

# Supplementary Material to wQFM: Highly Accurate Genome-scale Species Tree Estimation from Weighted Quartets

Mahim Mahbub<sup>1,†</sup>, Zahin Wahab<sup>1,†</sup>, Rezwana Reaz<sup>1,†,\*</sup>,  
M. Saifur Rahman<sup>1</sup>, and Md. Shamsuzzoha Bayzid<sup>1,\*</sup>

<sup>1</sup>Department of Computer Science and Engineering  
Bangladesh University of Engineering and Technology  
Dhaka-1205, Bangladesh

<sup>†</sup>These authors contributed equally to this work

\*Corresponding author: shams\_bayzid@cse.buet.ac.bd,  
rimpi@cse.buet.ac.bd

## Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Overview</b>   | <b>2</b>  |
| <b>2</b> | <b>Datasets</b>   | <b>2</b>  |
| <b>3</b> | <b>Additional Results</b>   | <b>2</b>  |
| 3.1      | Results on 11-taxon simulated dataset . . . . .                           | 2         |
| 3.2      | Results on 101-taxon simulated dataset . . . . .                          | 4         |
| 3.3      | Results on avian biological dataset: Exon, Intron and UCE trees . . . . . | 4         |
| 3.4      | Results on mammalian biological dataset . . . . .                         | 6         |
| 3.5      | Results on angiosperm dataset . . . . .                                   | 7         |
| <b>4</b> | <b>Impact of using weighted quartets: comparison between wQFM and QFM</b> | <b>8</b>  |
| <b>5</b> | <b>Quartet scores</b>   | <b>12</b> |
| <b>6</b> | <b>Running time</b>   | <b>17</b> |
| <b>7</b> | <b>Wilcoxon Signed-Rank Test</b>  | <b>19</b> |

# 1 Overview

These supplementary materials present additional details about the datasets analyzed in this study (Sec. 2), results on simulated and real biological datasets (Sec. 3), running time information (Sec. 6), and also additional tables and figures (Secs. 5 and 7). We also present a comparison between QFM and wQFM, showing the positive impact of including weights in estimating trees by amalgamating quartets (Sec. 4).

## 2 Datasets

The 37-taxon mammalian simulated dataset was simulated by taking the species tree estimated by MP-EST on the biological dataset studied in Song *et al.* [1]. This species tree had branch lengths in coalescent units, which we used to produce a set of gene trees under the coalescent model. Thus, the model tree has an ILS level based on a coalescent analysis of the biological mammalian dataset, and other properties of the simulation that are set to reflect the biological sequences they studied. We explored the impact of varying numbers of genes (25 - 800), varying amounts of gene tree estimation error (i.e., the amount of phylogenetic signal by varying the sequence length for the markers: 250bp - 1500bp). In both cases, the levels of ILS were varied (shorter branches increases ILS) by multiplying or dividing all internal branch lengths in the model species tree by two. Thus, we have three model conditions that are referred to as 1X (moderate ILS), 0.5X (high ILS) and 2X (low ILS). The 48-taxon avian simulated dataset is based on the species tree estimated using MP-EST on the avian dataset of [2], and was simulated by following a similar procedure as the mammalian dataset. Similar to the mammalian dataset, it has three different ILS levels (1X, 0.5x and 2X), albeit the ILS levels are higher than the mammalian dataset (i.e., more discordance between the true gene trees and the species tree).

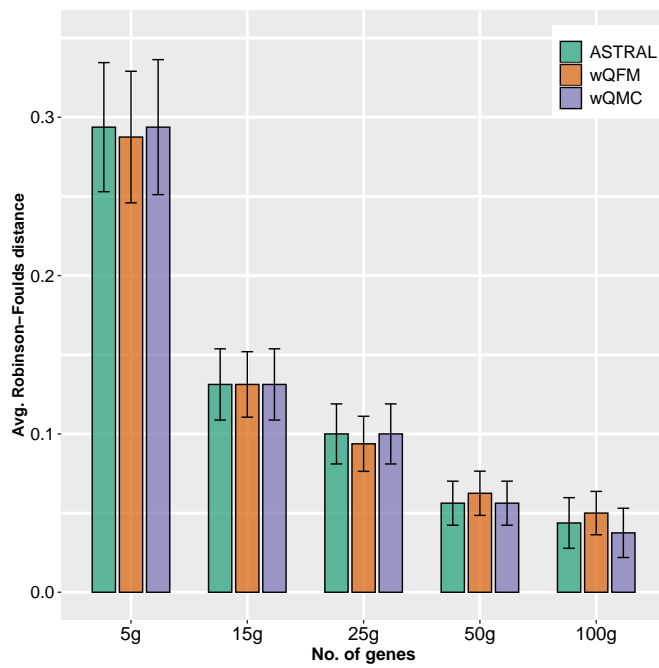
We analyzed the high-ILS 11-taxon datasets from [3] (as the model condition with lower amount of ILS is very easy to analyze [4]) which varies in the number of genes and amount of gene tree estimation error. 15-taxon datasets contain a high level of ILS and vary in sequence lengths and numbers of genes. Thus, the simulated datasets provide a wide range of challenging and practical model conditions in which we explore the performance of wQFM.

## 3 Additional Results

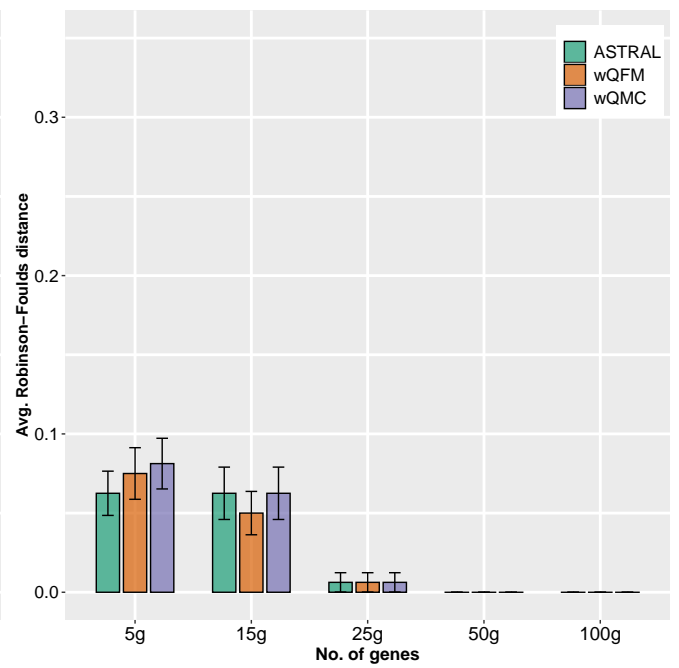
### 3.1 Results on 11-taxon simulated dataset

The performance of various methods on 11-taxon high-ILS dataset with varying numbers of estimated and true gene trees is shown in Fig. S1. On this dataset, wQFM, ASTRAL and wQMC achieved comparable tree accuracy with no statistically significant difference.

As was expected, the accuracy of these methods improved with the increase in the number of genes and they returned highly accurate species trees when true gene trees were used (even with only 25 genes).



(a) Analysis of estimated gene trees



(b) Analysis of true gene trees

Figure S1: **Comparison of ASTRAL, wQFM and wQMC on 11-taxon high-ILS dataset.** We varied the number of genes (5 genes to 100 genes) for both estimated and true gene trees. We show the average RF rates with standard errors over 20 replicates.

### 3.2 Results on 101-taxon simulated dataset

We show average RF rates of wQFM, ASTRAL and wQMC on 50 replicates of 101-taxon dataset with 1000 true gene trees (see Fig. S2). All of these methods produced highly accurate trees with around 1.3% ~3.1% tree error. wQFM was the most accurate method followed by ASTRAL. The difference between wQFM and ASTRAL was very small and was not statistically significant. However, the improvement of wQFM over wQMC was statistically significant.

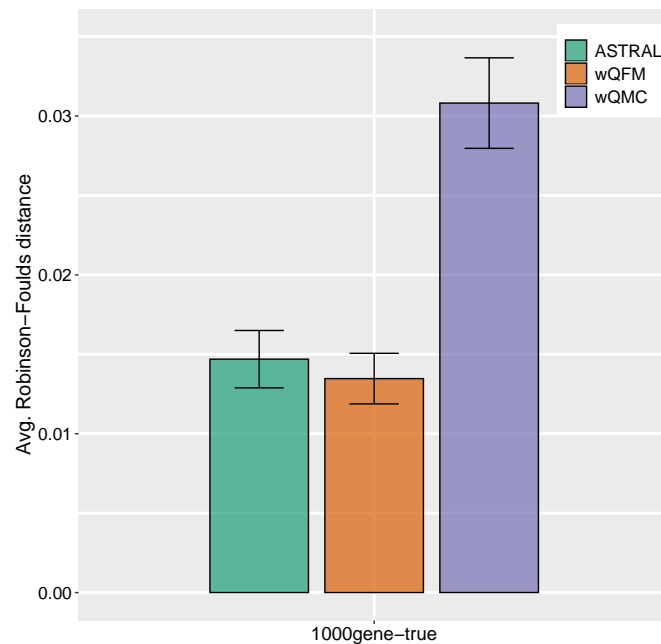


Figure S2: **Comparison of ASTRAL, wQFM and wQMC on 101-taxon dataset.** We show average RF rates with standard error bars over 50 replicates.

### 3.3 Results on avian biological dataset: Exon, Intron and UCE trees

We estimated species trees on 8,251 exon, 2,516 intron and 3,679 UCE gene trees from [2] using wQFM, ASTRAL and wQMC. Fig. S3 shows the trees estimated by different methods.

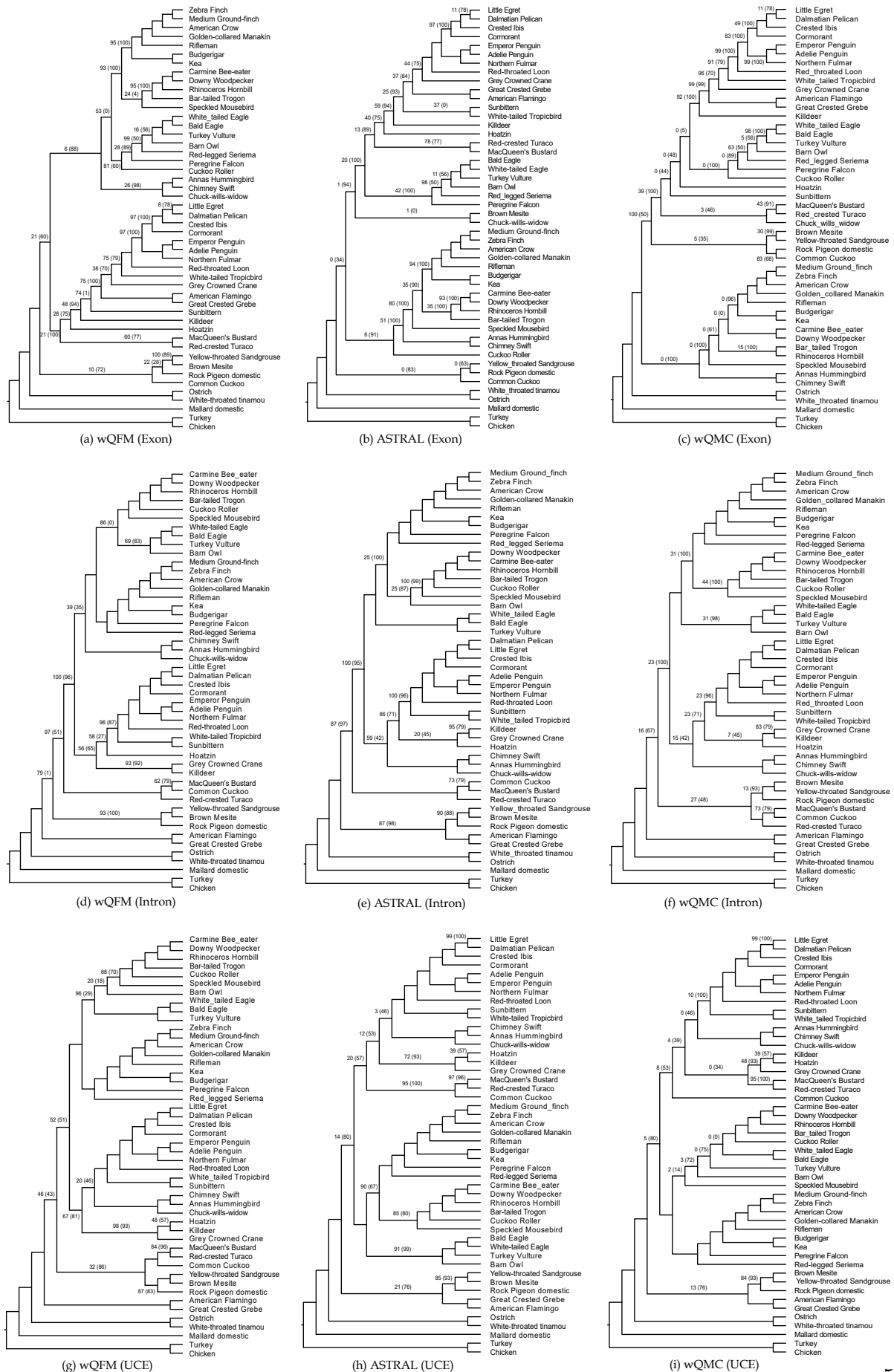


Figure S3: **Estimated trees on exon, intron and UCE gene trees.** Two branch support (BS) values are shown: the first value (without parentheses) is based on site-only MLBS and the second value (with parentheses) is quartet based local posterior probability (multiplied by 100). All BS values are 100% except where noted.

### 3.4 Results on mammalian biological dataset

Figure S4 shows the trees produced by wQFM, ASTRAL and wQMC on the mammalian dataset.

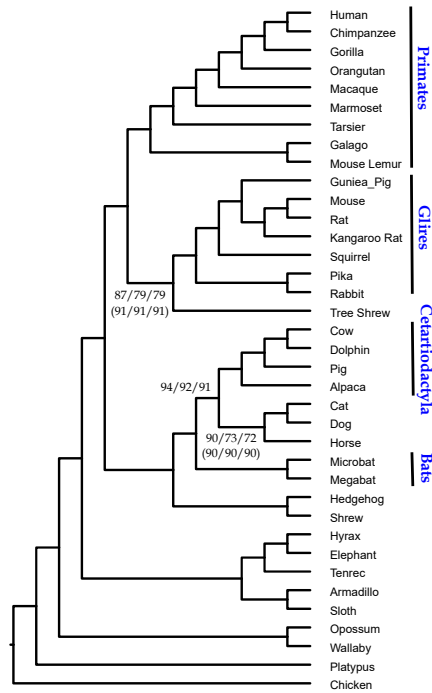


Figure S4: **Analyses of the mammalian dataset using ASTRAL, wQFM and wQMC.** All of these methods reconstructed the same topology. Two branch support values are shown: the values without parentheses are based on site-only MLBS and the values with parentheses are quartet based local posterior probabilities (multiplied by 100). For each of these two types of support, we show the BS for three methods (wQMF, ASTRAL and wQMC) in the following order: wQFM/ASTRAL/wQMC. All BS values are 100% except where noted.

### 3.5 Results on angiosperm dataset

Our analyses with ASTRAL, wQFM and wQMC support the placement of *Amborella* as sister to water lilies (i.e., Nymphaeales) and rest of the angiosperms (see Fig. S5). This placement of *Amborella* is congruent to the CA-ML analysis in Xi *et al.* [5] and other molecular studies [6, 7, 8]. An alternate hypothesis, which supports a clade containing *Amborella* plus water lilies that is sister to all other angiosperms, has also been observed [5, 9, 10]. wQFM and wQMC differ from ASTRAL on a single edge (the placement of Sapindales (Citrus)).

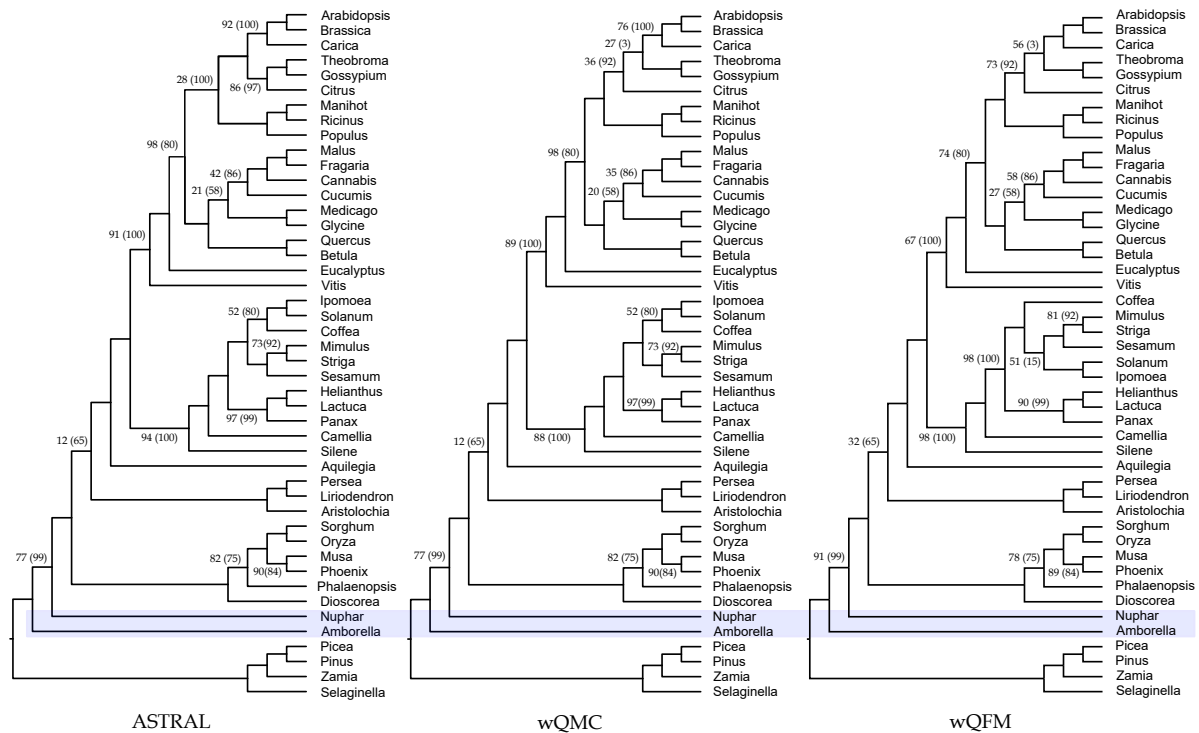


Figure S5: Analyses of the angiosperm dataset using ASTRAL, wQFM and wQMC. All of these methods support the placement of *Amborella* alone as sister to all other extant angiosperms. Two branch support (BS) values are shown: the first value (without parentheses) is based on site-only MLBS and the second value (with parentheses) is quartet based local posterior probability (multiplied by 100). All BS values are 100% except where noted.

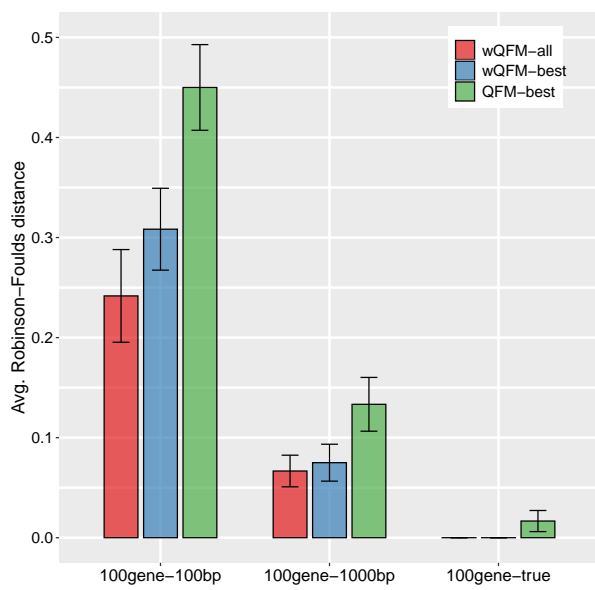
## 4 Impact of using weighted quartets: comparison between wQFM and QFM

A direct comparison between wQFM and QFM using all possible quartets (weighted quartets for wQFM and unweighted quartet for QFM) sampled from the input set of gene trees is not meaningful. This is because using all possible  $3\binom{n}{4}$  quartets *without any* weight as input and seeking a tree by maximizing the number of consistent quartets will just produce a random tree. Therefore, when using the unweighted setting – for each set of four taxa – the dominant quartets out of the three alternate quartet topologies are usually used [11]. In this study, we consider the support/weight for quartet tree  $ab|cd$  to be the number of the trees in  $\mathcal{G}$  that induce  $ab|cd$  on set  $a, b, c, d$ . Therefore, for a set of four taxa  $a, b, c, d$ , the dominant quartet (out of three possible quartets:  $ab|cd$ ,  $ac|bd$ , and  $bc|ad$ ) is defined to be the quartet with the highest weight. In order to show the efficacy of the weighted setting, we compare the following three variants of quartet amalgamation.

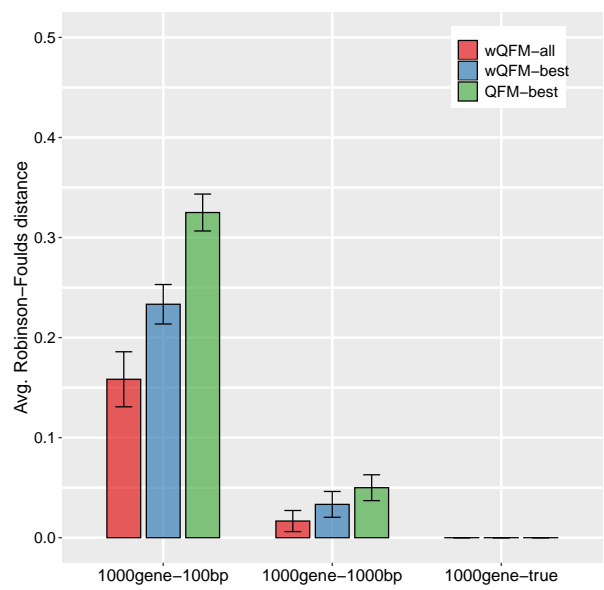
1. wQFM-all: wQFM with all possible weighted quartets.
2. wQFM-best: wQFM with weighted best (dominant) quartets. That means,  $\binom{n}{4}$  best weighted quartets (one quartet for each set of four taxa) are used instead of  $3\binom{n}{4}$  possible weighted quartets.
3. QFM-best: QFM with unweighted best (dominant) quartets (one quartet for each set of four taxa).

Figures S6-S8 show the comparison among these three variants on various simulated datasets. These results show the superiority of weighted setting over the unweighted setting as, in most of the cases, wQFM with all possible weighted quartets outperformed the other two variants. Moreover, wQFM with weighted best quartets outperformed QFM with unweighted best quartets in many of the model conditions on these datasets – another evidence that assigning weights to the quartets can improve phylogenetic analyses.





(a) Analysis on 100 genes



(b) Analysis on 1000 genes

Figure S6: Comparison of wQFM and QFM on 15-taxon dataset. We show the average RF rates over 10 replicates.

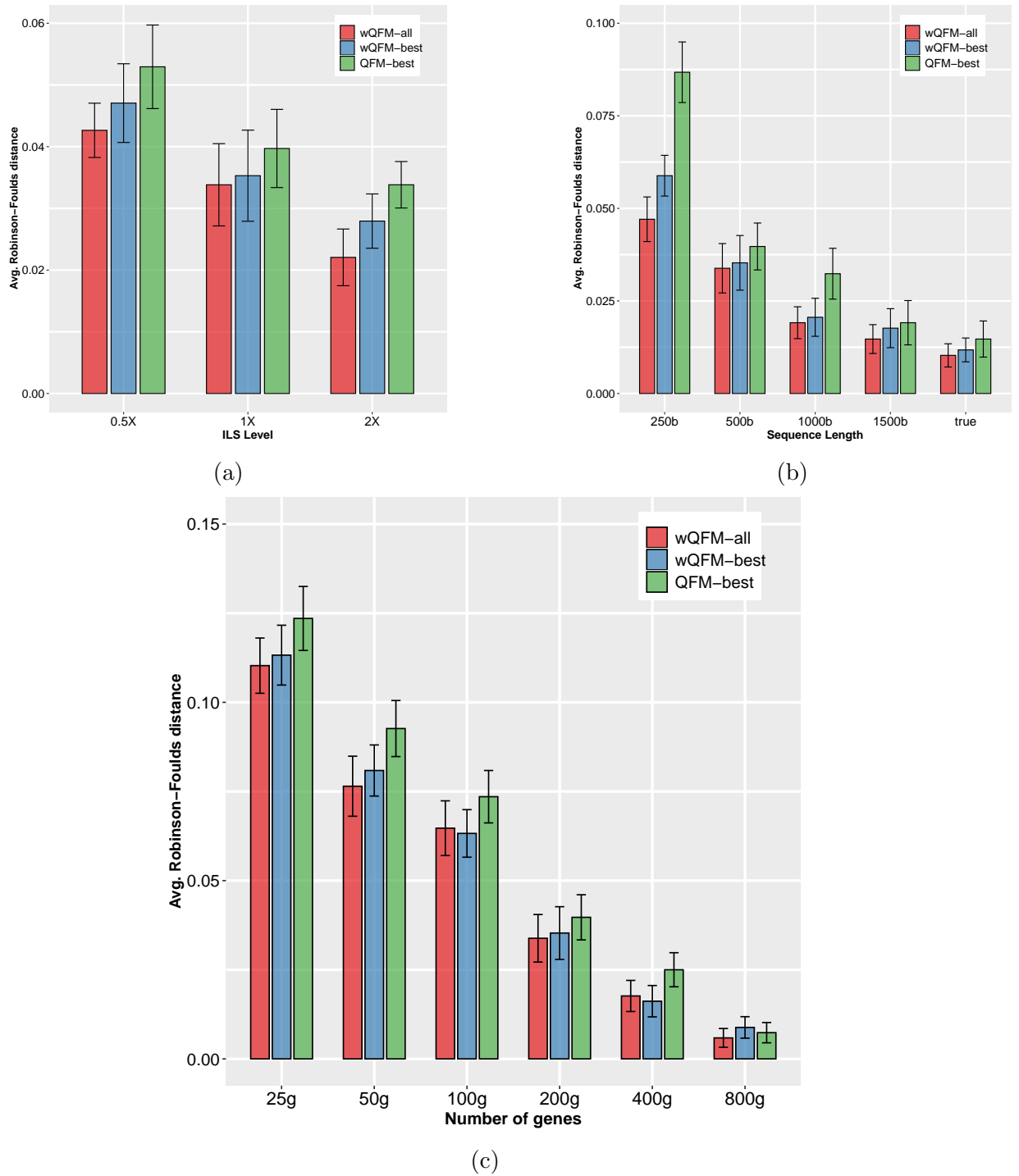
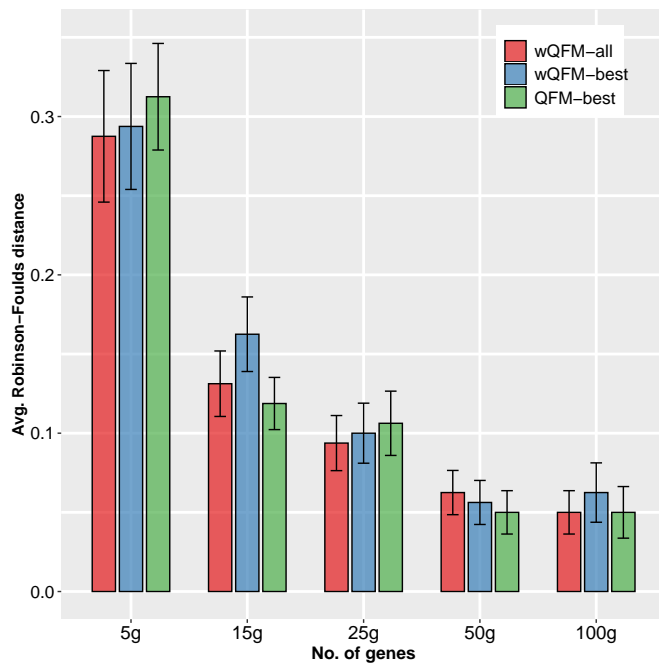
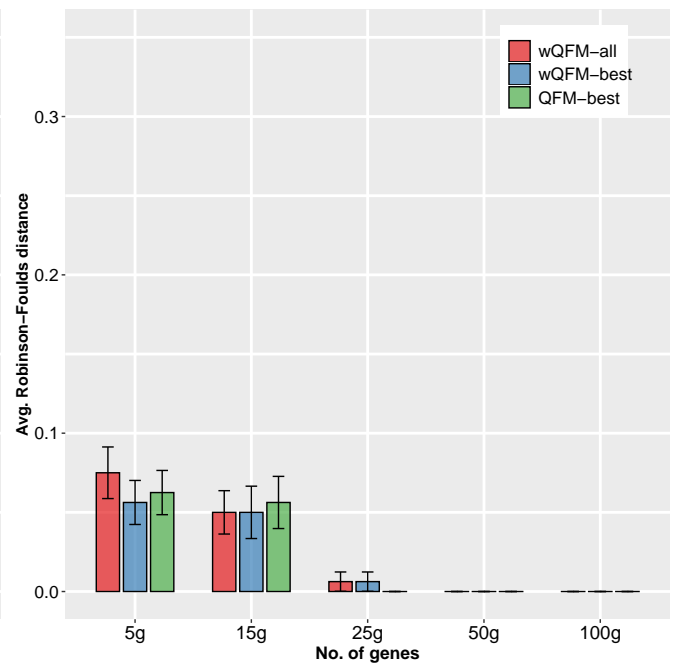


Figure S7: Comparison of wQFM and QFM on 37-taxon simulated dataset over 20 replicates per model condition. (a) The level of ILS was varied from 0.5X (highest) to 2X (lowest) amount, keeping the sequence length fixed at 500bp and the number of genes at 200; (b) The sequence length was varied from 250bp to 1500bp, keeping the number of genes fixed at 200, and ILS at 1X (moderate ILS); (c) The number of genes was varied from 25g to 800g, with 500bp sequence length and moderate (1X) ILS.



(a) Analysis of estimated gene trees



(b) Analysis of true gene trees

Figure S8: Comparison of wQFM and QFM on 11-taxon high-ILS dataset. We show the average RF rates over 20 replicates.

## 5 Quartet scores

Tables S1-S6 show the quartet scores of different methods on various simulated (11-, 15-, 37-, 48- and 101-taxon dataset) and biological datasets. For each model condition across various datasets, average quartet scores, average total weight of weighted quartets in the input gene tree, and average normalized quartets scores are shown for wQFM, wQMC and ASTRAL.

Table S1: **Quartet scores on 15-taxon dataset.** We show average (over 20 replicates) quartet scores (sum of the weights of the satisfied quartets) of various methods, total weight of the quartets in the input gene trees, and their respective ratios.

| Model Condition | Tree       | Quartet score | Total weight (input quartets) | Proportion (%) |
|-----------------|------------|---------------|-------------------------------|----------------|
| 100gene-100bp   | Model Tree | 69307.4       | 136500.0                      | 50.775         |
|                 | wQFM       | 69776.6       | 136500.0                      | 51.118         |
|                 | wQMC       | 69930.0       | 136500.0                      | 51.231         |
|                 | ASTRAL     | 69933.8       | 136500.0                      | 51.234         |
| 100gene-1000bp  | Model Tree | 82099.2       | 136500.0                      | 60.146         |
|                 | wQFM       | 82129.1       | 136500.0                      | 60.168         |
|                 | wQMC       | 82166.1       | 136500.0                      | 60.195         |
|                 | ASTRAL     | 82166.6       | 136500.0                      | 60.195         |
| 100gene-true    | Model Tree | 84634.9       | 136500.0                      | 62.004         |
|                 | wQFM       | 84634.9       | 136500.0                      | 62.004         |
|                 | wQMC       | 84634.9       | 136500.0                      | 62.004         |
|                 | ASTRAL     | 84634.9       | 136500.0                      | 62.004         |
| 1000gene-100bp  | Model Tree | 690268.0      | 1365000.0                     | 50.569         |
|                 | wQFM       | 692949.2      | 1365000.0                     | 50.766         |
|                 | wQMC       | 693656.1      | 1365000.0                     | 50.817         |
|                 | ASTRAL     | 693656.1      | 1365000.0                     | 50.817         |
| 1000gene-1000bp | Model Tree | 817937.3      | 1365000.0                     | 59.922         |
|                 | wQFM       | 817993.9      | 1365000.0                     | 59.926         |
|                 | wQMC       | 818022.2      | 1365000.0                     | 59.928         |
|                 | ASTRAL     | 818022.2      | 1365000.0                     | 59.928         |
| 1000gene-true   | Model Tree | 844184.3      | 1365000.0                     | 61.845         |
|                 | wQFM       | 844184.3      | 1365000.0                     | 61.845         |
|                 | wQMC       | 844184.3      | 1365000.0                     | 61.845         |
|                 | ASTRAL     | 844184.3      | 1365000.0                     | 61.845         |

Table S2: **Quartet scores on 37-taxon mammalian simulated dataset.** We show average (over 20 replicates) quartet scores (sum of the weights of the satisfied quartets) of the true and estimated trees, total weight of the quartets in the input gene trees, and their respective ratios. Various model conditions are defined by different ILS levels (1X, 0.5X, 2X), numbers of genes (100g, 200g, etc.) and sequence lengths (500b, 100b, etc.).

| Model Condition | Tree       | Quartet score | Total weight (input quartets) | Proportion (%) |
|-----------------|------------|---------------|-------------------------------|----------------|
| 1X-100g-500b    | Model Tree | 5630260.55    | 6604500.0                     | 85.249         |
|                 | wQFM       | 5635762.55    | 6604500.0                     | 85.332         |
|                 | wQMC       | 5636032.85    | 6604500.0                     | 85.336         |
|                 | ASTRAL     | 5636032.85    | 6604500.0                     | 85.336         |
| 1X-200g-1000b   | Model Tree | 11584969.95   | 13209000.0                    | 87.705         |
|                 | wQFM       | 11586354.60   | 13209000.0                    | 87.716         |
|                 | wQMC       | 11586641.45   | 13209000.0                    | 87.718         |
|                 | ASTRAL     | 11586641.45   | 13209000.0                    | 87.718         |
| 1X-200g-1500b   | Model Tree | 11657249.90   | 13209000.0                    | 88.252         |
|                 | wQFM       | 11658311.05   | 13209000.0                    | 88.260         |
|                 | wQMC       | 11658432.55   | 13209000.0                    | 88.261         |
|                 | ASTRAL     | 11658445.25   | 13209000.0                    | 88.261         |
| 1X-200g-250b    | Model Tree | 10557321.60   | 13209000.0                    | 79.925         |
|                 | wQFM       | 10560454.40   | 13209000.0                    | 79.949         |
|                 | wQMC       | 10561734.40   | 13209000.0                    | 79.959         |
|                 | ASTRAL     | 10561825.45   | 13209000.0                    | 79.959         |
| 1X-200g-true    | Model Tree | 11744078.75   | 13209000.0                    | 88.910         |
|                 | wQFM       | 11746076.85   | 13209000.0                    | 88.925         |
|                 | wQMC       | 11746177.35   | 13209000.0                    | 88.926         |
|                 | ASTRAL     | 11746203.50   | 13209000.0                    | 88.926         |
| 1X-25g-500b     | Model Tree | 1410250.70    | 1651125.0                     | 85.412         |
|                 | wQFM       | 1414398.50    | 1651125.0                     | 85.663         |
|                 | wQMC       | 1414450.05    | 1651125.0                     | 85.666         |
|                 | ASTRAL     | 1414559.25    | 1651125.0                     | 85.672         |
| 1X-400g-500b    | Model Tree | 22531906.20   | 26418000.0                    | 85.290         |
|                 | wQFM       | 22534417.70   | 26418000.0                    | 85.299         |
|                 | wQMC       | 22534424.55   | 26418000.0                    | 85.300         |
|                 | ASTRAL     | 22534424.55   | 26418000.0                    | 85.300         |
| 1X-50g-500b     | Model Tree | 2816708.35    | 3302250.0                     | 85.297         |
|                 | wQFM       | 2820640.95    | 3302250.0                     | 85.416         |
|                 | wQMC       | 2820892.25    | 3302250.0                     | 85.423         |
|                 | ASTRAL     | 2821123.85    | 3302250.0                     | 85.430         |
| 1X-800g-500b    | Model Tree | 45096639.80   | 52836000.0                    | 85.352         |
|                 | wQFM       | 45096858.95   | 52836000.0                    | 85.353         |
|                 | wQMC       | 45096874.45   | 52836000.0                    | 85.353         |
|                 | ASTRAL     | 45096874.45   | 52836000.0                    | 85.353         |
| 0.5X-200g-500b  | Model Tree | 10003929.55   | 13209000.0                    | 75.736         |
|                 | wQFM       | 10006593.65   | 13209000.0                    | 75.756         |
|                 | wQMC       | 10006510.55   | 13209000.0                    | 75.755         |
|                 | ASTRAL     | 10007425.55   | 13209000.0                    | 75.762         |
| 2X-200g-500b    | Model Tree | 11943759.05   | 13205697.75                   | 90.444         |
|                 | wQFM       | 11944408.40   | 13205697.75                   | 90.449         |
|                 | wQMC       | 11944654.40   | 13205697.75                   | 90.451         |
|                 | ASTRAL     | 11944654.40   | 13205697.75                   | 90.451         |
| 1X-200g-500b    | Model Tree | 11266716.15   | 13209000.0                    | 85.296         |
|                 | wQFM       | 11270240.85   | 13209000.0                    | 85.322         |
|                 | wQMC       | 11270657.55   | 13209000.0                    | 85.326         |
|                 | ASTRAL     | 11270657.55   | 13209000.0                    | 85.326         |

Table S3: **Quartet scores on 48-taxon simulated avian dataset with sequence length of 500 bp.** We show average (over 20 replicates) quartet scores (sum of the weights of the satisfied quartets) of the true and estimated trees, total weight of the quartets in the input gene trees, and their respective ratios. Various model conditions are defined by different ILS levels (1X, 0.5X, 2X), numbers of genes (100g, 200g, etc.) but by keeping the sequence length fixed to 500 bp.

| Model Condition | Tree       | Quartet score | Total weight (input quartets) | Proportion (%) |
|-----------------|------------|---------------|-------------------------------|----------------|
| 0.5X-1000g-500b | Model Tree | 110526512.90  | 194580000.0                   | 56.803         |
|                 | wQFM       | 110590453.55  | 194580000.0                   | 56.835         |
|                 | wQMC       | 109888348.30  | 194580000.0                   | 56.475         |
|                 | ASTRAL     | 110583148.55  | 194580000.0                   | 56.832         |
| 1X-1000g-500b   | Model Tree | 122945091.00  | 194580000.0                   | 63.185         |
|                 | wQFM       | 122994498.35  | 194580000.0                   | 63.21          |
|                 | wQMC       | 122865927.30  | 194580000.0                   | 63.144         |
|                 | ASTRAL     | 123024937.95  | 194580000.0                   | 63.226         |
| 2X-1000g-500b   | Model Tree | 131567523.05  | 194580000.0                   | 67.616         |
|                 | wQFM       | 131619669.80  | 194580000.0                   | 67.643         |
|                 | wQMC       | 131457328.10  | 194580000.0                   | 67.56          |
|                 | ASTRAL     | 131631393.90  | 194580000.0                   | 67.649         |
| 1X-50g-500b     | Model Tree | 6109255.45    | 9729000.0                     | 62.794         |
|                 | wQFM       | 6191081.00    | 9729000.0                     | 63.635         |
|                 | wQMC       | 6209354.00    | 9729000.0                     | 63.823         |
|                 | ASTRAL     | 6204883.70    | 9729000.0                     | 63.777         |
| 1X-100g-500b    | Model Tree | 12250259.15   | 19458000.0                    | 62.957         |
|                 | wQFM       | 12324222.00   | 19458000.0                    | 63.338         |
|                 | wQMC       | 12342740.10   | 19458000.0                    | 63.433         |
|                 | ASTRAL     | 12349310.60   | 19458000.0                    | 63.466         |
| 1X-200g-500b    | Model Tree | 24553380.50   | 38916000.0                    | 63.093         |
|                 | wQFM       | 24606333.80   | 38916000.0                    | 63.229         |
|                 | wQMC       | 24618794.95   | 38916000.0                    | 63.261         |
|                 | ASTRAL     | 24632823.80   | 38916000.0                    | 63.297         |
| 1X-500g-500b    | Model Tree | 61438944.15   | 97290000.0                    | 63.150         |
|                 | wQFM       | 61489325.10   | 97290000.0                    | 63.202         |
|                 | wQMC       | 61440254.85   | 97290000.0                    | 63.152         |
|                 | ASTRAL     | 61528828.05   | 97290000.0                    | 63.243         |

Table S4: **Quartet scores on 48-taxon simulated avian dataset with true gene trees.** We show average (over 20 replicates) quartet scores (sum of the weights of the satisfied quartets) of the true and estimated trees, total weight of the quartets in the input gene trees, and their respective ratios. Various model conditions are defined by different ILS levels (1X, 0.5X, 2X) and numbers of genes (100g, 200g, etc.).

| Model Condition | Tree       | Quartet score | Total weight (input quartets) | Proportion (%) |
|-----------------|------------|---------------|-------------------------------|----------------|
| 0.5X-1000g-true | Model Tree | 119669366.65  | 194580000.0                   | 61.501         |
|                 | wQFM       | 119694341.30  | 194580000.0                   | 61.514         |
|                 | wQMC       | 119579409.30  | 194580000.0                   | 61.455         |
|                 | ASTRAL     | 119700568.10  | 194580000.0                   | 61.517         |
| 1X-1000g-true   | Model Tree | 136107789.65  | 194580000.0                   | 69.950         |
|                 | wQFM       | 136103523.40  | 194580000.0                   | 69.947         |
|                 | wQMC       | 135951451.45  | 194580000.0                   | 69.869         |
|                 | ASTRAL     | 136134360.40  | 194580000.0                   | 69.963         |
| 2X-1000g-true   | Model Tree | 148466575.75  | 194580000.0                   | 76.301         |
|                 | wQFM       | 148487134.80  | 194580000.0                   | 76.312         |
|                 | wQMC       | 148409168.15  | 194580000.0                   | 76.272         |
|                 | ASTRAL     | 148498766.90  | 194580000.0                   | 76.318         |
| 1X-50g-true     | Model Tree | 6789683.00    | 9729000.0                     | 69.788         |
|                 | wQFM       | 6839237.20    | 9729000.0                     | 70.297         |
|                 | wQMC       | 6853615.75    | 9729000.0                     | 70.445         |
|                 | ASTRAL     | 6861504.80    | 9729000.0                     | 70.526         |
| 1X-100g-true    | Model Tree | 13610324.80   | 19458000.0                    | 69.947         |
|                 | wQFM       | 13664518.95   | 19458000.0                    | 70.226         |
|                 | wQMC       | 13659704.75   | 19458000.0                    | 70.201         |
|                 | ASTRAL     | 13675471.15   | 19458000.0                    | 70.282         |
| 1X-200g-true    | Model Tree | 27229536.95   | 38916000.0                    | 69.970         |
|                 | wQFM       | 27270064.25   | 38916000.0                    | 70.074         |
|                 | wQMC       | 27242135.85   | 38916000.0                    | 70.002         |
|                 | ASTRAL     | 27287354.90   | 38916000.0                    | 70.119         |
| 1X-500g-true    | Model Tree | 68033826.75   | 97290000.0                    | 69.929         |
|                 | wQFM       | 68055804.00   | 97290000.0                    | 69.951         |
|                 | wQMC       | 67981642.70   | 97290000.0                    | 69.875         |
|                 | ASTRAL     | 68078813.70   | 97290000.0                    | 69.975         |

Table S5: **Quartet scores on 11-taxon dataset.** We show average (over 20 replicates) quartet scores (sum of the weights of the satisfied quartets) of the true and estimated trees, total weight of the quartets in the input gene trees, and their respective ratios.

| Model Condition    | Tree       | Quartet score | Total weight (input quartets) | Proportion (%) |
|--------------------|------------|---------------|-------------------------------|----------------|
| estimated-5genes   | Model Tree | 1321.30       | 1650.0                        | 80.079         |
|                    | wQFM       | 1354.90       | 1650.0                        | 82.115         |
|                    | wQMC       | 1355.05       | 1650.0                        | 82.124         |
|                    | ASTRAL     | 1357.40       | 1650.0                        | 82.267         |
| estimated-15genes  | Model Tree | 4042.90       | 4950.0                        | 81.675         |
|                    | wQFM       | 4067.15       | 4950.0                        | 82.165         |
|                    | wQMC       | 4068.25       | 4950.0                        | 82.187         |
|                    | ASTRAL     | 4068.25       | 4950.0                        | 82.187         |
| estimated-25genes  | Model Tree | 6679.45       | 8250.0                        | 80.963         |
|                    | wQFM       | 6693.30       | 8250.0                        | 81.131         |
|                    | wQMC       | 6695.40       | 8250.0                        | 81.156         |
|                    | ASTRAL     | 6695.40       | 8250.0                        | 81.156         |
| estimated-50genes  | Model Tree | 13362.85      | 16500.0                       | 80.987         |
|                    | wQFM       | 13383.00      | 16500.0                       | 81.109         |
|                    | wQMC       | 13383.10      | 16500.0                       | 81.110         |
|                    | ASTRAL     | 13383.10      | 16500.0                       | 81.110         |
| estimated-100genes | Model Tree | 26686.55      | 33000.0                       | 80.868         |
|                    | wQFM       | 26703.60      | 33000.0                       | 80.920         |
|                    | wQMC       | 26702.05      | 33000.0                       | 80.915         |
|                    | ASTRAL     | 26705.40      | 33000.0                       | 80.925         |
| true-5genes        | Model Tree | 1528.00       | 1650.0                        | 92.606         |
|                    | wQFM       | 1536.80       | 1650.0                        | 93.139         |
|                    | wQMC       | 1536.80       | 1650.0                        | 93.139         |
|                    | ASTRAL     | 1536.80       | 1650.0                        | 93.139         |
| true-15genes       | Model Tree | 4619.50       | 4950.0                        | 93.323         |
|                    | wQFM       | 4637.90       | 4950.0                        | 93.695         |
|                    | wQMC       | 4637.90       | 4950.0                        | 93.695         |
|                    | ASTRAL     | 4637.90       | 4950.0                        | 93.695         |
| true-25genes       | Model Tree | 7708.25       | 8250.0                        | 93.433         |
|                    | wQFM       | 7708.45       | 8250.0                        | 93.436         |
|                    | wQMC       | 7708.45       | 8250.0                        | 93.436         |
|                    | ASTRAL     | 7708.45       | 8250.0                        | 93.436         |
| true-50genes       | Model Tree | 15412.55      | 16500.0                       | 93.409         |
|                    | wQFM       | 15412.55      | 16500.0                       | 93.409         |
|                    | wQMC       | 15412.55      | 16500.0                       | 93.409         |
|                    | ASTRAL     | 15412.55      | 16500.0                       | 93.409         |
| true-100genes      | Model Tree | 30844.25      | 33000.0                       | 93.467         |
|                    | wQFM       | 30844.25      | 33000.0                       | 93.467         |
|                    | wQMC       | 30844.25      | 33000.0                       | 93.467         |
|                    | ASTRAL     | 30844.25      | 33000.0                       | 93.467         |



Table S6: **Quartet scores on 101-taxon simulated dataset.** We show average (over 50 replicates) quartet scores (sum of the weights of the satisfied quartets) of the true and estimated trees, total weight of the quartets in the input gene trees, and their respective ratios.

| Model Condition | Tree       | Quartet score | Total weight (input quartets) | Proportion (%) |
|-----------------|------------|---------------|-------------------------------|----------------|
| 1X-1000g-true   | Model Tree | 2732512131.34 | 4082925000.0                  | 66.925         |
|                 | wQFM       | 2732660047.02 | 4082925000.0                  | 66.929         |
|                 | wQMC       | 2728234405.08 | 4082925000.0                  | 66.821         |
|                 | ASTRAL     | 2732767772.08 | 4082925000.0                  | 66.932         |

Table S7: **Quartet scores of various methods on biological datasets.** We show the quartet scores (sum of the weights of the satisfied quartets) of various methods, total weight of the quartets in the input gene trees, and their respective ratios.

| Dataset    | Method | Quartet score  | Total weight (input quartets) | Proportion of total weight (%) |
|------------|--------|----------------|-------------------------------|--------------------------------|
| Amniota_aa | wQFM   | 83,604         | 1,25,412                      | 66.66                          |
|            | wQMC   | 83,604         |                               | 66.66                          |
|            | ASTRAL | 83,604         |                               | 66.66                          |
| Amniota_nt | wQFM   | 97,890         | 1,25,412                      | 78.06                          |
|            | wQMC   | 97,890         |                               | 78.06                          |
|            | ASTRAL | 97,890         |                               | 78.06                          |
| Mammal     | wQFM   | 2,55,26,915    | 2,80,03,080                   | 91.16                          |
|            | wQMC   | 2,55,26,915    |                               | 91.16                          |
|            | ASTRAL | 2,55,26,915    |                               | 91.16                          |
| Avian      | wQFM   | 1,22,87,19,113 | 2,46,21,11,516                | 49.91                          |
|            | wQMC   | 1,22,35,78,787 |                               | 49.70                          |
|            | ASTRAL | 1,23,19,92,828 |                               | 50.04                          |
| Angiosperm | wQFM   | 1,15,51,819    | 1,44,99,592                   | 79.67                          |
|            | wQMC   | 1,15,51,948    |                               | 79.67                          |
|            | ASTRAL | 1,15,53,053    |                               | 79.68                          |

## 6 Running time

We performed the experiments on a Linux machine with 8 GB RAM and i7 2.50 GHz processor. We ran the exact version of ASTRAL-III (version: 5.7.3) on smaller datasets (11 ~ 15 taxa), and used the heuristic version to analyze larger datasets. For wQFM and wQMC, we report the running time for amalgamating the weighted quartets (given as input), which excludes the time for computing the weighted quartets. We used a custom script for generating weighted quartets, by computing the frequency of each quartet, from a collection of gene trees. However, weight/confidence of a quartet can be generated in different ways, e.g., using the quartet frequency (as used in this study), likelihood of a quartet, and various algebraic and statistical approaches. Thus, the time required to generate weighted quartets may differ depending on what types of weights are being used.

For smaller datasets, these three methods wQFM, wQMC and ASTRAL took very small amounts of time. Both wQFM and wQMC took only a fraction of a second for 11 and 15-taxon datasets. ASTRAL also took around a second to analyze 11-taxon datasets. For 15-taxon datasets however, it took ASTRAL (exact version) 1 to  $\sim 12$  min (depending on various numbers of genes), which is much longer than wQFM and wQMC. Note that, since the input to wQFM and wQMC are weighted quartets embedded in the input gene trees, their running times are not much sensitive to the number of genes.

For 37-taxon dataset, heuristic version of ASTRAL was used, which led to much smaller running times per replicates, ranging from 2 to 6 s. The running time of ASTRAL decrease from 6 s to 2 s as we decrease the level of ILS from higher ILS (0.5X) to lower ILS (2X). This decrease in running time may be due to the fact that as the amount of discordance (due to ILS) decreases in the gene trees, the number of bipartitions in the gene trees also decrease. This leads to a smaller search space for ASTRAL (heuristic version) to explore. ASTRAL took around 1 s (on 50 genes) to 10 s (on 800 genes) to analyze various numbers of genes. wQMC was the fastest method which took only a second to analyze a single replicate of this dataset. wQFM was also fast, taking only 4 - 12 s.

On 48-taxon simulated avian dataset, wQMC took around 2 s. ASTRAL's running time ranges from 6 to 150 s, depending on various numbers of genes and ILS levels. wQFM took around 15 - 100 s per replicate on this avian simulated dataset. The most significant difference in running times was observed on the avian biological dataset with 14K gene trees, where ASTRAL took around 32 hours to run. wQMC and wQFM, on the other hand, finished within 2 and 20 s, respectively. This is due to the fact that ASTRAL's running time increases as we increase the number of genes (hence the number of unique bipartitions), but wQMC and wQFM takes as input a set of weighted quartets and thus their running times are not sensitive to the number of genes. On a relatively larger dataset with 101 taxa, the running time of wQFM ranges from 25 to 40 min. ASTRAL and wQMC were much faster, taking around 2 - 3 min and 5 s, respectively.

Note that the running time results for wQFM and wQMC observed here may not scale to relatively larger data sets with hundreds or thousands of taxa as the number of quartets grow rapidly as we increase the number of taxa (there are  $3\binom{n}{4}$  possible quartets on  $n$  taxa). Moreover, enumerating induced quartets in the input gene trees to compute gene tree frequency based weights is time consuming. For example, generating weighted quartets on 37-, 48- and 101-taxon data sets using our custom script takes 4 - 120s (for 25-800 genes), 50 s to  $\sim 4$  hours (for 50 - 14,446 genes), 2.5 - 3 hours (1000 genes) respectively. Note that generating weighted quartets from different gene trees can be done in parallel and hence, given a sufficient number of processors, the running time of enumerating quartets will not increase as we increase the number of genes. For example, generating weighted quartets in a gene tree of the avian dataset with 48 taxa takes only 1 - 2 s while it takes around 4 hours (14,400 s) when we *sequentially* consider 14,446 trees, which can be substantially reduced by using multiple processors. Moreover, wQFM is a divide-and-conquer based algorithm and the divide steps (on the set of taxa) are independent of each other (i.e., embarrassingly parallel). Therefore, the running time of wQFM can be substantially improved by implementing an efficient multi-threaded version (similar to ASTRAL-MP [12]).

## 7 Wilcoxon Signed-Rank Test

Tables S8 and S9 show the  $p$ -values using the Wilcoxon signed-rank test (with  $\alpha = 0.05$ ) on various model conditions of 11-, 15-, 37-, 48- and 101-taxon datasets, comparing the RF rates obtained by wQFM with those of wQMC and ASTRAL, respectively.

Table S8: Statistical significance of the differences in RF rates between wQFM and wQMC on 11-, 15-, 37-, 48- and 101-taxon datasets. The  $p$ -values indicating statistically significant differences (i.e.,  $p \leq 0.05$ ) are shown in bold.

| <i>Dataset</i>   | <i>Model condition</i> |                      |                  | <i>p-value</i> |
|------------------|------------------------|----------------------|------------------|----------------|
|                  | <i>No. of genes</i>    | <i>Sequence len.</i> | <i>ILS Level</i> |                |
| <b>11-taxon</b>  | 5 g                    | true                 | high             | 0.9569         |
|                  | 15 g                   | true                 | high             | 0.1573         |
|                  | 5 g                    | estimated            | high             | 0.5498         |
|                  | 15 g                   | estimated            | high             | 1.0            |
|                  | 25 g                   | estimated            | high             | 0.5637         |
|                  | 50 g                   | estimated            | high             | 0.3173         |
|                  | 100 g                  | estimated            | high             | 0.3173         |
| <b>15-taxon</b>  | 100 g                  | 100 bp               | high             | <b>0.047</b>   |
|                  | 100 g                  | 1000 bp              | high             | 0.8154         |
|                  | 1000 g                 | 100 bp               | high             | <b>0.0109</b>  |
|                  | 1000 g                 | 1000 bp              | high             | 0.1573         |
| <b>37-taxon</b>  | 25 g                   | 500 bp               | 1X               | 0.5772         |
|                  | 50 g                   | 500 bp               | 1X               | 0.4659         |
|                  | 100 g                  | 500 bp               | 1X               | 0.3766         |
|                  | 200 g                  | 500 bp               | 1X               | 0.3192         |
|                  | 200 g                  | 500 bp               | 0.5X             | <b>0.0272</b>  |
|                  | 200 g                  | 500 bp               | 2X               | 0.3173         |
|                  | 400 g                  | 500 bp               | 1X               | 0.1573         |
|                  | 800 g                  | 500 bp               | 1X               | 0.3173         |
|                  | 200 g                  | 250 bp               | 1X               | <b>0.0065</b>  |
|                  | 200 g                  | 1000 bp              | 1X               | 0.1573         |
|                  | 200 g                  | 1500 bp              | 1X               | 0.1651         |
|                  | 200 g                  | true                 | 1X               | 0.1797         |
| <b>48-taxon</b>  | 1000 g                 | 500 bp               | 0.5X             | <b>0.0001</b>  |
|                  | 1000 g                 | 500 bp               | 1X               | <b>0.0001</b>  |
|                  | 1000 g                 | 500 bp               | 2X               | <b>0.0001</b>  |
|                  | 1000 g                 | true                 | 0.5X             | <b>0.0001</b>  |
|                  | 1000 g                 | true                 | 1X               | <b>0.0005</b>  |
|                  | 1000 g                 | true                 | 2X               | <b>0.0147</b>  |
|                  | 50g                    | 500 bp               | 1X               | <b>0.0002</b>  |
|                  | 100g                   | 500 bp               | 1X               | <b>0.0002</b>  |
|                  | 200g                   | 500 bp               | 1X               | <b>0.0003</b>  |
|                  | 500g                   | 500 bp               | 1X               | <b>0.0001</b>  |
|                  | 50                     | true                 | 1X               | <b>0.0002</b>  |
|                  | 100                    | true                 | 1X               | <b>0.0002</b>  |
|                  | 200                    | true                 | 1X               | <b>0.0002</b>  |
|                  | 500                    | true                 | 1X               | <b>0.0007</b>  |
| <b>101-taxon</b> | 1000 g                 | true                 | 1X               | $\ll 10^{-4}$  |

Table S9: Statistical significance of the differences in RF rates between wQFM and AS-TRAL on 11-, 15-, 37-, 48- and 101-taxon datasets. The  $p$ -values indicating statistically significant differences (i.e.,  $p \leq 0.05$ ) are shown in bold.

| <i>Dataset</i>   | <i>Model condition</i> |                      |                  | <i>p-value</i> |
|------------------|------------------------|----------------------|------------------|----------------|
|                  | <i>No. of genes</i>    | <i>Sequence len.</i> | <i>ILS Level</i> |                |
| <b>11-taxon</b>  | 5 g                    | true                 | high             | 0.4142         |
|                  | 15 g                   | true                 | high             | 0.1573         |
|                  | 5 g                    | estimated            | high             | 0.5637         |
|                  | 15 g                   | estimated            | high             | 1.0            |
|                  | 25 g                   | estimated            | high             | 0.5637         |
|                  | 50 g                   | estimated            | high             | 0.3173         |
|                  | 100 g                  | estimated            | high             | 0.5637         |
| <b>15-taxon</b>  | 100 g                  | 100 bp               | high             | <b>0.0472</b>  |
|                  | 100 g                  | 1000 bp              | high             | 0.9121         |
|                  | 1000 g                 | 100 bp               | high             | <b>0.0109</b>  |
|                  | 1000 g                 | 1000 bp              | high             | 0.1573         |
| <b>37-taxon</b>  | 25 g                   | 500 bp               | 1X               | 0.8315         |
|                  | 50 g                   | 500 bp               | 1X               | 0.7827         |
|                  | 100 g                  | 500 bp               | 1X               | 0.3766         |
|                  | 200 g                  | 500 bp               | 1X               | 0.3192         |
|                  | 200 g                  | 500 bp               | 0.5X             | 0.2076         |
|                  | 200 g                  | 500 bp               | 2X               | 0.3173         |
|                  | 400 g                  | 500 bp               | 1X               | 0.1573         |
|                  | 800 g                  | 500 bp               | 1X               | 0.3173         |
|                  | 200 g                  | 250 bp               | 1X               | <b>0.0423</b>  |
|                  | 200 g                  | 1000 bp              | 1X               | 0.1573         |
|                  | 200 g                  | 1500 bp              | 1X               | 0.3173         |
|                  | 200 g                  | true                 | 1X               | 0.3173         |
| <b>48-taxon</b>  | 1000 g                 | 500 bp               | 0.5X             | <b>0.0042</b>  |
|                  | 1000 g                 | 500 bp               | 1X               | <b>0.0127</b>  |
|                  | 1000 g                 | 500 bp               | 2X               | 0.3872         |
|                  | 1000 g                 | true                 | 0.5X             | 0.6692         |
|                  | 1000 g                 | true                 | 1X               | 0.5859         |
|                  | 1000 g                 | true                 | 2X               | 0.5847         |
|                  | 50g                    | 500 bp               | 1X               | 0.1936         |
|                  | 100g                   | 500 bp               | 1X               | 0.1021         |
|                  | 200g                   | 500 bp               | 1X               | 0.0915         |
|                  | 500g                   | 500 bp               | 1X               | <b>0.0274</b>  |
|                  | 50                     | true                 | 1X               | 0.1356         |
|                  | 100                    | true                 | 1X               | <b>0.0203</b>  |
|                  | 200                    | true                 | 1X               | <b>0.048</b>   |
|                  | 500                    | true                 | 1X               | 0.5292         |
| <b>101-taxon</b> | 1000 g                 | true                 | 1X               | 0.3635         |

## References

- [1] Sen Song, Liang Liu, Scott V Edwards, and Shaoyuan Wu. Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. *Proceedings of the National Academy of Sciences*, 109(37):14942–14947, 2012.
- [2] Erich D Jarvis, Siavash Mirarab, Andre J Aberer, Bo Li, Peter Houde, Cai Li, Simon YW Ho, Brant C Faircloth, Benoit Nabholz, Jason T Howard, et al. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science*, 346(6215):1320–1331, 2014.
- [3] Y Chung and C Ané. Comparing two Bayesian methods for gene tree/species tree reconstruction: A simulation with incomplete lineage sorting and horizontal gene transfer. *Systematic Biology*, 60(3):261–275, 2011.
- [4] Md Shamsuzzoha Bayzid and Tandy Warnow. Naive binning improves phylogenomic analyses. *Bioinformatics*, 29(18):2277–2284, 2013.
- [5] Zhenxiang Xi, Liang Liu, Joshua S Rest, and Charles C Davis. Coalescent versus concatenation methods and the placement of amborella as sister to water lilies. *Systematic Biology*, 63(6):919–932, 2014.
- [6] Norman J Wickett, Siavash Mirarab, Nam Nguyen, Tandy Warnow, Eric Carpenter, Naim Matasci, Saravanaraj Ayyampalayam, Michael S Barker, J Gordon Burleigh, Matthew A Gitzendanner, et al. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences*, 111(45):E4859–E4868, 2014.
- [7] Ning Zhang, Liping Zeng, Hongyan Shan, and Hong Ma. Highly conserved low-copy nuclear genes as effective markers for phylogenetic analyses in angiosperms. *New Phytologist*, 195(4):923–937, 2012.
- [8] Siavash Mirarab and Tandy Warnow. Astral-ii: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics*, 31(12):i44–i52, 2015.
- [9] Bryan T Drew, Brad R Ruhfel, Stephen A Smith, Michael J Moore, Barbara G Briggs, Matthew A Gitzendanner, Pamela S Soltis, and Douglas E Soltis. Another look at the root of the angiosperms reveals a familiar tale. *Systematic Biology*, 63(3):368–382, 2014.
- [10] Vadim V Goremykin, Svetlana V Nikiforova, Patrick J Biggs, Bojian Zhong, Peter Delange, William Martin, Stefan Woetzel, Robin A Atherton, Patricia A McLenachan, and Peter J Lockhart. The evolutionary root of flowering plants. *Systematic Biology*, 62(1):50–61, 2013.
- [11] Julia Chifman and Laura Kubatko. Quartet from snp data under the coalescent model. *Bioinformatics*, 30(23):3317–3324, 2014.

- [12] John Yin, Chao Zhang, and Siavash Mirarab. ASTRAL-MP: scaling ASTRAL to very large datasets using randomization and parallelization. *Bioinformatics*, 35(20):3961–3969, 2019.