

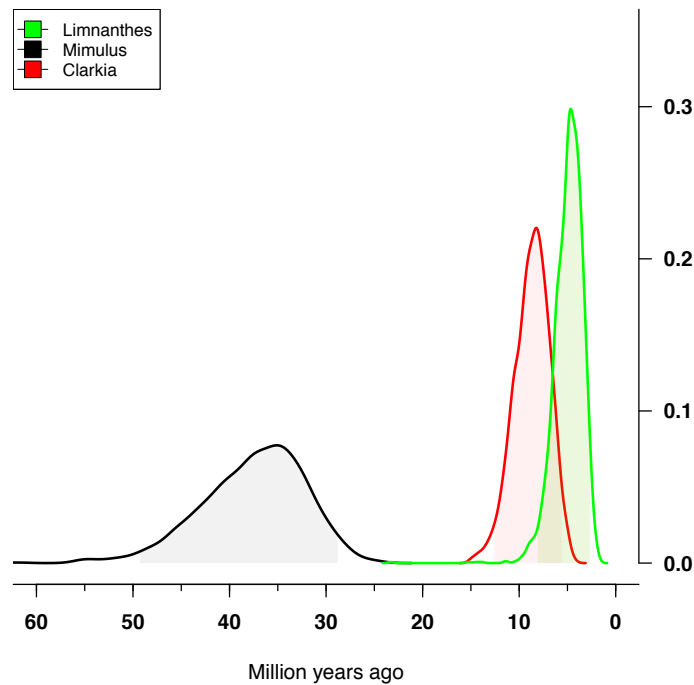
Estimating clade ages

This work was published in *Journal of Evolutionary Biology* by Runquist et al. (2016). Note that here *Mimulus* = Phrymaceae.

Because *Mimulus* has a worldwide distribution while *Clarkia* and *Limnanthes* are restricted to western North America, we hypothesized that *Mimulus* may be older than either *Clarkia* or *Limnanthes*. To test this hypothesis we estimated the time since the most recent common ancestor of all species in our study region for each clade, i.e. clade age.

Using the methods outlined below, we found that time since the most recent common ancestor for *Mimulus* was estimated to be more than twice that of either *Clarkia* or *Limnanthes* (95% highest posterior density: *Mimulus* 28.6 – 49.8 Ma, *Clarkia* 5.5 – 12.5 Ma, and *Limnanthes* 2.7 – 8.1 Ma).

Figure 1. Estimated clade age for *Limnanthes*, *Mimulus*, and *Clarkia*. The shaded areas indicate 95% highest posterior densities, estimated using kernel density functions.



Methods:

Fossils are not known for these clades, therefore, to estimate clade ages we used the mean and range nrITS substitution rate for herbaceous plants (mean = 4.13×10^{-9} subs/site/yr; range $1.72\text{--}8.34 \times 10^{-9}$ substitution/site/year) (Kay et al 2006). Molecular clock analysis of the ages and confidence intervals was performed using BEAST v1.8.0 (Drummond and Rambaut 2007) using the following nrITS data for each clade:

<i>Genus</i>	<i>nrITS base pair number</i>	<i>Substitution model</i>
<i>Limnanthes</i>	679	GTR + gamma model
<i>Mimulus</i>	654	GTR + gamma model
<i>Clarkia</i>	609	GTR + gamma model

To cover the nrITS substitution rate range (Kay et al 2006), the substitution rate was set to a normally distributed prior with a mean of 4.13×10^{-9} subs/site/yr and a standard deviation of 1.8077946. To accommodate heterogeneity in the molecular evolutionary rate among branches, we used an uncorrelated log-normal (UCLN) relaxed clock model. We used a Yule process prior on branching times and conducted four runs, each consisting of 20 million MCMC generations. Posterior samples of parameter values were summarized and assessed for convergence and mixing using Tracer v. 1.5 (included as a supplement) and we combined posterior distributions for all runs, resampling at a lower frequency using LogCombiner v. 1.5.3 (<http://beast.bio.ed.ac.uk/LogCombiner>). This was performed on each clade separately. We thank Peter Biro and Justen Whittall for advice with this analysis.

To determine whether the three clades differ in their ages, we used a kernel density function to estimate the probability density function of clade ages (i.e. time since the most recent common ancestor of all species in our study region) using a sample of 2222 trees from the posterior distribution of trees and assessed whether the 95% confidence intervals were overlapping among clades.

REFERENCES

- Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC evolutionary biology*, 7(1), 214.
- Kay, K. M., Whittall, J. B., & Hodges, S. A. (2006). A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects. *BMC Evolutionary Biology*, 6(1), 36.
- Briscoe Runquist, R., D. Grossenbacher, S. Porter, K. Kay, and J. Smith. 2016. Pollinator-mediated assemblage processes in California wildflowers. *Journal of Evolutionary Biology* 29: 1045–1058.