Received: 17 August 2023 | Accepted: 23 November 2023

DOI: 10.1111/syen.12618

ORIGINAL ARTICLE

Systematic Entomology

Royal Entomological Society

Phylogenomic species delimitation of the twisted-winged parasite genus *Stylops* (Strepsiptera)

Meri Lähteenaro[1,2] | Jakub Straka[3] | Mattias Forshage[1] | Rasmus Hovmöller[1,4] | Yuta Nakase[5] | Anders L. Nilsson[6] | John T. Smit[7,8] | Johan A. A. Nylander[9] | Johannes Bergsten[1,2]

[1]Department of Zoology, Swedish Museum of Natural History, Stockholm, Sweden

[2]Department of Zoology, Faculty of Science, Stockholm University, Stockholm, Sweden

[3]Department of Zoology, Faculty of Science, Charles University, Prague, Czech Republic

[4]SLU Swedish Species Information Centre, Uppsala, Sweden

[5]Faculty of the Arts, Kyoto University of the Arts, Kyoto, Japan

[6]Museum of Evolution, Uppsala University, Uppsala, Sweden

[7]Naturalis Biodiversity Center, Leiden, The Netherlands

[8]European Invertebrate Survey, Leiden, The Netherlands

[9]Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden

Correspondence
Meri Lähteenaro, Department of Zoology, Swedish Museum of Natural History, Box 50007, SE-104 05 Stockholm, Sweden.
Email: meri.lahteenaro@nrm.se

Funding information
Czech Science Foundation, Grant/Award Number: 20-14872S; The Swedish Taxonomy Initiative, Grant/Award Numbers: 2019.4.3-7, 2019.4.3-218

Abstract

The twisted-winged parasite genus *Stylops* has a history of different species concepts with varying host specificity resulting in diverse species diversity estimates in different regions of the Holarctic. The adoption of a supergeneralist species concept in Europe, proposing synonymization of all Western Palaearctic *Stylops* species, did not facilitate taxonomic clarity and obscured the available life-history data in the region for decades. Lack of molecular data has allowed divergent opinions on species hypotheses and little opportunity for evaluating them in this morphologically challenging genus. To solve these discrepancies and gain novel information about host associations, we applied whole-genome sequencing to 163 specimens, representing a significant portion of putative European species. We evaluate the existing and conflicting species hypotheses with molecular species delimitation using Species bOundry Delimitation using Astral (SODA) and use a maximum likelihood phylogeny to investigate host associations of the species. Furthermore, we evaluate the effect of a number of loci used in SODA for the number of inferred species. We find justification for synonymization of multiple species and indications of undescribed species, as well as new host–parasite relationships. We show that the number of inferred species in SODA is exceedingly and positively correlated with the number of loci used, urging for cautious application. The results of our study bring clarity to the Western Palaearctic species diversity of *Stylops*. Furthermore, the comprehensive molecular dataset generated in this study will be a valuable resource for future studies on *Stylops* and the evolution of parasites in general.

KEYWORDS
phylogenomics, SODA, species delimitation, Strepsiptera, Stylops, whole-genome sequencing

INTRODUCTION

Although parasites represent a considerable portion of global biodiversity, the diversity of many parasite groups is still unknown (Nadler & Pérez-Ponce de León, 2011). For instance, based on extrapolation from case studies of genus-specific parasite–host ratios, it appears that parasitic wasps would by conservative measures be two to three times more diverse than beetles, the current most species-rich order based on described species (Forbes et al., 2018). Furthermore, numerous studies have shown that cryptic species complexes exist in many parasite groups (e.g., Heraty et al., 2007; Pérez-Ponce de León & Nadler, 2016). Cryptic species do not always have enough interspecific morphological variation

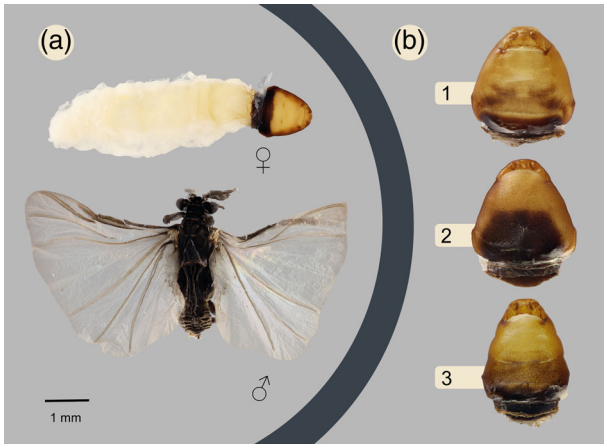

FIGURE 1 Representative adults of *Stylops*. (A) High sexual dimorphism in *Stylops ater* resulting from adult males (bottom) being free-living in contrast to adult females (top) remaining as endoparasites in the host (B) Morphological variation present in adult female cephalothoraxes. Despite the reduced morphological characters, interspecific variation is present among some *Stylops* species, as demonstrated by (1) *Stylops nassonowi* Pierce, 1909 (2) *Stylops thwaitesi* Perkins, 1918 and (3) *Stylops praecocis* Luna de Carvalho, 1974. Scale bar applies to A only.

for inferring species limits based on morphology alone. When this is the case, molecular data can help with inferring species limits (Cerca et al., 2020; Sudasinghe et al., 2020; Williams et al., 2022). In parasites, host usage is another attribute with taxonomic value. Parasites can be broadly classified into generalists and specialists, depending on their host range. The assumed host specificity can have vast implications on the estimated species diversity (Smith et al., 2007, 2008). Perhaps one of the best examples of this can be found in the insect order twisted-winged parasites (Strepsiptera). Strepsiptera is a small order of obligate entomophagous endoparasites, characterized by extreme sexual dimorphism and complicated life-cycles (Kathirithamby, 2018). The winged adult males are free-living and possess typical adult insect characteristics. The larviform adult females continue as endoparasites in the host (except in the family Mengenillidae) and remain in a structure comprised of puparial exuvia and exuvia of the second larval stage with highly reduced features (Löwe et al., 2016). Currently, Strepsiptera has less than 630 recognised species worldwide (Cook, 2019). However, multiple molecular studies have suggested that this is an underestimation due to cryptic species complexes (Benda et al., 2020; Hayward et al., 2011; Jůzová et al., 2015; Nakase & Kato, 2013).

The Holarctic genus *Stylops* Kirby is the most species-rich genus of Strepsiptera, whose members are parasites of *Andrena* Fabricius mining bees (Kathirithamby, 2018; Kathirithamby & Engel, 2014; Kinzelbach, 1971). They exhibit the typical sexual dimorphism present in Strepsiptera (Figure 1). Since the short-lived adult males are rarely encountered and the females are difficult to identify morphologically, host specificity has played a major role in the classification and estimated species diversity of *Stylops* (Straka et al., 2015). Over the years, species classifications have varied dramatically between different taxonomists from assumptions of highly specialist species with only single host associations (e.g., Kifune & Hirashima, 1985; Perkins, 1918; Pierce, 1911) to a supergeneralist species with more than 150 host associations (Kinzelbach, 1978). The supergeneralist classification strategy was in use for decades in the Western Palearctic (WP) region after a highly influential work synonymized all European species into a single species, *Stylops melittae* Kirby (Kinzelbach, 1978). Kinzelbach concluded that the data available at the time were not sufficient to divide WP *Stylops* into several natural species and should be instead treated as one species with multiple subspecies. This classification was recently challenged in a molecular study by Jůzová et al. (2015). They used two mitochondrial and one nuclear gene fragment to test different hypotheses of host specialisation in *Stylops*. Their results indicated that host specificity is prevalent among *Stylops*, although not always to a single species but typically to a few closely related host species, often belonging to the same subgenus of *Andrena*. They concluded that *S. melittae* is likely a species complex. Based on these findings, Straka et al. (2015) constructed a preliminary world checklist for the genus, where 30 species names in Europe were reinstated, largely based on hosts and DNA barcode distances (mitochondrial COI). Additional support came from a recent interspecific mating experiment study that included three species from the WP region (Jandausch et al., 2022). This study found that interspecific variation in female paragenital organs was correlated with male penis shape, allowing only conspecific matings to succeed.

However, there is still considerable uncertainty over the diversity within the genus *Stylops* and their host relationships. Jůzová et al. (2015) did not use any quantitative species delimitation method and distances between DNA barcodes alone are often insufficient for reliable species delimitation (Brower, 2006; DeSalle et al., 2005; Ortiz et al., 2021). Furthermore, in a recent World Catalogue of the order Strepsiptera, Cook (2019) reinstated multiple species synonymized by Straka et al. (2015) and the host associations presented in the catalogue differ from those listed by Straka et al. (2015) (Figure 2). There are also multiple *Andrena* species that are known to be parasitized by *Stylops* but are not linked to any specific *Stylops* species in either the World Catalogue or in the preliminary world checklist.

Here, we attempt to resolve the controversies in the species diversity estimates of WP *Stylops* with the use of phylogenomic species delimitation methods. We use whole-genome sequencing to generate a comprehensive dataset to test the hypothesised species boundaries in the *Stylops melittae* species complex. We evaluate reported and suggested host associations within each putative species, as well as present novel knowledge of unpublished host–parasite relationships. Finally, the challenges with molecular species delimitation using few genes up to genomic-size datasets are evaluated with the novel species delimitation method, Species bOundry Delimitation using Astral (SODA) (Rabiee & Mirarab, 2021).

MATERIALS AND METHODS

Taxon sampling

We obtained genomic data from 163 female individuals representing 22 putative species included in the preliminary world checklist of the

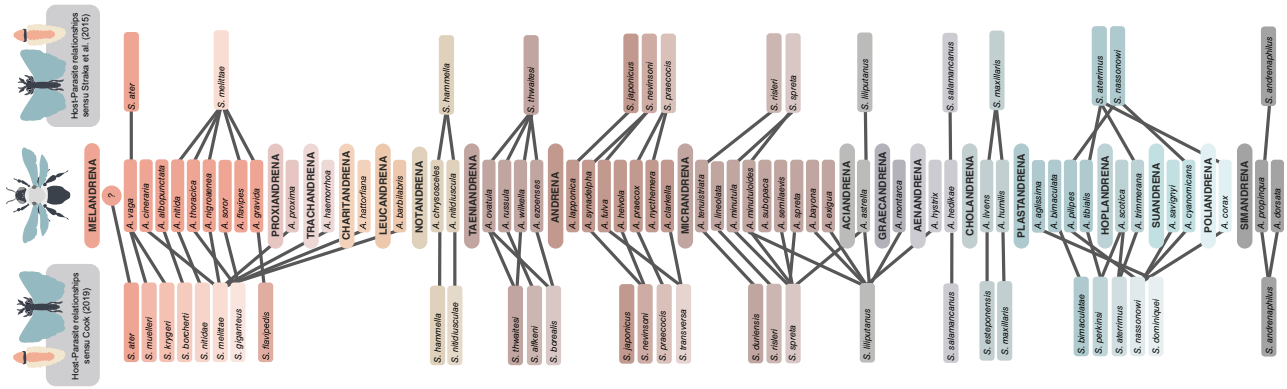

FIGURE 2 Discrepancies in Western Palaearctic *Stylops* taxonomy and host associations between the recent World Catalogue of the Order Strepsiptera (left) (Cook, 2019) and the Preliminary World Checklist of *Stylops* (right) (Straka et al., 2015). *Andrena* mining bee host species and their subgenera (following Kuhlmann et al. (2023) and Pisanty et al. (2022)) are listed in the middle. Host associations are represented with black lines.

genus *Stylops* (Straka et. al., 2015). Preliminary species identification of the specimens was largely based on host associations. If the host species of a specimen was not associated with any putative *Stylops* species, the host subgenus and morphology of the female were used to assign them to the most likely species. The classification of *Andrena* into subgenera follows Kuhlmann et al. (2023) and Pisanty et al. (2022). A total of 39 samples could not be assigned to any species. Each of the included putative species was represented by 1–17 individuals. We focused the geographical sampling on the WP region but included samples from the Eastern Palaearctic (EP) region as well. The material contained both freshly collected material stored in ethanol and pinned museum specimens. In addition, some DNA extracts were reutilised from the study of Jůzová et al. (2015). The specimens came from both natural history museum collections and private collectors. All material included in this study is listed in Table S1. We included *Halictoxenos tumulorum* Perkins, representing another genus in Stylopidae, as an outgroup.

Molecular methods and phylogenetic inference

The age and preservation method of included samples were taken into account in the data acquisition and processing pipeline (Figure 3). All *Stylops* females were removed from the host bees using forceps cleaned with chlorine, water and 80% EtOH. For the pinned specimens, the bees were relaxed in water vapour for 30 min before removal. The entire female body was used for genomic DNA extraction with QIAamp DNA Micro kit (Qiagen, Inc). After extraction, the intact cephalothorax was recovered as a morphological voucher. The manufacturer's protocol was followed, except 20 μL of DTT (1 M) was added to samples during the lysis step and lysed samples were kept 10 min at 72°C after adding Buffer AL. DNA concentrations were quantified with a Qubit fluorometer (HS Assay Kit, Life Technologies Inc.). Illumina libraries were prepared for old museum samples and samples with low DNA concentrations (a total of 43), following the museomics protocol by Irestedt et al. (2022). Equimolarly pooled libraries were sent to Science for Life Laboratory (SciLifeLab) in Stockholm, National Genomics Infrastructure Sweden, for paired-end sequencing (read length: 100 bp) using Illumina NovaSeq sequencing platform (Illumina Inc.). For the remaining 124 samples, Illumina Nextera DNA Flex libraries were constructed at SciLifeLab Stockholm, and they were paired-end sequenced (read length: 150 bp) using Illumina NovaSeq sequencing platform.

Nf-core/eager v2.4.0 (Yates et al., 2021) pipeline was used to pre-process raw sequence data and create sequence assemblies. Within the pipeline, adapter clipping was performed with AdapterRemoval v2.3.2 (Schubert et al., 2016), mapping against a *Stylops ater* reference genome (Podsiadlowski et al., unpubl.) with BWA mem v0.7.17 (Li & Durbin, 2009) and variant calling with angsd v0.935 (Korneliussen et al., 2014) with default settings of the pipeline. Possible contaminants (e.g., from a host) were removed by the mapping procedure. A complexity filtering step was included for the Illumina two-colour chemistry data to trim the poly-G tails from the short fragments. The quality of the assemblies was assessed with MultiQC v1.11 (Ewels et al. 2016) and with BUSCO v.5.0.0 (Manni et al. 2021) against endopterygota_odb10-dataset. Assemblies were searched for 3913 orthologous nuclear genes of Strepsiptera taxa included in the dataset of McKenna et al. (2019): *Triozocera* (Corioxenidae), *Mengenilla* (Mengenillidae), *Xenos* (Xenidae) and *Stylops* (Stylopidae). After extracting the genes of the four taxa from the McKenna et al. dataset with grepfasta (https://github.com/nylander/grepfasta), gaps and sequences shorter than 100 bp were removed from the gene sequences with fastagap (https://github.com/nylander/fastagap). Extracted genes were then used as nucleotide baits against nucleotide assemblies, only extracting the single best hit regions with the programme Alibaseq v1.2 (Knyshov et al., 2021).

Multiple sequence alignment (MSA) was performed with MAFFT v.7.490 (Katoh et al., 2002; Katoh & Standley, 2013) for the obtained genes, and a sanity check was conducted for the produced alignments with RaxML-NG (Kozlov et al., 2019). Phylogenetically informative regions were selected from the MSA with BMGE (Criscuolo & Gribaldo, 2010). Because of the large size of our dataset, ParGenes (Morel et al., 2019) was used for model selection and gene tree

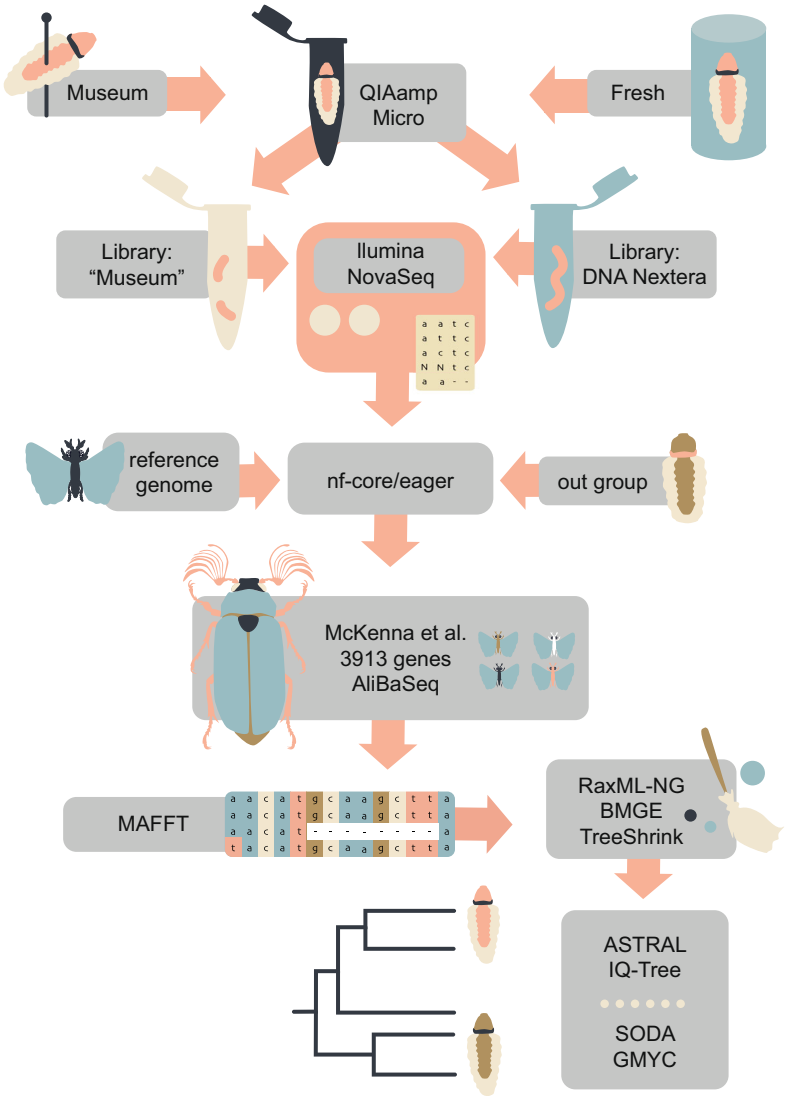

FIGURE 3 An overview of the data acquisition, data processing and analyses steps used.

inference. The programme allows both analyses to be done in parallel for thousands of gene MSAs. First ParGenes was used to infer maximum likelihood (ML) trees using RAxML-NG with a fixed model GTR + G8 + F, and outlier branches were filtered out with TreeShrink (Mai & Mirarab, 2018). The filtered data were then re-aligned and ParGenes ran again. This time, a model test was included, where the best-fit model of evolution for each gene MSA was selected based on the Bayesian information criterion. Resulting gene trees were used for subsequent inferences. ASTRAL-III v.5.6.3 (Zhang et al., 2018) was used to infer a species tree from the gene trees. Phylogeny under a concatenated ML approach was inferred with IQ-TREE2 v.2.2.0 (Minh et al., 2020) including branch support estimation with ultrafast bootstrap (−bb 1000) (Hoang et al., 2018). Model selection was done with ModelFinder (Kalyaanamoorthy et al., 2017), followed by tree construction (Nguyen et al., 2015) using the estimated best partitioning scheme (−m MFP + MERGE). To lower the computational load, a relaxed clustering algorithm was used (−rcluster 10) (Lanfear et al., 2014). First, all sequences were included to place critical taxa. We then removed terminals with eight or more times longer branches than adjacent terminals and alignments that were shorter than 100 bp or had less than four sequences to avoid spurious effects in the phylogenetic estimation and species delimitation. The remaining loci were then used in subsequent analyses. The ML inference was run a total of six times using the best model to find the optimal tree with highest log-likelihood. The produced trees did not have major topological differences (not shown). Furthermore, we repeated the ML tree inference using the second and third best models to test inference sensitivity to the selected model. These produced nearly identical trees to the best model (not shown). The generated species trees were visualised with FigTree v.1.4.4 (Rambaut, 2018).

Species delimitation analyses were conducted with SODA (Rabiee & Mirarab, 2021). SODA is a computationally powerful topology-based method that can process phylogenomic datasets rapidly under the multispecies coalescent (MSC). Simulations have shown that SODA's performance is close to popular methods such as Bayesian Phylogenetics and Phylogeography (BPP) (Yang & Rannala, 2010),

but computationally less demanding, which allows analysing the entire dataset without subsampling (Rabiee & Mirarab, 2021). We ran SODA using gene trees from RAxML-NG as input and a *p* value cut-off set to 0.05 (default in SODA). We let SODA infer the guide tree using ASTRAL-III. However, empirical data can have properties, which make MSC-methods less accurate. Other MSC-based species delimitation methods, such as BPP (Yang, 2015; Yang & Rannala, 2010), have been shown to be sensitive to population structure and hence over-splitting species, especially when analysing multiple loci from allopatric populations (Leaché et al., 2019; Sukumaran & Knowles, 2017). Since the cut-off value impacts how easily species are split, SODA was run again with a higher significance level of 0.001. Both analyses were repeated to confirm consistency between runs. Note that populations were not defined a priori; hence, each individual is a candidate population and species. Unlike BPP (Huang, 2018), how the number of loci affects the number of species delimited in SODA has not been previously studied with an empirical dataset. To further investigate the effect of number of loci, we subsampled sets of randomly selected loci (25, 50,100, 250, 500, 1000, 1500 and 2000). Ten replicates were produced for each set and SODA was run for each replicate with cut-off value 0.001. To further examine the sensitivity of SODA for over-splitting, we compared the results of SODA to the results of the generalised mixed Yule coalescent (GMYC) species delimitation method (Fujisawa & Barraclough, 2013; Pons et al., 2006). The GMYC method uses transitions between distinct branching patterns in an ultrametric gene tree to delimit species (Fujita et al., 2012). Although strictly appropriate only for single locus gene trees, the GMYC method is often used on trees from concatenated loci as well by postulating a shared genealogical history (Luo et al., 2018). In this respect, because GMYC looks for patterns in a single input tree, it should be robust against the particular phenomena of over-splitting due to a high number of loci included in the analyses. The ML-tree from IQ-TREE2 was converted into an ultrametric tree using chronos function (calibration = makeChronosCalib, model = 'discrete') in the APE package (Paradis & Schliep, 2019) in R software (R Core Team, 2022), a crude post hoc smoothing method (Talavera et al., 2013; Tang et al., 2014) on par with applying GMYC on a concatenated tree. The GMYC method was applied using the splits package (Ezard et al., 2009) implemented in R software and the single threshold option.

When interpreting the species delimitation results, we applied distinctness-in-sympatry criteria to alleviate the impact of geographical variation (Brandvain & Matute, 2018). We also treated the delimited species as 'MSC-species' (Bravo et al., 2019; Heled & Drummond, 2010); rather than corresponding to a traditional taxonomic rank, they represent populations with periods of ceased gene exchange that may be permanent.

RESULTS

Data assembly and phylogenetic inference

The mean coverage of the 163 assemblies was 12×, ranging from <1× to 52×. This variation was expected due to varying degrees of DNA degradation. The reference genome had a BUSCO score of 75, whereas the mean BUSCO score for the assemblies was 63 (<1–75). Of the 3913 orthologous nuclear genes of Strepsiptera from our reference dataset, on average, 3333 were obtained from the assemblies (53–3535). The average number of loci per specimen following the filtering steps was 2141 (1831–2210), and the average number of specimens per gene was 139 (4–150). Specimen age and the used library protocol did not have a significant effect on the number of loci recovered (Figure 4). After the removal of long branches and short alignments, our dataset had 151 terminals, including the outgroup, and 2315 loci.

The ML analysis with IQ-TREE2 resulted in a phylogeny with the majority of nodes fully supported (Figure 5). All except one backbone node were maximally supported. Mostly within putative species there were 26 nodes (17%) with less than strong support, here defined as <95% ultrafast bootstrap value. The resulting tree had two major clades, which could be further divided into seven subclades (Figure 5). Within the clades, most of the putative species formed monophyletic groups, often associated with a single host subgenus. Two putative species were polyphyletic: *Stylops aterrimus* Newport and *S. nassonowi* Pierce. Some species clades had structure within them, particularly the clade representing *S. spreta* Perkins, which was divided into three smaller subclades. WP specimens did not show strong geographical patterns; however, specimens from the EP region and the bordering areas composed separate clades within the putative species clades.

Molecular species delimitation

With the entire dataset of 2315 loci, in the majority of cases, SODA split single putative species into multiple MSC-species (Figure 6). With the default *p* value of 0.05, SODA split the entire dataset into 99 MSC-species and with the more conservative value of 0.001 into 84 MSC-species. The 20 preliminarily identified species included in the dataset were delimited into 62 MSC-species. When all representatives of a putative species were collected from the same country, it was more likely that those were delimited as one species. The host species within putative species did not seem to have a large effect on the inferred species limits. For example, putative *S. thwaitesi* Perkins from three different host species were delimited as one species, when all specimens were collected from the same country. However, the number of loci included in the SODA analysis had a significant effect on the number of inferred species in the sub-sampled replications (Kruskal–Wallis chi-square = 76.463, df = 7, $p < 0.0001$) (Figure 7). With 25 loci, the average number of delimited species was 23, whereas with 2000 loci, it was 74. With fewer loci, SODA species corresponded more with putative taxonomic species, except in some cases where species were merged together. The three putative species *S. japonicus* Kifune & Hirashima, *S. praecocis* Luna de Carvalho and *S. nevinsoni* Perkins were merged into one species, as well as the two species *S. obsoletus* Luna de Carvalho and *S. madrilensis* Luna de Carvalho. With the GMYC method, the number of delimited species was 23 (2 log likelihood confidence interval: 21–28), and the single

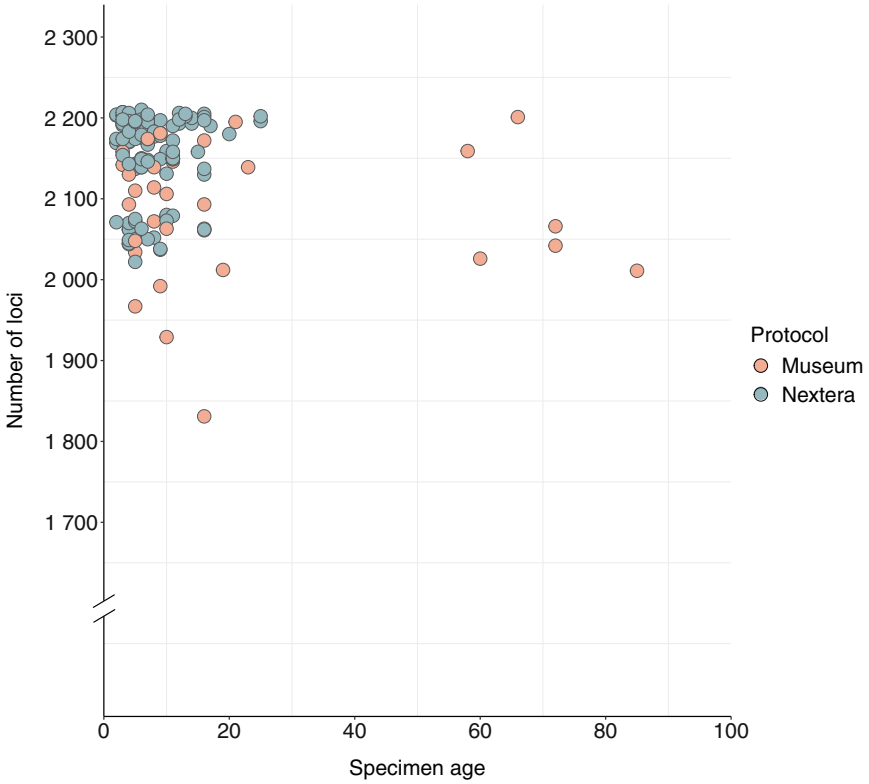

FIGURE 4 Number of recovered loci after filtering steps for the two protocols as a function of specimen age. Pink are samples prepared by the museum library protocol (Irestedt et al., 2022), blue represent samples prepared using the Nextera Flex library protocol (Illumina, 2018).

species null model was significantly rejected (null: 511.9, GMYC: 535.4, LR-test: 6.4×10^{-11}) (Figure 6). Species delimited with GMYC were almost the same as with SODA using only 25 loci.

Species diversity and host associations of Western Palaearctic *Stylops*

Here, we present the evaluation of the existing and conflicting species hypotheses of WP *Stylops* (Cook, 2019; Straka et al., 2015) based on our molecular analysis, summarised in Table 1. We also assess possible hidden diversity. Some additional host associations have been suggested after the publication of the catalogues (e.g., Smit et al., 2020). We report new host associations and provide information on the host-sharing between species, compiled in Table 2. If not stated otherwise, SODA results refer to the results obtained from analysis of the entire dataset. We refrain from describing new species in this study. However, the data provided here may be informative for future species descriptions.

Clade A

The four putative species within this clade, *Stylops lusohispanicus* Luna de Carvalho, 1974, *S. gwynanae* Günther & Šedivi, 1957, *S. moniliaphagus* Luna de Carvalho, 1974, and *S. salamancanus* Luna de Carvalho, 1974, were monophyletic, and all but *S. gwynanae* were delimited as single species. SODA split *S. gwynanae* into four species; however, the splits most likely reflect geographical distance instead of actual species limits. GMYC did not split *S. gwynanae*. The phylogeny implies *Andrena symphyti* Schmiedeknecht, 1883, and *A. granulosa* Pérez, 1902, to be likely host species of S. *gwynanae* besides *A. bicolor* Fabricius, 1775.

Two species are associated with hosts from subgenera *Aciandrena* Warncke, 1968, and *Graecandrena* Warncke, 1968: *S. lusohispanicus* and *S. liliputanus* Luna de Carvalho, 1974. However, both the phylogeny and SODA results suggest that more species than *S. liliputanus* and *S. lusohispanicus* have hosts in the subgenera *Aciandrena* and *Graecandrena* since specimens from those host species are in more than two clades. Some of the unidentified specimens might be *S. liliputanus* yet could not be reliably identified as such based on host association and morphology. It is also possible that these species exhibit host sharing. The identity of *Stylops* from *Aciandrena* and *Graecandrena* is not settled and requires further study. To our knowledge, the included *S. moniliaphagus* is the first DNA sequence obtained from *Stylops* from a host in the subgenus *Orandrena* Warncke, 1968.

Clade B

Putative *S. thwaitesi* Perkins, 1918, *S. maxillaris* Pasteels, 1949, and *S. spreta* Perkins, 1918, were assigned to this clade together with

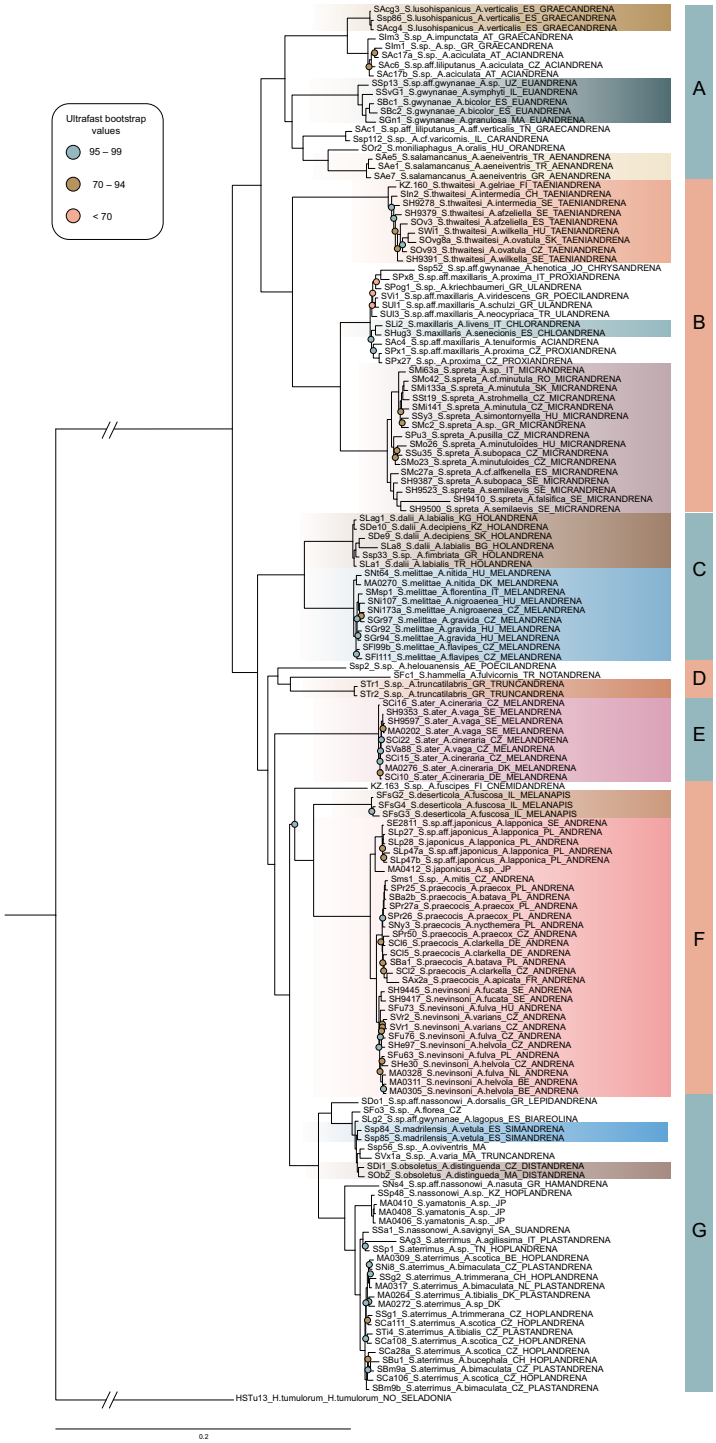

FIGURE 5 Maximum likelihood tree obtained from IQ-TREE2 based on the 2315 loci data-set. Parts of the tree highlighted with different colours correspond to host subgenera. All nodes were maximally (100) supported by ultrafast bootstrap, except those denoted with coloured circles. The scale bar indicates the expected number of substitutions per site. Terminal labels contain the following information separated by underscores: Voucher code, *Stylops* species, Host species (*Andrena*), sampling country in ISO 3166-1 alpha-2 abbreviation code and the host subgenus.

unclassified specimens from multiple host subgenera with varying degrees of support of the internal nodes. *Stylops thwaitesi* was monophyletic and included all specimens from the *Taeniandrena* Hedicke, 1933, host subgenus. Even though SODA split the clade into six species, sympatric specimens from *A. afzeliella* (Kirby, 1802), *A. wilkella* (Kirby, 1802) and *A. intermedia* Thomson, 1870, are delimited as the same species. Both SODA using only 25 loci and GMYC supported a single species hypothesis. Based on this and groupings in the ML tree,

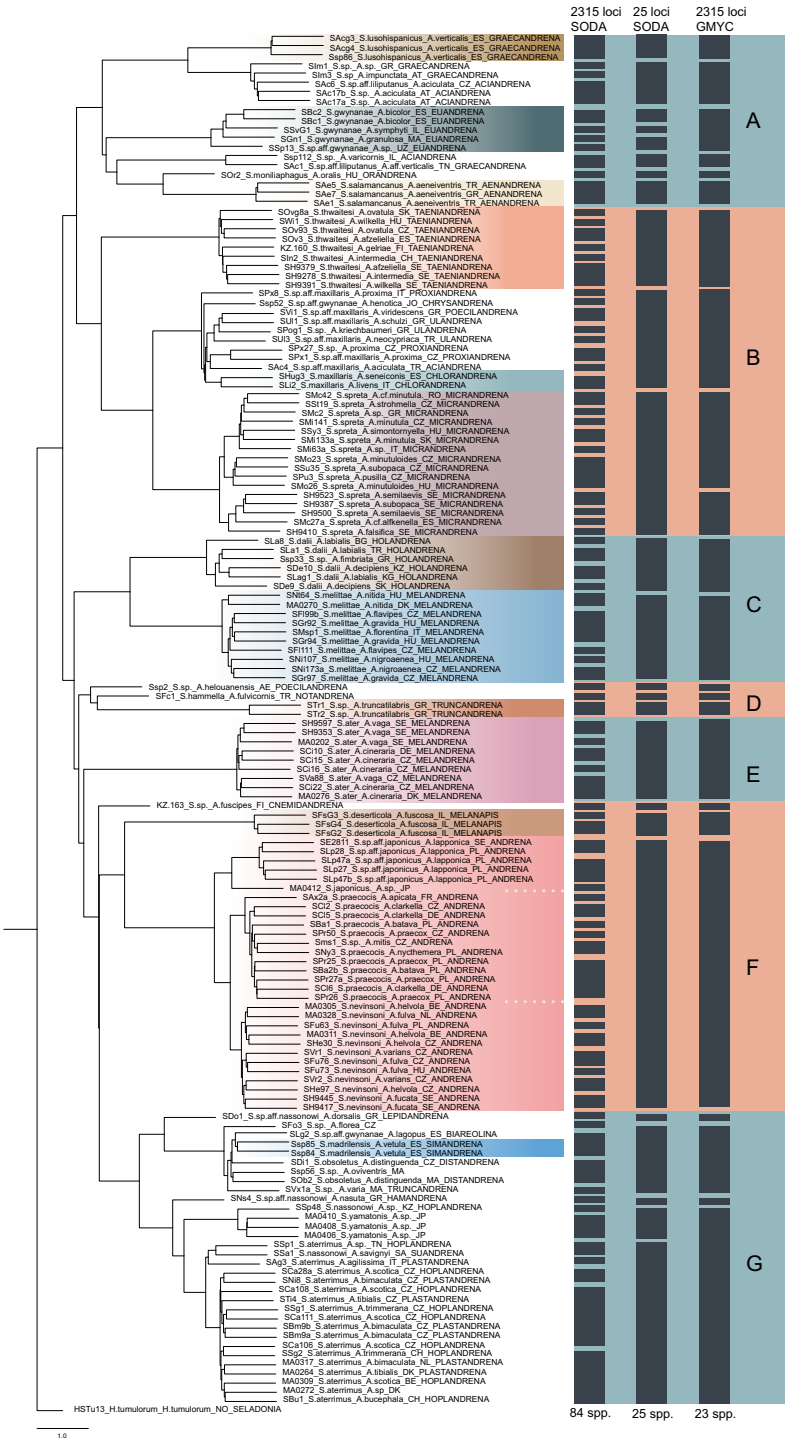

FIGURE 6 Results of species delimitation analyses for Western Palaearctic *Stylops*. The inferred species limits from (i) SODA using gene trees of the entire dataset (2315 loci), (ii) SODA using a random subset of 25 loci and (iii) GMYC using the IQ-TREE2 phylogeny (ultra-metricised) from 2315 concatenated loci are denoted by black bars mapped on a species tree from ASTRAL-III. Parts of the tree highlighted with different colours correspond to host subgenera. Scale bar in coalescent units. Terminal labels as in Figure 5.

the results of this study indicate that *A. gelriae* van der Vecht, 1927, *A. russula* Lepeletier, 1841, *A. afzeliella* and *A. intermedia* are hosts of *S. thwaitesi* in addition to previously reported hosts. *Stylops alfkeni* Hofeneder, 1939, which has a single reported host association with *A. russula*, was proposed to be a junior subjective synonym of *S. thwaitesi* by Straka et al., (2015). Cook (2019) considered the synonymization unjustified and reinstated the species. In the present study, *Stylops* specimens from *A. russula* came out within the *S. thwaitesi* clade in the ML tree (Figure S1). This suggests that synonymization with *S. thwaitesi* is justified, or that the two species exhibit

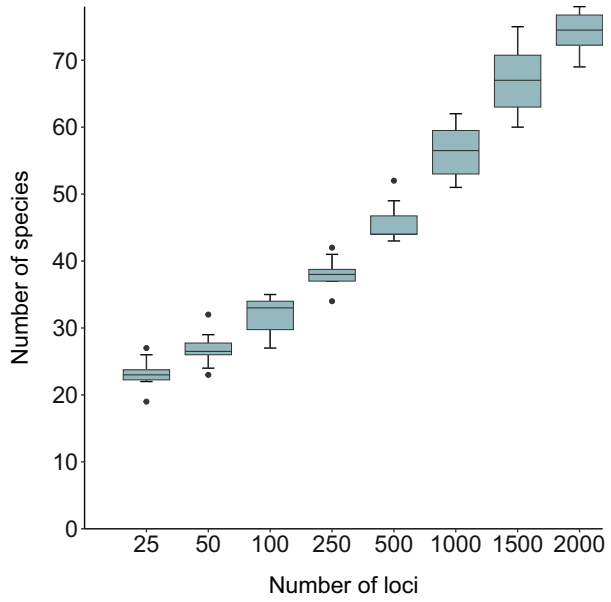

FIGURE 7 Boxplot with the effect of number of loci on the number of delimited species using SODA (p value cut-off = 0.001). Each set of loci number derive from 10 replicates of randomly selected loci.

host-sharing. *Stylops thwaitesi* females have a very distinct pigmentation; typically, they have a rectangular dark marking extending from the first abdominal segment into mesothorax. Unfortunately, the description of *S. alfkeni* focused more on the measurements of the female cephalothorax and did not include information on pigmentation patterns. This makes comparing the two putative species based on female morphology difficult, because the measurements of the cephalothorax are affected by host size (Cook, 2014). Furthermore, the illustration of the male penis in the description of *S. alfkeni* differs from that of *S. thwaitesi*. Thus, inspection of type specimens is required before making taxonomic conclusions.

Putative *S. maxillaris* was recovered in a subclade that contained specimens associated with hosts in subgenera *Proxiandrena* Schmid-Egger, 2005, *Aciandrena*, *Chlorandrena* Pérez, 1890, *Ulandrena* Warncke, 1968, and *Poecilandrena* Hedicke, 1933. Within this subclade, SODA delimited eight species including *S. maxillaris*. Cook (2019) reinstated *S. esteponensis* Luna de Carvalho, 1974, which was synonymized with *S. maxillaris* by Straka et al., (2015). *Stylops esteponensis* is associated with *A. livens* Pérez, 1895, and was described from females only. Our results support the synonymization since *S. maxillaris* from both *A. senecionis* Pérez, 1895, and *A. livens* were assigned to the same species. That being said, *A. senecionis* is a new host association of *S. Maxillaris*, and ideally, specimens from a previously reported host, *A. humilis* Imhoff, 1832, should be included in the analysis. Furthermore, there is a lot of uncertainty in this clade. Multiple nodes had low support and many specimens remain unidentified. It is possible that the subclade contains multiple species, some undescribed, and a revision is needed. *Andrena proxima* (Kirby, 1802) is likely to be a host of at least one such undescribed species. Cook (2019) listed *A. proxima* as a host of *S. melittae*, but all included specimens from *A. proxima* are in this clade and not in clade C with *S. melittae*. Furthermore, two of them were morphologically similar, whereas one was morphologically different. The morphologically similar specimens were delimited to belong to the same species. GMYC and 25 loci dataset SODA results both supported single species hypotheses for this subclade, but given the morphological differences between the included specimens, this is likely an incorrect lumping.

Based on the results from the present study, there might be either strong population structure or cryptic species in *S. spreta*. Intraspecific structure was visible in the phylogenetic trees, resulting in three clades within the species. Furthermore, SODA divided putative *S. spreta* into 10 species. However, since it is a geographically widespread species, the population structure and number of 'MSC-species' might be explained by isolation by distance and geographic sampling (Mason et al., 2020; McKay et al., 2013). The fact that all specimens from Sweden, the northern extreme of the species distribution area, were clustered together, suggests an isolation-by-distance explanation. Furthermore, with only 25 loci, SODA supported the single species hypothesis, whereas GMYC delimited the Swedish specimens as a separate species from the remaining *S. spreta*. Given this, we consider it premature to split *S. spreta* into several species, though more data may necessitate another solution in the future. Cook (2019) reinstated *S. duriensis* Luna de Carvalho, 1974, which was placed as a supposed junior synonym of *S. spreta* by Straka et al., (2015). Our dataset did not have any representatives from the host of *S. duriensis* (*A. tenuistriata* Pérez, 1895). Thus, the identity of *S. duriensis* remains uncertain.

Clade C

Stylops dalii Curtis, 1828, and *S. melittae* Kirby, 1802, were assigned to this clade with well-supported nodes. Three species that were left as a synonym of *S. melittae* by Straka et al., (2015) were reinstated by Cook (2019) with an argument that there was not enough evidence for the synonymy: *S. flavipedis* Hofeneder, 1923, *S. nitidae* Pasteels, 1954, and *S. giganteus* Luna de Carvalho, 1974. *Stylops flavipedis* is associated with *A. flavipes* Panzer, 1799, whereas *S. nitidae* is associated with *A. nitida* (Müller, 1776). *Stylops* specimens from both of these hosts were recovered within the *S. melittae* subclade and assigned to the same species as *S. melittae* by SODA. Thus, our data suggests that these species are synonyms of *S. melittae*. Our data did not include specimens from the host of *S. giganteus* (*A. thoracica* (Fabricius, 1775)). Cook (2019) listed *A. cineraria* (Linnaeus, 1758) as host of *S. melittae*. No specimen from *A. cineraria* was recovered in this clade. However, the results of the present study imply *A. gravida* Imhoff, 1832, and *A. florentina* Magretti, 1883, to be new host associations of *S. melittae*.

Previously, the only confirmed host association of *S. dalii* has been *A. labialis* (Kirby, 1802). Both the phylogeny and SODA results supported the suggested new host association, *A. decipiens* Schenck, 1861, which was based on host subgenus. *Stylops dalii* from

TABLE 1 Summary of the species delimitation results and status of the species hypotheses of Western Palaearctic *Stylops*.

Species	Synonyms	New synonyms	Clade	SODA 2315 loci	GMYC	Status
Stylops ater Reichert, 1914		*Stylops muelleri* Borchert, 1971, *Stylops krygeri* Pierce, 1919, *Stylops ovinae* Noskiewicz & Poluszyński, 1927	E	Split	Single	Unambiguous
Stylops aterrimus Newport, 1851	*Stylops trimmerana* Smith, 1857, *Neostylops trimmerana* Pierce 1919, *Stylops niger* Beneden, 1875, *Stylops trimmeranea* Kinzelbach, 1978	*Stylops perkinsi* Pasteels, 1949, *Stylops bimaculatae* Perkins, 1918	G	Split	Merged	Ambiguous
Stylops dalii Curtis, 1828			C	Split	Single	Unambiguous
Stylops deserticola Medvedev, 1970	*Stylops desertorum* Medvedev, 1970		F	Split	Single	Unambiguous
Stylops gwynanae Günther & Šedivi, 1957			A	Split	Single	Unambiguous
Stylops hammella Perkins, 1918	*Stylops hammelae* Kinzelbach, 1978		D	N/A	N/A	Unambiguous
Stylops lusohispanicus Luna de Carvalho, 1974			A	Single	Single	Unambiguous
Stylops madrilensis Luna de Carvalho, 1974			G	Merged	Merged	Ambiguous
Stylops maxillaris Pasteels, 1949		*Stylops esteponensis* Luna de Carvalho, 1974	B	Single	Merged	Unambiguous
Stylops melittae Kirby, 1802	*Stylops kirbii* Leach, 1815, *Stylops haworthi* Stephens, 1829a, *Stylops spencii* Pickering, 1836	*Stylops flavipedis* Hofeneder, 1924, *Stylops nitidae* Pasteels, 1954	C	Split	Single	Unambiguous
Stylops moniliaphagus Luna de Carvalho, 1974			A	N/A	N/A	Unambiguous
Stylops nassonowi Pierce, 1909	*Stylops melittae* Nassonow, 1893, *Stylops savignyi* Hofeneder, 1924		G	Split	Single	Ambiguous
Stylops nevinsoni Perkins, 1918		*Stylops transversa* Pasteels, 1949	F	Split	Merged	Unambiguous
Stylops obsoletus Luna de Carvalho, 1974			G	Single	Merged	Ambiguous
Stylops praecocis Luna de Carvalho, 1974		*Stylops nycthemerae* Noskiewicz & Poluszyński, 1927	F	Split	Merged	Unambiguous
Stylops salamancanus Luna de Carvalho, 1974			A	Single	Single	Unambiguous
Stylops sp. *aff. liliputanus ex A. varicornis*			A	Single	Single	Ambiguous
Stylops sp. *aff. liliputanus ex A. impunctata & A. aciculata*			A	Split	Single	Ambiguous
Stylops sp. *nov. ex A. truncatilabris*			D	Single	Single	Unambiguous
Stylops sp. *nov. ex A. varia*			G	N/A	N/A	Ambiguous
Stylops sp. *nov. ex A. helouanensis*			D	N/A	N/A	Unambiguous
Stylops sp. *nov. aff. japonicus ex A. lapponica*			F	Split	Merged	Unambiguous
Stylops sp. *nov. aff. maxillaris ex A. proxima*			B	Split	Merged	Ambiguous
Stylops sp. *nov. aff. nassonowi ex A. drosalis*			G	N/A	N/A	Unambiguous

(Continues)

TABLE 1 (Continued)

Species	Synonyms	New synonyms	Clade	SODA 2315 loci	GMYC	Status
Stylops sp. nov. *aff. nassonowi ex A. nasuta*			G	N/A	N/A	Unambiguous
Stylops sp. nov. *ex A. fuscipes*			F	N/A	N/A	Unambiguous
Stylops spreta Perkins, 1918			B	Split	Split	Species complex
Stylops thwaitesi Perkins, 1918	*Stylops* sp. Thwaites, 1841, *Stylops thwaitei* Saunders, 1872, *Stylops wilkellae* Perkins, 1918, *Stylops albofasciatae* Günther & Šedivi, 1957		B	Split	Single	Unambiguous

Note: Synonyms following Cook (2019).
Abbreviations: GMYC, generalised mixed Yule coalescent method; SODA, Species bOundry Delimitation using Astral.

both *A. labialis* and *A. decipens* were mixed in a clade and delimited as the same species. In addition, *Stylops* from *A. fimbriata* Brullé, 1832, ended up in the clade and assigned to the same species as *S. dalii*. Thus, we also report *A. fimbriata* to be a new host of *S. dalii*.

Clade D

This clade contained *Stylops* from three host subgenera: *Poecilandrena* Hedicke, 1933, *Notandrena* Pérez, 1890, and *Truncandrena* Warncke, 1968. Only one specimen was identified beforehand, *S. hamella* Perkins, 1918. *Andrena fulvicornis* Schenck, 1861, is a new host association of *S. hamella*. Cook (2019) reinstated *S. nitidiusculae* Poluszynski, 1927, which was synonymized with *S. hamella* by Straka et al., (2015). Our dataset did not include individuals from the host of *S. nitidiusculae* (*A. nitidiuscula* Schenck, 1853). There are no *Stylops* species associated with host subgenus *Poecilandrena* and the *Stylops* from *A. helouanensis* Friese, 1899, is likely an undescribed species. Likewise, the *Stylops* from *A. truncatilabris* Morawitz, 1877, is probably a new species and no described *Stylops* species are associated with hosts from subgenus *Truncandrena*. Interestingly, *Stylops* from *A.* (*Truncandrena*) *varia* Pérez, 1895, was recovered in a different clade (Clade G), suggesting that two different *Stylops* species may utilise hosts from subgenus *Truncandrena*.

Clade E

The *Stylops* species within this clade is probably one of the best studied WP species (e.g., Fraulob et al., 2015; Löwe et al., 2016; Peinert et al., 2016). Ironically, it is also one with the most controversial nomenclature, detailed in Straka, Alqarni, et al. (2015). Currently, two names are in use for the *Stylops* parasitizing *A. vaga* Panzer, 1799: *S. ater* Reichert, 1914 (e.g., Smit et al., 2020; Straka et al., 2015) and *S. ovinae* Noskiewicz & Poluszynski, 1928 (e.g., Fraulob et al., 2015; Jandausch et al., 2022; Löwe et al., 2016). Furthermore, Cook (2019) reinstated two more species associated with *A. vaga*: *S. muelleri* Borchert, 1971, and *S. krygeri* Perkins, 1918. However, the present study does not find justification for the reinstatement of those species. Based on our results, there is only one *Stylops* species with *A. vaga* as a principal host, *S. ater* (following the nomenclature of Straka, Alqarni, et al. (2015). Another host of *S. ater* is *A. cineraria* (Linnaeus, 1758) (Smit et al., 2020).

Clade F

This clade was divided into three subclades: one with a single unidentified specimen from *A. fuscipes*, which was sister to the rest of the clade, one containing *S. deserticola* Medvedev, 1970, and the last containing specimens parasitizing hosts of subgenus *Andrena*. The specimen from *A. fuscipes* (Kirby, 1802) is most likely an undescribed species. No known *Stylops* species from the West Palaearctic region

TABLE 2 Summary of the host associations of Western Palaearctic *Stylops*.

Species	Host associations	New host associations	Host subgenera
Stylops ater Reichert, 1914	*Andrena vaga* Panzer, 1799, A. cineraria (Linnaeus, 1758)		*Melandrena*
Stylops aterrimus Newport, 1851	*Andrena tibialis* (Kirby, 1802), *A. trimmerana* (Kirby, 1802), *A. agilissima* (Scopoli, 1770), *A. scotica* Perkins, 1916	*Andrena bimaculata* (Kirby, 1802), *A. bucephala* Stephens, 1846	*Plastandrena*, *Hoplandrena*, *Agandrena*
Stylops dalii Curtis, 1828	*Andrena labialis* (Kirby, 1802)	*Andrena decipiens* Schenck, 1861, *A. fimbriata* Brullé, 1832	*Holandrena*
Stylops deserticola Medvedev, 1970	*Andrena fuscosa* Erichson, 1835		*Melanapis*
Stylops gwynanae Günther & Šedivi, 1957	*Andrena bicolor* Fabricius, 1775	*Andrena symphyti* Schmiedeknecht, 1883, *A. granulosa* Pérez, 1902	*Euandrena*
Stylops hamella Perkins, 1918	*Andrena chrysosceles* (Kirby, 1802)	*Andrena fulvicornis* Schenck, 1861	*Notandrena*
Stylops lusohispanicus Luna de Carvalho, 1974	*Andrena verticalis* Pérez, 1895		*Graecandrena*
Stylops madrilensis Luna de Carvalho, 1974	*Andrena vetula* Lepeletier, 1841	*Andrena lagopus* Latreille, 1809	*Simandrena*, *Biareolina*
Stylops maxillaris Pasteels, 1949	*Andrena humilis* Imhoff, 1832, *A. livens* Pérez, 1895		*Chlorandrena*
Stylops melittae Kirby, 1802	*Andrena soror* Dours, 1872, *A. nitida* (Müller, 1776), *A. thoracica* (Fabricius, 1775), *A. nigroaenea* (Kirby, 1802), *A. flavipes* Panzer, 1799	*Andrena gravida* Imhoff, 1832, *A. florentina* Magretti, 1883	*Melandrena*
Stylops moniliaphagus Luna de Carvalho, 1974	*Andrena monilia* Warncke, 1967	*Andrena oralis* Morawitz, 1876	*Orandrena*
Stylops nassonowi Pierce, 1909	*Andrena pilipes* Fabricius, 1781, *A. savignyi* Spinola, 1838		*Plastandrena*, *Suandrena*
Stylops nevinsoni Perkins, 1918	*Andrena synadelpha* Perkins, 1914, *A. fulva* (Müller, 1766)	*Andrena fucata* Smith, 1847, *A. varians* (Kirby, 1802), *A. helvola* (Linnaeus, 1758)	*Andrena*
Stylops obsoletus Luna de Carvalho, 1974	*Andrena distinguenda* Schenck, 1871		*Distandrena*
Stylops praecocis Luna de Carvalho, 1974	*Andrena praecox* (Scopoli, 1763), *A. nycthemera* Imhoff, 1868	*Andrena batava* Pérez, 1902, *A. clarkella* (Kirby, 1802), *A. apicata* Smith, 1847, *A. mitis* Schmiedeknecht, 1883	*Andrena*
Stylops salamancanus Luna de Carvalho, 1974	*Andrena hedikae* Jaeger, 1934	*Andrena aeneiventris* Morawitz, 1872	*Aenandrena*
Stylops sp. *nov.*		*Andrena truncatilabris* Morawitz, 1877	*Truncandrena*
Stylops sp. *nov.*		*Andrena varia* Pérez, 1895	*Truncandrena*
Stylops sp. *nov.*		*Andrena fuscipes* (Kirby, 1802)	*Cnemidandrena*
Stylops sp. *nov.*		*Andrena helouanensis* Friese, 1899	*Poecilandrena*
Stylops sp. *nov. aff. japonicus*		*Andrena lapponica* Zetterstedt, 1838	*Andrena*
Stylops sp. *aff. liliputanus*		*Andrena varicornis* Pérez, 1895, *A. aff. verticalis* Pérez, 1895	*Aciandrena*, *Graecandrena*
Stylops sp. *aff. liliputanus2*		*Andrena impunctata* Pérez, 1895, *A. aciculata* Morawitz, 1886	*Aciandrena*, *Graecandrena*
Stylops sp. *n. aff. nassonowi*		*Andrena nasuta* Giraud, 1863	*Hamandrena*
Stylops sp. *n. aff. nassonowi*		*Andrena dorsalis* Brullé, 1832	*Lepiandrena*
Stylops sp. *aff. maxillaris*		*Andrena proxima*	*Proxiandrena*
Stylops spreta Perkins, 1918	*Andrena minutula* (Kirby, 1802), *A. tenuistriata* Pérez, 1895, *A. strohmella E. Stöckhert*, 1928, *A. subopaca* Nylander, 1848, *A. minutuloides* Perkins, 1914, *A. falsifica* Perkins, 1915	*Andrena simontornyella* Noskiewicz, 1939, *A. pusilla* Pérez, 1903, *A. semilaevis* Pérez, 1903	*Micrandrena*

(Continues)

TABLE 2 (Continued)

Species	Host associations	New host associations	Host subgenera
Stylops thwaitesi Perkins, 1918	*Andrena ovatula* (Kirby, 1802), *A. russula* Lepeletier, 1841, *A. wilkella* (Kirby, 1802), *A. intermedia* Thomson, 1870	*Andrena afzeliella* (Kirby, 1802), *A. gelriae* an der Vecht, 1927	*Taeniandrena*

are associated with hosts from subgenus *Cnemidandrena* Hedicke, 1933. Putative *S. deserticola* specimens were assigned to the same species by SODA with 25 loci and GMYC, whereas a two-species hypothesis was supported by the soda analysis when using the entire dataset.

The *Andrena* subclade contained four species: *S. japonicus* Kifune & Hirashima, 1985, *S. praecocis* Luna de Carvalho, 1974, *S. nevinsoni* Perkins, 1918, and one undescribed species closely related to *S. japonicus*. All putative species within the clade were split into multiple species by SODA when the entire dataset was included. However, representatives of the species were sampled from a geographically wide area, which may explain the over-splitting. Interestingly, the species were merged by both GMYC and the 25 loci SODA analysis. This suggests that the speciation might be relatively young between species utilising subgenus *Andrena*. The close relatedness of the species has been noted before (Jůzová et al., 2015; Straka et al., 2015). Nevertheless, since the species repeatedly form distinct clades, there is no overlap in utilised host species and they should be treated as separate species.

Stylops nycthemerae Noskiewicz & Poluszynski, 1928, parasitizing *A. nycthemera* Imhoff, 1868, was placed as a supposed synonym of *S. praecocis* by Straka et al., (2015) but reinstated by Cook (2019). Cook's argument was that since the synonymization of the species was based on close relatedness of hosts alone, there was insufficient evidence for the synonymization. However, in the present study, *Stylops* from *A. nycthemerae* is delimited as the same species as *S. praecocis*. Thus, our results support the synonymization. The placement of specimens in the same clade in the ML tree supports four new host associations of *S. praecocis*: *A. batava* Pérez, 1902, *A. clarkella* (Kirby, 1802), *A. apicata* Smith, 1847, and *A. mitis* Schmiedeknecht, 1883.

Straka et al., (2015) classified *S. transversa* Pasteels, 1949, as a synonym of *S. nevinsoni*. *Stylops transversa* was later reinstated by Cook (2019). *Stylops transversa* has two reported hosts: *A. fulva* (Müller, 1766) and *A. clarkella*. The former is also known as the host of *S. nevinsoni* and the latter of *S. praecocis*. However, based on the species description of *S. transversa*, *A. fulva* is its principal host (Pasteels, 1949), whereas *A. clarkella* is an auxiliary host of it (shared with *S. praecocis*). Since *Stylops* from the principal host *A. fulva* were assigned to the same species as *S. nevinsoni* in the analysis, our data support the synonymization by Straka et al. (2015).

Clade G

Five species were assigned to this clade, along with multiple unidentified specimens. *Stylops obsoletus* Luna de Carvalho, 1974, is found within the same subclade as *S. madrilensis* Luna de Carvalho, 1974, alongside unidentified specimens from various host subgenera. The specimen from *A. varia* (*Truncandrena*) surprisingly also ended up in this clade and not together with the specimen from the second included *Truncandrena* (*A. truncatlilabris*) in clade D. *Stylops* specimens from host subgenera *Lepidandrena* Hedicke, 1933, and *Hamandrena* are probably undescribed species. The identities of the other unidentified specimens need to be further inspected before drawing any conclusions of their relatedness to the putative species.

Stylops aterrimus Newport, 1851, and *S. nassonowi* Pierce, 1909, are closely related species, which are nearly indistinguishable morphologically (Straka, Alqarni, et al., 2015). Geographically, *S. aterrimus* represents the Northern lineage, and *S. nassonowi* the Southern lineage. Host sharing is likely to occur between these closely related species in the contact zone. In the present study, *S. nassonowi* from Kazakhstan formed a separate clade with Japanese *S. yamatonis* Kifune & Hirashima, 1985, whereas *S. nassonowi* from Saudi Arabia grouped with two specimens identified as *S. aterrimus*. *Stylops yamatonis* belongs to the same species group as *S. nassonowi* and *S. aterrimus* (Straka et al. 2015). Since differentiating species in this species group morphologically can be difficult, it is probable that the *S. nassonowi* from Kazakhstan actually represents an undescribed species closely related to *S. yamatonis*, or belongs to that species. The *Stylops nassonowi* from Saudi Arabia most likely represents the true *S. nassonowi*. It is possible that the two *S. aterrimus* within the same subclade are *S. nassonowi* too, since the two species are difficult to tell apart based on morphology. The SODA analysis of 25 loci supported a two-species hypothesis, where specimens from Kazakhstan and Japan formed one species and all other *S. nassonowi* and *S. aterrimus* specimens another one. The GMYC analysis, on the other hand, supported a one-species hypothesis, where all specimens from *Hoplandrena* Pérez, 1890, *Plastandrena* Hedicke, 1933, and *Suandrena* Warncke, 1968, were lumped together. The relationship between *S. nassonowi* and *S. aterrimus* requires further investigation.

Cook (2019) reinstated three species that were synonymized with *S. aterrimus* by Straka et al., (2015): *S. bimaculatae* Perkins, 1918, *S. perkinsi* Pasteels, 1954, and *S. dominiquei* Pierce, 1909. *Stylops* from the host of *S. bimaculatae*, *A. bimaculata* (Kirby, 1802), are mixed within the *S. aterrimus* clade and were delimited as the same species as *S. aterrimus* in SODA. Therefore, the results of this study support this synonymization. With *Stylops* from the reported hosts of *S. perkinsii*, the situation was the same: both the position in the ML tree and SODA results support the synonymization. Furthermore, the illustration of *S. perkinsii* in the species description shares the same characteristic coloration with *S. aterrimus*. With *S. dominiquei* the results are more ambiguous. Since the taxon from the reported host species of *S.*

dominiquei (*A. agilissima* (Scopoli, 1770)) was assigned to the proximity of *S. nassonowi*, it may in fact be a synonym of *S. nassonowi*. More representatives of *S. dominiquei* are needed to draw a compelling conclusion about the relationship of these species.

DISCUSSION

Stylops is by far the largest genus in Strepsiptera, but previous efforts to chart the species diversity in the genus have resulted in incongruent estimates (Cook, 2019; Jůzová et al., 2015; Kinzelbach, 1978; Straka et al., 2015). Multiple sources of information are relevant to distinguish separately evolving lineages (de Queiroz, 2007); hence, species delimitation attempts are optimally based on a combination of data types. Besides morphological characteristics and ecological associations such as host usage, an integrative approach could also consider, for example, life history traits, geographic distributions and genomic similarities (e.g., Ailán-Choke & Pereira, 2021; Brunet et al., 2017; Padial et al., 2010). More research is needed for all these aspects in *Stylops*, but comparative genomic data have been lacking entirely until now. With our comprehensive genomic dataset, we were able to fill this gap for the species diversity question of most *Stylops* in WP.

There are several reasons behind the taxonomic difficulties in the genus *Stylops*. In the pre-molecular era, it was difficult to link the morphological variable but rarely found males to the larviform females due to the extreme sexual dimorphism (Kinzelbach, 1978). Species descriptions based on females have usually not been sufficient to render the new species recognisable morphologically, and tended to emphasise host species. However, the host species emphasis was in essence an assumption of single-host specificity without any supplementary evidence (Straka et al., 2015). Furthermore, stylopisation can induce morphological changes on the host, making identification of the host error-prone (Salt, 1927, 1931). Using too narrow host range led to an inflation of species diversity estimates; inversely, a too broad host range assumption such as Kinzelbach's (1978) led to an underestimated diversity (Jůzová et al., 2015). In our assessment adding comparative genomic data, we found multiple examples of both types of erroneous assumptions.

First, the assumption of a single *Stylops* species in the WP (Kinzelbach, 1978) is not supported by the new molecular data. Here, we add comparative genomic confidence to the conclusion reached based on single gene analyses (Jůzová et al., 2015). The winged *Stylops* males are poor fliers with a very short lifespan (Balzer & Davis, 2020; Kathirithamby, 2009). Phoretic dispersal by the first instar larva with the stronger flying *Andrena* hosts is likely the most important mode of dispersal. However, even the hosts show considerable genetic population structure over WP (Černá et al., 2013; Davis et al., 2010). A single parasite species with isolation-by-distance population structure would show geographically structured clades irrespective of host relationships. Instead, what our genomic data show is significant clade structure based on host subgenera or closely related hosts, even in geographic sympatry. This is incompatible with a single-species hypothesis. The lineage-through-time plot shows a distinct branching rate transition within the confidence interval of the GMYC analysis, interpreted as slower speciation branching followed by faster within-species coalescent branching after the transition point (Figure S2). A minimum of 21 species is delimited by the GMYC analysis based on the 2-log likelihood confidence interval, a conservative estimate compared to the SODA analysis with the full dataset.

Second, the early-era tradition of describing a new *Stylops* species for every new *Andrena* species found to be stylopised can likewise be confidently rejected as a wise strategy for accurate parasite species recognition. Our 151-sample dataset includes parasites of ca 100 different *Andrena* host species (some samples not identified to species). Even the SODA analysis with the full dataset merges samples from different host species into a single parasite species, and the more conservative and perhaps reasonable estimates (see discussion below) give around 27 species. This shows that a single-host species specificity hypothesis of *Stylops* is as incompatible with the genomic data as was the supergeneralist hypothesis, and not only applies to early-era taxonomy. We found multiple cases in the latest world catalogue (Cook, 2019) where a species was assumed to have overly narrow host range. For example, we were able to associate host species *A. bimaculata* and *A. scotica* with *S. aterrimus* and provide evidence for synonymization of the two species (*S. bimaculatae* and *S. perkinsi*) formerly associated with those host species with *S. aterrimus*.

In between the two extremes, many *Stylops* species seem to be limited to exploit a set of closely related hosts, often classified in the same subgenus of *Andrena* (Jůzová et al., 2015). For instance, *S. dalii* is restricted to species within *Holandrena*, *S. spreta* to species within *Micrandrena*, and *S. thwaitesi* to species within *Taeniandrena*. Even though the host subgenus is a good indicator of probable host species, additional evidence is needed to determine the host associations of any given species. Members of the nominotypical subgenus *Andrena* for instance are parasitized by multiple but closely related *Stylops* species. In addition, several subgenera of *Andrena* are poly- or paraphyletic as currently defined (Pisanty et al., 2022). These uncertainties in the host taxonomy complicate host specificity inference in *Stylops*. On the other hand, our results may be helpful for the revision of *Andrena* subgenera. Some closely related *Stylops* taxa, such as *S. nassonowi* and *S. aterrimus*, parasitize hosts from multiple subgenera. Because parasites often exploit closely related species (Poulin, 2010; Tschopp et al., 2013), this might indicate close relatedness of the host subgenera. In our phylogeny, *S. nassonowi* specimen from a host belonging to subgenus *Suandrena* was within a clade with specimens from subgenus *Plastandrena*. It has indeed been argued that *Suandrena* should be synonymized with *Plastandrena* (Pisanty et al., 2022), host subgenus of both *S. nassonowi* and *S. aterrimus*. Interestingly, *Hoplandrena*, the third subgenus that these two species utilise, is phylogenetically distant from the other two host subgenera (Pisanty et al., 2022). This indicates that at least some *Stylops* species are able to utilise even distantly related hosts although it seems to be the exception rather than the rule for the genus.

Effects of number of loci used in SODA

Coalescent-based species delimitation methods emerged over a decade ago in an era when datasets typically consisted of a handful of loci (Fujita et al., 2012). The advances in sequencing have made it feasible to acquire thousands of genes from almost any organism (Goodwin et al., 2016). It is important to understand how these increasing dataset sizes affect downstream analyses such as species delimitation. There are many species delimitation methods available for molecular data but the field is heavily dominated by those relying on the MSC model framework with gene trees evolving within the constraint of a species tree (Rannala & Yang, 2003; Yang & Rannala, 2010). Essentially, MSC-based species delimitation methods detect genetic structure (Sukumaran et al., 2021; Sukumaran & Knowles, 2017), since an assumption that coalescence within species should follow that of a single panmictic population is a central tenet of the original model. This is an assumption that is very often violated in real datasets, instead of showing genetic signs of population structure. In addition, the larger the datasets, the finer population genetic structure can be detected. There is, therefore, an increasing concern that the power of genomic-sized datasets will result in an inflation of delimited species if used incautiously, with vast implications for affected fields using species as units such as ecology, evolution and conservation (Sukumaran & Knowles, 2017). Indeed, simulations have found that the use of more loci increases the probability of evolutionary lineages being split, even in the face of significant gene flow between populations. The posterior support for splits increased with the number of RAD-loci used for several nodes in an empirical study of Hercules beetles (Huang, 2018). Leaché et al. (2019) provided numerical evidence that a two-species model will dominate the posterior when moving towards infinitely many loci, even when the migration between the two populations is so high that they would be considered a single species by any species definition. Hence, the concern is valid. These studies all used the popular BPP software (Yang & Rannala, 2010), but the risk of over-splitting is inherent in all methods assuming a neutral coalescence process within species.

SODA builds upon the success of quartet-based species-tree inference method ASTRAL (Rabiee & Mirarab, 2021). It makes no claim to be the most accurate method, but the strength lies in the ultrafast capacity to analyse big genomic datasets under the MSC while being not far behind competing slower methods in accuracy (Rabiee & Mirarab, 2021). Nodes in a guide tree (inferred by ASTRAL or provided by the user) are tested under the null hypothesis of zero length in coalescent units. Under the null hypothesis, each of three possible resolutions of a quartet should be equally frequent among gene trees. If the null hypothesis is rejected, the node remains as a split, but if zero length cannot be rejected, the node is collapsed into intraspecific branches (Rabiee & Mirarab, 2021). The assumption of equal quartet resolution frequencies under the null hypothesis follows from the MSC model assuming neutral coalescent processes within species. Detectable population structure will cause gene trees to deviate from equifrequent quartet resolutions and, hence, reject the null and inflate species splits. As admitted by the authors, SODA, thus, shares the same problematic assumptions that are causing methods like BPP to over-split (Rabiee & Mirarab, 2021). Our loci number resampling exercise is basically reiterating for SODA what has already been shown for BPP (Huang, 2018; Jackson et al., 2017; Leaché et al., 2019): adding loci tends to cause more species to split. The steady increase in the number of delimited species in our dataset from below 30 with 25 loci to over 70 with 2000 loci is striking (Figure 7).

The closest to a panacea solution for handling these problems is to not rely solely on molecular species delimitation but to use such analyses in an integrative taxonomic framework along with other lines of evidence (Chambers & Hillis, 2020; Dayrat, 2005; Fujita et al., 2012; Yang & Rannala, 2010; Zhang et al., 2011). Often, as in this study, this can be done a posteriori where we evaluate the results from SODA in light of evidence from hosts, geography, morphology and loci-number indifferent GMYC to make conservative conclusions. We preferred this as a first take on the complex *Stylops* situation in West Palaearctic. This meant not assigning individuals to populations a priori using for instance hosts, since that would prevent us from detecting host sharing of multiple parasite species (Smit et al., 2020). In the future, more detailed studies on different subclades could attempt to input more a priori information in a guide tree where individuals are assigned to populations which SODA has support for (Rabiee & Mirarab, 2021). A more informed prior on the candidate species model (the guide tree) is one of the best safety nets against exorbitant over-splitting. SODA also requires that the analysis includes at least two individuals per species to be accurate (Rabiee & Mirarab, 2021). However, including singletons as in our analysis only risks erroneous lumping (false positives, fig. 1 in Rabiee & Mirarab, 2021) and would not contribute to over-splitting. In addition, singletons may only be erroneously lumped if forming a cherry. For instance, we had only one individual of the inferred undescribed species on *Cnemidandrena* but this is still delimited as separate because of its position in the topology (sister to a clade with multiple delimited species; Figures 5 and 6).

Sampling more than one individual per species is one important aspect for species delimitation analyses (Zhang et al., 2011), and another is the geographical distribution of samples. The geographical sampling of species can also inflate species numbers if intermediate regions are omitted due to isolation-by-distance effects (Chambers & Hillis, 2020; Mason et al., 2020; Rousset, 1997; Wright, 1943). This may be a concern for the dataset analysed here. For example, *S. thwaitesi* sampled from Central and Northern Europe, but not continuously along the latitudinal axis, delimited six species. It is almost a catch 22 that widespread sampling across a species' range is optimal to appreciate the total genetic and phenotypic variation of a species (Bergsten et al., 2012), yet in practice, for phylogenomic species delimitation that same sampling strategy may be a contributing factor to over-splitting. With limited funds for sequencing, maximising the geographical spread of samples omits intermediate regions and introduces isolation-by-distance effects that fool current methods designed to delimit species when detecting genetic structure. Future studies should, preferably with a denser strategic sampling, focus on subgroups within WP *Stylops* to test still outstanding questions.

We hope our study provides a framework for the identification of subgroups as targets for in-depth future research.

CONCLUSIONS

Using phylogenomic analyses, we evaluated the conflicting historical and presented species diversity estimates in WP *Stylops* and inferred a minimum of 27 different species. We are able to add 35 new host associations for *Stylops* species based on the molecular identification of specimens, some of which have previously been hypothesised yet omitted from checklists and catalogues. Finally, we were able to confirm 10 cases where *Stylops* species associated with hosts from the same subgenus were synonyms. It is likely that multiple EP and Eastern Nearctic species parasitizing host from the same subgenus should be synonymized as well, as proposed by Straka et al., 2015. Many of these species were reinstated by Cook (2019) and revision is needed also for the EP and Eastern Nearctic faunas. Our sensitivity analysis of species delimitation with SODA using various numbers of loci showed a concerning pattern of a strong positive correlation between the number of delimited species and the number of loci used in the analysis. Our integrative approach of evaluating supported species and comparison with a crude loci-number indifferent method (GMYC) show datasets much smaller than today's commonly used genomic datasets that seem to give most reasonable estimates. We, therefore, urge that species delimitation results using genomic datasets with SODA or other methods showing similar tendencies are interpreted with great caution and in an integrative framework. We hope the molecular dataset generated in this study will be a valuable asset for future research on *Stylops* and parasite evolution in general.

AUTHOR CONTRIBUTIONS

Meri Lähteenaro: Investigation; data curation; formal analysis; writing – original draft; visualization. **Jakub Straka:** Conceptualization; resources; supervision. **Mattias Forshage:** Resources; writing – review and editing. **Rasmus Hovmöller:** Resources; writing – review and editing. **Yuta Nakase:** Resources; writing – review and editing. **Anders L. Nilsson:** Resources; writing – review and editing. **John T. Smit:** Resources; writing – review and editing. **Johan A. A. Nylander:** Supervision; formal analysis; writing – review and editing. **Johannes Bergsten:** Conceptualization; supervision; funding acquisition; writing – original draft; project administration.

ACKNOWLEDGEMENTS

This study was supported with funding from The Swedish Taxonomy Initiative to JB (Dha 2019.4.3-7 and Dha 2019.4.3-218) and Czech Science Foundation grant 20-14872S to JS. The National Academic Infrastructure for Supercomputing in Sweden (NAISS) and the Swedish National Infrastructure for Computing (SNIC) provided resources for the computing at UPPMAX. We are grateful for technical support and assistance from the National Genomics Infrastructure, SciLifeLab (NGI) and the Swedish Museum of Natural History DNA laboratory. We are especially grateful to Martin Irestedt for his guidance with the laboratory protocols. This study would not have been possible without the generous specimen loans and donations from private persons and institutions: Sven Hellqvist, Jörgen Persson, Michael Tholin, MZH (Helsinki) and curator Heidi Viljanen, ZMUC (Copenhagen) and curators Lars Vilhelmsen and Henning Bang Madsen, MZLU (Lund) and curator Rune Bygebjerg, GNM (Gothenburg) and curator Charlotte Jonsson, Jan Batelka, Daniel Benda, and Marek and Jiri Halada. The authors would like to thank Tom Wood for providing identifications to some included mining bee specimens.

FUNDING INFORMATION

This study was supported with funding from The Swedish Taxonomy Initiative to JB (Dha 2019.4.3-7 and Dha 2019.4.3-218) and Czech Science Foundation grant 20-14872S to JS.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The 2315 loci dataset that supports the findings of this study are openly available in SciLifeLab Data Repository at http://doi.org/10.17044/scilifelab.23865126.

ETHICS STATEMENT

Ethical approval was not required for the execution of this research owing to the fact that the research organisms comprised of invertebrates not affected by any guidelines.

ORCID

Meri Lähteenaro https://orcid.org/0000-0001-9868-3134
Jakub Straka https://orcid.org/0000-0002-8987-1245
Rasmus Hovmöller https://orcid.org/0009-0008-3232-7944
Yuta Nakase https://orcid.org/0000-0001-5925-4564
John T. Smit https://orcid.org/0000-0002-1568-5183
Johan A. A. Nylander https://orcid.org/0000-0001-8940-456X
Johannes Bergsten https://orcid.org/0000-0002-6153-4431

REFERENCES

Ailán-Choke, L.G. & Pereira, F.B. (2021) Deep in the systematics of Camallanidae (Nematoda): using integrative taxonomy to better understand the phylogeny and consistency of diagnostic traits. *Parasitology*, 148(8), 962–974. Available from: https://doi.org/10.1017/S0031182021000706

Balzer, Z.S. & Davis, A.R. (2020) Life history traits and interactions of *Stylops advarians* (Strepsiptera) with its bee host, Andrena milwaukeensis. *Parasitology*, 147(4), 410–417. Available from: https://doi.org/10.1017/S0031182020000037

Benda, D., Votýpková, K., Nakase, Y. & Straka, J. (2020) Unexpected cryptic species diversity of parasites of the family Xenidae (Strepsiptera) with a constant diversification rate over time. *Systematic Entomology*, 46(1), 252–265. Available from: https://doi.org/10.1111/syen.12460

Bergsten, J., Bilton, D.T., Fujisawa, T., Elliott, M., Monaghan, M.T., Balke, M. et al. (2012) The effect of geographical scale of sampling on DNA barcoding. *Systematic Biology*, 61(5), 851–869.

Borchert, H.M. (1971) Benennung unseres Beliner *Stylops* aus Sandbiene *Andrena vaga* Pz. *Stylops muelleri* sp. N. (Insecta, Strepsiptera). *Sitzungsberichte der Gesellschaft Naturforschender Freunde zu Berlin*, 1, 18–27.

Brandvain, Y. & Matute, D.R. (2018) When genes move, genomes collide. *PLoS Genetics*, 14(4), e1007286. Available from: https://doi.org/10.1371/journal.pgen.1007286

Bravo, G.A., Antonelli, A., Bacon, C.D., Bartoszek, K., Blom, M.P.K., Huynh, S. et al. (2019) Embracing heterogeneity: coalescing the tree of life and the future of phylogenomics. *PeerJ*, 7, e6399. Available from: https://doi.org/10.7717/peerj.6399

Brower, A.V.Z. (2006) Problems with DNA barcodes for species delimitation: 'Ten species' of *Astraptes fulgerator* reassessed (Lepidoptera: Hesperiidae). *Systematics and Biodiversity*, 4(2), 127–132. Available from: https://doi.org/10.1017/S147720000500191X

Brullé, A. (1832) Insectes: expedition scientifique de Morée. *Zoologie*, Section des Sciences Physiques, Tome 3, Partie 1, Zoologie, Sect. 2, Des Animaux Articulés 17–29, 64–395.

Brunet, B.M.T., Blackburn, G.S., Muirhead, K., Lumley, L.M., Boyle, B., Lévesque, R.C. et al. (2017) Two's company, three's a crowd: new insights on spruce budworm species boundaries using genotyping-by-sequencing in an integrative species assessment (Lepidoptera: Tortricidae). *Systematic Entomology*, 42(2), 317–328. Available from: https://doi.org/10.1111/syen.12211

Cerca, J., Meyer, C., Purschke, G. & Struck, T.H. (2020) Delimitation of cryptic species drastically reduces the geographical ranges of marine interstitial ghost-worms (Stygocapitella; Annelida, Sedentaria). *Molecular Phylogenetics and Evolution*, 143, 106663. Available from: https://doi.org/10.1016/j.ympev.2019.106663

Černá, K., Straka, J. & Munclinger, P. (2013) Population structure of pioneer specialist solitary bee *Andrena vaga* (Hymenoptera: Andrenidae) in Central Europe: the effect of habitat fragmentation or evolutionary history? *Conservation Genetics*, 14(4), 875–883. Available from: https://doi.org/10.1007/s10592-013-0482-y

Chambers, E.A. & Hillis, D.M. (2020) The multispecies coalescent oversplits species in the case of geographically widespread taxa. *Systematic Biology*, 69(1), 184–193. Available from: https://doi.org/10.1093/sysbio/syz042

Cook, J.L. (2014) Review of the biology of parasitic insects in the order Strepsiptera. *Comparative Parasitology*, 81(2), 134–151. Available from: https://doi.org/10.1654/4723.1

Cook, J.L. (2019) Annotated catalog of the order Strepsiptera of the world. *Transactions of the American Entomological Society*, 145(2), 121–267. Available from: https://doi.org/10.3157/061.145.0202

Criscuolo, A. & Gribaldo, S. (2010) BMGE (block mapping and gathering with entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evolutionary Biology*, 10(1), 210. Available from: https://doi.org/10.1186/1471-2148-10-210

Curtis, J. (1828) Stylops dalii. *British Entomology*, 5, 195–241.

Davis, E.S., Murray, T.E., Fitzpatrick, Ú., Brown, M.J.F. & Paxton, R.J. (2010) Landscape effects on extremely fragmented populations of a rare solitary bee, Colletes floralis. *Molecular Ecology*, 19(22), 4922–4935. Available from: https://doi.org/10.1111/j.1365-294X.2010.04868.x

Dayrat, B. (2005) Towards integrative taxonomy. *Biological Journal of the Linnean Society*, 85(3), 407–417.

de Queiroz, K. (2007) Species concepts and species delimitation. *Systematic Biology*, 56, 8–886.

DeSalle, R., Egan, M.G. & Siddall, M. (2005) The unholy trinity: taxonomy, species delimitation and DNA barcoding. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1462), 1905–1916. Available from: https://doi.org/10.1098/rstb.2005.1722

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32(19), 3047–3048.

Ezard, T., Fujisawa, T. & Barraclough, T. (2009) splits: SPecies' LImits by Threshold Statistics. R package version 1.0-14/r31.

Fabricius, J.C. (1775) *Systema entomologiae: sistens insectorum classes, ordines, genera, species, adiectis synonymis, locis, descriptionibus, observationibus*. In Officina Libraria Kortii.

Forbes, A.A., Bagley, R.K., Beer, M.A., Hippee, A.C. & Widmayer, H.A. (2018) Quantifying the unquantifiable: why Hymenoptera, not Coleoptera, is the most speciose animal order. *BMC Ecology*, 18(1), 21. Available from: https://doi.org/10.1186/s12898-018-0176-x

Fraulob, M., Beutel, R.G., Machida, R. & Pohl, H. (2015) The embryonic development of *Stylops ovinae* (Strepsiptera, Stylopidae) with emphasis on external morphology. *Arthropod Structure & Development*, 44(1), 42–68. Available from: https://doi.org/10.1016/j.asd.2014.10.001

Friese, H. (1899) Die Bienen Europa's (Apidae europaeae) nach ihren Gattungen, Arten und Varietäten auf vergleichend morphologisch-biologischer Grundlage. Theil V: Solitäre Apiden: Genus *Lithurgus*, Genus *Megachile (Chalicodoma)*. C. Lampe, Innsbruck und Imst, 228.

Fujisawa, T. & Barraclough, T.G. (2013) Delimiting species using single-locus data and the generalized mixed yule coalescent approach: a revised method and evaluation on simulated data sets. *Systematic Biology*, 62(5), 707–724.

Fujita, M.K., Leaché, A.D., Burbrink, F.T., McGuire, J.A. & Moritz, C. (2012) Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution*, 27(9), 480–488. Available from: https://doi.org/10.1016/j.tree.2012.04.012

Goodwin, S., McPherson, J.D. & McCombie, W.R. (2016) Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, 17(6), 333–351. Available from: https://doi.org/10.1038/nrg.2016.49

Günther, V. & Šedivi, J. (1957) Řád Řasnokřídlí—Strepsiptera., 407–417. In: *Klíč zvířeny ČSR. Díl II. [Key to the fauna of Czechoslovakia]*. Praha: ČSAV, p. 746.

Hayward, A., Mcmahon, D.P. & Kathirithamby, J. (2011) Cryptic diversity and female host specificity in a parasitoid where the sexes utilize hosts from separate orders: cryptic diversity and female host specificity. *Molecular Ecology*, 20(7), 1508–1528. Available from: https://doi.org/10.1111/j.1365-294X.2011.05010.x

Hedicke, H. (1933) Beiträge zur Systematik der Gattung *Andrena* F. (Hym. Apid.). *Mitteilungen aus dem Zoologischen Museum in Berlin*, 19, 199–220.

Heled, J. & Drummond, A.J. (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, 27(3), 570–580. Available from: https://doi.org/10.1093/molbev/msp274

Heraty, J.M., Woolley, J.B., Hopper, K.R., Hawks, D.L., Kim, J.-W. & Buffington, M. (2007) Molecular phylogenetics and reproductive incompatibility in a complex of cryptic species of aphid parasitoids. *Molecular Phylogenetics and Evolution*, 45(2), 480–493. Available from: https://doi.org/10.1016/j.ympev.2007.06.021

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q. & Vinh, L.S. (2018) UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35(2), 518–522. Available from: https://doi.org/10.1093/molbev/msx281

Hofeneder, K. (1923) *Stylops* in copula. *Verhandlungen Der Zoologische-Botanischen Gesellschaft*, 74, 128–134.

Hofeneder, K. (1924) Wissenschaftliche Ergebnisse der 2. Expedition nach dem Anglo-ägyptischen Sudan 1914. XVI. Hymenoptera. I. Anhang: Strepsiptera. *Denkschriften—Akademie Der Wisseschaften, Mathematisch-Naturwissenschaftliche Klasse*, 99, 254.

Hofeneder, K. (1939) Über zwei neue Strepsipteren aus *Andrena similis* F. Smith und *Halictus malachurus* Kirby. *Veröffentlichungen Aus Dem Deutschen Kolonial-Und Übersee-Museum in Bremen*, 2, 181–200.

Huang, J.-P. (2018) What have been and what can be delimited as species using molecular data under the multi-species coalescent model? A case study using Hercules beetles (*Dynastes*; Dynastidae). *Insect*

Systematics and Diversity, 2(2), 3. Available from: https://doi.org/10.1093/isd/ixy003

Illumina. (2018) Nextera™ DNA Flex Library Prep Reference Guide # 1000000025416 v01. Illumina, Inc. San Diego, California, USA.

Imhoff, L. (1832) Entomologica. *Isis*, 1832, 1198–1208.

Imhoff, L. (1868) Die schweizerischen Arten der Gattung *Andrena* F.(partim), Latr.(maxima parte), Leach. *Mitteilungen der Schweizerischen Entomologischen Gesellshaft*, 2, 33–48.

Irestedt, M., Thörn, F., Müller, I.A., Jønsson, K.A., Ericson, P.G.P. & Blom, M.P.K. (2022) A guide to avian museomics: insights gained from resequencing hundreds of avian study skins. *Molecular Ecology Resources*, 22(7), 2672–2684. Available from: https://doi.org/10.1111/1755-0998.13660

Jackson, N.D., Carstens, B.C., Morales, A.E. & O'Meara, B.C. (2017) Species delimitation with gene flow. *Systematic Biology*, 66(5), 799–812. Available from: https://doi.org/10.1093/sysbio/syw117

Jandausch, K., Michels, J., Kovalev, A., Gorb, S.N., van de Kamp, T., Beutel, R.G. et al. (2022) Have female twisted-wing parasites (Insecta: Strepsiptera) evolved tolerance traits as response to traumatic penetration? *PeerJ*, 10, e13655. Available from: https://doi.org/10.7717/peerj.13655

Jůzová, K., Nakase, Y. & Straka, J. (2015) Host specialization and species diversity in the genus *Stylops* (Strepsiptera: Stylopidae), revealed by molecular phylogenetic analysis: host specialization in *Stylops*. *Zoological Journal of the Linnean Society*, 174(2), 228–243. Available from: https://doi.org/10.1111/zoj.12233

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A. & Jermiin, L.S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6), 587–589. Available from: https://doi.org/10.1038/nmeth.4285

Kathirithamby, J. (2009) Host-parasitoid associations in Strepsiptera. *Annual Review of Entomology*, 54(1), 227–249. Available from: https://doi.org/10.1146/annurev.ento.54.110807.090525

Kathirithamby, J. (2018) Biodiversity of Strepsiptera. In: Foottit, R.G. & Adler, P.H. (Eds.) *Insect biodiversity*. New York: John Wiley & Sons, Ltd., pp. 673–704. Available from: https://doi.org/10.1002/9781118945582.ch22

Kathirithamby, J. & Engel, M.S. (2014) A revised key to the living and fossil families of Strepsiptera, with the description of a new family, Cretostylopidae. *Journal of the Kansas Entomological Society*, 87(4), 385–388. Available from: https://doi.org/10.2317/JKES140407.1

Katoh, K., Misawa, K., Kuma, K. & Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30(14), 3059–3066. Available from: https://doi.org/10.1093/nar/gkf436

Katoh, K. & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772–780. Available from: https://doi.org/10.1093/molbev/mst010

Kifune, T. & Hirashima, Y. (1985) Nine new species of the genus *Stylops* (Strepsiptera: Stylopidae) parasitic on the genus *Andrena* (Hymenoptera: Andrenidae) of Japan (studies on the Japanese Strepsiptera X). *Esakia*, 23, 45–57. Available from: https://doi.org/10.5109/2467

Kinzelbach, R. (1971) *Morphologische Befunde an Fächerflüglern und ihre phylogenetische Bedeutung (Insecta: Strepsiptera)*. Stuttgart: Schweizerbart'sche Verlagsbuchhandlung, p. 256.

Kinzelbach, R. (1978) Fächerflügler (Strepsiptera). In: Senglaub, K. & Hannemann, H.J. (Eds.) *Die Tierwelt Deutschlands*, Vol. 65. Jena, Germany: Gustav Fischer. pp. 1–166.

Kirby, W. (1802) *Monographia Apum Angliae; or, an attempt to divide into their natural genera and families, such species of the Linnean genus Apis as have been discovered in England: With descriptions and observations. To which are prefixed some introductory remarks upon the class Hymenoptera, and a synoptical table of the nomenclature of the external parts of these Insects*. Ipswich, Printed for the author by J. Raw, 1802.

Knyshov, A., Gordon, E.R.L. & Weirauch, C. (2021) New alignment-based sequence extraction software (ALiBaSeq) and its utility for deep level phylogenetics. *PeerJ*, 9, e11019. Available from: https://doi.org/10.7717/peerj.11019

Korneliussen, T.S., Albrechtsen, A. & Nielsen, R. (2014) ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics*, 15(1), 356. Available from: https://doi.org/10.1186/s12859-014-0356-4

Kozlov, A.M., Darriba, D., Flouri, T., Morel, B. & Stamatakis, A. (2019) RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, 35(21), 4453–4455. Available from: https://doi.org/10.1093/bioinformatics/btz305

Kuhlmann, M., Ascher, J.S., Dathe, H.H., Ebmer, A.W., Hartmann, P., Michez, D. et al. (2023) Checklist of the Western Palaearctic Bees (Hymenoptera: Apoidea: Anthophila). Available from: http://westpalbees.myspecies.info

Lanfear, R., Calcott, B., Kainer, D., Mayer, C. & Stamatakis, A. (2014) Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evolutionary Biology*, 14(1), 82. Available from: https://doi.org/10.1186/1471-2148-14-82

Leaché, A.D., Zhu, T., Rannala, B. & Yang, Z. (2019) The spectre of too many species. *Systematic Biology*, 68(1), 168–181. Available from: https://doi.org/10.1093/sysbio/syy051

Lepeletier, A.L.M. (1841) Histoire Naturelle des Insectes. *Hyménoptères*, 2. Librairie Encyclopédique de Roret, 680.

Li, H. & Durbin, R. (2009) Fast and accurate short read alignment with burrows–wheeler transform. *Bioinformatics*, 25(14), 1754–1760. Available from: https://doi.org/10.1093/bioinformatics/btp324

Linnaeus, C. (1758) Systema naturae per regna tria naturae, secundum classes, ordinus, genera, species, cum characteribus, differentiis, synonymis, locis. *Editio decima, reformata. Holmiae (Laurentii Salvii)*. 1-4, 1–824.

Löwe, S., Beutel, R. & Pohl, H. (2016) The female cephalothorax of *Stylops ovinae* Noskiewicz & Poluszyński, 1928 (Strepsiptera: Stylopida). *Arthropod Systematics & Phylogeny*, 74, 65–81.

Luna de Carvalho, E. (1974) Contribuição para o estudo dos *Stylops* da Peninsula Iberica (Streps. Stylopidae). *Eos: Revista Española de Entomologia*, 48, 301–365.

Luo, A., Ling, C., Ho, S.Y.W. & Zhu, C.-D. (2018) Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Systematic Biology*, 67(5), 830–846. Available from: https://doi.org/10.1093/sysbio/syy011

Mai, U. & Mirarab, S. (2018) TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. *BMC Genomics*, 19(5), 272. Available from: https://doi.org/10.1186/s12864-018-4620-2

Manni, M., Berkeley, M. R., Seppey, M., & Zdobnov, E. M. (2021). BUSCO: assessing genomic data quality and beyond. *Current Protocols*, 1(12), e323.

Mason, N.A., Fletcher, N.K., Gill, B.A., Funk, W.C. & Zamudio, K.R. (2020) Coalescent-based species delimitation is sensitive to geographic sampling and isolation by distance. *Systematics and Biodiversity*, 18(3), 269–280. Available from: https://doi.org/10.1080/14772000.2020.1730475

McKay, B.D., Mays, H.L., Jr., Wu, Y., Li, H., Yao, C., Nishiumi, I. et al. (2013) An empirical comparison of character-based and coalescent-based approaches to species delimitation in a young avian complex. *Molecular Ecology*, 22(19), 4943–4957. Available from: https://doi.org/10.1111/mec.12446

McKenna, D.D., Shin, S., Ahrens, D., Balke, M., Beza-Beza, C., Clarke, D.J. et al. (2019) The evolution and genomic basis of beetle diversity. *Proceedings of the National Academy of Sciences*, 116(49), 24729–24737. Available from: https://doi.org/10.1073/pnas.1909655116

Medvedev, L.N. (1970) A family of Strepsiptera new to the fauna of the USSR and new species from south-eastern Kazakhstan. *Entomological Review*, 49, 113–114.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A. et al. (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37(5), 1530–1534. Available from: https://doi.org/10.1093/molbev/msaa015

Morawitz, F. (1877) Nachtrag zur Bienenfauna Caucasiens. *Horae Societatis Entomologicae Rossicae*, 14(1), 3–112.

Morel, B., Kozlov, A.M. & Stamatakis, A. (2019) ParGenes: a tool for massively parallel model selection and phylogenetic tree inference on thousands of genes. *Bioinformatics*, 35(10), 1771–1773. Available from: https://doi.org/10.1093/bioinformatics/bty839

Nadler, S.A. & Pérez-Ponce de León, G. (2011) Integrating molecular and morphological approaches for characterizing parasite cryptic species: implications for parasitology. *Parasitology*, 138(13), 1688–1709. Available from: https://doi.org/10.1017/S003118201000168X

Nakase, Y. & Kato, M. (2013) Cryptic diversity and host specificity in giant *Xenos* Strepsipterans parasitic in large *Vespa* hornets. *Zoological Science*, 30(4), 331–336. Available from: https://doi.org/10.2108/zsj.30.331

Newport, G. (1851) The natural history, anatomy and development of *Meloe* (continued). Second memoir. The history and general anatomy of *Meloe*, and its affinities, compared with those of the Strepsiptera and Anoplura, with reference to the connection which exists between structure, function, and instinct. *Transactions of the Linnean Society of London*, 20, 321–357.

Nguyen, L.-T., Schmidt, H.A., von Haeseler, A. & Minh, B.Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32(1), 268–274. Available from: https://doi.org/10.1093/molbev/msu300

Ortiz, D., Pekár, S., Bilat, J. & Alvarez, N. (2021) Poor performance of DNA barcoding and the impact of RAD loci filtering on the species delimitation of an Iberian ant-eating spider. *Molecular Phylogenetics and Evolution*, 154, 106997. Available from: https://doi.org/10.1016/j.ympev.2020.106997

Padial, J.M., Miralles, A., De la Riva, I. & Vences, M. (2010) The integrative future of taxonomy. *Frontiers in Zoology*, 7(1), 16. Available from: https://doi.org/10.1186/1742-9994-7-16

Panzer, G.W.F. (1799) Faunae Insectorum Germanicae initia oder Deutschlands Insecten. *64*, Felseckersche Buchhandlung.

Paradis, E. & Schliep, K. (2019) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35(3), 526–528. Available from: https://doi.org/10.1093/bioinformatics/bty633

Pasteels, J. (1949) Enquêtes sur les Strepsiptères (1re série). *Bulletin et Annales de La Société Entomologique*, 85, 185–196.

Pasteels, J. (1954) Pasteels, J. (1954). Enquêtes sur les Strepsiptères (4e série)—Strepsiptères du Musée zoologique de Lausanne. *Bulletin et Annales de La Société Entomologique de Belgique*, 90, 349–355.

Peinert, M., Wipfler, B., Jetschke, G., Kleinteich, T., Gorb, S.N., Beutel, R.G. et al. (2016) Traumatic insemination and female counter-adaptation in Strepsiptera (Insecta). *Scientific Reports*, 6(1), 25052. Available from: https://doi.org/10.1038/srep25052

Pérez, J. (1895) *Espèces nouvelles de mellifères de Barbarie. (Diagnoses préliminaires)*. Gounouilhou, Bordeux, p. 64.

Pérez, J. (1902) Espèces nouvelles de Mellifères (paléarctiques). *Procès-verbaux Société Linnean de Bordeaux*, 57, 174–180.

Pérez-Ponce de León, G. & Nadler, S.A. (2016) The importance of recognising parasite cryptic diversity for research programmes on food-borne trematodiases. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 110(1), 4–5. Available from: https://doi.org/10.1093/trstmh/trv090

Perkins, R.C.L. (1918) Synopsis of British Strepsiptera of the genera *Stylops* and *Halictoxenus*. *Entomologist's Monthly Magazine*, 54, 67–76.

Pierce, W.D. (1909) A monographic revision of the twisted winged insects comprising the order Strepsiptera Kirby. *Bulletin of the United States National Museum*, 66(I–XII), 1–232.

Pierce, W.D. (1911) Notes on insects of the order Strepsiptera, with descriptions of new species. *Proceedings of the United States National Museum*, 40(1834), 487–511. Available from: https://doi.org/10.5479/si.00963801.1834.487

Pisanty, G., Richter, R., Martin, T., Dettman, J. & Cardinal, S. (2022) Molecular phylogeny, historical biogeography and revised classification of andrenine bees (Hymenoptera: Andrenidae). *Molecular Phylogenetics and Evolution*, 170, 107151. Available from: https://doi.org/10.1016/j.ympev.2021.107151

Poluszynski, G. (1927) Stylops nitidiusculae n. sp. *Bulletin Entomologique de Pologne*, 6, 92–99.

Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S. et al. (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, 55(4), 595–609. Available from: https://doi.org/10.1080/10635150600852011

Poulin, R. (2010) Decay of similarity with host phylogenetic distance in parasite faunas. *Parasitology*, 137(4), 733–741. Available from: https://doi.org/10.1017/S0031182009991491

R Core Team. (2022) *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available from: https://www.R-project.org/

Rabiee, M. & Mirarab, S. (2021) SODA: multi-locus species delimitation using quartet frequencies. *Bioinformatics*, 36(24), 5623–5631. Available from: https://doi.org/10.1093/bioinformatics/btaa1010

Rambaut, A. (2018) FigTree v1.4.4. Available from: http://tree.bio.ed.ac.uk/software/figtree/

Rannala, B. & Yang, Z. (2003) Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*, 164(4), 1645–1656. Available from: https://doi.org/10.1093/genetics/164.4.1645

Reichert, A. (1914) Aus der Wochenstube einer Strepsipteren-Mutter. *Entomologisches Jahrbuch*, 23, 148–152.

Rousset, F. (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, 145(4), 1219–1228. Available from: https://doi.org/10.1093/genetics/145.4.1219

Salt, G. (1927) The effects of stylopization on aculeate hymenoptera. *Journal of Experimental Zoology*, 48(1), 223–331. Available from: https://doi.org/10.1002/jez.1400480107

Salt, G. (1931) A further study of the effects of stylopization on wasps. *Journal of Experimental Zoology*, 59(1), 133–166. Available from: https://doi.org/10.1002/jez.1400590107

Schenck, A. (1853) Nachtrag zu der Beschreibung nassauischer Bienenarten. *Jahrbücher des Vereins für Naturkunde im Herzogthum Nassau*, 9, 88–306.

Schenck, A. (1861) Die nassauischen Bienen. Revision und Ergänzung der früheren Bearbeitungen. *Jahrbuch des Vereins für Naturkunde Nassau*, 14(1859), 1–414.

Schmid-Egger, C. (2005) Proxiandrena subgen. nov. und Revision der west-und zentralpaläarktischen Arten der *Andrena proxima*-Gruppe (Hymenoptera, Apidae). *Revue Suisse de Zoologie*, 112(4), 1029–1044.

Schmiedeknecht, O. (1883) *Apidae Europaeae (Die Bienen Europas) per genera, species et varietates dispositae atque descriptae*. Tomus I: *Nomada*, *Bombus*, *Psithyrus* et *Andrena*. Privately published.

Schubert, M., Lindgreen, S. & Orlando, L. (2016) AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Research Notes*, 9(1), 88. Available from: https://doi.org/10.1186/s13104-016-1900-2

Scopoli, J.A. (1770) *Dissertatio de Apibus*. Annus IV Historico-Naturalis, Lipsiae: Christ. Gottlob. Hilscher, p. 130.

Smit, J.T., Smit, J., Raemakers, I.P. & van der Hoorn, B. (2020) The Strepsiptera of the Netherlands revisited (Insecta). *Entomologische Berichten*, 80, 8–30.

Smith, F. (1847) Description of British bees belonging to the genus *Andrena* of Fabricius. *Zoologist*, 5, 1732–1753.

Smith, M.A., Rodriguez, J.J., Whitfield, J.B., Deans, A.R., Janzen, D.H., Hallwachs, W. et al. (2008) Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. *Proceedings of the National Academy of Sciences*, 105(34), 12359–12364. Available from: https://doi.org/10.1073/pnas.0805319105

Smith, M.A., Wood, D.M., Janzen, D.H., Hallwachs, W. & Hebert, P.D.N. (2007) DNA barcodes affirm that 16 species of apparently generalist tropical parasitoid flies (Diptera, Tachinidae) are not all generalists. *Proceedings of the National Academy of Sciences*, 104(12), 4967–4972. Available from: https://doi.org/10.1073/pnas.0700050104

Straka, J., Alqarni, A., Juzova, K., Hannan, M., Hinojosa-Díaz, I. & Engel, M. (2015) Rediscovered parasitism of *Andrena savignyi* Spinola (Hymenoptera, Andrenidae) by *Stylops* (Strepsiptera, Stylopidae) and revised taxonomic status of the parasite. *ZooKeys*, 519, 117–139. Available from: https://doi.org/10.3897/zookeys.519.6035

Straka, J., Jůzová, K. & Nakase, Y. (2015) Nomenclature and taxonomy of the genus *Stylops* (Strepsiptera): an annotated preliminary world checklist. *Acta Entomologica Musei Nationalis Pragae*, 55(1), 305-332.

Sudasinghe, H., Pethiyagoda, R. & Meegaskumbura, M. (2020) A molecular phylogeny of the genus *Laubuka* (Teleostei: Cyprinidae) in Sri Lanka reveals multiple origins and a cryptic species. *Systematics and Biodiversity*, 18(6), 592–613. Available from: https://doi.org/10.1080/14772000.2020.1771468

Sukumaran, J., Holder, M.T. & Knowles, L.L. (2021) Incorporating the speciation process into species delimitation. *PLoS Computational Biology*, 17(5), e1008924. Available from: https://doi.org/10.1371/journal.pcbi.1008924

Sukumaran, J. & Knowles, L.L. (2017) Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences*, 114(7), 1607–1612. Available from: https://doi.org/10.1073/pnas.1607921114

Talavera, G., Dincă, V. & Vila, R. (2013) Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods in Ecology and Evolution*, 4(12), 1101–1110.

Tang, C.Q., Humphreys, A.M., Fontaneto, D. & Barraclough, T.G. (2014) Effects of phylogenetic reconstruction method on the robustness of species delimitation using single-locus data. *Methods in Ecology and Evolution*, 5(10), 1086–1094.

Thomson, C. (1870) *Opuscula entomologica*. *Hakan Ohlson*, Lund, p. 222.

Tschopp, A., Riedel, M., Kropf, C., Nentwig, W. & Klopfstein, S. (2013) The evolution of host associations in the parasitic wasp genus *Ichneumon* (Hymenoptera: Ichneumonidae): convergent adaptations to host pupation sites. *BMC Evolutionary Biology*, 13(1), 74. Available from: https://doi.org/10.1186/1471-2148-13-74

van der Vecht, J. (1927) Über eine neue holländische *Andrena*-art aus der *A. ovatula* K.-Gruppe, *Andrena gelriae n. sp*. *Zoologische Mededelingen*, 10(7), 87–89.

Warncke, K. (1968) Die Untergattungen der westpaläarktischen Bienengattung *Andrena* F. *Memórias e Estudos do Museu Zoológico da Universidade de Coimbra*, 307, 1–110.

Williams, J.L., Zhang, Y.M., LaPolla, J.S., Schultz, T.R. & Lucky, A. (2022) Phylogenomic delimitation of morphologically cryptic species in globetrotting *Nylanderia* (Hymenoptera: Formicidae) species complexes. *Insect Systematics and Diversity*, 6(1), 10. Available from: https://doi.org/10.1093/isd/ixab027

Wright, S. (1943) Isolation by distance. *Genetics*, 28(2), 114–138.

Yang, Z. (2015) The BPP program for species tree estimation and species delimitation. *Current Zoology*, 61(5), 854–865. Available from: https://doi.org/10.1093/czoolo/61.5.854

Yang, Z. & Rannala, B. (2010) Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences*, 107(20), 9264–9269. Available from: https://doi.org/10.1073/pnas.0913022107

Yates, J.A.F., Lamnidis, T.C., Borry, M., Valtueña, A.A., Fagernäs, Z., Clayton, S. et al. (2021) Reproducible, portable, and efficient ancient genome reconstruction with nf-core/eager. *PeerJ*, 9, e10947. Available from: https://doi.org/10.7717/peerj.10947

Zetterstedt, J.W. (1838) *Insecta Lapponica*. *Sectio secunda*. *Hymenoptera*. *Lipsiae*, 358–408.

Zhang, C., Rabiee, M., Sayyari, E. & Mirarab, S. (2018) ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, 19(6), 153. Available from: https://doi.org/10.1186/s12859-018-2129-y

Zhang, C., Zhang, D.-X., Zhu, T. & Yang, Z. (2011) Evaluation of a Bayesian coalescent method of species delimitation. *Systematic Biology*, 60(6), 747–761. Available from: https://doi.org/10.1093/sysbio/syr071

SUPPORTING INFORMATION

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

Figure S1: Maximum likelihood tree including long branches that were removed from downstream analyses. All nodes were maximally (100) supported by ultrafast bootstrap except those denoted with coloured circles. The scale bar indicates the expected number of substitutions per site. Terminal labels contain the following information separated by underscores: Voucher code, *Stylops* species, Host species (*Andrena*), sampling country in ISO 3166-1 alpha-2 abbreviation code and the host subgenus.

Figure S2: Lineage-through-time plot of ultrametric tree used for the GMYC analysis. The maximum likelihood solution is indicated with a red line and the 2-log likelihood confidence interval is denoted by a blue area.

Table S1: List of studied material.

How to cite this article: Lähteenaro, M., Straka, J., Forshage, M., Hovmöller, R., Nakase, Y., Nilsson, A.L. et al. (2024) Phylogenomic species delimitation of the twisted-winged parasite genus *Stylops* (Strepsiptera). *Systematic Entomology*, 49(2), 294–313. Available from: https://doi.org/10.1111/syen.12618