

**Final Report for Research Grant
July 15, 2008**

Provided to	Save-the-Redwoods League
Prepared by	Drs. Deborah Rogers, Barnaly Pande, and David Neale
Research title	Development of Genomic Tools in Coast Redwood
CoPIs	Dr. David B. Neale, University of California, Davis Dr. Barnaly Pande, University of California, Davis Dr. Deborah L. Rogers, University of California, Davis and Center for Natural Lands Management
Collaborators	Ms. Ellen Luong, University of California, Davis Dr. Andrew Eckert, University of California, Davis

Abstract This pilot study has resulted in the generation of the first sequence data for coast redwood. Forty gene fragments were sequenced and resulted in 12 fragments with high quality sequence. These 12 resequenced amplicons (“gene sequences”) were used to survey genetic diversity in a panel of 11 redwood individuals. Analysis of these data has revealed, for the first time, variation at the level of redwood DNA sequence (SNPs). Although the number of individuals, gene sequences, and redwood range sampled were all small, the nucleotide diversity detected was high relative to some other conifers. This result provides the ‘proof of concept’ that hexaploid redwood genes can be successfully resequenced in many cases and research impetus to continue to develop sequence data for coast redwood. In the longer term, results from such genomic discovery can reveal genetic variation in the adaptive response of coast redwood and allow comparison of genetic variation in gene sequences across different conifer species.

Project Context

This pilot study was proposed in response to and based upon the following principles, assumptions, and conditions concerning coast redwood and genetic discovery:

- Genetic discovery in coast redwood has been largely restricted because of the sampling, field study, and laboratory challenges of this very large, very long-lived, hexaploid.
- The rate of climate change and its potential impacts of this species *in situ* has increased the urgency for discovering the adaptive potential of this species.
- The field of genomics has progressed sufficiently with conifers (particularly loblolly pine and Douglas-fir) to allow the leveraging of this information towards discovery in coast redwood.
- Risk and uncertainty with a genomics approach are related to the untested resequencing ability using sequencing primers from other species and the potential effects of paralogs.
- A pilot study is critical in determining the merit of concerns over paralogs and in determining the investment needed for further genomics discovery in coast redwood.

Overall Objectives of Pilot Study

The overall goals were to produce the first genomic tools for genetic discovery in coast redwood and to generate preliminary data that will provide compelling support for a larger investigation of adaptation in this species.

Outcome: These goals were achieved.

Specific Project Objectives

- 1. Identification of PCR (Polymerase Chain Reaction) primers to generate high quality DNA sequences from 40 genes putatively involved in plant adaptation to the environment, and optimize the experimental protocols for their use.**

Outcome: Based on a coincident pilot study screening a diverse set of conifers for resequencing success using loblolly pine primers, 177 primer pairs were identified that generated detectable DNA sequence in coast redwood. From these, 40 primer pairs were selected that provided high- (27 primer pairs) to good- (13 primer pairs) quality sequence.

- 2. To resequence 40 genes in coast redwood and discover SNP (Single Nucleotide Polymorphism) and allelic (haplotype) variation in these sequences.**

Outcome: Of the 40 good- to high-quality primer pairs identified in Objective #1, 24 were successfully resequenced. Of those, 12 were determined to have high-quality data, amenable to analysis (see Figure 1). These 12 resequenced amplicons (“gene sequences”) were used to survey genetic diversity in a panel of 11 redwood individuals. Variation was detected in 10 out of 12 resequenced amplicons. In total, 109 single nucleotide polymorphisms (SNPs) were detected in these 10 amplicons across 11 individual redwoods. Between one and five haplotypes were observed per resequenced amplicon, with the average of three haplotypes per sequence. Further analysis of SNP diversity in this small redwood sample suggests that diversity may be high relative to other conifers thus far assessed (e.g., loblolly pine, Monterey pine, Douglas-fir). However, different sequences and numbers of sequences were assessed in each species, and thus the comparison is not standardized. Also, paralogs could (artificially) inflate the diversity statistics. This initial insight into SNP diversity in coast redwood, however, is significant in that it is based on a small number of individuals, all from a narrow portion of the species range (northern California), suggesting that a larger and broader sample would result in even higher diversity results.

Of further interest is the nature of the genes studied. Gene sequences were selected for this study based on practical considerations rather than gene function. Many of the 12 gene sequences are related to photosynthesis.

Finally, these candidate gene sequences were tested for ‘neutrality’ (i.e., evidence of natural selection). Six of the 10 resequenced amplicons appear to exhibit patterns consistent with natural selection using one or more tests. Further comparison of the redwood sequences with the original loblolly candidate gene sequences indicates evidence for a recent selective sweep in some of these resequenced amplicons.

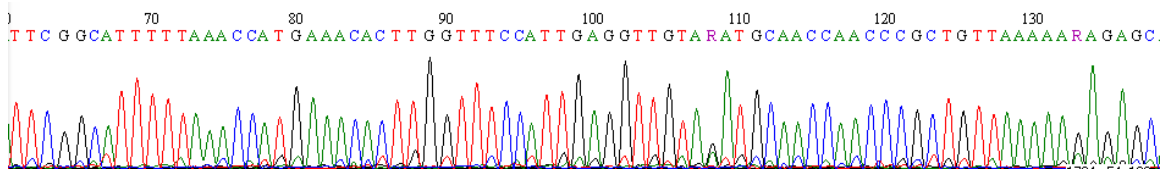


Figure 1. The resequenced amplicon *O_8850*, homologous to photosystem I P700 apoprotein A2 in a number of wide-ranging species: cycad, *Cycas micronesia* (EU016879); ginkgo, *Ginkgo biloba* (DQ069680); banana, *Musa acuminata* (EU017060); buttercup, *Ranunculus macranthus* (DQ069682).

Applications and Implications

- This study has removed some of the initial technical barriers and attitudinal limitations towards genomic discovery in coast redwood. It has provided impetus to continue to generate quality sequence data. Based on the results of the test DNA, all forty primers could be used without additional modification of loblolly-based protocols.

- There are still potential issues with paralogs.
- This study's main and substantial contribution is in its proof of concept for resequencing in coast redwood. The specific results on amount of diversity in coast redwood at the nucleotide level should not be over-interpreted. These results are based on a small number of sequences, in a small number of redwoods from a narrow part of the species' natural range.
- The results from this study can provide the foundation for longer term and more comprehensive studies to determine genetic variation in the adaptive response of coast redwood and allow comparison of genetic variation in gene sequences across different conifer species.

Funds Leveraged

The project was enhanced and broadened considerably by building on results from a recently completed conifer genomic study. In that study, a coast redwood sample was included in an array of other conifers, and tested for resequencing success with 8,000 loblolly pine primers. From this screening, 177 primer pairs produced some product that were then further screened for utility in the current project.

Laboratory costs were subsidized (highly leveraged by) by existing infrastructure.

Inkind contributions of time and expertise were provided by Rogers and Neale (Eckert, Lee, Wegrzyn)

Next steps

1. These encouraging results can be used to improve and validate our draft research proposal for federal funding (initially proposed in 2006) re: genetic variation and adaptation in coast redwood.
2. We will prepare a paper for publication that describes this success in redwood resequencing. The support of the Save-the-Redwoods League will be acknowledged in that paper (and a copy of the paper provided, upon publication).
3. We need to collect a diverse sample of coast redwood as the basis of a rangewide survey in adaptive variation. Existing collections (e.g., UC Russell Reservation) do not represent the full-range of environmental, and expected adaptive, variation. Such a (comprehensive) collection was initiated in 2007 with seed collections from approximately 20 redwoods (Rogers) and will continue as funding and redwood seed production *in situ* permit.
4. Additional development of high-quality sequence is needed, expanding the number of sequences and the inclusion of those involved in adaptive functions such as drought tolerance.
5. Our long-term goal is to discover the genetic variation present in redwood across its natural range at the DNA sequence level and to use this to understand adaptive variation in natural populations. This fundamental understanding will then have direct and important applications in restoration and conservation of this species.