DNA barcodes affirm that 16 species of apparently generalist tropical parasitoid flies (Diptera, Tachinidae) are not all generalists

M. Alex Smith*†, D. Monty Wood‡, Daniel H. Janzen†§, Winnie Hallwachs§, and Paul D. N. Hebert*

*Biodiversity Institute of Ontario, University of Guelph, Guelph, ON, Canada N1G 2W1; ‡Diptera Unit, Canadian National Collection of Insects, Agriculture and Agri-Food Canada, Ottawa, ON, Canada K1A 0C6; and §Department of Biology, University of Pennsylvania, Philadelphia, PA 19104-6018

Contributed by Daniel H. Janzen, December 30, 2006 (sent for review December 26, 2006)

Many species of tachinid flies are viewed as generalist parasitoids because what is apparently a single species of fly has been reared from many species of caterpillars. However, an ongoing inventory of the tachinid flies parasitizing thousands of species of caterpillars in Area de Conservación Guanacaste, northwestern Costa Rica, has encountered >400 species of specialist tachinids with only a few generalists. We DNA-barcoded 2,134 flies belonging to what appeared to be the 16 most generalist of the reared tachinid morphospecies and encountered 73 mitochondrial lineages separated by an average of 4% sequence divergence. These lineages are supported by collateral ecological information and, where tested, by independent nuclear markers (28S and ITS1), and we therefore view these lineages as provisional species. Each of the 16 apparently generalist species dissolved into one of four patterns: (*i*) a single generalist species, (*ii*) a pair of morphologically cryptic generalist species, (*iii*) a complex of specialist species plus a generalist, or (*iv*) a complex of specialists with no remaining generalist. In sum, there remained 9 generalist species among the 73 mitochondrial lineages we analyzed, demonstrating that a generalist lifestyle is possible for a tropical caterpillar parasitoid fly. These results reinforce the emerging suspicion that estimates of global species richness are likely underestimates for parasitoids (which may constitute as much as 20% of all animal life) and that the strategy of being a tropical generalist parasitic fly may be yet more unusual than has been envisioned for tachinids.

28S | Area de Conservación Guanacaste | cytochrome *c* oxidase 1 | internal transcribed spacer 1 | species diversity

Parasitoid insects are currently believed to comprise as much as one quarter of all insect species (1, 2) and, because insects comprise ≈80% of all named animal species, up to 20% of all animal life (3). However, accurate evaluations of parasitoid species richness (2), and subsequent determinations of parasitoid host-specificity, are impeded by the very large number of morphologically similar species and the resultant difficulty in identifying them. This situation further complicates the determination of host–parasitoid relationships. There may be many more species of insect parasitoids than currently believed if host-specificity has been underestimated (4, 5). After the Hymenoptera, Diptera (flies) are the most species-rich group of parasitoids, and the obligate parasitoid family Tachinidae is among the most species-rich of Diptera families, with nearly 10,000 described species (1, 6–8). Within this diversity, many described species of Tachinidae are extremely similar morphologically, and it is a taxonomically challenging family.

It is a widely held view that many species of tachinid parasitoids are relatively generalist (polyphagous) in the species of hosts they parasitize (7–10). However, a 29-year inventory of >400 species of tachinids reared from >390,000 wild-caught caterpillars of >3,500 species in Area de Conservación Guanacaste (ACG) in northwestern Costa Rica indicates that at least 90% of the tachinid species from this tropical site are host-specific to one or a few related species (specialists) (ref. 11 and http://janzen.sas.upenn.edu). However, there are conspicuous exceptions. To ascertain whether these exceptions are truly generalists, we cytochrome *c* oxidase 1 (CO1) DNA-barcoded (e.g., as in ref. 12) the 16 most generalist morphospecies, all being species that have been reared by the inventory from many species of caterpillars in a few to many families and all being species reared tens to hundreds of times [see supporting information (SI) Table 1 and *SI Appendices 1* and *2*] . When barcoded, this select group of exceptionally generalist morphospecies dissolved into 64 species of specialists and 9 generalists (Fig. 1 and SI Table 1). This outcome mirrors and magnifies the result recorded when we barcoded the 20 relatively specialist morphospecies of *Belvosia*, another tachinid genus living in the same ACG habitats; its three somewhat generalist species were each found to be complexes of specialists (12).

As of October 2006, the ACG caterpillar and parasitoid inventory had reared tachinid flies from 16,500 of >390,000 wild-caught caterpillars (≈4.2% rate of infection). For the past 17 years, D.M.W. (8, 13) has iteratively assigned these flies to morphospecies and identified them with scientific names when available. The first taxonomic assignment was completed without knowledge of the host caterpillar. Taxonomic assignments were flagged for reexamination when the fly's host caterpillar species did not match what appeared to be the host-use pattern of that fly species. Although all flies were placed to genus (described or undescribed), <10% of the morphospecies appeared to match a known type specimen and thus could be presently "named" with any confidence. Inasmuch as it is believed that only 7–20% of insect species have been scientifically described (3, 14), such a low level of taxonomic allocation is not surprising. However, all of the 16 most generalist morphospecies were sufficiently distinctive among the hundreds of species-, genus-, or family-level specialists that they did receive a tentative scientific name through this process; on average, these names are all >100 years old (SI Table 2). We then added sequence data from CO1 barcoding to the 16 morphologically defined and host-checked units already recognized to determine whether each generalist morphospecies comprised specimens with little intraspecific barcode variability. Such a protocol determines whether these 16 species can

Author contributions: M.A.S., D.H.J., W.H., and P.D.N.H. designed research; M.A.S., D.M.W., D.H.J., and W.H. performed research; M.A.S., D.H.J., and W.H. analyzed data; and M.A.S., D.H.J., W.H., and P.D.N.H. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Abbreviations: ACG, Area de Conservación Guanacaste, Costa Rica; CO1, cytochrome *c* oxidase 1; NJ, neighbor-joining.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (CO1: accession nos. EF180450–EF182583; 28S and ITS1: accession nos. EF183546–EF184019 and EF189688–EF189703 and two representative sequences of *Chetogena* scutellarisDHJ01 *Wolbachia*, accession nos. EF192042 and EF192043).

See Commentary on page 4775.

†To whom correspondence may be addressed. E-mail: salex@uoguelph.ca or djanzen@sas.upenn.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0700050104/DC1.

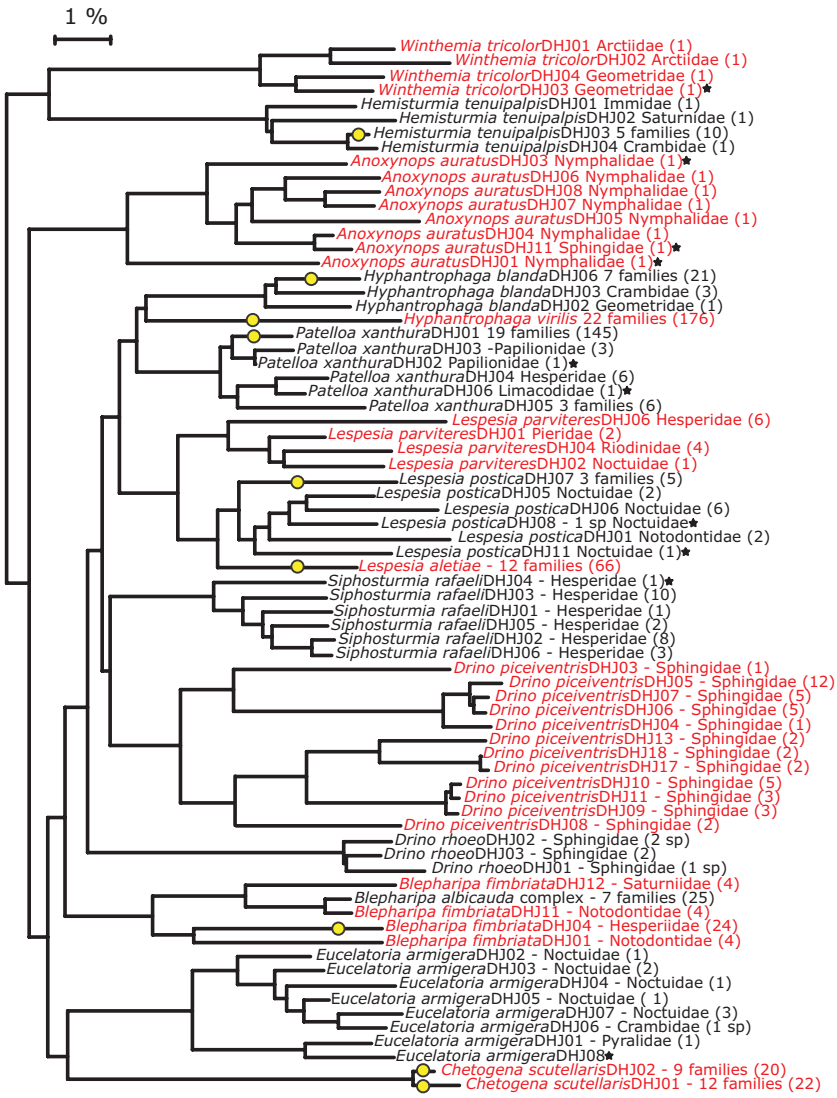

Fig. 1. NJ tree of genetic distance (K2P) for 73 specimens, each representing one of the barcode clusters (provisional species) encountered among 16 generalist morphospecies. See *SI Appendix 1* for an NJ tree containing all >500-bp barcoded individuals. The number of species of Lepidoptera parasitized by each parasitoid is shown in parentheses. Yellow circles flag species that remained generalist after analysis, and a star indicates those lineages represented by only one specimen. Note that the "*Blepharia albicauda*" species complex is here represented by only one specimen due to extremely restricted barcode divergences among its potential provisional species. Alternating red and black text does not represent any taxonomic pattern and is only to facilitate visual discrimination between adjacent morphospecies.

be easily distinguished by their barcodes, as has been effective for other taxa (12, 15–21), and is also the first step for using barcodes in the discovery of cryptic species (12, 15).

Results

We successfully barcoded one fly from each of 2,134 generalist tachinid rearings (among >4,500 barcoded flies of >100 morphospecies) and found that 14 of the 16 generalist morphospecies were readily distinguishable from all others by their DNA barcodes (Fig. 1, SI Table 1, and *SI Appendix 1*) but that barcodes of two morphospecies were only very slightly divergent from each other (although they were easily distinguishable from all others, i.e., *Blepharia albicauda* and *B. fimbriata*). Each of the 14 morphospecies is represented by a distinct, nonoverlapping cluster of sequences in the neighbor-joining (NJ) tree (*SI Appendix 1*). These barcode clusters had >5% sequence divergence among them, but the sequence divergence within each morphospecies was quite high, in some cases even greater than 10%. This finding suggested that the applied morphospecies name actually referenced multiple cryptic species. We then overlaid an array of ecological correlates [host caterpillar, caterpillar food plant, ecosystem (see *SI Appendix 2*)] and independent nuclear genetic markers (ITS1, 28S) to determine whether these data as a whole supported the hypothesis that each of these within-morphospecies barcode clusters represents a morphologically cryptic provisional species. This was unambiguously the case. Sixteen generalists became 73 cryptic provisional species. Only two of these, *Hyphantrophaga virilis* and *Lespesia aletiae*, could be matched confidently to an established name, and these two also displayed no internal barcode divergence. The clusters within the remainder of the morphospecies were then labeled with alphanumeric interim names (Fig. 1, SI Table 1, and *SI Appendices 1* and *2*), where the naming reflects the order in which they were encountered by the inventory. The name originally applied to a morphospecies is retained within the interim name as a reference point but is not meant as a firm scientific identification. It is unlikely that any of these cryptic species are actually conspecific with the holotypes that reference the 16 names, inasmuch as none of these holotypes are Costa Rican (see SI Table 2).

In those cases in which a barcode sequence cluster within 1 of the 16 morphospecies did not correlate with ecological information, or alternatively, a barcode cluster was only slightly distinct from its immediate neighbor, we also sequenced two independent nuclear markers. As with barcoding of *Belvosia*, we used the first nuclear rRNA internal transcribed spacer region (ITS1) (12), as well as the D2 region of 28S. A nuclear marker was not sequenced for all specimens for which we have CO1 barcodes because our purpose was species identification and the detection of cryptic provisional species within generalist morphospecies, which was often achieved with the standardized CO1 barcode alone.

When the CO1 barcode information was added to the ecological and nuclear sequence information, what appear to be 73 species fell into four patterns. SI Table 3 enumerates the intramorphospecific CO1 DNA barcode variation.

Pattern 1: A Barcoded Generalist Morphospecies Remains a Generalist. Just 2 of the 16 most generalist morphospecies encountered by the ACG tachinid inventory, *H. virilis* and *L. aletiae*, barcoded as single biological units. Because there is no evidence that they are more than two species, we feel confident in applying a scientific name to these Costa Rican specimens, even though these widespread species were described from the United States. We barcoded flies from 135 *H. virilis* rearings, which represent collections of at least 153 caterpillar species from 15 families (SI Table 1 and *SI Appendix 2*). We found essentially no barcode variation. The average CO1 intraspecific divergence was 0.01% (SE = 0.04). Likewise, *L. aletiae* barcoded as a single biological unit from 221 rearings spread across at least 55 species of caterpillars in 12 families. Here, the average CO1 intraspecific divergence was 0.07% (SE = 0.11). We cannot eliminate the possibility that cryptic tachinid species remain hidden among the flies of these two species reared from this morass of caterpillar diversity, but if they exist, they are yet more hidden than are the other cryptic species discussed below.

Pattern 2: The Barcoded Generalist Becomes Two Generalists. "*Chetogena scutellaris*" contained two barcode groups: *C.* scutellarisDHJ01 and *C.* scutellarisDHJ02. These groups differ by eight (1.23%) characteristic nucleotide substitutions within the CO1 barcode (one C–G transversion and seven transitions; six of the transitions are synonymous third-position substitutions, and one is a nonsynonymous first position) (Fig. 1, SI Table 1, and *SI Appendices 1* and *2*). These two provisional species are sympatric within the ACG dry forest, and each uses a multifamily list of hosts that overlap substantially. These two provisional species have been encountered with equal frequency in the caterpillar inventory.

Could there be cryptic species that are not distinguishable by their CO1 barcodes hidden within these two generalists? Specimens of *C.* scutellarisDHJ01 and *C.* scutellarisDHJ02 were also sequenced for 28S and six for ITS1 (*SI Appendix 3*). There is just 1 bp difference between *C.* scutellarisDHJ01 and *C.* scutellarisDHJ02 within 28S (*SI Appendix 3b*), and the sequences are homogeneous within each of the two provisional species. Additionally, there are five heteroplasmic individuals at that locus, which suggests that the two provisional species are capable of producing hybrids. The more-variable rRNA gene region (ITS1; *SI Appendix 3a*) supports the divergence between *C.* scutellarisDHJ01 and *C.* scutellarisDHJ02 (≈7% divergent) and further suggests that there is a division within *C.* scutellarisDHJ01 that is not apparent to the CO1 barcode.

The absence of mtDNA variation with evident nuclear variation, as displayed here by *C.* scutellarisDHJ01, can sometimes be caused by the presence of the cytoplasmic bacteria *Wolbachia*. *Wolbachia* are obligate intracellular endosymbiotic bacteria that cause reproductive incompatibility between infected and uninfected lineages, which results in an increased proportion of infected maternal lineages that cannot reproduce (22, 23). If closely related species hybridize and one member of this pair is infected with *Wolbachia*, the bacteria-caused cytoplasmic incompatibility can result in the lack of interspecific mtDNA variability with little effect on nuclear genes (24). Using a standard PCR diagnostic for the presence of *Wolbachia* (25), we determined that at least 73.9% of 23 specimens of *C.* scutellarisDHJ01 contained *Wolbachia*. Until more is known about the relationships within the "*C. scutellaris*" provisional species, we hypothesize that a *Wolbachia* infection may have caused the barcode to be identical between two morphologically cryptic species that can otherwise be differentiated by their nuclear DNA. In short, the two clearly cryptic generalists within "*C. scutellaris*" might even be three cryptic generalists.

Pattern 3: The Barcoded Generalist Becomes Multiple Specialists and One Generalist.

- **"*Patelloa xanthura*":** *P.* xanthuraDHJ01 was found to be one provisional species reared from at least 145 species of caterpillars in 19 families, unambiguously a generalist, and five specialists. Each of these specialists parasitizes a very restricted number of species in a different family of caterpillars (Fig. 1, SI Table 1 and *SI Appendices 1* and *2*).
- **"*Hyphantrophaga blanda*":** *H.* blandaDHJ06 was found to be one barcode group reared from at least 21 species of caterpillars in seven families-unambiguously a generalist- and two specialists: *H.* blandaDHJ02 and *H.* blandaDHJ03. Each of the latter use a very narrow and nonoverlapping range of caterpillar species, species that are not attacked by *H.* blandaDHJ06 (see SI Table 1 and *SI Appendix 2*). These three provisional species are supported by distinct haplotypes within both ITS1 and 28S (Fig. 2 and *SI Appendix 4 a* and *b*).
- **"*Hemisturmia tenuipalpis*":** *H.* tenuipalpisDHJ03 was found to be one generalist reared from at least 10 species and five families. However, *H.* tenuipalpisDHJ01 is a specialist on Immidae; *H.* tenuipalpisDHJ02 specializes on one species, *Hylesia umbrata* (Saturniidae); and *H.* tenuipalpisDHJ04 specializes on two members of one genus, *Cliniodes* (Crambidae). All of these species of caterpillars occur at high densities, but only on a few individuals of their food plants. *H.* tenuipalpisDHJ03 also contains shallow CO1 divergences that we hypothesized might be two marginally separated species, but this was not supported by ecological information or independent genetic markers [ITS1, 28S (see *SI Appendix 5 a* and *b*)]. We therefore conclude that *H.* tenuipalpisDHJ03 is one provisional species.
- **"*Lespesia postica*":** One morphospecies splits into six when barcoded, and each provisional species parasitizes a different species or a small closely related group of species of caterpillars.
- **"*Blepharipa fimbriata*" and "*Blepharipa albicauda*":** Upon barcoding, "*B. fimbriata*" becomes four distinct groups, with *B.* fimbriataDHJ04 appearing to be a generalist feeding on at least 24 species in two families. However, all but one of the *B.* fimbriataDHJ04 host records is from a large hesperiine Hesperiidae caterpillar feeding on a monocot (from grasses to palms) (*SI Appendices 1* and *2*). We therefore conclude that this species is a generalist but only within one subfamily of hosts. The single exceptional host species record is *Manataria maculata* (Nymphalidae), which feeds side-by-side with the usual hosts. The other three species of "*B. fimbriata*" each specialize on very taxonomically restricted groups of caterpillars. *B.* fimbriataDHJ11 displays more intraspecific variation than all of the other provisional specialist barcode species, including some instances of heteroplasmy, and the COI barcodes of all specimens are mingled among those of "*B. albicauda.*" *B. albicauda* is a putative generalist species that is morphologically quite distinctive because of a whitish tip (tergite 5) to an otherwise mostly black abdomen, whereas "*B. fimbriata*" has the classical gray and black striped/banded pattern of tachinids (*SI Appendix 6d*). The CO1 barcodes of some individuals of *B.* fimbriataDHJ11 are too similar to those

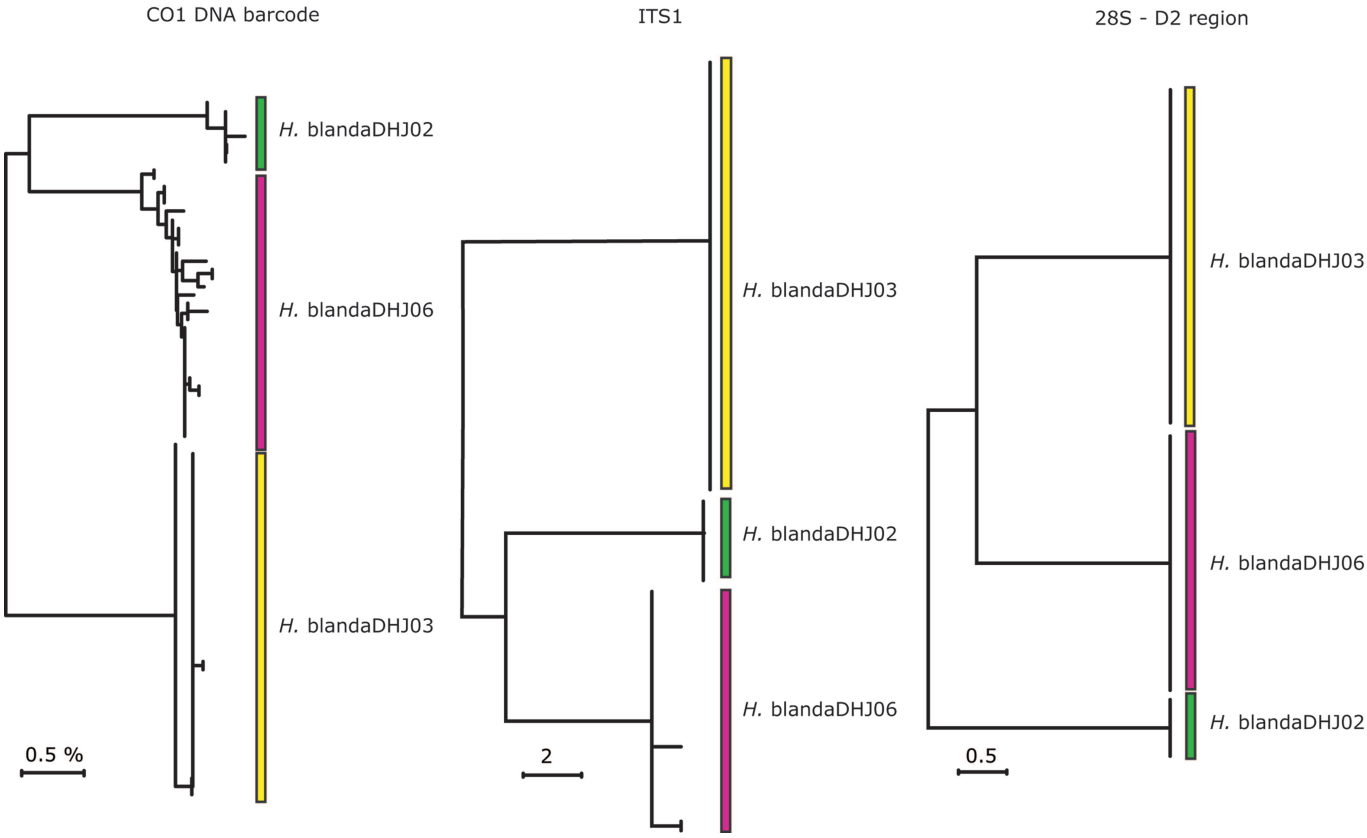

Fig. 2. CO1 barcodes facilitate species discovery in the absence of established alpha taxonomy. In the exemplary case of the apparently generalist morphospecies "*H. blanda*," barcodes separated the flies into two cryptic specialists (*H.* blandaDHJ02 and *H.* blandaDHJ03) and one generalist (*H.* blandaDHJ06). rRNA sequences (both ITS1 and 28S) reveal the same patterns of diversity. This barcode NJ tree is made with complete deletion K2P distances. The rRNA NJ trees are made with pairwise deletion and total number of differences.

of "*B. albicauda*" to be separated from them (Fig. 1 and *SI Appendices 1* and *6d*). There is no overlap between the two arrays of caterpillars that are parasitized by these two totally sympatric and barcode-indistinguishable provisional species (*SI Appendix 1*). The morphospecies "*B. albicauda*" is a morphologically distinct generalist that has been reared from 18 species in seven families of caterpillars (but see below).

Because "*B. albicauda*" and *B.* fimbriataDHJ11 are a special case from both a taxonomic and a molecular perspective, they were investigated in more detail. ITS1 and 28S were sequenced for 137 individuals from "*B. albicauda*" and *B.* fimbriataDHJ11. The two morphologically defined species showed <0.05% divergence at ITS (three transitions), reinforcing their close affinities because such variation fell within that seen in "*B. albicauda*" (including indels, which range from 0 to 11 differences). Furthermore, within the D2 region of 28S, *B.* fimbriataDHJ11 possesses the same haplotype as approximately half of the "*B. albicauda*" (there are two haplotypes within *B. albicauda* differing by a 1-bp insertion) (*SI Appendix 6*).

Considering only molecular data, it appears that the morphospecies "*B. albicauda*" is one internal member of the "*B. fimbriata*" species complex. However, both the absence of overlap in host use (none of the *B.* fimbriataDHJ11 hosts, i.e., all *Lirimiris*, *Heterocampa*, and *Farigia* in the Notodontidae, are parasitized by *B. albicauda*) and the distinctive color pattern suggest that "*B. albicauda*" can be comfortably viewed as its own evolutionary lineage.

We do not here consider "*B. albicauda*" to be a true generalist because 28S and ITS1 divergences are correlated with the fly's host species. The D2 region for 28S supports a division between flies that parasitize the *Hemiceras pallidula*/*Encruphion leena* group of hosts and those that primarily parasitize *Hemiceras nigrescens*, in that there is an insertion characteristic of the latter (*SI Appendix 6 b* and *d*). Furthermore, there is a different ITS1 haplotype for flies reared from *Phoebis sennae* (Pieridae), another for *Hylesia lineata* (Saturniidae), and a third for *Pentobesa pinna* and *Moresa valkeri* (Notodontidae) together (*SI Appendix 6*). We think it likely that there are cryptic species within *B. albicauda* (*SI Appendix 6d*) and that these are characterized by extremely shallow CO1 divisions. However, elucidation of this possibility will require larger sample sizes and other data to confirm. The observed heteroplasmy could be caused by the differential amplification of nuclear pseudogenes (26), by paternal leakage (27), or by somatic mutation (28). If these CO1 sequences are pseudogenes, it might explain why the evident diversity within ITS1 is not reflected within the barcode, given that nucleotide substitutions are not expected to occur as rapidly in the nucleus as in the mitochondrial genome (29). However, as pseudogenes are freed from functional constraint, nonsynonymous and synonymous substitutions should occur at the same rate. However, in this case the alternative base at each heteroplasmic site represents a synonymous substitution. This observation does not support the pseudogene hypothesis. Additional information that makes unintentional amplification of pseudogenes unlikely includes uniform electrophoretic bands, lack of indels, stop codons, or a transversion bias. The CO1 trace files of those likely hybrid individuals (see panel *B* of *SI Appendix 6 b* and *d*) display examples in which the evident CO1 heteroplasmy was the nucleotide of the provisional species that is the

other inferred parent. This finding supports the hypothesis that there has been leakage of small amounts of paternal mtDNA. Irrespective of whether this heteroplasmy is a pseudogene or the outcome of paternal leakage, from the perspective of species identification by barcoding, the single mtDNA barcode alone is not sufficient to differentiate these two species. None of the specimens tested positive for *Wolbachia.*

Pattern 4: The Barcoded Generalist Is a Complex of Specialists. Although the above eight morphospecies that appeared to be generalists retained at least one generalist lineage (except for the suspiciously complex "*B. albicauda*"), the other eight morphospecies that were barcoded appear to be constituted entirely of complexes of very similar species distinguishable by barcode and host caterpillar correlates (Fig. 1, SI Table 1, and *SI Appendices 1* and *2*). The scientific names of the remainder of the morphospecies are given in quotation marks because we have no idea which, if any, of the barcode sequence clusters truly match the holotype.

- **"*Winthemia tricolor*":** *W.* tricolorDHJ01 and *W.* tricolorDHJ02 are each specialists on a different species of Arctiidae, whereas *W.* tricolorDHJ03 and *W.* tricolorDHJ04 are specialists on each of two species of Geometridae. Specific ITS1 and 28S haplotypes are associated with each of the CO1-distinctive provisional species identified here, including the two provisional species represented by a single collection (*W.* tricolorDHJ03) (see *SI Appendix 7*).
- **"*Anoxynops auratus*":** Four of the specialists (*A.* auratusDHJ01, *A.* auratusDHJ03, *A.* auratusDHJ04, *A.* auratusDHJ05) parasitize four quite different sets of Nymphalidae. *A.* auratusDHJ01 and *A.* auratusDHJ03 are represented by single specimens but are further supported by having characteristic ITS1 and 28S haplotypes. The three remaining specialists (*A.* auratusDHJ06, *A.* auratusDHJ07, and *A.* auratusDHJ08) each parasitize a different member (and microhabitat) of the nymphalid genus *Opsiphanes* eating different food plants: broadleaf monocots, rainforest palms, and dry forest palms, respectively. Although these parasitoids of *Opsiphanes* are not variable within 28S, each bears a distinctive haplotype within the more variable ITS1 gene region (see *SI Appendix 8*).
- **"*Siphosturmia rafaeli*":** This morphospecies splits into four with barcoding. Each parasitizes a different taxonomic and ecological group of hesperiid butterfly caterpillars.
- **'*Lespesia parviteres*":** One morphospecies from 178 rearings barcodes into four specialists, each using a different family of caterpillars.
- **"*Eucelatoria armigera*":** One morphospecies barcodes into eight, and each provisional species parasitizes a different species or closely related small group of species.
- **"*Drino rhoeo*":** One morphospecies barcodes into three specialists on different species of *Manduca*; these specialists are ≈2% different in their CO1 barcodes.
- **"*Drino piceiventris*":** One morphospecies splits into 12 host-specific lineages, each of which specializes on different groups of Sphingidae. One (*D.* piceiventrisDHJ05) is a specialist on seven species of *Manduca* and a few other large similar macroglossine sphingid caterpillars, whereas three others (*D.* piceiventrisDHJ07, *D.* piceiventrisDHJ10, and *D.* piceiventrisDHJ13) specialize on different sphingid caterpillars that feed on Dilleniaceae. Many of the barcode divisions within "*D. piceiventris*" are very slight but consistent, sometimes as little as a single base pair divergence (a value of note because it is within sequencing error). Members of the "*D. piceiventris*" species complex were also sequenced for ITS1 and 28S (see *SI Appendix 9*). These gene regions did not contain haplotypes that provide independent support for the lesser barcode divergences (*D.* piceiventrisDHJ09, *D.* piceiventrisDHJ10, and *D.* piceiventrisDHJ11) but did support the greater CO1 divergences (e.g., between *D.* piceiventrisDHJ03 and *D.* piceiventrisDHJ06). On the basis of these results, in a certain sense we have maintained the recognition of these slightly separate provisional species more on the basis of ecological information than based strictly on barcode data.

Discussion

CO1 DNA barcoding has shown that what were thought to be 16 morphospecies of apparently generalist tachinid fly parasitoids are in fact a complex of at least 73 species, and that except for two (and potentially several more within "*B. albicauda*") all can be identified by their barcodes. Barcoding is not only an effective identification tool for these small and similar parasitoids, but it has also played a major role in discovering the existence of many provisional species among them. It has helped to bring clarity to the degree of host specificity within the 16 morphospecies of flies and suggests instances in which seemingly small variation in morphology reflects distinguishing traits of cryptic lineages. We say this because subsequent iterative morphological examination of the provisional tachinid species located with our barcoding is finding that some of these provisional species do indeed have distinguishing morphological traits, traits that were previously ascribed to intraspecific rather than interspecific variation.

Additionally, in this understudied group of tropical insects of expected high diversity, we encountered just two cases in which we might have overlooked a provisional species if we had used barcodes alone to analyze the 16 species of what appeared initially to be generalist morphospecies. In one case ("*B. albicauda*"), where ecological information and an independent nuclear marker support very slight CO1 differences, there also are diagnostic CO1 base pair substitutions. In the other case (*C.* scutellarisDHJ01), where an evident nuclear divergence was shown to be invariant for CO1, most of these specimens tested positive for the bacteria *Wolbachia.* Because females uninfected by *Wolbachia* can only breed successfully with uninfected males of the same species, the stage would be set for a sweep of the infected species' mtDNA through the uninfected species (as discussed in refs. 30 and 31).

CO1 provides an attractive genetic barcode for species identification because of its high copy number, rapid rate of mutation, and ease of amplification/sequencing and alignment for intra- and interspecific comparisons. However, with very young species, or species that can hybridize, a secondary independent molecular marker to solidify or confirm identification may be needed. This is especially true for extraordinarily species-rich and often morphologically very similar groups such as parasitoids, for which alpha taxonomic description based on morphology and behavior alone lags far behind existing diversity. Both of the secondary nuclear markers used here are rRNA and are therefore attractive because of their great abundance and relatively conserved flanking regions, which allows the design of primers of wide utility. However, successfully sequencing from regions with large indels, without cloning and subsequent aligning, is difficult (32). Because of this, we do not use or suggest the rRNA data as a substitute for the CO1 barcode. We apply it here to species complexes or pairs with slight CO1 differentiation in cases where ecological data suggested that the slight divergence was meaningful. It would be computationally and methodologically complex to conduct taxonomically broad sequencing and subsequent comparisons/identifications with ITS1 sequence data, and there is unlikely to be sufficient resolving power within 28S for many species. However, ITS1 and 28S are useful independent (from CO1) genetic covariates to help interpret hybridization and branching patterns of young species when a mitochondrial marker alone is insufficient. Nuclear sequence divergence correlated with CO1 barcode divergence is also particularly useful for demonstrating that two sympatric CO1

barcode lineages are two separate breeding entities rather than simply two haplotypes in one interbreeding population.

With the finding that 14 of 16 apparently generalist tachinid flies are rich in morphologically cryptic provisional species, these species can now be seen as conforming in large part to the pattern of strong specialization displayed by hundreds of other species reared by the ACG caterpillar inventory. In addition to the nearly 5-fold increase in species richness facilitated by DNA barcoding, perhaps five additional species remain to be characterized; these are provisional species that cannot be distinguished from closely related species by simple threshold interpretations of their CO1 barcodes. At the other end of the scale, the 9–10 species that appear to be genuine generalists pose difficult questions. How do they manage to use so many different taxa as hosts? Why are generalist parasitoids no more abundant than are most of the hundreds of host specialists? Are generalists to be thought of as an evolutionary source of specialists (as opposed to specialists evolutionarily begetting specialists)? Some of these questions will disappear if further molecular probing and larger sample sizes expose yet more cryptic diversity invisible to the level of examination we used. It is also possible that the generalists contain extremely rare specialists that have not yet been well represented by collections in the caterpillar inventory. For instance, see "*P. xanthura*" (Fig. 1 and *SI Appendices 1* and *2*). In this case, of 337 specimens sequenced, 83% were the generalist *P.* xanthuraDHJ01, whereas *P.* xanthuraDHJ02 and *P.* xanthuraDHJ06 are represented by only one specimen each (0.3%) in the caterpillar inventory to date.

Reviews of the nature of tachinid host use have concluded that, as a whole, the family is generalist, and most variation on this pattern is due to a small subset of even more extremely generalist species (33). However, this conclusion is based almost entirely on host records from temperate regions and on generalist "species" as defined by their morphology. Our results suggest that barcoding a large number of presumed generalist temperate tachinids reared from many species and families of carefully identified caterpillars might modify this conclusion. Irrespective of what happens in a tropical–temperate comparison, our results suggest that combining barcoding with morphology and natural history is very likely to increase global estimates of species richness and to expose the tropics as being yet more complex than currently appreciated. Species that are "so different from each other, and dependent on each other in so complex a manner" (34) may result in interactions of tropical species that are even more complex than we have yet realized. Correctly detecting and identifying these species will greatly facilitate our ability to unravel these entangled banks of ecological and evolutionary interactions.

Materials and Methods

All methods (field and molecular biology) were completed as described in the DNA barcoding of *Belvosia* (12), and slight modifications are detailed in *SI Materials and Methods* and SI Table 4. Sequences, trace files, and field data are available in the ACG Generalist Tachinidae file in the Completed Projects section of the Barcode of Life Database (BOLD; www.barcodinglife.org). Additional collection information is deposited at http://janzen.sas.upenn.edu, and all sequences have been deposited in the GenBank database (CO1: accession nos. EF180450–EF182583; 28S and ITS1: accession nos. EF183546–EF184019 and EF189688–EF189703 and two representative sequences of *C.* scutellarisDHJ01 *Wolbachia*, accession nos. EF192042 and EF192043).

We thank our many colleagues at Guelph, especially those at the Canadian Centre for DNA Barcoding and the Biodiversity Institute of Ontario. We thank Tanya Dapkey and Cathy Hulshof for delegging fly specimens and processing data and the 21 ACG parataxonomists for decades of collecting, rearing, and databasing caterpillars and parasitoids. We thank the Diptera Unit of the Canadian National Collection of Insects, Agriculture and Agri-Foods, Ottawa, for ongoing housing and caring for the ACG tachinid voucher collection deposited there and for facilities provided for this study. The manuscript benefited from the constructive comments of Dirk Steinke, Scott Miller, Jim Whitfield, Grace Wood, and Al Herre. This study would never have occurred, nor could the analysis have been conducted, without the taxonomic and identification support of more than 150 taxonomists who have identified Lepidoptera and plants for the ACG caterpillar and parasitoid inventory. This work was supported by grants from The Gordon and Betty Moore Foundation, the Natural Sciences and Engineering Research Council of Canada, and the Canada Research Chairs program (all to P.D.N.H.); a Fonds québécois de la Recherche sur la Nature et les Technologies B3 postdoctoral fellowship (to M.A.S.); U.S. National Science Foundation Grants BSR 9024770 and DEB 9306296, 9400829, 9705072, 0072730, and 0515699 (to D.H.J.); and grants from Guanacaste Dry Forest Conservation Fund and Area de Conservación Guanacaste (to D.H.J.) and from INBio (to D.M.W. and D.H.J.).

1. Godfray HCJ (1994) *Parasitoids: Behavioral and Evolutionary Ecology* (Princeton Univ Press, Princeton).
2. Godfray HCJ, Shimada M (1999) *Res Popul Ecol* 41:3–10.
3. Samways MJ (1993) *Biodivers Conserv* 2:258–282.
4. Bensch S, Perez-Tris J, Waldenstrom J, Hellgren O (2004) *Evolution (Lawrence, Kans)* 58:1617–1621.
5. Westenberger SJ, Sturm NR, Yanega D, Podlipaev SA, Zeledon R, Campbell DA, Maslov DA (2004) *Parasitology* 129:537–547.
6. Gaston KJ (1991) *Conserv Biol* 5:283–296.
7. Stireman JO, O'Hara JE, Wood DM (2006) *Annu Rev Entomol* 51:525–555.
8. Wood DM (1987) in *Manual of Nearctic Diptera*, ed McAlpine JF (Agric Canada, Ottawa), Vol 2, pp 1193–1269.
9. Eggleton P, Gaston KJ (1992) *Entomol Gaz* 43:139–143.
10. Stireman JO, III, Dyer LA, Janzen DH, Singer MS, Lill JT, Marquis RJ, Ricklefs RE, Gentry GL, Hallwachs W, Coley PD, *et al.* (2005) *Proc Natl Acad Sci USA* 102:17384–17387.
11. Janzen DH (1995) *The Tachinid Times* 8:2–5.
12. Smith MA, Woodley NE, Janzen DH, Hallwachs W, Hebert PDN (2006) *Proc Natl Acad Sci USA* 103:3657–3662.
13. Wood DM (1985) *Mem Entomol Soc Can* 132:1–130.
14. Godfray HCJ, Lewis OT, Memmott J (1999) *Philos Trans R Soc London B* 354:1811–1824.
15. Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) *Proc Natl Acad Sci USA* 101:14812–14817.
16. Ball SL, Hebert PDN, Burian SK, Webb JM (2005) *J North Am Benthol Soc* 24:508–524.
17. Hogg ID, Hebert PDN (2004) *Can J Zool* 82:749–754.
18. Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) *Philos Trans R Soc London B* 360:1847–1857.
19. Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003) *Proc R Soc London Ser B* 270:313–321.
20. Smith MA, Fisher BL, Hebert PDN (2005) *Philos Trans R Soc London B* 360:1825–1834.
21. Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004) *PLoS Biol* 2:e312.
22. Haine ER, Martin J, Cook JM (2006) *BMC Evol Biol* 6:83.
23. Moran N, Baumann P (1994) *Trends Ecol Evol* 9:15–20.
24. Hurst GDD, Jiggins FM (2005) *Proc R Soc London Ser B* 272:25–1534.
25. Braig HR, Zhou W, Dobson SL, O'Neill SL (1998) *J Bacteriol* 180:2373–2378.
26. Bensasson D, Zhang D, Hartl DL, Hewitt GM (2001) *Trends Ecol Evol* 16:314–321.
27. Kvist L, Martens J, Nazarenko AA, Orell M (2003) *Mol Biol Evol* 20:243–247.
28. Moum T, Bakke I (2001) *Curr Genet* 39:198–203.
29. Brown WM, George M, Wilson AC (1979) *Proc Natl Acad Sci USA* 76:1967–1971.
30. Narita S, Nomura M, Kato Y, Fukatsu T (2006) *Mol Ecol* 15:1095–1108.
31. Giordano R, Jackson JJ, Robertson HM (1997) *Proc Natl Acad Sci USA* 94:11439–11444.
32. Gillespie JJ, Yoder MJ, Wharton RA (2005) *J Mol Evol* 61:114–137.
33. Stireman I, John O, Singer MS (2003) *Oecologia* 135:629–638.
34. Darwin C (1859) *The Origin of Species* (John Murray, London).