



Genome Resources

The reference genome of a Sierra Nevada endemic, the cut-leaved monkeyflower, *Mimulus laciniatus* (syn. *Erythranthe lacinata*)

Jesús Martínez-Gómez^{1,2,†} , Merly Escalona^{3,†} , Jack M. Colicchio¹, Lauren N. Hamm¹, Mohan P.A. Marimuthu⁴, Oanh Nguyen⁴, Noravit Chumchim⁴, William Seligmann⁵, Rachel S. Meyer⁵, Jason P. Sexton⁶, Benjamin K. Blackman^{1,2,*} 

¹Department of Plant & Microbial Biology, University of California Berkeley, Berkeley, CA, United States, ²Department of Integrative Biology, University of California Berkeley, Berkeley, CA, United States, ³Department of Biomolecular Engineering, University of California Santa Cruz, Santa Cruz, CA, United States, ⁴DNA Technologies and Expression Analysis Core Laboratory, Genome Center, University of California, Davis, CA, United States, ⁵Department of Ecology & Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA, United States, ⁶Department of Life and Environmental Sciences, University of California Merced, Merced, CA, United States

*Corresponding author: Department of Plant & Microbial Biology, University of California Berkeley, 111 Koshland Hall #3102, Berkeley, CA 94609, United States. Email: bblackman@berkeley.edu

†Authors contributed equally.

Abstract

Mimulus laciniatus (syn. *Erythranthe lacinata*) is an annual plant endemic to the Sierra Nevada region of California. *M. laciniatus* is notable for its specialized ecological niche, thriving in granite outcrops of alpine environments characterized by shallow soils that dry out rapidly as the snowpack is exhausted during season-ending droughts. Due to its narrow habitat range and sensitivity to environmental change, this species serves as an important model for studying adaptation and survival in marginal habitats. As part of the California Conservation Genomics Project, here we report the sequencing and assembly of a high-quality nuclear genome and chloroplast genome of *M. laciniatus*. The primary assembly is 309.97 Mb and consists of 104 scaffolds with a scaffold N50 of 20.99 Mb, a largest contig size of 24.29 Mb, and a contig N50 of 11.09 Mb. The alternate haplotype assembly consists of 194 scaffolds spanning 213.84 Mb. BUSCO completeness of the primary assembly is 98.6%. This high quality genome adds a valuable resource to the expanding collection of sequenced genomes of the monkeyflowers (*Mimulus* sensu lato), which have become a model clade for studying ecological adaptation, speciation, and evolutionary genetics.

Key words: California Conservation Genomics Project, California endemic plants, CCGP, conservation, monkeyflower, speciation

Introduction

A foundational goal in evolutionary biology, ecology, and conservation biology is to understand the processes that cause lineages to speciate, persist, and go extinct. The neoendemic plants of the California floristic province (CA-FP) are an excellent system to study the processes that create, maintain, and diminish biodiversity. The CA-FP is the only member of the 36 globally recognized biodiversity hotspots (Mittermeier et al. 2011) located in the continental United States and Canada (reviewed in Baldwin 2014). Endemic species—i.e. species with a distribution restricted to a particular geographic range—are an important contributor to the biodiversity of the CA-FP. Of the 6143 minimally ranked vascular plant taxa (i.e. species, subspecies, varieties) native to the CA-FP, 42% are endemic to the CA-FP (Baldwin and Goldman 2012, Burge et al. 2016). Furthermore, more than half of the endemics (63.6%) have an annual life history (Raven and Axelrod 1974). Endemic plants of California have been used as models to understand speciation, plasticity, and local adaptation

(Hodges et al. 2004, Hufford and Mazer 2012, Pettengill and Moeller 2012, Gould et al. 2014, Sobel 2014, Hernández and Specht 2023), and understanding the mechanisms that shape diversity is necessary to develop conservation plans for endemics.

The yellow monkeyflower species complex *Mimulus* sect. *Simiolus* (syn. *Erythranthe* sect. *Simiola*) and the larger clade of monkeyflowers as a whole (reviewed in Yuan 2019), have long served as model lineages to study the evolutionary and ecological processes that drive adaptation and speciation (Vickery 1966, Wu et al. 2008, Twyford et al. 2015). The yellow monkeyflower clade has drawn interest due to its striking variation, including the colonization of diverse stressful edaphic environments (MacNair 1983, MacNair 1993, Lekberg et al. 2012, Sianta and Kay 2019), mating system evolution (Fenster and Ritland 1994, Martin and Willis 2007, Fishman et al. 2015), and transitions between annual and perennial life history strategies (Lowry et al. 2008, Coughlan et al. 2021). Members of this group are

Received on 22 May 2025; revised on 18 August 2025; accepted on 26 August 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of The American Genetic Association.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.



Fig. 1. *Mimulus laciniatus* and its growth habit at Olmstead point. A) *M. laciniatus* inflorescence, displaying dissected leaves (photo credit: Benjamin Blackman). B) Source population of sequenced individual OPN: Olmstead point in Yosemite National Park, Mariposa County CA (photo credit: Diana Tataru).

particularly well suited for the study of narrowly distributed taxa, especially those endemic to the CA-FP. About half of the ca. 50 minimally ranked monkeyflower taxa in sect. *Simiolus* are narrowly distributed (Nesom 2014), and 9 of these are endemic to CA-FP (Table S1).

Mimulus laciniatus A. Gray (syn. *Erythranthe laciniata* (A. Gray) G.L. Nesom) has garnered special attention in the study of narrowly distributed taxa (Fig. 1A; Note: *Erythranthe laciniata* is the appellation in the most recent taxonomic revision of this group (Barker et al. 2012). We continue to apply the older taxonomy for consistency with prior studies, predominance of use in the research community (Lowry et al. 2019), and public genomic resources including additional monkeyflower reference genome assemblies (Goodstein et al. 2012)). *Mimulus laciniatus* is restricted to the western slope of the Sierra Nevada and their western foothills. Across its range, *M. laciniatus* can be found in a wide range of climates and habitats (i.e. foothill woodland, montane mixed-conifer, subalpine, and alpine communities) and spans a broad elevational range (~900 to 3,270 m), but populations are generally restricted to seeps on granite outcrops, often co-occurring with mosses and spikemosses (Sexton and Dickman 2016, Fraga 2018; Fig. 1B). *Mimulus laciniatus* is largely self-fertilizing, reflected by its low observed heterozygosity ($h_o = 0.07$; Ferris et al. 2014) and high inbreeding coefficient ($F_{IS} = 0.93$; Sexton et al. 2016). The taxon is distinguished by its dissected leaves that differentiate it from nearly all other closely related members of the species complex except *Mimulus filicifolius* (syn. *Erythranthe filicifolia*; Sexton et al. 2013) and rare populations of *Mimulus guttatus* (syn. *E. guttata* and *E. microphylla*; Ferris et al. 2015). However, in the lab, *M. laciniatus* can readily be crossed with other members of the complex, and the presence of natural hybrid zones indicates low or lack of intrinsic reproductive barriers (Vickery 1978, Tataru et al. 2025). Given the degree of morphological and ecological differentiation with respect to the widely distributed outcrosser *M. guttatus*, this taxon has been used to study traits related to adaptive divergence and reproductive isolation, specifically the genetic basis of flowering time and its plasticity to environmental cues (Friedman and Willis 2013, Ferris 2016, Leinonen et al. 2020), leaf shape (Ferris et al. 2015, Ferris et al. 2017), and floral morphology (Ferris et al. 2017). Many of these traits are hypothesized or shown to be adaptive and contribute to local adaptation of the species (DeMarche et al. 2013, Ferris and Willis 2018, Tataru et al. 2023).

In addition to this work on the ecology and genetics of speciation and adaptation, *M. laciniatus* has provided insight into the factors that limit species distributions (i.e. range

limits) and on conservation strategies for narrow endemics of the CA-FP. Research has demonstrated that populations at the warmer edges of the species range (i.e. foothill habitats) benefit significantly from gene flow originating from populations occupying similar warm-edge climates (Sexton et al. 2011). Conversely, gene flow from centrally located populations (i.e. montane habitats) has been shown to have a detrimental effect, indicating the importance of the gene flow source. Moreover, patterns of genetic isolation by climate environment, rather than isolation by distance, have been found across the species range (Sexton et al. 2016). This finding underscores the importance of carefully evaluating spatial genetic relationships when developing conservation strategies, such as assisted gene flow programs to promote genetic or evolutionary rescue. Additionally, studies investigating adaptive responses to climate change have revealed substantial genetic potential for drought tolerance within this group (Dickman et al. 2019). However this resilience is relative. Compared to the widely distributed *M. guttatus*, *M. laciniatus* tolerates a narrower range of temperatures—a pattern consistent with other monkeyflower species pairs where distribution breadth correlates with thermal tolerance (Sheth et al. 2014; Sheth and Angert 2014). The contrast between widely distributed species and narrowly distributed species highlights the distinct dynamics and unique vulnerabilities of narrowly distributed species, reinforcing the need for targeted study and conservation strategies.

Genomics has been instrumental in advancing many of these studies, providing critical insights and enabling novel approaches to answering fundamental questions in evolutionary biology, ecology, and conservation. For instance, DNA sequencing supports efforts to map the genetic basis of adaptive trait variation, infer evolutionary relationships, and provide necessary context and tools for conservation strategies. Prior genomic studies have largely been performed by aligning *M. laciniatus* sequences to a *M. guttatus* genome assembly as a reference, but reference genome bias, gene presence-absence variation, and structural variation create caveats or limitations that can impact results (Günther and Nettelblad 2019). Here we report a haplotype resolved genome of *M. laciniatus*. Produced as a part of the California Genomics Conservation Project (Fiedler et al. 2022, Shaffer et al. 2022) and used in concert with additional resources for more broadly distributed taxa in *Mimulus* sect. *Simiolus* (Lovell et al. 2025; Kollar et al. 2025), this genome will empower new, more rigorous research into how annual endemics of the CA-FP evolve and how to develop effective conservation strategies for these taxa.

Methods

Biological material

Dried fruits were collected from *M. laciniatus* located at site OPN (Olmstead Point; 37.81073°N, 119.48518°W; Elev. 2,530 m) in Yosemite National Park, Mariposa County, California (Leinonen et al. 2020). Seeds from this individual were grown in the lab and selfed for multiple generations; tissue was collected from a single descendant of this line. Multiple floral buds were collected in a 50 mL conical tube and frozen.

Nucleic acid library preparation

PacBio HiFi

We extracted high molecular weight (HMW) genomic DNA (gDNA) from 130 mg of buds using the

cetyltrimethylammonium bromide (CTAB) method as described in Inglis et al. (2018), with the following modifications: 1) we used sodium metabisulfite (1% w/v) instead of 2-mercaptoethanol (1% v/v) in the sorbitol wash buffer and CTAB solution; 2) we repeated the tissue homogenate wash steps until the supernatant turned clear; 3) we performed the CTAB lysis step at 45 °C; and 4) the chloroform extraction step twice using ice-cold chloroform. The DNA purity was estimated by using the NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA). The DNA yield (1.3 µg) was quantified using a Quantus Fluorometer (QuantiFluor ONE dsDNA Dye assay; Promega, Madison, WI), and the size distribution of the DNA was estimated using the Femto Pulse system (Genomic DNA 165 kb kit, Agilent, Santa Clara, CA), where 70% of the DNA fragments were found to be 30 kilobases (kb) or longer.

The HiFi SMRTbell library was constructed using the SMRTbell prep kit 3.0 (Pacific Biosciences—PacBio, Menlo Park, CA; Cat. #102-182-700) according to the manufacturer's instructions. HMW gDNA was sheared to a target DNA size distribution between 15 and 18 kb using Diagenode's Megaruptor 3 system (Diagenode, Belgium; cat. B06010003). The sheared gDNA was concentrated using 1× of SMRTbell cleanup beads provided in the SMRTbell prep kit 3.0 for the repair and a-tailing incubation at 37 °C for 30 min and 65 °C for 5 min, followed by ligation of overhang adapters at 20 °C for 30 min, clean-up using 1X SMRTbell cleanup beads, and nuclease treatment at 37 °C for 15 min. The SMRTbell library was size selected using 3.1X of 35% v/v diluted AMPure PB beads (PacBio, Cat. #100-265-900) to progressively remove SMRTbell templates < 5 kb. The 15 to 18 kb average HiFi SMRTbell library was sequenced at UC Davis DNA Technologies Core (Davis, CA) using one 8 M SMRT cell (PacBio, Cat #101-389-001), Sequel II sequencing chemistry 2.0, and 30-h movies each on a PacBio Sequel IIe sequencer.

Omni-C

The Omni-C library was prepared using the Dovetail™ Omni-C™ Kit (Dovetail Genomics, Scotts Valley, CA) according to the manufacturer's protocol with slight modifications. First, tissue from floral buds and meristems was thoroughly ground with a mortar and pestle while cooled with liquid nitrogen. Nuclear isolation was then performed using published methods (Workman et al. 2018). Subsequently, chromatin was fixed in place in the nucleus and digested under various conditions of DNase I until a suitable fragment length distribution of DNA molecules was obtained. Chromatin ends were repaired and ligated to a biotinylated bridge adapter followed by proximity ligation of adapter containing ends. After proximity ligation, crosslinks were reversed and the DNA was purified from proteins. Purified DNA was treated to remove biotin that was not internal to ligated fragments. A next-generation sequencing library was generated using an NEB Ultra II DNA Library Prep kit (New England Biolabs, Ipswich, MA) with an Illumina compatible y-adaptor. Biotin-containing fragments were then captured using streptavidin beads. The post-capture product was split into two replicates prior to PCR enrichment to preserve library complexity with each replicate receiving unique dual indices. The library was sequenced at Vincent J. Coates Genomics Sequencing Lab (Berkeley, CA) on an Illumina NovaSeq 6000 platform

(Illumina, CA) to generate approximately 100 million 2 × 150 bp read pairs per gigabase (Gb) of genome size.

Nuclear genome assembly

We assembled the genome of a *M. laciniatus* individual following the CCGP assembly pipeline Version 5.0, as outlined in Table 1 listing the tools and non-default parameters used in the assembly process. We removed the remnants of adapter sequences from the PacBio HiFi dataset using HiFiAdapterFilter (Sim et al. 2022) and generated an initial diploid phased assembly using HiFiasm (Cheng et al. 2022) in HiC mode, with the filtered PacBio HiFi reads and the Omni-C short-reads, a process that generates two assemblies, one per haplotype. We aligned the Omni-C data to each assembly following the Arima Genomics Mapping Pipeline (https://github.com/ArimaGenomics/mapping_pipeline) and then scaffolded the assemblies with SALSA (Ghurye et al. 2017, Ghurye et al. 2019).

The assemblies for both haplotypes were manually curated by iteratively generating and analyzing their corresponding Omni-C contact maps. Briefly, to generate the contact maps we aligned the Omni-C data with BWA-MEM (Li 2013), identified ligation junctions, and generated Omni-C pairs (Lee et al. 2022) using pairtools (Open2C et al. 2024). Then, we generated multi-resolution Omni-C matrices with Cooler (Abdennur and Mirny 2020) and balanced them with hicExplorer (Ramírez et al. 2018). We used HiGlass (Kerpedjiev et al. 2018) and the PretextView (https://github.com/wtsi-hpag/PretextView; https://github.com/wtsi-hpag/PretextViewMap; https://github.com/wtsi-hpag/PretextViewSnapshot) to visualize the contact maps. We identified misassemblies and misjoins in these contact maps, and modified the assemblies using the Rapid Curation pipeline from the Wellcome Trust Sanger Institute, Genome Reference Informatics Team (https://gitlab.com/wtsi-grit/rapid-curation). Some of the remaining gaps (joins generated during scaffolding and/or curation) were closed using the PacBio HiFi reads and YAGCloser (https://github.com/merlyescalona/yagcloser). We checked for contamination using the BlobToolKit Framework (Challis et al. 2020).

Genome quality assessment

We generated k-mer counts from the Omni-C short-read and PacBio HiFi long-read datasets. For the Omni-C short read data, we first removed adapters using cutadapt (Martin 2011), following the indications from Dovetail Genomics (https://omni-c.readthedocs.io/en/latest/assembly.html). For the long-read data, we used the adapter-trimmed dataset. Then, we generated the k-mer counts for each dataset using meryl (https://github.com/marbl/meryl). The k-mer counts were then used in GenomeScope2.0 (Ranallo-Benavidez et al. 2020) to estimate genome features including genome size, heterozygosity, and repeat content. To obtain general contiguity metrics, we ran QUAST (Gurevich et al. 2013). To evaluate genome quality and functional completeness, we used BUSCO (Manni et al. 2021) with the Embryophyta ortholog database (embryophyta_odb10) which contains 1,614 genes. Assessment of base level accuracy (QV) and k-mer completeness was performed using the previously generated meryl database and merqury (Rhie et al. 2020). We further estimated genome assembly accuracy via BUSCO gene set frameshift analysis using the pipeline described in Korlach et al. (2017). Measurements of the size of the phased blocks is based on the

Table 1. Assembly pipeline and software used.

Assembly step	Software and any non-default options	Version	Reference
Initial assembly			
Filtering PacBio HiFi adapters	HiFiAdapterFilt	Commit 64d1c7b	Sim et al. 2022
K-mer counting	Meryl (k = 21)	1	https://github.com/marbl/meryl
Estimation of genome size and heterozygosity	GenomeScope (–l.50)	2	Ranallo-Benavidez et al. 2020
<i>De novo assembly (contigging)</i>	HiFiasm (Hi-C Mode, -l0, -primary, output hic.hap1.p_ctg, hic.hap2.p_ctg)	0.19.5-r592	Cheng et al. 2022
Scaffolding			
Omni-C data alignment	Arima Genomics Mapping Pipeline	Commit 2e74ea4	https://github.com/ArimaGenomics/mapping_pipeline
Arima Genomics Mapping Pipeline (AGMP)	BWA-MEM	0.7.17-r1188	Li 2013
	samtools	1.11	Danecek et al. 2021
	filter_five_end.pl (AGMP)	Commit 2e74ea4	https://github.com/ArimaGenomics/mapping_pipeline
	two_read_bam_combiner.pl (AGMP)	Commit 2e74ea4	https://github.com/ArimaGenomics/mapping_pipeline
	picard	2.27.5	https://broadinstitute.github.io/picard/
Omni-C Scaffolding	SALSA (-DNASE, -i 20, -p yes)	2	Ghurye et al. 2017, Ghurye et al. 2019
Omni-C Contact map generation			
Short-read alignment	BWA-MEM (-SSP)	0.7.17-r1188	Li 2013
SAM/BAM processing	samtools	1.11	Danecek et al. 2021
SAM/BAM filtering	pairtools	0.3.0	Open2C et al. 2024
Pairs indexing	pairix	0.3.7	Lee et al. 2022
Matrix generation	cooler	0.8.10	Abdennur and Mirny 2020
Matrix balancing	hicExplorer (hicCorrectmatrix correct –filterThreshold –2 4)	3.6	Ramírez et al. 2018
Contact map visualization			
	HiGlass	2.1.11	Kerpedjiev et al. 2018
	PretextView	0.1.4	https://github.com/wtsi-hpag/PretextView
	PretextView	0.1.5	https://github.com/wtsi-hpag/PretextView
	PretextViewSnapshot	0.0.3	https://github.com/wtsi-hpag/PretextViewSnapshot
Manual curation tools	Rapid curation pipeline (Wellcome Trust Sanger Institute, Genome Reference Informatics Team)	Commit 7acf220c	https://gitlab.com/wtsi-grit/rapid-curation
Genome quality assessment			
Basic assembly metrics	QUAST (–est-ref-size)	5.0.2	Gurevich et al. 2013
Assembly completeness	BUSCO (-m geno, -l embryophyta)	5.0.0	Manni et al. 2021
	Merqury	2020 January 29	Rhie et al. 2020
Contamination screening			
Local alignment tool	BLAST+ (-db nt, -outfmt ‘6 qseqid staxids bitscore std’, -max_target_seqs 1, -max_hsp 1, -evalue 1e-25)	2.15	Camacho et al. 2009
General contamination screening	BlobToolKit (HiFi coverage, BUSCO = embryophyta, NCBI Taxa ID = 4156)	2.3.3	Challis et al. 2020
Chloroplast genome assembly			
<i>de novo</i> assembler of organelle genomes	Oatk	1.0 (Commit 6591c42)	https://github.com/c-zhou/oatk
Organelle genome annotation	GeSeq	https://chlorobox.mpi-goettingen.de/geSeq.html	Tillich et al. 2017

size of the contigs generated by HiFiasm on HiC mode. We follow the quality metric nomenclature established by Rhie et al. (2021), with the genome quality code x.y.P.Q.C, where, x = log₁₀[contig NG50]; y = log₁₀[scaffold NG50]; P = log₁₀[phased block NG50]; Q = Phred base accuracy QV (quality value); C = % genome represented by the first “n” scaffolds, following a karyotype of 2n=28 for this species (Vickery

1995). Quality metrics for the notation were calculated on the primary haplotype assembly.

Chloroplast genome assembly

The chloroplast sequence for *M. laciniatus* was generated with Oatk (<https://github.com/c-zhou/oatk>). We used GeSeq (Tillich et al. 2017) to generate a draft genome annotation.

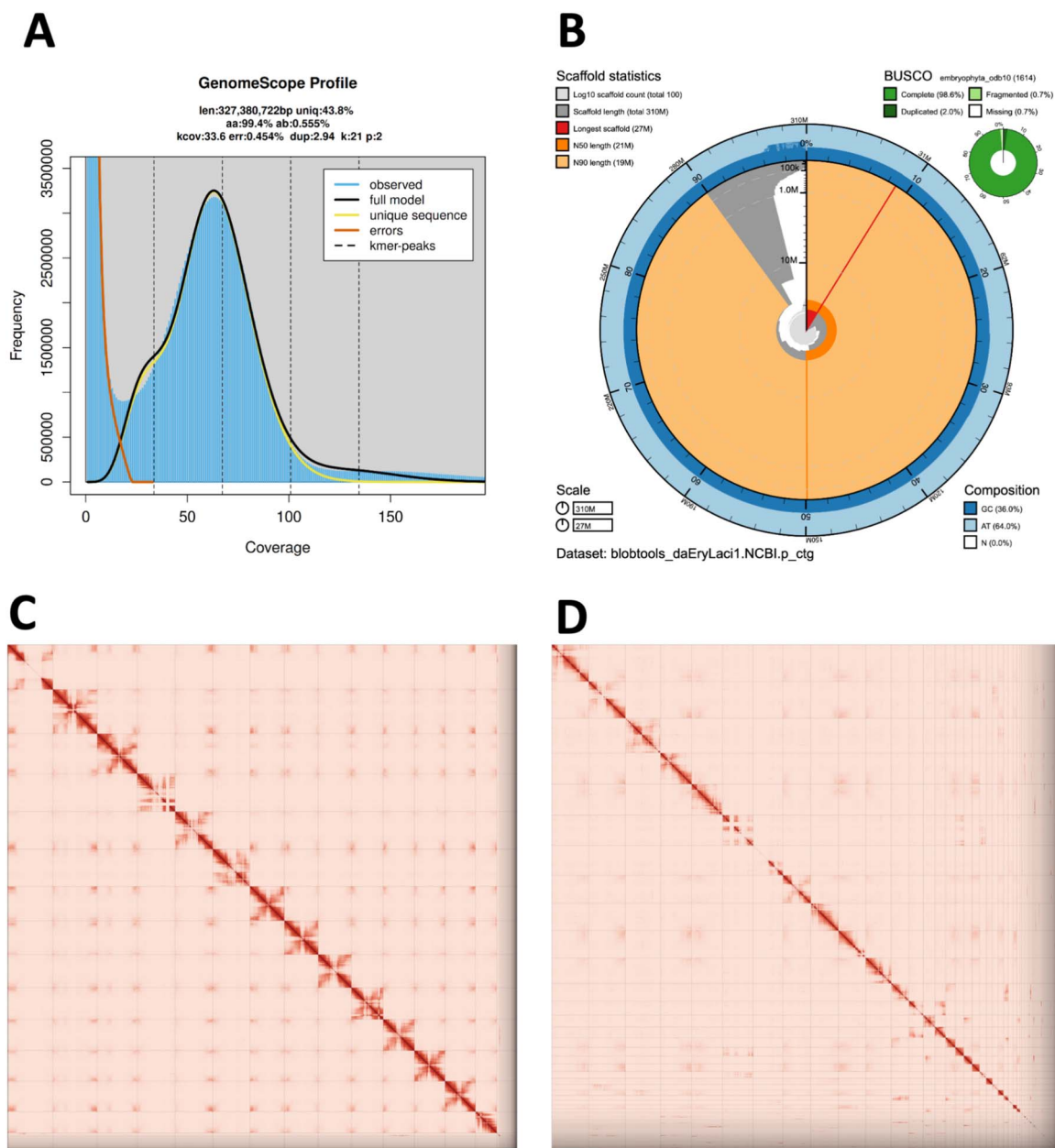


Fig. 2. *Mimulus laciniatus* genome metrics. A) K-mer spectrum of Omni-C short read data without adapters using Genomescope 2.0. A single peak consistent with a selfer. B) BlobToolKit snail plot showing visualization of metrics of the primary assembly (daEryLaci1.0.p). C, D) Hi-C contact maps for the primary (C) and alternate (D) genome.

After completion of the nuclear genome, we searched for matches of the resulting chloroplast assembly sequence in the nuclear genome assembly using BLAST+ (Camacho et al. 2009) and filtered out contigs and scaffolds from the nuclear genome with a percentage of sequence identity > 99% and size smaller than the chloroplast assembly sequence.

Results

Sequencing results

The Omni-C library generated 278.87 million read pairs, and the PacBio HiFi library generated 1.99 million reads. The PacBio HiFi sequences yielded $\sim 95\times$ genome coverage and had an N50 read length of 16,599 bp; a minimum read

length of 137 bp; a mean read length of 15,173 bp; and a maximum read length of 55,779 bp (see Fig. S1 for read length distribution). Based on k-mer analysis of both short- and long-read data, we estimated a genome size of ~ 317.72 Mb. The estimations vary depending on the dataset used. Omni-C short-read data estimated a heterozygous rate of 0.56%, a sequencing error of 0.454%, and a genome size of 327.38 Mb (Fig. 2A), while the PacBio HiFi long-read data estimated a heterozygous rate of 0.26%, a sequencing error of 0.124, and a genome size of 308.07 Mb (Fig. S2a). Using default parameters for these estimations shows poor model fitting (Table S2), especially for the long-read data, so we modified the parameterization of the underlying model by using $l = 50$ (corresponding to the 1n peak, which is half of the coverage of our major peak; Fig. S2).

Table 2. Sequencing and assembly statistics, and accession numbers.

Bio Projects and vouchers	CCGP NCBI BioProject		PRJNA720569				
	Genera NCBI BioProject		PRJNA765633				
	Species NCBI BioProject		PRJNA777196				
	NCBI BioSample		SAMN38285827				
	Specimen identification		CCGP_10_BB_OPN_1				
	NCBI Genome accessions		Haplotype 1	Haplotype 2			
	Assembly accession		JBCDMK000000000	JBCDML000000000			
	Genome sequences		GCA_040207155.1	GCA_040207145.1			
Genome Sequence	PacBio HiFi reads	Run	1 PACBIO_SMRT (Sequel Iie) run: 2 M spots, 30.2G bases, 18.5Gb				
		Accession	SRX26067539				
	Omni-C Illumina reads	Run	2 ILLUMINA (Illumina NovaSeq 6000) runs: 107.1 M spots, 32.3G bases, 7.7Gb				
		Accession	SRX26067540-1				
Genome Assembly Quality Metrics	Assembly identifier (Quality code ^c)		daEryLaci1(7.7.P.Q49.C98)				
	HiFi Read coverage ^a		95X				
			Haplotype 1	Haplotype 2			
	Number of contigs		148	316			
	Contig N50 (bp)		11,090,440	2,363,584			
	Contig NG50 ^a		11,090,440	838,590			
	Longest Contigs		24,297,547	9,135,367			
	Number of scaffolds		104	194			
	Scaffold N50		20,993,862	11,746,579			
	Scaffold NG50 ^a		20,993,862	4,460,205			
	Largest scaffold		27,399,035	15,572,277			
	Size of final assembly		309,971,392	213,857,377			
	Phased block NG50 ^a		11,220,445	1,195,452			
	Gaps per Gbp (# Gaps)		142(44)	570(122)			
	Indel QV (Frame shift)		51.35164467	49.71322245			
	Base pair QV		61.0895	60.3083			
	k-mer completeness		Full assembly = 60.7531				
			98.8371	62.4651			
			Full assembly = 99.2715				
	BUSCO completeness ^d (embryophyta) <i>n</i> = 1614	C	S	D	F	M	
		H1 ^b	98.60%	96.60%	2.00%	0.70%	0.70%
		H2 ^b	65.80%	63.60%	2.20%	1.10%	33.10%
	Organelles		1 Partial chloroplast sequence		JBCDMK010000104.1		

^aRead coverage and NGx statistics have been calculated based on the genome size of 317.72 Mb. The genome size is the average of the estimated genome sizes from PacBio HiFi reads and Omni-C Illumina short reads. ^b(H1) Haplotype 1 and (H2) Haplotype 2 assembly values. ^cAssembly quality code x.y.P.Q.C derived notation, from (Rhie et al. 2021). x = log10[contig NG50]; y = log10[scaffold NG50]; P = log10 [phased block NG50]; Q = Phred base accuracy QV (Quality value); C = % genome represented by the first “n” scaffolds, following a known karyotype for this species of 2n = 92 (Genome on a Tree, query (*Mimulus laciniatus*); Challis et al. 2023). Quality code for all the assembly denoted by primary assembly (daEryLaci1.0.p). ^dBUSCO Scores. Complete BUSCOs (C). Complete and single-copy BUSCOs (S). Complete and duplicated BUSCOs (D). Fragmented BUSCOs (F). Missing BUSCOs (M).

Nuclear genome assembly

The final genome assembly (daEryLaci1) consists of two phased haplotypes. Both assemblies are similar in size, but not equal to the estimated genome size from GenomeScope2.0, as has been observed in other taxa (see e.g. Pflug et al. 2020).

The primary haplotype assembly (daEryLaci1.0.p) consists of 104 scaffolds spanning 309.97 Mb with a contig N50 of 11.09 Mb, a scaffold N50 of 20.99 Mb, the largest contig size of 24.29 Mb, and the largest scaffold size of 27.39 Mb. The alternate haplotype assembly (daEryLaci1.0.a) consists of 194 scaffolds spanning 213.85 Mb with a contig N50 of 2.36 Mb, a scaffold N50 of 11.74 Mb, the largest contig size of 9.13 Mb and the largest scaffold size of 15.56 Mb. The primary haplotype assembly has a BUSCO completeness score for the Embryophyta gene set of 98.6%, a base pair quality value (QV) of 61.08, a kmer completeness of 98.83%, and a frameshift indel QV of 51.35. The alternate haplotype assembly has a BUSCO completeness score for the Embryophyta gene set of 65.8%, a base pair QV of 60.3, a

kmer completeness of 62.46%, and a frameshift indel QV of 49.71.

During manual curation we made a total of 86 joins (56 on the primary haplotype and 30 on the alternate haplotype) and 26 breaks (11 on the primary haplotype and 15 on the alternate haplotype) based on the Omni-C contact map signal. We closed a total of 18 gaps (7 on the primary haplotype and 11 on the alternate haplotype) and filtered out 5 contigs corresponding to chloroplast contamination. No other contigs were removed or modified. The Omni-C contact maps show highly contiguous assemblies, with chromosome-length scaffolds (Fig. 2B). Assembly statistics are reported in Table 2 and represented graphically in Fig. 2C. We have deposited the genome assembly on NCBI GenBank (See Table 2 and Data Availability for details).

Chloroplast genome assembly

We assembled a partial chloroplast genome for *M. laciniatus* with Oatk. The final chloroplast sequence has a size

Table 3. Genome sizes in *Mimulus*.

Section	Species	Population/identifier	Size (MB)	Reference
<i>Simolus</i> (syn. <i>Simiola</i>)	<i>Mimulus laciniatus</i> (syn. <i>Erythranthe lacinata</i>)	OPN	309.97	This study
	<i>Mimulus nasutus</i> (syn. <i>Erythranthe nasutua</i>)	SF	312.9	Lovell et al. 2025
	<i>Mimulus tilingii</i> (syn. <i>Erythranthe tilingii</i>)	LVR1.1	315.4	Lovell et al. 2025
	<i>Mimulus guttatus</i> (syn. <i>Erythranthe microphylla</i>) ^b	IM767 (V2)	314.6	Lovell et al. 2025
	<i>Mimulus guttatus</i> (syn. <i>Erythranthe microphylla</i>) ^b	IM62 (V3)	339.7	Lovell et al. 2025
	<i>Mimulus guttatus</i> (syn. <i>Erythranthe microphylla</i>) ^b	IM155	295	Veltsos and Kelly 2024
	<i>Mimulus guttatus</i> (syn. <i>Erythranthe microphylla</i>) ^b	IM444	323	Veltsos and Kelly 2024
	<i>Mimulus guttatus</i> (syn. <i>Erythranthe microphylla</i>) ^b	IM502	263	Veltsos and Kelly 2024
	<i>Mimulus guttatus</i> (syn. <i>Erythranthe microphylla</i>) ^b	IM541	327	Veltsos and Kelly 2024
	<i>Mimulus guttatus</i> (syn. <i>Erythranthe microphylla</i>) ^b	IM664	300	Veltsos and Kelly 2024
	<i>Mimulus guttatus</i> (syn. <i>Erythranthe microphylla</i>) ^b	IM909	311	Veltsos and Kelly 2024
	<i>Mimulus guttatus</i> (syn. <i>Erythranthe microphylla</i>) ^b	IM1034	325	Veltsos and Kelly 2024
	<i>Mimulus guttatus</i> (syn. <i>Erythranthe microphylla</i>) ^b	IM1192	284	Veltsos and Kelly 2024
	<i>Mimulus guttatus</i> (syn. <i>Erythranthe microphylla</i>) ^b	LMC-L1	277	Kollar et al. 2025
	<i>Mimulus guttatus</i> (syn. <i>Erythranthe grandis</i>) ^b	SWB-S1	278	Kollar et al. 2025
	<i>Mimulus sookensis</i> (no name under <i>Erythranthe</i>)	FAN	574	Whitener et al. 2024
	<i>Mimulus glabratus</i> (syn. <i>Erythranthe glabrata</i>)	–	391 ^a	Bai et al. 2012
	<i>Mimulus luteus</i> var. <i>luteus</i> (syn. <i>Erythranthe lutea</i> var. <i>lutea</i>)	EY7	599	Cooley et al. 2022
<i>Erythranthe</i> (syn. <i>Erythranthe</i>)	<i>Mimulus lewisii</i> (syn. <i>Erythranthe lewisii</i>)	LF10 (V2)	423	http://mimubase.org/
	<i>Mimulus cardinalis</i> (syn. <i>Erythranthe cardinalis</i>)	CE10 (V2)	448	http://mimubase.org/
	<i>Mimulus parishii</i> (syn. <i>Erythranthe parishii</i>)	Mpar (v1.1)	380	http://mimubase.org/
	<i>Mimulus verbenaceus</i> (syn. <i>Erythranthe verbenacea</i>)	MvBL (v1.1)	453	http://mimubase.org/
	<i>Mimulus aurantiacus</i> ssp. <i>puniceus</i> (syn. <i>Diplacus aurantiacus</i> ssp. <i>puniceus</i>)	UCSD	207	Stankowski et al. 2019
–	<i>Mimulus ringens</i>	–	630 ^a	Bai et al. 2012

^aConversion of C-value estimates based on flow cytometry. ^bSee Nesom (2012) on taxonomic identity of these samples.

of 155,214 bp, with base composition of A = 30.69%, C = 19.18%, G = 18.54%, T = 31.54%.

Discussion

Genome size distribution in the monkeyflowers

The assembly of the *M. laciniatus* genome represents the first reference genome assembly from a narrowly distributed member of the larger Phrymaceae family. The genome makes *M. laciniatus* the 11th member of genus *Mimulus* (sensu lato) and the 6th in section *Simiola* (pre Barker et al. 2012 taxonomic revision; Table 3) with a reference genome assembly. While we did not obtain a chromosome-level assembly, the first 14 scaffolds are by far the largest (Fig. 2C), consistent with prior chromosome counts (2n = 28; Vickery 1995). Of

the other plant genomes published in the CCGP, spanning liverworts, ferns and vascular plants, the *M. laciniatus* genome is the smallest to date (Anghel et al. 2022, Chaturvedi et al. 2024, González-Ramírez et al. 2024, Song et al. 2025). The genome size of the primary assembly is 309.97 Mb (Fig. S2), consistent with prior flow cytometry-based estimates (i.e. 360 to 368 Mb; Ferris et al. 2014). Our genome size range falls well within the range of other genomes in section *Simiola* (average = 338 Mb; Table 3). Genome sizes, of the assemblies or as determined by flow cytometry, are relatively stable in the genus overall (Range: 207 to 630 Mb), with the exceptions of recent allotetraploid *Mimulus sookensis* (574 Mb; Whitener et al. 2024), the allotetraploid *Mimulus luteus* var. *luteus* (599 Mb; Cooley et al. 2022), and *Mimulus ringens* (630 Mb; Bai et al. 2012). *Mimulus luteus* var. *luteus* belongs to a larger

clade that experienced an ancestral polyploidization event, resulting in an approximate doubling of the base chromosome number ($n = 30$; Beardsley et al. 2004). In contrast, *M. ringens*, which is not known to be a polyploid and has a smaller but variable base chromosome number ($n = 8, 11, 12$; Iyer 1980, Vickery 1995), is distantly related to the species in section *Simiola*. Nonetheless, the sequenced monkeyflower genomes likely represent a bias towards diploids, as polyploidy is common especially in section *Simiola* (Vickery 1995; Beardsley et al. 2004).

Genomics provides insight into the process of speciation

Mimulus laciniatus has been recognized as a morphologically and ecologically distinct species, yet its speciation history has remained unclear. Differences in flowering time and habitat differentiation have been identified as important prezygotic reproductive isolation barriers between *M. laciniatus* and *M. guttatus* (Ferris et al. 2017, Tataru et al. 2023). Yet, early phylogenetic analyses with few nuclear markers were unable to resolve the evolutionary relationship of *M. laciniatus*, particularly whether or not it represents a selfing lineage derived from *M. guttatus* independently from the origin of *M. nasutus* (Ritland et al. 1993, Ferris et al. 2014). A recent study using whole genome resequencing data has found that *M. laciniatus* forms a well-supported clade with *M. nasutus*, and this clade is sister to the, previously named, “southern *M. guttatus* clade” that includes both annual and perennial *M. guttatus* populations (Ivey et al. 2023). Additional investigation into these relationships is needed, however, to account for any complicating signals resulting from historic and ongoing introgression; natural hybrid zones between *M. laciniatus*, *M. nasutus*, *M. guttatus* likely introduce phylogenetic ambiguity (Vickery 1978; Tataru et al. 2025). The accession we have sequenced, OPN, comes from a population that is not sympatric with either *M. nasutus* or *M. guttatus*. Thus, the new *M. laciniatus* genome we have assembled provides an excellent resource that may be used in conjunction with the newly published *M. nasutus* (Lovell et al. 2025) and southern *M. guttatus* clade (Kollar et al. 2025) genome assemblies to continue disentangling the complicated evolutionary history of the yellow monkeyflower species complex as well as to investigate how gene flow and introgression have contributed to differentiation and homogenization of these lineages.

Importance of conservation strategies tailored to endemics of the CA-FP

In an era defined by rapid biodiversity loss and climate instability, understanding how species adapt at the genomic level has become increasingly crucial for effective conservation. This new *M. laciniatus* reference genome provides a foundation for future evolutionary and conservation genetics studies to uncover genetic mechanisms underpinning resilience in extreme environments. Due to its restricted range and specialized habitat requirements, *M. laciniatus* faces inherent vulnerability to climate change, reflected by its automatic classification as California Rare Plant Rank 4. Even so, *M. laciniatus* populations possess adaptive variation that can respond to strong selection pressures. For instance, resurrection experiments using seed collected before and after the multi-year drought experienced in the Western US in the 2010s found that seed from post-drought collections germinated more quickly than seed from pre-drought collections, indicative of

adaptation for drought escape (Dickman et al. 2019). Climate conditions also appear to affect the extent of gene flow with other taxa. The strength of reproductive isolation between *M. laciniatus* and *M. guttatus* is stronger in drought years compared to wetter years (Tataru et al. 2023), suggesting that drought years promote *M. laciniatus* to differentiate more from *M. guttatus*, but also limiting potential contributions of adaptive introgression to genetic or evolutionary rescue. Understanding the phenotypic and genomic diversity of *M. laciniatus* is therefore critical for predicting evolutionary and ecological responses to environmental change and guiding conservation interventions. This reference genome will allow evolutionary and landscape genomics studies that can identify adaptive genetic variants or centers of genetic diversity and guide actions such as habitat preservation, assisted migration, and genetic rescue (Shay et al. 2021, Fiedler et al. 2022; Shaffer et al. 2022). Such insights may be extendable to the > 40 additional monkeyflowers (*Mimulus* sensu lato) in the California Rare Plant Inventory (Table S1) or to California endemic plants more broadly.

Acknowledgments

PacBio Sequel II/HiE library prep and sequencing was carried out at the DNA Technologies and Expression Analysis Core at the UC Davis Genome Center, supported by NIH Shared Instrumentation Grant 1S10OD010786-01. Deep sequencing of Omni-C libraries used the Novaseq S4 sequencing platforms at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant. We thank the staff at the UC Davis DNA Technologies and Expression Analysis Core and the UC Santa Cruz Paleogenomics Laboratory for their diligence and dedication to generating high quality sequence data. We thank Brad Shaffer for establishing and administering the CCGP, Erin Toffelmier and Courtney Miller for their administrative support of CCGP Projects, Päivi Leinonen and John Willis for providing the seed collection that formed the basis for this work, Kathleen Ferris and Diana Tataru for initial guidance on population selection, Yaowu Yuan for providing genome size estimates for members of section *Erythranthe*, and Christina Wistrom for assistance with plant husbandry at the UC Berkeley Oxford Tract Facility.

Author contributions

Jesús Martínez-Gómez (Investigation, Methodology, Project administration, Writing—original draft, Writing—review & editing), Merly Escalona (Data curation, Formal analysis, Methodology, Visualization, Writing—review & editing), Jack M. Colicchio (Conceptualization, Investigation, Writing—review & editing), Lauren N. Hamm (Investigation, Methodology, Writing—review & editing), Noravit Chumchim (Investigation, Methodology, Writing—review & editing), Mohan P.A. Marimuthu (Investigation, Methodology, Writing—review & editing), Oanh Nguyen (Investigation, Methodology, Writing—review & editing), William E. Seligmann (Investigation, Methodology, Writing—review & editing), Rachel S. Meyer (Conceptualization, Writing—review & editing), Jason P. Sexton (Conceptualization, Supervision, Writing—review & editing), and Benjamin K. Blackman (Conceptualization, Supervision, Writing—review & editing)

Supplementary material

Supplementary material is available at *Journal of Heredity* online.

Funding

This work was supported by the California Conservation Genomics Project, with funding provided to the University of California by the

State of California, State Budget Act of 2019 [UC Award ID RSI-19-690224]. This material is based on work supported by a National Science Foundation (NSF) Postdoctoral Research Fellowships in Biology Grant to JMG (IOS-2209159), a National Institutes of Health Ruth L. Kirschstein National Research Service Award Individual Postdoctoral Fellowship to J.M.C. (F32 GM125244-01), and NSF grants to B.K.B. and J.P.S. (IOS-1558035 and IOS-2222464).

Data availability

Data generated for this study are available under NCBI BioProject PRJNA777196. Raw sequencing data for sample OPN (NCBI BioSample SAMN38285827) are deposited in the NCBI Short Read Archive (SRA) under SRR30645747 for PacBio HiFi sequencing data, and SRR30645745, SRR30645746 for the Omni-C Illumina sequencing data. GenBank accessions for both primary and alternate assemblies are GCA_040207155.1 and GCA_040207145.1; and for genome sequences JBCDMK000000000 and JBCDML000000000. The GenBank genome assembly for the chloroplast genome is JBCDMK010000104.1. Assembly scripts and other data for the analyses presented can be found at the following GitHub repository: www.github.com/ccgproject/ccgp_assembly.

References

- Abdennur N, Mirny LA. Cooler: scalable storage for hi-C data and other genomically labeled arrays. *Bioinformatics*. 2020;36:311–316. <https://doi.org/10.1093/bioinformatics/btz540>
- Anghel IG, Jacobs SJ, Escalona M, Marimuthu MP, Fairbairn CW, Beraut E, Nguyen O, Toffelmier E, Shaffer HB, Zapata F. Reference genome of the color polymorphic desert annual plant sandblossoms, *Linanthus parryae*. *J Hered*. 2022;113:712–721. <https://doi.org/10.1093/jhered/esac052>
- Bai C, Alverson WS, Follansbee A, Waller DM. New reports of nuclear DNA content for 407 vascular plant taxa from the United States. *Annals of Botany*. 2012;110:1623–1629. <https://doi.org/10.1093/aob/mcs222>
- Baldwin BG, Goldman DH eds. *The Jepson manual: vascular plants of California*. Oakland (CA): University of California Press; 2012. <https://doi.org/10.1525/9780520951372>
- Baldwin BG. Origins of plant diversity in the California Floristic Province. *Annu Rev Ecol Evol Syst*. 2014;45:347–369. <https://doi.org/10.1146/annurev-ecolsys-110512-135847>
- Barker WR, Nesom GL, Beardsley PM, Fraga NS. A taxonomic conspectus of Phrymaceae: a narrowed circumscription for *Mimulus*, new and resurrected genera, and new names and combinations. *Phytoneuron*. 2012;39:1–60.
- Beardsley PM, Schoenig SE, Whittall JB, Olmstead RG. Patterns of evolution in western north American *Mimulus* (Phrymaceae). *Am J Bot*. 2004;91:474–489. <https://doi.org/10.3732/ajb.91.3.474>
- Brandvain Y, Kenney AM, Flagel L, Coop G, Sweigart AL. Speciation and introgression between *Mimulus nasutus* and *Mimulus guttatus*. *PLoS Genet*. 2014;10:e1004410. <https://doi.org/10.1371/journal.pgen.1004410>
- Burge DO, Thorne JH, Harrison SP, O'Brien BC, Rebman JP, Shevock JR, Alverson ER, Hardison LK, Rodríguez JD, Junak SA, et al. Plant diversity and endemism in the California Floristic Province. *Madroño*. 2016;63:3–206. <https://doi.org/10.3120/madr-63-02-3-206.1>
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009;10:421. <https://doi.org/10.1186/1471-2105-10-421>
- Challis R, Richards E, Rajan J, Cochrane G, Blaxter M. BlobToolKit—interactive quality assessment of genome assemblies. *G3: Genes, Genomes, Genetics*. 2020;10:1361–1374. <https://doi.org/10.1534/g3.119.400908>
- Challis R, Kumar S, Sotero-Caio C, Brown M, Blaxter M. Genomes on a Tree (GoaT): A versatile, scalable search engine for genomic and sequencing project metadata across the eukaryotic tree of life. *Wellcome Open Research*. 2023;8:24. <https://doi.org/10.12688/wellcomeopenres.18658.1>
- Chaturvedi S, Escalona M, Marimuthu MP, Nguyen O, Chumchim N, Fairbairn CW, Seligmann W, Miller C, Shaffer HB, Whiteman NK. A draft reference genome assembly of California pipevine, *Aristolochia californica* Torr. *J Hered*. 2024;115:507–515. <https://doi.org/10.1093/jhered/esae023>
- Cheng H, Jarvis ED, Fedrigo O, Koepfli KP, Urban L, Gemmell NJ, Li H. Haplotype-resolved assembly of diploid genomes without parental data. *Nat Biotechnol*. 2022;40:1332–1335. <https://doi.org/10.1038/s41587-022-01261-x>
- Cooley AM, Schlutius C, Matthews M, Zheng X, Thomas D, Edger PP, Platts AE, LaFountain A, George L, Williams A, et al. Genetic architectures of floral pigment and patterning in hybrid monkeyflowers. *bioRxiv*. 2022.04.29.490035, 2023, preprint: not peer reviewed <https://doi.org/10.1101/2022.04.29.490035>
- Coughlan JM, Brown MW, Willis JH. The genetic architecture and evolution of life-history divergence among perennials in the *Mimulus guttatus* species complex. *Proc R Soc B*. 2021;288:20210077
- [CNPS] California Native Plant Society, Rare Plant Program. 2025. *Inventory of rare and endangered plants of California* (online edition, v9.5.1). [accessed 2025 May 2]. <http://www.rareplants.cnps.org>
- Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. Twelve years of SAMtools and BCFtools. *GigaScience*. 2021;10. <https://doi.org/10.1093/gigascience/giab008>
- Ghurye J, Pop M, Koren S, Bickhardt D, Chin C-S. Scaffolding of long read assemblies using long range contact information. *BMC Genomics*. 2017;18:527. <https://doi.org/10.1186/s12864-017-3879-z>
- Ghurye J, Rhie A, Walenz BP, Schmitt A, Selvaraj S, Pop M, Phillippy AM, Koren S. Integrating hi-C links with assembly graphs for chromosome-scale assembly. *PLoS Comput Biol*. 2019;15:e1007273. <https://doi.org/10.1371/journal.pcbi.1007273>
- González-Ramírez IS, Escalona M, Miller C, Chumchim N, Marimuthu MP, Nguyen O, Song MJ, Mishler BD. A genome assembly for the California endemic liverwort *Calasterella californica*. *J Hered*. 2024;116:389–396. <https://doi.org/10.1093/jhered/esae065>
- Grant AL. A monograph of the genus *Mimulus*. *Ann Mo Bot Gard*. 1924;11:99–388. <https://doi.org/10.2307/2394024>
- Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*. 2013;29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
- Dickman EE, Pennington LK, Franks SJ, Sexton JP. Evidence for adaptive responses to historic drought across a native plant species range. *Evol Appl*. 2019;12:1569–1582. <https://doi.org/10.1111/eva.12803>
- DeMarche ML, Rice KJ, Sexton JP. Niche partitioning between close relatives suggests trade-offs between adaptation to local environments and competition. *Ecology and Evolution*. 2013;3:512–522. <https://doi.org/10.1002/ece3.462>
- Fiedler PL, Erickson B, Esgro M, Gold M, Hull JM, Norris JM, Shapiro B, Westphal M, Toffelmier E, Shaffer HB. Seizing the moment: the opportunity and relevance of the California conservation genomics project to state and federal conservation policy. *J Hered*. 2022;113:589–596. <https://doi.org/10.1093/jhered/esac046>
- Farnitano MC, Karoly K, Sweigart AL. Fluctuating reproductive isolation and stable ancestry structure in a fine-scaled mosaic of hybridizing *Mimulus* monkeyflowers. *PLoS Genet*. 2025;21:e1011624. <https://doi.org/10.1371/journal.pgen.1011624>
- Fenster CB, Ritland K. Evidence for natural selection on mating system in *Mimulus* (Scrophulariaceae). *Int J Plant Sci*. 1994;155:588–596. <https://doi.org/10.1086/297197>

- Ferris KG, Sexton JP, Willis JH. Speciation on a local geographic scale: the evolution of a rare rock outcrop specialist in *Mimulus*. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2014;369:20140001. <https://doi.org/10.1098/rstb.2014.0001>
- Ferris KG, Rushton T, Greenlee AB, Toll K, Blackman BK, Willis JH. Leaf shape evolution has a similar genetic architecture in three edaphic specialists within the *Mimulus guttatus* species complex. *Ann Bot*. 2015;116:213–223. <https://doi.org/10.1093/aob/mcv080>
- Ferris KG. A hot topic: the genetics of adaptation to geothermal vents in *Mimulus guttatus*. *Mol Ecol*. 2016;25:5605–5607. <https://doi.org/10.1111/mec.13869>
- Ferris KG, Barnett LL, Blackman BK, Willis JH. The genetic architecture of local adaptation and reproductive isolation in sympatry within the *Mimulus guttatus* species complex. *Mol Ecol*. 2017;26:208–224. <https://doi.org/10.1111/mec.13763>
- Ferris KG, Willis JH. Differential adaptation to a harsh granite outcrop habitat between sympatric *Mimulus* species. *Evolution*. 2018;72:1225–1241. <https://doi.org/10.1111/evo.13476>
- Fishman L, Beardsley PM, Stathos A, Williams CF, Hill JP. The genetic architecture of traits associated with the evolution of self-pollination in *Mimulus*. *New Phytol*. 2015;205:907–917. <https://doi.org/10.1111/nph.13091>
- Friedman J, Willis JH. Major QTLs for critical photoperiod and vernalization underlie extensive variation in flowering in the *Mimulus guttatus* species complex. *New Phytol*. 2013;199:571–583. <https://doi.org/10.1111/nph.12260>
- Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, et al. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res*. 2012;40:D1178–D1186. <https://doi.org/10.1093/nar/gkr944>
- Gould B, Moeller DA, Eckhart VM, Tiffin P, Fabio E, Geber MA. Local adaptation and range boundary formation in response to complex environmental gradients across the geographical range of *Clarkia xantiana* ssp. *xantiana*. *J Ecol*. 2014;102:95–107. <https://doi.org/10.1111/1365-2745.12188>
- Günther T, Nettelblad C. The presence and impact of reference bias on population genomic studies of prehistoric human populations. *PLoS Genet*. 2019;15:e1008302. <https://doi.org/10.1371/journal.pgen.1008302>
- Hernández AI, Specht CD. Putative adaptive loci show parallel clinal variation in a California-endemic wildflower. *Mol Ecol*. 2023;32:4298–4312. <https://doi.org/10.1111/mec.17029>
- Hufford KM, Mazer SJ. Local adaptation and the effects of grazing on the performance of *Nassella pulchra*: implications for seed sourcing in restoration. *Restor Ecol*. 2012;20:688–695. <https://doi.org/10.1111/j.1526-100X.2011.00843.x>
- Hodges SA, Fulton M, Yang JY, Whittall JB. Verne Grant and evolutionary studies of aquilegia. *New Phytol*. 2004;161:113–120. <https://doi.org/10.1046/j.1469-8137.2003.00950.x>
- Inglis PW, Pappas MDCR, Resende LV, Grattapaglia D. Fast and inexpensive protocols for consistent extraction of high quality DNA and RNA from challenging plant and fungal samples for high-throughput SNP genotyping and sequencing applications. *PLoS One*. 2018;13:e0206085. <https://doi.org/10.1371/journal.pone.0206085>
- Ivey CT, Habecker NM, Bergmann JP, Ewald J, Frayer ME, Coughlan JM. Weak reproductive isolation and extensive gene flow between *Mimulus glaucescens* and *M. guttatus* in northern California. *Evolution*. 2023;77:1245–1261. <https://doi.org/10.1093/evolut/qp404>
- Iyer NP. *Cytogenetics of Mimulus alatus and Mimulus ringens*. American University; ProQuest Dissertations & Theses, 1980, 1315268.
- Kenney AM, Sweigart AL. Reproductive isolation and introgression between sympatric *Mimulus* species. *Mol Ecol*. 2016;25:2499–2517. <https://doi.org/10.1111/mec.13630>
- Kerpedjiev P, Abdennur N, Lekschas F, McCallum C, Dinkla K, Strobelt H, Luber JM, Ouellette SB, Azhir A, Kumar N, et al. HiGlass: web-based visual exploration and analysis of genome interaction maps. *Genome Biol*. 2018;19:1–12. <https://doi.org/10.1186/s13059-018-1486-1>
- Kollar LM, Stanley LE, Kenchanmane Raju SK, Lowry DB, Niederhuth CE. The evolution of locally adaptive chromosome inversions in *Mimulus guttatus*. *Mol Ecol*. 2025;e17708. <https://doi.org/10.1111/mec.17708>
- Korlach J, Gedman G, Kingan SB, Chin CS, Howard JT, Audet JN, Cantin L, Jarvis ED. De novo PacBio long-read and phased avian genome assemblies correct and add to reference genes generated with intermediate and short reads. *GigaScience*. 2017;6:gix085.
- Lee S, Bakker CR, Vitzthum C, Alver BH, Park PJ. Pairs and Pairix: a file format and a tool for efficient storage and retrieval for hi-C read pairs. *Bioinformatics*. 2022;38:1729–1731. <https://doi.org/10.1093/bioinformatics/btab870>
- Leinonen PH, Salmela MJ, Greenham K, McClung CR, Willis JH. Populations are differentiated in biological rhythms without explicit elevational clines in the plant *Mimulus laciniatus*. *J Biol Rhythm*. 2020;35:452–464. <https://doi.org/10.1177/0748730420936408>
- Lekberg Y, Roskilly B, Hendrick MF, Zabinski CA, Barr CM, Fishman L. Phenotypic and genetic differentiation among yellow monkeyflower populations from thermal and non-thermal soils in Yellowstone National Park. *Oecologia*. 2012;170:111–122. <https://doi.org/10.1007/s00442-012-2297-9>
- Li H. *Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM*. *arXiv*. 2013. <http://arxiv.org/abs/1303.3997>.
- Lowry DB, Rockwood RC, Willis JH. Ecological reproductive isolation of coast and inland races of *Mimulus guttatus*. *Evolution*. 2008;62:2196–2214. <https://doi.org/10.1111/j.1558-5646.2008.00457.x>
- Lowry DB, Sobel JM, Angert AL, Ashman TL, Baker RL, Blackman BK, Brandvain Y, Byers KJ, Cooley AM, Coughlan JM, et al. The case for the continued use of the genus name *Mimulus* for all monkeyflowers. *Taxon*. 2019;68:617–623. <https://doi.org/10.1002/tax.12122>
- Lovell J, Walstead R, Lawrence A, Stark-Dykema E, Farnitano M, Harder A, Bruna T, Barry K, Goodstein D, Jenkins J, et al. *Comparative analyses of four reference genomes reveal exceptional diversity and weak linked selection in the yellow monkeyflower (Mimulus guttatus) complex*. *Evol Appl*. 2025;e70012. <https://doi.org/10.1111/1755-0998.70012>
- Manni M, Berkeley MR, Seppey M, Simao FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. 2021, arXiv 2106.11799 [q-bio] [Internet]. <http://arxiv.org/abs/2106.11799>
- MacNair MR. The genetic control of copper tolerance in the yellow monkey flower, *Mimulus guttatus*. *Heredity*. 1983;50:283–293. <https://doi.org/10.1038/hdy.1983.30>
- MacNair MR. The genetics of metal tolerance in vascular plants. *New Phytol*. 1993;124:541–559. <https://doi.org/10.1111/j.1469-8137.1993.tb03846.x>
- Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal*. 2011;17:10–12. <https://doi.org/10.14806/ej.17.1.200>
- Martin NH, Willis JH. Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric *Mimulus* species. *Evolution*. 2007;61:68–82. <https://doi.org/10.1111/j.1558-5646.2007.00006.x>
- Mittermeier RA, Turner WR, Larsen FW, Brooks TM, Gascon C. Global biodiversity conservation: the critical role of hotspots. In: Zachos FE, Habel JC, editors. *Biodiversity hotspots: distribution and protection of conservation priority areas*. Springer-Verlag: Heidelberg Berlin; 2011. p. 3–22. https://doi.org/10.1007/978-3-642-20992-5_1
- Fraga NS. *Erythranthe laciniata*. In: Jepson Flora project, editor. *Jepson eFlora, Revision 6*. 2018. [accessed 2025 March 31]. https://ucjeps.berkeley.edu/eflora/eflora_display.php?tid=99142
- Nesom GL. Taxonomy of *Erythranthe* sect. *Simiola* (Phrymaceae) in the USA and Mexico. *Phytoneuron*. 2012;40:1–123.
- Nesom GL. Commentary: speciation on a local geographic scale in *Mimulus*. *Phytoneuron*. 2014;72:1–4.

- Open2C, Abdennur N, Fudenberg G, Flyamer IM, Galitsyna AA, Goloborodko A, Imakaev M, Venev SV. Pairtools: from sequencing data to chromosome contacts. *PLoS Comput Biol*. 2024;20:e1012164. <https://doi.org/10.1371/journal.pcbi.1012164>
- Pettengill JB, Moeller DA. Phylogeography of speciation: allopatric divergence and secondary contact between outcrossing and selfing *Clarkia*. *Mol Ecol*. 2012;21:4578–4592. <https://doi.org/10.1111/j.1365-294X.2012.05715.x>
- Pflug JM, Holmes VR, Burrus C, Johnston JS, Maddison DR. Measuring genome sizes using read-depth, k-mers, and flow cytometry: methodological comparisons in beetles (Coleoptera). *G3*. 2020;10:3047–3060. <https://doi.org/10.1534/g3.120.401028>
- Ramírez F, Bhardwaj V, Arrigoni L, Lam KC, Grüning BA, Villaveces J, Habermann B, Akhtar A, Manke T. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat Commun*. 2018;9:189.
- Ranallo-Benavidez TR, Jaron KS, Schatz MC. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat Commun*. 2020;11:1432.
- Raven PH, Axelrod DI. Angiosperm biogeography and past continental movements. *Ann Mo Bot Gard*. 1974;61:539–673. <https://doi.org/10.2307/2395021>
- Ritland CE, Ritland K, Straus NA. Variation in the ribosomal internal transcribed spacers (ITS1 and ITS2) among eight taxa of the *Mimulus guttatus* species complex. *Mol Biol Evol*. 1993;10:1273–1288.
- Rhie A, Walenz BP, Koren S, Phillippy AM. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol*. 2020;21:245. <https://doi.org/10.1186/s13059-020-02134-9>
- Rhie A, McCarthy SA, Fedrigo O, Damas J, Formenti G, Koren S, Uliano-Silva M, Chow W, Fungtammasan A, Kim J, et al. Towards complete and error-free genome assemblies of all vertebrate species. *Nature*. 2021;592:737–746. <https://doi.org/10.1038/s41586-021-03451-0>
- Sexton JP, Strauss SY, Rice KJ. Gene flow increases fitness at the warm edge of a species' range. *Proc Natl Acad Sci*. 2011;108:11704–11709. <https://doi.org/10.1073/pnas.1100404108>
- Sexton JP, Ferris KG, Schoenig SE. The fern-leaved monkeyflower (Phrymaceae), a new species from the northern Sierra Nevada of California. *Madroño*. 2013;60:236–242. <https://doi.org/10.3120/0024-9637-60.3.236>
- Sexton JP, Dickman EE. What can local and geographic population limits tell us about distributions? *Am J Bot*. 2016;103:129–139. <https://doi.org/10.3732/ajb.1500224>
- Sexton JP, Hufford MB, Bateman AC, Lowry DB, Meimberg H, Strauss SY, Rice KJ. Climate structures genetic variation across a species' elevation range: a test of range limits hypotheses. *Mol Ecol*. 2016;25:911–928. <https://doi.org/10.1111/mec.13528>
- Sianta SA, Kay KM. Adaptation and divergence in edaphic specialists and generalists: serpentine soil endemics in the California flora occur in barer serpentine habitats with lower soil calcium levels than serpentine tolerators. *Am J Bot*. 2019;106:690–703. <https://doi.org/10.1002/ajb2.1285>
- Sim SB, Corpuz RL, Simmonds TJ, Geib SM. HiFiAdapterFilt, a memory efficient read processing pipeline, prevents occurrence of adapter sequence in PacBio HiFi reads and their negative impacts on genome assembly. *BMC Genomics*. 2022;23:157. <https://doi.org/10.1186/s12864-022-08375-1>
- Shaffer HB, Toffelmier E, Corbett-Detig RB, Escalona M, Erickson B, Fiedler P, Gold M, Harrigan RJ, Hodges S, Luckau TK, et al. Landscape genomics to enable conservation actions: the California conservation genomics project. *J Hered*. 2022;113:577–588. <https://doi.org/10.1093/jhered/esac020>
- Shay JE, Pennington LK, Mandussi Montiel-Molina JA, Toews DJ, Hendrickson BT, Sexton JP. Rules of plant species ranges: applications for conservation strategies. *Front Ecol Evol*. 2021;9:700962. <https://doi.org/10.3389/fevo.2021.700962>
- Sheth SN, Angert AL. The evolution of environmental tolerance and range size: a comparison of geographically restricted and widespread *Mimulus*. *Evolution*. 2014;68:2917–2931. <https://doi.org/10.1111/evo.12494>
- Sheth SN, Jiménez I, Angert AL. Identifying the paths leading to variation in geographical range size in western north American monkeyflowers. *J Biogeogr*. 2014;41:2344–2356. <https://doi.org/10.1111/jbi.12378>
- Sobel JM. Ecogeographic isolation and speciation in the genus *Mimulus*. *Am Nat*. 2014;184:565–579. <https://doi.org/10.1086/678235>
- Song MJ, Rizzieri YC, Li FW, Freund F, Escalona M, Toffelmier E, Miller C, Shaffer HB, Nguyen O, Marimuthu MP, et al. The genome assembly of the duckweed fern, *Azolla caroliniana*. *J Hered*. 2025;5:691–701. <https://doi.org/10.1093/jhered/esaf022>
- Stankowski S, Chase MA, Fuiten AM, Rodrigues MF, Ralph PL, Streisfeld MA. Widespread selection and gene flow shape the genomic landscape during a radiation of monkeyflowers. *PLoS Biology*. 2019;17:e3000391. <https://doi.org/10.1371/journal.pbio.3000391>
- Tataru D, Wheeler EC, Ferris KG. Spatially and temporally varying selection influence species boundaries in two sympatric *Mimulus*. *Proc R Soc B*. 2023;290:20222279. <https://doi.org/10.1098/rspb.2022.2279>
- Tataru D, De Leon M, Dutton S, Machado Perez F, Rendahl A, Ferris KG. Fluctuating selection in a monkeyflower hybrid zone. *Evolution Letters*. 2025;9:77–88. <https://doi.org/10.1093/evlett/qrae050>
- Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Res*. 2017;45:W6–W11. <https://doi.org/10.1093/nar/gkx391>
- Twyford, AD, Streisfeld, MA, Lowry, DB, Friedman, J. *Genomic studies on the nature of species: adaptation and speciation in Mimulus*. *Mol Ecol*. 2015;24:2601, 2609. <https://doi.org/10.1111/mec.13190>
- Veltsos P, Kelly JK. The quantitative genetics of gene expression in *Mimulus guttatus*. *PLoS Genet*. 2024;20:e1011072. <https://doi.org/10.1371/journal.pgen.1011072>
- Vickery RK. Speciation and isolation in section Simiolus of the genus *Mimulus*. *Taxon*. 1966;15:55–63. <https://doi.org/10.2307/1217588>
- Vickery Jr, RK. Case studies in the evolution of species complexes in *Mimulus*. In: Hecht MK, Steere WC, Wallace B, editors. *Evolutionary biology*. Boston (MA): Springer; 1978. p. 405–507. https://doi.org/10.1007/978-1-4615-6956-5_7
- Vickery RK. Speciation by aneuploidy and polyploidy in *Mimulus* (Scrophulariaceae). *Gt Basin Nat*. 1995;55:174–176.
- Whitener MR, Mangelson H, Sweigart AL. Patterns of genomic variation reveal a single evolutionary origin of the wild allotetraploid *Mimulus sookensis*. *Evolution*. 2024;78:1464–1477. <https://doi.org/10.1093/evolut/qpae079>
- Workman R, Timp W, Fedak R, Kilburn D, Hao S, Liu K. *High molecular weight DNA extraction from recalcitrant plant species for third generation sequencing*. 2019. protocols.io <https://dx.doi.org/10.17504/protocols.io.4vbgw2n>
- Wu CA, Lowry DB, Cooley AM, Wright KM, Lee YW, Willis JH. *Mimulus* is an emerging model system for the integration of ecological and genomic studies. *Heredity*. 2008;100:220–230. <https://doi.org/10.1038/sj.hdy.6801018>
- Yuan YW. Monkeyflowers (*Mimulus*): new model for plant developmental genetics and evo-devo. *New Phytol*. 2019;222:694–700. <https://doi.org/10.1111/nph.15560>