Targeted gene sequencing and selection detection in ex situ collections of *Brighamia insignis*

Christina Shehata[1], Lauren Audi[1,2], Hilary Noble[2], Colby Witherup[1,2], Jeremie Fant[2], Norm Wickett[2]

[1]Northwestern University, Evanston, IL 60201; [2]Chicago Botanic Garden, Glencoe, IL 60022

Introduction

Brighamia insignis (family: Campanulaceae, subfamily: Lobelioideae) is a species of caudiciform succulent endemic to Hawaii. *Brighamia insignis* is extinct in the wild, due to habitat loss, small population sizes, and the loss of its native pollinator, *Tinostoma smaragditis*[1]. Fortunately, due to their novel morphology, *Brighamia insignis* have been cultivated as ornamental plants and still exist in ex situ collections around the world.

Figure 1. *Brighamia insignis* collection at the Chicago Botanic Garden.

Genomic data obtained via next-generation sequencing methods provide valuable information about putatively expressed regions of the genome. Identifying those genes under natural selection can help conservation scientists understand the variation that influences the fitness of individuals. This information may help guide breeding efforts to make ensure that extant individuals are genetically diverse, which is essential for future reintroduction efforts of critically endangered species.

Objectives

- Evaluate variation and relatedness among Hawaiian lobelioids kept in botanic collections around the world.
- Sequence protein-coding genes described from the Campanulaceae family using Hyb-Seq[2], a sequencing method that can capture targeted genes in non-model organisms.
- Identify genes that are adaptive candidates in populations of *Brighamia insignis* and *Brighamia rockii*.

Methods

Data Collection

DNA from collected lobeliad samples was extracted following the standard CTAB protocol. Samples were sheared using the Covaris Focused-Ultrasonicator. DNA libraries were prepared and amplified following the KAPA protocol. Libraries were pooled into groups of 12 samples. Campanulaceae baits were hybridized to the pools, which were then amplified using PCR. The Bioanalyzer was used to evaluate pool concentrations and calculate average fragment size. Samples were sequenced using Illumina's MiSeq System.

Post-Sequencing Analysis

Sequenced reads were trimmed with Trimmomatic then processed and assembled with HybPiper[3]. Orthologous gene sequences were aligned using MACSE and trimmed with Trimal. FastTree was used to construct a phylogenetic tree for each gene of interest. Astral was used to construct a population tree from the individual gene trees. LSD[4] was used to detect signatures of positive selection among the population tree branch leading to *Brighamia insignis*.

Results

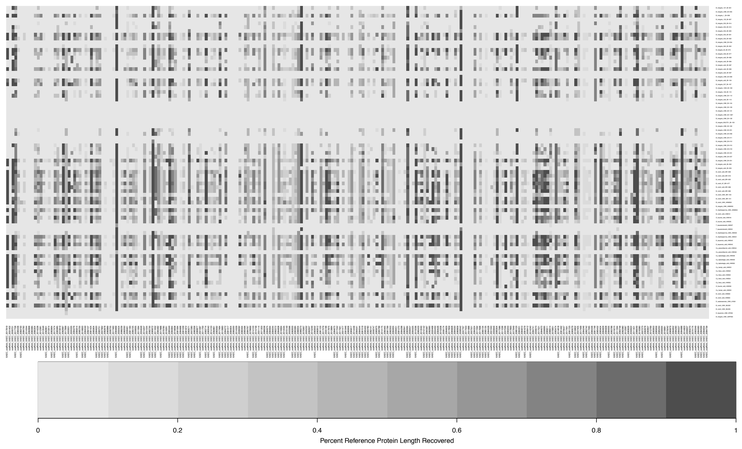

Figure 2. Heat map representing HybPiper's success at recovering genes at each locus, across 83 samples.

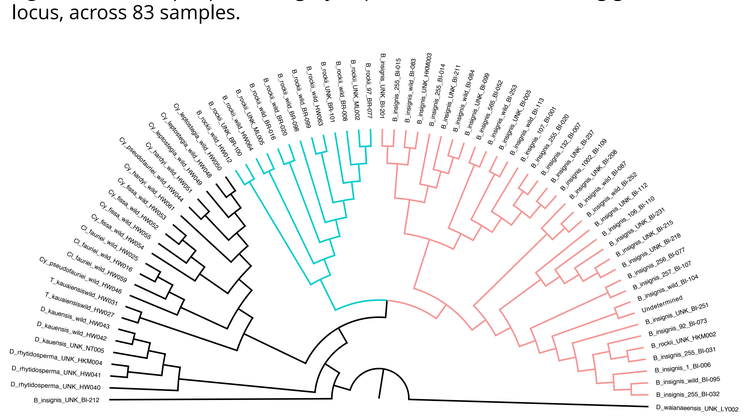

Figure 3. Phylogenetic tree constructed from individual gene trees. LSD was used to detect signatures of selection along the branches leading to *Brighamia insignis* (pink) and *Brighamia rockii* (blue).

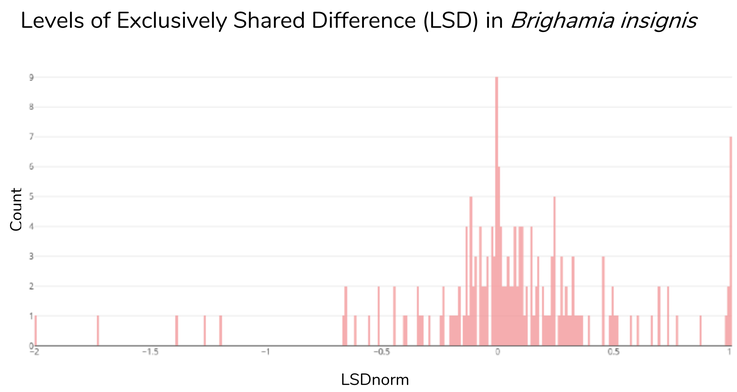

Figure 4. Histogram representing the distribution of normalized LSD values along the branch leading to *Brighamia insignis*.

Discussion

While 206 genes were recovered during assembly, some genes were not captured by the Campanulaceae probes in any of 83 samples (Figure 2). The probes were designed from *Campanula erinus*, a relatively distant species from *Brighamia insignis*, possibly explaining the lower success of target capture for some particular probes. The overall diversity among the samples of *Brighamia insignis* was greater than expected, with multiple points of divergence observed on the population tree along the branch leading to the populations of *Brighamia insignis* (Figure 2).

93 genes were found to have an LSDnorm value greater than 0 (Figure 3), suggesting that these genes are adaptive candidates within the sampled populations of *Brighamia insignis*. While traditional population genetic work focuses on putatively neutral loci, genomic data collected by next-generation sequencing methods provide important information regarding functional genetic variation, which could be incorporated into future breeding programs for *Brighamia*. If the species has very low levels of diversity, one of the suggested ways to increase diversity is to cross *Brighamia insignis* with its sister species, *Brighamia rockii*. However, if there are genes that are locally adapted to the sampled populations of *Brighamia insignis*, the information gained from identifying adaptive candidate genes can help ensure that these genes are preserved in future breeding efforts.

Future Directions

- Sequence the remaining *Brighamia* samples to complete a genomic dataset with representation of all *Brighamia insignis* individuals.
- Use the tested probes on other Hawaiian lobelioids to answer phylogenetic questions pertaining to the Campanulaceae family.
- Test other bioinformatic methods to detect signatures of positive selection (i.e. dN/dS).

Literature Cited

1. Gemmill CEC, Ranker TA, Ragone D, Perlman SP, Wood KR. 1998. Conservation genetics of the endangered Hawaiian genus *Brighamia* (Campanulaceae). American Journal of Botany 85(4): 528-539.
2. Weitemier K, Straub SCK, Cronn RC, Fishbein M, Schmickl R, McDonnell A, Liston A. 2014. Hyb-seq: Combining target enrichment and genome skimming for plant phylogenomics. Applications in Plant Sciences 2(9): 1400042.
3. Johnson MG, Gardner EM, Liu Y, Medina R, Goffinet B, Shaw AJ, Zerega NJC, Wickett NJ. 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. Applications in Plant Sciences 4(7): 1600016.
4. Librado P, Orland L. 2018. Detecting signatures of positive selection along defined branches of a population tree using LSD. Molecular Biology and Evolution 35(6): 1520–1535.

Contact Information

christinashehata2019@u.northwestern.edu

https://github.com/ChristinaShehata/

REU Site: *Plant Biology & Conservation Research Experiences for Undergraduates - From Genes to Ecosystems* (Supported by NSF award DBI-1757800).

Acknowledgements: We would like to thank NSF-REU grant DBI-1757800, Northwestern University's Undergraduate Research Grant Program, and the Chicago Botanic Garden for support and for making this summer experience possible. I would also like to thank my mentors Colby Witherup, Lindsey Bechen, Lauren Audi, Jeremie Fant, and Norm Wickett for their continuous guidance and support. Also, thank you to Hilary Noble and Jordan Wood for their support in the laboratory.