Wolbachia strain diversity in a complex group of sympatric cryptic parasitoid wasp species

Federica Valerio[1] , Cornelia Martel[2], Constanti Stefanescu[3] , Saskya van Nouhuys[4] , Maaria Kankare[5] and Anne Duplouy[1,2]*

Abstract

Background Maternally-inherited symbionts can induce pre-mating and/or post-mating reproductive isolation between sympatric host lineages, and speciation, by modifying host reproductive phenotypes. The large parasitoid wasp genus *Cotesia* (Braconidae) includes a diversity of cryptic species, each specialized in parasitizing one to few related Lepidoptera host species. Here, we characterized the infection status of an assemblage of 21 *Cotesia* species from 15 countries by several microbial symbionts, as a first step toward investigating whether symbionts may provide a barrier to gene flow between these parasitoid host lineages.

Results The symbiotic microbes *Arsenophonus*, *Cardinium*, Microsporidium and *Spiroplasma* were not detected in the *Cotesia* wasps. However, the endosymbiotic bacterium *Wolbachia* was present in at least eight *Cotesia* species, and hence we concentrated on it upon screening additional DNA extracts and SRAs from NCBI. Some of the closely related *Cotesia* species carry similar *Wolbachia* strains, but most *Wolbachia* strains showed patterns of horizontal transfer between phylogenetically distant host lineages.

Conclusions The lack of co-phylogenetic signal between *Wolbachia* and *Cotesia* suggests that the symbiont and hosts have not coevolved to an extent that would drive species divergence between the *Cotesia* host lineages. However, as the most common facultative symbiont of *Cotesia* species, *Wolbachia* may still function as a key-player in the biology of the parasitoid wasps. Its precise role in the evolution of this complex clade of cryptic species remains to be experimentally investigated.

Keywords Microbial symbiosis, Hymenoptera, Melitaeini butterflies, *Cotesia*, Cytoplasmic incompatibility

*Correspondence:
Anne Duplouy
anne.duplouy@helsinki.fi
[1]Organismal and Evolutionary Biology Research Programme, University of Helsinki, Helsinki, Finland
[2]Department of Biology, Lund University, Lund, Sweden
[3]Natural Sciences Museum of Granollers, Granollers, Spain
[4]Centre for Ecological Sciences, Indian Institute of Science, Bangalore, India
[5]Department of Biological and Environmental Science, University of Jyväskylä, Jyväskylä, Finland

Background

At least 40% of all insect species worldwide are associated with endosymbiotic microbes, including *Arsenophonus*, *Cardinium*, Microsporidium, *Rickettsia*, *Spiroplasma*, and possibly the most common one: *Wolbachia* [1]. To enhance their own fitness through transmission in their host population, these microbes can manipulate their host reproduction and other life-history traits [2–4]. For example, many symbiont species can induce cytoplasmic incompatibility (CI), in which infected males are incompatible with females that are uninfected or infected with another incompatible symbiotic strain [5, 6]. Some

endosymbiotic microbes can also manipulate behaviour of their host, such that infected and uninfected individuals have different mate or host preferences [7, 8, reviewed in 9]. These symbiont-induced reproductive and behavioural alterations have thus long been proposed as key drivers of host speciation and diversity [10], via post-mating isolation [11, 12], and/or pre-mating reproductive isolation between lineages of different infection status [13, 14]. For example, Shoemaker et al. [15] showed that in the *Drosophila subquinaria* species group, *Wolbachia* induces unidirectional CI, which, coupled with mate choice preferences, established a reproductive barrier between *D. recens* and *D. subquinaria*. While divergence between insect species often occurs independently of any symbiotic infection [16], the relative importance of microbial symbionts in this process is likely underestimated as the prevalence, diversity, and role of symbionts remain unknown for many insect systems.

Biogeographic studies of symbiotic diversity and prevalence, combined with phylogenetic analyses, can provide clues to the ecological and evolutionary roles of symbionts in their host species clade. For example, obligate symbionts transmitted exclusively maternally are likely to show high prevalence within their host species [17, 18], and their interactions can exhibit phylogenetic concordance between the symbiont and their host over long evolutionary periods. This pattern was, for example, observed between beneficial *Wolbachia* strains and their bedbug or nematode hosts [19, 20]. Similarly, post-mating and pre-mating reproductive isolation induced by facultative symbionts can also lead to co-divergence of the hosts and the symbionts [21], at least over short evolutionary periods required for cryptic species to diverge. However, although facultative endosymbionts are predominantly transmitted vertically from mothers to offspring, these bacteria can also be horizontally transferred between host lineages and species [22–27]. Horizontal transfer events might occur between interacting species, including between parasitoids and their prey, between prey attacked by the same parasitoid species, or between predators or parasitoids sharing the same prey [28, 29]. Transfer may also occur between herbivores sharing the same host plants [23, 24], or between hybridizing species [30]. These events allow the symbiotic strains to colonize divergent host species, which could obscure the evolution of patterns of phylogenetic concordance between host and symbiont.

Parasitoid wasps in the genus *Cotesia* (Hymenoptera: Braconidae) parasitize Lepidoptera by laying a single or multiple eggs in their host caterpillars. The parasitoid wasp larvae grow while feeding on the developing caterpillar's haemolymph, and then pupate in conspicuous silken cocoons outside the body of the host [31]. The whole genus *Cotesia* accounts over 1000 named species

worldwide, which parasitize many Lepidoptera species [32, 33]. In some cases, the *Cotesia* wasps can have dramatic effects on their host population dynamics [34]. For example, even by only infecting on average 10% of the caterpillars of *Melitaea cinxia* (Lepidoptera: Nymphalidae: Melitaeini) in the Åland Islands, Finland, *Cotesia melitaearum* has been found to cause localized decline within the larger host metapopulation [35, 36]. Furthermore, multiple *Cotesia* species can co-occur, where their host species occur together in a landscape. In Northeastern Spain, seven cryptic *Cotesia* species each use only one to two of the local eight related *Melitaea* and two *Euphydryas* (Melitaeini) butterfly species, which share some host plant species, and live in shared meadow habitats [37, 38].

To date, *Wolbachia* is, to our knowledge, the only endosymbiont that has been previously screened for, and detected from *Cotesia* species. The bacterium has been found in *C. glomerata* (Linnaeus) and *C. vestalis* (Haliday) (synonym of *C. plutellae* (Kurdjumov)) [39], and in *C. sesamiae* from Cameron and Kenya [40, 41]. Branca et al. [41] demonstrated that *Wolbachia* induces unidirectional CI in *C. sesamiae* from Sub-Saharan Africa, which is associated with host specialization, genetic structure, and biogeography. In the *C. melitaearum* clade, molecular characterizations based on small number of genes have shown that specialization and competitive interactions in local *Cotesia* are associated with the emergence of several cryptic sympatric *Cotesia* species [37, 42]. In this parasitoid wasp clade, however, the role of symbiont-induced pre-mating and/or post-mating isolation between host lineages remains to be investigated. In this study, we aimed at identifying whether common insect endosymbiotic microorganisms, including *Arsenophonus*, *Cardinium*, Microsporidium, *Spiroplasma*, and *Wolbachia*, were present in 15 *Cotesia* species and cryptic species parasitizing Melitaeini butterfly species across different geographic locations. After identifying *Wolbachia* as the only detectable symbiont in these *Cotesia* species, we characterized the diversity and phylogeny of the *Wolbachia* strains, to investigate any potential inter-species transfer of the symbionts. For this, we additionally included strains identified from six *Cotesia* species from which genomic data was publicly available on NCBI, and strains previously detected in diverse Lepidoptera species known to be attacked by *Cotesia* wasps (reviewed in [43]). Our study provides an overview of the prevalence of diverse endosymbiotic microbes in *Cotesia* wasps, as well as a more thorough description of the diversity and phylogeny of the *Wolbachia* strains detected from *Cotesia* wasps. These are the first steps towards evaluating the role such symbionts might play in the evolutionary ecology of parasitoid wasps.

Materials and methods

Material

We recovered 323 DNA-extracts from *Cotesia* specimens (Table S1) stored in the −20 °C freezers at the University of Helsinki. These samples were originally collected in the early 2000's, and used for earlier studies of the phylogeny and butterfly-host specialization of *Cotesia* species associated with checkerspot/fritillary butterflies (*Melitaea* and *Euphydryas*) [37, 42, 44]. They were previously characterized as representing at least 16 *Cotesia* species and cryptic species, and were predominantly collected from Finland (*N*=94 from 4 species) and from Spain (*N*=153 from 10 species). Moreover, samples from China (*N*=8 specimens from 3 species), Estonia (*N*=13 from 1 species), France (*N*=24 from 6 species), Hungary (*N*=3 from 1 species), Italy (*N*=3 from 1 species), Russia (*N*=11 from 1 species), Sweden (*N*=4 from 3 species), UK (*N*=3 from 2 species), and USA (*N*=7 from 2 species), were also included. No tissue nor DNA material were preserved from the original butterfly hosts of these wasps.

Molecular work on lab-stored DNA extracts

The DNA from all field collected wasps was extracted using NucleoSpin Tissue Kit (Macherey-Nagel) for the purpose of phylogenetic studies of the *Cotesia* wasp species in the early 2000s by Kankare and colleagues [37, 42, 44, 45]. The DNA extracts have since been preserved in the freezer (-20 °C) at the University of Helsinki, Finland. The quality of each DNA extract was tested by PCR amplification of the mitochondrial cytochrome C oxidase subunit I gene (*COI* - primer pair LCO/HCO) [46]. The DNA extracts that did not amplify with the primers LCO/HCO after two PCRs were removed from further analyses.

To identify which potential symbionts could be found in this parasitoid wasp system, we first screened 56 *Cotesia* specimens from six species, and from four countries (Estonia, Finland, Spain, Sweden), for infection with five microbial symbionts (Table S1) known for manipulating other insect species' reproductive systems. We screened for the bacteria *Spiroplasma* and *Cardinium* using the 16 S ribosomal RNA (*16 S r*RNA) gene [47–49], for the bacterium *Arsenophonus* by targeting the *23 S r*RNA gene [50], for *Wolbachia* using *Wolbachia*-specific primers amplifying the *wsp* (*Wolbachia surface protein*) gene and up to five conserved *Wolbachia* Multilocus Sequence Typing (MLST) markers: *coxA, fbpA, ftsZ, gatB* and *hcpA* [51], and for the fungal symbiont *Microsporidium* by amplifying the *18 S r*RNA gene [52]. We did not test for the presence of other microbial reproductive manipulators, such as *Rickettsia* [4]. We later screened all remaining *Cotesia* specimens for infection by the only detected symbiont: *Wolbachia.* One negative control (water

sample) and one positive control were included in each PCR. Positive controls derived from DNA extractions of either a *Wolbachia*-infected *Ischnura elegans* specimen [53], an *Arsenophonus*-infected two-spot ladybird, *Spiroplasma*-infected *Drosophila* flies, *Cardinium*-infected midges (graciously provided by Prof. Hurst from Liverpool University), or a Microsporidium-infected *Melitaea* butterfly (Duplouy's own lab). All primer sequences are given in Table S2. We Sanger sequenced the amplified genes on an ABI-3730 DNA Sequencer (Applied Biosystems) at the University of Helsinki, Finland, using only the forward primers for each gene. All *Wolbachia* MLST loci and *wsp* gene sequences were identified by comparing the resulting assemblies against the PubMLST database (https://pubmlst.org) with BLAST [54].

Extracting, cleaning, and processing additional sequence material from NCBI repository

To put our findings on the most common symbiotic infection in *Cotesia* wasps in a larger phylogenetic context, we expanded the host species range and *Wolbachia* strain diversities of our study by screening for *Wolbachia* genomic material in the genomic data from *Cotesia* sequencing projects publicly available from NCBI Sequence Read Archive (SRA) database (https://www.ncbi.nlm.nih.gov/sra), and from the NCBI nucleotide database (https://www.ncbi.nlm.nih.gov/nucleotide). To do so, we first searched the SRA database using the keyword "*Cotesia*", selecting "DNA" as source and "Genome" as strategy. With this approach, we identified 28 genome sequencing projects including both short-read and long-read data and representing six different *Cotesia* species (*C. congregata*, *C. flavipes*, *C. glomerata*, *C. rubecula*, *C. sesamiae*, and *C. vestalis* as synonym of *C. plutellae*) (Table S3).

We processed the short-read (Illumina) sequencing samples with Prinseq-lite (version 0.20.4) [55] to remove all sequences with at least one ambiguous nucleotide. The resulting reads were adapter trimmed and quality filtered using Trimmomatic (version 0.39) [56]. Quality assessment reports were obtained with FastQC [57] and summarized by MultiQC [58]. In contrast, the long-read sequences (Oxford Nanopore, ONT) were quality filtered using NanoFilt (version 2.7.1) [59], which excluded sequences with a mean base quality lower than ten and lengths lower than 1 kb. The quality of the processed specimens was evaluated with NanoStat (version 1.5.0) [59].

We then screened the Illumina samples for *Wolbachia* infection using Kraken2 (version 2.0.8) provided with a custom database of 142 *Wolbachia* publicly available reference genomes [60] (140 reference genomes from GenBank and two (*w*Di and *w*Ls) from http://nematodes.org/) (See Table S4). Samples with at least 1000 reads

classified as *Wolbachia* according to Kraken2 [61] were then mapped against our *Wolbachia* reference genomes database using Bowtie2 (version 2.4.4) [62]. In contrast, the ONT sequencing data were directly aligned to *Wolbachia* reference genomes by Minimap2 (version 2.21) [63]. We used SAMtools (version 1.13) [64] to extract, merge, and sort reads properly mapped as pairs (mapping quality of 20) from the SAM file generated in the alignment step. For each alignment, the per-base read depth across two *Wolbachia* reference genomes (*w*MelPop strain GenBank CP046921.1 and *w*PipPel strain GenBank AM999887.1) was calculated using the SAMtools depth function and plotted in R with ggplot2 [65] (Fig. S2-S3). Mapped reads belonging to samples from the same Bio-Sample were also processed as merged reads.

Finally, we built *Wolbachia* genome assemblies by individually assembling mapped reads from short- and long-read sequencing using the Unicycler pipeline (version 0.4.9) [66]. The quality and the completeness of the resulting genome assemblies were estimated by QUAST (version 5.0.2) [67] and BUSCO (version 5.4.3, Rickett-siales odb10 database) [68]. The assemblies, along with the two *Wolbachia* reference genomes mentioned above, were analysed using FastANI (version 1.3) [69]. FastANI estimates the Average Nucleotide Identity (ANI) metric, enabling the clustering of genomes from different individuals/organisms. This method facilitates the inference of the supergroup placement of *Wolbachia* strains by utilizing their entire genomes, and is a more comprehensive approach compared to using a limited set of markers. Annotation of *Wolbachia* assemblies and reference genomes was performed with Prokka (version 1.4.6) [70] using default settings. Subsequently, the protein sequences predicted by Prokka were uploaded into the OrthoVenn3 web server (https://orthovenn3.bioinfoto-olkits.net); accessed date: 15 July 2023) for identification and comparison of orthologous clusters (Fig. S4). All final assemblies are available from Zenodo at https://doi.org/10.5281/zenodo.8422079.

Identifying the CI-associated genes

To explore whether the *Wolbachia* strains analysed here may be causing CI in their hosts, we searched for the CI-associated genes, *cifA* and *cifB*, in the newly assembled *Wolbachia* genomes, using BLAST. For this purpose, we downloaded the amino acid sequences of CifA and CifB from various *Wolbachia* strains in the NCBI database (Organism: *Wolbachia*, Source: RefSeq only) using the keywords "cytoplasmic incompatibility CifA" or "cytoplasmic incompatibility CifB". Subsequently, these amino acid sequences served as queries in two distinct searches: TBLASTN against *Wolbachia* assemblies and BLASTP against proteomes derived from the same assemblies. Homologs covering at least 70% of the length

of the query, with an identity of at least 50%, and having an E-value cut-off of 10^{-10} were aligned to the intact Cif homologs identified by Martinez et al. [71] using MUS-CLE [72].

Phylogenetic analyses

We inferred the phylogenetic relationships between the different *Wolbachia* strains of *Cotesia* wasps using the characterized *Wolbachia* MLST (*coxA*, *fbpA*, *ftsZ*, *gatB* and *hcpA*) and *wsp* genes sequences, from both DNA extracts and genomic sequence material from NCBI. To increase the list of *Wolbachia* strains included in our phylogeny, we screened the NCBI nucleotide database for any of the five *Wolbachia* MLST markers [51] and *wsp* gene [73] from any *Cotesia*, and some of their known Lepidoptera host species. As of the 10th January 2023, there was no record of *Wolbachia* strain from *Cotesia* species in the PubMLST database, however, we still recovered an additional 40 MLST and *wsp* sequences from four *Cotesia* species (*C. flavipes*, *C. glomerata*, *C. sesamiae* and *C. vestalis*), and four Lepidoptera host species (*Melitaea didyma*, *Chilo partellus*, *Pieris rapae*, *Plutella xylostella*) from NCBI. The sample size and geographic sampling locations are provided in Table S1 and illustrated by Figure S1. Additionally, we also recovered the sequences of 12 reference *Wolbachia* strains belonging to A-, B-, D- or F- supergroups and originating from different host species (https://pubmlst.org, *w*Au, *w*Bm, *w*Bol1, *w*Clec-F, *w*Ha, *w*Irr, *w*MelPop, *w*No, *w*Pel, *w*Ri, *w*Stri, *w*Vit), one strain from the butterfly *Danaus chrysippus (Nymphalidae)*, and three from the parasitoid wasp *Hyposoter horticola (Hymenoptera: Ichneumonidae)* (Table S1).

Individual MLST and *wsp* genes, and their concatenated alignments were produced using MAFFT [74], and manually curated in AliView [75]. We performed the phylogenetic analyses using RAxML [76] in raxml-GUI 2.0 [77] applying a general time reversible model with gamma-distributed rate variation across sites and a proportion of invariable sites (GAMMAGTR+I) on individual genes and concatenated alignments (Fig. S5 and S6). In all cases, node support was calculated by the rapid bootstrap feature of RAxML (100 replicates). The *Wolbachia* reference strains *w*Bm [78] and *w*Cle [79], which belong to the D- and F-supergroup, respectively, were used as outgroups to root the *Wolbachia* trees.

To infer whether *Wolbachia* strain diversification is concordant with the *Cotesia* wasp species diversification, we inferred the phylogenetic relationships between *Cotesia* species using the *COI* sequences of *Cotesia* species, employing a maximum likelihood approach. We sampled all *COI* sequences deposited in GenBank for 39 different *Cotesia* species, including sequences from our *Cotesia* specimens previously deposited by [42, 44] (Table S5).

As outgroups, we selected three species belonging to the *Microgaster* genus (Hymenoptera, Braconidae), namely *Microgaster nobilis*, *M. deductor* and *M. subcompletus* (see Table S5). We generated a *COI* sequence alignment of 606 bp with MAFFT, that we manually curated for misaligned regions using AliView. We constructed a maximum likelihood phylogeny from this alignment using IQTREE [80] under the best-fit model automatically selected by ModelFinder [81] (Fig. 1 and Fig. S7). Node support was estimated using ultrafast bootstrapping with 1,000 replicates [82].

The CifA and CifB proteins have previously been classified into at least five distinct phylogenetic clades (types I–V) with different degrees of compatibility [71, 83–85]. To determine the group to which the annotated Cif homologs from *Wolbachia* found in *Cotesia* hosts belong, we performed a phylogenetic analysis. The best-fit substitution model for the protein multiple sequence alignment was estimated using Modeltest-NG [86] in raxmlGUI 2.0 and based on the Akaike information criterion (AIC), it was determined to be a JTT+G4+F model. A maximum likelihood phylogenetic tree was built using RAxML in raxmlGUI 2.0 software with 100 rapid bootstraps (Fig. S8). Tree visualization and figures were obtained with ITOL [87] using the bipartitions output trees produced by RAxML and the bootstrap consensus tree from IQTREE analysis.

Results

Detection of endosymbionts in *Cotesia* DNA extracts
Out of the 323 DNA extracts selected for *Wolbachia* screening, 282 were of good quality based on *COI* amplification, suggesting most of the specimens had been sufficiently preserved since extraction [42, 44].

The PCR amplifications for *Arsenophonus*, *Spiroplasma* or *Microsporidium* from 56 *Cotesia* specimens from four countries were negative (Table S1). There was one amplification using the *Cardinium 16S*rRNA primers in one unique specimen of *C. melitaearum* cryptic sp. H from Spain. However, our several attempts at sequencing this amplicon were not successful and hence we could also not confirm the presence of *Cardinium* in this *Cotesia* sample. In contrast, out of the 282 *Cotesia* samples of good quality, 50 (17.7%) carried the symbiotic bacterium *Wolbachia* (Tables 1, S1), representing at least eight *Cotesia* species parasitizing Melitaeini butterfly species (Fig. 1).

Detection of endosymbionts in genome projects available in NCBI
By screening the 28 *Cotesia* genome projects (i.e. SRA projects) available on NCBI, we also identified 14 specimens (50%) containing at least 1000 reads classified as *Wolbachia* (Table S6). Ten specimens (six specimens from *C. glomerata*, one from *C. sesamiae* and three from *C. vestalis*) included *Wolbachia* reads distributed throughout the *Wolbachia* reference genomes (Figs. S2-S3), while the last four specimens only included reads with patchy coverage across the *Wolbachia* reference genomes. These last four projects were considered as potential false positive results for *Wolbachia* infection, with the *Wolbachia* reads representing potential contamination, or insertions of *Wolbachia* sequences in the *Cotesia* host genomes.

Wolbachia strain diversity
Using the ten *Cotesia* genome projects found infected with *Wolbachia*, we partially assembled nine *Wolbachia* genomes. Three assemblies isolated from *C. glomerata*, exhibited BUSCO completeness of 86.8% (SRR13990441), 87.7% (SRR13990442), and 41.8% (SAMEA7283786) with corresponding total sizes of 1.10 Mbp, 1.08 Mbp, and 0.52 Mbp, respectively (See Tables S7-S8), while all other *Wolbachia* assemblies had a low number of BUSCO genes and were <0.1 Mbp in size (Tables S7-S8). We were only able to extract between three and six MLST and *wsp* markers from the three largest *Wolbachia* assemblies.

Combining results obtained by direct amplification of the *Wolbachia* markers by PCRs and by screening the *Wolbachia* genomic assemblies built from *Cotesia* genomic sequences available on NCBI for those same markers, we obtained sequences from one to six markers for 38 (out of 61) *Wolbachia*-infected specimens (Table S1). We identified a total of 14 alleles for the *ftsZ* gene, nine for the *hcpA* gene, five for the *coxA* gene, six for *fbpA*, and six for *gatB* (See Table S1 for further details). This resulted in a concatenated alignment of 2559 bp, which allowed us to discriminate ten *Wolbachia* strains from ten *Cotesia* species (Table 1). We did not detect multiple infections in any of the individual *Cotesia* specimens, but two species carried several *Wolbachia* strains. Specimens of *C. koebelei* reared from *E. editha* from western North America carried either a supergroup A or a B *Wolbachia* strain, and Spanish specimens of *C. bignellii* carried a A-supergroup strain, while French specimens of the same species carried one of two B-supergroup strains (Fig. 1).

Analyses of the *Wolbachia* genomic assemblies
By comparing the predicted proteomes of our two largest *Wolbachia* assemblies with an >50% BUSCO completeness against those of the two *Wolbachia* reference genomes (*w*MelPop and *w*PipPel) using Prokka, we identified 954 protein-coding genes, 30 tRNAs, and one rRNA in the SRR13990441 assembly, and 996 protein-coding genes, 32 tRNAs, and three rRNAs in the SRR13990442 assembly (Table S9). In contrast, the two reference genomes, *w*MelPop and *w*PipPel, contained
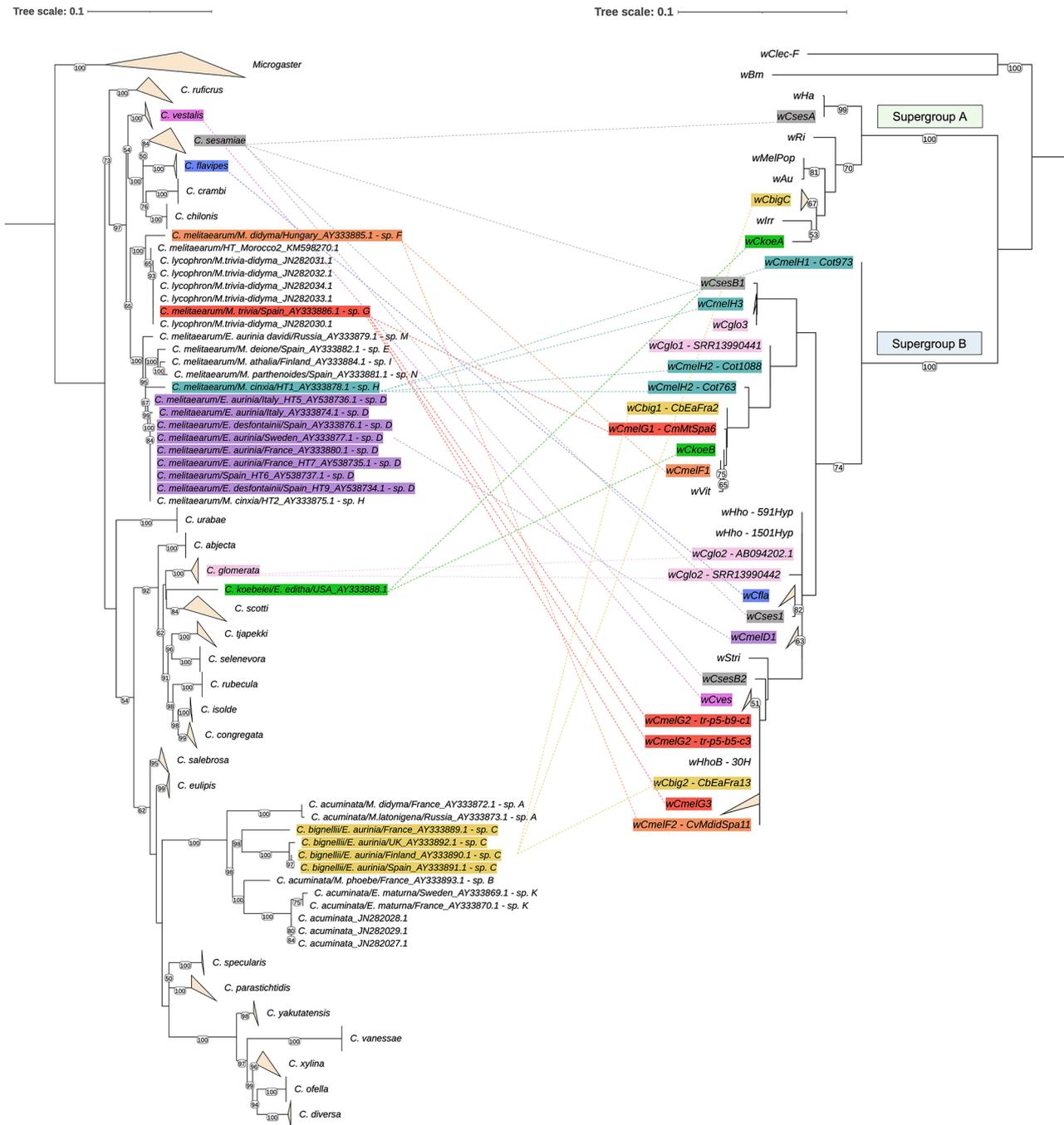
Valerio *et al. BMC Microbiology*   (2024) 24:319

Page 6 of 12

Fig. 1 Comparison between *Cotesia* parasitoid lineages and *Wolbachia* strains from *Cotesia* species. The *Cotesia* maximum likelihood phylogenetic tree was inferred from the nucleotide sequence alignment (606 bp) of the mitochondrial *COI* gene. The *Wolbachia* maximum likelihood tree was based on concatenated alignment (2,559 bp) of the MLST and *wsp* markers and rooted using reference genomes from *Wolbachia* strains *w*Bm and *w*Clec belonging to the supergroups D and F, respectively. *Cotesia* species labelled A through N correspond to cryptic species described in [44]. The coloured lines link *Cotesia* host species to their respective *Wolbachia* strain infections; with a unique colour for each host species. Branches corresponding to different sequences obtained from different specimens within the same species, and sequences from different species but within the same genus (only in the case of the outgroup *Microgaster*), were collapsed and represented as orange triangles for visual clarity. Bootstrap support values > 50 are displayed at each node

1304 and 1410 protein-coding genes, 34 tRNAs, and three rRNAs, respectively (Table S9). The comparison using the Orthovenn 3 web server showed a total of 1057 conserved orthologs in all four strains, with 590 of these being single copy. All four strains shared 639 ortholog clusters (Fig. S4). The SRR13990442 assembly contains 876 orthologs, while SRR13990441 has 875, and they both share 71 unique orthologs with the reference

Table 1 Metadata for the *Cotesia* species and cryptic species found to be infected with *Wolbachia*: their butterfly host species, country of origin, and *Wolbachia* prevalence. Rows in grey highlight the specimens that were screened for all five symbionts (table S1), while rows in white include the specimens screened for *Wolbachia* only

Species	Host species (reared from)	Country	Infection rate (infected/total)	Strains detected
C. acuminata cryptic sp. B	*Melitaea phoebe*	Spain	24.4% (5/17)	Uncharacterized
C. bignellii	*Euphydryas aurinia*	France	100.0% (2/2)	*w*Cbig
C. bignellii cryptic sp. C	*Euphydryas aurinia*	Spain	50.0% (3/6)	*w*CbigC
C. koebelei	*Euphydryas editha*	USA	100.0% (2/2)	*w*CkoeA, *w*CkoeB
C. melitaearum cryptic sp. D	*Euphydryas aurinia*	Spain	10.8% (4/37)	*w*CmelD
C. melitaearum cryptic sp. F	*Melitaea didyma*	Spain	100% (12/12)	*w*CmelF
C. melitaearum cryptic sp. G	*Melitaea trivia*	Spain	92.9% (13/14)	*w*CmelG
C. melitaearum cryptic sp. H	*Melitaea cinxia*	Finland	11.1% (6/54)	*w*CmelH1
C. melitaearum cryptic sp. H	*Melitaea cinxia*	Russia	27.3% (3/11)	Uncharacterized
C. glomerata	*Pieris sp.*		75% (6/8)	*w*Cglo
C. sesamiae	Stem boring moths		100% (1/1)	Uncharacterized
C. vestalis	*Plutella sp.*		75% (3/4)	Uncharacterized

B-supergroup *Wolbachia w*PipPel, but only 21 with the A-supergroup *Wolbachia w*MelPop reference (Fig. S4). Similarly, the ANI analysis, which calculates the average nucleotide identity among orthologous gene pairs shared between two genomes, revealed a high similarity between *w*PipPel, SRR13990441, and SRR13990442, with ANI values around 98% in pairwise comparisons (Table S10). In contrast, *w*MelPop displayed a lower ANI (~85%) in pairwise comparisons with *w*PipPel, SRR13990441, and

SRR13990442 (Table S10). Altogether, these results suggest the two *Wolbachia* assemblies from *Cotesia* belong to the B-supergroup *Wolbachia*.

Finally, we partially extracted the CI-associated genes from our *Wolbachia* assemblies. With this, we identified one copy of a Type I CifA in the SRR13990441 assembly (Table S11, Fig. S8), and a truncated/partial copy of *cifB* in both the SRR13990441 assembly (contig 109, position 1492–3201) and the SRR13990442 assembly (contig 221, position 1-1624). The sequences of the *cif*B gene from our *Wolbachia* assemblies were highly similar to that previously characterized from the fig wasp *Kradibia gibbosae* (Hymenoptera: Chalcidoidea) (WP_275944372.1), without any report of the role played by the symbiont in this host species [88].

Phylogenetic analyses
Our phylogenetic tree of the *COI* mitochondrial gene of 39 *Cotesia* species shows that the *Cotesia* wasps parasitizing Melitaeini butterflies belong to three distinct clades (See Fig. 1, S5-6:

- Clade 1 includes *C. melitaearum* cryptic species (D, E, F, G, H, I, M, N),
- Clade 2 includes *C. koebelei*,
- Clade 3 includes *C. bignellii* cryptic species C, and *C. acuminata* cryptic species (A, B, and K).

The *Wolbachia* phylogeny confirms that all *Wolbachia* strains characterized from *Cotesia* belonged to the A- and B-supergroups, with the majority (49/53, 92.4%) belonging to the B-supergroup (Fig. 1). Despite fewer representative taxa per phylogeny and lower resolution, phylogenies based only on individual gene alignments maintained similar sample groupings, with conserved strain assignment to supergroups A and B (Fig. S5), thus suggesting no recombination has occurred between the strains of the two supergroups in these *Cotesia* species. A visual comparison supports the lack of congruence and co-phylogeny between the maximum likelihood trees of *Cotesia* and their symbiotic strains. Phylogenetically close *Wolbachia* strains were found in phylogenetically distant *Cotesia* host species (Fig. 1, & S5-6), or in both a butterfly host and their *Cotesia* parasitoid (i.e. the *w*Mdid from a *M. didyma* butterfly and the *w*CmelF strain from *Cotesia* wasps emerging from Spanish *M. didyma* butterflies, Fig. S5).

Discussion
We built up from early studies to bring some light on the possible role(s) of endosymbionts on the evolutionary history of *Cotesia* parasitoid wasps, particularly on the divergence between sympatric cryptic species. The complex of *Cotesia* wasps parasitizing Melitaeini butterflies

belongs to three distinct clades, as previously shown by [44]. These three *Cotesia* clades are mostly specialists to the Melitaeini butterflies, but clustering based on the Lepidoptera host is not conserved across the genus *Cotesia*. Indeed, closely related *Cotesia* species to each of the three clades have been described as parasitoid wasps of divergent Lepidoptera. For example, *C. glomerata* parasitizes *Pieris sp.* butterflies (Pieridae), while *C. specularis* emerges from *Lampides boeticus* (Lycaenidae), and other *Cotesia* species are known to attack diverse moths (i.e. *Chilo sp.* for *C. flavipes*, or *Plutella sp.* for *C. vestalis*), with each of these Lepidoptera species feeding on a wide diversity of host plants. In their early studies, Kankare et al. [37, 42] suggested that direct competition between *Cotesia* wasps for the Melitaeini butterfly host species might have driven the divergence between their parasitoid wasp species and cryptic species. We did not detect the symbionts *Arsenophonus*, *Cardinium*, *Microsporidium* and *Spiroplasma*, and did not test for infection with *Rickettsia*, however the endosymbiotic bacterium *Wolbachia* was found in 61 (17.9%) of all our samples, covering 11 *Cotesia* species (52.4%) out of 21 included in the study. As in previous studies on *Cotesia* wasps species [39–41], and in insects in general [89, 90], such *Wolbachia* prevalence is still likely an underestimate of the true infection prevalence in the entire *Cotesia* genus. This is because our study covers only a small number of *Cotesia* species, populations and individuals representing only part of their geographic distributions [33].

Although the commonly used MLST markers have been criticised for being too conserved to allow reliable strain differentiation or infer precise phylogenetic relationships of closely related *Wolbachia* strains [91], comparison of the phylogenetic trees from *Cotesia* hosts and their *Wolbachia* symbionts clearly showed that distantly related *Cotesia* species share similar *Wolbachia* strains. Such lack of concordance between the host and the symbiont phylogenies has been previously described in diverse systems [22, 29]. This pattern suggests that *Wolbachia* strains have transferred horizontally, and not strictly vertically between mothers and their offspring. Because many *Cotesia* species occur in sympatry, sharing either their geographical ranges, their local habitats, their hosts, which in some cases also share the same host plants [37, 92, 93], the *Cotesia* species complex offers plausible opportunities for *Wolbachia* to transfer horizontally:

First, divergent *Cotesia* wasps might have acquired their *Wolbachia* infections while parasitising infected caterpillars. Between others, studies by Vavre et al. [29], and Qi et al. [94] provide evidence of such horizontal transfers between *Drosophila* flies and their parasitoid wasps, and between whiteflies and their parasitoid wapss, respectively. Although *Wolbachia* was previously

detected in *M. didyma* [95, 96], *M. athalia*, *M. britomartis* [97], *M. phoebe*, *M. ornata* [98] and *M. cinxia* [22], genetic sequences for most of those strains were not publicly available (Nov. 2023). However, we did find that a sequence from the *wsp* gene of the *Wolbachia* strain infecting *C. melitaearum*, parasitizing *M. didyma* caterpillars, was very similar to the *wsp* sequence from a strain infecting *M. didyma*. This suggests that this *Wolbachia* strain might have transferred between the *Cotesia* wasp and the Lepidoptera host, as also shown in *Nasonia* wasps and their *Drosophila* hosts [30]. But this is not always the case, as other strains characterized from other Lepidoptera species (i.e. *w*Cpar from *Chilo partellus*, *w*Prap from *Pieris rapae*, and *w*Pxyl from *Plutella xylostella*) were phylogenetically divergent from the strains found in the *Cotesia* wasps infecting those Lepidoptera (*C. flavipes*, *C. glomerata*, *C. vestalis*, respectively).

Second, *Wolbachia* could be exchanged between parasitoid wasps simultaneously parasitising the same host caterpillar. Such hypothesis was previously tested by [29], who found that *Leptopilina*, *Trichopria* and *Asobara* parasitoids of *Drosophila* flies can share identical *Wolbachia* bacteria, at least based on their *wsp* gene sequences. In Åland, *M. cinxia* is commonly parasitized by several parasitoid wasps [99]. Out of these, *Hyposoter horticola* is known to carry *Wolbachia* [100], and we showed that this *Wolbachia* strain (*w*Hho) is phylogenetically closely related to the *Wolbachia* characterized from *C. melitaearum*. These results suggest that at least some *Wolbachia* might also transfer horizontally between divergent parasitoid species sharing the same Lepidoptera hosts, or between the parasitoid wasps and their Lepidoptera hosts. In the future, hypotheses presented above could be more comprehensively tested in the *Cotesia* system by simultaneously screening *Cotesia* wasps and their Lepidoptera hosts for endosymbiotic infections.

The lack of co-divergence between the *Wolbachia* strains and their associated *Cotesia* lineages does not allow us to fully reject the hypothesis that the symbiont might be involved in the restricted gene flow between at least some of the sympatric *Cotesia* lineages [101]. For example, CI could occur between *Cotesia* lineages that carry divergent *Wolbachia* strains, such as *C. melitaearum* sp. F and G, or between *Cotesia* lineages of different infection status, such as *C. melitaearum* sp. D and E. In the future, experimental rearing [102, 103] and crossing between lineages, associated with microscopy imagery, should confirm the expression of CI between different host lineages. Indeed, because CI causes visible morphological abnormalities of sperm in the testes of infected males [104, 105], or cytological embryonic defects, microscopic approaches may be used to confirm post-mating isolation between *Cotesia* lineages, as shown previously in *Culex pipiens*, *Drosophila simulans* and the

parasitoid wasp, *Nasonia* [106–109]. Here, we isolated a complete homolog of the *cifA* gene and a partial homolog of *cifB* gene, which code for the *Wolbachia*-induced CI phenotype in other species [83, 110, 111]. However, with the growing general interest for whole genome sequencing of parasitoid wasps, especially of species used as agricultural pest control agents, we expect more *Cotesia* genomic projects to be completed in the near future. These projects will hopefully assemble and analyse the whole genomes of *Wolbachia* and other endosymbiotic species associated with *Cotesia* hosts, and soon provide material for holistic estimate of the endosymbionts patterns of diversity, and their function(s) in this complex of parasitoid wasp species.

Abbreviations
CI   Cytoplasmic Incompatibility
MLST   Multi Locus Sequence Typing

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12866-024-03470-7.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Supplementary Material 6

Supplementary Material 7

Supplementary Material 8

Supplementary Material 9

Supplementary Material 10

Supplementary Material 11

Supplementary Material 12

Supplementary Material 13

Supplementary Material 14

Supplementary Material 15

Supplementary Material 16

Supplementary Material 17

Supplementary Material 18

Supplementary Material 19

Supplementary Material 20

Supplementary Material 21

Data availability
All sequence alignments and final assemblies are available from Zenodo at https://doi.org/10.5281/zenodo.8422079. All sequences were deposited on NCBI with the accession numbers OR597552-OR597565 (wsp), OR608488 - OR608500 (hcpA), OR640995-OR641006 (coxA), OR641007-OR641026 (ftsZ), OR641031- OR641044 (gatB), and OR641048-OR641057 (fbpA).

Declarations

Ethics approval and consent to participate
The data does not include any personal data. The collection of the samples dated prior to 2004, and the author list includes a member of a Spanish institution, which is in line with the current requirement of the Nagoya protocol on Access and Benefit-sharing.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References
1. Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH. How many species are infected with *Wolbachia*?- a statistical analysis of current data. FEMS Microbiol Lett. 2008;281(2):215–20. https://doi.org/10.1111/j.1574-6968.2008.01110.x.
2. Engelstädter J, Hurst GD. The impact of male-killing bacteria on host evolutionary processes. Genetics. 2007;175(1):245–54. https://doi.org/10.1534/genetics.106.060921.
3. O'Neill SL, Hoffmann AA, Werren JH. Influential passengers: inherited microorganisms and arthropod reproduction. Oxford; New York: Oxford University Press; 1997.
4. Perlman SJ, Hunter MS, Zchori-Fein E. The emerging diversity of *Rickettsia*. Proc Biol Sci. 2006;273(1598):2097–106. https://doi.org/10.1098/rspb.2006.3541.
5. Cordaux R, Bouchon D, Greve P. The impact of endosymbionts on the evolution of host sex-determination mechanisms. Trends Genet. 2011;27(8):332–41. https://doi.org/10.1016/j.tig.2011.05.002.
6. Pollmann M, Moore LD, Krimmer E, D'Alvise P, Hasselmann M, Perlman SJ, et al. Highly transmissible cytoplasmic incompatibility by the extracellular insect symbiont *Spiroplasma*. iScience. 2022;25(5):104335. https://doi.org/10.1016/j.isci.2022.104335.
7. Richard F-J. Symbiotic bacteria influence the odor and mating preference of their hosts. Front Ecol Evol. 2017;5. https://doi.org/10.3389/fevo.2017.00143.
8. Vala F, Egas M, Breeuwer JA, Sabelis MW. *Wolbachia* affects oviposition and mating behaviour of its spider mite host. J Evol Biol. 2004;17(3):692–700. https://doi.org/10.1046/j.1420-9101.2003.00679.x.
9. Engl T, Kaltenpoth M. Influence of microbial symbionts on insect pheromones. Nat Prod Rep. 2018;35(5):386–97. https://doi.org/10.1039/C7NP00068E.
10. Werren JH. Biology of *Wolbachia*. Annu Rev Entomol. 1997;42:587–609. https://doi.org/10.1146/annurev.ento.42.1.587.
11. Bordenstein SR, O'Hara FP, Werren JH. *Wolbachia*-induced incompatibility precedes other hybrid incompatibilities in *Nasonia*. Nature. 2001;409(6821):707–10. https://doi.org/10.1038/35055543.

12. Telschow A, Flor M, Kobayashi Y, Hammerstein P, Werren JH. *Wolbachia*-induced unidirectional cytoplasmic incompatibility and speciation: mainland-island model. PLoS ONE. 2007;2(8):e701. https://doi.org/10.1371/journal.pone.0000701.

13. Chafee ME, Zecher CN, Gourley ML, Schmidt VT, Chen JH, Bordenstein SR, et al. Decoupling of host-symbiont-phage coadaptations following transfer between insect species. Genetics. 2011;187(1):203–15. https://doi.org/10.1534/genetics.110.120675.

14. Miller WJ, Ehrman L, Schneider D. Infectious speciation revisited: impact of symbiont-depletion on female fitness and mating behavior of *Drosophila paulistorum*. PLoS Pathog. 2010;6(12):e1001214. https://doi.org/10.1371/journal.ppat.1001214.

15. Shoemaker DD, Katju V, Jaenike J. *Wolbachia* and the evolution of reproductive isolation between *Drosophila recens* and *Drosophila subquinaria*. Evolution. 1999;53(4):1157–64. https://doi.org/10.1111/j.1558-5646.1999.tb04529.x.

16. Mayhew PJ. Why are there so many insect species? Perspectives from fossils and phylogenies. Biol Rev Camb Philos Soc. 2007;82(3):425–54. https://doi.org/10.1111/j.1469-185X.2007.00018.x.

17. Takiya DM, Tran PL, Dietrich CH, Moran NA. Co-cladogenesis spanning three phyla: leafhoppers (Insecta: Hemiptera: Cicadellidae) and their dual bacterial symbionts. Mol Ecol. 2006;15(13):4175–91. https://doi.org/10.1111/j.1365-294X.2006.03071.x.

18. Chen X, Li S, Aksoy S. Concordant evolution of a symbiont with its host insect species: molecular phylogeny of genus *Glossina* and its bacteriome-associated endosymbiont, *Wigglesworthia glossinidia*. J Mol Evol. 1999;48(1):49–58. https://doi.org/10.1007/pl00006444.

19. Balvin O, Roth S, Talbot B, Reinhardt K. Co-speciation in bedbug *Wolbachia* parallel the pattern in nematode hosts. Sci Rep. 2018;8(1):8797. https://doi.org/10.1038/s41598-018-25545-y.

20. Bandi C, Anderson TJ, Genchi C, Blaxter ML. Phylogeny of *Wolbachia* in filarial nematodes. Proc Biol Sci. 1998;265(1413):2407–13. https://doi.org/10.1098/rspb.1998.0591.

21. van Opijnen T, Baudry E, Baldo L, Bartos J, Werren JH. Genetic variability in the three genomes of *Nasonia*: nuclear, mitochondrial and *Wolbachia*. Insect Mol Biol. 2005;14(6):653–63. https://doi.org/10.1111/j.1365-2583.2005.00595.x.

22. Ahmed MZ, Breinholt JW, Kawahara AY. Evidence for common horizontal transmission of *Wolbachia* among butterflies and moths. BMC Evol Biol. 2016;16(1):118. https://doi.org/10.1186/s12862-016-0660-x.

23. Chrostek E, Pelz-Stelinski K, Hurst GDD, Hughes GL. Horizontal transmission of intracellular insect symbionts via plants. Front Microbiol. 2017;8:2237. https://doi.org/10.3389/fmicb.2017.02237.

24. Li SJ, Ahmed MZ, Lv N, Shi PQ, Wang XM, Huang JL, et al. Plant mediated horizontal transmission of *Wolbachia* between whiteflies. ISME J. 2017;11(4):1019–28. https://doi.org/10.1038/ismej.2016.164.

25. Sintupachee S, Milne JR, Poonchaisri S, Baimai V, Kittayapong P. Closely related *Wolbachia* strains within the pumpkin arthropod community and the potential for horizontal transmission via the plant. Microb Ecol. 2006;51(3):294–301. https://doi.org/10.1007/s00248-006-9036-x.

26. Stahlhut JK, Desjardins CA, Clark ME, Baldo L, Russell JA, Werren JH, et al. The mushroom habitat as an ecological arena for global exchange of *Wolbachia*. Mol Ecol. 2010;19(9):1940–52. https://doi.org/10.1111/j.1365-294X.2010.04572.x.

27. Zug R, Koehncke A, Hammerstein P. Epidemiology in evolutionary time: the case of *Wolbachia* horizontal transmission between arthropod host species. J Evol Biol. 2012;25(11):2149–60. https://doi.org/10.1111/j.1420-9101.2012.02601.x.

28. Ke F, You S, Huang S, Chen W, Liu T, He W, et al. Herbivore range expansion triggers adaptation in a subsequently-associated third trophic level species and shared microbial symbionts. Sci Rep. 2019;9(1):10314. https://doi.org/10.1038/s41598-019-46742-3.

29. Vavre F, Fleury F, Lepetit D, Fouillet P, Bouletreau M. Phylogenetic evidence for horizontal transmission of *Wolbachia* in host-parasitoid associations. Mol Biol Evol. 1999;16(12):1711–23. https://doi.org/10.1093/oxfordjournals.molbev.a026084.

30. Raychoudhury R, Baldo L, Oliveira DC, Werren JH. Modes of acquisition of *Wolbachia*: horizontal transfer, hybrid introgression, and codivergence in the *Nasonia* species complex. Evolution. 2009;63(1):165–83; https://doi.org/10.1111/j.1558-5646.2008.00533.x

31. Quicke DLJ. Phylogeny and Systematics of the Ichneumonidae. In: The Braconid and Ichneumonid parasitoid wasps. 2014. pp. 341–449.

32. Fernandez-Triana J, Shaw MR, Boudreault C, Beaudin M, Broad GR. Annotated and illustrated world checklist of *Microgastrinae* parasitoid wasps (Hymenoptera, Braconidae). Zookeys. 2020;920:1–1090. https://doi.org/10.3897/zookeys.920.39128.

33. Mason WRM. The polyphyletic nature of A*panteles foerster* (Hymenoptera: Braconidae): a phylogeny and reclassification of *Microgastrinae*. Mem Entomol Soc Can. 1981;113(S115):1–147. https://doi.org/10.4039/entm113115fv.

34. Shaw MR, Stefanescu C, van Nouhuys S. Parasitoids of European butterflies. In: Settele J, Shreeve TG, Konvicka M, Van H, editors. Ecology of butterflies of Europe. Cambridge: Cambridge University Press; 2009. pp. 130–56.

35. Lei G-C, Hanski I. Metapopulation structure of *Cotesia melitaearum*, a specialist parasitoid of the butterfly *Melitaea Cinxia*. Oikos. 1997;78(1):91–100. https://doi.org/10.2307/3545804.

36. Opedal Ø, Ovaskainen O, Saastamoinen M, Laine A-L, van Nouhuys S. Host plant availability drives the spatio-temporal dynamics of interacting metapopulations across a fragmented landscape. Ecology. 2020;101(12):e03186. https://doi.org/10.1002/ecy.3186.

37. Kankare M, Stefanescu C, Van Nouhuys S, Shaw MR. Host specialization by *Cotesia* wasps (Hymenoptera: Braconidae) parasitizing species-rich Melitaeini (Lepidoptera: Nymphalidae) communities in north-eastern Spain. Biol J Linn Soc. 2005;86(1):45–65. https://doi.org/10.1111/j.1095-8312.2005.00523.x.

38. Ehrlich PR, Hanski I. On the wings of checkerspots: a model system for population biology. Oxford University Press; 2004.

39. Rattan RS, Hadapad AB, Reineke A, Gupta PR, Zebitz CPW. Molecular evidence for the presence of the endosymbiontic bacteria *Wolbachia* in *Cotesia* populations (Hymenoptera: Braconidae). J Asia Pac Entomol. 2011;14(2):183–5. https://doi.org/10.1016/j.aspen.2010.12.009.

40. Mochiah MB, Ngi-Song AJ, Overholt WA, Stouthamer R. *Wolbachia* infection in *Cotesia sesamiae* (Hymenoptera: Braconidae) causes cytoplasmic incompatibility: implications for biological control. Biol Control. 2002;25(1):74–80. https://doi.org/10.1016/S1049-9644(02)00045-2.

41. Branca A, BP LER, Vavre F, Silvain JF, Dupas S. Intraspecific specialization of the generalist parasitoid *Cotesia sesamiae* revealed by polyDNAvirus polymorphism and associated with different *Wolbachia* infection. Mol Ecol. 2011;20(5):959–71. https://doi.org/10.1111/j.1365-294X.2010.04977.x.

42. Kankare M, Van Nouhuys S, Hanski I. Genetic divergence among host-specific cryptic species in *Cotesia melitaearum* Aggregate (Hymenoptera: Braconidae), parasitoids of Checkerspot butterflies. Ann Entomol Soc Am. 2005;98(3):382–94. https://doi.org/10.1603/0013-8746(2005)098[0382:GDAHCS]2.0.CO;2.

43. Duplouy A, Hornett EA. Uncovering the hidden players in Lepidoptera biology: the heritable microbial endosymbionts. PeerJ. 2018;6:e4629. https://doi.org/10.7717/peerj.4629.

44. Kankare M, Shaw MR. Molecular phylogeny of *Cotesia* Cameron, 1891 (Insecta: Hymenoptera: Braconidae: Microgastrinae) parasitoids associated with Melitaeini butterflies (Insecta: Lepidoptera: Nymphalidae: Melitaeini). Mol Phylogenet Evol. 2004;32(1):207–20. https://doi.org/10.1016/j.ympev.2003.11.013.

45. Kankare M, van Nouhuys S, Gaggiotti O, Hanski I. Metapopulation genetic structure of two coexisting parasitoids of the Glanville Fritillary butterfly. Oecologia. 2005;143(1):77–84. https://doi.org/10.1007/s00442-004-1782-1.

46. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial *cytochrome c oxidase subunit I* from diverse metazoan invertebrates. Mol Mar Biol Biotechnol. 1994;3(5):294–9.

47. Alexeeva I, Elliott EJ, Rollins S, Gasparich GE, Lazar J, Rohwer RG. Absence of *Spiroplasma* or other bacterial 16s rRNA genes in brain tissue of hamsters with scrapie. J Clin Microbiol. 2006;44(1):91–7. https://doi.org/10.1128/JCM.44.1.91-97.2006.

48. Duron O, Bouchon D, Boutin S, Bellamy L, Zhou L, Engelstadter J, et al. The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. BMC Biol. 2008;6:27. https://doi.org/10.1186/1741-7007-6-27.

49. Gotoh T, Noda H, Ito S. *Cardinium* symbionts cause cytoplasmic incompatibility in spider mites. Heredity (Edinb). 2007;98(1):13–20. https://doi.org/10.1038/sj.hdy.6800881.

50. Thao ML, Baumann P. Evidence for multiple acquisition of *Arsenophonus* by Whitefly species (Sternorrhyncha: Aleyrodidae). Curr Microbiol. 2004;48(2):140–4. https://doi.org/10.1007/s00284-003-4157-7.

51. Baldo L, Dunning Hotopp JC, Jolley KA, Bordenstein SR, Biber SA, Choudhury RR, et al. Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. Appl Environ Microbiol. 2006;72(11):7098–110. https://doi.org/10.1128/AEM.00731-06.

52. Terry RS, Smith JE, Bouchon D, Rigaud T, Duncanson P, Sharpe RG, et al. Ultrastructural characterisation and molecular taxonomic identification of *Nosema granulosis* n. sp., a transovarially transmitted feminising (TTF)

microsporidium. J Eukaryot Microbiol. 1999;46(5):492–9. https://doi.org/10.1111/j.1550-7408.1999.tb06066.x.

53. Deng J, Assandri G, Chauhan P, Futahashi R, Galimberti A, Hansson B, et al. *Wolbachia*-driven selective sweep in a range expanding insect species. BMC Ecol Evol. 2021;21(1):181. https://doi.org/10.1186/s12862-021-01906-6.

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

55. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. Bioinformatics. 2011;27(6):863–4. https://doi.org/10.1093/bioinformatics/btr026.

56. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20. https://doi.org/10.1093/bioinformatics/btu170.

57. Andrews S. FastQC: A quality control tool for high throughput sequence data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc. In.; 2010.

58. Ewels P, Magnusson M, Lundin S, Kaller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32(19):3047–8. https://doi.org/10.1093/bioinformatics/btw354.

59. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. Bioinformatics. 2018;34(15):2666–9. https://doi.org/10.1093/bioinformatics/bty149. NanoPack: visualizing and processing long-read sequencing data.

60. Valerio F, Twort V, Duplouy A. Screening host genomic projects for *Wolbachia* infections. In: Fallon AM, editor. Methods in Molecular Biology - *Wolbachia*. In press: Springer; 2023.

61. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol. 2019;20(1):257. https://doi.org/10.1186/s13059-019-1891-0.

62. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–9. https://doi.org/10.1038/nmeth.1923.

63. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34(18):3094–100. https://doi.org/10.1093/bioinformatics/bty191.

64. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078–9. https://doi.org/10.1093/bioinformatics/btp352.

65. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-; 2016.

66. Wick RR, Judd LM, Gorrie CL, Holt KE, Unicycler. Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol. 2017;13(6):e1005595. https://doi.org/10.1371/journal.pcbi.1005595.

67. Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. Versatile genome assembly evaluation with QUAST-LG. Bioinformatics. 2018;34(13):i142–50. https://doi.org/10.1093/bioinformatics/bty266.

68. Manni M, Berkeley MR, Seppey M, Simao FA, Zdobnov EM. BUSCO Update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol Biol Evol. 2021;38(10):4647–54. https://doi.org/10.1093/molbev/msab199.

69. Jain C, Rodriguez RL, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun. 2018;9(1):5114. https://doi.org/10.1038/s41467-018-07641-9.

70. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068–9. https://doi.org/10.1093/bioinformatics/btu153.

71. Martinez J, Klasson L, Welch JJ, Jiggins FM. Life and death of selfish genes: comparative genomics reveals the dynamic evolution of cytoplasmic incompatibility. Mol Biol Evol. 2021;38(1):2–15. https://doi.org/10.1093/molbev/msaa209.

72. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792–7. https://doi.org/10.1093/nar/gkh340.

73. Zhou W, Rousset F, O'Neil S. Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. Proc Biol Sci. 1998;265(1395):509–15. https://doi.org/10.1098/rspb.1998.0324.

74. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80. https://doi.org/10.1093/molbev/mst010.

75. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics. 2014;30(22):3276–8. https://doi.org/10.1093/bioinformatics/btu531.

76. Stamatakis A. Bioinformatics. 2014;30(9):1312–3. https://doi.org/10.1093/bioinformatics/btu033. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies.

77. Edler D, Klein J, Antonelli A, Silvestro D. raxmlGUI 2.0: a graphical interface and toolkit for phylogenetic analyses using RAxML. Methods Ecol Evol. 2021;12(2):373–7. https://doi.org/10.1111/2041-210x.13512.

78. Foster J, Ganatra M, Kamal I, Ware J, Makarova K, Ivanova N, et al. The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode. PLoS Biol. 2005;3(4):e121. https://doi.org/10.1371/journal.pbio.0030121.

79. Nikoh N, Hosokawa T, Moriyama M, Oshima K, Hattori M, Fukatsu T. Evolutionary origin of insect-*Wolbachia* nutritional mutualism. Proc Natl Acad Sci U S A. 2014;111(28):10257–62. https://doi.org/10.1073/pnas.1409284111.

80. 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(5):1530–4. https://doi.org/10.1093/molbev/msaa015.

81. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14(6):587–9. https://doi.org/10.1038/nmeth.4285.

82. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018;35(2):518–22. https://doi.org/10.1093/molbev/msx281.

83. LePage DP, Metcalf JA, Bordenstein SR, On J, Perlmutter JI, Shropshire JD, et al. Prophage *WO* genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. Nature. 2017;543(7644):243–7. https://doi.org/10.1038/nature21391.

84. Lindsey ARI, Rice DW, Bordenstein SR, Brooks AW, Bordenstein SR, Newton ILG. Evolutionary genetics of cytoplasmic incompatibility genes *cifA* and *cifB* in prophage *WO* of *Wolbachia*. Genome Biol Evol. 2018;10(2):434 – 51; https://doi.org/10.1093/gbe/evy012.

85. Bing XL, Zhao DS, Sun JT, Zhang KJ, Hong XY. Genomic analysis of *Wolbachia* from *Laodelphax striatellus* (Delphacidae, Hemiptera) reveals insights into its Jekyll and Hyde mode of infection pattern. Genome Biol Evol. 2020;12(2):3818–31. https://doi.org/10.1093/gbe/evaa006.

86. Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T. ModelTest-NG: a New and scalable tool for the selection of DNA and protein evolutionary models. Mol Biol Evol. 2020;37(1):291–4. https://doi.org/10.1093/molbev/msz189.

87. Letunic I, Bork P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49(W1):W293–6. https://doi.org/10.1093/nar/gkab301.

88. Miao Y-h, Xiao J-h, Huang D-w. Distribution and evolution of the bacteriophage *WO* and its antagonism with *Wolbachia*. Front Microbiol. 2020;11:595629. https://doi.org/10.3389/fmicb.2020.595629.

89. Sazama EJ, Bosch MJ, Shouldis CS, Ouellette SP, Wesner JS. Incidence of *Wolbachia* in aquatic insects. Ecol Evol. 2017;7(4):1165–9. https://doi.org/10.1002/ece3.2742.

90. Weinert LA, Araujo-Jnr EV, Ahmed MZ, Welch JJ. The incidence of bacterial endosymbionts in terrestrial arthropods. Proceedings of the Royal Society B: Biological Sciences. 2015;282(1807):20150249; https://doi.org/10.1098/rspb.2015.0249.

91. Bleidorn C, Gerth M. A critical re-evaluation of multilocus sequence typing (MLST) efforts in *Wolbachia*. FEMS Microbiol Ecol. 2018;94(1). https://doi.org/10.1093/femsec/fix163.

92. Mutamiswa R, Machekano H, Chidawanyika F, Nyamukondiwa C. Thermal resilience may shape population abundance of two sympatric congeneric *Cotesia* species (Hymenoptera: Braconidae). PLoS ONE. 2018;13(2):e0191840. https://doi.org/10.1371/journal.pone.0191840.

93. Bredlau JP, Kuhar D, Gundersen-Rindal DE, Kester KM. The parasitic wasp, *Cotesia congregata* (say), consists of two incipient species isolated by asymmetric reproductive incompatibility and hybrid inability to overcome host defenses. Front Ecol Evol. 2019;7. https://doi.org/10.3389/fevo.2019.00187.

94. Qi LD, Sun JT, Hong XY, Li YX. Diversity and phylogenetic analyses reveal horizontal transmission of endosymbionts between whiteflies and their parasitoids. J Econ Entomol. 2019;112(2):894–905. https://doi.org/10.1093/jee/toy367.

95. Dinca V, Lee KM, Vila R, Mutanen M. The conundrum of species delimitation: a genomic perspective on a mitogenetically super-variable butterfly. Proc Biol Sci. 2019;286(1911):20191311. https://doi.org/10.1098/rspb.2019.1311.

96. Russell JA, Funaro CF, Giraldo YM, Goldman-Huertas B, Suh D, Kronauer DJC, et al. A veritable menagerie of heritable bacteria from ants, butterflies, and beyond: broad molecular surveys and a systematic review. PLoS ONE. 2012;7(12):e51027. https://doi.org/10.1371/journal.pone.0051027.

97.   Ilinsky Y, Kosterin OE. Molecular diversity of *Wolbachia* in Lepidoptera: preva-lent allelic content and high recombination of MLST genes. Mol Phylogenet Evol. 2017;109:164–79. https://doi.org/10.1016/j.ympev.2016.12.034.

98.   Tóth JP, Varga Z, Verovnik R, Wahlberg N, Váradi A, Bereczki J. Mito-nuclear discordance helps to reveal the phylogeographic patterns of *Melitaea ornata* (Lepidoptera: Nymphalidae). Biol J Linn Soc. 2017;121(2):267–81. https://doi.org/10.1093/biolinnean/blw037.

99.   Lei GC, Vikberg V, Nieminen M, Kuussaari M. The parasitoid complex attacking Finnish populations of the Glanville Fritillary *Melitaea Cinxia* (Lep: Nymphali-dae), an endangered butterfly. J Nat Hist. 1997;31(4):635–48. https://doi.org/10.1080/00222939700770301.

100. Duplouy A, Couchoux C, Hanski I, van Nouhuys S. *Wolbachia* infection in a natural parasitoid wasp population. PLoS ONE. 2015;10(8):e0134843. https://doi.org/10.1371/journal.pone.0134843.

101. Rokas II. *Wolbachia* as a speciation agent. Trends Ecol Evol. 2000;15(2):44–5. doi: 10.1016/s0169-5347(99)01783-8.

102. Ngi-Song AJ, Kimani-Njogu S, Overholt WA. Multiple parasitism by Cotesia sesamiae and Cotesia flavipes (Hymenoptera: Braconidae) on Busseola Fusca (Lepidoptera: Noctuidae). Biocontrol Sci Technol. 2001;11(3):381–90. https://doi.org/10.1080/09583150120055790.

103. van Nouhuys S, Niemikapee S, Hanski I. Variation in a host-parasitoid interac-tion across independent populations. Insects. 2012;3(4):1236–56. https://doi.org/10.3390/insects3041236.

104. Champion de Crespigny FE, Wedell N. *Wolbachia* infection reduces sperm competitive ability in an insect. Proc Biol Sci. 2006;273(1593):1455–8. https://doi.org/10.1098/rspb.2006.3478.

105. Ferree PM, Aldrich JC, Jing XA, Norwood CT, Van Schaick MR, Cheema MS, et al. Spermatogenesis in haploid males of the jewel wasp *Nasonia Vitripennis*. Sci Rep. 2019;9(1):12194. https://doi.org/10.1038/s41598-019-48332-9.

106. Bonneau M, Landmann F, Labbe P, Justy F, Weill M, Sicard M. The cellular phenotype of cytoplasmic incompatibility in *Culex pipiens* in the light of *cidB* diversity. PLoS Pathog. 2018;14(10):e1007364. https://doi.org/10.1371/journal.ppat.1007364.

107. Landmann F, Orsi GA, Loppin B, Sullivan W. *Wolbachia*-mediated cytoplasmic incompatibility is associated with impaired histone deposition in the male pronucleus. PLoS Pathog. 2009;5(3):e1000343. https://doi.org/10.1371/journal.ppat.1000343.

108. Reed KM, Werren JH. Induction of paternal genome loss by the paternal-sex-ratio chromosome and cytoplasmic incompatibility bacteria (*Wolba-chia*): a comparative study of early embryonic events. Mol Reprod Dev. 1995;40(4):408–18. https://doi.org/10.1002/mrd.1080400404.

109. Tram U, Fredrick K, Werren JH, Sullivan W. Paternal chromosome segregation during the first mitotic division determines *Wolbachia*-induced cytoplasmic incompatibility phenotype. J Cell Sci. 2006;119(Pt 17):3655–63. https://doi.org/10.1242/jcs.03095.

110. Beckmann JF, Fallon AM. Detection of the *Wolbachia* protein WPIP0282 in mosquito spermathecae: implications for cytoplasmic incompatibility. Insect Biochem Mol Biol. 2013;43(9):867–78. https://doi.org/10.1016/j.ibmb.2013.07.002.

111. Shropshire JD, On J, Layton EM, Zhou H, Bordenstein SR. One prophage *WO* gene rescues cytoplasmic incompatibility in *Drosophila melanogaster*. Proc Natl Acad Sci U S A. 2018;115(19):4987–91. https://doi.org/10.1073/pnas.1800650115.