Systematic Entomology

Systematic Entomology (2018), DOI: 10.1111/syen.12296

Towards holomorphology in entomology: rapid and cost-effective adult–larva matching using NGS barcodes

DARREN YEO[1], JAYANTHI PUNIAMOORTHY[2], ROBIN WEN JIANG NGIAM[3] and RUDOLF MEIER[1]

[1]Department of Biological Sciences, National University of Singapore, Singapore, Singapore, [2]Lee Kong Chian Natural History Museum, Singapore, Singapore and [3]National Parks Board, Singapore, Singapore

Abstract. In many taxa the morphology of females and immatures is poorly known because species descriptions and identification tools have a male bias. The root causes are problems with matching life-history stages and genders belonging to the same species. Such matching is time-consuming when conventional methods are used (e.g. rearing) and expensive when the stages are matched with DNA barcodes. Unfortunately, the lack of associations is not a trivial problem because it renders a large part of the phenome of insects unexplored, although larvae and females are useful sources of characters for descriptive and phylogenetic purposes. In addition, many collectors intentionally avoid females and immature stages, which skews survey results, interferes with collecting life-history information, and makes it less likely that rare species are discovered. These problems even exist for well-studied taxa like Odonata, where obtaining adult–larva matches relies largely on rearing. Here we demonstrate how the matching problem can be addressed with cost-effective tagged amplicon sequencing of a 313-bp segment of *cox1* with next-generation sequencing (NGS) ('NGS barcoding'). We illustrate the value of this approach based on Singapore's odonate fauna which is of a similar size as the European fauna (Singapore, 122 extant species; Europe, 138 recorded species). We match the larvae and adults of 59 species by first creating a barcode database for 338 identified adult specimens representing 83 species. We then sequence 1178 larvae from a wide range of sources. We successfully barcode 1123 specimens, which leads to adult–larva matches for 59 species based on our own barcodes (55) and online barcode databases (4). With these additions, 84 of the 131 species recorded in Singapore have now been associated with a species name. Most common species are now matched (83%), and good progress has been made for vulnerable/near-threatened (55%), endangered (53%), and critically endangered species (38%). We used nondestructive DNA extraction methods in order to be able to use high-resolution imaging of matched larvae to establish a publicly available digital reference collection for odonates which is incorporated into 'Biodiversity of Singapore' (https://singapore.biodiversity.online/). We suggest that the methods described here are suitable for many insect taxa because NGS barcoding allows for fast and low-cost matching of well-studied life-history stages with neglected semaphoronts (eggs, larvae, females). We estimate that the specimen-specific amplicons in this study (*c*. 1500 specimens) can now be obtained within eight working days and that the laboratory and sequencing cost is *c*. US$600 (< US$0.40 per specimen).

Correspondence: Meier Rudolf, Department of Biological Sciences. National University of Singapore, 14 Science Drive 4, #05–20 Singapore 117543. E-mail: meier@nus.edu.sg

Introduction

One of the major challenges in entomology is obtaining a complete and comprehensive inventory of all insect species and their life-history stages. This task is particularly overwhelming in the tropics (Basset *et al.*, 2012) with its high diversity, shorter history of study and insufficient resources for taxonomic research. While the species diversity challenge is frequently discussed in the literature, another challenge receives much less attention, i.e. the complex phenomes of most insect species. For the vast majority of species, the immatures are not miniature adults. Instead they are morphologically so disparate that the morphological distinction between life-history stages of the same species is larger than the morphological disparity between the same stages of closely related species. Yet, many species descriptions focus on adult males only because they tend to have species-specific differences in their sclerotized genitalia (Eberhard, 1985). This leaves the morphological diversity of females and immatures underexplored although they can be morphologically very diverse (Meier, 1995; Puniamoorthy *et al.*, 2010) and can be a rich source of characters for phylogenetic analyses (Beutel, 1993; Meier & Lim, 2009; Beutel *et al.*, 2010; Aspöck *et al.*, 2012). The male bias in entomology has led to a lack of morphological identification tools for immatures and females, which in turn has affected the efficacy of arthropod surveys. In such surveys, female and immature specimens are often not even collected and/or they are ignored/discarded during sorting. This is unfortunate because life-history theory predicts that the abundance of immatures will be higher than the abundance of males and females combined (Pianka, 1970), i.e. only identifying male specimens will significantly weaken the conclusions that can be drawn from survey data. It will also exacerbate the rarity problem in arthropod biodiversity surveys. A large number of species are so rare that they are only known or described based on a single specimen or collecting event (Novotný & Basset, 2000; Lim *et al.*, 2011). Yet a substantial proportion of specimens obtained in surveys is not processed, i.e. the commonness-of-rarity problem is partially caused by not assessing diversity based on all life-history stages.

Including females and immatures in surveys is also important for our knowledge of insect natural history and for conserving species. In many species, females and larvae play a larger role in biomass turnover than males. For example, females tend to have higher caloric requirements as they have to build larger gametes. Similarly, immatures have to build the biomass of the species and are therefore often ecologically more important than adults. At the same time, larvae tend to be more sedentary and long-lived than adult stages. This often means that larvae are more likely to be affected by anthropogenic threats such as habitat loss and pollution (Kraus & Secor, 2005). But entomologists cannot identify the immatures of many species, which becomes an impediment to conservation because the habitat preferences of the immature stages remain unknown. The lack of identification tools is a particularly significant problem in freshwater biomonitoring, which is heavily dependent on information gathered from immature insects. The immatures are frequently only identified to taxa above the species rank, which often leads to coarse habitat assessments (Gaston, 2000).

The lack of adult–larva associations in many insect groups is partially due to the fact that matching life-history stages is time-consuming. Traditionally, associations have been obtained by rearing larvae to adulthood and then identifying the adults to species (van Gossum *et al.*, 2003). However, this approach has drawbacks. First, not all species are amenable to rearing in captivity because they have narrow and unknown dietary and environmental requirements. Secondly, even when rearing is technically feasible, the life cycles of some species are so long that rearing is time-consuming. Lastly, the process of rearing partially destroys the larval morphology unless living larvae can be imaged in sufficient detail for morphological descriptions and/or the exuviae preserve the necessary information. These problems can sometimes be overcome by rearing only some individuals belonging to the same clutch to adulthood. However, association via rearing rarely starts with eggs and more likely involves multiple larvae collected at the same time and site.

More recently, DNA barcoding (Hebert *et al.*, 2003) has been used to obtain species-level adult–larva associations. This technique has been widely applied across a variety of insect groups, such as Coleoptera (Miller *et al.*, 2005; Ahrens *et al.*, 2007; Curiel & Morrone, 2012), Diptera (Trivinho-Strixino *et al.*, 2012; Pramual & Wongpakam, 2014) and EPT taxa (Zhou *et al.*, 2007; Gattolliat & Monaghan 2010; Ruiter *et al.*, 2013; Avelino-Capistrano *et al.*, 2014). Barcodes have also been used for associating anisomorphic developmental stages of other invertebrates (e.g. trematodes; Jousson *et al.*, 1999; Blasco-Costa *et al.*, 2016) and vertebrates (Thomas *et al.*, 2005; Victor 2007). Indeed, DNA barcoding overcomes many of the problems encountered during rearing. It is faster and does not require prior knowledge of the species' life-history requirements (e.g. feeding preferences). Furthermore, nondestructive DNA extraction allows for morphological study of the specimens that have been barcoded. Lastly, while specialist expertise is useful for sorting or identification of larval or adult forms, it is not strictly required as long as all specimens are barcoded (Wang *et al.*, 2018).

However, DNA barcodes also have drawbacks. They can be misleading when closely related species have intraspecific distances that are lower than expected (0–2%) or when species have deep intraspecific splits (Meier *et al.*, 2006). In addition, obtaining barcodes with Sanger sequencing is so expensive and time-consuming that it is often not feasible for all available specimens in specimen- and species-rich samples. This means that the existing material has to be presorted into putative species based on morphological characters. This is time-consuming and can be very difficult if different instars are in the sample. We believe that these problems can be overcome through the use of 'next-generation sequencing (NGS) barcodes'. Here, NGS of tagged amplicons is used to obtain a short barcode (313 bp) at low cost (US$0.40 per specimen; Meier *et al.*, 2016). This low cost means that large numbers of immatures and adults can be sequenced without any morphological pre-sorting (Wang *et al.*, 2018). Indeed, the molecular pre-sorting becomes a technical exercise, which means that entomologists can focus

on working through specimen clusters that were assembled based on NGS barcodes and in most cases represent species. Lastly, the comparatively short NGS barcode fragment (313 bp) amplifies well even for fairly degraded material that has been collected and preserved in suboptimal conditions because it was collected, for example, for the purpose of biomonitoring. The low cost of NGS barcodes also means that it becomes feasible to carry out large-scale biodiversity surveys that deliberately target and/or include larvae and females. Such targeting would be desirable because it has the potential to yield much natural history information.

In this study, we document the potential of NGS barcoding for life-history stage matching by sequencing > 1100 specimens of Singaporean odonates. Singapore's odonates exhibit some of the typical problems of studying insect phenomes in the tropics. The adults are better known than the larvae and covered by species-level identification guides (Orr, 2005; Tang *et al.*, 2010) and species inventories (Ngiam & Cheong, 2016). This work has revealed that the fauna remains surprisingly rich although Singapore has lost most of its original forest cover. As of 2016, 131 species of odonates have been recorded in Singapore, with 122 species being considered extant. This means that the fauna is of a similar size to the entire European fauna (138 species: Kalkman *et al.*, 2010), albeit being compressed into an area of 700 km^2. Note that Singapore's fauna is also very respectable when compared to the fauna of other tropical islands; e.g. Borneo is 1000 times larger but only has twice the number of species (275 spp.; Orr & Hämäläinen, 2003), and Madagascar has approximately 175 species while being ca. 800 times larger (Dijkstra & Clausnitzer, 2004). Unfortunately, the Singaporean fauna is also typical in that much less is known about the larval morphology, ecology or distribution than about the adult stages. This is a common problem that is well exemplified by the 500 odonate species of India for which fewer than 100 have known larval forms (Prasad & Varshney, 1995; Andrew *et al.*, 2008).

New larval forms for odonate species living in Singapore are regularly described, but prior to this study the material was mostly obtained via rearing of targeted species or descriptions based on exuviae (Lok & Orr, 2009; Ngiam, 2011; Ngiam *et al.*, 2011; Orr & Ngiam, 2011; Ngiam & Leong, 2012; Ngiam & Dow, 2013). This approach has led to the larval identification and description of 71 of the 131 species of odonates in Singapore. However, it took more than 80 years to reach this stage (first larval description was in Lieftinck, 1932) and only the larvae of six species were described based on material from Singapore. All the remaining descriptions are based on larvae obtained from surrounding countries (Table S2).

Here, we document how NGS barcoding can not only accelerate the life-history stage matching process, but also radically alter the approach to larval identification across entomology. Instead of employing coarse parataxonomist sorting with subsequent DNA barcoding of few specimens, our approach uses NGS barcoding to directly associate all available specimens for a taxon. Larval specimens can thereby be barcoded and eventually sorted and associated en masse without any morphospecies pre-sorting. In our study, most specimens were obtained serendipitously over many years via miscellaneous surveys of freshwater bodies across Singapore. The larvae from Singapore were stored in different laboratories and collections, so that they were of widely varying preservation quality. Note that such haphazardly collected larval material is available for many taxa across the natural history museums of the world. Here we match the Singapore specimens to an adult sample that was obtained for the purpose of this study. After matching, the larval stages were imaged and made available online in the form of a digital reference collection (Ang *et al.*, 2013a), while the specimens were added to the collection of the Lee Kong Chian Natural History Museum. A useful by-product of the matching campaign was an odonate DNA barcode database for Singapore that has already been used for matching eDNA signatures obtained directly from water (Lim *et al.*, 2016).

Materials and methods

Sampling and identification

Most specimens were collected between 2011 and 2014. Some specimens were directly killed and preserved in 100% ethanol, while other specimens had unfortunately been anaesthetized first with carbonated water (the acidity probably affected DNA quality) before they were preserved in 70% methylated ethanol. Adults were obtained via targeted sweep-netting with the aim of maximizing species coverage. Adult specimens for some species were also serendipitously obtained from Malaise traps (Singapore). Some larvae were collected with kick-nets, sieves and coconut brush samplers (Loke *et al.*, 2010). Many of the samples were collected for biomonitoring purposes by the Tropical Marine Science Institute (TMSI). The freshwater bodies sampled include streams (natural and canalized), reservoirs, ponds and swamps across the country. Specimens were individualized and preserved in 70–100% ethanol. Only adult specimens were identified (Meier, 2017) using identification guides that include detailed drawings and photographs (Orr, 2005; Tang *et al.*, 2010). Most adults were identified by the first authors, but identifications for some specimens were verified by RWJN. Additional confirmation came from BLAST and Barcode of Life Data Systems (BOLD) hits. The specimens are deposited in the Lee Kong Chian Natural History Museum (National University of Singapore).

DNA extraction, amplification and sequencing

Regardless of extraction method, a leg or piece of tissue (0.5–1 cm) was removed from the specimen. Early on in the project, DNA was extracted using either a phenol chloroform extraction protocol (Kutty *et al.*, 2007) or Qiagen's (Singapore) DNeasy Blood & Tissue kit following the manufacturer's instructions. Later in the study, we used directPCR for larval specimens (Wong *et al.*, 2014). In directPCR, a small amount of tissue is used directly as DNA template in the PCR reaction.

We used the following *cox1* primers from Leray *et al.* (2013) for obtaining amplicons (mlCO1intF, 50-GGWAC

WGGWTGAACWGTWTAYCCYCC-30; and jgHCO2198, 50-TAIACYTCIGGRTGICCRAARAAYCA-30), which amplify a 313-bp region of the barcoding portion of the gene. The primers were labelled with a 9-bp tag at the 5′ ends. The tags were generated with a barcode generator script (Henry *et al.*, 2014) to differ by at least 3 bp. The tagged primers were also screened for secondary structure formation. Each specimen was assigned a unique combination of primer tags during amplification to allow recovered sequences to be assigned to specific specimen after sequencing. PCR amplification of both gDNA and dissected tissue was performed in a 25 μL reaction with 2.5 μL of 10X BioReady rTaq buffer (with 15 mM $MgCl_2$), 2.0 μL of 2.5 mM dNTP mixture, 1.0 μL of 1.0 mM BSA solution, 0.2 μL of BioReady rTaq and 2.0 μL of 10 μM forward and reverse primer, under the following conditions: initial denaturation at 95°C for 3 min, 40 cycles of denaturation at 94°C for 30 s, annealing at 50°C for 1 min and extension at 72°C for 1 min, with a final extension at 72°C for 5 min.

Successful amplification was verified for a few products for each 96-well PCR plate via gel electrophoresis in order to ensure there was no plate-wide amplification failure. Afterwards, 2.0 μL of PCR product was obtained from each sample and pooled. The pooled sample was purified using SureClean Plus (Bioline Inc., U.K.), following the manufacturer's protocol. The final pool was quantified via a Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, U.S.A.) and submitted to AITbiotech (Singapore) along with accompanying samples for library preparation and sequencing on an Illumina Miseq platform (Illumina Inc., San Diego, CA, U.S.A.) with a 2×300-bp paired end kit. The specimens were sequenced in this manner across four different Miseq runs alongside other NGS-based projects.

Read processing and barcode determination

Raw paired-end reads were first merged with PEAR (Zhang *et al.*, 2013) using the default parameters. A custom Python script was then used to demultiplex the data, count the number of reads per sample, identify identical reads, merge and count them, as well as compare the number of reads in the largest identity set and with the second largest identity set (Meier *et al.*, 2016). The barcoding of a specimen was only considered successful if the total amplicon coverage for the sample was $> 50\times$, the number of reads in the largest set of identical reads was $> 10\times$, and the read coverage in the largest identical set was more than five times the coverage of the second largest set (Meier *et al.*, 2016). If these criteria were met, the sequence of the largest read set was considered the barcode for the specimen. Lastly, we used BLAST against NCBI's nucleotide database for removing any nonodonate sequences.

mOTU estimation for adult–larva association

The *cox1* barcodes were aligned using MAFFT v.7 (http://mafft.cbrc.jp/alignment/software/) under default parameters to confirm open reading frames. Larval matching was achieved by estimating molecular operational taxonomic units (mOTUs) via objective clustering of uncorrected p-distances using SPECIESIDENTIFIER (TAXONDNA 1.6.2; Meier *et al.*, 2006). We also used the Automatic Barcoding Gap Discovery (ABGD) algorithm (Puillandre *et al.*, 2012) to test how robust the mOTU assignments were when a different species-delimitation technique was utilized. For objective clustering, a range of distance thresholds (2–4%) were applied that are commonly used in the barcoding literature (Ratnasingham & Hebert, 2013). For ABGD, we used a range of priors ($P = 0.002\,783$ to $P = 0.007\,743$). We then used the threshold/prior that maximized congruence with the adult identification, given that species limits in odonates are overall well understood based on adult morphology. We considered those mOTU clusters that contained sequences from both adults and larvae as having yielded a life-history stage association. Congruence between the different delimitation methods was assessed by pairwise match ratios (Ahrens *et al.*, 2016), which are defined as: $2 \times \text{Nmatch}/(\text{N1} + \text{N2})$, where Nmatch is the number of clusters identical across both mOTU delimitation methods/thresholds (N1 and N2).

Imaging and databasing

In order to facilitate the widespread use of the newly obtained associations, we imaged the larvae with distinct morphological features and created a digital reference collection with high-quality images (Ang *et al.*, 2013a). For all matched clusters, a representative specimen was selected for imaging based on morphological integrity. The selected specimens were then imaged with a Visionary Digital BK Lab System to obtain habitus photographs for the dorsal, ventral and lateral sides in addition to detailed images for labium and caudal lamellae (Zygoptera only). The images were stacked with HELICON FOCUS, edited in Adobe PHOTOSHOP CS6 and uploaded online to the Biodiversity of Singapore digital reference collection (https://singapore.biodiversity.online), with each species being represented by high-resolution specimen images and links to relevant publications. The DNA barcodes have been submitted to NCBI GenBank.

Review of adult–larva association literature

In order to review which techniques have been traditionally used for life-history stage matching in odonates, we conducted a literature search for odonate larval descriptions that were published between 2000 and 2016, using the following search terms on ISI Web of Science: (odonat$ OR zygopter$ OR anisopter$ OR anisozygopt$) AND (larv$ OR nymph& OR immature&). We determined the number of larval species described, country of larval origin, adult–larva association method, and sex/life-history stage of specimens described. We included publications describing larvae of species that were already described based on adult specimens as well as larval descriptions that were part of new species descriptions.

Table 1. Sequencing success rates and number of reads obtained.

Run number	No. of samples included	Total read counts	No. of samples that failed first filter[a]	No. of samples that failed second filter[b]	No. of samples that failed third filter[c]	No. of samples that failed BLAST filter	No. of samples that passed all filters	Sequencing success rate
1	826	13 742 608	12	5	32	4	773	93.58%
2	558	150 344	238	2	12	35	271	48.57%
3	88	132 724	17	14	5	0	52	59.09%
4	44	53 696	10	0	7	0	27	61.36%
Total	**1516**	**14 079 372**	**277**	**21**	**56**	**39**	**1123**	**74.08%**

[a] Amplicons with read counts < 50 are discarded.
[b] Amplicons with a dominant, unique read count < 10 are discarded.
[c] Amplicons where a read count ratio of the second-most dominant read to the most dominant read exceeds 0.2 are discarded.

Results

Sequencing and initial processing

A total of 1516 specimens were processed (338 adults and 1178 larvae). In the initial stage, we used formal DNA extraction via Qiagen DNeasy blood and tissue kits and phenol chloroform protocols to obtain DNA for 826 specimens. After developing the more time- and cost-effective directPCR methods, we used directPCR for the remaining 690 specimens. We initially generated DNA barcodes with Sanger sequencing, but the amplification success rates for *cox1* were fairly low and we thus switched to NGS barcoding via tagged amplicon sequencing. All specimens were sequenced using 14 079 372 paired reads (Table 1). This is equivalent to around 1.5 Miseq runs (300 bp paired-end) or roughly 10% of a HiSeq 2500 (250 bp paired-end) lane.

After applying quality control filters, sequences for 392 specimens were excluded, leaving 1123 barcodes for subsequent analyses (sequences are available on NCBI GenBank: MG884602–MG885724). Of the successfully barcoded specimens, 319 were adults and 804 were larvae. The overall sequencing recovery rate was unusually low (74.08%) which was due to specimens that had been anaesthetized with acidic, carbonated water before being stored in 70% methylated ethanol at room temperature for variable amounts of time. The success rates for those samples varied from 48.6% to 61.4% (Table 1), whereas the success rates for amplicons obtained from formally extracted gDNA was 93.6%.

Clustering and life-history stage matching

The number of mOTUs is overall stable across 2–4% p-distance thresholds and yields 95–99 mOTUs (Table 2). Of these mOTUs, 83 match known species based on the presence of pre-identified adult specimens. Note that specimens for one nominal species (*Neurothemis fluctuans*) are consistently found in two mOTUs, which diverge by a 6% p-distance, thereby suggesting the presence of a cryptic species (Bickford *et al.*, 2007). Our study covers 68% of Singapore's extant species and we obtained adult–larva matches for 46% (55 species). Twenty-seven of the remaining mOTUs contained adults only, whereas another 12–14 contained only larvae. BLAST searches of the larval *cox1* barcodes against NCBI GenBank and BOLD were used in an attempt to identify these latter larvae to species. This was successful for four species (*Tetracanthagyna plagiata*, *Diplacodes trivialis*, *Hydrobasileus croceus*, *Ceriagrion chaoi*), which brings the total number of matched species in this study to 59. Overall, the study adds life-history stage matches for 13 species to the existing list of described larvae of species occurring in Singapore (Table S2).

Overall, BLAST searches were able to provide species matches with reasonable certainty (> 97% match similarity) for 50 mOTUs, i.e. the publicly available barcode databases remain fairly incomplete even for such a charismatic and well-studied taxon as odonates. However, GenBank and BOLD yield identification for a large proportion of all specimens because common species tend to be overrepresented in the databases (1123 successfully identified specimens; GenBank, 827; BOLD, 848 at

Table 2. Molecular operational taxonomic units delimited via objective clustering and Automatic Barcoding Gap Discovery (ABGD) approaches.

	Objective clustering			ABGD		
	2% p-distance	3% p-distance	4% p-distance	$P = 0.002783$	$P = 0.004642$	$P = 0.007743$
Total no. of clusters	99	96	95	102	97	95
No. of adult clusters split	3	1	1	6	1	1
No. of adult clusters lumped	0	0	0	0	0	0
No. of clusters with adult–larva matches	55	56	56	55	55	56
No. of adult-only clusters	30	27	27	33	28	27
No. of larva-only clusters	14	13	12	14	14	12

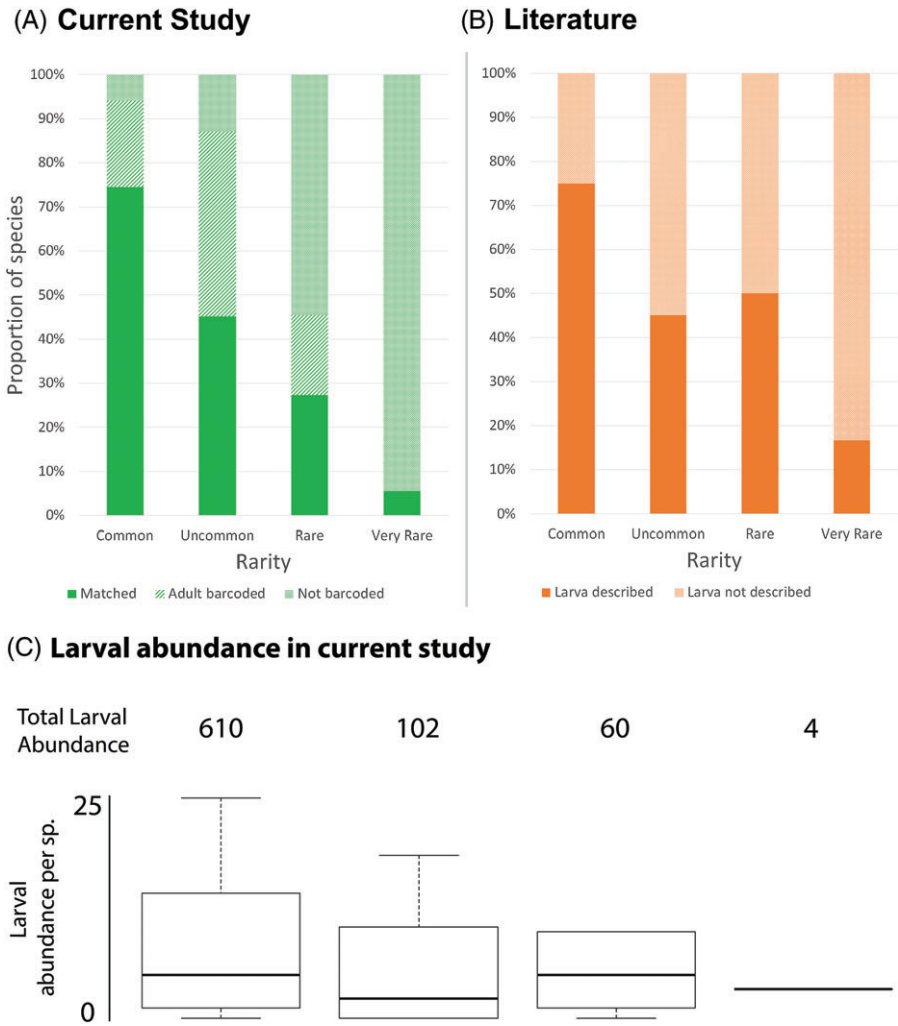

Fig. 1. Rarity and probability of adult–larva matching for Singapore's odonate species. (A) Current study; (B) literature (numbers = number of species); (C) larval abundance across rarity classes. [Colour figure can be viewed at wileyonlinelibrary.com].

97.5–100.0% similarity). The remaining 46 mOTUs without GenBank/BOLD matches contribute only 296 specimens in our study.

The checklist of Singapore's odonates by Ngiam & Cheong (2016) includes an assessment of rarity for each species in Singapore. Most of the species that are frequently encountered have been matched (common, 74.5%; uncommon, 45.2%) and barcoded (common, 94.5%; uncommon, 87.1%). 'Rare' and 'very rare' species are underrepresented in this study (Fig. 1a), but we nevertheless matched seven of the 40 (17.5%). The likelihood of a species being matched or barcoded correlates well with the abundance in Singapore, but the number of larvae collected across rarity classes and species is not significantly different (Kruskal–Wallis, $P = 0.62$) (Fig. 1a). In contrast, the proportions of species with larvae described prior to this study (mostly based on nonmolecular methods) do not match the abundance of the species in Singapore (Fig. 1c), which is presumably due to the fact that most species matches were obtained outside of Singapore (Fig. 1b).

mOTU stability

In order to determine which distance thresholds maximize congruence with adult morphology, we tested different threshold and priors (objective clustering of uncorrected p-distances and ABGD). The clustering thresholds and prior values used were 2–4% and $P = 0.002\,783$–$0.007\,743$ respectively, which correspond to what is typically employed in the literature (Ratnasingham & Hebert, 2013). This resulted in 95–102 mOTUs (Table 2), depending on the choice of clustering method and parameter. As the adults can be reliably identified to species, we used the clustering method and parameters that maximized congruence with adult morphology (3% and 4% for objective clustering; $P = 0.007\,743$ for ABGD).

Table 3. Pairwise match ratios of clusters delimited by objective clustering and Automatic Barcoding Gap Discovery (ABGD).

	Objective clustering			ABGD		
	2%	3%	4%	$P = 0.002\,783$	$P = 0.004\,642$	$P = 0.007\,743$
Objective clustering						
2%	1.00					
3%	0.96	1.00				
4%	0.95	0.98	1.00			
ABGD						
$P = 0.002\,783$	0.97	0.95	0.92	1.00		
$P = 0.004\,642$	0.98	0.98	0.97	0.95	1.00	
$P = 0.007\,743$	0.95	0.99	1.00	0.92	0.97	1.00

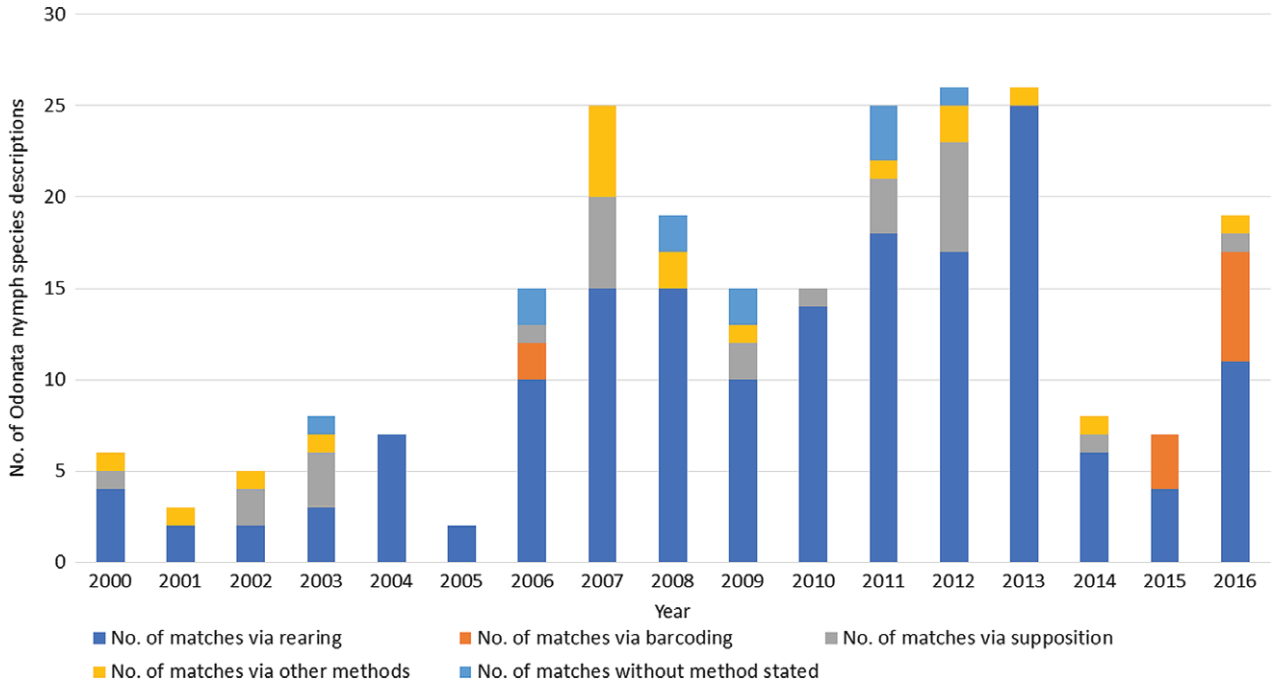

Fig. 2. Adult–larva association methods in descriptive odonate literature (2000–2016). [Colour figure can be viewed at wileyonlinelibrary.com].

The pairwise match ratios (> 0.92) across the different clustering algorithms and parameters (Table 3) indicate high stability, i.e. most species are genetically distant enough to be identifiable across the delimitation methods and thresholds. Only one species represented by adults is consistently split into two clusters.

Literature survey

Our literature survey revealed larval descriptions for 229 species that were published between 2000 and 2016 (Table S3). Of these, only 11 were based on DNA barcodes (Fig. 2) while the vast majority were obtained via rearing (164 species) or identifying tenerals next to their exuviae. Overall, there seems to be a slight increase of adult–larva associations via DNA barcodes over time. Note that some adult–larva associations are still made via supposition (i.e. identifying larvae based on the putative presence of adults, which is sometimes determined based on checklists).

Discussion

Expediting adult–larva associations with NGS barcoding

Singapore is a small island with a surface area of 719 km^2. The country nevertheless used to be the home of 131 species of Odonata (Ngiam & Cheong, 2016) of which 122 are still considered extant. This is a very large number of species given that all of Europe only has 138 species (Kalkman *et al.*, 2010), i.e. the odonate fauna of Singapore poses a significant enough challenge to be typical for other life-history matching problems. Here, we barcoded 1123 adult and larval specimens that belonged to 95–96 mOTUs, of which 83 were identifiable to species due to the presence of a pre-identified adult in the cluster.

To our knowledge, this is currently the largest adult–larva association study in the literature. A very large proportion of the specimens in our study can be associated (1012 of 1123, 90.1%) and yield morphological and habitat information for 59 species. Most of these associations were obtained during our study (55) and only four were established via BLAST matches to NCBI GenBank. Our study adds a total of 13 species to the list of species with associated larvae. This was accomplished via occasional collecting spread across a 2-year period. In particular, we targeted habitats that were suspected to have high diversity. In comparison to collecting, the molecular work was not time-consuming. Based on the techniques applied to our latest batch of specimens, we could now obtain amplicons for all 1123 specimens in *c*. 8 days of molecular work (Wang *et al.*, 2018). In contrast, matching via rearing has resulted in only six new larval descriptions between 2009 and 2013 – i.e. NGS barcoding has the potential to quickly yield life-history associations for a large number of species, especially if a large amount of material is already in collections and an adult barcode database can be obtained quickly.

Immature associations via DNA sequences have been available since the late 1990s (DeSalle & Birstein, 1996; Miller *et al.*, 1997), but most studies are small-scale and only associate the life-history stages for a few species at time. Such small-scale application of DNA barcoding is particularly common in odonates where the first study was published in 2006 (Etscher *et al.*, 2006; Fleck *et al.*, 2006). Indeed, even recent publications involving odonate larval matching via DNA barcoding only involve one to two species and are based on < 50 specimens. Other insect orders have seen larger-scale studies. The largest we found in the literature included 250 larvae and adult beetles (Ahrens *et al.*, 2007), but we believe that it is now time to become more ambitious. Large-scale immature sampling in arthropod surveys is now because NGS barcodes can be used to match life-history stages.

The small number of DNA association studies and small number of specimens in most adult-larva matching studies are most likely due to the substantial per-specimen cost of Sanger sequencing. Sanger sequencing is expensive (anywhere from US$7–34; from http://ccdb.ca/pricing), but the cost of sequencing can be dramatically cut by using NGS approaches (Meier *et al.*, 2016; Hebert *et al.*, 2017). Here, we sequenced 1123 specimens using 14 079 372 reads of 300 bp PE Illumina Miseq sequencing (Table 1). The average coverage of each specimen was 12 537x; this is very excessive. It would have been safe to reduce the average number of reads per specimen to one-tenth (1254 reads per specimen). This implies that $> 11\,000$ specimens can be barcoded in a single Illumina MiSeq lane. The cost of the lane would be *c*. US$2000 (see http://medicine.yale.edu/keck/ycga/services/illuminaprices.aspx), thereby reducing the sequencing cost of each specimen to < US$0.20. Indeed, the main cost and challenge are not sequencing, but obtaining specimens and generating a sufficiently large number of amplicons with PCR to fill a lane or flowcell. However, the latter problem can be overcome because the same sequencing run can be shared by multiple projects and obtaining amplicons with direct PCR is cost-effective (US$0.16). The process is also sufficiently simple that even inexperienced personnel can process 200 specimens per day (Wang *et al.*, 2018). This means that the wet-lab cost per specimen is *c*. US$0.40 and that a study like the one whose results we are here reporting costs only *c*. US$600 (1516 specimens × 0.40). We estimate that all specimens in our study could be processed within eight working days.

Tackling the problem of rarity and elusive larvae

As discussed earlier, insect larvae are absent or underrepresented in many species descriptions. The most likely reasons are the difficulty of finding larvae and associating them with adults. As documented here, the latter problem can be addressed but the larvae of some species will remain elusive because they have very specific habitat requirements (e.g. phytotelma: Ngiam & Leong, 2012). Indeed, the likelihood of being matched and barcoded in this study is overall positively correlated with the species' abundance/rarity (Fig. 1a) although it appears from our study that 'rare' species are more likely to have been associated in the literature than 'uncommon' ones (Fig. 1b). This anomaly is probably due to the fact that most of the published larval descriptions were prepared based on material collected outside of Singapore, while the ranking into 'rare' or 'uncommon' is based on the abundance of the species in the country. Note that the larval abundances in this study also do not correspond closely to the rarity classes. This is probably due to nonstandardized method of larva collection. Entomologists tend to oversample habitats that are likely to yield larvae for rare or unsampled species.

We suggest that with the advent of NGS barcoding future insect surveys may want to deliberately target larvae. Such larval screens should be a priority in habitats that have a large number of species with unknown larval stages. Many odonate larval association studies use targeted sampling of larvae in promising habitats in order to find the larval stages for rare species. However, it is important to remember that larval and adult habitats can differ considerably for species where the territorial behaviour of adults force individuals to forage away from their optimal breeding zones (Wolf & Waltz, 1984). This is why we believe that such targeted collecting should be complemented with more large-scale sampling of potential larval habitats that are otherwise neglected. Processing samples from such localities will be fast and affordable once NGS barcoding is used for barcoding the specimens. Currently, most adult–larva associations for odonates are still obtained via rearing (Fig. 2). This is complemented by studying the morphology of pharates (Carvalho, 2000; Pérez-Gutiérrez & Montes-Fontalvo, 2011) and obtaining exuviae collected after the teneral adult has emerged (Xu, 2012; Del Palacio & Muzon, 2014; Domenico *et al.*, 2016). These are chance events that should, of course, be utilized whenever available, but they are probably too infrequent to constitute a significant method for obtaining adult–immature associations. Some authors have also identified larvae via supposition, i.e. based on the presence of adults in a particular habitat (Tennessen, 2010; Müller *et al.*,

2012; Novelo-Gutierrez *et al.*, 2014). However, this practice can lead to incorrect results.

Larval matching via NGS barcodes is particularly promising for habitats with a large number of species without larval associations, but it should also be considered for well-studied areas especially if there are existing and underutilized collections of larvae and the specimens are relevant for freshwater biomonitoring. Identifying larval specimens to species can yield significant benefits when using the taxa as bioindicator (Lenat & Resh, 2001). Fortunately, NGS barcoding is also suitable for many of the specimens in museum collections given that it only involves a 313-bp piece of *cox1* that is more likely to amplify for specimens that were not preserved for molecular purposes and may have been stored for a long time under problematic conditions. Lastly, the standardization of the barcoding gene for animals allows for the creation of large databases and promotes information sharing. Locally rare species might be common abroad, and comparisons of local barcodes against global databases can yield new matches, as was the case for four species in our study.

Barcode reliability and potential sources of error

In our study, we find that standard clustering thresholds applied to barcodes yield mOTUs that are overall consistent with species identified based on morphology (Table 2). In addition, the mOTUs are overall stable across delimitation algorithms and thresholds (Table 3). This goes quite a long way to alleviating the theoretically justified concerns that barcodes can contain misleading signal. After all, genetic distances between barcodes do not always correspond to species boundaries (Meier *et al.*, 2006) because *cox1* is not a speciation gene (Kwong *et al.*, 2012a). Based on theoretical consideration, one should therefore always scrutinize results obtained with barcodes for two potential sources of error: low interspecific variability leading to lumping, and high intraspecific variability leading to splitting. The former is not observed here, but such cases have been reported in the literature. In other cases, high genetic variability within morphological clusters can be indicative of the presence of cryptic species (Damm *et al.*, 2010). In our study, there may be two cases: *Amphicnemis gracilis* had unusually high *cox1* pairwise distances of > 4% between individuals collected at the same site. However, the largest genetic distances were observed within *Neurothemis fluctuans* (> 6%). Unfortunately, our current sample size is not large enough for both cases to clarify the species boundaries.

However, overall there is good congruence between NGS barcodes and morphotypes. Congruence between morphology and DNA sequences is often discussed at the species level. However, it is arguably equally important to consider how many specimens can be confidently placed into mOTUs that are stable across clustering methods. In our study, we find that almost all specimens (1099 of 1123; 97.9%) can be placed into such stable mOTUs (mOTUs that are insensitive to algorithms and thresholds). This means that only the assignment of a very small number of specimens is uncertain because they pertain to species with barcodes that are not decisive.

The value of both morphology and molecules

Odonate larvae are typically described based solely on the final instar; presumably this is partly due to the fact that morphological characters are most clearly visible at this stage and exuviae of final instars are often used for morphological descriptions. As DNA barcoding can be applied to all instars, it also allows for the association of early-instar larvae. This might pose problems when using the morphological literature that focuses on the morphology of late-instar larva, but it is arguably time to also study the morphology of early instars. Because NGS barcodes can identify all instars, barcodes are also more useful for identifying larval habitats of species because early instars are more common than late instars. If a late instar is needed for descriptive purposes, follow-up work can concentrate on finding a late-instar larva in a habitat where an early instar had been identified based on barcoding.

The pipeline described here also paves the way for a more extensive use of larvae in phylogenetics. Larval morphology is a potentially rich yet frequently underutilized source of characters (Beutel, 1993; Meier & Lim, 2009; Beutel *et al.*, 2010; Aspöck *et al.*, 2012). Phylogenomics is a powerful tool, but it is not the 'magic bullet' that can resolve all problems with confidence (Bordenstein *et al.*, 2008; Kutty *et al.*, 2018). In the era of genomics, morphology still has an important role to play in phylogenetics (Giribet, 2015). Congruence between morphology and genes increases the confidence in phylogenetic results, while incongruence can lead to novel discoveries of body plan evolution. There are several cases in the recent literature where larval morphology provided important sets of characters that were used to resolve problematic relationships (Heikkilä *et al.*, 2014; Badano *et al.*, 2017; Michat *et al.*, 2017). However, more larvae that can be identified to species are needed in order to incorporate larval characters in phylogenetic analyses. A systematic effort to associate larval and adult forms via NGS barcoding can overcome this bottleneck.

Description, illustration, or both?

One of the aims of the project was to create a lasting and publicly available resource for researchers or naturalists interested in the odonates of Singapore. Such a resource should cover adult and larval stages. For this reason, we prepared a digital reference collection (Ang *et al.*, 2013a) by documenting the species with high-resolution images and making the images available online as part of the 'The Biodiversity of Singapore' initiative (https://singapore.biodiversity.online). For diagnostic purposes, not only the habitus, but also images of the labium and caudal lamellae are included (Fig. 3). All images can be enlarged to reveal details such as the number of setae on the specimens. This resource is for odonate specialists, but is also very useful for biologists involved in freshwater assessment

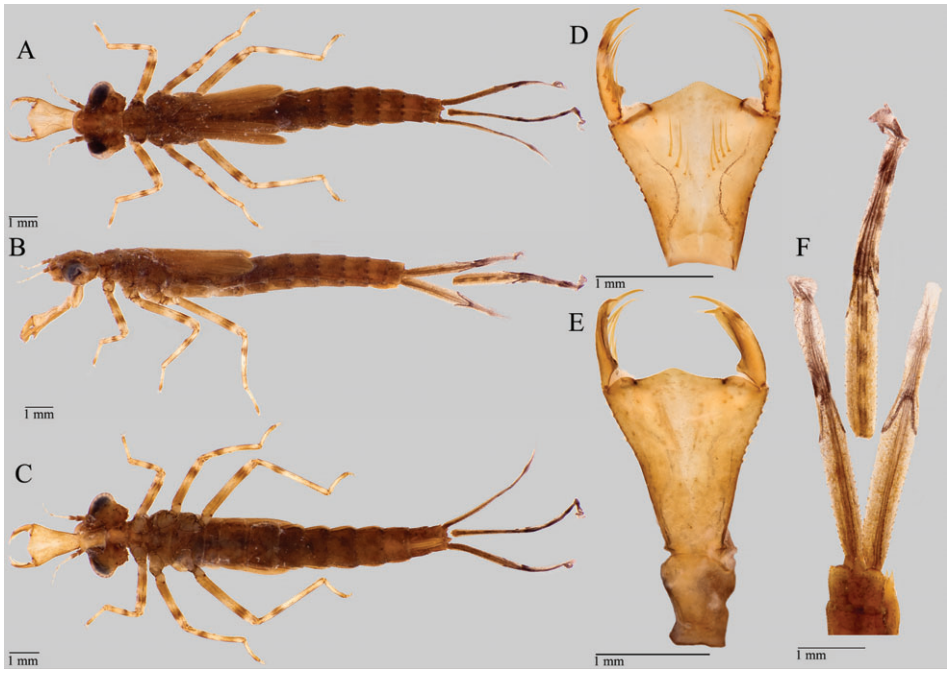

Fig. 3. Species entry in digital reference collection for larva of *Mortonagrion arthuri*: dorsal (A), lateral (B) and ventral (C) views; enlarged images of dorsal (D) and ventral (E) views of the labium and caudal lamellae (F). Scale = 1 mm. All images can be enlarged in browser. [Colour figure can be viewed at wileyonlinelibrary.com].

via macroinvertebrates. Lastly, it communicates to taxonomic experts that material is available for the formal description of new larvae or larval instars (specimens are deposited in the Lee Kong Chian Natural History Museum). The digital reference collection now has entries for 110 species of Odonata, with 59 having associated larvae. It is expected to grow as new specimens are collected, sequenced and imaged. Due the low technical complexity of NGS barcoding, regular updates can be obtained via undergraduate student projects every few years.

The gold standard in taxonomy is providing high-quality descriptions that are supported by high-quality images that illustrate all characters that are known to be important for species identification (Ang *et al.*, 2013b). At the other end of the scale is the kind of short descriptions that are common in the old literature (see Lamb, 1924 for an odonate example) and that are not supported by illustrations and/or supported by poorly executed or even incorrect drawings. In this study we have opted for the middle path, i.e. high-resolution images of the larvae (Fig. 3) which illustrate the features that are currently used for describing and identifying odonate larvae. This includes photographs of the habitus, but also of the modified labium, anal pyramid, lateral and dorsal abdominal spines (Anisoptera) and caudal lamellae (Zygoptera). We considered providing more comprehensive descriptions, but this would be premature given that much fewer than half of all known odonate species in Southeast Asia have known larvae, i.e. this would make it very difficult to identify those characters that will eventually be revealed to be species-specific. With the current level of knowledge, we would thus argue that a digital reference collection consisting of high-quality images of potentially diagnostic features is preferable.

Barcode databases

In this study, four larval clusters could initially not be identified to species because they lacked an adult match. This illustrates why it is important to have comprehensive and publicly available barcode databases. This study generated a large number of such (mini) barcodes. Approximately half of the 95 species barcoded here (46 spp.) lacked barcodes in NCBI GenBank or BOLD. Barcode databases are extremely valuable once they are sufficiently comprehensive, but they tend to be very incomplete for most invertebrate groups (Kwong *et al.*, 2012b) so that continued adult and immature collecting and barcoding remain necessary. Comprehensive and well-curated barcode databases are not only needed for specimen-based barcoding, but they are also important for metabarcoding studies that yield sequence information from gut content, faecal matter, soil or water (Taberlet *et al.*, 2012; Srivathsan *et al.*, 2015; Lim *et al.*, 2016). For instance, biomonitoring of freshwater macroinvertebrates occasionally utilizes a ‘soup’ approach to species identification. Homogenized samples are studied with metabarcoding (Yu *et al.*, 2012), but the metabarcoding data are much more meaningful when many sequence signatures can be matched to species. Routine or exploratory biomonitoring via metabarcoding of eDNA can also yield new larval matches if

the metabarcodes are identified as belonging to odonates that are not present in available databases.

Concluding remarks

We hope that our matching study will inspire more work on adult–larva matching in entomology. NGS barcoding provides the means for identifying larvae to species, finding breeding sites and reconstructing habitat preferences; i.e. it enables species-specific conservation projects that focus on larval habitats. In addition, NGS barcoding can significantly aid in the study of insect phenomes. Indeed, we know next to nothing about most of the species described (Greene, 2005). This is problematic because observing and detailing natural processes can help to generate novel and idiographic hypotheses, as well as generate interest and support in biodiversity research (Willson & Armesto, 2006; Cotterill & Foissner, 2010). Odonates in Singapore have fortunately received more attention than other invertebrates, with updated species checklists (Ngiam & Cheong, 2016), updates of new records and existing populations (Ngoi *et al.*, 2011; Ngiam *et al.*, 2011), larval descriptions (Ngiam & Leong, 2012; Ngiam & Dow, 2013), as well as excellent notes on species life histories (Ngiam, 2009; Orr *et al.*, 2010). However, only very few insect taxa have received so much attention.

Here we describe a pipeline for the large-scale identification of larvae of an ecologically important and charismatic insect taxon via association through NGS barcodes. While DNA sequences have been used before to associate different life-history stages, it is only through the use of NGS technologies that the cost can be reduced sufficiently to justify sequencing all specimens in large samples. This approach ('reverse workflow' in Wang *et al.*, 2018) renders morphological pre-sorting of specimen samples unnecessary. Instead, biologists can focus on confirming the morphological integrity of specimen clusters that were assembled according to NGS barcodes and then studying the morphology of the associated life-history stages. The method described here is particularly relevant for the study of insects with morphologically very different life stages from poorly explored environments and habitats. It is particularly suitable for taxa with high specimen abundance and species diversity. Finally, we hope that the method will eventually allow for the inclusion of more immature data in phylogenetics, habitat surveys and environmental impact assessments.

Supporting Information

Additional Supporting Information may be found in the online version of this article under the DOI reference: 10.1111/syen.12296

Figure S1. Larval abundance distributions across rarity classes (C, common; UC, uncommon; R, rare; VR, very rare; left, full dataset; right, single 'common' outlier excluded) (Tang *et al.*, 2010).

Table S2. Checklist of Odonata recorded in Singapore, with distribution, rarity and conservation status, the abundance of specimens collected and barcoded in this paper, as well as whether descriptions exist for the larvae preceding this paper.

Table S3. List of odonate larvae description references (2000–2016) used in this study to assess the frequency and method of adult–larva association used.

Acknowledgements

We would like to thank Dr Esther Clews and her team from the TMSI for providing larval specimens and expertise, Foo Maosheng and Siti Maimon Bte Hussin for providing assistance with the larval imaging, Tang Hung Bun for providing photographs for the graphical abstract, as well as Dr Yuchen Ang and Jonathan Ho Kit Ian for their work on the Biodiversity of Singapore digital reference collection. We would like to acknowledge grant support from NUS SEABIG (R-154-000-648-646 and R-154-000-648-733).

The authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

References

Ahrens, D., Monaghan, M.T. & Vogler, A.P. (2007) DNA-based taxonomy for associating adults and larvae in multi-species assemblages of chafers (Coleoptera: Scarabaeidae). *Molecular Phylogenetics and Evolution*, **44**, 436–449.

Ahrens, D., Fujisawa, T., Krammer, H.J., Eberle, J., Fabrizi, S. & Vogler, A.P. (2016) Rarity and incomplete sampling in DNA-based species delimitation. *Systematic Biology*, **65**, 478–494.

Andrew, R.J., Subramanian, K.A. & Tiple, A.D. (2008) *A Handbook on Common Odonates of Central India*. Hislop College, Nagpur.

Ang, Y., Puniamoorthy, J., Pont, A.C. *et al.* (2013a) A plea for digital reference collections and other science-based digitization initiatives in taxonomy: Sepsidnet as exemplar. *Systematic Entomology*, **38**, 637–644.

Ang, Y., Wong, L.J. & Meier, R. (2013b) Using seemingly unnecessary illustrations to improve the diagnostic usefulness of descriptions in taxonomy–a case study on *Perochaeta orientalis* (Diptera, Sepsidae). *ZooKeys*, **355**, 9–27.

Aspöck, U., Haring, E. & Aspöck, H. (2012) The phylogeny of the Neuropterida: long lasting and current controversies and challenges (Insecta: Endopterygota). *Arthropod Systematics & Phylogeny*, **70**, 119–129.

Avelino-Capistrano, F., Nessimian, J.L., Santos-Mallet, J.R. & Takiya, D.M. (2014) DNA-based identification and descriptions of immatures of Kempnyia Klapálek (Insecta: Plecoptera) from Macaé River Basin, Rio de Janeiro State, Brazil. *Freshwater Science*, **33**, 325–337.

Badano, D., Aspöck, U., Aspöck, H. & Cerretti, P. (2017) Phylogeny of Myrmeleontiformia based on larval morphology (Neuropterida: Neuroptera). *Systematic Entomology*, **42**, 94–117.

Basset, Y., Cizek, L., Cuénoud, P. *et al.* (2012) Arthropod diversity in a tropical forest. *Science*, **338**, 1481–1484.

Beutel, R.G. (1993) Phylogenetic analysis of Adephaga (Coleoptera) based on characters of the larval head. *Systematic Entomology*, **18**, 127–147.

Beutel, R.G., Friedrich, F. & Aspöck, U. (2010) The larval head of Nevrorthidae and the phylogeny of Neuroptera (Insecta). *Zoological Journal of the Linnean Society*, **158**, 533–562.

Bickford, D., Lohman, D.J., Sodhi, N.S. *et al.* (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, **22**, 148–155.

Blasco-Costa, I., Poulin, R. & Presswell, B. (2016) Species of *Apatemon* Szidat, 1928 and *Australapatemon* Sudarikov, 1959 (Trematoda: Strigeidae) from New Zealand: linking and characterising life cycle stages with morphology and molecules. *Parasitology Research*, **115**, 271–289.

Bordenstein, S.R., Paraskevopoulos, C., Dunning Hotopp, J.C. *et al.* (2008) Parasitism and mutualism in *Wolbachia*: what the phylogenomic trees can and cannot say. *Molecular Biology and Evolution*, **26**, 231–241.

Carvalho, A.L. (2000) Descriptions of the last instar larva and some structures in the pharate male adult of *Praeviogomphus proprius* Belle, 1995, with notes on the occurrence and taxonomic status of the species (Anisoptera: Gomphidae, Octogomphinae). *Odonatologica*, **29**, 239–246.

Cotterill, F.P. & Foissner, W. (2010) A pervasive denigration of natural history misconstrues how biodiversity inventories and taxonomy underpin scientific knowledge. *Biodiversity and Conservation*, **19**, 291.

Curiel, J. & Morrone, J.J. (2012) Association of larvae and adults of Mexican species of *Macrelmis* (Coleoptera: Elmidae): a preliminary analysis using DNA sequences. *Zootaxa*, **3361**, 56–62.

Damm, S., Schierwater, B. & Hadrys, H. (2010) An integrative approach to species discovery in odonates: from character-based DNA barcoding to ecology. *Molecular Ecology*, **19**, 3881–3893.

Del Palacio, A. & Muzon, J. (2014) Description of the final instar larva of *Limnetron antarcticum* Förster and notes on its female (Anisoptera: Aeshnidae). *Zootaxa*, **3884**, 89–94.

DeSalle, R. & Birstein, V.J. (1996) PCR identification of black caviar. *Nature*, **381**, 197.

Dijkstra, K.D.B. & Clausnitzer, V. (2004) Critical species of Odonata in Madagascar. *International Journal of Odonatology*, **7**, 219–228.

Domenico, M.D., Dijkstra, K.D. & Carchini, G. (2016) Redescription of the larva of *Gynacantha cylindrata* Karsch (Insecta: Odonata: Aeshnidae). *Zootaxa*, **4078**, 78–83.

Eberhard, W.G. (1985) *Sexual Selection and Animal Genitalia*, Vol. **244**. Harvard University Press, Cambridge, Massachusetts.

Etscher, V., Miller, M.A. & Burmeister, E.G. (2006) The larva of *Polythore spaeteri* Burmeister & Börzsöny, with comparison to other polythorid larvae and molecular species assignment (Zygoptera: Polythoridae). *Odonatologica*, **35**, 127–142.

Fleck, G., Brenk, M. & Misof, B. (2006) DNA Taxonomy and the identification of immature insect stages: the true larva of *Tauriphila argo* (Hagen 1869) (Odonata: Anisoptera: Libellulidae). *Annales de la Société Entomologique de France*, **42**, 91–98.

Gaston, K.J. (2000) Biodiversity: higher taxon richness. *Progress in Physical Geography*, **24**, 117–127.

Gattolliat, J.L. & Monaghan, M.T. (2010) DNA-based association of adults and larvae in Baetidae (Ephemeroptera) with the description of a new genus *Adnoptilum* in Madagascar. *Journal of the North American Benthological Society*, **29**, 1042–1057.

Giribet, G. (2015) Morphology should not be forgotten in the era of genomics–a phylogenetic perspective. *Zoologischer Anzeiger-A Journal of Comparative Zoology*, **256**, 96–103.

van Gossum, H., Sánchez, R. & Rivera, A.C. (2003) Observations on rearing damselflies under laboratory conditions. *Animal Biology*, **53**, 37–45.

Greene, H.W. (2005) Organisms in nature as a central focus for biology. *Trends in Ecology & Evolution*, **20**, 23–27.

Hebert, P.D., Cywinska, A. & Ball, S.L. (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*, **270**, 313–321.

Hebert, P. D., Braukmann, T. W., Prosser, S. W., Ratnasingham, S., Ivanova, N. V., Janzen, D. H., ... & Zakharov, E. V. (2017) A sequel to sanger: Amplicon sequencing that scales. *bioRxiv* 191619. https://doi.org/10.1101/191619

Heikkilä, M., Mutanen, M., Kekkonen, M. & Kaila, L. (2014) Morphology reinforces proposed molecular phylogenetic affinities: a revised classification for Gelechioidea (Lepidoptera). *Cladistics*, **30**, 563–589.

Henry, I.M., Nagalakshmi, U., Lieberman, M.C. *et al.* (2014) Efficient genome-wide detection and cataloging of EMS-induced mutations using exome capture and next-generation sequencing. *The Plant Cell*, **26**, 1382–1397.

Jousson, O., Bartoli, P. & Pawlowski, J. (1999) Molecular identification of developmental stages in Opecoelidae (Digenea). *International Journal for Parasitology*, **29**, 1853–1858.

Kalkman, V.J., Boudot, J., Bernard, R. *et al.* (2010) *European Red List of Dragonflies*. Publications Office of the European Union, Luxembourg.

Kraus, R.T. & Secor, D.H. (2005) Application of the nursery-role hypothesis to an estuarine fish. *Marine Ecology Progress Series*, **291**, 301–305.

Kutty, S.N., Bernasconi, M.V., Šifner, F. & Meier, R. (2007) Sensitivity analysis, molecular systematics and natural history evolution of Scathophagidae (Diptera: Cyclorrhapha: Calyptratae). *Cladistics*, **23**, 64–83.

Kutty, S.N., Wong, W.H., Meusemann, K., Meier, R. & Cranston, P.S. (2018) A phylogenomic analysis of Culicomorpha (Diptera) resolves the relationships among the eight constituent families. *Systematic Entomology*. https://doi.org/10.1111/syen.12285.

Kwong, S., Srivathsan, A., Vaidya, G. & Meier, R. (2012a) Is the COI barcoding gene involved in speciation through intergenomic conflict? *Molecular Phylogenetics and Evolution*, **62**, 1009–1012.

Kwong, S., Srivathsan, A. & Meier, R. (2012b) An update on DNA barcoding: low species coverage and numerous unidentified sequences. *Cladistics*, **28**, 639–644.

Lamb, L. (1924) A tabular account of the differences between the earlier instars of *Pantala flavescens* (Odonata: Libellulidae). *Transactions of the American Entomological Society*, **50**, 289–312.

Lenat, D.R. & Resh, V.H. (2001) Taxonomy and stream ecology – the benefits of genus-and species-level identifications. *Journal of the North American Benthological Society*, **20**, 287–298.

Leray, M., Agudelo, N., Mills, S.C. & Meyer, C.P. (2013) Effectiveness of annealing blocking primers versus restriction enzymes for characterization of generalist diets: unexpected prey revealed in the gut contents of two coral reef fish species. *PLoS One*, **8**, e58076.

Lieftinck, M.A. (1932) Notes on the larvae of two interesting Gomphidae (Odon.) from the Malay Peninsula. *The Bulletin of the Raffles Museum*, **7**, 102–115.

Lim, G.S., Balke, M. & Meier, R. (2011) Determining species boundaries in a world full of rarity: singletons, species delimitation methods. *Systematic Biology*, **61**, 165–169.

Lim, N.K., Tay, Y.C., Srivathsan, A. *et al.* (2016) Next-generation freshwater bioassessment: eDNA metabarcoding with a conserved metazoan primer reveals species-rich and reservoir-specific communities. *Royal Society Open Science*, **3**, 160635.

Lok, A.F.S.L. & Orr, A.G. (2009) The biology of *Euphaea impar* Selys (Odonata: Euphaeidae) in Singapore. *Nature in Singapore*, **2**, 135–140.

Loke, L.H., Clews, E., Low, E.W., Belle, C.C., Todd, P.A., Eikaas, H.S. & Ng, P.K. (2010) Methods for sampling benthic macroinvertebrates in tropical lentic systems. *Aquatic Biology*, **10**, 119–130.

Meier, R. (1995) Cladistic analysis of the Sepsidae (Cyclorrhapha: Diptera) based on a comparative scanning electron microscopic study of larvae. *Systematic Entomology*, **20**, 99–128.

Meier, R. (2017) Citation of taxonomic publications: the why, when, what and what not. *Systematic Entomology*, **42**, 301–304.

Meier, R. & Lim, G.S. (2009) Conflict, convergent evolution, and the relative importance of immature and adult characters in endopterygote phylogenetics. *Annual Review of Entomology*, **54**, 85–104.

Meier, R., Shiyang, K., Vaidya, G. & Ng, P.K. (2006) DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Systematic Biology*, **55**, 715–728.

Meier, R., Wong, W., Srivathsan, A. & Foo, M. (2016) $1 DNA barcodes for reconstructing complex phenomes and finding rare species in specimen-rich samples. *Cladistics*, **32**, 100–110.

Michat, M.C., Alarie, Y. & Miller, K.B. (2017) Higher-level phylogeny of diving beetles (Coleoptera: Dytiscidae) based on larval characters. *Systematic Entomology*, **42**, 734–767.

Miller, L.J., Graham, G.C., Allsopp, P.G. & Yeates, D.K. (1997) Use of DNA sequencing for species identification and phylogeography: canegrub examples (Scarabaeidae: Melolonthini). *Soil Invertebrates* (ed. by P.G. Allsopp, D.J. Rogers and L.N. Robertson), pp. 31–34. Bureau of Sugar Experiment Stations, Brisbane.

Miller, K.B., Alarie, Y., Wolfe, G.W. & Whiting, M.F. (2005) Association of insect life stages using DNA sequences: the larvae of *Philodytes umbrinus* Motschulsky (Coleoptera: Dytiscidae). *Systematic Entomology*, **30**, 499–509.

Müller, O., Taron, U., Jansen, A. & Schneider, T. (2012) Description of the larva of *Boyeria cretensis* Peters and comparison with *B. irene* (Fonscolombe)(Anisoptera: Aeshnidae). *Odonatologica*, **41**, 47.

Ngiam, R.W.J. (2009) The biology and distribution of *Pseudagrion rubriceps rubriceps* Selys, 1876 (Odonata: Zygoptera: Coenagrionidae) in Singapore. *Nature in Singapore*, **2**, 209–214.

Ngiam, R.W.J. (2011) *Heliogomphus* cf. *retroflexus* Ris, 1912, (Odonata: Anisoptera: Gomphidae), a possible new record for Singapore. *Nature in Singapore*, **3**, 221–225.

Ngiam, R.W.J. & Cheong, L.F. (2016) The dragonflies of Singapore: an updated checklist and revision of the national conservation statuses. *Nature in Singapore*, **9**, 149–163.

Ngiam, R.W.J. & Dow, R.A. (2013) The larva of *Leptogomphus risi* Laidlaw from Singapore with a comparison to *Leptogomphus williamsoni* Laidlaw from Sarawak and congeners (Odonata: Anisoptera: Gomphidae). *Nature in Singapore*, **6**, 307–312.

Ngiam, R.W.J. & Leong, T.M. (2012) Larva of the phytotelm-breeding damselfly, *Pericnemis stictica* Selys from forests in Singapore (Odonata: Zygoptera: Coenagrionidae). *Nature in Singapore*, **5**, 103–115.

Ngiam, R.W.J., Sun, S.W. & Sek, J.Y. (2011) An update on *Heliogomphus* cf. *retroflexus* Ris, 1912, with notes on *Microgomphus chelifer* Selys, 1858 in Singapore (Odonata: Anisoptera: Gomphidae). *Nature in Singapore*, **4**, 95–99.

Ngoi, P.S., Tan, J. & Ngiam, R.W.J. (2011) New record of a dragonfly, *Zyxomma obtusum* Albarda, 1881 in Singapore (Odonata: Anisoptera: Libellulidae). *Nature in Singapore*, **4**, 241–244.

Novelo-Gutierrez, R., Sites, R.W. & Vitheepradit, A. (2014) New province record of *Rhinagrion* for Thailand and description of the larva of *R. mima* (Odonata: Zygoptera: Philosinidae). *Zootaxa*, **3852**, 562–568.

Novotný, V. & Basset, Y. (2000) Rare species in communities of tropical insect herbivores: pondering the mystery of singletons. *Oikos*, **89**, 564–572.

Orr, A.G. (2005) *Dragonflies of Peninsular Malaysia and Singapore*. Natural History Publications (Borneo), Sabah.

Orr, A.G. & Hämäläinen, M. (2003) *Guide to the Dragonflies of Borneo*. Natural History Publications (Borneo), Sabah.

Orr, A.G. & Ngiam, R.W. (2011) A description of the larva of *Heliaeschna uninervulata* Martin (Odonata: Aeshnidae) from Singapore, with notes on its relationships. *International Journal of Odonatology*, **14**, 163–169.

Orr, A.G., Ngiam, R.W. & Leong, T.M. (2010) The larva of *Tetracanthagyna plagiata*, with notes on its biology and comparisons with congeneric species (Odonata: Aeshnidae). *International Journal of Odonatology*, **13**, 153–166.

Pérez-Gutiérrez, L.A. & Montes-Fontalvo, J.M. (2011) Description of the last stadium larvae of *Argia medullaris* Hagen in Selys and *A. variegata* Förster (Odonata: Coenagrionidae). *International Journal of Odonatology*, **14**, 217–222.

Pianka, E.R. (1970) On r-and K-selection. *The American Naturalist*, **104**, 592–597.

Pramual, P. & Wongpakam, K. (2014) Association of black fly (Diptera: Simuliidae) life stages using DNA barcode. *Journal of Asia-Pacific Entomology*, **17**, 549–554.

Prasad, M. & Varshney, R.K. (1995) A check-list of the Odonata of India including data on larval studies. *Oriental Insects*, **29**, 385–428.

Puillandre, N., Lambert, A., Brouillet, S. & Achaz, G. (2012) ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, **21**, 1864–1877.

Puniamoorthy, N., Kotrba, M. & Meier, R. (2010) Unlocking the" Black box": internal female genitalia in Sepsidae (Diptera) evolve fast and are species-specific. *BMC Evolutionary Biology*, **10**, 275.

Ratnasingham, S. & Hebert, P.D. (2013) A DNA-based registry for all animal species: the Barcode Index Number (BIN) system. *PLoS One*, **8**, e66213.

Ruiter, D.E., Boyle, E.E. & Zhou, X. (2013) DNA barcoding facilitates associations and diagnoses for Trichoptera larvae of the Churchill (Manitoba, Canada) area. *BMC Ecology*, **13**, 5.

Srivathsan, A., Sha, J., Vogler, A.P. & Meier, R. (2015) Comparing the effectiveness of metagenomics and metabarcoding for diet analysis of a leaf-feeding monkey (*Pygathrix nemaeus*). *Molecular Ecology Resources*, **15**, 250–261.

Taberlet, P., Coissac, E., Hajibabaei, M. & Rieseberg, L.H. (2012) Environmental DNA. *Molecular Ecology*, **21**, 1789–1793.

Tang, H.B., Wang, L.K. & Hämäläinen, M. (2010) *A Photographic Guide to the Dragonflies of Singapore*. Published and distributed by Raffles Museum of Biodiversity Research, Department of Biological Sciences, National University of Singapore, Singapore.

Tennessen, K.J. (2010) The madicolous larva of *Heteropodagrion sanguinipes* Selys (Odonata: Megapodagrionidae). *Zootaxa*, **2531**, 29–38.

Thomas, M., Raharivololoniaina, L., Glaw, F., Vences, M. & Vieites, D.R. (2005) Montane tadpoles in Madagascar: molecular identification and description of the larval stages of *Mantidactylus elegans*, *Mantidactylus madecassus*, and *Boophis laurenti* from the Andringitra Massif. *Copeia*, **2005**, 174–183.

Trivinho-Strixino, S., Pepinelli, M., Siqueira, T. & de Oliveira Roque, F. (2012) DNA barcoding of *Podonomus* (Chironomidae, Podonominae)

enables stage association of a named species and reveals hidden diversity in Brazilian inselbergs. *Annales de Limnologie-International Journal of Limnology*, **48**, 411–423.

Victor, B.C. (2007) *Coryphopterus kuna*, a new goby (Perciformes: Gobiidae: Gobiinae) from the western Caribbean, with the identification of the late larval stage and an estimate of the pelagic larval duration. *Zootaxa*, **1526**, 51–61.

Wang, W.Y., Srivathsan, A., Foo, M., Yamane, S. & Meier, R. (2018) Sorting specimen-rich invertebrate samples with cost-effective NGS barcodes: validating a reverse workflow for specimen processing. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.12751.

Willson, M.F. & Armesto, J.J. (2006) Is natural history really dead? Toward the rebirth of natural history. *Revista Chilena de Historia Natural*, **79**, 279–283.

Wolf, L.L. & Waltz, E.C. (1984) Dominions and site-fixed aggressive behavior in breeding male *Leucorrhinia intacta* (Odonata: Libellulidae). *Behavioral Ecology and Sociobiology*, **14**, 107–115.

Wong, W.H., Tay, Y.C., Puniamoorthy, J., Balke, M., Cranston, P.S. & Meier, R. (2014) 'Direct PCR' optimization yields a rapid, cost-effective, nondestructive and efficient method for obtaining DNA barcodes without DNA extraction. *Molecular Ecology Resources*, **14**, 1271–1280.

Xu, Q.H. (2012) A description of the final stadium larva of *Leptogomphus elegans* Lieftinck, with a discussion of taxonomic characters of the larvae of the genus *Leptogomphus* Selys (Odonata: Gomphidae). *International Journal of Odonatology*, **15**, 25–29.

Yu, D.W., Ji, Y., Emerson, B.C., Wang, X., Ye, C., Yang, C. & Ding, Z. (2012) Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, **3**, 613–623.

Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. (2013) PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*, **30**, 614–620.

Zhou, X., Kjer, K.M. & Morse, J.C. (2007) Associating larvae and adults of Chinese Hydropsychidae caddisflies (Insecta: Trichoptera) using DNA sequences. *Journal of the North American Benthological Society*, **26**, 719–742.

Accepted 9 February 2018