

Supplemental Materials

S.1 Distribution of indels of different lengths among introns (Table S1, Figure S1)

Since long indels are on average more parsimony consistent than short indels, it might therefore seem desirable to use only long indels in phylogenetic analyses. However, we show here that there are too few long indels to produce a sufficient number of well resolved gene trees needed for coalescent analysis. We grouped indels in nested length categories, and report their numbers per intron as well as the number of introns represented by each category (Table S1). Although there are a total of 141,375 intron indels >50bp in length in our dataset, they are spread unequally among 2,452 loci, only 24% of indels are parsimony informative, and they are distributed among lineages in proportion to branch length. Ergo, there are very few, if any, longer indels in the deep short branches of most gene trees.

We calculated Bootstrap Support (BS) on Maximum Likelihood (ML) trees using concatenated data in each of the nested length categories (Figure S1). ML trees were calculated using the BINGAMMA model in RAxML [1] starting from 1 (all indels and >1bp indel length categories) to 10 randomized MP trees (all other length categories), depending on computational limitations in relation to dataset size. The number of bootstrap replications was determined by the extended majority-rule boot-stopping criterion in RAxML (50–800 replications for indel classes >10bp), or truncated manually for the larger datasets of more inclusive indel length categories. Not surprisingly, mean nodal BS support is lower in trees generated using longer indel categories due to the relative dearth of long indels.

S.2 Maximum Parsimony (MP) Analyses (Figures S2–S5)

MP analyses of the intron indel, UCE, exon, and combined indel+UCE+exon datasets of Jarvis et al [2,3], not previously reported, are presented for comparative purposes (Figures S2–S5). Bootstrap MP trees of concatenated indels were calculated using TNT [4] with 1000 replications of tree bisection-reconnection branch-swapping and 10 trees saved per replication (equally most parsimonious trees were treated as strict consensus) with 10,000 bootstrap replications. We considered MP to be a reasonable analytical strategy given the relatively large state space of indels. When the state space for characters is sufficiently large the MP tree is the ML tree [5,6]; see Yuri et al. [7] and Braun et al. [8] for additional discussion of the utility of MP when analyzing low-homoplasy characters.

S.3 Maximum Likelihood (ML) Analyses (Figures S6, S7)

ML analyses of the intron indel and whole genome indel datasets of Jarvis et al [2,3], not previously reported, are presented for comparative purposes (Figures S6, S7). Best ML tree of concatenated intron indels was calculated using the BINGAMMA model in RAxML [1], from 10 randomized starting MP trees. Bootstrap values were calculated on one starting tree with 100 replicates.

S.4 Intron indel and nt gene tree ensemble RI as a function of number of characters per locus. (Figure S8)

Different loci were chosen for the indel and nt data to range from 1,000 to 3,500 characters, whether indels or nts. This character size interval was selected to capture the greatest amount of dataset size overlap without getting into the tail end of either the indel or nt distributions. The indels, of course, were scored from loci whose nt sequence length is much larger than the loci plotted for nts. Gene trees of both the indel and nt datasets were generated from the full locus. Intron indels exhibit higher RI than intron nts, with no character dataset size effect. Note that Figure 3 and Table 2 report RI, a measure of the behavior of individual characters, whereas ensemble RI is a collective measure of those characters, here per gene tree.

S.5 Robinson-Foulds (RF) distances (Figure S9)

RF distances were calculated between gene tree pairs of 1) indel and nt gene trees of the same locus, 2) indel gene trees of randomly chosen loci, and 3) nt gene trees of randomly chosen loci, for both intron and UCE loci. Frequency distributions of each the six type of RF comparisons are plotted as density of normalized RF distances. Randomly sampled nucleotide gene trees are more congruent with one another than indel and nucleotide gene trees of the same locus. Indel and nucleotide gene trees of the same locus are substantially more congruent with one another than randomly sampled indel gene trees.

S.6 Quartet Frequency Analyses (Figures S10–S13)

Quartet frequencies were calculated using ASTRAL III [10] on ML BS gene trees with 5% contraction, and presented as relative frequencies using DiscoVista [11].

S.7 ASTRAL polytomy tests (Figures S14–S22)

Polytomy tests were performed on each data partition separately and in various combinations using ASTRAL-III [11] as described in section 3.7 of main text.

Data and code used in this paper can be found at <http://doi.org/10.5281/zenodo.3237219>.

References

1. Stamatakis, A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **2006**, *22*, 2688–2690, doi:10.1093/bioinformatics/btl446.
2. Jarvis, E.D.; Mirarab, S.; Aberer, A.J.; Li, B.; Houde, P.; Li, C.; Ho, S.Y.W.; Faircloth, B.C.; Nabholz, B.; Howard, J.T.; et al. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* **2014**, *346*, 1320–1331, doi:10.1126/science.1253451.
3. Jarvis, E.D.; Mirarab, S.; Aberer, A.J.; Li, B.; Houde, P.; Li, C.; Ho, S.Y.W.; Faircloth, B.C.; Nabholz, B.; Howard, J.T.; et al. Avian Phylogenomics Consortium Phylogenomic analyses data of the avian phylogenomics project. *Gigascience* **2015**, *4*, 4, doi:10.1186/s13742-014-0038-1.
4. Goloboff, P.A.; Farris, J.S.; Nixon, K.C. TNT, a free program for phylogenetic analysis. *Cladistics* **2008**, *24*, 774–786, doi:10.1111/j.1096-0031.2008.00217.x.
5. Steel, M.; Penny, D. Two further links between MP and ML under the poisson model. *Appl. Math. Lett.* **2004**, *17*, 785–790, doi:10.1016/j.aml.2004.06.006.
6. Steel, M.; Penny, D. Maximum parsimony and the phylogenetic information in multistate characters. In *Parsimony, Phylogeny, and Genomics*; Albert, V. A., Ed.; Oxford University Press: Oxford, UK, 2005; pp. 163–178.
7. Yuri, T.; Kimball, R.T.; Harshman, J.; Bowie, R.C.K.; Braun, M.J.; Chojnowski, J.L.; Han, K.-L.; Hackett, S.J.; Huddleston, C.J.; Moore, W.S.; et al. Parsimony and model-based analyses of indels in avian nuclear genes reveal congruent and incongruent phylogenetic signals. *Biology* **2013**, *2*, 419–444, doi:10.3390/biology2010419.
8. Braun, E.L.; Cracraft, J.; Houde, P. Resolving the avian tree of life from top to bottom: The promise and potential boundaries of the phylogenomic era. In: *Kraus RHS (ed) Avian genomics in ecology and evolution – from the lab into the wild*. Springer, Cham **2019**.
9. Mirarab, S.; Bayzid, M.S.; Boussau, B.; Warnow, T. Statistical binning enables an accurate coalescent-based estimation of the avian tree. *Science* **2014**, *346*, 1250463, doi:10.1126/science.1250463.
10. Sayyari, E.; Mirarab, S. Testing for polytomies in phylogenetic species trees using quartet frequencies. *Genes* **2018**, *9*, doi:10.3390/genes9030132.
11. Mirarab, S.; Reaz, R.; Bayzid, M.S.; Zimmermann, T.; Swenson, M. S.; Warnow, T. ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics* **2014**, *30*, i541–i548, doi:10.1093/bioinformatics/btu462.

Table S1. Distribution of indels per intron per size class.

	All indels	>1bp	>10bp	>50bp	>100bp
Number of Indels					
Minimum	6	1	0	0	0
1 st Quartile	431.5	322.5	110.5	13	6
Median	1014	755	256	35	17
3 rd Quartile	2065	1541.5	521	75	39
Maximum	13612	9917	3525	648	409
Number of Introns					
Represented	2515	2515	2509	2452	2324
Absent	1	1	7	64	192

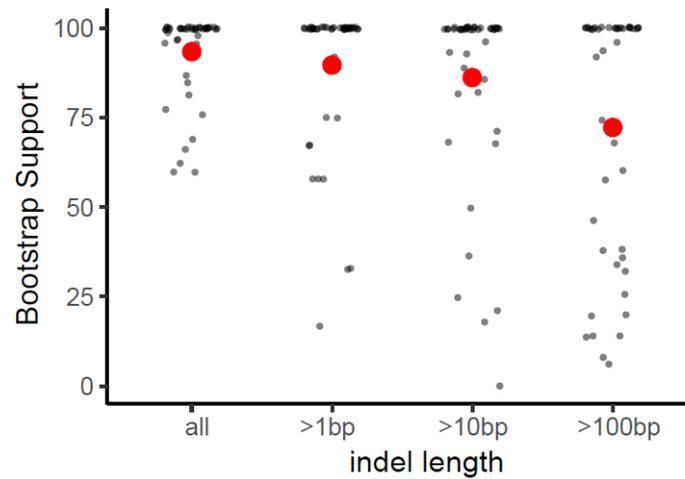


Figure S1. Bootstrap Support for Maximum Likelihood Intron Indel trees using indels of different size classes. Black dots are nodal BS values, large red circles are means.

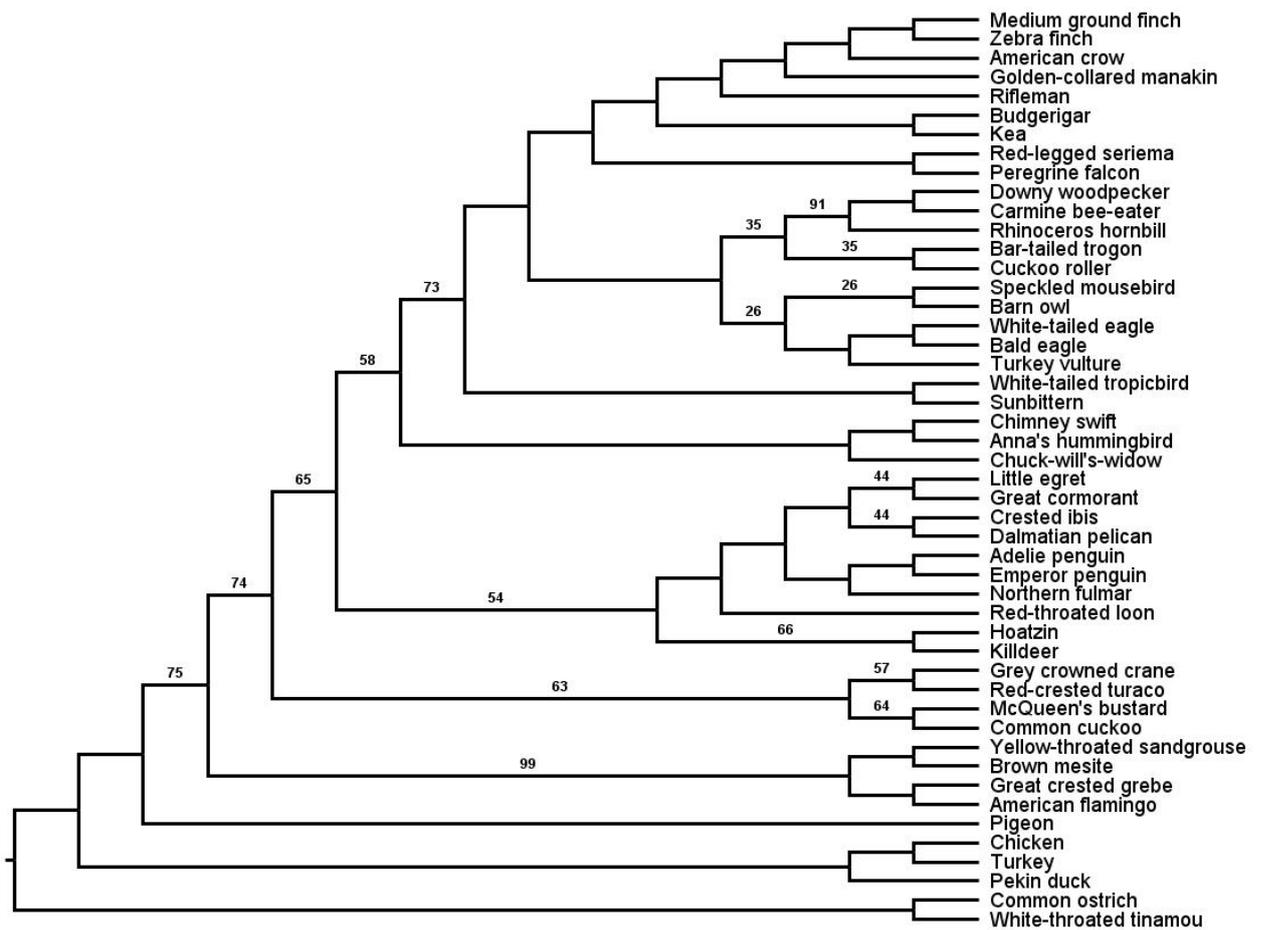


Figure S2. Bootstrap Maximum Parsimony tree of concatenated intron indels using TNT [4] with 1000 replications of tree bisection-reconnection branch-swapping and 10 trees saved per replication (equally most parsimonious trees were treated as strict consensus) and 10,000 bootstrap replications. Bootstrap scores are shown only for branches with <100% support.

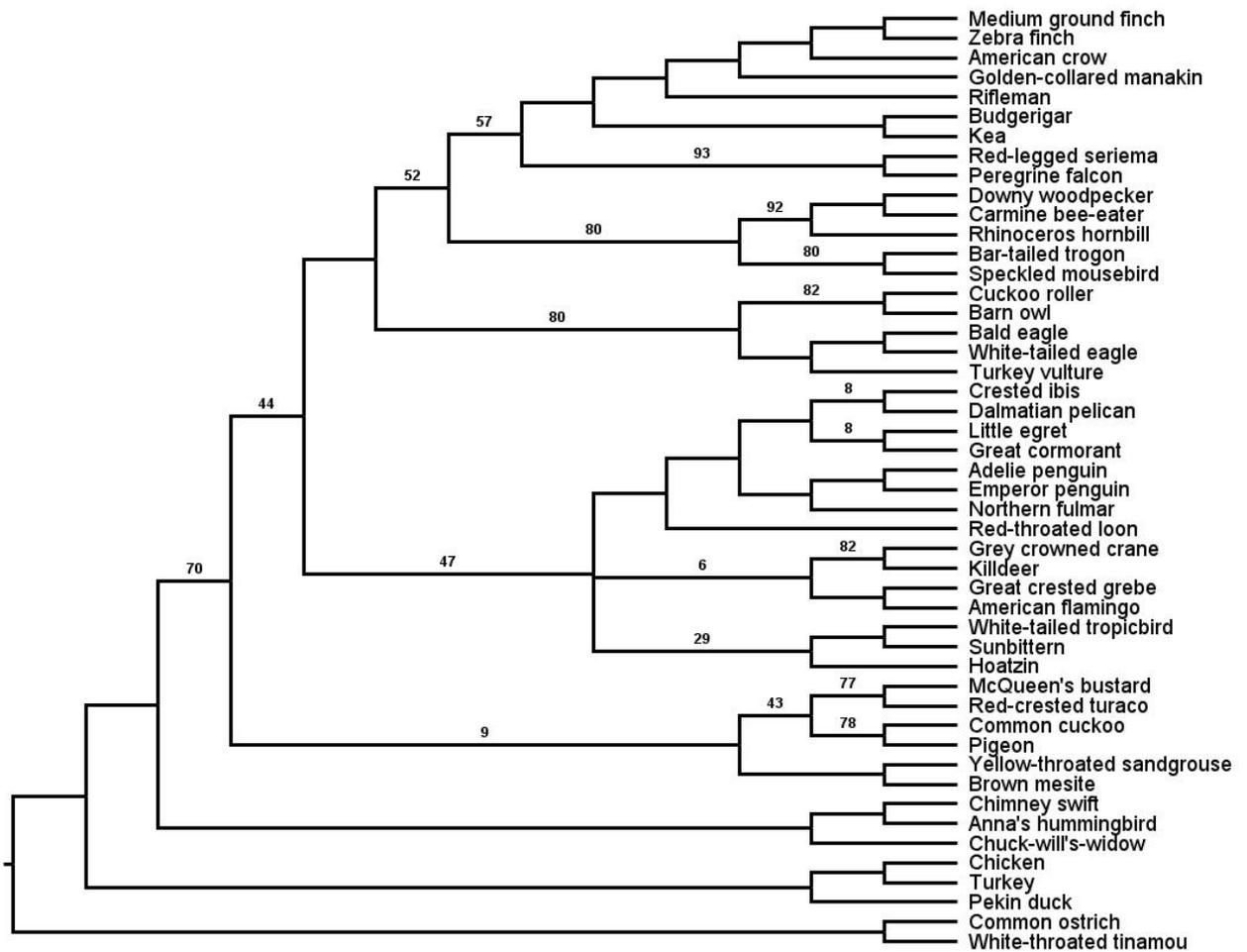


Figure S3. Bootstrap Maximum Parsimony tree of concatenated UCE indels. Tree estimation was performed as in Figure S2. Bootstrap scores are shown only for branches with < 100% support.

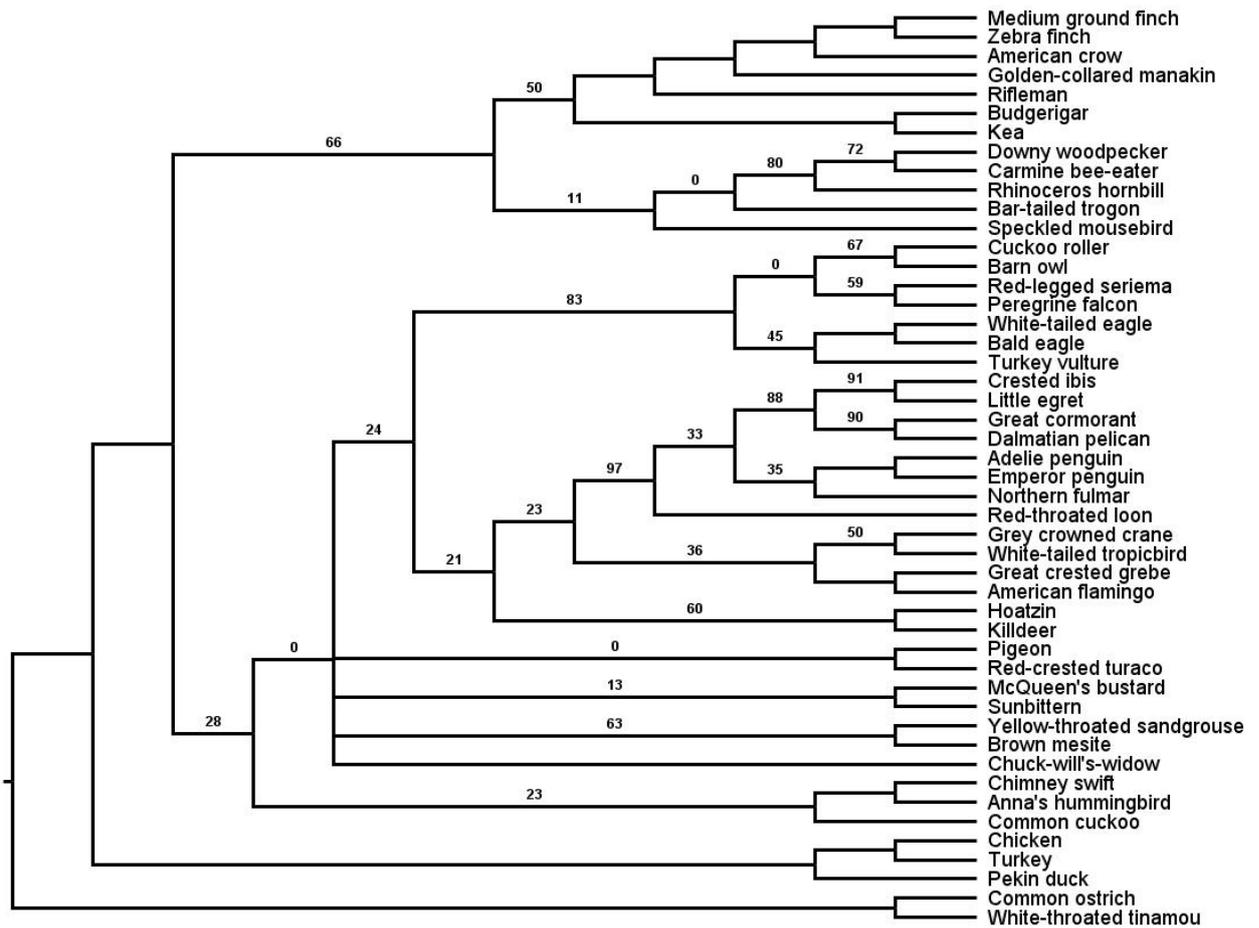


Figure S4. Bootstrap Maximum Parsimony tree of concatenated exon indels. Indels >29bp in length were omitted to avoid scoring missing exons as indels since indels were scored from exons that were concatenated by locus. Tree estimation was otherwise performed as in Figure S2. Bootstrap scores are shown only for branches with <100% support.

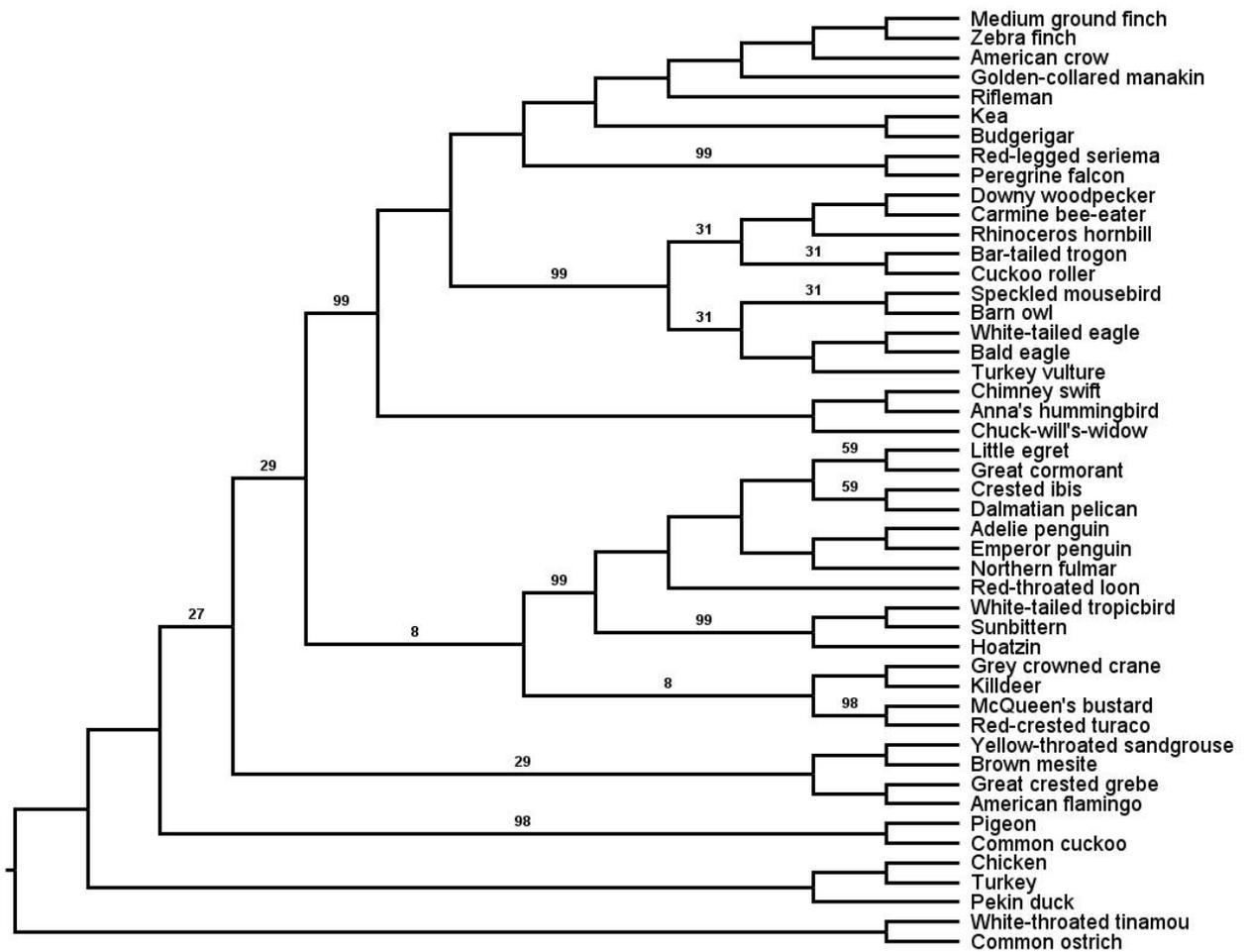


Figure S5. Bootstrap Maximum Parsimony tree of concatenated intron+UCE+exon indels. Tree estimation was performed as in Figures S2–S4. Bootstrap scores are shown only for branches with <100% support.

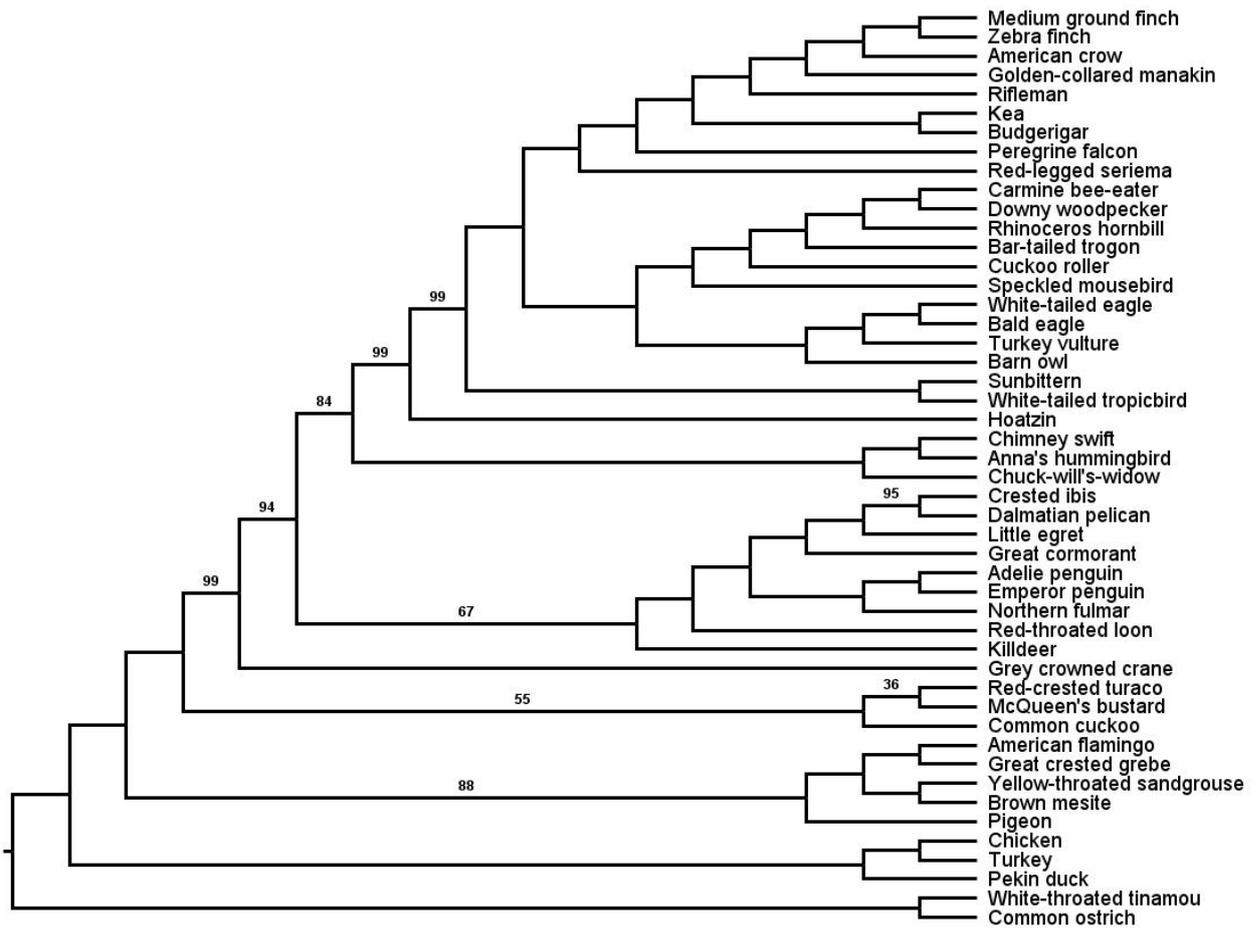


Figure S6. Best Maximum Likelihood tree of concatenated intron indels using the BINGAMMA model in RAxML [1], from 10 randomized starting MP trees (-10 option). Bootstrap values calculated on one starting tree with 100 replicates. Bootstrap scores are shown only for branches with <100% support.

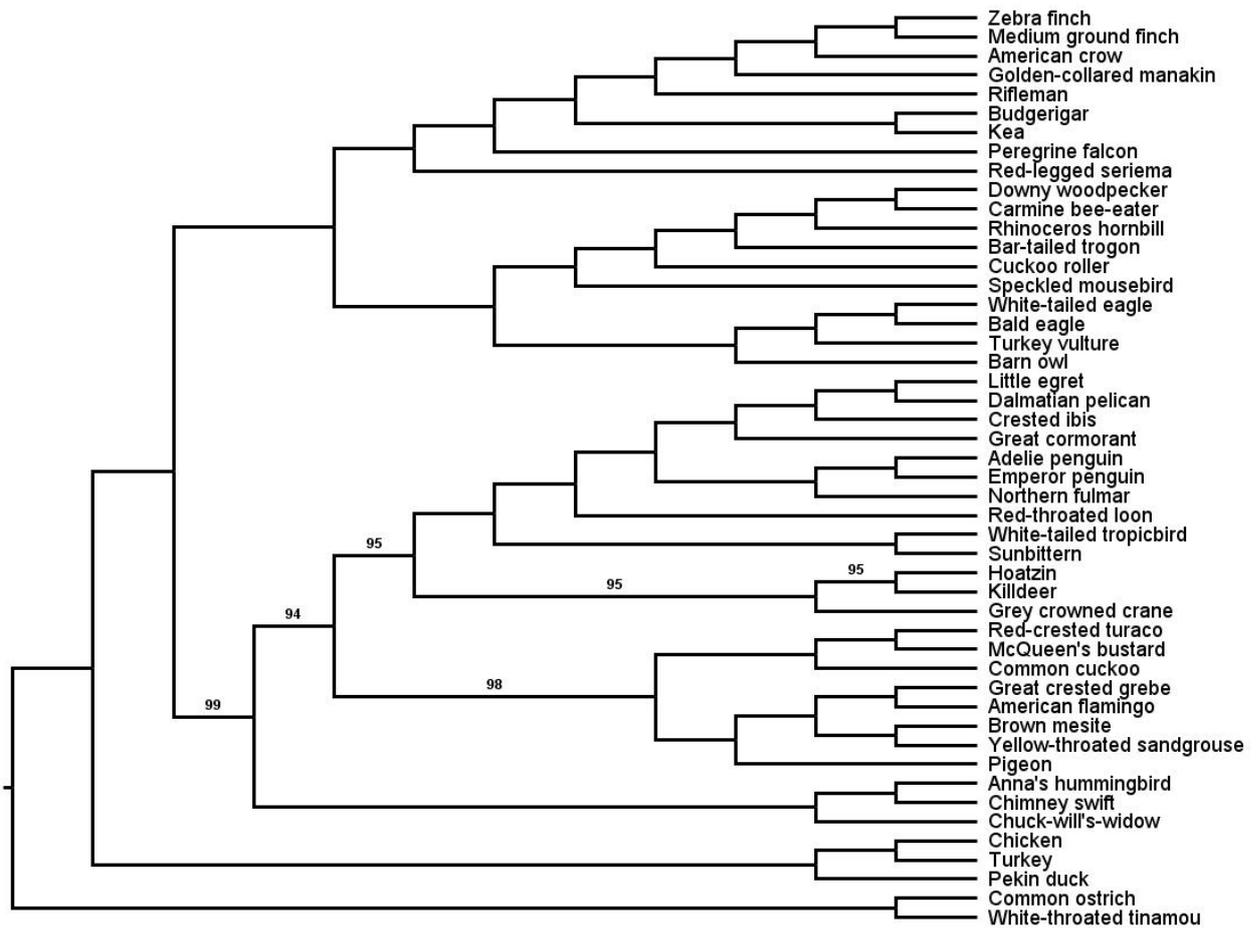


Figure S7. Best Maximum Likelihood tree of whole genome indels using the BINGAMMA model in RAxML [1]. Bootstrap values calculated on one starting tree with 100 replicates. Bootstrap scores are shown only for branches with <100% support.

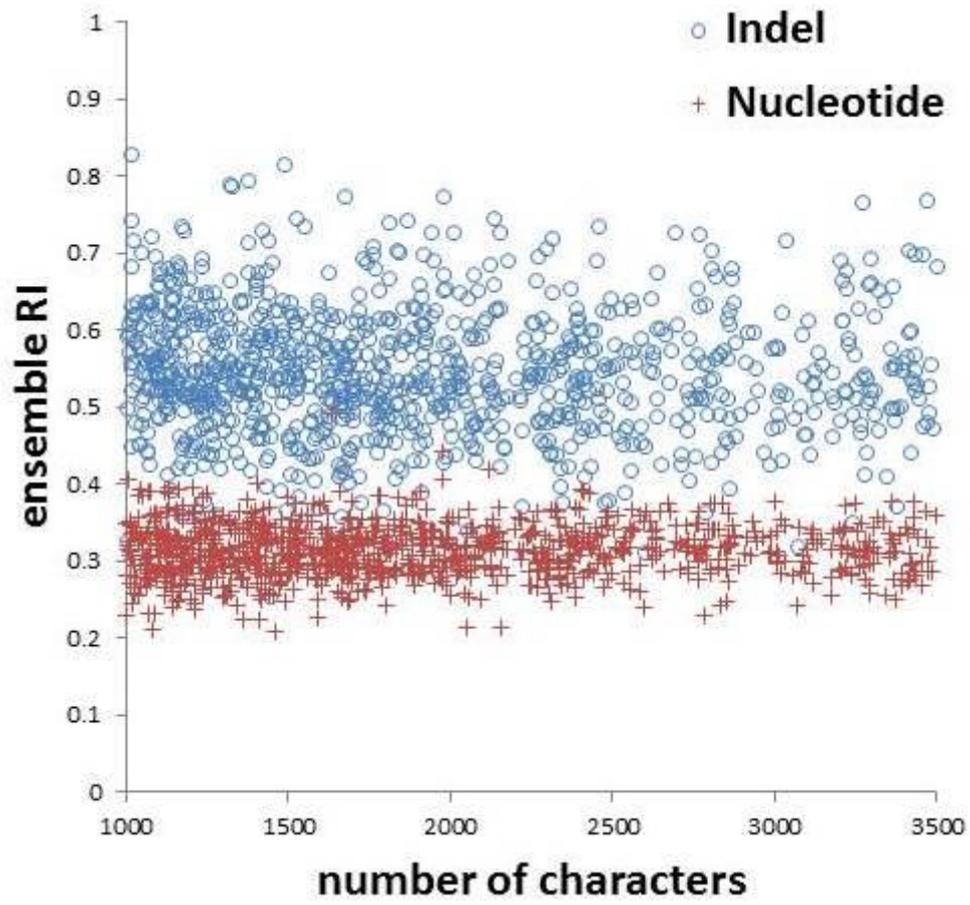


Figure S8. Intron indel and nt gene tree ensemble RI as a function of number of characters per locus.

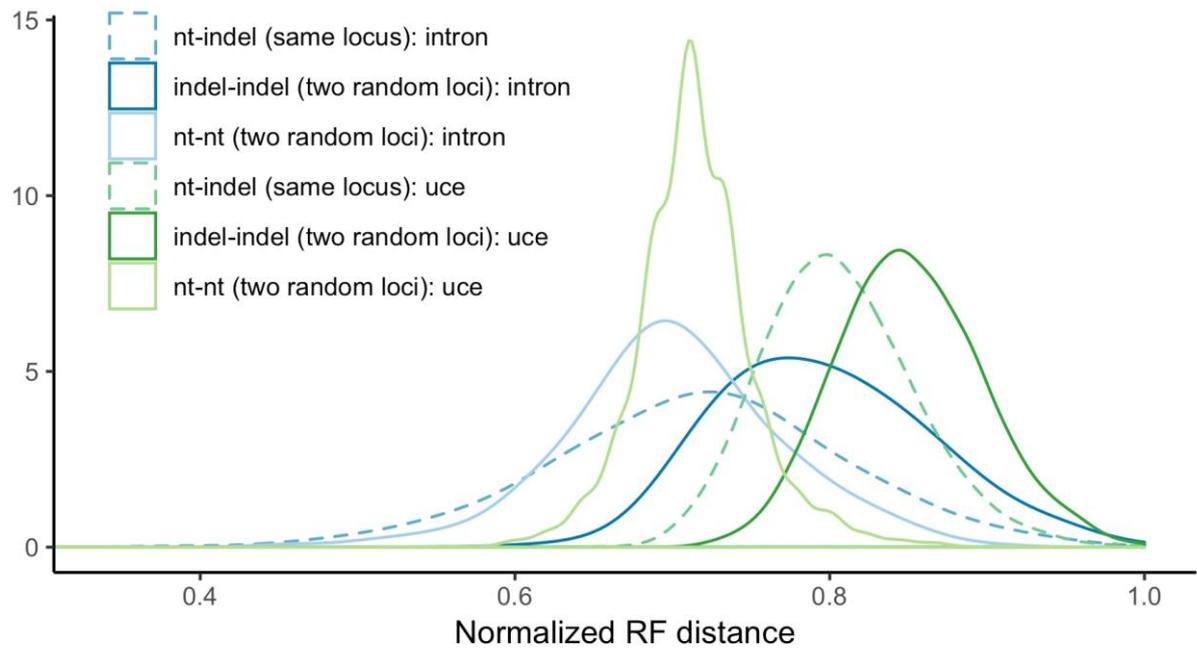


Figure S9. Gene tree Robinson-Foulds (RF) distances for intron and uce gene trees. Density of normalized RF distances between nucleotide and indel gene trees of the same locus (dashed lines) are compared to distances between randomly sampled indel gene trees and randomly sampled nucleotide gene trees (solid lines). Randomly sampled nucleotide gene trees are more congruent with one another than indel and nucleotide gene trees of the same locus. Indel and nucleotide gene trees of the same locus are substantially more congruent with one another than randomly sampled indel gene trees.

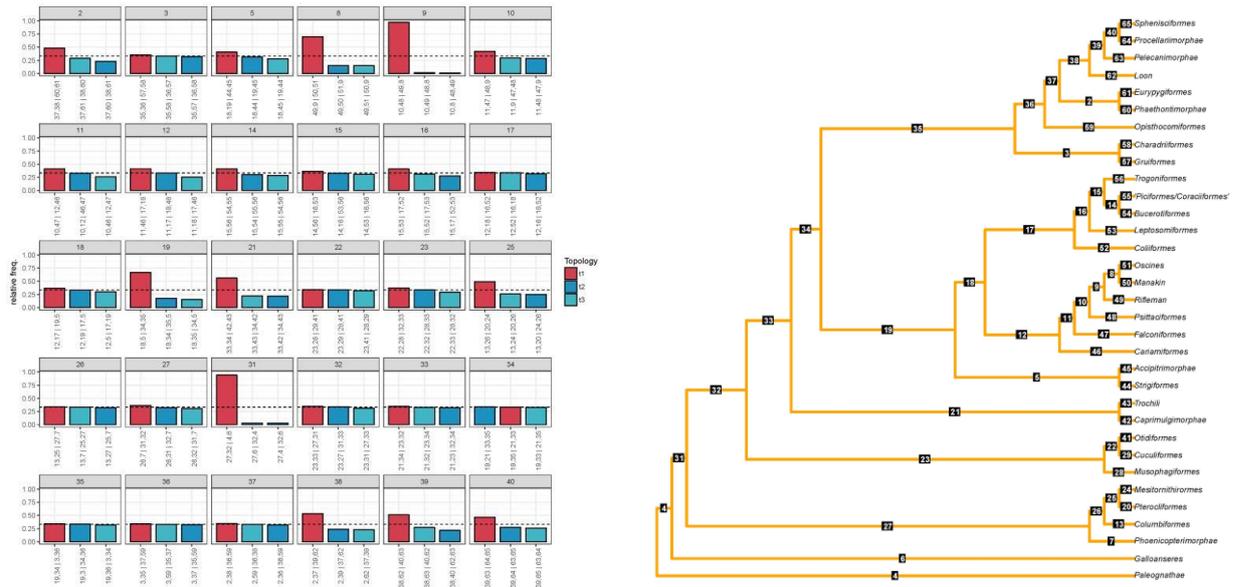


Figure S10. Quartet Scores of intron indel ASTRAL tree. Right panel: Collapsed intron indel 5% ASTRAL tree. Left panel: For all labeled internal branches, we show the quartet frequency of the main resolution (the red bar) and the two alternative resolutions around that branch (blue bars). Labels on the x-axis indicate the quartet topology. Quartet frequencies computed from nt gene trees are shown in lightertones and quartet frequencies based on indel gene trees are shown in darker tones. The horizontal dashed lines show the $\frac{1}{3}$ threshold, which corresponds to either a true hard polytomy or random resolutions in gene trees. Nodes with high support (top) and low support (bottom) are separated to allow two different scales. Note that many of the basal branches have quartet frequencies that are very close to $\frac{1}{3}$. Note that differences between indel and nt tree topologies render differences in node presence and outlay between Figures S10–S13.

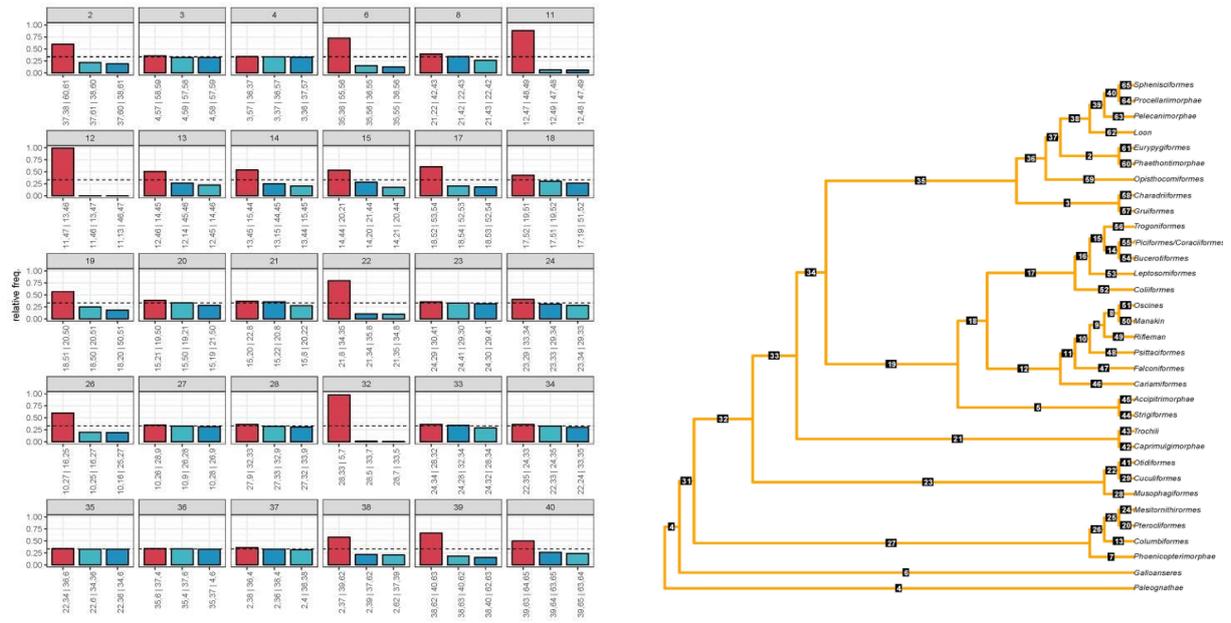


Figure S11. Quartet Scores of intron nt ASTRAL tree. Right panel: Collapsed intron nt 5% ASTRAL tree. Left panel: For all labeled internal branches, we show the quartet frequency of the main resolution (the red bar) and the two alternative resolutions around that branch (blue bars). Labels on the x-axis indicate the quartet topology. Quartet frequencies computed from nt gene trees are shown in lightertones and quartet frequencies based on indel gene trees are shown in darker tones. The horizontal dashed lines show the $\frac{1}{3}$ threshold, which corresponds to either a true hard polytomy or random resolutions in gene trees. Nodes with high support (top) and low support (bottom) are separated to allow two different scales. Note that many of the basal branches have quartet frequencies that are very close to $\frac{1}{3}$. Note that differences between indel and nt tree topologies render differences in node presence and outlay between Figures S10–S13.

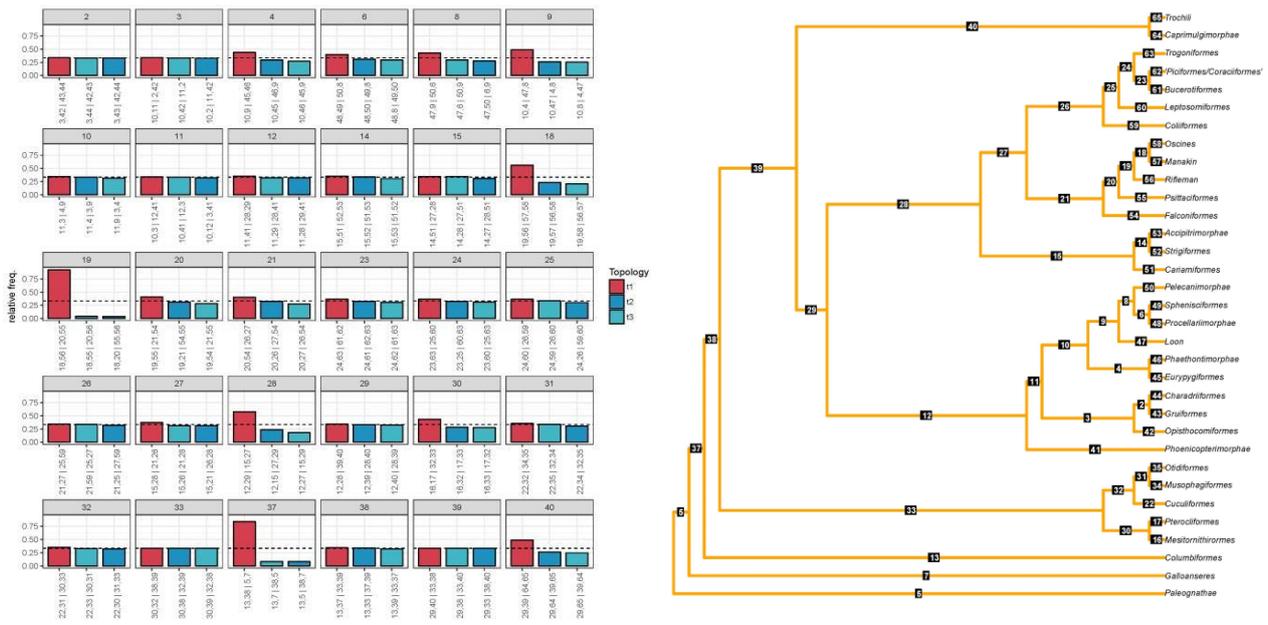


Figure S12. Quartet Scores of UCE indel ASTRAL tree. Right panel: Collapsed UCE indel 5% ASTRAL tree. Left panel: For all labeled internal branches, we show the quartet frequency of the main resolution (the red bar) and the two alternative resolutions around that branch (blue bars). Labels on the x-axis indicate the quartet topology. Quartet frequencies computed from nt gene trees are shown in lightertones and quartet frequencies based on indel gene trees are shown in darker tones. The horizontal dashed lines show the $\frac{1}{3}$ threshold, which corresponds to either a true hard polytomy or random resolutions in gene trees. Nodes with high support (top) and low support (bottom) are separated to allow two different scales. Note that many of the basal branches have quartet frequencies that are very close to $\frac{1}{3}$. Note that differences between indel and nt tree topologies render differences in node presence and outlay between Figures S10–S13.

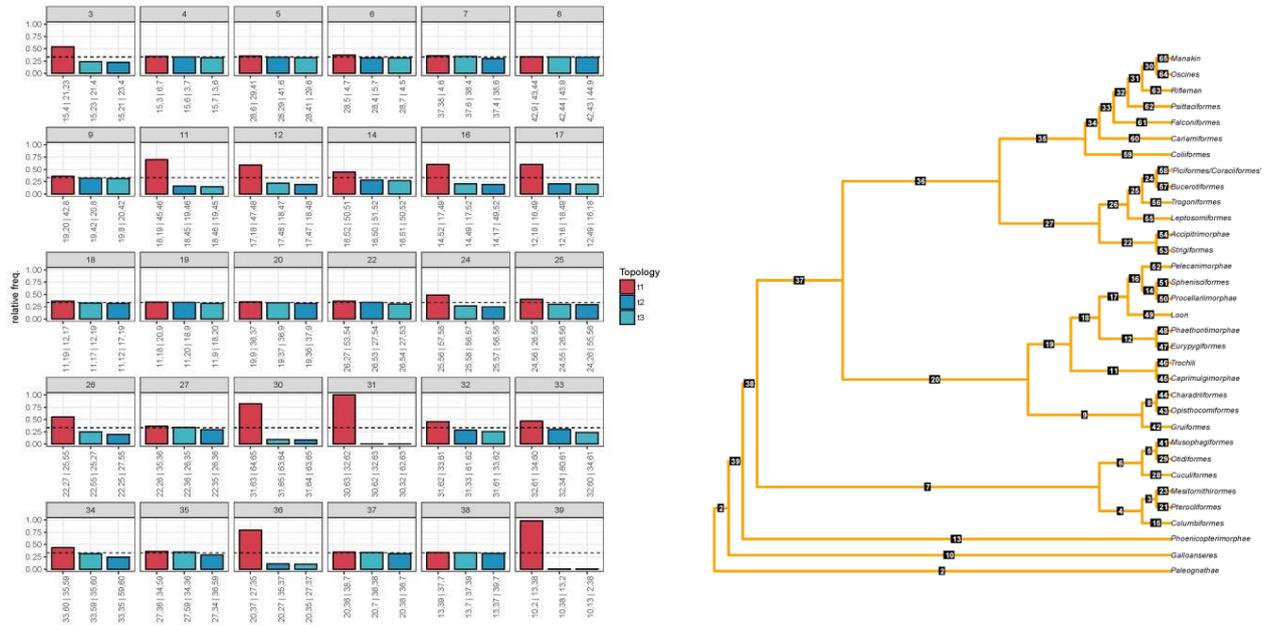


Figure S13. Quartet Scores of UCE nt ASTRAL tree. Right panel: Collapsed UCE nt 5% ASTRAL tree. Left panel: For all labeled internal branches, we show the quartet frequency of the main resolution (the red bar) and the two alternative resolutions around that branch (blue bars). Labels on the x-axis indicate the quartet topology. Quartet frequencies computed from nt gene trees are shown in lightertones and quartet frequencies based on indel gene trees are shown in darker tones. The horizontal dashed lines show the $\frac{1}{3}$ threshold, which corresponds to either a true hard polytomy or random resolutions in gene trees. Nodes with high support (top) and low support (bottom) are separated to allow two different scales. Note that many of the basal branches have quartet frequencies that are very close to $\frac{1}{3}$. Note that differences between indel and nt tree topologies render differences in node presence and outlay between Figures S10–S13.

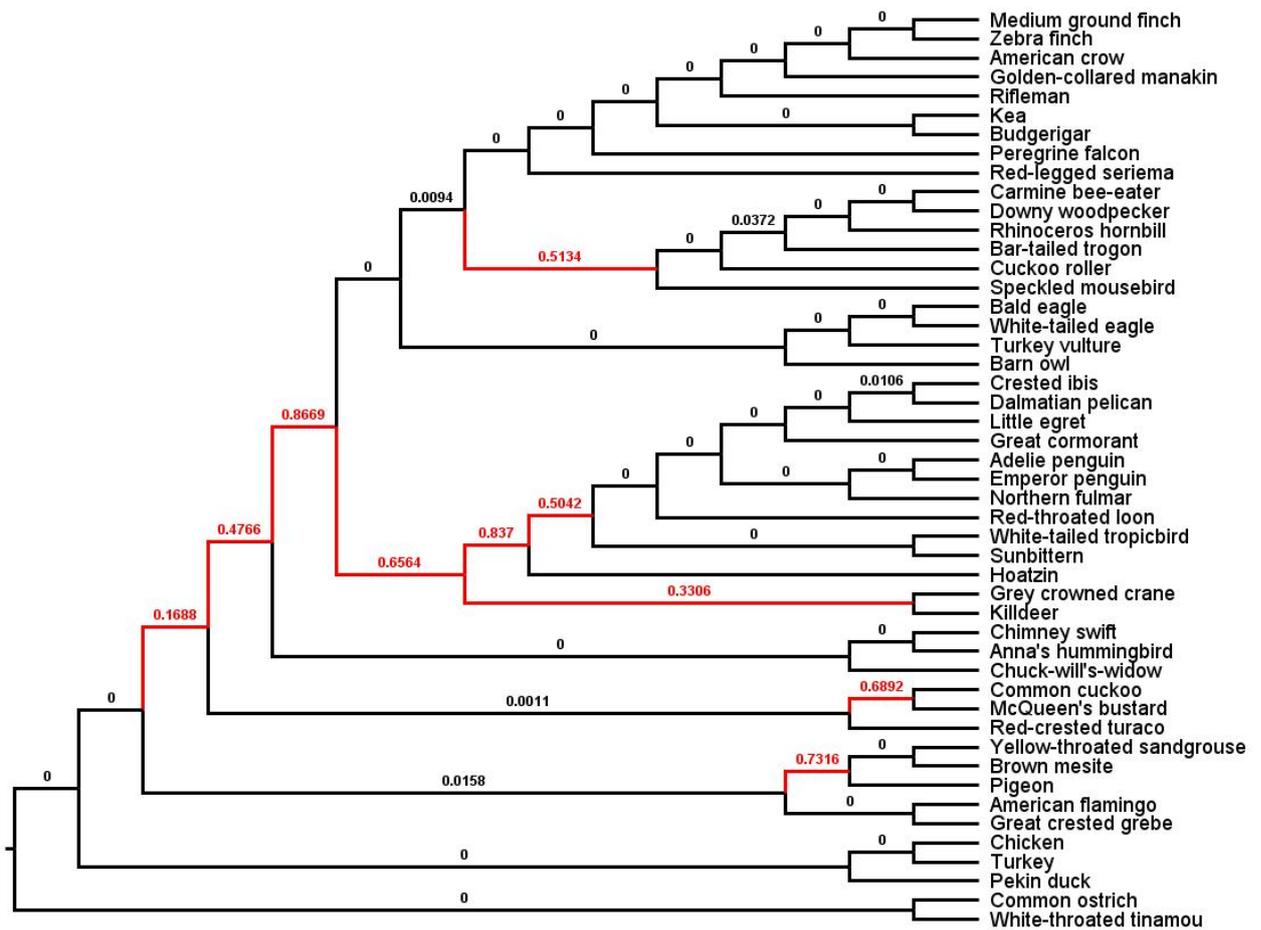


Figure S14. Polytomy test of intron indel dataset [11] conducted using ASTRAL III [12]. *p* values are shown on branches. Those highlighted in red indicate that a hard polytomy cannot be rejected with the data at hand at the $p \leq 0.05$ level.

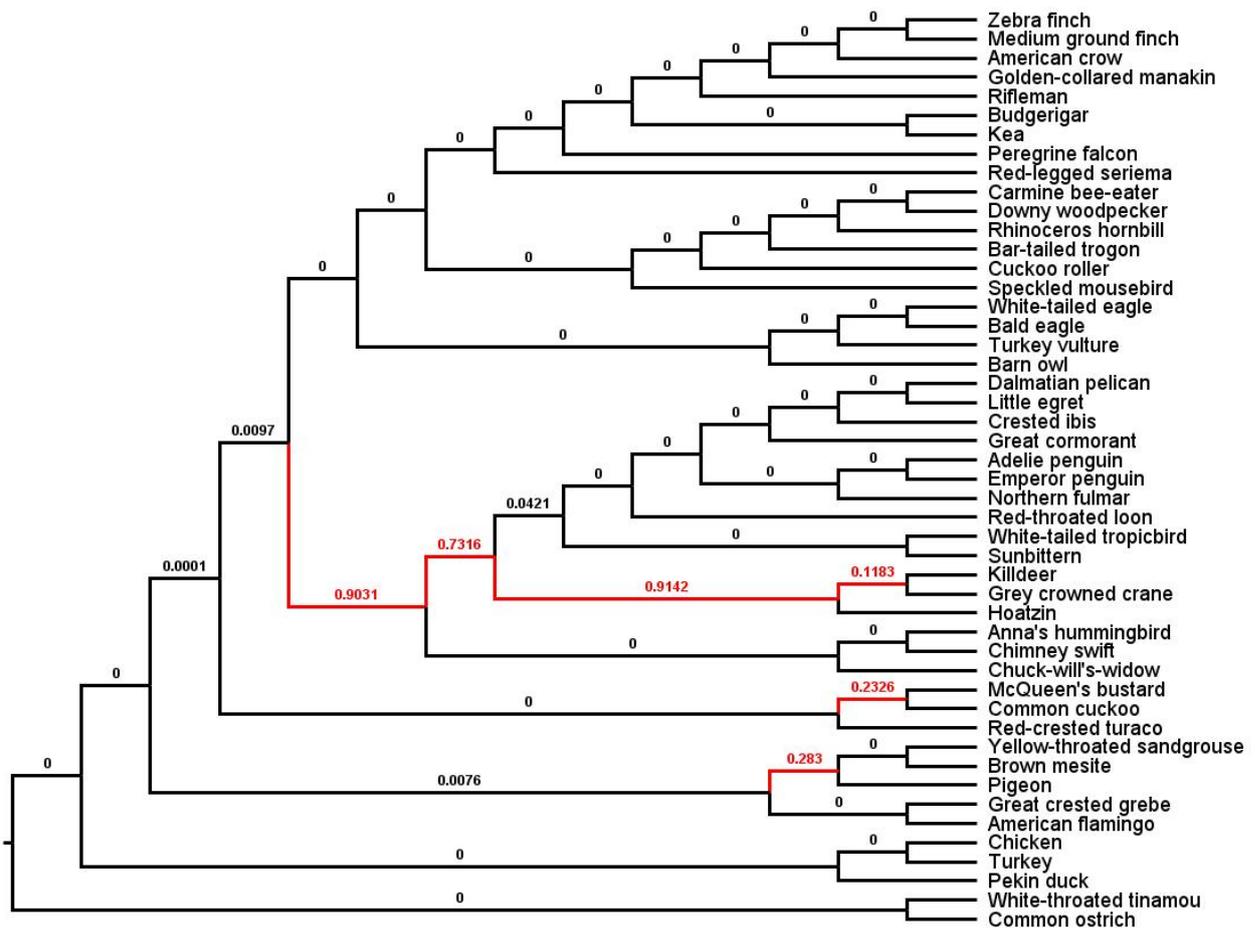


Figure S15. Polytomy test of intron nt dataset [11] conducted using ASTRAL III [12]. p values are shown on branches. Those highlighted in red indicate that a hard polytomy cannot be rejected with the data at hand at the $p \leq 0.05$ level.

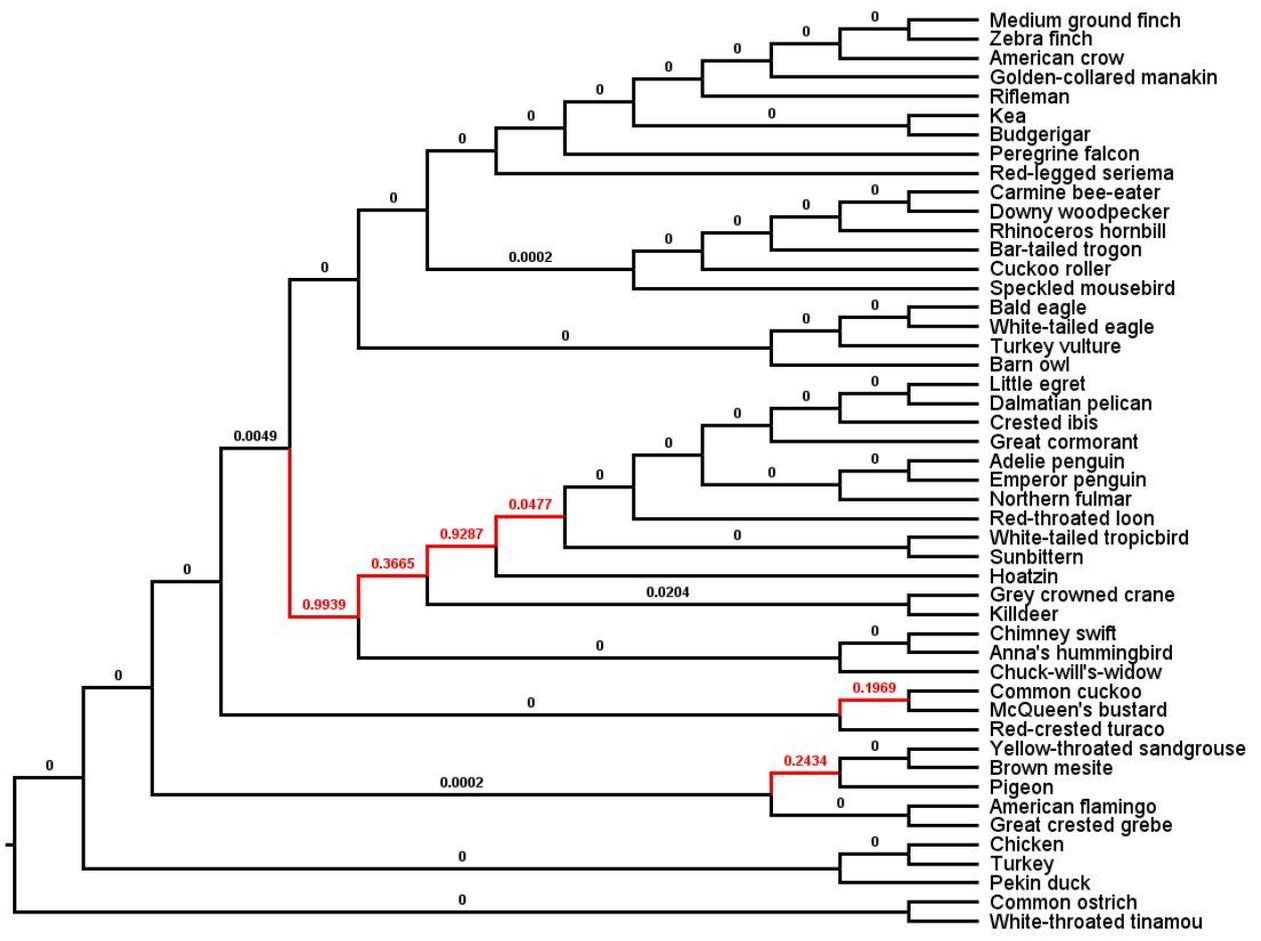


Figure S16. Polytomy test of combined intron indel+nt dataset [11] conducted using ASTRAL III [12]. *p* values are shown on branches. Those highlighted in red indicate that a hard polytomy cannot be rejected with the data at hand at the $p \leq 0.05$ level.

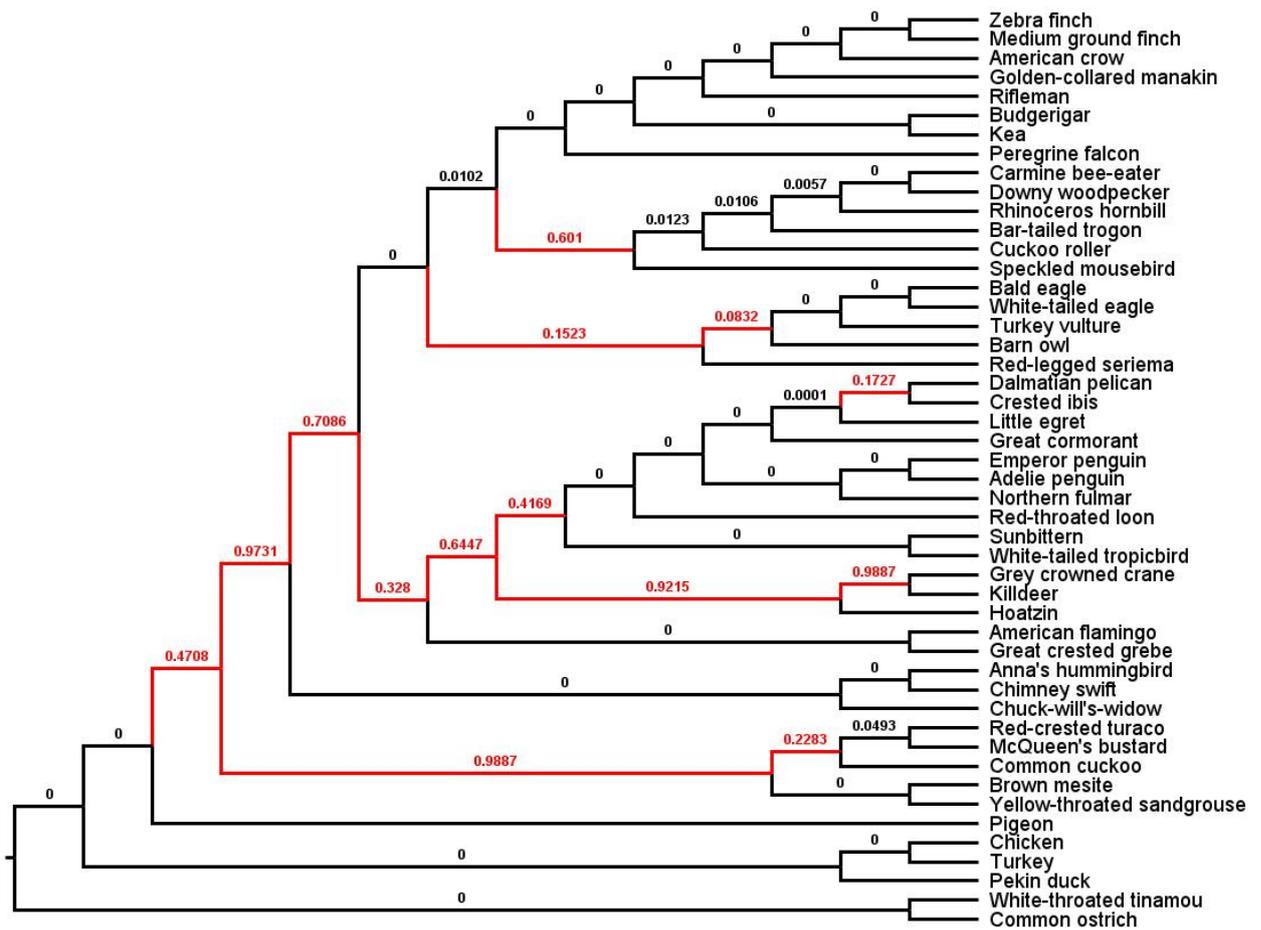


Figure S17. Polytomy test of UCE indel dataset [11] conducted using ASTRAL III [12]. p values are shown on branches. Those highlighted in red indicate that a hard polytomy cannot be rejected with the data at hand at the $p \leq 0.05$ level.

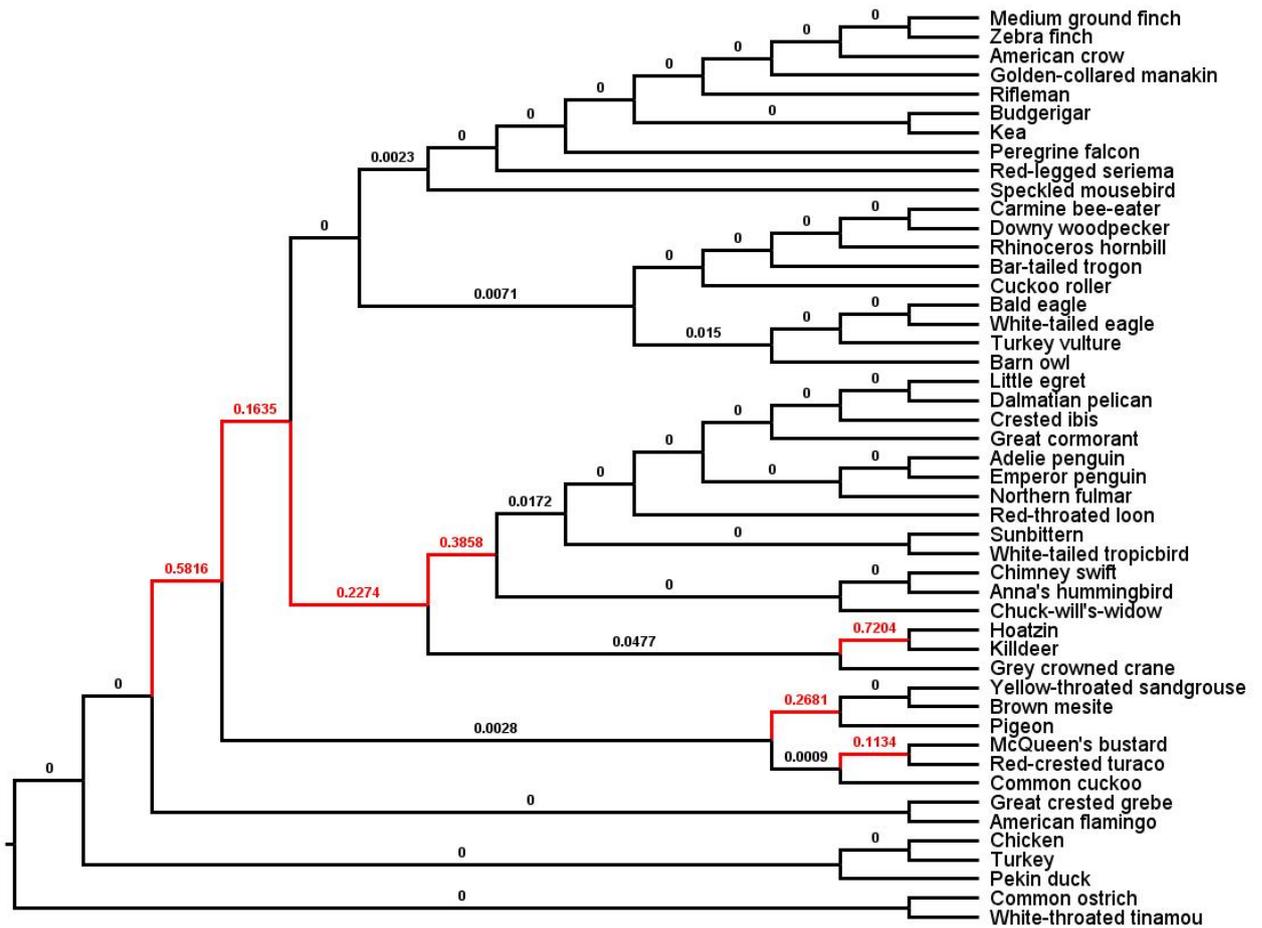


Figure S18. Polytomy test of the UCE nt dataset [11] conducted using ASTRAL III [12]. p values are shown on branches. Those highlighted in red indicate that a hard polytomy cannot be rejected with the data at hand at the $p \leq 0.05$ level.

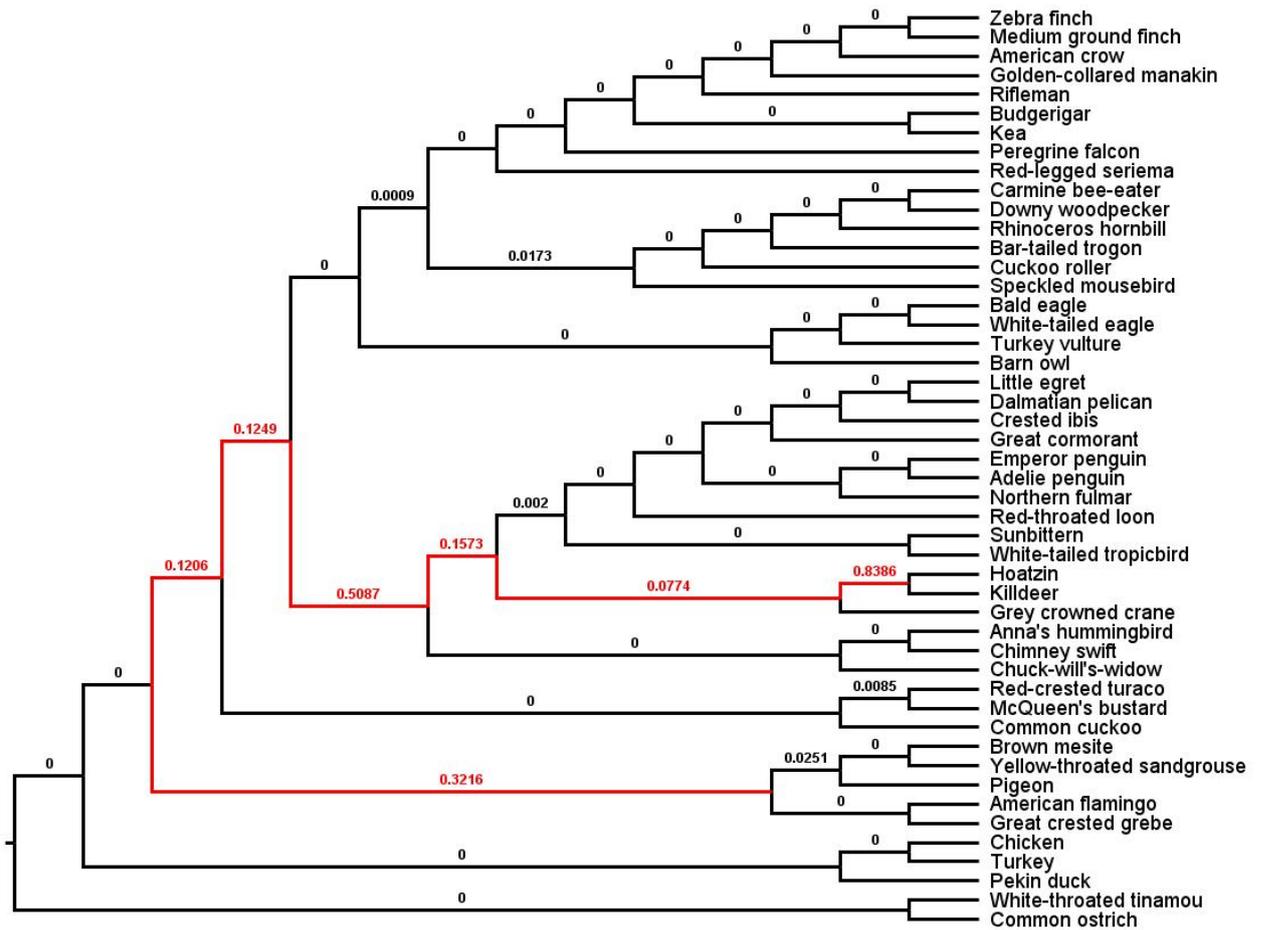


Figure S19. Polytomy test of the combined UCE indel+nt dataset [11] conducted using ASTRAL III [12]. p values are shown on branches. Those highlighted in red indicate that a hard polytomy cannot be rejected with the data at hand at the $p \leq 0.05$ level.

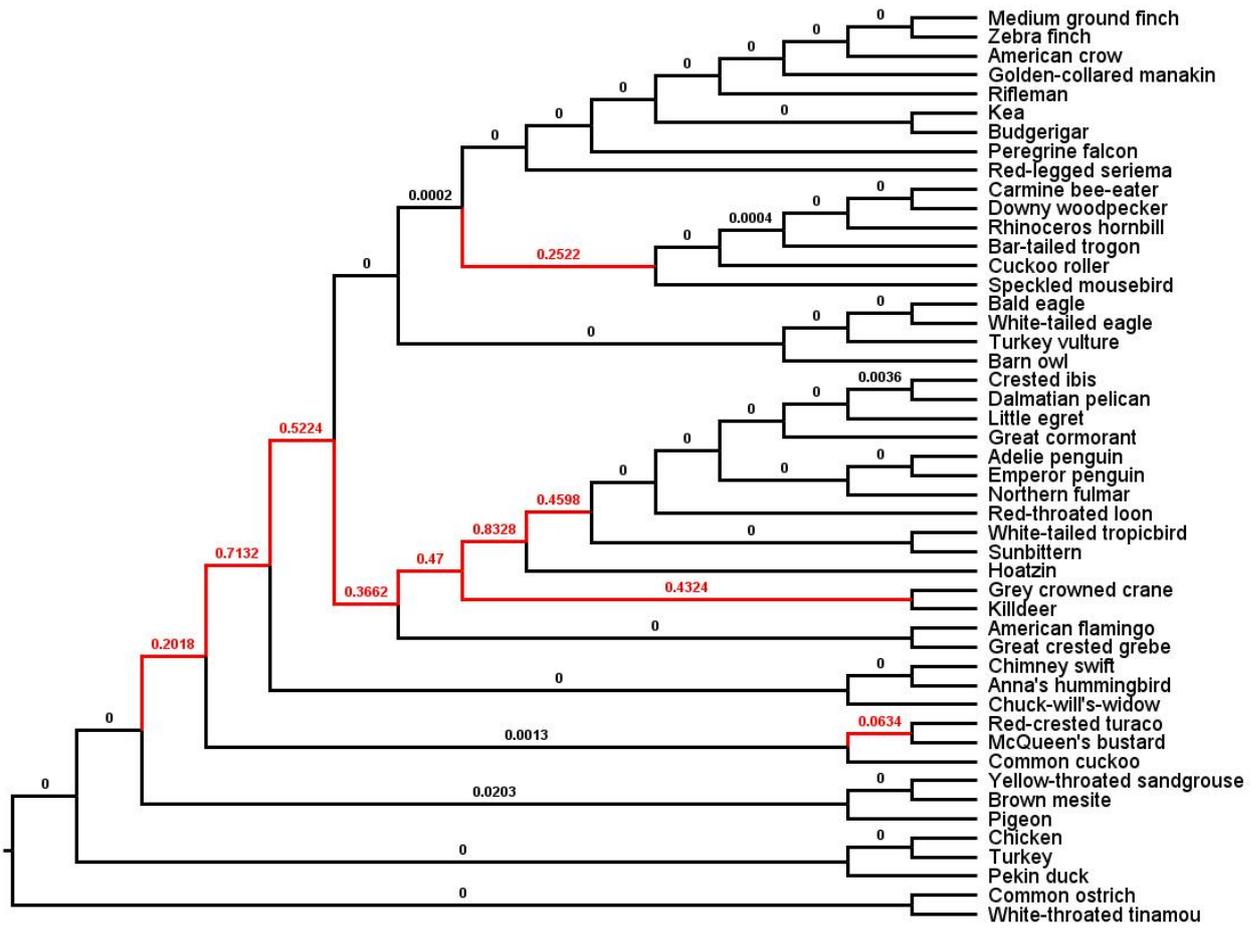


Figure S20. Polytomy test of the combined intron+UCE indel dataset [11] conducted using ASTRAL III [12]. p values are shown on branches. Those highlighted in red indicate that a hard polytomy cannot be rejected with the data at hand at the $p \leq 0.05$ level.

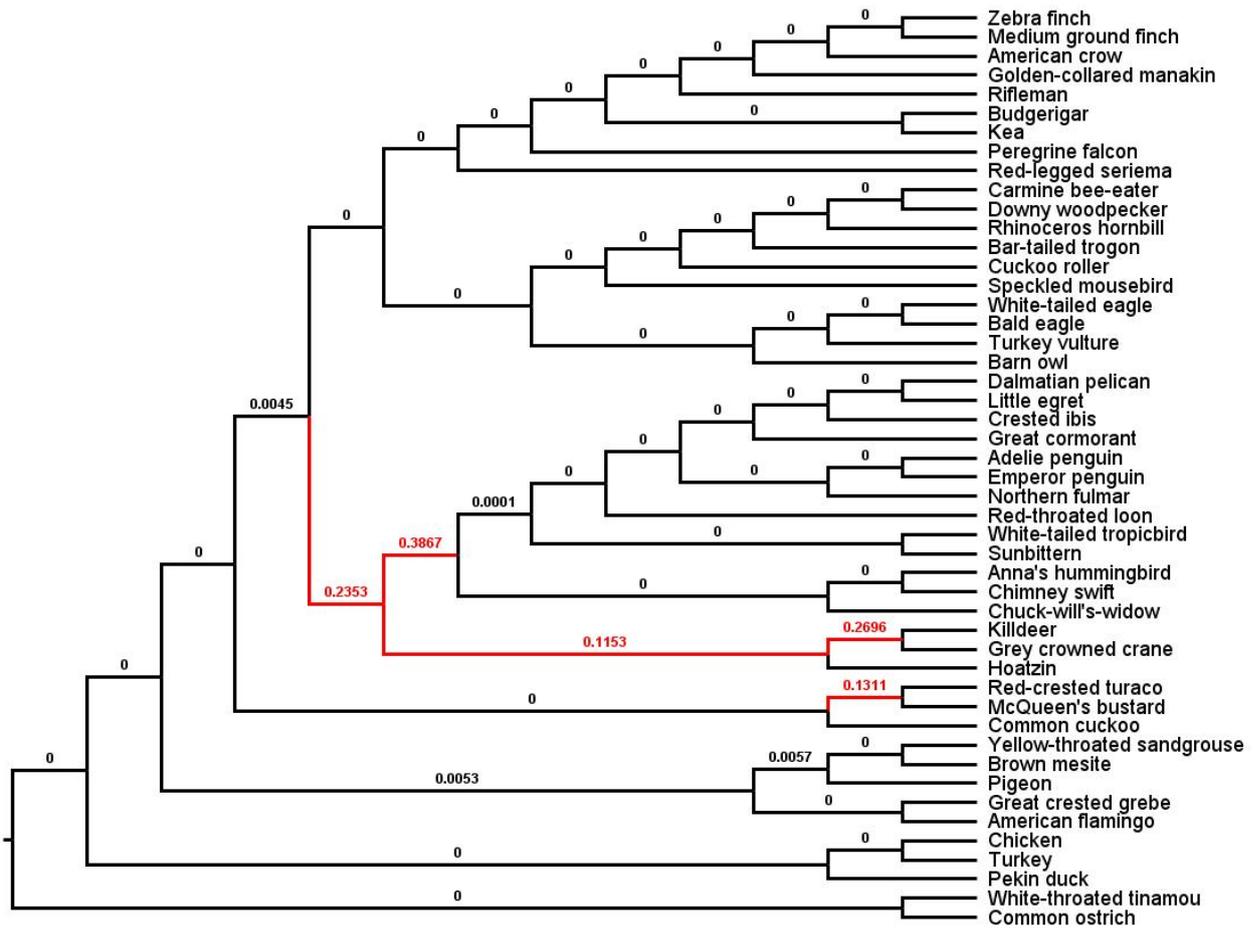


Figure S21. Polytomy test of the combined intron+UCE nt dataset [11] conducted using ASTRAL III [12]. p values are shown on branches. Those highlighted in red indicate that a hard polytomy cannot be rejected with the data at hand at the $p \leq 0.05$ level.

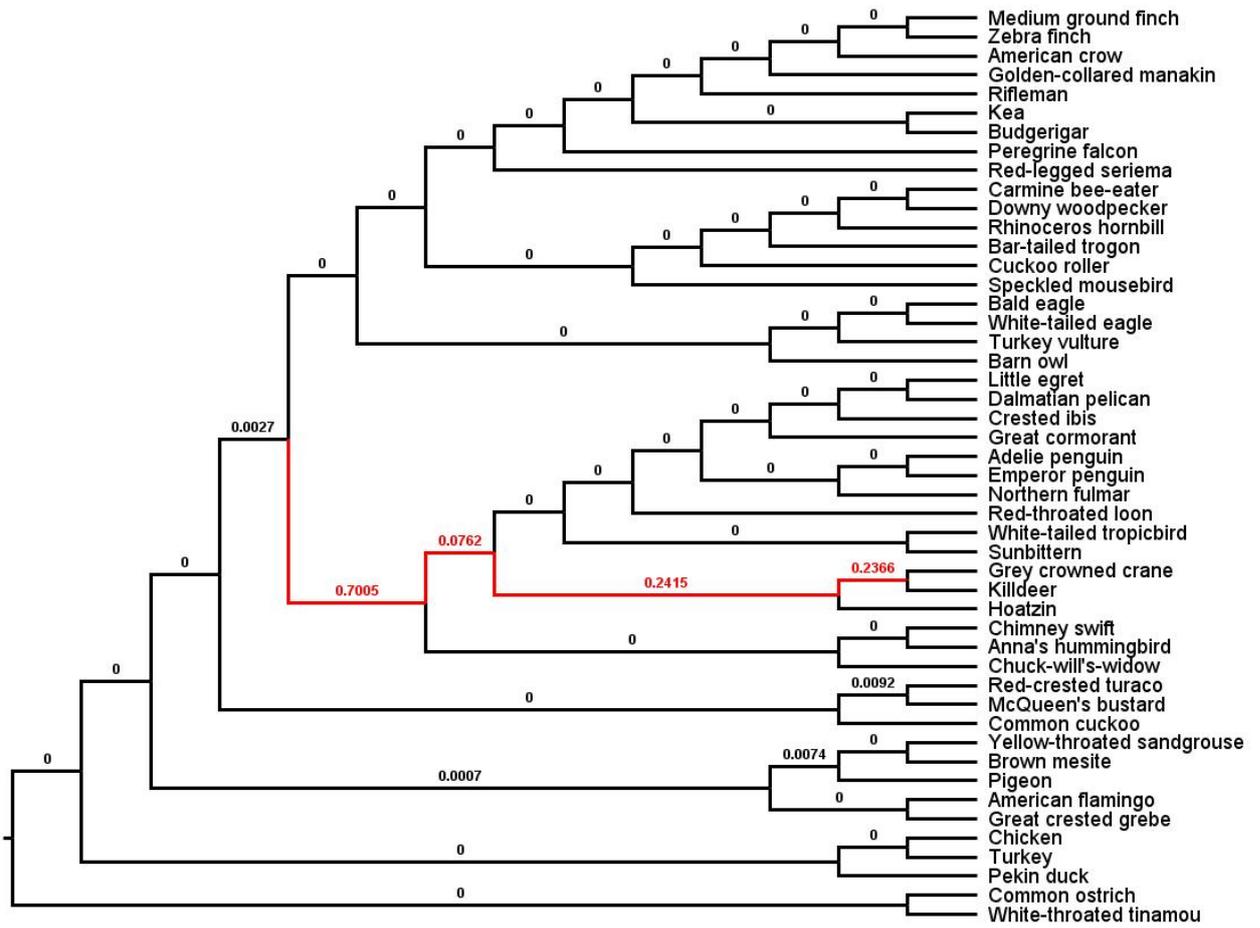


Figure S22. Polytomy test of the combined intron+UCE indel+nt dataset [11] conducted using ASTRAL III [12]. p values are shown on branches. Those highlighted in red indicate that a hard polytomy cannot be rejected with the data at hand at the $p \leq 0.05$ level.