Using high-throughput DNA sequence data to infer the phylogeny of *Lewisia* (Montiaceae) and evaluate a potentially new species from the Sierra Nevada foothills

Caitlin Hazelquist[1], C. Matt Guilliams[1], Kristen Hasenstab-Lehman[1], Reed Kenny[2], Wendy Boes[3]

[1] Santa Barbara Botanic Garden, [2] University of California, Davis, [3] Nevada City, CA

Santa Barbara Botanic GARDEN

Introduction

The genus *Lewisia* comprises between 25 and 30 taxa of perennial herbs in the Montiaceae. This variable, western North American genus has been the subject of considerable taxonomic and morphological study. Earlier phylogenetic studies based on Sanger sequencing revealed geographically and morphologically cohesive groupings across the genus, but with relatively low statistical support.

In 2017, Wendy Boes encountered a low elevation population of morphologically and ecologically distinctive lewisias in the northern Sierra Nevada Foothills, which were somewhat similar to the widespread, high elevation *L. nevadensis*. Sanger sequencing and analysis of a sample of the putative new *Lewisia* confirmed a close relationship with *L. nevadensis*, along with *L. kelloggii* and *L. triphylla*, although support values in that analysis were low in general (**Fig. 1**).

With our initial phylogenetic work providing only partial insight into the putative new *Lewisia*, here we aim to:

1) reexamine phylogenetic relationships across the genus with a focus on the putative new taxon from the Sierra Nevada Foothills and close relatives;
2) use population genomics to understand genetic grouping of the widespread *L. nevadensis* and putative new taxon;
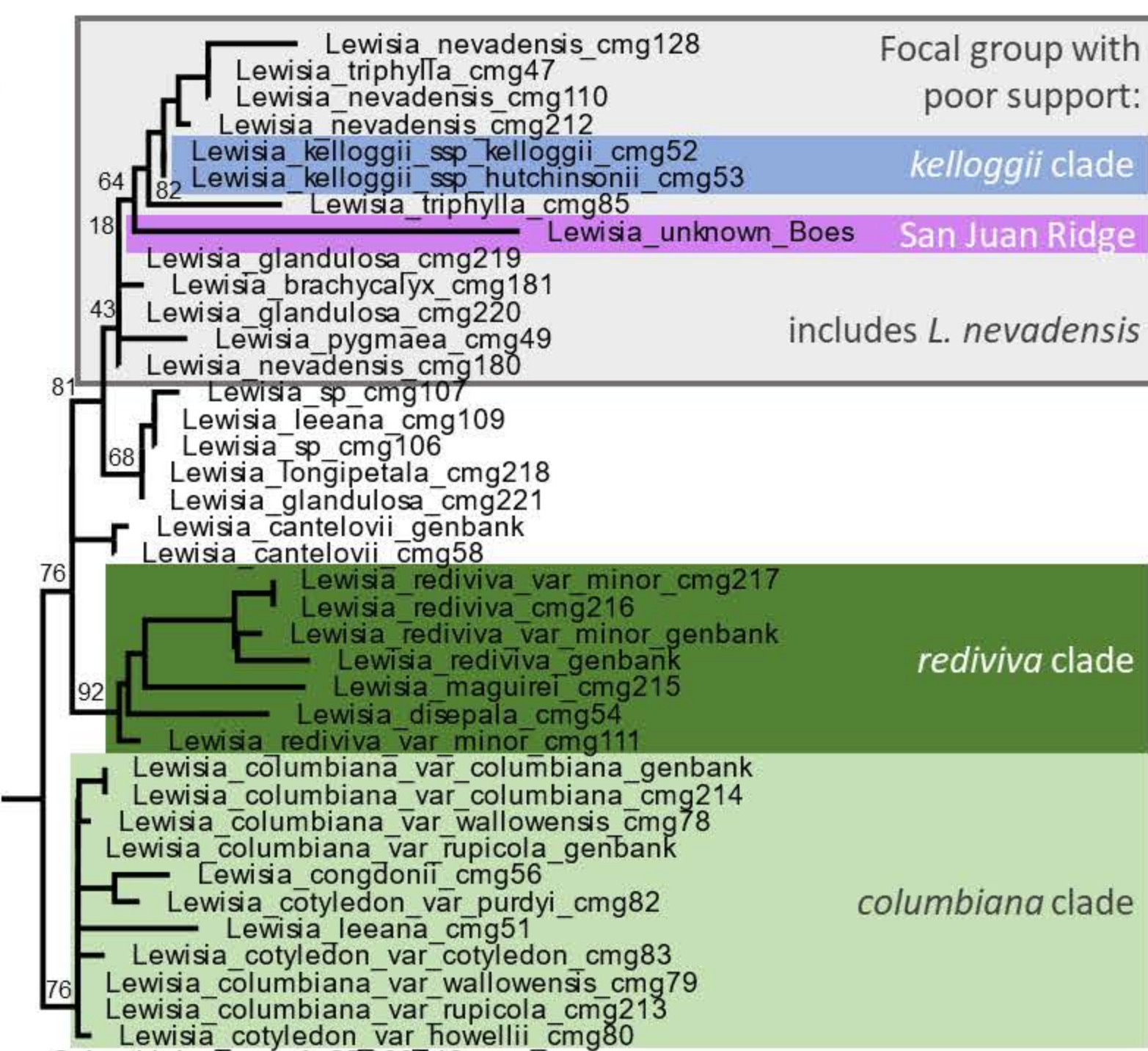3) Investigate observed morphological differences using quantitative morphometrics.

Figure 1. Unpublished phylogeny based on nrITS data. The *columbiana*, *rediviva*, and *kelloggii* clades have relatively strong support, with the single sample of the putative new taxon placed with *L. nevadensis*, *L. kelloggii*, and *L. triphylla*.

Methods

Silica dried material was collected from throughout its range, homogenized in bead beater tubes into a fine powder and extracted using a modified CTAB protocol (Doyle and Doyle 1987). Libraries were prepared for high throughput sequencing using a double digest Restriction site Associated DNA sequencing (ddRADseq) protocol following Tripp et al. (2017). Data were assembled through the ipyrad pipeline (Eaton 2014). Phylogenetic trees were inferred using maximum likelihood in RAxML (Stamatakis 2014) on the CIPRES Science Gateway v3.3 (Miller et al. 2010). Trees were visualized using the program FigTree v1.4.4 . Genetic groupings of *L. nevadensis* and the San Juan Ridge population were assessed and visualized using a neighbor joining dendrogram in adgenet (Jombart 2008). Morphometric data was collected from iNaturalist images of *L. nevadensis* and images of the *Lewisia* "San Juan Ridge" population gathered in the field. Some data were also gathered from herbarium specimens. Statistical analyses were performed in R (R core team 2018, Wickham 2016).

Results & Discussion

- ddRAD approach was effective at creating high-quality DNA sequence data in *Lewisia*
- DNA sequence dataset useful from population genomic level to phylogenomic level
- *kelloggii*, *columbiana*, and *rediviva* clades were recovered with strong support in ML tree (**Fig. 2**)
- Samples of putative new taxon form a clade with strong support, nested within *L. nevadensis*
- The neighbor-joining dendrogram shows clear groupings of *L. nevadensis* and putative new taxon (**Fig. 4**)
- Morphometric analyses show that putative new taxon is morphologically distinct from *L. nevadensis* (Fig 5A-B)
- Preliminary genetic and morphometric data appear to support the recognition of the San Juan Ridge *Lewisia* as a new species
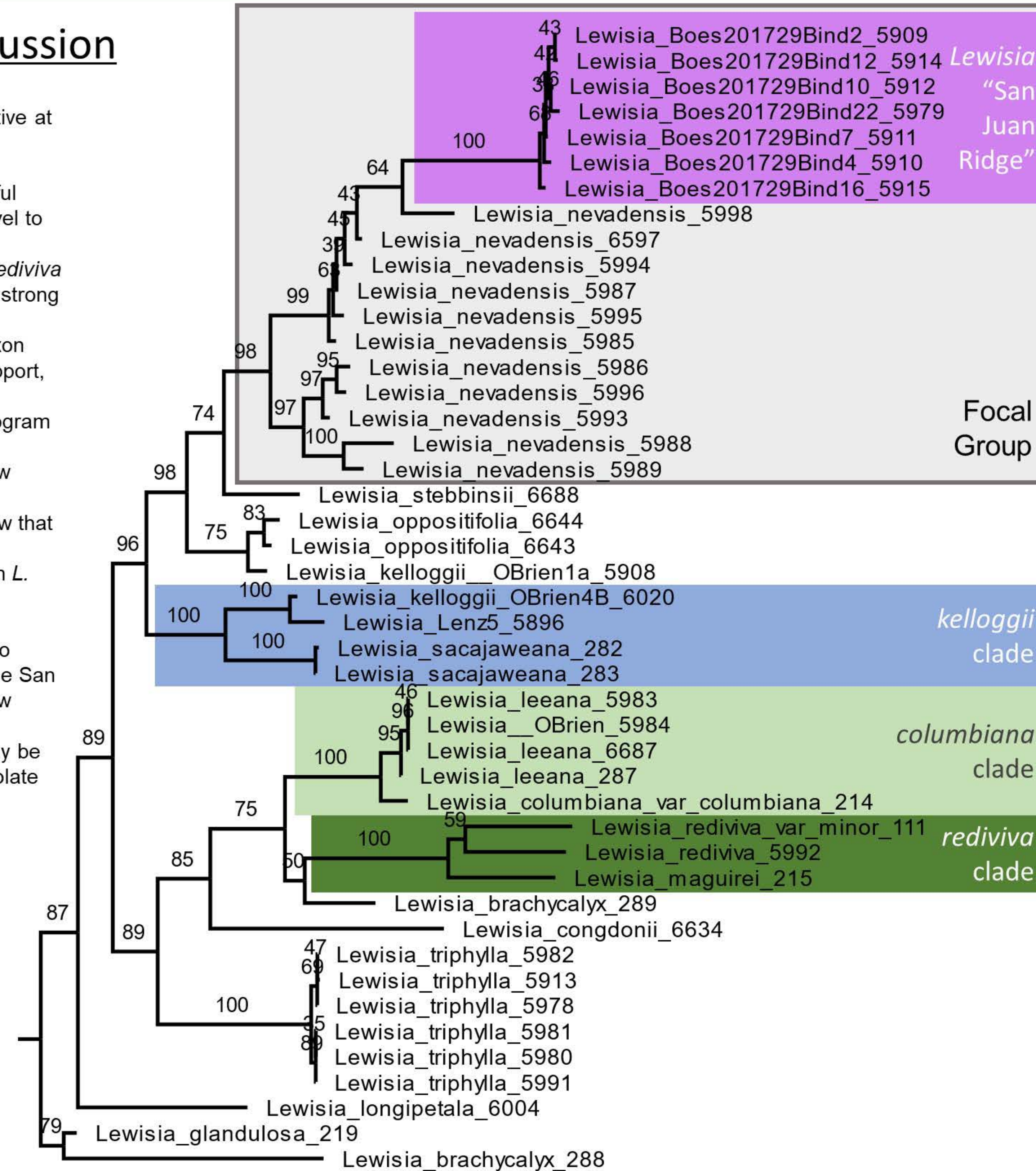- San Juan Ridge *Lewisia* may be an example of peripheral isolate speciation.

Figure 2. ddRAD phylogeny of *Lewisia*. There is continued strong support for the *kelloggii*, *columbiana*, and *rediviva* clades previously identified with Sanger sequencing. Samples of the putative new taxon form a clade with strong support, nested within *L. nevadensis*.
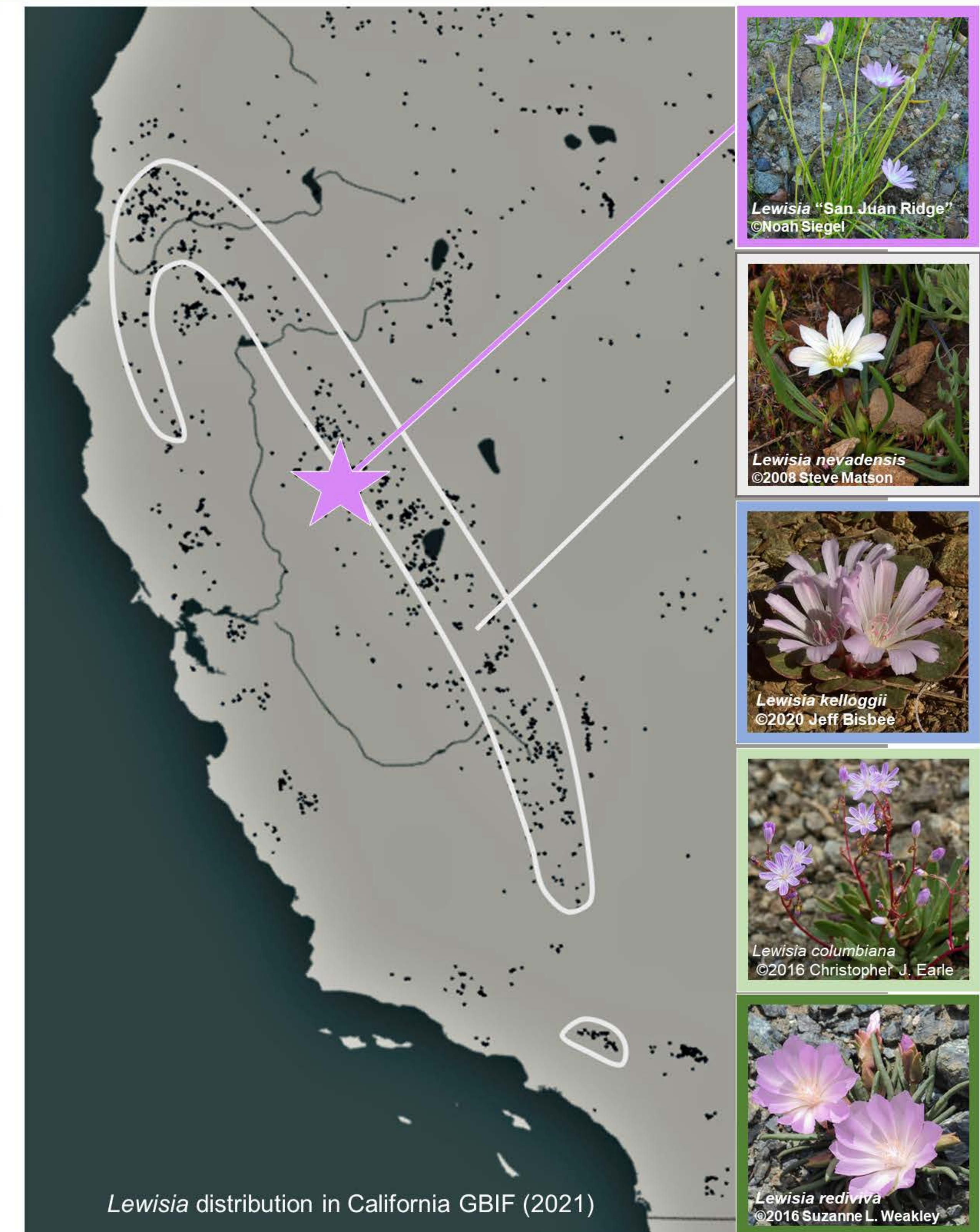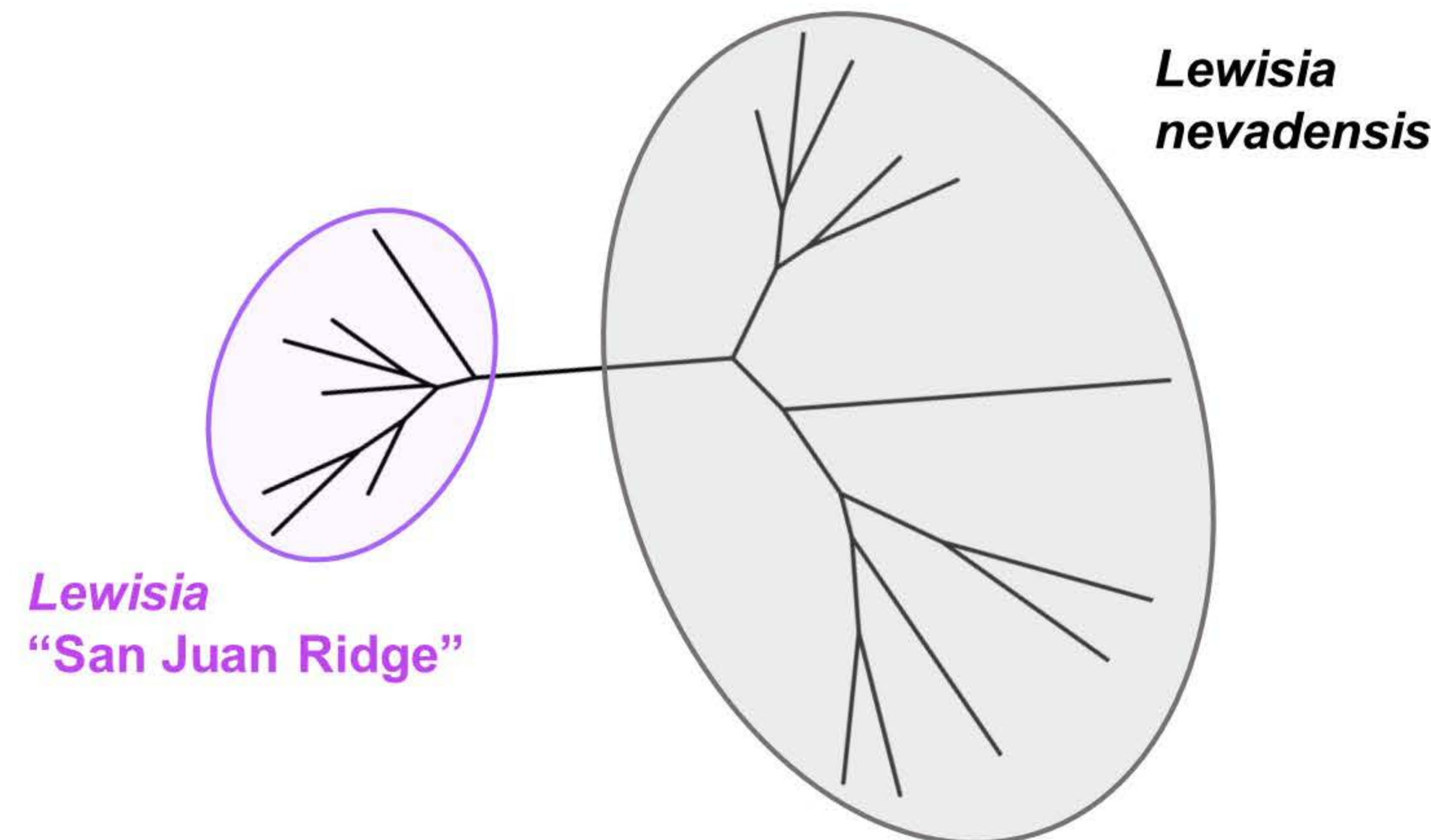

Figure 3. Map showing the geographic distribution of the genus *Lewisia* and examples of taxa represented in the phylogeny. Area outlined in light grey represents the distribution of *L. nevadensis*. Starred region represents the San Juan Ridge population.

Lewisia distribution in California GBIF (2021)

Figure 4. Unrooted ddRAD neighbor-joining dendrogram of samples of *L. nevadensis* and the putative new taxon.
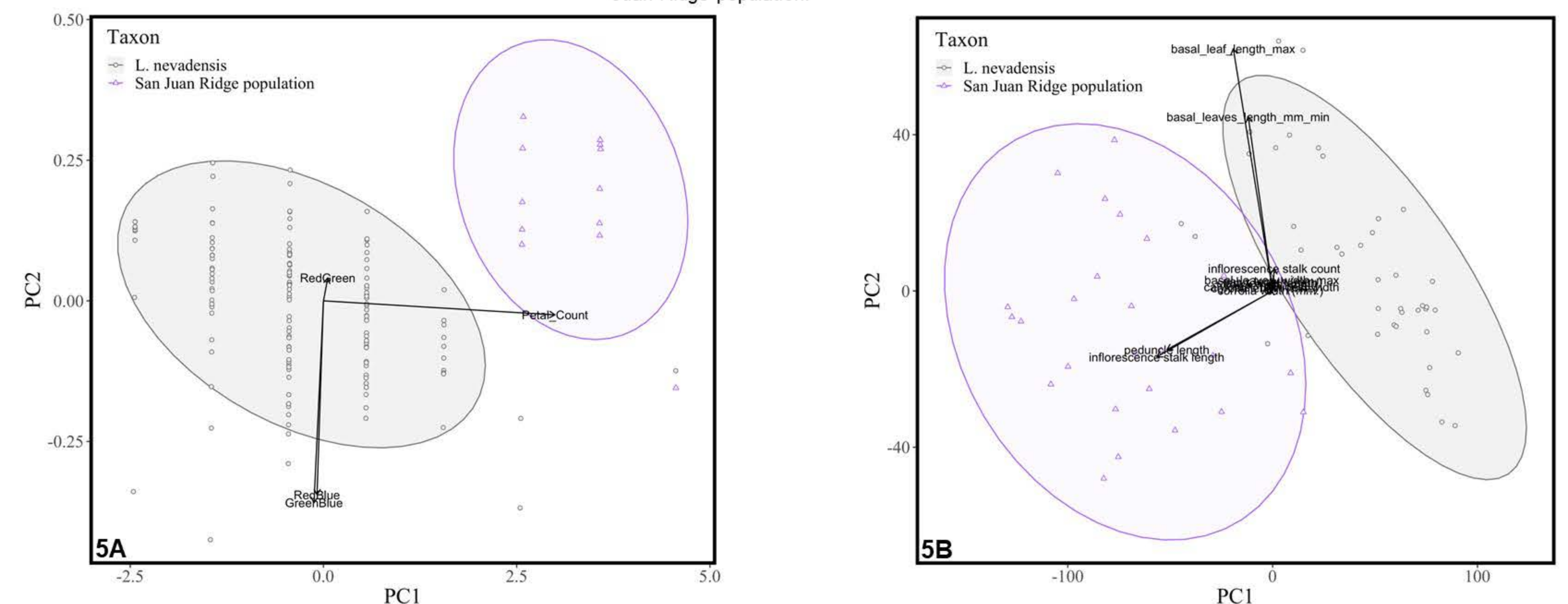

Figure 5A-B. PCA scatterplots from morphological analyses. **A.** Scatterplot of PC2 versus PC1 for floral data gathered from digital images; **B.** Scatterplot of PC2 versus PC1 for whole plant data gathered from herbarium specimens.

References

Baird NA, Etter PD, Atwood TS, et al (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS One. doi: 10.1371/journal.pone.0003376. Davidson BL (2000). *Lewisias*. Timber Press, Portland. Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small amounts of fresh leaf tissue. Phytochem. Bull. 19: 11-15. Eaton DAR (2014) PyRAD: Assembly of de novo RADseq loci for phylogenetic analysis. Bioinformatics. Elliot RC (1966). The Genus *Lewisia*. Alpine Garden Society, Great Britain. Etter PD, Preston JL, Bassham S, et al (2011) Local de novo assembly of rad paired-end contigs using short sequencing reads. PLoS One. doi: 10.1371/journal.pone.0018561. Jombart T (2008) Adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics. *Lewisia* Pursh in GBIF Secretariat (2021) GBIF Backbone Taxonomy. Checklist dataset https://doi.org/10.15468/39omei accessed via GBIF.org on 2021-10-13. Mathew B (1989). The Genus *Lewisia*. Timber Press, Portland. Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA.Tripp EA, Tsai YHE, Zhuang Y, Dexter KG (2017) RADseq dataset with 90% missing data fully resolves recent radiation of *Petalidium* (Acanthaceae) in the ultra-arid deserts of Namibia. Ecol Evol. doi: 10.1002/ece3.3274. R Core Team (2018): R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria https://www.R-project.org/. Stamatakis A (2014) RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9) 1312-1313. Wickham H (2016). Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4 https://ggplot2.tidyverse.org