Barcoding and Comparison of DNA Extractions of ITS Region Between Angiosperm and Gymnosperm Herbarium Specimens

Ethan Rosati[1]
Erin Riggs[2] UVSC Director -Curator
Utah Valley University, Biology Department

Abstract

Goals in this investigation are to extract DNA from the angiosperm and gymnosperm clades from exsiccate herbarium specimens. The most valuable goal in this investigation is the addition of the sequenced data to the international barcoding body of knowledge (“Education and the Barcode of Life”, 2020). The addition of this data will help biologists have intellectual control of gene sequence data aiding in the conservation and protection of the biota of Utah and the world. (de Vere, et. al, 2012). I will amplify and sequence the ITS gene region then conduct BLAST comparison for plant identification confirmation. This will expand the foundation of my lab experience and bring big data into the herbarium by making a sequence reference library. This sequence library will be used for researchers conducting research outside curation and phylogeny such as biofuel (Zhang, 2012). Finally I will then compare angiosperm and gymnosperm physical ease of DNA extraction, and the quality and quantity of DNA. A major outcome for this investigation is to lay the groundwork for other UVU students in herbarium barcoding for future engaged learning opportunities in comparative species work, rare species or lab techniques. Overall this will benefit the students in the College of Science and UVU as a whole.

Methods

First steps are the extraction process, this is conducted through a destructive method. Twenty mg herbarium plant fragments will be collected from 100 specimens, then ground into a fine powder, this powder is placed in 1.5ml tube and are extracted via geneJet extraction kits. The extracted DNA is then assessed using a nanodrop to test for initial concentration and purity. A 4ul aliquot of DNA, targeting the ITS gene region (Cheng, 2016), is then amplified via polymerase chain reaction (PCR). The PCR product is again assessed on a nanodrop 2000 measuring concentration and purity. Those that show ample DNA concentration are purified through a Qiagen purification protocol to remove heavy salts. The samples are then prepared for ITS gene region sequencing and sent to Psomagen per their protocol. In addition to barcode targeting ITS gene region. Geneious software will be used to edit DNA and measure quality and quantity to assess DNA yield. ANOVA was the intended analysis, due to COVID my work was shortened on this project and my sample size is unusually small therefore no ANOVA will be conducted. Sequences were run through BLAST in order to compare species percent identification.

Figure 1 a.

Figure 1 b.

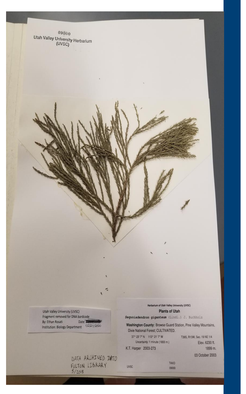

Figure 1a. is a gymnosperm sample and is observed to have a waxy and tough exterior. Figure 1b. shows an angiosperm sample that has a more fragile and soft exterior.

Results

Figure 2 a. Angiosperm Voucher Sample

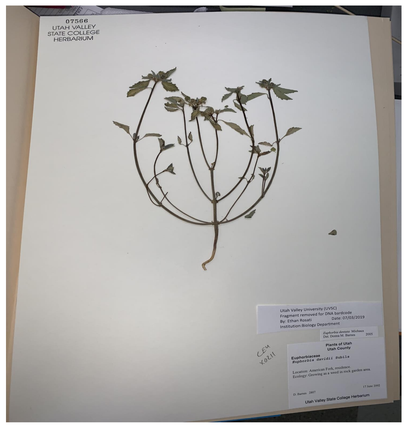

Figure 2 b. Nanodrop DNA Concentrations

Sample # Angiosperm	DNA Concentration (ng/ul)
2	4.9
5	33.3
7	5.1
12	1.9
14	1.7
15	4.3
16	7.5
20	6.8
21	3.8
22	6.8
23	16.5
24	7.6
25	7.4
29	19.3
31	12.7
Average	9.3

Sample # Gymnosperm	DNA Concentration (ng/ul)
1	7.1
2	0.4
3	1.3
4	2.7
5	5.6
Average	3.42

Figure 2 b. Due to the small sample size a statistical analysis was not conducted.

Figure 3a. Angiosperm BLAST Results

Angiosperm Sample #	Identity Match Name	Percent Identity
2	Shepherdia canadensis internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and 26S ribosomal RNA gene, partial sequence	81.82
5	Clematis ladakhiana voucher TUR:601932 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1 and 5.8S ribosomal RNA gene, complete sequence; and internal transcribed spacer 2, partial sequence	93.74
7	Prunus padus isolate 239 internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene, complete sequence; and internal transcribed spacer 2, partial sequence	96.35
12	Actaea asiatica voucher Q110 small subunit ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequence	97.87
15	Thalictrum venulosum isolate T90VEUm 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and 26S ribosomal RNA gene, partial sequence	93.5
21	Solidago houghtonii voucher PJL1211 clone type 2 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and 26S large subunit ribosomal RNA gene, partial sequence	95.86
22	Ericameria discoidea var. linearis RR313 internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence	94.77
23	Ericameria nauseosus voucher PJL1186 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and 26S large subunit ribosomal RNA gene, partial sequence	91.06
24	Disperis sp. CPG 29216 isolate ITS14 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and 26S ribosomal RNA gene, partial sequence	90.7
29	Baccharis neglecta internal transcribed spacer 1, 5.8S ribosomal RNA gene; and internal transcribed spacer 2, complete sequence	85.78

Figure 3b. Gymnosperm BLAST Results

Gymnosperm Sample #	Identity Match Name	Percent Identity
1	Cloning vector PBTH-mwtGFP	96.55
2	Expression vector Sos_ZsG_pol I-mg	92.93
4	Expression vector Sos_ZsG_pol I-mg	91.39

When the angiosperm samples were blasted in NCBI BLAST 10 of the 15 samples identified with an ITS region species with the lowest percent identification at 81.82 and the highest at 96.35. When the gymnosperm samples were blasted 2 of 5 had the highest percent identification with an expression vector with a percent identification average of 92.16. One sample did identify with a cloning vector with the percent identification value at 96.55.

This is preliminary data laying the foundation for a larger comparative project. When more data is documented ANOVA analysis will be conducted.

Discussion

I am in the early stages of this investigation, a total of 20 samples have been extracted. Gymnosperm samples have been physically more difficult to extract than angiosperm due to the heavy waxy leaf cuticle and fibrous sclerenchyma.

When total DNA concentration was measured the angiosperm samples had a low value of 1.7ng/ul and the highest sample measuring at 33.3 ng/ul. Gymnosperm samples had a low value of 0.4 ng/ul and the highest value measuring 7.1 ng/ul.

One comparison that can be made from this small set of data between the two groups is that no gymnosperm sample sequences when blasted came back with no identity for the ITS region. Whereas the angiosperm samples did have the highest identity with ITS region containing species. This could be a cause from poor DNA quality because of difficulty extracting the DNA.

I am able to edit and record the resulting ITS gene regions into barcoding initiatives and add to the UVSC, UVU herbarium, barcoding database (“Education and the Barcode of Life”, 2020). With a barcoding library on site different research avenues can be done outside of phylogeny and curation. Furtherment of scientific education can be obtained from the knowledge gained from barcoding herbarium samples.

More work will be done in comparing the two groups including an ANOVA test. I will also be conducting next-gen sequencing in the future for the dual purpose of adding that data to our herbarium as part of my undergraduate learning experience.

References

1. Cheng, T., Xu, C., Lei, L., Li, C., Zhang, Y., & Zhou, S. (2015, July 3). Barcoding the kingdom Plantae: new PCR primers for ITS regions of plants with improved universality and specificity. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/1755-0998.12438
2. de Vere N, Rich TCG, Ford CR, Trinder SA, Long C, et al. (2012) DNA Barcoding the Native Flowering Plants and Conifers of Wales. PLoS ONE 7(6): e37945. doi:10.1371/journal.pone.0037945
3. Zhang, G., Liu, X., Quan, Z., Cheng, S., Xu, X., Pan, S., … Wang, J. (2012, May 13). Genome sequence of foxtail millet (Setaria italica) provides insights into grass evolution and biofuel potential. Retrieved from https://www.nature.com/articles/nbt.2195
4. Quang Ong, Phuc Nguyen, Nguyen Phuong Thao and Ly Le, “Bioinformatics Approach in Plant Genomic Research”, Current Genomics (2016) 17: 368. https://doi.org/10.2174/1389202917666160331202956
5. Hollingsworth, P. M., et al. “A DNA Barcode for Land Plants.” Proceedings of the National Academy of Sciences, vol. 106, no. 31, 2009, pp. 12794–12797., doi:10.1073/pnas.0905845106.
6. Biolabs, C. M. (n.d.). Education & the Barcode of Life. Retrieved from http://www.educationandbarcoding.org/

Acknowledgements

Funding for this project was granted by Utah Valley University SAC committee and the Office of Engaged Learning.

Thank you to the Utah Valley University Office of Engaged Learning and to Erin Riggs, project PI and UVSC herbarium curator, for being a mentor and oracle of knowledge through this project.

Contact
1. Ethan Rosati
Biotechnology Researcher
emrosati@gmail.com
1. Erin Riggs
PI and Curator
ErinRiggs@uvu.edu

Barcoding and Comparison of DNA Extractions of ITS Region Between Angiosperm and Gymnosperm Herbarium Specimens