

Rangewide Genetic Variation in Coast Redwood Populations at a Chloroplast Microsatellite Locus¹

Chris Brinegar²

Abstract

Old growth and second growth populations of coast redwood (*Sequoia sempervirens*) were sampled at 10 locations throughout its range and analyzed at a highly variable chloroplast microsatellite locus. Very low F_{ST} values indicated that there was no significant genetic differentiation between adjacent old growth and second growth populations at each location. Genetic diversity was moderate to high in populations north of the San Francisco Bay but low to very low in more southerly populations. Phylogenetic analysis produced a neighbor-joining tree with one clade composed of all populations north of San Francisco Bay and the other composed of the remaining southern populations. Differentiation of these two groups is consistent with the loss of rare alleles in the smaller, more fragmented southern populations by genetic drift.

Key words: chloroplast microsatellite, coast redwood, genetic drift, *Sequoia sempervirens*

Introduction

Ancestors of the modern coast redwood (*Sequoia sempervirens*) were once dominant forest species of the temperate northern latitudes. Glaciation ultimately forced the ancient redwood forests to retreat into the coastal fog belts of Oregon and California where they reached their extreme southern range limit near Santa Barbara in the late Pleistocene (Sawyer and others 2000). Genetic analysis of extant populations throughout the current 450 mile range could provide insights into the nature of historical expansions and contractions of redwood forests while shedding light on the evolutionary forces that will shape the composition of redwood forests in the future.

Three rangewide studies have provided data on the diversity and genetic structure of coast redwood populations. Hall and Langenheim (1987) detected a statistically significant geographic variation in leaf monoterpene composition. Cluster analysis revealed three major groups with the more divergent southern group separated from the more related central and northern groups by the San Francisco Bay. Growth performance differences in common garden experiments (Anekonda 1992) also supported a division between redwoods north and south of the San Francisco Bay. Douhovnikoff and Dodd (2011) detected a weak divergence between redwood lineages near the Sonoma/Mendocino border using nuclear microsatellite markers.

¹ This paper was presented at coast redwood forests in a changing California, June 21-23, 2011, University of California, Santa Cruz.

² Division of Natural Sciences, University of Maine Farmington, Farmington, ME 04961. (chris.brinegar@maine.edu).

Molecular markers in haploid genomes have smaller effective population sizes compared to nuclear markers (Weising et al. 2005). Therefore, chloroplast microsatellites are typically more sensitive in detecting founder effects, bottlenecks, and differentiation by genetic drift (Provan et al. 2001). Among a set of coast redwood microsatellite markers developed by Bruno and Brinegar (2004) was the chloroplast *Seq21E5* locus, a highly variable tetranucleotide repeat. Preliminary data from the *Seq21E5* locus suggested a greater genetic diversity in northern coast redwood populations compared to populations in Santa Cruz County (Brinegar et al. 2007). However, the northern and southern populations were sampled from old growth and second growth forests, respectively. Whether the observed differences represented a north-south divergence or were due to regeneration artifacts could not be determined in that limited study.

The major goal of the current study was to determine whether a chloroplast microsatellite analysis at the *Seq21E5* locus could provide independent confirmation of a north-south split in the genetic structure of coast redwood populations. A secondary objective was to ascertain if, in fact, there are detectable differences in genetic composition between old growth and second growth populations.

Methods

Sampling sites are shown in *figure 1*. At most sites, old growth and second growth populations were sampled in different areas within the boundaries of the same park or reserve. Exceptions were in the old growth/second growth collections at Armstrong Redwoods State Reserve/Austin Creek State Recreation Area, Big Basin Redwoods State Park/Castle Rock State Park, and Miller Creek/Lime Kiln State Park. In all cases, paired old growth/second growth populations were within five miles of each other. Redwood branchlets were collected from 40 trees spaced at least 10 m apart along transects. Tissue was dried in desiccant and stored at room temperature.

DNA extraction and PCR conditions were performed as described by Bruno and Brinegar (2004). PCR products were separated on 6.5 percent polyacrylamide gels (29:1 acrylamide:bisacrylamide, 40 mM Tris-acetate [pH 8.3], 1 mM EDTA) supplemented with 1X Spreadex Polymer NAB (Elchrom Scientific, Cham, Switzerland). Samples were electrophoresed for 3 h at 200 V and stained with 0.5 mg/L ethidium bromide. PCR product sizes were estimated using M3 markers (Elchrom Scientific). *Seq21E5* size variants are referred to in this study as “alleles.”

Allele frequencies and Wright’s F_{ST} values, which represent the variance of allele frequencies among populations (Avice 2004), were calculated using HAPLOTYPE ANALYSIS (Eliades and Eliades 2009). Shannon diversity indices were calculated for populations from the formula: $S = -\sum p_i \log_2 p_i$ (Allnut and others 1999) where p_i is the frequency of the i th allele. Standardized Shannon indices were calculated by dividing S by its maximum value for a given sample size. A neighbor-joining tree was constructed in POPTREE2 (Takezaki and others 2010) using Nei’s D_A distance (Nei et al. 1983).

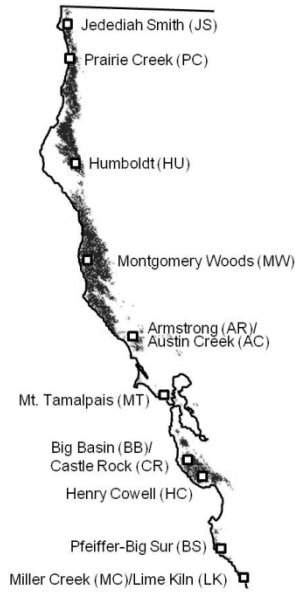


Figure 1—Map of coast redwood distribution in California with sampling locations.

Results

Comparison of old growth and second growth populations

The numbers of individuals with a given allele in the paired old growth and second growth populations are shown in *table 1*. Histograms of allele distributions based either on old growth or second growth allele frequencies are compared in *fig. 2*.

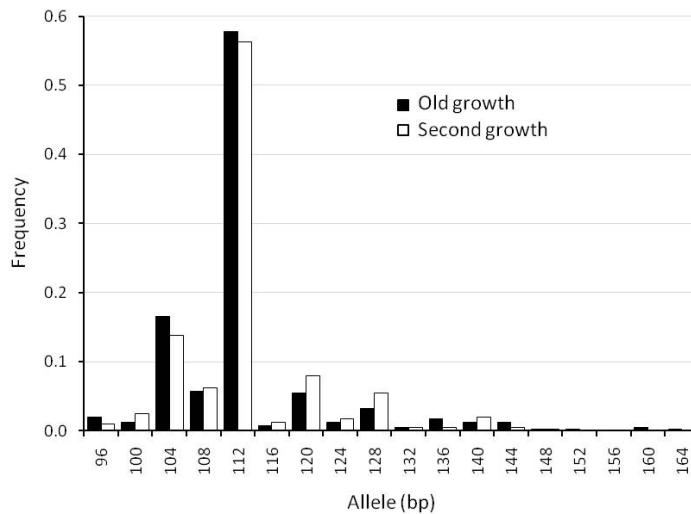


Figure 2—Allele frequencies in combined old growth populations and combined second growth populations.

Table 1—Allele distributions (number of individuals) in paired old growth and second growth redwood populations.^a

Population ^b	Allele size (bp)																	
	96	100	104	108	112	116	120	124	128	132	136	140	144	148	152	156	160	164
JS-OG	4	0	6	7	11	1	3	1	2	1	0	3	1	0	0	0	0	0
JS-SG	1	0	6	4	19	1	2	0	4	0	0	3	0	0	0	0	0	0
PC-OG	1	1	8	4	15	0	5	1	2	0	3	0	0	0	0	0	0	0
PC-SG	1	0	7	1	14	0	8	1	5	1	0	1	0	1	0	0	0	0
HU-OG	1	0	1	2	28	0	1	1	2	0	0	1	1	1	1	0	0	0
HU-SG	0	2	6	6	18	1	3	1	2	0	0	1	0	0	0	0	0	0
MW-OG	0	1	6	6	11	2	5	1	3	1	1	0	0	0	0	0	2	1
MW-SG	2	2	5	5	8	1	8	2	1	1	0	3	2	0	0	0	0	0
AR-OG	2	1	6	4	16	0	7	0	2	0	1	0	1	0	0	0	0	0
AC-SG	0	3	5	3	17	1	6	0	3	0	2	0	0	0	0	0	0	0
MT-OG	0	2	6	0	26	0	1	1	2	0	0	1	1	0	0	0	0	0
MT-SG	0	3	2	4	21	0	1	3	6	0	0	0	0	0	0	0	0	0
BB-OG	0	0	15	0	22	0	0	0	0	0	2	0	1	0	0	0	0	0
CR-SG	0	0	8	0	29	0	2	0	1	0	0	0	0	0	0	0	0	0
HC-OG	0	0	12	0	28	0	0	0	0	0	0	0	0	0	0	0	0	0
HC-SG	0	0	10	2	26	0	2	0	0	0	0	0	0	0	0	0	0	0
BS-OG	0	0	5	0	35	0	0	0	0	0	0	0	0	0	0	0	0	0
BS-SG	0	0	4	0	36	0	0	0	0	0	0	0	0	0	0	0	0	0
MC-OG	0	0	1	0	39	0	0	0	0	0	0	0	0	0	0	0	0	0
LK-SG	0	0	2	0	37	1	0	0	0	0	0	0	0	0	0	0	0	0

^a n = 40 for each population.^b Population abbreviations are defined in *figure 1*. OG = old growth; SG = second growth.

The two distributions, each based on data from 400 individuals, are extremely similar and include alleles ranging from 96 to 164 bp (corresponding to 4 to 21 repeat units).

Allele number and F_{ST} comparisons between paired old growth and second growth populations are displayed in *table 2*. Allele number ranged from two in some of the southernmost populations to 12 in the MW populations. In general, allele numbers in old growth populations are mirrored in their second growth counterparts. When differences in allele numbers occur, they are due to rare alleles.

Table 2—*Allele number and F_{ST} comparisons of paired old growth and second growth redwood populations.*

	Population ^a									
	JS	PC	HU	MW	AR/ AC	MT	BB/ CR	HC	BS	MC/ LK
OG ^b alleles	11	9	11	12	9	8	4	2	2	2
SG alleles	8	10	9	12	8	7	4	4	2	3
Normalized percent F_{ST}	0.7	0.3	1.6	0.4	0.1	0.8	1.3	0.2	0.0	0.2

^aPopulation abbreviations are from *fig. 1*.

^bOG = old growth; SG = second growth.

Wright’s F_{ST} , when normalized to the maximum value possible, can be expressed as the percent of genetic diversity due to allele frequency differences among populations. Pairwise F_{ST} values for each old growth/second growth pair (*table 2*) are all extremely low (0.0 to 1.6 percent) indicating that there is negligible genetic differentiation between adjacent old growth and second growth populations throughout the coast redwood range. Therefore, for subsequent analyses, allele frequency data from the adjacent old growth and second growth populations at each location were combined in order to provide a larger population sample size (n = 80).

Rangewide variation in genetic diversity

The standardized Shannon diversity index is a measure of genetic diversity that takes both allele number and frequency into account. Values of 0 and 1 indicate no genetic diversity and maximum genetic diversity, respectively, within populations. *Figure 3* shows how genetic diversity at the ten sampling sites (combined old growth and second growth data) varies with distance from the northern range limit of coast redwood (approximately eight miles north of the Oregon/California border).

Populations north of the San Francisco Bay have standardized Shannon indices of 0.51 to 0.81 with the MW population being the most diverse. North of MW, diversity drops at the HU population owing to its high frequency of the 112 bp allele relative to the other populations in Mendocino, Humboldt and Del Norte Counties. Genetic diversity decreases almost linearly with distance in populations south of MW, falling to 0.08 in the southernmost population (MC/LK) where the 112 bp allele is nearly fixed (95 percent).

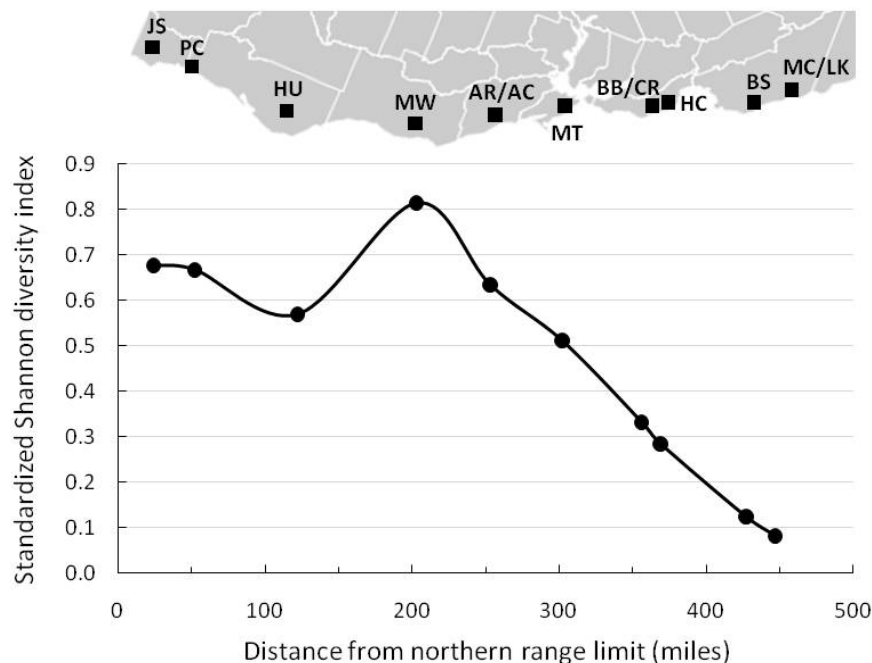


Figure 3—Standardized Shannon diversity index of combined old growth and second growth populations at each sampling location (indicated on map inset) vs. distance from the northern range limit.

Interpopulation genetic structure

Using the allele frequency data from combined old growth/second growth paired populations, an unrooted neighbor-joining tree based on Nei's D_A distances was constructed (*fig. 4*). The populations segregated into two clades: one containing the six populations north of the San Francisco Bay and one containing the four southernmost populations. Resampling of data by bootstrapping (Felsenstein 1985) is typically performed in order to determine the degree of support for the topography of phylogenetic trees. Unfortunately, bootstrapping values cannot be assigned to this tree which is based on data from a single locus.

Trees were also constructed using data from the old growth and second growth populations separately (not shown). The tree from second growth population data was essentially the same as the combined population tree. The tree from old growth population data differed only in the placement of the MT population which appeared on an interior branch of the southern clade. This suggests that the allele composition of the MT population, in Marin County, is teetering on the “borderline” between the northern and southern clades, indicating that it, too, may be succumbing to the effects of genetic drift.

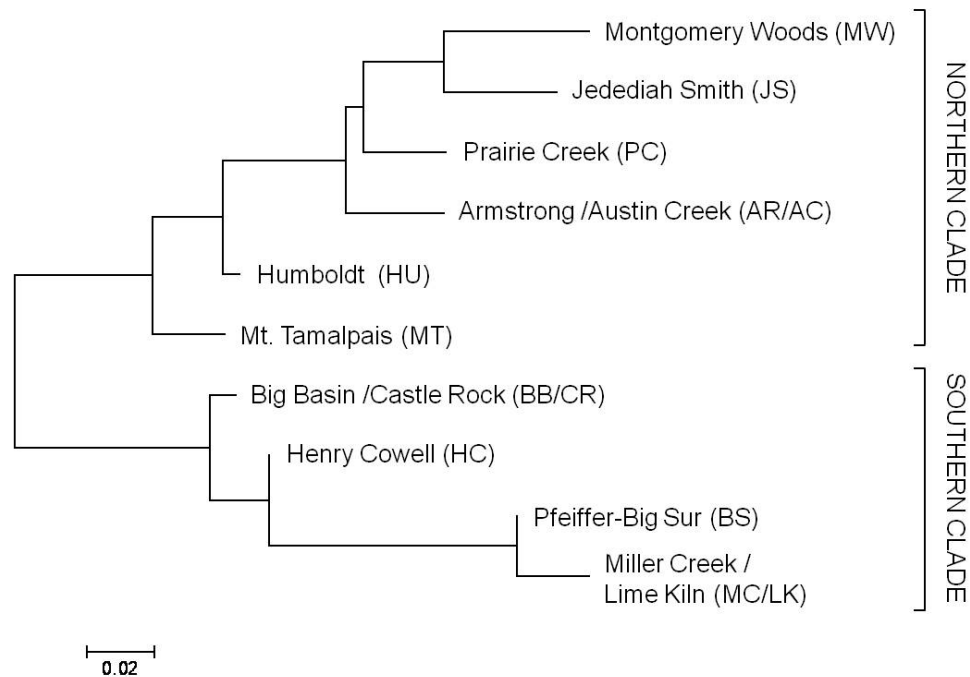


Figure 4—Neighbor-joining tree based on Nei's D_A distances calculated from combined old growth and second growth populations at each sampling location.

Discussion

In all 10 locations sampled throughout the range, second growth populations were not significantly different in their allelic compositions from adjacent old growth populations. Apparently, neither the decreased rate of stump sprouting in redwoods over the age of 400 years (Powers and Wiant 1970) nor the higher ratio of ramets per genet in second growth forests (Douhovnikoff et al. 2004) altered the genetic architecture of naturally regenerated forests to any significant extent.

Genetic diversity of coast redwood populations at the *Seq21E5* locus varied considerably, with higher diversity levels in the northern populations and lower levels in the southern populations. The MW population, the most genetically diverse, is centrally located in the large tract of continuous coast redwood forest ranging from central Sonoma County to just north of the Mendocino County-Humboldt County border. It is likely that long distance gene flow via pollen would be substantial in this region, leading to higher diversity levels than would be present in smaller, more fragmented forests. An alternative explanation for the high diversity in the MW population is that this region could be closer to the center of expansion of the modern range's founding population and, being older, would be more diverse.

The two southernmost populations along the Monterey County coast, BS and MC/LK, are the least genetically diverse and are classic examples of how fragmented populations are susceptible to fixation through genetic drift. These and other populations along the Big Sur coastline are separated from each other by a severe canyon topography and therefore have little or no exchange of pollen with each other or with the more diverse populations to the north. Furthermore, Viers (1996)

proposed that the warmer and drier conditions in the southern part of the range would cause greater difficulty in seedling establishment resulting in even higher rates of clonal reproduction than observed in northern forests. The high incidence of wildfires along the Big Sur coast would also lower diversity by promoting clonal over sexual reproduction. One can assume that the same evolutionary forces are at work on nuclear loci as well, although fixation will occur more slowly in a polyploid genome.

Although the Santa Cruz County populations (BB/CR and HC) had lower genetic diversity values and allele numbers compared to the northern populations, a larger sample study ($n = 530$) in Big Basin Redwoods State Park (Glavas 2006) detected 10 *Seq21E5* alleles compared to six in this study's combined BB/CR population ($n = 80$). Between the two studies, 11 total alleles were observed in this general location. Although many of these alleles were found in only one or a few individuals, their presence raises the possibility that this region may have once harbored a forest as diverse as its neighbors to the north but has since become less diverse through a reduction in rare allele frequency through genetic drift.

The division of populations in the neighbor-joining tree (*fig. 4*) into northern and southern clades with a boundary at or near the San Francisco Bay is in keeping with results from previous biochemical, physiological and genetic studies (Anekonda 1992, Douhovnikoff and Dodd 2011, Hall and Langenheim 1987). Certainly, there is no substitute for using multiple loci in phylogenetic studies, but the large population sample sizes used in this study, the high variability of the *Seq21E5* locus, and the low effective population size of the chloroplast genome all contributed to the detection of a north-south genetic division despite the use of only one locus.

While the major genetic transition zone between the northern and southern redwood populations has been identified by this and previous studies, the evolutionary relationships between redwood populations north of the San Francisco Bay remain unclear. The characterization and use of more chloroplast microsatellite loci would allow for a much more powerful phylogenetic analysis of that very important part of the redwood range.

Acknowledgments

I would like to thank the USDA Forest Service, the National Park Service, and the California Department of Parks and Recreation for issuing collecting permits; Rueben Bolton, Greg Brinegar, Michele Conrad and Steve Glavas for collecting assistance; Dr. Robert Fowler (San Jose State University) for the use of his laboratory for DNA extractions; and Bonnie Brinegar for generous financial support.

References

- Allnut, T.R.; Newton, A.C.; Lara, A.; Premoli, A.; Armesto, J.J.; Vergara, R.; Gardner, M. 1999. **Genetic variation in *Fitzroya cupressoides* (alerce), a threatened South American conifer.** *Molecular Ecology* 8: 975-987.
- Anekonda, T.S. 1992. **A genetic architecture study of coast redwood.** Berkeley, CA: University of California, Berkeley. Ph.D. dissertation. 285 p.
- Brinegar, C.; Bruno, D.; Kirkbride, R.; Glavas, S.; Udranszky, I. 2007. **Applications of redwood genotyping by using microsatellite markers.** In: Standiford, R.B.; Giusti,

G.A.; Valachovic, Y.; Zielinski, W.J.; Furniss, M.J., technical editors. Proceedings of the redwood region forest science symposium: What does the future hold? 2004 March 15-17; Rohnert Park, CA. Gen. Tech. Rep. PSW-GTR-194. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station: 47-56.

- Bruno, D.; Brinegar, C. 2004. **Microsatellite markers in coast redwood (*Sequoia sempervirens*)**. Molecular Ecology Notes 4: 482-484.
- Douhovnikoff, V.; Cheng, A.; Dodd, R.S. 2004. **Incidence, size and spatial structure of clones in second growth stands of coast redwood, *Sequoia sempervirens* (Cupressaceae)**. American Journal of Botany 91: 1140-1146.
- Douhovnikoff, V.; Dodd, R.S. 2011. **Lineage divergence in coast redwood (*Sequoia sempervirens*), detected by a new set of nuclear microsatellite loci**. The American Midland Naturalist 165: 22-37.
- Eliades N.-G.; Eliades D.G. 2009. **HAPLOTYPE ANALYSIS: software for analysis of haplotype data**. Distributed by the authors. Forest Genetics and Forest Tree Breeding, Georg-August University, Goettingen, Germany. <http://www.uni-goettingen.de/de/134935.html>.
- Felsenstein, J. 1985. **Confidence limits on phylogenies: an approach using the bootstrap**. Evolution 39: 783-791.
- Glavas, C.S. 2006. **Genotyping a second growth coast redwood forest: a high throughput methodology**. San Jose, CA: San Jose State University. M.S. thesis. 46 p.
- Hall, G.D.; Langenheim, J.H. 1987. **Geographic variation in leaf monoterpenes of *Sequoia sempervirens***. Biochemical Systematics and Ecology 15: 31-43.
- Nei, M.; Tajima, F.; Tateno, Y. 1983. **Accuracy of estimated phylogenetic trees from molecular data**. Journal of Molecular Evolution 19: 153-170.
- Powers, R.F.; Wiant, H.V. 1970. **Sprouting of old-growth coastal redwood stumps**. Forest Science 16: 339-341.
- Provan, J.; Powell, W.; Hollingsworth, P.M. 2001. **Chloroplast microsatellites: new tools for studies in plant ecology and evolution**. Trends in Ecology and Evolution 16: 142-147.
- Sawyer, J.O.; Gray, J.; West, G.J.; Thornburgh, D.A.; Noss, R.F.; Engbeck, J.H. Jr.; Marcot, B.G.; Raymond, R. 2000. **History of redwood and redwood forests**. In: Noss, R.F., editor. The redwood forest: history, ecology, and conservation of the coast redwoods. Washington, DC: Island Press: 7-38.
- Takezaki, N.; Nei, M.; Tamura, K. 2010. **POPTREE2: software for constructing population trees from allele frequency data and computing other population statistics with Windows interface**. Molecular Biology and Evolution 27: 747-752. <http://www.kms.ac.jp/~genomelb/takezaki/poptree2/index.html>.
- Veirs, D.D. 1996. **Ecology of the coast redwood**. In: LeBlanc, J., editor. **Proceedings of the conference on coast redwood forest ecology and management**. University of California Cooperative Extension Forestry Publication. Arcata, CA: Humboldt State University: 9-12.
- Weising, K.; Nybom, H.; Wolff, K.; Kahl, G. 2005. **DNA fingerprinting in plants: principles, methods, and applications**. 2nd edition. Boca Raton: CRC Press. 444 p.