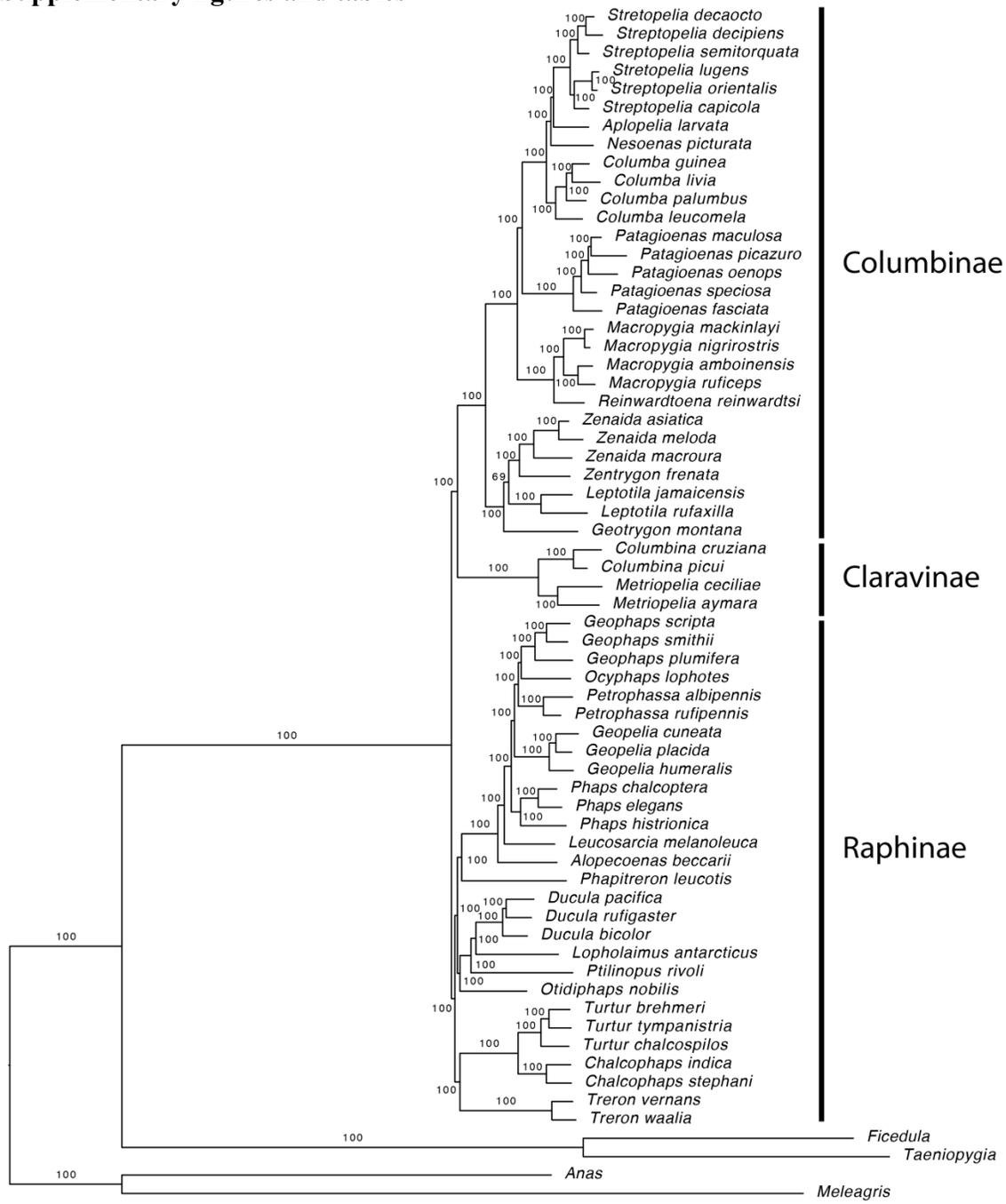
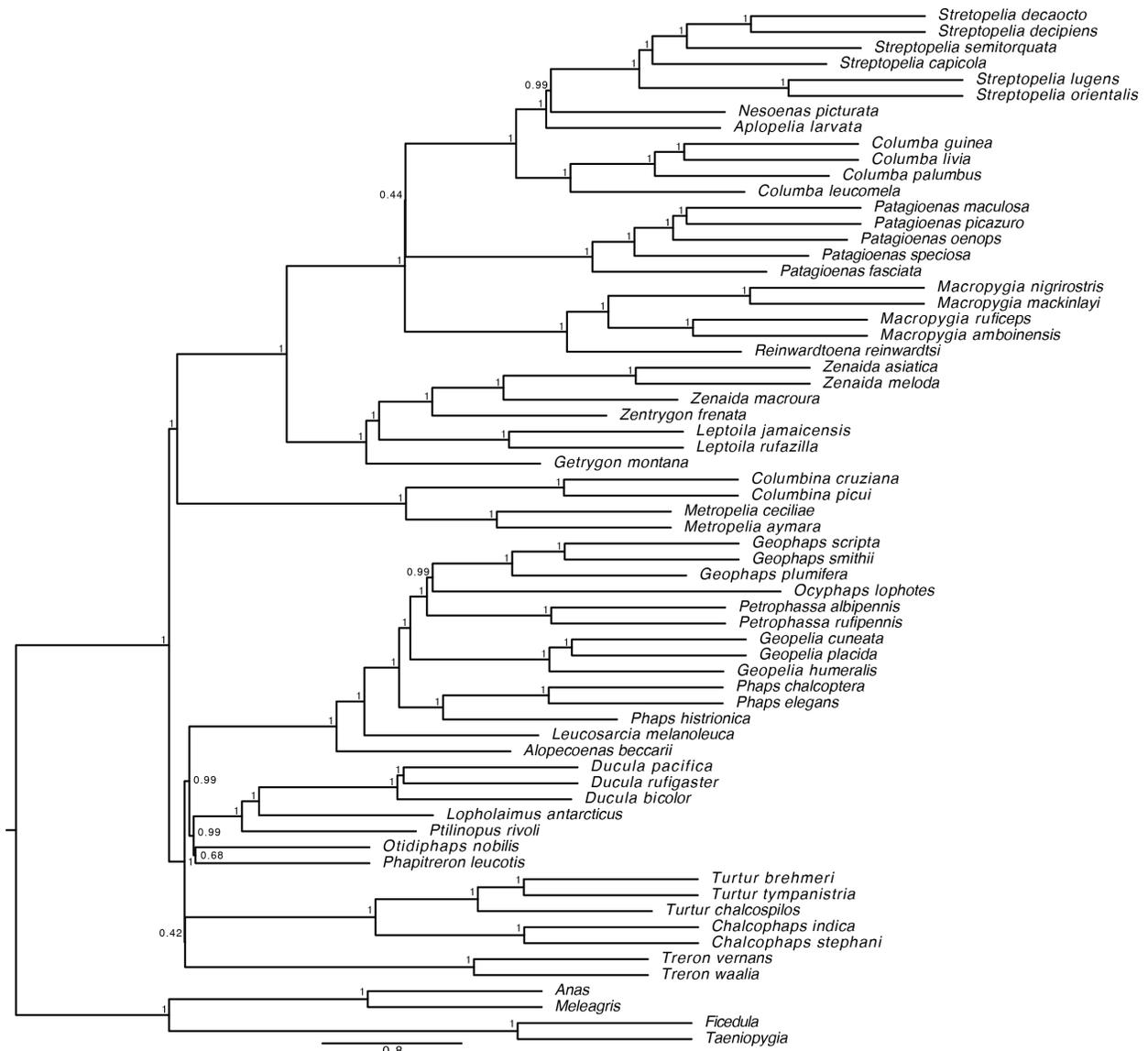


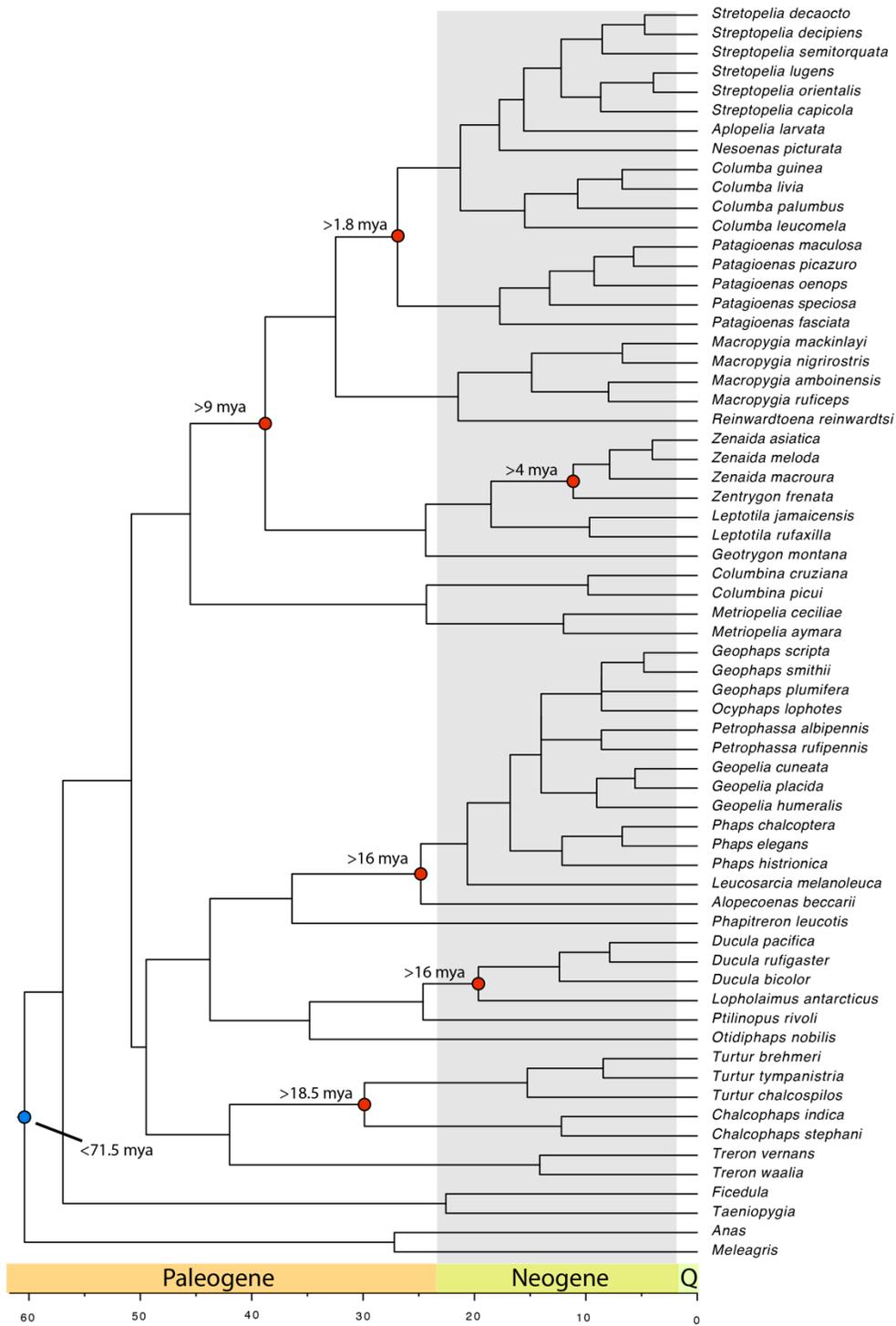
1 **Supplementary figures and tables**



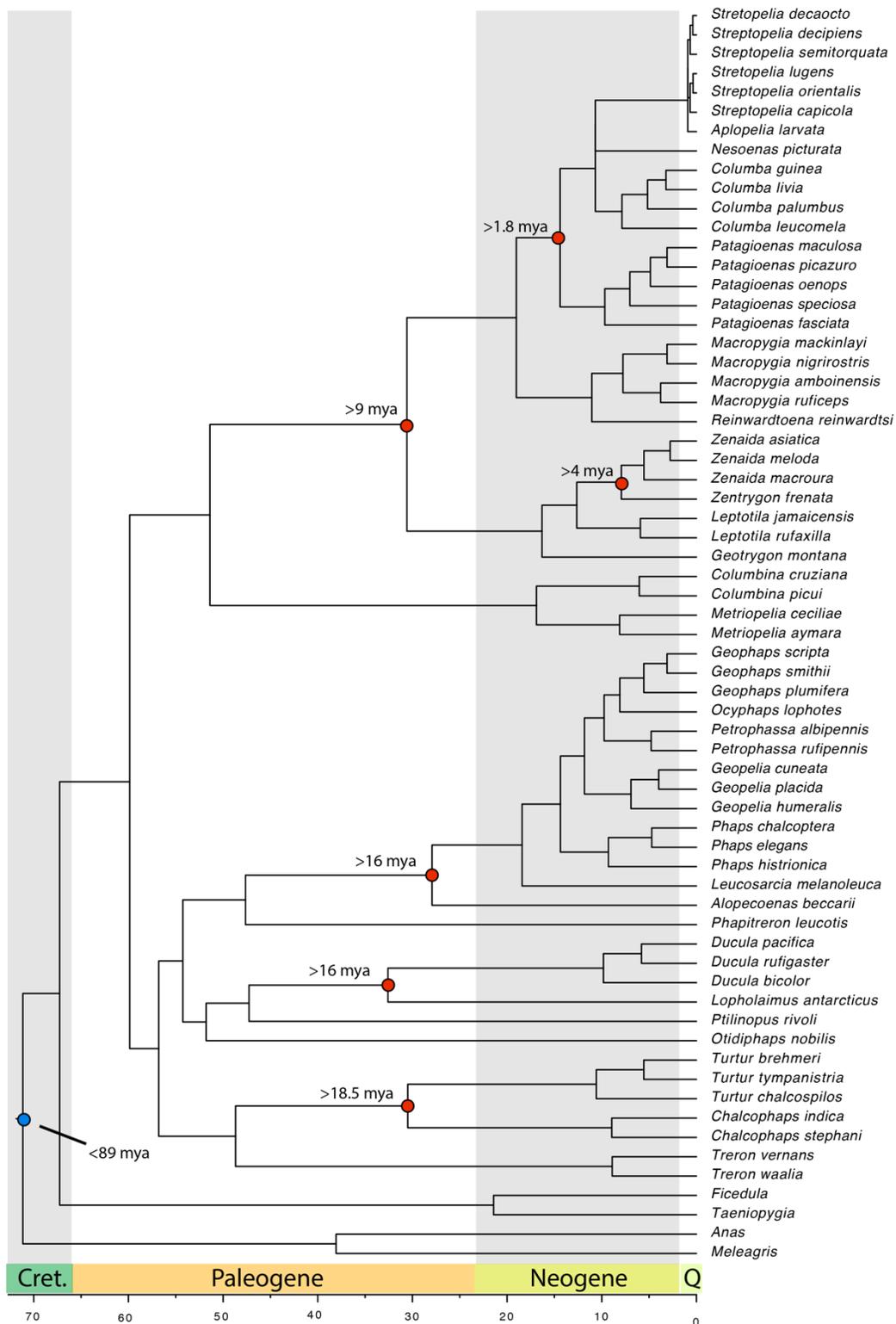
2  
 3 **Figure S1.** Maximum-likelihood tree showing the relationships of dove species based on the  
 4 concatenation of 6,363 single-copy orthologous gene sequence alignments. The tree was  
 5 calculated using RAxML and support was calculated from bootstrap replicates. Numbers at nodes  
 6 indicate support as percent of 100 bootstrap replicates that also recovered the same node. Vertical  
 7 bars and names indicate subfamilies.



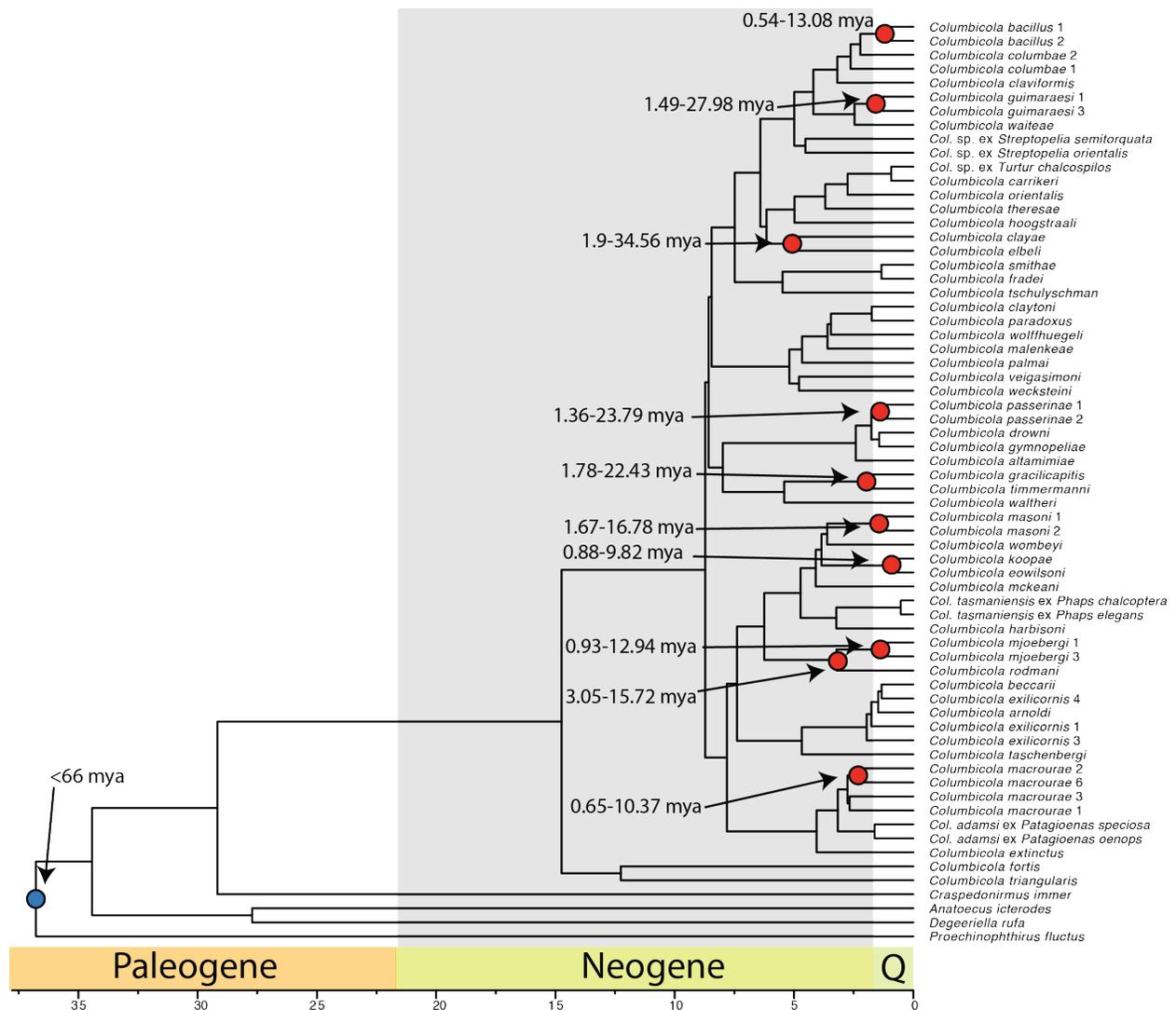
8  
 9 **Figure S2.** Coalescent tree showing relationships of dove species based on 6,363 single-copy  
 10 ortholog gene trees. Tree was calculated using ASTRAL from gene trees calculated using RAxML  
 11 and numbers at nodes indicate support based on local probabilities.



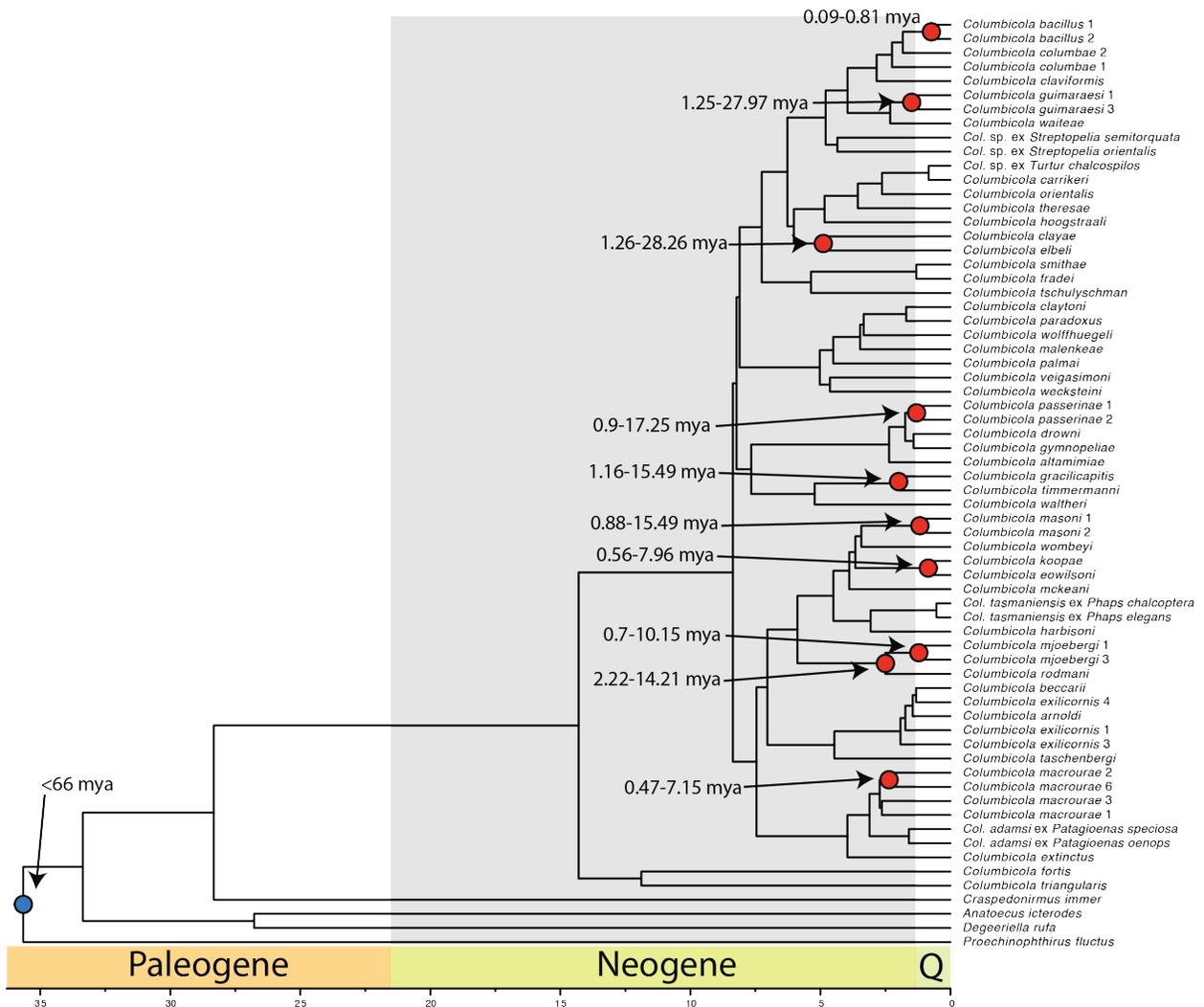
12  
 13 **Figure S3.** Time calibrated maximum-likelihood tree showing the relationships of dove species  
 14 based on the concatenation of 6,363 single-copy orthologous gene sequence alignments. Tree  
 15 calibrated with a root of 74.5 mya.



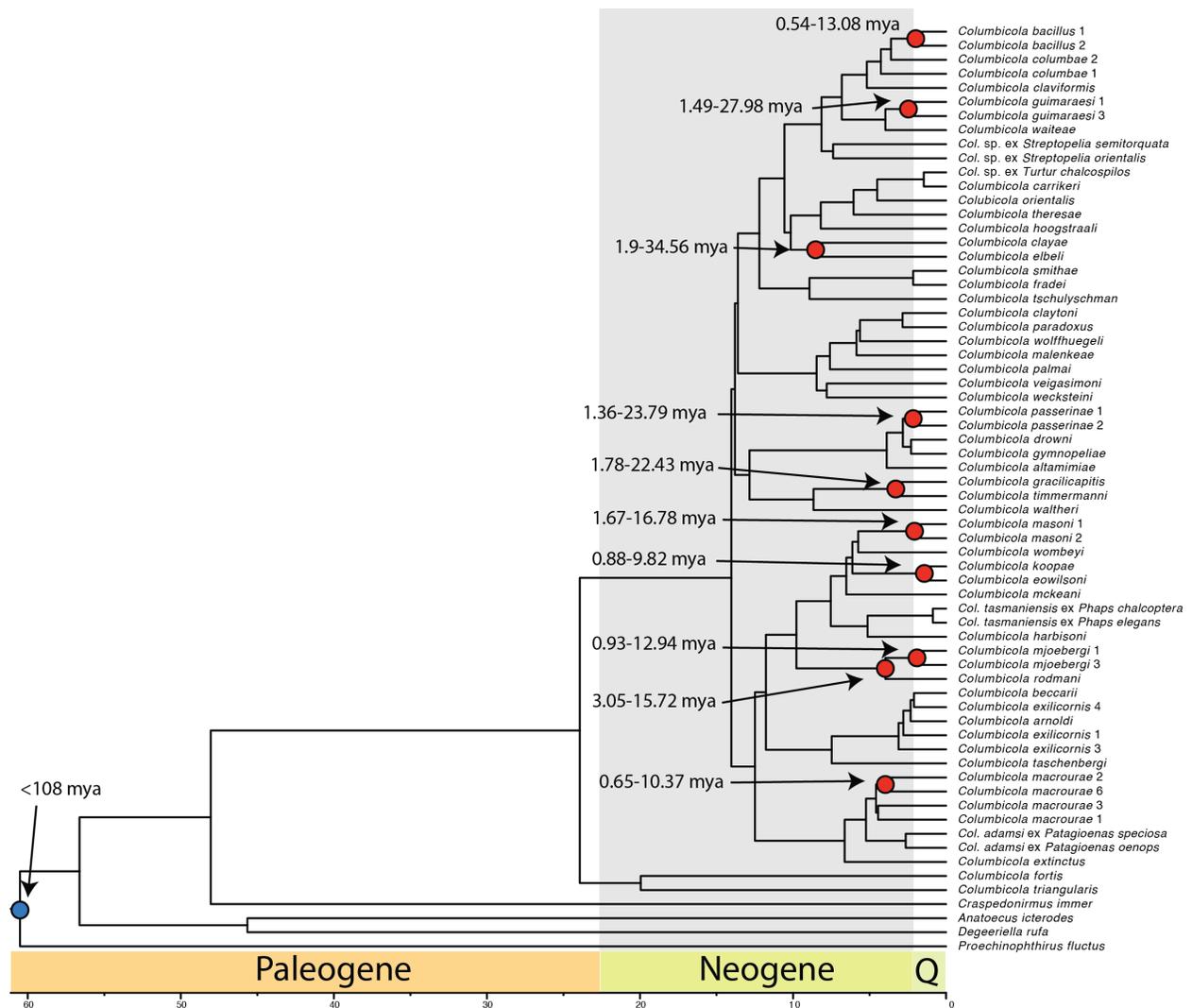
16  
 17 **Figure S4.** Time calibrated maximum-likelihood tree showing the relationships of dove species  
 18 based on the concatenation of 6,363 single-copy orthologous gene sequence alignments. Tree  
 19 calibrated with a root of 89 mya.



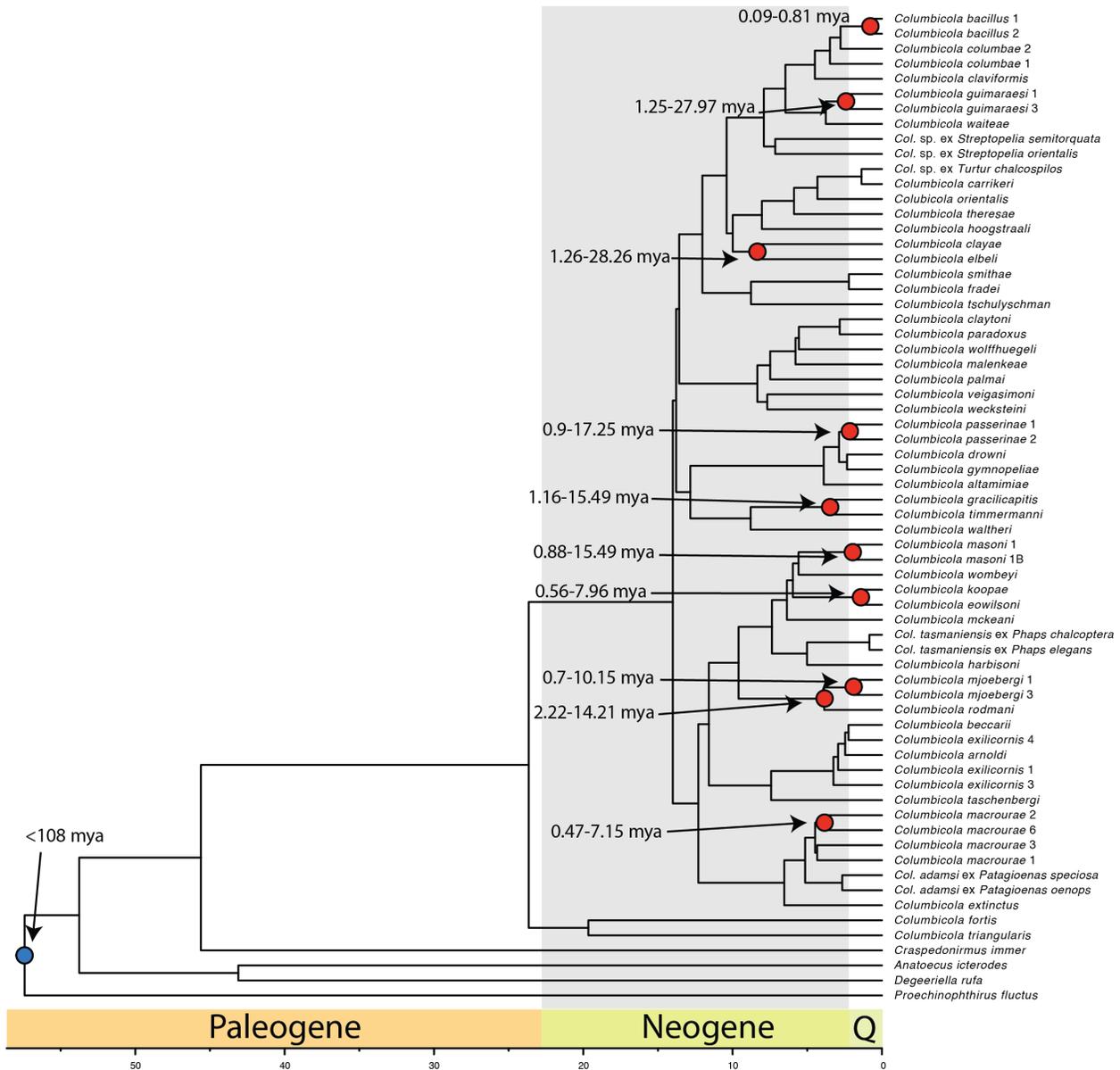
20  
 21 **Figure S5.** Time calibrated maximum-likelihood tree showing the relationships of wing louse  
 22 species based on the concatenation of 977 single-copy orthologous gene sequence alignments.  
 23 Tree calibrated with a root of 66 mya and candidate cospeciation events identified from divergence  
 24 estimates in the dove tree in figure S3.



25  
 26 **Figure S6.** Time calibrated maximum-likelihood tree showing the relationships of wing louse  
 27 species based on the concatenation of 977 single-copy orthologous gene sequence alignments.  
 28 Tree calibrated with a root of 66 mya and candidate cospeciation events identified from divergence  
 29 estimates in the dove tree in figure S4.

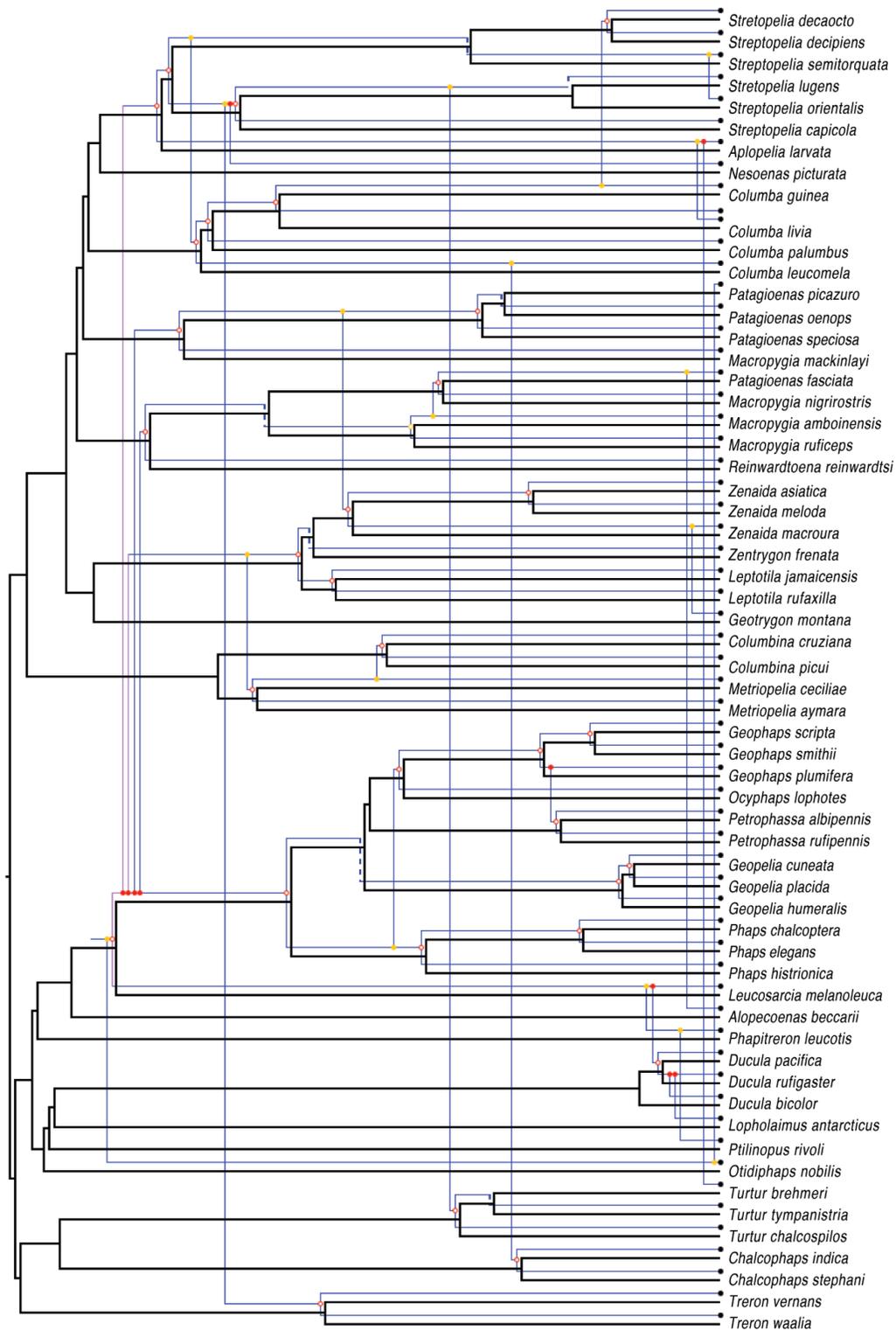


30  
 31 **Figure S7.** Time calibrated maximum-likelihood tree showing the relationships of wing louse  
 32 species based on the concatenation of 977 single-copy orthologous gene sequence alignments.  
 33 Tree calibrated with a root of 108 mya and candidate cospeciation events identified from  
 34 divergence estimates in the dove tree in figure S3.

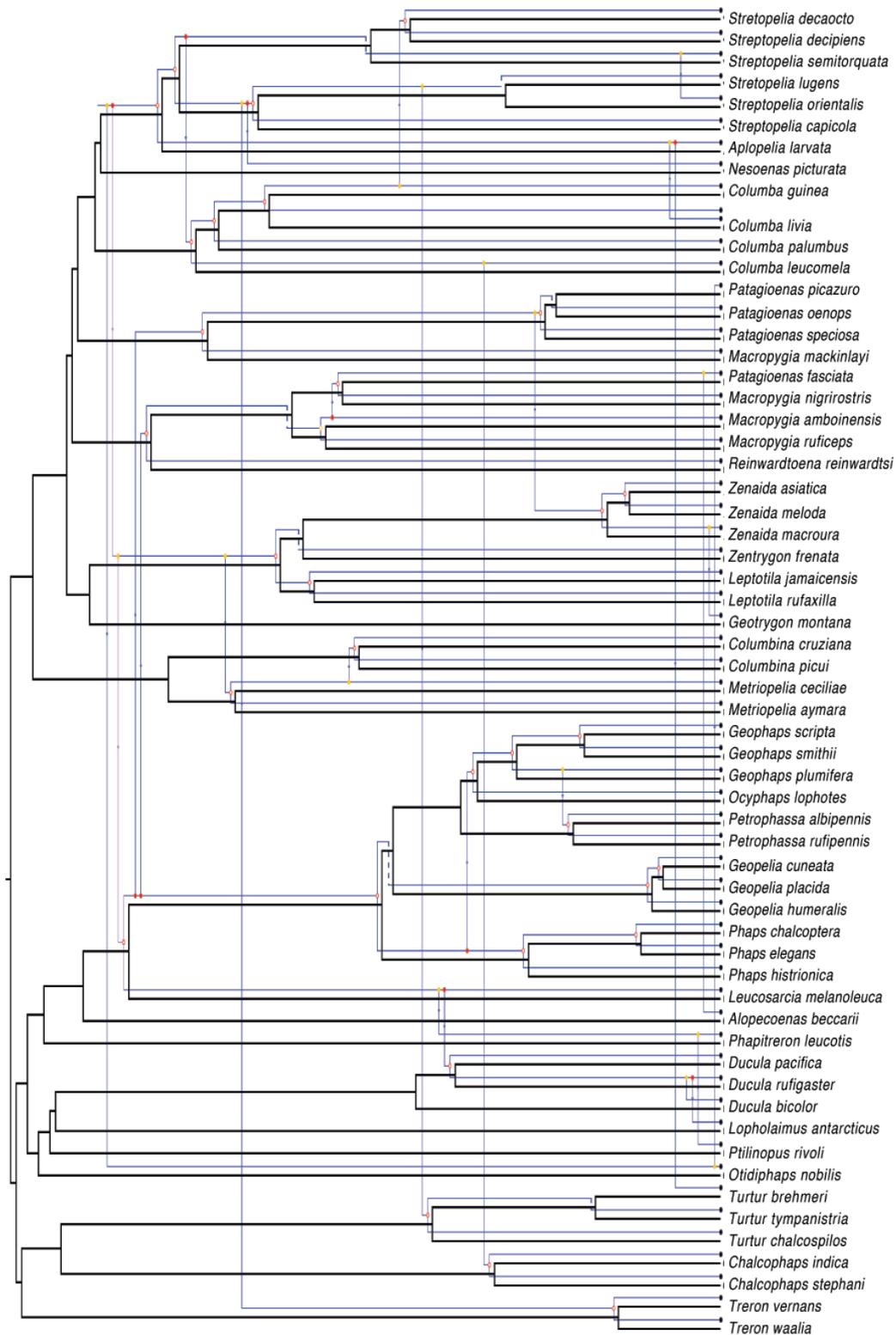


35  
 36 **Figure S8.** Time calibrated maximum-likelihood tree showing the relationships of wing louse  
 37 species based on the concatenation of 977 single-copy orthologous gene sequence alignments.  
 38 Tree calibrated with a root of 108 mya and candidate cospeciation events identified from  
 39 divergence estimates in the dove tree in figure S4.

40  
 41

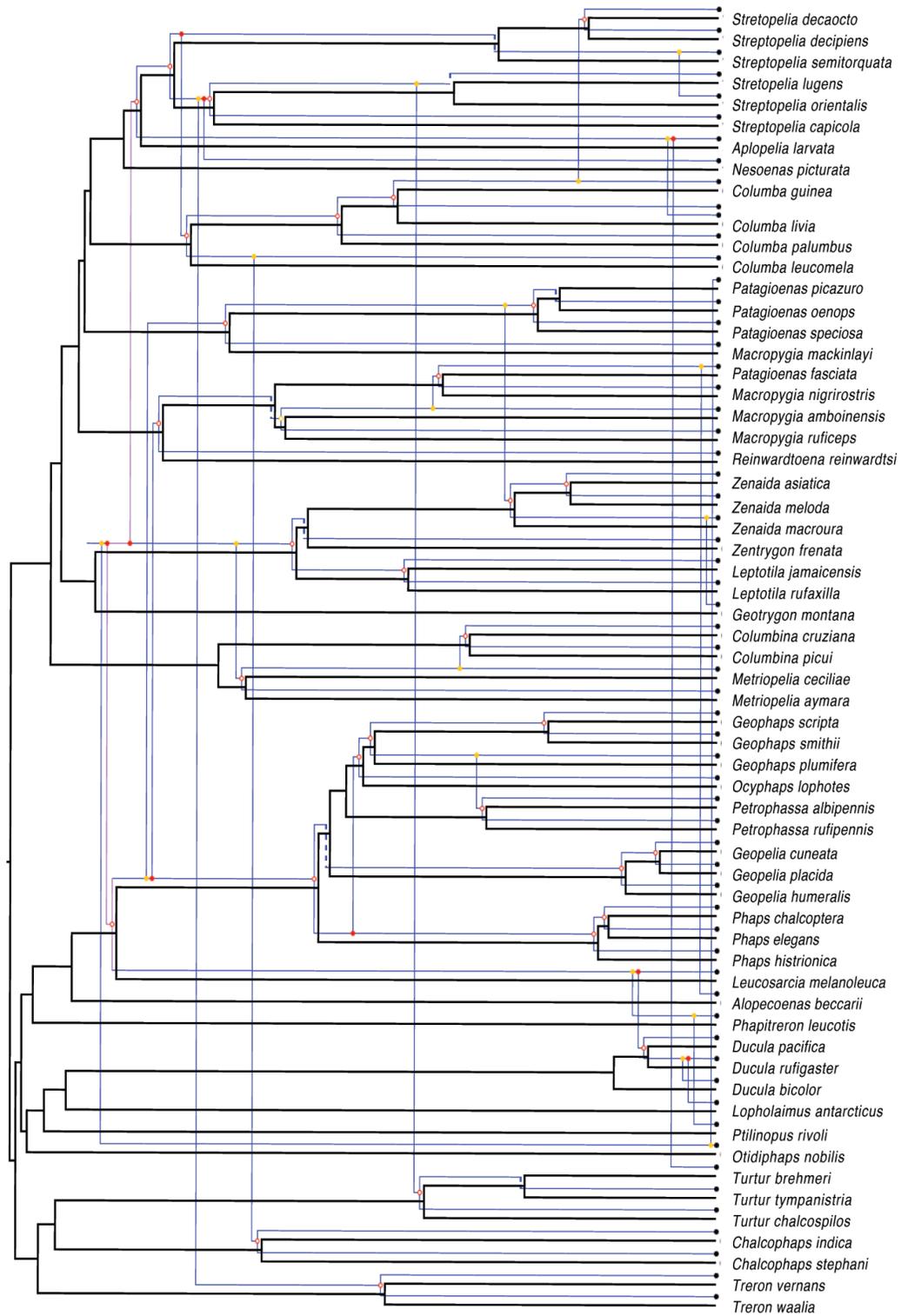


42  
 43 **Figure S9.** JANE cophylogeny reconstruction of doves and wing lice with an Australasian origin  
 44 of wing lice. Host phylogeny is black and parasite phylogeny is blue.



45  
46  
47

**Figure S10.** JANE cophylogeny reconstruction of doves and wing lice with an African origin of wing lice. Host phylogeny is black and parasite phylogeny is blue.

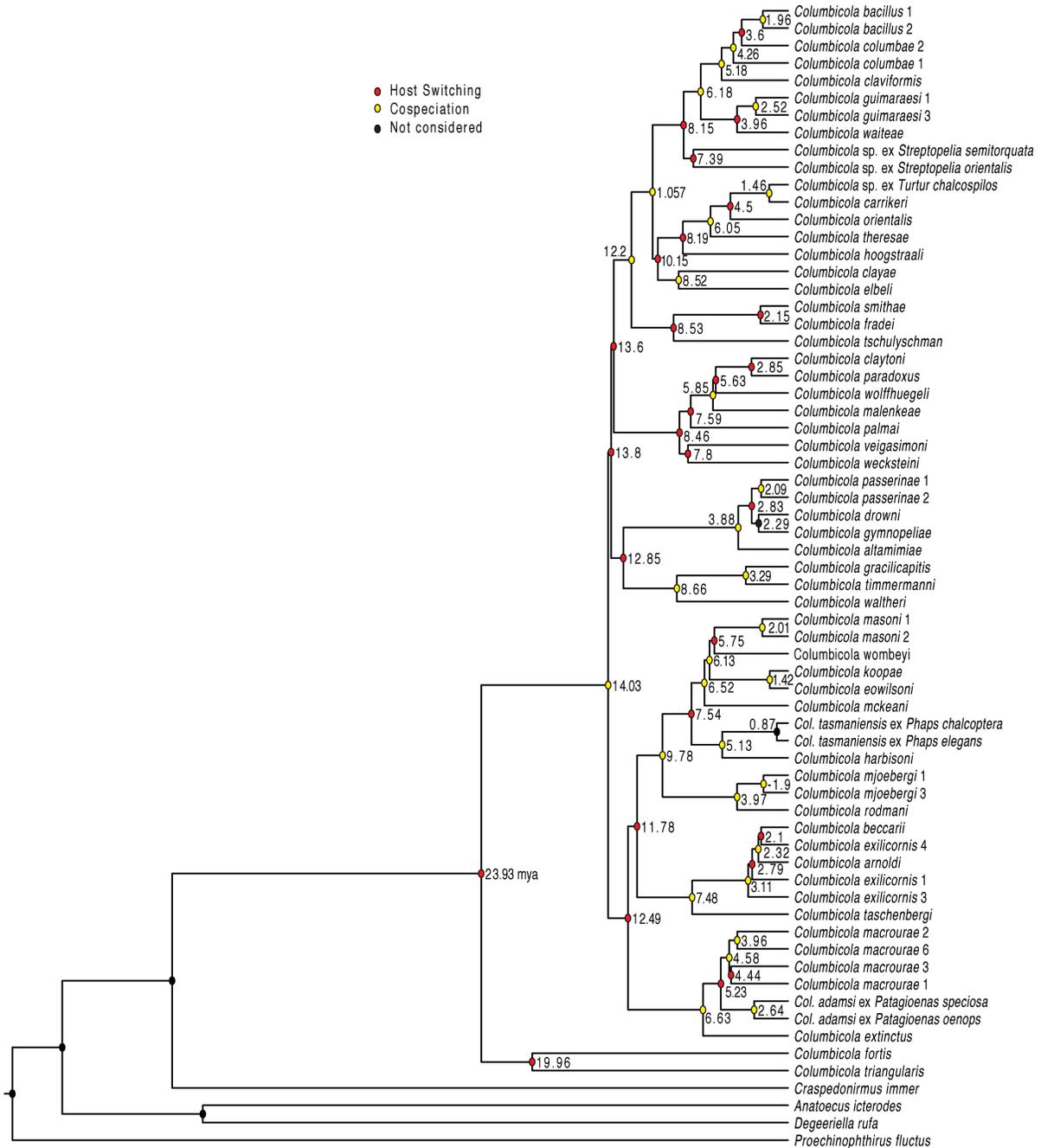


48

49 **Figure S11.** JANE cophylogeny reconstruction of doves and wing lice with a New World origin  
 50 of wing lice. Host phylogeny is black and parasite phylogeny is blue.

51

52



53  
 54 **Figure S12.** Time calibrated maximum-likelihood tree showing the relationships of wing louse  
 55 species based on the concatenation of 977 single-copy orthologous gene sequence alignments.  
 56 Tree calibrated with a root of 108 mya and candidate cospeciation events identified from  
 57 divergence estimates in the dove tree in figure S3. Node labels indicate node age in millions of  
 58 years. Node color indicates cospeciation or host-switching events, based on an Australasian origin  
 59 of wing lice described in figure S9.

60  
 61

62 **Table S1.** Calibration points used to determine species divergence times for dove and wing louse  
 63 species.

| Tree     | Constraint   | Age in mya | Data type | Source         |
|----------|--|------------|-----------|----------------|
| host     | <i>Turtur</i> and <i>Chalcophaps</i>               | 18.5       | fossil    | [6]            |
| host     | <i>Lopholaimus</i> and <i>Ducula</i>               | 16         | fossil    | [7]            |
| host     | <i>Alopecoenas</i> and <i>phabines</i>             | 20         | fossil    | [8]            |
| host     | <i>Zenaida</i> and <i>Zentrygon</i>                | 4          | fossil    | [9]            |
| host     | <i>Geotrygon</i> and NWGD                          | 9          | fossil    | DWS per. comm. |
| host     | <i>Patagioenas</i> and <i>Columba</i>              | 1.8        | fossil    | [10]           |
| host     | root   | 71.5       | inference | [11]           |
| host     | root   | 89         | inference | [12]           |
| parasite | <i>C. macrourae</i> 2 and <i>C. macrourae</i> 6*   | variable^  | inference | this study     |
| parasite | <i>C. rodmani</i> and <i>C. mjoebergi</i>          | variable^  | inference | this study     |
| parasite | <i>C. koopae</i> and <i>C. eowilsoni</i>           | variable^  | inference | this study     |
| parasite | <i>C. masoni</i> 1 and <i>C. masoni</i> 2*         | variable^  | inference | this study     |
| parasite | <i>C. guimaraesi</i> 1 and <i>C. guimaraesi</i> 3* | variable^  | inference | this study     |
| parasite | <i>C. bacillus</i> 1 and <i>C. bacillus</i> 2*     | variable^  | inference | this study     |
| parasite | <i>C. clayae</i> and <i>C. elbeli</i>              | variable^  | inference | this study     |
| parasite | <i>C. gracilicapitis</i> and <i>C. timmermanni</i> | variable^  | inference | this study     |
| parasite | <i>C. passerinae</i> 1 and <i>C. passerinae</i> 2* | variable^  | inference | this study     |
| parasite | root   | 108        | inference | [13]           |
| parasite | root   | 66         | inference | [14]           |

64  
 65 NWGD = New World Ground-doves, DWS = David Steadman, *C.* = *Columbicola*. \* designates  
 66 divergent lineages of lice that were recognized and designated by numbers following Boyd et al.  
 67 [16], ^ designates that the age of calibration changes based on alternative host tree calibrations  
 68 (figures S5-S8).

69  
 70 **Table S2.** Comparison of host and parasite trees under different coevolutionary scenarios

| Host diversification | Time zones | Possible origins | Expected cost |    | Observed cost | Cospeciation events | H-S triggered event |
|----------------------|------------|------------------|---------------|----|---------------|---------------------|---------------------|
|                      |            |                  | M             | SD |               |                     |                     |
| not used             | not used   | 1                | 107           | 2  | 61            | 35                  | 25                  |
| 51 mya               | 3          | 3                | 107           | 2  | 63            | 32                  | 28                  |
| 60 mya               | 3          | 4                | 107           | 3  | 63            | 32                  | 28                  |

71  
 72 M = mean, SD = standard deviation. H-S triggered events = parasite speciation triggered by a host  
 73 switch. The “possible origins” column denotes the sum of mutually exclusive events.

74  
 75

76 **Supplementary methods**

77

78 **Sequencing and quality control:** Kapa Library Preparation kit was used to prepare gDNA  
79 libraries. Four libraries were sequenced on one lane on the Illumina HiSeq 2500 using the HiSeq  
80 SBS Sequence Kit v2 for 2x160 cycles. The remaining libraries received 6-base barcodes and  
81 were pooled into groups of two for sequencing on the same instrument for 2x251 cycles. Resulting  
82 sequence reads were trimmed to remove bases with a phred score <30 from the 5' end using  
83 Trimmomatic PE (v0.22) [1].

84

85 **Identification of within clade paralogs:** A total of 6,923 genes were identified as single-copy  
86 orthologs using OrthoDB v8 [2]. The *Columba livia* ortholog set was indexed using Bowtie2  
87 (v2.1.0) [3] and genome sequence reads from *Columba rupestris* were mapped onto the orthologs.  
88 Average sequencing depth of each ortholog was calculated using SAMtools and BEDtools  
89 (v0.1.19) [4]. Summary statistics were calculated for sequencing depth across orthologs and a z-  
90 score of 2.5 standard deviations was established to identify outliers. Those genes with a mean  
91 sequencing depth of  $-2.5 < Z > 2.5$  were removed from the ortholog set.

92

93 **Ortholog assembly:** Sequence reads from each library were assembled using the *C. livia* orthologs  
94 as a reference sequence using Bowtie2 (v2.3.4.1) [3]. Results of read to gene alignments were  
95 converted to a binary alignment map and sorted using SAMtools (v1.7). Finally, consensus  
96 sequences were created using SAMtools and vcfutils.pl (vcf2fq) by taking the consensus base at  
97 each position, yielding ortholog sequences that could be used for phylogenetics. We conducted a  
98 final check of overall sequencing coverage for all genes in all taxa to ensure there were no high-  
99 coverage sequences, which might signify repeat elements.

100

101 **Model fitting and quality control in supermatrix generation:** To control for substantial  
102 variation in base composition in the concatenated ML tree, we estimated individual GTR+ $\Gamma$  rate  
103 parameters for all codon positions across all genes and removed any position that had a rate  
104 parameter that deviated more than 10 standard deviations from the mean. We then clustered the  
105 codon positions based upon GTR+ $\Gamma$  parameters using k-means clustering to generate partitions for  
106 the ML estimates that would group codon partitions with similar rate matrices into the same  
107 partition. After varying the number of partitions to be between 2 and 30, we found that seven  
108 partitions explained more than 75% of the total variance of the clustering.

109

110 **Dove MCMCtree parameters:** Bayesian MCMC searches were conducted for 10 million  
111 generations, sampling every 100 generations with the following parameters: 71.5 calibration (alpha  
112 = 0.19563, and two parameters for rgene\_gamma (1 12.01) and sigma2\_gamma (1 10)); 89  
113 calibration (alpha = 0.19, rgene\_gamma (1 15.49) sigma\_2gamma (1 10)) as estimated from  
114 baseml. The first 50,000 trees were excluded as burn-in for each run. Implemented in PAML  
115 (v4.9) [5].

116

117 **Louse MCMCtree parameters:** Bayesian MCMC searches were conducted for 5 million  
118 generations, sampling every 50 generations with the following parameters: 108 calibration (alpha  
119 = 0.275, and two parameters for rgene\_gamma (680 1000 1.0 0) and sigma2\_gamma (0.77 100  
120 1.0)); 66 calibration (alpha = 0.275, rgene\_gamma (1115.7 1000 1.0) sigma\_2gamma (0.01 100  
121 1.0)) as estimated from baseml. The first 50,000 trees were excluded as burn-in for each run.

122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167

## Supplementary data

The following is included in the figshare database 10.6084/m9.figshare.12730337: raw coding sequences, gene trees, ASTRAL tree, concatenation alignment, concatenation partition, concatenation RAXML-bipartition tree, base frequency by codon image, MCMC trees, Jane input files, sample collection data, NCBI-SRA identifiers, and list of sample collections identifiers.

## References

1. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinf.* **30**, 2114-2120.
2. Kriventseva EV, Tegenfeldt F, Petty TJ, Waterhouse RM, Simao FA, Pozdnyakov IA, Ioannidis P, Zdobnov EM, 2015. OrthoDB v8: update of the hierarchical catalog of orthologs and underlying free software. *Nucl. Acids Res.* **43**, D250-D256.
3. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods.* **9**, 357-359.
4. Li H, Handsaker B, Wysoker A, Fennel T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinf.* **25**, 2078-2079.
5. Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586-1591.
6. Steadman DW. 2008. Doves (Columbidae) and cuckoos (Cuculidae) from the early Miocene of Florida. *Bull. Flor. Mus. Natl. Hist.* **48**, 1-16.
7. De Pietri VL, Scofield RP, Tennyson AJD, Hand SJ, Worthy TH. 2017. The diversity of Early Miocene pigeons (Columbidae) in New Zealand. *Cont. del Museo Argentino de Ciencias Natural.* **7**, 49-68.
8. Worthy, TH. 2012. A phabine pigeon (Aves: Columbidae) from Oligo-Miocene of Australia. *Emu.* **112**, 23-31.
9. Brodkorb P. 1969. An ancestral mourning dove from Rexroad, Kansas. *Quart. J. Florida Acad. Sci.* **31**, 173-176.
10. Emslie SD. 1998. Avian community, climate, and sea-level changes in the Plio-Pliocene of the Florida peninsula. *Ornith. Monograph.* **50**, 1-113.
11. Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon M, Lemmon AR. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature.* **526**, 569-573.
12. Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C, Ho SYW, Faircloth BC, Nabholz , Howard JT, et al. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science.* **346**, 1320-1331.
13. Smith VS, Ford T, Johnson KP, Johnson PCD, Yoshizawa K, Light JE. 2011. Multiple lineages of lice pass through the K-Pg boundary. *Biol. Lett.* **7**, 782-785.
14. Johnson KP, Nguyen NP, Sweet AD, Boyd BM, Warnow T, Allen JM. 2018. Simultaneous radiation of bird and mammal lice following the K-Pg boundary. *Biol. Lett.* **14**, 20180141.