

# Identification of sawflies and hornails (Hymenoptera, ‘Symphyta’) through DNA barcodes: successes and caveats

STEFAN SCHMIDT,\* ANDREAS TAEGER,† JÉRÔME MORINIÈRE,\* ANDREW LISTON,†  
STEPHAN M. BLANK,† KATJA KRAMP,† MANFRED KRAUS,‡ OLGA SCHMIDT,\* ERIK HEIBO,§  
MARKO PROUS,†¶ TOMMI NYMAN,\*\* TOBIAS MALM†† and JULIE STAHLHUT‡‡

\*SNSB-Zoologische Staatssammlung München, Münchhausenstr. 21, 81247 Munich, Germany, †Senckenberg Deutsches Entomologisches Institut, Eberswalder Str. 90, 15374 Müncheberg, Germany, ‡Manfred Kraus, Fallrohrstr. 27, 90480 Nürnberg, Germany, §Erik Heibo, Rypeveien 34 A, 3420 Lierskogen, Norway, ¶Department of Zoology, Institute of Ecology and Earth Sciences, University of Tartu, Vanemuise 46, 51014 Tartu, Estonia, \*\*Department of Environmental and Biological Sciences, University of Eastern Finland, PO Box 111, Joensuu FI-80101 Finland, ††Department of Zoology, Swedish Museum of Natural History, P.O. Box 50007, SE-104 05 Stockholm, Sweden, ‡‡Biodiversity Institute of Ontario, University of Guelph, 50 Stone Rd E, Guelph, ON N1G 2W1, Canada

## Abstract

The ‘Symphyta’ is a paraphyletic assemblage at the base of the order Hymenoptera, comprising 14 families and about 8750 species. All have phytophagous larvae, except for the Orussidae, which are parasitoids. This study presents and evaluates the results of DNA barcoding of approximately 5360 specimens of ‘Symphyta’, mainly adults, and 4362 sequences covering 1037 species were deemed of suitable quality for inclusion in the analysis. All extant families are represented, except for the Anaxyelidae. The majority of species and specimens are from Europe, but approximately 38% of the species and 13% of the specimens are of non-European origin. The utility of barcoding for species identification and taxonomy of ‘Symphyta’ is discussed on the basis of examples from each of the included families. A significant level of cryptic species diversity was apparent in many groups. Other attractive applications include the identification of immature stages without the need to rear them, community analyses based on metabarcoding of bulk samples and association of the sexes of adults.

**Keywords:** DNA barcoding, hornails, Insects, morphology, sawflies, taxonomy

Received 6 June 2016; revision received 8 September 2016; accepted 16 September 2016

## Introduction

The Hymenoptera is a megadiverse order of insects with more than 155 000 described species (Aguiar *et al.* 2013). The ‘Symphyta’, commonly known as sawflies and hornails, comprises about 8750 species in 818 genera (of these, respectively, 8335 and 631 are extant), or ca. 6.5% of the Hymenoptera species [data based on SDEI (Senckenberg Deutsches Entomologisches Institut) database, accessed December 2015, and Taeger *et al.* 2010]. The group, which in the extant fauna comprises 14 families, is a paraphyletic assemblage basal to the Apocrita (bees, stinging wasps and ants; Sharkey *et al.* 2012). Phytophagy of larvae is the common denominator of all symphytan species except the Orussidae, which are parasitoids of wood-boring insects. Many sawfly species are of economic importance because their larvae can cause

considerable damage to agricultural crops, forests and ornamental plants (Pschorn-Walcher 1982; Hill 1983, 1987). Sawflies are often regarded as a neglected group of insects, and many taxa are difficult to identify due to morphological uniformity or because their taxonomy has not been sufficiently investigated. The fact that identification keys for many groups are outdated provides a further incentive for applying DNA barcodes for species identification and for revealing cryptic species diversity.

The present DNA barcode release represents the first step towards creating a comprehensive global DNA barcode library for the suborder ‘Symphyta’. The initial phase of this work was part of the ‘Barcoding Fauna Bavarica’ project of the SNSB-Zoologische Staatssammlung München (ZSM), Germany, which commenced in 2009 (Hendrich *et al.* 2010; Hausmann *et al.* 2012, 2013a, b). The project aims to assemble DNA barcodes for all Bavarian animal species. In 2010, the ‘Barcodes of Symphyta’ project started at the SDEI, aiming to assemble

Correspondence: Stefan Schmidt, Fax: +49 89 8107300;  
E-mail: stefan.schmidt@zsm.mwn.de

DNA barcodes of 'Symphyta' species with a global scope. The 'German Barcode of Life' (GBOL) project provided additional sequences after it commenced in 2012. All projects operated in close cooperation with the Biodiversity Institute of Ontario within the framework of the International Barcode of Life (iBOL) project. All sequences and associated specimen data are available through the Barcode of Life Database (BOLD, [www.boldsystems.org](http://www.boldsystems.org)).

Despite a strong emphasis on European species of 'Symphyta', specimens from elsewhere were used whenever material suitable for DNA barcoding was available. The sequences of non-European material, accounting for about 38% of the species, are included in the present release because their open access status is potentially useful for identification purposes and can be an incentive for further DNA barcoding efforts on all continents.

## Materials and methods

### Sampling

Specimens used for barcoding were primarily taken from the collections of the ZSM and the SDEI. Between 2009 and 2014, about 7850 specimens of about 1330 species of 'Symphyta' were processed, including about 290 species awaiting formal description. Species definitions were based on traditional morphological characters. This barcode collection represents about one quarter of the 'Symphyta' specimens and half of the 'Symphyta' species present in BOLD at the time of this release (24.5% and 49.0%, respectively, January 2016).

Typically, a single leg was removed and isolated from each adult specimen, after which the leg samples were sent to the Canadian Centre for DNA Barcoding (CCDB) in Guelph, Canada, for DNA extraction and barcode sequencing. In a few cases, parts of larvae were used for analysis. Specimens were identified using the most recent literature for each taxon and, if required, by comparing specimens with doubtful identifications with correctly identified individuals from the collections of the SDEI and the ZSM. A complete list of voucher specimens included in the current release is given in Appendix S1 (Supporting information).

The nomenclature and classification used in the present release are based on the World Catalogue of 'Symphyta' (Taeger *et al.* 2010). Since publication of this catalogue, extensive phylogenetic analyses have led to the proposal of several changes in the classification of 'Symphyta', including the elevation of the Heptamelini (Tenthredinidae) to family level, and the Athaliini to subfamily level (Malm & Nyman 2015; see also Schulmeister 2003a who suggested a raise to family level). Additional changes affected the Nematinae, with several genera

now being subsumed under *Euura* (Prous *et al.* 2014), and the Pergidae, with the Euryinae now included in the Perreyiinae, and the Phylacteophaginae placed within the Acordulecerinae (Schmidt & Walter 2014). These changes have not been adopted in the present study because they have yet to be implemented in BOLD. A comparison between the classification currently used in BOLD and the most recent including all major modifications is given in Appendix S2 (Supporting information). Scientific names throughout the text are cited without author and year (except for species that are mentioned but not present in the data set). A list of all species included in this release and the currently accepted name for each species with authorship information is given in Appendix S1 (Supporting information).

### DNA sequencing

DNA extraction, PCR amplification and sequencing were conducted at the CCDB using standardized high-throughput protocols (Ivanova *et al.* 2006; deWaard *et al.* 2008), available online under [www.ccdb.ca/resources.php](http://www.ccdb.ca/resources.php). The sequenced fragment starts from the 5' end of the mitochondrial cytochrome oxidase *c* (COI) gene and includes the standard 658-bp barcode region of the animal kingdom (Hebert *et al.* 2003). The DNA extracts are stored at the CCDB, but aliquots of ZSM vouchers will eventually be deposited in the ZSM DNA-Bank facility that is part of the DNA-Bank Network (see [www.dna-bank-network.org](http://www.dna-bank-network.org)). Similarly, aliquots of SDEI vouchers are deposited in the DNA storage facility of the SDEI. Successfully sequenced specimens are listed in Appendix S1 (Supporting information), together with sequence lengths and numbers of unresolved bases. Specimen data are accessible in BOLD and include collecting locality, geographic coordinates, elevation, collector, one or more digital images, identifier and voucher depository. All specimen data are accessible on BOLD through the following doi: 10.5883/DS-RSYM. Sequences obtainable through BOLD include a detailed Laboratory Information Management System (LIMS) report, primer information and trace files. These data are also available through GenBank (for Accession nos, see Appendix S1, Supporting information).

### Data Analysis

Distances among barcode sequences were calculated and a neighbour-joining tree built using the Kimura two-parameter model. Barcode Index Numbers (BINs) were assigned by the BOLD system, representing globally unique identifiers for clusters of sequences that correspond closely to biological species (Ratnasingham & Hebert 2013). BINs can be regarded as molecular

operational taxonomic units (MOTUs), and they provide an interim taxonomic system, that is a way to signify genetic units prior to detailed taxonomic studies including morphology. Specimens not assigned to any BIN were excluded from the calculations. For BIN assignment in BOLD, a minimum sequence length of 500 bp is required. Sequences between 300 and 500 bp can join an existing BIN, but will not create or split BINs. Sequences were aligned using the BOLD Aligner (amino acid-based hidden Markov models). Genetic distances and summary indices were calculated using analytical tools in BOLD and are given as mean and maximum pairwise distances for intraspecific variation, and as minimum pairwise distances for interspecific variation. The analyses are based on sequences with a minimum length of 500 bp and <1% ambiguous bases.

## Results and discussion

About 7500 specimens of 'Symphyta' were processed, of which 5360 specimens yielded sequences, representing 1126 species or subspecies. Country of origin of most specimens is Germany (2388), followed by France (420), Switzerland (306), Greece (281), Austria (250), China (236), Italy (198) and 57 other countries (Appendix S1, Supporting information). The number of specimens and species with sequences is given in Table 1 for each family of 'Symphyta'. The age of specimens ranged from 1 to 34 years, with about two-thirds (63%) being between 1 and 5 years old (Fig. 1).

Eighty-one per cent (4362) of the voucher specimens yielded sequences meeting the requirements for inclusion in further analyses, that is a minimum length of 500 bp and less than 1% ambiguous bases. The excluded sequences ranged from 400 to 499 bp (879 sequences), 300 to 399 bp (44), 200 to 299 (70) and <200 bp (six sequences). The 4362 sequences used for the analysis cover 1037 species from 170 genera and 13 families of 'Symphyta' (only the Anaxyelidae with a single extant species is not represented in the data set, though it is present in BOLD through GenBank).

The highest pairwise sequence divergence within a species, 14.72%, was observed in the Xyelidae, followed by 13.24% in the tenthredinid subfamily Allantinae (Table 2). The mean within-species distance across all family-level taxa ranged from 0.04 to 1.66% (Table 2). About 15% of the species exhibit BIN divergence, with conspecific individuals placed in two (102 species), three (37 species), four (12 species) or even five BINs (nine species, Appendix S3, Supporting information). Within-species divergence in excess of 2% is frequent in the Argidae and Pamphiliidae (11 species each), and in the tenthredinid subfamilies Allantinae (25 species), Nematiniae (63 species), Selandriinae (25 species) and

**Table 1** Number of world species of 'Symphyta' (extant species, without subspecies and unplaced taxa), and specimens and species (including subspecies) in the present release with sequences (including sequences <500 bp)

Family	Number of world species	Release	
		Specimens	Species
Argidae	918	209	49
Blasticotomidae	12	9	1
Cephidae	171	94	29
Cimbicidae	184	103	32
Diprionidae	145	51	12
Megalodontesidae	42	97	22
Orussidae	88	5	2
Pamphiliidae	297	135	43
Pergidae	451	68	47
Siricidae	150	14	5
Tenthredinidae	5686	4467	853
Allantinae	[754]	[538]	[95]
Blennocampinae	[647]	[318]	[53]
Heterarthrinae	[247]	[126]	[39]
Nematinae	[1251]	[1161]	[295]
Selandriinae	[1012]	[605]	[76]
Tenthredininae	[1775]	[1719]	[295]
Xiphydriidae	120	10	4
Xyelidae	72	98	27
<b>Total</b>	<b>8336</b>	<b>5360</b>	<b>1126</b>

Numbers in brackets are totals for the subfamilies of Tenthredinidae are hence not included in the totals of the bottom row.

Tenthredininae (69 species; Table 2). On the other hand, comparatively low interspecific distance (<2%) is found in many taxa and is particularly common in the Selandriinae (37 species) and Nematiniae (105 species), and also in the Xyelidae (12 species, Table 2). The maximum within-family divergence varied considerably across families and ranged from 9.5% in the Xiphydriidae to 35.7% in the Siricidae, with intermediate values in the Megalodontesidae (17%), Pamphiliidae (23.2%), Diprionidae (23.4%), Cimbicidae (26.0%), Cephidae (28.9%), Argidae (29.8%), Xyelidae (29.2%), Pergidae (31.9%) and Tenthredinidae (32.9%).

In the following discussion, families are arranged following the phylogenies proposed by Sharkey *et al.* (2012) for the 'Symphyta' and Schulmeister (2003a) for the Tenthredinoidea. Subfamilies of the Tenthredinidae are dealt with in alphabetical order.

### Xyelidae

The Xyelidae comprises 72 extant species worldwide (Taeger *et al.* 2010; Blank *et al.* 2013). This group has received attention because of its position as sister to the rest of the Hymenoptera (Sharkey *et al.* 2012) and its old age (Ronquist *et al.* 2012). Larval hosts are either

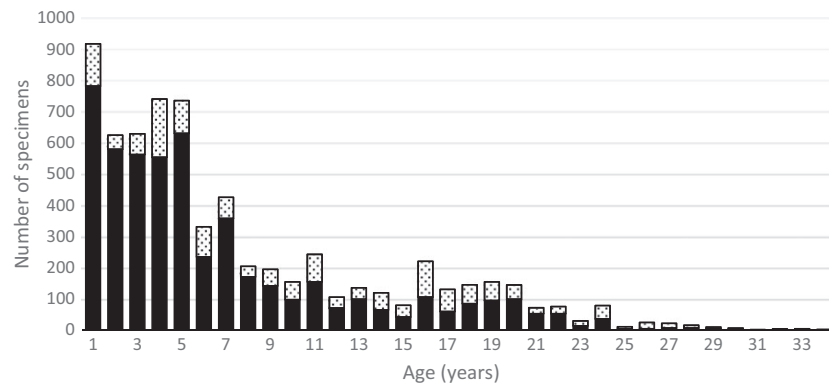


Fig. 1 Specimen age (in years since collection) for all sequences (dotted areas indicate specimens that did not produce any sequences).

Table 2 Distance summary statistics for families (and subfamilies for Tenthredinidae) of 'Symphyta' based on sequences  $\geq 500$  bp

Taxon	Sequences analysed	Species with high (>2%) intraspecific distance	Species with low (<2%) distance to another species	Minimum between-species distance (%)	Maximum within-species distance (%)	Mean within-species distance (%)
Argidae	203	11	13	0	8.87	0.94
Blasticotomidae	9	0	0	N/a	0.16	0.04
Cephidae	79	3	4	0	7.87	1.17
Cimbicidae	91	1	4	0.32	3.59	0.43
Diprionidae	35	2	2	0.81	6.09	1.41
Megalodontesidae	92	4	5	0.46	8.78	0.87
Orussidae	4	0	0	19.87	0.15	0.10
Pamphiliidae	122	11	11	0	7.09	1.53
Pergidae	68	0	0	4.23	4.90	0.64
Siricidae	12	0	0	18.6	2.10	0.81
Tenthredinidae:						
Allantinae	439	25	28	0	13.24	1.50
Blennocampinae	256	9	11	0	7.74	0.84
Heterarthrinae	98	1	8	0.15	7.57	0.47
Nematinae	902	63	105	0	11.28	1.08
Selandriinae	504	25	37	0	8.73	0.69
Tenthredininae	1350	69	96	0	10.45	0.87
Xiphydriidae	9	0	0	5.49	0.96	0.65
Xyelidae	89	11	12	0	14.72	1.66
Total	4362	235	338			0.93

coniferous trees (Xyelinae, comprising the majority of species worldwide, and all European ones) or angiosperm trees (Macroxyelinae). No reports of economic damage caused by Xyelidae have been made.

Among the West Palearctic taxa, barcode sequences are available for eight of the 13 species, and six of these are diagnostic at the species level (Appendix S3, Supporting information). Only *Xyela julii* and *X. obscura* share the same BIN, but can be discriminated by colour and morphological characters of adults, as well as by their host plants, *Pinus sylvestris* and *P. mugo*, respectively (Blank 2002). *Xyela* species are generally poor in taxonomically useful characters, and morphologically, highly similar species may be associated with the same

host (Blank *et al.* 2005; Appendix S3, Supporting information). Even in the comparatively well-known European fauna, the existence of two different BINs in *X. menelaus* and the separation of *X. sp. 012* Portugal from *X. graeca* by a distance of 1.86% (Appendix S3, Supporting information) suggests the presence of undescribed, cryptic species. It was possible to confirm records of larval host plants by matching sequence data of adults with larvae extracted from staminate cones of pines, including *X. alpigena* (*Pinus cembra*) and *X. bakeri* (*P. sabiniana*) (Blank 2002; Blank *et al.* 2013). Palearctic *Xyela* species have been found to be generally monophagous (Blank *et al.* 2013), but BIN sharing of *Xyela* larvae extracted from different pine species in the western United



States suggests oligophagy for some Nearctic species (Appendix S3, Supporting information).

### Pamphiliidae

The Pamphiliidae ('web-spinning sawflies') is a moderately sized family with nearly 300 extant species worldwide (Taeger *et al.* 2010). It consists of the Cephalciinae associated with coniferous trees and the Pamphiliinae associated, with few exceptions, with deciduous trees. Some species of Cephalciinae may occasionally reach pest status and can cause economic damage to spruce and pine (Pschorn-Walcher 1982; Viitasaari 2002).

The barcode sequence data include 41 species, represented by 46 BINs. The comparatively large proportion of species with low interspecific distances (Table 2, Appendix S3, Supporting information) is mostly caused by the high number of such cases in *Cephalcia* (9 of the 11 species). This genus has been regarded as taxonomically difficult because of the low degree of morphological variation among species and the high level of intraspecific colour variation (Enslin 1912; Battisti *et al.* 1998). Within the genus, several BINs consist of two or more species; for example, BOLD:AAK0589 comprises four species (Appendix S3, S4, Supporting information), viz. *C. arvensis*, *C. lariciphila*, *C. erythrogaster* and *C. intermedia*. At the same time, some of these species exhibit BIN divergence, including *C. arvensis* occurring in five BINs and *C. lariciphila* in three BINs. The genus *Cephalcia* requires further study with a larger data set (21 of the 46 BINs are represented by singletons), re-evaluation of morphological characters and the use of additional genetic markers.

### Megalodontesidae

Currently, 42 species of the exclusively Palearctic Megalodontesidae are considered to be valid (Taeger *et al.* 2010). However, several additional species are still undescribed and about 50 species are expected altogether (Taeger 2002 and unpublished data). Known larval hosts are Apiaceae; records of Rutaceae and Lamiaceae require confirmation (Liston & Späth 2011). No species has ever been regarded as a pest. On the other hand, records of *Megalodontes* in Western Europe can be interpreted as indicative of the above average value to nature conservation of the sites where they occur.

Barcode sequences are available for 91 specimens, representing 21 nominal taxa. High intraspecific distances (>2%) were found in four taxa, including *Megalodontes flabellicornis*, *M. thor* and *M. flavicornis*. *Megalodontes phaenicius* represents a group of taxonomically uncertain species (Taeger 2002). The species *M. plagiocephalus*, *M. panzeri* and *M. turcicus* are morphologically very similar and also share barcodes. *Megalodontes quinquecinctus* and *M. spiraeae* share the

same BIN. However, they represent morphologically well-separated species despite being separated genetically by a distance of only 0.77% (Appendix S3, Supporting information, Taeger 1998a). The remaining taxa are well supported by the barcoding data, and even morphologically very similar species like *M. cephalotes/thor* and *M. eversmanni/reitteri* are clearly distinguished by their barcodes.

### Blasticotomidae

*Blasticotoma filiceti*, the single analysed species of this small family with 14 extant species in two genera worldwide (Taeger *et al.* 2010), can be identified unambiguously using DNA barcodes. All Blasticotomidae species are thought to be associated with ferns. Apart from occasional damage to ferns grown in gardens, the species have no economic significance.

### Argidae

The Argidae is, with about 920 known species, the second largest family of 'Symphyta' after the Tenthredinidae (Taeger *et al.* 2010). The collective host plant spectrum of Argidae is very wide, if the world fauna is considered. However, in Europe, a large number of species, particularly of *Arge*, feed upon woody angiosperms, while all known larvae of *Aprosthemina* are attached to herbaceous Fabaceae (e.g. Taeger *et al.* 1998; Vikberg 2004). Species such as *Arge pullata* on birch sometimes defoliate their hosts, but the main economic risk resulting from outbreaks of this species is poisoning of farm livestock after ingestion of larvae (Thamsborg *et al.* 1987).

Using DNA barcodes, 35 of the 47 morphologically defined argid species included in our data set can be identified to species level, whereas eight species belonging to two species complexes exhibit barcode sharing (Appendix S4, Supporting information). The *Arge clavicornis* group is represented by eight nominal species, viz. *Arge ciliaris*, *A. expansa*, *A. fuscipes*, *A. nigripes*, *A. pullata*, *A. shawi*, *A. sorbi* and *A. ustulata*. Of these, three species show a comparatively high intraspecific distance of 2.82% (*A. fuscipes*), 2.43% (*A. nigripes*) and 2.18% (*A. ustulata*), whereas their interspecific distance (i.e. minimum distance to the next neighbour) ranges from 0.00 to 0.17% (Appendix S3, Supporting information). Low levels of genetic differentiation in the *A. clavicornis* group are often paralleled by small differences in morphology (Schedl & Pschorn-Walcher 1984; Smith 1989), which sometimes are regarded as intraspecific variability (see, e.g. treatment of *A. ustulata* by Zhelochovtsev & Zinovjev 1994), but often the larvae are associated with different host plants (Taeger *et al.* 1998).

Another species group exhibiting low genetic differentiation includes the widespread Palearctic species

*A. melanochra* and specimens of *A. cingulata* from Iran. Specimens from sympatric populations of the two species in Iran show low interspecific differences, and they may therefore be conspecific colour forms.

Six species of Argidae are associated with more than one BIN, including *Arge cyanocrocea* and *A. pagana* with two BINs each, *A. ochropus* with four BINs, *Sterictiphora angelicae* with three BINs, and *A. rustica* and *A. scita* with five BINs each (Appendix S3, Supporting information). Maximum intraspecific differences in these taxa range from 3.30% (*A. cyanocrocea*) to 8.87% (*A. pagana*). In at least two cases (*A. cyanocrocea* and *A. scita*), these differences occur in sympatric populations and may therefore indicate the presence of cryptic species rather than intraspecific variation of widespread species as proposed earlier (Gussakovskij 1935; Pesarini 2002).

*Aproceros leucopoda*, a severe defoliator of elms that is invasive in Europe at least since 2003 (Blank *et al.* 2014), can be identified unambiguously by its barcode (Appendix S3, Supporting information). Sequences of females from Austria, Romania and Russia and of a larva collected from elm in Yunnan, China, correspond well (0.92% intraspecific variation) and confirm the East Asian origin of this exclusively parthenogenetic species that can occur as pest (Blank *et al.* 2010).

### Pergidae

The Pergidae is a moderately sized family with 441 described species (Schmidt & Smith 2016) and, after the Tenthredinidae and Argidae, the third largest family of 'Symphyta'. The family has a 'Gondwanan' distribution with most species occurring in Australia and South America. Food plants are extremely diverse, but data are lacking for most species, and better known for the Australian fauna than for the Neotropics. Many Australian species feed on species of *Eucalyptus* (*sensu lato*), others have such divergent food sources as dead or dying leaves, aquatic ferns, or fungi (Schmidt & Smith 2006).

Barcode sequence data are available for 47 species of Pergidae, covering all subfamilies except the South American Parasyzygoniinae. All species can unambiguously be identified using barcodes. In a few species, the barcoding data revealed the presence of undescribed species, supporting the notion that the family contains several groups of taxonomically unresolved species complexes that need to be clarified (Benson 1939; Macdonald & Ohmart 1993; Schmidt *et al.* 2006).

### Diprionidae

While the Diprionidae is a relatively small family with an estimated 145 species worldwide (Taeger *et al.* 2010), it ranks among the economically most important groups

of sawflies. Diprionid larvae feed on coniferous trees, in particular spruce and pine, and can cause considerable damage during outbreaks (Pschorn-Walcher 1982).

The present release includes 12 species in six genera, including the major pest species *Diprion pini* and *Neodiprion sertifer*. All species can be identified using barcodes, although *D. similis* is represented only by three shorter sequences of 418–419 bp. High intraspecific diversity occurs in *Monoctenus juniperi* (five BINs, 6.09% intraspecific distance), *Gilpinia frutetorum* (three BINs, 3.47% intraspecific variation) and *Gilpinia polytoma* (two BINs, 1.32%).

### Cimbicidae

Host plant data for the ca. 190 species in this family are mostly available only for Palearctic and Nearctic taxa. Species of *Cimbex*, *Trichiosoma*, *Praia* and *Pseudoclavellaria* (Cimbicinae) are associated with woody angiosperms, *Abia* (Abiinae) is mainly on herbaceous or woody Dipsacales, and *Corynis* (Corynidinae) larvae feed on diverse families of herbaceous eudicots. Economically, Cimbicidae are seldom of much significance, except for *Cimbex quadrimaculatus* as an orchard pest in Southern Europe (Cingovski 1965; Özbek 2014).

Barcode sequences were obtained for 28 species in six genera and, with the exception of *Trichiosoma tibiale*, all species can be unambiguously identified using barcodes. The single specimen of *T. tibiale* is placed in one of the three different BIN clusters representing *T. lucorum*, which reflects the taxonomically chaotic state of *Trichiosoma*. Species of *Cimbex* and *Trichiosoma* are morphologically variable, and their identification by morphological characters is often problematic (Gussakovskij 1947; Taeger 1998b). The barcode data will facilitate correct identification of species, although some of them possess high genetic variation (e.g. 3.72% in *T. lucorum*, Appendix S3, Supporting information). The taxonomy of problematic cimbicid genera needs to be re-evaluated using morphological and molecular data and broader taxon sampling, as the current data set includes only two *Trichiosoma* species. Analyses would probably also benefit from incorporation of specimen-level data on larval host plants.

### Tenthredinidae: Allantinae

The phylogeny of Allantinae, Blennocampinae and Heterarthrinae is still poorly understood. The classification and nomenclature used here follows the traditional views, with Athaliini treated as a tribe within the Allantinae, although recent phylogenetic analyses suggest that the Athaliini is only distantly related to the rest of Allantinae. Schulmeister (2003b) suggested that additional substantial changes to the classification of these taxa will be required (but see Malm & Nyman 2015).

The Allantinae includes worldwide about 110 genera and 900 species (Taeger *et al.* 2010). The host plant spectrum includes a wide range of herbaceous and woody eudicots. Economic damage by some species has been recorded in horticulture, for example *Allantus cinctus* on *Fragaria* and *Rosa* (Martelli 1941; Scheibelreiter 1973), as well as to forest trees, for example *Monsoma pulveratum* as an invasive pest on *Alnus* in North America (Kruse *et al.* 2010).

Barcode sequences are available for 15 genera and 90 morphologically identified species, of which 80% (73, including singletons and possibly undescribed species) can be recognized based on their barcodes and have a >2% distance to their nearest neighbour. The results indicate that many genera contain species groups that include cryptic species (e.g. *Monsoma pulveratum*, *Xenapates similis* and *Allantus viennensis*, cf. Appendix S3, Supporting information), species that are genetically so close that they cannot be distinguished using barcodes (two species groups in *Empria* and one in *Apethymus*), or the sources of discordance between barcoding and morphology-based taxonomy can only be resolved after major taxonomic revision (e.g. for two species groups of *Allantus*, involving 10 species). In addition, several genera include possibly undescribed species that were recognized from morphological data and that are supported by barcoding results (two species in each of the genera *Allantus*, *Ametastegia* and *Empria*). *Allantus* (16 species barcoded) is, according to our barcoding results, taxonomically the most challenging taxon in the Allantinae. The need for taxonomic revision is most obvious in two species complexes, the first containing eight species (*A. basalis*, *A. calceatus*, *A. cinctus*, *A. cingillum*, *A. cingulatus*, *Allantus* 002 nr. *cingulatus* Iran, *A. rufocinctus* and *A. truncatus*) and the other two species (*A. didymus* and *A. laticinctus*; cf. Appendix S5, Supporting information). *Allantus ariadne* from Cyprus, a species close to *A. didymus* and *A. laticinctus*, had already been described as new based on barcoding data (Liston & Jacobs 2012). *Empria*, with 20 species barcoded, includes two species groups with altogether seven species that cannot be distinguished using barcoding (Appendix S3, Supporting information). For separation of these species, ITS1 and ITS2 sequences have been shown to work much better (Prous *et al.* 2011).

The genus *Athalia* is, together with related genera, nowadays treated as a separate subfamily Athaliinae, following Schulmeister (2003b) and Malm & Nyman (2015). Larval hosts are herbaceous eudicots, particularly Brassicaceae and Lamiales (Opitz *et al.* 2011). Economically significant damage to crops of cultivated Brassicaceae is frequently caused by some *Athalia* species (Benson 1962). *Athalia* and *Hypsathalia*, with 21 species and 115 specimens barcoded, contain a large number of putatively

cryptic species, in particular in *A. ancilla*, *A. circularis*, *A. cordata*, *A. cornubiae*, *A. incompta*, as well as in the turnip sawfly *A. rosae*, a species of which populations can reach pest status (Hill 1987, Appendix S3, Supporting information).

#### *Tenthredinidae: Blennocampinae, Heterarthrinae*

The Blennocampinae is a morphologically and biologically highly diverse assemblage, currently regarded to comprise about 650 species in one hundred genera (Taeger *et al.* 2010). Larval host plants include a broad spectrum of herbaceous and woody eudicots, and less frequently monocots. A number of species cause significant damage to trees, for example *Tomostethus nigratus* on *Fraxinus* (Austarå 1991). In contrast to the Blennocampinae, with mostly exophytic larvae, the leaf-mining Heterarthrinae is a monophylum (when *Caliroa* and *Endelomyia* are moved to the Blennocampinae; Leppänen *et al.* 2014; Malm & Nyman 2015) and contains about 170 species (Taeger *et al.* 2010).

*Blennocampinae*. Two European *Ardis* species are recognized at present, but their barcodes segregate under two shared BINs (Appendix S3, Supporting information) that do not correlate with the colour characters currently used to separate the species (Blank & Taeger 1998). Within *Claremontia*, barcodes are available for most of the European species that are currently treated as valid, and the barcode differences are in general sufficient for reliable identification. Since identification based on morphology is sometimes difficult, barcoding should be very useful in this genus. *Claremontia puncticeps* exhibits genetic divergence and is split into three BINs that require taxonomic reappraisal. Barcoding results for *Periclista* also seem promising, but a number of species have not yet been sampled. Specimens identified as *Monophadnus pallescens* and *Monophadnus* sp. show BIN divergence and confirm the notion that the taxonomy of this species complex is inadequately resolved (Prieto *et al.* 2007; Blank *et al.* 2009). The three sampled European *Eurhadinoceraea* species (the two remaining species are extremely rare) display very divergent barcodes. While the adults are easily distinguished using colour characters, barcoding is essential for separating their larvae, which share the same host plant genus (*Clematis*). *Eutomostethus ephippium ephippium* and *E. nigrans*, which exhibit no structural differences (Blank & Taeger 1998), share the same BIN and lack a barcode gap (Appendix S3, S4, Supporting information), and thus appear to be colour forms of the same species. Contrastingly, *E. ephippium vopiscus* differs by 3.61% from *E. ephippium ephippium*, and they may in the future be treated as separate species. Barcodes of *E. gagathinus*, *E. luteiventris* and *E. punctatus* differ very clearly from each other, but *E. gagathinus* is



divided into two BINs, with one BIN containing Central European specimens and the second BIN consisting of a single specimen from Greece (Appendix S3, Supporting information). Four sampled European *Caliroa* species exhibit large barcode divergences, which is encouraging in view of the difficulty in identifying some species using morphological characters. The two described European *Endelomyia* species, which are morphologically hard to distinguish (Lacourt 1998), segregate in two barcode groups with an interspecific distance of 3.96% (Appendix S3, Supporting information).

*Heterarthrinae*. Most of the included species are clearly identifiable by their barcodes. Exceptions include *Fenella judaica/nigrita* (1.71% interspecific distance), *Fenusa ulmi/altenhoferi* (0.81%) and *Heterarthrus wuestneii/imbrosensis* (0.65%), although there is an indication of a barcode gap that may allow species separation (cf. Appendix S5, Supporting information). Noteworthy is the large minimum distance between *Parna kamijoi* and *P. apicalis* of about 9.57%, strongly supporting the notion that they are separate species (Liston 1993; Liston *et al.* 2006), rather than conspecific as suggested by Byun & Shinohara (1999).

#### *Tenthredinidae: Nematinae*

The Nematinae, containing about 1250 species worldwide (Taeger *et al.* 2010), is particularly species-rich in the cooler parts of the Northern Hemisphere, even dominating the sawfly fauna of subarctic areas. Large deficits in taxonomy often make their identification very difficult, especially among the species-rich 'higher Nematinae' (Nyman *et al.* 2006). The classification of Nematinae has recently been revised based on a comprehensive phylogenetic analysis (Prous *et al.* 2014). However, the nomenclatural changes that have been proposed are not yet implemented in BOLD, and we therefore use the traditional classification for the following discussion of results.

The majority of hosts are shrubs or trees, both gymnosperm and angiosperm, although many species are attached partly or exclusively to herbaceous eudicots or monocots. Particularly, some of the Nematinae on coniferous trees are known to undergo outbreaks resulting in significant economic damage, for example *Pachynematus montanus* on *Picea abies* (Schafellner & Schopf 2014). The latter sawfly species can be unambiguously identified using barcoding (cf. Appendix S3, Supporting information).

Barcode-based identification of most European species in many of the smaller genera of the 'basal Nematinae' seems likely to be successful. This applies to *Hoplocampa*, *Mesoneura*, *Pseudodineura* and *Stauronematus*. However, some individual cases need further investigation. One of the two specimens of *Hoplocampa minuta*

shares a BIN with *H. fulvicornis*, whereas the other *H. minuta* is placed in its own BIN. *Mesoneura opaca* has two BINs and an intraspecific divergence of 2.02% (Appendix S3, Supporting information). The eight *Nematinus* species for which barcode data are available (representing all seven currently recognized European species and a possible previously unrecognized species) show clear differences, except for *N. luteus* and *N. fuscipennis*, which share the same BIN and do not exhibit a barcode gap (Appendix S4, S5, Supporting information). Barcoding results are promising for *Anoplonyx*, in which the three sampled species segregate clearly in three BINs, but two additional European species remain without barcodes. In *Dineura*, with all five known European morphospecies sampled, *D. virididorsata* and *D. pullior* share the same BIN, as do *D. stilata* and *D. testaceipes*. In the latter case, the status of the species has long been a subject of controversy (e.g. compare treatment by Lindqvist (1955) with that by Benson (1958)). On the other hand, *D. parvicornis* seems to be clearly identifiable by its barcode and host plant (Liston 2015), despite being morphologically difficult to separate from *D. testaceipes*.

In *Cladius*, a number of long-standing taxonomic problems are reflected also in the barcoding data. Particularly regarding *C. pectinicornis*, there is an ongoing discussion concerning whether this represents a single species (e.g. Zhelochovtsev 1952) or at least two species, *C. pectinicornis* and *C. difformis* Panzer, 1799 (e.g. Benson 1958; Wei 2001). A high level of intraspecific variability, suggesting existence of cryptic species, is indicated in *C. pilicornis* (three BINs), *C. pectinicornis* (three BINs), *C. brullei* (four BINs) and *C. compressicornis* (four BINs; Appendix S3, Supporting information).

Barcoding results for the 'higher Nematinae' exhibit clear differences between a majority of morphospecies in the species-rich genera *Pristiphora*, *Nematus* and *Amauronematus* [the latter now included within *Euura* (Prous *et al.* 2014)]. In each of these genera, there are several cases of shared BINs, despite the fact that the species involved are morphologically and biologically relatively well characterized. Examples of these are *Pristiphora* BOLD:AAK9450 containing *P. abietina*, *P. compressa*, *P. pseudodecipiens*, *P. gerula*, *P. decipiens* and *P. saxesenii*; and *Amauronematus* BOLD:ABU5508 containing *A. histrio*, *A. mimus*, *A. arvii* and *A. stenogaster*. Remarkable is the sharing of BOLD:ABU5509 by *Nematus flavescens* and *N. reticulatus*, species that superficially (colour and size) are very different from each other.

A case where barcoding divergence may lead to revision of previous taxonomic opinions is based on the large sequence divergence between *Nematus desantisi* (species revocata) and *N. oligospilus*. The former species is adventive in South America, South Africa, Australia and New Zealand (Schmidt & Smith 2009; Caron *et al.*



2014). It was synonymized with *N. oligospilus* by Koch & Smith (2000). Morphologically, these taxa are certainly very similar, although the colour differences in living adults had already led to doubts that they are conspecific (female *N. desantisi* with yellow-brown and partly whitish pale body parts; female *N. oligospilus* with pale body parts green).

In some nominal species, the existence of multiple BINs underlines a still inadequate state of taxonomic knowledge that specialists have long been aware of. Examples are *Pachynematus fallax* (four BINs (Taeger & Blank 1998)) and *Amauronematus viduatus* [two BINs; Appendix S3 (Supporting information), V. Vikberg, unpublished results]. At least in these two cases, the existence of additional species seems probable. The hitherto problematic association of sexes, which particularly in *Pachynematus* is of taxonomic significance, will most likely be facilitated by use of barcode data.

Poorly resolved by barcoding are species of the gall-making 'Euurina'. Large divergences are apparent between species groups that have sometimes been treated as separate genera or subgenera (Zinovjev & Vikberg 1999; e.g. Vikberg 2010), and barcode divergence between the four sequenced species of *Tubpontania* is rather pronounced (Appendix S3, Supporting information) but, within other groups, many of nominal species share the same BIN. Examples are BOLD:AAR6800, which includes at least nine nominal species of *Pontania* subgenus *Eupontania*, and BOLD:ABV1036, which includes six species of the *Phyllocolpa oblita* group (Appendix S4, Supporting information). In Northern European leaf (sg. *Eupontania*) and bud gallers (*Euura*), differences in haplotype and allele frequency exist despite low genetic divergence in mitochondrial COI and nuclear ITS2 sequences (Leppänen *et al.* 2014).

### *Tenthredinidae: Selandriinae*

Following Taeger *et al.* (2010), the Selandriinae here includes also the Dolerinae and the Heptamelidae sensu Malm & Nyman (2015). Hosts include ferns and horse-tails, the monocot Poales (Taeger *et al.* 1998), dicot Boraginaceae (Macek 2013) and, for one species, even mosses (Vikberg & Nuorteva 1997). Selandriinae species have rarely been reported as economically significant pests.

The Dolerini, and in particular its major genus *Dolerus*, has long been known to be taxonomically problematic because of the morphological similarity among adults and convergences in shape and colour patterns (Goulet 1986; Heidemaa 2004). This situation is reflected in our barcoding results, in which several species of *Dolerus* exhibit BIN sharing with two, three, four or even eight species in a single BIN (Appendix S4, Supporting information). Some species are found in two or three

BINs, adding to the discordance between morphological and molecular data. Examples are *D. frigidus*, *D. gessneri*, *D. liogaster* and *D. stygius*, which belong to clusters containing two or more species, but at the same time exhibit genetic divergence and consist of two BINs each (Appendix S3, S4, Supporting information). It appears that combined analyses of morphological, ecological and molecular data are necessary to resolve the taxonomic problems that are inherent in this genus (Heidemaa *et al.* 2004).

Genetic divergence was observed in a few other species of the Selandriinae, including *Stromboceros delicatulus* (Appendix S3, Supporting information). A single specimen from Sicily seems to represent a different species, whereas specimens from Central and Northern Europe show an intraspecific divergence of 2.21% and were assigned to two different BINs (Appendix S3, Supporting information), indicating the possible presence of two sympatrically occurring species. Similarly, two species in the genus *Strongylogaster* show BIN divergence with two (*S. mixta*) or even four BINs (*S. macula*), all of which appear to occur sympatrically in Central Europe (Appendix S3, Supporting information).

The genera *Heptamelus* and *Pseudoheptamelus* are, according to Malm & Nyman (2015), better treated as a separate family Heptamelidae, with about 50 known species in six genera. All known larvae feed on ferns (Vikberg & Liston 2009). No reports exist of economic damage caused by Heptamelidae. The two rather similar European *Heptamelus* species are well separated by their barcodes (distance 8.25%).

### *Tenthredinidae: Tenthredininae*

The Tenthredininae is by far the largest subfamily of the Tenthredinidae and the 'Symphyta' in general, with about 1775 species worldwide (Taeger *et al.* 2010). Most species are attached to herbaceous eudicots, rarely to monocots, but woody angiosperms are used by a significant minority of species, such as *Tenthredo* and *Rhogogaster* (Taeger *et al.* 1998). Reports of economic damage caused by Tenthredininae are rare.

The present study includes 1350 specimens, representing 263 species in 12 genera, equalling almost one-third of the specimens and one quarter of the species analysed. The number of BINs is, with 248, lower than the number of species, indicating BIN sharing in several species. About one-third of the species (96 species, 37%) show a minimum interspecific distance of less than two per cent (Table 2). This is particularly the case in *Tenthredo*, with nine species pairs sharing the same BIN each, three BINs with three species each, two BINs with five species each and 12 species of the *T. arcuata* species complex sharing the same BIN. Intraspecific distances

within this group range from 0.00% to 4.28% (Appendix S3, Supporting information).

Nominal species of the *T. arcuata* complex are separated by very low interspecific genetic distances. Taxonomists have traditionally regarded this species group as being notoriously difficult (Taeger 1985, 1988). The neighbour-joining tree of five species of the complex, represented by 25 specimens that were all collected in the same area, shows (with the exception of one specimen of *T. korabica*) a distinct separation of species (Fig. 2). All specimens were collected in 2010 and 2011 between July 1 and July 26. The intraspecific distance ranged from 0.00 to 0.50%, with a minimum distance to the nearest neighbour species ranging from 0.15 between *T. algoviensis* and *T. korabica* to 1.81% between *T. n. notha* and *T. korabica*. The separation of species into distinct clusters disappears when specimens from a wider geographic area are included. Intraspecific distances increase with geographic distance and then exceed the interspecific distances, a problem that has also been documented in other taxa (Bergsten *et al.* 2012; Huemer *et al.* 2014).

BIN divergence is prevalent in several species of *Tenthredo*, with two (16 species), three (two species) or four (one species) BINs within a species (Appendix S3, Supporting information). Some of these species, like *T. mesomela*, rank among the most commonly collected sawflies in Central Europe.

Apart from species of *Tenthredo*, several other genera exhibit BIN sharing, including species of *Macrophya*,

*Tenthredopsis*, *Siobla* and *Rhogogaster*. *Macrophya blanda* and *M. oedipus* are separated by an interspecific distance of only 0.15% (Appendix S3, Supporting information), which is relevant to the identification of specimens from Greece, where the species occur sympatrically (Appendix S3, Supporting information).

Species of *Tenthredopsis* exhibit a high level of intraspecific colour variation, which has led taxonomists to describe numerous species and varieties (Blank & Ritau 1998; and discussion therein). Our barcoding results reflect the difficult taxonomic situation in this genus: specimens of *T. friesei*, *T. nassata*, *T. litterata*, *T. scutellaris* and *T. sordida* are placed in two or more genetic clusters (Appendix S3, Supporting information) and show high levels of intraspecific variation, in particular *T. friesei* (5.81%, merging with *T. coquebertii*), *T. nassata* (5.56%, merging with *T. sordida*), *T. scutellaris* (5.91%, merging with *T. nassata*) and *T. sordida* (6.24%, merging with *T. nassata*) (Appendix S3, Supporting information). On the other hand, the barcode results support the recent removal of *T. andrei* and *T. corcyrensis* from synonymy (Pesarini 2002) and reveal potential undescribed species from Iran and Israel (e.g. BOLD:ABV9487 with 7.60% distance to *T. tarsata*, BOLD:ABV9486 with 8.09% distance to *T. andrei*). Two species complexes of *Rhogogaster*, a moderately sized genus with about 40 species worldwide that is distributed in the Northern Hemisphere, show BIN sharing, that is the *R. viridis* species group (with *R. chlorosoma* and *R. viridis*) and the *R. genistae*



**Fig. 2** Neighbour-joining tree of barcode sequence data of five sympatrically occurring species of the *Tenthredo arcuata* species complex, illustrating separation of species by barcode gaps despite low interspecific divergence. Map insert shows collecting localities of specimens that were used to create the tree (map data: Google, DigitalGlobe). Terminal branch information includes species, sample ID, country, province, exact location, latitude, longitude and elevation in metres a.s.l. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

group (with *R. chambersi*, *R. genistae* and *R. picta*) (Appendix S3, Supporting information). Despite a limited number of species in any particular area, identification of the aforementioned species has been highly problematic (Taeger & Viitasaari 2015). In particular, identifications of *R. viridis* have to be treated with care, and the name '*R. viridis*' (now *R. scalaris*) is here used in its traditional sense (cf. Taeger & Viitasaari 2015). The high intraspecific variation in *R. punctulata* (5.05%) suggests the presence of two or more species separated by a clear barcode gap (Appendix S5, Supporting information), and BIN divergence was observed also in *R. chambersi* and *R. dryas* (Appendix S3, Supporting information).

### Cephidae

The Cephidae comprises 171 species distributed in the Northern Hemisphere (Taeger *et al.* 2010) and exceptionally in Australia (Smith & Schmidt 2009) and Madagascar (Benson 1935). The larvae are internal feeders: those of Cephini in Poaceae, those of Hartigiinae in mostly woody dicots and those of Pachycephini in Asteraceae and Papaveraceae (Zhelochovtsev 1968; Scheibelreiter 1978; Taeger *et al.* 1998).

Barcodes enable the identification of 25 described species, an undescribed *Phylloecus* (= *Hartigia*, Liston & Prous 2014) species from Iran (BOLD:ABW9690), and a new genus and species from China (BOLD:ACG2240). The barcode data set includes the 'European wheat stem sawfly' *Cephus pygmeus* and the 'black grain stem sawfly' *Trachelus tabidus*, both of which are invasive in the Nearctic region and well-known pests of cereals (e.g. Gahan 1920; Miller *et al.* 1993), as well as the 'rose shoot sawfly' *Syrista parreyssii*, with larvae damaging roses used for perfume production (Nikolova & Natskova 1970; Liston 2012). *Calameuta haemorrhoidalis* exhibits 5.31% intraspecific variation and includes two geographically separated clusters (Greece and Sicily) (Appendix S3, S5, Supporting information), but it remains unclear if this reflects geographic variation or if different species are involved. Similarly, in *C. pallipes* (4.96% intraspecific variation), a specimen from Italy is assigned to a separate BIN (Appendix S3, Supporting information). The occurrence of one of five specimens of *Hartigia linearis* under BIN BOLD:ABW9043, which otherwise comprises *H. xanthotoma*, reflects the difficulty in distinguishing males of these two species morphologically (Jansen 1998).

### Siricidae and Xiphydriidae

The two families of woodwasps include about 150 species (Siricidae) and 120 species (Xiphydriidae) worldwide (Taeger *et al.* 2010). Several species damage

economically important conifers or woody dicots, because females infect host trees with harmful fungi during oviposition, and the larvae bore into wood (Eichhorn 1982; Schiff *et al.* 2012).

For Siricidae, DNA barcoding has been shown to be a reliable identification tool, in particular for larvae, which cannot be identified based on morphology (Schiff *et al.* 2012). The present study includes five species of Siricidae and four species of Xiphydriidae, all of which are separated by distinct barcode gaps (Appendix S3, Supporting information). However, several species are represented by one or two specimens only, so further material is needed to assess intraspecific variation.

### Orussidae

The enigmatic Orussidae includes 89 extant species worldwide. These are the only parasitoids among the sawflies and horntails (Vilhelmsen *et al.* 2013). Orussidae are rarely collected, and only two species, *Orussus abietinus* and *O. unicolor*, are represented in the data set (Appendix S3, Supporting information). At 0.15%, intraspecific variation in *O. abietinus* is very low, despite the fact that the collecting localities of the specimens (eastern Germany and Epirus Mountains in Greece) are separated by a distance of 1500 km.

### Concluding remarks

The present barcode release provides the foundation for a comprehensive DNA barcode library of sawflies and woodwasps. It covers about 12% of all species of 'Symphyta', with representatives of all extant families. Only the Anaxyelidae is not included in the present study, but sequences are available in BOLD and GenBank. Species with low levels of taxonomic uncertainty can be readily identified using DNA barcodes. These taxa include, for example, the Cephidae, Diprionidae, Megalodontesidae, Siricidae, Xiphydriidae, most Argidae and most species of the tenthredinid subfamilies Tenthredininae, Blennocampinae and Heterarthrinae. For these species, the present barcode library will provide a valuable identification tool for ecologists and applied entomologists. It allows not only the identification of adults but can also be used for reliable identification of larvae, most of which are external (the majority) or internal feeders on a wide range of herbaceous or woody eudicot plants.

The OTU designation process in BOLD employs a two-algorithm process, comprising a single linkage clustering analysis using a 2.2% threshold, followed by Markov clustering (Ratnasingham & Hebert 2013). It has been shown to yield clusters that match species that were delimited using traditional taxonomic methods, which is the primary goal for automated species delimitation



algorithms. In addition, it performs much better than other methods, most of which are not scalable for use with the barcode library that currently contains over five million barcode sequences.

The slightly higher number of BINs (1058) compared to the number of species (1037) may indicate the presence of cryptic diversity. In fact, our results indicate the presence of cryptic diversity in numerous sawfly taxa. Although the status of most putative cryptic species still needs to be evaluated, detailed examination of specimens often revealed parallels between genetic and morphological variation.

In many cases, the finding of subtle but consistent morphological differences suggests that the apparent cryptic species might have been resolved using traditional approaches. Nevertheless, the barcode data highlighted the existence of taxonomically critical taxa, and this awareness will help to prioritize groups in need of revision. The relatively high number of such cases is surprising in the comparably well-known fauna of Central Europe. However, sawflies have received rather little attention from hymenopteran taxonomists, compared to more prominent groups like bees and wasps (but much more attention than the parasitoids, which represent the vast majority of Central European Hymenoptera; Quicke 2012). In addition, certain groups of sawflies clearly have been neglected because they are taxonomically difficult, species-rich and/or require special preparation techniques (genitalia). Consequently, cryptic species were expected or even known to exist in, for example a large proportion of species of the Nematinae, and species complexes like *Tenthredo arcuata*, *Rhogogaster viridis* and *Arge clavicornis* (see Results and discussion section). However, cryptic diversity came as a surprise in many groups such as *Athalia rosae* or *Aglaostigma fulvipes*, which are among the most common species in Central Europe. The taxonomy of these easily recognizable species was thought to be resolved, so they attracted no taxonomic interest, even though some were used as objects in chemo-ecological studies (e.g. *A. rosae*, Opitz *et al.* 2011, and references therein) or may occur as pests. In cases like this, routine barcoding can safeguard against compromising ecological or applied studies by mixing up species.

In a recent study, Mutanen *et al.* (2016) concluded that, in European Lepidoptera, most cases of nonmonophyly in DNA barcode sequence data can be attributed to methodological issues, including misidentifications, oversplitting or overlooking species, and subjectivity of species delimitation (in particular with respect to allopatric populations). It remains to be seen to what extent these factors explain nonmonophyly in 'Symphyta', and case-by-case studies are needed because the degree to which methodological factors have contributed to

nonmonophyly can be expected to vary considerably across taxa. For example, there is evidence that oversplitting has occurred in gall-making nematine sawflies on willows (A. Liston *et al.*, in preparation). However, we do not expect this to be more common in 'Symphyta' than in other groups of insects. Lack of taxonomic treatments and usable keys for identification have invariably led to a certain degree of misidentification, but this cannot explain the problematic situation found in the genus *Allantus* or in the *Arge clavicornis* species complex, where apparent nonmonophyly affects species that can reliably be distinguished using morphology.

It is to be expected that misidentification is a more frequent cause of nonmonophyly in groups that are taxonomically problematic. For example, in web-spinning sawflies, a family with two subfamilies, species identification is straightforward in the subfamily Pamphiliinae, and barcoding results are congruent with traditional species. By contrast, the second pamphiliid subfamily, Cephalciinae, and in particular its major genus *Cephalcia*, exhibits incongruence between molecular and morphological data (see Results and discussion section). The identification of *Cephalcia* species has long been known to be problematic, so nonmonophyly may, at least partly, be caused by misidentification of specimens. The likelihood of misidentifications is elevated by the fact that available keys rely, to a large extent, on colour characters that in some species may be more variable than expected.

However, barcode sharing among species may in many cases be 'real' and reflect retention of ancestral polymorphisms (incomplete lineage sorting) in recently diverged lineages and/or introgression of mitochondrial genes through occasional hybridization and backcrossing of hybrids to the parental forms (Mardulyn *et al.* 2011). Leppänen *et al.* (2014) showed that, despite extensive sharing of mitochondrial COI and nuclear ITS2 sequences among species within Northern European sg. *Eupontania* leaf gallers and *Euura* bud gallers, most species exhibit marked differences in haplotype and allele frequencies. As the authors pointed out, this finding could be explained by incomplete lineage sorting or by occasional introgression, so further studies with more powerful multilocus markers are needed. High levels of interspecific gene flow have been documented in *Neodiprion* (Linnen & Farrell 2007, 2008). As Patten *et al.* (2015) showed, the haplodiploid sex determination system of Hymenoptera facilitates mitochondrial introgression in relation to cross-species transfer of nuclear genes. Indeed, species of *Empria* that cannot be distinguished using barcodes can be separated with ITS1 and ITS2 sequences (Prous *et al.* 2011), and in *Neodiprion*, results of phylogenetic analyses of nuclear sequences agreed with morphology (Linnen & Farrell 2008). It remains to be



studied whether interspecific barcode sharing is more common within the Hymenoptera than in insect taxa with chromosomally based sex determination (XY or ZW systems).

Other factors that may account for the lack of agreement between morphology-based taxonomy and DNA barcodes includes the presence of bacterial endosymbionts like *Wolbachia*. Infection with different strains within a species can lead to overestimation of diversity (Whitworth *et al.* 2007; Xiao *et al.* 2012), whereas cross-species infections with the same *Wolbachia* strain through hybridization and subsequent 'mitochondrial sweeps' may lead to underestimating diversity (Whitworth *et al.* 2007; Raychoudhury *et al.* 2009).

Possibly, species in our data set with discordant results between morphological and molecular data suffer from *Wolbachia*-induced conflation of infected lineages. High levels of introgression, that is hybridization and subsequent repeated backcrossing, across species has been documented in *Neodiprion* (Linnen & Farrell 2007, 2008). Movement of genes across species by introgression is a widespread phenomenon and extremely variable across species, but the underlying causes are little understood (Patten *et al.* 2015). However, in these cases, taxa should be diagnosable with nuclear sequences. Species of *Empria* that cannot be distinguished using barcodes can be separated with ITS1 and ITS2 sequences (Prous *et al.* 2011), and in *Neodiprion*, results of phylogenetic analyses of nuclear sequences agreed with morphology (Linnen & Farrell 2008).

The effect of geographic scale of sampling on barcoding results has been discussed by Bergsten *et al.* (2012) and Huemer *et al.* (2014), and the effect is in many cases highly significant also in sawflies. This is illustrated by the frequently large divergences between barcodes of Central and North European specimens and those of the same morphospecies from Iran and the Mediterranean islands Sicily and Cyprus. Taxonomic evaluation of the differences is in most cases not at present possible because of a lack of samples from intervening areas. Therefore, our barcode data can generally be expected to be more suitable as references for the identification of Central and North European specimens, while caution should be used when determining specimens with provenances outside these areas.

Knowledge of sawfly larvae and host plant associations is generally incomplete, and existing data often require corroboration (e.g. Lorenz & Kraus 1957; Taeger *et al.* 1998). Furthermore, larvae of many species are difficult to rear through to the adult stage. Barcode-based identification of larvae has the potential to rapidly improve biological knowledge on sawflies by reducing the need for risky and time-consuming rearing (cf. Nyman *et al.* 2015). In a similar way, correct association

of the sexes will in many cases be made easier. However, this and other applications of DNA barcoding (e.g. metabarcoding) rely heavily on the existence of comprehensive barcode libraries. Future efforts should therefore aim at complementing the reference library of 'Symphyta' barcodes, but should also examine the causes for barcode incongruence in some groups. In addition to addressing methodological issues, an integrative taxonomic approach is called for in order to disentangle problematic species and species groups of 'Symphyta'.

## Acknowledgements

The collection and processing of specimens were funded by the Bavarian Ministry of Education and Culture, Science and Art (Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst, project 'Barcoding Fauna Bavarica') and by the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, project 'German Barcode of Life'). The sequence analyses for this study were supported, in part, by Genome Canada through the Ontario Genomics Institute, while informatics support was provided through a grant from the Ontario Ministry of Research and Innovation. The following specialists provided specimens: Ewald Altenhofer, Wolfgang Bäse, Keith Bland, Frank Burger, Hans-Joachim Flügel, Hans-Joachim Jacobs, Ewald Jansen, Guy Knight, Frank Koch, Klaus Mandery, Gisela Merkel-Wallner, Bruno Peter, Henri Savina, Mark Shaw, Jochen Späth and Johannes Voith. The contribution of MP was funded by the Swedish Taxonomy Initiative.

## References

- Aguiar AP, Deans AR, Engel MS *et al.* (2013) Order Hymenoptera. In: *Animal Biodiversity: An Outline of Higher-level Classification and Survey of Taxonomic Richness (Addenda 2013)* Zootaxa (ed Zhang Z.-Q.), pp. 1–82. Magnolia Press, Auckland, New Zealand.
- Austarå Ø (1991) Severe outbreaks of the ash sawfly *Tomostethus nigratus* F. (Hymenoptera: Tenthredinidae) on ornamental trees in Oslo. *Anzeiger für Schädlingskunde, Pflanzenschutz, Umweltschutz*, **64**, 70–72.
- Battisti A, Boato A, Zanocco D (1998) Two sibling species of the spruce web-spinning sawfly *Cephalcia fallenii* (Hymenoptera: Pamphiliidae) in Europe. *Systematic Entomology*, **23**, 99–108.
- Benson RB (1935) On the Genera of Cephidae, and the erection of a new Family Syntexidae (Hymenoptera, Symphyta). *Annals and Magazine of Natural History, including Zoology, Botany, and Geology*, 10th series, **16**, 535–553.
- Benson RB (1939) A revision of the Australian sawflies of the genus *Perga* Leach, sens. lat. (Hymenoptera Symphyta). *Australian Zoologist*, **9**, 324–357.
- Benson RB (1958) Hymenoptera, Symphyta. *Handbooks for the Identification of British Insects*, **6**, 139–258.
- Benson RR (1962) A revision of the Athaliini (Hymenoptera: Tenthredinidae). *Bulletin of the British Museum (Natural History) Entomology*, **11**, 333–382.
- Bergsten J, Bilton DT, Fujisawa T *et al.* (2012) The effect of geographical scale of sampling on DNA barcoding. *Systematic Biology*, **61**, 851–869.
- Blank SM (2002) The Western Palaearctic Xyelidae (Hymenoptera). In: *Sawflies (Hymenoptera, Symphyta) I. A review of the suborder, the Western Palaearctic taxa of Xyeloidea and Pamphilioidea* (ed Viitasaari M.), pp. 197–233. Tremex, Helsinki.

- Blank SM, Ritzau C (1998) Die Tenthredopsini Deutschlands (Hymenoptera, Tenthredinidae). In: *Pflanzenwespen Deutschlands (Hymenoptera, Symphyta)* (eds Taeger A., Blank S. M.), pp. 227–246. Goecke & Evers, Keltern.
- Blank SM, Taeger A (1998) Comments on the taxonomy of Symphyta (Hymenoptera) (Preliminary studies for a catalogue of Symphyta, part 4). In: *Pflanzenwespen Deutschlands (Hymenoptera, Symphyta). Kommentierte Bestandsaufnahme* (eds Taeger A., Blank S. M.), pp. 141–174. Goecke & Evers, Keltern.
- Blank S, Shinohara A, Byun B-K (2005) The East Asian *Xyela* species (Hymenoptera: Xyelidae) associated with Japanese Red Pine (*Pinus densiflora*; Pinaceae) and their distribution history. *Insect Systematics & Evolution*, **36**, 259–278.
- Blank SM, Taeger A, Liston AD *et al.* (2009) Studies toward a World Catalog of Symphyta (Hymenoptera). *Zootaxa*, **2254**, 1–96.
- Blank SM, Hara H, Mikulas J *et al.* (2010) *Aproceros leucopoda* (Hymenoptera: Argidae): an East Asian pest of elms (*Ulmus* spp.) invading Europe. *European Journal of Entomology*, **107**, 357–367.
- Blank SM, Shinohara A, Altenhofer E (2013) The Eurasian species of *Xyela* (Hymenoptera, Xyelidae): taxonomy, host plants and distribution. *Zootaxa*, **3629**, 1–106.
- Blank SM, Köhler T, Pfannenstill T *et al.* (2014) Zig-zagging across Central Europe: recent range extension, dispersal speed and larval hosts of *Aproceros leucopoda* (Hymenoptera, Argidae) in Germany. *Journal of Hymenoptera Research*, **41**, 57–74.
- Byun B-K, Shinohara A (1999) Notes on a Linden Leaf Mining Sawfly, *Parna kamijoi* (Hymenoptera, Tenthredinidae), new to Korea with brief life history notes. *Korean Journal of Applied Entomology*, **38**, 81–84.
- Caron V, Ede F, Sunnucks P, O'Dowd DJ (2014) Distribution and rapid range expansion of the introduced willow sawfly *Nematus oligospilus* Förster (Hymenoptera: Tenthredinidae) in Australasia. *Austral Entomology*, **53**, 175–182.
- Cingovski J (1965) Biološki i ekološki proučavan'a na bademovata lisna osa (*Palaeocimbex quadrimaculata* Müll.) vo Makedonii. *Posebno izdanje. Prirodnauchen Muzej na Makedonija, Skopje*, **3**, 1–106.
- Eichhorn O (1982) Familienreihe Siricoidea. In: *Die Forstschädlinge Europas* (ed Schwenke W.), pp. 196–231. Paul Parey, Hamburg, Berlin.
- Enslin E (1912) Die Tenthredinoidea Mitteleuropas. *Deutsche entomologische Zeitschrift, Beihefte*, **1–7** (1912–1917), 1–790.
- Gahan AB (1920) Black grain-stem sawfly of Europe in the United States. *United States Department of Agriculture, Bulletin*, **834**, 1–18.
- Goulet H (1986) The genera and species of the Nearctic Dolerini (Symphyta, Tenthredinidae, Selandriinae): classification and phylogeny. *Memoirs of the Entomological Society of Canada*, **135**, 1–208.
- Gussakovskij VV (1935) Insectes Hyménoptères, Chalastogastra 1. *Fauna SSSR, Moskva and Leningrad*, **2**, 1–453.
- Gussakovskij VV (1947) Insectes Hyménoptères, Chalastogastra 2. In: *Faune de l'URSS (eds Pavlovsky E. N., Stackelberg A. A.)*, pp. 1–235. Leningrad, Moscow.
- Hausmann A, Balke M, Haszprunar G *et al.* (2012) Barcoding Fauna Bavarica: Taking Barcoding of Central European Animals to a new level. *113. Jahrestagung der Gesellschaft für Biologische Systematik, Programm und Abstracts*, 34.
- Hausmann A, Godfray HCJ, Huemer P *et al.* (2013a) Genetic patterns in European geometrid moths revealed by the Barcode Index Number (BIN) system. *PLoS ONE*, **8**, 1–11.
- Hausmann A, Hendrich L, Balke M *et al.* (2013b) A major contribution of the ZSM to the genetic library of life: 35,000 insect species barcoded in 7 years. In: *Abstract Volume NOBIS* (eds Kroh A., Berning B., Haring E., *et al.*), p. 86. Gesellschaft für Biologische Systematik, Vienna, Austria.
- Hebert PD, Cywinska A, Ball SL *et al.* (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*, **270**, 313–321.
- Heidema M (2004) Systematic studies on sawflies of the genera *Dolerus*, *Empria*, and *Caliroa* (Hymenoptera: Tenthredinidae). *Dissertationes biologicae Universitatis Tartuens*, **99**, 1–162.
- Heidema M, Nuorteva M, Hantula J, Saarna U (2004) *Dolerus asper* Zaddach, 1859 and *Dolerus brevicornis* Zaddach, 1859 (Hymenoptera: Tenthredinidae), with notes on their phylogeny. *European Journal of Entomology*, **101**, 637–650.
- Hendrich L, Balke M, Haszprunar G *et al.* (2010) Barcoding Fauna Bavarica—capturing Central European animal diversity. In: *Proceedings of the International Congress Tools for Identifying Biodiversity: Progress and Problems* (eds Nimis P., Vignes Lebbe R.), p. 347. EUT Edizioni Università di Trieste, Trieste.
- Hill DS (1983) *Agricultural Insect Pests of the Tropics and Their Control*. Cambridge University Press, Cambridge.
- Hill DS (1987) *Agricultural Insect Pests of Temperate Regions and Their Control*. Cambridge University Press, Cambridge.
- Huemer P, Mutanen M, Sefc KM, Hebert PDN (2014) Testing DNA barcode performance in 1000 species of European Lepidoptera: large geographic distances have small genetic impacts. *PLoS ONE*, **9**, 1–21.
- Ivanova NV, Dewaard JR, Hebert PDN (2006) An inexpensive, automation-friendly protocol for recovering high-quality DNA: technical note. *Molecular Ecology Notes*, **6**, 998–1002.
- Jansen E (1998) Die Gattung *Hartigia* Schiödte, 1838 in Europa (Hymenoptera: Cephidae). In: *Pflanzenwespen Deutschlands (Hymenoptera, Symphyta). Kommentierte Bestandsaufnahme* (eds Taeger A., Blank S. M.), pp. 301–318. Goecke & Evers, Keltern.
- Koch WR, Smith DR (2000) *Nematus oligospilus* Förster (Hymenoptera: Tenthredinidae), an introduced willow sawfly in the Southern Hemisphere. *Proceedings of the Entomological Society of Washington*, **102**, 292–300.
- Kruse J, Zogas K, Hard J, Llsuzzo N (2010) *New Pest in Alaska and Washington. The Green Alder Sawfly—Monsoma pulveratum* (Retzius). *Pest Alert*. United States Department of Agriculture, Forest Service, State and Private Forestry, Alaska Region, **R10-PR-022**.
- Lacourt J (1998) *Endelomyia filipendulae*, nouvelle espèce du sud de la France (Hymenoptera, Tenthredinidae). *Bulletin de la Société entomologique de France*, **103**, 393–394.
- Leppänen SA, Malm T, Värrä K, Nyman T (2014) A comparative analysis of genetic differentiation across six shared willow host species in leaf- and bud-galling sawflies. *PLoS ONE*, **9**, 1–19.
- Lindqvist E (1955) Beitrag zur Kenntnis einiger nordischer Blattwespen (Hym., Tenthredinoidea). *Notulae entomologicae*, **35**, 137–144.
- Linnen CR, Farrell BD (2007) Mitonuclear discordance is caused by rampant mitochondrial introgression in *Neodiprion* (Hymenoptera: Diprionidae) sawflies. *Evolution*, **61**, 1417–1438.
- Linnen C, Farrell B (2008) Comparison of methods for species-tree Inference in the sawfly genus *Neodiprion* (Hymenoptera: Diprionidae). *Systematic Biology*, **57**, 876–890.
- Liston AD (1993) A new European *Parna*, *reseri* n. sp. (Hymenoptera, Symphyta, Tenthredinidae). *Entomologische Berichte Luzern*, **29**, 73–76.
- Liston AD (2012) The identity of a stem sawfly recorded as a pest of *Rosa damascena* in Iran. *Archives of Phytopathology and Plant Protection*, **45**, 1039–1041.
- Liston AD (2015) New records and host plants of Symphyta (Hymenoptera) for Germany, Berlin and Brandenburg. *Beiträge zur Entomologie*, **65**, 383–391.
- Liston AD, Jacobs H-J (2012) Review of the sawfly fauna of Cyprus, with descriptions of two new species. *Zoology in the Middle East*, **56**, 67–84.
- Liston A, Prous M (2014) Sawfly taxa (Hymenoptera, Symphyta) described by Edward Newman and Charles Healy. *ZooKeys*, **398**, 83–98.
- Liston AD, Späth J (2011) Aspects of the biology of *Megalodontes thor* Taeger, 2002 (Hymenoptera: Symphyta: Megalodontesidae). *Beiträge zur Entomologie*, **61**, 259–270.
- Liston AD, Taeger A, Blank SM (2006) Comments on European Sawflies (Hymenoptera: Symphyta). In: *Recent Sawfly Research: Synthesis and Prospects* (eds Blank S. M., Schmidt S., Taeger A.), pp. 245–263. Goecke & Evers, Keltern.
- Lorenz H, Kraus M (1957) Die Larvalsystematik der Blattwespen (Tenthredinoidea und Megalodontoidea). *Abhandlungen zur Larvalsystematik der Insekten*, **1**, i–ix, 1–339.

- Macdonald J, Ohmart CP (1993) Life history strategies of Australian pergid sawflies and their interactions with host plants. *Sawfly Life History Adaptations to Woody Plants*, pp. 485–502. Academic Press, San Diego.
- Macek J (2013) Descriptions of larvae of *Birka annulitarsis* and *B. cinereipes* (Hymenoptera: Symphyta: Tenthredinidae). *Acta Entomologica Musei Nationalis Pragae*, **53**, 815–819.
- Malm T, Nyman T (2015) Phylogeny of the symphytan grade of Hymenoptera: new pieces into the old jigsaw (fly) puzzle. *Cladistics*, **31**, 1–17.
- Martelli M (1941) Comportamenti biologici e danni dell'*Emphytus cinctus* L. (Hymenoptera Tenthredinidae) sulla Fragola nell'Emilia. *Bolletino dell' Istituto di Entomologia della R. Università degli Studi di Bologna*, **12**, 171–178.
- Mardulyn P, Othmezzouri N, Mikhailov YE, Pasteels JM (2011) Conflicting mitochondrial and nuclear phylogeographic signals and evolution of host-plant shifts in the boreo-montane leaf beetle *Chrysomela lapponica*. *Molecular Phylogenetics and Evolution*, **61**, 686–696.
- Miller RH, Masri SE, Jundi KA (1993) Plant density and wheat stem sawfly (Hymenoptera: Cephidae) resistance in Syrian wheats. *Bulletin of Entomological Research*, **83**, 95–102.
- Mutanen M, Kivelä SM, Vos RA *et al.* (2016) Species-level para- and poly-phyly in DNA barcode gene trees: strong operational bias in European Lepidoptera. *Systematic Biology*, **65**, 1024–1040.
- Nikolova V, Natskova V (1970) The rose stem borer—a dangerous pest of young and rejuvenated rose gardens. *Rastitelna Zashchita*, **18**, 31–34.
- Nyman T, Zinovjev AG, Vikberg V, Farrell BD (2006) Molecular phylogeny of the sawfly subfamily Nematinae (Hymenoptera: Tenthredinidae). *Systematic Entomology*, **31**, 569–583.
- Nyman T, Leppänen SA, Várkonyi G *et al.* (2015) Determinants of parasitoid communities of willow-galling sawflies: habitat overrides physiology, host plant and space. *Molecular Ecology*, **24**, 5059–5074.
- Opitz SEW, Mix A, Winde IB, Müller C (2011) Desulfation followed by sulfation: metabolism of benzylglucosinolate in *Athalia rosae* (Hymenoptera: Tenthredinidae). *ChemBioChem*, **12**, 1252–1257.
- Özbek H (2014) Ichneumonid parasitoids of the sawfly *Cimbex quadrimaculata* (Müller) feeding on almonds in Antalya, along with a new parasitoid and new record. *Turkish Journal of Zoology*, **38**, 657–659.
- Patten MM, Carioscia SA, Linnen CR (2015) Biased introgression of mitochondrial and nuclear genes: a comparison of diploid and haplodiploid systems. *Molecular Ecology*, **24**, 5200–5210.
- Pesarini F (2002) Contributo alla conoscenza dei sinfiti della regione balcanico-egaea (Hymenoptera, Symphyta). *Bolletino del Museo Regionale di Scienze Naturali*, **19**, 121–183.
- Prieto JM, Schaffner U, Barker A *et al.* (2007) Sequestration of furostanol saponins by *Monophadnus* sawfly larvae. *Journal of Chemical Ecology*, **33**, 513–524.
- Prous M, Heidema M, Soon V (2011) *Empria longicornis* species group: taxonomic revision with notes on phylogeny and ecology (Hymenoptera, Tenthredinidae). *Zootaxa*, **2756**, 1–39.
- Prous M, Blank SM, Goulet H *et al.* (2014) The genera of Nematinae (Hymenoptera, Tenthredinidae). *Journal of Hymenoptera Research*, **40**, 1–69.
- Pschorn-Walcher H (1982) Unterordnung Symphyta, Pflanzenwespen. In: *Die Forstschädlinge Europas* (ed Schwenke W.), pp. 4–196, 232–234. Parey, Hamburg, Berlin.
- Quicke DLJ (2012) We know too little about parasitoid wasp distributions to draw any conclusions about latitudinal trends in species richness, body size and biology (R DeSalle, Ed.). *PLoS ONE*, **7**, 1–9.
- Ratnasingham S, Hebert PDN (2013) A DNA-based registry for all animal species: the barcode index number (BIN) system. *PLoS ONE*, **8**, 1–16.
- Raychoudhury R, Baldo L, Oliveira DCSG, Werren JH (2009) Modes of acquisition of *Wolbachia*: Horizontal transfer, hybrid introgression, and codivergence in the *Nasonia* species complex. *Evolution*, **63**, 165–183.
- Ronquist F, Klopstein S, Vilhelmsen L *et al.* (2012) A total-evidence approach to dating with fossils, applied to the early radiation of the hymenoptera. *Systematic Biology*, **61**, 973–999.
- Schaffellner C, Schopf A (2014) Massenaufreten der Fichtengebirgsblattwespe in Tieflagen als Folge des Klimawandels? *Forstschutz Aktuell*, **60**(61), 12–19.
- Schedl W, Pschorn-Walcher H (1984) Ein Beitrag zu schwarzen Bürstenhornblattwespen aus Mitteleuropa (Insecta: Hymenoptera, Argidae). *Berichte des Naturwissenschaftlich-Medizinischen Vereins in Innsbruck*, **71**, 173–179.
- Scheibelreiter G (1973) Die Tenthrediniden der Rose (*Rosa spec.*). *Zeitschrift für Angewandte Entomologie*, **72**, 225–259.
- Scheibelreiter GK (1978) The poppy-cephid, *Pachycephus smyrnensis* Stein (Hymenoptera: Cephidae). *Zeitschrift für Angewandte Entomologie*, **86**, 19–25.
- Schiff NM, Goulet H, Smith DR *et al.* (2012) Siricidae (Hymenoptera: Symphyta: Siricoidea) of the Western Hemisphere. *Canadian Journal of Arthropod Identification*, **21**, 1–305.
- Schmidt S, Smith DR (2006) An annotated systematic world catalogue of the Pergidae (Hymenoptera). *Contributions of the American Entomological Institute*, **34**, 1–207.
- Schmidt S, Smith DR (2009) Selandriinae, a subfamily of Tenthredinidae new to Australia, and a review of other Australian Tenthredinidae (Hymenoptera: Symphyta). *Australian Journal of Entomology*, **48**, 305–309.
- Schmidt S, Smith DR (2016) Pergidae of the World - An online catalogue of the sawfly family Pergidae (Hymenoptera, Symphyta). *World Wide Web electronic publication*. <http://pergidae.hymis-zsm.de>.
- Schmidt S, Walter GH (2014) Young clades in an old family: major evolutionary transitions and diversification of the eucalypt-feeding pergid sawflies in Australia (Insecta, Hymenoptera, Pergidae). *Molecular Phylogenetics and Evolution*, **74**, 111–121.
- Schmidt S, Walter GH, Grigg J, Moore CJ (2006) Sexual communication and host plant associations of Australian Pergid Sawflies (Hymenoptera, Symphyta, Pergidae). In: *Recent Sawfly Research: Synthesis and Prospects* (eds Blank S. M., Schmidt S., Taeger A.), pp. 173–193. Goecke & Evers, Keltern.
- Schulmeister S (2003a) Simultaneous analysis of basal Hymenoptera (Insecta): introducing robust-choice sensitivity analysis. *Biological Journal of the Linnean Society*, **79**, 245–275.
- Schulmeister S (2003b) Review of morphological evidence on the phylogeny of basal Hymenoptera (Insecta), with a discussion of the ordering of characters. *Biological Journal of the Linnean Society*, **79**, 209–243.
- Sharkey MJ, Carpenter JM, Vilhelmsen L *et al.* (2012) Phylogenetic relationships among superfamilies of Hymenoptera. *Cladistics*, **28**, 80–112.
- Smith DR (1989) The Sawfly Genus *Arge* (Hymenoptera: Argidae) in the Western Hemisphere. *Transactions of the American Entomological Society* (1890–), **115**, 83–205.
- Smith DR, Schmidt S (2009) A new subfamily, genus, and species of Cephidae (Hymenoptera) from Australia. *Zootaxa*, **60**, 56–60.
- Taeger A (1985) Zur Systematik der Blattwespengattung *Tenthredo* (s. str.) L. (Hymenoptera, Symphyta, Tenthredinidae). *Entomologische Abhandlungen*, **48**, 83–148.
- Taeger A (1988) Dritter Beitrag zur Kenntnis der Blattwespengattung *Tenthredo* L. (Hymenoptera: Symphyta: Tenthredinidae). *Beiträge zur Entomologie*, **38**, 337–359.
- Taeger A (1998a) Die Megalodontesidae Europas (Hymenoptera: Symphyta). In: *Pflanzenwespen Deutschlands* (Hymenoptera, Symphyta). *Kommentierte Bestandsaufnahme* (eds Taeger A., Blank S. M.), pp. 175–192. Goecke & Evers, Keltern.
- Taeger A (1998b) Bestimmungsschlüssel der Keulhornblattwespen Deutschlands (Hymenoptera: Cimbicidae). In: *Pflanzenwespen Deutschlands* (Hymenoptera, Symphyta). *Kommentierte Bestandsaufnahme* (eds Taeger A., Blank S. M.), pp. 193–205. Goecke & Evers, Keltern.
- Taeger A (2002) The Megalodontesidae of Europe (Hymenoptera, Symphyta). In: *Sawflies I. A Review of the Suborder, the Western Palaearctic Taxa of Xyeloidea and Pamphilioidea* (ed Viitasaari M.), pp. 461–480. Tremex, Helsinki.
- Taeger A, Blank SM (1998) Beitrag zur Kenntnis einiger Nematinae (Hymenoptera: Tenthredinidae). In: *Pflanzenwespen Deutschlands* (Hymenoptera, Symphyta). *Kommentierte Bestandsaufnahme* (eds Taeger A., Blank S. M.), pp. 247–277. Goecke & Evers, Keltern.



- Taeger A, Viitasaari M (2015) European *Rhogogaster* s. str., with notes on several Asian species (Hymenoptera: Tenthredinidae). *Zootaxa*, **4013**, 369–398.
- Taeger A, Altenhofer E, Blank SM *et al.* (1998) Kommentare zur Biologie, Verbreitung und Gefährdung der Pflanzenwespen Deutschlands (Hymenoptera, Symphyta). In: *Pflanzenwespen Deutschlands* (Hymenoptera, Symphyta). *Kommentierte Bestandsaufnahme* (eds Taeger A., Blank S. M.), pp. 49–135. Goecke & Evers, Keltern.
- Taeger A, Blank SM, Liston AD (2010) World catalog of symphyta (Hymenoptera). *Zootaxa*, **1064**, 1–1064.
- Thamsborg SM, Jørgensen RJ, Brummerstedt E (1987) Sawfly poisoning in sheep and goats. *The Veterinary Record*, **121**, 253–255.
- Viitasaari M (2002) The Northern European taxa of Pamphiliidae (Hymenoptera). In: *Sawflies I. A Review of the Suborder, the Western Palaearctic Taxa of Xyeloidae and Pamphilioidae* (ed Viitasaari M.), pp. 235–358. Tremex, Helsinki.
- Vikberg V (2004) Seasonal head dimorphism and taxonomy of some European species of *Aprostema* (Hymenoptera: Symphyta: Argidae). *Beiträge zur Entomologie*, **54**, 107–125.
- Vikberg V (2010) European species of *Tubpontania* gen. nov., a new genus for species of the *Pontania crassispina* group. *Zootaxa*, **2620**, 1–28.
- Vikberg V, Liston A (2009) Taxonomy and biology of European Heptamelini (Hymenoptera, Tenthredinidae, Selandriinae). *Zootaxa*, **2112**, 1–24.
- Vikberg V, Nuorteva M (1997) On the rearing of *Nesoselandria morio* (Fabricius) and *Birka cinereipes* (Klug) (Hymenoptera, Tenthredinidae), with descriptions of their larvae. *Entomologica Fennica*, **8**, 27–38.
- Wilhelmsen L, Blank SM, Costa VA, Alvarenga TM, Smith DR (2013) Phylogeny of the ophrynopine clade revisited: review of the parasitoid sawfly genera *Ophrella* Middlekauff, *Ophrynopus* Konow and *Stirocorsia* Konow (Hymenoptera: Orussidae). *Invertebrate Systematics*, **27**, 1–9.
- deWaard JR, Ivanova NV, Hajibabaei M, Hebert PDN (2008) Assembling DNA Barcodes. In: *Environmental Genomics* (ed Martin CC), pp. 275–294. Humana Press, Totowa, NJ.
- Wei M (2001) Revision of Cladiini (Hymenoptera: Tenthredinomorpha: Nematidae) From China (I): Notes on Cladiini and *Driocampus* Zhang With Revision of *Cladius* Illiger from China. *Journal of Jishou University (Natural Science Edition)*, **22**, 34–39.
- Whitworth TL, Dawson RD, Magalon H, Baudry E (2007) DNA barcoding cannot reliably identify species of the blowfly genus *Protophormia* (Diptera: Calliphoridae). *Proceedings of the Royal Society B-Biological Sciences*, **274**, 1731–1739.
- Xiao J-H, Wang N-X, Murphy RW *et al.* (2012) *Wolbachia* infection and dramatic intraspecific mitochondrial DNA divergence in a fig wasp. *Evolution; International Journal of Organic Evolution*, **66**, 1907–1916.
- Zhelochovtsev AN (1952) Review of the sawflies of the subfamily Cladiinae (Hymenoptera, Tenthredinidae) of the fauna of the USSR. *Zoologicheskii Zhurnal*, **31**, 257–269.
- Zhelochovtsev AN (1968) New species of Symphyta (Hymenoptera) of the fauna of the USSR. *Sbornik trudov Zoologicheskogo Muzeja MGU*, **11**, 47–56.
- Zhelochovtsev AN, Zinovjev AG (1994) Hymenoptera. Part VI. Symphyta. In: *Keys to the Insects of the European Part of the USSR*, vol. III (ed Medvedev G. S.), pp. 1–432. Oxonian Press, New Delhi.
- Zinovjev AG, Vikberg V (1999) The sawflies of the *Pontania crassispina* group with a key for the genera of the subtribe Eurina (Hymenoptera: Tenthredinidae, Nematinae). *Entomologica Scandinavica*, **30**, 281–298.

S.S. and A.T. coordinated the work of the other co-authors and edited the final draft of the manuscript including the supporting information. S.S., A.T., A.L.,

S.B., E.H. and M.K. identified or checked the identification of specimens. A.T. and O.S. organized photography and data entry. J.S. was initially responsible for project management in BOLD. Barcoding statistics were prepared by J.M. and S.S.. Main responsibility for text sections on families/subfamilies is Allantinae (M.P.), Argidae (S.S.), Blasticotomidae (S.S.), Blennocampinae and Heterarthrinae (A.L.), Cephidae (S.S., S.B.), Cimbicidae (S.S., A.T.), Diprionidae (S.S.), Megalodontesidae (A.T.), Nematinae (A.L.), Orussidae (S.S.), Pamphiliidae (S.S., A.T.), Pergidae (S.S.), Selandriinae (S.S., A.T.), Sircidae and Xiphydriidae (S.S.), Tenthredininae (A.T., S.S.), Xyelidae (S.B.). T.N., T.M. and K.K. advised on technical and theoretical aspects of barcoding and the interpretation of results. All authors contributed improvements during the final stages of manuscript preparation.

### Data accessibility

All specimen data are accessible on BOLD (www.boldsystems.org) through the following doi: 10.5883/DS-RSYM. The data include collection locality, geographic coordinates, altitude, collector, one or more images, identifier and voucher depository. Sequence data are available on BOLD and include a detailed LIMS report, primer information and trace files, and sequences have been deposited also in GenBank.

### Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Appendix S1** List of voucher specimens with country of origin, specimen ID, depository, GenBank accession code, sequencing success (COI-fragment in bp, in square brackets number of unresolved bases), and, if present, Binary Index Number (BIN).

**Appendix S2** Classification of the suborder Symphyta in BOLD at the time of the release and current classification (differences in bold, for details see Materials and Methods section (differences in bold).

**Appendix S3** Barcoding statistics of Symphyta species with distribution, Barcode Index Number (BIN), number of specimens, mean intraspecific distance, maximum intraspecific distance, nearest neighbour species and distance to nearest species.

**Appendix S4** List of species that share the same Barcode Index Number (BIN).

**Appendix S5** Neighbour-joining tree of sequences  $\geq 500$  bp and less the 1% ambiguous bases, or, for sequences  $< 500$  bp, if the sequence was assigned to a BIN.