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

Resource article

Word count: 7850 excluding references

Authors

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

Matthew L. Buffington (MLB)

Sonja Scheffer (SS)

Matthew Lewis (ML)

Tyler A. Elliott (TAE)

Amelia R. I. Lindsey (AL)

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

Anna Jandova (AJ)

Masahito T. Kimura (MTK)

Yves Carton (YC)

Robert R. Kula (RRK)

Todd A. Schlenke (TAS)

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

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

Julien Varaldi (JV)

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

Massimo Giorgini (MG), (0000-0001-8670-0945)

Xingeng Wang (XW)

Kim Hoelmer (KH)

Kent M. Daane (KMD)

Paul K. Abram (PKA)

Nicholas A. Pardikes (NAP), (0000-0002-9175-4494)

Joel J. Brown (JJB), (0000-0002-3608-6745)

Melanie Thierry (MT)

Marylène Poirié (MP)

Paul Goldstein (PG), (0000-0002-1443-7030)

Scott E. Miller (SEM), (0000-0002-4138-1378)

W. Daniel Tracey (WDT)

Jeremy S. Davis (JSD), (0000-0002-5214-161X)

Francis M. Jiggins (FMJ)

Bregje Wertheim (BW)

Owen T. Lewis (OTL)

Jeff Leips (JL)

Phillip P. A. Staniczenko (PPAS)

Jan Hrcek (JH), (0000-0003-0711-6447), janhrcek@gmail.com

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

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

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

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

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

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

(MP) Université "Côte d'Azur", INRAE, CNRS. and Evolution and Specificity of Multitrophic Interactions (ESIM) Sophia Agrobiotech Institute, 400 Route des Chappes, BP 167, 06903 Sophia Antipolis, France

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

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

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

(JSD) Biology Department, University of Kentucky, 101 T. H. Morgan Building, Lexington, 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 (*Dro*sophila 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; Gariepy 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.

153 Although the BOLD database function “BOLD-IDS” allows considerable database
154 curation (e.g., sequences are used for identification and/or flagging of
155 misidentified/contaminated records), it also automatically includes sequences from
156 GenBank, and may perpetuate the shortcomings previously mentioned since these
157 cannot be curated from within BOLD. As such, the quality of sequences and the
158 reliability of identifications obtained from BOLD-IDS can vary, and depends on the
159 curation by systematists focusing on individual taxa (Meiklejohn et al., 2019). BOLD-IDS
160 works well for taxa where qualified taxonomists have been involved with assuring data
161 quality; some insect examples include beetles (Hendrich et al., 2015), butterflies
162 (Escalante et al., 2010), geometrid moths (Hausmann et al., 2011, 2016; Miller et al.,
163 2016), true bugs (Raupach et al., 2014), and microgastrine wasps (Smith et al., 2013).
164 Unfortunately, this is not the case of parasitoids (Insecta: Hymenoptera) of
165 *Drosophila* flies (Insecta: Drosophilidae). There are vast numbers of *Drosophila*
166 parasitoid sequences readily available in GenBank and BOLD, as these parasitoids and
167 their hosts are important model organisms in biology. As of this writing, there are
168 88,666 nucleotide sequences deposited in GenBank for *Leptopilina heterotoma*
169 (Thomson) and *L. boulardi* (Barbotin, Carton & Kelner-Pillault) (Hymenoptera: Figitidae)
170 alone. However, less than 1 % of the identifications associated with these sequences
171 have been confirmed by taxonomists or are associated with voucher specimens
172 deposited in museum collections. With sequencing shifting from individual genes to
173 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

196 is especially pressing given current efforts to use parasitoids in biological control efforts,
197 such as those of the invasive pest spotted wing *Drosophila*, *Drosophila suzukii* (Abram et
198 al., 2020; Daane et al., 2016; Giorgini et al., 2019; Wang et al., 2020 a&b).
199 Properly executed molecular identification has the potential to be much more
200 efficient for the majority of researchers, and many laboratory strains are commonly
201 identified using DNA sequences alone. While it is practical for researchers to assign
202 species names based on a match to sequence records in genetic databases, this practice
203 often causes a cascade of inaccuracies. To illustrate the extent of the problem, we note
204 the example of *Ganaspis,* a genus of parasitoids commonly used in laboratories that
205 includes both superficially indistinguishable species with highly divergent sequences
206 that are often treated as conspecific, as well as specimens with identical sequences
207 identified under different names (Figure 1).
208 To address these issues, we introduce a newly curated molecular reference database
209 for *Drosophila* parasitoids —DROP— in which sequences are either linked to voucher
210 specimens identified by taxonomists or have a traceable provenance (Figure 2). The first
211 aim of DROP is to provide a reliable DNA sequence library for molecular identification of
212 *Drosophila* parasitoids that enables cross-referencing of original taxonomic concepts
213 with those of subsequent studies. We pay special attention to live parasitoid strains
214 which are available for future experiments. The second aim is to standardize and
215 expedite the linkage between specimens and available sequence data; we place a
216 premium on museum vouchers as they allow for repeatable scientific research. In DROP,
217 this goal is facilitated through a consolidated digital infrastructure of data associated

218 with laboratory strains, offering the opportunity for researchers to re-examine past
219 experimental results in a permanent context. The third aim is to provide an up-to-date
220 catalogue of the diversity of *Drosophila* parasitoids as a foundation for advancing the
221 understanding of their taxonomy. Finally, the fourth aim of DROP is for our collaborative
222 effort to serve as an inspiration to communities of researchers studying other groups of
223 organisms who are experiencing difficulties with the reliability of molecular reference
224 databases.
225

226 **Materials and Methods**

227 ***Drosophila species and their parasitoids***

228 The phylogenetic and subgeneric structure within *Drosophila* and related genera is
229 not yet fully resolved (O'Grady & DeSalle, 2018). Various subgenera, including
230 *Scaptomyza*, *Zaprionus*, *Lordiphosa* and *Samoaia*, have been treated as both genera and
231 subgenera, and researchers have yet to achieve consensus on these various hypotheses
232 (O'Grady & DeSalle, 2018; Remsen & O'Grady, 2002; Yassin, 2013; Yassin & David,
233 2010). Species in *Drosophila* subgenera and genera closely related to *Drosophila*
234 commonly share niche space and natural histories and, as a result, are often attacked by
235 overlapping or identical groups of parasitoids. For instance, the invasive African fig fly,
236 *Zaprionus indianus* Gupta is attacked by *Pachycrepoideus vindemiae* (Rondani, 1875)
237 and *Leptopilina boulardi* (Pfeiffer et al., 2019; Santos et al., 2016), all of which have been
238 recorded from *Drosophila*. Therefore, we also include these groups within the contents
239 of DROP.

240 Parasitoids of *Drosophila* belong to four superfamilies of Hymenoptera
241 (Chalcidoidea, Cynipoidea, Ichneumonoidea, Diaprioidea) which evolved parasitism of
242 *Drosophila* flies independently (Carton et al., 1986; Prévost, 2009). All the parasitoids
243 known to attack *Drosophila* are solitary and attack either the larval or pupal stage; in
244 both cases, they emerge from the fly's puparium. The known *Drosophila* larval
245 parasitoids belong to two families, Braconidae (including the genera *Asobara,*
246 *Aphaereta, Phaenocarpa, Tanycarpa, Aspilota, Opius*) and Figitidae (*Leptopilina,*
247 *Ganaspis, Leptolamina, Kleidotoma*); all are koinobionts that allow the host to continue
248 development while the parasitoid grows within it. The known *Drosophila* pupal
249 parasitoids belong to three other families, Diapriidae (*Trichopria, Spilomicrus*),
250 Pteromalidae (*Pachycrepoideus, Spalangia, Trichomalopsis, Toxomorpha*) and Encytidae
251 (*Tachinaephagus*); they are all idiobionts that terminate host development immediately.
252 Host-specificity across the *Drosophila* parasitoids is poorly characterized—while some
253 can parasitize other families of Diptera (e.g., *Aphaereta aotea*) (Hughes & Woolcock,
254 1976), most are thought to be limited to *Drosophila* hosts.
255
256 ***Data sources***
257 To assemble the DROP database, we targeted 20 genera that potentially parasitize
258 frugivorous *Drosophila* species. We compiled DNA sequence and voucher data from four
259 sources: 1) museum collections, 2) publications, 3) molecular biodiversity inventories
260 publicly available in BOLD and GenBank, for which we managed to secure inspection of

261 the vouchers by taxonomists, and 4) conducted a sequencing and taxonomic inventory
262 of laboratory strains.
263 We first gathered species information into a catalogue of *Drosophila* parasitoid
264 species (Table 1) from 212 references (see DROP database reference table) and 36
265 institutes (Table S2). To ensure reliable names for nominal species (sequences identified
266 by a species name) in our database, we confirmed their taxonomic validity using the
267 Ichneumonoidea 2015 digital catalogue (Yu et al., 2016;
268 https://web.archive.org/web/20161022093945/http:/ichneumonoidea.name/global.ph
269 p) and Hymenoptera Online (HOL; http://hol.osu.edu/), both of which are curated by
270 taxonomic experts. To obtain reliable molecular identification data, we harvested 8,298
271 DNA sequences from GenBank and BOLD (all compiled into BOLD system: DS-DROPAR
272 dataset). These sequences represent 443 Barcode Index Numbers (BINs – a form of
273 provisionary taxa in BOLD) and 520 taxa, for a total of 963 operational taxonomic units
274 (OTUs). We use the term "OTU" as a general and neutral designation encompassing
275 described species, provisional species, undescribed species, cryptic species, and mis-
276 identified species.
277 The majority of the harvested sequences were Braconidae (6690), Diapriidae
278 (967), Figitidae (622), and Pteromalidae (19). Because of the concerns with generic
279 databases (noted above and in Figure 1 and Table S1), we assembled a list of sequences
280 with valid species names that could either be traceably linked to vouchers examined by
281 taxonomists or referred to directly in publications authored by a recognized expert in
282 the relevant taxon group. We then cross-checked species names with their

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

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

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

305 Where possible, we identified OTU strains using a combination of morphological and
306 sequence data, and characterized provisional species or species clusters using neighbor-
307 joining trees (Figure S1) based on the COI gene sequences (Supplemental material). For
308 establishing BIN limits in the context of DROP, we have adopted an initial percent cutoff
309 at 2%. As Ratnasingham & Hebert (2013) pointed out, this is a good starting point for
310 many taxa, but it also may be adjusted as more samples are acquired and compared.
311

312 ***Drosophila parasitoid database—DROP***

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

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

327 vouchers. DROP will be continued to be maintained by C-HL until further notice at the
328 Zenodo repository and sequences generated in the future will also be deposited in BOLD
329 (DROP project).
330
331 ***Species, provisional species, and OTU designations***
332 In addition to the inherent value of a formal taxonomic name, a reliable provisional
333 taxon label can also be used for exchanging scientific information and conveying
334 experimental results among researchers (Schindel & Miller, 2010). Based on the amount
335 of sequence divergence between described species, we observed what appears to be a
336 significant number of provisional OTUs in the initial dataset we compiled. Furthermore,
337 among the data linked to a valid species name, some of these provisional OTUs are
338 actively being used in research and have sequences available to the public. We
339 therefore provide a list of potential new species with their molecular vouchers.
340 We use the following designation format for OTUs that refer to a provisional species:
341 "Drop_strainX_sp.1" or, when no other information is known, "DROP_sp.1". Where
342 possible, these OTUs are linked to BINs within BOLD and to a voucher USNM specimen
343 label number. If the genus of the OTU is known, the "Drop_Leptopilina_sp.1" format is
344 followed. These designations can facilitate species identification as well as discovery and
345 description of new species without compromising the existing taxonomy of the
346 described OTUs in question. As more complete species descriptions become available,
347 this provisional species framework can be updated while keeping the link to previous
348 provisional species name.

349 **Results**

350 ***Overview of DROP***

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

361

362 ***Molecular Vouchers***

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

371 generated by this project (dataset: DS-LABS, DS-AUSPTOID) and came from 121 OTUs
372 (12 provisional species and 101 lab strains).
373 The DROP database is largely made up of standard barcode COI sequences (340
374 sequences), which includes 77 OTUs: 43 described species and 33 provisional species.
375 We aimed to supplement COI with secondary markers (28SD2, 18S, ITS2) when possible,
376 resulting in an additional 120 sequences from 26 OTUs: 15 described species and 11
377 provisional species. There are currently 19 OTUs that have sequences from more than
378 one genetic marker.
379
380 ***Species Delimitation in Laboratory Strains***
381 We used 298 COI sequences to resolve the identification of each laboratory
382 strain, and where possible, indicated potential species clusters for *Drosophila*
383 parasitoids (Fig. S1 and Table S3). Using a fixed 2% divergence cutoff, a total of 31 lab
384 strain OTUs were assignable to a valid species name, and the remaining 70 strain OTUs
385 were assigned to a provisional species. The taxonomic status of several of these
386 provisional species is also being investigated using an integrative taxonomic approach
387 involving morphological identification, genomic data, or other genetic data.
388
389 **Discussion**
390 In this paper, we introduce and describe a free and open-access database for the
391 reliable molecular identification of *Drosophila* parasitoids. The guiding principle of DROP

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

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

403

404 ***Molecular (mis-)identification***

405 We observe that 17% of the described *Drosophila* parasitoid OTUs in BOLD and
406 GenBank (dataset: DS-DROPAR) are associated with more than one BIN; these are
407 examples of BIN-ID conflict. Roughly half of these OTUs are used as lab strains. This
408 latter observation is disturbing, because it demonstrates that the criteria used to
409 differentiate and reference species in active research programs are clouded. For
410 example, BIN-ID conflicts were observed in the *Drosophila* parasitoids *Ganaspis*
411 *brasiliensis* (Ihering) and *Asobara japonica* Belokobylskij (Table S1), both of which are in
412 active use in numerous research programs (e.g. Moreau et al., 2009; Nomano et al.,
413 2017; Reumer et al., 2012; Wang et al., 2020a & 2021) as well as in biological control

414 efforts against the invasive *D. suzukii* (e.g. Abram et al., 2020; Daane et al., 2016;
415 Giorgini et al., 2019). All the BINs from *G. brasiliensis* carry the name *G. xanthopoda*
416 (Figure 1). In such instances, assigning an identification by matching specimens to
417 barcode records in the genetic database is problematic, as two names are applied to the
418 same BIN. If sequences comprising the BIN are not linked to a voucher that can be
419 examined, teasing apart the two names and how they are applied is impossible.
420 Applying explicit, consistent criteria for species determination ensures that
421 experimental results can be reliably repeated, and that any potentially novel
422 observations will not be explained away as artifacts of identification. DROP addresses
423 these concerns by linking reliable reference sequences and vouchers for *G. brasiliensis*
424 (Figure 1) and from different studies: one with reference to the morphological
425 description (Buffington & Forshage, 2016) and the other with reference to the genome
426 (using voucher specimens from the morphological study; Blaimer et al., 2020).
427 We were not able to resolve all conflicts between BIN and species identity, for one
428 or more of the following three reasons: First, many records lack reliably identified
429 vouchers and have often been themselves used for molecular identification,
430 proliferating errors. Second, in some cases, it is not possible to verify whether the
431 genetic differences among BINs represent different species or simply intraspecific
432 genetic variation (Bergsten et al., 2012), because BINs themselves are not a species
433 concept. The only solution to this problem is to derive original sequence data from type
434 specimens (which is often either impractical or impossible for a number of technical
435 reasons), or from specimens whose conspecificity with the types has been corroborated.

436 Since species boundaries are always subject to testing, additional specimens from
437 multiple collecting events (e.g., representing different seasons and geographic regions)
438 may help provide the additional data to circumscribe a given species' limits. The third
439 difficulty in resolving BIN-ID conflict derives from the data themselves: Although the
440 mitochondrial COI gene is the locus most frequently chosen for identification of insects
441 and other animals, its effectiveness varies among insect groups (Brower & DeSalle,
442 2002; Gompert et al., 2008; Lin & Danforth, 2004). In part, this derives from gene-
443 tree/species-tree conflict as a function of mitochondrial DNA introgression (Gompert et
444 al., 2008; Klopfstein et al., 2016), parthenogenesis (Reumer et al., 2012), and/or
445 *Wolbachia* infection (Wachi et al., 2015; Xiao et al., 2012), any of which may lead to
446 complications in species delimitation using mitochondrial loci. Ideally, studies should
447 apply multiple loci, genomes, and comparative taxonomic data to clarify species
448 boundaries. As *Drosophila* parasitoids are often maintained in laboratory cultures, it is
449 also possible to use mating experiments to explore species boundaries under the
450 paradigm of the biological species concept (Seehausen et al., 2020).
451
452 ***DROP as a taxonomic tool***
453 DROP offers an empirical platform for species discovery and a useful tool for
454 taxonomic research. The fact that the number of BINs reported here exceeds the
455 number of described species (Table S1, Figure S3) highlights the need for taxonomic
456 work. But such work cannot proceed on the basis of BINs or barcodes, but requires
457 integrative taxonomic approach employing a combination of molecular and

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

466 Public genetic databases have adopted a longstanding convention in treating
467 undetermined OTUs and sequences, referring to provisional species with numbers, as
468 for example "sp. 1", and these are rarely linked to vouchers. For OTUs designated as
469 provisional species, DROP enables cross-indexing of specimens, sequences and
470 references with studies and provides researchers with valuable tools for taxonomic
471 revisions, including the means of discovery, corroboration, and description of new
472 species. For example, "drop_Gan1_sp.1" refers to voucher USNMENT01557320
473 deposited in the USNM, Washington DC, COI sequence (DROP sequence_id 2), BOLD
474 process ID: DROP143-21, BIN number: XXXXXX (will update in the revision), 28D1
475 sequences (DROP sequence_id 289), and 28D2 sequences (DROP sequence_id 303). In
476 the future, when "drop_Gan1_sp.1" is described as a new species with a formal specific
477 epithet, DROP will update the species status and holotype information while keeping
478 this provisional species name as an informal "synonym." We recognize tracking these

479 informal ‘tags’ through time can be problematic; however, linking these tags in DROP to
480 a vouchered specimen and unique identifier will minimize confusion.
481
482 ***From molecular mechanisms to ecosystem structure***
483 The use of molecular tools in insect biodiversity studies has gradually expanded from
484 barcoding single individuals to metabarcoding large environmental samples
485 representing entire food webs (Jeffs et al., 2020; Littlefair et al., 2016). *Drosophila* and
486 their parasitoids are among the few systems that currently allow us to explore
487 thoroughly the mechanisms of species interactions at scales ranging from the molecular
488 to the ecological. Here, we highlight two examples where information compiled in DROP
489 enables the study of the *Drosophila*-parasitoid system across multiple levels of biological
490 organization:

491 DROP includes a DNA reference library of Australian *Drosophila* parasitoids (DS-
492 AUSPTOID in BOLD) that connects laboratory experiments and field research. Molecular
493 vouchers of both hosts and parasitoids were collected along altitudinal gradients in the
494 rainforest of northern Queensland, Australia (Jeffs et al., 2020). With this DNA reference
495 library, researchers can detect interactions between *Drosophila* and their parasitoids
496 using PCR-based approaches and parasitized pupae (Hrcek & Godfray, 2015; Jeffs et al.,
497 2020). Surveying host-parasitoid interactions in this way will improve our understanding
498 of how environmental change alters the structure of host-parasitoid networks (Morris et
499 al., 2014; Staniczenko et al., 2017; Tylianakis et al., 2007) by accelerating data collection
500 in the field. In addition, JH established lab cultures of both hosts and their parasitoids

501 from the same Australian sampling sites with the aim of conducting laboratory
502 experiments (e.g. Thierry et al., 2021). Molecular vouchers of the lab strains were then
503 submitted to DROP as a reference database (DS-LABS in BOLD) to ensure that criteria for
504 species determination were applied consistently—and will continue to be applied
505 consistently—between the natural community studies and the laboratory experiments.
506 The presence of a foundational DNA reference library and species catalogue in
507 DROP will enable the process of exploring parasitoid biodiversity to become more
508 efficient. For example, DROP includes molecular vouchers from *Drosophila* parasitoids
509 that were collected across seasons and along latitudinal gradients in the eastern Unites
510 States (Lue et al., 2016, 2018). These data proved to be extremely useful for identifying
511 species in a more recent exploration of native parasitoid biodiversity across North
512 America (e.g., Abram et al., 2020). There are additional uses for DROP: curated
513 specimen collections may be used to document species distributions, phenology,
514 understand micro-evolutionary patterns, observe the effects of climate change, and
515 detect and track biological invasions (Funk, 2018; Schilthuizen et al., 2015; Tarli et al.,
516 2018).
517
518 ***Taxonomic accuracy for biocontrol studies***
519 Unfortunately, the history of biological control includes many examples of
520 misidentifications that have resulted in failures to employ or establish the expected
521 control agent, thus hindering eventual success (Buffington et al., 2018; Rosen, 1986;
522 Huffaker et al. 1962). In the context of biological control research on *Drosophila* pest

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

534 In addition to species identification, reference sequences from DROP may be used to
535 create species-specific primers for the accurate identification of parasitoids, design
536 multiplex PCR assays that rapidly distinguish species in natural or agricultural
537 ecosystems (Ye et al., 2017), and apply high-throughput molecular identification
538 diagnostics (Fagan-Jeffries et al., 2018). Applications of such specific primers have been
539 used in bacteria, fungi, oomycetes and insect pests (Liu et al., 2017; Tedersoo et al.,
540 2019; Tsai et al., 2020).

541

542 ***Long-term molecular voucher preservation***

543 During the curation of DROP, we found that holotype specimens were missing from
544 museums for several iconic *Drosophila* parasitoid species: *Asobara tabida* (Nees von
545 Esenbeck), *Leptopilina clavipes* (Hartig), and *Leptopilina longipes* (Hartig). This is not
546 uncommon and impedes future taxonomic revisions regardless of whether molecular
547 data are used. To avoid contributing to this problem, DROP uses museums as
548 depositories for ensuring that sequenced vouchers of both described species and
549 provisional species are permanently stored. In order to stabilize nomenclature, we
550 further advocate the designation of neotypes (a replacement specimen for a missing
551 holotype) that have museum-vouchered DNA barcodes and additional genomic
552 extractions in storage.
553 Natural history museums are designed to maintain vouchers (including types) for
554 long-term preservation, and increasingly they implement institutionalized workflows
555 that link DNA sequences to specimens and specimen metadata (Prendini et al., 2002).
556 We strongly encourage the deposition of voucher specimens from field surveys and
557 experimental studies in museum collections, as has been urged by the Entomology
558 Collections Network (ECN) and required in many PhD programs. No matter how quickly
559 new molecular techniques are developed or refined, there is no substitute for a reliable
560 database of voucher specimens when it comes to ensuring the repeatability of biological
561 research (Funk et al., 2005; Lendemer et al., 2020).
562 Our results show that species richness of the parasitic wasps that attack *Drosophila*
563 is severely underestimated, and only a fraction of them have been described. In DROP,
564 38% of the OTUs are provisional species, and more than 46% of the named OTUs have

565 synonyms. Remarkably, *Leptopilina heterotoma*, one of the world's most studied
566 parasitoids, has more than 20 synonyms! As is generally the case, the rate of species
567 description and revision of *Drosophila* parasitoids lags far behind that with which
568 molecular sequence data are generated. Ensuring a consistent application of OTU
569 recognition is therefore essential. With DROP, researchers may ensure consistency is
570 their application of scientific names, and that those names are valid, making the
571 daunting process of describing *Drosophila* parasitoids more accurate and efficient.
572 In addition to the collection of physical museum resources, a central role
573 taxonomists play in DROP and its curation is that of fostering better integration of
574 taxonomy with experimental and biodiversity research. Our intention is to perpetuate
575 DROP beyond this introductory publication. We hope that experts in all areas of
576 *Drosophila*-parasitoid biology and related fields will join us in this effort.
577
578 ***Conclusion***
579 Taxonomic confusion presents many obstacles in experimental and biodiversity
580 studies. One way of addressing this impediment is to provide a reliable DNA library with
581 traceable vouchers (Astrin et al., 2013). We developed DROP as a resource and platform
582 for gathering and sharing reliable genomic sequence data for *Drosophila* parasitoids. We
583 hope it will serve as a model for researchers working with organisms which present
584 similar difficulties. While compiling DROP, we found that the high number of provisional
585 *versus* named OTUs suggests that the diversity of parasitic wasps attacking *Drosophila* is
586 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

Bergsten, J., Bilton, D. T., Fujisawa, T., Elliott, M., Monaghan, M. T., Balke, M., Hendrich, … Vogler, A. P. (2012). The effect of geographical scale of sampling on DNA barcoding. *Systematic Biology,* 61 (5), 851-869. https://doi.org/10.1093/sysbio/sys037

Blaimer, B. B., Gotzek, D., Brady, S. G., & Buffington, M. (2020). Comprehensive phylogenomic analyses re-write the evolution of parasitism within cynipoid wasps. *BMC Ecology and Evolution*, 20 (155). https://doi.org/10.1186/s12862-020-01716-2

Brower, A. V. Z., & DeSalle, R. (2002). Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies: The utility of wingless as a source of characters for phylogenetic inference. *Insect Molecular Biology*, 7 (1), 73-82. https://doi.org/10.1046/j.1365-2583.1998.71052.x

Buffington, M., &Forshage, M. (2016). Redescription of *Ganaspis brasiliensis* (Ihering, 1905), new combination, (Hymenoptera: Figitidae) a natural enemy of the Invasive *Drosophila suzukii* (Matsumura, 1931) (Diptera: Drosophilidae). *Proceedings of the Entomological Society of Washington*, 118(1), 1–13. https://doi.org/10.4289/0013-8797.118.1.1

Buffington, M., Talamas, E. J., & Hoelmer, K. A. (2018). Team Trissolcus: Integrating taxonomy and biological control to combat the brown marmorated stink bug. *American Entomologist*, 64 (4), 224–232

Carton, Y., Boulétreau, M., van Alphen, J. J. M., & van Lenteren, J. C. (1986). The *Drosophila* parasitic wasps. In Ashburner M, Carson HL, Thompson JN (Eds), *The genetics and biology of Drosophila*, (3),348–394.

Carton, Y., Haouas, S., Marrakchi, M., & Hochberg, M. (1991). Two competing parasitoid species coexist in sympatry. *Oikos*, 60, 222-230. https://doi.org/10.2307/3544869

Condon, M. A., Scheffer, S. J., Lewis, M. L., Wharton, R., Adams, D.C., & Forbes, A. A. (2014). Lethal interactions between parasites and prey increase niche diversity in a tropical community. *Science*, *343*(6176), pp.1240-1244.

Daane, K. M., Wang, X.-G., Biondi, A., Miller, B. E., Miller, J. C., Riedl, H., Shearer, P. W., … Walton, V. M. (2016). First exploration of parasitoids of *Drosophila suzukii* in South Korea as potential classical biological agents. *Journal of Pest Science* 89, 823–835, doi:10.1007/s10340-016-0740-0.

Dolphin, K., & Quicke, D. L. J. (2001). Estimating the global incompletely described parasitoid wasps. *Biological Journal Of The Linnean Society*, 73 (3), 279-286, https://doi.org/10.1006

Escalante, P., Ibarra-Vazquez, A., & Rosas-Escobar, P. (2010). Tropical montane nymphalids in Mexico: DNA barcodes reveal greater diversity. *Mitochondrial DNA*, 21, 30-37, https://doi.org/10.3109/19401736.2010.535527

Fagan-Jeffries, E. P., Cooper, S. J. B., Bertozzi, T., Bradford, T. M., & Austin, A. D. (2018). DNA barcoding of microgastrine parasitoid wasps (Hymenoptera: Braconidae) using high-throughput methods more than doubles the number of species known for Australia. *Molecular Ecology Resources*, 18(5), 1132–1143. https://doi.org/10.1111/1755-0998.12904

Ferguson, J. W. H. (2002). On the use of genetic divergence for identifying species. 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. doi: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

Hebert, P. D. N., Ratnasingham, S., & DeWaard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences*, 270 (Suppl.), 96-99, https://doi.org/10.1098/rsbl.2003.0025

Hendrich, L., Morinière, J., Haszprunar, G., Hebert, P. D. N., Hausmann, A., Köhler, F., & Balke, M. (2015). A comprehensive DNA barcode database for Central European beetles with a focus on Germany: Adding more than 3500 identified species to BOLD. *Molecular Ecology Resources*, 15(4), 795-818, https://doi.org/10.1111/1755-0998.12354

Hrček, J., & Godfray, H. C. J. (2015). What do molecular methods bring to host-parasitoid food webs? *Trends in Parasitology*, 31(1), 30–35. https://doi.org/10.1016/j.pt.2014.10.008

Hrcek, J., Miller, S. E., Quicke, D. L. J., & Smith, M. A. (2011). Molecular detection of trophic links in a complex insect host-parasitoid food web. *Molecular Ecology Resources*, 11(5), 786–794. https://doi.org/10.1111/j.1755-0998.2011.03016.x

Huffaker, C. B., Kennett, C. E., Finney, G. L. (1962). Biological control of olive scale, Pwrlatoria oleae (Cohree), in California by imported Aphytis maculicornis (Masi) (Hymenoptera: Aphelinidae). *Hilgardia,* 32 (13): 541-636. DOI: 10.3733/hilg.v32n13p541

Hughes, R. D., Woolcock, L. T. (1976). Aphaereta aotea sp. N. (Hymenoptera: Braconidae), an Alysiine parasite of dung breeding flies. J*ournal of Australian Entomological Society*, 15, 191-196.

Jeffs, C. T., Terry, J. C. D., Higgie. M., Jandová, A., Konvičková. H., Brown. J. J., Lue. C.-H., … Lewis, O. T. (2020). Molecular analyses reveal consistent food web structure with elevation in rainforest *Drosophila* - parasitoid communities. *Ecography*, 43, 1-11, https://doi.org/10.1111/ecog.05390

Kim-Jo, C., Gatti, J. L., & Poirié, M. (2019). *Drosophila* cellular immunity against parasitoid wasps: A complex and time-dependent process. *In Frontiers in Physiology*, https://doi.org/10.3389/fphys.2019.00603

Kimura, M. T. (2015). Prevalence of exotic frugivorous *Drosophila* species, *D. simulans* and *D. immigrans* (Diptera: Drosophilidae), and its effects on local parasitoids in Sapporo, northern Japan. *Applied Entomology and Zoology*, 50(4), 509–515. https://doi.org/10.1007/s13355-015-0361-8

Kimura, M. T., & Mitsui, H. (2020). *Drosophila* parasitoids (Hymenoptera) of Japan. I*n Entomological Science*, 23(4), 359-368, https://doi.org/10.1111/ens.12432

Klopfstein, S., Kropf, C., & Baur, H. (2016). Wolbachia endosymbionts distort DNA barcoding in the parasitoid wasp genus Diplazon (Hymenoptera: Ichneumonidae). *Zoological Journal of the Linnean Society*, 177(3), 541–557. https://doi.org/10.1111/zoj.12380

Kraaijeveld, A. R., & Godfray, H. C. J. (1997). Trade-off between parasitoid resistance and larval competitive. *Nature*, 389, 278-280, https://doi.org/10.1038/38483

Lefèvre, T., De Roode, J. C., Kacsoh, B. Z., & Schlenke, T. A. (2012). Defence strategies against a parasitoid wasp in Drosophila: Fight or flight? *Biology Letters*, 8(2), 230-233, https://doi.org/10.1098/rsbl.2011.0725

Lendemer, J., Thiers, B., Monfils, A. K., Zaspel, J., Ellwood, E. R., Bentley, A., LeVan, K., … Aime, M. C. (2020). The extended specimen network: A strategy to enhance US biodiversity collections, promote research and education. *BioScience*, 70(1), 23-30, https://doi.org/10.1093/biosci/biz140

Lin, C. P., & Danforth, B. N. (2004). How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analyses of combined datasets. *Molecular Phylogenetics and Evolution*, 30(3), 686-702, https://doi.org/10.1016/S1055-7903(03)00241-0

Littlefair, J. E., Clare, E. L., & Naaum, A. (2016). Barcoding the food chain: From Sanger to high-throughput sequencing1. *Genome*, 59(11), 946–958. https://doi.org/10.1139/gen-2016-0028

Liu, L. J., Pang, A. H., Feng, S. Q., Cui, B. Y., Zhao, Z. H., Kučerová, Z., Stejskal, V., … Li, Z. H. (2017). Molecular Identification of ten species of stored-product psocids through microarray method based on ITS2 rDNA. *Scientific Reports*, 7(1): 16694, https://doi.org/10.1038/s41598-017-16888-z

Lue, C.-H., Borowy, D., Buffington, M. L., & Leips, J. (2018). Geographic and seasonal variation in species diversity and community composition of frugivorous *Drosophila* (Diptera: Drosophilidae) and their *Leptopilina* (Hymenoptera: Figitidae) parasitoids. *Environmental Entomology*, 47(5): 1096-1106. https://doi.org/10.1093/ee/nvy114

Lue, C.-H., Driskell, A. C., Leips, J., & Buffington, M. L. (2016). Review of the genus *Leptopilina* (Hymenoptera, Cynipoidea, Figitidae, Eucoilinae) from the Eastern United States, including three newly described species. J*ournal of Hymenoptera Research*, 53: 35-76. https://doi.org/10.3897/jhr.53.10369

Meiklejohn, K. A., Damaso, N., & Robertson, J. M. (2019). Assessment of BOLD and GenBank – Their accuracy and reliability for the identification of biological materials. *PLoS ONE,* 14(6): e0217084. https://doi.org/10.1371/journal.pone.0217084

Melk, J. P., & Govind, S. (1999). Developmental analysis of *Ganaspis xanthopoda*, a larval parasitoid of *Drosophila melanogaster*. *Journal of Experimental Biology*, 202, 1885-1896

Miller, S. E., Hausmann, A., Hallwachs, W., & Janzen, D. H. (2016). Advancing taxonomy and bioinventories with DNA barcodes. *Philosophical Transactions of the Royal Society Biological Sciences*, 371(1702): 20150339. doi: 10.1098/rstb.2015.0339

Mitsui, H., van Achterberg, K., Nordlander, G., & Kimura, M. T. (2007). Geographical distributions and host associations of larval parasitoids of frugivorous *Drosophilidae* in Japan. *Journal of Natural History*, 41(25–28), 1731–1738. https://doi.org/10.1080/00222930701504797

Moreau, S. J. M., Vinchon, S., Cherqui, A., & Prévost, G. (2009). Components of Asobara venoms and their effects on hosts. *In Advances in Parasitology*, Prévost G (Ed). 70, 217-232, https://doi.org/10.1016/S0065-308X(09)70008-9

Morris, R. J., Gripenberg, S., Lewis, O. T., & Roslin, T. (2014). Antagonistic interaction networks are structured independently of latitude and host guild. *Ecology Letters*, 17(3), 340-349, https://doi.org/10.1111/ele.12235

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

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

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,

https://doi.org/10.1038/hdy.2011.82
Rosen, D. (1986). The role of taxonomy in effective biological control programs. *Agriculture, Ecosystems & Environment*, 15(2-3), 121-129. https://doi.org/10.1016/0167-8809(86)90085-X
Santos, W. G. N., Fernandes, E. C., Souza, M. M., Guimarães, J. A., & Araujo, E. L. (2016). First record of Eucoilinae (Hymenoptera: Figitidae), parasitoids of African fig fly *Zaprionus indianus* Gupta (Diptera: Drosophilidae), in the Caatinga biome. *Semina:Ciencias Agrarias*, 37(5), 3055–3058. https://doi.org/10.5433/1679-0359.2016v37n5p3055
Schilthuizen, M., Vairappan, C. S., Slade, E. M., Mann, D. J., & Miller, J. A. (2015). Specimens as primary data: Museums and "open science." *Trends in Ecology and Evolution*, 30(5), 237–238. https://doi.org/10.1016/j.tree.2015.03.002
Schlenke, T. A., Morales, J., Govind, S., & Clark, A. G. (2007). Contrasting infection strategies in generalist and specialist wasp parasitoids of *Drosophila melanogaster*. *PLoS Pathogens*, 3(10):e158, https://doi.org/10.1371/journal.ppat.0030158
Schindel, D., & Miller, S. E. (2010). Provisional Nomenclature the on-ramp to taxonomic names. In: Polaszek, A. (Ed), *Systema Nature*, 250: The Linnaean Ark. CRC, Boca Raton, 109-115.
Seehausen, M. L., Ris, N., Driss, L., Racca, A., Girod, P., Warot, S., Borowiec, N., Tosevski, I., & Kenis, M. (2020). Evidence for a cryptic parasitoid species reveals its suitability as a biological control agent. *Scientific reports*, 10: 19096. https://doi.org/10.1038/s41598-020-76180-5
Shokralla, S., Spall, J. L., Gibson, J. F., & Hajibabaei, M. (2012). Next-generation sequencing technologies for environmental DNA research. *In Molecular Ecology*, 21(8), 1794-1805, https://doi.org/10.1111/j.1365-294X.2012.05538.x
Smith, M. A., Fisher, B. L., & Hebert, P. D. N. (2005). DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: The ants of Madagascar. *Philosophical Transactions of the Royal Society Biological Sciences*, 360(1462), 1825-1834, https://doi.org/10.1098/rstb.2005.1714
Smith, M. A., Fernandez-Triana, J. L., Eveleigh, E., Gomez, J., Guclu, C., Hallwachs, W., Hebert. P. D. N., … Zaldivar-Riveron, A. (2013). DNA barcoding and the taxonomy of Microgastrinae wasps (Hymenoptera, Braconidae): impacts after 8 years and nearly 20000 sequences. *Molecular Ecology Resource*s, 13, 168-276, https://doi.org/10.1111/1755-0988.12038
Staniczenko, P. P. A., Reed-Tsochas, F., Lewis, O. T., Tylianakis, J. M., Albrecht, M., Coudrain, V., & Klein, A. M. (2017). Predicting the effect of habitat modification on networks of interacting species. *Nature Communications*, 8, 792, https://doi.org/10.1038/s41467-017-00913-w
Tarli, V. D., Grandcolas, P., & Pellens, R. (2018). The informative value of museum collections for ecology and conservation: A comparison with target sampling in the Brazilian Atlantic forest. *PLoS ONE*, 13(11). https://doi.org/10.1371/journal.pone.0205710
Tedersoo, L., Drenkhan, R., Anslan, S., Morales-Rodriguez, C., & Cleary, M. (2019). High-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

923 pressure. *FEMS Microbiology Ecology*, 91(2), 1-11,
924 https://doi.org/10.1093/femsec/iu017
925 Yassin, A. (2013). Phylogenetic classification of the Drosophilidae Rondani (Diptera): The
926 role of morphology in the postgenomic era. *Systematic Entomology*,
927 https://doi.org/10.1111/j.1365-3113.2012.00665.x
928 Yassin, A., & David, J. R. (2010). Revision of the Afrotropical species of *Zaprionus*
929 (Diptera, Drosophilidae), with descriptions of two new species and notes on
930 internal reproductive structures and immature stages. *ZooKeys,* 51, 33-72,
931 https://doi.org/10.3897/zookeys.51.380
932 Ye, Z., Vollhardt, I. M. G., Girtler, S., Wallinger, C., Tomanovic, Z., & Traugott, M. (2017).
933 An effective molecular approach for assessing cereal aphid-parasitoid-
934 endosymbiont networks. *Scientific Reports*, 7(1), 1–12.
935 https://doi.org/10.1038/s41598-017-02226-w
936
937

938 **Data Accessibility**

939 The DROP database is freely accessible at Zenodo depository

940 (http://doi.org/10.5281/zenodo.4519656). New sequences have been deposited in

941 BOLD in datasets DS-LABS, and DS-AUSPTOID. [NOTE: the doi records will be update at

942 revision stage]

943

944 **Author Contributions**

945 The initial project idea was originated by C-HL, MLB, JH, MM, TS, JV, SG, and

946 PPAS. Molecular work was conducted by C-HL, SS, ML, AJ, and AD. BOLD and GenBank

947 data was harvested by TAE and C-HL. Figures were made by AL and C-HL. Laboratory and

948 field sample preparations were conducted by MTK, YC, TS, MM, SG, JV, EG, MG, XW,

949 KM, KMD, PA, NAP, MT, JJB, MP, FMJ, WDT, JSD, BW, OTL, PPAS, JL and AL. Taxonomic

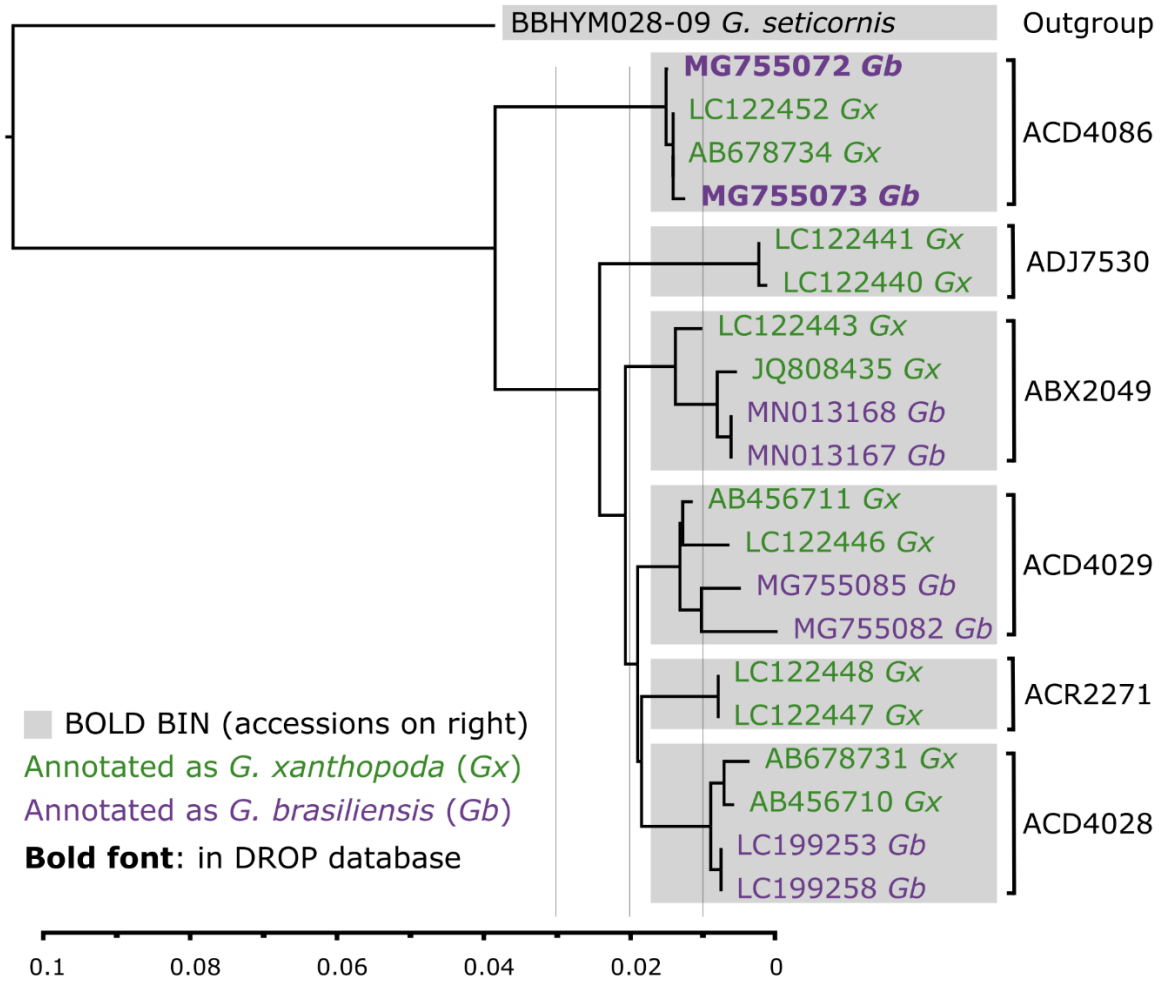
950 concepts and interpretations were conducted by RRK, MLB, CH-L, PG, and SEM. DROP

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.

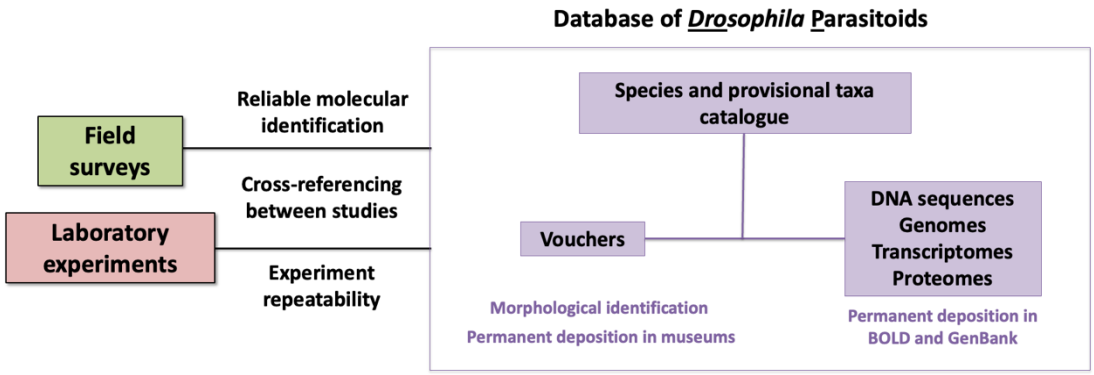


Figure 2: Concept of a centralized, vetted, integrated database for *Drosophila* Parasitoids (DROP) we developed. First, we provide a species and provisional species catalog with correct taxonomy. Second, to provide a reliable genetic reference library, genetic data (DNA sequences, genomes, transcriptomes, 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		*drop_Cha2_sp12*	
Chalcidoidea	Encyrtidae	*Tachinaephagus*	*drop_ IR1_sp41*	Kimura
Chalcidoidea	Encyrtidae	*Tachinaephagus*	*drop_BG1_sp42*	Kimura
Chalcidoidea	Encyrtidae	*Tachinaephagus*	*zealandicus*	Ashmead 1904
Chalcidoidea	Pteromalidae		*drop_Pte69_sp11*	
Chalcidoidea	Pteromalidae	*Pachycrepoideus*	*vindemmiae*	(Rondani, 1875)
Chalcidoidea	Pteromalidae	*Spalangia*	*drop_IR1_sp38*	Kimura
Chalcidoidea	Pteromalidae	*Spalangia*	*drop_NG1_sp39*	Kimura
Chalcidoidea	Pteromalidae	*Spalangia*	*drop_SK1_sp40*	Kimura
Chalcidoidea	Pteromalidae	*Spalangia*	*drosophilae*	Ashmead 1887
Chalcidoidea	Pteromalidae	*Spalangia*	*erythromera*	Foerster 1850
Chalcidoidea	Pteromalidae	*Trichomalopsis*	*dubia*	(Ashmead, 1896)
Chalcidoidea	Pteromalidae	*Trichomalopsis*	*microptera*	(Lindeman, 1887)
Chalcidoidea	Pteromalidae	*Trichomalopsis*	*nigricola*	Boucek

Chalcidoidea	Pteromalidae	*Trichomalopsis*	*sarcophagae*	(Gahan, 1914)
Chalcidoidea	Pteromalidae	*Vrestovia*	*brevior*	Boucek 1993
Chalcidoidea	Pteromalidae	*Vrestovia*	*fidenas*	(Walker, 1848)
Chalcidoidea	Pteromalidae		*drop_ PacAtl_sp46*	
Chalcidoidea	Pteromalidae		*drop_ PachyPort_sp45*	
Chalcidoidea			*drop_ CH_sp64*	
Cynipoidea	Figitidae	*Ganaspis*	*brasiliensis*	(Ihering, 1905)
Cynipoidea	Figitidae	*Ganaspis*	*drop_ Gan_sp51*	
Cynipoidea	Figitidae	*Ganaspis*	*drop_ Gan_sp52*	
Cynipoidea	Figitidae	*Ganaspis*	*drop_ Gan_sp53*	
Cynipoidea	Figitidae	*Ganaspis*	*drop_ Gsp1_sp67*	
Cynipoidea	Figitidae	*Ganaspis*	*drop_ Gsp2_sp68*	
Cynipoidea	Figitidae	*Ganaspis*	*drop_ Gsp50_sp66*	
Cynipoidea	Figitidae	*Ganaspis*	*drop_ IR1_sp25*	Kimura
Cynipoidea	Figitidae	*Ganaspis*	*drop_ IR2_sp26*	Kimura
Cynipoidea	Figitidae	*Ganaspis*	*drop_Gan1_sp1*	
Cynipoidea	Figitidae	*Ganaspis*	*drop_TK1_sp27*	Kimura
Cynipoidea	Figitidae	*Ganaspis*	*hookeri*	Craword 1913
Cynipoidea	Figitidae	*Ganaspis*	*mahensis*	Kieffer 1911
Cynipoidea	Figitidae	*Ganaspis*	*mellipes*	(Say, 1826)
Cynipoidea	Figitidae	*Ganaspis*	*mundata*	Forster 1869
Cynipoidea	Figitidae	*Ganaspis*	*seticornis*	(Hellen, 1960)
Cynipoidea	Figitidae	*Ganaspis*	*tenuicornis*	Kieffer 1904
Cynipoidea	Figitidae	*Ganaspis*	*xanthopoda*	(Ashmead, 1896)
Cynipoidea	Figitidae	*Kleidotoma*	*bicolor*	(Giraud, 1860)
Cynipoidea	Figitidae	*Kleidotoma*	*dolichocera*	Thomson 1877
Cynipoidea	Figitidae	*Kleidotoma*	*drop_TK1_sp28*	Kimura
Cynipoidea	Figitidae	*Kleidotoma*	*filicornis*	(Cameron, 1889)
Cynipoidea	Figitidae	*Kleidotoma*	*icarus*	(Quinlan, 1964)
Cynipoidea	Figitidae	*Kleidotoma*	*psiloides*	Westwood 1833
Cynipoidea	Figitidae	*Kleidotoma*	*tetratoma*	(Hartig, 1841)
Cynipoidea	Figitidae	*Leptolamina*	*drop_Fig64_sp5*	
Cynipoidea	Figitidae	*Leptolamina*	*drop_Lmn_sp6*	
Cynipoidea	Figitidae	*Leptolamina*	*drop_TK1_sp29*	Kimura
Cynipoidea	Figitidae	*Leptolamina*	*gressitti*	Yoshimoto & Yasumatsu 1965
Cynipoidea	Figitidae	*Leptolamina*	*papuensis*	Yoshimoto 1963
Cynipoidea	Figitidae	*Leptolamina*	*ponapensis*	Yoshimoto 1962
Cynipoidea	Figitidae	*Leptolamina*	*seychellensis*	(Kieffer, 1911)

Cynipoidea	Figitidae	*Leptopilina*	*atraticeps*	(Kieffer, 1911)
Cynipoidea	Figitidae	*Leptopilina*	*australis*	(Belizin, 1966)
Cynipoidea	Figitidae	*Leptopilina*	*boulardi*	(Barbotin, Carton & Kelner-Pillault, 1979)
Cynipoidea	Figitidae	*Leptopilina*	*clavipes*	(Hartig, 1841)
Cynipoidea	Figitidae	*Leptopilina*	*cupulifera*	(Kieffer, 1916)
Cynipoidea	Figitidae	*Leptopilina*	*decemflagella*	Lue & Buffington 2017
Cynipoidea	Figitidae	*Leptopilina*	*drop_ Lep_sp54*	
Cynipoidea	Figitidae	*Leptopilina*	*drop_ Lep_sp55*	
Cynipoidea	Figitidae	*Leptopilina*	*drop_ Lep_sp56*	
Cynipoidea	Figitidae	*Leptopilina*	*drop_ Lep_sp57*	
Cynipoidea	Figitidae	*Leptopilina*	*drop_ Lep_sp58*	
Cynipoidea	Figitidae	*Leptopilina*	*drop_ Lep_sp59*	
Cynipoidea	Figitidae	*Leptopilina*	*drop_ Lep_sp60*	
Cynipoidea	Figitidae	*Leptopilina*	*drop_ Lep_sp61*	
Cynipoidea	Figitidae	*Leptopilina*	*drop_ Lep_sp62*	
Cynipoidea	Figitidae	*Leptopilina*	*drop_BG1_sp34*	Kimura
Cynipoidea	Figitidae	*Leptopilina*	*drop_Fig059_sp4*	
Cynipoidea	Figitidae	*Leptopilina*	*drop_Fig124_sp2*	
Cynipoidea	Figitidae	*Leptopilina*	*drop_Fig58_sp3*	
Cynipoidea	Figitidae	*Leptopilina*	*drop_IR1_sp30*	Kimura
Cynipoidea	Figitidae	*Leptopilina*	*drop_NG1_sp33*	Kimura
Cynipoidea	Figitidae	*Leptopilina*	*drop_SK1_sp35*	Kimura
Cynipoidea	Figitidae	*Leptopilina*	*drop_STL_sp7*	
Cynipoidea	Figitidae	*Leptopilina*	*drop_TK2_sp31*	Kimura
Cynipoidea	Figitidae	*Leptopilina*	*drop_TK3_sp32*	Kimura
Cynipoidea	Figitidae	*Leptopilina*	*fimbriata*	(Kieffer, 1901)
Cynipoidea	Figitidae	*Leptopilina*	*freyae*	Allemand & Nordlander 2002
Cynipoidea	Figitidae	*Leptopilina*	*guineaensis*	Allemand & Nordlander 2002
Cynipoidea	Figitidae	*Leptopilina*	*heterotoma*	(Thomson, 1862)
Cynipoidea	Figitidae	*Leptopilina*	*japonica japonica*	Novkovic & Kimura 2011
Cynipoidea	Figitidae	*Leptopilina*	*lasallei*	Buffington & Guerrieri 2020
Cynipoidea	Figitidae	*Leptopilina*	*leipsi*	Lue & Buffington 2018
Cynipoidea	Figitidae	*Leptopilina*	*lonchaeae*	(Cameron, 1912)
Cynipoidea	Figitidae	*Leptopilina*	*longipes*	(Hartig, 1841)
Cynipoidea	Figitidae	*Leptopilina*	*mahensis*	(Kieffer, 1911)
Cynipoidea	Figitidae	*Leptopilina*	*maia*	Lue & Buffington 2016
Cynipoidea	Figitidae	*Leptopilina*	*maria*	(Girault, 1930)
Cynipoidea	Figitidae	*Leptopilina*	*orientalis*	Allemand & Nordlander 2002

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

Ichneumonoidea	Braconidae	*Asobara*	*orientalis*	Viereck 1913
Ichneumonoidea	Braconidae	*Asobara*	*persimilis*	(Prince, 1976)
Ichneumonoidea	Braconidae	*Asobara*	*pleuralis*	(Ashmead, 1905)
Ichneumonoidea	Braconidae	*Asobara*	*rossica*	Belokobylskij 1998
Ichneumonoidea	Braconidae	*Asobara*	*rufescens*	(Fˆrster, 1862)
Ichneumonoidea	Braconidae	*Asobara*	*tabida*	(Nees, 1834)
Ichneumonoidea	Braconidae	*Asobara*	*triangulata*	van Achterberg & Guerrieri 2016
Ichneumonoidea	Braconidae	*Asobara*	*turneri*	Peris-Felipo 2014
Ichneumonoidea	Braconidae	*Asobara*	*unicolorata*	van Achterberg & Guerrieri 2016
Ichneumonoidea	Braconidae	*Aspilota*	*albertica*	Berry 2007
Ichneumonoidea	Braconidae	*Aspilota*	*andyaustini*	Wharton 2002
Ichneumonoidea	Braconidae	*Aspilota*	*angusta*	Berry 2007
Ichneumonoidea	Braconidae	*Aspilota*	*concolor*	Nees 1812
Ichneumonoidea	Braconidae	*Aspilota*	*parecur*	Berry 2007
Ichneumonoidea	Braconidae	*Aspilota*	*villosa*	Berry 2007
Ichneumonoidea	Braconidae	*Dinotrema*	*barrattae*	Berry 2007
Ichneumonoidea	Braconidae	*Dinotrema*	*longworthi*	Berry 2007
Ichneumonoidea	Braconidae	*Dinotrema*	*philipi*	Berry 2007
Ichneumonoidea	Braconidae		*drop_Aso_sp8*	
Ichneumonoidea	Braconidae	*Opiognathus*	*pactus*	(Haliday, 1837)
Ichneumonoidea	Braconidae	*Opius*	*bellus*	Gahan 1930
Ichneumonoidea	Braconidae	*Opius*	*cinerariae*	Fischer
Ichneumonoidea	Braconidae	*Opius*	*crenuliferus*	Li & van Achterberg 2013
Ichneumonoidea	Braconidae	*Opius*	*monilipalpis*	Li & van Achterberg 2013
Ichneumonoidea	Braconidae	*Opius*	*ocreatus*	(Papp)
Ichneumonoidea	Braconidae	*Opius*	*pallipes*	Wesmael 1835
Ichneumonoidea	Braconidae	*Opius*	*pteridiophilus*	Wharton & Austin 1990
Ichneumonoidea	Braconidae	*Opius*	*pterus*	Wharton & Austin 1990
Ichneumonoidea	Braconidae	*Opius*	*trimaculatus*	Spinola
Ichneumonoidea	Braconidae	*Opius*	*youi*	Li & van Achterberg 2013
Ichneumonoidea	Braconidae	*Phaenocarpa*	*conspurcator*	(Haliday, 1838)
Ichneumonoidea	Braconidae	*Phaenocarpa*	*drop_ IR1_sp22*	Kimura
Ichneumonoidea	Braconidae	*Phaenocarpa*	*drop_TK1_sp21*	Kimura
Ichneumonoidea	Braconidae	*Phaenocarpa*	*tacita*	Stelfox 1941
Ichneumonoidea	Braconidae	*Phaenocarpa*	*drosophilae*	(Fischer 1975)
Ichneumonoidea	Braconidae	*Tanycarpa*	*bicolor*	(Nees, 1814)
Ichneumonoidea	Braconidae	*Tanycarpa*	*chors*	Belokobylskij 1998
Ichneumonoidea	Braconidae	*Tanycarpa*	*drop_NG1_sp23*	Kimura
Ichneumonoidea	Braconidae	*Tanycarpa*	*punctata*	van Achterberg 1976

Ichneumonoidea	Braconidae		*drop_ Aly_sp47*	
Ichneumonoidea	Braconidae		*drop_ Aly_sp48*	
Ichneumonoidea	Braconidae		*drop_ Aly_sp49*	
Ichneumonoidea	Braconidae		*drop_ Aly_sp50*	
Ichneumonoidea	Braconidae		*drop_ Aly_sp63*	
Ichneumonoidea	Braconidae		*drop_ Aso_sp69*	
Diaprioidea	Diapriidae	*Trichopria*	*anastrephae*	Costa Lima 1940
Diaprioidea	Diapriidae	*Trichopria*	*drop_ BG1_sp37*	Kimura
Diaprioidea	Diapriidae	*Trichopria*	*drop_ Dia70_sp65*	
Diaprioidea	Diapriidae	*Trichopria*	*drop_ Tri_sp44*	
Diaprioidea	Diapriidae	*Trichopria*	*drop_Bdia_sp10*	
Diaprioidea	Diapriidae	*Trichopria*	*drop_Dia127_sp9*	
Diaprioidea	Diapriidae	*Trichopria*	*drop_TK1_sp36*	Kimura
Diaprioidea	Diapriidae	*Trichopria*	*drosophilae*	(Kieffer, 1912)
Diaprioidea	Diapriidae	*Trichopria*	*modesta*	(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
Ganaspis	*brasiliensis*	19	8	868	GCA_009823575.1
Ganaspis	*brasiliensis*	19	16	872	SRX8882993
Ganaspis	*brasiliensis*	19	17	871	SRX8882992
Ganaspis	*drop_ Gsp1_sp67*	182	15	873	SRX8882994
Ganaspis	*drop_ Gsp2_sp68*	183	14	874	SRX8882995
Ganaspis	*drop_ Gsp50_sp66*	181	9	869	GCA_011057455.1
Leptolamina	*ponapensis*	48	13	875	SRX8883008
Leptopilina	*boulardi*	4	5	865	GCA_011634795.1
Leptopilina	*boulardi*	4	6	866	GCA_003121605.1
Leptopilina	*boulardi*	4	12	876	SRX8883009
Leptopilina	*clavipes*	5	7	867	GCA_001855655.1
Leptopilina	*heterotoma*	6	1	861	GCA_010016045.1
Leptopilina	*heterotoma*	6	2	862	GCA_009602685.1
Leptopilina	*heterotoma*	6	3	863	GCA_009026005.1
Leptopilina	*heterotoma*	6	4	864	GCA_009025955.1
Leptopilina	*japonica japonica*	13	11	877	SRX8883011

993 **Table 2.2:** *Drosophila* parasitoid transcriptome data included in DROP.
994

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

995

996 **Table 2.3:** *Drosophila* parasitoid proteomes data included in DROP
997

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

998