit Fly. *Insects* **15**, 538 (2024).
51. Burmester, T. Evolution and function of the insect hexamerins. *Eur. J. Entomol.* **96**, 213–215 (1999).
52. Eliautout, R. et al. Immune response and survival of Circulifer haematoceps to Spiroplasma citri infection requires expression of the gene hexamerin. *Dev. Comp. Immunol.* **54**, 7–19 (2016).
53. Mao, M., Yang, X. & Bennett, G. M. Evolution of host support for two ancient bacterial symbionts with differentially degraded genomes in a leafhopper host. *Proc. Natl. Acad. Sci.* **115**, E11691 (2018).
54. Purcell, A. H. & Suslow, K. G. Pathogenicity and effects on transmission of a mycoplasmalike organism of a transovarially infective bacterium on the leafhopper *Euscelidius variegatus* (Homoptera: Cicadellidae). *J. Invertebr. Pathol.* **50**, 285–290 (1987).
55. Kang, S., Shields, A. R., Jupatanakul, N. & Dimopoulos, G. Suppressing Dengue-2 infection by chemical inhibition of *Aedes aegypti* host factors. *PLoS Negl. Trop. Dis.* **8**, e3084 (2014).
56. Dong, S., Kang, S. & Dimopoulos, G. Identification of anti-flaviviral drugs with mosquitocidal and anti-Zika virus activity in Aedes aegypti. *PLoS Negl. Trop. Dis.* **13**, e0007681 (2019).
57. Rossi, M. et al. Genetic diversity of flavescence dorée phytoplasmas at the vineyard scale. *Appl. Environ. Microbiol.* **85**, e03123 (2019).
58. Rossi, M. et al. Competition among flavescence dorée phytoplasma strains in the experimental insect vector *Euscelidius variegatus*. *Insects* **14**, 575 (2023).
59. Marzachì, C., Veratti, F. & Bosco, D. Direct PCR detection of phytoplasmas in experimentally infected insects. *Ann. Appl. Biol.* **133**, 45–54 (1998).
60. Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P. A. Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**, 540–546 (2019).
61. Walker, B. J. et al. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* **9**, e112963 (2014).
62. Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D. & Gurevich, A. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* **34**, i142–i150 (2018).
63. Cantarel, B. L. et al. MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **18**, 188–196 (2008).
64. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
65. Almagro Armenteros, J. J. et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423 (2019).

Acknowledgements

The authors would like to thank Luca Bordone for providing plants for insect rearing, Luca Bucci for maintaining the insect rearing and Stefano Ghignone for providing technical assistance with the bioinformatic server. This work has been performed with the contribution of the research infrastructure IMPreSA (Metrological Infrastructure for the food safety), at INRiM, supported by Regione Piemonte.

Author contributions

MR, SA and MV designed the experiments; BA, SC, FR, LG and MR performed the experiments; SA performed the bioinformatic analyses; SA, MR and MV drafted the manuscript; MR, SA, CM and DB funded the work. All authors revised the manuscript draft and contributed to and accepted its final version.

Funding

MR discloses support for the research of this work from Consiglio Nazionale delle Ricerche—Dipartimento di Scienze Bio-Agroalimentari "Premio per Progetti di Ricerca 2023, Project Promise". SA discloses support for the research of this work from Ministero dell'agricoltura, della sovranità alimentare e delle foreste, Azione di ricerca n. 1 (D.M. 419782, 14/08/2023), for the Project "MICOTI", CUP B17G23000320005. CM discloses support for the research of this work from Regione Piemonte, under the frame of Direzione Agricoltura e cibo Programma regionale (2025–2026), "Metodi Innovativi per la DIfesa e gestione della Flavescenza Dorata della vite 2 (MIDI-FD3)". DB discloses support for the research of this work from PRIN 2022 Italian Ministry of University and Research, 2022BPB5A8, "RNA interference for the control of hemipteran pests (HEMipteran INTerference, HEMINT)".

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-30920-7.

Correspondence and requests for materials should be addressed to S.A.

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

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2025