

Uncovering the origin and movement of serpentine-adapted alleles in monkeyflowers

Background: Understanding the genetic basis of adaptation is a longstanding goal of evolutionary biology. However, the relative importance of standing variation, *de novo* mutations, and introgression in creating the adaptive variation that allows plants to colonize harsh environments has remained elusive. Additionally, as humans rapidly alter climate and soil nutrient levels, it is essential to disentangle the proximate sources and relative timescales of polygenic adaptation. Tracking the origins and movements of adaptive alleles across complex natural environments in closely related species can be used to elucidate the importance and timing of each process.

Serpentine is a globally occurring, patchily distributed, ultramafic soil type characterized by low nutrient content and phytotoxic levels of metals^{1,2}. Serpentine soils are also often extensively weathered, dry, and high temperature^{1,2}. Many plants have evolved polygenic adaptations to serpentine¹, with varied allelic origins. For example, genome scans of *Arabidopsis arenosa* identified dozens of alleles highly associated with serpentine-tolerance, some of which were recently introgressed from *A. lyrata*, while others originated independently in *A. arenosa*².

Four related monkeyflower species within the genus *Mimulus* (Phrymaceae) occur on serpentine soils in California, creating a natural experiment for testing the relative importance and repeatability of different proximate sources of adaptive variation. Two species, *M. guttatus* and *M. glaucescens*, occur both on and off serpentine soils; however, individuals from off-serpentine backgrounds do not survive on serpentine, suggesting a genetic basis of adaptation³. In contrast, *M. nudatus* and *M. pardalis* are strictly endemic to serpentine patches in California's Coastal Range and in the foothills of the Sierra Nevada, respectively (Fig. 1). The small genome (430Mb) of *M. guttatus* is assembled and annotated. Gene flow within *M. guttatus* is widespread⁴, and crosses between serpentine-tolerant species are possible, although pre- and post-zygotic barriers to hybridization exist between some pairings⁵. Common garden experiments indicate that on-serpentine *M. guttatus* plants flower earlier, take up more calcium, and display increased drought tolerance^{1,6}. A locus on chromosome 13 of *M. guttatus* is highly associated with serpentine survival³, and selective sweep analyses highlight serpentine-associated differentiation at 26 loci spread throughout the genome. However, the genetic basis of serpentine-tolerance in other *Mimulus* species and the history of each allele are poorly understood.

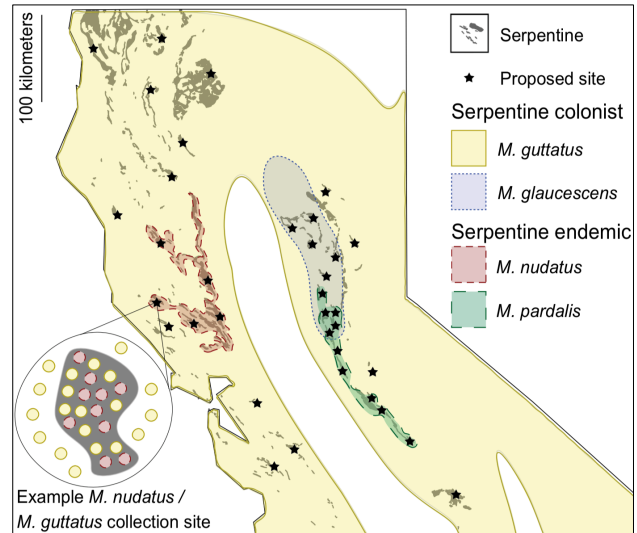


Fig. 1: Serpentine-tolerant *Mimulus* species ranges in California and proposed collection sites.

Aim 1: Identify the genetic basis of serpentine-tolerance in *Mimulus*. Because strong selective pressures are present across serpentine sites, I hypothesize that 1) adaptation to serpentine in *Mimulus* is polygenic, and 2) some large effect loci are common across *Mimulus* species. I will collect tissue and seed from 1050 *Mimulus* individuals distributed across 35 distinct serpentine patches (Fig. 1). At each site, I will sample 10 on- and 10 off-serpentine *M. guttatus*. Depending on the site's location, I will also collect from 10 on- and 10 off-serpentine *M. glaucescens* and/or from 10 serpentine endemic *Mimulus* individuals. I will conduct low-coverage (~5×) whole genome resequencing for each species and align sequences to the

annotated *M. guttatus* assembly. I will also collect soil from each site to confirm a serpentine origin (Calcium:Magnesium < 0.35) and to characterize chemical differences². To identify the genetic basis of serpentine tolerance in *M. nudatus* and *M. pardalis*, I will create hybrid mapping populations with crosses to local off-serpentine *M. guttatus* (or, if this fails, crosses to serpentine intolerant *M. marmoratus*, which is fully compatible with *M. pardalis*⁶). I will transplant F₂ seedlings to serpentine and non-serpentine sites in the field, sequence survivors, and scan for ancestry distortion QTLs³. For *M. guttatus* and *M. glaucescens*, I will conduct genome-wide tests of serpentine soil association, leveraging comparisons between local on- and off-serpentine conspecifics to control for population structure without swamping adaptive differences between patches. I will test if regions are under selection by calculating Tajima's D, F_{ST}, and π , treating common outliers as candidate regions². I will use the annotated *M. guttatus* genome to identify putative functions of these genes. My preliminary findings from 13 *M. guttatus* populations indicate that differentiated regions are enriched in sequences associated with drought stress, ion channels, and cofactor binding sites. Finding similar regions under selection in other *Mimulus* species would indicate that evolution of serpentine-tolerance is constrained to a few pathways.

Aim 2: Characterize the origin timing and movement of serpentine-tolerant alleles. Because reproductive isolation is incomplete between serpentine colonizers^{5,6}, and serpentine represents a strong selective pressure², I hypothesize that 1) some parallelism results from introgression. To compare allelic ages, I will quantify haplotype sizes in putatively serpentine adapted regions in *M. guttatus* and *M. glaucescens*. Short haplotypes and evidence of soft selective sweeps suggest that adaptive alleles arose from standing variation, while large haplotypes and hard selective sweeps provide evidence for a recent, *de novo* origin or introgression. To test for introgression, I will create gene trees for regions associated with serpentine colonization. I will distinguish introgression from incomplete lineage sorting by calculating D and F₃ statistics, and by using coalescent simulations^{2,7}. To further quantify timing and levels of gene flow, I will build replicated population trees with variable migration nodes to see which is best supported by the data^{2,7}. Finding repeated instances of shared alleles with a common origin in multiple *Mimulus* species would indicate that introgression is a powerful source of adaptive variation.

Intellectual Merit: This study will uncover the extent to which evolutionary trajectories are parallel and repeatable, as well as the relative contribution of standing variation, *de novo* mutations, and introgression to adaptive variation. In California, 45% of endemic serpentine taxa are listed as threatened¹, and understanding the genetic basis will allow managers to make informed decisions about species translocations and genetic rescue. Alleles identified in *Mimulus* can also inform crop breeding to cope with water stress, soil quality degradation, and thermal extremes.

Broader Impacts: I plan to recruit student research assistants from Napa Valley College (NVC), a Hispanic-serving, two-year institution near serpentine-adapted *M. guttatus* and *M. nudatus*. I will build on my previous mentorship at UC Berkeley to provide well-rounded, interdisciplinary experiences to NVC students that include sampling, sequencing, analysis, and data presentation. I plan to collaborate with the NVC students, leveraging their connections to the Napa Valley community, to create evolutionary biology lessons for local high school students. Specifically, I will facilitate in-class experiments attempting to grow off-serpentine *Mimulus* on serpentine soil by adding different nutrient combinations. These activities will frame serpentine-adapted plants as a powerful example of evolutionary processes occurring in these students' own backyards.

References: (1) Anacker *et al.* 2011. *Evolution*, 65: 365-376. (2) Arnold *et al.* 2016. *PNAS*, 113: 8320-8325. (3) Selby *et al.* 2018. *Mol Ecol*, 23: 5073-5087. (4) Puzey *et al.* 2017. *Mol Ecol*, 26: 519-535. (5) Vickery 1964. *Evolution*, 18: 52-69. (6) Hughes, *et al.* 2001. *S Afr J Sci*, 97: 581-586. (7) Chen *et al.* 2018. *Nat Commun*, 9: 2337.