

RESEARCH ARTICLE

Progress in the molecular phylogeny of *Cotesia acuminata* and *C. melitaearum* cryptic species complexes

Julian Schach^{1*}, Federica Valerio¹, Anne Duploux^{1,2*}

1 Department of Organismal and Evolutionary Biology, Faculty of Biological and Environmental Sciences, The University of Helsinki, Helsinki, Finland, **2** Institute of Life Sciences (HiLIFE), The University of Helsinki, Helsinki, Finland

* julian.schach@helsinki.fi (JS); anne.duploux@helsinki.fi (AD)



OPEN ACCESS

Citation: Schach J, Valerio F, Duploux A (2026) Progress in the molecular phylogeny of *Cotesia acuminata* and *C. melitaearum* cryptic species complexes. PLoS One 21(5): e0348973. <https://doi.org/10.1371/journal.pone.0348973>

Editor: Petr Heneberg, Charles University: Univerzita Karlova, CZECHIA

Received: January 15, 2026

Accepted: April 23, 2026

Published: May 27, 2026

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0348973>

Copyright: © 2026 Schach et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Cryptic species present major challenges for biodiversity and evolutionary research due to the morphological similarity and frequent lack of available genetic and ecological data. In parasitoid wasps, this is especially true for lineages with limited taxonomic and genetic sampling and unclear geographic ranges. Here we reconstruct the phylogeny of *Cotesia* wasps parasitizing Melitaeini butterflies, including two cryptic species complexes (*Cotesia acuminata* agg. and *C. melitaearum* agg.). Using a ten-gene dataset from samples collected in Europe, Asia, and North America, we inferred relationships among 22 *Cotesia* species using maximum likelihood. We also included non-Melitaeini associated *Cotesia* to assess whether the parasitoids of Melitaeini form a monophyletic group. Our analyses yielded a highly supported phylogeny, revealing four major clades, three of which included the Melitaeini associated species. This confirms that the *Cotesia* parasitizing Melitaeini butterflies are polyphyletic, likely resulting from independent host shifts across the genus. Each clade is further subdivided into subclades corresponding to the different cryptic species complexes, clarifying previously unresolved relationships. These results provide a robust framework for future studies on the evolution, ecology, and host use dynamics of *Cotesia* wasps and highlight the utility of multi-locus data for resolving phylogenies in morphologically cryptic taxa.

Introduction

Species have historically been described using morphological data, and species relationships based on degrees of similarities of morphological characteristics. More recently, the advent of molecular techniques has refined many taxonomic/phylogenetic classifications, by allowing the study of species relationships and biodiversity with much higher resolution [1]. Cryptic species are two or more molecularly distinct species, occasionally accompanied by rapid speciation, which are or have been

Data availability statement: All sequence data is available from NCBI. Accession numbers: PX523132 - PX523219, PX559195 - PX559573, PX642706 - PX642796, and PX648620 - PX649011.

Funding: The work was funded by the Research Council of Finland to Dr. Anne Duploux grant number 328944. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

erroneously classified as a single species because they are morphologically indistinguishable [2]. Such cryptic species are known to span the tree of life [3–6]. Their hidden genetic diversity has however challenged our ability to estimate the true biodiversity richness on Earth, as many remain to be discovered and characterized [7]. Furthermore, with the current biodiversity crisis, we are likely losing many such species at alarming rates even before they have been taxonomically characterized, or before their ecosystem roles were understood [8].

Parasitoid wasps make up one of the most diverse animal groups in the world [9]. They lay their eggs within or on their hosts' body, and the parasitoid larvae develop feeding on the host, until emergence of the adult wasp [10]. These parasitoids can be very efficient at killing their hosts and can act as key evolutionary pressures on their host species population growth [11] and host population dynamics. For example, it has been shown that the specialist species *Cotesia melitaeorum* (Wilkinson), parasitizes only 10% of *Melitaea cinxia* (Linnaeus) caterpillars but can still cause localized extinctions within the butterfly host metapopulation [12]. Because of their narrow host ranges and highly efficient killing of their hosts, several species of parasitoids have become popular biocontrol agents in sustainable agriculture, contributing significantly to the \$417 billion USD global value of insect biocontrol [13]. For example, *Cotesia ruficrus* (Haliday), parasitizes the fall armyworm (*Spodoptera frugiperda*, Lepidoptera: Noctuidae, Smith), which is contributing to food insecurity in Africa, Asia, and the Pacific [14].

Most of the world's parasitoid diversity occurs within the superfamily Ichneumonoidea, which includes the two families Ichneumonidae (Latreille) and Braconidae (Nees von Esenbeck), with approximately 60,000 described species and many more estimated [15]. Within the Braconidae, the subfamily Microgastrinae is the most diverse. For example, the genus *Cotesia* (Cameron) contains an estimated 1500–2000 species worldwide [16] of which over 300 have been named as species [17]. *Cotesia* wasps usually exhibit narrow host ranges, and closely related species often differ in their host species preferences, attacking one or two closely related lepidopteran host species [18]. For example, although previous studies suggested that seven species of *Cotesia* were generalist parasites of the Melitaeini (Newman) butterfly species found across Europe and Asia [19,20], more recent molecular analyses accompanied by rearing experiments provided evidence that the originally considered generalist parasitoid species, *Cotesia acuminata* (Reinhard) and *C. melitaeorum*, were in fact comprised of numerous cryptic species, each with narrow Melitaeini butterfly host specificity [18,21].

In the complex species community of Melitaeini butterflies occurring in Catalonia in northeastern Spain (Catalan Butterfly Monitoring Scheme, CMBS), 10 local Melitaeini butterfly species were identified as hosts to at least 10 *Cotesia* (cryptic) species [18]. Unfortunately, the original barriers to geneflow between the current cryptic lineages of the wasps remain unclear, as studies exploring isolation by distance or endosymbiotic bacteria failed to explain the sympatric speciation of the parasitoid species [18,22,23]. Currently, the phylogeny of the Catalanian *C. acuminata* complex and *C. melitaeorum* complex are based on 12 microsatellite loci [18], while a wider Eurasian

phylogeny of these groups combined two mitochondrial (*COI* and *ND1*) and one nuclear (*ITS2*) gene sequences [19]. As a consequence of the limited available genetic data included in these phylogenies, the relationships between *Cotesia* species still contain several unresolved nodes (i.e., nodes with low bootstrap support), and we still lack a clear description of the diversity and the speciation scenarios between *Cotesia* wasps attacking Melitaeini butterflies [18,19,21].

Here, we aimed to refine the phylogeny of *Cotesia* species associated with 16 species of Melitaeini butterflies by using ten markers and a wider sample diversity than previous studies. We analysed sequences of two mtDNA and eight nDNA markers from both freshly collected specimens from Finland, Italy, Spain, and Switzerland, as well as DNA extracts from samples collected between 1979–2002 from China, Finland, France, Hungary, Russia, Spain, Sweden, and USA [19]. With this, we expected to resolve the previously unclear nodes in the *Cotesia* phylogeny.

Materials and methods

Samples

This study includes *Cotesia* parasitoid wasp specimens that had emerged from the caterpillars of 16 species of Melitaeini butterflies. These include *Euphydryas aurinia* (Rottemburg), *E. aurinia davidi* (Wahlburg), *E. cynthia* (Dennis and Schifferrmueller), *E. desfontainii* (Godart), *E. editha* (Boisduval), *E. maturna* (Linnaeus), *E. phaeton* (Drury), *Melitaea athalia* (Rottemburg), *M. cinxia*, *M. deione* (Geyer), *M. didyma* (Esper), *M. latonigena* (Eversmann), *M. parthenoides* (Kefenstein), *M. phoebe* (Dennis and Schifferrmueller), *M. scotosia* (Butler), and *M. trivia* (Dennis and Schifferrmueller) (Table 1).

Among the *Cotesia* wasps, two cryptic species complexes are known to consist of multiple species: *C. acuminata* and *C. melitaeorum*. Following previously introduced convention by Kankare *et al.* [21], we refer to these as *C. acuminata* agg. and *C. melitaeorum* agg., with individual lineages denoted by species-level designations (e.g., *Sp. A* to *Sp. H*).

The Melitaeini caterpillars were collected over two time periods: 1979–2002 and 2022–2023. The DNA from early samples collected for the study of *Cotesia* wasps' phylogeny and butterfly-host specialization in earlier studies was extracted in the early 2000's. These DNA extracts represent 10 species from eight different countries, including China (N=3 from one species), Finland (N=5 from one species), France (N=7 from three species), Hungary (N=3 from one species), Russia (N=6 from two species), Spain (N=2 from one species), Sweden (N=3 from one species), and USA (N=14 from two species) (Table 1). Additionally, caterpillars from one locality in Finland (Åland) were also collected in the fall of 2022. Finally, parasitized Melitaeini caterpillars were collected in early April 2023 from five localities in Catalonia, Spain (Coll d'Estenalles, El Brull, Fígols-El Forn, Pyrenees), two localities in Switzerland (Vaud and Valais), and one locality in Italy (Aosta) (Fig 1). After collection, the caterpillars were reared in the laboratory on excess of their respective host plant, until the adult wasps emerged. Freshly emerged adult wasps were killed by placing them in a freezer for 24 hours. All samples were stored in 90% ethanol at –20°C until DNA extraction.

To test whether the *Cotesia* species parasitizing Melitaeini butterflies form a monophyletic group, we also included gene sequences retrieved from *Cotesia chilonis* (Munakata), *C. congregata* (Say), *C. glomerata* (Linnaeus), *C. ruficrus*, *C. typhae* (Fernández-Triana), and *C. vestalis* (Haliday), for which full genomes are publicly available from the NCBI database. These *Cotesia* species naturally parasitize species of moths (e.g., *Spodoptera* (Guenée), or *Manduca* (Hübner)) or butterflies (e.g., *Pieris* (Schrank)) [24,25]. We also included respective gene sequences from two Microgastrine species (*Microplitis demolitor* (Wilkinson), and *Microplitis mediator* (Haliday)), for which full genomes are publicly available from the NCBI database (Table 2), to use as outgroup in the phylogenetic analyses described below. We used BLAST search against the respective genomes to extract the sequences of the genes used in the phylogeny [26]. All NCBI references can be found in Table 2.

Depending on availability, we selected one to ten (on average five) parasitoid specimens per caterpillar species, and per locality sampled. As expected, the cryptic nature of the parasitoids, and the fact that some caterpillar species can host multiple species of parasitoids makes their identification difficult using only morphology. The species identity of the parasitoids collected in the early 2000's was previously confirmed through morphology, *COI*, and host associations

Table 1. Sample summary.

Cotesia species	Butterfly host species	# of Ind.	Identified host plant	Country	Locality	Collector	Collection year
<i>C. acuminata</i> Sp. A	<i>Melitaea latonigena</i>	3	–	Russia	Siberia	N.W. I.H.	1998
<i>C. acuminata</i> Sp. A	<i>Melitaea didyma</i>	5	<i>Plantago lanceolata</i>	Spain	Coll d'Estenalles	F.V.	2023
<i>C. acuminata</i> Sp. A	<i>Melitaea didyma</i>	5	–	Switzerland	Conthey	Y.C.	2023
<i>C. acuminata</i> Sp. B	<i>Melitaea scotosia</i>	3	–	China	Peking	I.H., L.G.	1999
<i>C. acuminata</i> Sp. B	<i>Melitaea phoebe</i>	5	–	Italy	Aosta	Y.C.	2023
<i>C. acuminata</i> Sp. K	<i>Euphydryas maturna</i>	1	–	France,	Cote d'Or, Malloy,	P.J.C.R.	1999
<i>C. acuminata</i> Sp. K	<i>Euphydryas maturna</i>	3	–	Sweden	Lindesberg	C.U.E.	1999
<i>C. acuminata</i> Sp. L	<i>Melitaea athalia</i>	1	–	France	Var, St. Paul-en-Forêt	P.W.C.	1982
<i>C. bignellii</i>	<i>Euphydryas aurinia</i>	5	–	Switzerland	La Vraconnaz	Y.C.	2023
<i>C. cynthiae</i>	<i>Euphydryas cynthia</i>	4	–	France	Alps, Laus de Cervieres	M.S.	2001
<i>C. euphydryidis</i>	<i>Euphydryas phaeton</i>	7	–	USA	Warren Co., Front Royal	N.S.	1979
<i>C. koebelei</i>	<i>Euphydryas editha</i>	7	–	USA	Yosemite National Park	M.S., B.W.	2001
<i>C. melitaeorum</i> Sp. D	<i>Euphydryas desfontainii</i>	5	<i>Cephalaria leucantha</i>	Spain	Fígols-El Forn	CS	2023
<i>C. melitaeorum</i> Sp. D	<i>Euphydryas aurinia</i>	5	<i>Lonicera implexa</i>	Spain	Cornudella de Montsant	C.S.	2023
<i>C. melitaeorum</i> Sp. D	<i>Euphydryas aurinia</i>	5	<i>Lonicera implexa</i>	Spain	La Mussara	C.S.	2023
<i>C. melitaeorum</i> Sp. E	<i>Melitaea deione</i>	5	<i>Antirrhinum majus</i>	Spain	Les Illes	CS	2023
<i>C. melitaeorum</i> Sp. E	<i>Melitaea deione</i>	5	<i>Antirrhinum majus</i>	Spain	Poboleda	CS	2023
<i>C. melitaeorum</i> Sp. F	<i>Melitaea didyma</i>	3	–	Hungary	Örseg	M.R.S.	2001
<i>C. melitaeorum</i> Sp. F	<i>Melitaea didyma</i>	2	–	Spain	Cantallops	C.S.	2002
<i>C. melitaeorum</i> Sp. G	<i>Melitaea trivia</i>	10	<i>Verbascum pulverulentum</i>	Spain	Pyrenees	C.S.	2023
<i>C. melitaeorum</i> Sp. H	<i>Melitaea cinxia</i>	8	<i>Plantago lanceolata</i>	Finland	Åland	S.I.	2022
<i>C. melitaeorum</i> Sp. I	<i>Melitaea athalia</i>	5	–	Finland	Åland	L.G., S.vN.	1997
<i>C. melitaeorum</i> Sp. M	<i>Euphydryas aurinia davidi</i>	3	–	Russia	Siberia	N.W., I.H.	1998
<i>C. melitaeorum</i> Sp. N	<i>Melitaea parthenoides</i>	5	<i>Plantago media</i>	Spain	El Brull	C.S.	2023

Cotesia identity, host species, number of individuals, caterpillar host plant, country, locality, collector ID, and finally collection year. “–” indicates the information was unavailable.

Collectors. Y.C., Yannick Chitarro; P.W.C., Peter W. Cribb; C.U.E., Clae U. Eliasson; L.G., Lei Guanchung; I.H., Ilkka Hanski; S.I., Suvi Ikonen; P.J.C.R., Peter J. C. Russel; M.R.S., Mark R. Shaw; M.S., Michael Singer; N.S., Nancy Stamp; C.S., Constanti Stefansescu; S. vN., Saskya van Nouhys; N.W. Niklas Wahlberg; B.W., Brian Wee; F.V. Federica Valerio.

<https://doi.org/10.1371/journal.pone.0348973.t001>

[19]. For the samples collected in 2022 and 2023, species identities were assigned after blasting their respective *COI* sequences against the NCBI database, and by considering their host associations, following previously established protocol [19].

DNA extraction

For the samples collected in the early 2000's the DNA was extracted from whole insects using the NucleoSpin Tissue Kit (Macherey-Nagel) following the manufacturers protocol and eluted in 50 µL of MilliQ water (18). The samples have since been stored in the freezer (−20°C) at the University of Helsinki, Finland. For the samples collected in 2022 and 2023, the DNA was extracted destructively from each whole insects, except in two cases where we pooled damaged adult wasps that emerged from the same host caterpillar (*M. trivia*). The Macherey-Nagel NucleoSpin Tissue Kit (Düren, Germany) was used following the manufacturer's protocol. The DNA was eluted twice in 40 µL of elution buffer to increase final yield and concentration.

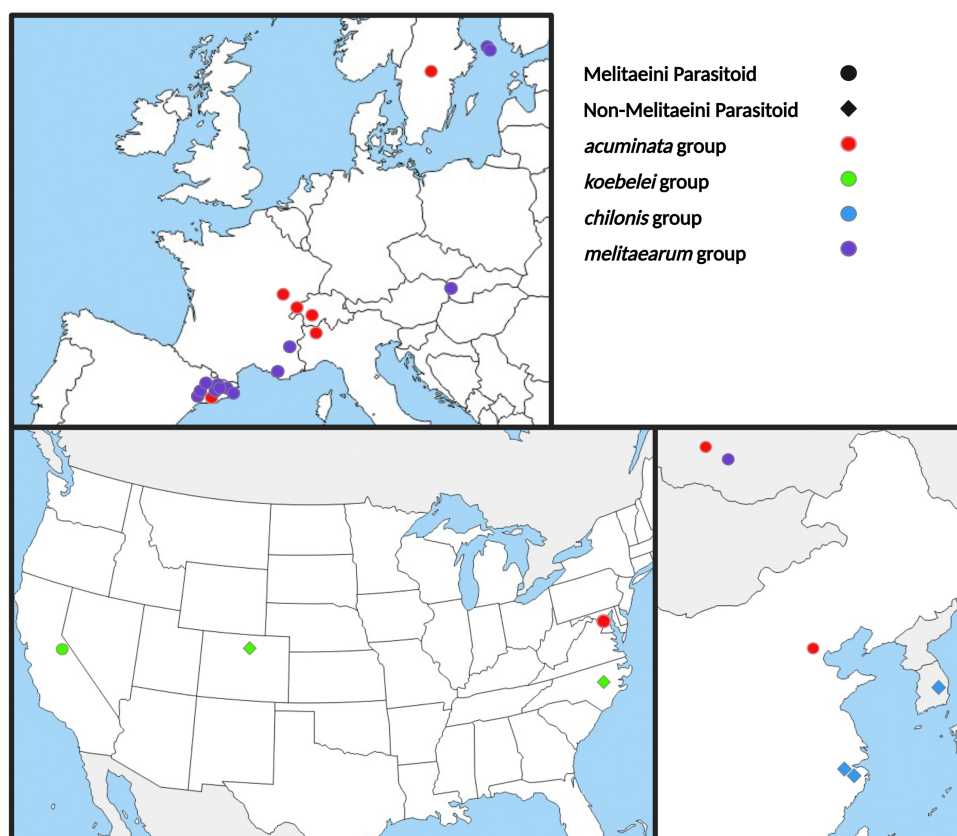


Fig 1. Sample locations. Sampling locations of *Cotesia* specimens used in this study. Diamonds represent species that do not parasitize Melitaeini, and circles represent species that do. Colors correspond to the different clades. Each point represents a sampling location. The African *C. typhae* sample is not included. Modified from maps freely available here: Europe (https://www.d-maps.com/carte.php?num_car=2232&lang=en), China (https://www.d-maps.com/carte.php?num_car=168&lang=en), USA (https://www.d-maps.com/carte.php?num_car=1682&lang=en). Created in BioRender. Schach, J. (2026) <https://BioRender.com/futh3ay>.

<https://doi.org/10.1371/journal.pone.0348973.g001>

Table 2. Genbank reference genomes of additional species.

Cotesia Species	Host	GenBank accession #	Country	Reference
<i>Cotesia chilonis</i>	<i>Chilo suppressalis</i>	GCA_018835575.1	China	[27]
<i>Cotesia congregata</i>	<i>Manduca sexta</i>	GCA_905319865.3	USA	[28]
<i>Cotesia glomerata</i>	—	GCA_020080835.1	USA	[29]
<i>Cotesia ruficrus</i>	<i>Mythimna separata</i>	GCA_048537615.1	China	—
<i>Cotesia typhae</i>	<i>Sesamia nonagrioides</i>	GCA_013202065.2	Kenya	[30]
<i>Cotesia vestalis</i>	<i>Plutella xylostella</i>	GCA_000956155.1	South Korea	—
<i>Microplitis demolitor</i>	<i>Pseudoplusia includens</i>	GCA_026212275.2	USA	[31]
<i>Microplitis mediator</i>	—	GCA_029852145.1	USA	—

Species, host species, GenBank reference genome accession number, sampling location and reference for six species added to the *Cotesia* ingroup and *Microplitis* outgroup of the phylogenetic tree. "—" indicates the information was unavailable.

<https://doi.org/10.1371/journal.pone.0348973.t002>

Gene selection

Genetic markers and primer pairs designed and tested by previous studies, were selected by reviewing the available literature on *Cotesia* and Braconidae phylogenies [32–38] (Table 3). All primer pairs were tested for specificity to *Cotesia* using Primer-Blast [39]. Additionally, 16 BUSCO genes were randomly selected from the *C. congregata* genome (GCA_905319865.3) using BUSCO v 5.7.1 [40]. Using Primer-Blast, we designed an additional set of primer pairs to amplify these BUSCO genes. The best pairs were selected based on the lowest self-complimentary score given by Primer-Blast. In total, two mtDNA markers and 23 nDNA markers were tested, but only two and eight, respectively, were successfully used in this study, as 13 other nDNA primer pairs designed failed to amplify through gradient PCR (Table E in S1 Data). All primer pairs, from the literature and designed for this study, targeted genetic regions with a length ranging from 150 to 850 bp (Table 3).

PCR

Each 10 µL PCR reaction included 1 µL of Buffer solution, 0.8 µL of DNTPs (25nM) mix, 0.5 µL of forward primers (10nM), 0.5 µL of reverse primers (10nM), 0.5 µL Mg2 reagent, 0.05 µL KAPA Taq polymerase, 6.7 µL of H₂O (sterile), and 1 µL of DNA extract. The DNA extracts ranged from 90–470 ng/µL and were not normalized. The PCR reactions were initiated by incubation at 95°C for 3 min, followed by 30 cycles of 95°C for 40 s, the optimal annealing temperature of the respective primer pair (Table 2) for 40 s, and 72°C for 40 s. The PCR reaction ended with a final elongation step at 72°C for 5 min, and a final hold at 18°C.

Table 3. Primer pairs.

Gene	Primer name	Sequence (5'-3')	Amplicon size	Ta (°C)	Primer source
<i>COI</i>	LCO2198	5'-GGTCAACAATCATAAAGATATTGG	658	49	[32]
	HCO1490	5'-TAACTTCAGGGTGACCAAAAAATCA			
<i>16s</i>	16SWb	5'-CACCTGTTTATCAAAAACAT	375	58	[33,34]
	16S outer	5'-CTTATTCAACATCGAGGTC			
<i>18s</i>	18s - F	5'-CATTGGAGGGCAAGTCTGGTGCCA	240	58	[35]
	18s - R	5'-AGTAAACGTACCGGCCCTCCTCG			
<i>n28s rDNA</i>	28s	5'-AAGAGAGAGTTCAAGAGTACGTG	618	56	[36]
	28s-PM	5'-TAGTTCACCATCTTTCGGGTCCC			
<i>EF1a</i>	EF1A1 - F	5'-AGATGGGYAARGTTCTTCAA	418	56	[37]
	EF1A1 - R	5'-AACATGTTGTCDCCGTGCCATCC			
<i>ITS2</i>	NG02955	5'-ATGAACATCGACATTTGCAACGCACA	153-341	60	[38]
	AB052895	5'-TTCTTTTCTCCGCTTAGTAATATGCTTAA			
<i>LW-Rh</i>	OpsFor2	5'-GGATGTASCTCCATTTGGTC	432-712	58	[41]
	Ops3'2	5'-AVHGATGCRACRTTCATTTTCT			
<i>Elob</i>	TCEB - F	5'-TCTTAAAAATTCTCCAACGCATC	250	57	This study
	TCEB - R	5'-GATGCCTGCTCTTGTCATT			
<i>SLD5</i>	SLD5 - F	5'-CCTAAGAACGCGTTGGAGA	327	57	This study
	SLD5 - R	5'-TGAACAATATGCTGCGAGCC			
<i>Innexin</i>	INX-F	5'-CGCTCTTTGTGTGCTTGCTT	342	54	This study
	INX-R	5'-GGTACAGAAGTCGGTGAGGC			

Characteristics of the primer pairs used in this study. Including the name of the marker targeted (gene), primer name, primer sequence, amplicon size, annealing temperature, and source.

<https://doi.org/10.1371/journal.pone.0348973.t003>

The PCR products were run through a 1% agarose gel electrophoresis for 30 minutes at 100 volts, with SYBR-safe stain (Thermo Scientific, USA), and visualized under UV light. For each gel, 5 μ L of 100 bp DNA ladder (#SM0321, Thermo Scientific, USA) was loaded into the first well, and each subsequent well was loaded with a mixture of 4 μ L of both loading dye (bromophenol blue) and PCR product. The PCR amplification was repeated once for any failed sample, to identify potential false negative results.

Preparation for sanger sequencing

The successfully amplified PCR products were individually cleaned using Exonuclease I (ExoI) to remove any single-stranded DNA (Thermo Scientific, USA). For each reaction 0.5 μ L of ExoI was added to 5 μ L of PCR product, then the samples were incubated at 37°C for 15 minutes, followed by 15 minutes at 85°C to terminate the reaction. Cleaned PCR products were sent to Macrogen Europe for Sanger Sequencing using the company Multiple Primer/Plate option.

Sequence cleaning and alignments

The chromatograms of each sequence were manually quality-checked by using Geneious Prime 2025.01 (<https://www.geneious.com>). Low-quality sequences (error probability limit 0.05) and highly incomplete sequences (less than half of target length) were removed from further analysis.

We included the respective gene sequences from two Microgastrine species (*Microplitis demolitor* and *M. mediator*), as outgroup in the phylogenies described below. Additionally, six species of *Cotesia* that do not parasitize Melitaeini were added to the ingroup (Table 2). For each marker, all sequences were aligned with MUSCLE using default settings [42], and all alignments were checked by eye with AliView V1.28 [43]. The accession numbers for the sequences submitted to NCBI are PX523132-PX523219, PX559195-PX559573, PX642706-PX642796, and PX648620-PX649011 (S1 File).

Cotesia phylogeny

For the phylogenetic analyses we used maximum likelihood (ML) with IQ-TREE V2.4.0 [44]. Phylogenetic analyses were conducted for: (I) all ten genes concatenated (mtDNA and nDNA), (II) eight concatenated nuclear genes, (III) two concatenated mitochondrial genes, and (IV) all genes independently. Specimens that amplified for less than half of the markers were excluded from concatenated analysis but kept for the individual gene tree analyses. Phylogenetic trees were visualized using iTOLv7.5.1 [45].

For ML analysis, we used ModelFinder [46], and the Bayesian Information Criterion (BIC) to select for the best models. The partition finding algorithm [47] was run for the concatenated data (ten-genes, nDNA, mtDNA) partitioned by genes using the MFP+MERGE option. Support for nodes was evaluated with 1000 nonparametric bootstrap replicates, where we considered values under 60 to represent unsupported nodes. Additionally, gene concordance factors (gCF) were calculated in IQ-TREE as the proportion of gene trees supporting each bipartition in the concatenated tree [48–50].

Coalescent-based inference

To account for potential incomplete lineage sorting (ILS) we performed a quartet-based analysis in ASTRAL-III (version 5.7.8) [51], which estimates a species tree from unrooted gene trees. The input consisted of the gene trees produced by IQtree2 (see above) and a species map to account for the multiple individuals per species. We ran the ASTRAL analysis with default settings. Branch lengths were estimated in coalescent units, where shorter internal branch lengths indicated higher levels of ILS. Node support was assessed using local posterior probabilities based on quartet frequencies.

Species delimitation

To test presence of cryptic species, we tested two methods of species delimitation. First, we used a coalescent based species delimitation with our sequence data using Bayesian Phylogenetics and Phylogeography model (BPP version

4.8.7) [52]. Then, we used a Bayesian implementation of a Poisson Tree Processes model (bPTP), which performs species delimitation on a rooted phylogenetic tree [53].

For the BPP analysis, we performed species delimitation on a guide tree (analysis A10) which uses a reversible-jump MCMC (rjMCMC) algorithm to move between species delimitation models that are compatible with a fixed tree. BPP requires prior distributions for ancestral population (θ) and the age of the root (τ). Because these priors have not yet been estimated for our system, we used four combinations of θ and τ to encompass a variety of demographic scenarios ranging from small populations with old divergence to large populations with recent divergence [54,55]. Specifically, we used the following combinations: $\theta = G(2, 2000)$ and $\tau = G(2, 200)$, $\theta = G(2, 2000)$ and $\tau = G(2, 1000)$, $\theta = G(2, 200)$ and $\tau = G(2, 200)$, and $\theta = G(2, 200)$ and $\tau = G(2, 2000)$. We used both the ASTRAL species tree and the ten-gene concatenated phylogeny as the guide tree to test different evolutionary hypotheses. We ran the MCMC chain for 200k generations with sampling every five generations and a burn-in of 50k generations. For the bPTP analysis, we used the bPTP web server to model species delimitation for each of our concatenated ML trees, conserving the default parameters, with outgroups removed and 500k generations.

Results

Parasitoid identification

The N=42 specimens collected from 1979 to 2002 include 10 species of *Cotesia*: four cryptic species of *C. acuminata* agg., *C. cynthiae* (Nixon), *C. euphydryidis*, *C. koebelei* (Riley), and three cryptic species of *C. melitaearum* agg., either as previously identified by earlier studies [18,19] or by comparing their COI sequences to NCBI. The sequencing results of the 2022 and 2023 collections confirmed that these N=68 specimens include eight species of *Cotesia*: *C. bignellii* (Marshall), two cryptic species of *C. acuminata* agg., and five cryptic species of *C. melitaearum* agg. (Table 1).

Sequencing results

Of the 110 specimens used in this study, 41% were successfully sequenced for all 10 genes, 66% for at least 9 genes, 75% for 8 genes, 85% for 7 genes, 88% for 6 genes, and 90% for 5 genes.

Cotesia phylogenies

(I) Ten-gene dataset. The ten-gene dataset included 110 wasp specimens (102 Melitaeini associated, and eight from NCBI). The ten-gene alignment length was 3486 bp long, of which 1498 sites (43%) represented unique site-patterns. The model partitioning algorithm revealed five partitions in the ten-gene dataset (Table A in S1 Data).

The ML analysis revealed four major clades with moderate to high bootstrap support (97, 100, 70, 97), but low to moderate gCF (50, 55, 0, 30, respectively), indicating high conflict between gene trees and the concatenated tree. The phylogeny also showed that the species attacking Melitaeini butterflies are polyphyletic (Fig 2). The node containing the *acuminata*, *vestalis*, and *melitaearum* groups was poorly supported (bootstrap < 60, gCF = 0).

- **The *koebelei* group** contained *C. koebelei*, *C. congregata* and *C. glomerata* parasitoid wasps. *Cotesia congregata* and *C. glomerata* do not use Melitaeini butterflies as hosts, but *C. koebelei* wasps parasitize the Melitaeini species *E. editha*. The nodes in this clade were highly supported (bootstrap = 91 and above)
- **The *acuminata* group** was formed by the *C. acuminata* complex, *C. bignellii*, and *C. euphydryidis* parasitoids from Melitaeini butterflies. *Cotesia euphydryidis* was placed at the base of the clade. The next division was between *C. bignellii* and *C. acuminata* agg. The clade revealed *C. acuminata* agg. as polyphyletic, where *Sp. B* and *K* cluster together and are sister taxa to *C. bignellii*. The divergence between *C. acuminata* *Sp. A* and *C. bignellii*, *C. acuminata* *Sp. B* and *K* was well supported (bootstrap = 91) but the relationship between *C. bignellii*, *C. acuminata* *Sp. B* and *K* was not.

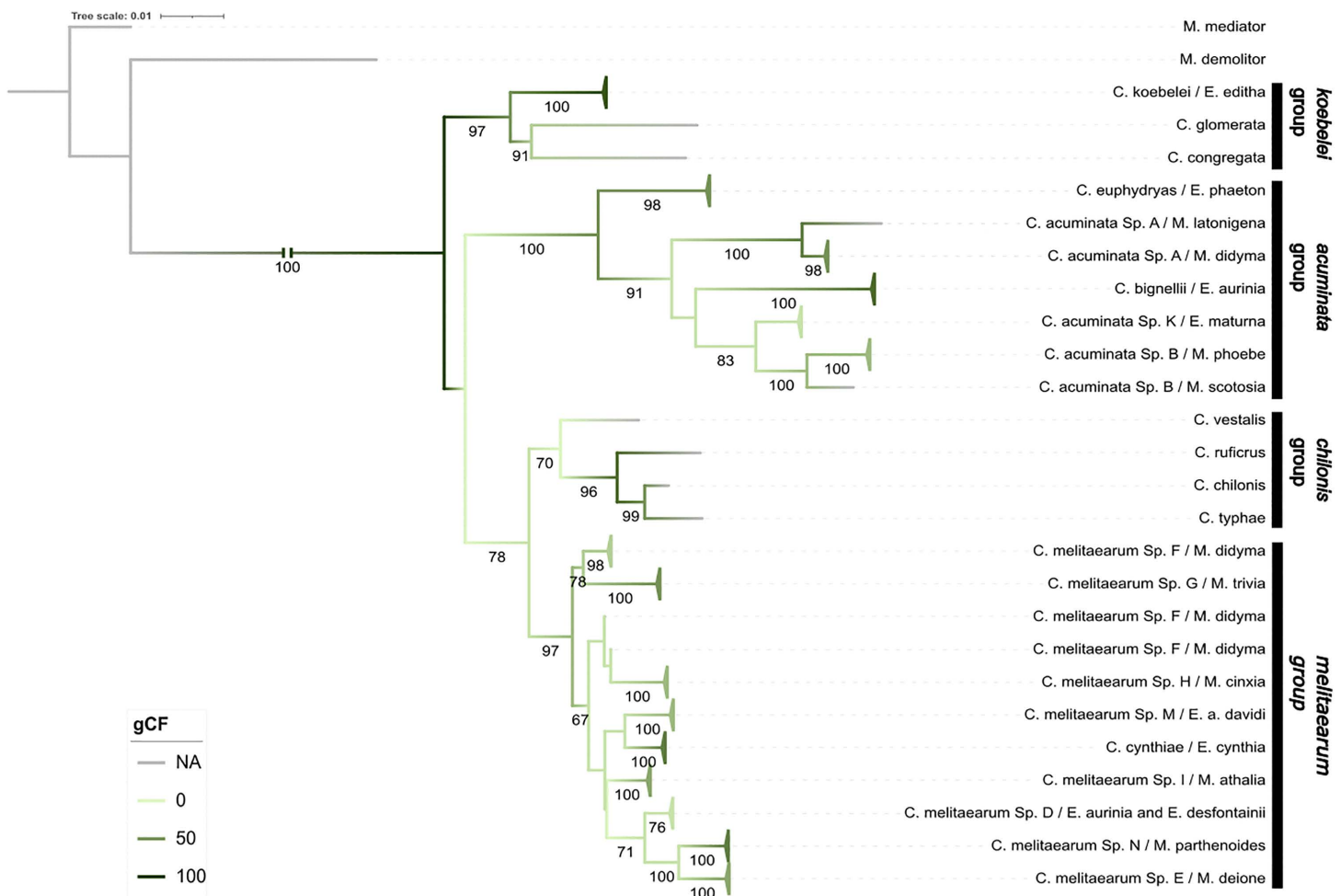


Fig 2. Ten-gene maximum likelihood phylogeny. Maximum likelihood phylogeny of *Cotesia* species parasitizing different Melitaeini host species and their relatives based on the ten concatenated mitochondrial and nuclear genes (16s, 18s, 28s, COI, EF1a, INX, ITS2, LW-Rh, SLD5, and Elob). The specimens are labelled by the names of the *Cotesia* species and Melitaeini host caterpillar species. Bootstrap support values (1000 replicates) are indicated for supported branches (≥ 60). The branches are colored according to gene concordance factor (gCF). Branches corresponding to the same species were collapsed into triangles for visual clarity.

<https://doi.org/10.1371/journal.pone.0348973.g002>

- **The *chilonis* group** contained *C. chilonis*, *C. ruficrus*, *C. typhae*, and *C. vestalis* parasitoids, which all parasitize non-Melitaeini hosts. The relationships in this clade were well supported.
- **The *melitaeorum* group** contained *C. melitaeorum* agg. and *C. cynthiae*, which parasitize Melitaeini butterflies. Ten subclades were revealed, represented by the individual cryptic species and *C. cynthiae* with varying support (bootstrap 55–100). Here, *C. melitaeorum* agg. formed a paraphyletic group. Most of the individual species within this clade grouped together as monophyletic, except *C. melitaeorum* Sp. F. For this particular species, our phylogeny placed the specimens from Spain in a sister group to *C. melitaeorum* Sp. G also from Spain and the specimens from Hungary were paraphyletic and sister to the *C. melitaeorum* Sp. H from Finland. Additionally, *C. melitaeorum* Sp. D was grouped together and not clustered by host species.

(II) Nuclear-gene dataset. The nuclear-gene dataset included 112 wasp specimens (104 Melitaeini-associated, and eight from NCBI). The concatenated alignment was 2413 bp long, of which 1035 sites (43%) represented unique site-patterns. The model partitioning algorithm revealed two partitions in the nuclear-gene dataset (Table B in [S1 Data](#)).

The ML analysis produced a moderately supported phylogeny that recovered the same four clades as the ten-gene dataset with varying bootstrap support (100, 96, 51, 84) and low gCF (57, 50, 0, 12.5). The placements of the *acuminata* and *koebelei* groups were switched when compared to the ten-gene concatenated tree, however the node containing the *koebelei*, *chilonis*, and *melitaeum* groups was poorly supported (bootstrap < 60, gCF = 12.5) ([Fig 3](#)). Additionally, there were some differences within the *acuminata* and *melitaeum* groups between the two trees.

- **The *acuminata* group** recovered *C. acuminata* agg. as monophyletic, although poorly supported (bootstrap = 39).

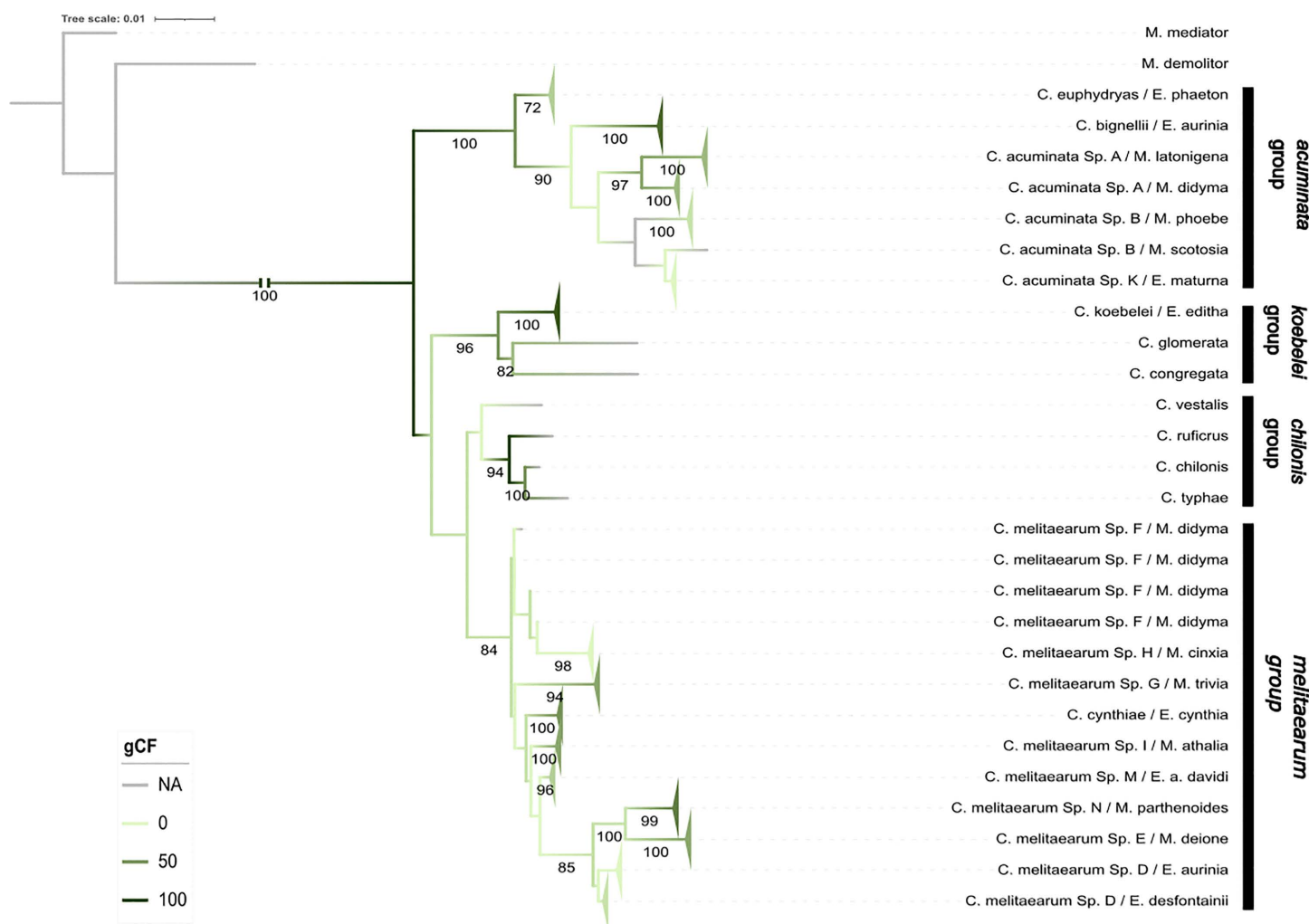


Fig 3. Nuclear maximum likelihood phylogeny. Maximum likelihood phylogeny of *Cotesia* species parasitizing different Melitaeini host species and their relatives based on the nuclear genes (*18s*, *28s*, *EF1a*, *INX*, *ITS2*, *LW-Rh*, *SLD5*, and *Elob*). The specimens are labelled by the names of the *Cotesia* species and Melitaeini host caterpillar species. Branches corresponding to the same species were collapsed for visual clarity and represented by triangles. Bootstrap support values (1000 replicates) are indicated for supported branches (≥ 60). The branches are colored according to gene concordance factor (gCF). Branches corresponding to the same species were collapsed into triangles for visual clarity.

<https://doi.org/10.1371/journal.pone.0348973.g003>

- The **melitaeorum group** showed some discordance from the ten-gene phylogeny. For example, *C. melitaeorum Sp. F* was grouped together with *C. melitaeorum Sp. H* however as paraphyletic, thus resulting in nine subclades. Many of which were well supported, but the relationships between them were not and have low gCF values. Additionally, *C. cynthiae* and *C. melitaeorum Sp. M* were no longer sister groups.

(III) Mitochondrial-gene dataset. The mitochondrial-gene dataset included the partial sequences of the 16s and COI genes from 92 specimens (86 Melitaeini associated, six from NCBI). The mitochondrial alignment was 1072 bp long, of which 451 (42%) represented unique site-patterns. The model partitioning algorithm revealed two partitions in the mitochondrial gene dataset, which correspond to the two markers (Table C in [S1 Data](#)). No specimens from *C. euphydryidis* and *C. melitaeorum Sp. F* from Hungary amplified for either mitochondrial marker, and no sequences were found from *C. typhae* and *C. vestalis* genomic projects, thus they were missing from the mitochondrial phylogenetic analysis. The ML analysis revealed the same four major clades as the ten-gene and nuclear gene phylogenies, however with varying bootstrap support (62, 91, 98, 58) and gCF (50, 100, 50, 100) ([Fig 4](#)). The relationships within the *koebelei* and *chilonis* groups were congruent to the ten-gene and nuclear datasets, however, two species were missing from the *chilonis* group.

- The **acuminata group** recovered the *C. acuminata* complex as monophyletic, similar to the nuclear dataset, however the support was low (bootstrap < 60).
- The **melitaeorum group** formed eight well supported subgroups corresponding to the different cryptic species (bootstrap 76–99) and maintained the paraphyletic nature of *C. melitaeorum* agg. however, many of the relationships between clades were unsupported (bootstrap < 60).

(IV) Individual genes. All genes were also analysed individually using ML to estimate gene concordance factors, and to build a coalescent based species tree. The substitution models for the ML analyses are listed in Table D in [S1 Data](#). Rather expectedly, the individual gene phylogenies had much lower branch supports than either of the other phylogenies, likely because of the small nucleotide sample size included in each analysis. Alone, the individual gene trees are not phylogenetically informative. The major clades were only recovered in the 16s, COI, and SLD5 ML trees, but the relationships within the clades, particularly within the *melitaeorum* group, remain incongruent ([S7–S16 Figs](#)).

Coalescent-based inference

Individual gene trees were also analysed together under the multispecies coalescent using ASTRAL-III. While the ASTRAL analysis recovered the same major clades as the concatenated ML phylogenies, the relationships within the clades were more often weakly supported and topologically incongruent from the concatenated analyses. The coalescent analysis placed the *acuminata* group as the ancestral clade and did not propose the *chilonis* group as monophyletic, as *C. vestalis* diverges from the split between the *chilonis* and *melitaeorum* groups. These differences might be caused by the lower amount of sequence data for *C. vestalis* (only five loci were available for that particular species). Overall, internal nodes within the coalescent tree exhibited low support (LPP < 0.7) with short branches, indicating substantial gene tree discordance ([S1 Fig](#)).

Species delimitation

The Bayesian species delimitation provided strong support for the *a priori* defined species, and the results were robust to the guide tree, and priors on population size (θ) and root age (τ) ([S2](#) and [S3 Fig](#)). Each of our cryptic species had high posterior support (1.00), however we observed some over-splitting, for example, *C. acuminata Sp. B* and *C. melitaeorum Sp. D* from different host species/populations were recovered as independent species. The lowest supported split was between *C. chilonis* and *C. typhae*, however they were still maintained as separate species. The bPTP analysis of the

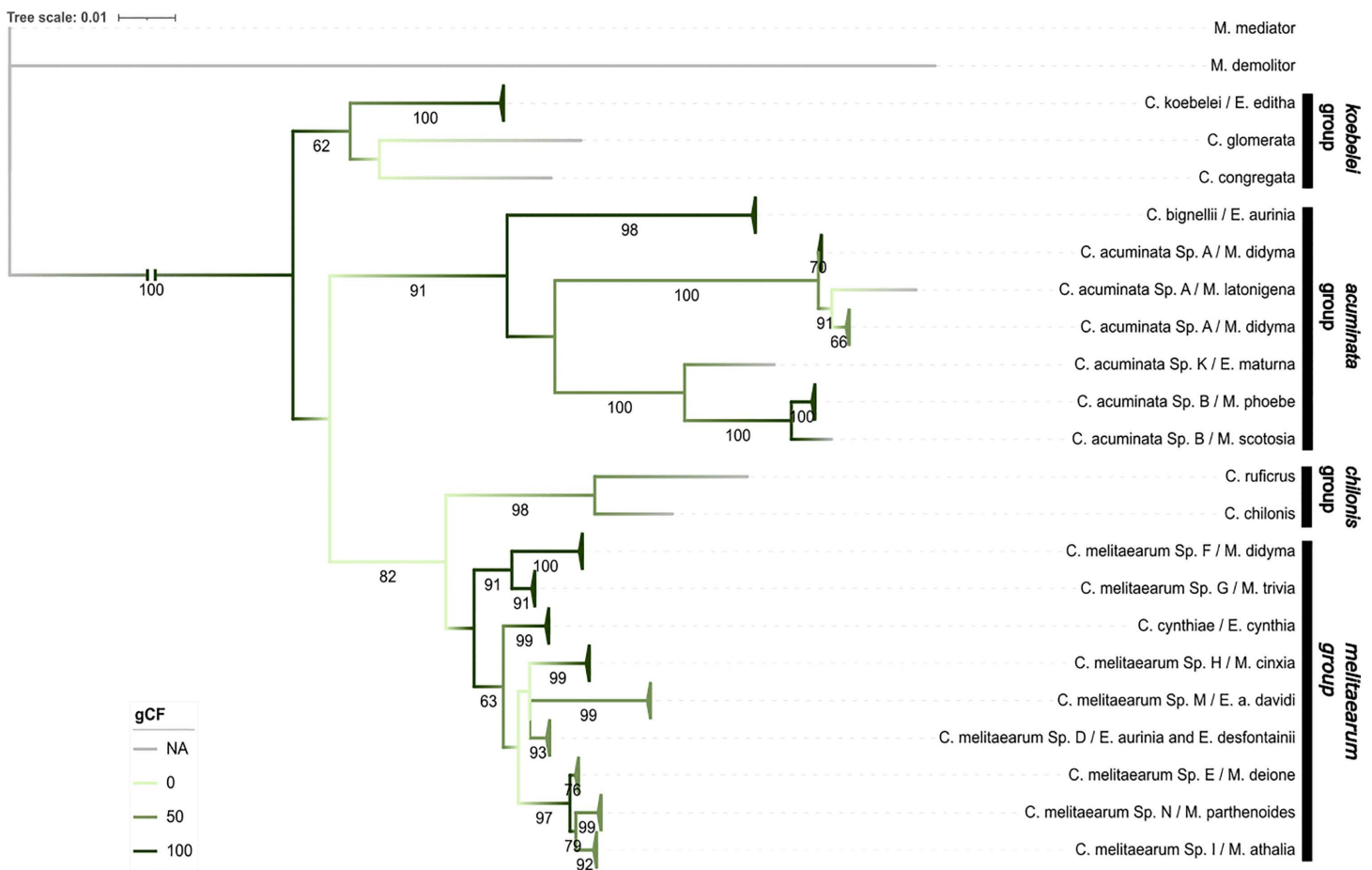


Fig 4. Mitochondrial maximum likelihood phylogeny. Maximum likelihood phylogeny of *Cotesia* species parasitizing different Melitaeini host species and their relatives based on the mitochondrial genes (*16s* and *COI*). The specimens are labelled by the names of the *Cotesia* species and Melitaeini host caterpillar species. Branches corresponding to the same species were collapsed for visual clarity and represented by triangles. Bootstrap support values (1000 replicates) are indicated for supported branches (≥ 60). The branches are colored according to gene concordance factor (gCF). Branches corresponding to the same species were collapsed and represented by triangles for visual clarity.

<https://doi.org/10.1371/journal.pone.0348973.g004>

concatenated phylogenies supported the presence of cryptic species and distinct clades (S4–S6 Figs). Although the bPTP never lumped together cryptic species into a single species, this analysis suggested more over-splitting of cryptic species, than previous analyses. For example, the three *C. melitaeum* Sp. I individuals were categorized as separate species in the bPTP analysis.

Discussion

Resolving species phylogenies is essential for studying trait evolution, diversification, and ecological interactions, but can remain challenging in non-model cryptic species, due to limited access to sampling and molecular resources. By designing new primers and increasing the number of available molecular markers from two to ten, we were able to produce the most robust phylogeny of the *Cotesia* species parasitizing Melitaeini butterflies to date. This improved phylogeny informs on previously hidden species diversity within this genus and is an important resource for future research on Braconidae population genetics, evolutionary dynamics and ecology.

Our analyses revealed that all specimens group in four highly supported clades (*i.e.*, *koebelei* group, *acuminata* group, *chilonis* group, and *melitaeorum* group), each subdivided into distinct subclades of (cryptic) species. While our results showed strong support for the species groups, the relationships between them are still unresolved, similar to other *Cotesia* multi-gene phylogenies [19,56]. By including genetic sequences from *Cotesia* species that do not attack Melitaeini butterflies, we confirm that the Melitaeini-specialized species are polyphyletic, appearing in three of the four main *Cotesia* clades. This suggests that *Cotesia* wasps may have shifted hosts multiple times throughout their evolutionary history. Such patterns of host shifts have been observed in other parasitoid lineages, including Ichneumonoidea [57], and parasitoids of leaf miners [58], indicating that parasitoid evolution often involves ecological switching rather than strict co-divergence with hosts.

The presence of cryptic species was confirmed using two methods of species delimitation (*i.e.*, BPP and bPTP analyses). The BPP model strongly supported the predefined species taxonomy, with the lowest supported split being between *C. chilonis* and *C. typhae* that do not attack Melitaeini butterflies. However, this may be due to the presence of only one individual for these species, making it impossible to estimate within species population parameters [52]. The coalescent based delimitation tended to over-split especially when a single species was differentiated by their host species, likely capturing population structure as distinct species, and which can be exacerbated by the presence of speciation with gene flow [59]. Similarly, the bPTP over-split the species taxonomy, which is maybe not surprising as this method relies entirely on branch length to distinguish between intra- and interspecific variation, and tends to over-split species particularly when there is geneflow between species and deep ILS [60], which is likely the case for these rapidly diverging cryptic species. Additionally, missing data in the concatenated alignments could have distorted branch lengths furthering over-splitting in bPTP [61].

Many of the Melitaeini-associated *Cotesia* species occur in sympatry, yet the drivers of radiation and barriers to gene flow remain unclear [18,22]. The butterfly host evolution could, at least partially, affect the divergence between close *Cotesia* lineages and cryptic species. Indeed, while most *C. acuminata* species parasitize *Melitaea* butterflies, the ancestral species parasitize *Euphydryas* caterpillars, which have diverged from *Melitaea* 42.68 MYA [62]. Similarly, in the *koebelei* group, the *C. koebelei* species parasitizes *E. editha* and has also been tentatively reported on *Melitaea* [63], this may suggest the parasitoid is a rare generalist able to attack both butterfly genera. Alternatively, if a preference would be detected between the *C. koebelei* lineages parasitizing *Euphydryas* versus *Melitaeini* butterflies, the host difference could simply indicate the cryptic diversity in this species complex. The question of which of these two hypotheses is the most plausible could potentially be resolved through future intensive sampling of North American Melitaeini butterflies and their *Cotesia* parasitoids.

In contrast, the *melitaeorum* group shows an opposite pattern to the *acuminata* group: with the ancestral species parasitizing *Melitaea*, followed by a shift to *Euphydryas*. This may reflect a retained ability to use *Euphydryas*, or recolonization event of *Euphydryas*, the possible host of the ancestor shared between the *melitaeorum* group and the *acuminata* group. In contrast, species from the *chilonis* group are not known to use Melitaeini butterflies but span a wider range of Lepidoptera host taxa. For example, *C. vestalis* parasitizes *Plutella xylostella*, (Linnaeus) (Yponomeutoidea), a lineage near the base of the Lepidopteran tree [64], while more recently diverged *Cotesia* in this clade parasitize more recently diverged groups such as Pyraloidea, (Latreille) and Noctuidae, (Latreille).

Although the relationships within and between cryptic species are better supported in our phylogeny compared to previous studies [19], comparisons of the tree topologies, concordance factors, and ASTRAL suggested conflicts between nuclear and mitochondrial reconstructions and gene trees. The discordance within clades may reflect biological processes such as geneflow between lineages or rapid radiation resulting in incomplete lineage sorting [65,66]. The non-random absence of mitochondrial data for several taxa in two clades can bias phylogenetic inference, especially when relatively few markers are used [61], and could partially affect the differences observed between datasets used here. Furthermore, the ten loci and missing data may not provide enough signal for a comprehensive ASTRAL analysis, resulting in the poorly supported species ASTRAL tree, which exhibit very short branches.

While we present an improved *Cotesia* phylogeny, future work may require additional genetic data, broader geographic sampling, and inclusion of other species to fully evaluate the evolutionary history of the genus *Cotesia*. Although whole genome sequencing approaches might provide more detailed information on hybridization and introgression events, or on lineage sorting in the genus *Cotesia*, the integration of new experimental and field data on host fidelity (as in Kankare et al [18]) and/or reproductive isolation would reveal elements of the ecological barriers that have facilitated the speciation of *Cotesia* wasps. Many parasitoid wasp females preferentially oviposit on the same host species from which they emerged, reinforcing such reproductive barriers and potentially promoting speciation [67–69]. Alternatively, the use of large network-based approaches could offer useful frameworks to further investigate host-parasitoid coevolution [70], testing hypotheses such as adaptive radiation [71], or host use oscillations [72]. Our improved *Cotesia* phylogeny offers a great first step towards the evaluation of *Cotesia* species diversity. Combined with a clearer understanding of other aspects of *Cotesia*'s ecology, such as ecological niches and host range breadth, this data will support conservation of its cryptic diversity, as many species with narrow host ranges may face elevated extinction risk. Our study questions the need to revise the taxonomy of *Cotesia* to highlight the cryptic diversity, and the possible introduction of new sub-genera by taxonomists.

Conclusion

Our study provides a robust phylogenetic framework for *Cotesia* parasitizing Melitaeini butterflies, clarifying uncertainties in their evolutionary relationships and cryptic diversity. The results show that the species attacking Melitaeini are paraphyletic, grouping in three of the four major clades, with indications of independent host shifts. Although our sampling does not allow for inference of which lineage first adapted to Melitaeini hosts, the confirmation of distinct cryptic species in *C. acuminata* agg. and *C. melitaeorum* agg. highlights the extent of hidden diversity in this system. Together, these findings establish a foundation for future taxonomic revisions, comparative genomic studies, and exploration of the ecological barriers that initiated and maintain their reproductive isolation.

Supporting information

S1 Data. Tables. Table A. Ten-gene partitioning. Ten-gene dataset partitions determined by partitioning algorithm, and best fit model as determined by ModelFinder using BIC. **Table B. Nuclear partitioning.** Nuclear gene dataset partitions determined by partitioning algorithm, and best fit model determined by ModelFinder using BIC. **Table C. Mitochondrial partitioning.** Mitochondrial-gene dataset partitions determined by partitioning algorithm, and best fit model determined by ModelFinder using BIC. **Table D. Substitution models.** Substitution models used for individual genes, determined by ModelFinder, and selected based on BIC. **Table E. Primer pairs.** Characteristics of the additional primer pairs created for this study. Includes the marker targeted (gene), primer name, primer sequence, amplicon size.

(DOCX)

S1 File. Sequence information. Species ID, sample ID, approximate GPS location, Host ID, and genbank accession number for each sample sequenced in this study. “-” represents the individual was not sequenced for that locus.

(XLSX)

S1 Fig. Coalescent based phylogeny. Species tree inferred using ASTRAL-III relatives based on the ten mitochondrial and nuclear gene trees (16s, 18s, 28s, *COI*, *EF1a*, *INX*, *ITS2*, *LW-Rh*, *SLD5*, and *Elob*). Branch lengths are expressed in coalescent units. Node support values represent local posterior probabilities (LPP).

(TIF)

S2 Fig. BPP ASTRAL guide tree. Bayesian Phylogenetics and Phylogeography (BPP) results using the Coalescent based species tree estimations. Speciation probabilities from BPP for each node are shown in boxes: Top, $\theta = G(2, 2000)$

and $\tau = G(2, 200)$; second, $\theta = G(2, 2000)$ and $\tau = G(2, 1000)$; third $\theta = G(2, 200)$ and $\tau = G(2, 200)$; bottom, $\theta = G(2, 200)$ and $\tau = G(2, 2000)$.

(TIF)

S3 Fig. BPP ML guide tree. Bayesian Phylogenetics and Phylogeography (BPP) results using the ML tree estimations. Speciation probabilities from BPP for each node are shown in boxes: Top, $\theta = G(2, 2000)$ and $\tau = G(2, 200)$; second, $\theta = G(2, 2000)$ and $\tau = G(2, 1000)$; third $\theta = G(2, 200)$ and $\tau = G(2, 200)$; bottom, $\theta = G(2, 200)$ and $\tau = G(2, 2000)$.

(TIF)

S4 Fig. Ten-gene dataset bPTP species delimitation tree. Ten-gene maximum likelihood phylogeny showing bPTP results. Red clades represent putative species.

(TIF)

S5 Fig. Nuclear dataset bPTP species delimitation tree. Nuclear DNA maximum likelihood phylogeny showing bPTP results. Red clades represent putative species.

(TIF)

S6 Fig. Mitochondrial dataset bPTP species delimitation tree. Mitochondrial DNA maximum likelihood phylogeny showing bPTP results. Red clades represent putative species.

(TIF)

S7 Fig. 16s gene tree. Maximum likelihood phylogeny of *Cotesia* species parasitizing different Melitaeini host species and their relatives based on the mitochondrial gene 16s. The specimens are labelled by the names of the *Cotesia* species and Melitaeini host caterpillar species. Bootstrap support values (1000 replicates) are indicated for supported branches (≥ 60).

(TIF)

S8 Fig. 18s gene tree. Maximum likelihood phylogeny of *Cotesia* species parasitizing different Melitaeini host species and their relatives based on the mitochondrial gene 18s. The specimens are labelled by the names of the *Cotesia* species and Melitaeini host caterpillar species. Bootstrap support values (1000 replicates) are indicated for supported branches (≥ 60).

(TIF)

S9 Fig. 28s gene tree. Maximum likelihood phylogeny of *Cotesia* species parasitizing different Melitaeini host species and their relatives based on the mitochondrial gene 28s. The specimens are labelled by the names of the *Cotesia* species and Melitaeini host caterpillar species. Bootstrap support values (1000 replicates) are indicated for supported branches (≥ 60).

(TIF)

S10 Fig. COI gene tree. Maximum likelihood phylogeny of *Cotesia* species parasitizing different Melitaeini host species and their relatives based on the mitochondrial gene COI. The specimens are labelled by the names of the *Cotesia* species and Melitaeini host caterpillar species. Bootstrap support values (1000 replicates) are indicated for supported branches (≥ 60).

(TIF)

S11 Fig. EF1a gene tree. Maximum likelihood phylogeny of *Cotesia* species parasitizing different Melitaeini host species and their relatives based on the mitochondrial gene EF1a. The specimens are labelled by the names of the *Cotesia* species and Melitaeini host caterpillar species. Bootstrap support values (1000 replicates) are indicated for supported branches (≥ 60).

(TIF)

S12 Fig. *INX* gene tree. Maximum likelihood phylogeny of *Cotesia* species parasitizing different Melitaeini host species and their relatives based on the mitochondrial gene *INX*. The specimens are labelled by the names of the *Cotesia* species and Melitaeini host caterpillar species. Bootstrap support values (1000 replicates) are indicated for supported branches (≥ 60).

(TIF)

S13 Fig. *ITS2* gene tree. Maximum likelihood phylogeny of *Cotesia* species parasitizing different Melitaeini host species and their relatives based on the mitochondrial gene *ITS2*. The specimens are labelled by the names of the *Cotesia* species and Melitaeini host caterpillar species. Bootstrap support values (1000 replicates) are indicated for supported branches (≥ 60).

(TIF)

S14 Fig. *LW-Rh* gene tree. Maximum likelihood phylogeny of *Cotesia* species parasitizing different Melitaeini host species and their relatives based on the mitochondrial gene *LW-Rh*. The specimens are labelled by the names of the *Cotesia* species and Melitaeini host caterpillar species. Bootstrap support values (1000 replicates) are indicated for supported branches (≥ 60).

(TIF)

S15 Fig. *SLD5* gene tree. Maximum likelihood phylogeny of *Cotesia* species parasitizing different Melitaeini host species and their relatives based on the mitochondrial gene *SLD5*. The specimens are labelled by the names of the *Cotesia* species and Melitaeini host caterpillar species. Bootstrap support values (1000 replicates) are indicated for supported branches (≥ 60).

(TIF)

S16 Fig. *Elob* gene tree. Maximum likelihood phylogeny of *Cotesia* species parasitizing different Melitaeini host species and their relatives based on the mitochondrial gene *Elob*. The specimens are labelled by the names of the *Cotesia* species and Melitaeini host caterpillar species. Bootstrap support values (1000 replicates) are indicated for supported branches (≥ 60).

(TIF)

Acknowledgments

Thanks to the members of the ISEE lab and life-history evolution group at the University of Helsinki for the many discussions and input. Thank you to Dr. Victoria Twort for her input on the manuscript. Thanks to everyone involved in the sample collection, including Constantí Stefanescu and his team in Spain, Yannick Chittaro in Switzerland, and in Finland Suvi Ikonen and Krista Raveala, along with those involved in the Åland butterfly survey.

Author contributions

Conceptualization: Anne Duploux.

Data curation: Julian Schach.

Formal analysis: Julian Schach.

Funding acquisition: Anne Duploux.

Investigation: Julian Schach.

Methodology: Julian Schach, Federica Valerio.

Project administration: Anne Duploux.

Supervision: Federica Valerio, Anne Duploux.

Writing – original draft: Julian Schach.

Writing – review & editing: Julian Schach, Federica Valerio, Anne Duploup.

References

1. Kjer KM, Simon C, Yavorskaya M, Beutel RG. Progress, pitfalls and parallel universes: a history of insect phylogenetics. *J R Soc Interface*. 2016;13(121):20160363. <https://doi.org/10.1098/rsif.2016.0363> PMID: 27558853
2. Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, et al. Cryptic species as a window on diversity and conservation. *Trends Ecol Evol*. 2007;22(3):148–55. <https://doi.org/10.1016/j.tree.2006.11.004> PMID: 17129636
3. Kon T, Yoshino T, Mukai T, Nishida M. DNA sequences identify numerous cryptic species of the vertebrate: a lesson from the gobioid fish *Schindleria*. *Mol Phylogenet Evol*. 2007;44(1):53–62. <https://doi.org/10.1016/j.ympev.2006.12.007> PMID: 17275344
4. Shin CP, Allmon WD. How we study cryptic species and their biological implications: a case study from marine shelled gastropods. *Ecol Evol*. 2023;13(9):e10360. <https://doi.org/10.1002/ece3.10360> PMID: 37680961
5. Grundt HH, Kjølner S, Borgen L, Rieseberg LH, Brochmann C. High biological species diversity in the arctic flora. *Proc Natl Acad Sci U S A*. 2006;103(4):972–5. <https://doi.org/10.1073/pnas.0510270103> PMID: 16418291
6. Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc Natl Acad Sci U S A*. 2004;101(41):14812–7. <https://doi.org/10.1073/pnas.0406166101> PMID: 15465915
7. Li X, Wiens JJ. Estimating global biodiversity: the role of cryptic insect species. *Syst Biol*. 2023;72(2):391–403. <https://doi.org/10.1093/sysbio/syac069> PMID: 36301156
8. Koh LP, Dunn RR, Sodhi NS, Colwell RK, Proctor HC, Smith VS. Species coextinctions and the biodiversity crisis. *Science*. 2004;305(5690):1632–4. <https://doi.org/10.1126/science.1101101> PMID: 15361627
9. Forbes AA, Bagley RK, Beer MA, Hippee AC, Widmayer HA. Quantifying the unquantifiable: why Hymenoptera, not Coleoptera, is the most speciose animal order. *BMC Ecol*. 2018;18(1):21. <https://doi.org/10.1186/s12898-018-0176-x> PMID: 30001194
10. Whitfield JB, Austin AD, Fernandez-Triana JL. Systematics, biology, and evolution of microgastrine parasitoid wasps. *Ann Rev Entomol*. 2018;63(1):389–406. <https://doi.org/10.1146/annurev-ento-020117-043405>
11. Cornell HV, Hawkins BA. Patterns of parasitoid accumulation on introduced herbivores. In: Hawkins BA, Sheehan W, editors. *Parasitoid Community Ecology*. Oxford University Press; 1994. pp. 77–89. <https://doi.org/10.1093/oso/9780198540588.003.0005>
12. Lei G-C, Hanski I. Metapopulation structure of *Cotesia melitaeorum*, a specialist parasitoid of the butterfly *Melitaea cinxia*. *Oikos*. 1997;78(1):91. <https://doi.org/10.2307/3545804>
13. Brock RE, Cini A, Sumner S. Ecosystem services provided by aculeate wasps. *Biol Rev Camb Philos Soc*. 2021;96(4):1645–75. <https://doi.org/10.1111/brv.12719> PMID: 33913243
14. Li X, Jia J-J, An J-L, Meng F-X, Liu T-X, Zhang S-Z. Effect of *Cotesia ruficrus* parasitization on diversity and community composition of intestinal bacteria in *Spodoptera frugiperda*. *Insects*. 2024;15(8):570. <https://doi.org/10.3390/insects15080570> PMID: 39194775
15. Quicke DLJ. The braconid and ichneumonid parasitoid wasps. 2014.
16. Mason WRM. The polyphyletic nature of *Apanteles foerster* (Hymenoptera: Braconidae): a phylogeny and reclassification of microgastrinae. *Mem Entomol Soc Can*. 1981;113(S115):1–147. <https://doi.org/10.4039/entm113115fv>
17. 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> PMID: 32390740
18. 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>
19. 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> PMID: 15186808
20. Wahlberg N, Kullberg J, Hanski I. Natural history of some Siberian *Melitaeine* butterfly species (Nymphalidae: *Melitaeini*) and their parasitoids. *Entomol Fennica*. 2001;12(2). <https://doi.org/10.33338/ef.84102>
21. Kankare M, Van Nouhuys S, Hanski I. Genetic divergence among host-specific cryptic species in *Cotesia melitaeorum* Aggregate (Hymenoptera: Braconidae), Parasitoids of Checkerspot Butterflies. *an*. 2005;98(3):382–94. [https://doi.org/10.1603/0013-8746\(2005\)098\[0382:gdahcs\]2.0.co;2](https://doi.org/10.1603/0013-8746(2005)098[0382:gdahcs]2.0.co;2)
22. Valerio F, Martel C, Stefanescu C, van Nouhuys S, Kankare M, Duploup A. *Wolbachia* strain diversity in a complex group of sympatric cryptic parasitoid wasp species. *BMC Microbiol*. 2024;24(1):319. <https://doi.org/10.1186/s12866-024-03470-7> PMID: 39223450
23. Schach J. Molecular phylogeny of *Cotesia* wasps parasitizing *Melitaeini* butterflies and the possible role of *Wolbachia* endosymbionts. University of Helsinki; 2024.

24. Miles CI, Chen WP, Adamo SA, Kester KM, Miller DW. *Manduca sexta* caterpillars parasitized by the wasp *Cotesia congregata* stop chewing despite an intact motor system. *J Exp Biol.* 2023;226(16):jeb245716. <https://doi.org/10.1242/jeb.245716> PMID: 37534841
25. Blažytė-Čereškienė L, Aleknavičius D, Apšegaitė V, Būda V. Response of parasitic wasp *Cotesia glomerata* L. (Hymenoptera: Braconidae) to cabbage plants of two varieties: olfactory spectra of males and females. *J Econ Entomol.* 2022;115(5):1464–71. <https://doi.org/10.1093/jee/toac135> PMID: 36062941
26. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215(3):403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) PMID: 2231712
27. Ye X, Xiong S, Teng Z, Yang Y, Wang J, Yu K, et al. Retraction: Amino acid synthesis loss in parasitoid wasps and other hymenopterans. *Elife.* 2020;9:e65123. <https://doi.org/10.7554/eLife.65123> PMID: 33258773
28. Gauthier J, Boulain H, van Vugt JJFA, Baudry L, Persyn E, Aury J-M, et al. Chromosomal scale assembly of parasitic wasp genome reveals symbiotic virus colonization. *Commun Biol.* 2021;4(1):104. <https://doi.org/10.1038/s42003-020-01623-8> PMID: 33483589
29. Pinto BJ, Weis JJ, Gamble T, Ode PJ, Paul R, Zaspel JM. A Chromosome-level genome assembly of the Parasitoid Wasp, *Cotesia glomerata* (Hymenoptera: Braconidae). *J Hered.* 2021;112(6):558–64. <https://doi.org/10.1093/jhered/esab032> PMID: 34043785
30. Muller H, Chebbi MA, Bouzar C, Périquet G, Fortuna T, Calatayud P-A, et al. Genome-wide patterns of bracovirus chromosomal integration into multiple host tissues during parasitism. *J Virol.* 2021;95(22):e0068421. <https://doi.org/10.1128/JVI.00684-21> PMID: 34319152
31. Burke GR, Strand MR. Deep sequencing identifies viral and wasp genes with potential roles in replication of *Microplitis demolitor* Bracovirus. *J Virol.* 2012;86(6):3293–306. <https://doi.org/10.1128/JVI.06434-11> PMID: 22238295
32. 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. PMID: 7881515
33. Dowton M, Austin AD. Molecular phylogeny of the insect order Hymenoptera: apocritan relationships. *Proc Natl Acad Sci U S A.* 1994;91(21):9911–5. <https://doi.org/10.1073/pnas.91.21.9911> PMID: 7937916
34. Whitfield JB. Molecular and morphological data suggest a single origin of the polydnaviruses among braconid wasps. *Naturwissenschaften.* 1997;84(11):502–7. <https://doi.org/10.1007/s001140050434>
35. Quicke DLJ, Austin AD, Fagan-Jeffries EP, Hebert PDN, Butcher BA. Recognition of the Trachypetidae stat.n. as a new extant family of Ichneumonoidea (Hymenoptera), based on molecular and morphological evidence. *Syst Entomol.* 2020;45(4):771–82. <https://doi.org/10.1111/syen.12426>
36. Mardulyn P, Whitfield JB. Phylogenetic signal in the COI, 16S, and 28S genes for inferring relationships among genera of Microgastrinae (Hymenoptera; Braconidae): evidence of a high diversification rate in this group of parasitoids. *Mol Phylogenet Evol.* 1999;12(3):282–94. <https://doi.org/10.1006/mpev.1999.0618> PMID: 10413623
37. Belshaw R, Quicke DL. A molecular phylogeny of the Aphidiinae (Hymenoptera: Braconidae). *Mol Phylogenet Evol.* 1997;7(3):281–93. <https://doi.org/10.1006/mpev.1996.0400> PMID: 9187088
38. Neto MSR, Barros RP, Morais MAJ, Balbino VQ, Loreto V. ITS2 as a molecular marker for the identification of *Diatraea saccharalis* and *D. flavipennis* and possible infection with *Cotesia* spp. *Genet Mol Res.* 2017;16(3). <https://doi.org/10.4238/gmr16039732> PMID: 28873209
39. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics.* 2012;13:134. <https://doi.org/10.1186/1471-2105-13-134> PMID: 22708584
40. Manni M, Berkeley MR, Seppey M, Zdobnov EM. BUSCO: assessing genomic data quality and beyond. *Curr Protoc.* 2021;1(12):e323. <https://doi.org/10.1002/cpz1.323> PMID: 34936221
41. Banks JC, Whitfield JB. Dissecting the ancient rapid radiation of microgastrine wasp genera using additional nuclear genes. *Mol Phylogenet Evol.* 2006;41(3):690–703. <https://doi.org/10.1016/j.ympev.2006.06.001> PMID: 16854601
42. 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> PMID: 15034147
43. 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> PMID: 25095880
44. 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> PMID: 32011700
45. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* 2024;52(W1):W78–82. <https://doi.org/10.1093/nar/gkae268> PMID: 38613393
46. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* 2017;14(6):587–9. <https://doi.org/10.1038/nmeth.4285> PMID: 28481363
47. Chernomor O, von Haeseler A, Minh BQ. Terrace aware data structure for phylogenomic inference from supermatrices. *Syst Biol.* 2016;65(6):997–1008. <https://doi.org/10.1093/sysbio/syw037> PMID: 27121966
48. Lanfear R, Hahn MW. The meaning and measure of concordance factors in phylogenomics. *Mol Biol Evol.* 2024;41(11):msae214. <https://doi.org/10.1093/molbev/msae214> PMID: 39418118
49. Minh BQ, Hahn MW, Lanfear R. New methods to calculate concordance factors for phylogenomic datasets. *Mol Biol Evol.* 2020;37(9):2727–33. <https://doi.org/10.1093/molbev/msaa106> PMID: 32365179

50. Mo YK, Lanfear R, Hahn MW, Minh BQ. Updated site concordance factors minimize effects of homoplasy and taxon sampling. *Bioinformatics*. 2023;39(1):btac741. <https://doi.org/10.1093/bioinformatics/btac741> PMID: [36383168](#)
51. Zhang C, Rabiee M, Sayyari E, Mirarab S. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*. 2018;19(Suppl 6):153. <https://doi.org/10.1186/s12859-018-2129-y> PMID: [29745866](#)
52. Flouri T, Jiao X, Rannala B, Yang Z. Species tree inference with BPP using genomic sequences and the multispecies coalescent. *Mol Biol Evol*. 2018;35(10):2585–93. <https://doi.org/10.1093/molbev/msy147> PMID: [30053098](#)
53. Zhang J, Kapli P, Pavlidis P, Stamatakis A. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*. 2013;29(22):2869–76. <https://doi.org/10.1093/bioinformatics/btt499> PMID: [23990417](#)
54. Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C. Coalescent-based species delimitation in an integrative taxonomy. *Trends Ecol Evol*. 2012;27(9):480–8. <https://doi.org/10.1016/j.tree.2012.04.012> PMID: [22633974](#)
55. Fujita MK, Leaché AD. A coalescent perspective on delimiting and naming species: a reply to Bauer et al. *Proc R Soc B*. 2010;278(1705):493–5. <https://doi.org/10.1098/rspb.2010.1864>
56. Michel-Salzat A, Whitfield JB. Preliminary evolutionary relationships within the parasitoid wasp genus *Cotesia* (Hymenoptera: Braconidae: Microgasterinae): combined analysis of four genes. *Syst Entomol*. 2004;29(3):371–82. <https://doi.org/10.1111/j.0307-6970.2004.00246.x>
57. Quicke DLJ. We know too little about parasitoid wasp distributions to draw any conclusions about latitudinal trends in species richness, body size and biology. *PLoS One*. 2012;7(2):e32101. <https://doi.org/10.1371/journal.pone.0032101> PMID: [22355411](#)
58. Leppänen SA, Altenhofer E, Liston AD, Nyman T. Ecological versus phylogenetic determinants of trophic associations in a plant-leafminer-parasitoid food web. *Evolution*. 2013;67(5):1493–502. <https://doi.org/10.1111/evo.12028> PMID: [23617924](#)
59. Sukumaran J, Knowles LL. Multispecies coalescent delimits structure, not species. *Proc Natl Acad Sci U S A*. 2017;114(7):1607–12. <https://doi.org/10.1073/pnas.1607921114> PMID: [28137871](#)
60. Luo A, Ling C, Ho SYW, Zhu C-D. Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Syst Biol*. 2018;67(5):830–46. <https://doi.org/10.1093/sysbio/syy011> PMID: [29462495](#)
61. Xi Z, Liu L, Davis CC. The impact of missing data on species tree estimation. *Mol Biol Evol*. 2016;33(3):838–60. <https://doi.org/10.1093/molbev/msv266> PMID: [26589995](#)
62. Long EC, Thomson RC, Shapiro AM. A time-calibrated phylogeny of the butterfly tribe Melitaeini. *Mol Phylogenet Evol*. 2014;79:69–81. <https://doi.org/10.1016/j.ympev.2014.06.010> PMID: [24952320](#)
63. Moore SD. Regulation of host diapause by an insect parasitoid. *Ecol Entomol*. 1989;14(1):93–8. <https://doi.org/10.1111/j.1365-2311.1989.tb00758.x>
64. Kawahara AY, Plotkin D, Espeland M, Meusemann K, Toussaint EFA, Donath A, et al. Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. *Proc Natl Acad Sci U S A*. 2019;116(45):22657–63. <https://doi.org/10.1073/pnas.1907847116> PMID: [31636187](#)
65. Zadra N, Rizzoli A, Rota-Stabelli O. Chronological Incongruences between Mitochondrial and Nuclear Phylogenies of *Aedes* Mosquitoes. *Life (Basel)*. 2021;11(3):181. <https://doi.org/10.3390/life11030181> PMID: [33669100](#)
66. Maddison WP. Gene trees in species trees. *Syst Biol*. 1997;46(3):523–36. <https://doi.org/10.1093/sysbio/46.3.523>
67. Forbes AA, Devine SN, Hippee AC, Tvedte ES, Ward AKG, Widmayer HA, et al. Revisiting the particular role of host shifts in initiating insect speciation. *Evolution*. 2017;71(5):1126–37. <https://doi.org/10.1111/evo.13164> PMID: [28052326](#)
68. Hood GR, Forbes AA, Powell THQ, Egan SP, Hamerlinck G, Smith JJ, et al. Sequential divergence and the multiplicative origin of community diversity. *Proc Natl Acad Sci U S A*. 2015;112(44):E5980–9. <https://doi.org/10.1073/pnas.1424717112> PMID: [26499247](#)
69. Kimura MT, Novković B. Local adaptation and ecological fitting in host use of the *Drosophila* parasitoid *Leptopilina japonica*. *Ecol Res*. 2015;30(3):499–505. <https://doi.org/10.1007/s11284-015-1244-8>
70. Braga MP, Guimarães PR Jr, Wheat CW, Nylin S, Janz N. Unifying host-associated diversification processes using butterfly-plant networks. *Nat Commun*. 2018;9(1):5155. <https://doi.org/10.1038/s41467-018-07677-x> PMID: [30514925](#)
71. Ehrlich PR, Raven PH. Butterflies and plants: a study in coevolution. *Evolution*. 1964;18(4):586. <https://doi.org/10.2307/2406212>
72. Janz N, Nylin S. The oscillation hypothesis of host-plant range and speciation. *Specialization, Speciation, and Radiation. The Evolutionary Biology of Herbivorous Insects*. 2008. pp. 203–15.