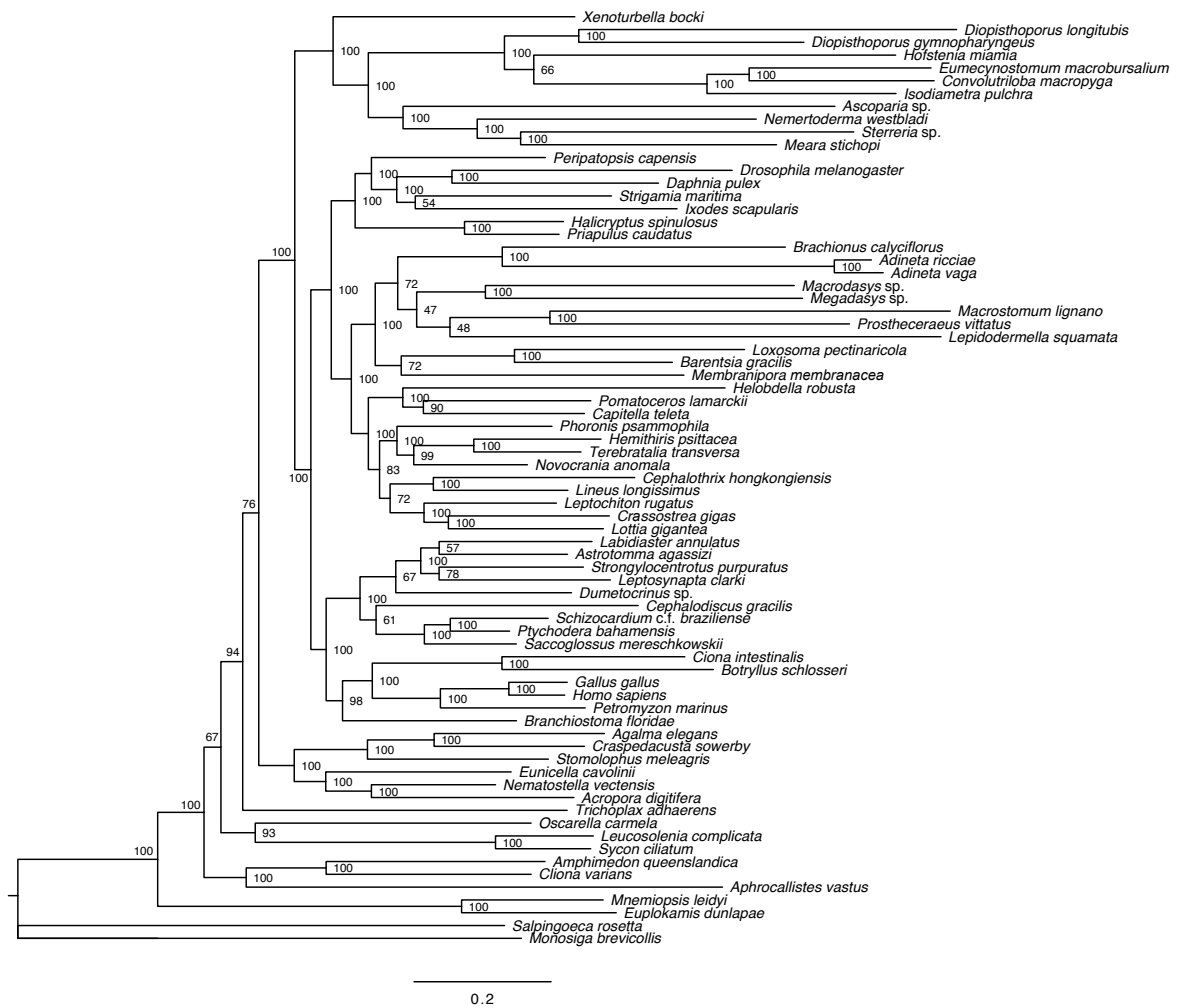
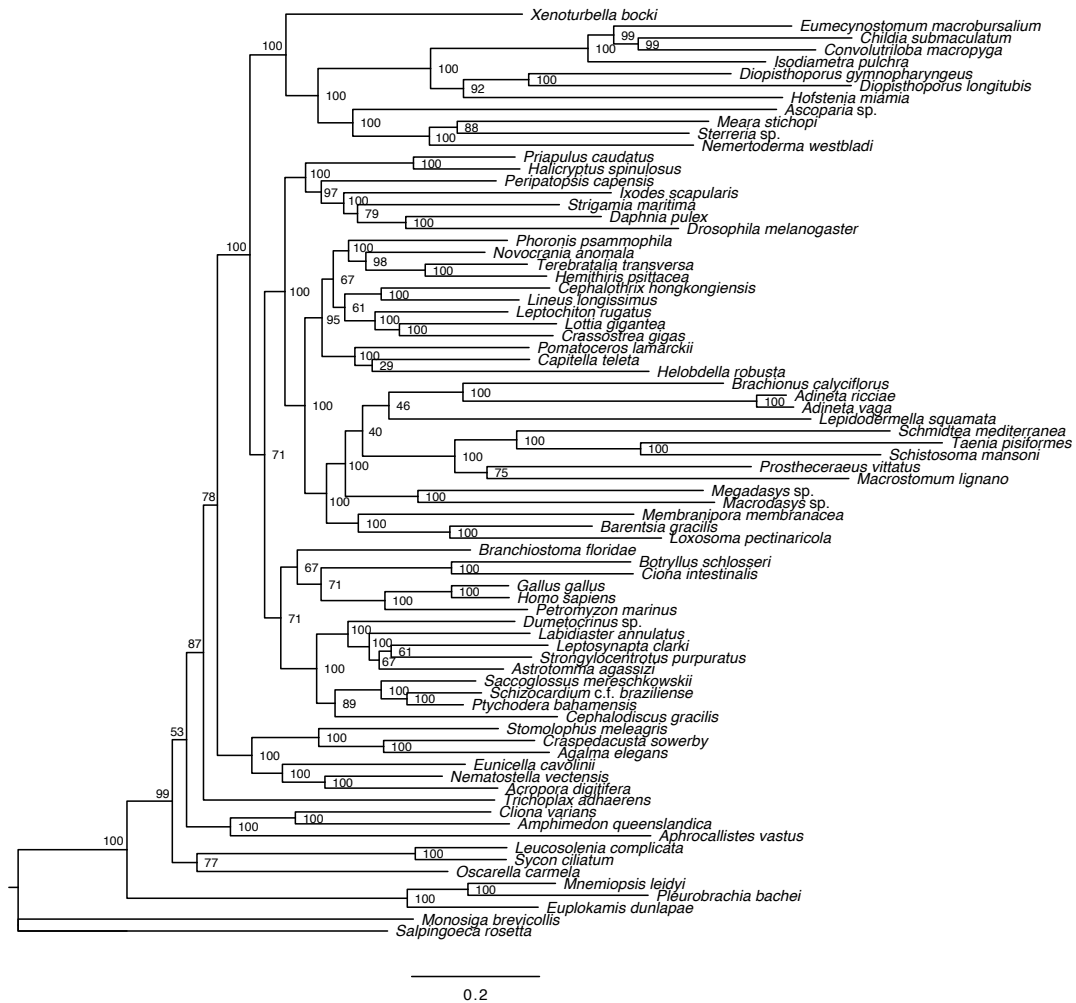


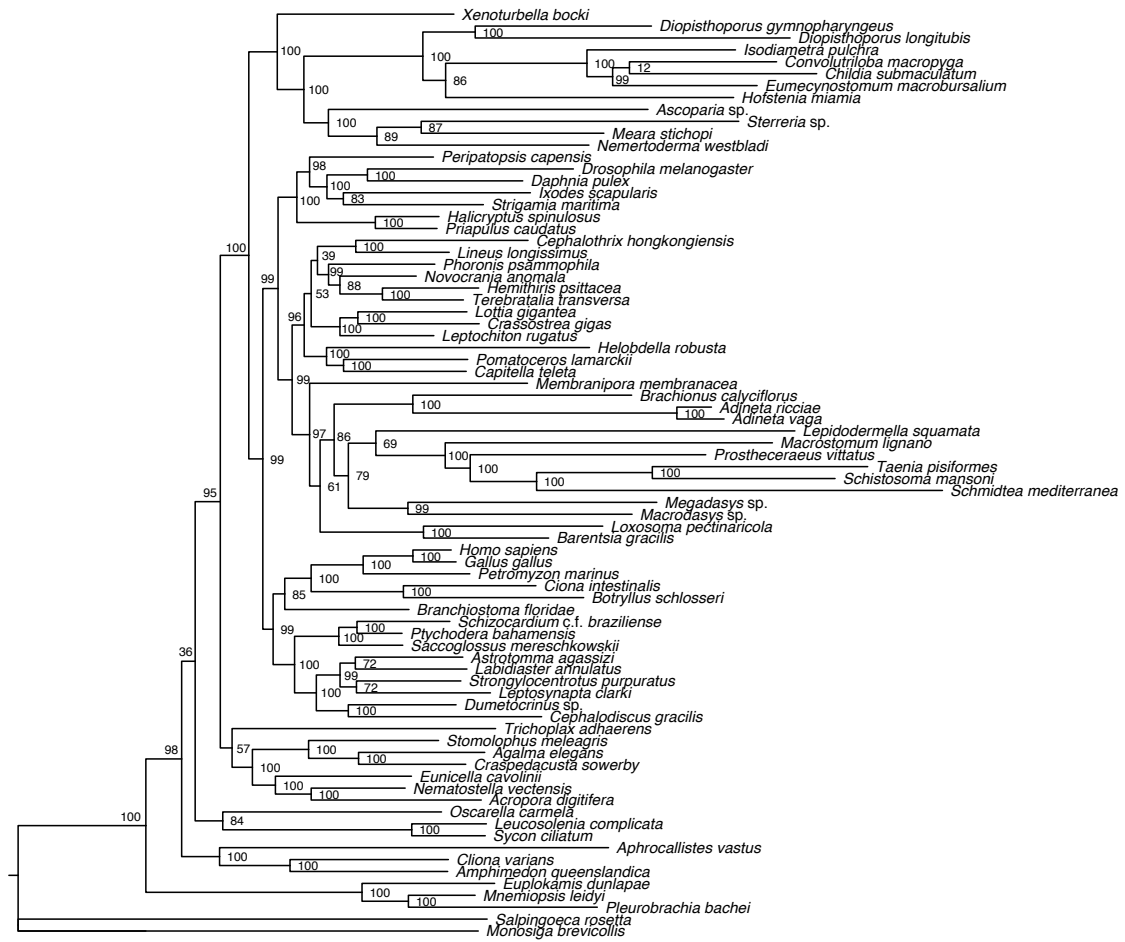
Supplementary Figure 14. Maximum likelihood topology based on the primary 212-gene matrix with taxa with LB scores (calculated using TreSpex) over 13 excluded. The most likely tree sampled in RAxML using the LG+I+ Γ substitution model for each gene is shown with ML bootstrap (bs) support values from 100 replicates indicated at each node. The length of the matrix is 43247 AAs, and overall matrix completeness is 70 percent.



Supplementary Figure 15. Maximum likelihood topology based on the primary 212-gene matrix with taxa with LB scores (calculated using TreSpex) over 30 excluded. The most likely tree sampled in RAxML using the LG+I+ Γ substitution model for each gene is shown with ML bootstrap (bs) support values from 100 replicates indicated at each node. The length of the matrix is 44260 AAs, and overall matrix completeness is 70 percent.



Supplementary Figure 16. Maximum likelihood topology of the 106-genes with the best LB scores as calculated by TreSpeX. The most likely tree sampled in RAXML using the LG+I+Γ substitution model for each gene is shown with ML bootstrap (bs) support values from 100 replicates indicated at each node. The length of the matrix is 22295 AAs, and overall matrix completeness is 71 percent.

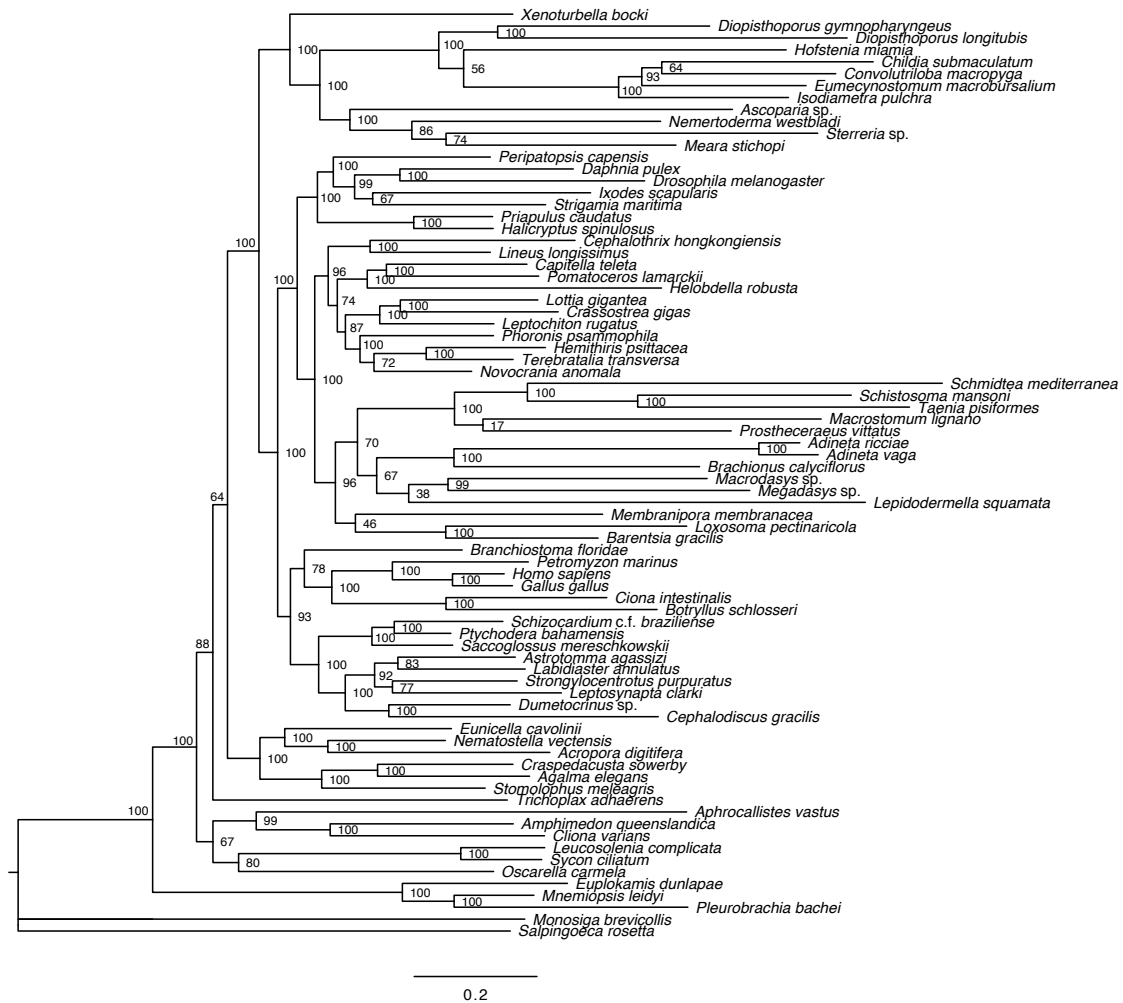


0.2

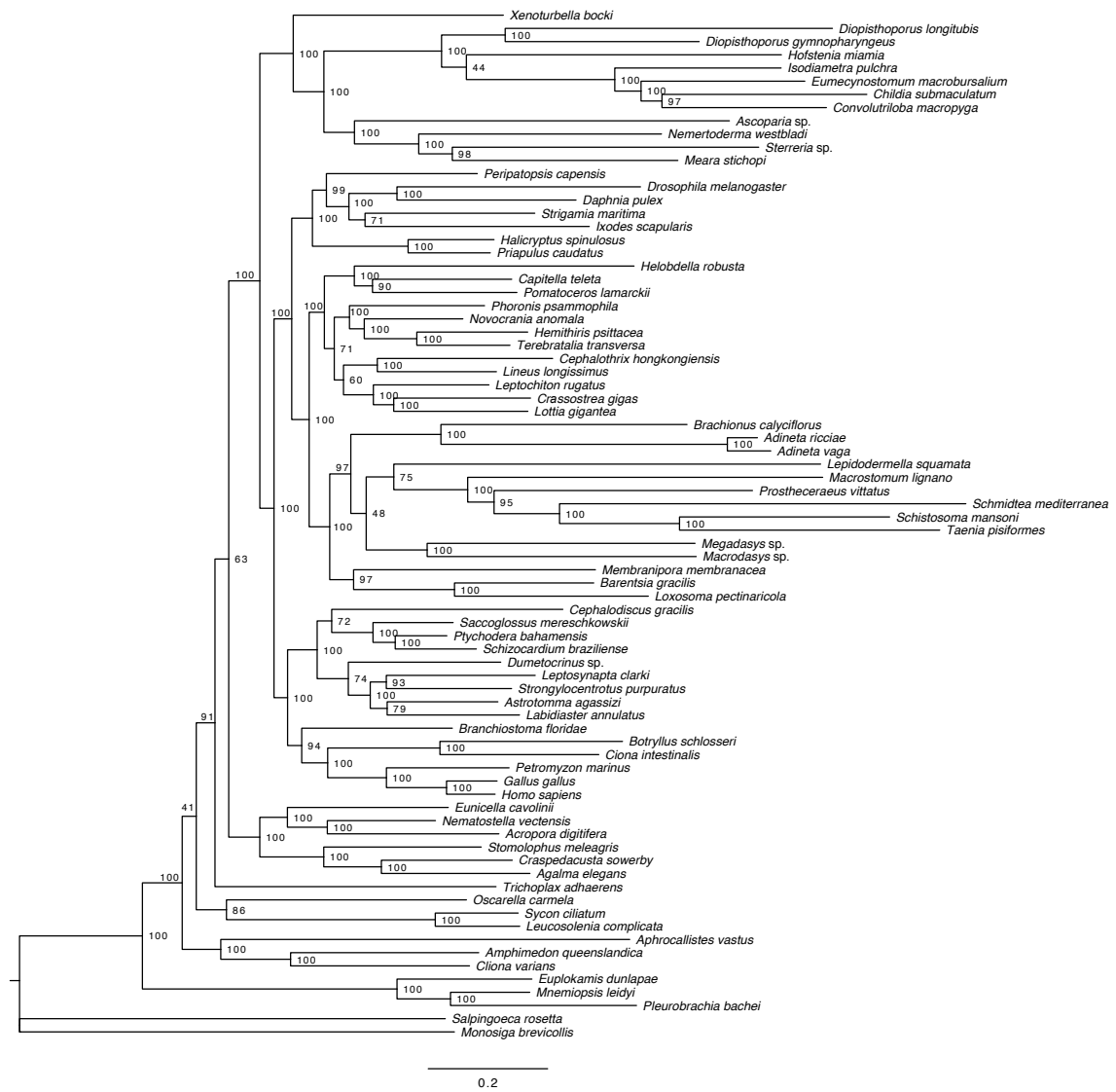
Supplementary Figure 17. Maximum likelihood topology of the 106-genes with the worst LB scores as calculated by TreSpeX. The most likely tree sampled in RAxML using the LG+I+Γ substitution model for each gene is shown with ML bootstrap (bs) support values from 100 replicates indicated at each node. The length of the matrix is 22601 AAs, and overall matrix completeness is 68 percent.



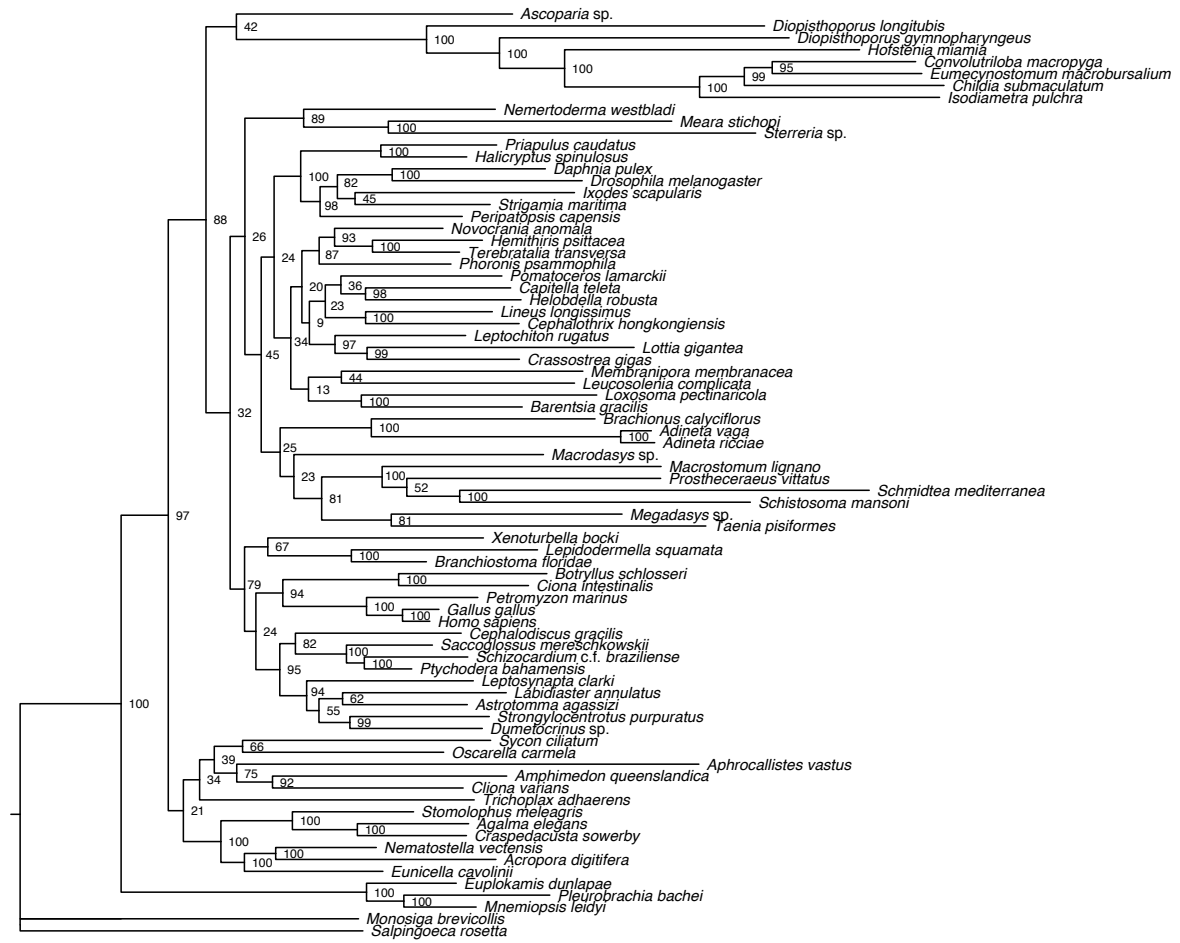
Supplementary Figure 18. Maximum likelihood topology of the 106 least saturated genes as calculated by TreSpeX. The most likely tree sampled in RAxML using the LG+I+ Γ substitution model for each gene is shown with ML bootstrap (bs) support values from 100 replicates indicated at each node. The length of the matrix is 23414 AAs, and overall matrix completeness is 71 percent.



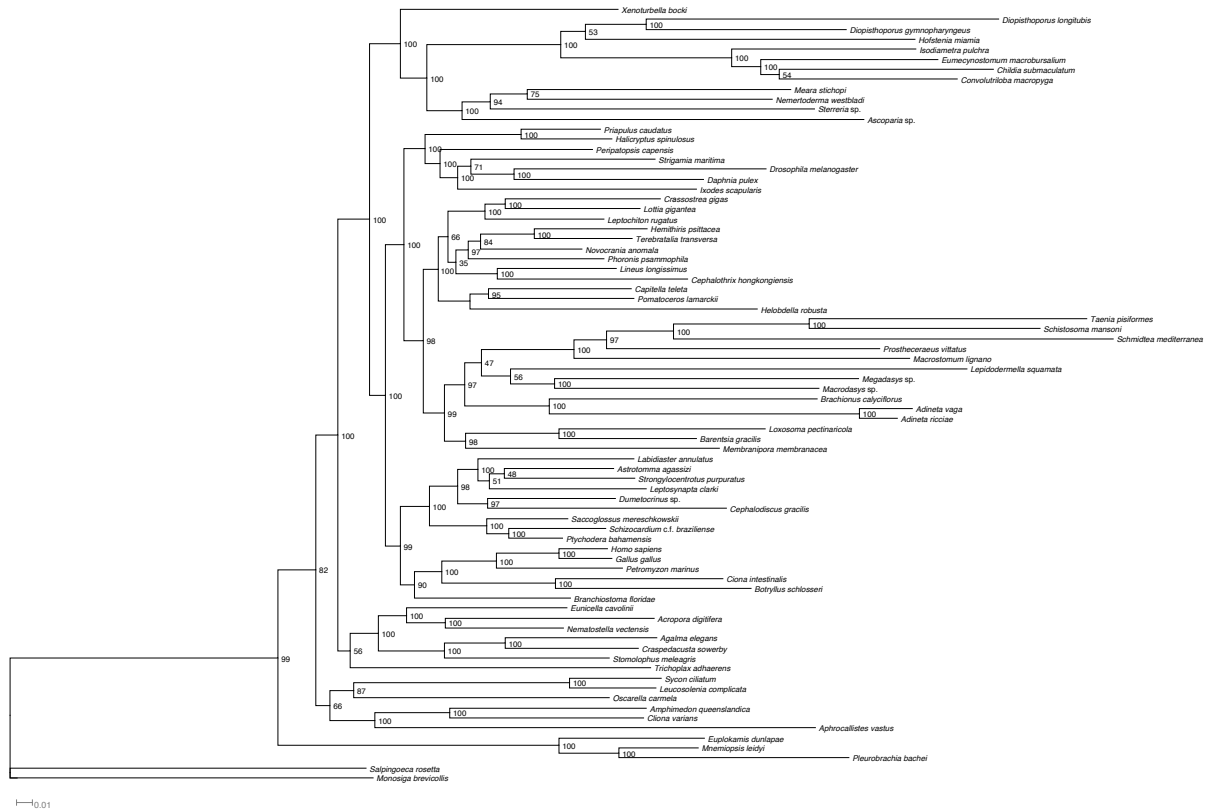
Supplementary Figure 19. Maximum likelihood topology of the 106 most saturated genes as calculated by TreSpeX. The most likely tree sampled in RAxML using the LG+I+ Γ substitution model for each gene is shown with ML bootstrap (bs) support values from 100 replicates indicated at each node. The length of the matrix is 21482 AAs, and overall matrix completeness is 66 percent.



Supplementary Figure 20. Maximum likelihood topology of 207 non-ribosomal protein coding genes. The most likely tree sampled in RAxML using the LG+I+ Γ substitution model for each gene is shown with ML bootstrap (bs) support values from 100 replicates indicated at each node. The length of the matrix is 44715 AAs, and overall matrix completeness is 68 percent.



Supplementary Figure 21. Maximum likelihood topology of 53 ribosomal protein genes. The most likely tree sampled in RAxML using the LG+I+ Γ substitution model for each gene is shown with ML bootstrap (bs) support values from 100 replicates indicated at each node. The length of the matrix is 9010 AAs, and overall matrix completeness is 81 percent.



Supplementary Figure 22. Maximum likelihood topology of BMGE-trimmed data matrix. The most likely tree sampled in RAxML using the LG4X substitution model for each gene is shown with ML bootstrap (bs) support values from 100 replicates indicated at each node. The length of the matrix is 33323 AAs, and overall matrix completeness is 66 percent.