Elucidating the evolutionary history of evening primroses (Onagraceae) using exon-capture and next-gen sequencing

Amanda Patsis[1], Jeremie Fant[2], Rachel Levin[1], Michael Moore[3], Rick Overson[2], Krissa Skogen[2], Norm Wickett[2]

[1]Amherst College, Amherst, MA 01002; [2]Chicago Botanic Garden, Glencoe, IL 60022, [3]Oberlin College, Oberlin, OH 44074

Objectives:

- Generate an unprecedented amount of sequence data for evening primroses
- Create a well supported phylogeny that reveals the evolutionary history of Onagraceae

Photo : Rick Overson

Introduction:

Onagraceae, the evening primrose family, comprises 655 species worldwide, with a majority located in western North America (Levin et al. 2004). Throughout the evolutionary history of this family there have been several shifts in pollinator syndromes (e.g. from bee to moth) and life history traits, making Onagraceae an ideal system for exploring floral evolution and the role of pollinators, and possibly floral parasites, as drivers of speciation (Raven 1979). However, to understand floral variation and its role in diversification, an accurate understanding of the evolutionary history of this family is necessary.

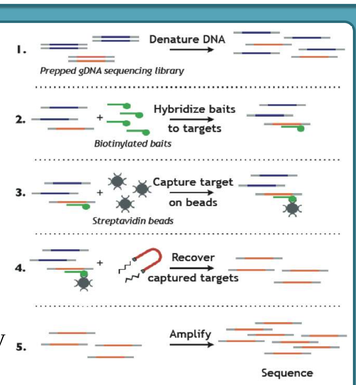
Research focused on the delineation of species and their evolutionary relationships within Onagraceae is ongoing and dynamic. For example, in a study by Levin et al. (2004), one nuclear region and two chloroplast regions were sequenced to elucidate evolutionary relationships, focusing on the tribe Onagreae. Johnson et al. (2009) used two plastid and three nuclear regions to reconstruct these same relationships. The resulting phylogenies, however, suggest significantly different evolutionary histories, most likely due to differences in the loci on which they are based.

With the advent of next-generation sequencing and accompanying advances in bioinformatics, unprecedented amounts of DNA sequence data can now be gathered and analyzed. By using hundreds to thousands of loci to resolve phylogenetic relationships, rather than only a few loci, these new methods have the potential to clarify species boundaries, and to illuminate the evolutionary history of many unresolved lineages (Cronn et al. 2012, Straub et al. 2012). Using Hyb-Seq, a technique that efficiently enriches genomic DNA for targeted regions, we will reconstruct relationships in Onagraceae using a number of phylogenetic markers that far exceeds any existing analysis. We will use the resulting data to produce a robust phylogenetic resource that will be used as a framework to understand the role of floral traits, pollinators, and parasites in the evolution of this family.

What is Hyb-Seq ?

- Custom “baits” (short, 120bp “primer” sequences) are designed based on target exons that are conserved across the group of interest.
- DNA is extracted from each species being analyzed, and an Illumina library is constructed for each DNA sample, with each sample characterized by a unique “barcode” sequence.
- For target enrichment, each library is hybridized, in solution, to the baits, which are then immobilized on a bead.
- Removing DNA that is not captured on the beads, the sample becomes enriched for the target loci. This prevents wasted sequence reads, as only phylogenetically informative markers are sequenced.

(Weitemier et al. 2014)



MYbaits User Manual Version 1.3.8, MYcroarray

Methods:

DNA Extraction:

- 59 species were extracted with a modified Qiagen DNeasy Kit
- Leaf tissue was immersed in liquid nitrogen and shaken in a homogenizer to grind tissue
- Samples were incubated in lysis buffer at 60°C overnight
- Extracted DNA was quantified using a Qubit 2.0 Fluorometer

Library Preparation:

- A TruSeq Nano DNA Library Prep Kit was used for library prep
- DNA was sonicated into 550bp fragments with a Covaris S220 Ultrasonicator (Figure 1)
- Barcoded adaptors were ligated to the ends of fragments
- Sample Purification Beads were used to size select library
- Library size distribution was validated using an Agilent 2100 Bioanalyzer (Figure 2)

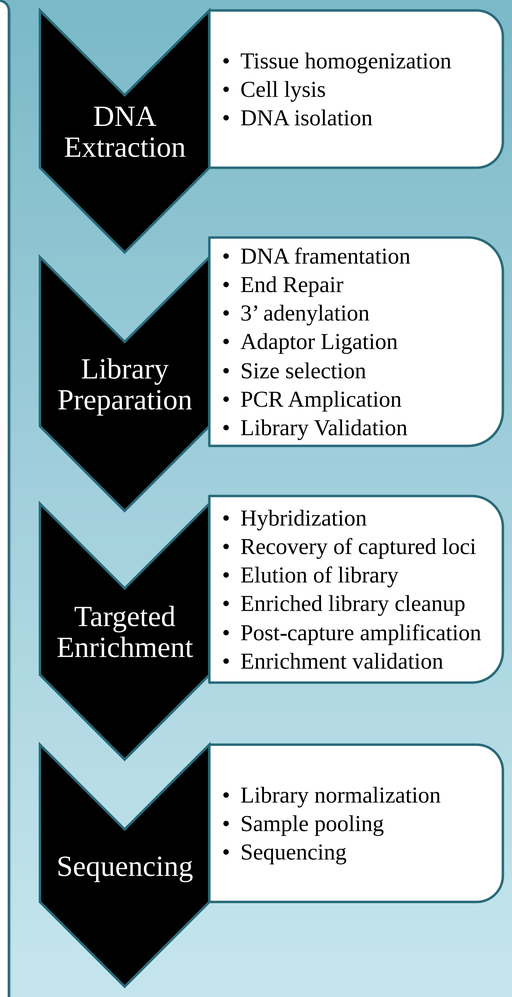
*Target Enrichment:

- A Mybaits Target Enrichment Kit was used for hybridization-based enrichment
- 11-13 libraries were pooled in a single hybridization; pools were based on relatedness to limit capture bias
- 360 target loci were captured using biotinylayted RNA baits and magnetic streptavidin beads

*Sequencing and Data Analysis:

- Barcoded, enriched Illumina libraries will be sequenced on the Illumina MiSeq platform using 2x300bp reads
- Sequence data will be assembled using a custom Python pipeline (M. Johnson; https://github.com/mossmatters/HybSeqPipeline)
- For each gene, a multiple sequence alignment will be created
- Phylogenetic analyses will be carried out using supermatrix and coalescent-based methods

* Not yet completed!



Future Directions:

350 species of Onagraceae, focusing on the Tribe Onagreae, will be sampled to offer a more complete representation of the family and to construct a well supported phylogeny. This analysis will provide a framework for mapping functional traits, improving our understanding of the evolution of pollinator syndromes across this diverse family.

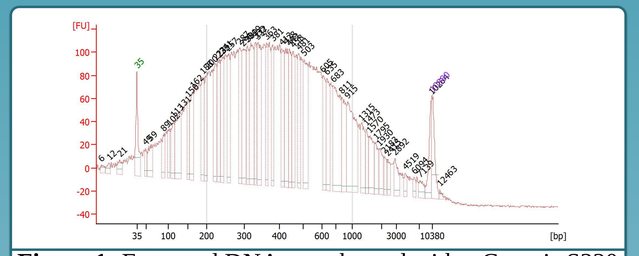


Figure 1: Extracted DNA was sheared with a Covaris S220 Ultrasonicator, with a target fragment size of 500bp. The results were validated using an Agilent 2100 Bioanalyzer, shown here for *Oenothera cespitosa marginata*.

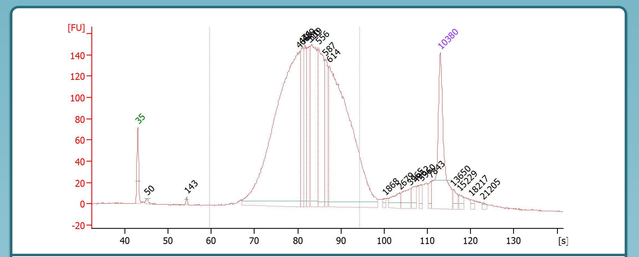


Figure 2: Illumina libraries were validated using an Agilent 2100 Bioanalyzer. A narrow size distribution and sufficient quantity of DNA fragments indicates successful size selection and PCR amplification. Results shown here are for *Oenothera cespitosa navajoensis*.

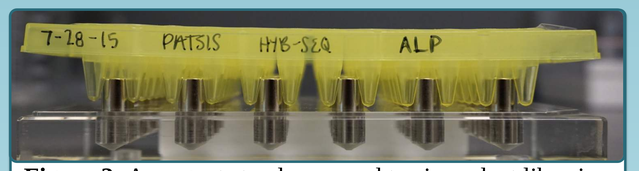


Figure 3: A magnet stand was used to size select libraries and to recover captured DNA fragments during target gene enrichment.

References:

Cronn, R., B. J. Knaus, A. Liston, P.J. Maughan, M. Parks, J. V. Syring, & J. Udall. 2012. Targeted enrichment strategies for next-generation plant biology. American Journal of Botany 99:291-311.

Levin, R. A., W. L. Wagner, P. C. Hoch, W. J. Hahn, A. Rodriguez, D. A. Baum, L. Katinas, E. A. Zimmer, & K. J. Sytsma. 2004. Paraphyly in tribe Onagreae: insights into phylogenetic relationships of Onagraceae based on nuclear and chloroplast sequence data. Systematic Botany 147-164.

Meier, Rudolf. 2000. Species concepts and phylogenetic theory: a debate. Columbia University Press, New York, New York, USA.

Raven, P. H. 1979. A survey of reproductive biology in Onagraceae. New Zealand Journal of Botany 17:575-593.

Straub, S. C., M. Parks, K. Weitemier, M. Fishbein, R. C. Cronn, & A. Liston. 2012. Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. American Journal of Botany 99:349-364.

Acknowledgements:

I would like to thank NSF grant DEB-1342873 for funding this project, and NSF grant DBI-1461007 for additional aid. I would also like to thank the Amherst College Department of Biology for their continued encouragement and support. We thank Warren Wagner for providing additional onagrad samples. All Covaris and Bioanalyzer runs were performed at the DNA Discovery Center at the Field Museum.