Supplementary Figure Legends

Figure A5.1. Phylogram depicting the relationship among 33 *Pimelea* taxa and eight outgroup taxa, as estimated using maximum-likelihood analysis of 75 protein-coding genes in IQ-TREE. The dataset was partitioned by the optimal scheme identified using the ModelFinder option of IQ-TREE, and support is indicated for each node in the form of SH-aLRT/ultrafast bootstrap support. The scale is in estimated substitutions per site.

Figure A5.2. Phylogram depicting the relationship among 33 *Pimelea* taxa and eight outgroup taxa, as estimated using Bayesian inference of 75 protein-coding genes in MrBayes. The dataset was partitioned by the optimal scheme identified using the ModelFinder option of IQ-TREE, and support is indicated for each node in the form of posterior probabilities. The scale is in estimated substitutions per site.

Figure A5.3. Phylogram depicting the relationship among 33 *Pimelea* taxa and eight outgroup taxa, as estimated using maximum-likelihood analysis of 59 non-coding molecular markers in IQ-TREE. The dataset was partitioned by the optimal scheme identified using the ModelFinder option of IQ-TREE, and support is indicated for each node in the form of SH-aLRT/ultrafast bootstrap support. The scale is in estimated substitutions per site.

Figure A5.4. Phylogram depicting the relationship among 33 *Pimelea* taxa and eight outgroup taxa, as estimated using Bayesian inference of 59 non-coding molecular markers in MrBayes. The dataset was partitioned by the optimal scheme identified using the ModelFinder option of IQ-TREE, and support is indicated for each node in the form of posterior probabilities. The scale is in estimated substitutions per site.

Figure A5.5. Phylogram depicting the relationship among 270 taxa from Thymelaeaceae and one outgroup taxon from Malvaceae, as estimated using maximum-likelihood analysis of three protein-coding genes in IQ-TREE. The dataset was partitioned with a separate GTR+Γ substitution model for each codon position and for the non-coding marker, and support is indicated for each node in the form of SH-aLRT/ultrafast bootstrap support. The scale is in estimated substitutions per site. Taxon names with the “F16” suffix refer to those generated in Chapter 5; those with the “OGB” suffix are all others sourced from GenBank.

Figure A5.6. Phylogram depicting the relationship among 270 taxa from Thymelaeaceae and one outgroup taxon from Malvaceae, as estimated using Bayesian inference of three protein-coding genes in MrBayes. The dataset was partitioned with a separate GTR+Γ substitution model for each codon position and for the non-coding marker, and support is indicated for each node in the form of posterior probabilities. The scale is in estimated substitutions per site. Taxon names with the “F16” suffix refer to those generated in Chapter 5; those with the “OGB” suffix are all others sourced from GenBank.

Figure A5.7. The results of a topology-clustering analysis in which the Robinson-Foulds distance between gene trees was calculated, and then clustered with the CLARA clustering algorithm. (a) According to the gap statistic, there are three clusters of tree topologies in the chloroplast genes of *Pimelea*. (b) A multi-dimensional scaling plot showing the three clusters of gene tree topologies. The black circles correspond to Cluster 1, the red circles correspond to Cluster 2, and the green circles correspond to Cluster 3.

Figure A5.8. Boxplot comparing the lengths of gene trees in the three clusters identified by our topology-clustering analysis. Filled circles correspond to the 95% confidence intervals. The asterisk indicates a significant difference in tree length between Cluster 1 and the other two clusters.

Figure A5.9. Boxplot comparing the mean GC content of genes in the three clusters identified by our topology-clustering analysis. Notches in the boxes correspond to the 95% confidence intervals. The asterisk indicates a significant difference in GC content between Cluster 2 and the other two clusters.

Figure A5.10. Boxplot comparing the ratio of nonsynonymous to synonymous substitution rates for each gene in the three clusters identified by our topology-clustering analysis. Filled circles correspond to the 95% confidence intervals. There was no significant difference between clusters.

Figure A5.11. Phylogram depicting the relationship among 27 *Pimelea* taxa, as estimated using maximum-likelihood analysis of 11 protein-coding genes in IQ-TREE. The genes included in this analysis are referred to as 'Cluster 1' in this study, as assigned in a topology-clustering analysis. The dataset was partitioned by the optimal scheme identified using the ModelFinder option of IQ-TREE, and support is indicated for each node in the form of SH-aLRT/ultrafast bootstrap support. The scale is in estimated substitutions per site.

Figure A5.12. Phylogram depicting the relationship among 27 *Pimelea* taxa, as estimated using Bayesian inference of 11 protein-coding genes in MrBayes. The genes included in this analysis are referred to as 'Cluster 1' in this study, as assigned in a topology-clustering analysis. The dataset was partitioned by the optimal scheme identified using the ModelFinder option of IQ-TREE, and support is indicated for each node in the form of posterior probabilities. The scale is in estimated substitutions per site.

Figure A5.13. Phylogram depicting the relationship among 27 *Pimelea* taxa, as estimated using maximum-likelihood analysis of 18 protein-coding genes in IQ-TREE. The genes included in this analysis are referred to as 'Cluster 2' in this study, as assigned in a topology-clustering analysis. The dataset was partitioned by the optimal scheme identified using the ModelFinder option of IQ-TREE, and support is indicated for each node in the form of SH-aLRT/ultrafast bootstrap support. The scale is in estimated substitutions per site.

Figure A5.14. Phylogram depicting the relationship among 27 *Pimelea* taxa, as estimated using Bayesian inference of 18 protein-coding genes in MrBayes. The genes included in this analysis are referred to as 'Cluster 2' in this study, as assigned in a topology-clustering analysis. The dataset was partitioned by the optimal scheme identified using the ModelFinder option of IQ-TREE, and support is indicated for each node in the form of posterior probabilities. The scale is in estimated substitutions per site.

Figure A5.15. Phylogram depicting the relationship among 27 *Pimelea* taxa, as estimated using maximum-likelihood analysis of 24 protein-coding genes in IQ-TREE. The genes included in this analysis are referred to as 'Cluster 3' in this study, as assigned in a topology-clustering analysis. The dataset was partitioned by the optimal scheme identified using the ModelFinder option of IQ-TREE, and support is indicated for each

node in the form of SH-aLRT/ultrafast bootstrap support. The scale is in estimated substitutions per site.

Figure A5.16. Phylogram depicting the relationship among 27 *Pimelea* taxa, as estimated using Bayesian inference of 24 protein-coding genes in MrBayes. The genes included in this analysis are referred to as 'Cluster 3' in this study, as assigned in a topology-clustering analysis. The dataset was partitioned by the optimal scheme identified using the ModelFinder option of IQ-TREE, and support is indicated for each node in the form of posterior probabilities. The scale is in estimated substitutions per site.

Figure A5.17. Chronogram depicting the evolutionary timescale of 33 *Pimelea* taxa and eight outgroup taxa, as estimated using Bayesian inference of 134 molecular markers (75 protein-coding and 59 non-coding) in MCMCTree. We used four data partitions, one for each codon position of the protein-coding genes and one for the non-coding data, and implemented a maximum age of 126.7 Ma.

Figure A5.18. Chronogram depicting the evolutionary timescale of 33 *Pimelea* taxa and eight outgroup taxa, as estimated using Bayesian inference of 134 molecular markers (75 protein-coding and 59 non-coding) in MCMCTree. We used four data partitions, one for each codon position of the protein-coding genes and one for the non-coding data, and implemented a conservative maximum age of 77.61 Ma.