Phylogeography and conservation of *Pedicularis* (Orobanchaceae) in the Hengduan Mountains of SW China and Tibet

Jared B. Meek

Submitted in partial fulfillment of the requirements for the degree of Master of Arts in Ecology, Evolution and Conservation Biology Department of Ecology, Evolution and Environmental Biology

COLUMBIA UNIVERSITY

May 19, 2020

Phylogeography and conservation of *Pedicularis* **(Orobanchaceae) in the Hengduan Mountains of SW China and Tibet**

Jared Meek1, Sandra Hoffberg1, Kasielis Molina-Velez1,2, Richard Ree3, Deren Eaton1

¹Department of Ecology, Evolution and Environmental Biology, Columbia University, New York, NY

² Lehman College, The City University of New York, Bronx, NY

³ The Field Museum and The University of Chicago, Chicago, IL

Abstract

The Hengduan Mountains Region (HMR) of southwest China is a temperate biodiversity hotspot characterized by high rates of plant endemism, including approximately 300 endemic species of *Pedicularis* (Orobanchaceae). Intersecting processes of mountain uplift during the early Oligocene (\sim 30 Mya), followed by monsoon intensification during the mid-Miocene (\sim 8-10 Mya) and glacial cycles throughout the Quaternary, have contributed to high rates of speciation within this genus due to population isolation in allopatry. In order to accurately identify the major geographic barriers influencing this speciation process, genomic data (RADseq) and comparative phylogeographic methods were used to characterize the history of population divergence among six widespread species of *Pedicularis*. Three biogeographic regions, which are delineated by four main tributaries of the Upper Yangtze River system, were identified within the HMR. In recent decades, infrastructure development has substantially altered this mountain landscape - connecting previously isolated areas by roads and tunnels - with the potential to threaten endemic biodiversity through genetic homogenization. This inferred phylogeographic history of *Pedicularis* within the HMR can be used in future research to determine whether recent gene flow has occurred across historical barriers due to human influence, thus aiding plant conservation efforts in this biodiversity hotspot.

Introduction

Mountain ranges are among the most dynamic and heterogeneous landscapes on Earth. Extended periods of orogeny and erosion, along with climatic and glacial fluctuations, have dramatically shaped their attendant biodiverse communities over vast time scales (Antonelli et al., 2018; Rahbek et al., 2019). Phylogeographic research has highlighted the impact of mountain uplift on rapid plant radiations in young mountain systems (Hughes & Eastwood, 2006; Xing & Ree, 2017; Drummond et al., 2012; Ye et al., 2019), while others have demonstrated that glaciation cycles and climatic shifts can promote exceptional rates of plant diversification

regardless of mountain age (Vasconcelos et al., 2020). In recent years, the majority of mountain biodiversity research has focused on mountain ranges in the tropics, especially the recently uplifted northern Andes (Nevado et al., 2018; Lagomarsino et al., 2016), but there is growing interest in comparing the patterns of biodiversity accumulation and evolutionary radiation in mountain ranges among varying latitudes across the globe (Hughes & Atchison, 2015).

The Tibeto-Himalayan Region (THR) (Mosbrugger et al., 2018) is a geographic complex in western China composed of three unique but interacting high-elevation systems: the Qinghai-Tibet Plateau (QTP), the Himalaya, and the Hengduan Mountains Region (HMR) (for comprehensive reviews, see Favre et al., 2015 and Muellner-Riehl, 2019). While there is considerable debate concerning the timing of initial orogenic activity between these three systems (Spicer et al., 2020; Renner, 2016), the north-south trending Hengduan Mountains have consistently been described as younger in origin (most recent estimates at ~30 Mya) than the east-west facing Himalaya (~40 Mya) and the Qinghai-Tibet Plateau (formation beginning ~40- 55 Mya). For plants, a portion of this system is also referred to as the Sino-Himalaya floristic subregion, which more specifically encompasses the southeast QTP, eastern Himalaya, and the Hengduan Mountains (Zhou et al., 2013; Wu & Wang, 1983). Due to the intersecting influences of tectonic uplift and the Asian monsoon, this region has been described as a 'cradle' of angiosperm evolution, with much of the contemporary flora of western China diverging within the last 15-19 Mya (Lu et al., 2018). This study focuses on biennial plant populations within the HMR (Figure 1), a temperate biodiversity hotspot connecting the QTP with the Yunnan-Guizhou Plateau and lowland tropics of southern China, Myanmar (Burma), and India (Boufford, 2014). The north-south directionality of the Hengduan Mountains is an important climatic factor

because it enables the region to receive warm, moist air from monsoons in the Indian and Pacific Oceans (Luo et al., 2017).

Over the past decade, multiple phylogeographic scenarios have emerged to describe the impacts of geologic upheaval, Quaternary glacial cycles, and climatic transitions on plant diversification across the THR. Specifically, the "contraction/recolonization" and "microrefugia" patterns (Gao et al. 2016) describe species that retreated into the lower-elevation Hengduan Mountains during glacial periods, followed by expansion back onto the QTP during interglacials, with some lineages becoming isolated and forming local endemics. While many studies have focused on these large-scale dynamics for multiple species across the entire THR, descriptions of spatially-explicit geographic barriers are few, and investigations into Quaternary glacial processes on population structure on a local-scale within the Hengduan Mountains are needed.

Of primary interest to this study is the phylogeographic scenario in which tectonic activity and river dynamics in the Hengduan Mountains have led to population isolation and endemic species formation (Qiu, Fu & Comes, 2011). Processes of mountain uplift, glaciation, and climatic transitions have formed a rich, heterogeneous landscape in the Hengduan Mountains region, with elevational gradients contributing to niche specialization and high levels of endemic plant species due to *in situ* speciation (Xing & Ree, 2017). An outstanding question is how the Last Glacial Maximum (ending 19-20 ka; Clark et al., 2009) interacted with pre-existing topography to influence infraspecific population demographics, and whether different plant species share spatio-temporally similar population patterns across this heterogeneous landscape.

Detailed case studies of endemic plant clades can help describe the processes driving diversification in biodiversity hotspots. *Pedicularis* (Orobanchaceae) is a genus of hemi-parasitic plants with highly diverse floral morphologies. It includes ca. 600 species worldwide, more than

half of which are native to China and endemic to the Hengduan Mountains, where their greatest diversity can be found (Xing & Ree, 2017; Eaton & Ree, 2013; Jie et al., 2004; Zhou et al., 2013). The crown age for *Pedicularis* has been estimated from pollen fossil-calibrated molecular phylogenies to be ~ 39 Mya (Wang et al., 2015; Song et al., 2004), placing its origin within the same timeframe as the putative origins of the QTP, Himalaya, and HMR (Spicer et al., 2020). While a comprehensive phylogeny of this group is under development (Yu et al., 2015), the phylogeographic histories of *Pedicularis* species endemic to the Hengduan Mountains remains to be explored.

It is common to find multiple *Pedicularis* species co-occurring and flowering synchronously throughout the Hengduan Mountains (Eaton et al., 2012; Armbruster et al., 2014) (Figure 1), which makes it likely that these species have experienced parallel histories of migration, expansion and gene flow. This hypothesis can be tested by identifying barriers of geographic isolation that have led to speciation within *Pedicularis* since the Hengduan Mountains started forming ~30 Mya, along with evidence of more recent glacial and monsoon influences on population structure within the past 8-10 Mya. The concentration of *Pedicularis* diversity in the Hengduan Mountains provides an opportunity to connect population genetic scale processes with phylogenetic scale patterns of diversification by treating each species as a replicated experiment. Reconstructing the phylogeographic history of many species can provide information about how the landscape has shaped and influenced the diversification of *Pedicularis*. If closely related species have historically been influenced by similar landscape barriers and abiotic factors, then it is expected that replicate analyses across these species will show similar timing and order of phylogeographic splits among populations sampled from different geographic regions.

Figure 1: A) Maximum likelihood phylogenetic tree of the six *Pedicularis* species used in this study; B) Map of *Pedicularis* sampling locations from expeditions in 2018 and 2019 in the Hengduan Mountains (Yunnan, Sichuan, Qinghai and Gansu Provinces), demonstrating significant co-occurrence of these six species. Inset shows the location of the Hengduan Mountains Region (HMR) within the larger Tibeto-Himalayan Region (THR) and Sino-Himalaya floristic subregion.

Methods

Sample collection

733 samples of 96 different *Pedicularis* species were collected from the Hengduan Mountains in Yunnan, Sichuan, Qinghai and Gansu Provinces during expeditions in July 2018 and 2019. Leaf tissue was dried and preserved in silica gel and stored at room temperature at Columbia University, and replicate vouchers for all specimens were deposited at the Kunming Institute of Botany, Harvard University Herbaria, and the William & Lynda Steere Herbarium at the New York Botanical Garden. In addition to these contemporary samples, leaf tissue and DNA extracts from samples collected on previous expeditions to the Hengduan Mountains by HUH and FM herbaria researchers [\(http://hengduan.huh.harvard.edu/fieldnotes\)](http://hengduan.huh.harvard.edu/fieldnotes) were obtained to increase the spatial and temporal extent of *Pedicularis* sampling.

To increase the likelihood of uncovering parallel population structure among multiple taxa, six widespread species were selected from these collections: *Pedicularis lachnoglossa*, *Pedicularis rhinanthoides*, *Pedicularis integrifolia*, *Pedicularis kansuensis*, *Pedicularis cranolopha*, and *Pedicularis longiflora* (Table 1; listed in order of phylogenetic relatedness). Some have subspecies descriptions, but these are only roughly associated with geography (Flora of China, Yang, Holmgren, & Mill, 1998). After these samples were narrowed to species of direct interest for this study, the final dataset contained 153 samples belonging to 6 species for phylogeographic analysis (Supplement 1).

Extraction and Sequencing

Genomic DNA was extracted from silica-preserved leaf tissue using a modified CTAB protocol. Samples were quantified using a Qubit 4 Fluorometer (dsDNA broad range) to assess quality of extraction. All samples at or above 10 ng/ul of gDNA were suspended in TLE buffer and stored at -80˚C.

For DNA sequencing, the 3RAD (Bayona-Vásquez et al., 2019) library preparation protocol, a version of ddRAD (Andrews et al., 2016) which uses three enzymes to randomly fragment the entire genome at specific cut sites in all samples, was implemented to increase the number of loci that can be compared among species. This 3RAD protocol enabled the incorporation of many more genomic loci than previous studies of plant phylogeography in the Hengduan Mountains.

Samples were sequenced on three lanes of an Illumina HiSeqX in pooled libraries containing other samples, resulting in 306M paired-end reads. After demultiplexing, an average of 1.28M paired-end reads (stdev=915.2K) per sample were recovered.

Table 1: Description of species analyzed for this study, with information gathered from Yu et al. (2015) and the Flora of China (Yang, Holmgren, & Mill, 1998).

RAD-seq data assembly

Sequenced reads from the three Illumina lanes were merged together into one dataset and assembled using the *de novo* pipeline in ipyrad v.0.9.50 (Eaton & Overcast, 2020; [https://ipyrad.readthedocs.io/en/latest/\)](https://ipyrad.readthedocs.io/en/latest/). The 'datatype' parameter was set to 'pair3rad', and the 'mindepth majrule' parameter was set to '1' ($1x = low$ depth) in order to retain the maximum number of loci for each individual sample. All other parameters were left at their default values.

After assembling all samples to low depth, the 'branch' function was used to cluster and format a subset of samples corresponding to each of the six *Pedicularis* species (i.e. retain all *P. cranolopha* + outgroup samples and re-run ipyrad steps 6 & 7, repeated for each species). After clustering and formatting for each species, any samples that retained fewer than 1,000 loci were removed from the dataset.

Phylogenomic analyses

A series of analyses were performed using the ipyrad-analysis toolkit (Eaton & Overcast, 2020), which provides a convenient wrapper around many evolutionary inference tools while also providing a means to filter data for each analysis to limit the influence of missing data, which is a common feature of RAD-seq. Samples from each of the six *Pedicularis* species were run through a pipeline of statistical methods designed to visualize genetic similarity, define population structure, and reconstruct phylogeographic history through estimates of divergence times and demographics of population lineages.

Maximum likelihood phylogenies were inferred for each species dataset with RAxML v.8 (Stamatkis, 2014), using concatenated alignments that included all loci shared across at least four samples. As an initial hypothesis of population structure, samples were grouped into discrete

populations according to the clades from each ML tree topology. Population structure and genetic similarity of samples was then inferred and visualized using principal component analysis (PCA) and STRUCTURE v.2.3.4 (Pritchard et al., 2000). For PCA and STRUCTURE analyses, the ML topology-based population hypotheses were used to create an 'imap' dictionary with a 'minmap' filter of 0.5 to include only genomic loci that achieved at least 50% coverage in each population. An additional 'mincov' filter of 0.75 was applied to remove any loci that were not present in at least 75% of all samples. Missing data was imputed for PCA analyses using the sample imputation method. Burn-in length for STRUCTURE analyses was set to 1e5, number of replicates was set to 5e5, and five replicates were run for each value of 'K'. Initial results from PCA (amount of genetic variance between hypothesized populations) and STRUCTURE plots ('K' parameter corresponding to number of populations) were used to reassign samples to different population groups for reanalysis through an iterative process until population assignments in each 'imap' showed congruence with the population discretization in PCA and STRUCTURE results.

After putative population assignments were validated, ancestral population sizes and species divergence estimates were inferred with BPP v.4.0 (Yang, 2015; Flouri et al., 2018), a Bayesian MCMC program for analyzing multi-locus genomic sequence data under the multispecies coalescent. BPP utilizes an 'imap' of sample populations and a fixed tree topology, along with prior probabilities of generation time, mutation rate, 'theta' (effective population size), and 'tau' (divergence time). The same 'imap' dictionaries from PCA and STRUCTURE analyses were used in the BPP analysis for each corresponding species, but with a 'minmap' set to 0.75 to sample only loci with at least 75% coverage in each population. Simplified tree topologies based on the aforementioned population assignments were used as fixed trees for the

 $'A00'$ inference method ('speciesdelimitation' = 0 and 'speciestree' = 0). Identical priors for generation time (min = 1, max = 3), mutation rate (min = 5e-9, max = 1e-8), and 'theta' $(3, 1)$ 0.007) parameters were used for all species. Priors for the 'tau' parameter varied for each species, and were assigned based on the distance between species in the most recently published *Pedicularis* phylogeny (Yu et al., 2015) as follows: *P. lachnoglossa* (3, 0.007), *P. rhinanthoides* (3, 0.006), *P. integrifolia* (3, 0.0035), *P. kansuensis* (3, 0.0035), *P. cranolopha* (3, 0.002), and *P. longiflora* (3, 0.002). Additional parameters for maximum loci (=500), MCMC burn-in length $(=1,000)$, number of samples $(=10,000)$, and sample frequency $(=10)$ were set identically for all species. Four replicates of the 'A00' inference method were run for each species, and results were calibrated to geological time estimates using a Python implementation of R code developed for converting BPP units for 'theta' and 'tau' into Ne and divergence times (dos Reis, 2017; Yoder et al., 2016).

Results

RAxML

RAxML tree results for all six species had high bootstrap values ($> 70\%$ along the backbone) between clades, providing evidence that samples from all six species are segregating into distinct populations across the landscape (Figures 2-7). *P. rhinanthoides* (13.4 Mbp, 911.6 K SNPs, 79.61% missing data) was the only species with complete bootstrap support (100%), followed by >75 % bootstrap support along the backbone for *P. integrifolia* (8.7 Mbp, 468 K SNPs, 73.77% missing data), *P. lachnoglossa* (9.5 Mbp, 560 K SNPs, 74.5% missing data), and *P. longiflora* (10.1 Mbp, 643 K SNPs, 78.96% missing data), and <75 % bootstrap support along

the backbone for *P. cranolopha* (10.3 Mbp, 628 K SNPs, 76.23% missing data) and *P. kansuensis* (21.5 Mbp, 1.3 M SNPs, 83.64% missing data).

PCA and STRUCTURE

The number of SNPs used for the analysis of genetic similarity between samples varied across species (Supplement 2). PCA variance for each species was best explained as follows: *P. lachnoglossa* by PC axes 0, 1 (15.9%, 10.8%), *P. rhinanthoides* by PC axes 0, 1 (16.1%, 10.5%), *P. integrifolia* by PC axes 1, 2 (16.9%, 9.0%), *P. kansuensis* by PC axes 0, 1 (21.9%, 8.9%), *P. cranolopha* by PC axes 1, 2 (10.6%, 8.8%), and *P. longiflora* by PC axes 2, 3 (10.4%, 8.5%).

After an iterative population assignment process (see Methods), the final PCA and STRUCTURE plots corresponded with 'population' clades in the ML tree topologies. Based on the most consistent population structure across all three analyses, the number of putative populations per species were as follows: *P. lachnoglossa* (K=8), *P. rhinanthoides* (K=5), *P. integrifolia* (K=5), *P. kansuensis* (K=7), *P. cranolopha* (K=7), and *P. longiflora* (K=6).

BPP

Estimates of divergence times and effective population sizes varied between all six species. While BPP produced parameter estimates for each node in the fixed phylogenetic trees it was given, the most directly comparable are the estimates for the ancestral populations of each species (i.e., first node after divergence from the *P. bidentata* outgroup). Divergence times and effective population sizes were as follows: *P. lachnoglossa* $(.185 \text{ Mya} \pm .059; 283 \text{K} \pm 54 \text{K})$, *P. rhinanthoides* (.368 Mya \pm .122; 123K \pm 111K), *P. integrifolia* (.547 Mya \pm .217; 128K \pm 128K), *P. kansuensis* (.256 Mya \pm .085; 151K \pm 105K), *P. cranolopha* (.558 Mya \pm .192; 249K \pm 267K), and *P. longiflora* (.670 Mya \pm .213; 137K \pm 96K).

Figure 2: Results for *Pedicularis lachnoglossa*. Species 'population' tree was generated by RAxML, and populations were assigned based on PCA and STRUCTURE analyses. Estimates of divergence time and effective population size were inferred with BPP. Samples were color-coded to match ML tree visualization, and plotted onto a map of the Hengduan Mountains region using geographic coordinates from fieldnotes (Supplement 1).

Figure 3: Results for *Pedicularis rhinanthoides*. Species 'population' tree was generated by RAxML, and populations were assigned based on PCA and STRUCTURE analyses. Estimates of divergence time and effective population size were inferred with BPP. Samples were color-coded to match ML tree visualization, and plotted onto a map of the Hengduan Mountains region using geographic coordinates from fieldnotes (Supplement 1).

Figure 4: Results for *Pedicularis integrifolia*. Species 'population' tree was generated by RAxML, and populations were assigned based on PCA and STRUCTURE analyses. Estimates of divergence time and effective population size were inferred with BPP. Samples were color-coded to match ML tree visualization, and plotted onto a map of the Hengduan Mountains region using geographic coordinates from fieldnotes (Supplement 1).

Figure 5: Results for *Pedicularis kansuensis*. Species 'population' tree was generated by RAxML, and populations were assigned based on PCA and STRUCTURE analyses. Estimates of divergence time and effective population size were inferred with BPP. Samples were color-coded to match ML tree visualization, and plotted onto a map of the Hengduan Mountains region using geographic coordinates from fieldnotes (Supplement 1).

Figure 6: Results for *Pedicularis cranolopha*. Species 'population' tree was generated by RAxML, and populations were assigned based on PCA and STRUCTURE analyses. Estimates of divergence time and effective population size were inferred with BPP. Samples were color-coded to match ML tree visualization, and plotted onto a map of the Hengduan Mountains region using geographic coordinates from fieldnotes (Supplement 1).

Figure 7: Results for *Pedicularis longiflora*. Species 'population' tree was generated by RAxML, and populations were assigned based on PCA and STRUCTURE analyses. Estimates of divergence time and effective population size were inferred with BPP. Samples were color-coded to match ML tree visualization, and plotted onto a map of the Hengduan Mountains region using geographic coordinates from fieldnotes (Supplement 1).

Discussion

Phylogeography and Speciation

The population structure identified among all six *Pedicularis* species highlights three distinct regions within the Hengduan Mountains landscape that are divided by major elevational differences or river valleys. Currently, the four river valleys formed by the Jinsha, Yajiang, Dadu and Min Rivers appear to be the clearest geographic barriers inhibiting gene flow between populations of *Pedicularis* species analyzed in this study (Figures 2-7, 9).

The Yangtze River system is a dominant feature of the landscape in Sichuan Province (Sun et al., 2013). The three aforementioned rivers are tributaries to the Yangtze that are important for understanding population divergence among *Pedicularis* species. The Jinsha River (Jinshaxiang) is the upper portion of the Yangtze River, and forms a boundary between eastern Tibet and western Sichuan. The Yajiang River starts on the QTP and passes through the center of the sampled region between the cities of Garze to Yajiang before flowing into the Yangtze River from the north. This is a putative biogeographic barrier between the 'Jinsha-Yajiang' populations and the 'Yajiang-Dadu' populations. The Dadu River is east of the Yajiang River, starting in southern Qinghai and traveling south until it nears Mount Gongga, from where it flows east to join the Yangtze River. The Dadu River may be an important factor separating the northeast populations on the QTP from those within the HMR proper. The Min River begins in northern Sichuan and flows south, passing east of Chengdu (capital of Sichuan) and into the Yangtze flowing east. The regions where discrete populations are clustering have been named based on the rivers they are bounded by. Thus, there is a 'Jinsha-Yajiang' region in the west, a 'Yajiang-Dadu' region in the center, and a 'Dadu-Min' region in the northeast, which extends onto the eastern edge of the Qinghai-Tibet Plateau.

Although the populations for each species cluster within similar regions throughout the HMR landscape, the putative origins (i.e., ancestral biogeographic area) and divergence time estimates of these populations imply that each species has had a unique migration history. Considering that these *Pedicularis* species originated within separate time periods, it was expected that closely related species (i.e., *P. cranolopha* and *P. longiflora*) would share more similar phylogeographic histories than distantly related species (i.e., *P. cranolopha* and *P. lachnoglossa*). The estimates of ancestral population size and divergence times generated from BPP provide some evidence for this hypothesis. For example, the divergence time estimates for the ancestral populations of the sister species *P. cranolopha* (.37 - .75 Mya) and *P. longiflora* (.46 - .88 Mya) overlap considerably (Figure 8). However, the correlation of divergence time and species relatedness does not hold across all relationships in the simplified six-species tree, and the divergence time estimates for all species are considerably younger (.1-.8 Mya) than was expected.

A major interest of this study was to determine if these closely related and widespread species shared similar biogeographic barriers across this dynamic landscape. The divergence time estimates for populations on opposite sides of the Yajiang and Dadu Rivers indicate that population isolation across these barriers occurred at similar historical periods for many species. For example, there is significant overlap among divergence estimates for populations split by the Yajiang River (.09 - .4 Mya) for all species except *P. longiflora*. Additionally, divergence times across the putative barriers in the northeast (i.e., Dadu River and elevated range between Luhuo and Zamthang) are similar for *P. integrifolia* (.33 - .76 Mya) and *P. longiflora* (.40 - .77 Mya), and for *P. lachnoglossa* (.13 - .24 Mya) and *P. cranolopha* (.098 - .24 Mya). Additional barriers to gene flow surely exist in this landscape, but these patterns were the most clearly replicated.

While a similar pattern of population structure was visible for all species, this pattern was less clearly defined for *P. cranolopha* and *P. kansuensis*. The disjunct population structure of *P. cranolopha* in the 'Yajiang-Dadu' region may be explained by the Xianshui River, which passes just east of the town of Luhuo and connects to the Yajiang River further south. The populations in this area show distinct morphological differentiation, transitioning between flowers with and without a forked galea multiple times. Based on molecular evidence, it is plausible that the Xianshui River is facilitating seed transport from a population in the north to a closely related population in the south, bypassing two genetically distinct populations in between (Figure 9). The less strict population structure of *P. kansuensis* may be explained by its ability to grow abundantly as a weedy species in poor conditions on roadsides. This habit may facilitate easier migration between less discrete populations by human vectors, a pattern that has been demonstrated between the QTP and Tianshan Mountains (Li et al., 2016).

Figure 9: Map of Upper Yangtze River system in Sichuan Province, with delineated biogeographic regions bounded by the tributaries of the Yangtze River. The Jinsha River is the natural border between eastern Tibet and western Sichuan, the Yajiang River flows through the center of the sampled region, the Dadu River forms an eastern boundary between the HMR and the QTP, and the Min River (not shown) flows south into the Yangtze further east. Population structure for *P. cranolopha* is displayed, highlighting the disjunct between 'forked galea' (red arrow) and 'normal galea' populations. The smaller Xianshui River is outlined to demonstrate that it may be transporting seeds from the northern population to a southern population (yellow markers), bypassing genetically distinct populations (green and blue markers). Black circles indicate barriers across which multiple species share population divergence estimates.

This study provides insight into the evolutionary influences of tectonic, glacial, and river dynamics in a temperate mountain system. While the unique patterns of mountain biodiversity have received increased attention in recent years (Perrigo, Hoorn, & Antonelli, 2019), few studies have attempted to explicitly define regions of population structure at this scale of resolution (Xu et al., 2010). The molecular evidence generated from these six *Pedicularis* species emphasizes the importance of topographically complex mountain systems in generating intraspecific variation, even within a relatively short time scale $(< 1$ Mya) and small geographical radius (800 km2, about the size of New York City). This finding validates and

builds upon recent work involving *P. longiflora* across a wider range (Yang et al., 2008), which emphasized that topographical isolation, combined with low dispersal ability, has contributed to the rapid speciation of *Pedicularis*.

When making evolutionary inferences based on molecular evidence, it is vital to consider the time scale which the organisms have evolved within. According to Muellner-Riehl (2019), the HMR was mostly uplifted by the beginning of the Quaternary (2.58 Mya), so population divergence within the HMR since that time is more likely to be driven by glaciation and monsoon cycles, and only indirectly by the lingering results of tectonic uplift (i.e., mountain barriers and river valleys already existing in the landscape). While uplift of the HMR drove *in situ* speciation starting ~30 Mya, glaciation, the Asian monsoons, and migration among river corridors has clearly had a more recent impact on population structure at the intraspecific level since populations can become isolated by landscape barriers at rapid rates as their ranges contract and expand during climate oscillations.

Finally, it is important to consider the variable modes of speciation that reveal themselves across a dynamic landscape. Although allopatric speciation is an intuitive description for plant species in topographically complex mountain ranges, introgression has also been shown to promote speciation (Marques, Meier, & Seehausen, 2019). Many studies have clearly demonstrated the ability of the HMR to act as a refuge for plant species during periods of glacial advance on the QTP (Hauenschild et al., 2017; Sun et al., 2017; Xu et al., 2010; Yang et al., 2008; Meng et al., 2007). As populations of these species moved in and out of the HMR during glacial cycles, they would have experienced pronounced periods of introgression. The recently proposed 'mixing-isolation-mixing' (MIM) model of speciation (He et al., 2019), originally formulated to explain speciation of mangroves between the Indian and Pacific Oceans, may

prove influential in understanding the rapid generation of population genetic diversity within plant species that have migrated back and forth along climatic gradients in the Hengduan Mountains since the Last Glacial Maximum (also see "flickering connectivity" concept from Flantua et al., 2019). However, it is important to note that sympatric speciation has likely occurred multiple times within *Pedicularis* (Eaton et al., 2012), so pronounced speciation due to periods of introgression may be more likely in other plant taxa within the HMR.

Conclusion and Future Directions

The "mountain-geobiodiversity hypothesis" emphasizes the importance of three conditions for the formation of biodiversity hotspots in mountains: full altitudinal zonation, climatic oscillations, and high-relief terrain (Mosbrugger et al., 2019, Muellner-Riehl et al., 2019). Such conditions have clearly played an important role in the speciation processes with the HMR temperate biodiversity hotspot. While populations isolated by river barriers were detected, it would be useful to further identify the role of elevational gradients in shaping discrete populations (e.g., high elevation QTP in north-northeast vs. low elevation river valleys). While this study has focused on abiotic drivers of population divergence, with increased resolution on the population structure of multiple *Pedicularis* species in the HMR, additional questions can now be asked concerning the influence that biotic and trait-based factors have on the phylogeographic history of plants throughout this landscape (Papadopoulou & Knowles, 2016; Eaton & Ree, 2013). Additionally, considering the importance of rivers as barriers and transporters of seeds in the HMR, the effects of climate change on the Yangtze River system is an important area of research (Su et al., 2017).

A future extension of this study will aim to determine the impacts of recent human infrastructure development (Dai et al., 2018) on gene flow between endemic populations

throughout the HMR. Human influences, such as climate change and habitat destruction, have accelerated global biodiversity loss in the last century, and mountain plant populations are especially susceptible to these environmental disturbances because they have limited space and mobility to escape increasing temperatures and loss of habitat (Lamprecht et al., 2018). Uncovering the phylogeographic history of these micro-endemic *Pedicularis* species will enable improved assessments of past and present landscape connectivity and species distribution models (Wiens et al., 2009), as well as inform future conservation priorities.

By identifying some of the major phylogeographic barriers to *Pedicularis* migration within the HMR, these results make it possible to test whether recent gene flow has occurred across these historical barriers due to contemporary human influence. Using phylogenomic methods to detect a signature of recent human impact in the genomes of contemporary plant populations would be a novel application for conservation genomics (Supple & Shapiro, 2018), enabling researchers to understand how human settlement and infrastructure development is affecting the evolutionary trajectories of endemic plants in a biodiversity hotspot. Thus, future research should expand sampling efforts to generate sequence data for plants that have existed further out in space and further back in time, as well as incorporate remote sensing landscape analysis to quantify the impact of increased human settlement in the HMR over the past century. These results also highlight the unique factors affecting montane and subalpine plant speciation, demonstrating that increased efforts to discover the phylogeographic histories of mountaindwelling taxa will lead to more robust comparisons of diversification rates in these dynamic and biodiverse regions of the world.

References

- Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81.
- Antonelli, A., Kissling, W. D., Flantua, S. G., Bermúdez, M. A., Mulch, A., Muellner-Riehl, A. N., ... & Fritz, S. A. (2018). Geological and climatic influences on mountain biodiversity. *Nature Geoscience*, *11*(10), 718-725.
- Armbruster, W. S., Shi, X. Q., & Huang, S. Q. (2014). Do specialized flowers promote reproductive isolation? Realized pollination accuracy of three sympatric Pedicularis species. *Annals of Botany*, *113*(2), 331-340.
- Bayona-Vásquez, N. J., Glenn, T. C., Kieran, T. J., Pierson, T. W., Hoffberg, S. L., Scott, P. A., ... & Diaz-Jaimes, P. (2019). Adapterama III: Quadruple-indexed, double/triple-enzyme RADseq libraries (2RAD/3RAD). *PeerJ*, *7*, e7724.
- Boufford, D. (2014). Biodiversity Hotspot: China's Hengduan Mountains. *The Magazine of the Arnold Arboretum*, 24.
- Clark, P. U., Dyke, A. S., Shakun, J. D., Carlson, A. E., Clark, J., Wohlfarth, B., ... & McCabe, A. M. (2009). The last glacial maximum. *science*, *325*(5941), 710-714.
- Dai, E., Wang, Y., Ma, L., Yin, L., & Wu, Z. (2018). 'Urban-Rural' Gradient Analysis of Landscape Changes around Cities in Mountainous Regions: A Case Study of the Hengduan Mountain Region in Southwest China. *Sustainability*, *10*(4), 1019.
- dos Reis, M. (2017). https://github.com/mariodosreis/mousies/blob/master/R/analysis.R.
- Drummond, C. S., Eastwood, R. J., Miotto, S. T., & Hughes, C. E. (2012). Multiple continental radiations and correlates of diversification in Lupinus (Leguminosae): testing for key innovation with incomplete taxon sampling. *Systematic biology*, *61*(3), 443-460.
- Eaton, D. A., Fenster, C. B., Hereford, J., Huang, S. Q., & Ree, R. H. (2012). Floral diversity and community structure in Pedicularis (Orobanchaceae). *Ecology*, *93*(sp8), S182-S194.
- Eaton, D. A., & Overcast, I. (2020). ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, *36*(8), 2592-2594.
- Eaton, D. A., & Ree, R. H. (2013). Inferring phylogeny and introgression using RADseq data: an example from flowering plants (Pedicularis: Orobanchaceae). *Systematic biology*, *62*(5), 689-706.
- Favre, A., Päckert, M., Pauls, S. U., Jähnig, S. C., Uhl, D., Michalak, I., & Muellner‐Riehl, A. N. (2015). The role of the uplift of the Qinghai‐Tibetan Plateau for the evolution of Tibetan biotas. *Biological Reviews*, *90*(1), 236-253.
- Flantua, S. G., O'dea, A., Onstein, R. E., Giraldo, C., & Hooghiemstra, H. (2019). The flickering connectivity system of the north Andean páramos. *Journal of Biogeography*, *46*(8), 1808-1825.
- Flouri, T., Jiao, X., Rannala, B., & Yang, Z. (2018). Species tree inference with BPP using genomic sequences and the multispecies coalescent. *Molecular Biology and Evolution*, *35*(10), 2585-2593.
- Gao, Q. B., Zhang, F. Q., Xing, R., Gornall, R. J., Fu, P. C., Li, Y., ... & Chen, S. L. (2016). Phylogeographic study revealed microrefugia for an endemic species on the Qinghai–Tibetan Plateau: Rhodiola chrysanthemifolia (Crassulaceae). *Plant Systematics and Evolution*, *302*(9), 1179-1193.
- Hauenschild, F., Favre, A., Schnitzler, J., Michalak, I., Freiberg, M., & Muellner-Riehl, A. N. (2017). Spatiotemporal evolution of Allium L. in the Qinghai–Tibet-Plateau region: Immigration and in situ radiation. *Plant diversity*, *39*(4), 167-179.
- He, Z., Li, X., Yang, M., Wang, X., Zhong, C., Duke, N. C., ... & Shi, S. (2019). Speciation with gene flow via cycles of isolation and migration: Insights from multiple mangrove taxa. *National science review*, *6*(2), 275-288.
- Hughes, C. E., & Atchison, G. W. (2015). The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *New Phytologist*, *207*(2), 275-282.
- Hughes, C., & Eastwood, R. (2006). Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences*, *103*(27), 10334-10339.
- Jie, C., Wang, H., Gu, Z., Mill, R. R., & Li, D. (2004). Karyotypes of thirteen species of Pedicularis (Orobanchaceae) from the Hengduan Mountains region, NW Yunnan, China. *Caryologia*, *57*(4), 337-347.
- Lagomarsino, L. P., Condamine, F. L., Antonelli, A., Mulch, A., & Davis, C. C. (2016). The abiotic and biotic drivers of rapid diversification in A ndean bellflowers (Campanulaceae). *New Phytologist*, *210*(4), 1430- 1442.
- Lamprecht, A., Semenchuk, P. R., Steinbauer, K., Winkler, M., & Pauli, H. (2018). Climate change leads to accelerated transformation of high‐elevation vegetation in the central Alps. *New Phytologist*, *220*(2), 447- 459.
- Li, W. J., Sui, X. L., Kuss, P., Liu, Y. Y., Li, A. R., & Guan, K. Y. (2016). Long-distance dispersal after the Last Glacial Maximum (LGM) led to the disjunctive distribution of Pedicularis kansuensis (Orobanchaceae) between the Qinghai-Tibetan Plateau and Tianshan region. *PloS one*, *11*(11).
- Lu, L. M., Mao, L. F., Yang, T., Ye, J. F., Liu, B., Li, H. L., ... & Niu, Y. T. (2018). Evolutionary history of the angiosperm flora of China. *Nature*, *554*(7691), 234-238.
- Luo, D., Xu, B., Li, Z. M., & Sun, H. (2017). The 'Ward Line–Mekong–Salween Divide'is an important floristic boundary between the eastern Himalaya and Hengduan Mountains: evidence from the phylogeographical structure of subnival herbs Marmoritis complanatum (Lamiaceae). *Botanical Journal of the Linnean Society*, *185*(4), 482-496.
- Marques, D. A., Meier, J. I., & Seehausen, O. (2019). A combinatorial view on speciation and adaptive radiation. *Trends in ecology & evolution*.
- Meng, L., Yang, R. U. I., Abbott, R. J., Miehe, G., Hu, T., & Liu, J. (2007). Mitochondrial and chloroplast phylogeography of Picea crassifolia Kom.(Pinaceae) in the Qinghai‐Tibetan Plateau and adjacent highlands. *Molecular Ecology*, *16*(19), 4128-4137.
- Mosbrugger, V., Favre, A., Muellner-Riehl, A. N., Päckert, M., & Mulch, A. (2018). Cenozoic evolution of geobiodiversity in the Tibeto‐ Himalayan Region. In C. Hoorn, A. Perrigio, & A. Antonelli (Eds.), Mountains, climate, and biodiversity (pp. 429–448). Oxford, UK: Wiley‐Blackwell.
- Muellner-Riehl, A. N. (2019). Mountains as evolutionary arenas: patterns, emerging approaches, paradigm shifts, and their implications for plant phylogeographic research in the Tibeto-Himalayan Region. *Frontiers in plant science*, *10*, 195.
- Muellner‐Riehl, A. N., Schnitzler, J., Kissling, W. D., Mosbrugger, V., Rijsdijk, K. F., Seijmonsbergen, A. C., ... & Favre, A. (2019). Origins of global mountain plant biodiversity: Testing the 'mountain‐geobiodiversity hypothesis'. *Journal of Biogeography*, *46*(12), 2826-2838.
- Nevado, B., Contreras‐Ortiz, N., Hughes, C., & Filatov, D. A. (2018). Pleistocene glacial cycles drive isolation, gene flow and speciation in the high‐elevation Andes. *New Phytologist*, *219*(2), 779-793.
- Papadopoulou, A., & Knowles, L. L. (2016). Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. *Proceedings of the National Academy of Sciences*, *113*(29), 8018-8024.
- Perrigo, A., Hoorn, C., & Antonelli, A. (2019). Why mountains matter for biodiversity. *Journal of Biogeography*.
- Pickrell, J., & Pritchard, J. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *Nature Precedings*, 1-1.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945-959.
- Qiu, Y. X., Fu, C. X., & Comes, H. P. (2011). Plant molecular phylogeography in China and adjacent regions: tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. *Molecular phylogenetics and evolution*, *59*(1), 225-244.
- Rahbek, C., Borregaard, M. K., Antonelli, A., Colwell, R. K., Holt, B. G., Nogues-Bravo, D., ... & Fjeldså, J. (2019). Building mountain biodiversity: Geological and evolutionary processes. *Science*, *365*(6458), 1114- 1119.
- Renner, S. S. (2016). Available data point to a 4-km-high Tibetan Plateau by 40 Ma, but 100 molecular-clock papers have linked supposed recent uplift to young node ages. *Journal of Biogeography*, *43*(8), 1479-1487.
- Song, Z.-C., Wang, W.-M. & Huang, F. 2004. Fossil pollen records of extant angiosperms in China. Bot. Rev. (Lancaster) 70: 425–458.
- Spicer, R. A., Su, T., Valdes, P. J., Farnsworth, A., Wu, F. X., Shi, G., ... & Zhou, Z. (2020). Why the 'Uplift of the Tibetan Plateau' is a myth. *National Science Review*.
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312-1313.
- Su, B., Huang, J., Zeng, X., Gao, C., & Jiang, T. (2017). Impacts of climate change on streamflow in the upper Yangtze River basin. *Climatic change*, *141*(3), 533-546.
- Sun, C., Shen, Z., Liu, R., Xiong, M., Ma, F., Zhang, O., ... & Chen, L. (2013). Historical trend of nitrogen and phosphorus loads from the upper Yangtze River basin and their responses to the Three Gorges Dam. *Environmental Science and Pollution Research*, *20*(12), 8871-8880.
- Sun, H., Zhang, J., Deng, T., & Boufford, D. E. (2017). Origins and evolution of plant diversity in the Hengduan Mountains, China. *Plant Diversity*, *39*(4), 161.
- Supple, M. A., & Shapiro, B. (2018). Conservation of biodiversity in the genomics era. *Genome biology*, *19*(1), 1- 12.
- Vasconcelos, T. N., Alcantara, S., Andrino, C. O., Forest, F., Reginato, M., Simon, M. F., & Pirani, J. R. (2020). Fast diversification through a mosaic of evolutionary histories characterizes the endemic flora of ancient Neotropical mountains. *Proceedings of the Royal Society B*, *287*(1923), 20192933.
- Wang, H. J., Li, W. T., Liu, Y. N., Yang, F. S., & Wang, X. Q. (2015). Range-wide multilocus phylogenetic analyses of Pedicularis sect. Cyathophora (Orobanchaceae): Implications for species delimitation and speciation. *Taxon*, *64*(5), 959-974.
- Wiens, J. A., Stralberg, D., Jongsomjit, D., Howell, C. A., & Snyder, M. A. (2009). Niches, models, and climate change: assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences*, *106*(Supplement 2), 19729-19736.
- Wu, Z.Y., Wang, H.S., 1983. Plant Geography of China. Science Press, Beijing.
- Xing, Y., & Ree, R. H. (2017). Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. *Proceedings of the National Academy of Sciences*, *114*(17), E3444-E3451.
- Xu, T., Abbott, R. J., Milne, R. I., Mao, K., Du, F. K., Wu, G., ... & Liu, J. (2010). Phylogeography and allopatric divergence of cypress species (Cupressus L.) in the Qinghai-Tibetan Plateau and adjacent regions. *BMC evolutionary biology*, *10*(1), 194.
- Yang, F. S., Li, Y. F., Ding, X. I. N., & Wang, X. Q. (2008). Extensive population expansion of Pedicularis longiflora (Orobanchaceae) on the Qinghai‐Tibetan Plateau and its correlation with the Quaternary climate change. *Molecular Ecology*, *17*(23), 5135-5145.
- Yang, H., Holmgren, N. H., & Mill, R. R. (1998). Pedicularis Linnaeus. Flora of China 18: 97–209.
- Yang, Z. (2015). The BPP program for species tree estimation and species delimitation. *Current Zoology*, *61*(5), 854-865.
- Ye, X. Y., Ma, P. F., Yang, G. Q., Guo, C., Zhang, Y. X., Chen, Y. M., ... & Li, D. Z. (2019). Rapid diversification of alpine bamboos associated with the uplift of the Hengduan Mountains. *Journal of Biogeography*, *46*(12), 2678-2689.
- Yoder, A. D., Campbell, C. R., Blanco, M. B., Dos Reis, M., Ganzhorn, J. U., Goodman, S. M., ... & Ralison, J. M. (2016). Geogenetic patterns in mouse lemurs (genus Microcebus) reveal the ghosts of Madagascar's forests past. *Proceedings of the National Academy of Sciences*, *113*(29), 8049-8056.
- Yu, W. B., Liu, M. L., Wang, H., Mill, R. R., Ree, R. H., Yang, J. B., & Li, D. Z. (2015). Towards a comprehensive phylogeny of the large temperate genus Pedicularis (Orobanchaceae), with an emphasis on species from the Himalaya-Hengduan Mountains. *BMC Plant Biology*, *15*(1), 176.
- Zhou, Z., Hong, D., Niu, Y., Li, G., Nie, Z., Wen, J., & Sun, H. (2013). Phylogenetic and biogeographic analyses of the Sino-Himalayan endemic genus Cyananthus (Campanulaceae) and implications for the evolution of its sexual system. *Molecular Phylogenetics and Evolution*, *68*(3), 482-497.

Supplementary Materials

Supplement 1: Field notes for *Pedicularis* samples used in this study, include date of collection and latitude/longitude for all samples.

Supplement 2: Data retained for PCA and STRUCTURE analyses after filtering loci to 50% coverage within each 'imap' population and 75% coverage across all samples.

Supplement 3: Jupyter notebooks for ipyrad assembly, RAxML tree inference, PCA, STRUCTURE, and BPP analyses [\(https://j-meek.github.io/thesis/\)](https://j-meek.github.io/thesis/).