INTRODUCTION

Enriching subsets of the genome prior to sequencing (target enrichment, Mamanova et al., 2010) allows effort to be concentrated on genomic regions that are relevant to answer specific research questions. Using this approach also contributes to reducing cost, as experimental design can be adapted to sequence many samples simultaneously. In the fields of ecology and evolution, target enrichment has been used for genotyping or phylogenomics of plant and animal species (Gasc, Peyretaillade, & Peyret, 2016; Lemmon & Lemmon, 2013), to characterize phenotypic traits (e.g., Muraya, Schmutzer, Ulpinnis, Scholz, & Altmann, 2015) or to explore microbial ecosystems (Gasc & Peyret, 2018).

However, routine target enrichment by research laboratories is limited both by the complexity of current protocols and low input DNA quantity. Thus, working with tiny organisms such as microarthropods can be challenging. Here, we propose easy to set up optimizations for DNA extraction and library preparation prior to target enrichment. Prepared libraries were used to capture 1,432 ultraconserved elements (UCEs) from microhymenoptera (Chalcidoidea), which are among the tiniest insects on Earth and the most commercialized worldwide for biological control purposes. Results show no correlation between input DNA quantities (1.8–250 ng, 0.4 ng with an extra whole genome amplification step) and the number of sequenced UCEs on an Illumina MiSeq. Phylogenetic inferences highlight the potential of UCEs to solve relationships within the families of chalcid wasps, which has not been achieved so far. The protocol (library preparation + target enrichment) allows processing 96 specimens in five working days, by a single person, without requiring the use of expensive robotic molecular biology platforms, which could help to generalize the use of target enrichment for minute specimens.

KEYWORDS
Chalcidoidea, library construction, low DNA quantity, microarthropods, target enrichment, UCEs
microarthropods <1 mm) and/or old/rare (museum) specimens. Indeed, current protocols (DNA extraction, library preparation, target enrichment) are time-consuming and require handling expertise. They have been initially developed to work with large amounts of input DNA (e.g., vertebrates or large/medium-size insects; Faircloth, Branstetter, White, & Brady, 2015; McCormack et al., 2013) and include many purification steps that increase DNA loss. Working on hyperdiverse groups of microarthropods is challenging, as it requires one to perform the extraction on (a) a large number of specimens/species to be representative of the overall diversity of the group, without the possibility of using pipetting robots that increase DNA loss, (b) single individuals because species complexes are frequent (Al Khatib et al., 2014; Kenyon, Buerki, Hansson, Alvarez, & Benrey, 2015; Mottern & Heraty, 2014), (c) the whole insect without destruction for vouchering and often prior to species identification, (d) rare species that have been collected once and may be represented in collections by a few specimens or only one specimen and, sometimes, (e) old and dry museum specimens used for species description (types).

In this study, we propose optimized protocols for DNA extraction and library preparation for target enrichment purposes, as well as a custom pipeline to analyse the sequence data obtained. We used these protocols and customized pipeline to capture and analyse ultraconserved elements (UCEs) in minute wasps, the chalcids (Insecta: Hymenoptera: Chalcidoidea, Heraty et al., 2013; Noyes, 2018), that are key components of terrestrial ecosystems. Chalcids are key models for basic and applied research. With an estimated diversity of more than 500,000 species, these microhymenoptera have colonized almost all extant terrestrial habitats. Many of them develop as parasitoids of arthropod eggs, larvae or pupae. As such, they are both key regulators of the populations of many other arthropod species in natural ecosystems and are increasingly used worldwide as biocontrol agents (e.g., Consoli et al., 2010; Heraty, 2009). A few of them, especially Nasonia (Pteromalidae) or Trichogramma (Trichogrammatidae) species, are also used as model systems to answer challenging questions about sex determination, genetics of speciation, host-symbiont interactions or behavioural ecology (e.g., Pinto, Stouthamer, Platner, & Oatman, 1991; Stouthamer, Luck, & Hamilton, 1990; Warren & Loehlin, 2009). Chalcidoidea has undergone a spectacular radiation resulting in a huge diversity of morphologies and sizes (Gibson, Heraty, & Woolley, 1999; Heraty et al., 2013), but are generally small insects (<2 mm long). Among them, Kiki kuna Huber (Mymaridae) at 158 µm long is the smallest winged insect currently known, and the wingless male of Dicopomorpha echmepterygis Mockford at 139 µm is the smallest insect currently known (Huber & Noyes, 2013). Notably, most species used for biological control, belonging mainly to five families (Aphelinidae, Encyrtidae, Eulophidae, Mymaridae and Trichogrammatidae), are among the tiniest wasps on earth (<1 mm).

Their small size, huge diversity and widespread morphological convergence make chalcid wasps difficult to identify to species by nonexpert taxonomists, which limits their use in biological control. Attempts have been made to resolve the phylogeny of the whole superfamily (Heraty et al., 2013; Munro et al., 2011) or a few families (Burks, Heraty, Gebiola, & Hansson, 2011; Chen, Hui, Fu, & Huang, 2004; Cruaud et al., 2012; Desjardins, Regier, & Mitter, 2007; Janštá et al., 2018; Owen, George, Pinto, & Heraty, 2007), but none has succeeded. Indeed, the few markers that could be targeted with Sanger sequencing were not informative enough to solve deeper relationships. A study based on transcriptome data (3,239 single-copy genes) obtained from 37 species of chalcids and 11 outgroups also failed to solve relationships within the superfamily (Peters et al., 2018). As only a representative sampling in both markers and taxa will allow one to draw accurate conclusions on the history of this hyperdiverse group, target enrichment approaches appear relevant. More specifically, targeting UCEs and their flanking regions that have been proven useful to solve ancient and recent divergences (Faircloth et al., 2012; Smith, Harvey, Faircloth, Glenn, & Brumfield, 2014) seems pertinent. Indeed, a set of probes has been developed to target UCEs in Hymenoptera (Faircloth et al., 2015). This set and an enriched one (Branstetter, Longino, Ward, & Faircloth, 2017c) were successfully used to solve the phylogeny of a few groups of ants, wasps and bees for which the amount of DNA was not limiting (Blaimer et al., 2015; Blaimer, LaPolla, Branstetter, Lloyd, & Brady, 2016; Blaimer, Lloyd, Guilory, & Brady, 2016; Bossert, Murray, Blaimer, & Danforth, 2017; Branstetter, Danforth et al., 2017a; Branstetter, Josovnik et al., 2017b; Josovnik et al., 2017; Prebus, 2017; Ward & Branstetter, 2017). Thus, contributing to the global effort to solve the Hymenoptera tree of life while addressing the challenge of the phylogeny of chalcid wasps seemed sound.

Here, we provide a detailed description of the optimized protocol for DNA extraction and library preparation, followed by a description of the phylogenetic trees obtained through target enrichment of UCEs from 96 species belonging to seven families and one subfamily of chalcids used for biological control (Aphelinidae, Azotidae, Encyrtidae, Eulophidae, Mymaridae, Pteromalidae: Euvotinae, Signiphoridae, Trichogrammatidae) as well as three outgroups in Mymarommatidae, the putative sister group to Chalcidoidea (Gibson, Read, & Huber, 2007; Heraty et al., 2013).

2 | MATERIALS AND METHODS

2.1 | Sampling

Samples were taken from the personal collections of the co-authors of this study or borrowed from the Queensland Museum (Australia) or the Australian National Insect Collection, Canberra. Details of the 99 samples included in the analysis are presented in Supporting Information Table S1. Most specimens sampled in the field were placed directly into ethanol for storage. On average, specimens spent 3.5 years in ethanol before being processed (maximum storage time in alcohol = 34 years). Two specimens were critical point dried 25 or 34 years ago. UCE data for three specimens were retrieved from a previous study (Branstetter, Danforth et al., 2017a): Euplectrus sp. (empirical data); Copidosoma floridanum and Trichogramma pretiosum (in silico extraction of UCEs from genomes).
2.2 | DNA extraction

DNA was extracted using the Qiagen DNeasy Blood and Tissue Kit. All extractions were conducted without destruction of the specimens’ external (and certain internal) structures, with digestion and lysis of just the soft tissues. In this way, actual or potential type specimens are preserved. An often-essential feedback to the morphology is also preserved which is critical in this difficult group. Extractions were done from single specimens (sample codes in Supporting Information Table S1 ending with 01) or from a pool of several specimens (sample codes ending with 89).

The following modifications were made to manufacturer’s protocol. Samples were incubated overnight in an Eppendorf thermomixer (temperature = 56°C, mixing frequency = 300 rpm). To obtain DNA yield, two successive elutions (50 μl each) were performed with heated buffer AE (56°C) and an incubation step of 15 min followed by centrifugation (6000 g for 1 min at room temperature). Eppendorf microtubes LoBind 1.5 ml were used for elution and to store DNA at −20°C until library preparation. DNA was quantified with a Qubit® 2.0 Fluorometer (Invitrogen). The final version of the DNA extraction protocol is available as a Supporting Information Additional File S1. Vouchers were deposited at CBGP, Montferrier-sur-Lez, France, or returned to their owner.

2.3 | Whole genome amplification

DNA extracted from two specimens was subjected to ethanol precipitation and whole genome amplification (WGA) using the GenomiPhi™ V2 DNA Amplification Kit (GE Healthcare) as described in Cruaud et al. (2018). One microlitre of concentrated DNA was used as input (i.e., 4 ng or 0.4 ng, Supporting Information Table S1).

2.4 | Library preparation

Our starting point was the protocol described in http://ultraconserved.org and Faircloth et al. (2015). The final goal was to obtain a standardized protocol that could be implemented by one person and which, in 5 working days, would allow the manual preparation of libraries and the capture of UCEs from 96 samples in parallel. Step-by-step optimizations were made specially to remove time-consuming purification steps, and different reagents were tested. The final version of the protocol is available as Supporting Information Additional File S2. Briefly, input DNA was sheared to a size of ca 400 bp using the Bioruptor® Pico (Diagenode). End repair, 3′-end adenylation, adapters ligation and PCR enrichment were then performed with the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB). We used barcoded adapters that contained amplification and Illumina sequencing primer sites, as well as a nucleotide barcode of 5 or 6 bp long for sample identification (Supporting Information Additional File S3). Pools of 16 samples were made at equimolar ratio. We enriched each pool using the 2,749 probes designed by Faircloth et al. (2015) using MyBaits kits (MYcroarray, Inc.). We followed manufacturer’s protocol (MyBaits, user manual version 3, http://www.mycroarray.com/pdf/Mybaits-manual-v3.pdf). The hybridization reaction was run for 24 hr at 65°C. Postenrichment amplification was performed on beads with the KAPA HiFi HotStart ReadyMix. The enriched libraries were quantified with Qubit, an Agilent Bioanalyzer and qPCR with the Library Quantification Kit—Illumina/Universal from KAPA (KK4824). They were then pooled at equimolar ratio. Paired-end sequencing (2*300 bp) was performed on an Illumina MiSeq platform at UMR AGAP (Montpellier, France) to get longer flanking regions and, as a consequence, more information to differentiate closely related species.

2.5 | Raw data cleaning

The analytical workflow is summarized in Supporting Information Figure S1. In the next paragraph, chosen parameter values (different from default value) are provided between parentheses. Quality control checks were performed on raw sequence data with FASTQC version 0.11.2 (Andrews, 2010). Quality filtering and adapter trimming were performed with TRIMMOMATIC-0.36 (Bolger, Lohse, & Usadel, 2014) (LEADING:20 TRAILING:20 SLIDINGWINDOW:4:20 MINLEN:180, with PrefixPE/1 = AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT and PrefixPE/2 = CAAGCAGAAGACCGACATCGATCGGTCTCGGATCTTCCTGCTGAACCGCTCTTCCGATCT). Overlapping reads were merged using FLASH 1.2.11 (-M 300) (Magoc & Salzberg, 2011). Demultiplexing was performed using a bash custom script (no mismatch in barcode sequences was allowed, Supporting Information Additional File S4). Assembly of cleaned reads was performed using CAP3 (-i 25 -o 25 -s 400) (Huang & Madan, 1999). The 2,749 probes designed by Faircloth et al. (2015) were assembled into nonoverlapping UCEs (hereafter called reference UCEs, n = 1,432, Supporting Information Additional File S5 and S6) using GENEIOUS 8.1.8 (Kearse et al., 2012), and contigs were aligned to this set of reference UCEs using LASTZ Release 1.02.00 (Harris, 2007). Contigs that aligned with more than one reference UCE and different contigs that aligned with the same reference UCE were filtered out using GENEIOUS 8.1.8.

2.6 | Data analysis

UCEs for which sequences were available for more than 25% of the taxa were kept in the next steps of the analysis. Alignments were performed with MAFFT version 7.245 (Katoh & Standley, 2013) (linsi option). Ambiguously aligned blocks were removed using Gblock_0.91b with relaxed constraints (−t = d −b2 = b1 −b3 = 10 −b4 = 2 −b5 = h) (Talavera & Castresana, 2007). The final data set was analysed using supermatrix approaches and coalescent-based summary methods. Two gene tree reconciliation approaches were used: ASTRAL-III version 5.6.1 (Zhang, Rabiee, Sayyari, & Mirarab, 2018), which computes the phylogeny that agrees with the largest number of quartet trees induced by the set of input gene trees, and ASTRID (Vachaspati & Warnow, 2015), which takes a set of gene trees, computes a distance matrix (ca sum of number of edges in the path between two samples divided by the number of
gene trees in which the two samples are represented) and infers a phylogeny from this distance matrix. Following recommendations for incomplete distance matrices, BioNJ was used to compute the phylogeny. Individual trees were inferred from each UCE using RAxMLHPC -p threads -avx (Stamatakis, 2014) (version 8.2.4; -f a -x 12345 -p 12345 -# 100 -m GTRGAMMA). ASTRAL and ASTRID analyses were performed with 100 multilocus bootstrapping (MLBS, site-only resampling [Seo, 2008]). Phylogenetic trees were estimated from the concatenate, unpartitioned data set using maximum-likelihood (ML) approaches as implemented in RAxML and IQ-TREE version 1.6.4 (Nguyen, Schmidt, Haeseler, & Minh, 2015). For the RAxML analysis, a rapid bootstrap search (100 replicates) followed by a thorough ML search (-m GTRGAMMA) was performed. For the IQ-TREE analysis, a ML search with the best-fit substitution model automatically selected was performed with branch supports assessed with ultrafast bootstrap (Minh, Nguyen, & Haeseler, 2013) and SH-aLRT test (Guindon et al., 2010) (1,000 replicates).

Summary statistics for all data sets (alignment length, number of samples, number of variable sites, number of parsimony informative sites, etc.) were calculated using AMAS (Borowiec, 2016). Tree annotation was performed with TREEGRAPH 2.13 (Stöver & Müller, 2010).

Linear correlation between the number of UCEs and the quantity of DNA used to build the library was tested with the Pearson correlation coefficient in R (R Core Team, 2015). Analyses were performed on a Dell PowerEdge T630 with 10 Intel Xeon E5-2687 dual-core CPUs (3.1 GHz, 9.60 GT/s), 125 Go RAM and 13 To hard drive and on the Genotoul Cluster (INRA, Toulouse).

3 | RESULTS

Optimizations made for DNA extraction are detailed in the Supporting Information Additional File S1. The final version of the library preparation protocol is available as Supporting Information Additional Files S2 and S3. Hereafter, the range of values provided between parentheses refers to the range of data that fall between the 2.5th percentile (LB = lower bound) and 97.5th percentile (UP = upper bound). Supporting Information Table S1 contains sequencing information for all samples. The median amount of input DNA was 25 ng (LB = 1.8 ng; UP = 250 ng). An average of 76,330 reads (cleaned and merged) per sample was obtained (LB = 3,359; UP = 348,326). The average number of contigs was 3,454 (LB = 546; UP = 14,012), and the average sequencing depth was 18× (LB = 3×; UP = 44.0×). The average number of UCEs obtained per sample after filtering of problematic contigs was 687 (LB = 193-UP = 1,082) with a length comprised between LB = 315 and UP = 816 bp (mean = 603 bp).

Figure 1 and Supporting Information Figure S2 show the variation of the number of UCEs obtained with regard to the amount of input DNA. No significant correlation was observed (Pearson’s correlation coefficient = 0.096, p-value = 0.36).

The final data set (25%-complete matrix; i.e., at least 25 taxa on the 99 should have a sequence to keep the locus in the analysis) comprised 1,139 UCEs and 340,286 bp (missing data = 47.0%; parsimony informative sites = 72.3%; GC content = 42.6%).

Specimens retrieved from a previous study (Branstetter, Danforth et al., 2017a) that were either represented by empirical data (Euplectrus sp.) or UCEs extracted from published genomes (Copidosoma floridanum and Trichogramma pretiosum) displayed a number of UCEs comparable to what was obtained for other specimens. Their placement in the trees (Figure 2, Supporting Information Figures S3–S5) was in accordance with their morphology. Whatever the method used (supermatrix approaches vs. coalescent-based summary methods), all families, except for Aphelinidae, were recovered as monophyletic with high support. Aphelinids were split into three groups: (a) a monophyletic Aphelininae + Eretmocerinae; (b) Coccophaginae; and (c) Cales sp. (Calesinae). The position of Cales was ambiguous. Cales was either recovered as sister to Trichogrammatidae (RAxML, low support) or as a lineage distinct from all other chalcidoids (all other analyses). Except for Mymaridae that was strongly placed as sister to all other Chalcidoidea in all analyses, the tree backbone remained poorly resolved. Statistical support was much higher within families. In all analyses, Azotidae clustered with Signiphoridae, with strong support.

4 | DISCUSSION

To our knowledge, this study is the second after Sproul and Maddison (2017) to demonstrate success in library preparation from such low input using commercial kits, and the first to report successful sequencing of >1,000 low copy genes in 96 specimens in parallel, from such low input and processing time. Our optimizations differ from what was proposed by Sproul and Maddison (2017). First, we tried to optimize DNA extraction itself by using overnight
lysis with gentle mixing to preserve fragile specimens, heated elution buffer and increased elution time before elution. Second, instead of increasing the number of time-consuming purification steps, we decreased them. It is noteworthy that in this library, only two historical specimens were included. This may have masked challenges posed by adapter dimers (Burrell, Disotell, & Bergey, 2015; Sproul & Maddison, 2017; Tin, Economo, & Mikheyev, 2014) that led Sproul and Maddison (2017) to add a second bead clean-up prior to library amplification. However, we have already used this protocol on hundreds of chalcid and moth species, including historical specimens that were processed the same way as fresh ones, and we never had such an issue. Finally, instead of increasing the number of amplification cycles prior to target enrichment, we used a new generation mastermix including a hot start, processive and high-fidelity polymerase (NEBNext Ultra II Q5 Master Mix). Our protocol also allows one to back-up DNA at several steps that allows for multiple attempts without delay in case the first attempt fails. Finally, sequencing was performed on a MiSeq to get longer flanking regions and, as a consequence, more information to differentiate closely related species.

The protocol was successfully used on minute chalcid wasps widely used for biological control purposes. Up to 1,165 valid UCEs were captured from 25 ng DNA (median amount of DNA used for this study). No correlation was observed between the quantity of DNA used for library preparation and the number of captured UCEs. The average number of captured UCEs validated by our quality control workflow was 687, and 685 valid UCEs were captured from a tiny aphelinid (1.8 ng of input DNA). The number of UCEs obtained per individual appears to drop within the basal clades (i.e., Mymaridae and Trichogrammatidae, Supporting Information Figure S2), a result probably linked to the relatively long branches observed in these groups, and that could reduce the efficiency of the probes that were designed from the genome of Nasonia (Faircloth et al., 2015). Trees were well resolved at the family level, with high statistical support, showing the potential of UCEs to solve long-standing taxonomic issues. However, the tree backbone remained unresolved, a pattern that confirms the rapid diversification of the group (Heraty et al., 2013; Peters et al., 2018). Understanding the evolutionary history of the group was not the purpose of this paper. Indeed, only a representative sampling in both markers and taxa as well as cutting-edge data analysis will allow drawing accurate conclusions. By providing suitable tools for a fast, easy and affordable acquisition of data, this paper is a first step.

Interestingly, as it has been shown previously on RADseq data (Cruaud et al., 2018), WGA does not seem to bias the results even when input DNA is below the recommendation on the manufacturer kit (here 0.4 and 4 ng when the GenomiPhi Kit V2 requires 10 ng). It is noteworthy that, after completion of this paper and using the same protocol, we were able to capture ca 300 UCEs from native DNA with a concentration that falls below the detection limit of the Qubit, without WGA. Reducing the amount of DNA required for library preparation allows one to use extracted DNA for different approaches in parallel (RADseq, amplicon, Shotgun, etc.). This also allows one to send DNA back to museums from which specimens were borrowed, for archival purposes. When DNA quantities are too low, an extra WGA step can be performed. We definitely agree with Sproul and Maddison (2017) who emphasize how important it is not to waste DNA obtained from irreplaceable specimens (whether fresh or historical). It is even more important to capitalize on existing collections as collecting samples for large-scale studies may be more and more difficult, given that many countries have imposed restrictive access regulations, even to academic researchers, to reduce the risk of supposed biopiracy (Divakaran Prathapan, Pethiyagoda, Bawa, Raven, & Rajan, 2018).

All the elements discussed above indicate that this protocol may be of great help to reconstruct phylogenetic hypotheses in multiple groups of tiny arthropods, for example, springtails, sandflies, lice, whiteflies and mites. Coupled or not with WGA, some steps of our protocol may also contribute to simplify and hasten construction of other reduced-representation libraries (RADseq, ddRADSeq, etc.). These methods may then be used to analyse species relationships, hybridization or population genomics (Cruaud et al., 2014; Eaton & Ree, 2013; Emerson et al., 2010; Gagnaire, Normandeau, Pavey, & Bernatchez, 2013; Hipp et al., 2014) on minute arthropods or on endangered species using noninvasive DNA sampling (Vila, Auger-Rozenberg, Goussard, & Lopez-Vaamonde, 2009).

ACKNOWLEDGEMENTS

We are grateful to Nicole Fisher (CSIRO, National Museum of Australia, Canberra), Chris Burwell, Christine Lambkin and Susan Wright (Queensland Museum, Brisbane, Australia) for facilitating access to samples and to the Genotoul bioinformatics platform Toulouse Midi-Pyrenees, France, for providing computing resources. AC thanks Lisa Schlée, Philippe Cruaud and Maya for their patience in the field and Jean-Pierre Rossi for his help with R. AG thanks V. Fursov (Schmalhausen Institute of Zoology, Ukraine) for his help with specimen identification. We thank two anonymous reviewers for their careful review of our manuscript. This work is part of a large NSF project (Award #1555808) led by John Heraty (UC Riverside USA) that attempts to solve the phylogeny of Chalcidoidea with NGS approaches, and was partly funded by the ANR project
TriPTIC (ANR-14-CE18-0002) led by JYR and the SPE department of the INRA (French National Agronomic Institute). Collection permit in Australia - Queensland: WITK18248017-WITK18278817 to AC - JYR. LF travels to France were supported by the grants IntegPar (89BM/2017) and NEVPIT (PN-III-P4-ID-PCE-2016-0233) awarded by the Romanian National Authority for Scientific Research, CNCS—UEFISCDI. This paper is dedicated to the memory of our friend Dr John La Salle, whose kindness and sense of humour will be remembered forever.

AUTHOR CONTRIBUTIONS

A.C. and J.Y.R. designed the study; A.C., L.F., A.G., J.H., A.P. and J.Y.R. contributed to samples collection and identification; S.N. with the help of A.C., J.Y.R. and P.A. optimized the protocol; S.N. with the help of A.C., J.Y.R. and L.F. performed laboratory work; A.W. sequenced libraries; A.C. and J.Y.R. analysed data; and A.C. and J.Y.R. drafted the manuscript. All authors commented on the manuscript.

DATA ACCESSIBILITY

Demultiplexed cleaned reads merged with FLASH are available as a NCBI Sequence Read Archive (ID# PRJNA495844).

ORCID

Astrid Cruaud https://orcid.org/0000-0001-8932-4199

REFERENCES

Al Khatib, F., Fusu, L., Cruaud, A., Gibson, G., Borowiec, N., Rasplus, J.-Y., ... Delvare, G. (2014). An integrative approach for species discrimination in the Eupelmus urazonus complex (Hymenoptera, Eupelmidae), with the descriptions of eleven new species from the West Palearctic. Systematic Entomology, 39, 806–862. Andrews, S. (2010). FastQC—a quality control application for FastQ files. Available from http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Blaimer, B. B., Brady, S. G., Schultz, T. R., Lloyd, M. W., Fisher, B. L., & Ward, P. S. (2015). Phylogenomic methods outperform traditional multi-locus approaches in resolving deep evolutionary history: A case study of formicine ants. BMC Evolutionary Biology, 15, 271. https://doi.org/10.1186/s12862-015-0552-5 Blaimer, B. B., LaPolla, J. S., Branstetter, M. G., Lloyd, M. W., & Brady, S. G. (2016). Phylogenomics, biogeography and diversification of obligate myalybug-tending ants in the genus Acropyga. Molecular Phylogenetics and Evolution, 102, 20–29. Blaimer, B. B., Lloyd, M. W., Guillory, W. X., & Brady, S. G. (2016). Sequence capture and phylogenetic utility of genomic ultraconserved elements obtained from pinned insect specimens. PLoS ONE, 11, e0161531. Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics, 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170 Borowiec, M. L. (2016). AMAS: A fast tool for alignment manipulation and computing of summary statistics. PeerJ, 4, e1660. https://doi.org/10.7717/peerj.1660

Bossert, S., Murray, E. A., Blaimer, B. B., & Danforth, B. N. (2017). The impact of GC bias on phylogenetic accuracy using targeted enrichment phylogenomic data. Molecular Phylogenetics and Evolution, 111, 149–157. https://doi.org/10.1016/j.ympev.2017.03.022

Branstetter, M. G., Danforth, B. N., Pitts, J. P., Faircloth, B. C., Ward, P. S., Buffington, M. L., ... Brady, S. G. (2017a). Phylogenomic insights into the evolution of stinging wasps and the origins of ants and bees. Current Biology, 27, 1019–1025.

Branstetter, M. G., Ješovnik, A., Sosa-Calvo, J., Lloyd, M. W., Faircloth, B. C., Brady, S. G., & Schultz, T. R. (2017b). Dry habitats were crucibles of domestication in the evolution of agriculture in ants. Proceedings of the Royal Society B-Biological Sciences, 284, 20170095.

Branstetter, M. G., Longino, J. T., Ward, P. S., & Faircloth, B. C. (2017c). Enriching the ant tree of life: Enhanced UCE bait set for genome-scale phylogenetics of ants and other Hymenoptera. Methods in Ecology and Evolution, 8, 768–776.

Burks, R. A., Heraty, J. M., Gebiola, M., & Hansson, C. (2011). Combined molecular and morphological phylogeny of Eulophidae (Hymenoptera: Chalcidoidea), with focus on the subfamily Entedoninae. Cladistics, 27, 1–25. https://doi.org/10.1111/j.1096-0031.2011.00358.x

Burrell, A. S., Disotell, T. R., & Bergey, C. M. (2015). The use of museum specimens with high-throughput DNA sequencers. Journal of Human Evolution, 79, 35–44. https://doi.org/10.1016/j.jhevol.2014.10.015

Chen, Y., Hui, X. A., Fu, J. Z., & Huang, D. W. (2004). A molecular phylogeny of eurytomid wasps inferred from DNA sequence data of 28S, 18S, and COI genes. Molecular Phylogenetics and Evolution, 31, 300–307. https://doi.org/10.1016/S1055-7903(03)00282-3

Consoli, F. L., J. R. P. Parra, & R. A. Zucchi (Eds.). (2010). Egg parasitoids in agroecosystems with emphasis on Trichogramma. Dordrecht, The Netherlands: Springer.

Cruaud, A., Gautier, M., Galan, M., Foucaud, J., Sauné, L., Genson, G., ... Rasplus, J. Y. (2014). Empirical assessment of RAD sequencing for interspecific phylogeny. Molecular Biology and Evolution, 31, 1272–1274. https://doi.org/10.1093/molbev/msu063

Cruaud, A., Groussetier, G., Genson, G., Sauné, L., Polaszek, A., & Rasplus, J.-Y. (2018). Pushing the limits of whole genome amplification: Successful sequencing of RADseq library from a single microhymenopteran (Chalcidoidea, Trichogramma). PeerJ, 6, e5640. https://doi.org/10.7717/peerj.5640

Cruaud, A., Rønsted, N., Chantarasuwon, B., Chou, L. S., Clement, W. L., Couloux, A., ... Savolainen, V. (2012). An extreme case of plant-insect co-diversification: Figs and fig-pollinating wasps. Systematic Biology, 61, 1029–1047. https://doi.org/10.1093/sysbio/sys068

Desjardins, C. A., Regier, J. C., & Mitter, C. (2007). Phylogeny of pteromalid parasitic wasps (Hymenoptera : Pteromalidae): Initial evidence from four protein-coding nuclear genes. Molecular Phylogenetics and Evolution, 45, 454–469. https://doi.org/10.1016/j.ympev.2007.08.004

Divakaran Prathapan, K., Pethiyagoda, R., Bawa, K. S., Raven, P. H., & Rajan, P. D. (2018). When the cure kills—CBD limits biodiversity research. National laws fearing biopiracy squelch taxonomy studies. Science, 360, 1405–1406. https://doi.org/10.1126/science.aat9844

Eaton, D. A. R., & Ree, R. H. (2013). Inferring phylogeny and introgression using RADseq data: An example from flowering plants (Pedicularis: Orobanchaceae). Systematic Biology, 62, 689–706. https://doi.org/10.1093/sysbio/sy032

Faircloth, B. C., Branstetter, M. G., White, N. D., & Brady, S. G. (2015). Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among...
Formicidae). BMC Evolutionary Biology, 17, 250. https://doi.org/10.1186/s12862-017-1095-8
R Core Team. (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. http://www.r-project.org/.
Seo, T. K. (2008). Calculating bootstrap probabilities of phylogeny using multilocus sequence data. Molecular Biology and Evolution, 25, 960-971. https://doi.org/10.1093/molbev/msn043
Smith, B. T., Harvey, M. G., Faircloth, B. C., Glenn, T. C., & Brumfield, R. T. (2014). Target capture and massively parallel sequencing of ultra-conserved elements for comparative studies at shallow evolutionary time scales. Systematic Biology, 63, 83–95. https://doi.org/10.1093/sysbio/syt061
Sproul, J. S., & Maddison, D. R. (2017). Sequencing historical specimens: Successful preparation of small specimens with low amounts of degraded DNA. Molecular Ecology Resources, 17, 1183–1201. https://doi.org/10.1111/1755-0998.12660
Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, 30, 1312-1313. https://doi.org/10.1093/bioinformatics/btu033
Stouthamer, R., Luck, R. F., & Hamilton, W. D. (1990). Antibiotics cause parthenogenetic Trichogramma (Hymenoptera/Trichogrammatidae) to revert to sex. Proceedings of the National Academy of Sciences, USA, 87, 2424-2427. https://doi.org/10.1073/pnas.87.7.2424
Stöver, B. C., & Müller, K. F. (2010). TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. BMC Bioinformatics, 11, 7. https://doi.org/10.1186/1471-2105-11-7
Talavera, G., & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Systematic Biology, 56, 564–577. https://doi.org/10.1080/10635150701472164
Tin, M.-Y., Economo, E. P., & Mikheyev, A. S. (2014). Sequencing degraded DNA from non-destructively sampled museum specimens for RADtagging and low-coverage shotgun phylogenetics. PLoS ONE, 9, e96793.
Vachaspati, P., & Warnow, T. (2015). ASTRID: Accurate species TRees from internode distances. BMC Genomics, 16(Suppl. 10), 53.
Vila, M., Auger-Rozenberg, M. A., Goussard, F., & Lopez-Vaamonde, C. (2009). Effect of non-lethal sampling on life-history traits of the protected moth Groeßia isabellea (Lepidoptera: Saturniidae). Ecological Entomology, 34, 356–362.
Ward, P. S., & Branstetter, M. G. (2017). The acacia ants revisited: convergent evolution and biogeographic context in an iconic ant/plant mutualism. Proceedings of the Royal Society B-Biological Sciences, 284, 20162569.
Werren, J. H., & Loehlin, D. W. (2009). The parasitoid wasp Nasonia: An emerging model system with haploid male genetics. Cold Spring Harbor protocols, 2009, pdb.emo134. https://doi.org/0.1101/pdb.emo134
Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics, 19, 153. https://doi.org/10.1186/s12859-018-2129-y
SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Cruaud A, Nidelet S, Arnal P, et al. Optimized DNA extraction and library preparation for minute arthropods: Application to target enrichment in chalcid wasps used for biocontrol. Mol Ecol Resour. 2019;19:702-710. https://doi.org/10.1111/1755-0998.13006