

Applications of Redwood Genotyping by Using Microsatellite Markers¹

Chris Brinegar,² Dan Bruno, Ryan Kirkbride, Steven Glavas, and Ingrid Udranszky

Abstract

A panel of polymorphic microsatellite markers have been developed in coast redwood (*Sequoia sempervirens*). Two loci in particular (*Seq18D7-3* and *Seq21E5*) demonstrate the potential of microsatellite genotyping in the assessment of genetic diversity and inheritance in redwoods. The highly polymorphic *Seq18D7-3* marker provided evidence for the planting of non-local redwood seedlings in a restored area of a state park and was used to genotype progeny from a controlled cross. A putative chloroplast microsatellite locus (*Seq21E5*) was shown to be paternally inherited in two controlled crosses and can therefore be used as a marker for alleles carried in pollen. Significant differences in allelic diversity were observed at the *Seq21E5* locus in redwood populations with different ecological and management histories. Advantages of microsatellite analysis and its limitations (including the problem of null alleles) are discussed as well as its potential as a tool for forest management and conservation.

Key words: genotyping, microsatellites, redwood, Sequoia, simple sequence repeats

Introduction

Data from common garden experiments (Anekonda and Libby 1996), monoterpene characterization (Hall and Langenheim 1987), and allozyme analysis (Rogers 1997, 1999, 2000) show that coast redwood [*Sequoia sempervirens* (D. Don) Endl.] possesses significant genetic variation. Much of its genetic diversity might be attributed to the fact that it is a hexaploid organism. The potential origins of hexaploidy in redwood have been reviewed by Ahuja and Neale (2002). There have been limited molecular genetics studies on redwoods using DNA-based approaches. Neale and others (1989) used restriction fragment length polymorphism (RFLP) to show that both chloroplast and mitochondrial DNA are inherited paternally in redwood. Evidence for variation in the redwood chloroplast genome has also been demonstrated by RFLP analysis (Ali and others 1991). In order to understand the complexities of the redwood genome and its mechanisms of inheritance in greater detail, it will be necessary to have many DNA markers in hand that are amenable to rapid and relatively inexpensive genetic analyses. Microsatellites, also known as simple sequence repeat (SSR) markers, are non-coding, co-dominant loci with high mutation rates that have been used extensively for inheritance and population diversity studies in plants and animals (Parker and others 1998).

We have characterized a panel of five polymorphic microsatellite loci from a

¹ This paper was presented at the Redwood Science Symposium: What does the future hold? March 15-17, 2004, Rohnert Park, California.

² Conservation Genetics Laboratory, Department of Biological Sciences, San Jose State University, San Jose, CA 95192. email: cbrinega@email.sjsu.edu

redwood genomic library that have shown utility as genetic markers (Bruno 2002, Bruno and Brinegar 2004). This paper presents preliminary data that demonstrate how these markers can be used to assess inheritance in controlled crosses and characterize genetic diversity within and between populations. The potential of microsatellite genotyping in forest management will also be addressed.

Methods

Redwood branchlets were collected in coin envelopes and subsequently dried with desiccant at room temperature prior to DNA extraction. Redwood genomic library construction, screening, microsatellite clone characterization, primer sequences, DNA extraction, PCR conditions and gel electrophoresis are described elsewhere (Bruno and Brinegar 2004). PCR products were separated on non-denaturing, high-resolution Spreadex EL 300 or EL 500 gels (Elchrom Scientific, Cham, Switzerland). M3 DNA markers (Elchrom) were used to estimate allele sizes. Three microsatellite markers were used in this report, but sequences for all five clones [*Seq37G5* (CT_n), *Seq21E5* (CTTA_n), *Seq20E5*(CA_n), *Seq18D7-3* (CTT_n), and *Seq8E8* (CA_n)] are available in the GenBank database (accession numbers AY562165, AY562166, AY562167, AY562168, and AY562169, respectively).

Results and Discussion

Redwood microsatellite polymorphism

Microsatellite *Seq18D7-3* is the most polymorphic locus in the current panel of markers with 13 alleles observed so far in several test populations. Eight alleles were observed in the genotypes of only 12 individual trees from Big Basin Redwoods State Park (*fig. 1*). In this example, each individual had one to three alleles with the microsatellite repeats ranging in size from (CTT)₅ to (CTT)₂₁. Allele sizes differed by multiples of three base pairs (bp) as expected, and sequencing of individual alleles eluted from the gel confirmed that size polymorphisms were, in fact, due only to changes in the number of tandem CTT repeats.

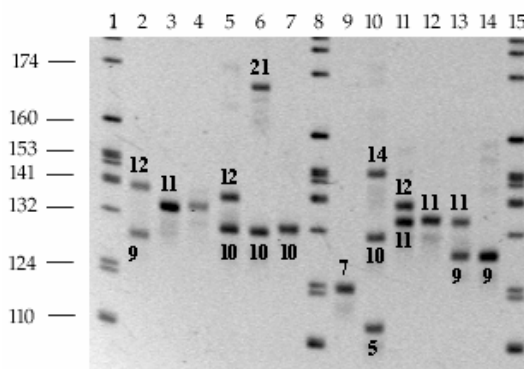


Figure 1—Size polymorphism of alleles at the *Seq18D7-3* microsatellite locus in 12 randomly selected redwoods from Big Basin Redwoods State Park. Numbers above or below bands indicate the number of CTT repeat units in that allele. Lanes 1, 8 and 15 are DNA markers. Marker sizes (base pairs) are indicated at left.

Seq8E8 is also very polymorphic with 10 known alleles while *Seq37G5* and *Seq20E5* have three and four known alleles, respectively. However, these three markers have not been tested on as many populations as the other two and may actually have higher degrees of polymorphism. *Seq21E5* is unique in that it shows very high polymorphism, but each individual (out of several hundred tested) possesses only one allele. This observation suggests a cytoplasmic origin, presumably from chloroplast DNA, since microsatellite sequences are common in chloroplasts (Provan and others 2001). Further data on the allelic diversity and inheritance of the *Seq21E5* marker are provided later in this report.

In theory, since redwoods are hexaploid, the number of alleles at a nuclear locus could range from one to six per individual. Redwoods may be autoallohexaploid (Saylor and Simons 1970) with an AAAABB type of genome ancestry. If so, it is possible that some microsatellite loci may appear only in the A or B chromosome sets. This could restrict the maximum number of alleles per individual to two or four. We have yet to observe more than four alleles per locus in any of our markers. Though the true nature of redwood hexaploidy is still unresolved (Ahuja and Neale 2002), microsatellite DNA clones have potential as probes for *in situ* chromosome hybridization to help resolve this question.

Detecting non-indigenous redwoods

In a study to test the utility of our microsatellite markers, a group of closely spaced redwoods in Big Basin Redwoods State Park were genotyped to assess their genetic relationship. The trees are along a 50 m edge of a restored meadow with 12 young redwoods (0.5 to 7 m tall) growing between two older fairy ring groves to the north and south. *Figure 2A* shows the *Seq18D7-3* genotypes of these 12 small redwoods (trees 7-18) and six redwoods each from the southern grove (trees 1-6) and northern grove (trees 19-24). Both groves had different genotypes, but all trees within a grove shared the same allelic pattern. Allozyme data (Rogers 2000) indicate that redwoods in a fairy ring may not all be clones of the parent stump, so identification of trees within these groves as clonal is tentative until more loci are genotyped.

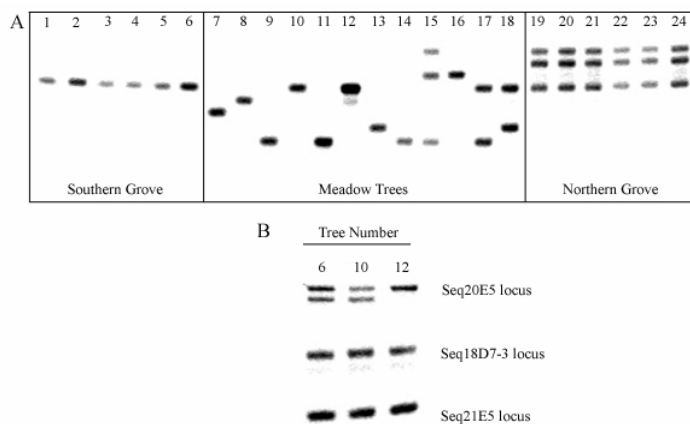


Figure 2—Microsatellite genotyping to assess relationships between redwoods in a restored area. A. *Seq18D7-3* genotypes of redwoods in two fairy ring groves and along a meadow's edge between the groves. B. Genotypes at three loci (*Seq21E5*, *Seq18D7-3*, *Seq20E5*) of redwoods 6, 10 and 12. See text for discussion.

None of the small trees between the groves displayed the three-allele pattern of the northern grove, and only two (trees 10 and 12) had the same single-allele pattern as the southern grove. Seven of the small redwoods had no alleles in common with either of the groves. Comparison of trees 10 and 12 with tree 6 from the southern grove at two additional microsatellite loci differentiated tree 12 from the others by its extra allele at *Seq20E5* (fig. 2B).

Due to a lack of shared alleles, it is apparent that most of the small trees (7, 8, 9, 11, 13, 14 and 16) were neither clones of either grove nor derived from seed produced by either grove. Also, because of a stream on one side and a meadow on the other, there were no other seed-producing redwoods within 50 m of this site. Tree 10 appears to be the only one that could possibly be a clone of one of the groves (southern). Alternatively, it could be a highly inbred relative.

Although this limited microsatellite data suggests that the two groves are clonal and that few of the small redwoods are closely related to the groves, an analysis with more markers and sampling of other trees in the vicinity would be required for definitive proof. However, after researching the recent history of this section of Big Basin, the interpretation of the data became easier. As part of prior restoration projects, Girl Scout troops had planted seedlings from nursery stock in several areas near the park headquarters—including along the meadow between the northern and southern groves! Based on this information and the microsatellite data, there is a high probability that most of the 12 small redwoods are not indigenous to Big Basin.

The practice of planting non-local redwood seedlings in parks and private timber land could potentially lead to the “genetic pollution” of populations that are well adapted to their local environment. This potential problem has been addressed previously by Millar and Libby (1989). A thorough microsatellite analysis could be used to assess the presence of non-native trees in areas suspected of being planted with less well adapted nursery stock.

Inheritance in controlled crosses

In angiosperms chloroplasts are maternally or biparentally inherited (Provan and others 2001) whereas paternal inheritance of chloroplasts has been reported in a number of conifers, such as pine (Dyer and Sork 2001). Neale and others (1989) provided restriction fragment length polymorphism (RFLP) evidence that chloroplasts, as well as mitochondria, are paternally inherited in coast redwood. The use of chloroplast genetic markers should, therefore, allow alleles from the pollen donor to be monitored in controlled crosses. Since redwood microsatellite marker *Seq21E5* had properties of a chloroplast locus (one allele per individual and typically higher PCR product yield than nuclear loci due to more copies per cell), we tested its possible paternal inheritance using tissue from known redwood families supplied by the UC-Berkeley Russell Plantation.

Three parents were used in two crosses with Arc154 being the female parent in one cross and the male parent in the other. In Arc154 x Arc28 (fig. 3A), the female parent’s allele size was 120 bp and the male parent's allele size was 140 bp. All five of the tested sibs had the 140 bp allele. All five of the S1 x Arc154 sibs inherited the 120 bp paternal allele (fig. 3B). These results support the paternal inheritance of chloroplasts in redwoods, but more work is needed to determine if maternal “leakage” of chloroplasts occurs during fertilization. In the conifer *Chamaecyparis obtusa*, chloroplast DNA was inherited primarily through the pollen but a small amount (2.5 percent) of maternal leakage was detected (Shiraishi and others 2001).

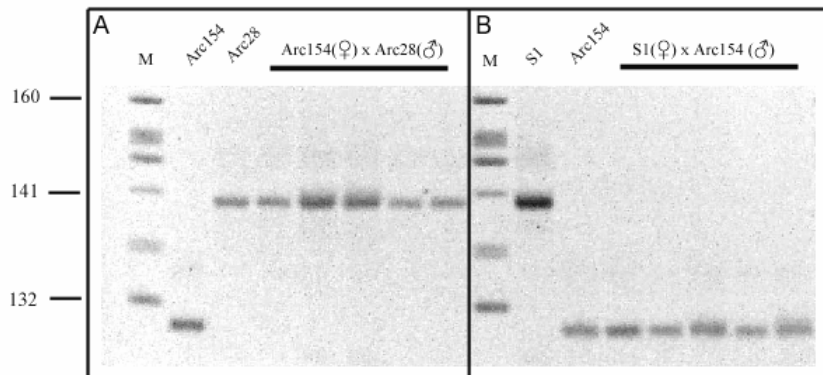


Figure 3—Paternal inheritance of the *Seq21E5* locus in two controlled crosses. A. Genotypes of parents Arc154 and Arc28 with genotypes of five sibs. B. Genotypes of parents S1 and Arc154 with genotypes of five sibs. DNA markers are in lanes labeled M with their sizes (base pairs) indicated at left.

The *Seq18D7-3* locus (presumably a nuclear marker) was also tested on the S1 x Arc154 family (*fig. 4*). The female parent (S1) had one allele at 141 bp while the male parent (Arc154) had a 138 bp allele (one less repeat unit). The five sibs from this cross displayed three different genotypes: both parental alleles (sibs A and B), no visible alleles (sib C), or only the maternal allele (sibs D and E). The absence of the 138 bp allele in sibs D and E could be explained by the donation of a paternal null allele at this locus. The absence of alleles in sib C could be due to null allele contributions from both parents or a failure in the amplification reaction; however, the *Seq21E5* locus amplified well using the same DNA extract from sib C.

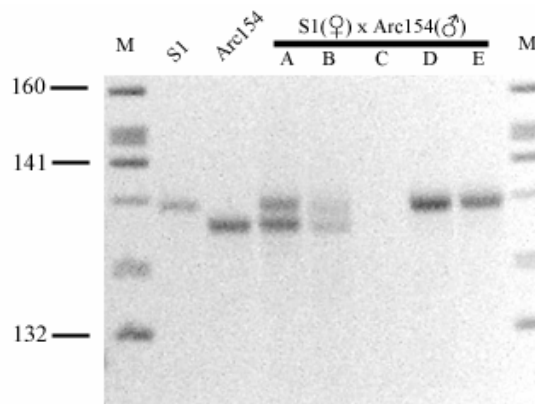


Figure 4—Inheritance of *Seq18D7-3* alleles in a controlled cross (S1 x Arc154). DNA markers are in lanes labeled M with their sizes (base pairs) indicated at left.

Null alleles are major obstacles in the interpretation of inheritance data, especially in polyploids where there could be a higher probability of their occurrence (Weeden and Wendel 1989). Their presence can be inferred from the absence of an amplifiable allele from one parent (as above) or from a low observed population heterozygosity (H_o) at the locus. (Unfortunately, due to the difficulty in estimating allele dosage in polyploids from allele staining intensities, expected heterozygosities

cannot be easily estimated from microsatellite data.) Indications of null alleles at three of our four presumptive nuclear microsatellite markers were found in a population of second-growth redwoods in the Santa Cruz Mountains. Their observed H_o values were relatively low, ranging from 0.28 to 0.44 (Bruno and Brinegar 2004). *Seq18D7-3*, however, had a higher H_o of 0.80, indicating that null allele frequency is low in that southern population. S1 and Arc154 are clones from northern populations where the null allele frequency could be higher. In a redwood allozyme study (Rogers 1997) there was no evidence of null alleles in the six enzyme systems tested. Microsatellite alleles, however, being non-coding sequences, do not impart any known selective advantages to the organism and will have no deleterious consequences if lost.

These results demonstrate that microsatellite markers can be useful in characterizing the results of controlled crosses. Libby and others (1981) provided evidence of inbreeding depression in redwoods. Microsatellite genotyping might also have potential in assessing inbreeding in redwood breeding programs.

Genetic diversity within and between populations

With highly polymorphic genetic markers, the allelic richness of populations—a measure of genetic diversity (Petit and others 1998)—can be measured and compared. To avoid the problem of null alleles, we examined the allelic frequencies at the *Seq21E5* locus in a total of 194 trees from two northern old-growth populations (Humboldt County) and two southern second-growth populations (Santa Cruz County). Histograms of allele frequencies are shown in *figure 5*.

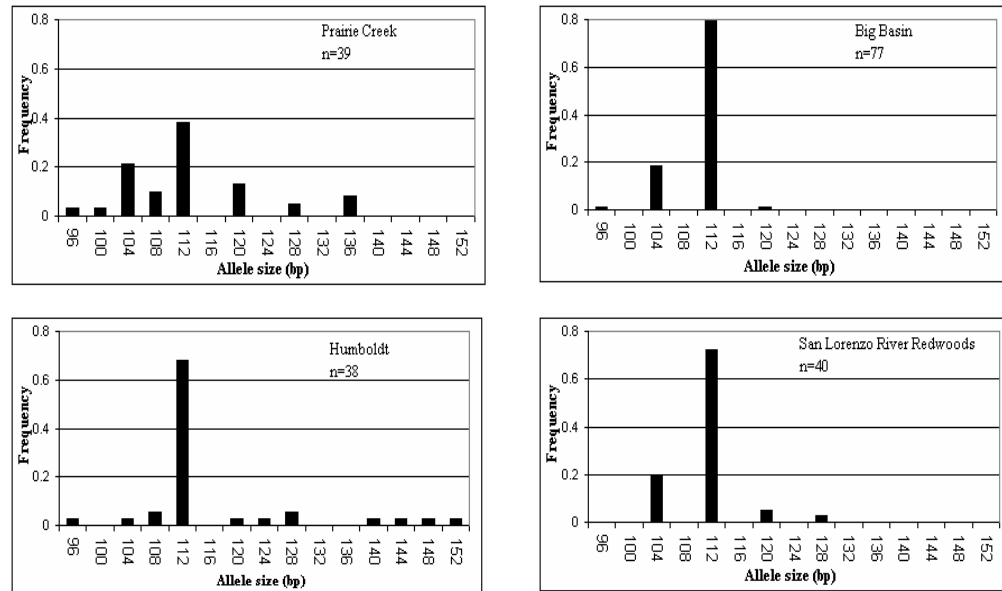


Figure 5—*Seq21E5* allele frequencies in four redwood populations. Individuals in Prairie Creek Redwoods State Park and Humboldt Redwoods State Park were from old-growth populations. Individuals from Big Basin Redwoods State Park and the San Lorenzo River Redwoods (a private reserve near Castle Rock State Park in Santa Cruz County) were from second-growth populations.

The 112 bp allele was by far the most common, found in nearly 68 percent of all trees tested. Frequency of this allele in the four populations ranged from a low of almost 40 percent in Prairie Creek to a high of 80 percent in Big Basin. However, the allelic richness (the number of different alleles observed) in the two old-growth populations was significantly higher than that of the second-growth populations. The Prairie Creek and Humboldt Redwoods State Park old-growth populations had eight and 11 different alleles, respectively, while the Big Basin Redwoods State Park and the San Lorenzo River Redwoods second-growth populations had only four alleles each.

There are several possible explanations for the lower allele numbers in the southern populations. Collections were made to avoid the sampling of clones, but it is possible that the second-growth population samples included some trees that sprouted from the same parent, thereby reducing the chance of sampling other trees with different alleles. However, the low allelic richness of the Big Basin population, in spite of the high sample number ($n = 77$), argues against a significant sampling bias. It has been proposed that there is a higher percentage of clonal reproduction in the southern redwood forests due to the warmer, drier climate that makes seedling establishment more difficult (Veirs 1996). This would reduce the creation and spread of new alleles by limiting the number of trees produced through sexual reproduction. Some of the loss in diversity at this locus could also be caused by the harvesting of the old-growth forests. Powers and Wiant (1970) found that stump sprouting decreased significantly when redwoods over the age of 400 years were cut. Therefore, the primarily clonal second-growth trees that replaced the clear cut old-growth forests in the southern region might under-represent the diversity of the original forest. Finally, differences in founding populations may account for the observations.

Since *Seq21E5* is a putative chloroplast locus, the allelic diversity it displays in various populations may not necessarily correlate with diversity at nuclear loci. Also, one should not draw too many conclusions by comparing genotypes of old-growth northern forests with second-growth southern forests. More valid comparisons would be to genotype adjacent old-growth and second-growth populations or northern and southern old-growth populations. However, the many questions that these observations create are amenable to being answered by more thorough genetic analysis using allozymes or microsatellites.

Conclusions

The preliminary data we have presented on microsatellite genotyping of redwoods demonstrates the potential of this approach for the study of redwood inheritance and diversity. Using microsatellite genotyping to detect non-local nursery stock planted in parks or timberland will ensure that native redwood populations maintain their adaptive traits. By tracking paternal inheritance with chloroplast microsatellite loci, pollination can be studied in more detail. In spite of problems caused by null alleles, genotyping progeny from controlled crosses can provide useful information to breeders.

The development and characterization of polymorphic microsatellite markers can be a long and expensive undertaking, but once available the PCR-based analysis is rapid and cost effective. Using our current system, one person can extract, amplify and genotype 40 redwoods at two loci in two days at a reagent and supply cost of

approximately \$70. We are working on a 96-well plate format for both DNA extraction and PCR that will increase the two-day output to at least 180 samples. Some loci might be amenable to PCR multiplexing which could eventually double or triple that output. Using a multi-capillary electrophoresis system (with fluorescent detection of labeled multiplexed PCR products) will further increase data production while reducing labor and analysis time. Such high throughput will make large-scale population studies much more feasible than they are at present. We will continue to characterize additional redwood microsatellite loci which will allow a full genetic analysis of individual trees and populations in controlled and natural environments.

Microsatellite genotyping, or other DNA-based methods, will no doubt become another routine tool for redwood geneticists, forest managers, and conservationists of the future. Decisions to conserve redwood forests could, in part, be dictated by an analysis of their genetic structure. The redwood genome is complex but it will eventually, and no doubt begrudgingly, provide us with the information we need to make good forest management decisions.

Acknowledgments

We thank the California Department of Parks and Recreation, the National Park Service, and the Sempervirens Fund for permission to collect redwood tissue. We are also grateful to Dr. William Libby and the UC-Berkeley Russell Plantation for providing tissue from controlled crosses. This work was partially funded by the Graduate Research Fellowship program in the Department of Biological Sciences, San Jose State University.

References

- Ahuja, M.R.; Neale, D.B. 2002. **Origins of polyploidy in coast redwood (*Sequoia sempervirens* (D. Don) Endl.) and relationship of coast redwood to other genera of Taxodiaceae.** *Silvae Genetica* 51: 93-100.
- Ali, I.F.; Neale, D.B.; Marshall, K.A. 1991. **Chloroplast DNA restriction fragment length polymorphism in *Sequoia sempervirens* (D. Don) Endl., *Pseudotsuga menziesii* (Mirb.) Franco, *Calocedrus decurrens* (Torr.), and *Pinus taeda* L.** *Theoretical and Applied Genetics* 81: 83-89.
- Anekonda, T.S.; Libby, W.J. 1996. **Effectiveness of nearest-neighbor data adjustment in a clonal test of redwood.** *Silvae Genetica* 45: 46-51.
- Bruno, D.; Brinegar, C. 2004. **Microsatellite markers in coast redwood (*Sequoia sempervirens*).** *Molecular Ecology Notes* 4: 482-484.
- Bruno, D. 2002. **Isolation and characterization of microsatellite loci in *Sequoia sempervirens*.** San Jose, CA: San Jose State University; 97 p. M.S. thesis.
- Dyer, R.J.; Sork, V.L. 2001. **Pollen pool heterogeneity in shortleaf pine, *Pinus echinata* Mill.** *Molecular Ecology* 10: 859-866.
- Hall, G.D.; Langenheim, J.H. 1987. **Geographic variation in leaf monoterpenes of *Sequoia sempervirens*.** *Biochemical Systematics and Ecology* 15: 31-43.
- Libby, W.J.; McCutchan, B.G.; Millar, C.I. 1981. **Inbreeding depression in selfs of redwood.** *Silvae Genetica* 30: 15-25.

- Millar, C.I.; Libby, W.J. 1989. **Restoration: Disneyland or a native ecosystem? A question of genetics.** Restoration Management Notes 7: 18-24.
- Neale, D.B.; Marshall, K.A.; Sederoff, R.R. 1989. **Chloroplast and mitochondrial DNA are paternally inherited in *Sequoia sempervirens* (D. Don) Endl.** Proceedings of the National Academy of Sciences of the USA 86: 9347-9349.
- Parker, P.G.; Snow, A.A.; Schug, M.D.; Booton, G.C.; Fuerst, P.A. 1998. **What molecules can tell us about populations: choosing and using a molecular marker.** Ecology 79: 361-382.
- Petit, R.J.; El Mousadik, A.; Pons, O. 1998. **Identifying populations for conservation on the basis of genetic markers.** Conservation Biology 12: 844-855.
- 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.
- Rogers, D.L. 1997. **Inheritance of allozymes from seed tissues of the hexaploid gymnosperm, coast redwood, *Sequoia sempervirens* (D. Don) Endl. (Coast redwood).** Heredity 78: 166-175.
- Rogers, D.L. 1999. **Allozyme polymorphisms discriminate among coast redwood (*Sequoia sempervirens*) siblings.** Journal of Heredity 90: 429-433.
- Rogers, D.L. 2000. **Genotypic diversity and clone size in old-growth populations of coast redwood (*Sequoia sempervirens*).** Canadian Journal of Botany 78: 1408-1419.
- Saylor, L.C.; Simons, H.A. 1970. **Karyology of *Sequoia sempervirens*: Karyotype and accessory chromosomes.** Cytologia 35: 294-303.
- Shiraishi, S.; Maeda, H.; Toda, T.; Seido, K.; Sasaki, Y. 2001. **Incomplete paternal inheritance of chloroplast DNA recognized in *Chamaecyparis obtusa* using an intraspecific polymorphism of the trnD-trnY intergenic spacer region.** Theoretical and Applied Genetics 102: 935-941.
- Veirs, D.D. 1996. **Ecology of the coast redwood.** In: LeBlanc, J., ed. 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.
- Weeden, N.F.; Wendel, J.F. 1989 **Genetics of plant isozymes.** In: Soltis, D.E.; Soltis, P.S., eds. Isozymes in plant biology, Advances in Plant Sciences Series, Vol. 4. Portland, OR: Dioscorides Press; 46-72.