Abstract. The parasitoid wasp family Ichneumonidae is among the most diverse groups of organisms, with conservative estimates suggesting that it contains more species than all vertebrates together. However, ichneumonids are also among the most severely understudied groups, and our understanding of their evolution is hampered by the lack of a robust higher-level phylogeny of this group. Based on newly generated transcriptome sequence data, which were filtered according to several criteria of phylogenetic informativeness, we developed target DNA enrichment baits to capture 93 genes across species of Ichneumonidae. The baits were applied to DNA of 55 ichneumonids, with a focus on Pimpliformes, an informal group containing nine subfamilies. Phylogenetic trees were inferred under maximum likelihood and Bayesian approaches, at both the nucleotide and amino acid levels. We found maximum support for the monophyly of Pimpliformes but low resolution and very short branches close to its base, strongly suggesting a rapid radiation. Two genera and one genus-group were consistently recovered in unexpected parts of the tree, prompting changes in their higher-level classification: *Pseudorhyssoa* Merrill, currently classified in the subfamily Poemeniinae, is transferred to the tribe Delomeristini within Pimplinae, and *Hemiphanes* Förster is moved from Orthocentrinae to Cryptinae. Likewise, the tribe Theroniini is resurrected for the *Theronia group* of genera (stat. rev.). Phylogenetic analyses, in which we gradually increased the numbers of genes, revealed that the initially steep increase in mean clade support slows down at around 40 genes, and consideration of up to 93 genes still left various nodes in the inferred phylogenetic tree poorly resolved. It remains to be shown whether more extensive gene or taxon sampling can resolve the early evolution of the pimpliform subfamilies.
Introduction

Ichneumonidae constitute the largest family of parasitoid wasps, with more than 25,000 species described (Yu et al., 2016) and with conservative estimates suggesting that several-fold more are yet to be discovered (Townes, 1969; Gauld et al., 2002a). Ichneumonid wasps in the vast majority of cases attack immature stages of holometabolous insects and, to a lesser extent, immature and mature spiders (Gauld & Dubois, 2006). They can also be predators of arachnid egg sacs (Austin, 1985) and, even more rarely, facultatively herbivorous (Gauld et al., 2002a). They play a crucial role in most terrestrial ecosystems by regulating the abundance of other species, including pests in agriculture and forestry (Narendran, 2001). Despite the fact that there are probably many more ichneumonid wasps than vertebrates (about 68,600 species; IUCN, 2017) the group has received only limited attention from systematists and evolutionary biologists. Indeed, the circumscriptions of, and phylogenetic relationships among, higher-level taxa within the family are still very poorly understood (Quicke, 2015). These limitations hinder inferences pertaining to evolutionary pathways of parasitoid lifestyles, major changes in host ranges, and coevolution with their hosts. The current classification of the family is based mostly on morphological data (e.g. Townes, 1969; Wahl, 1986, 1990; Gauld et al., 2000), and the only available molecular phylogeny of the family is based on a single gene (28S rRNA) that is highly sensitive to alignment parameters (Quicke et al., 2009). The status of ichneumonid taxonomy and phylogeny was summarized by Quicke (2015) and Broad et al. (2018). These authors pointed out that the current classification recognizing up to 42 extant subfamilies will probably be subject to future changes. Seven informal subfamily groups are currently recognized, mostly based on research on larval morphology and results from a phylogeny obtained from the combined analysis of the 28S rRNA gene and a morphological character matrix (Wahl, 1986, 1990; Quicke et al., 2009; Quicke, 2015). Three of these seven groups contain most of the ichneumonid species, i.e., Ichneumoniformes, Opiioniformes, and Pimpliformes (Wahl, 1986; Gauld, 1991; Wahl & Gauld, 1998; Quicke et al., 2009).

We focus here on the clade Pimpliformes, which includes nine subfamilies: Acaentinae, Collyriinae, Cylloceriniinae, Dia- criditinae, Diplazonitinae, Orthocentrinae, Pimplinae, Poemeni- inae and Rhyssinae. There is great variety in their exploited host groups and life-history strategies. Specifically, they include parasitoids of, in descending order of importance, Lepidoptera, Hymenoptera, Coleoptera, Diptera and Araneae; they act as ectoparasitoids and endoparasitoids, and as idiobionts and koinobionts (Wahl & Gauld, 1998; Quicke, 2015). Idiobionts stop the development of their hosts at the time of parasitization, with the wasp female usually paralysing the host larva or pupa permanently before laying her egg on top of or, more rarely, inside it (Askew & Shaw, 1986). Koinobionts, by contrast, allow their hosts to continue their development, often over several molting cycles; they are typically endoparasitoids, although the subfamily Tryphoninae and the Polysphincta genus-group in Pimplinae are conspicuous exceptions in being koinobiont ectoparasitoids (Gauld, 1988). All four strategies are represented by several subfamilies in Pimpliformes, and several transitions have probably taken place, both between idiobiosis and koinobiosis and between ecto- and endoparasitism within this group. Also, a single transition to Diptera as a host group has been postulated (Wahl & Gauld, 1998), but the phylogenetic relationships among the pimpliform subfamilies are too uncertain to draw any robust conclusions (Quicke et al., 2009).

In addition to studying subfamily relationships within Pimpliformes, we also use this empirical example to pose operational questions regarding molecular phylogenetic analyses. Namely, we ask how many and which genes are needed to resolve the pimpliform tree. A much-cited phylogenomics study by Rokas et al. (2003) examined the relationships among eight species of yeast as inferred from 106 genes (comprising 127 kb) and concluded that ‘very large concatenated data sets of ~20 genes were required to provide 95% bootstrap support for all nodes’ (Gatesy et al., 2007, p. 355). This number might seem low now, given that hundreds to thousands of genes are regularly sequenced to infer large phylogenies (dos Reis et al., 2012; Missot et al., 2014; Prum et al., 2015; Bank et al., 2017; Branstetter et al., 2017; e.g. Peters et al., 2017; Peters et al., 2018; Sann et al., 2018). However, the discussion triggered by this study (Phillips et al., 2004; Collins et al., 2005; Gatesy et al., 2007) focused on two questions that are now more timely than ever: what are the causes of systematic biases in phylogenomics, and how many genes are really needed if they are chosen carefully?

The importance of these questions is underlined by the increasing number of phylogenomic studies that arrive at contradictory conclusions, despite including hundreds to thousands of genes (e.g. McCormack et al., 2013a; Sharma et al., 2014; Prum et al., 2015; Streicher et al., 2016). The identification of systematic biases (Philippe et al., 2011) often requires reanalysis of the entire dataset, which can be prohibitive for very large datasets. The same applies to analyses that aim to study the influence of different taxon sampling schemes to assess the robustness of phylogenetic inference (Hedtke et al., 2006; Heath et al., 2008). Thus, increasing dataset size often comes at the cost of reduced rigour in the analytical methodology. The computational limitations imposed by many of the more advanced analytical methods, such as certain Bayesian inference methods, and the use of complex evolutionary models, represent another important argument for carefully choosing markers instead of simply increasing dataset size.

Several criteria have been suggested to aid in identifying the phylogenetically most informative genes, such as: (i) previous performance in a different, typically smaller taxon set (Salichos & Rokas, 2013); (ii) evolutionary rate and its distribution among sites (Yang, 1998; Townsend, 2007; Klopfstein et al., 2010; Townsend et al., 2012; Klopfstein et al., 2017); (iii) clock-like evolution (Doyle et al., 2015); and (iv) stationarity of nucleotide or amino acid composition (Collins et al., 2005). However, different studies of empirical datasets have arrived at conflicting conclusions about the performances of these criteria (Regier et al., 2008; Doyle et al., 2015), with the consequence that the pathway to efficient experimental design in phylogenomics remains obscure. It is even likely that identification of the best-performing approach will remain
developed scripts, which are provided here as supplementary material (File S1). We mostly used shell scripting (bash) to create a pipeline of programs for specific tasks and the ape package in the program r (Paradis et al., 2004; R Core Team, 2014) to manipulate alignments and trees. All calculations were performed on UBELIX (http://www.id.unibe.ch/hpc), the HPC cluster at the University of Bern, Switzerland.

Transcriptome data, orthology prediction and transcriptome phylogenies

Transcriptomes of two pimpliform wasp species were sequenced and published by Peters et al. (2017). We additionally sequenced whole-body transcriptomes of adult specimens for eight species from five pimpliform subfamilies and three other ichneumonids, including Xorides praecatorius Fabricius (Xoridinae), the hypothesized sister to all remaining ichneumonids (Quicke et al., 2009). For a full list of specimens used in this study, see Table 1 (and the more detailed File S2). As we sampled a single species per genus in most cases, we henceforth refer to genera only. The specimens for the transcriptome analysis were freshly collected in a mountainous area in the canton of Bern, Switzerland, and either identified alive after a cooling period in the fridge, or collected alongside other, identical-looking individuals at the same location, with the latter kept in ethanol as ‘pseudo-vouchers’ or ‘paragenophores’ for later confirmation of the preliminary identifications (File S2). Specimens chosen for the transcriptome analysis were either ground directly in RNAlater solution (Merck KGaA, Germany) or shock-frozen at −80°C. Total RNA was extracted using the ReliaPrep™ Tissue Miniprep system (Promega AG, Switzerland). Library preparation was done using the Illumina TruSeq stranded RNA method (Illumina Inc., Albany, NY, U.S.A.), which includes poly-A selection, followed by size selection. The resulting 11 libraries were indexed, multiplexed and sequenced on one Illumina HiSeq3000 lane (150 bp paired-end, 250 million reads) at the Institute of Genetics of the University of Bern (VetSuisse).

The raw reads from transcriptome sequencing were demultiplexed and trimmed using TRIMMOMATIC 0.33 (Bolger et al., 2014), which removed adapters, leading and trailing bases below a phred score of 10, and sections in a sliding window of size 4 with an average score <15. Read quality was assessed with FASTQC 0.11.2 (Andrews, 2014). Transcriptomes were assembled with TRINITY 2.1.1 (Grabherr et al., 2011) in paired-end mode on six CPUs and with a minimum memory setting of 24 Gb of RAM under default settings. Combined with transcriptomes from the 1KITE project (Peters et al., 2017), this resulted in 19 ichneumonid and six braconid transcriptomes (Table 1). The raw reads of the newly sequenced taxa are archived at the National Center for Biotechnology Information (GenBank) under BioProject PRJNA450386. To identify orthologous transcripts of putative single-copy protein-coding genes among these 25 transcript libraries, we searched them for 3260 genes suggested by the OrthoDB database (version 7, Waterhouse et al., 2013) to be single-copy in Hymenoptera (and in Neuroptera; note that the applied orthologue set was

Materials and methods

This study included five major steps: (i) transcriptome sequencing and analysis; (ii) estimation of phylogenetic informativeness of genes; (iii) hybrid capture laboratory work and data analysis; (iv) development and testing of oligonucleotide primers for PCR and Sanger sequencing, and (v) phylogenetic analyses. These individual steps are detailed in the following sections. Most of the bioinformatics for this study were conducted using newly developed scripts, which are provided here as supplementary

In any case, increasing the number of genes sampled can be expected to have a positive effect on the robustness of phylogenetic inference, at least up to a point, for instance by allowing identification of regions of uncertainty in a phylogenetic hypothesis that are due to incongruences among the individual gene trees (Salichos & Rokas, 2013). Next-generation sequencing facilitates the acquisition of unprecedented amounts of data, and techniques such as hybrid capture (also called anchored/hybrid enrichment or target DNA enrichment) enable cost-effective application even to nonmodel taxa (McCormack et al., 2013b). Hybrid capture approaches typically use RNA baits c. 100−150 bp in length, which show a close match to the target DNA sequences, in order to enrich the target DNA in a genomic extract (Mamanova et al., 2010). The enrichment step facilitates the pooling of large numbers of samples on a single sequencing lane, thus decreasing sequencing costs considerably. The technique has been applied to taxonomic groups with various levels of divergence, with capture success often still high in distantly related groups (Bragg et al., 2016; Mayer et al., 2016; Haddad et al., 2018; Percy et al., 2018). Given that only a handful of markers have been established for ichneumonid phylogenetics in the past (Quicke et al., 2000; Quicke et al., 2009; Klopstein et al., 2011; Rouse et al., 2016; Santos, 2017), and divergences within Ichneumonidae probably date back to the Late Cretaceous (Kopylov, 2010, 2012), a hybrid capture approach is very promising for inferring a robust phylogeny for the family.

In this study, we aimed to resolve the higher-level relationships within Pimpliformes to gain insights into their evolution, especially with respect to parasitoid strategy and host range. To this end, we used previously (Peters et al., 2017) and newly generated transcriptome data to identify candidate genes, filtered them by several criteria of phylogenetic informativeness, and created a set of baits to enrich 152 exons of a total of 93 genes. We used this bait set to enrich the target DNA in 55 species of ichneumonids and two distantly related outgroup taxa. We applied several methods for phylogenetic reconstruction and examined them for congruence and conflict. Furthermore, based on both random and nonrandom gene sampling strategies, we studied subsets of our data to assess strategies to improve phylogenetic studies of the group in the future.

highly context-specific, with small differences in taxon sampling and analytical approach having a big impact on the preferred strategy. Thus, only a few criteria might emerge that have general applicability.

© 2018 The Royal Entomological Society, *Systematic Entomology*, 44, 361–383
| Higher grouping (Sub)family | Genus | Species | Specimen ID | Data type | Country | Repository |
|-----------------------------|-------|---------|-------------|-----------|---------|------------|
| Outgroup: Chalcidoidea      | Pteromalidae | Thaumasura | Ichn-2467 | hybEnr | AU | WC |
| Outgroup: Gasteruptiidae    | Gasteruption | sp.       | Ichn-2470 | hybEnr | AU | WC |
| Outgroup: Alysiinae         | Dacnusa | sibirica | 1KITE | transcr | DE | – |
| Outgroup: Aphidiinae        | Aphidius | colemani | 1KITE | transcr | DE | – |
| Outgroup: Diaterus          | essigellae | 1KITE | transcr | AU | – |
| Outgroup: Helconinae        | Gen. | sp.       | Ichn-2469 | hybEnr | AU | – |
| Outgroup: Macrocenetrinae   | Macrocentrus | marginator | 1KITE | transcr | SE | – |
| Outgroup: Rogadinae         | Aleiodes | testaceus | 1KITE | transcr | FR | – |
| Outgroup: Anomalomalinae    | Heteropelma | amictum | 1KITE | transcr | CH | – |
| Outgroup: Banchinae         | Apophaia | evanescent | transcr-SK09 | transcr | CH | NMBS |
| Outgroup: Campopleginae     | Campoplex | capitator | Ichn-2519 | hybEnr | FR | NMBS |
| Outgroup: Hyposoter         | dudunator | 1KITE | transcr | FR | – |
| Outgroup: Dimorphora        | kentmartini | Ichn-2465 | hybEnr | AU | WC |
| Outgroup: Netelia           | Xenoschesis | fulvicornis | Ichn-2359 | hybEnr | CH | NMBS |
| Outgroup: Mesochorinae      | Astiphromma | sp.       | Ichn-2453 | hybEnr | US | KYU |
| Outgroup: Colpotrochia      | cineta | transcr-SK10 | transcr | CH | – |
| Outgroup: Netelia           | Leptophion | anici | Ichn-2462 | hybEnr | AU | WC |
| Outgroup: Netelia           | Nitella | cf. melanura | 1KITE | transcr | SE | – |
| Outgroup: Netelia           | Nitella | sp.       | 1KITE | transcr | HN | – |
| Ichnemoniformes             | Alomyina | sp.       | Ichn-2387 | hybEnr | CH | NMBS |
| Ichnemoniformes             | Cryptinae | Buathra | 1KITE | transcr | SE | NMBS |
| Ichnemoniformes             | Ichneumoninae | Ichneumon | 1KITE | transcr | CH | – |
| Labeniformes                | Labeninae | Ploeicryptus | Ichn-2466 | hybEnr | AU | WC |
| Pimpliformes                | Acaenitinae | Aores | Ichn-2341 | hybEnr | JP | NMBS |
| Pimpliformes                | Acaenitinae | Colocentrus | transcr-SK01 | transcr | CH | NMBS |
| Pimpliformes                | Acaenitinae | Jezarotes | Ichn-2338 | hybEnr | JP | NMBS |
| Pimpliformes                | Acaenitinae | Spiloterum | Ichn-2507 | hybEnr | US | KYU |
| Pimpliformes                | Acaenitinae | Yezoeryx | Ichn-2509 | hybEnr | TH | KYU |
| Pimpliformes                | Collyriinae | Collyria | transcr-SK11 | transcr | CH | – |
| Pimpliformes                | Cyllacrylinae | Cyllacryla | Ichn-2131 | hybEnr | SE | NMBS |
| Pimpliformes                | Cyllacrylinae | Cyllacryla | Ichn-2524 | hybEnr | US | KYU |
| Pimpliformes                | Diacritinae | Diacritus | Ichn-2522 | hybEnr | SE | NMBS |
| Pimpliformes                | Diplozontinae | Homotropus | Ichn-2365 | hybEnr | CH | NMBS |
| Pimpliformes                | Diplazontinae | Sussaba | Ichn-2513 | hybEnr | CH | NMBS |
| Pimpliformes                | Diplazontinae | Syphophillus | Ichn-2364 | hybEnr | CH | – |
| Pimpliformes                | Diplazontinae | Syphophillus | Ichn-2460 | hybEnr | CH | NMBS |
| Pimpliformes                | Diplazontinae | Meryngophus | Ichn-2502 | hybEnr | SE | SMTP |
| Pimpliformes                | Diplazontinae | Hemiphases | Ichn-2523 | hybEnr | SE | NMBS |
| Pimpliformes                | Diplazontinae | Hemiphases | Ichn-2499 | hybEnr | SE | NMBS |
| Pimpliformes                | Diplazontinae | Megastylus | Ichn-2183 | hybEnr | SE | NMBS |
| Pimpliformes                | Diplazontinae | Neurates | Ichn-2526 | hybEnr | SE | NMBS |
| Pimpliformes                | Diplazontinae | Orthocentrus | Ichn-2500 | hybEnr | SE | NMBS |
| Pimpliformes                | Diplazontinae | Picrostigus | Ichn-2498 | hybEnr | SE | NMBS |
| Pimpliformes                | Diplazontinae | Plectisidea | Ichn-2503 | hybEnr | SE | NMBS |
| Pimpliformes                | Diplazontinae | Proclitica | Ichn-2510 | hybEnr | SE | NMBS |
| Pimpliformes                | Diplazontinae | Acroclita | Ichn-2137 | hybEnr | SE | NMBS |
| Pimpliformes                | Diplazontinae | Anamophila | Ichn-2455 | hybEnr | EC | UTU |
| Pimpliformes                | Diplazontinae | Acanthoptila | Ichn-2454 | hybEnr | NI | UTU |
| Pimpliformes                | Diplazontinae | Aphecithis | transcr-SK05 | transcr | CH | NMBS |
| Pimpliformes                | Diplazontinae | Atractoacanthus | Ichn-2278 | hybEnr | SE | SMTP |
| Pimpliformes                | Diplazontinae | Camptotus | Ichn-2434 | hybEnr | CG | NHM |
| Pimpliformes                | Diplazontinae | Delomerista | Ichn-2528 | hybEnr | SE | KYU |
| Pimpliformes                | Diplazontinae | Dolichomitus | transcr-SK04 | transcr | CH | NMBS |
| Pimpliformes                | Diplazontinae | Dolichomitus | transcr-SK15 | transcr | CH | – |
| Pimpliformes                | Diplazontinae | Echthromorpha | Ichn-2464 | hybEnr | AU | WC |

© 2018 The Royal Entomological Society, *Systematic Entomology*, **44**, 361–383
The alignments obtained were used as blueprints to align the DNA sequences of the following species, which were part of the orthologue set designed by Peters et al. (2017), except that we removed the DNA sequences of the jewel wasp *Nasonia vitripennis* (Walker), and the corresponding nucleotide sequences with a modified version of PAUP* version 4.0a (Swofford, 2003), at both the amino acid and nucleotide levels. To examine the phylogenetic signal in the transcriptome data, we conducted phylogenetic analyses on this gene set with RAxML 8.1.2 (Stamatakis, 2014; see Misof et al., 2014 for details on modification). We used a custom R script to trim leading and trailing amino acid positions in the resulting multiple DNA sequence alignments of each gene to ensure that the flanking positions had at least 90% coverage across the 26 included taxa. To obtain candidate loci for which to assess phylogenetic informativeness, we used a coverage filter, retaining only genes with at least 600bp in total, at least 300bp present in every taxon, and at most 5% of the alignment consisting of gaps or missing data. Only 723 genes fulfilled these criteria and their alignments are provided as supplementary material (File S3). In order to ensure that there was no cross-contamination between taxa resulting from the pooled Illumina sequencing procedure, we checked the alignments of the 723 genes for raw pairwise distances <1%, and found only a single case that can be explained by a true close relationship (0.87% between *Dolichomitus imperator* and *Dolichomitus quericus* in gene Nasvi2EG0030079). To examine the phylogenetic signal in the transcriptome data, we conducted phylogenetic analyses on this gene set with RAxML 8.1.2 (Stamatakis, 2014), at both the amino acid and nucleotide levels.

© 2018 The Royal Entomological Society, *Systematic Entomology*, 44, 361–383

### Table 1. Continued

| Higher grouping | (Sub)family | Genus | Species | Specimen ID | Data type | Country | Repository |
|-----------------|-------------|-------|---------|-------------|-----------|---------|-----------|
| Pimpliformes    | Pimplinae   | *Flavopimpla* | nigromaculata | Ichn-2458 | hybEnr | TH | KYU |
| Pimpliformes    | Pimplinae   | *Liotryphon* | cf. punctulatus | Ichn-2361 | hybEnr | CH | NMBS |
| Pimpliformes    | Pimplinae   | *Nomosphacia* | pyramida | Ichn-2093 | hybEnr | TH | KYU |
| Pimpliformes    | Pimplinae   | *Parema* | nigrobalteata | Ichn-2497 | hybEnr | TH | KYU |
| Pimpliformes    | Pimplinae   | *Perithous* | secura | Ichn-2525 | hybEnr | SE | SMTP |
| Pimpliformes    | Pimplinae   | *Pimpla* | rufipes | transcr-SK03 | transcr | CH | – |
| Pimpliformes    | Pimplinae   | *Pimpla* | flavicosis | 1KITE | transcr | SE | NMBS |
| Pimpliformes    | Pimplinae   | *Pioaster* | pilosator | Ichn-2144 | hybEnr | CH | NMBE |
| Pimpliformes    | Pimplinae   | *Polysphincta* | tuberosa | Ichn-2518 | hybEnr | SE | NMBS |
| Pimpliformes    | Pimplinae   | *Reclinervellus* | nielseni | Ichn-2505 | hybEnr | CZ | NMBS |
| Pimpliformes    | Pimplinae   | *Schizopyga* | frigida | Ichn-2136 | hybEnr | SE | NMBS |
| Pimpliformes    | Pimplinae   | *Theronia* | hilaris | Ichn-2433 | hybEnr | US | NMBS |
| Pimpliformes    | Pimplinae   | *Xanthopimpla* | varmaculata | Ichn-2448 | hybEnr | TH | KYU |
| Pimpliformes    | Pimplinae   | *Zatypota* | percontatoria | Ichn-2139 | hybEnr | SE | NMBS |
| Pimpliformes    | Pimplinae   | *Zonopimpla* | sp. | Ichn-2457 | hybEnr | GU | UTU |
| Pimpliformes    | Poemeniinae | *Cnastis* | vulgaris | Ichn-2354 | hybEnr | JP | NMBS |
| Pimpliformes    | Poemeniinae | *Deuteroxorides* | elevator | transcr-SK07 | transcr | CH | – |
| Pimpliformes    | Poemeniinae | *Eugalta* | cf. strigosa | Ichn-2442 | hybEnr | TH | KYU |
| Pimpliformes    | Poemeniinae | *Poemenia* | hectica | Ichn-2234 | hybEnr | SE | NMBS |
| Pimpliformes    | Poemeniinae | *Pseudorhyssa* | nigricornis | Ichn-2344 | hybEnr | JP | NMBS |
| Pimpliformes    | Rhyssinae   | *Rhyssa* | amoena | transcr-SK08 | transcr | CH | – |
| Pimpliformes    | Rhyssinae   | *Rhyssina* | sp. | Ichn-2074 | hybEnr | TH | KYU |
| Pimpliformes    | Rhyssinae   | *Rhyssina* | sp. | Ichn-2343 | hybEnr | JP | NMBS |
| Xoridiformes    | Xoridinae   | *Xorides* | praecatorius | transcr-SK12 | transcr | CH | NMBS |

*Data type: transcr, transcriptome; hybEnr, hybrid enrichment data.*

*Country of origin of the specimen in ISO 3166-1 alpha-2 codes.*

*Repositories of voucher specimens: KYU, Department of Entomology, University of Kentucky, U.S.A.; NHM, Natural History Museum London; NMBE, Natural History Museum Bern, Switzerland; NMBS, Natural History Museum Basel, Switzerland; SMTP, Swedish Malaise Trap Project, Station Linné on Öland, Sweden; UTU, University of Turku, Finland; WC, Waite Insect collection, University of Adelaide, Australia. A dash indicates that no voucher specimens are available as they were ground for RNA extractions (if a museum collection is mentioned for specimens used for transcriptome sequencing, it refers to a pseudo-voucher/para- or syngenophore, i.e. identical morphospecies collected at the same locality as the specimen whose transcriptome was sequenced).
Partitioning and amino acid substitution models for each partition were obtained using partitionfinder2 (Lanfear et al., 2016) with the rcluster algorithm using equal weights on the clustering criteria rate, state frequencies, model parameters, and among-site rate variation. As raxml only allows for a single model across partitions when analysing nucleotide sequence data, even though the model parameters can be estimated separately for each partition, we applied the GTRCAT model (Stamatakis, 2006) when analysing the nucleotide alignments. When analysing the amino acid sequences, we applied the identified fixed-rate model from partitionfinder in combination with the CAT model of among-site rate variation. To assess strength in signal, we ran 100 nonparametric bootstrap replicates. As all nodes obtained maximum bootstrap support, we also conducted gene jack-knifing by randomly sampling sets of ten genes 100 times. Each of these datasets was analysed in raxml as amino acid alignments (under the PROTCATJTT substitution model, the most commonly selected model by partitionfinder on the full dataset) and as first plus second codon positions (using the GTRCAT nucleotide substitution model), partitioning by gene and/or codon position.

Estimating phylogenetic informativeness

To predict ordered phylogenetic informativeness for the 723 candidate genes, we first trimmed each down to 600bp that showed the best taxon coverage. We then obtained five predictors: achieved mean clade support per node, maximum likelihood (ML) quartet mapping score, sites with optimal evolutionary rates, clock-likeness, and the stationarity of nucleotide frequencies. First, to get a measure of clade support, we ran single-gene ML analyses under the GTR + Γ + I model (Yang, 1996), partitioning by codon position and with 50 bootstrap replicates. We then recorded the sum of bootstrap support values for 12 nodes which were unequivocally retrieved in all phylogenetic analyses of the concatenated transcriptome data and that are in agreement with taxonomy and with previous phylogenetic studies (Quicke et al., 2009): Braconidae, Ichneumonidae, Ichneumonidae without Xorides Latreille, Ophioniformes, Ichneumoniformes, Pimpliformes, (Pimplinae + Rhyssinae + Poemeniinae), Pimplinae (Rhyssinae + Poemeniinae), Pimplini, and the genera Pimpla Fabricius and Dolichomitus Smith. We refer to this measure as the ‘taxonomic bootscore’. We also separately recorded if the rooting of Ichneumonidae was in accordance with a previous molecular and morphological analyses (Quicke et al., 1999), i.e. whether Xorides was recovered as the sister group to all other ichneumonids.

Maximum likelihood quartet mapping (Strimmer & von Haeseler, 1997) has been shown to perform well as a predictor of phylogenetic informativeness in datasets simulated at varying evolutionary rates (Klopfstein et al., 2017). We obtained tree-likeness scores under the HKY model (Hasegawa et al., 1985) for 1000 randomly chosen taxon quartets, using a custom \texttt{r} script (File S1).

To estimate evolutionary rates at each site of the alignment, we first transformed the tree obtained from the ML analysis of the complete amino acid dataset into an ultrametric tree using the penalized likelihood function ‘chronopli’ from the ape package in \texttt{r} (Paradis et al., 2004; R Core Team, 2014). An ML estimate of the transition/transversion ratio was obtained for each gene using this tree and ape’s ‘optim.pml’ function. Transition/transversion ratios and an ultrametric tree were both used as input for the program DNAtRates (Olsen, 1996), which calculates site-specific evolutionary rates under ML. The number of sites that evolve at a rate between 0.1 and 0.4 expected substitutions per root-tip distance was recorded for each gene, a range found to result in near-optimal performance in a simulation study (Klopfstein et al., 2017). Clock-likeness of the genes was measured by running single-gene relaxed-clock analyses under the independent gamma rates (IGR) model (Lepage et al., 2007) in MRBAYES 3.2 (Ronquist et al., 2012b). The data were partitioned into first + second versus third codon positions and analysed under the GTR + Γ + I model, with the topology constrained to include the taxonomic groups that comprised the taxonomic bootscore (see earlier) to avoid convergence issues and obtain comparable results. The median estimate of the variance of the clock model (‘igrvar’ parameter) was used as a measure of clock-likeness.

To test for deviations from the stationarity hypothesis of nucleotide composition, we used the chi-squared test of homogeneity as implemented in \texttt{PAUP*} 4.0 (Swofford, 2002). This test does not account for phylogenetic signal in the data, but we assume that it is conservative and thus sufficient for our purpose of identifying and excluding genes that show a large variation in nucleotide composition across taxa. As homogeneity tests such as the one implemented in \texttt{PAUP*} have been criticized as biased (Jermiin et al., 2004), we also conducted pairwise sequence comparisons using Bowker’s matched-pairs tests of symmetry (Bowker, 1948) implemented in symtest version 2.0.47 (https://github.com/ottmi/symtest) (Ababneh et al., 2006). This test found some taxon pairs in every single gene that departed from the stationarity assumption, at least when all three codon positions were included in the test; as the median P-values of the homogeneity and of Bowker’s matched-pairs tests were highly correlated (cor = 0.56, P < 2.2e−16), we continued with the previous test for choosing among genes.

Aiming for a good balance between the number of enriched genes and pooled taxa, we developed baits of 120bp length each for 100 candidate genes using the program \texttt{BAITFISHER} (Mayer et al., 2016). In order to span the sequence variation that can be expected in such a taxonomically broad sample, we included separate baits for each cluster of sequences with at least 0.12 nucleotide sequence divergence for a specific bait region. We used the intron–exon boundaries from the Nasonia genome (ogs 2.0; Rago et al., 2016) to inform bait positions and, in a second step, filtered the obtained baits against the genome of Microplitis demolitor Wilkinson (annotation release 101, https://i5k.nal.usda.gov/Microplitis_demolitor; i5K Consortium, 2013), a member of Braconidae and, at least at the time of our bait design analyses, the closest relative to our ingroup with an annotated genome. Baits were only retained if
they had a single hit and 95% overlap in the *Microplitis* genome. We applied different tiling designs: we tiled five baits with an offset of 20 bp across a target coding exon in 51 genes (68 exons), three baits with an offset of 20 bp across a target coding exon in 36 genes (77 exons), and a single bait to capture a given target coding exon in six additional genes (eight exons). In total, we designed 5419 baits to capture 153 exons of a total of 93 genes. Bait nucleotide sequences are provided in *FASTA* format in the supplementary material (File S4) and were ordered as RNA baits from the company MYcroarray (now Arbor Biosciences, Ann Arbor, MI, U.S.A.).

**Hybrid enrichment laboratory work**

The wet laboratory procedure was as follows: genomic DNA was extracted using the Gentra Pure-gene extraction kit (Gentra Systems, Minneapolis, MN, U.S.A.). Extractions were from either single legs or whole bodies, with the extracted specimens kept as vouchers in a number of museum collections (File S2). DNA concentration was measured using fluorometric quantification (Invitrogen Qubit; Thermo Fisher Scientific, Waltham, MA, U.S.A.) for all samples and for about a quarter of the samples also with the Agilent 2200 TapeStation System (Agilent Technologies, Santa Clara, CA, U.S.A.). DNA was sheared to an average fragment distribution of about 400 bp using a Bioruptor sonicator (Diagenode, Liège, Belgium), with the shearing success subsequently verified on the TapeStation. Libraries were constructed following Meyer & Kircher (2010), using reagents from Thermo Fisher Scientific (Waltham, MA, U.S.A.) and unique combinations of dual i7 and i5 index sets (Hugall *et al.*, 2015; Glenn *et al.*, 2016) to allow for subsequent pooling. BeckmanCoulter AMPure XP beads (Brea, CA, U.S.A.) were used for all purifications following the bead ratios specified by Meyer & Kircher (2010). We assessed the success of library preparation by measuring DNA concentration with fluorometric quantification and for about a third of the samples also with a standard quantitative PCR run using the LightCycler 96 Real-Time PCR System (Roche Diagnostics, Switzerland). In some of the libraries, we detected adapter dimers after running them on a standard electrophoresis gel. We cut these from an agarose gel, which had been stained with GelGreen (Biotium, Fremont, CA, U.S.A.), and purified the gel slices excluding the adapter dimers using Qiagen MinElute Gel Extraction columns (Qiagen, Switzerland) and the manufacturer’s protocol. Prior to capture, we divided the target baits into half-capture reactions. We enriched one to four samples per half-capture depending on the importance of the taxa and the similarity of the DNA concentration in the libraries. The target capture was performed with a 24-h hybridization at 65°C following the MYbaits protocol v3 from Microarray (Arbor Bioscience, Ann Arbor, MI, U.S.A.). The post-capture amplification was completed using KAPA HiFi DNA Polymerase (Kapa Biosystems, Wilmington, MA, U.S.A.) with the following PCR protocol: 98°C for 30 s, followed by 12 cycles of 98°C for 20 s, 60°C for 30 s, and 72°C for 1 min, with a final extension of 5 min at 72°C. Illumina IS5 and IS6 primers were used to attach to the i7 and i5 ends of the library. We verified enrichment success with a final run on the Agilent 2200 TapeStation System. Size selection was performed using AMPure XP beads in order to remove the small nucleotide fragments that may have been introduced during the capture. Samples were multiplexed into three pools of 14, 24 and 23 samples at equal concentrations and were run on one lane each on the Illumina MiSeq (pool 1) and the Illumina Next-Seq platforms (pools 2 and 3), obtaining 150-bp paired-end reads (Australian Genome Resource Facility, Adelaide, Australia).

**From raw reads to alignments**

We closely followed the strategy described by Bragg *et al.* (2016), which is tailored for hybrid capture data from highly divergent taxa, but developed our own scripts. All scripts used in our bioinformatics pipeline are publicly available, including detailed comments (File S1). The paired-end reads were demultiplexed and partial adapter sequences removed using the program **cutadapt** (Martin, 2011). We used **TRIMMOMATIC** (Bolger *et al.*, 2014) to remove leading and trailing bases below a phred score of 30 and sections in a sliding window of size 4 with an average score < 25. Read quality was ascertained with **fastqc** 0.11.2 (Andrews, 2014). blastx (Altschul *et al.*, 1990) was used to search for reads in each sample that showed a significant match with each of the enriched genes, a strategy making use of the more conserved amino acid regions in comparison with nucleotide sequences across taxa separated by long divergence times (Bragg *et al.*, 2016). Only after this step were reads assembled into contigs separately per gene and sample, first, by using the program **velvet** (Zerbino & Birney, 2008) to generate contigs under a range of **Kmer** values (25–75 bp), and second, by merging (some of) the resulting contigs with the software **cap3** (Huang & Madan, 1999). To ascertain that the identified contigs indeed represent the targeted single-copy genes, we used **BLASTX** to align the contigs to the predicted mRNA of the *Nasonia* official gene set (ogs 2.0; Rago *et al.*, 2016), only retaining best matches. For most genes, this resulted in more than a single contig, as introns were often too long to be fully captured. To identify intron–exon boundaries, we used the ‘protein2genome’ method in the aligner **exonerate** (Slater & Birney, 2005) with all the transcriptome sequences as queries, keeping the longest match. **Exonerate** outputs the sequence identities of its matches, and we removed those that were < 60%, a threshold obtained by comparing the transcriptome sequences among each other. After translation into amino acids, the remaining matching contigs were aligned to all transcriptomes and to each other in the multiple-sequence aligner **MAFFT** using the E-INS-i algorithm (Katoh & Standley, 2013). The contigs were then translated back into nucleotide sequences using a custom **r** script. As **exonerate** only finds complete introns, and contigs often started and/or ended in intronic sections, we developed an **r** script that identifies leading and trailing intron portions based on a drop in the pairwise distance when intron–exon boundaries are crossed. This script identified many but not all remaining intronic sequences, leaving out mostly those overhangs which were too short to calculate a meaningful distance. Thus, we had to correct some alignments manually (a log of these changes is given in File S5). The remaining contigs were used as queries
Table 2. Established oligonucleotide primer pairs for Sanger sequencing of eight exons. Gene numbers correspond to the Nasonia genome annotation and are preceded by ‘Nasvi2EG’ on GenBank.

| Gene   | Primer pair | Primer sequence 5′ – 3′ | Annealing temperature (°C) | Fragment length         |
|--------|-------------|-------------------------|----------------------------|-------------------------|
| 014339 | F           | GCC AAG AGG GTC AAA TTA CA | 51 – 52                   | 142 bp + intron + 190 bp |
|        | R           | CTC CAC AAT CTT TCC ATG CC |                           |                         |
| 011390 | F           | GAT GCT CGA AGT TAC AGC TA | 49 – 50                   | 357 bp                  |
|        | R           | GGA GTG CCG GTC ATA AAG TC |                           |                         |
| 013087 | F1          | CTG AAG ACC ATT TTC CTG CG | 51 – 52                   | 389 bp                  |
| R1     |             | GCG ACC TTG GAA GCA TCT TG |                           |                         |
| F2     |             | CAG CAG CAA CGT GGA ATG G  | 51                        | 413 bp                  |
| R2     |             | GCG ACC TTG GAA GCA TCT TG |                           |                         |
| 001408 | F           | ACC GGT GGT CAT CAT GC   | 51                        | 377 bp                  |
|        | R           | CCA CAA TTT GAA TGT TCG ATC |                         |                         |
| 003698 | F           | GGA AGA AGT CGT ATT TGA GGA A | 51 – 52            | 536 bp                  |
|        | R           | CGA GCA GAG AGA GTA CAT G  |                           |                         |
| 010812 | F           | GAT TAT CGT CCG TAC TTC AC | 48 – 49                   | 140 bp + intron + 205 bp |
|        | R           | TAT CTC TCG AGG GGA ATC AT |                           |                         |
| 008805 | F           | GAG GAT GTC TCG AGC GC    | 52 – 53                   | 328 bp                  |
|        | R           | CTA GCG AAA TCA TAA GCG TAT C |                   |                         |
| 021461 | F           | GGA GTC GAC ATA AAC GGA TA | 50 – 51                   | 404 bp                  |
|        | R           | GTG ATG TAG CCG TGT CTG AT |                           |                         |

in a BLASTN search against the original reads with a restrictive e-value of 10^{-10} to obtain coverage values and perform variant calling. Sequences were retained from a minimum coverage of eight identical bases, with variants called as ambiguity codes at a minimum coverage of four or at least 10% of the total coverage, with most positions obtaining much higher coverage (File S6). The alignments of all 93 genes are provided as supplementary material (File S7).

Primer development for sanger sequencing

One of the main hindrances for developing PCR oligonucleotide primers for amplification and direct Sanger sequencing of nuclear protein-coding genes is the short average length of exons in most arthropod genes. This is problematic because both the length variability and high AT content of introns can pose problems for successful PCR and sequencing. We initially attempted to develop primer pairs for the eight genes with the highest phylogenetic informativeness, but in almost all cases found several introns ranging up to more than 1000 bp in length within each of the target segments. To overcome this issue, we used the intron–exon boundaries identified in the Nasonia genome, together with the data obtained from the hybrid enrichment, to search for long exons (> 300 bp) or for exons separated by an intron that was relatively short (< c. 200 bp in most of the taxa whose genes we captured via hybrid enrichment; Table 2). Primer design was done manually in Geneious 7.1.3 based on the Nasonia genome and the obtained ichneumonid transcriptome and hybrid capture sequences, while primer synthesis was by Microsynth (Balgach, Switzerland). The primer pairs (Table 2) were tested in a three-step procedure: (i) touchdown PCR with a provisional annealing temperature on eight taxa sampled across ichneumonid species to ascertain broad applicability.

Phylogenetic analyses

The final set of alignments covering the 93 enriched genes was subjected to various phylogenetic analyses. First, we analysed it under the ML criterion in RAXML 8.1.2 (Stamatakis, 2014) after determining the partitioning scheme in PARTITIONFINDER2 (Lanfear et al., 2016), as described for the phylogenetic analyses of the transcriptomes. As RAXML only allows for a single substitution model for all nucleotide partitions, we ran the GTRCAT model for the nucleotide alignments, while the amino acid alignments were analysed under the partition-specific substitution model chosen by PARTITIONFINDER2 and the CAT model of among-site rate variation. To assess signal strength, we ran 1000 nonparametric bootstrap replicates. The concatenated genes were analysed at the nucleotide level by including: (i) all codon positions; and (ii) first and second codon positions only. We additionally analysed the concatenated dataset at the amino acid level. The amino acid alignment was also subjected to Bayesian analyses using fixed amino-acid substitution models.
(as obtained from partitionfinder2) in mrbayes 3.2 (Ronquist et al., 2012b) and the empirical profile mixture model implemented in phylobayes-mpi 1.7 (Lartillot & Philippe, 2004). In the analyses with mrbayes, we increased the relative probability of the topology and branch-length moves to 200 in order to counterbalance the large number of substitution model parameters estimated separately for the 56 partitions. In both Bayesian analyses, we conducted four independent runs to assess convergence.

Reconstructing the evolution of parasitoid lifestyle

As an example of an important aspect of the parasitoid lifestyle, we used ancestral state reconstruction to examine transitions between idiobiosis and koinobiosis in ichneumonids. Biological data for the included genera were extracted from several sources (Wahl & Gauld, 1998; Gauld et al., 2002b; Broad et al., 2018). For many ichneumonid genera, no sufficiently detailed records of their parasitoid strategy exist, but we assume that it typically follows a pattern shared by closely related genera and even members of the same subfamilies. Thus, we extrapolated a ‘ground-plan’ from related taxa in several cases (see File S8 for details). Ancestral states were reconstructed under maximum parsimony using the ‘hsp_max_parsimony’ function in the R package ‘castor’ (R Core Team, 2014; Louca & Doebeli, 2018), which allows for reconstruction of missing tip data. We assessed the impact of topological uncertainty on the number and direction of evolutionary transitions by including consensus and ML trees from the various Bayesian and ML analyses, plus 1000 trees each from the Bayesian posterior distributions as obtained from the analyses of the amino acid alignments with phylobayes and mrbayes, respectively.

Increase of tree resolution with dataset size

To assess the relationship between the number of genes included and the mean clade support of the resulting tree, we randomly sampled two to 90 genes, with 20 replicates for up to 20 genes and 10 replicates from 30 genes. Each concatenated gene set was analysed in raxml at the nucleotide level, partitioned by codon position and using the GTRCAT model with 50 bootstrap replicates. Mean clade support was measured as the sum of bootstrap support values > 50%. In order to re-evaluate the success of the criteria ‘clade support’, ‘clock-likeness’ and ‘evolutionary rates’, we estimated their scores again under the full taxon sampling for each gene separately. For the clade support measure, raxml analyses were conducted at nucleotide level, partitioning by codon position and under a GTRCAT model. Genes were ranked according to decreasing median clock variance. To identify regions in the tree with lower support, we performed gene jack-knifing based on 100 datasets of 10 randomly chosen genes each, analysed at the amino acid level or including first and second codon positions. These analyses mostly, but not always, recovered the same nodes that were in conflict between datasets.

order to force a reasonable rooting in the clock analyses (Ronquist et al., 2012a): Ichneumonidae, Braconidae, and Ichneumonidae + Braconidae. Analyses were run for 10 million generations, a value that was increased to 15 million generations if the average standard deviation of split frequencies (ASDSF), a measure for topology convergence, had not dropped below 0.02. The genes were then sorted according to descending median clock variance. As a third criterion, we estimated site-specific rates in the program dnarates (Olsen, 1996) as done previously for the transcriptome sequence data, and ranked the genes according to the number of sites in the ‘near-optimal’ rates bracket (according to Klopstein et al., 2017).

Results

Transcriptomes and phylogenetic informativeness

An average of 10^8 bp (8.2 × 10^7–1.3 × 10^8) were sequenced from each transcriptome library, which assembled into c. 10^6 contigs (7.2 × 10^5–1.5 × 10^6). Both these numbers are about 3.6 times larger than those of the transcriptome data published by Peters et al. (2017) (File S9), which can be explained by the fact that we pooled fewer samples per lane. Searching for 3260 single-copy genes, we found comparable numbers in our transcriptomes and those published by Peters et al. (2017) (3059 vs 2969). However, different genes were missing in different taxa, so that only 1908 genes were present across all 25 transcriptomes. Filtering these for sections at least 600 bp long and good coverage across all taxa resulted in 723 genes spanning more than 568 kbp in length.

Analysing the above 723 genes as concatenated amino acids, first and second codon positions, and full nucleotide alignments under ML, we recovered partly conflicting trees (Fig. 1), each of which obtained maximal bootstrap support at every node. Such a result might be expected from a taxon sampling as sparse as this one, even when a large number of genes are analysed, and thus we focus on the few well-supported relationships and on identifying the main regions of conflict among the transcriptome trees. As expected from previous studies, Xorides was recovered as sister to the remaining ichneumonids (Quicke et al., 1999), and the higher taxonomic groupings Ophioniformes, Ichneumoniformes and Pimpliformes (as defined in Quicke et al., 2009) were recovered in all three datasets. However, the relationships among the pimpliform subfamilies showed conflicts, and the subfamily Pimplinae, which was recovered as monophyletic in the amino acid analysis, was split into two groups in the nucleotide analyses, both when including and when excluding third codon positions. The two analyses at the nucleotide level showed a single conflict with regard to the placement of Rhyssa Gravenhorst either as sister to Deuterexorides Viereck + Pimplinae II (Fig. 1) or as sister to (Collyrina Schiötte + Coleocentrus Gravenhorst) + Dolichomitus. In order to identify regions in the tree with lower support, we performed gene jack-knifing based on 100 datasets of 10 randomly chosen genes each, analysed at the amino acid level or including first and second codon positions. These analyses mostly, but not always, recovered the same nodes that were in conflict between datasets.

© 2018 The Royal Entomological Society, Systematic Entomology, 44, 361–383
Hybrid capture success and Sanger sequencing primers

All 93 target genes were enriched successfully, with an average length of the captured coding target sequence of 376 bp per taxon after quality and coverage filtering (across genes: median = 415 bp, min = 150 bp, max = 1197 bp; see File S9). Calculated across all 59 taxa in the hybrid capture approach, the minimum coverage across the target genes was 331 bp,

### Table 3. Summary of gene choice criteria.

| Criterion                  | Mean/ median | 95% range | No. of genes |
|----------------------------|--------------|-----------|--------------|
| Stationarity               | P > 0.05     | 0.365/0.212 | 0.000–0.999 | 449 |
| Clock-likeness             | igvar < 0.015 | 0.045/0.017 | 0.004–0.067 | 305 |
| Root correct               | Bootstrap > 0.5 | 0.52/0.52     | 0.04–0.90 | 380 |
| #SiteRates                 | > 80            | 117.7/115.5 | 73–178       | 683 |
| Taxonomic bootstrap score  | for final ranking | 259.2/263.5 | 137–369      | – |

*Criteria: Stationarity, homogeneity of nucleotide composition across taxa, tested by chi-squared test, with genes with P-values > 0.05 retained; clock-likeness, variance of independent gamma rates (IGR) relaxed-clock model in single-gene analysis, genes retained above 0.015; root correct, bootstrap support for placement of Xorides as sister to all other Ichneumonidae, gene retained if it was above 50%; #SiteRates, number of sites evolving at near-optimal rate (according to Klopfeinstein et al., 2017: between 0.1 and 0.4 substitutions from root to tip), genes retained if they had more than 80; taxonomic bootstrap, sum of bootstrap support for 12 groupings chosen as unequivocal in both our and previous analyses.

*Number of genes that would have been retained under each of the criteria if applied individually.
and on average 90 genes were recovered with more than 100 bp. The exception was the outgroup sample PTE_Thaumasura from the superfamily Chalcidoidea, which had much lower enrichment success (108 bp, on average, across genes, and only 19 genes with > 100 bp). It is unclear if this was due to its taxonomically more distant position from the taxa on whose DNA the bait design was based, or due to the low quality of the DNA extract from the rather small specimen that we studied. Another even more distantly related outgroup, Gasteruption Latreille (Evanoidae), showed only slightly lower success when compared with the ingroup samples (337 bp, on average, per gene, 83 genes with > 100 bp). Comparing the coverage of the hybrid capture with the transcriptome sequence data, we found that the latter performed better, with an average of 436 bp across taxa and genes. The full alignment of 93 genes for 84 taxa contained 41 565 bp with a gap proportion of 14%, a value that dropped to 12.4% when only the 75 ichneumonid taxa were included.

Trying to identify comparatively long exons in the entire 93-gene dataset, we found only four exons that were longer than 450 bp and eight more that were longer than 350 bp. Nevertheless, we developed oligonucleotide primers for eight genes that amplified and sequenced with good consistency (Table 2, File S10). In one case (gene NasViEG013087), a combination of two primer pairs were needed to obtain target amplicons from DNA of the majority of taxa. In two cases, the amplified fragments spanned one intron, which was almost consistently about 100 bp in length, even in distantly related taxa. The exon portion of the fragments ranged between 328 and 536 bp (Table 2) and was amplified in species across the entire ichneumonid tree, representing most of the tested subfamilies. Success was lower in the two specimens from different families (Braconidae and Gasteruptiidae), but additional testing is needed to see if that was due to low quality of the extracts or divergent sequences.

Phylogeny of Pimpliformes

The five different phylogenetic analyses that we performed recovered different trees depending on data type and analysis method, with most incongruence located close to the base of the pimpliform radiation (Figs 2, 3). This area of the tree also exhibited very short branches. The incongruent nodes were mostly the same as those conflicting between the nucleotide versus amino acid versions of the transcriptome analyses, but appeared much more complex because of the inclusion of additional subfamilies in the hybrid capture taxon sampling (Fig. 1). For several nodes at the base of the pimpliform radiation, no analysis retrieved a significant signal (clade support < 0.95 posterior probability or 85% bootstrap support: Fig. 3), and most of the gene trees were also unresolved at these nodes (File S11). By contrast, the base of the inferred tree was well resolved. Xorides as the representative of Xoridinae consistently appeared at the base of a monophyletic Ichneumonidae, and the three informal subfamily groups, Ophioniformes, Ichneumoniformes, and Pimpliformes, were recovered with maximum support in all analyses. The only exception from this consistently well-supported resolution was the placement of the only representative of the Labeninae (Poecilocryptus Cameron), which branched off either right after the branch leading to Xorides (Fig. 3b, c), Ophioniformes (Fig. 3a, d, e) or Ichneumoniformes (Fig. 3f). The latter would imply a sister-group relationship between Labeninae and Pimpliformes.

The monophyly of most pimpliform subfamilies was well supported by all analyses. Exceptions were Pimplinae (because the genus Xanthopimpla Saussure, and sometimes also the tribe Pimplini, took up a position closer to the base of the pimpliform radiation) and Acaenitinae (with the Acaenitini recovered separately from Coleocentridae). Pimplinae (excluding Xanthopimpla) retrieved maximum support only in the amino acid analyses (Fig. 3a–d), whereas Orthocentrinae were only supported by the Bayesian analyses of the amino acid data (Fig. 3a–c). The only relationships among higher groups of Pimpliformes that were recovered consistently were sister-group relationships between Rhysinae and Poemeniinae, between the pimpline tribe Delomeristini and the Theronia genus-group, and (in all but the ML analysis of codon positions 1 and 2) between Orthocentrinae and Diacritinae. The pimpline tribe Ephialtini was consistently recovered as monophyletic, which was also true for Pimplini (excluding the Theronia genus-group) and for Delomeristini (including Pseudorhyssa Merrill).

Three taxa were consistently recovered in places that are in conflict with their current classification (Figs 2, 3). The genus Pseudorhyssa, regarded as a member of Poemeniinae since Gauld (1991), was placed with maximum support in all analyses in the tribe Delomeristini (Pimplinae). The Theronia group of genera (represented in our analyses by one species each of Theronia Holmgren, Nomosphecia Gupta and Parema Gupta) appeared as the sister to Delomeristini, including Pseudorhyssa. And Hemiphanes Förster, which has been classified as a member of the informal Helictes genus-group within Orthocentrinae, grouped with Ichneumoniformes with maximum support, even though its position within Ichneumoniformes is not resolved. In the cases of both Hemiphanes and Pseudorhyssa, the new placements were also recovered in most of the single-gene trees, but they were never placed with the subfamilies in which they are currently classified (File S11). The placement of Pseudorhyssa with Delomeristini was supported in 89 out of 93 gene trees, 44 of them with bootstrap support of at least 85%, and placement of Hemiphanes with Ichneumoniformes was supported by 81 gene trees, 22 of them with at least 85% bootstrap support. The Theronia group showed more variation in its placement among the single-gene trees, appearing with Delomeristini in 25 and with Pimplini in 15, whereas it showed various placements in the remaining gene trees (File S11).

Inferring the evolution of parasitoid lifestyles

Despite the low resolution at the base of the pimpliform radiation, the maximum parsimony reconstructions of parasitoid ecology retrieved very similar numbers of transitions across analyses (Fig. 3, Table 4). However, the directions of change within Pimpliformes varied strongly among analyses, with the most recent common ancestor of Pimpliformes recovered as
Fig. 2. Consensus tree resulting from the Bayesian analysis of the amino acid sequence data under an empirical profile mixture model as implemented in PHYLOBAYES, with support values added from the partitioned Bayesian analysis in MRBAYES, maximum likelihood analysis in RAxML, and the nucleotide analyses in RAxML. Support values represent Bayesian posterior probabilities and bootstrap supports, respectively. Note the very short branch lengths at the base of the Pimpliformes. For taxa with an asterisk after the name, we have analysed the transcriptomes; *Nasonia vitripennis* was included based on its published genome data. Current subfamily classification of the pimpliform genera is indicated by coloured taxon labels, which start with three letters corresponding to the subfamily name or to the family name (upper-case letters) in the outgroups as follows: Aca, Acaenitinae; Alo, Alomyinae; Ano, Anomaloninae; Ban, Banchinae; BRA, Braconidae; Cam, Camphylopilinae; Col, Collyriinae; Cre, Cremastinae; Crypt, Cryptinae; GAS, Gasteruptiidae; Ich, Ichneumoninae; Lab, Labeninae; Mes, Mesochorinae; Met, Metopini; Oph, Ophioninae; Ort, Orthocentrinae; Pim, Pimplinae; Poe, Poemeniinae; PTE, Pteromalidae; Rhy, Rhyssinae; Ter, Tersioliinae; Try, Tryphoninae; Xor, Xoridinae. [Colour figure can be viewed at wileyonlinelibrary.com].

© 2018 The Royal Entomological Society, Systematic Entomology, 44, 361–383
Fig. 3. Legend on next page
Table 4. Number of transitions between idiobiont and koinobiont parasitoid strategy as inferred from trees from the Bayesian posterior sample or from the maximum likelihood bootstrap analyses, along with the average of the estimated probabilities that the ancestor of Pimpliformes was an idiobiont.

| Analysis     | Number of transitions | Probability (MRCA = idiobiont) |
|--------------|-----------------------|--------------------------------|
| PhyloBayes (AA) | 5.0 5 5 6 0.99 |                                  |
| MrBayes (AA)   | 5.8 6 5 6 0.18 |                                  |
| Raxml (AA)     | 5.5 6 5 7 0.58 |                                  |
| Raxml (12)     | 5.2 5 5 7 0.98 |                                  |
| Raxml (123)    | 5.1 5 5 6 0.99 |                                  |

MRCA, most recent common ancestor.

an idiobiont with high probability in the PhyloBayes and nucleotide analyses, whereas it was retrieved as a koinobiont in most trees for the MrBayes analysis of the amino acid data (Table 4; the bootstrap trees from the ML analysis at the amino acid level were equivocal). Nevertheless, it is interesting to note that the idiobiont and koinobiont Pimpliformes tended to cluster together in most of the analyses, with the exception of Xanthopimpla (in most trees) and the derived Polysphincta group of genera. For two included taxa, the mode of parasitoidism is unknown; the reconstruction for Diacritinae was always in favour of it being a koinobiont, whereas it was sometimes equivocal for Coleocentrus.

How many genes do we need?

Random subsampling of the full dataset of 93 genes resulted in a curve of increasing mean clade support with increasing dataset size (Fig. 4). The curve first showed a very steep increase up to about 40 genes, after which the improvement slowed down. Even at 93 genes, only a 90% mean clade support was achieved, with lower values mostly due to the basal pimpliform nodes (Figs 2, 3). However, it is unclear whether saturation was reached, which suggests that additional gene sampling could lead to a further increase in tree resolution even under the same taxon sampling and analytical method.

When not sampling genes randomly, but instead ranking them according to different criteria, we found a clear improvement for up to about 30 genes under all criteria except the ranking by clock-likeness (Fig. 4). The rankings of the different criteria were very highly correlated (File S12), with the highest correlation coefficients between gene length and the number of sites evolving at a near-optimal rate, while clock-likeness showed the weakest correlation with the other criteria. Thus, gene length seems to be the main factor determining differences in single-gene success in our hybrid capture dataset, even though the evolutionary rate and clade support criteria performed slightly better between five and 10–30 genes (Fig. 4).

When comparing the ranking obtained from the full hybrid capture dataset with the gene rankings obtained from the transcriptome sequence data, we found no significant correlation between gene rankings, even though the bootstrap score and mean clade support measures were weakly correlated with each other (Spearman, $\rho = 0.18$, $P = 0.08$; File S12). Ranking genes according to their performance in the transcriptome sequence data also did not lead to any improvement over random gene sampling (results not shown). Besides the obviously different taxon sampling, the two datasets also differed strongly in gene length, as the hybrid capture was only enriching some of the exons, whereas nearly complete genes were available in the transcriptome sequence dataset.

Discussion

New molecular resources for Ichneumonoidea

The bait set for hybrid enrichment proposed here proved highly efficient in enriching 152 exons of 93 genes, not only across the target group, the ichneumonids, but also in the sister family Braconidae and in the genus Gasteruption, a member of a distantly related group of apocritan wasps (Evanioidea: Gasteruptiidae). Given the small number of genes previously used in ichneumonid phylogenetics (the maximum being seven for the subfamily Cryptinae; Santos, 2017), this bait set represents a significant step forward. We also provide oligonucleotide primers for PCR and Sanger sequencing of eight markers that were tested successfully across Ichneumonidae and several outgroups. This set complements a recent study very well, which suggested oligonucleotide primers for apocritan Hymenoptera, but did not include any representatives of Ichneumonoidae (Hartig et al., 2012). The largest obstacle for developing PCR primers is the large density of introns and corresponding short exon lengths found in many arthropod taxa. Despite the fact that the divergence between Nasonia and Ichneumonidae probably dates back to the Triassic (Peters et al., 2017), most of the intron–exon boundaries were found to be conserved, which might be because introns typically insert at certain sequence motifs (Dibb & Newman, 1989; Klopfstein & Ronquist, 2013). This high conservation of intron positions across groups encourages the use of distantly related models for

Fig. 3. Phylogenetic uncertainty at the base of Pimpliformes shown by comparison of consensus phylogenies from the Bayesian analyses (a, c), and maximum likelihood (ML) trees from the ML analyses (d–f). Support values in the form of posterior probabilities (a, c) or bootstrap percentages (d–f) are only shown if $< 1.0$ or $100\%$, respectively. Ancestral states of idiobiosis versus koinobiosis are shown as pies next to nodes, and stars indicate state transitions as inferred under maximum parsimony. Panel (b) shows one of the trees from the PhyloBayes analysis that shows an alternative reconstruction, with the reconstruction of the most recent common ancestor of Pimpliformes in favour of it being a koinobiont (arrow), as is also the case in panel (c). Outgroups and additional taxa within Ophioniformes and Ichneumoniformes were suppressed here for clarity of the figures. [Colour figure can be viewed at http://wileyonlinelibrary.com].

© 2018 The Royal Entomological Society, Systematic Entomology, 44, 361–383
intron–exon boundaries for designing phylogenetic markers, be it for PCR amplification or, in bait design, for hybrid captures.

Of the eight genes for whose amplification we developed PCR primers, two actually included an intron, but one that was found to be short in almost all the taxa we sequenced. Preliminary analyses showed that the captured intronic sequences could not be aligned at the deep phylogenetic divergences found in our dataset, but they might prove useful in future studies as markers for species delimitation. Given previous reports of shortcomings of the standard mitochondrial barcode locus cytochrome oxidase subunit 1 (COI) in insects, and especially ichneumonids (Virgilio et al., 2010; Bergsten et al., 2012; Nicholls et al., 2012; Klopstein, 2014; Klopstein et al., 2016), additional nuclear markers that are sufficiently variable at the species level, such as introns, are sorely needed. However, it remains to be shown how useful our intronic markers are at delimiting species.

**Taxonomic implications of the pimpliform tree**

Our combined transcriptome and hybrid capture data represent an important step forward from the previous molecular datasets that targeted higher ichneumonid relationships, all of which included only a single gene (28S rRNA) and were highly sensitive to alignment strategies (Belshaw et al., 1998; Quicke et al., 2009). Even the addition of morphological data in Quicke et al. (2009) left many parts of the ichneumonid tree inconsistent between different analytical approaches; this might be due to high levels of homoplasy found in these parasitoid wasps, which can partly be explained by convergent adaptations to the same host groups (Gauld & Mound, 1982).

Given the sparse taxon sampling in the transcriptome data, which omitted members of several pimpliform subfamilies, we focus our discussion here on the main analyses of the combined dataset, but mention results from the transcriptome trees where appropriate. Many previously suggested higher relationships, most of which were based on morphology, were confirmed here. This is the case for the placement of Xoridinae (Quicke et al., 1999), which is unequivocally recovered as sister to the remaining ichneumonids here for the first time (Quicke, 2015). The informal subfamily groupings, Ophioniformes, Ichneumoniformes and Pimpliformes, and the monophyly of most subfamilies were also recovered with good support. With respect to relationships among the pimpliform subfamilies, our results of the analyses of amino acid data (Figs 2, 3a–d) are similar to the ‘best guess’ phylogeny shown in Quicke (2015, fig. 13.1, which is a summary of the analyses in Quicke et al., 2009), at least in terms of the grouping of Poemeniinae, Rhyssinae and Pimplinae (although with the exception of Xanthopimpla; see later). However, this grouping was not found in the nucleotide analyses (Fig. 3c, d), where we recovered different (but partly weakly supported) relationships among the remaining subfamilies. The positions of Acaenitinae, Collyriinae and Diplazontinae are highly inconsistent among our analyses (e.g. compare Fig. 3a, e) and thus remain unclear, but the low support in this part of each tree precludes rejection of the hypothesis that they all branch off close to the base of a group containing the koinobiont, endoparasitoid pimpliform subfamilies (Quicke et al., 2009; Quicke, 2015). In contrast to previous suggestions (Wahl, 1990; Wahl & Gauld, 1998), Diplazontinae and Orthocentrinae were never recovered as sister groups; instead, Diplazontinae were repeatedly recovered as sister to the remaining pimpliform subfamilies (Figs 1, 3c, e, f), even though this placement was, again, not always recovered. Instead of grouping with Diplazontinae, the Orthocentrinae formed a clade with Diacritinae and Cylloceriinae in almost all analyses,
usually with good to even maximum support. A close relationship between Cyllocerinae and Orthocentrinae has long been suggested based on morphological and 28S evidence (Townes, 1971; Quicke et al., 2009), and some authors (e.g. Humala, 2007) maintain one subfamily for the Cyllocerinae and Orthocentrinae. Wahl (1990) defined a clade of koinobiont endoparasitoids of Diptera comprising Cyllocerinae, Diplazoninae and Orthocentrinae based mainly on a larval character, the distinctive fused hypostomal-stipital plate, not otherwise found in Ichneumonidae. Given our results and the lack of convincing apomorphies in adult morphology, it is possible that this plate has evolved more than once in ichneumonid larvae feeding within dipteran larvae.

Diacritinae and Orthocentrinae have also been grouped together previously, especially by Perkins (1940) and Townes (1957), who both included Diacritus Förster in their Plectiscinae (roughly equivalent to Orthocentrinae), even though Townes did so only provisionally and later classified Diacritinae as a separate tribe within an expanded Pimplinae that also included Rhyssinae and Poemeniinae (Townes, 1969). Morphologically, a close relationship between Orthocentrinae and Diacritinae is not without merit. Diacritus shares a fringe of dense setae along the inner side of the apex of the hind tibia with Orthocentrinae (even though it is somewhat less pronounced), and the first metasomal segment, with its elongate shape and fusion between sternite and tergite, is reminiscent of some orthocentrine genera such as Proclitus Förster, Dialipsis Förster and Symplecis Förster. Considering that both Orthocentrinae and Cyllocerinae are parasitoids of Diptera (File S6), this might provide a clue as to where to look for hosts of Diacritinae, for which no host data are available to date.

Wahl & Gauld (1998) suggested that Acaenitinae consists of a grade of genera previously classified in the tribe Coleocentrinia and the monophyletic former Acaenitini, but only referred to an unpublished, preliminary cladistics analysis to support this claim. It is a distinct possibility that the former grade is not only paraphyletic but indeed polyphyletic with respect to other pimpliform subfamilies (Quicke et al., 2009). Morphological similarities between Acaenitinae and other subfamilies have been pointed out in the past, especially with Collyrinae (Sheng et al., 2012), with which many acaenitine genera share the median tubercle on the clypeus, the propodeum lacking any transverse carinae, and the nervellus of the hindwing intercepted high up. In our analyses, the four genera of the former Acaenitini were always recovered together, but almost never sharing a unique common ancestor with the sole representative of the former Coleocentrinia included here, Coleocentrus (Figs 2, 3). In the analyses of the transcriptome sequence data, Coleocentrus always clustered with Collyria (Fig. 1; no other acaenitine genera were included), and this sister-group relationship also appeared in some of the analyses of the full dataset (e.g. Fig. 3d). However, Coleocentrus clustered consistently with various pimpliform groups, depending on data type and analysis, and support for a separation of the two acaenitine groups was weak in most analyses. Furthermore, Acaenitinae including Coleocentrus were recovered as monophyletic in some of the single-gene trees (File S9). Thus, we await an increased taxon sampling of the Coleocentrus group of genera (especially including specimens of Procinetus Förster) before suggesting any taxonomic changes.

The subfamily Pimplinae (excluding Xanthopimpla) was only recovered as monophyletic in the analyses of the amino acid alignment (Figs 2, 3a–d) and not those of nucleotide sequence data (Fig. 3e–f). The placement of Xanthopimpla might have been obscured by the long subtending branch of our exemplar species (Xanthopimpla varimaculata Cameron), an issue that could be remedied in future by a more extensive sampling of this species-rich genus. Xanthopimpla shares at least two unique characters with Lissopimpla Kriechbaum of Pimplini, namely the transversely divided clypeus and (with some exceptions in Xanthopimpla) a transverse carina at the anterior end of the notauli (unfortunately, no representatives of Lissopimpla were included in our study). In contrast with the findings of various authors (e.g. Townes, 1969), Echthromorpha Holm gren does not share the transversely divided clypeus, and its mandibles, although narrowed with the lower tooth very short, are not twisted as in Lissopimpla and Xanthopimpla. Xanthopimpla also seems to be among the oldest extant pimpline genera when considering the fossil record (Spasojevic et al., 2018b), which could support an early branching of this genus. Together with Lissopimpla, Xanthopimpla might have to be included in a new tribe or even subfamily in the future, but we await an analysis with more extensive taxon sampling before formalizing such a change. As for the relationships among the other Pimplini genera, it is interesting that the results recovered here are strongly reminiscent of the tribal classification suggested by Townes (1969): he combined today’s Delomeristini with the Theronia group of genera and Pseudorhyssa in the tribe Theroniini, but expressed uncertainty in both cases. Later cladistic analyses based on morphological characters moved Pseudorhyssa to Poemeniinae and the Theronia group into Pimplini (Eggleton, 1989; Gauld, 1991). Our analyses provide support for Townes’s classification, which was largely intuitive. The grouping of Pseudorhyssa in Delomeristini was recovered with maximum support in all analyses of the full dataset (Fig. 2) and was consistent even across single-gene trees (File S11); we therefore move this genus to the tribe Delomeristini (Pimplinae), with the result that the tribe Pseudorhyssini Wahl & Gauld becomes a junior synonym of Delomeristini Hellén (syn.n.). We note that the continuation of the epomia along the ventral edge of the pronotum in Pseudorhyssa is similar to that found in Delomerista Förster and Atractogaster Kriechbaum (the epomia is much shorter in Perithous, the other genus of Delomeristini) and dissimilar to the more pronounced bulging of the pronotum along the carina in Poemeniinae. Additionally, Pseudorhyssa has a weak, ventrally placed glymma on the first metasomal tergite in common with other Delomeristini and unlike Poemeniinae. The placement of the Theronia genus-group with Delomeristini was not recovered in a majority of gene trees, and the current placement within Pimplini also occurred in some gene trees. However, the group was consistently recovered as monophyletic here and is well defined on morphological grounds (Pham et al., 2013); thus, we resurrect Theronia as a tribe for the
Theronia genus-group (stat. rev.), as was advocated by Carlson (1979).

Our results clearly placed Hemiphanes in the Ichneumoniformes group of subfamilies, both in the concatenated analyses and in the vast majority of the gene trees, but, on the basis of our taxon sampling, they did not firmly suggest a subfamily placement; herein we transfer Hemiphanes to Cryptinae. Whilst the morphology of Hemiphanes does not present any unambiguous synapomorphies with Cryptinae, it is a better fit than Orthocentrinae, as is borne out by our molecular characters. In support of this placement include the fused first metasomal sternite and tergite, absence of a glymma, anteriorly narrow and posteriorly much wider first tergite, and strongly sclerotized remaining tergites. There is a distinct sternaulus on the anterior 0.3–0.4 of the mesopleuron, hindwing vein M+Cu is straight and the nervellus (distant abscissa of Cu) is strong. Furthermore, the forewing vein 2 m-cu has one bulla, the ovipositor has a nodus, and the male aedeagus is roughly ‘drop-shaped’ distally, not dorsoventrally flattened as in Pimpliformes. Features such as the twisted mandibles and fringe of setae on the inner aspect of the hind tibia do not strongly support the placement of Hemiphanes in Orthocentrinae, as the mandibles are more robust and the fringe less dense than in most orthocentrids. The sum of characters, particularly the wing venation and reduced propodeal carinae, suggest that Hemiphanes belongs in Cryptinae rather than Phygadeuontinae. In several respects, including the shape of the clypeus, large hypopygium, short first metasomal sternite, spiral of the first tergite anterior to the mid-length, wing venation and general structure, Hemiphanes is surprisingly similar to Sphecophaga Westwood (tribe Cryptini), a parasitoid of eusocial vespid wasps; the two key out to the same couplet in a recent key to ichneumonid subfamilies (Broad et al., 2018). However, the pronounced triangular expansions of the metanotum are a unique feature of the tribe Aptesini rather than Cryptini, so we place Hemiphanes in Aptesini, despite its overall similarity with Sphecophaga.

Rapid radiation of Pimpliformes

Rapid radiations have been suggested in several groups of parasitoid wasps and are mostly attributed to the formation of new ecological niches through host switches. A rapid radiation after host switching is possibly the case for the microgastrid subfamilies of the Braconidae, which changed to lepidopteran hosts and at the same time acquired poly-DNA viruses to overcome their hosts’ immune systems (Whitfield et al., 2018). However, previous studies inferring rapid radiations (e.g. Murphy et al., 2007; Murphy et al., 2008; Zaldivar-Riverón et al., 2008) were based on the analysis of small sets of genes, usually no more than three or four, which could explain the low resolution in the phylogenetic inferences. In Chalcidoidea, a recent transcriptome study based on more than 3200 genes suggested a rapid radiation of some of its families during the Late Cretaceous, although the inferred branch lengths of their dated tree still mostly span about 5–10 Ma and are thus not that rapid after all (Peters et al., 2018). Our study might therefore be viewed as the most convincing example to date of a rapid radiation in parasitoid wasps, supported by both considerable gene sampling and very short inferred branch lengths (Fig. 2). As a very rough estimate, assuming the most recent common ancestor of Pimpliformes existed about 66 Ma (see later) and using average branch lengths as clock lengths, the shortest branches at the base of the radiation would be less than 1 Ma long, but this result should be confirmed by a properly calibrated molecular clock analysis.

Our ancestral-state reconstructions inferred at least one switch between idiobiosis and koinobiosis during the most rapid phase of the radiation (Fig. 3). In addition, the switch to Diptera as hosts by Diplazontinae, Orthocentrinae and Cylloceriinae (and possibly Diacritinae) probably occurred in two separate events (but see the previous section where we discuss that we cannot exclude that they are indeed monophyletic, which would imply a single switch). Diptera are rather rare as hosts for Ichneumonidae, and the effect of this host switch on diversification rates and patterns deserves further study. The same is true for the switches between parasitoid strategies, including between ecto- and endoparasitoidism. We refrain here from a detailed discussion of the implications of our results, given the inconsistency between analyses (Fig. 3). A better-resolved tree is needed to address such questions.

The timing of the rapid radiation of the pimpliform subfamilies remains unclear as no dated tree for the group is yet available. However, the fossil record indicates that it might have taken place at around or just after the Cretaceous–Paleogene boundary. Fossil ichneumonids from the Cretaceous are all classified in now extinct subfamilies (Kopylov, 2009, 2010), with the exception of a species from Canadian amber that was tentatively placed in Labeninae (McKellar et al., 2013), a group consistently branching earlier than Pimpliformes in our analyses. On the other hand, several clear representatives of the subfamilies Acaenitinae, Orthocentrinae, Pimplinae and Rhyssinae are known from the late and middle Eocene (Spasojevic et al., 2018a; Spasojevic et al., 2018b). If a dated phylogenetic analysis confirms the timing of the pimpliform radiation as directly after the K–Pg boundary, this group’s history can be compared to other rapid radiations that have presumably taken place around the same time, such as that of Neoaves (Ericson et al., 2014; Ksepka & Phillips, 2015), placental mammals (dos Reis et al., 2012) and legumes (Koenen et al., 2013).

Do we need more genes, more taxa or better methods?

Rapid radiations that happened in the deep past constitute some of the most difficult phylogenetic problems. The low probability of any substitution happening along the short basal branches, combined with the risk that such true signal is masked by subsequent changes along the typically long branches leading from there, might mean that very large amounts of data need to be generated in order to observe sufficient signal for resolving deep, rapid radiations (Goldman, 1998; Fischer & Steel, 2009). Biases in the data arising from model misspecification,
nonrandom patterns of missing data, or among-lineage rate variation and associated effects such as long-branch attraction might easily obscure any true phylogenetic signal from such radiations (Collins et al., 2005; Rodríguez-Ezpeleta et al., 2007; Waegle & Mayer, 2007; Philippe et al., 2011; Doyle et al., 2015). Consequently, several studies targeting old radiations using genome-scale datasets have failed to completely resolve the underlying species tree (Misof et al., 2014; Prum et al., 2015; Peters et al., 2018), and others retrieved conflicting hypotheses depending on dataset or analytical methods (Meusemann et al., 2010; Jarvis et al., 2014; Sharma et al., 2014; Reddy et al., 2017).

Is it possible to distinguish between a lack of signal due to an extremely rapid radiation (or, in other words, a truly hard polytomy), insufficient gene or taxon sampling, and systematic biases in a dataset? Analyses of subsets of a dataset can, to a certain extent, exclude undersampling as a major disruptive factor. Our phylogenetic inferences based on gene sets of different sizes indicate a slowdown in the increase of mean clade support with increasing gene sampling and indications of a levelling-off at an average bootstrap per node of about 0.9, even though it is not entirely clear if a further increase might still be possible if more than 93 genes were sampled (Fig. 4).

It also remains unclear whether our approach to marker choice has in fact improved the inference of the phylogeny, given that none of the criteria clock-likeness, number of sites at near-optimal rates, or single-gene bootstrap supports clearly outperformed mere gene length in the analyses of subsets of genes (Fig. 4). However, it is possible that the initial choice of markers from the transcriptome gene set already eliminated most of the inferior genes and thus left us with insufficient power to detect any remaining effect of these predictors. The efficiency of some of the criteria we applied is also not beyond doubt. For instance, clade support values for certain clades might be high despite very low resolution in other parts of the tree, especially if there is a difference in node depth (Klopfstein et al., 2017). Using sums of clade support might therefore favour genes that resolve the more numerous, shallow nodes well, while not accounting sufficiently for the deeper nodes. Furthermore, we used a combined approach with the sequential application of several different criteria for choosing genes, and the relative performance and combinaibility of these criteria still need to be evaluated (but see Regier et al., 2008; Doyle et al., 2015; Klopfstein et al., 2017).

Tests for the amount and extent of taxon sampling that are necessary to resolve a deep radiation are even more difficult to perform, as the addition of taxa always changes the nature of the phylogenetic problem. In our case, the trees derived from the 723-gene transcriptome dataset of 26 taxa was fully resolved in terms of maximum bootstrap support in all analyses, while the combined hybrid capture data of 93 genes of 84 taxa exposed the weakly resolved rapid radiation at the base of Pimpliformes. The larger dataset included the additional subfamilies Cylocorini, Diacritinae and Orthocentrinae, and several additional tribes. Thus, it included additional nodes which were difficult to resolve. The most striking topological difference between the trees was the placement of Diplazontinae: breaking down the branch leading to Syrphophilus Dasch, the only diplazontine included in the transcriptome dataset, obviously had the effect of moving this subfamily further up the pimpliform tree (compare Figs 2 and 3). Consideration of phylogenetic informativeness led to the conclusion that targeted taxon sampling should focus on species that would branch off as closely to the unresolved nodes as possible (Geuten et al., 2007). In our case, a promising approach would be to sample multiple genera per subfamily, such as in Diacritinae and Collyriinae, but also focusing on more basal-branching members of other subfamilies such as Acaenitinae. While such an approach of course cannot increase the length of the most basal branches of the pimpliform radiation, it breaks down their long subtending branches and can therefore ameliorate possible long-branch attraction phenomena that could obscure the true phylogenetic signal.

Systematic biases such as nucleotide composition biases and other mismatches of the substitution model can be exposed by analytical methods (e.g. the test for uneven nucleotide composition performed here; Swofford, 2002), but also by more advanced methods falling into the class of model adequacy testing (Brown, 2014; Doyle et al., 2015; Duchêne et al., 2018). In most cases, such biases may be datatype-specific and will thus translate into conflicts between different dataset types, such as amino acids versus nucleotides or all three versus the first two codon positions. Such conflicts were especially apparent in our transcriptome sequence analyses, while the complete data typically exhibited low support at those nodes that were in conflict between analyses. The absence of supported conflict between data types might indicate that substitution-model misspecification was not the major factor preventing resolution of the pimpliform radiation. Also, in contrast to previous analyses that recovered strongly supported conflict among the single-gene trees (Rokas et al., 2003; Song et al., 2012), those here were mostly unresolved at the base of Pimpliformes, which indicates that gene-tree/species-tree analyses would not yield improved resolution (Edwards et al., 2007; Liu et al., 2008). Nevertheless, better evolutionary models might facilitate the extraction of what little signal is left in sequence data from the early and rapid pimpliform radiation. Another alternative is the analysis of different data types that show comparatively low levels of homoplasy, for instance certain ecological or morphological characters or characters pertaining to genome morphology (Nylander et al., 2004; Dowton et al., 2009; Henricson et al., 2010).

In conclusion, we postulate here one of the most rapid radiations uncovered to date in parasitoid wasps and speculate that the most promising approach to improve its resolution lies in increased taxon sampling, even though the sampling of additional markers or even other data types remains an alternative. The molecular resources and analysis pipeline we present here represent a significant step forward in phylogenetics of ichneumonoids and beyond and will in future hopefully allow a more robust inference of the relationships among these wasps and of the evolutionary history of their varied parasitoid ecologies.
Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

File S1. Collection of scripts used in the bioinformatics pipeline. The scripts are either in shell script or R script format or text format input files and include details of the program versions used. An overview is given commenting on the individual steps.

File S2. Taxon sampling including detailed collection data.

File S3. Alignment of 723 genes from the transcriptome analyses in PHYLIP format, followed by gene and codon position partitioning.

File S4. Baits used for the hybrid capture in FASTA format.

File S5. Log detailing all manual changes performed during the quality control steps of the hybrid capture data.

File S6. Coverage graphs of hybrid capture taxa. Each page shows one gene, including seven genes not included in the capture. Vertical lines indicate intron positions.

File S7. Alignment of 93 genes resulting from the hybrid capture and transcriptome analyses in PHYLIP format, followed by gene and codon position partitioning.

File S8. Parasitoid strategies of included taxa. Table detailing the parasitoid strategies as assumed in our analyses, including numbers of supporting references for the genus and for the subfamily and notes explaining our interpretations.

File S9. Statistics transcriptomes and hybrid capture. Multi-sheet Excel file containing details on the coverage per gene and per taxon both in the transcriptome and in the hybrid capture part of our study.

File S10. Success of the newly designed primer pairs for eight genes. Symbols in the table stand for the proportion of specimens successfully sequenced: +, more than half; +/−, at least one but less than half; −, none. Number of taxa (genera) indicates the total number of specimens used when testing whether or not more than two genes are successfully amplified (filters out possible low-quality extractions); usually this is also the number of specimens used for testing each primer pair, but for second primer pairs (F2R2), fewer specimens were used and mostly those from which no amplicons were obtained with the first primer pairs.

File S11. Single-gene trees. Each page shows the maximum likelihood tree resulting from a RAXML analysis of all three codon positions. Trees were rooted with Nasonia, and bootstrap support values are given next to nodes. Taxon names are coloured according to subfamilies or in dark red for Hemiphanes and Pseudorhyssa.

File S12. Spearman correlations between gene ranking criteria. Rho values are given above the diagonal, and P-values below.

Acknowledgements

We are very grateful to everyone who provided rare specimens for this study: Eric Chapman (University of Kentucky, U.S.A.), the Swedish Malaise Trap Project (www.stationlinne.se, Sweden), Masato Ito (Kobe University, Japan), Ilari Saikkola (University of Turku, Finland), Stanislav Korenko (Czech University of Life Sciences Prague), Philippe Loupope (Université de Bourgogne, France) and Marc Neumann (Walterswil SO, Switzerland). Muriel Fragnière (University of Bern, Switzerland) helped with the planning of the transcriptome part of this study, prepared the RNA libraries and conducted the transcriptome sequencing. Gary Gibson (Canadian National Collection of Insects, Ottawa, Canada) identified the Australian pteromalid outgroup. We also thank Kathy Saint and Tessa Bradford (University of Adelaide, Australia) for helpful advice on the hybrid-enrichment laboratory work. Hugh Robertson (University of Illinois, U.S.A.) provided us with access to the genome of Microplitis demolitor before its release for assessing intron–exon boundaries for bait design. Shaun Winterton (California Department of Food and Agriculture, Sacramento, U.S.A.), Andrew Bennett (Canadian National Collection of Insects, Ottawa, Canada) and an anonymous reviewer provided very valuable feedback on a previous version of our manuscript. This study was funded by the Swiss National Science Foundation (grant PZ00P3_154791 to SK). The authors declare no conflict of interest.

References

Ababneh, F., Jerumi, L., Ma, C. & Ronbinson, J. (2006) Matched-pairs tests of homogeneity with applications to homologous nucleotide sequences. Bioinformatics, 22, 1225–1231.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) Basic local alignment search tool. Journal of Molecular Biology, 215, 403–410.

Andrews, S. (2014) FastQC Version 0.11.2. Babraham Institute, Cambridge.

Askew, R.R. & Shaw, M.R. (1986) Parasitoid communities: their size, structure and development. Insect Parasitoids (ed. by J.K. Waage and D. Greathead), pp. 225–264, Academic Press, London.

Austin, A.D. (1985) The function of spider egg sacs in relation to parasitoids and predators, with special reference to the Australian fauna. Journal of Natural History, 19, 359–376.

Bank, S., Sann, M., Mayer, C. et al. (2017) Transcriptome and target DNA enrichment sequence data provide new insights into the phylogeny of vespid wasps (Hymenoptera: Aculeata: Vespidae). Molecular Phylogenetics and Evolution, 116, 213–226.

Belshaw, R., Fitton, M.G., Herniou, E., Gimeno, C. & Quicke, D.L.J. (1998) A phylogenetic reconstruction of the Ichneumonoidea (Hymenoptera) based on the D2 variable region of 28S rivosomal RNA. Systematic Entomology, 23, 109–123.

Bergsten, J., Bilton, D.T., Fujisawa, T. et al. (2012) The effect of geographical scale of sampling on DNA barcoding. Systematic Biology, 61, 851–869.
Rapid radiation of pimpliform parasitoid wasps

Louca, S. & Doebeli, M. (2018) Efficient comparative phylogenetics on large trees. Bioinformatics, 34, 1053–1055.

Mamanova, L., Coffey, A.J., Scott, C.E. et al. (2010) Target-enrichment strategies for next-generation sequencing. Nature Methods, 7, 111–118.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.Journal, 17, 10–12.

Mayer, C., Sann, M., Donath, A. et al. (2016) BaitFisher: a software package off multispecies target DNA enrichment probe design. Molecular Biology and Evolution, 33, 1875–1886.

McCormack, J.E., Harvey, M.G., Faircloth, B.C., Crawford, N.G., Glenn, T.C. & Brumfield, R.T. (2013a) A phylogeny of birds based on over 1,500 loci collected by target enrichment and high-throughput sequencing. PLoS One, 8, e54858.

McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C. & Brumfield, R.T. (2013b) Applications of next-generation sequencing to phyleogeography and phylogenetics. Molecular Phylogenetics and Evolution, 66, 526–538.

McKellar, R.C., Kopylov, D.S. & Engel, M.S. (2013) Ichneumonidae (Insecta: Hymenoptera) in Canadian Late Cretaceous amber. Fossil Record, 16, 217–227.

Meusemann, K., von Reumont, B.M., Simon, S. et al. (2010) A phylogenomic approach to resolve the arthropod tree of life. Molecular Biology and Evolution, 27, 2451–2464.

Meyer, M. & Kircher, M. (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harbor Protocols, 2010, 1–20.

Misof, B., Liu, S.L., Meusemann, K. et al. (2014) Phylogenomics resolves the timing and pattern of insect evolution. Science, 346, 763–767.

Murphy, N.P., Carey, D., Castro, L.R., Dowton, M. & Austin, A.D. (2007) Phylogeny of the platygastroid wasps (Hymenoptera) based on sequences from the 18S rRNA, 28S rRNA and cytochrome oxidase I genes: implications for the evolution of the ovipositor system and host relationships. Biological Journal of the Linnean Society, 91, 653–669.

Murphy, N.P., Banks, J.C., Whitfield, J.B. & Austin, A.D. (2008) Phylogeny of the parasitic microgastroid subfamilies (Hymenoptera: Braconidae) based on sequence data from seven genes, with an improved time estimate of the origin of the lineage. Molecular Phylogenetics and Evolution, 47, 378–395.

Narendran, T.C. (2001) Parasitic Hymenoptera and biological control. Biocontrol Potential and its Exploitation in Sustainable Agriculture (ed. by R.K. Upadhyay, K.G. Mukerji and B.P. Chamola). Kluwer Academic / Plenum Publishers, New York.

Nicholls, J.A., Challis, R.J., Mutun, S. & Stone, G.N. (2012) Mitochondrial barcodes are diagnostic of shared refugia but not species in hybridizing oak gallwasp. Molecular Ecology, 21, 4051–4062.

Nylander, J.A.A., Ronquist, F., Huelsenbeck, J.P. & Nieves-Aldrey, J.I. (2004) Bayesian phylogenetic analysis of combined data. Systematic Biology, 53, 47–67.

Olsen G.L. (1996) DNArates - maximum likelihood site rate calculation. Published by the author.

Paradis, E., Claude, J. & Strimmer, K. (2004) APE: analyses of phylogenetics and evolution in R language. Bioinformatics, 20, 289–290.

Percy, D.M., Crampton-Platt, A., Sveinsson, S., Lemmon, A.R., Lemmon, E.M., Ouvrard, D. & Burckhardt, D. (2018) Resolving the psyllid tree of life: phylogenomic analyses of the superfamily Psylloidea (Hemiptera). Systematic Entomology, 43, 762–776.

Perkins, J.F. (1940) Notes on the synonymy of some genera of European Pimplinae (s.l.) (Hym. Ichneumonidae). The Entomologist, 73, 54–56.
Peters, R.S., Krogmann, L., Mayer, C. et al. (2017) Evolutionary history of the Hymenoptera. Current Biology, 27, 1013–1018.

Peters, R.S., Niehuis, O., Gunkel, S. et al. (2018) Transcriptome sequence-based phylogeny of chalcidoid wasps (Hymenoptera: Chalcidioidea) reveals a history of rapid radiations, convergence, and evolutionary success. Molecular Phylogenetics and Evolution, 120, 286–296.

Petersen, M., Meusemann, K., Donath, A. et al. (2017) Orthograph: a versatile tool for mapping coding nucleotide sequences to clusters of orthologous genes. BMC Bioinformatics, 18, 111.

Pham, N.T., Broad, G.R. & Waagele, J.W. (2013) A review of the Thermonia genus-group (Hymenoptera: Ichneumonidae: Pimplinae) from Vietnam with descriptions of five new species. Journal of Natural History, 47, 1501–1538.

Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Wörheide, G. & Baurain, D. (2011) Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biology, 9, e1000602.

Phillips, M.J., Deluc, F. & Penny, D. (2004) Genome-scale phylogeny and the detection of systematic biases. Molecular Biology and Evolution, 21, 1455–1458.

Prum, R.O., Berv, J.S., Dornburg, A., Field, D.J., Townsend, J.P., Moriarty Lemon, E. & Lemon, A.R. (2015) A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature, 526, 569–573.

Quicke, D.L.J. (2015) Phylogeny and systematics of the Ichneumonidae. The Braconid and Ichneumon Parasitoid Wasps: Biology, Systematics, Evolution and Ecology (ed. by Q. DLI). John Wiley & Sons, Ltd, Chichester.

Quicke, D.L.J., Lopez-Vaamonde, C. & Belshaw, R. (1999) The basal Ichneumonidae (Insecta, Hymenoptera): 28S D2 rDNA considerations of the Brachycrytinae, Labeninae, Paxylommatinae and Xoridiinae. Zoologica Scripta, 28, 203–210.

Quicke, D.L.J., Fitton, M.G., Notton, D.G., Belshaw, R., Broad, G.R. & Dolphin, K. (2000) Phylogeny of the Ichneumonidae (Hymenoptera). A simultaneous molecular and morphological analysis. Hymenoptera: Evolution, Biodiversity and Biological Control (ed. by A.D. Austin and M. Dowton), pp. 74–83. CSIRO, Canberra.

Quicke, D.L.J., Lauret, N.M., Fitton, M.G. & Broad, G.R. (2009) A thousand and one wasps: a 28S DNA and morphological phylogeny of the Ichneumonidae (Insecta: Hymenoptera) with an investigation into alignment parameter space and elision. Journal of Natural History, 43, 1305–1421.

R Core Team (2014) R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna.

Rago, A., Gilbert, D.G., Choi, J.H. et al. (2016) OGS2: genome re-annotation of the jewel wasp Nasonia vitripennis. BMC Genomics, 17, 678.

Reddy, S., Kimball R.T., Pandey A., Hosner P.A., Braun M.J., Hackett S.J., Han K.-L., Harshman J., Heddleston C.J., Kingston S., Marks B.D., Miglia K.J.M., W.S., Sheldon F.H., Witt C.C., Yuri T., Braun E.L. (2017) Why do phylogenomic data sets yield conflicting trees? Data type influences the avian tree of life more than taxon sampling. Systematic Biology, 66, 857–879.

Regier, J.C., Shultz, J.W., Ganley, A.R.D. et al. (2008) Resolving arthropod phylogeny: exploring phylogenetic signal within 41 kb of protein-coding nuclear gene sequence. Systematic Biology, 57, 920–938.

Rodríguez-Ezpeleta, N., Brinkmann, H., Roure, B., Dartillot, N., Lang, B.F. & Philippe, H. (2007) Detecting and overcoming systematic errors in genome-scale phylogenies. Systematic Biology, 56, 389–399.

Rokas, A., Williams, B.L., King, N. & Carroll, S.B. (2003) Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature, 425, 798–804.

Ronquist, F., Klop夫stein, S., Vilhelmsen, L., Schulmeister, S., Murray, D.L. & Rasmussen, A.P. (2012a) A total-evidence approach to dating with fossils. Applied to the early radiation of the Hymenoptera. Systematic Biology, 61, 973–999.

Ronquist, F., Teslenko, M., Van der Mark, P. et al. (2012b) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology, 61, 539–542.

Rousse, P., Quicke, D.L.J., Matthee, C.A., Lefèvre, P. & Van Noort, S. (2016) A molecular and morphological reassessment of the phylogeny of the subfamily Ophioninae (Hymenoptera: Ichneumonidae). Zoological Journal of the Linnean Society, 178, 128–148.

Salichos, L. & Rokas, A. (2013) Inferring ancient divergences requires genes with strong phylogenetic signals. Nature, 497, 327–331.

Sann, M., Niehuis, O., Peters, R.S. et al. (2018) Phylogenomic analysis of Apoidea sheds new light on the sister group of bees. BMC Evolutionary Biology, 18, 71.

Santos, B. (2017) Phylogeny and reclassification of Cryptini (Hymenoptera, Ichneumonidae, Cryptinae), with implications for ichneumonid higher-level classification. Systematic Entomology, 42, 650–676.

Sharma, P.P., Kaluziak, S.T., Pérez-Porro, A.R., Gonzáles, V.L., Hormiga, G., Wheeler, W.C. & Giribet, G. (2014) Phylogenetic interrogation of Arachnida reveals systemic conflicts in phylogenetic signal. Molecular Biology and Evolution, 31, 2963–2984.

Sheng, M.-L., Broad, G.R. & Sun, S.-P. (2012) A new genus and species of Collyriinae (Hymenoptera, Ichneumonidae). Journal of Hymenoptera Research, 23, 103–125.

Slater, G.S. & Birney, E. (2005) Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics, 6, 31.

Song, S., Liu L., Edwards S.V., Wu S. (2012) Resolving conflict in the eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. Proceedings of the National Academy of Sciences of the United States of America, 109, 14942–14947.

Spasojevic, T., Broad, G.R., Bennett, A.M.R. & Klop夫stein, S. (2018a) Ichneumonid parasitoid wasps from the Early Eocene Green River Formation: five new species and a revision of the known fauna (Hymenoptera, Ichneumonidae). Paläontologische Zeitschrift, 92, 35–63.

Spasojevic, T., Wedmann, S. & Klop夫stein, S. (2018b) Seven remarkable new fossil species of parasitoid wasps (Hymenoptera, Ichneumonidae) from the Eocene Messel Pit. *PLoS One*, 13, e0197477.

Stamatakis A. (2006) Phylogenetic models of rate heterogeneity: A high performance computing perspective. Proceedings of IPDPS2006, Rhodes, Greece.

Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, 30, 1312–1313.

Streicher, J.W., Schulte, J.A.I. & Wiens, J.J. (2016) How should genes and taxon be sampled for phylogenomic analyses with missing data? An empirical study in iguanian lizards. Systematic Biology, 65, 128–145.

Strimmer, K. & von Haeseler, A. (1997) Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. Proceedings of the National Academy of Sciences of the United States of America, 94, 6815–6819.

Suyama, M., Torres, D. & Bork, P. (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Research, 34, W609–W612.

Swoford, D.L. (2002) PAUP*: Phylogenetic Analysis using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, MA.

© 2018 The Royal Entomological Society, Systematic Entomology, 44, 361–383
Rapid radiation of pimpliform parasitoid wasps

Townes, H.K. (1957) A revision of the genera of Poemeniini and Xoridini (Hymenoptera, Ichneumonidae). *Proceedings of the Entomological Society of Washington*, 59, 15–23.

Townes, H.K. (1969) *The Genera of Ichneumonidae, Part 1*, Vol. 11, pp. 1–300. *Memoirs of the American Entomological Institute*. AEI, Gainesville, Florida.

Townes, H.K. (1971) *The Genera of Ichneumonidae, Part 4*, Vol. 17, pp. 1–372. *Memoirs of the American Entomological Institute*. AEI, Gainesville, Florida.

Townsend, J.P. (2007) Profiling phylogenetic informativeness. *Systematic Biology*, 56, 222–231.

Townsend, J.P., Su, Z. & Tekle, Y.I. (2012) Phylogenetic signal and noise: predicting the power of a dataset to resolve phylogeny. *Systematic Biology*, 61, 835–849.

Virgilio, M., Backeljau, T., Nevada, B. & De Meyer, M. (2010) Comparative performances of DNA barcoding across insect orders. *BMC Bioinformatics*, 11, 206.

Waagele, J.W. & Mayer, C. (2007) Visualizing differences in phylogenetic information content of alignments and distinction of three classes of long-branch effects. *BMC Evolutionary Biology*, 7, 147.

Wahl, D.B. (1986) Larval structures of oxytorines and their significance for the higher classification of some Ichneumonidae (Hymenoptera). *Systematic Entomology*, 11, 117–127.

Wahl, D.B. (1990) A review of the mature larvae of Diplazontinae, with notes on larvae of Acaenitinae and Orthocentrinae and proposal of two new subfamilies (Insecta: Hymenoptera, Ichneumonidae). *Journal of Natural History*, 24, 27–52.

Wahl, D.B. & Gauld, I.D. (1998) The cladistics and higher classification of the Pimpliformes (Hymenoptera: Ichneumonidae). *Systematic Entomology*, 23, 265–298.

Waterhouse, R.M., Tegenfeldt, F., Li, J., Zdobnov, E.M. & Kriventseva, E.V. (2013) OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs. *Nucleic Acids Research*, 41, D358–D365.

Werren, J.H., Richards, S., Desjardins, C.A., Niehuis, O., Gadau, J., Colbourne, J.K. & Group, T.N.G.W. (2010) Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. *Science*, 327, 334–348.

Whitfield, J.B., Austin, A.D. & Fernandez-Triana, J. (2018) Systematics, biology, and evolution of microgastrine parasitoid wasps. *Annual Review of Entomology*, 63, 389–406.

Yang, Z. (1996) Among-site rate variation and its impact on phylogenetic analyses. *Trends in Ecology & Evolution*, 11, 367–372.

Yang, Z. (1998) On the best evolutionary rate for phylogenetic analysis. *Systematic Biology*, 47, 125–133.

Yu D.S., Van Achterberg C., Horstmann K. (2016) Taxapad 2016. Ichneumonoidea 2015 (Biological and taxonomical information), Taxapad Interactive Catalogue Database on flash-drive. Nepean [WWW document]. URL www.taxapad.com.

Zaldivar-Riverón, A., Shaw, S.R. & Quicke, D.L.J. (2008) Evolution of the parasitic wasp subfamily Rogadinae (Braconidae): phylogeny and evolution of lepidopteran host ranges and mummy characteristics. *BMC Evolutionary Biology*, 8, 329.

Zerbino, D.R. & Birney, E. (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, 18, 821–829.

Accepted 17 August 2018
First published online 31 October 2018

© 2018 The Royal Entomological Society, *Systematic Entomology*, 44, 361–383