Medical and Veterinary Entomology

Medical and Veterinary Entomology (2016) **30**, 200–208 doi: 10.1111/mve.12157

Patterns of cryptic host specificity in duck lice based on molecular data

G. C. ESCALANTE[1], A. D. SWEET[1], K. G. MCCRACKEN[2,3,4], D. R. GUSTAFSSON[4], R. E. WILSON[5] and K. P. JOHNSON[1]

[1]Illinois Natural History Survey, University of Illinois, Champaign, IL, U.S.A., [2]Department of Biology, University of Miami, Miami, FL, U.S.A., [3]Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, FL, U.S.A., [4]Department of Biology, University of Utah, Salt Lake City, UT, U.S.A. and [5]Institute of Arctic Biology, University of Alaska, Fairbanks, AK, U.S.A.

Abstract. Documenting patterns of host specificity in parasites relies on the adequate definition of parasite species. In many cases, parasites have simplified morphology, making species delimitation based on traditional morphological characters difficult. Molecular data can help in assessing whether widespread parasites harbour cryptic species and, alternatively, in guiding further taxonomic revision in cases in which there is morphological variation. The duck louse genus *Anaticola* (Phthiraptera: Philopteridae), based on current taxonomy, contains both host-specific and widespread species. Mitochondrial and nuclear DNA sequences of samples from this genus were used to document patterns of host specificity. The comparison of these patterns with morphological variations in *Anaticola* revealed a general correspondence between the groups identified by DNA sequences and morphology, respectively. These results suggest that a more thorough taxonomic review of this genus is needed. In general, the groups identified on the basis of molecular data were associated with particular groups of waterfowl (e.g. dabbling ducks, sea ducks, geese) or specific biogeographic regions (e.g. North America, South America, Australia, Eurasia).

Key words. Anatidae, parasites, phylogeny, waterfowl, wing lice.

Introduction

Parasites often have simplified morphologies because they have adapted to living on or in a host. Many parasites are also small-bodied in comparison with their hosts, which also leads to the simplification of morphological features (Lafferty & Kuris, 2002). These features make the occurrence of cryptic species a common phenomenon in parasites (Poulin & Morand, 2000; Miura *et al.*, 2005). Such species can often only be distinguished using molecular data (Wilkerson *et al.*, 1995; Jousson *et al.*, 2000; Perkins, 2000; Gustafsson & Olsson, 2012).

In parasites, it is important to be able to distinguish cryptic species because it is then possible to uncover distinctive patterns of pathogen transmission (Besansky *et al.*, 2003; Miura *et al.*, 2006; Panzera *et al.*, 2006) or host association (Le Gac *et al.*, 2007; Zhou *et al.*, 2012). Molecular data in the form of DNA sequences or other variable genetic markers are often used to detect these cryptic species (Wilkerson *et al.*, 1995; Blouin, 2002; Vilas *et al.*, 2005; Nadler & Perez-Ponce de Leon, 2011). In some cases, a molecular study can reveal unexpected patterns of variation that provide insight into population structure or differentiation. These patterns, in turn, can lead to re-evaluation of morphological differences, which reveals characters that had not been recognized previously as diagnostic (Hellgren *et al.*, 2007; Lee *et al.*, 2009).

Lice (Phthiraptera) are one group of parasites in which cryptic species have been documented. These wingless insects are ectoparasites occurring on most groups of birds and mammals. In birds, preening is the main defence against feather lice (Philopteridae), which eat the downy parts of body feathers.

Correspondence: Gabriela C. Escalante, Illinois Natural History Survey, University of Illinois, 1816 South Oak Street, Champaign, IL 61820, U.S.A. Tel.: + 1 217 244 9267; Fax: + 1 217 244 0802; E-mail: escalnt2@illinois.edu

These lice, in turn, escape from host preening defences in a variety of ways. One of these escape mechanisms is to hide between the feather barbs of wing feathers. Wing lice have a long and slender body form that is similar in width to the spaces between feather barbs, and the body size of these lice is correlated with host feather dimensions, which, in turn, is related to overall host body size (Johnson *et al.*, 2005).

Wing lice are typically host-specific and are found on only one or a few species of host, although some louse species can be found on a large number of host species (Price *et al.*, 2003). For example, one species of wing louse, *Columbicola macrourae* (Wilson, 1941) (Phthiraptera: Philopteridae), occurs on 15 species of pigeons and doves (Columbiformes) in the New World. A taxonomic revision of this genus based on morphology (Clayton & Price, 1999) concluded that the *Columbicola* Ewing, 1929 on all these hosts were the same morphospecies. However, comparisons of mitochondrial gene cytochrome oxidase subunit I (COI) sequences revealed sequence divergence of up to 21% between louse individuals on different host species (Johnson *et al.*, 2002; Malenke *et al.*, 2009). In addition, experimental transfers of genetically distinct lice between two of these hosts (*Zenaida asiatica* and *Zenaida macroura*) revealed that although these lice represented the same morphospecies, they had greatly reduced fitness (survival and reproduction) when transferred to the alternative host (Malenke *et al.*, 2009). This difference is most likely related to subtle differences in body size between these two host species and the correlated small differences in louse size.

Another genus of wing lice (*Anaticola* Clay, 1936) has a high potential to harbour cryptic species. This genus is widespread on ducks, geese and swans (i.e. waterfowl; Anseriformes: Anatidae) and comprises 33 species (Price *et al.*, 2003), but there is a lack of reliable diagnostic features based on morphology for most of these species, and the species limits adopted by Price *et al.* (2003) are poorly circumscribed. Price *et al.* (2003) considered some *Anaticola* species to be host-specific and found on only one species of host, but considered others to be extremely widespread. For example, *Anaticola crassicornis* (Scopoli, 1763) as circumscribed by Price *et al.* (2003) has been recorded from 24 species of waterfowl, and *Anaticola mergiserrati* (De Geer, 1778) and *Anaticola anseris* (L., 1758) occur on 16 and 10 species of host, respectively. However, no thorough morphological revision of the genus has been performed, and even distinguishing these three widespread species from one another is difficult using published descriptions based on only small dimensional differences (R. D. Price, personal communication, year). This taxonomic uncertainty is made even more difficult by the fact that, in the past, louse taxonomy was often heavily reliant on host taxonomy, and new host associations were considered to be new louse species despite the absence of detailed documentation of morphological differences (Eichler & Vasjukova, 1980). Stragglers found on new hosts were sometimes described as new species only because they showed novel host associations, with no regard to morphology (Tandan & Hajela, 1962). Therefore any evaluation of species limits in this genus must account for both the potential of cryptic species and taxonomic uncertainties.

Although most species of louse, in general, are very host-specific, the lice of waterfowl seem to be an exception to this rule. Two other duck louse genera, *Anatoecus* Cummings, 1916, and *Trinoton* Nitzsch, 1818, (both: Phthiraptera: Philopteridae) also have species that occur on very large numbers of host species (69 and 68, respectively). Aspects of waterfowl ecology may be influencing these patterns. For example, many species of duck co-occur on the same bodies of water, providing potential opportunities for host transfer. In addition, many species in the northern hemisphere migrate southwards, where they form extremely large mixed-species wintering flocks (Bellrose, 1976). This migration pattern brings parasites from more northern regions into southern regions. In addition, these dense wintering aggregations may provide even more opportunities for louse transmission between species through direct host contact. However, some species of waterfowl are extremely territorial or live in isolation from other species, such that transmission opportunities for lice between host species may be limited. The interplay of all these factors may lead to the patterns of host specificity observed.

The goal of this study was to document the patterns of host specificity in the wing lice (*Anaticola*) of waterfowl. Given the difficulties of using morphology for diagnosing species in the genus *Anaticola*, and the lack of good published descriptions for most of the species within this genus, the assessment of patterns of host specificity in this genus was begun using genetic data. For the purposes of exploring these patterns, species names following the host associations described in Price *et al.* (2003) were used. The species delimitations of Price *et al.* (2003) have the potential to be overly conservative because they attempt to account for previous descriptions of new species that do not refer to clear morphological differences. As part of this study, a preliminary examination of the morphology of the samples in the study was conducted, but a complete taxonomic revision of the genus *Anaticola* was beyond the scope of this study. These morphological features were compared with results of DNA sequencing of both mitochondrial and nuclear genes to elucidate patterns of host specificity in *Anaticola* across a wide variety of host species. Patterns of host association were evaluated with respect to host body size, ecology and biogeography.

Materials and methods

Data collection

Lice were collected from waterfowl that were either killed using ethyl acetate fumigation or body washed while isolated in separate containers or from live birds kept isolated in paper bags dusted with pyrethrum powder (Clayton & Drown, 2001). The study targeted 11 species of *Anaticola* [as delimited by Price *et al.* (2003)] for a total of 82 samples (Table S1, online). Louse DNA was extracted using a Qiagen Blood and Tissue Kit (Qiagen, Inc., Valencia, CA, U.S.A.) using standard protocols, and the exoskeleton was mounted as a voucher on a microslide for the examination of morphology. The protocols recommended for New England Biolab's (Ipswich, MA, U.S.A.) 5× Master Mix (5 μL of Master Mix, 0.5 μL of each primer for a 25-μL reaction) and previously documented polymerase chain reaction (PCR)

protocols (Johnson *et al.*, 2002; Sweet *et al.*, 2014) were used to amplify three genes (one mitochondrial and two nuclear). The mitochondrial gene COI was amplified in two fragments, using the primer pairs L6625 and H7005 for the shorter (COIS) fragment (Hafner *et al.*, 1994) and LCOI-490 and HCO2-198 for the longer (COIL) fragment (Folmer *et al.*, 1994). The nuclear elongation factor 1 (EF-1α) gene was amplified using the primers For3 and CHO10 for EF-1α (Danforth & Ji, 1998). For a protein of unknown function [*hyp*; see Sweet *et al.* (2014)], the primers BR50-181L and BR50-621R were used (Sweet *et al.*, 2014). The success of the PCR was confirmed using gel electrophoresis with a 1% agarose gel and GelGreen (Biotium, Inc., Hayward, CA, U.S.A.). The resulting products were purified using the standard procedures of both ExoSAP-IT (Affymetrix, Inc., Santa Clara, CA, U.S.A.) and a Qiagen PCR Purification Kit (Qiagen, Inc.), and purified products were sequenced using ABI Prism BigDye Terminators (Applied Biosystems, Inc., Foster City, CA, U.S.A.) and Sanger sequencing on an AB 3730xl DNA Analyzer at the Roy J. Carver Biotechnology Center (University of Illinois, Champaign, IL, U.S.A.). From resulting files, base calls within complementary chromatograms were reconciled by eye using the programs SEQUENCHER Version 5.0.1 (Gene Codes Corp., Ann Arbor, MI, U.S.A.) or GENEIOUS Version 6 (Biomatters Ltd, Auckland, New Zealand). All novel sequences have been submitted to GenBank (pending acceptance).

Voucher slides were examined for potential morphological differences. Features examined included the pre-antennal structure, head shape, shape and arrangement of setae, genitalia, pigmentation patterns, and antennal structure. Some distinctive variation was observed in the pre-antennal head region (frons) and therefore the present study has focused on illustrating these differences.

Phylogenetic analysis

Alignment. The resulting sequence data were used to create alignments of each locus using default parameters in MUSCLE (Edgar, 2004). Each locus alignment was then checked by eye. As all four loci are coding regions, each indel was checked for base call errors against the original chromatograms. When the presence or absence of indels in the initial alignments had been confirmed, neighbour-joining trees (NJ) were created in GENEIOUS to check for conspicuously long branches. Long branches may indicate a misalignment, miscalled bases, or misidentification of sequences. From the NJ trees, several candidates for potential misidentification, sequencing error or contamination were identified. Based on subsequent National Center for Biotechnology Information (NCBI) BLAST searches and confirmation of poor sequence quality, the appropriate sequences were subsequently removed from all future analyses. Additionally, as the study protocol included the intended use of the COIS as a potential aid in the identification of cryptic species, any taxon that did not have this locus was removed from the final data matrix. Finally, an uncorrected pairwise distance matrix for COIS was computed using the R package *ape* (R Foundation for Statistical Computing, Vienna, Austria) (Paradis *et al.*, 2004).

For each locus alignment, trees were inferred with RA×ML Version 7.0.4 (Stamatakis, 2006) using the GTRGAMMA (GTR + Γ) model and 500 bootstrap replicates. The *Anaticola* sample from *Phoenicopterus ruber* (American flamingo) was used as the outgroup for each locus. This rooting decision was based on the previous phylogenetic placement of the flamingo *Anaticola* louse as sister to the rest of the genus (Johnson *et al.*, 2006). The resulting RA×ML trees were compared to identify well-supported discrepancies between individual gene trees. A well-supported difference in topology (based on bootstrap values) between two trees indicates conflicting evolutionary histories for those two loci, which would perhaps give misleading results from a concatenated alignment. If there are no well-supported discrepancies between individual gene trees, this indicates the evolutionary histories from each locus are consistent, and therefore concatenation is appropriate. In the present study, individual gene trees did not produce any well-supported conflicts and therefore GENEIOUS was used to concatenate the data into a single super-matrix for all future phylogenetic analyses.

Concatenated phylogenetic analysis. Both maximum likelihood (ML) and Bayesian analyses were run on the concatenated matrix. For both methods, the matrix was partitioned according to locus and the flamingo *Anaticola* louse was used as an outgroup taxon. PartitionFinder Version 1.1.1 (Lanfear *et al.*, 2014) was used to identify the best-fitting models for each locus according to the Akaike information criterion corrected for small sample size (AICc) (Sugiura, 1978). Based on the PartitionFinder results, COIS and COIL were treated as a single locus, and *hyp* and EF-1α as separate loci.

For the ML analysis, GARLI Version 2.0 (Zwickl, 2006) was used. A GTR + I + Γ model was applied for both the COIS/COIL and *hyp* partitions, and a TVMef + Γ model was applied for the EF-1α partition. GARLI was run for 500 bootstrap replicates. MrBayes Version 3.2 (Ronquist & Huelsenbeck, 2003) was used for the Bayesian analysis. Once again, GTR + I + Γ models were applied to the COIS/COIL and *hyp* partitions, but an SYM + Γ model was applied to the EF-1α partition. MrBayes does not implement a TVMef + Γ model (the best AICc score), and SYM + Γ had the next best AICc score. Two independent runs of MrBayes with four chains for 20 million generations using Markov chain Monte Carlo methods (MCMC), sampling every 1000 generations, were run, and the resulting trace files were viewed in Tracer Version 1.4 (Rambaut & Drummond, 2007) to ensure mixture and stationarity of the MCMC chains. Based on the trace files, the first 2000 trees (10%) were discarded as representing a burn-in.

Results

Within *Anaticola*, sequences of the mitochondrial COIS gene evolved relatively rapidly, differing by 0.0–20.0% between individuals (Table S2, online). By contrast, sequences of the nuclear genes EF-1α and *hyp* differed by up to only 3.3% and 6.0%, respectively. Analyses of each gene independently revealed no significant conflict between tree topologies derived

from different genes. That is, no nodes were in conflict above 75% bootstrap between the various gene trees. Therefore, gene sequences were combined into a concatenated alignment for subsequent analyses.

The combined analyses had strongly supported terminal groups of closely related individuals, whereas the deeper phylogenetic relationships were generally less strongly supported (Fig. 1). In many cases these terminal groups corresponded to defined species of louse, although several species were separated into multiple genetically distinct clusters. For example, individuals identified as *A. crassicornis* based on host association fell into six different groups (Fig. 1). For mitochondrial COIS, divergence within these groups amounted to < 2.6% and averaged 0.84%. By contrast, COIS divergence ranged from 4.6% to 13.8% between groups within *A. crassicornis.* This level of divergence was similar to those observed between different species of *Anaticola* (2.1–20.0%) described by Price *et al.* (2003). In general, there was also substantial variation in the shape of the pre-antennal region of the head (Fig. 1), both between species and within species [based on the current classification of Price *et al.* (2003)] containing highly divergent genetic lineages.

These different groups correspond to distinct patterns of either host association or geography (i.e. different continents). In *A. crassicornis*, the six groups identified in the tree (Fig. 1) were each restricted to particular host species or particular geographic regions, which suggests they may be different species. *Anaticola crassicornis* group 1 was widespread on eight species of dabbling duck (*Anas*, *Lophonetta* and *Speculanas*) throughout South America and the Falkland Islands. Within this group, COIS sequences diverged from one another by an average of only 0.6%. *Anaticola crassicornis* groups 2, 3 and 4 formed a clade that was sister to group 1. Group 2 was found on dabbling ducks (*Anas*) in Australia. Groups 3 and 4 were found only on mallards (*Anas platyrhynchos*) and a mallard relative (*Anas undulata*). Group 3 was restricted to the mallard group in the Old World, whereas group 4 was found only on mallards in the New World. Interestingly, the louse from a mallard relative in Australia (*Anas superciliosa*) fell into group 2, with those from other Australian dabbling ducks.

Anaticola crassicornis group 5 was quite widespread and found on 10 species of dabbling duck (*Anas*) throughout North America, Eurasia and Africa. Within this group COIS sequences averaged only 0.5% difference from one another. However, one individual, which, based on host association would be classified as *A. mergiserrati*, from *Aythya australis* in Australia was 2.2% different on average from the remainder of *A. crassicornis* group 5, suggesting some genetic distinctiveness, but this divergence was much smaller than the average divergence from *A. crassicornis* group 6 (7.9%), the next nearest relative of group 5. In addition, the morphology of this specimen from *Ay. australis* was highly similar to that of specimens of *A. crassicornis* group 5 (Fig. 1).

Anaticola specimens from host species (*Biziura lobata*, *Oxyura australis* and *Merganetta armata*) for which there was no prior host association generally grouped near *A. crassicornis* but were highly genetically distinct (11.8–14.9% COIS sequence divergence), which suggests they are separate species. These samples also had distinct pre-antennal morphologies (Fig. 1).

Interestingly, lice that would be considered *A. mergiserrati* based on host relationships [based on Price *et al.* (2003)] were placed into five genetically distinct lineages. The louse from *Ay. australis* (*A. mergiserrati* 1), as mentioned previously, was genetically very close to one group of *A. crassicornis* (group 5). The other lice (*A. mergiserrati* 2) from pochards (*Aythya*) and mergansers (*Mergus*) were placed very close to *A. crassicornis* 6 from stiff-tailed ducks (*Oxyura*). The louse (*A. mergiserrati* 3) from pink-eared duck (*Malacorhynchus membranaceus*) was genetically distinct from all others. The louse (*A. mergiserrati* group 4) from the common eider (*Somateria mollissima*) was very close to other species of *Anaticola* (*Anaticola tamarae*, *Anaticola constrictus* and *Anaticola branderi*) (2.6% COIS sequence divergence), and although the louse (*A. mergiserrati* 5) from the white-winged scoter (*Melanitta fusca*) clustered with these other species from sea ducks, it was quite genetically divergent (average 10.5% COIS).

On a genetically distinct long branch just outside the clade containing all the *A. crassicornis* lineages (marked as Main *crassicornis* Clade) and the *A. mergiserrati* 3 was the louse from the ringed teal (*Callonetta leucophrys*), which had no previously recorded host association with *Anaticola*. Just outside this sample were lice from shelducks (*Tadorna* and *Alopochen*), all of which came out together in one clade, reflecting host phylogeny (Gonzalez *et al.*, 2009). Each of the three host species sampled in this group had distinct louse species. Further down the tree, a louse from the grey teal (*Anas gracilis*) was very close to lice from *Chenonetta jubata* (0.0% COIS), but another specimen from the grey teal clustered with lice from other dabbling ducks in Australia (*A. crassicornis* group 2). This suggests the grey teal harbours two different genetically distinct lineages, and these lice shared with *Ch. jubata* may be distinct species.

Another lineage with potentially cryptic species was the lice of geese (*Branta* and *Anser*), which are classified as *A. anseris*. One group (*A. anseris* group 1) was found on two geese that are sister taxa (*Branta hutchinsii* and *Branta leucopsis*, Paxinos *et al.*, 2002), whereas the others (*A. anseris* 2 and 3) were found on two other goose species (*Branta canadensis* and *Anser caerulescens*). *Anaticola rheinwaldi*, which occurs on the brant goose (*Branta bernicla*) was most closely related to *A. anseris* 3, making *A. anseris* paraphyletic. The louse *Anaticola australis* from a goose relative (*Cereopsis novaehollandiae*) was sister to the *A. anseris* clade.

The species of *Anaticola* from sheldgeese (*Chloephaga*) were sisters to groups 4 and 5 of *A. mergiserrati*. Each of the two hosts sampled had a genetically distinct louse; in one case this represented a new host association.

Discussion

Sequences of mitochondrial and nuclear genes revealed evidence of potentially cryptic species in the wing lice (*Anaticola*) of waterfowl. In particular, the most widespread species [based on Price *et al.* (2003)], *A. crassicornis* (24 host species), *A. mergiserrati* (16 host species) and *A. anseris* (10 host

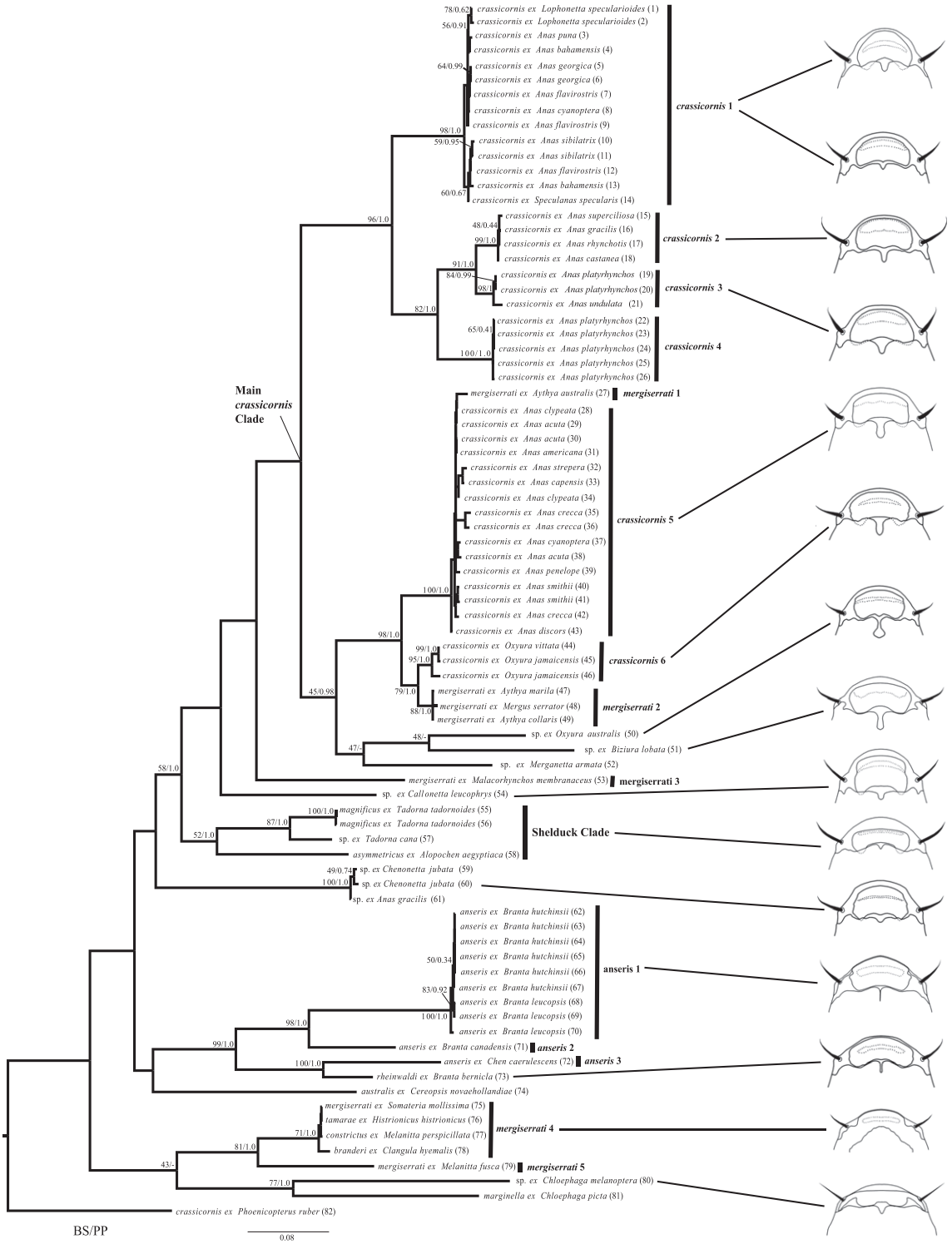

Fig. 1. Most likely maximum likelihood (ML) phylogeny of *Anaticola* species generated from 500 bootstrap replicates in GARLI Version 2.0 (Zwickl, 2006). Over each node, the first value is the bootstrap value (BS) from the ML analysis, and the second value is the posterior probability (PP) from a Bayesian analysis. Values below 40% BS and/or PP are not indicated. The scale bar indicates nucleotide substitutions per site along the branch lengths. Taxon names indicate *Anaticola* species followed by the host name. Numbers in parenthesis to the right of each name refer to reference numbers in Table S1 (online). Major clades are indicated. Illustrations are diagrams of pre-antennal head morphology (frons) of female representatives of various groups in the present study. If only a male specimen was available, the head morphology is not shown.

species), showed the most evidence of harbouring multiple distinct lineages based on genetic divergence within them. The level of genetic divergence, particularly in mitochondrial sequences (e.g. up to 13.8% within *A. crassicornis*), was similar to that observed between different species of *Anaticola*. These potentially cryptic species have their own distinct patterns of host association and biogeographic distribution that correlate with these genetic differences. In addition, morphological variation among the samples that corresponded to these genetic differences was identified. Further taxonomic study of additional slide-mounted material in museum collections may indicate that these species are not in fact cryptic, but diagnosable based on morphological features that correspond with the genetic differences.

For example, individuals of *A. crassicornis* fell into six distinct groups, typically with restricted host or biogeographic distributions. In addition, morphological features of the pre-antennal region varied among these groups (Fig. 1). *Anaticola crassicornis* is generally found on dabbling ducks (*Anas* and related genera), but also occurs on stiff-tailed ducks (*Oxyura*). These groups, rather than being distinguished by the host group on which they occur, are more limited to specific biogeographic regions. In particular, the widespread *A. crassicornis* group 1 occurs on dabbling ducks in South America, whereas *A. crassicornis* group 5 is widespread on dabbling ducks in North America, Eurasia and Africa. These two groups are not one another's closest relatives, although both contain individuals that parasitize related host species (Johnson & Sorenson, 1999). Some examples include pintail hosts in both the South American restricted *A. crassicornis* group 1 (*Anas georgica* and *Anas bahamensis*) and in the more widespread *A. crassicornis* group 5 (*Anas acuta*). Similarly wigeons (*Anas sibilatrix*) in South America host *A. crassicornis* group 1, whereas wigeons in North America (*Anas americana*) and Eurasia (*Anas penelope*) host *A. crassicornis* group 5.

Interestingly, one group of dabbling ducks, the mallard group (Johnson & Sorenson, 1999), hosts a combination of possibly cryptic louse species correlated with both host species and geography. *Anaticola crassicornis* groups 3 and 4 are restricted to mallards (*Anas platyrhynchos*) and mallard relatives. Based on the present sampling, group 3 is found only on mallards in Eurasia and on an African relative (*Anas undulata*), whereas group 4 is found only on mallards from North America. Although the geographic distribution of the mallard group overlaps with that of many other species of dabbling duck, the group appears not to share lice with these other sympatric dabbling duck species. One reason for this may be that species in the mallard group (which weigh ~ 1000 g) are much larger in body size than other dabbling ducks (~ 300–900 g) (Dunning, 1993). Generally louse size is correlated with host size, particularly for wing lice (Johnson *et al.*, 2005). As species of *Anaticola* are wing lice, it may be that a louse species of a given size cannot survive on both larger mallards and smaller dabbling ducks. This limitation may explain why *A. crassicornis* from the mallard group are genetically distinct from *A. crassicornis* on other dabbling ducks.

Given this explanation, it is surprising that the louse on the mallard representative (*Anas superciliosa*) from Australia was genetically extremely similar to *A. crassicornis* samples that occurred on small-bodied dabbling ducks in Australia (*Anas gracilis*, *Anas rhynchotis* and *Anas castanea*). In view of the position of *A. crassicornis* group 2 within the tree, it seems likely that this represents a host switch from the mallard group to other smaller-bodied dabbling ducks in Australia. Because Australia is very isolated, smaller dabbling duck lineages may have lost their lice when they colonized Australia [i.e. 'missing the boat' (Page & Charleston, 1998)]. Hence, these smaller-bodied dabbling ducks may have represented an open niche that allowed wing lice from other ducks to switch hosts. As the mallard group has very recently radiated around the world (Johnson & Sorenson, 1999), this may have been a very recent host-switching event; this suggestion is supported by the very low genetic divergence (~ 5.3%) between the *A. crassicornis* group 3 lice of Eurasian/African mallards and *A. crassicornis* group 2 from Australia.

One of the other species of lice that showed evidence for highly divergent lineages within it was *A. anseris*. Like *A. crassicornis*, *A. anseris* [as delimited by Price *et al.* (2003)] is restricted to hosts that are similar in size (large geese), which also happen to be phylogenetically closely related, making it difficult to disentangle these two factors. Five different goose species were sampled and showed evidence for four divergent lineages of lice on them: three divergent lineages within *A. anseris*, and also *A. rheinwaldi*, which was phylogenetically embedded within *anseris* (making *A. anseris*, as currently defined, paraphyletic). Although many species overlap during winter, it appears they do not share lice. In the one instance in which it appears that different species of goose share the same louse (*A. anseris* group 1, shared by *B. hutchinsii* and *B. leucopsis*) (Fig. 1), these two host species do not overlap geographically as *B. hutchinsii* is restricted to North America and *B. leucopsis* is restricted to Europe and Greenland. Although these two host species are sister species (Paxinos *et al.*, 2002), there is no evidence for genetic differentiation between their lice. Interestingly, *B. hutchinsii* and *B. canadensis* overlap in North America during the winter, forming mixed flocks [indeed, these two species were once considered to represent the same species (Paxinos *et al.*, 2002)], yet they do not share genetically similar lice. This may be because these two hosts differ dramatically in body size by approximately 2000 g (Fox *et al.*, 1996). By contrast, although they do not overlap geographically, *B. hutchinsii* and *B. leucopsis* are very similar in body size (Paxinos *et al.*, 2002). However, there are occasional records of *B. leucopsis* from North America in winter (Sherovny, 2008), as well as records of *B. hutchinsii* in Europe, where they overlap with *B. leucopsis*, and in some cases even hybridize (Hobbs, 2015), which may provide opportunities for transfer of lice between them. Therefore, geography plays less of a role in structuring the patterns of cryptic species and diversification within *A. anseris* than it does in *A. crassicornis*.

Lice that were placed as *A. mergiserrati* based on usual host associations (Price *et al.*, 2003) occurred in five different places in the tree. However, in most of these cases these lice were embedded within other species, often nearly genetically identical to them (*A. mergiserrati* 1 and group 2 within *A. crassicornis*; *A. mergiserrati* groups 4 and 5 close to three other species of *Anaticola* from sea ducks). Instances of *A. mergiserrati* that fell within *A. crassicornis* were collected from pochards (*Aythya*) and mergansers (*Mergus*), which

overlap in both breeding and winter distributions with dabbling ducks (the main hosts of *A. crassicornis*). Although these *A. mergiserrati* lice were embedded within *A. crassicornis*, they were somewhat genetically distinct (~2.2% COI for *A. mergiserrati* 1 from *A. crassicornis* group 5 and ~4.4% COI for *A. mergiserrati* 2 from *A. crassicornis* group 6), suggesting an older host-switching event rather than ongoing gene flow. Similarly, *A. mergiserrati* groups 4 and 5, which occur on sea ducks, are closely related to other species of lice from sea ducks [based on published host associations (Price *et al.*, 2003)], with which they share wintering locations. Finally, *A. mergiserrati* 3 was genetically distinct from anything else and was found only on the pink-eared duck (*Malacorhynchos membranaceus*) from Australia, which has no close relatives (Gonzalez *et al.*, 2009). In general, therefore, it appears that *A. mergiserrati* needs extensive re-evaluation from a taxonomic perspective.

One clade in which it appears that host phylogeny plays a major role in structuring the phylogeny of *Anaticola* refers to the lice of shelducks (Fig. 1; shelduck clade), from which three host species (*Tadorna tadornoides*, *Tadorna cana* and *Alopochen aegyptiaca*) were sampled. Among the waterfowl sampled, these species are one another's closest relatives (Gonzalez *et al.*, 2009). The lice from these hosts together formed a well-supported clade, yet lice from each host species were all genetically distinct from one another. Shelducks in general are extremely territorial in comparison with many other ducks (Young, 1970) and thus may have fewer opportunities to share lice with other host species.

The division of *A. crassicornis*, *A. anseris* and *A. mergiserrati* into several genetic clades suggests that the genus is presently over-lumped. The morphological variation identified in this study also supports this explanation. No attempt was made during this study to compare the present material with type material from the many proposed species names within *Anaticola* (Price *et al.*, 2003); however, a more thorough morphological review of the genus is warranted.

One of the themes to emerge from the present results concerns the importance of biogeography in structuring host–parasite associations. Previous studies have used molecular data to discover the role of biogeography in organizing host–parasite relationships, noting how geography influences various events, such as host switching and cospeciation (Hoberg & Brooks, 2008). For example, genetic differences observed among populations of the tapeworm *Ligula intestinalis* (L.) (Pseudophyllidea: Diphyllobothriidae) corresponded to distinct geographical areas of its three hosts, rather than being structured among host species (Stefka *et al.*, 2009). Similar patterns have been found in toucan lice [*Austrophilopterus* Ewing, 1929 (Phthiraptera: Philopteridae)], in which geographic barriers, such as the Amazon River, appear to impact the speciation of the lice (Weckstein, 2004).

In summary, based on mitochondrial and nuclear DNA sequences, the waterfowl wing louse genus *Anaticola* is more diverse than predicted based on published host associations. This genus shows evidence of several divergent lineages within species (as currently delimited), and there is distinctive morphological variation correlating with these genetic differences. In some cases, these genetic lineages were restricted to particular host taxonomic or ecological groups (e.g. dabbling ducks, sea ducks, geese). However, both geographic overlap and host body size appear to play large roles in the patterns of host specificity in this genus.

Supporting Information

Additional Supporting Information may be found in the online version of this article under the DOI reference: DOI: 10.1111/mve.12157

Table S1. Matrix of all species used in evaluating possible cryptic species in the louse genus *Anaticola*, along with relevant host association information. The extract code refers to the louse species and is followed by host information and extract date. The host voucher refers to specific museum specimens or collection voucher numbers. Louse specimens are listed according to Fig. 1. The last specimen (no. 82) is the outgroup taxon. Loci targeted for amplification and sequencing are listed to the right. Successful sequences are listed with their respective GenBank accession numbers. Missing data are represented by a dash.

Table S2. Uncorrected pairwise distance matrix of locus of mitochondrial gene cytochrome oxidase subunit I (COI) shorter fragment (COIS). Values were generated from the alignment of COIS in the *ape* package in R (R Foundation for Statistical Computing, Vienna, Austria). Number labels for each row and column refer to the specific taxa reference numbers in Table S1.

Acknowledgements

The authors thank T. Chesser, T. Galloway, L. Gijsen, I. Mason, P. J. McCracken, M. Meyer, E. Osnas, R. Palmer, T. Valqui, J. Weckstein, C. Witt, and J. Wombey for assistance in collecting louse samples. This work was supported by University of Illinois Research Board and National Science Foundation grants (DEB-0107891, DEB-1050706 and DEB-1239788) to KPJ.

References

Bellrose, F.C. (1976) *Ducks, Geese, and Swans of North America*. Stackpole, Harrisburg, PA.

Besansky, N.J., Severson, D.W. & Ferdig, S.M. (2003) DNA barcoding of parasites and invertebrate disease vectors: what you don't know can hurt you. *Trends in Parasitology*, **19**, 545–546.

Blouin, M.S. (2002) Molecular prospecting for cryptic species of nematodes: mitochondrial DNA versus internal transcribed spacer. *International Journal for Parasitology*, **32**, 527–531.

Clayton, D.H. & Drown, D.M. (2001) Critical evaluation of five methods for quantifying chewing lice (Insecta: Phthiraptera). *Journal of Parasitology*, **87**, 1291–1300.

Clayton, D.H. & Price, R.D. (1999) Taxonomy of New World *Columbicola* (Phthiraptera: Philopteridae) from the Columbiformes (Aves), with descriptions of five new species. *Annals of the Entomological Society of America*, **92**, 675–685.

Danforth, B.N. & Ji, S. (1998) Elongation factor-1 alpha occurs as two copies in bees: implications for phylogenetic analysis of EF-1 alpha sequences in insects. *Molecular Biology and Evolution*, **15**, 225–235.

Dunning, J.B. (1993) *CRC Handbook of Avian Body Masses*. CRC Press, Boca Raton, FL.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**, 1792–1797.

Eichler, W. & Vasjukova, T.T. (1980) Die Mallophagengattung *Anaticola* (Phthiraptera, Mallophaga). *Deutsche Entomologische Zeitschrift*, **27**, 335–375.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**, 294–299.

Fox, A.D., Glahder, G., Mitchell, C.R., Stroud, D.A., Boyd, H. & Frikke, J. (1996) North American Canada geese (*Branta canadensis*) in West Greenland. *The Auk*, **1**, 231–233.

Gonzalez, J., Duttmann, H. & Wink, M. (2009) Phylogenetic relationships based on two mitochondrial genes and hybridization patterns in Anatidae. *Journal of Zoology*, **279**, 310–318.

Gustafsson, D.R. & Olsson, U. (2012) Flyway homogenization or differentiation? Insights from the phylogeny of the sandpiper (Charadriiformes: Scolopacidae: Calindrinae) wing louse genus *Lunaceps* (Phthiraptera: Ischnocera). *International Journal for Parasitology*, **42**, 93–102.

Hafner, M.S., Sudman, P.D., Villablanca, F.X., Spradling, T.A., Demastes, J.W. & Nadler, S.A. (1994) Disparate rates of molecular evolution in cospeciating hosts and parasites. *Science*, **265**, 1087–1090.

Hellgren, O., Krizanauskiene, A., Valkiunas, G. & Bensch, S. (2007) Diversity and phylogeny of mitochondrial cytochrome *b* lineages from six morphospecies of avian *Haemoproteus* (Haemosporida: Haemoproteidae). *Journal of Parasitology*, **93**, 889–896.

Hobbs, J. (2015) A list of Irish birds. http://www.southdublinbirds.com/nimages/fyles/List-of-Irish-Birds-print.pdf [accessed on 21 December 2015].

Hoberg, E.P. & Brooks, D.R. (2008) A macroevolutionary mosaic: episodic host-switching, geographical colonization and diversification in complex host–parasite systems. *Journal of Biogeography*, **35**, 1533–1550.

Johnson, K.P. & Sorenson, M.D. (1999) Phylogeny and biogeography of dabbling ducks (genus: *Anas*): a comparison of molecular and morphological evidence. *The Auk*, **116**, 792–805.

Johnson, K.P., Williams, B.L., Drown, D.M., Adams, R.J. & Clayton, D.H. (2002) The population genetics of host specificity: genetic differentiation in dove lice (Insecta: Phthiraptera). *Molecular Ecology*, **11**, 25–38.

Johnson, K.P., Bush, S.E. & Clayton, D.H. (2005) Correlated evolution of host and parasite body size: tests of Harrison's rule using birds and lice. *Evolution*, **59**, 1744–1753.

Johnson, K.P., Kennedy, M. & McCracken, K.G. (2006) Reinterpreting the origins of flamingo lice: cospeciation or host-switching? *Biology Letters*, **2**, 275–278.

Jousson, O., Bartoli, P. & Pawlowski, J. (2000) Cryptic speciation among intestinal parasites (Trematoda: Digenea) infecting sympatric host fishes (Sparidae). *Journal of Evolutionary Biology*, **13**, 778–785.

Lafferty, K.D. & Kuris, A.M. (2002) Trophic strategies, animal diversity and body size. *Trends in Ecology & Evolution*, **17**, 507–513.

Lanfear, R., Calcott, B., Kainer, D., Mayer, C. & Stamatakis, A. (2014) Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evolutionary Biology*, **14**, 82.

Lee, K.S., Cox-Singh, J. & Singh, B. (2009) Morphological features and differential counts of *Plasmodium knowlesi* parasites in naturally acquired human infections. *Malaria Journal*, **8**, 73.

Le Gac, M., Hood, M.E., Fournier, E. & Giraud, T. (2007) Phylogenetic evidence of host-specific cryptic species in the anther smut fungus. *Evolution*, **61**, 15–26.

Malenke, J.R., Johnson, K.P. & Clayton, D.H. (2009) Host specialization differentiates cryptic species of feather-feeding lice. *Evolution*, **63**, 1427–1438.

Miura, O., Kuris, A.M., Torchin, M.E., Hechinger, R.F., Dunham, E.J. & Chiba, S. (2005) Molecular-genetic analyses reveal cryptic species of trematodes in the intertidal gastropod, *Batillaria cumingi* (Crosse). *International Journal for Parasitology*, **35**, 793–801.

Miura, O., Torchin, M.E., Kuris, M.E., Hechinger, R.F. & Chiba, S. (2006) Introduced cryptic species of parasites exhibit different invasion pathways. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 19818–19823.

Nadler, S.A. & Perez-Ponce de Leon, G. (2011) Integrating molecular and morphological approaches for characterizing parasite cryptic species implications for parasitology. *Parasitology*, **138**, 1688–1709.

Page, R.D.M. & Charleston, M.A. (1998) Trees within trees: phylogeny and historical associations. *Trends in Ecology and Evolution*, **13**, 356–359.

Panzera, F., Ferrandis, I., Ramsey, J. *et al.* (2006) Chromosomal variation and genome size support existence of cryptic species of *Triatoma dimidiata* with different epidemiological importance as Chagas disease vectors. *Tropical Medicine & International Health*, **11**, 1092–1103.

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

Paxinos, E.E., James, H.F., Olson, S.L., Sorenson, M.D., Jackson, J. & Fletscher, R.C. (2002) mtDNA from fossils reveals a radiation of Hawaiian geese recently derived from the Canada goose (*Branta canadensis*). *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 1399–1404.

Perkins, S.L. (2000) Species concepts and malaria parasites: detecting a cryptic species of *Plasmodium*. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **267**, 2345–2350.

Poulin, R. & Morand, S. (2000) The diversity of parasites. *Quarterly Review of Biology*, **75**, 277–293.

Price, R.D., Hellenthal, R.A., Palma, R.L., Johnson, K.P. & Clayton, D.H. (eds) (2003) *The Chewing Lice: World Checklist and Biological Overview*, Special Publication no. 24. Illinois Natural History Survey, Champaign, IL.

Rambaut, A. & Drummond, A.J. (2007) Tracer v 1.4: MCMC trace analyses tool. http://beast.bio.ed.ac.uk/Tracer [accessed on 21 December 2015].

Ronquist, F. & Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Sherovny, D.F. (2008) Greenland geese in North America. *Birding*, **40**, 46–56.

Stamatakis, A. (2006) RA×ML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.

Stefka, J., Hypsa, V. & Scholz, T. (2009) Interplay of host specificity and biogeography in the population structure of a cosmopolitan endoparasite: microsatellite study of *Ligula intestinalis* (Cestoda). *Molecular Ecology*, **18**, 1187–1206.

Sugiura, N. (1978) Further analysts of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics – Theory and Methods*, **7**, 13–26.

Sweet, A.D., Allen, J.M. & Johnson, K.P. (2014) Novel primers form informative nuclear loci for louse molecular phylogenetics (Insecta: Phthiraptera). *Journal of Medical Entomology*, **51**, 1122–1126.

Tandan, B.K. & Hajela, K.P. (1962) The identity of *Abumarkub Koenigi* Eichler, 1959 (Insecta: Mallophaga). *Journal of Natural History*, **5**, 319–320.

Vilas, R., Criscione, C.D. & Blouin, M.S. (2005) A comparison between mitochondrial DNA and the ribosomal internal transcribed regions in prospecting for cryptic species of platyhelminth parasites. *Parasitology*, **131**, 839–846.

Weckstein, J.D. (2004) Biogeography explains cophylogenetic patterns in toucan chewing lice. *Systematic Biology*, **53**, 154–164.

Wilkerson, R.C., Parsons, T.J., Klein, T.A., Gaffigan, T.V., Bergo, E. & Consolim, J. (1995) Diagnosis by random amplified polymorphic DNA polymerase chain reaction of four cryptic species related to *Anopheles* (*Nyssorhynchus*) *albitarsis* (Diptera: Culicidae) from Paraguay, Argentina, and Brazil. *Journal of Medical Entomology*, **32**, 697–704.

Young, C.M. (1970) Territoriality in the common shelduck *Tadorna tadorna*. *Ibis*, **112**, 330–335.

Zhou, M.J., Xiao, J.H., Bian, S.N. *et al.* (2012) Molecular approaches identify known species, reveal cryptic species and verify host specificity of Chinese *Philotrypesis* (Hymenoptera: Pteromalidae). *Molecular Ecology Resources*, **12**, 598–606.

Zwickl, D.J. (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD Dissertation. University of Texas, Austin, TX.

Accepted 22 August 2015
First published online 11 January 2016