nature ecology & evolution

ARTICLES

https://doi.org/10.1038/s41559-022-01803-1

Phylogenomics reveals the origin of mammal lice out of Afrotheria

Kevin P. Johnson[1]✉, Conrad Matthee[2] and Jorge Doña[1,3]✉

Mammals host a wide diversity of parasites. Lice, comprising more than 5,000 species, are one group of ectoparasites whose major lineages have a somewhat patchwork distribution across the major groups of mammals. Here we explored patterns in the diversification of mammalian lice by reconstructing a higher-level phylogeny of these lice, leveraging whole genome sequence reads to assemble single-copy orthologue genes across the genome. The evolutionary tree of lice indicated that three of the major lineages of placental mammal lice had a single common ancestor. Comparisons of this parasite phylogeny with that for their mammalian hosts indicated that the common ancestor of elephants, elephant shrews and hyraxes (that is, Afrotheria) was the ancestral host of this group of lice. Other groups of placental mammals obtained their lice via host-switching out of these Afrotherian ancestors. In addition, reconstructions of the ancestral host group (bird versus mammal) for all parasitic lice supported an avian ancestral host, indicating that the ancestor of Afrotheria acquired these parasites via host-switching from an ancient avian host. These results shed new light on the long-standing question of why the major groups of parasitic lice are not uniformly distributed across mammals and reveal the origins of mammalian lice.

Mammals are a prominent and conspicuous group of animals and are a main component of the Earth's currently living megafauna. They occupy diverse regions, from polar to tropical, and diverse habitats from terrestrial to marine[1]. Despite a much earlier origin, much of the diversification of mammals occurred after the Cretaceous–Palaeogene (K–Pg) boundary, following the extinction of the dinosaurs[2–7]. Like most vertebrate groups, mammals also host a diversity of parasite lineages that radiated along with them over their evolutionary history[8,9]. Understanding patterns of parasite diversification can shed light on host diversification[10–15], as well as provide insights into the processes of coevolution more broadly[15–17].

Lice are one familiar group of parasites that are widespread across the diversity of mammals, including humans and birds[9,18,19]. These wingless ectoparasitic insects spend their entire lifecycle on the body of the host[15]. Five major groups of lice are recognized[20] and all of these have at least some representation on mammals. These include species with chewing mouthparts (Amblycera, Ischnocera, Trichodectera, Rhynchophthirina) and sucking mouthparts (Anoplura). Most, but not all, major mammalian orders are host to one or more of these major groups of lice. Curiously, the presence of each of these groups is somewhat patchy across the diversity of mammals. For example, rodents broadly host three of these major groups (Amblycera, Trichodectera, Anoplura). In contrast, marsupials host only one group of chewing lice (Amblycera), while the great apes (including humans) host only sucking lice (Anoplura). While the evolutionary history of the blood-feeding sucking lice of great apes is well understood[21,22], the relationships among mammal lice more broadly are less clear[22].

Recent higher-level phylogenetic studies of lice[4,20] have revealed that sucking lice (Anoplura) are clustered together with two groups of mammalian chewing lice (Trichodectera and Rhynchophthirina). This arrangement had not previously been suggested and identifies an expanded group of lice comprising over 1,000 species[18] exclusive to mammals that also radiated after the K–Pg boundary[9]. Curiously, each of the major groups within this newly identified lineage occurs on at least one member of Afrotheria (elephants, elephant shrews and hyraxes, among others), a group of mammals of primarily African distribution, which together with Xenarthra (anteaters, armadillos and sloths), is sister to all other placental mammals[6,7]. For example, elephants host only members of Rhynchophthirina (of which there are only three species[18]). Elephant shrews host only members of Anoplura, and hyraxes host members of both Trichodectera and Anoplura[18,19]. These patterns of host distribution raise the question: what was the original host group of placental mammal lice? Furthermore, are there broad patterns of codiversification between mammals and their lice at deep taxonomic scales, or is the patchy distribution of lice across mammals the result of parasite extinction and/or host-switching?

To address these questions, we expand previous taxonomic sampling[9] of mammalian louse genomes to include additional key lineages hosted by members of Afrotheria (including lice from elephant shrews and hyraxes) and additional sampling across the major mammalian louse clade. We leverage these genomic sequences to construct a phylogenomic dataset for lice more broadly, but with focus on the major mammal louse clade. The evolutionary history of these mammal lice is compared with that for their mammal hosts to identify the ancestral host lineage and to understand the dynamics of diversification of this major group of parasites.

Mammal louse cophylogenetics

The phylogeny of parasitic lice (Fig. 1) based on a target set of 2,395 single-copy nuclear genes (3,921,975 aligned base pairs; bp) recovered all major groups as monophyletic (Amblycera, Ischnocera, Trichodectera, Anoplura; note Rhynchophtirina was sampled by only one species, so its monophyly could not be tested). Importantly for this study, a large mammal louse group (Trichodectera, Rhychophthirina, Anoplura) had a single common ancestor, and

[1]Illinois Natural History Survey, Prairie Research Institute, University of Illinois, Champaign, IL, USA. [2]Department of Botany and Zoology, Stellenbosch University, Stellenbosch, South Africa. [3]Departamento de Biología Animal, Universidad de Granada, Granada, Spain. ✉e-mail: kpjohnso@illinois.edu; jorged@illinois.edu

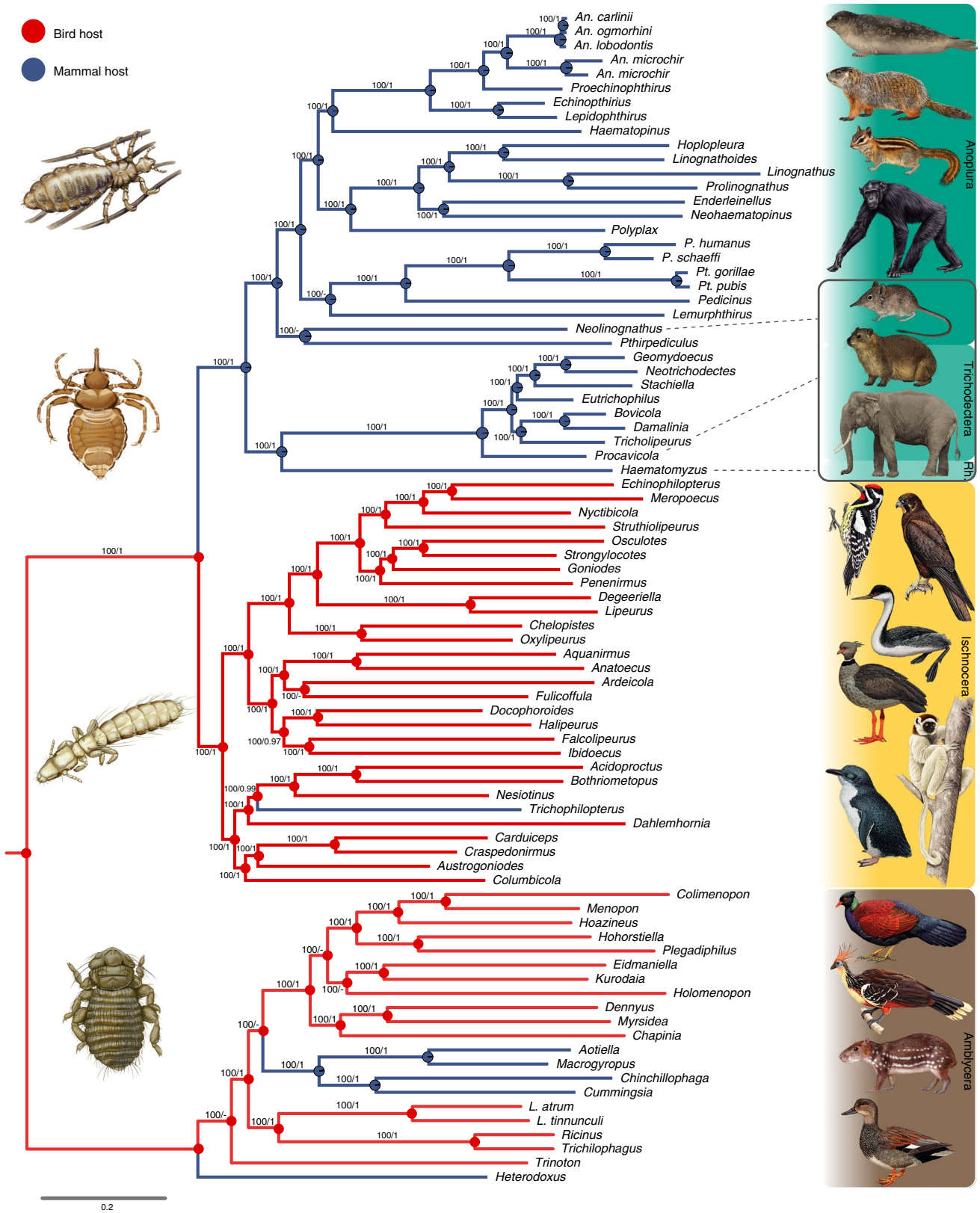


Fig. 1 | Phylogenetic tree from maximum likelihood analysis of the concatenated alignment of 3,921,975 bp from 2,395 single-copy target nuclear gene orthologues. The number associated with branches to the left of the slash is ultrafast bootstrap support. The number to the right of the slash is local posterior probability from ASTRAL coalescent gene tree/species tree analysis. Dashes indicate nodes not present in the coalescent tree. Pie charts on nodes represent the relative likelihoods from maximum likelihood reconstruction of bird (red) or mammal (blue) host under the all-rates-different model (root age: 92 Ma). Major louse groups are indicated with coloured shading on the right (Rh., Rhynchophthirina) and louse images on the left. Bird and mammal images are representative hosts for louse parasites in the tree. Grey box and dashed lines indicate key louse genera from Afrotheria hosts. Credit: louse images on the left, ©Jacqueline Mahannah; bird and mammal images on the right, ©Lynx Edicions. Scale bar indicates number of substitutions per site.

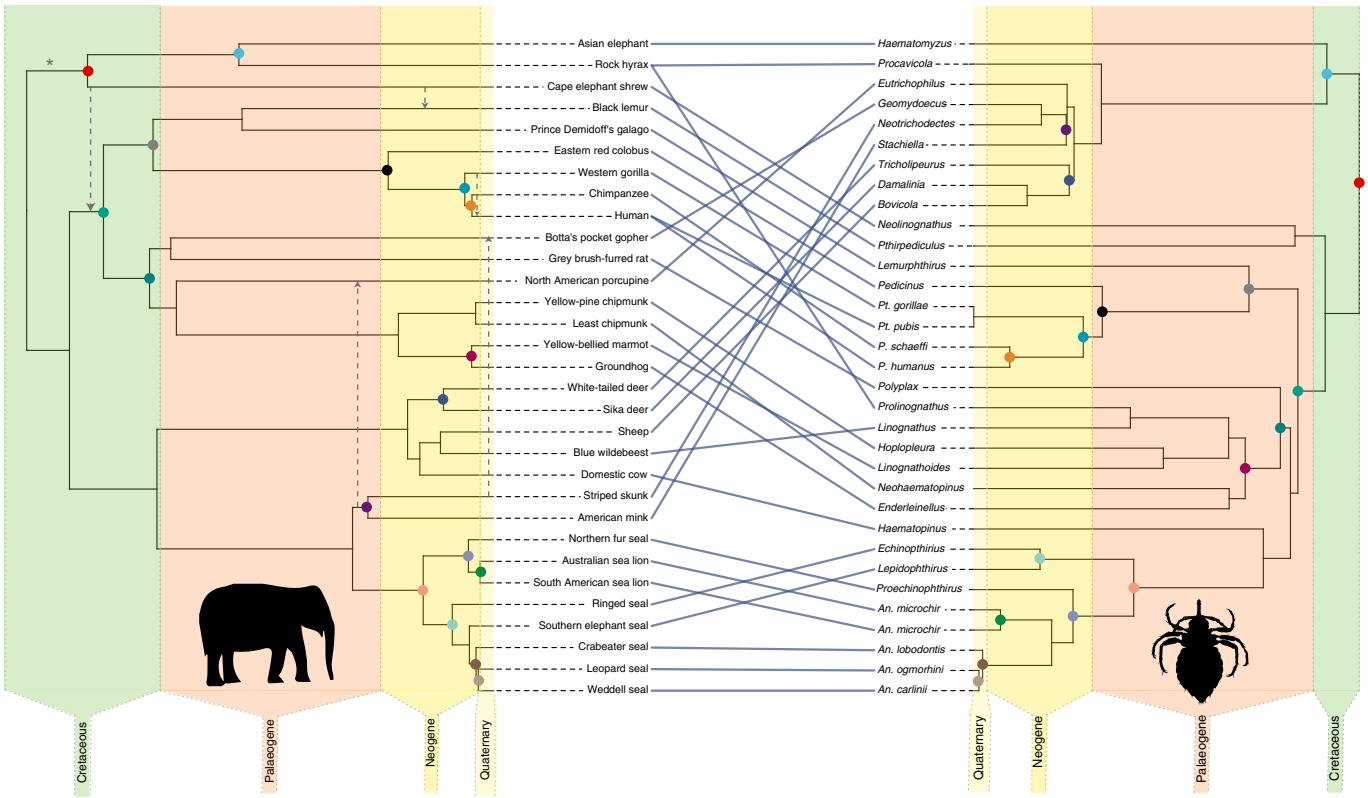


Fig. 2 | Cophylogenetic comparison of dated mammal (left) and mammal louse (right) phylogenies. Mammal tree and dates from a comprehensive tree of mammals[5]. Louse tree is dated using a combination of fossil and codivergence calibrations and the residual least squares method (Methods). Blue lines link lice with their respective mammal hosts. Coloured circles link louse and mammal nodes that are shared cospeciation events between mammal and louse trees. Geological timescale is indicated by coloured shading (Quartenary 0–2.6 Ma, Neogene 2.6–23 Ma, Palaeogene 23–66 Ma and Cretaceous 66–145 Ma) to the same scale for both mammal and louse trees. Dashed lines with arrows indicate key host-switching events mentioned in the main text. Asterisk indicates the mammal lineage (Afrotheria) on which mammal lice were inferred to have originated. Elephant silhouette extracted from Phylopic (www.phylopic.org).

this group received maximal support by both concatenated (100% bootstrap) and coalescent (1.0 local posterior probability) analyses. In general, the trees from concatenated and coalescent analyses were highly similar, with only seven branches not present in both trees (Fig. 1).

Within the mammal louse clade, lice from Afrotherian hosts (elephant, hyrax, elephant shrew) were among the earliest diverging branches. Specifically, the elephant louse (Rhynchophthirina: *Haematomyzus*) was sister to all of Trichodectera, and within Trichodectera, the hyrax louse (*Procavicola*) was the earliest diverging lineage. Within sucking lice (Anoplura), the elephant shrew louse (*Neolinognathus*) was a member of the earliest diverging lineage (concatenated tree) or sister to all other Anoplura (coalescent tree).

The early diverging phylogenetic position of these Afrotherian lice resulted in the common ancestor of this mammalian group being reconstructed as the ancestral host for the entire mammal louse clade (Fig. 2). This reconstruction was stable across both louse trees (concatenated and coalescent) and across all maximum parsimony reconciliations (MPRs) and evaluated cost schemes in the eMPRess analysis. Furthermore, the divergence between elephant lice (*Haematomyzus*) and hyrax lice (*Procavicola*) is also a cospeciation event under all reconstructions, indicating that the common ancestor of elephants and hyraxes also inherited its lice from the ancestor of all Afrotheria. Subsequent radiation of lice across mammals was a result of host-switching from Afrotheria to other mammal lineages. Overall, there were either 17 (concatenated tree) or 15 (coalescent tree) reconstructed cospeciation events, and this was many more than expected by chance ($P<0.01$), although host-switching of lice among major mammal lineages was also common (15 or 17 host-switching events; Extended Data Figs. 1 and 2). Distance-based cophylogenetic methods also recovered statistically significant congruence ($P<0.001$) between host and parasite trees, suggesting an extensive history of codivergence. In addition, 17 host–parasite links had mean squared residual values (Extended Data Fig. 3) greater than or equal to the overall median squared residual value, and these are interpreted as links with topological conflict[23]. A total of 14 of 17 (82%) of these associations were directly involved in host-switching events as inferred by eMPRess, either as donor or recipient lineages, including the two lineages involved in switches out of Afrotheria (elephant shrew louse: *Neolinognathus* and rock hyrax louse: *Procavicola*). These results indicated general concordance between event-based[24] and distance-based methods[23] regarding the occurrence of host-switching events.

Cospeciation at the node of earliest divergence in mammal lice implies the divergence in the ancestral host, that is, the earliest divergence within Afrotheria, also occurred at that time. The confidence intervals on the dates we estimated for this earliest divergence in mammal lice broadly overlap with published estimates of the earliest divergence within Afrotheria (Supplementary Table 1, except for the analysis with the youngest root age constraint, which would imply slightly delayed cospeciation). For maximum likelihood estimation of the ancestral host (bird or mammal) over these dated trees, the all-rates-different model was favoured by the Akaike information criterion (AIC; 92 Myr ago (Ma): equal rates AIC = 38.18, all-rates-different AIC = 37.11; 131 Ma: equal rates

AIC = 38.0, all-rates-different AIC = 37.0; 171 Ma: equal rates AIC = 37.97, all-rates-different AIC = 36.98). Under this model, an avian host was reconstructed at the ancestor for all lice with 100% relative likelihood (Fig. 1) over all dated trees. Other louse ancestors uniting parasite lineages with descendent lineages of bird versus mammal hosts were also reconstructed as an avian host ancestor with 100% relative likelihood. This result indicates that there were four transitions (that is, host switches) from avian to mammalian hosts over the evolutionary history of lice, including the ancestor of the large mammal louse clade originating on Afrotheria.

Discussion

Cophylogenetic analysis of a major group of mammalian parasites indicated that early divergences in these parasitic lice occurred on ancestors of a mammalian group with primarily African distribution: Afrotheria (elephants, hyraxes, elephant shrews and relatives). This result sheds new light on the origins and diversification of mammal lice. After initial codivergence with ancestral Afrotheria, these lice went on to colonize, through host-switching, other lineages of placental mammals and later codiversified with them (Fig. 2).

For example, the ancestor of lice from elephant shrews colonized either the common ancestor of rodents and primates (concatenated louse tree) or the ancestor of a lineage within primates (galagos and Madagascan lemurs, coalescent louse tree). This colonization event, in either case, was followed by additional diversification of lice within primates by both cospeciation and further host-switching of lice among primates, leading to the presence of two genera (*Pediculus* and *Pthirus*) of lice on humans. Interestingly, Madagascan lemurs, or their direct ancestors, seem to be involved in acquiring lice from ancestral elephant shrews. Another louse genus (*Trichophilopterus*) that occurs on Madagascan lemurs is derived from within avian feather lice (Ischnocera; Fig. 1), again suggesting a host switch to these primates, in this case from birds to mammals. It seems that the ancestor of Madagascan lemurs was possibly free of parasitic lice and represented an open niche, facilitating host-switching[25].

The genera of sucking lice (*Polyplax*, *Neohaematopinus*, *Enderleinellus*, *Linognathoides* and *Hoplopleura*) of rodents fall within a single group of Anoplura. In many cases, single species of rodents can be infested with multiple genera of sucking lice, suggesting the diversity of these lice on rodents arose by speciation of parasites within this host lineage (that is, duplication). The presence of sucking lice within the largely rodent louse clade on hyraxes and wildebeest appears to have been the result of host-switching of lice from rodents to these other mammal lineages in Africa. Some New World rodent lineages, such as porcupines and pocket gophers, are not host to any sucking lice, and these rodents acquired their lice (Trichodectera) by host-switching from carnivorous mammals (Carnivora). Although initially somewhat unexpected, host-switching of lice from predators to prey has also been shown for the feather lice of avian raptors, which also appear to have switched to their prey[26]. It may be that if initial capture followed by escape of prey occurs frequently, host-switching of lice from predators to prey could be facilitated.

Even though there are prominent cases of host-switching of lice between major lineages of placental mammals, a large proportion (47–53%) of parasite divergence events were found to be the result of cospeciation. This is also the case in the percentage of mammalian host branching points, with 50–57% of nodes in the host tree associated with a cospeciation event by their parasitic lice. The dominance of cospeciation of the lice of mammals stands in contrast to a comparable study of avian feather lice (Ischnocera)[27], in which across a similarly broad diversity of lice, only 17% of avian hosts nodes were associated with a cospeciation event by their parasitic feather lice. Perhaps the reduced mobility of mammals compared with birds makes host-switching more difficult, and this has been suggested as an ecological factor facilitating the strong signatures of codivergence, and lack of host-switching, observed in gopher lice[28,29].

In summary, the mammalian lineage Afrotheria played a major role as an ancestral host for a major lineage of mammalian parasites. Much of the subsequent diversification of this group of lice can be attributed to an 'out-of-Afrotheria' scenario, in which lice from ancestral Afrotheria colonized other placental mammals. This raises the question of how the ancestor of Afrotheria acquired their lice in the first place. Through maximum likelihood character state reconstruction, we inferred that the ancestors of Afrotheria acquired their lice from an avian host, leading to the diversification of this major group of mammalian parasites.

Methods

Taxon sampling. We sampled 33 louse species (across 27 genera) from across the major mammal louse clades (Trichodectera, Rhynchophthirina, Anoplura) for genome sequencing (Supplementary Table 2). These included 14 species newly sequenced for this study. Most notably, our new sequences included one species of Anoplura from an elephant shrew and two species (representing both Anoplura and Trichodectera) from a hyrax, which expanded our sampling of these parasites from Afrotheria beyond the elephant louse (Rhynchophthirina) already sequenced. We also included broad outgroup sampling of 29 species (across 29 genera) of Ischnocera and 21 species (across 20 genera) of Amblycera, including representatives of all major louse lineages with bird and mammals hosts within each of these outgroups.

DNA extraction and genome sequencing. For newly sequenced samples, total genomic DNA extractions were prepared from a single specimen that had been stored in 95% ethanol at −80 °C, which was first photographed as a voucher. Prior to extraction, the sample was ground using a plastic pestle in a 1.5 ml tube. Extractions followed manufacturers' protocols for the QIAamp DNA Micro Kit (Qiagen), with initial incubation in ATL buffer with proteinase K at 55 °C for 48 hours. Final elution was in 50 µl buffer AE, and DNA was quantified with the high-sensitivity kit using a Qubit 2.0 Fluorometer (Invitrogen).

From total genomic DNA extractions, the Hyper library kit (Kapa Biosystems) was used to prepare Illumina libraries. An Illumina NovaSeq 6000 S4 lane multiplexed with 48 libraries tagged with unique dual-end adaptors was used to sequence 150 bp paired-end reads to achieve at least 30–60× coverage of the nuclear genome. To generate fastq files for each library, files were demultiplexed and adaptors trimmed using bcl2fastq v.2.20, and raw reads were deposited in the National Center for Biotechnology Information Sequence Read Archive (NCBI SRA; Supplementary Table 2). Raw reads from previously sequenced samples were downloaded from the NCBI SRA (Supplementary Table 2 for details).

Gene assembly and phylogenomics. We used aTRAM v.2.0[30] to assemble a target set of 2,395 single-copy orthologue protein coding genes, using reference amino acid sequences from the human louse *Pediculus humanus*. These amino acid sequences were used in tblastn searches of the genomic raw read libraries to locally assemble each orthologue gene. Specific assembly, alignment and phylogenomic analysis processing steps, parameters and commands followed previous studies[31]. After assembly, exon sequences were identified and stitched together to remove intron sequences using an Exonerate-based[32] stitching pipeline (atram_stitcher[30]).

Nucleotide sequences were translated to amino acids and aligned using MAFFT v.7.471[33]. After back-translation to nucleotide sequences, individual gene alignments were trimmed using trimAL v.1.4.rev22[34] with a 40% gap threshold. Any gene present for less than four taxa was discarded. Gene alignments were then concatenated into a supermatrix and analysed under maximum likelihood using IQ-TREE 2 v.2.1.2[35] in a partitioned analysis that included model selection for each partition. All trees were rooted on Amblycera based on prior studies[9]. Support was estimated using ultrafast bootstrapping. To account for the potential that incomplete lineage sorting could result in gene trees not reflecting the species tree[36], a coalescent analysis using ASTRAL-III[37] was conducted on individual gene trees estimated by maximum likelihood in IQ-TREE 2. As a measure of branch support, local posterior probability for each branch was also computed using ASTRAL-III.

Cophylogenetics and molecular dating analyses. To estimate the ancestral host of the mammal louse clade (Trichodectera, Rhynchophthirina, Anoplura), we used eMPRess[24] to compare the phylogenies (both concatenated and coalescent) estimated for this group against a phylogeny for their mammal hosts. For the mammal phylogeny, we used the most comprehensive species-level phylogeny to date[5], because this tree included all the mammal host species corresponding to our louse parasite sampling. The major branches in this mammal host tree are generally consistent across a number of recent studies[2–4,6,7,38] and reflect the best current knowledge of the mammal phylogeny. For the cophylogenetic analysis, we used a cost scheme of 1 for duplication and sorting events, and 2 for host-switching. We

also explored the six adjacent cost scheme zones to this scheme in eMPRess, which included host-switching costs ranging from zero to three times the cost of sorting events, and duplication costs ranging from zero to ten times the cost of sorting events. We followed previously outlined procedures for summarizing the MPRs[31]. The parasite tree was randomized 100 times compared with the host tree to test whether the number of cospeciation events was more than expected by chance. In addition, we ran two distance-based methods, PACo[23] (default parameters, 100,000 permutations) and ParaFit[39] (default parameters), to evaluate if there was an overall signature of cospeciation in our dataset.

For each median MPR cluster reconstruction, we ascertained the reconstructed ancestral host of the mammal clade of lice. To evaluate whether the age of this host ancestor was compatible between host (mammal) and parasite (louse) trees, we performed a dating analysis on the concatenated louse tree. For this method, we used the residual least squares method in IQ-TREE and the same calibration points for internal nodes used by previous molecular dating analyses of lice (split between human and chimp lice 5–7 Ma, split between the lice from Old World primates and Great Apes 20–25 Ma, minimum age for Menoponidae of 44 Ma based on fossil[2]). We performed this dating analysis using three possible root ages (that is, split between Amblycera and other parasitic lice) that span estimates from prior studies: 92 Ma[9], 131 Ma[40] and 171 Ma[41].

For lice more broadly, we also reconstructed whether the ancestral host was a bird or mammal using maximum likelihood. Each louse was coded as being hosted by a bird or mammal, and the time-calibrated trees resulting from the three possible root ages were used for the reconstruction. For this analysis, we used the ace function (equal rates and all-rates-different models) in the R package APE v.5.5[42]. We evaluated whether the equal rates model could be rejected in favour of the all-rates-different model using AIC.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Information. Phylogenomic data generated in this study are available at figshare (https://doi.org/10.6084/m9.figshare.18737423).

Received: 28 January 2022; Accepted: 19 May 2022;
Published online: 04 July 2022

References

1. Wilson, D. E., Mittermeier, R. A. & Cavallini, P. *Handbook of the Mammals of the World* Vol. 1 (Lynx Edicions, 2009).
2. Meredith, R. W. et al. Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* **334**, 521–524 (2011).
3. dos Reis, M. et al. Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. *Proc. R. Soc. B* **279**, 3491–3500 (2012).
4. O'Leary, M. A. et al. The placental mammal ancestor and the post-K–Pg radiation of placentals. *Science* **339**, 662–667 (2013).
5. Upham, N. S., Esselstyn, J. A. & Jetz, W. Inferring the mammal tree: species-level sets of phylogenies for questions in ecology, evolution, and conservation. *PLoS Biol.* **17**, e3000494 (2019).
6. Murphy, W. J., Foley, N. M., Bredemeyer, K. R., Gatesy, J. & Springer, M. S. Phylogenomics and the genetic architecture of the placental mammal radiation. *Annu. Rev. Anim. Biosci.* **9**, 29–53 (2021).
7. Álvarez-Carretero, S. et al. A species-level timeline of mammal evolution integrating phylogenomic data. *Nature* **602**, 263–267 (2022).
8. Kim, K. C. *Coevolution of Parasitic Arthropods and Mammals* (Wiley, 1986).
9. Johnson, K. P. et al. Simultaneous radiation of bird and mammal lice following the K–Pg boundary. *Biol. Lett.* **14**, 20180141 (2018).
10. Hafner, M. S. & Nadler, S. A. Phylogenetic trees support the coevolution of parasites and their hosts. *Nature* **332**, 258–259 (1988).
11. Huelsenbeck, J. P. & Rannala, B. Phylogenetic methods come of age: testing hypotheses in an evolutionary context. *Science* **276**, 227–232 (1997).
12. Page, R. D. M. & Charleston, M. A. Trees within trees: phylogeny and historical associations. *Trends Ecol. Evol.* **13**, 356–359 (1998).
13. Nieberding, C. M. & Olivieri, I. Parasites: proxies for host genealogy and ecology? *Trends Ecol. Evol.* **22**, 156–165 (2007).
14. Allen, J. M., Worman, C. O., Light, J. E. & Reed D. L. in *Developments in Primatology: Progress and Prospects* Vol. 38 (eds Brinkworth, J. & Pechenkina, K.) 161–186 (Springer, 2013).
15. Clayton, D. H., Bush, S. E. & Johnson, K. P. *Coevolution of Life on Hosts: Integrating Ecology and History* (Univ. Chicago Press, 2015).
16. Althoff, D. M., Segraves, K. A. & Johnson, M. T. J. Testing for coevolutionary diversification: linking pattern with process. *Trends Ecol. Evol.* **29**, 82–89 (2014).
17. Hayward, A., Poulin, R. & Nakagawa, S. A broadscale analysis of host-symbiont cophylogeny reveals the drivers of phylogenetic congruence. *Ecol. Lett.* **24**, 1681–1696 (2021).
18. Price, R. D., Hellenthal, R. A., Palma, R. L., Johnson, K. P. & Clayton, D. H. *The Chewing Lice: World Checklist and Biological Overview* (Illinois Natural History Survey, 2003).
19. Durden, L. A. & Musser, G. G. *The Sucking Lice (Insecta, Anoplura) of the World: A Taxonomic Checklist with Records of Mammalian Hosts and Geographical Distributions* (American Museum of Natural History, 1994).
20. de Moya, R. S. et al. Phylogenomics of parasitic and non-parasitic lice (Insecta: Psocodea): combining sequence data and exploring compositional bias solutions in next generation datasets. *Syst. Biol.* **70**, 719–738 (2021).
21. Reed, D. L., Light, J. E., Allen, J. M. & Kirchman, J. J. Pair of lice lost or parasites regained: the evolutionary history of anthropoid primate lice. *BMC Biol.* **5**, 7 (2007).
22. Light, J. E., Smith, V. S., Allen, J. M., Durden, L. A. & Reed, D. L. Evolutionary history of mammalian sucking lice (Phthiraptera: Anoplura). *BMC Evol. Biol.* **10**, 292 (2010).
23. Balbuena, J. A., Míguez-Lozano, R. & Blasco-Costa, I. PACo: a novel Procrustes application to cophylogenetic analysis. *PLoS ONE* **8**, e61048 (2013).
24. Santichaivekin, S. et al. eMPRess: a systematic cophylogeny reconciliation tool. *Bioinformatics* https://doi.org/10.1093/bioinformatics/btaa978 (2020).
25. Boyd, B. M. et al. Long-distance dispersal of pigeons and doves generated new ecological opportunities for host-switching and adaptive radiation by their parasites. *Proc. R. Soc. B* https://doi.org/10.1098/rspb.2022.0042 (2022).
26. Catanach, T. A. & Johnson, K. P. Independent origins of the feather lice (Insecta: *Degeeriella*) of raptors. *Biol. J. Linn. Soc.* **114**, 837–847 (2015).
27. de Moya, R. S. et al. Extensive host-switching of avian feather lice following the Cretaceous–Paleogene mass extinction event. *Commun. Biol.* **2**, 445 (2019).
28. Hafner, M. S. et al. Disparate rates of molecular evolution in cospeciating hosts and parasites. *Science* **265**, 1087–1090 (1994).
29. Page, R. D. M. (ed.) *Tangled Trees: Phylogeny, Cospeciation, and Coevolution* (Univ. Chicago Press, 2003).
30. Allen, J. M., LaFrance, R., Folk, R. A., Johnson, K. P. & Guralnick, R. P. aTRAM 2.0: an improved, flexible locus assembler for NGS data. *Evol. Bioinform.* **14**, 1176934318774546 (2018).
31. Johnson, K. P., Weckstein, J. D., Herrera, S. V. & Doña, J. The interplay between host biogeography and phylogeny in structuring diversification of the feather louse genus *Penenirmus*. *Mol. Phylogenet. Evol.* **165**, 107297 (2021).
32. Slater, G. S. C. & Birney, E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinform.* **6**, 31 (2005).
33. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
34. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
35. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
36. Degnan, J. H. & Rosenberg, N. A. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* **24**, 332–340 (2009).
37. Zhang, C., Rabiee, M., Sayyari, E. & Mirarab, S. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinform.* **19**, 153 (2018).
38. Ronquist, F., Lartillot, N. & Phillips, M. J. Closing the gap between rocks and clocks using total-evidence dating. *Phil. Trans. R. Soc. B* **371**, 20150136 (2016).
39. Legendre, P., Desdevises, Y. & Bazin, E. A statistical test for host–parasite coevolution. *Syst. Biol.* **51**, 217–234 (2002).
40. Smith, V. S. et al. Multiple lineages of lice pass through the K–Pg boundary. *Biol. Lett.* **7**, 782–785 (2011).
41. Johnson, K. P. et al. Phylogenomics and the evolution of hemipteroid insects. *Proc. Natl Acad. Sci. USA* **115**, 12775–12780 (2018).
42. Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2018).

Acknowledgements

We thank J. Allen, K. C. Bell, I. Beveridge, S. Bush, T. Chesser, D. Clayton, C. Floyd, R. Furness, T. D. Galloway, S. Goodman, P. James, R. E. Junge, A. Lawrence, S. Leonardi, J. Malenke, S. Matthee, K. McCracken, M. Meyer, L. Mugisha, T. Nyman, B. O'Shea, E. Osnas, R. Palma, J. Scherer, V. Smith, T. Spradling, O. Sychra, W. Veronesi, D. Verrier, J. Weckstein and R. Wilson for assistance in obtaining specimens for this study. Funding was provided by US NSF DEB-1239788, DEB-1925487 and DEB-1926919 grant awards to K.P.J, and European Commission grant H2020-MSCA-IF-2019 (INTROSYM:886532) to J.D. We thank R. de Moya, S. Virrueta Herrera and K. K. O. Walden for assistance with DNA extraction. We thank A. Hernandez and the Roy J. Carver Biotechnology Center at the University of Illinois for assistance with genome sequencing. We thank K. K. O.

Walden for assistance with submission of data to NCBI. Images copyright Lynx Edicions originally illustrated by H. Burn, F. Jutglar, T. Llobet, F. Peacock, L. Sanz, L. Solé and I. Velikov. We thank L. J. Revell for assistance with Fig. 2. We thank CSIRC personnel (Universidad de Granada, Spain) for assistance and providing computational resources (Alhambra supercomputer).

Author contributions

K.P.J. designed the study, obtained funding and wrote the manuscript draft. C.M. provided critical samples and edited the manuscript. J.D. designed the study, conducted the analyses, prepared figures and edited the manuscript.

Competing interests

The authors declare no competing interests.

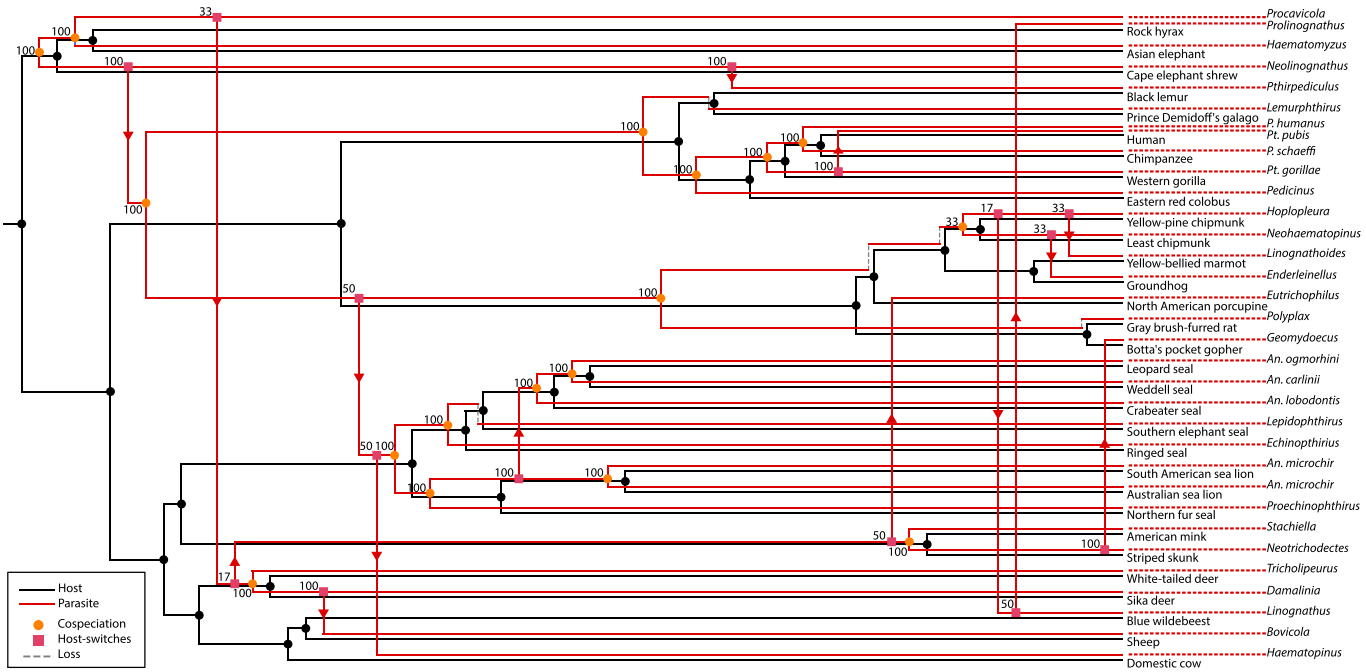
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41559-022-01803-1.

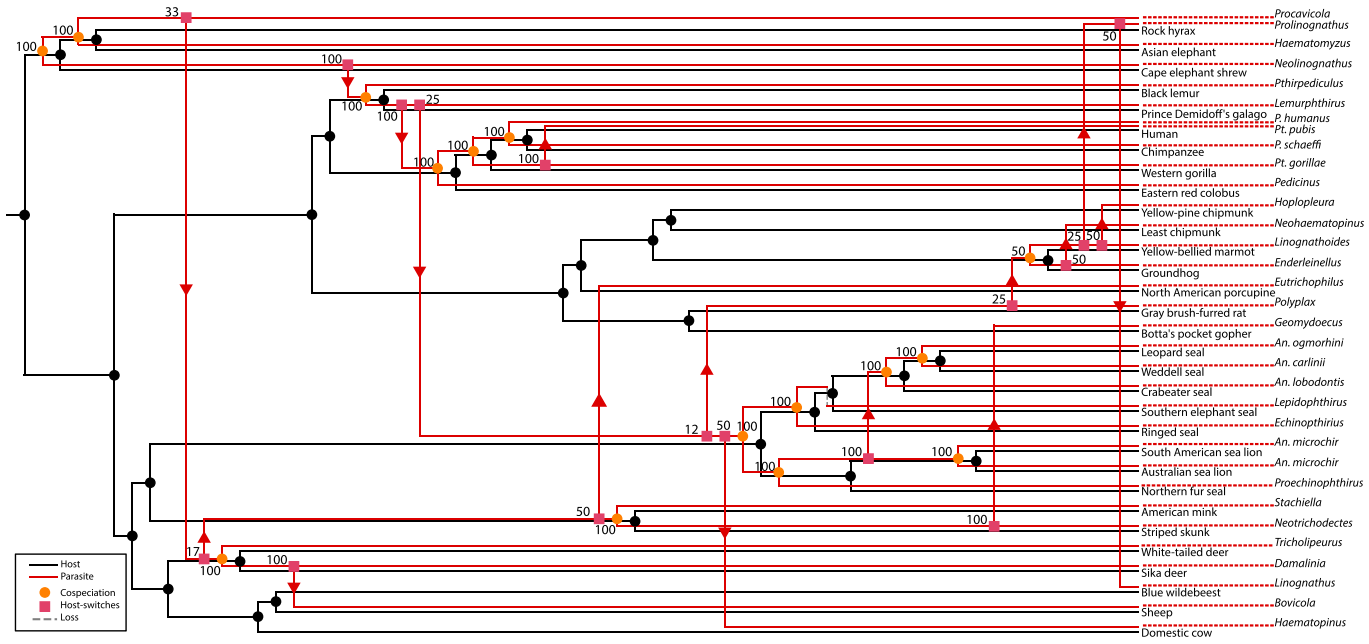
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41559-022-01803-1.

Correspondence and requests for materials should be addressed to Kevin P. Johnson or Jorge Doña.

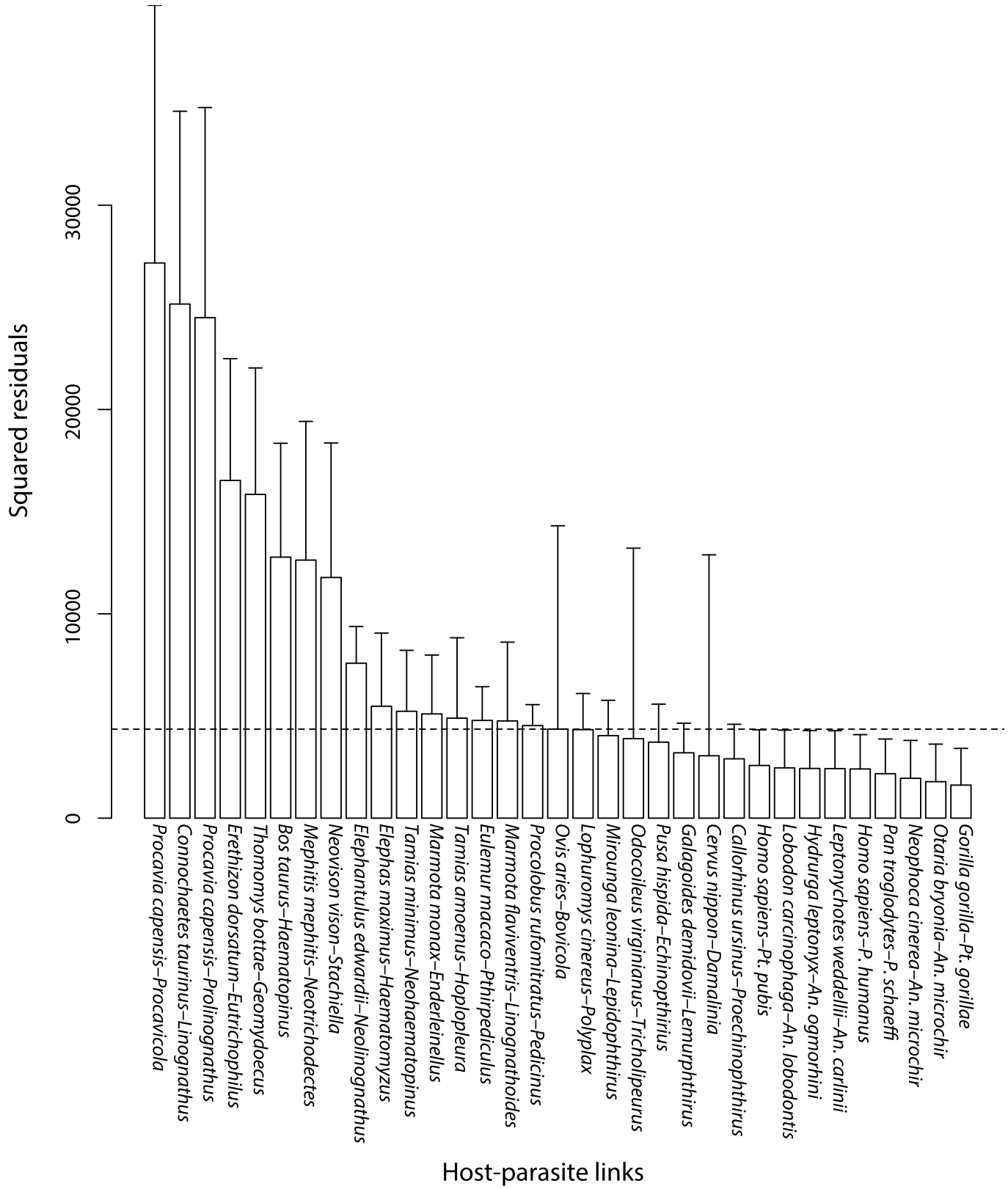
Peer review information *Nature Ecology & Evolution* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.



Extended Data Fig. 1 | Summary of cophylogenetic reconstruction of optimal MPRs from eMPRess comparison (cost scheme duplication: 1, sorting: 1, and host-switching: 2) of the louse (concatenated) tree with the mammal host tree. Arrows indicate direction of host-switches. Numbers associated with events are the percentage of MPRs with that event.



Extended Data Fig. 2 | Summary of cophylogenetic reconstruction of optimal MPRs from eMPRess comparison (cost scheme duplication: 1, sorting: 1, and host-switching: 2) of the louse (coalescent) tree with the mammal host tree. Arrows indicate direction of host-switches. Numbers associated with events are the percentage of MPRs with that event.



Extended Data Fig. 3 | Jack-knifed squared residuals (bars) and upper 95% confidence interval (error bars) associated with each mammal-louse association (link). Dashed line indicates the overall median squared residual value (n = 33 biologically independent samples).