## **Ecological Diversification in an Adaptive Radiation of Plants:** The Role of De Novo Mutation and Introgression

Benjamin W. Stone 10 1,\* and Carolyn A. Wessinger 10 1

<sup>1</sup>Department of Biological Sciences, University of South Carolina, Columbia, SC 29208-3401, USA

\*Corresponding author: E-mail: bs66@mailbox.sc.edu.

Associate editor: Emily Josephs

## **Abstract**

Adaptive radiations are characterized by rapid ecological diversification and speciation events, leading to fuzzy species boundaries between ecologically differentiated species. Adaptive radiations are therefore key systems for understanding how species are formed and maintained, including the role of de novo mutations versus preexisting variation in ecological adaptation and the genome-wide consequences of hybridization events. For example, adaptive introgression, where beneficial alleles are transferred between lineages through hybridization, may fuel diversification in adaptive radiations and facilitate adaptation to new environments. In this study, we employed whole-genome resequencing data to investigate the evolutionary origin of hummingbird-pollinated flowers and to characterize genome-wide patterns of phylogenetic discordance and introgression in Penstemon subgenus Dasanthera, a small and diverse adaptive radiation of plants. We found that magenta hummingbird-adapted flowers have apparently evolved twice from ancestral blue-violet bee-pollinated flowers within this radiation. These shifts in flower color are accompanied by a variety of inactivating mutations to a key anthocyanin pathway enzyme, suggesting that independent de novo loss-of-function mutations underlie the parallel evolution of this trait. Although patterns of introgression and phylogenetic discordance were heterogenous across the genome, a strong effect of gene density suggests that, in general, natural selection opposes introgression and maintains genetic differentiation in gene-rich genomic regions. Our results highlight the importance of both de novo mutation and introgression as sources of evolutionary change and indicate a role for de novo mutation in driving parallel evolution in adaptive radiations.

Key words: phylogenomic, loss-of-function, pollination syndrome, anthocyanin, Penstemon.

## Introduction

Adaptive radiation—the proliferation of ecological roles and adaptations in different species within a lineage—is an important process considered to play a central role in the generation of biodiversity on Earth (Simpson 1953; Schluter 2000; Givnish 2015). This phenomenon, which often coincides with an increase in the rate of lineage diversification (and is thus termed a "rapid" or "explosive" adaptive radiation), encompasses a number of well-known and charismatic taxonomic groups, including Darwin's finches, Anolis lizards, and cichlid fishes. Sustained interest in adaptive radiations from evolutionary biologists has, over time, revealed several genomic features characteristic of rapidly radiating clades. Principal among these is an abundance of phylogenetic discordance that, in depicting species' complicated evolutionary histories, may more closely resemble a web than a bifurcating tree. This discordance may be chiefly caused by incomplete lineage sorting (ILS), which is expected to be abundantpotentially overwhelmingly so-when speciation occurs rapidly (Pamilo and Nei 1988; Maddison 1997; Rosenberg

2002; Whitfield and Lockhart 2007). Rapid adaptive radiations are also commonly characterized by introgression between closely related species that have not had much time to develop barriers to reproduction. Given that introgressed genetic material may confer fitness effects and can be subject to strong selection, particularly in functionally important genomic regions (Moran et al. 2021), signatures of introgression are expected to be distributed heterogeneously across the genome (Harrison and Larson 2016; Martin and Jiggins 2017). Introgression may provide a source of genetic diversity in adaptive radiations (Edelman and Mallet 2021), and although its importance in this context has historically been somewhat controversial (Seehausen 2004), many recent efforts have highlighted the myriad ways in which introgression may effectively broaden the genomic substrate for adaptation through novel combinations of haplotypes (e.g. McGee et al. 2020; Ferreira et al. 2021; Ronco et al. 2021; De-Kayne et al. 2022; Suvorov et al. 2022; Cicconardi et al. 2023; Marcionetti and Salamin 2023).

In addition to widespread ILS and introgression, radiations also often feature patterns of repeated evolution, which may

Received: November 01, 2023. Revised: January 01, 2024. Accepted: January 10, 2024 © The Author(s) 2024. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution. 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

Open Access

be either the result of the adaptive introgression of beneficial alleles or independent (convergent) adaptation in disparate lineages to similar ecological conditions.

While adaptive divergence progresses in adaptive radiations through complex interplays of standing genetic variation, introgression, and de novo mutation, it is standing or preexisting variation (often introduced through introgression) that is often implicated as the most important source of genetic variation underlying repeated or parallel evolution (Berner and Salzburger 2015; Haenel et al. 2019; Marques et al. 2019; Marques et al. 2022). Yet, it is clear that the repeated generation of similar phenotypes in ecologically similar habitats—a hallmark of adaptive radiation is also driven by de novo mutation. For example, both processes appear to be important in the radiation of hares (genus Lepus), in which the repeated evolution of nonwhite winter coats has proceeded through both adaptive introgression (Jones et al. 2018; Giska et al. 2019) and independent loss-of-function (LOF) mutations at the Agouti pigmentation gene (Jones et al. 2020). Whether adaptation more often involves de novo versus standing variation informs whether trait evolution is constrained by mutation events or by the opportunity for hybridization between differentiated lineages (Marques et al. 2019). Empirical data from diverse lineages across the tree of life are needed to identify common genetic features of adaptive radiations.

The genus Penstemon is an adaptive radiation of North American angiosperms, which exhibits an astonishing array of phenotypic and ecological diversity across its nearly 300 described species. Many Penstemon species are interfertile and form hybrids in nature, fueling speculation that interspecific hybridization has played a major role in the diversification of the group (Straw 1955; Wolfe et al. 2006; Wessinger et al. 2016). Penstemon is also recent and rapid radiation, estimated to have originated 1.419 to 6.416 MYA (Wolfe et al. 2021), and this rapid speciation has been highlighted as a potential major source of genealogical discordance (due to ILS) in addition to discordance caused by interspecific hybridization (Wolfe et al. 2006; Wessinger et al. 2016). Within Penstemon, the subgenus Dasanthera (hereafter, "Dasanthera") includes 9 ecologically diverse species and represents an adaptive radiation in microcosm. This group includes several species that hybridize in nature, making it an attractive system in which to study the importance of introgression and de novo mutation on adaptive divergence. A primary axis of variation in Dasanthera is floral pollination syndrome, wherein 2 hummingbird-adapted species (Penstemon newberryi and Penstemon rupicola) have evolved from the ancestral state of hymenopteran-adapted flowers (Datwyler and Wolfe 2004; Wilson et al. 2004; Kimball 2008). This mirrors a genus-wide trend involving the repeated evolution of specialist hummingbird-adapted flowers from more general hymenopteran-adapted progenitors (Wilson et al. 2007). The hummingbird-adapted species P. newberryi and P. rupicola each form hybrid zones with the hymenopteran-adapted, high-elevation Penstemon davidsonii; in each case, the hummingbird-adapted species occupies a lower elevation zone, and hybrids are formed

and persist at intermediate elevations in the Cascades and Sierra Nevada Mountains where the parent species' distributions overlap. Previous studies indicate that *P. newberryi* and *P. rupicola* are not sister species, suggesting parallel divergence in elevational tolerance and pollinator adaptation within *Dasanthera* (Stone and Wolfe 2021; Stone et al. 2023). An alternative scenario is that hummingbird pollination syndrome in *Dasanthera* could have a single evolutionary origin shared between species through introgression. Such a scenario has occurred in other plant groups; for example, the introgression of a *MaMyb2* allele has led to repeated origins of red-flowered phenotypes in the *Mimulus aurantiacus* species complex (Stankowski and Streisfeld 2015; Short and Streisfeld 2023).

Floral syndrome divergence in Dasanthera involves changes in floral morphology (hummingbird-pollinated: long, narrow flowers with exserted stamens and styles; hymenopteran-pollinated: short, wide flowers with inserted stamens and styles), and perhaps most strikingly, the evolution of bright magenta flowers in hummingbird-pollinated species from ancestral hymenopteran-pollinated blue-violet flowers. Whereas little is currently known about the underlying genetic mechanisms for floral morphological shifts, the evolution of red flowers tends to involve predictable changes to the genetic pathway for anthocyanin production. In other groups within Penstemon, shifts toward redder floral hues often involve the functional inactivation of the anthocyanin pathway enzyme flavonoid 3',5'-hydroxylase (F3'5'H) (Wessinger and Rausher 2014; Wessinger and Rausher 2015). F3'5'H belongs to the functionally diverse superfamily of cytochrome P450-dependent monooxygenases. This enzyme catalyzes the hydroxylation of the B-ring of anthocyanins, leading to the production of blue-colored delphinidin-based pigments (Tanaka and Brugliera 2013). The degeneration of F3'5'h prohibits the production of delphinidin and has been shown in Penstemon to result in flowers that produce red pelargonidin-based anthocyanins; importantly, the evolution of red flowers in multiple Penstemon species appears to be the result of repeated de novo LOF mutations to the F3'5'h coding sequence (CDS) (Wessinger and Rausher 2014; Wessinger and Rausher 2015).

In this study, we used whole-genome resequencing data to trace the evolutionary history and source of genetic variation for hummingbird-pollinated flowers in Dasanthera and characterize genome-wide patterns of introgression. Signatures of introgression were uneven across the genome and were generally lower in gene-rich regions, suggesting that even in this very recent radiation, introgression often may be maladaptive. We found strong support for independent evolutionary origins of magenta floral pigments in the 2 hummingbird-pollinated species. These parallel shifts in floral pigment production are accompanied by distinct degenerative mutations to the F3'5'h CDS. While introgression is prevalent within Dasanthera, we found little evidence that the 2 hummingbird-pollinated species share a history of introgression; this suggests that despite the importance of introgression across Dasanthera generally, it is de novo mutation, and not adaptive introgression, that likely fuels parallel shifts to hummingbird pollination in this adaptive radiation.

## **Methods**

## Study System and Species Sampling

We generated whole-genome resequencing data for 18 individual plants, each from a unique locality, representing 7 of the 9 total Dasanthera species (supplementary table \$1, Supplementary Material online). We included multiple samples for each of the species in the Cascades/Sierra Nevada Mountain clade (Penstemon cardwellii, P. davidsonii, Penstemon fruticosus, P. newberryi, and P. rupicola) and a single individual for each of the Northern Rocky Mountain exclusive taxa Penstemon montanus and Penstemon Ivallii for use as outgroups, following results from previous phylogenetic inquiries into the group (Stone and Wolfe 2021; Stone et al. 2023). We extracted DNA from silica-dried leaf tissues using a modified CTAB protocol and performed whole-genome Illumina library preparations and sequencing to ~10× depth using 150-bp paired-end reads generated on the NovaSeq (Illumina Inc.) platform at Duke University. Raw Illumina reads were quality filtered with fastp (Chen et al. 2018), enabling auto-detection of adapters, limiting read length to 30 bp, filtering out unpaired reads, enabling base correction for overlapping reads, and enabling poly-x trimming on 3' ends of reads. Final sequence quality was verified with fastqc (Andrews 2010) and multiqc (Ewels et al. 2016). Quality-filtered reads were mapped to the annotated P. davidsonii reference genome (Ostevik et al. 2023) with bwa-mem (Li 2013), and reads with low mapping quality (q < 20) were removed. Duplicated reads were marked and removed with SAMtools markdup (Li et al. 2009), and overlapping paired-end reads were clipped with bamutil clipOverlap (Jun et al. 2015).

Filtered reads were then used to call genotypes to produce an "all sites" .vcf file using bcftools v1.15.1 (Li 2011). Variant sites were soft filtered for minimum genotype-call quality (20), minimum and maximum allelic depths, which were chosen based on genome-wide averages of coverage depth (3 and 60, respectively), minimum individual allelic depth (2), and missingness (50%). Sites failing the minimum genotype-call quality filter were changed to match the reference allele (changed to invariant sites); otherwise, filters produced missing data for sites or individual genotypes. The resulting filtered all sites .vcf was then used to generate consensus whole-genome sequences and individual gene CDSs.

#### Species Relationships

We inferred species trees in 3 ways to compare genomic sampling strategies and analysis methods. First, we used custom Python scripts to split whole-genome sequences into nonoverlapping 10-kb windows, removing windows with more than 75% missing data (Ns) for any sample. We then estimated "window trees" for each genomic window with

IQ-TREE 2 v2.2.0.3 (Nguyen et al. 2015), enabling the ModelFinder (-m MFP) option (Kalyaanamoorthy et al. 2017) to select the best substitution model for each window. We then used these window trees to infer a species tree with ASTRAL-III v.5.7.8 (Zhang et al. 2018), using the full annotation option (-t 2). Second, for each sample, we used gffread v0.12.7 (Pertea and Pertea 2020) and the P. davidsonii reference genome annotations to generate fasta files with spliced exons (CDS), retaining the single longest isoform for each transcript. We then used a custom Python script to filter individuals with >50% missing data for each CDS. Gene trees were then estimated for each CDS, and a species tree was constructed in ASTRAL in the same manner as for the window alignments. For both of the ASTRAL species trees, we also calculated gene and site concordance factors (gCF and sCF; Minh et al. 2020) in IQ-TREE, using as input the window or gene trees and sequence alignments used to generate the species tree. Third, we estimated a concatenated maximum likelihood (ML) species tree in IQ-TREE, specifying the GTR + I + R model, estimating rates for each of the 12 largest scaffolds, and performing 1,000 ultrafast bootstrap (UFBoot) replicates (Hoang et al. 2018).

## Floral Pigment Production

We determined the identity of anthocyanidin pigments produced in floral tissues, which are responsible for floral color in *Penstemon*. We collected intact flowers from voucher specimens that were preserved in a herbarium press at the time of collection (samples were collected in 2017 and 2018). We used an isoamyl alcohol extraction protocol outlined in Harborne (1984) and separated pigments using TLC on Cellulose F glass plates (Millipore, Burlington, MA, USA). We compared extracted anthocyanidin pigments to a multi-anthocyanidin standard solution containing pelargonidin, cyanidin, and delphinidin each in concentrations of ~0.013 mg/mL, as described in Stevens et al. (2023).

#### Genome-Wide Average Signatures of Introgression

We investigated genome-wide signatures of introgression using Dsuite v0.5 (Malinsky et al. 2021). Specifically, we used Dsuite to calculate genome-wide D statistics (ABBA-BABA) for all possible trios of samples and to implement and visualize the "f-branch" metric  $f_b$  (Malinsky et al. 2018) using the 10-kb window ASTRAL species tree as the guide tree and P. montanus as the outgroup.

## Genomic Landscape of Divergence, Discordance, and Introgression

We examined patterns of differentiation among species and patterns of phylogenetic discordance across the genome. We calculated average  $F_{ST}$  (Weir and Cockerham 1984) and average nucleotide distance within  $(\pi)$  and between  $(d_{xy})$  species in 10- and 50-kb nonoverlapping windows with pixy v1.2.7 (Korunes and Samuk 2021), using the filtered all sites .vcf files as input. We used BEDTools v2.27.1 (Quinlan and Hall 2010) to calculate genic fractions (the proportion of sites that are CDS) per window. We identified  $d_{xy}$  outliers between 2

species pairs that form hybrid zones:  $P.\ davidsonii-P.\ newberryi$  and  $P.\ davidsonii-P.\ rupicola$ . Highly differentiated genomic regions in the face of gene flow may be especially resistant to introgression and could play a key role in the formation or maintenance of species boundaries. We first filtered all sliding windows with more missing counts than count comparisons and then identified outliers as 10-kb sliding windows with average  $d_{xy}$  values exceeding the genomewide mean by 3 standard deviations (Z-score > 3). We then identified shared outliers as sliding windows with outlier  $d_{xy}$  values in both species pairs.

To explore how topological discordance varies across the genome, we first calculated normalized Robinson-Foulds (RF) distances (Robinson and Foulds 1981) between each 10-kb window tree and the species tree using the R package phangorn v2.10.0 (Schliep 2011). In addition, we examined topological discordance affecting 3 main internal branches of the species tree using TWISST (Martin and Van Belleghem 2017). TWISST quantifies the proportion (or "weighting") of different tree topologies across the genome. For this approach, we used scripts provided by the developers (available at https://github.com/simonhmartin/twisst), using the 10-kb window trees as input. We summed topology weights for all trees consistent with the following internal branches of the species tree: the branch leading to P. newberryi, P. cardwellii, and P. rupicola (termed IB1); the branch leading to P. newberryi and P. cardwellii (IB2); and the branch leading to P. davidsonii and P. fruticosus (IB3). In each case, P. montanus served as the outgroup.

Although the D statistic is a powerful metric to detect genome-wide patterns of introgression, it is less useful to identify specific introgressed genomic regions because it is a poor estimator of the amount of introgression and it has a high degree of variance when applied to genomic windows of smaller sizes (Martin et al. 2015). For this reason, to explore signatures of introgression along the genome, we employed the  $f_{dM}$  statistic, which is specifically designed to estimate introgression between pairs of taxa in genomic windows along chromosomes and, like many statistics related to the f4 ratio, performs better than the D statistic (exhibits less variance and is a better estimator of the amount of introgression) when applied to small genomic windows (Malinsky et al. 2015; Malinsky et al. 2021). We estimated  $f_{dM}$  in 10,000 SNP windows, sliding every 2,500 SNPs. For each of these  $f_{dM}$  windows, we also calculated the genic fraction to assess the relationship between the genic fraction and  $f_{dM}$ . We conducted these calculations for all triplets of taxa consistent with the species tree.

#### Introgression of F3'5'h

We examined several lines of evidence to assess whether F3'5'h—one of the key genetic components underlying transitions to hummingbird pollination syndrome—could be shared between red-flowered *Dasanthera* through introgression. First, we calculated  $f_{dM}$  in sliding windows along the genome to investigate signatures of introgression between *P. newberryi* (P2) and *P. rupicola* (P3), setting

*P. cardwellii* as P1. Here, we estimated  $f_{dM}$  at 4 different window sizes: (i) in 10,000 SNP windows, sliding every 2,500 SNPs, (ii) in 5,000 SNP windows, sliding every 1,250 SNPs, (iii) in 1,000 SNP windows, sliding every 250 SNPs, and (iv) in 500 SNP windows, sliding every 125 SNPs. If F3'5'h is introgressed, values of  $f_{dM}$  should be elevated compared to both genome-wide estimates and the local genomic region.

Because F3'5'h is potentially pseudogenized in our magenta-flowered samples (leading to relaxed selection against indels), we re-called variants for reads mapping to this genomic region while additionally calling and normalizing indels with bcftools. We performed downstream variant calling and filtering as described for genome-wide variant calling steps. We then aligned the CDS for each sample with muscle v.5.1 (Edgar 2004) and re-estimated the gene tree for F3'5'h in IQ-TREE, performing 1,000 UFBoot replicates and 1,000 bootstrap replicates for the Shimodaira—Hasegawa-like approximate likelihood ratio test (SH-aLRT) to assess branch support (Guindon et al. 2010). If F3'5'h is introgressed, this gene tree should be discordant with the species tree and place *P. newberryi* and *P. rupicola* as sister.

We examined F3'5'h nucleotide and amino acid sequence alignments to identify the positions of potential indels and nonsynonymous mutations. F3'5'h CDSs in other hummingbird-adapted Penstemon species exhibit a variety of functional inactivating mutations, including premature stop codon mutations, frameshift mutations, and large deletions (Wessinger and Rausher 2015). We searched for similar LOF mutations present in our magenta-flowered samples. Likewise, mutations in any of the highly conserved domains, which included the heme-binding domain and 6 substrate recognition sites (SRSs) putatively involved in substrate contact (Gotoh 1992), would likely signal a dysfunctional enzyme. If F3'5'h is introgressed, our magenta-flowered samples should share the same LOF mutation(s), be they through structural changes (e.g. indels) or through mutations in conserved domains or SRSs.

#### Results

Whole-Genome Data Confirm Species Relationships and Suggest 2 Separate Origins of Hummingbird Pollination in *Dasanthera* 

Our whole-genome sequencing of 18 *Dasanthera* samples produced ~13× average coverage per sample across approximately 70% of the 8 main scaffolds (assumed to represent the 8 *Penstemon* chromosomes). All information pertaining to read quantities, genome-wide coverage, and per-scaffold coverage can be found in supplementary table S2, Supplementary Material online.

All 3 species tree inference methods produced very similar topologies and support the same main relationships among species within the Cascades and Sierra Nevada Mountains clade. These results corroborate previous inferences made with restriction enzyme-based approaches (Stone and Wolfe 2021; Stone et al. 2023). While the 10-kb window

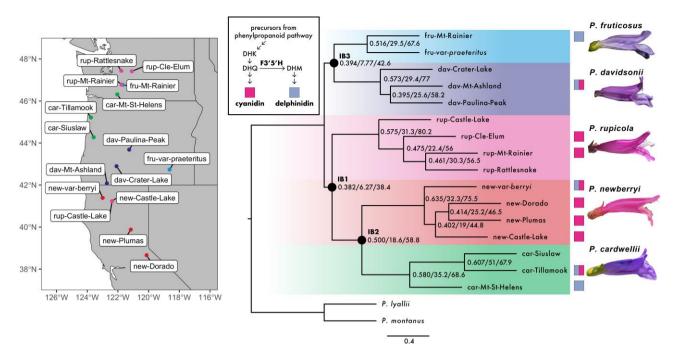


Fig. 1. Sampling sites, species relationships, and anthocyanin pigmentation data for sampled *Dasanthera* taxa. Left: sampling localities for all ingroup samples included in this study. Right: 10-kb genomic window ASTRAL species tree, where tips are individual samples as shown on the map. Numbers on nodes indicate ASTRAL quartet support/gCFs/sCFs, respectively. All ASTRAL local posterior probabilities are 1. Circles at select nodes denote focal internal branches. Magenta and blue-violet squares indicate floral anthocyanidins produced by each sample; pigment extraction was unsuccessful for samples without squares. Inset shows a simplified cartoon version of the anthocyanin biosynthesis pathway where F3'5'H functions to convert the precursor of cyanidin (dihydroquercetin [DHQ]) into the precursor of delphinidin (dihydromyricetin [DHM]). Arrows indicate enzymatic reactions.

ASTRAL-based species tree and the concatenated ML tree produced identical topologies, the CDS ASTRAL-based tree differed in the placement of 1 branch, placing P. rupicola from Cle Elum Lake in central Washington as sister to the remainder of the P. rupicola clade, instead of the sample from Castle Lake in northern California (supplementary fig. S1, Supplementary Material online). From this point forward, we refer to the 10-kb window ASTRAL tree as the species tree (Fig. 1). Despite agreement among alternative species tree topologies and high bootstrap values, quartet support and concordance factors (gCFs and SCFs) were relatively low (Fig. 1). This suggests conflicting phylogenetic signal among window trees (due to ILS and/or introgression) at many branches, especially for relationships deeper in the tree. While the overall level of discordance was relatively high, the most strongly supported internal branch (with the least discordance) is the branch placing P. newberryi and P. cardwellii as sister (Fig. 1; IB2: quartet support = 50%, gCF = 18.6%, and sCF = 58.8%), suggesting strongly that the 2 hummingbird-pollinated species (P. newberryi and P. rupicola) are not sister species.

Anthocyanin pigment data for the 10 available samples confirm that the magenta flowers associated with bird pollination in *P. newberryi* and *P. rupicola* contain pink cyanidin-based anthocyanin pigments only, whereas bee-adapted species produce blue-delphinidin-based pigments, sometimes in combination with cyanidin (results summarized in supplementary table S3, Supplementary Material online). The only exception is the *P. newberryi* var. *berryi* individual that produced both cyanidin and delphinidin. Flowers of

this taxon are not nearly as apparently specialized to bird pollination: the anthers and stigma are inserted within the flower, and the corolla tube may appear blue-violet in color, rather than magenta. Overall, these results align with a previous study reporting a single *P. newberryi* sample produced cyanidin and a single *P. montanus* sample produced both cyanidin and delphinidin (Scogin and Freeman 1987).

Because the production of cyanidin is biochemically simpler than the production of delphinidin, *P. rupicola* and *P. newberryi* (with the exception of the var. *berryi* sample) have experienced a loss of delphinidin production, relative to the bee-pollinated species. Taking the species tree at face value, these results would suggest one of the following scenarios: (i) parallel losses of the ability to produce delphinidin (resulting in magenta flowers) in *P. newberryi* and *P. rupicola*, (ii) a single loss of delphinidin production along the branch leading to *P. newberryi*, *P. cardwellii*, and *P. rupicola*, followed by at least 1 reversal in *P. cardwellii* (and additionally introgression into *P. newberryi* var. *berryi* or a second reversal in this taxon), or (iii) a single loss of delphinidin production, which is shared by *P. newberryi* and *P. rupicola* through either hemiplasy (ILS) or introgression.

CDS Variation at F3'5'h Suggests Separate LOF De Novo Mutations Are Responsible for the Repeated Evolution of Magenta Flowers

Nucleotide sequence alignments of the F3'5'h gene sequence revealed that a diversity of LOF mutations to the CDS

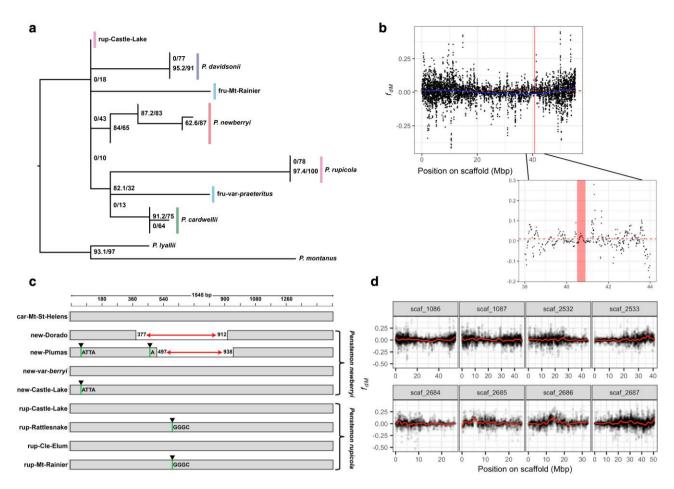


Fig. 2. Patterns of genetic variation at the candidate flower color gene F3'5'h. a) ML gene tree estimated for F3'5'h gene sequence. Numbers at nodes signify the SH-aLRT support (%)/UFBoot support (%). Branches with no SH-aLRT support are collapsed. b) Values of the modified  $f_d$  introgression statistic ( $f_{dM}$ ) for the triplet P. cardwellii-P. newberryi-P. rupicola, calculated in 500 SNP windows, sliding every 250 SNPs. Positive values are consistent with introgression between P. newberryi and P. newberryi in P. newberryi in P. newberryi in P. newberryi in P. newberryi value. c) Visualization of CDS for the P. newberryi in P. newberryi and P.

accompany loss of delphinidin production in P. newberryi and P. rupicola. We identified 2 distinct large deletions (~500 bp) present in 2 P. newberryi samples from Eldorado and Plumas National Forests but absent in all other samples of P. newberryi and P. rupicola (Fig. 2c). These deletions either partially or entirely remove the second exon, undoubtedly abolishing the activity of F3'5'H (Seitz et al. 2007). A third P. newberryi sample from Castle Lake shares a 4-bp insertion with the sample from Plumas that produces a premature stop codon (Fig. 2c; supplementary fig. S2, Supplementary Material online). All 3 of these samples only produce cyanidin in their floral tissues (supplementary table S3, Supplementary Material online). The F3'5'h sequence for P. newberryi var. berryi lacks inactivating mutations, consistent with observed delphinidin production in its floral sample. Two P. rupicola samples from Mt. Rainier and Rattlesnake Ridge shared a 4-bp insertion that results in premature stop codons. We were unable to obtain pigment data for the P. rupicola sample from Rattlesnake Ridge, but the sample from Mt. Rainier

produces only cyanidin in floral tissues (supplementary table S3, Supplementary Material online). The remaining 2 *P. rupicola* samples are less clear. One of these samples (from Cle Elum Lake) produces only cyanidin in flowers and is heterozygous for sites that cause amino acid substitutions in SRS1 and the hydroxylation activity site (supplementary fig. S2, Supplementary Material online)—these replacement mutations could reduce the functionality of F3'5'H. The *P. rupicola* sample from Castle Lake has no suggestive functional mutations in F3'5'h, but we were unable to extract anthocyanidins due to poor tissue quality.

Values of  $f_{dM}$  assessing introgression between *P. newberryi* and *P. rupicola* surrounding the F3'5'h locus were near 0 and were not significantly elevated above the scaffold-wide mean (1-tailed *t*-test: P=0.18), suggesting F3'5'h is unlikely to have been shared between the 2 species through introgression (Fig. 2b). This pattern is apparent regardless of the genomic window size used (supplementary fig. S3, Supplementary Material online). The ML F3'5'h gene tree was discordant

from the species tree, but not in a manner consistent with introgression between *P. newberryi* and *P. rupicola* (Fig. 2a). However, we note that many internal branches of the tree were quite poorly supported, likely due to low information content. In summary, we identified several distinct LOF mutations in *F3′5′h*. Because LOF mutations are difficult to reverse through gene repair, we find a scenario involving a single loss of delphinidin followed by reversals through gene repair (in *P. cardwellii* and *P. newberryi* var. *berryi*) to be highly unlikely. Further, we find no evidence for a history of introgression at the *F3′5′h* locus. Our data are consistent with parallel, independent LOF mutations as the mechanism behind the repeated evolution of magenta flowers in *Dasanthera*.

## Introgression Is Prevalent in Dasanthera but Does Not Suggest a History of Gene Flow Between Hummingbird-Pollinated Species

Our analysis of genome-wide introgression found that many taxa exhibit a history of introgression and admixture, consistent with previous inquiries into the group (Stone and Wolfe 2021; Stone et al. 2023); 64% of ABBA-BABA tests for all trios of samples in our data set were significantly different from 0 (after a Bonferroni correction at  $\alpha = 0.05$ ), providing strong support for widespread gene flow between species (supplementary table S4, Supplementary Material online). The f-branch statistics corroborated these results and identified individual samples with elevated levels of introgression with another taxon, suggesting populationspecific admixture that in some cases may be recent (Fig. 3). Included among these results are (i) the P. rupicola sample from Cle Elum Lake, with introgression from P. fruticosus, (ii) the P. rupicola sample from Castle Lake, with introgression from P. newberryi, and (iii) the P. davidsonii sample from Crater Lake, with introgression from the clade containing P. newberryi, P. cardwellii, and P. rupicola (Fig. 3). In each of these examples, the taxa in question are geographically close enough to exchange genes. As such, these may represent lineages that have experienced relatively recent bouts of introgression. Additionally, we identified significant signatures of introgression between P. davidsonii and both P. newberryi and P. rupicola, a result consistent with the presence of longstanding hybrid zones between these species.

While our F3′5′h sequence analyses suggest independent parallel evolution of magenta flowers in the 2 humming-bird-adapted species, the repeated evolution of other floral traits that make up the hummingbird syndrome (e.g. narrow floral tubes, increased nectar, and elongated reproductive organs) might have evolved through adaptive introgression. In this scenario, we might expect to see a prevailing genomic signal of admixture between P. newberryi and P. rupicola. Contrary to this, our f-branch results find only limited support for introgression between P. newberryi and P. rupicola. Introgression between these species appears to be limited to samples collected together at the same site near Castle Lake in the Trinity Mountains of northern California. This region represents the northernmost and southernmost extent of the distributions of P. newberryi and P. rupicola respectively.

Genome-wide estimates of the D statistic suggest a slightly positive but statistically insignificant signature of introgression between P. newberryi and P. rupicola (D=0.003, P=0.409; supplementary table S5, Supplementary Material online). Moreover,  $f_{dM}$  estimated in windows across the genome failed to identify any clear signatures of introgression between these species (Fig. 2d). After removing the 2 samples from Castle Lake, genome-wide estimates of the D statistic suggest a significant but weak signature of introgression between P. P0. P1. P2. P3. P3. P3. P4. P3. P5. P4. P5. P5. P5. P6. P8. P8. P9. P9.

# Genome-Wide Patterns of Genomic Variation Indicate Selection Against Introgression

Genome-wide patterns of genetic divergence and discordance in Dasanthera are strongly shaped by local gene density (supplementary figs. S4 to S7, Supplementary Material online). We found that both genetic distance  $(d_{xy})$  and differentiation  $(F_{ST})$  between species are significantly positively associated with gene density (supplementary table S6, Supplementary Material online), a pattern that is evident in both 10- and 50-kb windows (supplementary fig. S8, Supplementary Material online). Conversely, normalized RF distance (a measure of topological discordance) of trees calculated from genomic windows relative to the species tree was negatively related to gene density, such that RF distance decreases (window trees become more similar to the species tree) as genic fraction increases (Fig. 4). The topology weights for the main internal branches of the species tree were higher in genomic regions with a larger proportion of CDS (Fig. 4), suggesting a higher fraction of taxa have data conforming to the species tree in gene-dense regions. Moreover, topology weights for the internal branch supporting P. newberryi and P. cardwellii as sister taxa (IB2) were consistently higher across the genome compared to the topology weights for the other 2 internal branches. This pattern is consistent with metrics of branch support (quartet weights and concordance factors) that find stronger support overall for IB2 than for either of the other internal branches (Fig. 4).

In general, we found that introgression is lower in more gene-dense regions of the genome (supplementary fig. S9, Supplementary Material online), consistent with selection against introgression. For 4 of the 8 possible interspecific comparisons, there was a significant negative relationship between  $f_{dM}$  and genic fraction. Curiously, this negative relationship was absent for all 3 comparisons including P. fruticosus: 2 comparisons showed no relationship, and 1 comparison showed a significantly positive relationship. Two species pairs are known to form hybrid zones in secondary contact: P. davidsonii-P. newberryi and P. davidsonii-P. rupicola. We found strong and significant negative relationships between  $f_{dM}$  and genic fraction in both cases. This signal was stronger in P. davidsonii–P. newberryi (regression coefficient = -0.119;  $R^2 = 0.115$ ) than in P. davidsonii-P. rupicola (regression coefficient = -0.065;  $R^2 = 0.054$ ) and, in both cases, is consistent with selection against introgression. We additionally searched for genomic windows (10-kb sliding intervals) exhibiting exceptionally high levels of genetic differentiation. We found a total of 70 outlier  $d_{xy}$  windows

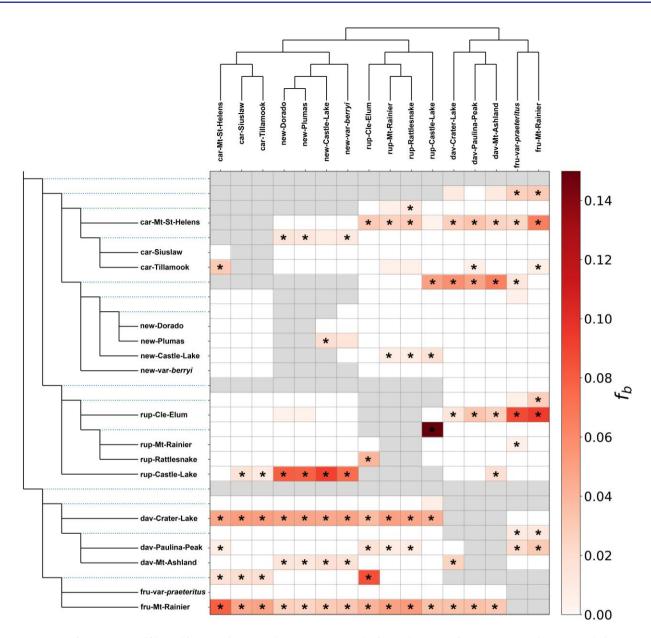


Fig. 3. Patterns of introgression (f-branch) across the *Dasanthera* species tree. The f-branch statistic  $f_b$  is a genome-wide metric to help interpret a system of correlated f4 ratio results and is interpreted as excess allele sharing (i.e. introgression) between branches b (y axis) and tips (x axis) in the species tree. The color gradient represents the  $f_b$  score; gray boxes represent tests that are inconsistent with the species tree topology (tip is descended from sister branch of focal branch b). Statistically significant results (Z-score > 3) are indicated by an asterisk.

that were present in both species pairs (shared outlier windows). These shared outlier windows were distributed across each of the 8 largest scaffolds in the *P. davidsonii* reference genome (supplementary fig. S10, Supplementary Material online). Given their high degree of divergence in the face of hybridization, these regions may be particularly resistant to gene flow and may represent genomic loci underlying reproductive barriers in both hybrid zones.

#### Discussion

Independent LOF Mutations Contribute to Repeated Adaptation in an Adaptive Radiation

Shifts from blue-violet flowers to magenta flowers in 2 Dasanthera species are most likely the result of independent

LOF mutations to a gene in the anthocyanin biosynthesis pathway, F3'5'h. First, this gene is required for the production of delphinidin-based pigments and nonfunctional alleles are present in the magenta-flowered species that lack delphinidin pigments. This result strongly suggests shifts to magenta flowers involve LOF alleles at this gene. Second, distinct LOF alleles are found in *P. newberryi* and *P. rupicola*. Finally, we found no signal of introgression between the 2 magenta-flowered species near the F3'5'h gene, based on the genealogy and values of  $f_{dM}$  surrounding the gene. These findings highlight the role of de novo mutation in the repeated evolution of adaptive phenotypes in an adaptive radiation of plants, which is in contrast to a growing body of research suggesting that parallel evolution in adaptive radiations is often derived from the adaptive introgression of preexisting genetic variation

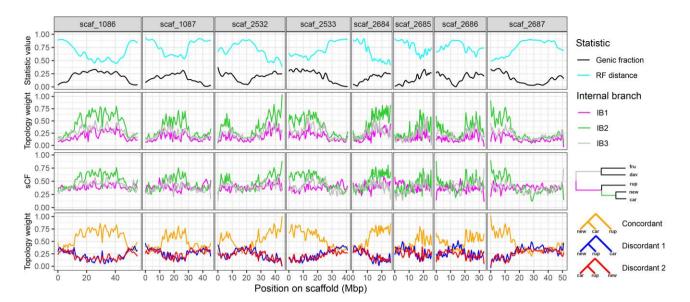


Fig. 4. Patterns of topological discordance across the genome. Top panel: smoothed values for the proportion of sites in coding regions (genic fraction), and the normalized RF distance between 10-kb window trees and the species tree. Second panel: smoothed topology weights for 5-taxon window trees consistent with the 3 main internal branches of the species tree. Third panel: smoothed sCFs averaged over 100 quartets for 10-kb window trees at the 3 main internal branches of the species tree. Bottom panel: smoothed topology weights for the triplet: *P. newberryi* (new), *P. cardwellii* (car), and *P. rupicola* (rup).

(Colosimo et al. 2005; Schluter and Conte 2009; Berner and Salzburger 2015; Marques et al. 2019). It is worth considering that the type of mutation—in this case, one conferring LOF may strongly influence the frequency of de novo mutation in driving parallel phenotypic evolution. LOF mutations are an abundant source of variation in natural populations and may act as hotspots for repeated evolutionary transitions (Martin and Orgogozo 2013; Xu and Guo 2020). For example, LOF mutations are particularly prevalent in the evolution of plant-pollinator interactions; the loss of scent compounds, associated with transitions to hummingbird pollination and self-pollination, has repeatedly and independently evolved in several plant groups, including in Petunia (Segatto et al. 2014; Amrad et al. 2016), Capsella (Sas et al. 2016), and monkeyflowers (Peng et al. 2017; Liang et al. 2023). While there are many examples of LOF mutations conferring adaptive benefit, their occurrence is often limited to genes with few pleiotropic effects or to regulatory elements that affect gene expression only in specific tissues (Xu and Guo 2020). Although transcription factors in the flavonoid pigment pathway have been shown to generally have higher rates of molecular evolution than the structural genes (like F3'5'h) they regulate (Wheeler et al. 2022), LOF mutations to F3'5'h seem to incur minimal pleiotropic effects, perhaps explaining their apparent preponderance. F3'5'h functions to produce tri-hydroxylated flavonoid compounds, including delphinidin. Yet, tri-hydroxylated flavonoids seem to be restricted to floral tissue—many species that have delphinidin-producing flowers produce di-hydroxylated flavonoids (products of the upstream gene F3'h) in vegetative tissues (Wessinger and Rausher 2012). Furthermore, several plant species have evolved red floral pigments through LOF mutations in F3'5'h (Smith and Rausher 2011;

Ishiguro et al. 2012), including other species of *Penstemon* (Wessinger and Rausher 2015), and many other angiosperm taxa lack this gene entirely (Wessinger and Rausher 2012), suggesting the defunctionalization of *F3'5'h* can underlie adaptation without deleterious side effects. While transitions to magenta flowers in *P. newberryi* and *P. rupicola* appear to involve parallel de novo mutations, this does not preclude a role for adaptive introgression fueling other aspects of the hummingbird floral syndrome. However, we found no clear evidence of gene flow between *P. newberryi* and *P. rupicola* genome wide or near the *F3'5'h* locus. While these results do not rule out the potential for adaptive introgression of other floral traits associated with hummingbird pollination syndrome (e.g. exserted stigma and anthers), at this time, we have no evidence to support this hypothesis.

It is also worth considering that the degree of color shift we observe in Dasanthera is not as severe as in other Penstemon taxa that have transitioned to hummingbird pollination. Most other hummingbird-pollinated Penstemon species produce pelargonidin-based anthocyanins (Scogin and Freeman 1987; Wessinger and Rausher 2015), whereas the hummingbird-adapted Dasanthera species produce the more-hydroxylated cyanidin. This difference is observable by the naked eye, such that flowers with cyanidin are pinkmagenta, while flowers with pelargonidin are red. It is possible that the absence of pelargonidin in Dasanthera flowers reflects a more generalized floral phenotype that is thus less specialized to hummingbird pollination. Although hummingbirds are frequent visitors to P. newberryi, they are only present at low and mid elevations throughout their range; a diverse community of potential pollinators visits both P. newberryi and P. davidsonii, which suggests a somewhat generalist phenotype in P. newberryi and explains why hybridization is so common between these species (Kimball 2008). And, while P. newberryi has many characteristics of hummingbird specialization, other features, including small quantities of nectar (Scogin and Freeman 1987; personal observation) and the aforementioned magenta corolla, suggest it has not fully transitioned to a specialized hummingbirdpollinated phenotype. The same may be true for P. rupicola, where (to the authors' knowledge) the only pollinator observation study found that while hummingbirds do visit P. rupicola flowers, the most common floral visitors to P. rupicola were pollen-collecting halictid bees, which are unlikely to be effective pollinators (Datwyler 2001). Conversely to P. newberryi, however, P. rupicola produces large quantities of nectar (personal observation), consistent with specialization to hummingbird pollination. In both cases, the apparent lack of complete pollinator specialization likely plays a large role in facilitating interspecific gene flow, an apparently common phenomenon in Dasanthera.

## Patterns of Discordance and Introgression Are Consistent with Selection Driving Divergence

We found widespread signatures of introgression, corroborating previous inquiries into hybridization dynamics in Dasanthera (Stone and Wolfe 2021; Stone et al. 2023). However, patterns of introgression and phylogenetic discordance were heterogeneous across the genome. The genomic landscape of discordance and introgression are consistent with selection driving species divergence and opposing introgression in the Dasanthera radiation. Trees inferred from genomic windows tended to be more similar to the species tree topology in regions of higher gene density (Fig. 4). This relationship between discordance and gene content is consistent with stronger purifying selection acting in gene-dense regions: purifying selection acts to reduce effective population size hence reducing discordance caused by ILS and acts to oppose introgression, a further source of discordance. Moreover, gene-dense regions exhibited greater differentiation and reduced signals of introgression, consistent with reduced gene flow in gene-rich regions. We expect that levels of differentiation and introgression are also shaped by variation in local recombination rate, a nearuniversal pattern in population genetic data sets (Begun and Aquadro 1992; Cruickshank and Hahn 2014; Martin and Jiggins 2017; Wolf and Ellegren 2017). Specifically, high recombination regions should experience reduced genetic differentiation and tolerate higher levels of introgression. Unfortunately, we currently lack a recombination map for the P. davidsonii reference genome. In a more distantly related Penstemon species (Penstemon barbatus), gene-dense regions tend to have higher recombination rates, resulting in an overall negative relationship between population differentiation and recombination rate (Wessinger et al. 2023).

The 2 hummingbird-adapted species in our study both hybridize with the bee-adapted species *P. davidsonii* in the Cascades and Sierra Nevada Mountains. These sets of species pairs provide an opportunity to examine whether similar genomic barriers to gene flow accompany parallel divergence in

floral syndrome. For each species pair, we found strongly differentiated genomic regions (supplementary fig. \$10, Supplementary Material online). Such localized genomic regions of elevated divergence are expected when there is strong divergent selection in the face of gene flow (Charlesworth et al. 1997; Feder and Nosil 2010). These genomic loci appear resistant to introgression and may potentially be involved in local adaptation to divergent habitats or reproductive isolation. One caveat is that these outlier loci do not truly represent independent regions of elevated divergence, because they include the same samples of P. davidsonii in both contrasts. Future genomic and phenotypic analyses of these replicate hybrid zones will be better suited to determine whether these genomic regions represent parallel loci of elevated divergence and will begin to address whether patterns of phenotypic variation and selection within hybrid zones show similar and/or predictable patterns.

## **Conclusion**

Penstemon subgenus Dasanthera represents a microcosm adaptive radiation, with its 9 species exhibiting a diversity of ecological niches and floral pollination strategies. Species boundaries remain porous in Dasanthera, allowing for interspecific hybridization and highlighting the potential for the adaptive introgression of ecologically important traits. In this study, we focused on the source of genetic variation for repeated origins of a novel flower color in hummingbirdpollinated species, as we currently lack information on the genetic basis for other ecological traits that have diversified in the group. Although our data suggest parallel de novo mutations underlie repeated flower color evolution, it is plausible that adaptive introgression has fueled other aspects of the radiation because evidence for introgression is prevalent in Dasanthera. We suspect that both parallel de novo mutations and introgression have contributed to diversification, as inferred in other rapidly radiating clades.

## **Supplementary Material**

Supplementary material is available at Molecular Biology and Evolution online.

## **Acknowledgments**

We thank members of the Wessinger Lab for discussions on the project and the USFS and NPS for sampling permits. We thank Kate Ostevik and Mark Rausher for access to the *P. davidsonii* genome prior to its publication. The Division of Information Technology at the University of South Carolina provided access to computing resources on the Hyperion HPC cluster. We thank Stacey Smith and an anonymous reviewer for helpful feedback on the manuscript.

## **Funding**

This work was funded by NSF IOS-2209128 (to B.W.S.), NSF DEB-2052904 (to C.A.W.), and NIH NIGMS R35GM142636

(to C.A.W.). Scripts used for this study are available at https://github.com/benstemon/MBE-23-0936.

## **Data Availability**

Raw sequencing data have been deposited into the NCBI Sequence Read Archive under the accession number PRJNA1057825. Other data underlying this article, including processed .vcf files, are available on figshare under doi: 10.6084/m9.figshare.24480499.

#### References

- Amrad A, Moser M, Mandel T, De Vries M, Schuurink RC, Freitas L, Kuhlemeier C. Gain and loss of floral scent production through changes in structural genes during pollinator-mediated speciation. *Curr Biol.* 2016:**26**(24):3303–3312. https://doi.org/10.1016/j.cub.2016.10.023.
- Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
- Begun DJ, Aquadro CF. Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature*. 1992:**356**(6369):519–520. https://doi.org/10.1038/356519a0.
- Berner D, Salzburger W. The genomics of organismal diversification illuminated by adaptive radiations. *Trends Genet.* 2015:**31**(9): 491–499. https://doi.org/10.1016/j.tig.2015.07.002.
- Charlesworth B, Nordborg M, Charlesworth D. The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genet Res.* 1997:**70**(2):155–174. https://doi.org/10.1017/S0016672397002954.
- Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018:**34**(17):i884–i890. https://doi.org/10.1093/bioinformatics/bty560.
- Cicconardi F, Milanetti E, Pinheiro De Castro EC, Mazo-Vargas A, Van Belleghem SM, Ruggieri AA, Rastas P, Hanly J, Evans E, Jiggins CD, et al. Evolutionary dynamics of genome size and content during the adaptive radiation of *Heliconiini* butterflies. *Nat Commun*. 2023:14(1):5620. https://doi.org/10.1038/s41467-023-41412-5.
- Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G Jr, Dickson M, Grimwood J, Schmutz J, Myers RM, Schluter D, Kingsley DM. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science*. 2005:**307**(5717):1928–1933. https://doi.org/10.1126/science.1107239.
- Cruickshank TE, Hahn MW. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol Ecol.* 2014;**23**(13):3133–3157. https://doi.org/10.1111/mec.12796.
- Datwyler SL. Evolution and dynamics of hybridization in *Penstemon* subgenus Dasanthera (Scrophulariaceae S.L.) [Ph.D. dissertation]. [Columbus (OH), USA]: Ohio State University; 2001.
- Datwyler SL, Wolfe AD. Phylogenetic relationships and morphological evolution in *Penstemon* subg. *Dasanthera* (Veronicaceae). *Syst Bot.* 2004:**29**(1):165–176. https://doi.org/10.1600/036364404772974077.
- De-Kayne R, Selz OM, Marques DA, Frei D, Seehausen O, Feulner PGD. Genomic architecture of adaptive radiation and hybridization in Alpine whitefish. *Nat Commun.* 2022:**13**(1):4479. https://doi.org/10.1038/s41467-022-32181-8.
- Edelman NB, Mallet J. Prevalence and adaptive impact of introgression. *Annu Rev Genet.* 2021:**55**(1):265–283. https://doi.org/10.1146/annurev-genet-021821-020805.
- Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004:**32**(5):1792–1797. https://doi.org/10.1093/nar/gkh340.
- Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report.

- Bioinformatics. 2016:**32**(19):3047–3048. https://doi.org/10.1093/bioinformatics/btw354.
- Feder JL, Nosil P. The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. *Evolution*. 2010:**64**(6): 1729–1747. https://doi.org/10.1111/j.1558-5646.2009.00943.x.
- Ferreira MS, Jones MR, Callahan CM, Farelo L, Tolesa Z, Suchentrunk F, Boursot P, Mills LS, Alves PC, Good JM, et al. The legacy of recurrent introgression during the radiation of hares. Syst Biol. 2021:**70**(3):593–607. https://doi.org/10.1093/sysbio/syaa088.
- Giska I, Farelo L, Pimenta J, Seixas FA, Ferreira MS, Marques JP, Miranda I, Letty J, Jenny H, Hackländer K, et al. Introgression drives repeated evolution of winter coat color polymorphism in hares. Proc Natl Acad Sci U S A. 2019:**116**(48):24150–24156. https://doi.org/10.1073/pnas.1910471116.
- Givnish TJ. Adaptive radiation versus 'radiation' and 'explosive diversification': why conceptual distinctions are fundamental to understanding evolution. *New Phytol.* 2015:**207**(2):297–303. https://doi.org/10.1111/nph.13482.
- Gotoh O. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *J Biol Chem.* 1992:**267**(1): 83–90. https://doi.org/10.1016/S0021-9258(18)48462-1.
- Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010:**59**(3):307–321. https://doi.org/10.1093/sysbio/syq010.
- Haenel Q, Roesti M, Moser D, MacColl ADC, Berner D. Predictable genome-wide sorting of standing genetic variation during parallel adaptation to basic versus acidic environments in stickleback fish. *Evol Lett.* 2019:**3**(1):28–42. https://doi.org/10.1002/evl3.99.
- Harborne JB. Methods of plant analysis. *Phytochemical methods: a guide to modern techniques of plant analysis*. Dordrecht: Springer Netherlands; 1984. p. 1–36.
- Harrison RG, Larson EL. Heterogeneous genome divergence, differential introgression, and the origin and structure of hybrid zones. *Mol Ecol.* 2016:**25**(11):2454–2466. https://doi.org/10.1111/mec.13582.
- Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018;35(2):518–522. https://doi.org/10.1093/molbev/msx281.
- Ishiguro K, Taniguchi M, Tanaka Y. Functional analysis of *Antirrhinum kelloggii* flavonoid 3'-hydroxylase and flavonoid 3',5'-hydroxylase genes; critical role in flower color and evolution in the genus *Antirrhinum*. *J Plant Res*. 2012:**125**(3):451–456. https://doi.org/10.1007/s10265-011-0455-5.
- Jones MR, Mills LS, Alves PC, Callahan CM, Alves JM, Lafferty DJR, Jiggins FM, Jensen JD, Melo-Ferreira J, Good JM. Adaptive introgression underlies polymorphic seasonal camouflage in snowshoe hares. Science. 2018:360(6395):1355–1358. https://doi.org/ 10.1126/science.aar5273.
- Jones MR, Mills LS, Jensen JD, Good JM. Convergent evolution of seasonal camouflage in response to reduced snow cover across the snowshoe hare range. Evolution. 2020:74(9):2033-2045. https://doi.org/10.1111/evo.13976.
- Jun G, Wing MK, Abecasis GR, Kang HM. An efficient and scalable analysis framework for variant extraction and refinement from population-scale DNA sequence data. *Genome Res.* 2015:25(6): 918–925. https://doi.org/10.1101/gr.176552.114.
- Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 2017:**14**(6):587–589. https://doi.org/10.1038/nmeth.4285.
- Kimball S. Links between floral morphology and floral visitors along an elevational gradient in a *Penstemon* hybrid zone. *Oikos*. 2008:**117**(7): 1064–1074. https://doi.org/10.1111/j.0030-1299.2008.16573.x.
- Korunes KL, Samuk K. Pixy: unbiased estimation of nucleotide diversity and divergence in the presence of missing data. Mol Ecol Resour. 2021:21(4):1359–1368. https://doi.org/10.1111/1755-0998. 13326.

- Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013. Available from: http://arxiv.org/abs/1303. 3997.
- Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*. 2011:**27**(21): 2987–2993. https://doi.org/10.1093/bioinformatics/btr509.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009:**25**(16):2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
- Liang M, Chen W, LaFountain AM, Liu Y, Peng F, Xia R, Bradshaw HD, Yuan Y-W. Taxon-specific, phased siRNAs underlie a speciation locus in monkeyflowers. *Science*. 2023:**379**(6632):576–582. https://doi.org/10.1126/science.adf1323.
- Maddison WP. Gene trees in species trees. Syst Biol. 1997:46(3): 523–536. https://doi.org/10.1093/sysbio/46.3.523.
- Malinsky M, Challis RJ, Tyers AM, Schiffels S, Terai Y, Ngatunga BP, Miska EA, Durbin R, Genner MJ, Turner GF. Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. Science. 2015:**350**(6267):1493–1498. https://doi.org/10.1126/science.aac9927.
- Malinsky M, Matschiner M, Svardal H. Dsuite—fast D-statistics and related admixture evidence from VCF files. *Mol Ecol Resour*. 2021:**21**(2):584–595. https://doi.org/10.1111/1755-0998.13265.
- Malinsky M, Svardal H, Tyers AM, Miska EA, Genner MJ, Turner GF, Durbin R. Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. *Nat Ecol Evol.* 2018:**2**(12):1940–1955. https://doi.org/10.1038/s41559-018-0717-x.
- Marcionetti A, Salamin N. Insights into the genomics of clownfish adaptive radiation: the genomic substrate of the diversification. *Genome Biol Evol.* 2023:**15**(7):evad088. https://doi.org/10.1093/gbe/evad088.
- Marques DA, Jones FC, Di Palma F, Kingsley DM, Reimchen TE. Genomic changes underlying repeated niche shifts in an adaptive radiation. *Evolution*. 2022:**76**(6):1301–1319. https://doi.org/10.1111/evo.14490.
- Marques DA, Meier JI, Seehausen O. A combinatorial view on speciation and adaptive radiation. *Trends Ecol Evol*. 2019:**34**(6): 531–544. https://doi.org/10.1016/j.tree.2019.02.008.
- Martin SH, Davey JW, Jiggins CD. Evaluating the use of ABBA-BABA statistics to locate introgressed loci. *Mol Biol Evol.* 2015:**32**(1): 244–257. https://doi.org/10.1093/molbev/msu269.
- Martin SH, Jiggins CD. Interpreting the genomic landscape of introgression. *Curr Opin Genet Dev.* 2017:**47**:69–74. https://doi.org/10.1016/j.gde.2017.08.007.
- Martin A, Orgogozo V. The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution*. 2013:**67**(5): 1235–1250. https://doi.org/10.1111/evo.12081.
- Martin SH, Van Belleghem SM. Exploring evolutionary relationships across the genome using topology weighting. *Genetics*. 2017:**206**(1):429–438. https://doi.org/10.1534/genetics.116.194720.
- McGee MD, Borstein SR, Meier JI, Marques DA, Mwaiko S, Taabu A, Kishe MA, O'Meara B, Bruggmann R, Excoffier L, et al. The ecological and genomic basis of explosive adaptive radiation. *Nature*. 2020:**586**(7827):75–79. https://doi.org/10.1038/s41586-020-2652-7.
- Minh BQ, Hahn MW, Lanfear R. New methods to calculate concordance factors for phylogenomic datasets. *Mol Biol Evol*. 2020:**37**(9): 2727–2733. https://doi.org/10.1093/molbev/msaa106.
- Moran BM, Payne C, Langdon Q, Powell DL, Brandvain Y, Schumer M. The genomic consequences of hybridization. *eLife*. 2021:**10**: e69016. https://doi.org/10.7554/eLife.69016.
- Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 2015:**32**(1):268–274. https://doi.org/10.1093/molbev/msu300.
- Ostevik KL, Alabady M, Zhang M, Rausher MD. Genome report: whole genome sequence and annotation of *Penstemon*

- davidsonii. G3 (Bethesda). 2023:jkad296. https://doi.org/10. 1093/g3journal/jkad296.
- Pamilo P, Nei M. Relationships between gene trees and species trees. *Mol Biol Evol.* 1988:**5**(5):568–583. https://doi.org/10.1093/oxfordjournals.molbev.a040517.
- Peng F, Byers KJRP, Bradshaw HD Jr. Less is more: independent loss-of-function *OCIMENE SYNTHASE* alleles parallel pollination syndrome diversification in monkeyflowers (*Mimulus*). *Am J Bot.* 2017:**104**(7):1055–1059. https://doi.org/10.3732/ajb.1700104.
- Pertea G, Pertea M. GFF utilities: GffRead and GffCompare. F1000Res. 2020:9:ISCB Comm J-304. https://doi.org/10.12688/f1000research. 23297.2.
- Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010:**26**(6):841–842. https://doi.org/10.1093/bioinformatics/btq033.
- Robinson DF, Foulds LR. Comparison of phylogenetic trees. *Math Biosci.* 1981:**53**(1-2):131–147. https://doi.org/10.1016/0025-5564(81)90043-2.
- Ronco F, Matschiner M, Böhne A, Boila A, Büscher HH, El Taher A, Indermaur A, Malinsky M, Ricci V, Kahmen A, et al. Drivers and dynamics of a massive adaptive radiation in cichlid fishes. *Nature.* 2021:**589**(7840):76–81. https://doi.org/10.1038/s41586-020-2930-4.
- Rosenberg NA. The probability of topological concordance of gene trees and species trees. *Theor Popul Biol.* 2002:**61**(2):225–247. https://doi.org/10.1006/tpbi.2001.1568.
- Sas C, Müller F, Kappel C, Kent TV, Wright SI, Hilker M, Lenhard M. Repeated inactivation of the first committed enzyme underlies the loss of benzaldehyde emission after the selfing transition in Capsella. *Curr Biol.* 2016:26(24):3313–3319. https://doi.org/10.1016/j.cub.2016.10.026.
- Schliep KP. Phangorn: phylogenetic analysis in R. *Bioinformatics*. 2011:**27**(4):592–593. https://doi.org/10.1093/bioinformatics/btq706.
- Schluter D. *The ecology of adaptive radiation*. Oxford: OUP; 2000.
- Schluter D, Conte GL. Genetics and ecological speciation. *Proc Natl Acad Sci U S A*. 2009:**106 Suppl 1**(Suppl 1):9955–9962. https://doi.org/10.1073/pnas.0901264106.
- Scogin R, Freeman CE. Floral anthocyanins of the genus *Penstemon*: correlations with taxonomy and pollination. *Biochem Syst Ecol.* 1987:**15**(3):355–360. https://doi.org/10.1016/0305-1978(87)90011-1.
- Seehausen O. Hybridization and adaptive radiation. *Trends Ecol Evol.* 2004:19(4):198–207. https://doi.org/10.1016/j.tree.2004.01.003.
- Segatto ALA, Cazé ALR, Turchetto C, Klahre U, Kuhlemeier C, Bonatto SL, Freitas LB. Nuclear and plastid markers reveal the persistence of genetic identity: a new perspective on the evolutionary history of *Petunia exserta*. *Mol Phylogenet Evol*. 2014:**70**: 504–512. https://doi.org/10.1016/j.ympev.2013.10.011.
- Seitz C, Ameres S, Forkmann G. Identification of the molecular basis for the functional difference between flavonoid 3'-hydroxylase and flavonoid 3',5'-hydroxylase. FEBS Lett. 2007:**581**(18): 3429–3434. https://doi.org/10.1016/j.febslet.2007.06.045.
- Short AW, Streisfeld MA. Ancient hybridization leads to the repeated evolution of red flowers across a monkeyflower radiation. *Evol Lett.* 2023:**7**(5):293–304. https://doi.org/10.1093/evlett/qrad024.
- Simpson GG. The major features of evolution. New York, USA: Columbia University Press; 1953.
- Smith SD, Rausher MD. Gene loss and parallel evolution contribute to species difference in flower color. *Mol Biol Evol*. 2011:**28**(10): 2799–2810. https://doi.org/10.1093/molbev/msr109.
- Stankowski S, Streisfeld MA. Introgressive hybridization facilitates adaptive divergence in a recent radiation of monkeyflowers. Proc Biol Sci. 2015:282 (1814):20151666. https://doi.org/10.1098/rspb.2015.1666.
- Stevens JTE, Wheeler LC, Williams NH, Norton AM, Wessinger CA. Predictive links between petal color and pigment quantities in natural Penstemon hybrids. Integr Comp Biol. 2023:63(6): 1340–1351. doi:10.1093/icb/icad073.
- Stone BW, Rodríguez-Peña RA, Wolfe AD. Testing hypotheses of hybrid taxon formation in the shrubby beardtongues (*Penstemon subgenus Dasanthera*). Am J Bot. 2023:**110**(1):e16118. https://doi.org/10.1002/ajb2.16118.



- Stone BW, Wolfe AD. Phylogeographic analysis of shrubby beardtongues reveals range expansions during the Last Glacial Maximum and implicates the Klamath Mountains as a hotspot for hybridization. *Mol Ecol.* 2021;**30**(15):3826–3839. https://doi.org/10.1111/mec.15992.
- Straw RM. Hybridization, homogamy, and sympatric speciation. *Evolution*. 1955:**9**(4):441–444. https://doi.org/10.2307/2405478.
- Suvorov A, Kim BY, Wang J, Armstrong EE, Peede D, D'Agostino ERR, Price DK, Waddell PJ, Lang M, Courtier-Orgogozo V, et al. Widespread introgression across a phylogeny of 155 Drosophila genomes. Curr Biol. 2022;32(1):111–123.e5. https://doi.org/10. 1016/j.cub.2021.10.052.
- Tanaka Y, Brugliera F. Flower colour and cytochromes P450. *Philos Trans R Soc Lond B Biol Sci.* 2013:**368**(1612):20120432. https://doi.org/10.1098/rstb.2012.0432.
- Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution. 1984:**38**(6):1358–1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x.
- Wessinger CA, Freeman CC, Mort ME, Rausher MD, Hileman LC. Multiplexed shotgun genotyping resolves species relationships within the North American genus *Penstemon. Am J Bot.* 2016:**103**(5):912–922. https://doi.org/10.3732/ajb.1500519.
- Wessinger CA, Katzer AM, Hime PM, Rausher MD, Kelly JK, Hileman LC. A few essential genetic loci distinguish *Penstemon* species with flowers adapted to pollination by bees or hummingbirds. *PLoS Biol.* 2023:**21**(9):e3002294. https://doi.org/10.1371/journal.pbio.3002294.
- Wessinger CA, Rausher MD. Lessons from flower colour evolution on targets of selection. *J Exp Bot.* 2012:**63**(16):5741–5749. https://doi.org/10.1093/jxb/ers267.
- Wessinger CA, Rausher MD. Predictability and irreversibility of genetic changes associated with flower color evolution in *Penstemon barbatus*. *Evolution*. 2014:**68**(4):1058–1070. https://doi.org/10.1111/evo.12340.
- Wessinger CA, Rausher MD. Ecological transition predictably associated with gene degeneration. *Mol Biol Evol*. 2015:**32**(2): 347–354. https://doi.org/10.1093/molbev/msu298.

- Wheeler LC, Walker JF, Ng J, Deanna R, Dunbar-Wallis A, Backes A, Pezzi PH, Palchetti MV, Robertson HM, Monaghan A, et al. Transcription factors evolve faster than their structural gene targets in the flavonoid pigment pathway. Mol Biol Evol. 2022:39(3): msac044. https://doi.org/10.1093/molbev/msac044.
- Whitfield JB, Lockhart PJ. Deciphering ancient rapid radiations. *Trends Ecol Evol*. 2007:**22**(5):258–265. https://doi.org/10.1016/j.tree.2007.01.012.
- Wilson P, Castellanos MC, Hogue JN, Thomson JD, Armbruster WS. A multivariate search for pollination syndromes among *Penstemons. Oikos.* 2004:**104**(2):345–361. https://doi.org/10.1111/j.0030-1299.2004.12819.x.
- Wilson P, Wolfe AD, Armbruster WS, Thomson JD. Constrained lability in floral evolution: counting convergent origins of hummingbird pollination in *Penstemon* and *Keckiella*. *New Phytologist*. 2007:**176**(4): 883–890. https://doi.org/10.1111/j.1469-8137.2007.02219.x.
- Wolf JBW, Ellegren H. Making sense of genomic islands of differentiation in light of speciation. *Nat Rev Genet*. 2017:**18**(2):87–100. https://doi.org/10.1038/nrg.2016.133.
- Wolfe AD, Blischak PD, Kubatko LS. Phylogenetics of a Rapid, Continental Radiation: Diversification, Biogeography, and Circumscription of the Beardtongues (*Penstemon*; Plantaginaceae); 2021. Available from: http://biorxiv.org/lookup/doi/10.1101/2021.04.20.440652
- Wolfe AD, Randle CP, Datwyler SL, Morawetz JJ, Arguedas N, Diaz J. Phylogeny, taxonomic affinities, and biogeography of *Penstemon* (Plantaginaceae) based on ITS and cpDNA sequence data. *Am J Bot* 2006:**93**(11):1699–1713. https://doi.org/10.3732/ajb.93.11. 1699.
- Xu Y-C, Guo Y-L. Less is more, natural loss-of-function mutation is a strategy for adaptation. *Plant Commun.* 2020:1(6):100103. https://doi.org/10.1016/j.xplc.2020.100103.
- Zhang C, Rabiee M, Sayyari E, Mirarab S. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics. 2018:**19**(Suppl 6):153. https://doi.org/10. 1186/s12859-018-2129-y.