



DROP: Molecular voucher database for identification of *Drosophila* parasitoids

Chia-Hua Lue, Matthew L Buffington, Sonja Scheffer, Matthew Lewis, Tyler A Elliott, Amelia R I Lindsey, Amy Driskell, Anna Jandova, Masahito T Kimura, Yves Carton, et al.

► To cite this version:

Chia-Hua Lue, Matthew L Buffington, Sonja Scheffer, Matthew Lewis, Tyler A Elliott, et al.. DROP: Molecular voucher database for identification of *Drosophila* parasitoids. *Molecular Ecology Resources*, 2021, 10.1111/1755-0998.13435 . hal-03275697

HAL Id: hal-03275697

<https://hal.science/hal-03275697>

Submitted on 1 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **DROP: Molecular voucher database for identification of *Drosophila* parasitoids**

2 Resource article

3 Word count: 7850 excluding references

4

5 **Authors**

6 Chia-Hua Lue (C-HL), (0000-0002-5245-603X), chiachia926@gmail.com, corresponding

7 author

8 Matthew L. Buffington (MLB)

9 Sonja Scheffer (SS)

10 Matthew Lewis (ML)

11 Tyler A. Elliott (TAE)

12 Amelia R. I. Lindsey (AL)

13 Amy Driskell (AD), (0000-0001-8401-7923)

14 Anna Jandova (AJ)

15 Masahito T. Kimura (MTK)

16 Yves Carton (YC)

17 Robert R. Kula (RRK)

18 Todd A. Schlenke (TAS)

19 Mariana Mateos (MM), (0000-0001-5738-0145)

20 Shubha Govind (SG), (0000-0002-6436-639X)

21 Julien Varaldi (JV)

22 Emilio Guerrieri (EG), (0000-0002-0583-4667)

- 23 Massimo Giorgini (MG), (0000-0001-8670-0945)
- 24 Xingeng Wang (XW)
- 25 Kim Hoelmer (KH)
- 26 Kent M. Daane (KMD)
- 27 Paul K. Abram (PKA)
- 28 Nicholas A. Pardikes (NAP), (0000-0002-9175-4494)
- 29 Joel J. Brown (JJB), (0000-0002-3608-6745)
- 30 Melanie Thierry (MT)
- 31 Marylène Poirié (MP)
- 32 Paul Goldstein (PG), (0000-0002-1443-7030)
- 33 Scott E. Miller (SEM), (0000-0002-4138-1378)
- 34 W. Daniel Tracey (WDT)
- 35 Jeremy S. Davis (JSD), (0000-0002-5214-161X)
- 36 Francis M. Jiggins (FMJ)
- 37 Bregje Wertheim (BW)
- 38 Owen T. Lewis (OTL)
- 39 Jeff Leips (JL)
- 40 Phillip P. A. Staniczenko (PPAS)
- 41 Jan Hrcek (JH), (0000-0003-0711-6447), janhrcek@gmail.com
- 42
- 43
- 44

Affiliations

(C-HL, AJ, NAP, JJB, MT, JH) Biology Centre of the Czech Academy of Sciences, Institute of Entomology, Branisovska 31, 37005 Ceske Budejovice, Czech Republic

(C-HL, PPAS) Department of Biology, Brooklyn College, City University of New York (CUNY), 2900 Bedford Ave, Brooklyn, NY11210, USA

(MLB, SS, ML, RRK, PG) Systematic Entomology Laboratory, ARS/USDA c/o Smithsonian Institution, National Museum of Natural History, 10th& Constitution Ave, NW, Washington DC 20560, USA

(TAE) Centre for Biodiversity Genomics, 50 Stone Road East, University of Guelph, Guelph, Ontario, N1G2W1, Canada

(AL) Department of Entomology, University of Minnesota, Saint Paul MN 55108.

(AD) Laboratories of Analytical Biology, Smithsonian Institution, National Museum of Natural History, 10th & Constitution Ave, NW, Washington DC 20560, USA

(MTK) Hokkaido University Museum, Sapporo, Hokkaido 060-0810, Japan

(YC) “Évolution, Génomes, Comportement, Écologie”, CNRS et Université Paris-Saclay

(TAS) Department of Entomology at the University of Arizona, Forbes 410, PO BOX 210036, Tucson, AZ 85721-0036.

(MM) Wildlife and Fisheries Sciences Department, Texas A&M University

(SG) The Graduate Center of the City University of New York, 365 Fifth Avenue, New York, NY10016, USA

(JV) Université de Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive UMR 5558, F-69622 Villeurbanne, France

67 (EG, MG) CNR- Institute for Sustainable Plant Protection (CNR-IPSP), National Research
68 Council of Italy, Piazzale E. Fermi 1, 80055 Portici, Italy.

69 (XW, KH) United States Department of Agriculture, Agricultural Research Services,
70 Beneficial Insects Introduction Research Unit. 501 S. Chapel St., Newark, DE 19713, USA

71 (KMD) Department of Environmental Science, Policy and Management. University of
72 California, Berkeley. Mulford Hall, 130 Hilgard Way, Berkeley, CA 94720.

73 (PKA) Agriculture and Agri-Food Canada, Agassiz Research and Development Centre,
74 6947 Hwy #7, Agassiz, V0M 1A0, British Columbia, Canada

75 (JJB, MT, JH) University of South Bohemia, Faculty of Science, Branisovska 31, 37005,
76 Czech Republic

77 (MP) Université “Côte d’Azur”, INRAE, CNRS. and Evolution and Specificity of
78 Multitrophic Interactions (ESIM) Sophia Agrobiotech Institute, 400 Route des Chappes,
79 BP 167, 06903 Sophia Antipolis, France

80 (SEM) Smithsonian Institution, National Museum of Natural History, 10th & Constitution
81 Ave, NW, Washington DC 20560, USA

82 (WDT, JSD) Department of Biology, Indiana University Bloomington. 702 N. Walnut
83 Grove, Bloomington, IN47405

84 (WDT) Gill Center for Biomolecular Science, Indiana University Bloomington. 702 N.
85 Walnut Grove, Bloomington, IN47405

86 (JSD) Biology Department, University of Kentucky, 101 T. H. Morgan Building, Lexington,
87 KY, 40506

(FMJ) Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2

3EH UK

(BW) Groningen Institute for Evolutionary Life Sciences, University of Groningen,

Nijenborgh 9, 9747 AG Groningen, the Netherlands

(OTL) Department of Zoology, University of Oxford. 11a Mansfield Road, Oxford OX1

3SZ, UK.

(JL) Department of Biological Sciences, University of Maryland Baltimore County. 1000

Hilltop circle, Baltimore, MD, 21250

Abstract

Molecular identification is increasingly used to speed up biodiversity surveys and laboratory experiments. However, many groups of organisms cannot be reliably identified using standard databases such as GenBank or BOLD due to lack of sequenced voucher specimens identified by experts. Sometimes a large number of sequences are available, but with too many errors to allow identification. Here we address this problem for parasitoids of *Drosophila* by introducing a curated open-access molecular reference database, DROP (*Drosophila* parasitoids). Identifying *Drosophila* parasitoids is challenging and poses a major impediment to realize the full potential of this model system in studies ranging from molecular mechanisms to food webs, and in biological control of *Drosophila suzukii*. In DROP (<http://doi.org/10.5281/zenodo.4519656>), genetic data are linked to voucher specimens and, where possible, the voucher

specimens are identified by taxonomists and vetted through direct comparison with primary type material. To initiate DROP, we curated 154 laboratory strains, 853 vouchers, 545 DNA sequences, 16 genomes, 11 transcriptomes, and 6 proteomes drawn from a total of 183 operational taxonomic units (OTUs): 113 described *Drosophila* parasitoid species and 70 provisional species. We found species richness of *Drosophila* parasitoids to be acutely underestimated and provide an updated taxonomic catalogue for the community. DROP offers accurate molecular identification and improves cross-referencing between individual studies that we hope will catalyze research on this diverse and fascinating model system. Our effort should also serve as an example for researchers facing similar molecular identification problems in other groups of organisms.

Key Words

Biodiversity, DNA sequences, Genomes, Integrative taxonomy, Molecular diagnostics, Biological control

Introduction

Building a knowledge base that encompasses ecology, evolution, genetics, and biological control is contingent on reliable taxonomic identifications. Molecular identification is commonly used in groups of organisms with cryptic species that are difficult to identify morphologically (Fagan-Jeffries et al., 2018; Miller et al., 2016; Novotny & Miller, 2014), for the molecular detection of species interactions (Baker et al., 2016; Condon et al., 2014; Garipey et al., 2019; Hrček & Godfray, 2015; Hrcek et al., 2011), and for identification of species from environmental DNA samples (Shokralla et al., 2012). The accuracy of molecular identification, however, depends on the accuracy of identifications associated with sequences databased in existing online depositories. The foundations of that accuracy are the voucher specimens which were sequenced and the collaboration of a taxonomic authority in the deposition of the sequence data.

GenBank serves as the most widely used sequence depository; however, deposition of sequences in GenBank, which is required by most peer-reviewed journals, does not require deposition of associated vouchers. The Barcode of Life Data System database (BOLD) (Ratnasingham & Hebert, 2007) explicitly aims to provide a framework for identifying specimens using single-locus DNA sequences (Hebert et al., 2003; Smith et al., 2005), and while these are associated with vouchers and metadata, the curation of these data is not consistently maintained by those submitting material. A recent study by Pentinsaari et al. (2020) showed misidentification in both databases caused by missteps in the protocols from query sequences to final determination.

Although the BOLD database function “BOLD-IDS” allows considerable database curation (e.g., sequences are used for identification and/or flagging of misidentified/contaminated records), it also automatically includes sequences from GenBank, and may perpetuate the shortcomings previously mentioned since these cannot be curated from within BOLD. As such, the quality of sequences and the reliability of identifications obtained from BOLD-IDS can vary, and depends on the curation by systematists focusing on individual taxa (Meiklejohn et al., 2019). BOLD-IDS works well for taxa where qualified taxonomists have been involved with assuring data quality; some insect examples include beetles (Hendrich et al., 2015), butterflies (Escalante et al., 2010), geometrid moths (Hausmann et al., 2011, 2016; Miller et al., 2016), true bugs (Raupach et al., 2014), and microgastrine wasps (Smith et al., 2013).

Unfortunately, this is not the case of parasitoids (Insecta: Hymenoptera) of *Drosophila* flies (Insecta: Drosophilidae). There are vast numbers of *Drosophila* parasitoid sequences readily available in GenBank and BOLD, as these parasitoids and their hosts are important model organisms in biology. As of this writing, there are 88,666 nucleotide sequences deposited in GenBank for *Leptopilina heterotoma* (Thomson) and *L. boulardi* (Barbotin, Carton & Kelner-Pillault) (Hymenoptera: Figitidae) alone. However, less than 1 % of the identifications associated with these sequences have been confirmed by taxonomists or are associated with voucher specimens deposited in museum collections. With sequencing shifting from individual genes to genomes we risk that the identification problems will soon apply to whole genomes.

There are around 4000 described species of Drosophilidae, and *Drosophila* contains more than a third of the family's described species (O'Grady & DeSalle, 2018). By contrast, although parasitic wasps are generally a species-rich group (Dolphin & Quicke, 2001; Quicke, 2015), the most recent catalogue of parasitoid species that attack *Drosophila* lists only 50 described species (Carton et al., 1986). This disparity suggests that the diversity of parasitic wasps attacking *Drosophila* is severely underestimated, an assertion supported by the results presented here. This is largely a consequence of the challenging nature of parasitoid taxonomy, in which morphological identification is intractable for many species, and the fact that taxonomic specialists are greatly outnumbered by the species they study.

Currently, only a few biological study systems have been characterized in sufficient breadth and depth to allow researchers to connect various levels of biological organization, from molecular mechanisms to food webs of interacting species. Parasitoids of *Drosophila* represent one such system (Prévost, 2009). Moreover, the practical feasibility of rearing parasitoids of *Drosophila* parasitoids under laboratory conditions has led to a number of fundamental discoveries in ecology (Carton et al., 1991; Terry et al. 2020; Thierry et al., 2021), evolution (Kraaijeveld & Godfray, 1997), immunology (Kim-Jo et al., 2019; Nappi & Carton, 2001; Schlenke et al., 2007), physiology (Melk & Govind, 1999), symbiosis (Xie et al., 2011, 2015), behavioral science (Lefèvre et al., 2012) and other fields. In contrast to this large body of laboratory studies, basic natural history of *Drosophila* parasitoids, especially their species richness (Kimura & Mitsui, 2020; Lue et al., 2018), is little known. Addressing this knowledge gap

is especially pressing given current efforts to use parasitoids in biological control efforts, such as those of the invasive pest spotted wing *Drosophila*, *Drosophila suzukii* (Abram et al., 2020; Daane et al., 2016; Giorgini et al., 2019; Wang et al., 2020 a&b).

Properly executed molecular identification has the potential to be much more efficient for the majority of researchers, and many laboratory strains are commonly identified using DNA sequences alone. While it is practical for researchers to assign species names based on a match to sequence records in genetic databases, this practice often causes a cascade of inaccuracies. To illustrate the extent of the problem, we note the example of *Ganaspis*, a genus of parasitoids commonly used in laboratories that includes both superficially indistinguishable species with highly divergent sequences that are often treated as conspecific, as well as specimens with identical sequences identified under different names (Figure 1).

To address these issues, we introduce a newly curated molecular reference database for *Drosophila* parasitoids —DROP— in which sequences are either linked to voucher specimens identified by taxonomists or have a traceable provenance (Figure 2). The first aim of DROP is to provide a reliable DNA sequence library for molecular identification of *Drosophila* parasitoids that enables cross-referencing of original taxonomic concepts with those of subsequent studies. We pay special attention to live parasitoid strains which are available for future experiments. The second aim is to standardize and expedite the linkage between specimens and available sequence data; we place a premium on museum vouchers as they allow for repeatable scientific research. In DROP, this goal is facilitated through a consolidated digital infrastructure of data associated

with laboratory strains, offering the opportunity for researchers to re-examine past experimental results in a permanent context. The third aim is to provide an up-to-date catalogue of the diversity of *Drosophila* parasitoids as a foundation for advancing the understanding of their taxonomy. Finally, the fourth aim of DROP is for our collaborative effort to serve as an inspiration to communities of researchers studying other groups of organisms who are experiencing difficulties with the reliability of molecular reference databases.

Materials and Methods

Drosophila species and their parasitoids

The phylogenetic and subgeneric structure within *Drosophila* and related genera is not yet fully resolved (O’Grady & DeSalle, 2018). Various subgenera, including *Scaptomyza*, *Zaprionus*, *Lordiphosa* and *Samoaia*, have been treated as both genera and subgenera, and researchers have yet to achieve consensus on these various hypotheses (O’Grady & DeSalle, 2018; Remsen & O’Grady, 2002; Yassin, 2013; Yassin & David, 2010). Species in *Drosophila* subgenera and genera closely related to *Drosophila* commonly share niche space and natural histories and, as a result, are often attacked by overlapping or identical groups of parasitoids. For instance, the invasive African fig fly, *Zaprionus indianus* Gupta is attacked by *Pachycrepoideus vindemiae* (Rondani, 1875) and *Leptopilina boulardi* (Pfeiffer et al., 2019; Santos et al., 2016), all of which have been recorded from *Drosophila*. Therefore, we also include these groups within the contents of DROP.

Parasitoids of *Drosophila* belong to four superfamilies of Hymenoptera (Chalcidoidea, Cynipoidea, Ichneumonoidea, Diaprioidea) which evolved parasitism of *Drosophila* flies independently (Carton et al., 1986; Prévost, 2009). All the parasitoids known to attack *Drosophila* are solitary and attack either the larval or pupal stage; in both cases, they emerge from the fly's puparium. The known *Drosophila* larval parasitoids belong to two families, Braconidae (including the genera *Asobara*, *Aphaereta*, *Phaenocarpa*, *Tanycarpa*, *Aspilota*, *Opius*) and Figitidae (*Leptopilina*, *Ganaspis*, *Leptolamina*, *Kleidotoma*); all are koinobionts that allow the host to continue development while the parasitoid grows within it. The known *Drosophila* pupal parasitoids belong to three other families, Diapriidae (*Trichopria*, *Spilomicrus*), Pteromalidae (*Pachycrepoideus*, *Spalangia*, *Trichomalopsis*, *Toxomorpha*) and Encyrtidae (*Tachinaephagus*); they are all idiobionts that terminate host development immediately. Host-specificity across the *Drosophila* parasitoids is poorly characterized—while some can parasitize other families of Diptera (e.g., *Aphaereta aotea*) (Hughes & Woolcock, 1976), most are thought to be limited to *Drosophila* hosts.

Data sources

To assemble the DROP database, we targeted 20 genera that potentially parasitize frugivorous *Drosophila* species. We compiled DNA sequence and voucher data from four sources: 1) museum collections, 2) publications, 3) molecular biodiversity inventories publicly available in BOLD and GenBank, for which we managed to secure inspection of

the vouchers by taxonomists, and 4) conducted a sequencing and taxonomic inventory of laboratory strains.

We first gathered species information into a catalogue of *Drosophila* parasitoid species (Table 1) from 212 references (see DROP database reference table) and 36 institutes (Table S2). To ensure reliable names for nominal species (sequences identified by a species name) in our database, we confirmed their taxonomic validity using the Ichneumonoidea 2015 digital catalogue (Yu et al., 2016; <https://web.archive.org/web/20161022093945/http://ichneumonoidea.name/global.php>) and Hymenoptera Online (HOL; <http://hol.osu.edu/>), both of which are curated by taxonomic experts. To obtain reliable molecular identification data, we harvested 8,298 DNA sequences from GenBank and BOLD (all compiled into BOLD system: DS-DROPAR dataset). These sequences represent 443 Barcode Index Numbers (BINs – a form of provisional taxa in BOLD) and 520 taxa, for a total of 963 operational taxonomic units (OTUs). We use the term “OTU” as a general and neutral designation encompassing described species, provisional species, undescribed species, cryptic species, and mis-identified species.

The majority of the harvested sequences were Braconidae (6690), Diapriidae (967), Figitidae (622), and Pteromalidae (19). Because of the concerns with generic databases (noted above and in Figure 1 and Table S1), we assembled a list of sequences with valid species names that could either be traceably linked to vouchers examined by taxonomists or referred to directly in publications authored by a recognized expert in the relevant taxon group. We then cross-checked species names with their

corresponding BINs in BOLD and flagged potential conflicts between species names and BINs (Table S1).

A core goal of DROP besides that of a tool for biodiversity research is to function as a platform that accommodates *Drosophila* parasitoids kept in culture (for experimental work) or in quarantine (for biological control applications). So far, there has been a lack of a coherent and reliable means of verifying species kept in laboratory settings, which can be a serious problem. Since lab cultures are routinely contaminated by neighboring cultures (e.g., through escapees), one species may be displaced by another even under a vigilant eye.

For lab and quarantine lines in DROP, we deposited DNA extractions and vouchers in the National Insect Collection, National Museum of Natural History, Smithsonian Institution (USNM; Washington, DC, USA). During their initial assembly for DROP, laboratory OTUs were designated by their strain name; most laboratory strains can be associated with provisional species but some cannot yet be assigned. Three females and three males of each strain were dry-mounted and individually assigned a USNMENT 'QR code' specimen label as representative vouchers. For each molecular voucher, three legs from a female wasp were removed for DNA extraction and sequencing (Supplementary Methods for details), and the rest of the body was assigned a USNMENT specimen label and preserved for morphological identification. Both DNA extraction and vouchers were entered into the database and uploaded to BOLD (DROP project: DS-LABS) with an associated GenBank ID [NOTE: the BOLD records will be pushed to GenBank at revision stage; these data are not embargoed].

Where possible, we identified OTU strains using a combination of morphological and sequence data, and characterized provisional species or species clusters using neighbor-joining trees (Figure S1) based on the COI gene sequences (Supplemental material). For establishing BIN limits in the context of DROP, we have adopted an initial percent cutoff at 2%. As Ratnasingham & Hebert (2013) pointed out, this is a good starting point for many taxa, but it also may be adjusted as more samples are acquired and compared.

Drosophila parasitoid database—DROP

To compile the above information, we built a simple Structured Query Language (SQL) database in sqlite3 format using SQLiteStudio. There are eight linked tables in the database—species, strain, voucher, sequence, genome, transcriptome, proteome and reference—along with additional tables for linking these to reference table (Figure S2). The database incorporates all sample fields used by BOLD for compatibility and includes a number of new fields to accommodate a catalogue of *Drosophila* parasitoid species, laboratory strain information, and links from the DROP database to BOLD and GenBank records.

DROP is available on Zenodo (<http://doi.org/10.5281/zenodo.4519656>) for permanent deposition and version control. In addition to the main database, the Zenodo repository includes additional files to facilitate easy use of the database. These files include: 1) the reference database in comma-separated text (.csv) and FASTA format ready to be used for molecular identification; 2) a species catalogue with taxonomic information; and 3) a list of laboratory strains with confirmed molecular

vouchers. DROP will be continued to be maintained by C-HL until further notice at the Zenodo repository and sequences generated in the future will also be deposited in BOLD (DROP project).

Species, provisional species, and OTU designations

In addition to the inherent value of a formal taxonomic name, a reliable provisional taxon label can also be used for exchanging scientific information and conveying experimental results among researchers (Schindel & Miller, 2010). Based on the amount of sequence divergence between described species, we observed what appears to be a significant number of provisional OTUs in the initial dataset we compiled. Furthermore, among the data linked to a valid species name, some of these provisional OTUs are actively being used in research and have sequences available to the public. We therefore provide a list of potential new species with their molecular vouchers.

We use the following designation format for OTUs that refer to a provisional species: “Drop_strainX_sp.1” or, when no other information is known, “DROP_sp.1”. Where possible, these OTUs are linked to BINs within BOLD and to a voucher USNM specimen label number. If the genus of the OTU is known, the “Drop_Leptopilina_sp.1” format is followed. These designations can facilitate species identification as well as discovery and description of new species without compromising the existing taxonomy of the described OTUs in question. As more complete species descriptions become available, this provisional species framework can be updated while keeping the link to previous provisional species name.

Results

Overview of DROP

We catalogued 182 OTUs in the DROP database with 113 described species of *Drosophila* parasitoids and 69 provisional species (Table 1). In total, we documented 154 laboratory strains (Table S3), 853 vouchers from 36 institutions (Table S2). Among the described species, 98 have voucher information, of which 61 are traceable to type specimens, including 45 to holotypes (i.e., specimen used to root a name to the taxonomic author's concept of the species). *Leptopilina* is represented by the highest number of species with 45 OTUs, followed by *Asobara* with 26 OTUs. Within the 154 catalogued lab strains, 86 were actively being used in ongoing research (i.e., a live strain being cultivated). These strains represent 39 OTUs: 11 described species and 28 provisional species (Table S3, Figure S1).

Molecular Vouchers

So far, DROP includes 545 DNA sequences and links to 16 genomes (Table 2.1), 11 transcriptomes (Table 2.2), and 6 proteomes (Table 2.3). From the total of 8298 DNA sequences (dataset: DS-DROPAR) collected from public databases, only 322 sequences (less than 4% of available sequences) satisfied the criteria for validity we imposed for molecular vouchers (see material and methods) included in DROP. The DS-DROPAR dataset initially referred to 520 taxa names, but only 52 names were valid, linked to vouchers, or linked to a publication with evidence that the specimens had been identified by taxonomists. The remaining 223 of 545 DROP DNA sequences were

generated by this project (dataset: DS-LABS, DS-AUSPTOID) and came from 121 OTUs (12 provisional species and 101 lab strains).

The DROP database is largely made up of standard barcode COI sequences (340 sequences), which includes 77 OTUs: 43 described species and 33 provisional species. We aimed to supplement COI with secondary markers (28SD2, 18S, ITS2) when possible, resulting in an additional 120 sequences from 26 OTUs: 15 described species and 11 provisional species. There are currently 19 OTUs that have sequences from more than one genetic marker.

Species Delimitation in Laboratory Strains

We used 298 COI sequences to resolve the identification of each laboratory strain, and where possible, indicated potential species clusters for *Drosophila* parasitoids (Fig. S1 and Table S3). Using a fixed 2% divergence cutoff, a total of 31 lab strain OTUs were assignable to a valid species name, and the remaining 70 strain OTUs were assigned to a provisional species. The taxonomic status of several of these provisional species is also being investigated using an integrative taxonomic approach involving morphological identification, genomic data, or other genetic data.

Discussion

In this paper, we introduce and describe a free and open-access database for the reliable molecular identification of *Drosophila* parasitoids. The guiding principle of DROP

is data credibility, based on the prerequisite that genetic data be associated with explicit criteria linking voucher specimens with taxonomic concepts of the original authors (Troudet et al., 2018). When incorporating information from public genetic databases, we include only sequences that have passed our filtering protocol. This protocol ensures each entry is associated with a valid scientific name, provisional name, or consistently applied OTU designation that can be used to integrate genetic and organismal data from independent studies.

The following discussion expands on the utility of DROP and how we hope it will benefit molecular species identification, connect research from various disciplines, support biological control applications, and serve as a long-term molecular voucher repository and clearinghouse for vetted data.

Molecular (mis-)identification

We observe that 17% of the described *Drosophila* parasitoid OTUs in BOLD and GenBank (dataset: DS-DROPAR) are associated with more than one BIN; these are examples of BIN-ID conflict. Roughly half of these OTUs are used as lab strains. This latter observation is disturbing, because it demonstrates that the criteria used to differentiate and reference species in active research programs are clouded. For example, BIN-ID conflicts were observed in the *Drosophila* parasitoids *Ganaspis brasiliensis* (Ihering) and *Asobara japonica* Belokobylskij (Table S1), both of which are in active use in numerous research programs (e.g. Moreau et al., 2009; Nomano et al., 2017; Reumer et al., 2012; Wang et al., 2020a & 2021) as well as in biological control

efforts against the invasive *D. suzukii* (e.g. Abram et al., 2020; Daane et al., 2016; Giorgini et al., 2019). All the BINs from *G. brasiliensis* carry the name *G. xanthopoda* (Figure 1). In such instances, assigning an identification by matching specimens to barcode records in the genetic database is problematic, as two names are applied to the same BIN. If sequences comprising the BIN are not linked to a voucher that can be examined, teasing apart the two names and how they are applied is impossible. Applying explicit, consistent criteria for species determination ensures that experimental results can be reliably repeated, and that any potentially novel observations will not be explained away as artifacts of identification. DROP addresses these concerns by linking reliable reference sequences and vouchers for *G. brasiliensis* (Figure 1) and from different studies: one with reference to the morphological description (Buffington & Forshage, 2016) and the other with reference to the genome (using voucher specimens from the morphological study; Blaimer et al., 2020).

We were not able to resolve all conflicts between BIN and species identity, for one or more of the following three reasons: First, many records lack reliably identified vouchers and have often been themselves used for molecular identification, proliferating errors. Second, in some cases, it is not possible to verify whether the genetic differences among BINs represent different species or simply intraspecific genetic variation (Bergsten et al., 2012), because BINs themselves are not a species concept. The only solution to this problem is to derive original sequence data from type specimens (which is often either impractical or impossible for a number of technical reasons), or from specimens whose conspecificity with the types has been corroborated.

Since species boundaries are always subject to testing, additional specimens from multiple collecting events (e.g., representing different seasons and geographic regions) may help provide the additional data to circumscribe a given species' limits. The third difficulty in resolving BIN-ID conflict derives from the data themselves: Although the mitochondrial COI gene is the locus most frequently chosen for identification of insects and other animals, its effectiveness varies among insect groups (Brower & DeSalle, 2002; Gompert et al., 2008; Lin & Danforth, 2004). In part, this derives from gene-tree/species-tree conflict as a function of mitochondrial DNA introgression (Gompert et al., 2008; Klopstein et al., 2016), parthenogenesis (Reumer et al., 2012), and/or *Wolbachia* infection (Wachi et al., 2015; Xiao et al., 2012), any of which may lead to complications in species delimitation using mitochondrial loci. Ideally, studies should apply multiple loci, genomes, and comparative taxonomic data to clarify species boundaries. As *Drosophila* parasitoids are often maintained in laboratory cultures, it is also possible to use mating experiments to explore species boundaries under the paradigm of the biological species concept (Seehausen et al., 2020).

DROP as a taxonomic tool

DROP offers an empirical platform for species discovery and a useful tool for taxonomic research. The fact that the number of BINs reported here exceeds the number of described species (Table S1, Figure S3) highlights the need for taxonomic work. But such work cannot proceed on the basis of BINs or barcodes, but requires integrative taxonomic approach employing a combination of molecular and

morphological data. Describing new species on the sole basis of a barcode or BIN, without the benefit of independent character data, should, in general, be avoided. It risks creating nomenclatural synonymy if it is later determined that a sequence can be attributed to a specimen that bears a valid, available name. Moreover, BINs are based on distance analyses which, by definition, are incompatible with diagnoses per se (Ferguson, 2002; Prendini et al., 2002; Goldstein & DeSalle, 2011). Therefore, in taxonomic treatments, it is critical to clarify the range of applicability of a given BIN and its overlap with a taxonomic name (see example in Figure 1).

Public genetic databases have adopted a longstanding convention in treating undetermined OTUs and sequences, referring to provisional species with numbers, as for example “sp. 1”, and these are rarely linked to vouchers. For OTUs designated as provisional species, DROP enables cross-indexing of specimens, sequences and references with studies and provides researchers with valuable tools for taxonomic revisions, including the means of discovery, corroboration, and description of new species. For example, “drop_Gan1_sp.1” refers to voucher USNMMENT01557320 deposited in the USNM, Washington DC, COI sequence (DROP sequence_id 2), BOLD process ID: DROP143-21, BIN number: XXXXXX (will update in the revision), 28D1 sequences (DROP sequence_id 289), and 28D2 sequences (DROP sequence_id 303). In the future, when “drop_Gan1_sp.1” is described as a new species with a formal specific epithet, DROP will update the species status and holotype information while keeping this provisional species name as an informal “synonym.” We recognize tracking these

informal ‘tags’ through time can be problematic; however, linking these tags in DROP to a vouchered specimen and unique identifier will minimize confusion.

From molecular mechanisms to ecosystem structure

The use of molecular tools in insect biodiversity studies has gradually expanded from barcoding single individuals to metabarcoding large environmental samples representing entire food webs (Jefferies et al., 2020; Littlefair et al., 2016). *Drosophila* and their parasitoids are among the few systems that currently allow us to explore thoroughly the mechanisms of species interactions at scales ranging from the molecular to the ecological. Here, we highlight two examples where information compiled in DROP enables the study of the *Drosophila*-parasitoid system across multiple levels of biological organization:

DROP includes a DNA reference library of Australian *Drosophila* parasitoids (DS-AUSPTOID in BOLD) that connects laboratory experiments and field research. Molecular vouchers of both hosts and parasitoids were collected along altitudinal gradients in the rainforest of northern Queensland, Australia (Jefferies et al., 2020). With this DNA reference library, researchers can detect interactions between *Drosophila* and their parasitoids using PCR-based approaches and parasitized pupae (Hrcek & Godfray, 2015; Jefferies et al., 2020). Surveying host-parasitoid interactions in this way will improve our understanding of how environmental change alters the structure of host-parasitoid networks (Morris et al., 2014; Staniczenko et al., 2017; Tylianakis et al., 2007) by accelerating data collection in the field. In addition, JH established lab cultures of both hosts and their parasitoids

from the same Australian sampling sites with the aim of conducting laboratory experiments (e.g. Thierry et al., 2021). Molecular vouchers of the lab strains were then submitted to DROP as a reference database (DS-LABS in BOLD) to ensure that criteria for species determination were applied consistently—and will continue to be applied consistently—between the natural community studies and the laboratory experiments.

The presence of a foundational DNA reference library and species catalogue in DROP will enable the process of exploring parasitoid biodiversity to become more efficient. For example, DROP includes molecular vouchers from *Drosophila* parasitoids that were collected across seasons and along latitudinal gradients in the eastern United States (Lue et al., 2016, 2018). These data proved to be extremely useful for identifying species in a more recent exploration of native parasitoid biodiversity across North America (e.g., Abram et al., 2020). There are additional uses for DROP: curated specimen collections may be used to document species distributions, phenology, understand micro-evolutionary patterns, observe the effects of climate change, and detect and track biological invasions (Funk, 2018; Schilthuizen et al., 2015; Tarli et al., 2018).

Taxonomic accuracy for biocontrol studies

Unfortunately, the history of biological control includes many examples of misidentifications that have resulted in failures to employ or establish the expected control agent, thus hindering eventual success (Buffington et al., 2018; Rosen, 1986; Huffaker et al. 1962). In the context of biological control research on *Drosophila* pest

species, a simple, reliable, and rapid identification tool for their natural enemies is essential (Wang et al. 2020b). By anchoring the criteria for determining identities of organisms being considered for biological control programs, DROP annotation enables the direct examination of centers of origin for parasitoid species, their co-occurrence with natural enemies, and the optimal timing for potential introductions of such enemies (Abram et al., 2020; Daane et al., 2016; Girod et al., 2018a and b; Kimura, 2015; Mitsui et al., 2007). Because most sequences from DROP are already vetted for reliability, they can be used to identify biological control agents rapidly, before or after being brought into quarantine facilities for safety and efficacy testing. This will decrease the risk of non-target ecological impacts arising from misidentifications and facilitate regulatory review for releases of effective and specific natural enemies.

In addition to species identification, reference sequences from DROP may be used to create species-specific primers for the accurate identification of parasitoids, design multiplex PCR assays that rapidly distinguish species in natural or agricultural ecosystems (Ye et al., 2017), and apply high-throughput molecular identification diagnostics (Fagan-Jeffries et al., 2018). Applications of such specific primers have been used in bacteria, fungi, oomycetes and insect pests (Liu et al., 2017; Tedersoo et al., 2019; Tsai et al., 2020).

Long-term molecular voucher preservation

During the curation of DROP, we found that holotype specimens were missing from museums for several iconic *Drosophila* parasitoid species: *Asobara tabida* (Nees von Esenbeck), *Leptopilina clavipes* (Hartig), and *Leptopilina longipes* (Hartig). This is not uncommon and impedes future taxonomic revisions regardless of whether molecular data are used. To avoid contributing to this problem, DROP uses museums as depositories for ensuring that sequenced vouchers of both described species and provisional species are permanently stored. In order to stabilize nomenclature, we further advocate the designation of neotypes (a replacement specimen for a missing holotype) that have museum-vouchered DNA barcodes and additional genomic extractions in storage.

Natural history museums are designed to maintain vouchers (including types) for long-term preservation, and increasingly they implement institutionalized workflows that link DNA sequences to specimens and specimen metadata (Prendini et al., 2002). We strongly encourage the deposition of voucher specimens from field surveys and experimental studies in museum collections, as has been urged by the Entomology Collections Network (ECN) and required in many PhD programs. No matter how quickly new molecular techniques are developed or refined, there is no substitute for a reliable database of voucher specimens when it comes to ensuring the repeatability of biological research (Funk et al., 2005; Lendemer et al., 2020).

Our results show that species richness of the parasitic wasps that attack *Drosophila* is severely underestimated, and only a fraction of them have been described. In DROP, 38% of the OTUs are provisional species, and more than 46% of the named OTUs have

synonyms. Remarkably, *Leptopilina heterotoma*, one of the world's most studied parasitoids, has more than 20 synonyms! As is generally the case, the rate of species description and revision of *Drosophila* parasitoids lags far behind that with which molecular sequence data are generated. Ensuring a consistent application of OTU recognition is therefore essential. With DROP, researchers may ensure consistency in their application of scientific names, and that those names are valid, making the daunting process of describing *Drosophila* parasitoids more accurate and efficient.

In addition to the collection of physical museum resources, a central role taxonomists play in DROP and its curation is that of fostering better integration of taxonomy with experimental and biodiversity research. Our intention is to perpetuate DROP beyond this introductory publication. We hope that experts in all areas of *Drosophila*-parasitoid biology and related fields will join us in this effort.

Conclusion

Taxonomic confusion presents many obstacles in experimental and biodiversity studies. One way of addressing this impediment is to provide a reliable DNA library with traceable vouchers (Astrin et al., 2013). We developed DROP as a resource and platform for gathering and sharing reliable genomic sequence data for *Drosophila* parasitoids. We hope it will serve as a model for researchers working with organisms which present similar difficulties. While compiling DROP, we found that the high number of provisional versus named OTUs suggests that the diversity of parasitic wasps attacking *Drosophila* is greatly underestimated. With this in mind, DROP represents the start of an important

knowledge base that will strengthen future studies of natural host-parasitoid interactions, population dynamics, biocontrol, and the impact of climate change on biodiversity and ecosystem services.

Acknowledgements

The DROP project was developed during the 2018 Entomology Society of America conference, during the symposium “Drosophila parasitoids: from molecular to ecosystem level”. We thank Dr. Elijah Talamas for valuable comments on earlier drafts. We also thank Chris Jeffs for providing Australian field samples. We are also thankful for funding support from the Czech Science Foundation (17-27184Y). Additional fund for sequencing was provided by MLB, OTL, and PPAS. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA. USDA is an equal opportunity provider and employer.

References

- Abram, P. K., Mcpherson, A. E., Kula, R., Hueppelsheuser, T., Perlman, S. J., Curtis, C. I., Fraser, J. L., ... Buffington, M. (2020). New records of *Leptopilina*, *Ganaspis*, and *Asobara* species associated with *Drosophila suzukii* in North America, including detections of *L. japonica* and *G. brasiliensis*. *Journal of Hymenoptera Research*, 78, 1-17, <https://doi.org/10.3897/jhr.78.55026>
- Astrin, J. J., Zhou, X., & Misof, B. (2013). The importance of biobanking in molecular taxonomy, with proposed definitions for vouchers in a molecular context. *ZooKeys*, 365(SPEC.ISSUE), 67–70. <https://doi.org/10.3897/zookeys.365.5875>
- Baker, C. C. M., Bittleston, L. S., Sanders, J. G., & Pierce, N. E. (2016). Dissecting host-associated communities with DNA barcodes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1702). <https://doi.org/10.1098/rstb.2015.0328>

- 615 Bergsten, J., Bilton, D. T., Fujisawa, T., Elliott, M., Monaghan, M. T., Balke, M.,
616 Hendrich, ... Vogler, A. P. (2012). The effect of geographical scale of sampling on
617 DNA barcoding. *Systematic Biology*, 61 (5), 851-869.
618 <https://doi.org/10.1093/sysbio/sys037>
- 619 Blaimer, B. B., Gotzek, D., Brady, S. G., & Buffington, M. (2020). Comprehensive
620 phylogenomic analyses re-write the evolution of parasitism within cynipoid wasps.
621 *BMC Ecology and Evolution*, 20 (155). <https://doi.org/10.1186/s12862-020-01716-2>
- 622 Brower, A. V. Z., & DeSalle, R. (2002). Patterns of mitochondrial versus nuclear DNA
623 sequence divergence among nymphalid butterflies: The utility of wingless as a
624 source of characters for phylogenetic inference. *Insect Molecular Biology*, 7 (1), 73-
625 82. <https://doi.org/10.1046/j.1365-2583.1998.71052.x>
- 626 Buffington, M., & Forshage, M. (2016). Redescription of *Ganaspis brasiliensis* (Ihering,
627 1905), new combination, (Hymenoptera: Figitidae) a natural enemy of the invasive
628 *Drosophila suzukii* (Matsumura, 1931) (Diptera: Drosophilidae). *Proceedings of the*
629 *Entomological Society of Washington*, 118(1), 1–13. <https://doi.org/10.4289/0013-8797.118.1.1>
- 631 Buffington, M., Talamas, E. J., & Hoelmer, K. A. (2018). Team Trissolcus: Integrating
632 taxonomy and biological control to combat the brown marmorated stink bug.
633 *American Entomologist*, 64 (4), 224–232
- 634 Carton, Y., Boulétreau, M., van Alphen, J. J. M., & van Lenteren, J. C. (1986). The
635 *Drosophila* parasitic wasps. In Ashburner M, Carson HL, Thompson JN (Eds), *The*
636 *genetics and biology of Drosophila*, (3), 348–394.
- 637 Carton, Y., Haouas, S., Marrakchi, M., & Hochberg, M. (1991). Two competing parasitoid
638 species coexist in sympatry. *Oikos*, 60, 222-230. <https://doi.org/10.2307/3544869>
- 639 Condon, M. A., Scheffer, S. J., Lewis, M. L., Wharton, R., Adams, D.C., & Forbes, A. A.
640 (2014). Lethal interactions between parasites and prey increase niche diversity in a
641 tropical community. *Science*, 343(6176), pp.1240-1244.
- 642 Daane, K. M., Wang, X.-G., Biondi, A., Miller, B. E., Miller, J. C., Riedl, H., Shearer, P. W.,
643 ... Walton, V. M. (2016). First exploration of parasitoids of *Drosophila suzukii* in
644 South Korea as potential classical biological agents. *Journal of Pest Science* 89, 823–
645 835, doi:10.1007/s10340-016-0740-0.
- 646 Dolphin, K., & Quicke, D. L. J. (2001). Estimating the global incompletely described
647 parasitoid wasps. *Biological Journal Of The Linnean Society*, 73 (3), 279-286,
648 <https://doi.org/10.1006>
- 649 Escalante, P., Ibarra-Vazquez, A., & Rosas-Escobar, P. (2010). Tropical montane
650 nymphalids in Mexico: DNA barcodes reveal greater diversity. *Mitochondrial DNA*,
651 21, 30-37, <https://doi.org/10.3109/19401736.2010.535527>
- 652 Fagan-Jeffries, E. P., Cooper, S. J. B., Bertozzi, T., Bradford, T. M., & Austin, A. D. (2018).
653 DNA barcoding of microgastrine parasitoid wasps (Hymenoptera: Braconidae) using
654 high-throughput methods more than doubles the number of species known for
655 Australia. *Molecular Ecology Resources*, 18(5), 1132–1143.
656 <https://doi.org/10.1111/1755-0998.12904>
- 657 Ferguson, J. W. H. (2002). On the use of genetic divergence for identifying species.
658 *Biological Journal of the Linnean Society*, 75, 509–16.

- Forbes, A. A., Bagley, R. K., Beer, M. A., Hippee, A. C., & Widmayer, H. A. (2018). Quantifying the unquantifiable: Why Hymenoptera, not Coleoptera, is the most speciose animal order. *BMC Ecology*, 18(1), 1–11. <https://doi.org/10.1186/s12898-018-0176-x>
- Funk, V. A. (2018). Collections-based science in the 21st Century. *Journal of Systematics and Evolution*, 56(3), 175–193. <https://doi.org/10.1111/jse.12315>
- Funk, V. A., Hoch, P. C., Prather, L. A., & Wagner, W. L. (2005). The importance of vouchers. *Taxon*, 54(1), 127–129. <https://doi.org/10.2307/25065309>
- Gariepy, T. D., Bruin, A., Konopka, J., Scott-Dupree, C., Fraser, H., Bon, M. C., & Talamas, E. (2019). A modified DNA barcode approach to define trophic interactions between native and exotic pentatomids and their parasitoids. *Molecular Ecology*, 28(2), 456–470. <https://doi.org/10.1111/mec.14868>
- Giorgini, M., Wang, X.-G., Wang, Y., Chen, F. S., Hougardy, E. , Zhang, H. M., Chen, Z. Q., ... Guerrieri, E. (2019). Exploration for native parasitoids of *Drosophila suzukii* in China reveals a diversity of parasitoid species and narrow host range of the dominant parasitoid. *Journal of Pest Science*, 92(2), 509–522. <https://doi.org/10.1007/s10340-018-01068-3>
- Girod, P., Borowiec, N., Buffington, M., Chen, G., Fang, Y., Kimura, M. T., Peris-Felipo, F. J., ... Kenis, M. (2018). The parasitoid complex of *D. suzukii* and other fruit feeding *Drosophila* species in Asia. *Scientific Reports*, 8(1), e11839. <https://doi.org/10.1038/s41598-018-29555-8>
- Girod, P., Lierhmann, O., Urvois, T., Turlings, T. C. J., Kenis, M., & Haye, T. (2018). Host specificity of Asian parasitoids for potential classical biological control of *Drosophila suzukii*. *Journal of Pest Science* 91,1241–1250, <https://doi.org/10.1007/s10340-018-1003-z>
- Goldstein, P. Z., & DeSalle, R. (2011). Integrating DNA barcode data and taxonomic practice: Determination, discovery, and description. *BioEssays*, 33(2),135-147, <https://doi.org/10.1002/bies.201000036>
- Gompert, Z., Forister, M. L., Fordyce, J. A., & Nice, C. C. (2008). Widespread mito-nuclear discordance with evidence for introgressive hybridization and selective sweeps in Lycaeides. *Molecular Ecology*, 17(24), 5231-5244, <https://doi.org/10.1111/j.1365-294X.2008.03988.x>
- Grissell, E. (1999). Hymenopteran biodiversity: some alien notions. *American Entomologist*, 45,236-244.
- Hardy, I. C., van Alphen, J. J. M., & Godfray, H. C. J. (1994). Parasitoids: Behavioral and evolutionary ecology. *The Journal of Animal Ecology*, 63(4), 1009-1010, <https://doi.org/10.2307/5282>
- Hausmann, A., Haszprunar, G., & Hebert, P. D. N. (2011). DNA barcoding the geometrid fauna of bavaria (Lepidoptera): Successes, surprises, and questions. *PLoS ONE*, 6(2), 1–9. <https://doi.org/10.1371/journal.pone.0017134>
- Hausmann, A., Miller, S. E., Holloway, J. D., Dewaard, J. R., Pollock, D., Prosser, S. W. J., & Hebert, P. D. N. (2016). Calibrating the taxonomy of a megadiverse insect family: 3000 DNA barcodes from geometrid type specimens (Lepidoptera, Geometridae). *Genome*, 59(9), 671–684. <https://doi.org/10.1139/gen-2015-0197>

- 703 Hebert, P. D. N., Ratnasingham, S., & DeWaard, J. R. (2003). Barcoding animal life:
704 Cytochrome c oxidase subunit 1 divergences among closely related species.
705 *Proceedings of the Royal Society B: Biological Sciences*, 270 (Suppl.), 96-99,
706 <https://doi.org/10.1098/rsbl.2003.0025>
- 707 Hendrich, L., Morinière, J., Haszprunar, G., Hebert, P. D. N., Hausmann, A., Köhler, F., &
708 Balke, M. (2015). A comprehensive DNA barcode database for Central European
709 beetles with a focus on Germany: Adding more than 3500 identified species to
710 BOLD. *Molecular Ecology Resources*, 15(4), 795-818, [https://doi.org/10.1111/1755-](https://doi.org/10.1111/1755-0998.12354)
711 0998.12354
- 712 Hrček, J., & Godfray, H. C. J. (2015). What do molecular methods bring to host-parasitoid
713 food webs? *Trends in Parasitology*, 31(1), 30–35.
714 <https://doi.org/10.1016/j.pt.2014.10.008>
- 715 Hrcek, J., Miller, S. E., Quicke, D. L. J., & Smith, M. A. (2011). Molecular detection of
716 trophic links in a complex insect host-parasitoid food web. *Molecular Ecology*
717 *Resources*, 11(5), 786–794. <https://doi.org/10.1111/j.1755-0998.2011.03016.x>
- 718 Huffaker, C. B., Kennett, C. E., Finney, G. L. (1962). Biological control of olive scale,
719 *Pwrlatoria oleae* (Cohree), in California by imported *Aphytis maculicornis* (Masi)
720 (Hymenoptera: Aphelinidae). *Hilgardia*, 32 (13): 541-636. DOI:
721 10.3733/hilg.v32n13p541
- 722 Hughes, R. D., Woolcock, L. T. (1976). *Aphaereta aotea* sp. N. (Hymenoptera:
723 Braconidae), an Alysiine parasite of dung breeding flies. *Journal of Australian*
724 *Entomological Society*, 15, 191-196.
- 725 Jeffs, C. T., Terry, J. C. D., Higgie, M., Jandová, A., Konvičková, H., Brown, J. J., Lue, C.-H.,
726 ... Lewis, O. T. (2020). Molecular analyses reveal consistent food web structure with
727 elevation in rainforest *Drosophila* - parasitoid communities. *Ecography*, 43, 1-11,
728 <https://doi.org/10.1111/ecog.05390>
- 729 Kim-Jo, C., Gatti, J. L., & Poirié, M. (2019). *Drosophila* cellular immunity against
730 parasitoid wasps: A complex and time-dependent process. *In Frontiers in*
731 *Physiology*, <https://doi.org/10.3389/fphys.2019.00603>
- 732 Kimura, M. T. (2015). Prevalence of exotic frugivorous *Drosophila* species, *D. simulans*
733 and *D. immigrans* (Diptera: Drosophilidae), and its effects on local parasitoids in
734 Sapporo, northern Japan. *Applied Entomology and Zoology*, 50(4), 509–515.
735 <https://doi.org/10.1007/s13355-015-0361-8>
- 736 Kimura, M. T., & Mitsui, H. (2020). *Drosophila* parasitoids (Hymenoptera) of Japan. *In*
737 *Entomological Science*, 23(4), 359-368, <https://doi.org/10.1111/ens.12432>
- 738 Klopstein, S., Kropf, C., & Baur, H. (2016). Wolbachia endosymbionts distort DNA
739 barcoding in the parasitoid wasp genus *Diplazon* (Hymenoptera: Ichneumonidae).
740 *Zoological Journal of the Linnean Society*, 177(3), 541–557.
741 <https://doi.org/10.1111/zoj.12380>
- 742 Kraaijeveld, A. R., & Godfray, H. C. J. (1997). Trade-off between parasitoid resistance and
743 larval competitive. *Nature*, 389, 278-280, <https://doi.org/10.1038/38483>
- 744 Lefèvre, T., De Roode, J. C., Kacsoh, B. Z., & Schlenke, T. A. (2012). Defence strategies
745 against a parasitoid wasp in *Drosophila*: Fight or flight? *Biology Letters*, 8(2), 230-
746 233, <https://doi.org/10.1098/rsbl.2011.0725>

- 747 Lendemer, J., Thiers, B., Monfils, A. K., Zaspel, J., Ellwood, E. R., Bentley, A., LeVan, K., ...
748 Aime, M. C. (2020). The extended specimen network: A strategy to enhance US
749 biodiversity collections, promote research and education. *BioScience*, 70(1), 23-30,
750 <https://doi.org/10.1093/biosci/biz140>
- 751 Lin, C. P., & Danforth, B. N. (2004). How do insect nuclear and mitochondrial gene
752 substitution patterns differ? Insights from Bayesian analyses of combined datasets.
753 *Molecular Phylogenetics and Evolution*, 30(3), 686-702,
754 [https://doi.org/10.1016/S1055-7903\(03\)00241-0](https://doi.org/10.1016/S1055-7903(03)00241-0)
- 755 Littlefair, J. E., Clare, E. L., & Naaum, A. (2016). Barcoding the food chain: From Sanger to
756 high-throughput sequencing1. *Genome*, 59(11), 946–958.
757 <https://doi.org/10.1139/gen-2016-0028>
- 758 Liu, L. J., Pang, A. H., Feng, S. Q., Cui, B. Y., Zhao, Z. H., Kučerová, Z., Stejskal, V., ... Li, Z.
759 H. (2017). Molecular Identification of ten species of stored-product psocids through
760 microarray method based on ITS2 rDNA. *Scientific Reports*, 7(1): 16694,
761 <https://doi.org/10.1038/s41598-017-16888-z>
- 762 Lue, C.-H., Borowy, D., Buffington, M. L., & Leips, J. (2018). Geographic and seasonal
763 variation in species diversity and community composition of frugivorous *Drosophila*
764 (Diptera: Drosophilidae) and their *Leptopilina* (Hymenoptera: Figitidae) parasitoids.
765 *Environmental Entomology*, 47(5): 1096-1106. <https://doi.org/10.1093/ee/nvy114>
- 766 Lue, C.-H., Driskell, A. C., Leips, J., & Buffington, M. L. (2016). Review of the genus
767 *Leptopilina* (Hymenoptera, Cynipoidea, Figitidae, Eucoilinae) from the Eastern
768 United States, including three newly described species. *Journal of Hymenoptera*
769 *Research*, 53: 35-76. <https://doi.org/10.3897/jhr.53.10369>
- 770 Meiklejohn, K. A., Damaso, N., & Robertson, J. M. (2019). Assessment of BOLD and
771 GenBank – Their accuracy and reliability for the identification of biological
772 materials. *PLoS ONE*, 14(6): e0217084.
773 <https://doi.org/10.1371/journal.pone.0217084>
- 774 Melk, J. P., & Govind, S. (1999). Developmental analysis of *Ganaspis xanthopoda*, a larval
775 parasitoid of *Drosophila melanogaster*. *Journal of Experimental Biology*, 202, 1885-
776 1896
- 777 Miller, S. E., Hausmann, A., Hallwachs, W., & Janzen, D. H. (2016). Advancing taxonomy
778 and bioinventories with DNA barcodes. *Philosophical Transactions of the Royal*
779 *Society Biological Sciences*, 371(1702): 20150339. doi: 10.1098/rstb.2015.0339
- 780 Mitsui, H., van Achterberg, K., Nordlander, G., & Kimura, M. T. (2007). Geographical
781 distributions and host associations of larval parasitoids of frugivorous *Drosophilidae*
782 in Japan. *Journal of Natural History*, 41(25–28), 1731–1738.
783 <https://doi.org/10.1080/00222930701504797>
- 784 Moreau, S. J. M., Vinchon, S., Cherqui, A., & Prévost, G. (2009). Components of Asobara
785 venoms and their effects on hosts. In *Advances in Parasitology*, Prévost G (Ed). 70,
786 217-232, [https://doi.org/10.1016/S0065-308X\(09\)70008-9](https://doi.org/10.1016/S0065-308X(09)70008-9)
- 787 Morris, R. J., Gripenberg, S., Lewis, O. T., & Roslin, T. (2014). Antagonistic interaction
788 networks are structured independently of latitude and host guild. *Ecology Letters*,
789 17(3), 340-349, <https://doi.org/10.1111/ele.12235>
- 790 Nappi, A. J., & Carton, Y. (2001). Immunogenetic aspects of the cellular immune

response of *Drosophila* against parasitoids. *Immunogenetics*, 52(3–4), 157–164.
<https://doi.org/10.1007/s002510000272>

Nomano, F. Y., Kasuya, N., Matsuura, A., Suwito, A., Mitsui, H., Buffington, M. L., & Kimura, M. T. (2017). Genetic differentiation of *Ganaspis brasiliensis* (Hymenoptera: Figitidae) from East and Southeast Asia. *Applied Entomology and Zoology*, 52(3), 429–437. <https://doi.org/10.1007/s13355-017-0493-0>

Novotny, V., & Miller, S. E. (2014). Mapping and understanding the diversity of insects in the tropics: Past achievements and future directions. *Austral Entomology*, 53(3), 259–267. <https://doi.org/10.1111/aen.12111>

O’Grady, P. M., & DeSalle, R. (2018). Phylogeny of the genus *Drosophila*. *Genetics*, 209(1), 1–25. <https://doi.org/10.1534/genetics.117.300583>

Pentinsaari, M., Ratnasingham, S., Miller, S. E., & Hebert, P. D. N. (2020). BOLD and GenBank revisited – Do identification errors arise in the lab or in the sequence libraries? *PLoS One*, 15(4): e0231814. <https://doi.org/10.1371/journal.pone.0231814>

Pfeiffer, D. G., Shrader, M. E., Wahls, J. C. E., Willbrand, B. N., Sandum, I., van der Linde, K., Laub, C. A., ... Day, E. R. (2019). African Fig Fly (Diptera: Drosophilidae): Biology, expansion of geographic range, and its potential status as a soft fruit pest. *Journal of Integrated Pest Management*, 10(1), 1–8. <https://doi.org/10.1093/jipm/pmz018>

Prendini, L., Hanner, R., & DeSalle, R. (2002). Obtaining, storing and archiving specimens and tissue samples for use in molecular studies. In *Techniques in Molecular Systematics and Evolution*. https://doi.org/10.1007/978-3-0348-8125-8_11

Prévost, G. (2009). Parasitoids of *Drosophila*. In *Advances in parasitology*. [https://doi.org/10.1016/S0065-308X\(09\)70018-1](https://doi.org/10.1016/S0065-308X(09)70018-1)

Quicke, D. L. J. (2015). The Braconid and Ichneumonid parasitoid wasps: Biology, systematics, evolution and ecology. Wiley-Blackwell, <https://doi.org/10.1002/9781118907085>

Ratnasingham, S., & Hebert, P. D. N. (2007). BARCODING: bold: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Molecular Ecology Notes*, 7(3), 355–364, <https://doi.org/10.1111/j.1471-8286.2007.01678.x>

Ratnasingham, S., & Hebert, P. D. N. (2013). A DNA-based registry for all animal species: the barcode index number (BIN) system. *PLoS ONE*, 8(7): e66213. <https://doi.org/10.1371/journal.pone.0066213>

Raupach, M. J., Hendrich, L., Kuchler, S. M., Deister, F., Moriniere, J., & Gossner, M. M. (2014). Building-Up of a DNA Barcode Library for true bugs (Insecta: Hemiptera: Heteroptera) of Germany reveals taxonomic uncertainties and surprises. *PLoS ONE*, 9(9), 1–13. <https://doi.org/10.1371/journal.pone.0106940>

Remsen, J., & O’Grady, P. (2002). Phylogeny of Drosophilinae (Diptera: Drosophilidae), with comments on combined analysis and character support. *Molecular Phylogenetics and Evolution*, 24(2), 249–264, [https://doi.org/10.1016/S1055-7903\(02\)00226-9](https://doi.org/10.1016/S1055-7903(02)00226-9)

Reumer, B. M., van Alphen, J. J. M., & Kraaijeveld, K. (2012). Occasional males in parthenogenetic populations of *Asobara japonica* (Hymenoptera: Braconidae): Low *Wolbachia* titer or incomplete coadaptation. *Heredity*, 108(3), 341–346,

835 <https://doi.org/10.1038/hdy.2011.82>

836 Rosen, D. (1986). The role of taxonomy in effective biological control programs.

837 *Agriculture, Ecosystems & Environment*, 15(2-3), 121-129.

838 [https://doi.org/10.1016/0167-8809\(86\)90085-X](https://doi.org/10.1016/0167-8809(86)90085-X)

839 Santos, W. G. N., Fernandes, E. C., Souza, M. M., Guimarães, J. A., & Araujo, E. L. (2016).

840 First record of Eucoilinae (Hymenoptera: Figitidae), parasitoids of African fig fly

841 *Zaprionus indianus* Gupta (Diptera: Drosophilidae), in the Caatinga biome.

842 *Semina: Ciencias Agrarias*, 37(5), 3055–3058. [https://doi.org/10.5433/1679-](https://doi.org/10.5433/1679-0359.2016v37n5p3055)

843 [0359.2016v37n5p3055](https://doi.org/10.5433/1679-0359.2016v37n5p3055)

844 Schilthuizen, M., Vairappan, C. S., Slade, E. M., Mann, D. J., & Miller, J. A. (2015).

845 Specimens as primary data: Museums and “open science.” *Trends in Ecology and*

846 *Evolution*, 30(5), 237–238. <https://doi.org/10.1016/j.tree.2015.03.002>

847 Schlenke, T. A., Morales, J., Govind, S., & Clark, A. G. (2007). Contrasting infection

848 strategies in generalist and specialist wasp parasitoids of *Drosophila melanogaster*.

849 *PLoS Pathogens*, 3(10):e158, <https://doi.org/10.1371/journal.ppat.0030158>

850 Schindel, D., & Miller, S. E. (2010). Provisional Nomenclature the on-ramp to taxonomic

851 names. In: Polaszek, A. (Ed), *Systema Nature*, 250: The Linnaean Ark. CRC, Boca

852 Raton, 109-115.

853 Seehausen, M. L., Ris, N., Driss, L., Racca, A., Girod, P., Warot, S., Borowiec, N., Tosevski,

854 I., & Kenis, M. (2020). Evidence for a cryptic parasitoid species reveals its suitability

855 as a biological control agent. *Scientific reports*, 10: 19096.

856 <https://doi.org/10.1038/s41598-020-76180-5>

857 Shokralla, S., Spall, J. L., Gibson, J. F., & Hajibabaei, M. (2012). Next-generation

858 sequencing technologies for environmental DNA research. *In Molecular Ecology*,

859 21(8), 1794-1805, <https://doi.org/10.1111/j.1365-294X.2012.05538.x>

860 Smith, M. A., Fisher, B. L., & Hebert, P. D. N. (2005). DNA barcoding for effective

861 biodiversity assessment of a hyperdiverse arthropod group: The ants of

862 Madagascar. *Philosophical Transactions of the Royal Society Biological Sciences*,

863 360(1462), 1825-1834, <https://doi.org/10.1098/rstb.2005.1714>

864 Smith, M. A., Fernandez-Triana, J. L., Eveleigh, E., Gomez, J., Guclu, C., Hallwachs, W.,

865 Hebert, P. D. N., ... Zaldivar-Riveron, A. (2013). DNA barcoding and the taxonomy of

866 Microgastrinae wasps (Hymenoptera, Braconidae): impacts after 8 years and nearly

867 20000 sequences. *Molecular Ecology Resources*, 13, 168-276,

868 <https://doi.org/10.1111/1755-0988.12038>

869 Staniczenko, P. P. A., Reed-Tsochas, F., Lewis, O. T., Tylianakis, J. M., Albrecht, M.,

870 Coudrain, V., & Klein, A. M. (2017). Predicting the effect of habitat modification on

871 networks of interacting species. *Nature Communications*, 8, 792,

872 <https://doi.org/10.1038/s41467-017-00913-w>

873 Tarli, V. D., Grandcolas, P., & Pellens, R. (2018). The informative value of museum

874 collections for ecology and conservation: A comparison with target sampling in the

875 Brazilian Atlantic forest. *PLoS ONE*, 13(11).

876 <https://doi.org/10.1371/journal.pone.0205710>

877 Tedersoo, L., Drenkhan, R., Anslan, S., Morales-Rodriguez, C., & Cleary, M. (2019). High-

878 throughput identification and diagnostics of pathogens and pests: Overview and

practical recommendations. *Molecular Ecology Resources*, 19(1), 47–76.
<https://doi.org/10.1111/1755-0998.12959>

Terry, J. C. D., Chen, J., & Lewis, O. T. (2020). The effect of natural enemies on the coexistence of competing species - an empirical test using Bayesian modern coexistence theory. *bioRxiv*: <https://doi.org/10.1101/2020.08.27.270389>

Thierry, M., Pardikes, N. A., Lue, C.-H., Lewis, O. L., & Hrcek, J. (2021). Experimental warming influences species abundances in a *Drosophila* host community through direct effects on species performance rather than altered competition and parasitism. *PLOS ONE* (In Press).

Troudet, J., Vignes-Lebbe, R., Grandcolas, P., & Legendre, F. (2018). The increasing disconnection of primary biodiversity data from specimens: How does it happen and how to handle it? *Systematic Biology*, 67(6), 1110–1119.
<https://doi.org/10.1093/sysbio/syy044>

Tsai, C.-L., Chu, I.-H., Chou, M.-H., Chareonviriyaphap, T., Chiang, M.-Y., Lin, P.-A., Lu, K.-H., & Yeh, W.-B. (2020). Rapid identification of the invasive fall armyworm *Spodoptera frugiperda* (Lepidoptera, Noctuidae) using species-specific primers in multiplex PCR. *Scientific Reports*, 10(1), 1–9. <https://doi.org/10.1038/s41598-020-73786-7>

Tylianakis, J. M., Tscharntke, T., & Lewis, O. T. (2007). Habitat modification alters the structure of tropical host-parasitoid food webs. *Nature*, 445(7124), 202–205.
<https://doi.org/10.1038/nature05429>

Wachi, N., Nomano, F. Y., Mitsui, H., Kasuya, N., & Kimura, M. T. (2015). Taxonomy and evolution of putative thelytokous species of Leptopilina (Hymenoptera: Figitidae) from Japan, with description of two new species. *Entomological Science*, 18(1), 41–54. <https://doi.org/10.1111/ens.12089>

Wang, X.-G., Biondi, A., & Daane, K. M. (2020). Functional responses of three candidate Asian larval parasitoids evaluated for classical biological control of *Drosophila suzukii*. *Journal of Economic Entomology*, 113(1): 73–80. doi: 10.1093/jee/toz265

Wang, X.-G., Biondi, A., Nance, A. N., Zappalà, L., Hoelmer, K. A., & Daane, K. M. (2021). Assessment of *Asobara japonica* as a potential biological control agent for the spotted wing drosophila, *Drosophila suzukii*. *Entomologia Generalis* (In Press) doi: 10.1127/entomologia/2020/1100

Wang, X.-G., Lee, J., Daane, K.M., Buffington, M., & Hoelmer, K. A. (2020). Biological control of *Drosophila suzukii*. *CAB Reviews* 54, 10.1079/PAVSNNR202015054

Xiao, J. H., Wang, N. X., Murphy, R. W., Cook, J., Jia, L. Y., & Huang, D. W. (2012). Wolbachia infection and dramatic intraspecific mitochondrial DNA divergence in a fig wasp. *Evolution*, 66, 1907–1916, <https://doi.org/10.1111/j.1558-5646.2011.01561.x>

Xie, J., Tiner, B., Vilchez, I., & Mateos, M. (2011). Effect of the *Drosophila* endosymbiont *Spiroplasma* on parasitoid wasp development and on the reproductive fitness of wasp-attacked fly survivors. *Evolutionary Ecology*, 25, 1065–1079, <https://doi.org/10.1007/s10682-010-9453-7>

Xie, J., Winter, C., Winter, L., & Mateos, M. (2015). Rapid spread of the defensive endosymbiont *Spiroplasma* in *Drosophila hydei* under high parasitoid wasp

pressure. *FEMS Microbiology Ecology*, 91(2), 1-11,
<https://doi.org/10.1093/femsec/iu017>
Yassin, A. (2013). Phylogenetic classification of the Drosophilidae Rondani (Diptera): The
role of morphology in the postgenomic era. *Systematic Entomology*,
<https://doi.org/10.1111/j.1365-3113.2012.00665.x>
Yassin, A., & David, J. R. (2010). Revision of the Afrotropical species of *Zaprionus*
(Diptera, Drosophilidae), with descriptions of two new species and notes on
internal reproductive structures and immature stages. *ZooKeys*, 51, 33-72,
<https://doi.org/10.3897/zookeys.51.380>
Ye, Z., Vollhardt, I. M. G., Girtler, S., Wallinger, C., Tomanovic, Z., & Traugott, M. (2017).
An effective molecular approach for assessing cereal aphid-parasitoid-
endosymbiont networks. *Scientific Reports*, 7(1), 1–12.
<https://doi.org/10.1038/s41598-017-02226-w>

Data Accessibility

The DROP database is freely accessible at Zenodo depository
(<http://doi.org/10.5281/zenodo.4519656>). New sequences have been deposited in
BOLD in datasets DS-LABS, and DS-AUSPTOID. [NOTE: the doi records will be update at
revision stage]

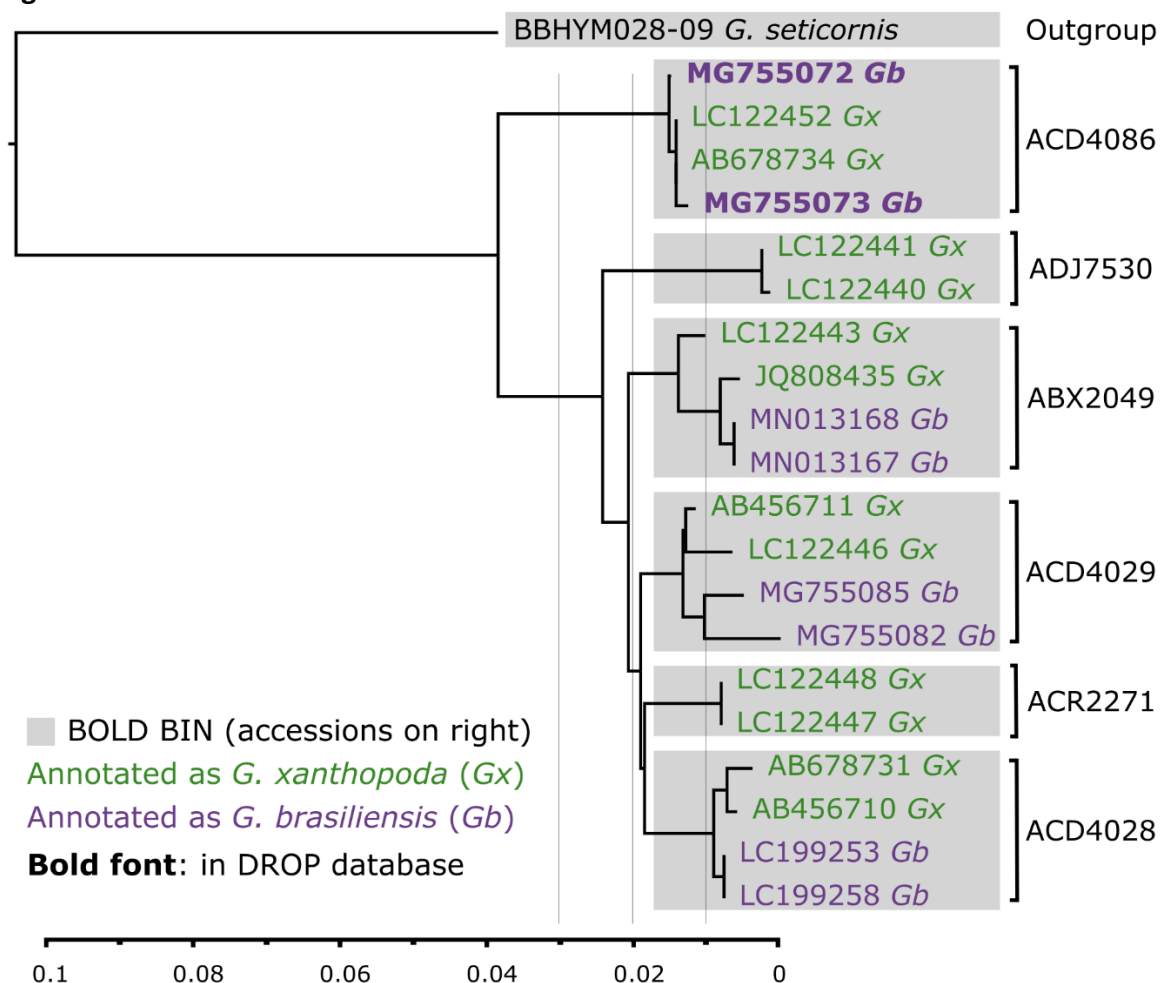
Author Contributions

The initial project idea was originated by C-HL, MLB, JH, MM, TS, JV, SG, and
PPAS. Molecular work was conducted by C-HL, SS, ML, AJ, and AD. BOLD and GenBank
data was harvested by TAE and C-HL. Figures were made by AL and C-HL. Laboratory and
field sample preparations were conducted by MTK, YC, TS, MM, SG, JV, EG, MG, XW,
KM, KMD, PA, NAP, MT, JJB, MP, FMJ, WDT, JSD, BW, OTL, PPAS, JL and AL. Taxonomic
concepts and interpretations were conducted by RRK, MLB, CH-L, PG, and SEM. DROP

951 database was built by JH and C-HL. All authors contributed to review and final revisions
952 of the manuscript, which was written primarily by C-HL, MLB and JH.
953

954

Figures:



955

956 Figure 1: An example of difficulties of molecular identification demonstrated on
 957 *Ganaspis xanthopoda* and *G. brasiliensis*. Only two sequences (in bold text) can be
 958 reliably used for identification and are included in DROP database. To select the
 959 sequences, we searched the BINs associated with the organism's name "*Ganaspis*
 960 *xanthopoda*" (green) or "*Ganaspis brasiliensis*" (purple) in BOLD. From each BIN, two
 961 sequences from each species were selected to build a neighbor-joining tree (bottom axis
 962 indicated % genetic divergence). There was a total of 6 BINs (gray boxes) in this
 963 sequence complex. Of these, 4 BINs contained both species names and without
 964 examination of vouchers identification would be impossible. In DROP, vouchers from
 965 two sequences, **MG755073** and **MG755072**, were deposited in CNR-IPSP (Table S2),
 966 examined by taxonomists and identify as *G. brasiliensis*. These two COI sequences can
 967 now be used to reliably identify *G. brasiliensis*. For *G. xanthopoda*, there were no
 968 available vouchers or reliable sequences that passed DROP standards to use for
 969 identification. Species delimitation between *G. brasiliensis* and *G. xanthopoda* is
 970 convoluted (see discussion), varies according to arbitrary % genetic divergence (gray
 971 vertical lines), and needs future an integrative taxonomic revision.

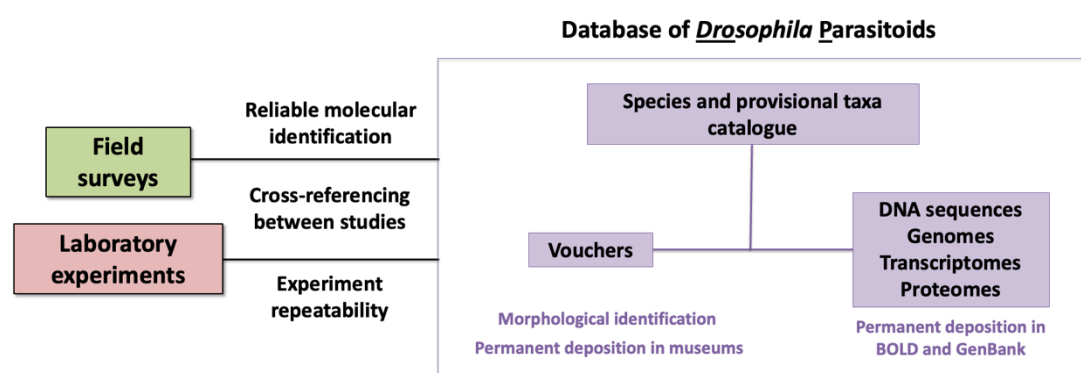


Figure 2: Concept of a centralized, vetted, integrated database for *Drosophila* Parasitoids (DROP) we developed. First, we provide a species and provisional species catalog with correct taxonomy. Second, to provide a reliable genetic reference library, genetic data (DNA sequences, genomes, transcriptomes, proteomes) link to a voucher connected to the species catalog. Third, we link the two primary sources of data (field surveys and laboratory experiments) by requiring a permanent deposition of vouchers and sequences in order to be included in DROP.

Tables:

Table 1: List of species and provisional species included in DROP. For additional taxonomic details, see DROP.

Superfamily	Family	Genus	Species_Name	Author
Chalcidoidea	Encyrtidae		<i>drop_Cha2_sp12</i>	
Chalcidoidea	Encyrtidae	<i>Tachinaephagus</i>	<i>drop_IR1_sp41</i>	Kimura
Chalcidoidea	Encyrtidae	<i>Tachinaephagus</i>	<i>drop_BG1_sp42</i>	Kimura
Chalcidoidea	Encyrtidae	<i>Tachinaephagus</i>	<i>zealandicus</i>	Ashmead 1904
Chalcidoidea	Pteromalidae		<i>drop_Pte69_sp11</i>	
Chalcidoidea	Pteromalidae	<i>Pachycrepoideus</i>	<i>vindemmiae</i>	(Rondani, 1875)
Chalcidoidea	Pteromalidae	<i>Spalangia</i>	<i>drop_IR1_sp38</i>	Kimura
Chalcidoidea	Pteromalidae	<i>Spalangia</i>	<i>drop_NG1_sp39</i>	Kimura
Chalcidoidea	Pteromalidae	<i>Spalangia</i>	<i>drop_SK1_sp40</i>	Kimura
Chalcidoidea	Pteromalidae	<i>Spalangia</i>	<i>drosophilae</i>	Ashmead 1887
Chalcidoidea	Pteromalidae	<i>Spalangia</i>	<i>erythromera</i>	Foerster 1850
Chalcidoidea	Pteromalidae	<i>Trichomalopsis</i>	<i>dubia</i>	(Ashmead, 1896)
Chalcidoidea	Pteromalidae	<i>Trichomalopsis</i>	<i>microptera</i>	(Lindeman, 1887)
Chalcidoidea	Pteromalidae	<i>Trichomalopsis</i>	<i>nigricola</i>	Boucek

Chalcidoidea	Pteromalidae	<i>Trichomalopsis</i>	<i>sarcophagae</i>	(Gahan, 1914)
Chalcidoidea	Pteromalidae	<i>Vrestovia</i>	<i>brevior</i>	Boucek 1993
Chalcidoidea	Pteromalidae	<i>Vrestovia</i>	<i>fidenas</i>	(Walker, 1848)
Chalcidoidea	Pteromalidae		<i>drop_PacAtl_sp46</i>	
Chalcidoidea	Pteromalidae		<i>drop_PachyPort_sp45</i>	
Chalcidoidea			<i>drop_CH_sp64</i>	
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>brasiliensis</i>	(Ihering, 1905)
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>drop_Gan_sp51</i>	
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>drop_Gan_sp52</i>	
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>drop_Gan_sp53</i>	
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>drop_Gsp1_sp67</i>	
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>drop_Gsp2_sp68</i>	
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>drop_Gsp50_sp66</i>	
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>drop_IR1_sp25</i>	Kimura
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>drop_IR2_sp26</i>	Kimura
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>drop_Gan1_sp1</i>	
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>drop_TK1_sp27</i>	Kimura
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>hookeri</i>	Crawford 1913
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>mahensis</i>	Kieffer 1911
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>mellipes</i>	(Say, 1826)
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>mundata</i>	Forster 1869
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>seticornis</i>	(Hellen, 1960)
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>tenuicornis</i>	Kieffer 1904
Cynipoidea	Figitidae	<i>Ganaspis</i>	<i>xanthopoda</i>	(Ashmead, 1896)
Cynipoidea	Figitidae	<i>Kleidotoma</i>	<i>bicolor</i>	(Giraud, 1860)
Cynipoidea	Figitidae	<i>Kleidotoma</i>	<i>dolichocera</i>	Thomson 1877
Cynipoidea	Figitidae	<i>Kleidotoma</i>	<i>drop_TK1_sp28</i>	Kimura
Cynipoidea	Figitidae	<i>Kleidotoma</i>	<i>filicornis</i>	(Cameron, 1889)
Cynipoidea	Figitidae	<i>Kleidotoma</i>	<i>icarus</i>	(Quinlan, 1964)
Cynipoidea	Figitidae	<i>Kleidotoma</i>	<i>psiloides</i>	Westwood 1833
Cynipoidea	Figitidae	<i>Kleidotoma</i>	<i>tetratoma</i>	(Hartig, 1841)
Cynipoidea	Figitidae	<i>Leptolamina</i>	<i>drop_Fig64_sp5</i>	
Cynipoidea	Figitidae	<i>Leptolamina</i>	<i>drop_Lmn_sp6</i>	
Cynipoidea	Figitidae	<i>Leptolamina</i>	<i>drop_TK1_sp29</i>	Kimura
Cynipoidea	Figitidae	<i>Leptolamina</i>	<i>gressitti</i>	Yoshimoto & Yasumatsu 1965
Cynipoidea	Figitidae	<i>Leptolamina</i>	<i>papuensis</i>	Yoshimoto 1963
Cynipoidea	Figitidae	<i>Leptolamina</i>	<i>ponapensis</i>	Yoshimoto 1962
Cynipoidea	Figitidae	<i>Leptolamina</i>	<i>seychellensis</i>	(Kieffer, 1911)

Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>atraticeps</i>	(Kieffer, 1911)
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>australis</i>	(Belizin, 1966)
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>boulardi</i>	(Barbotin, Carton & Kelner-Pillault, 1979)
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>clavipes</i>	(Hartig, 1841)
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>cupulifera</i>	(Kieffer, 1916)
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>decemflagella</i>	Lue & Buffington 2017
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_Lep_sp54</i>	
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_Lep_sp55</i>	
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_Lep_sp56</i>	
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_Lep_sp57</i>	
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_Lep_sp58</i>	
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_Lep_sp59</i>	
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_Lep_sp60</i>	
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_Lep_sp61</i>	
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_Lep_sp62</i>	
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_BG1_sp34</i>	Kimura
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_Fig059_sp4</i>	
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_Fig124_sp2</i>	
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_Fig58_sp3</i>	
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_IR1_sp30</i>	Kimura
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_NG1_sp33</i>	Kimura
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_SK1_sp35</i>	Kimura
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_STL_sp7</i>	
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_TK2_sp31</i>	Kimura
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>drop_TK3_sp32</i>	Kimura
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>fimbriata</i>	(Kieffer, 1901)
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>freyae</i>	Allemand & Nordlander 2002
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>guineaensis</i>	Allemand & Nordlander 2002
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>heterotoma</i>	(Thomson, 1862)
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>japonica japonica</i>	Novkovic & Kimura 2011
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>lasallei</i>	Buffington & Guerrieri 2020
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>leipsi</i>	Lue & Buffington 2018
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>lonchaeae</i>	(Cameron, 1912)
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>longipes</i>	(Hartig, 1841)
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>mahensis</i>	(Kieffer, 1911)
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>maia</i>	Lue & Buffington 2016
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>maria</i>	(Girault, 1930)
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>orientalis</i>	Allemand & Nordlander 2002

Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>pacifica</i>	Novkovic & Kimura 2011
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>rufipes</i>	(Cameron, 1908)
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>rugipunctata</i>	(Yoshimoto, 1962)
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>ryukyuensis</i>	Novkovic & Kimura 2011
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>tokioensis</i>	Wachi & Kimura 2015
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>tsushimaensis</i>	Wachi & Kimura 2015
Cynipoidea	Figitidae	<i>Leptopilina</i>	<i>victoriae</i>	Nordlander 1980
Cynipoidea	Figitidae	<i>Rhoptromeris</i>	<i>heptoma</i>	(Hartig, 1840)
Cynipoidea	Figitidae	<i>Rhoptromeris</i>	<i>nigriventris</i>	Nordlander 1978
Cynipoidea	Figitidae	<i>Rhoptromeris</i>	<i>rufiventris</i>	(Giraud, 1860)
Cynipoidea	Figitidae	<i>Rhoptromeris</i>	<i>villosa</i>	(Hartig, 1840)
Cynipoidea	Figitidae		<i>drop_Lg500_sp43</i>	
Ichneumonoidea	Braconidae	<i>Alysia</i>	<i>drop_SP1_sp24</i>	Kimura
Ichneumonoidea	Braconidae	<i>Aphaereta</i>	<i>aotea</i>	Hughes & Woolcock 1976
Ichneumonoidea	Braconidae	<i>Aphaereta</i>	<i>drop_SP1_sp15</i>	Kimura
Ichneumonoidea	Braconidae	<i>Aphaereta</i>	<i>drop_TK1_sp13</i>	Kimura
Ichneumonoidea	Braconidae	<i>Aphaereta</i>	<i>drop_TM1_sp14</i>	Kimura
Ichneumonoidea	Braconidae	<i>Aphaereta</i>	<i>minuta</i>	(Nees, 1811)
Ichneumonoidea	Braconidae	<i>Aphaereta</i>	<i>pallipes</i>	(Say, 1829)
Ichneumonoidea	Braconidae	<i>Aphaereta</i>	<i>scaptomyzae</i>	Fischer 1966
Ichneumonoidea	Braconidae	<i>Areotetes</i>	<i>striatiferus</i>	Li & van Achterberg 2013
Ichneumonoidea	Braconidae	<i>Areotetes</i>	<i>carinuliferus</i>	Li & van Achterberg 2013
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>ajbelli</i>	Berry 2007
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>albiclava</i>	Berry 2007
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>antipoda</i>	(Ashmead, 1900)
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>bactrocerae</i>	(Gahan, 1952)
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>brevicauda</i>	van Achterberg & Guerrieri 2016
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>citri</i>	(Fischer, 1963)
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>drop_KG1_sp16</i>	Kimura
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>drop_NG1_sp17</i>	Kimura
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>drop_SK2_sp20</i>	Kimura
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>drop_SP1_sp18</i>	Kimura
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>drop_Sp2_sp19</i>	Kimura
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>elongata</i>	van Achterberg & Guerrieri 2016
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>gahani</i>	(Papp, 1969)
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>japonica</i>	Belokobylskij 1998
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>kenyaensis</i>	Peris-Felipo 2014
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>leveri</i>	(Nixon, 1939)
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>mesocauda</i>	van Achterberg & Guerrieri 2016

Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>orientalis</i>	Viereck 1913
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>persimilis</i>	(Prince, 1976)
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>pleuralis</i>	(Ashmead, 1905)
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>rossica</i>	Belokobylskij 1998
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>rufescens</i>	(Förster, 1862)
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>tabida</i>	(Nees, 1834)
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>triangulata</i>	van Achterberg & Guerrieri 2016
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>turneri</i>	Peris-Felipo 2014
Ichneumonoidea	Braconidae	<i>Asobara</i>	<i>unicolorata</i>	van Achterberg & Guerrieri 2016
Ichneumonoidea	Braconidae	<i>Aspilota</i>	<i>albertica</i>	Berry 2007
Ichneumonoidea	Braconidae	<i>Aspilota</i>	<i>andyaustini</i>	Wharton 2002
Ichneumonoidea	Braconidae	<i>Aspilota</i>	<i>angusta</i>	Berry 2007
Ichneumonoidea	Braconidae	<i>Aspilota</i>	<i>concolor</i>	Nees 1812
Ichneumonoidea	Braconidae	<i>Aspilota</i>	<i>parecur</i>	Berry 2007
Ichneumonoidea	Braconidae	<i>Aspilota</i>	<i>villosa</i>	Berry 2007
Ichneumonoidea	Braconidae	<i>Dinotrema</i>	<i>barrattae</i>	Berry 2007
Ichneumonoidea	Braconidae	<i>Dinotrema</i>	<i>longworthi</i>	Berry 2007
Ichneumonoidea	Braconidae	<i>Dinotrema</i>	<i>philipi</i>	Berry 2007
Ichneumonoidea	Braconidae		<i>drop_Aso_sp8</i>	
Ichneumonoidea	Braconidae	<i>Opiognathus</i>	<i>pactus</i>	(Haliday, 1837)
Ichneumonoidea	Braconidae	<i>Opius</i>	<i>bellus</i>	Gahan 1930
Ichneumonoidea	Braconidae	<i>Opius</i>	<i>cinerariae</i>	Fischer
Ichneumonoidea	Braconidae	<i>Opius</i>	<i>crenuliferus</i>	Li & van Achterberg 2013
Ichneumonoidea	Braconidae	<i>Opius</i>	<i>monilipalpis</i>	Li & van Achterberg 2013
Ichneumonoidea	Braconidae	<i>Opius</i>	<i>ocreatus</i>	(Papp)
Ichneumonoidea	Braconidae	<i>Opius</i>	<i>pallipes</i>	Wesmael 1835
Ichneumonoidea	Braconidae	<i>Opius</i>	<i>pteridiophilus</i>	Wharton & Austin 1990
Ichneumonoidea	Braconidae	<i>Opius</i>	<i>pterus</i>	Wharton & Austin 1990
Ichneumonoidea	Braconidae	<i>Opius</i>	<i>trimaculatus</i>	Spinola
Ichneumonoidea	Braconidae	<i>Opius</i>	<i>youi</i>	Li & van Achterberg 2013
Ichneumonoidea	Braconidae	<i>Phaenocarpa</i>	<i>conspurcator</i>	(Haliday, 1838)
Ichneumonoidea	Braconidae	<i>Phaenocarpa</i>	<i>drop_IR1_sp22</i>	Kimura
Ichneumonoidea	Braconidae	<i>Phaenocarpa</i>	<i>drop_TK1_sp21</i>	Kimura
Ichneumonoidea	Braconidae	<i>Phaenocarpa</i>	<i>tacita</i>	Stelfox 1941
Ichneumonoidea	Braconidae	<i>Phaenocarpa</i>	<i>drosophilae</i>	(Fischer 1975)
Ichneumonoidea	Braconidae	<i>Tanycarpa</i>	<i>bicolor</i>	(Nees, 1814)
Ichneumonoidea	Braconidae	<i>Tanycarpa</i>	<i>chors</i>	Belokobylskij 1998
Ichneumonoidea	Braconidae	<i>Tanycarpa</i>	<i>drop_NG1_sp23</i>	Kimura
Ichneumonoidea	Braconidae	<i>Tanycarpa</i>	<i>punctata</i>	van Achterberg 1976

Ichneumonoidea	Braconidae		<i>drop_Aly_sp47</i>	
Ichneumonoidea	Braconidae		<i>drop_Aly_sp48</i>	
Ichneumonoidea	Braconidae		<i>drop_Aly_sp49</i>	
Ichneumonoidea	Braconidae		<i>drop_Aly_sp50</i>	
Ichneumonoidea	Braconidae		<i>drop_Aly_sp63</i>	
Ichneumonoidea	Braconidae		<i>drop_Aso_sp69</i>	
Diaprioidea	Diapriidae	<i>Trichopria</i>	<i>anastrephae</i>	Costa Lima 1940
Diaprioidea	Diapriidae	<i>Trichopria</i>	<i>drop_BG1_sp37</i>	Kimura
Diaprioidea	Diapriidae	<i>Trichopria</i>	<i>drop_Dia70_sp65</i>	
Diaprioidea	Diapriidae	<i>Trichopria</i>	<i>drop_Tri_sp44</i>	
Diaprioidea	Diapriidae	<i>Trichopria</i>	<i>drop_Bdia_sp10</i>	
Diaprioidea	Diapriidae	<i>Trichopria</i>	<i>drop_Dia127_sp9</i>	
Diaprioidea	Diapriidae	<i>Trichopria</i>	<i>drop_TK1_sp36</i>	Kimura
Diaprioidea	Diapriidae	<i>Trichopria</i>	<i>drosophilae</i>	(Kieffer, 1912)
Diaprioidea	Diapriidae	<i>Trichopria</i>	<i>modesta</i>	(Ratzeburg, 1848)

Table 2.1: *Drosophila* parasitoid whole-genome sequences included in DROP. For additional details, see DROP.

Genus	Species_Name	Species_id	Genome_id	Voucher_id	GenBank_id
<i>Ganaspis</i>	<i>brasiliensis</i>	19	8	868	GCA_009823575.1
<i>Ganaspis</i>	<i>brasiliensis</i>	19	16	872	SRX8882993
<i>Ganaspis</i>	<i>brasiliensis</i>	19	17	871	SRX8882992
<i>Ganaspis</i>	<i>drop_Gsp1_sp67</i>	182	15	873	SRX8882994
<i>Ganaspis</i>	<i>drop_Gsp2_sp68</i>	183	14	874	SRX8882995
<i>Ganaspis</i>	<i>drop_Gsp50_sp66</i>	181	9	869	GCA_011057455.1
<i>Leptolamina</i>	<i>ponapensis</i>	48	13	875	SRX8883008
<i>Leptopilina</i>	<i>boulardi</i>	4	5	865	GCA_011634795.1
<i>Leptopilina</i>	<i>boulardi</i>	4	6	866	GCA_003121605.1
<i>Leptopilina</i>	<i>boulardi</i>	4	12	876	SRX8883009
<i>Leptopilina</i>	<i>clavipes</i>	5	7	867	GCA_001855655.1
<i>Leptopilina</i>	<i>heterotoma</i>	6	1	861	GCA_010016045.1
<i>Leptopilina</i>	<i>heterotoma</i>	6	2	862	GCA_009602685.1
<i>Leptopilina</i>	<i>heterotoma</i>	6	3	863	GCA_009026005.1
<i>Leptopilina</i>	<i>heterotoma</i>	6	4	864	GCA_009025955.1
<i>Leptopilina</i>	<i>japonica japonica</i>	13	11	877	SRX8883011

Table 2.2: *Drosophila* parasitoid transcriptome data included in DROP.

Genus	Species_Name	Strain_id	Transcriptome_id	Voucher_id	Genbank_id
<i>Leptopilina</i>	<i>boulardi</i>	126	2	858	2183568
<i>Leptopilina</i>	<i>boulardi</i>	127	3	859	2183567
<i>Leptopilina</i>	<i>boulardi</i>	151	8	882	15642271
<i>Leptopilina</i>	<i>boulardi</i>	151	9	883	15642270
<i>Leptopilina</i>	<i>heterotoma</i>	147	1	857	2183569
<i>Leptopilina</i>	<i>heterotoma</i>	152	5	884	2046288
<i>Leptopilina</i>	<i>heterotoma</i>	61	6	880	11581553
<i>Leptopilina</i>	<i>heterotoma</i>	60	7	881	11662592
<i>Leptopilina</i>	<i>boulardi</i>	11	10	908	GAJA00000000.1
<i>Leptopilina</i>	<i>heterotoma</i>	14	11	909	GAJC00000000.1
<i>Ganaspis</i>	<i>hookeri</i>	7	12	910	GAIW00000000.1

Table 2.3: *Drosophila* parasitoid proteomes data included in DROP

Genus	Species_Name	Strain_id	Proteomes_id	Voucher_id	Assession_id
<i>Leptopilina</i>	<i>heterotoma</i>	152	1	885	PRIDE: PXD005639
<i>Leptopilina</i>	<i>heterotoma</i>	61	2	886	PRIDE: PXD005632
<i>Leptopilina</i>	<i>boulardi</i>	27	3	911	Upon request to julien.varaldi@univ-lyon1.fr
<i>Leptopilina</i>	<i>boulardi</i>	11	4	912	PRIDE: PDX023836
<i>Leptopilina</i>	<i>heterotoma</i>	14	5	913	PRIDE: PDX023824
<i>Ganaspis</i>	<i>hookeri</i>	7	6	914	PRIDE: PDX023825