2002

Isolation and characterization of microsatellite loci in Sequoia Sempervirens

Daniel P. Bruno
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation

Bruno, Daniel P., "Isolation and characterization of microsatellite loci in Sequoia Sempervirens" (2002).
Master's Theses. 2350.
DOI: <https://doi.org/10.31979/etd.6wjt-85w8>
https://scholarworks.sjsu.edu/etd_theses/2350

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

**ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600**

UMI[®]

ISOLATION AND CHARACTERIZATION OF MICROSATELLITE LOCI IN
SEQUOIA SEMPERVIRENS

A Thesis

Presented to

The Faculty of the Department of Biological Sciences

San Jose State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science

by

Daniel P. Bruno

December 2002

UMI Number: 1411603

Copyright 2002 by
Bruno, Daniel Paul

All rights reserved.

UMI[®]

UMI Microform 1411603

Copyright 2003 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© 2002

Daniel P. Bruno

ALL RIGHTS RESERVED

APPROVED FOR THE DEPARTMENT OF BIOLOGICAL SCIENCES

Dr. Chris Brinegar

Dr. Robert Fowler

Mr. Stephen Schramm, Milagen Inc.

APPROVED FOR THE UNIVERSITY

ABSTRACT

ISOLATION AND CHARACTERIZATION OF MICROSATELLITE LOCI IN *SEQUOIA SEMPERVIRENS*

by Daniel P. Bruno

The first microsatellite loci for California Coast Redwood, *Sequoia sempervirens*, are described. A redwood genomic DNA library was constructed using a plasmid vector. Microsatellite containing clones were isolated, sequenced, amplified, and characterized. The successful identification of 37 possible markers for *S. sempervirens* is described. Twenty of those markers are optimized for PCR and can be used for future studies. Four of those twenty were characterized in great detail and one was used in a simple study to show the usefulness of microsatellite markers for the species. Of the markers that were analyzed in detail, two show complex polymorphism patterns while the remaining two appear to be monomorphic or contain a low frequency of polymorphism. These genetic markers lay the foundation for population studies involving parentage analysis, pollen dispersal, genetic diversity, and can also be used in forensics or other applications requiring a specific identification of individual trees.

ACKNOWLEDGEMENTS

I want to dedicate this thesis to the brave environmental activists on the front lines in the forest that struggle daily to save the California coast redwood. Without your voice, more of this majestic species would have already been destroyed. Chris Brinegar taught me that environmental activism does not always involve a tree sit or civil disobedience; every conservation movement needs its share of scientists and lawyers. Chris is a source of inspiration for many young scientists to fight noble and just causes in their field. May no one ever underestimate his compassion, respect, and understanding of our environment.

Mr. Stephen Schramm gave me my first chance at being a scientist in “the real world.” His guidance and leadership have shaped me into the proactive scientist that I am today. Steve’s participation in the conservation community inspires me to stand up for the causes I believe in. He has been an excellent role model for both professional and personal reasons. Robert Fowler taught me my basic foundation in genetics as an undergraduate at San Jose State University. Although our communication was mostly inside the classroom, his impact on my discipline as a scientist is immeasurable.

The scale of this work would not have been possible without the generosity of Dr. Ronald Davis at the Stanford Genome Technology Center (SGTC) in Palo Alto, CA. At SGTC Dr. Richard Hyman was an invaluable resource for the library construction. Audrey Southwick provided detailed methodology throughout my work and Keith Anderson provided practical information on PCR optimization. To the many other minor

contributors at SGTC not mentioned, thank you for all the conversations and ideas that added to this work. A graduate research award from the Marianna Pisano Memorial Scholarship at San Jose State University provided the seed money for this project. The remainder of this project was funded by contributions from SGTC and myself.

My family has been the greatest support system anyone could ask for. My parents always encouraged me to follow my own desires and aspirations without limit. I will forever be grateful for their support and love. My wife Jennifer has encouraged me to move forward at times when the research seemed dull or difficult. She endured many weekends in the lab by my side and provided constant encouragement during my work.

Throughout my education at Silver Creek High School in San Jose, Sharon Metoyerjones taught me that learning is a lifelong process. The impact she had on my life to be a better person and fight the right causes will always be with me. To my first science teachers at Silver Creek, thank you for challenging me to be curious about science and the natural world; it has instilled in me a constant desire to want to learn more about nature.

Amanpreet Brar gave me confidence in my scholastic abilities to do a Master's degree at San Jose State University. Without her encouragement, I would have never entered into the program. Thanks to all the people in my life who in one way or another led to the culmination of this thesis. The many obstacles that I have run into along the way were eventually overcome thanks to your support.

TABLE OF CONTENTS

<u>Subject</u>	<u>Page</u>
LIST OF TABLES.....	ix
LIST OF FIGURES.....	x
INTRODUCTION.....	1
History of the organism.....	1
Physical characteristics.....	3
Redwood genetics.....	4
Use of microsatellite DNA (SSRs) as molecular markers.....	6
The importance of molecular DNA markers for coast redwood.....	7
MATERIALS AND METHODS.....	9
Tissue collection and DNA extraction.....	9
Insert preparation.....	9
Vector preparation.....	11
Ligation reaction.....	11
Electroporation of recombinant DNA.....	12
Validation of library.....	13
Isolation of recombinant clones in a high-throughput format.....	14
Stamping and fixing of clones on nylon in a high-throughput format.....	15
Hybridization of microsatellite probes to target DNA.....	16
Chemiluminescent detection of hybridized microsatellite probes.....	17
Isolation of positive clones and sequencing of recombinant DNA.....	18
PCR optimization of microsatellite loci.....	18
Validation and sequencing of amplified products.....	20
Validation of polymorphism for seven microsatellite loci.....	20
Practical application of microsatellite markers.....	21
RESULTS.....	22
Tissue collection and DNA extraction.....	22
Insert DNA preparation.....	23
Insert and vector quantitation using PicoGreen®.....	29
Determination of library titer.....	30
Validation of library.....	31
Isolation of recombinant clones in a high-throughput format.....	34
Hybridization to target DNA and positive detection.....	38
Sequencing results.....	41
PCR optimization of microsatellite loci.....	45
Validation of microsatellite allele sequence	49

RESULTS (continued)

Validation of polymorphism for seven microsatellite loci.....	52
Practical application of a microsatellite marker.....	57
DISCUSSION.....	59
LITERATURE CITED.....	63
APPENDIX A: Large Scale Isolation of DNA from Redwood Needles.....	67
APPENDIX B: Optimization of Restriction Endonuclease Digestion.....	69
APPENDIX C: Dye Terminator Cycle Sequencing Setup and Purification.....	71
APPENDIX D: Bacterial Colony Lysis and Fixation of Target DNA to Nylon.....	72
APPENDIX E: DNA Hybridization to Fixed DNA Targets on Nylon.....	74
APPENDIX F: Chemiluminescent Detection of Biotin Labeled Probes.....	75
APPENDIX G: Manual Processing of BioMax Light Film.....	77
APPENDIX H: Complete Flanking Sequence for Microsatellites Identified.....	78
APPENDIX I: PCR Amplification of Genomic DNA.....	96
APPENDIX J: Examples of False Positives from Library Screening.....	97

LIST OF TABLES

<u>Table</u>	<u>Page</u>
1. Titer of recombinant library glycerol cell stock.....	30
2. Summary of sequence alignments from 192 random clones.....	32
3. Summary of probes used in recombinant clone screening	38
4. Summary of clones containing perfect dinucleotide motifs.....	42
5. Summary of clones containing perfect trinucleotide motifs.....	43
6. Summary of clones containing the largest perfect motifs.....	43
7. Summary of clones containing compound, imperfect or complex motifs.....	44
8. Optimized universal PCR cycling condition.....	46
9. Prioritized list of 20 microsatellite candidates using a universal PCR condition...	47
10. Summary of sequenced alleles from amplified products.....	50
11. Summary of total clone failure for major stages in SSR development.....	57

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
1. Genomic DNA extraction from the Father of the Forest	22
2. <i>Sau3AI</i> restriction digest optimization on redwood genomic DNA	23
3. Insert DNA preparative agarose gel from 30 min <i>Sau3AI</i> digest	24
4. Insert DNA preparative agarose gel from 60 min <i>Sau3AI</i> digest	24
5. Excision of 0.5 - 3.0 kbp DNA fragments from agarose gel (30 min digest)	25
6. Excision of 0.5 - 3.0 kbp DNA fragments from agarose gel (60 min digest)	25
7. Both excised gel plugs inserted into a new agarose gel prior to electrophoresis ...	26
8. Concentrated DNA fragments in agarose gel after reversing polarity	26
9. Excision of concentrated DNA fragments from preparative agarose gel	27
10. Verification of enriched DNA fragments in agarose gel	28
11. Quantitation of insert and vector DNA for ligation.....	29
12. Purified clones digested with <i>HindIII</i> and <i>EcoRI</i> in agarose gel	31
13. Scatter plot of insert sizes based on sequence data	33
14. Photograph of custom colony picker	34
15. Photograph of high-throughput cassette for HiGro shaker	35
16. Photograph of HiGro high-capacity incubating shaker	36
17. Illustration of sample tracking procedure for hybridization screening	37
18. Scanned image of BioMax film showing a genuine positive signal	39
19. Scanned image of BioMax film showing false positive signals	40

20. PCR products from 20 markers using a universal condition	49
21. Validation of 18D7-3 marker across 12 individuals	53
22. Validation of 21E5 marker across 12 individuals	54
23. Validation of 23F9 marker across 12 individuals	55
24. Validation of 20B3 marker across 12 individuals	56
25. Simple reproductive study using 18D7-3 marker	58

INTRODUCTION

The California coast redwood (*Sequoia sempervirens*) is an ancient species of conifer that grows naturally in the foggy coastal mountain ranges of northern and central California. Their current range extends from the southwestern tip of Oregon to southern Monterey County, only penetrating the coast by about 30 miles inland. This narrow section of land is discontinuous and covers approximately 2500-3000 square miles (Olsen et al., 1990). The distribution of coast redwood within this area is fragmented, primarily due to geologic features, natural barriers, regional climate, and human disturbance. The limited and sparse populations of redwood groves seen today reflect the last natural remnants of a once greater distribution in North America that has receded significantly in just the past 3,000 years.

History of the Organism

The precise ancestry of *S. sempervirens* is a matter of debate, but Miller (1977, 1988) suggests a phylogenetic history of putative ancestors that date back to the early Mesozoic era, which began some 245 million years ago. The Taxodiaceae family in which *S. sempervirens* was once classified is recognizable as a family by the middle of the Triassic period based on seed cones from silicified peat (Yao et al., 1997). Seed cones resembling that of the *Sequoia* genus (Chaney, 1951; Miller, 1977) and fossilized leafy imprints show evidence that *Sequoia* may have existed by the Jurassic period, approximately 208-144 million years ago.

According to Miller (1977, 1988) the Taxodiaceae family underwent a great period of diversification into several lineages during the late Mesozoic era, and many of those lineages are now extinct. By the early to middle Tertiary period (66-36 million years ago), the range of the *Sequoia* genus is thought to have reached present day Wyoming, Colorado, Idaho, North Dakota and Canada (Chaney, 1951). Toward the end of the Tertiary period, the mid-latitude range of the genus in North America began to decline, probably due to increasing temperatures and a more arid climate caused by geologic uplifting of the Sierran and Cascade mountain ranges. The beginning of the Quaternary period (1.6 million years ago) brought about extreme cyclic variations in temperature that further reduced the range of the genus to primarily the coastal areas of present day California.

Today *S. sempervirens* represents the last living species of the *Sequoia* genus. It was previously suggested (Eckenwalder, 1976; Hart, 1987; Miller, 1988) that the Taxodiaceae family should be merged with the Cupressaceae family, and recent molecular DNA studies have led to the reclassification of the species by merging the groups together (Brunsfield et al., 1994; Tsumura et al., 1995; Stefanović et al., 1998; Kusumi et al., 2000). *S. sempervirens* is therefore currently classified as belonging to the Cupressaceae family. The exact lineage of the genus *Sequoia* is often a matter of conjecture due to the lack of fossilized evidence. Leaf imprints are not typically considered sufficient evidence to associate divergence with a specific time period, thus lineages are often difficult to distinguish. It is preferable to study the evolution of reproductive structures (seed cones) to track changes in development over time in relation

to the fossil record. Since *S. sempervirens* is such an ancient species, fossilized seed cones found intact are quite rare.

Physical Characteristics

California coast redwood is the tallest living organism on Earth, reaching heights in excess of 100 meters. The canopy of an average size tree (<200 years old) has a pyramidal shape, while “old growth” (>200 years old) usually has a broad, complex structure due to extended weathering. The lower branches are typically horizontal and somewhat droopy. Leaves that are highly exposed at the top of the tree are awl-like, while less exposed leaves exhibit a flattened or spreading appearance. The bark is a brownish-red color (shades can vary by age) and can grow as thick as 30 centimeters. The root system grows laterally and lacks a taproot (Lindquist, 1974).

The California coast redwood is a monoecious species, having both pollen and seed cones developing on different branches of the same tree. Seed cones are terminal and weigh approximately 2-4 grams. Their size is relatively small, measuring only 10-30 millimeters in length. Seed cones are dropped during the rainy season (September through March) with pollen being shed during rainless periods (November through March). The viability of redwood seed varies between locations (Ornduff, 1998) but x-ray examination (Roy, 1966) estimates the viability on average to be somewhere between 1 and 32 percent. Seed germination takes place in three to six weeks. In optimal conditions, saplings can grow up to 2 meters in height and 2.5 centimeters in diameter each year until mature.

As an alternative to sexual reproduction, redwood has the ability to clone itself asexually. Clones or “sprouts” can be seen growing from roots, stumps or dead wood. These sprouts sustain the stump root system (Cole, 1983) and eventually form their own root systems giving rise to the formation of circles of trees or “fairy rings” (Olson et al., 1990). The factors that affect the frequency of sprouting are not understood. In a 40-year-old stand, Lindquist (1979) reported more than 90 percent of cut stumps sprouted, while Powers and Wiant (1970) observed only 30 percent of cut stumps sprouted from an old growth stand. This signifies that tissue from older trees is more difficult to propagate, even by taking cuttings and rooting them (Olsen et al., 1990). Generally, it is thought that factors which induce stress on the organism will increase stump sprout frequency. The ability to live for centuries combined with the ability to propagate clonally account for the longevity of the species. The oldest confirmed redwood tree is at least 2,200 years old (Fritz, 1957) as determined by core samples.

Redwood Genetics

The genome of redwood is hexaploid ($2n=6x=66$) indicating six sets of eleven chromosomes (Hirayoshi and Nakamura, 1943; Saylor and Simons, 1970; Schlarbaum and Tsuchiya, 1984). In contrast most conifers are diploid, having 20-24 chromosomes. The modes of inheritance in *S. sempervirens* are not completely understood. Neale et al. (1989) concluded that chloroplast and mitochondrial DNA is paternally inherited. This is quite unique since most eukaryotes exhibit maternal inheritance of mitochondrial DNA. Aside from the study by Neale (1989), little is known about the segregation of

chromosomes in the species. In addition, nothing is known about how the hexaploid genome was formed. It was suggested by Li (1987, 1988) that the genome might have been formed by a hybridization event between two ancient species, which later gave rise to the genus *Sequoiadendron* and *Metasequoia*. Another possibility is that the genome underwent a duplication event, a common cause of polyploidy. Regardless of the reasons for the hexaploid condition of the genome, future population studies for conservation purposes would greatly benefit from a basic understanding of how the chromosomes segregate and where they came from.

Since redwood saplings are not sexually mature until they are 5-15 years old (Boe, 1968), most population studies involving *S. sempervirens* use stump sprouts or cloned shoots in regeneration experiments, making breeding programs that study sexual reproduction difficult to manage. These types of studies predominately use phenotypic data and visual assessment to identify clones. However, Rogers (2000) recently used allozymes to show that visual assessment is a poor predictor of clonal identity. Measuring genetic variance is a better method for the identification of individuals and population studies. Rogers (1994) was able to successfully identify 10 individuals within four self-pollinated families using 10 allozyme loci. Structural fairy rings are thought to have genotypic variation (Rogers, 2000) since it is often difficult to identify the root system from which a tree is growing. Additionally, Tufor (1973) suggests that a maturational gradient may exist in individuals that exhibit genetic variability along the length of the tree. Although this has not been examined at the molecular level, tools for measuring genetic variability at the DNA level need to be developed for the species to

make accurate assessments of population variance. A private timber company (Simpson Timber, Arcata, CA) is currently using amplified fragment length polymorphisms (AFLPs) to compare genetic variability with their phenotypic data. This technique rivals that of microsatellite analysis but is more expensive to perform. In addition, the scoring of gels can be quite complex often revealing too much variation. This may or may not be as informative as microsatellite marker data since the patterns of bands are a mixture of dominant and codominant markers (Robinson and Harris, 1999).

Many companies wish to improve the redwood in terms of growth rate, timber quality, drought tolerance, etc., whereas conservation groups aim to preserve genetic diversity. To date, there are no known DNA markers for coast redwood to use in population or conservation studies. Most molecular marker research focuses on the *Pinus* genus since it is the most economically important species in the conifer family.

Use of Microsatellite DNA (SSRs) as Molecular Markers

Microsatellite DNA is a term given to repetitive stretches of DNA that are composed of ≤ 6 repeated nucleotides or “repeat units” that reiterate for some length. These stretches of DNA serve no known function and can vary in the number of repeat units at a given locus (Lodish and Darnell, 2000). Microsatellite DNA regions are ideal as genetic markers for several reasons. They are cost effective to assay by using the polymerase chain reaction (PCR). In addition, microsatellite markers are typically codominant and polymorphic between individual trees (Scotti et al., 1999). They can show a high number of alleles in low frequencies, thus increasing their importance in the

grouping of individuals within a population (Robinson and Harris, 1999). Since microsatellite DNA is considered neutral, selection and environmental pressures do not influence their expression directly.

There are also some problems associated with microsatellite DNA as markers. Due to the repetitive nature of the genetic sequence, DNA polymerase tends to slip during strand replication in PCR, sometimes generating “stutter” bands detected during electrophoretic analysis. This can be controlled somewhat though positive controls and the correct DNA size standard. Another drawback is that heterozygosity calculations can lead to underestimations if null alleles are present since they will not be scored on a gel (Robinson and Harris, 1999). Yet despite these minor problems, microsatellite DNA markers are quite valuable tools for parentage analysis, pollen dispersal, genetic diversity, and forensics. Once developed their usefulness can be further shared with other researchers without the need for expensive equipment to perform the analysis.

The Importance of Molecular DNA Markers for Coast Redwood

The ratio of clonal versus sexually reproduced seedlings in a forest is often debated. It was suggested that in northern redwood forests, less than 10 percent of the trees were created from sprouting and that sprouting is more common in drier, warm climates where seedlings are less likely to propagate (Veirs, 1996). This debate could be analyzed more conclusively through the use of microsatellite markers in a parentage analysis, rather than looking at viable seed cones and studying germination rates in the forest.

By comparing the metabolic heat rates of individuals, Anekonda (1992) concluded that the largest component of redwood's genetic variation resides within stands of trees. They specifically found 11-21% genetic variation within stands, 7-13% among stands, and 4-6% among regions. Although this study was based on a physiologic trait, clone identification was made by comparison to morphological characters. Rather than infer genetic variation from physiologic traits or morphological characters, microsatellite markers could measure genetic variation directly and provide "DNA fingerprints" for tree identification.

Conservation efforts to preserve genetic diversity would be more effective if genotypes of trees could be determined. Certainly, conservation groups do not want to spend millions of dollars on conserving a stand of clonal redwood trees. Rather than strictly examine trees based on visual assessment, morphological characters, physiologic traits or allozymes, it is necessary that DNA markers be developed to complement the current scientific data on this species. The scope of this thesis is to isolate, identify, characterize and validate microsatellite loci that can be used as genetic markers for the California coast redwood.

MATERIALS AND METHODS

Tissue Collection and DNA Extraction

Needles were collected from "The Father of the Forest", an old growth redwood tree located on the Redwood Trail in Big Basin Redwoods State Park (Santa Cruz County, CA). Total genomic DNA was extracted using a modified CTAB method from Doyle and Doyle (1990). For a detailed protocol see Appendix A. Briefly, the procedure consisted of crushing redwood needles in liquid nitrogen and then suspending the tissue in hot CTAB (hexadecyltrimethylammonium bromide) extraction buffer. The suspension was extracted with chloroform:isoamyl alcohol (24:1), treated with RNase A, and then extracted again. The DNA was precipitated with isopropanol followed by an ethanol wash. The DNA pellet was then redissolved in TE buffer (10 mM Tris, 1 mM EDTA, pH 7.5). To calculate DNA yield, absorbance readings of a 1/10 dilution were taken at 260/280 nm. To estimate the purity and intactness of the genomic DNA, some of the sample was run on a 1% Seakem GTG agarose gel (BioWhittaker Molecular Applications, East Rutherford, NJ) against a 1 kbp DNA ladder (Invitrogen, Carlsbad, CA) for 1 hr at 10 V/cm in 1X TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.3) and then stained with 1 µg/ml ethidium bromide.

Insert Preparation

A genomic digest using *Sau3AI* was optimized to give the highest yield of 0.5-3.0 kbp size DNA fragments (see detailed protocol in Appendix B). Twenty-six micrograms

of genomic DNA were then digested at two different times bracketing the optimized condition to increase the randomness of the library. The two digests were run on a preparative wide format 0.5% Seakem GTG agarose gel with 1 µg/ml ethidium bromide for 2 hr at 10 V/cm in 0.5X TAE buffer against a 1 kbp DNA ladder. To increase the DNA concentration per area of agarose, a 3 x 12 cm gel slice was excised in the 0.5-3 kbp range using a sterile scalpel and placed into a new gel using a cutout of the same size. The orientation of the plug was retained and the interface between the plug and the gel was flushed with 0.5X TAE buffer to remove any air bubbles. The polarity was reversed from normal and the DNA fragments were allowed to run into the new gel in the opposite direction. After 30 min the smallest DNA fragments from the plug had migrated into the new gel and a smaller gel slice (about half the size) was excised. This second, smaller plug was then divided into 42 Qiaquick gel extraction columns (Qiagen, Valencia, CA) and processed according to the manufacturer's recommendation. The samples were eluted using 50 µl of Tris buffer (included in the kit) per column. All 42 samples were pooled together and concentrated using four Amicon YM-100 micro-concentrators (Millipore Corporation, Bedford, MA). The samples were then combined into a sterile microfuge tube and precipitated by adding one volume of isopropanol followed by one wash with two volumes of 70% ethanol. The ethanol was carefully decanted and the pellet was allowed to air-dry for 1 hr at room temperature. The pellet was then redissolved in 25 µl of a 10 mM Tris solution (pH 7.5) with gentle aspiration and dispensing. Some of the sample was run on a 1% agarose (Sigma, St. Louis, MO) mini-gel in 1X TAE buffer for 25 min at 10 V/cm against a 1 kbp DNA ladder to verify

the accuracy and enrichment of the excised size range. Absorbance readings of a 1/10 dilution were taken at 260 nm and 280 nm to estimate purity. In addition, a fluorescent quantitation method using PicoGreen® reagent (Molecular Probes, Eugene, OR) was used to compare against the accuracy of the absorbance measurements.

Vector Preparation

Five micrograms of pUC18 DNA (New England Biolabs, Beverly, MA) were digested using 8 U *Bam*HI per µg pUC18 DNA at 37°C for 1 hr. The digestion reaction was heat inactivated for 20 min at 80°C then briefly cooled down to 37°C. The vector was then treated with the addition of 2.5 µl of calf intestinal alkaline phosphatase (New England Biolabs) and incubated at 37°C for 1 hr. The tube was snap-cooled on ice for 5 min and 500 µl of 10 mM Tris (pH 7.5) were added. The treated vector was extracted with equilibrated phenol twice, subsequently followed by one chloroform:isoamyl alcohol (24:1) extraction. The vector DNA was precipitated by adding one volume of isopropanol and washed once with two volumes of 70% ethanol. The ethanol was carefully decanted and the pellet was allowed to air-dry for 1 hr. The pellet was redissolved in 25 µl of 10 mM Tris (pH 7.5) with gentle aspiration and dispensing. PicoGreen® quantitation reagent (Molecular Probes) was used to calculate the yield.

Ligation Reaction

Based on quantitation data (see Results), 1 µl of vector was combined with 2 µl of enriched insert in a microfuge tube containing 5 µl of ultra-pure sterile water. The

mixture was heated for 5 min at 65°C in a metal block, then snap-cooled on ice for 5 min. One microliter of 10X ligase buffer (supplied with the T4 DNA ligase) and 1 µl of T4 DNA ligase (New England Biolabs) were added to the ligation reaction. The tube was gently mixed, spun down and placed in a chilled water bath at 16°C overnight. After 18 hr, the tube was removed from the bath and the DNA was precipitated by adding 500 µl isopropanol. The DNA was washed once with 500 µl of 70% ethanol. The ethanol was carefully decanted and the pellet was allowed to air-dry for 30 min. The pellet was resuspended in 25 µl of 10 mM Tris buffer (pH 7.5).

Electroporation of Recombinant DNA

A BioRad Gene Pulser II (BioRad, Hercules, CA) was set according to the manufacturer's recommendation for Electromax™ electrocompetent DH10B cells (Invitrogen, Carlsbad, CA). Two microliters of the recombinant DNA were added to 100 µl of the electrocompetent DH10B cells and then subsequently divided into two 0.2 cm gap electroporation cuvettes (BioRad). After delivering the pulse for one cuvette, 1 ml of SOC medium (Invitrogen) was quickly added to the cuvette and gently flushed by aspirating and dispensing to rinse the cells off the sides of the metal contacts. Both samples were processed, combined into a 5 ml test tube, and incubated in a tube roller for 1 hr at 37°C. An equal volume of 50% glycerol was added to the tube and then gently mixed. Luria Bertani agar plates containing 0.1 mg/ml carbenicillin, 0.1 mM IPTG (isopropyl-β-D-1-thiogalactopyranoside) and 20 µg/ml X-GAL (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) were used to plate 5 µl, 25 µl and 100 µl respectively of

the glycerol cell mixture. The agar plates were incubated at 37°C for 14 hr and then moved to 4°C for 6 hr prior to determining the titer of the library. The glycerol cell stock was dispensed into 1 ml aliquots and stored at -80°C along with the insert, vector and recombinant DNA ligation.

Validation of Library

Upon establishing the titer of the library, 192 random clones were picked with sterile toothpicks into a pair of 96 deep-well blocks and grown up in 1 ml Luria Bertani medium per clone. The two deep-well blocks were covered with airpore tape (Qiagen) and grown at 37°C for 20 hr anchored in an orbital shaker set at 500 rpm. The plates were then processed using an alkaline extraction method on a vacuum manifold (R.E.A.L Prep, Qiagen). The purified clones were run on a 1% agarose gel (Sigma) in 1X TAE buffer for 1 hr at 10 V/cm against a 1 kbp DNA ladder. The purified clones were also digested to excise the inserts and confirm the range of fragment sizes that were cloned. Since the restriction site in the vector (*Bam*HI) was not retained in the cloning method used, *Eco*RI and *Hind*III were used according to the manufacturer's recommendations to digest the vector.

In addition to confirming the insert size, the clones were set up for fluorescent dideoxy cycle sequencing using Big-Dye™ Terminator v2.0 (Applied Biosystems, Foster City, CA) using universal primers to sequence in both directions (see detailed protocol in Appendix C). Cycling was performed using a GeneAmp® PCR system 9700 (Applied Biosystems). The sequencing products were purified using Sephadex G-50 resin

(Amersham Biosciences, Piscataway, NJ) hydrated in a 96-well HV membrane plate (Millipore Corporation). The purified sequencing reactions were then sequenced on an ABI PRISM® 3700 DNA Analyzer (Applied Biosystems) using POP6 polymer. The sequences were then BLASTed in Genbank and assembled against each other in Sequencher (Gene Codes, Ann Arbor, MI) to look for redundancy of the library.

Isolation of Recombinant Clones in a High-Throughput Format

Clones were plated on Luria Bertani agar plates containing 0.1 mg/ml carbenicillin, 0.1 mM IPTG and 20 µg/ml X-GAL in square agar plates using autoclaved glass beads to spread the cell stock. The agar plates were incubated at 37°C for 14 hr then moved to 4°C for 8 hr. The plates were then placed on a custom robotic picker that was developed and created at Stanford Genome Technology Center (SGTC, Palo Alto, CA) under the leadership of Dr. Ronald Davis. This instrument was designed and used to isolate single colonies from square agar plates into a 96-well plate output format. The instrument is able to isolate 1152 clones in a 1 hr period. For each week of screening, clones were isolated into 300 µl Luria Bertani medium containing 0.1 mg/ml carbenicillin across a total of forty-eight 96-well plates that were labeled numerically. To grow the clones in this high density format, another custom instrument developed at SGTC was used. This multi-plate shaking device commercially available under the name HiGro® (GeneMachines, San Carlos, CA) has four chambers that each hold a “cassette” which consists of twelve 96-well plates. Used at full capacity, this shaking device is able to grow forty-eight 96-well plates at once. After 18 hr of shaking, the 48 growth plates

were reduced to 16 by stamping clones from three plates into new culture plates containing 300 μ l Luria Bertani medium with 0.1 mg/ml carbenicillin. This new consolidated set of sixteen plates (containing three clones per well) were labeled alphabetically and grown in the shaker for 18 hr and then used directly in the high density screening. To summarize, clones were isolated on the picker, grown in batches in the shaker, then pooled together and grown again in the shaker. All plates were then stored at 4°C prior to screening on nylon.

Stamping and Fixing of Clones on Nylon in a High-Throughput Format

Two 12" x 18" autoclaved Pyrex baking dishes were prepared as agar plates, each being filled with 500 ml of Luria Bertani agar containing 0.1 mg/ml carbenicillin, 0.1 mM IPTG and 20 μ g/ml X-GAL. Once poured, the dishes were allowed to sit for 24 hr at room temperature prior to use. Clones taken from the consolidated set of sixteen 96-well plates were stamped on Hybond-N+ positively charged nylon (Amersham Biosciences) using 96-pin disposable plate replicators (Incyte Genomics, Palo Alto, CA). The nylon membrane was then carefully laid corner first onto the agar in the baking dishes using autoclaved forceps. Great care was taken to ensure that no air bubbles were trapped between the nylon and the agar. The baking dishes were then covered and incubated inverted at 37°C for 20 hr.

After growing the colonies on nylon, the baking dishes were allowed to sit at room temperature for 3-4 hr. The colonies were then lysed and fixed to the nylon (see detailed protocol in Appendix D). In short, three trays were each set up with a sheet of

blotting paper that was saturated with a different solution. The first tray contained a 10% SDS (sodium dodecyl sulfate) solution, the second contained an alkaline cell lysing solution, and the third contained a neutralizing solution. Using sterile forceps, the membrane was carefully laid on the saturated blotting paper for a set amount of time then transferred to the next saturated paper very carefully so the colonies were not disturbed. Once the nylon was laid through each subsequent solution, the membrane was then submerged in a tray filled with 2X SSPE (0.3 M NaCl, 0.02 M NaH₂PO₄, 2 mM EDTA, pH 7.4) and gently stirred using an orbital shaker. The nylon was finally laid on a new sheet of blotting paper that was saturated with 2X SSPE and exposed to 1200 µJ of UV light in a Stratalinker® 2400 (Stratagene, La Jolla, CA). To remove additional cellular debris, the membrane was immersed in solution of 100 µg/ml Proteinase K in 2X SSC (standard saline citrate) and heated with gentle rocking for 1 hr at 55°C. The Proteinase K was then inactivated by immersing the membrane in PMSF (phenylmethylsulfonyl fluoride, 1 mM in 2X SSC). If additional cellular material was still present, the membrane was sandwiched between two 2X SSPE saturated blotting papers and a round object was rolled over the top layer with light pressure. The top layer of the sandwich was then peeled away, physically pulling up any cellular debris off the membrane. The membrane was then thoroughly washed in 2X SSPE prior to hybridization.

Hybridization of Microsatellite Probes to Target DNA

The sequence reads from the 192 random genomic library clones were screened to assess the repetitive nature of the genome. In addition, the scientific literature was

searched and used as a basis for probe selection in conifers (Schmidt, et al., 1999). Noncomplementary di-, tri-, and tetranucleotide probes (26-44 bases in length) were designed (see Results) and purchased (Operon Technologies, Alameda, CA) with biotin attached to the 5' end to allow for the use of chemiluminescent detection. All hybridizations were performed in a Hybaid tube roller hybridization oven (ThermoHybaid, Franklin, MA). For a detailed protocol see Appendix E. Briefly, the membrane was prehybridized for 1 hr at the optimized hybridization temperature. The hybridization buffer (1 mM EDTA, 7% SDS, 0.25 M NaH_2PO_4) was replaced and the probe was then added at a concentration of 1-5 pmol/ml of each probe. The hybridization was allowed to proceed overnight for at least 18 hr. The next day the membrane was put through a series of SSC (standard saline citrate) washes with increasing stringency and taken directly through the detection process.

Chemiluminescent Detection of Hybridized Microsatellite Probes

Any probe that was bound to the membrane was detected using Phototope® Star Chemiluminescent Detection Kit (New England Biolabs). For a detailed protocol see Appendix F. In short, the nylon membrane was exposed to streptavidin, which specifically bound any biotinylated probe that was hybridized to the target DNA. Next, a biotinylated alkaline phosphatase was added which bound to the streptavidin, forming a conjugate between the alkaline phosphatase and the probe. In the final step, a CDP-Star substrate (phenylphosphate substituted 1, 2 dioxetane) was added, making the alkaline phosphatase catalyze the removal of phosphate from the substrate. This yielded a

moderately stable intermediate that decayed, emitting light at 461 nm. The emission of light from the membrane was detected using BioMax Light film (Eastman Kodak Company, New Haven, CT) and developed using GBX developer and fixer (see Appendix G for processing procedure).

Isolation of Positive Clones and Sequencing of Recombinant DNA

Spots that gave a positive signal were traced back to the original set of forty-eight 96-well plates. For every positive chosen on a film, there were three corresponding clones in the original set to select. Clones were identified by their plate number and corresponding well position, then grown up in test tubes containing 5 ml Luria Bertani medium containing 0.1 mg/ml carbenicillin. The test tubes were then placed in a tube roller for 18 hr at 37°C. Plasmid DNA was isolated using a Qiaprep Spin Miniprep Kit (Qiagen) and glycerol stocks of the bacterial clones were made. The purified clones were sequenced in both directions using Big-Dye™ Terminator v 2.0 (Applied Biosystems) and universal primers as described earlier. The DNA sequences for each clone in both directions were then assembled using Sequencher (Gene Codes) and examined manually for microsatellites. Once a large enough microsatellite was found (>20 repeat units), it was subsequently used as a positive control for future hybridizations.

PCR Optimization of Microsatellite Loci

Positive clones were categorized by the relative size and complexity of the microsatellite they contained. The online version of Primer3 (Rozen and Skaletsky,

2000) was used to pick primers that gave a product size in the range of 100-300 bp. For some of the larger microsatellites, the smallest possible product size was made. All primers were synthesized at SGTC using a custom automated oligo synthesizer (AMOS), which is now commercially available under the name PolyPlex® (GeneMachines). Positive clones were amplified by polymerase chain reaction (PCR) using AmpliTaq Gold® DNA Polymerase (Applied Biosystems).

A series of three optimization experiments were then performed (data not shown). The first compared the effect of primer annealing temperature to product size and specificity. The second experiment explored the effect of magnesium concentration, and the third experiment involved testing several sets of primers for a few clones to see if the actual primer design affected product specificity. The resulting optimized PCR protocol is described in Appendix H. PCR reactions were performed and products electrophoresed in 1% high grade analytical agarose (BioRad) in 0.5X TBE (Tris-borate EDTA) buffer for approximately 3 hr at 5 V/cm against a 1 kbp DNA ladder. The Father of the Forest DNA was used as a positive control to ensure that the microsatellites isolated were real and not artificially created during the library construction. If the amplification products were in the correct size range and showed a clear banding pattern, they were brought to the Conservation Genetics Laboratory at San Jose State University where Dr. Chris Brinegar ran the samples on non-denaturing 6.5% polyacrylamide gels with 1X Spreadex NAB polymer (Elchrom Scientific AG, Switzerland) in 1X TAE buffer for 3 hr at 16 V/cm against a special marker (M3 marker, Elchrom). Only through

running such a high resolution polyacrylamide gel with 1-2 bp resolution could the usefulness of a given marker and primer set be determined.

Validation and Sequencing of Amplified Products

After optimization of the PCR parameters, the usefulness of each microsatellite was determined and a final list of clones was distilled from all the possible positives. This final list included only those clones that had a working primer set at a universal condition and gave a decent banding pattern. The products from the Father of the Forest amplifications were run on a polyacrylamide gel (as previously described) and every band for each clone was cut out of the gel with a sterile razor blade and placed in a 1.5 ml microfuge tube. Twenty-five microliters of 10 mM Tris buffer was added to each gel slice and the gel was macerated with a sterile pipette tip. The samples were heated for 5 min at 95°C to elute the DNA from the gel. After centrifugation to remove the gel debris, 5 µl of each sample was used in a PCR reaction (as previously described) to amplify the target. Subsequently, 2 µl of the amplified product was used in a sequencing reaction (as previously described) except the primers unique to each clone were used to sequence in both directions. The resulting sequence data was analyzed using Sequencher (Gene Codes) by assembling the reads to the original clone's sequence.

Validation of Polymorphism for Seven Microsatellite Loci

Once the amplification products of each locus were confirmed as being the correct sequence of the intended clone, the best markers were then tested for polymorphisms

using several different individual trees. DNA was collected and extracted by Dr. Chris Brinegar and Debbie Kinsey from twelve random individual trees located along different tributaries in the Waddell Creek area of Big Basin Redwoods State Park (Santa Cruz County, CA). PCR was performed on the best 7 loci from the final list of 20 candidates. The samples were run on a polyacrylamide gel at SJSU, polymorphic bands were cut out of the gel and sequenced as described earlier. The resulting sequence reads were then analyzed for each microsatellite locus using Sequencher by comparing against the clone's sequence.

Practical Application of Microsatellite Markers

The best marker (18D7-3) was used in a simple study that was presented at the Big Basin Centennial Celebration on September 14th-15th 2002. The study was to determine if several small redwoods growing at the edge of a meadow near the park headquarters were related to stands of larger trees that flanked them. The original work for this study was presented as a Master's thesis by Ingrid Udranszky in the Spring of 2002 using a putative microsatellite locus characterized by Monica Chadha. Briefly, samples were taken from six trees thought to be clones of one another belonging to a grove on the southern side of the meadow. Samples were also taken from another six trees thought to be clones of one another belonging to a grove on the northern end of the meadow. Samples were then collected from the meadow trees in between. The DNA samples were set up for PCR as described earlier using the 18D7-3 locus and run on a polyacrylamide gel against a 25 bp step DNA ladder (Promega, Madison, WI).

RESULTS

Tissue Collection and DNA Extraction

The genomic DNA preparation from redwood needles of the Father of the Forest tree yielded DNA of relatively high molecular weight when compared to the DNA ladder (Figure 1). The majority of the DNA was >12 kbp with little degradation and appeared to be free of RNA.

Figure 1. Genomic DNA extraction from the Father of the Forest. Lanes 2 and 4 were loaded with 2 μ l and 4 μ l respectively of the genomic DNA preparation. Lanes 1 and 5 are a 1 kbp DNA ladder. Lane 3 was not loaded.

Insert DNA Preparation

To generate DNA fragments of known size reproducibly, the genomic DNA was digested for 30 min using decreasing concentrations of *Sau3*AI with a constant amount of genomic DNA (1 µg each) across 10 samples (Figure 2).

Figure 2. *Sau3*AI restriction digest optimization on redwood genomic DNA. Lanes 2 -11 represents decreasing restriction enzyme concentration (See Appendix B for exact concentrations). Lanes 1 and 12 are a 1 kbp DNA ladder.

The optimal enzyme concentration for the desired cloning size (0.5 - 3 kbp) using *Sau3*AI on redwood genomic DNA was no less than 1 unit of enzyme activity per microgram of genomic DNA (Figure 2, lane 2). Performing the digestion with a higher enzyme concentration did not increase the cutting rate (results not shown); therefore, the

time of digestion was used to control fragment size at the maximal cutting ratio of 1 unit of enzyme activity per microgram of DNA. Digestion was performed at 30 min and 60 min and then loaded on a wide format preparative agarose gel containing 0.5 μ g ethidium bromide per ml (Figures 3 and 4, respectively).

Figure 3. Insert DNA preparative agarose gel from 30 min *Sau3AI* digest. Lane 2 was loaded with approximately 13 μ g of genomic DNA. Lanes 1 and 3 are a 1 kbp DNA ladder.

Figure 4. Insert DNA preparative agarose gel from 60 min *Sau3AI* digest. Lane 2 was loaded with approximately 13 μ g of genomic DNA. Lanes 1 and 3 are a 1 kbp DNA ladder.

A plug was cut from both gels in the 0.5 – 3.0 kbp size range (Figures 5 and 6 respectively) and carefully moved aside retaining the original orientation.

Figure 5. Excision of 0.5 - 3.0 kbp DNA fragments from agarose gel (30 min digest). Lane 2 represents the remaining DNA fragments left in the gel. Lanes 1 and 3 are a 1 kbp DNA ladder.

Figure 6. Excision of 0.5 - 3.0 kbp DNA fragments from agarose gel (60 min digest). Lane 2 represents the remaining DNA fragments left in the gel. Lanes 1 and 3 are a 1 kbp ladder.

Both agarose gel plugs were placed into a new gel with the original orientation and the gel interface was flushed with running buffer to remove air bubbles (Figure 7).

Figure 7. Both excised gel plugs inserted into a new agarose gel prior to electrophoresis. The 30 min excised plug is on the left and the 60 min excised plug is on the right.

The polarity on the gel box was reversed from normal and the gel was run in the opposite direction. Since bromophenol blue migrates in agarose at about 500 bp, the gel was stabbed at the start of the smallest DNA fragments (bottom of the plug) with some loading buffer to estimate the migration of the DNA fragments (not shown). It was critical to not disturb the gel until all the fragments were electrophoresed into the new gel, otherwise the fragments would be lost in the running buffer. Conversely, running the gel too long would spread the fragments out in the opposite direction, defeating the purpose of concentrating the fragments (Figure 8).

Figure 8. Concentrated DNA fragments in agarose gel after reversing polarity. The DNA fragments from both plugs have fully electrophoresed into the new gel.

Once the fragments fully electrophoresed into the second gel, two new plugs containing the concentrated 0.5 - 3.0 kbp DNA fragments were excised and placed on the UV light box. In addition, the original plugs were removed to examine the reduction of agarose achieved by concentrating the DNA fragments in this manner (Figure 9).

Figure 9. Excision of concentrated DNA fragments from preparative agarose gel. The two plugs containing the concentrated 0.5 – 3.0 kbp size fragments are resting on a UV light source (bottom). The original plugs were also removed for comparative purposes (top).

By running the excised fragments into a new gel and concentrating them, there was roughly a 50% reduction in the amount of agarose that contained the DNA fragments. This represented about a two-fold increase in the amount of DNA per area of agarose, greatly reducing the amount of Qiaquick (Qiagen) gel extraction columns that needed to be processed. After extracting the DNA fragments from the agarose plugs (see

methods for more detail), the eluted DNA was combined and a gel was run to confirm the enriched fragment sizes (Figure 10).

Figure 10. Verification of enriched DNA fragments in agarose gel. Lane 2 was loaded with 1 μ l of uncut genomic DNA. Lanes 3 & 4 were loaded with 5 μ l and 1 μ l respectively of the enriched insert. Lane 5 is a negative control (micro-concentrator supernatant). Lanes 1 & 6 are a 1 kbp DNA ladder.

The enriched insert appeared very faint on the gel due to the amount of area covered by the excised range. Only 5 μ l was run on a gel so the rest could be used for the ligation reaction. Although faint, the DNA fragments appeared to be in the correct size range that was originally excised (0.5 – 3.0 kbp).

Insert and Vector Quantitation using PicoGreen®

PicoGreen® (Molecular Probes) was used to measure the DNA concentration of both the insert and the vector. A standard curve was generated using known concentrations of pUC18 plasmid DNA. The assay was performed at 1/10 the manufacturer's recommended volume and the result was extrapolated off the standard curve generated for pUC18 DNA (Figure 11).

Figure 11. Quantitation of insert and vector DNA for ligation. Standard curve for pUC18 plasmid DNA showing the extrapolated concentration values for the insert and vector DNA in ng/μl.

Based on the Picogreen assay, the insert DNA concentration was 65 ng/μl and the vector DNA concentration was 185 ng/μl. The DNA ligation was then set up (see methods for more detail) using 2 μl insert DNA and 1 μl vector DNA. For an average

insert size of 1,750 bp and a vector size of 2,686 bp this represented approximately a 1:1 molar ratio of insert to vector.

Determination of Library Titer

Three different volumes (5 µl, 25 µl and 100 µl) of the electroporated glycerol cell stock were plated on agar dishes containing carbenicillin and grown overnight to calculate the library titer (Table 1).

Table 1. Titer of recombinant library glycerol cell stock. The number of blue and white colonies obtained from plating three different volumes of a glycerol cell stock.

Volume of cell stock	5 µl	25 µl	100 µl
White colonies	37	106	>225
Blue colonies	16	50	>100

Based on the values for 5 µl of the glycerol cell stock, the titer of a single ligation reaction was calculated as:

$$\frac{37 \text{ colonies}}{5 \text{ µl of cell stock}} \times \frac{4000 \text{ µl cell stock}}{2 \text{ µl ligation DNA}} \times 25 \text{ µl ligation DNA} = 370,000 \text{ clones}$$

After DNA quantitation, approximately 15 µl of the enriched insert DNA remained for future ligations. Since this is the limiting factor in the ligation for which 2 µl of insert DNA is used, approximately seven more ligations can be performed. This equates to at least 2,590,000 more clones that can be used for future screening.

Validation of library

Random clones were purified using a rapid extraction alkaline lysis method (R.E.A.L. prep, Qiagen). All 192 purified plasmid clones were digested and run on a gel to estimate the size range of the inserts in the library. Out of all 192 clones digested, only 16 are shown in Figure 12.

Figure 12. Purified clones digested with *Hind*III and *Eco*RI in agarose gel. Lanes 2-17 were loaded with 8 μ l of digestion products. Lanes 1 and 18 are a 1 kbp DNA ladder.

Since the *Bam*HI restriction site is not retained in this cloning method, *Hind*III and *Eco*RI were used to cut out the inserts. In some cases the inserts were cut into multiple bands due to a *Hind*III or *Eco*RI site contained within the insert itself. The

linearized pUC18 vector ran at the appropriate size (2.686 kbp) for all clones confirming the plasmids were recombinant. Some clones were only partially digested therefore supercoiled and concatenated forms can be seen. The size range of the inserts was 298-1,200 bp with some bands only representing part of the complete insert.

After sequencing all 192 clones in both directions across the inserted sequence, the reads were aligned against one another using Sequencher (Gene Codes). A summary of the assembly results are shown in Table 2.

Table 2. Summary of sequence alignments from 192 random clones.

	Reads	Contigs
2 reads per contig	216	108
3 reads per contig	6	2
4 reads per contig	4	1
Singlets (Poor Quality)	151	0
Failed reactions	7	0
Total sequence reads	384	

Out of 384 sequence reads, 151 did not have any match and 7 failed completely. This can be attributed to poor sequencing reaction setup on one of the 96-well template plates. Of the reads that gave sufficient sequence, 216 formed contigs that assembled with the appropriate reverse complement of the same clone. There were 2 contigs that contained three reads each, but upon further scrutiny it was concluded that the overlap

was not true sequence homology but rather due to poor sequence data. Poorly called bases will assemble to each other if they are both basecalled "N". The 108 contigs that were composed of only two reads each were screened for vector sequence and then graphed based upon their size (Figure 13). This was more accurate in estimating the average size of the inserts in the library because it reflects contiguous sequence data across the cloning site in the vector. Based on these results the mean insert size for the library was determined to be 702 bp in length.

Figure 13. Scatter plot of insert sizes based on sequence data. 108 contigs assembled with 2 reads each. Clone sizes ranged from 400-1200 bp, with an average size of 702 bp.

Isolation of Recombinant Clones in a High-Throughput Format

Clones were isolated using a custom colony picker developed by Stanford Genome Technology Center (Figure 14). The instrument identifies colonies plated on square agar dishes and inoculates a 96-well titer plate that is stored in a high throughput cassette (Figure 15). After inoculation of all 48 plates, each cassette was then placed into a HiGro shaker, also developed by Stanford Genome Technology Center (Figure 16).

Figure 14. Photograph of custom colony picker. Square colony dishes are imaged under a lamp (left side of photo) then a rotating turret (center of photo) with pressure-actuated needles continuously picks colonies and simultaneously inoculate individual wells of a 96-well plate with the aid of an output server arm. On the backside of the turret (not shown in photo), needles are flushed with water in a wash station then heated in a heating element to maintain sterility.

Figure 15. Photograph of high-throughput cassette for HiGro shaker. This is the output format generated by the picker. Each cassette can hold up to twelve 96-well plates. An empty cassette is shown on the left and a full cassette is shown on the right. Also shown is the top view of an inoculated growth plate containing 300 μ l LB medium.

Figure 16. Photograph of HiGro high-capacity incubating shaker. The instrument is composed of four chambers that each hold one metal cassette (as seen in Figure 15). Three of the chambers are shown with the lids attached and one is removed to show how a cassette is placed into the device. Once fully secured with a lid, each chamber independently moves in an orbital motion and a thermostat controls the temperature (blue wires on the left side of the chamber).

The 48 plates were pooled into 16 plates by stamping with a disposable pintle (Incyte Genomics). The pooled set of culture plates were grown, then stamped in a staggered fashion on Hybond-N+ (Amersham Biosciences) nylon using a disposable pintle. Eight plates were stamped onto each nylon (an illustration is shown in Figure 17).

Figure 17. Illustration of sample tracking procedure for hybridization screening. Clones were stamped on nylon (not shown actual size) in a high-density array. 96-well plate orientation was maintained by clipping the corner of the nylon to indicate the A1 well position.

The plate letter indicated which pooled plate set the positive came from and the plate number corresponded to the individually isolated clone plate. Clones giving a positive signal were tracked back to the original set of 48 plates and all three clones for the pool were grown separately in test tubes.

Hybridization to Target DNA and Positive Detection

The nylon was then processed as described in Materials and Methods. Nine repetitive probes were designed with similar melting temperatures (63.8-71.1°C) and 5' biotin attachment (Table 3). The probes were pooled together and used at the same time for each screening.

Table 3. Summary of probes used in recombinant clone screening.

Probe	Repeat Unit	T _m (°C)
1	(CT)13	66.1
2	(CA)13	64.7
3	(CAA)12	64.7
4	(GAA)12	64.7
5	(CAG)8	71.1
6	(AAGA)11	63.8
7	(GATA)11	63.8
8	(GGAT)11	69.6
9	(ACCT)8	69.6

Hybridization was performed at 55°C and stringency was controlled by increasing the wash temperature until an optimal balance between real and false positives was

achieved. Nineteen films were processed, but only two are shown for comparison (Figures 18 and 19). The processed film in Figure 18 shows a positive signal that was later determined by sequencing to be a genuine repeat.

Figure 18. Scanned image of BioMax film showing a genuine positive signal. The target shown indicating a positive signal was later determined by DNA sequencing to be a genuine repeat (not shown actual size).

Figure 19. Scanned image of BioMax film showing false positive signals. Circles indicate scored positive signals that were later determined by DNA sequencing to be false positives (not shown actual size).

In contrast, the processed film in Figure 19 shows several positive signals that appear to be genuine, but upon sequencing were determined to be false positives. It was discovered that trace amounts of cellular debris remaining after the colony lysis step

interfered with the detection process by either binding the probes or the streptavidin nonspecifically. In addition, the repetitive nature of the genome complicated the hybridization procedure by duplexing the probes to clones that were only partial sequence matches.

Sequencing Results

Approximately 30,000 clones were screened and 465 positives were counted and sequenced. Of those 465 positives, only 1 in 3 were expected to be true positives since the samples were pooled together. After close scrutiny of each sequence read, 37 clones were identified as carrying a microsatellite (≥ 5 repeat units). The frequency of microsatellite containing clones in this library was therefore roughly 1 in 1000 clones using the selected probes. For the complete flanking sequence data for the 37 microsatellites identified see Appendix H. Clones were grouped based on the length of the repeat unit they contained and are summarized in Tables 4, 5, 6 and 7. Of the two longest repeat units found in Table 6, one of them is actually a minisatellite since it has a repeat unit of 16 nucleotides (clone 2D5).

Some microsatellites contained a high number of repeat units. Due to size limitations in the detection method that was subsequently used, any microsatellite (and the one minisatellite) that was too large to make a PCR product <300 bp in length were catalogued but not aggressively pursued in the primer optimization.

Table 4. Summary of clones containing perfect dinucleotide motifs. Each motif repeats 5 or more times. Reiterations under each clone include additional repeats found elsewhere in the clone. Contigs with gaps have minimum insert sizes indicated by ">".

Clone Name	Contig Status	Repeat Unit	Insert Size
7F11	Gap	(CT)53	>1012
20C5	Closed	(AT)26	530
15A4	Gap	(GT)>21	>1362
24E6	Closed	(AT)16	531
42B11	Closed	(CA)15	419
20E5	Closed	(CA)13	563
37G5	Closed	(CT)14	937
20B3	Closed	(CA)11	608
24H9	Closed	(CA)10	1348
11D11	Gap	(GA)>28 (CA)8	>625
39G9	Gap	(AC)16	>783
10H12	Closed	(TA)8 (TA)5	642
10D3	Gap	(GA)24 (GT)23	1173
8E8	Closed	(CA)12 (AT)4 GTATCT (AT)5	798
9A3	Closed	(TG)24 (GT)9 GC (GT)6 GC (GT)8 (GT)16 TTATGCTAT (GT)10 (TGTGTGT)4	1157

Table 5. Summary of clones containing perfect trinucleotide motifs. Each motif repeats 5 or more times. Reiterations under each clone include additional repeats found elsewhere in the clone. Contigs with gaps have minimum insert sizes indicated by ">". The letter "N" represents any of the four nucleotides.

Clone Name	Contig Status	Repeat Unit	Insert Size
31E2	Closed	(GAA)11	587
18D7-3	Closed	(CTT)13	1011
23E12	Closed	(CAG)7	686
2B2	Closed	(TGC)7	1072
21E3	Gap	(TGN)18	>1008
5H3	Closed	(TTG)9 (TTG)3 AGC (TTG)5 (GTT)3 GGGCAT (GTT)4 (GTT)3 (GTGTT) GCTAATGCTT (GTT)6	545
13G7	Closed	(GCT)6 (GAG)4 (GTGAT)5 (CTG)7 (CCTGCTG)2	964

Table 6. Summary of clones containing the largest perfect motifs. Each motif repeats 5 or more times. Clone 2D5 is a minisatellite.

Clone Name	Contig Status	Repeat Unit	Insert Size
21E5	Closed	(CTTA)8	1383
2D5	Closed	(GCAAAATCGAGCCGCAACT)16	1157

Table 7. Summary of clones containing compound, imperfect or complex motifs. Each motif repeats 5 or more times. Reiterations under each clone include additional repeats found elsewhere in the clone. Contigs with gaps have minimum insert sizes indicated by ">." The letter "N" represents any of the four nucleotides.

Clone Name	Contig Status	Repeat Unit	Insert Size
23G6	Gap	(AT)5 (GT)55	882
18D7	Closed	(AT)27(GT)59	253
23F9	Closed	(CA)27 (TA)6	574
28E5	Closed	(AT)7 (GT)19	672
11G8	Gap	(AT)3 (GT)11	>1500
3D4	Gap	(CT)4 (AT)8	>682
38H4	Closed	(AT)16 GTATAT (GT)32	820
4E2	Closed	(CT)20 GT (CT)13	808
27H5	Closed	(TA)5 (GATA)5 (TA)2	539
3A7	Closed	(GT)10 (AT)5 T (AT)3	1134
15B3	Gap	(AAAAGG)9 AAA (GAA)2 ATAA (GAA)48	>1311
4A8	Closed	(CT)2 (CTTT)2 (CT)9 CC (CT)5	936
1B5	Closed	(CATA)5 (CA)38 (TA)3 (GT)3 AT (GT)5	700

PCR Optimization of Microsatellite Loci

PCR amplification of the microsatellite loci proved to be a challenging task. The repetitive nature of the genome made the primers less specific for the intended target. In addition, it was difficult to design unique primers to amplify PCR products in the size range of optimal separation for the high resolution polyacrylamide gels (100-300 bp). The results of optimization experiments (data not shown) determined that a very stringent PCR was required to achieve the appropriate specificity. A temperature gradient experiment revealed that a higher melting temperature favored specificity, but a majority of clones would not amplify. When the temperature was reduced by a 3°C increment, several clones would amplify nonspecifically. In the magnesium gradient experiment, a lower magnesium concentration (1.5 mM) improved the specificity, with little effect on the whether or not the clone amplified. Therefore, the primers needed to have a high melting temperature (around 60°C) and the PCR conditions needed to be very stringent. This might have been expected given the high number of false positives during the clone hybridization screening.

In the multiple primer set experiment, typically 6 out of 8 primer sets for the same clone amplified with similar banding patterns. Either the clone had flanking sequence that made PCR conducive for the size restriction (100-300 bp) or it did not. Most clones that had products giving poor banding patterns were recalcitrant to improvement by increasing primer melting point or lower magnesium concentration. The use of a hot start DNA polymerase, AmpliTaq Gold® (Applied Biosystems) helped to maintain specificity and allowed for full denaturation of the genome in the initial cycle of the PCR. For the

sake of an easy and routine assay, one set of PCR conditions was chosen for all the clones. Although this is not optimal for all the clones, it does make the assay more amenable to routine screening (Table 8). For a detailed protocol of the PCR reaction setup see Appendix I.

Table 8. Optimized universal PCR cycling condition.

<u>Temp. (°C)</u>	<u>Time</u>	
95	10 min	
95	30 sec	} Cycle 30 times
60	1 min	
72	2 min	
4	Hold	

Out of a total of 37 microsatellites identified, 11 did not have suitable flanking sequence to select primers and 8 had nonspecific amplification across several primers tested. The primer selection process, PCR optimization, and analysis by electrophoresis (Figure 20) resulted in a distilled set of 20 possible microsatellites for routine use (Table 9). This list was prioritized based on several factors: (1) the correct amplification size, (2) the specificity of the amplification products, (3) the size range of the products, (4) the microsatellite they contained, and (5) the number of different primers that were available.

Table 9. Prioritized list of 20 microsatellite candidates using a universal PCR condition.

Priority	Clone	Repeat	Expected¹ Product Size (bp)	Observed² Product Size(s) (bp)
1	18D7-3	(CTT)13	150	150 153
2	21E5	(CTTA)8	118	118
3	20E5	(CA)13	197	197 195
4	31E2	(GAA)11	99	99
5	3D4	(CT)4(AT)8	147	147 153
6	11G8	(AT)3(GT)11	131	131 123 120
7	20B3	(CA)11	201	201 192
8	23F9	(CA)27(TA)6	309	309 285
9	10H12	(TA)5(N)37(TA)8	265	265

¹ Based on the sequence of the original cloned insert.

² Estimated sizes of alleles from the amplification of Father of the Forest (Figure 20).

Table 9 (continued). Prioritized list of 20 microsatellite candidates using a universal PCR condition. Parenthesis indicate estimates of product size.

Priority	Clone	Repeat	Expected ¹ Product Size (bp)	Observed ² Product Size(s) (bp)
10	28E5	(AT)7(GT)19	205	205 218
11	37G5	(CT)14	124	124 126
12	27H5	(TA)5(GATA)5(TA)2	240	240
13	4A8	CTTT(CT)2(CTTT)2(CT)9CC(CT)5	319	319 308
14	38H4	(AT)16GTATAT(GT)32	234	234
15	5H3	(TTG)3AGC(TTG)5	536	(>500)
16	11D11	(GA)>28	gap in clone	(>500)
17	3A7	(GT)10(AT)5T(AT)3	575	(>500) (450)
18	9A3	(GT)16(N)9(GT)10(TGTGTGT)4	203	380
19	24E6	(AT)16	134	155 145
20	4E2	(CT)20 GT(CT)13	181	124

¹ Based on the sequence of the original cloned insert.

² Estimated sizes of alleles from the amplification of Father of the Forest (Figure 20).

Validation of Microsatellite Allele Sequence

The Father of the Forest was amplified with all 20 markers and electrophoresed as mentioned earlier. Bands were cut out of the gel as shown in Figure 20 and sequenced as previously described.

Figure 20. PCR products from 20 markers using a universal condition. Numbers next to the bands indicate how many and which bands were cut from the gel for sequencing. Lanes 10, 12 and 14 were not cut from the gel since they became too faint to see upon prolonged exposure to UV light. The M3 size standard is indicated at the top of the gel.

The sequence reads obtained from each amplified product were aligned against the previous sequence data from each clone. Each microsatellite was then checked for accuracy of the amplified product and length disagreement between the fragments of different sizes. The resulting summary can be seen in Table 10.

Table 10. Summary of sequenced alleles from amplified products. All products were obtained from the Father of the Forest DNA. The letter "N" indicates that the repeat was found in the read but the sequence was too poor to accurately estimate length.

Lane #	Clone	Expected Repeat Motif ¹	Observed Repeat Motif ²	
			Band #	Sequence of Product
1	18D7-3	(CTT)13	1	(CTT)13
			2	(CTT)12
2	21E5	(CTTA)8	1	(CTTA)8
3	20E5	(CA)13	1	(CA)9(TA)4
			2	(CA)8(TA)4
4	31E2	(GAA)11	1	(GAA)N
5	3D4	(CT)4(AT)8	1	(CT)4(AT)8
			2	(CT)4(AT)8
6	11G8	(AT)3(GT)11	1	(AT)5(GT)12
			2	(AT)5(GT)10
			3	(AT)3(GT)11
7	20B3	(CA)11	1	(CA)11
			2	(CA)7
8	23F9	(CA)27(TA)6	1	(CA)N(TA)7
			2	(CA)N(TA)7

¹ Based on the sequence of the original cloned insert.

² Based on the sequence data obtained from the amplification products cut out of the gel.

Table 10 (continued). Summary of sequenced alleles from amplified products. All products were obtained from the Father of the Forest DNA. "N" indicates that the repeat was found in the read but the sequence was too poor to accurately estimate length. Amplification products from 28E5, 27H5, and 38H4 were not cut from the gel.

Lane #	Clone	Expected Repeat Motif ¹	Observed Repeat Motif ²	
			Band	Sequence of Product
9	10H12	(TA)5(N)37(TA)8	1	(TA)5(N)37(TA)8
10	28E5	(AT)7 (GT)19	-	Not sequenced
11	37G5	(CT)14	1	(CT)N
			2	No alignment
12	27H5	(TA)5(GATA)5(TA)2	-	Not sequenced
13	4A8	CTTT(CT)2(CTTT)2(CT)9CC(CT)5	1	(CT)N
			2	(CT)N
14	38H4	(AT)16GTATAT(GT)32	-	Not sequenced
15	5H3	(TTG)3AGC(TTG)5	1	(TTG)N
16	11D11	(GA)>28	1	(GA)N
17	3A7	(GT)16 TTATGCTAT (GT)10 (TGTGTGT)4	1	(GT)N
			2	(GT)N
18	9A3	(GT)16TTATGCTAT(GT)10	1	(GT)N
19	24E6	(AT)16	1	(AT)N
			2	
20	4E2	(CT)20 GT(CT)13	1	(CT)N

¹ Based on the sequence of the original cloned insert.

² Based on the sequence data obtained from the amplification products cut out of the gel.

For almost every sequenced amplified product, the microsatellite motif was identified. Sequence quality varied greatly due to possible contamination with degenerate or non-specific fragments. For 8 of the microsatellites, the length correlated directly to the size of the fragment as seen in Figure 20. In addition, variations in size on the gel correlated directly with variation in the length of the microsatellite allele.

Validation of Polymorphism for Seven Microsatellite Loci

A small study on 12 redwood trees collected randomly at Waddell Creek tributaries in Big Basin Redwood State Park was conducted to assess polymorphism for 7 of the best microsatellite loci characterized (18D7-3, 21E5, 20E5, 31E2, 20B3, 10H12, and 23F9). From that study, the results from four loci are shown in detail (18D7-3, 21E5, 20B3, and 23F9) since they display representative results. Bands were cut from the gel for most alleles to confirm the repeat number. The three that are not shown (31E2, 10H12 and 20E5) contained multiple fragments of different sizes on the gel, but were not polymorphic for the microsatellite loci. This could either be non-specific priming or polymorphism located somewhere else in the fragment. The quality of the sequence data was very poor so an accurate determination could not be made.

The best marker showing clear variation in the 12 individuals is 18D7-3 (Figure 21) which is a trinucleotide repeat. Sequence from amplified products for the 18D7-3 marker correlated extremely well to the fragment migration in the gel. Although some repeat differences do not match the size standard exactly, larger fragments did contain more repeats and vice-versa.

Figure 21. Validation of 18D7-3 marker across 12 individuals. Lanes 2-7 and 9-14 were loaded with the amplification products of 12 random individuals. Numbers indicate the number of repeat units identified in that band by sequencing. Lanes 1, 8, and 15 are the M3 size standard.

The cleanest looking marker is 21E5 (Figure 22) which is a tetranucleotide repeat showing limited variation. Sequence from amplified products for the 21E5 marker correlated exactly to the migration of fragments in the gel. The lower band is at about 104 bp whereas the upper band is two repeats larger (8 bases) at about 112 bp.

Figure 22. Validation of 21E5 marker across 12 individuals. Lanes 2-7 and 9-14 were loaded with the amplification products of 12 random individuals. Numbers indicate the number of repeat units identified in that band by sequencing. Lanes 1, 8, and 15 are the M3 size standard.

An example of how a marker can be misleading is 23F9 (Figure 23) which is a compound repeat that is much larger than the other three. The 23F9 marker is shown as an example of why compound repeats can be misleading for microsatellite analysis.

After sequencing the lower bands from lanes 6 and 7 of Figure 23, it was discovered that

a deletion in one microsatellite and an insertion in the adjacent one does not change the length of the products, thus two different alleles may be scored as being the same since they will migrate into the gel the same distance. This makes accurate analysis of alleles impossible if sequencing is not going to be used for confirmation. The upper band that appears in almost all the individuals did not produce good sequence and may be an artifact from the PCR amplification.

Figure 23. Validation of 23F9 marker across 12 individuals. Lanes 2-7 and 9-14 were loaded with the amplification products of 12 random individuals. Numbers indicate the number of repeat units identified in that band by sequencing (counting the CA repeat only). Lanes 1, 8, and 15 are the M3 size standard.

The most complex marker is 20B3 (Figure 24), which could not be properly identified by sequencing due to poor quality sequence reads, but represents a good example of how complex the patterns can be. The patterns observed for the 20B3 marker can be confusing. On the Father of the Forest DNA this marker gave only two fairly

clean amplification products (Lane 7, Figure 20). In addition this is unexpected since the repeat is a perfect dinucleotide repeat (refer to Table 4). There is a possibility that some of the bands are being caused by stutter or artifact from the PCR amplification. Nonetheless, the results were unexpected given the specificity seen from the amplification of the Father of the Forest DNA.

Figure 24. Validation of 20B3 marker across 12 individuals. Lanes 2-7 and 9-14 were loaded with the amplification products of 12 random individuals. All bands that were cut out of the gel gave poor sequence. Lanes 1, 8, and 15 are the M3 size standard.

Out of 37 potential microsatellite targets for redwood, four have been described here in full detail, two of which have good utility as markers (18D7-3 and 21E5). Due to the restrictions put upon the conditions for PCR, several clones that looked ideal never

made it into the final prioritized list (Table 9). A summary of where clones failed to make it to the next stage of development can be seen in Table 11.

Table 11. Summary of total clone failure for major stages in SSR development.

	Total
No suitable flanking sequence for PCR primers	10
Nonspecific products across several primer sets	8
Product size was too large for detection method	4
Compound or imperfect microsatellite	10
Nonpolymorphic (for 12 individuals tested)	1
Too many polymorphisms on gel	2
Ready to use for future studies	2

Practical Application of a Microsatellite Marker

The results of the preliminary reproductive study of redwoods in the meadow near Big Basin Park headquarters using the 18D7-3 marker is shown in Figure 25. Individuals within one fairy ring structure are expected to be similar at this marker. The trees in the southern grove (lanes 2-7) indeed look clonal based on the one marker (18D7-3). Likewise, all the trees from the northern grove appear clonal as well (lanes 20-25). A few trees in the meadow (lanes 8-19) have some alleles in common with the

northern grove, but none are identical to the northern grove trees. Two trees (lanes 11 and 13) have an identical pattern to the southern grove and could be clones of that grove.

Figure 25. Simple reproductive study using 18D7-3 marker. Lanes 2-7 are clones from the same fairy ring grove on the southern end of the meadow. Lanes 20-25 are clones from the same fairy ring grove on the northern end of the meadow. Lanes 8-19 are young saplings in the meadow. Lanes 1 and 26 are a 25 bp step DNA size standard.

The remainder of the trees do not share any alleles in common with either grove, indicating that the two flanking groves were not their clonal or sexual parents. These trees could have been derived from seed blown to that site from another nearby tree. Another explanation is that they were purposely planted there. Information from historians of Big Basin indicates that the meadow used to be a swimming pool and that trees were planted there by conservation groups in the 1970's and 1980's to fill in the area for restoration purposes. To add more certainty to these types of studies more markers from Table 9 will need to be optimized and considered in the analysis.

DISCUSSION

The creation of a *S. sempervirens* genomic library of a specified size using *Sau3AI* was successful. Alternative protocols suggest enrichment for fragments containing microsatellites by using hybridization methods prior to cloning. Given the repetitive nature of the genome not much would be gained by this method. A majority of the clones that were sequenced and did not contain microsatellites were highly repetitive.

The ability to use custom instrumentation at the Stanford Genome Technology Center (SGTC) allowed for the screening of thousands of clones quickly and cheaply compared to conventional methods. Designing the density of clones on the nylon and the size of the film took several attempts before achieving an efficient format that was amenable to the throughput of the instrumentation.

Hybridization of the probes to the DNA targets on nylon was challenging because a balance of sensitivity vs. specificity was difficult to achieve. Either the hybridization was too stringent, limiting the ability to achieve a good frequency of positive signals, or was not stringent enough, increasing the likelihood of false positives. The optimal balance between these two states was narrow (only about 5°C). After initial sequencing of some clones, it was easy to understand how false positives accounted for a majority of the positive signals. Close scrutiny of sequence data revealed that the genome is quite repetitive with several small clusters of quasi or pseudo-repeat motifs. Out of 155 clones initially identified as putative microsatellites, 75% were false positives. For two examples of sequence from false positives see Appendix J.

Nearly all of the clones recovered from the library that were thought to contain a microsatellite sequenced well. There were only three examples of clones that had sequencing gaps due to poor read length. This can be attributed to the length of the microsatellite they contained and the difficulty in amplifying these motifs. Some of the clones that had a sequence gap also contained fairly large inserts. These clones can only be closed by further walking on the clone since re-sequencing a majority of them did not give read lengths longer than the average of 500 bp.

Development of the PCR assay was the most complicated task. The problems associated with the hybridization stringency found earlier were further exacerbated by imposing more restrictions on a limited sequence (such as product size, primer melting temperature, GC content, etc.). Several attempts to optimize and unify the clones into a universal PCR condition led to the “weeding out” of more clones than anticipated. A more specialized approach for each clone may have been more appropriate, but the focus was to get as many primer sets as possible to work under the same conditions with potential for multiplexing. Performing PCR for a population study (with several samples) is difficult enough without having to run different markers at different specialized conditions.

Sequence reads from the amplification products of the Father of the Forest confirmed that the PCR assay works for a few selected markers. Obtaining good sequence from the amplification products cut from the gel was difficult, therefore the size of all the motifs could not be determined. The confirmation of most microsatellite motifs was discernable from the sequence reads, but the background was too high to make

correct estimations of length in all cases. The same was true for the polymorphism study on 12 individual trees. Although the amplification products were optimized using the Father of the Forest, one marker (20B3) amplified multiple sites yielding complex patterns that could not be corroborated by sequence data. The compound microsatellite detailed (23F9) is of little use since it showed a similar migration rate for two alleles consisting of different sequence. So far only two microsatellite markers (18D7-3 and 21E5) have been proven to vary in repeat number using actual sequence data.

The study done in the redwood meadow was a good example of the types of questions these markers can help answer, and it demonstrates that future studies involving populations should be researched for historical disturbance in the area. For example, logging companies might plant non-native seedlings after cutting, thus making a natural population study more difficult.

The two optimized markers reflect different extremes in polymorphism. The variation seen in 18D7-3 is more representative of the genetic variation expected in redwood given the hexaploid nature of the genome. Ten individuals in a close geographical area (3-4 square miles) accounted for nine different alleles at that locus. The lack of variation in 21E5 might be expected since it is a tetranucleotide repeat of fairly short length (Table 9). This marker may be more suitable for comparing differences between stands of trees because ten individuals only accounted for two alleles at this locus. It is interesting to note that the three markers that were supported with sequence data (18D7-3, 21E5 and 23F9) only show one, two, or three amplification products at each locus. No more than three bands were ever seen for any locus except in

the case of 20B3 (which has not been ruled out as nonspecific amplification). In theory, a hexaploid could show as many as six different alleles at a locus.

The development of these markers has proven that it is possible to analyze genetic variation at the DNA level for *Sequoia sempervirens*, despite the repetitive and hexaploid condition of the genome. This work will complement current research on the species and direct future studies to involve a genetic component at the DNA level such as those accomplished for several members of the *Pinaceae* family (Devey et al., 1996; Echt and Nelson, 1997; Paglia et al., 1998). These markers will better serve biologists, ecologists, foresters, conservationists, and private forest owners in understanding the value of what is being destroyed and what needs to be preserved.

LITERATURE CITED

- Anekonda, T.S. (1992). A genetic architecture study of coast redwood. Ph.D. diss., University of California, Berkeley.
- Boe, K.N. (1968). Cone production, seed dispersal, germination in old-growth redwood cut and uncut stands. USDA Forest Service, Research Note PSW-184. Berkeley, CA: Southwest Forest and Range Experiment Station.
- Brunsfeld, S.J., P.S. Soltis, D.E. Soltis, P.A. Gadek, C.J. Quinn, D.D. Streng, and T.A. Ranker. (1994). Phylogenetic relationships among the genera of Taxodiaceae and Cupressaceae: Evidence from rbcL sequences. *Systematic Biology* 19(2):253-262.
- Chaney, R.W. (1951). A revision of fossil *Sequoia* and *Taxodium* in western North America based on the recent discovery of *Metasequoia*. *Transactions of the American Philosophical Society*, n.s. 40(3):171-263.
- Cole, D.W. (1983). Redwood sprout growth three decades after thinning. *Journal of Forestry* 81:148-150, 157.
- Devey, M.E., J.C. Bell, D.N. Smith, D.B. Neale, and G.F. Moran. (1996). A genetic linkage map for *Pinus radiata* based on RFLP, RAPD, and microsatellite markers. *Theoretical and Applied Genetics* 93:673-679.
- Echt, C.S. and C.D. Nelson. (1997). Linkage mapping and genome length in eastern white pine (*Pinus strobes* L.). *Theoretical and Applied Genetics* 94:1031-1037
- Eckenwalder, J.E. (1976). Re-evaluation of Cupressaceae and Taxodiaceae: A proposed merger. *Madrono* 23(5):237-300.
- Fritz, E. (1931) The life and habits of redwood: The extraordinary are described by an authority. *Western Conservation Journal* 14:4-7,38.
- Hart, J.A. (1987). A cladistic analysis of conifers: Preliminary results. *Journal of Arnold Arboretum*. Harvard University. 68(3):269-308.
- Hirayoshi, T. and Y. Nakamura. (1943). Chromosome number of *Sequoia sempervirens*. *Bot. Zool. (Tokyo)* 2:73-75.

- Kusumi, J., Y. Tsumura, H. Yoshimaru, and H. Tachida. (2000). Phylogenetic relationships in Taxodiaceae and Cupressaceae sensu stricto based on *matK* gene, *chl* gene, *trnL-trnF* IGS region, and *trnL* intron sequences. *Am. Journal of Botany* 87(10):1480-1488.
- Li, L. (1987). The origin of *Sequoia sempervirens* (Taxodiaceae) based on karyotype. *Acta Botannica Yunnanica* 9(2):187-192.
- Li, L. (1988). The parents of *Sequoia sempervirens* (Taxodiaceae) based on morphology. *Acta Botannica Yunnanica* 10(1):33-37.
- Lindquist, J.L. (1974). *Redwood, an American wood*. USDA Forest Service Report FS-264. Washington DC: USDA Forest Service.
- Lodish, H. and J.E. Darnell. (2000). *Molecular Cell Biology*, 3rd Ed. New York, N.Y. W.H. Freeman Company.
- Miller, C.N. Jr. (1977). Mesozoic conifers. *Botanical Review* 43:217-280.
- Miller, C.N. Jr. (1988). The origin of modern conifer families, in C.B. Beck, ed., *Origin and Evolution of Gymnosperms*. New York; Columbia University Press, Pp. 448-486.
- Neale, D.B., K.A. Marshall, and R.R. Sederoff. (1989). Chloroplast and mitochondrial DNA are paternally inherited in *Sequoia sempervirens* (D. Don Endl.). *Proceedings of the National Academy of Sciences* 86:9347-9349.
- Olson, D.F., D.F. Roy, and G.A. Walters. (1990). *Sequoia sempervirens* (D. Don Endl.) in R.M. Barnes and B.H. Honkala, eds., *Silvics of North America. Vol 1. Conifers*. Agricultural Handbook 654. Washington D.C.: USDA Forest Service. Pp. 541-551.
- Ornduff, R. (1998). The *Sequoia sempervirens* (coast redwood) forest of the Pacific Coast, USA. in A.D. Laderman ed., *Coastally Restricted Forests*. New York, NY: Oxford University Press, Pp. 221-236.
- Paglia, G.P., A.M. Olivieri, and M. Morgante. (1998). Towards second-generation STS (sequence-tagged sites) linkage maps in conifers: a genetic map of Norway spruce (*Picea abies* K.). *Mol Gen Genet* 258:466-478
- Price, R.A. and J.M. Lowenstein. (1989). An immunological comparison of the Sciadopityaceae, Taxodiaceae, and Cupressaceae. *Systematic Botany* 14(2):141-149.
- Powers, R.F. and H.V. Wiant. (1970). Sprouting of old growth coastal redwood stumps. *Forest Science* 16:339-341.

Robinson, J.P. and S.A.Harris. (1999). Amplified fragment length polymorphisms and microsatellites: A phylogenetic perspective. Chapter 12. In: Gillet, E.M. (ed.). *Development, optimization, and validation of molecular tools for assessment of biodiversity in forest trees*. European Union DGXII Biotechnology FW IV Research Programme Molecular Tools for Biodiversity. Göttingen, Germany.

Rogers, D.L. (2000). Genotypic diversity and clone size in old-growth populations of coast redwood (*Sequoia sempervirens*). *Canadian Journal of Botany* 78(11): 1408-1419.

Rogers, D.L. (1999). Allozyme polymorphisms discriminate among coast redwood (*Sequoia sempervirens*) siblings. *The Journal of Heredity* 90(3):429-433.

Roy, D.F. (1966). Silvical characteristics of redwood. Research Paper PSW-28, Pacific Southwest Forest and Range Experiment Station. Berkeley, CA: USDA Forest Service.

Rozen, S. and H. Skaletsky. (2000). Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds.) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp 365-386.

Sambrook, J. and Russell, D.W., (2000). *Molecular Cloning: A Laboratory Manual 3rd ed.* Cold Spring Harbor, N.Y..Cold Spring Harbor Laboratory Press.

Sawyer, J.O., J. Gray, G.J. West, D.A. Thornburgh, R.F. Noss, J.H. Engbeck Jr., B.G. Marcot, and R. Raymond. (2000). History of redwood and redwood forests. Chapter 2. In R.F. Noss (ed.). *The Redwood Forest: History, Ecology and Conservation of the California Coast Redwood*. Washington, DC. Island Press.

Sawyer, J.O., S.C. Sillett, W.J. Libby, T.E. Dawson, J.H. Popenoe, D.L. Largent, R.V. Pelt, S.D. Veirs Jr., R.F. Noss, D.A. Thornburgh, and P.D. Tredici. (2000) Redwood trees, Communities and ecosystems: A closer look. Chapter 4. In R.F. Noss (ed.). *The Redwood Forest: History, Ecology and Conservation of the California Coast Redwood*. Washington, DC. Island Press.

Saylor, L.C. and H.A. Simons. (1970). Karyology of *Sequoia sempervirens*: Karyotype and accessory chromosomes. *Cytologia* 35:294-303.

Schlarbaum, S.E. and T. Tsuchiya. (1984). A chromosome study of coast redwood, *Sequoia sempervirens* ((D. Don) Endl.). *Silvae Genetica* 33:56-62.

Schmidt, A., R.L Doudrick, J.S Heslop-Harrison, and T. Schmidt. (2000). The contribution of short repeats of low sequence complexity to large conifer genomes. *Theoretical and Applied Genetics* 101:7-14.

Scotti, I., G. Paglia, F. Magni, and M. Morgante. (1999). Microsatellite markers as a tool for the detection of intra- and interpopulation genetic structure. Chapter 8. In: Gillet, E.M. (ed.). *Development, optimization, and validation of molecular tools for assessment of biodiversity in forest trees*. European Union DGXII Biotechnology FW IV Research Programme Molecular Tools for Biodiversity. Göttingen, Germany.

Stefanović, S., M. Jager, J. Deutsch, J. Broutin, and M. Masselot. (1998). Phylogenetic relationships of conifers inferred from partial 28S rRNA gene sequences. *Am. Journal of Botany* 85(5):688-697.

Tsumura, Y., K. Yoshimura, N. Tomaru, and H. Ohba. (1995). Molecular phylogeny of conifers using RFLP analysis of PCR-amplified specific chloroplast genes. *Theoretical and Applied Genetics* 91:1222-1236.

Tufor, K. (1973). Comparative growth performance of seedlings in vegetative propagules of *Pinus radiata* and *Sequoia sempervirens*. Ph.D. diss., University of California, Berkeley.

Veirs, S.D. (1996). Ecology of the coast redwood. in J. LeBlanc, ed., *Proceedings of the conference on coast redwood forest ecology and management*. University of California Cooperative Extension Forestry Publication. Arcata, CA: Humbolt State University. Pp. 9-12.

Yao, X., T.N. Taylor, and E.L. Taylor. (1997). A taxodiaceous seed cone from the Triassic of Antarctica. *Am. Journal of Botany* 84:342-354.

APPENDIX A

LARGE-SCALE ISOLATION OF DNA FROM REDWOOD NEEDLES

Based on the procedure of Doyle and Doyle, 1990 (Isolation of DNA from small amounts of plant tissue, BRL Focus 12:13-15). Modified for extracting DNA from needles by the Forest Biotechnology Laboratory at N.C. State University. (10/19/99, C. Brinegar)

2X Extraction Buffer (per 100 ml):

2.0% CTAB	2.0 g
1.4 M NaCl	8.12g
20 mM EDTA	4.0 ml of 0.5 M stock (pH 8.0)
100 mM Tris	10 ml of 1 M Tris base stock
1.0% PVP-10	1.0 g

Bring to <90 ml with ddH₂O. Heat at 60°C with magnetic stirring to dissolve.

No need to adjust pH (~8.5 @ r.t., ~ 8.0 @ 60°C).

Bring to final volume of 100 ml.

Autoclave 20 min and store at room temperature

Procedure

1. Add 15 ml of 2X Extraction buffer to a sterile 50 ml screw cap culture tube followed by 30 µl of 2-mercaptoethanol (in fume hood) and 0.75 ml of 10 mg/ml Proteinase K (in water).
2. Preheat the tube at 65°C in a shaking incubator (no need to shake yet).
3. Add 500 mg of freeze dried redwood needles to an autoclaved mortar. Add a little liquid nitrogen and crush needles to a powder with a pestle.
4. Using a sterile round-tipped nickel spatula that has been cooled in liquid nitrogen or on dry ice, transfer the powder to the 65°C Extraction buffer.
5. Immediately invert the tube to suspend the tissue in the buffer, then continue incubating at 65°C for 1 hr with shaking at 200 rpm.
6. In a fume hood, add 15 ml of chloroform:isoamyl alcohol (24:1), mix thoroughly by inversion until you see an emulsion form, then centrifuge at r.t. for 10 min in a tabletop at full speed to separate the phases.

7. Transfer the upper aqueous phase to a clean 50 ml tube. Be careful to avoid any of the insoluble material at the interface. Discard the interface and the lower organic phase into the appropriate waste container. Repeat the extraction then transfer the final aqueous phase to another 50 ml tube.
8. Add one volume of -20°C isopropanol. Mix gently by inversion and let sit at -20°C for at least 30 min (or overnight).
9. Centrifuge in a tabletop at r.t. for 15 min, remove and discard supernatant. Rinse the pellet with 2 ml of -20°C 70% ethanol (dislodge the pellet from the bottom of the tube with a stream of ethanol from the pipettor) and centrifuge again.
10. Rinse the pellet with 2 ml of 95% ethanol. (The sample may be stored at -20°C indefinitely at this point).
11. Centrifuge as in step 9, discard the ethanol supernatant and dry the DNA pellet in a vacuum dessicator for 15-30 min.
12. Add 0.4 ml of TE (pH 8) and incubate at 37°C for 1 hr in a heat block filled with water. Pellet should dissolve easily with minimal “flicking” of the tube. Total volume after dissolving should be ~ 0.5 ml.
13. Add 10 μl of 1 mg/ml DNase-free RNase A (for a 20 $\mu\text{g/ml}$ final concentration) then incubate at 37°C for 1 hr.
14. Perform another chloroform:isoamyl (24:1) alcohol extraction as in step 6. Recover as much of the top aqueous phase as possible, and transfer it to a 15 ml screw cap culture tube. Estimate volume then add 1/10 volume of 3 M sodium acetate (pH 5.5) and 2.5 volumes of -20°C 95% ethanol. Let sit at -20°C for at least 30 min.
15. Repeat step 9 then dry the pellet in a vacuum dessicator for 30 min.
16. Dissolve the pellet in 0.2 ml of TE by incubation at 37°C for 1 hr or let it sit at 4°C overnight. Remove any insoluble material by a 10 min centrifugation in a tabletop then transfer supernatant to a sterile microcentrifuge tube.
17. For quantitation, dilute 20 μl to 1 ml in TE and scan from 320-220 nm. Calculate the A₂₆₀/A₂₈₀ ratio (hopefully in the range of 1.8-1.9)
18. Label the tube with the date, your initials, and any other pertinent information and store at 4°C for a few days or at -20°C for longer periods.

APPENDIX B

OPTIMIZATION OF RESTRICTION ENDONUCLEASE DIGESTION

Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5399 USA

(Modified by D. Bruno 1999)

Perform the following small-scale reactions to establish the optimal enzyme concentration to generate a desired size range of fragments. In this example, Sau3A I is used; however, other restriction enzymes (e.g., MboI) may be used with their optimal restriction enzyme assay buffer, provided that they generate sticky ends compatible with the cloning site being used.

1. Prepare the following buffers:

Dilution buffer:

10X buffer B (Promega)	150 μ l
acetylated BSA (1 mg/ml)	+ 150 μ l
nuclease-free water to a final volume of	<u>1.50 ml</u>

Assay buffer:

genomic DNA	10 μ g
10X buffer B (Promega)	45 μ l
acetylated BSA (1 mg/ml)	+ 50 μ l
nuclease-free water to a final volume of	<u>450 μl</u>

2. Prepare the following dilutions on ice using Sau3A I (at 3 U/ μ l).

Tube #	Preparation	Dilution
1	10 μ l Sau3A I + 140 μ l dilution buffer	1:15
2	10 μ l 1:15 dilution + 90 μ l dilution buffer	1:150
3	10 μ l 1:150 dilution + 10 μ l dilution buffer	1:300
4	10 μ l 1:150 dilution + 30 μ l dilution buffer	1:600
5	10 μ l 1:150 dilution + 50 μ l dilution buffer	1:900
6	10 μ l 1:150 dilution + 70 μ l dilution buffer	1:1,200
7	10 μ l 1:150 dilution + 90 μ l dilution buffer	1:1,500
8	10 μ l 1:150 dilution + 110 μ l dilution buffer	1:1,800
9	10 μ l 1:150 dilution + 190 μ l dilution buffer	1:3,000
10	10 μ l 1:150 dilution + 290 μ l dilution buffer	1:4,500

3. After the dilutions have been made, assemble small-scale digestion reactions in separate tubes. This is done by numbering 10 individual tubes, adding 45 μ l of assay

buffer to each tube, and then adding 5 μ l of the appropriate Sau3A I dilution prepared in Step 2.

Final Enzyme Concentrations for Sau3A I Partial Digestion

<u>Tube #</u>	<u>Units of Enzyme per μg DNA</u>
1	1
2	0.1
3	0.05
4	0.025
5	0.017
6	0.0125
7	0.0100
8	0.0083
9	0.0050
10	0.0033

4. Incubate the 10 digestion reactions at 37°C for 30 min. Stop the reactions by adding 1 μ l of 0.5 M EDTA and 10 μ l of gel loading buffer.
5. Load approximately 20 μ l of each reaction on a 0.4% agarose gel along with the appropriate DNA markers. Run the gel at 2 V/cm for 16–20 hr or until the bromophenol blue has just migrated off the gel. Stain the gel with ethidium bromide or an alternative DNA-binding dye.
6. Photograph the gel and determine the amount of enzyme needed to produce the maximum intensity of fluorescence in the desired size range. Be sure to select more than one digestion time to increase the randomness of the library.

APPENDIX C

DYE TERMINATOR™ CYCLE SEQUENCING SETUP AND PURIFICATION

Applied Biosystems, 47 Wiggins Avenue, Bedford, MA 01730 USA

Millipore Corporation, 80 Ashby Road, Bedford, MA 01730.

(Modified by D. Bruno, 1999)

Preparation of Sequencing Reaction

Reagent	Quantity (Per reaction)	
Dye Terminator™ Ready Reaction Mix	2 µl	(1/4 reaction)
5X sequencing buffer (400 mM Tris, 10 mM MgCl ₂ , pH 9)	1 µl	
Primer	5 µl	(3.2 pmol)
Template DNA (for plasmids)	4 µl	(200-500 ng)
	<u>12 µl</u>	

Cycle using the manufacturer's recommendation.

Sequencing Reaction Purification

1. Load dry Sephadex® G-50 Superfine (Amersham Pharmacia Biotech) into all 96 wells of a MultiScreen HV plate using the 45 µl Column Loader (Millipore).
2. Using a multi-channel pipettor, add 300 µl ultra clean autoclaved water to each well to swell the resin. Incubate at room temperature for at least 3 hr.
3. Place the HV plate on top of a standard 96-well microplate, then centrifuge at 910 x g for 5 min to pack the mini-columns. Tape to secure if necessary

[Users of capillary sequencers, please note: for sample clean-up prior to sequencing on ABI 3700 systems, pre-rinse the mini-columns with 150 µL aliquots of Milli-Q water to remove contaminants from the resin. Centrifuge at 910 x g for 5 min.]

4. Carefully add sequencing reactions to the center of each well.
5. Place the HV plate on top of a NEW standard 96-well microplate, then centrifuge at 910 x g for 5 min to elute sample. Tape to secure if necessary.

[Users of capillary sequencers, please note: If using a capillary system, the sequencing products may be directly injected out of water without drying down. Store at -20°C if needed.

APPENDIX D

BACTERIAL COLONY LYSIS AND FIXATION OF TARGET DNA TO NYLON

Sambrook, J. and Russell, D.W., 2000. Protocol 31 in *Molecular Cloning: A laboratory manual* 3rd ed. Pp.1.135-1.137, *Cold Spring Harbor Laboratory Press*, NY
(Modified by D. Bruno 2000)

1. Prepare three separate trays with blotting paper saturated with the following solutions: 10% SDS, denaturing solution and neutralizing solution. Use a sterile round object to lightly flatten out the surface of the paper to remove any bubbles. Remove any excess liquid from the tray, but be sure the paper is fully saturated. (The paper should be virtually free of all bubbles).
2. Using autoclaved forceps and wearing gloves, carefully lay the membrane (edge first) on the 10% SDS saturated paper for 3 min. (For all transfers, try not to touch the membrane with your gloves).
3. Carefully transfer the membrane (picking it up by the edges) to the denaturing solution for 5 min.
4. Carefully transfer the membrane (picking it up by the edges) to the neutralization solution for 5 min.
5. Float the membrane in a tray filled with 2X SSPE, agitate on an orbital shaker for 10 minutes, allowing the membrane to sink to the bottom.
6. Pick up the membrane with forceps and hold vertically for a few seconds to allow excess liquid to drip off. Lay the membrane on blotting paper soaked with 2X SSPE (The membrane should be wet, but not dripping).
7. Spot positive DNA control onto the membrane. (optional)
8. UV crosslink in a Stratalinker® using the "Auto-crosslink" function (1200 microjoules)

PROTEINASE K TREATMENT OF FIXED DNA TARGETS ON NYLON

1. Immerse the membrane in Proteinase K solution for 1 hr at 55°C with mild agitation.
2. Rinse membrane once in 2X SSC at room temperature for 5 min.

3. Immerse the membrane in PMSF solution for 15 min at room temperature with mild agitation.

4. Rinse twice in 2X SSC at room temperature.

**** Add 5 ml of 1 M NaOH to PMSF solution after use and discard after 24 hours ****

Denaturation Solution

NaCl 1.5 M

NaOH 0.5 M

Neutralization Solution (pH 7.4)

Tris-Cl 0.5 M

NaCl 1 M

20X SSPE (pH 7.4 and adjust volume to 1 liter)

NaCl 175.3 g

NaH₂PO₄-H₂O 27.6 g

EDTA 7.4 g

25X SSC (pH 7.0 and adjust volume to 1 liter)

NaCl 219 g

Sodium Citrate dihydrate 110 g

Proteinase K stock solution (20 mg/ml)

Dissolve 200 mg Proteinase K in 10 ml sterile water. Store frozen aliquots at -20°C.

Proteinase K working solution (100 µg/ml Proteinase K in 2X SSC)

25X SSC 16 ml

ddH₂O 183 ml

Proteinase K stock 1 ml

PMSF stock solution (200 mM in isopropanol)

Dissolve 0.35 g solid PMSF in 10 ml isopropanol. Heat to 50°C to dissolve. Store in aliquots at -20°C. Redissolve by heating just before use.

PMSF working solution (1 mM PMSF in 2X SSC)

Prepare fresh as needed

25X SSC 16 ml

ddH₂O 183 ml

PMSF stock 1 ml

APPENDIX E

DNA HYBRIDIZATION TO FIXED DNA TARGETS ON NYLON

Adapted from Southern-Light™ and Southern-Star™ Systems protocol.

Applied Biosystems, 47 Wiggins Avenue, Bedford, MA 01730 USA.

(Modified by D. Bruno, 2001)

1. Rinse membrane in 0.25 M sodium phosphate (pH 7.2) for 3 min.
2. Prehybridize in hybridization solution for at least 1 hr at 55°C
3. Prepare probe by heating it to 95°C for 5 min, then snap cool on ice.
4. Add probe (1 pmol/ml each oligo) to a fresh batch of prewarmed hybridization solution and hybridize overnight at 55°C. (The actual temperature should be determined empirically).
5. Wash twice, 5 min each at room temperature in 1X SSC/1% SDS.
6. Wash twice, 15 min each at the determined hybridization temperature in 0.5X SSC/1% SDS. (Be sure to have the solution at temperature before adding the membrane).
7. Wash twice, 5 min each at room temperature in 0.5X SSC

Hybridization solution

0.2 M EDTA	0.25 ml
ddH ₂ O	7.25 ml
0.5 M sodium phosphate (pH 7.2)	25 ml
20% SDS	17.5 ml

25X SSC (pH 7.0 and adjust volume to 1 liter)

NaCl	219 g
Sodium Citrate dihydrate	110 g

0.5 M Sodium Phosphate stock (pH 7.2 and adjust volume to 1 liter)

Na ₂ HPO ₄ -7H ₂ O	134 g
85% Phosphoric Acid	4 ml

0.2 M EDTA (pH 8.0 and adjust volume to 1 liter)

EDTA dihydrate 74 g

APPENDIX F

CHEMILUMINESCENT DETECTION OF BIOTIN LABELED PROBES

Phototope® Star Detection Kit protocol, New England Biolabs, INC.,
32 Tozer Road, Beverly, MA 01915-5599 (Modified by D. Bruno 2001)

The following modifications reflect a reduction in the standard volumes and/or concentrations as recommended by the manufacturer. A different tray is used for each numbered step. The membrane is carefully lifted from one tray using autoclaved forceps, then placed in the next solution for the given time at that step. Trays can be rinsed with deionized water and reused for subsequent steps. Pay close attention that the sides and corners of the membrane receive a uniform treatment. All steps are done at room temperature.

1. Agitate in blocking solution for 5 min.
(36 ml blocking solution)
2. Cover thoroughly in streptavidin solution for 5 min.
(Add 10 µl streptavidin to 10 ml blocking solution)
3. Agitate in wash solution I for 5 min.
(Use 100 ml wash solution I, repeat 4 times)
4. Cover thoroughly in biotinylated alkaline phosphatase solution for 5 min.
(Add 10 µl BAP to 10 ml blocking solution)
5. Agitate in blocking solution for 5 min.
(100 ml blocking solution)
6. Agitate in wash solution II for 5 min.
(Use 100 ml wash solution II, repeat 4 times)
7. Cover thoroughly in CDP-Star substrate for 5 min.
(Add 10 µl CDP-Star substrate to 10 ml CDP-Star Diluent)

8. Do not let the membrane dry out. Take a clear sheet protector and cut three of the four sides. Tape one side of the sheet to your benchtop and open the top sheet. Lay membrane in between the sheets carefully, then roll the top sheet across the membrane so as to not introduce any air bubbles.
9. Take the membrane into a darkroom and expose to film with colony side facing film. Exposure times may vary but are typically < 5 min.

Blocking Solution (pH 7.2 and adjust volume to 1 liter)

NaCl	7.3 g
Na ₂ HPO ₄ (dibasic)	2.4 g
NaH ₂ PO ₄ (monobasic)	1.0 g
SDS	50 g

Wash Solution I (q.s. to 1 liter)

100 mls Blocking solution

10X Wash Solution II (pH 9.5 and adjust volume to 1 liter)

Tris	12 g
NaCl	5.8 g
MgCl ₂	2.0 g

Wash Solution II (q.s. to 1 liter)

10X Wash solution II 100 mls

APPENDIX G

MANUAL PROCESSING OF BIOMAX LIGHT FILM

Scientific Imaging Systems, Eastman Kodak Company,
4 Science Park, New Haven, CT 06511

The Tray Method includes at least three trays that are two centimeters or more larger than the sheet of film to be processed, and an adequate amount of running water for washing. The film is moved from tray to tray by hand using print tongs. Maintaining fresh processing chemistry is critical to obtaining high quality autoradiograms. Old developer and fixer will adversely affect the image quality of processed film, even if they are used infrequently. Kodak recommends that the developer and fixer chemistry be changed every month to ensure optimal processing conditions. In addition, the developer chemistry should be replenished during processing at a rate of 60 ml for each 35 x 43 cm film sheet processed. The fixer chemistry should be replenished at a rate of 75 ml for each 35 x 43 cm film sheet processed.

Step I: Developer

KODAK GBX Developer and Replenisher is recommended for developing Kodak Autoradiography Films. Do not agitate the film in the developer.

Step II: Rinse

Remove film from the developer and place in running water. Wash with continuous, moderate agitation for 30 seconds. Recommended rinse temperature is 16-24°C.

Step III: Fix

KODAK GBX Fixer and Replenisher is recommended for fixing of Kodak Autoradiography Films. Remove the film sheet from the rinse and place it in the fixer tank for 5-10 minutes. Recommended fixer temperature is 16 to 24°C. Moderate agitation of film in the fixer is recommended.

Step IV: Wash

Remove the film from the fixer and replace it in a running water wash for 5-10 minutes. Running water wash should be 16-24°C with a rate of 8 volume changes per hour.

Step V: Dry

Remove the film from the water wash and hang it with a clip attached to one corner of film sheet in a dust-free area at room temperature until the film is dry.

APPENDIX H

COMPLETE FLANKING SEQUENCE FOR MICROSATELLITES IDENTIFIED

Ordered by increasing number then alphabetically. Microsatellites are in **BOLD** typeface and any gaps in the contig assembly are represented by "...". Poor sequence data is basecalled "N". Some clones did not assemble into one contig, thus the corresponding forward and reverse sequence reads are shown. Some clones only contain repeats in one direction or highly repetitive sequence. All reads are shown 5'—> 3' except for those labeled "Reverse".

1B5 Consensus (699 bases)

AAAAAATATAGGCACCTTGAGAAAGATATCTCTTGAATATTGTTTTAAATT
TTAAGAGATATATATATGAGTAGTATACATATCTCAAATTACATTAATAATGCAT
GCATACATACATACATACATACACACACACACACACACACACACACACACAC
ACATAT
ATAGATATAGATATAGATATGTGTGTATGTGTGTGTGTGTACTCCCTATGAG
AGCTTCACTCCCATGGGGGCTACTGTCCCCTACAAGCCAAAGCACACAAAGG
GGTTTGCAACTTGCCATGTGAGTCAAAGCCACCTCAAGTTTATAATTATATT
TTATTCCCTATATATAGATATGTGTGTATGTGTGTGTGTGTGTATTCCCT
ATGAGAGCTTCGCTACCATGGGGGCTACTACCCCCTACAATCCAAAGCACAC
AAAGGGGTTTGCAACTTGCCATGTGAGTCAAAGCCACCTCAAGTTTATAATT
ATATTTTTTCATTCCCATCAATCAAGACAAAAGTGTATAATGCACACCACACA
TTCGTATGTTTTAAAAAATAATTTAAAAATCAATGAAACCCTATCTTAA
ATAAATCACAAGTATAAGGTGTATGCGTCTATTATAATGGAAAAAAGAAAAT
GAATATTTTTTAATAATGAGTGATC

2B2 Consensus (1072 bases)

GAATGTGGGACGTGGAGGGAGGAGGAGGACGACTGTCCGCTGTGACGAGAA
GAGGCTGTGTGAAGGAGACGTGGAGCTGGGGGGTGGTCTGAGTGTGGCTGA
GGACGTAAGTGGCGTCATACGAGTTGGGGGCGCCTGGCACCTCCATGTC
CTCAGCTGTCCGTGCTGGTGTACCTTCTGCTGCGTGGTGTGCTGCTGCTGCTG
TGCTGCTGCTGTAGCCTGTTCTGCTCTCTGCGTAGTCTCTCTTCTTCTGCTG
CTGCTGCAGCTGTTGCTCCTCCTCTTGTGCGTCCGATCTCAGCTCGTCTCC
GCTCCTCAGTCACGTCCACCACATCCGGAAGTGTCCCGCTCTCGTCCGCGACTG
GCCTCATCCGGACCGCCGATCGGAATCACACGTCGTCGTTCAACCTCACTCCT
CTGCCTCTGTGCACTTCCGGACGGCAGCGCCTGCTGTCCGTCGCTCGGCCAG
TCAGCCGCTCCGTCGTGCCGTCGTGCCTGTGCAGCACTGTGCTCTCCTCAGTG
CTCTGACGGCCGTTGCGGAAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
CCTGCTCGTTCGAACGAGGTGCGGCTCGACACGCTCACCGCATTGCGACCAACC
TGTCAGCGAGGGCCGTCCCATCCCCATCCCCAGCCCTCCCATGCCATCCCCA
TGCCGCCCGCAGCCAGCAGGCCCAATGGTCCAGCGCTGAACAGCGAAGGCC

AGCGCCCAGCATGCCAAAGCCGAACGGCATAGGCGCAATGCCAAAGCCGCC
CGCGTGAGGCTCGTCGAACGGCGAAGCGAAGGGGTCGTTGTCGCCGAAGAA
GTCGCGGAAGATGTCGTGGGCGGAGCGGAAGGTGAACACGGGCTGTCGGAA
GGCGAAGTGGCGTGGCGGGCCGAAGGCCATGAAAGGATGTGCTGCGAAGCC
TGCACTGGCGCCGAAGGGTGCAGTGGAGTGTGGCGGCTGCTGTGGGAGTGAG
GTGGCGAGGGAGGTGCGTCGAGGGCGTCATCGTACTCGCGGCGTTTGTCTGT
GTCGGAGAGGACGGAGAATGCTTCACTGTT

2D5 Consensus (1157 bases)

GCTTGCTGCCTGCAGGTCGACTCTAGAGGATCACATTGGTGAACCCGACTCTC
TGTTTTCTGTCATTTTTCCGTACAGTTGCATGTATTTTTGTCAATTTTGCAGG
AACAAGTACAGAGTGCCTGCAATCGTAGGCTCAAGTGCCGCAATCGCAA
AATCGAGCCGCAACTGCCAAATTGAGCCGCAACTGCAAAATCGAGTTGC
AACTGCAAAATCGAGTTGCAACTGCCAAATCGAGTTGCAACTGCCAAAT
CGAGCCGCAACTGCCAAATCGAGCCGCAACTGACGAATCGAGCCGCAAC
TGACGAATCGAGCCGCAACTGACAAATCGAGCCGCAACTGACAAATCGA
GCCGCAACTGATAAATCGAGCCGCAACTGATAAATCGAGCCGCAACTGA
TAAATCGAGCCGCAACTGATAAATCGAGCCGCAACTGACAAATCGAGCC
GCAACTGCCTGGTAAAATTTTTCTTTCAGTTTTGAGGTTCTGATTTTATAGATT
TTCGGGTTTTTCTATTTACTGGTTTTCTGGTTTTCTATTTCTTTCTGTGTCTTT
ATTCTTGCAAGTTAACTAGATTAGATTTTTCTTGTATTACTAATCTTTCTTGT
AGCGTGTGGCAATCCTCATTCTGTGTGCCAAGACCATTGCATAAGAGACTACT
AATTCATTGTTTGCAGATTCGAATGTCAAAGAAGGAAGAACAACATTGTGC
TTTTTTTTCTTTCATAGATTAGCAAACATTTATTTTGGGGCATATAAAAAGAA
CACATGTCTTCTTTGGTTGATTGTTCTTGCAAGATGGTAGGTGAGTATGCTCT
CACTTGGTGAAAACCTTGTGATCAGAAAACCAGACCCATCTTTATTTGTTAAT
TACCCTTCACGATTCGGTATATCAGAGTATCCTCTCCCCTTGCGCATTTCGTGC
GATCCTGGGTGTAAGAGAAACGGTTACTGACTCTAGAACTTGGGCCTGTTCT
GGTGGGAAAAAGTACTACCCCCAGGACGACTCAGACCAGTAGCACCCCTGAC
ACCGAGGGACTTCCAGCCGTATATAAAAGTGGGTTCTTGGGGAAAGCCCAA
GAATCACGAGCGTGGTTCCGACGAGATGGATCCCCGGGTACCGAGCTCGAAT
TCGTAATCATGGTCATAGCTG

3A7 Consensus (1134 bases)

AAACAATAAGCCAATTGGAATCTGGGTGCAAACACCACAAATGGCATGATGA
CCTTCATTTGAACTAAATTCATATATAAGAGAATGTAATTTCCCGTCTCAA
CTATTCTAGTTTTTAAAAATAAATGGCTTGTAGATTTAAAATTTGTCTACTTA
AACTTTGTGTTTGAAGAAAAGAAAAATGAGTATTGTTAAGGAGATTTCTTTGT
TAGTGTGTGTGTGTGTGTGTATATATATATTATATATGCTCGAAATGAA
AATTCAAGTTCCTTGTTTTAAATAATGATGTATGGAACTTTTAAGCCATTCTA
AACAAACAAATTATACATGAAAATATATATTTAAATATATATATACAATATT
GGAAAACATGATAAATAAGAATGGAACCATTTAAATACATATAGAAATACA
AACATAATATAAATGCATGCATATTCAATTTATATCAAATTGAAATTTATAAC

CGATATATTGTAAAATTGCTCTATTCGAAATGAAGAAGATATCATAAGGAAA
TCCACTAAGATAGTGGACTATCACCACCTTTATGGAAATAATCTAAATTTAGA
GGTAGTCAGCAATGGCATGGCAATAATGGCATCACTAGAAAAGCATGAAAA
ATAATGTATTGTATCTATCCTTTTGGGGGAAGATAGTTGTGTGAAAAAGCAG
GGAAGGAGGCAACAAATTGGGGGAAAATTTGAAAGGAAAAGGGTCAAATAA
AAATTCAAGAAAAATATAAGGAAAGAATGACATCATAGTTCAATTTGGGAGG
TAATTTGATCCCCATATTCTTGCAATGACTTAGCATTACTTCTATTATCGAATG
TTGTTAGAAGTTTATTCTTCATTTTTCTCAAGAGTATTCAAGGAATCATAGGA
AAATATGTCGTTGTTCTAGATGTCAAATGGCATGAAGCTGAATGAGTGAGTA
GTTATATGTATGTTGTTTCCACCACATGTAGGATGTGTGGCCTATGGCCTGCA
CAATGGGAAACTGACCTAGTTTAAAGCCTTCATGGCTAATGTTTAGAGCCAG
GACGAGGCTCGAGAGCCTCCAATGGCTAGGCCTCCACCTTTTTGTTGCATAAA
AATACTCTACCTTATGTAGGCTATCATAAATGATC

3D4 Forward (357 bases)

TGGGGCATGCATCATTGCTCTAGTTCCATTTCTCTCTCTATATATATATATA
TATCCAGAACTTTCGTACGAATAGTTAGTCAAGATACTGCTGGGTAGTAAAG
AGACTACTGTCAAGATACAACCGGAGATATAGGGATTTGGGTGGTCAGTTTT
TACCAACGGAAAGCATGGGGGAAGCTATGAGGGTGGGGGGGGCCCAACGGC
CCTACAGGATATAGGCGGCACTATAGTCCTTCTTTCTATAGTTGGTTCTATAC
TCAACTATAGAGTTGGCGTTTAGTAACGTCTAGTTTAGTGCTTAGGATTAATT
CCCAACCCATGGAACCTATAGAGAGGGGGGGGGCGGGGGGTTAT

3D4 Reverse (342 bases)

GAATACAATTTCTTGTTTCCGGTCTTCGGGAAGGACCGCTCGTCGACAAGG
AAAAGGAAGGGGGGAGATAGATGAAGAGTTGGAACGCCTGTGGCACAACCA
CTGGCTGTAAGGCCCTCAACAACCTGCACCCTGTAGAGAGGGAACTGAATAT
AGAATTGTACAAGCTGGGCGAGGGGCAGGTAGCCTCCAAAATGTCTTATACT
TTTGAAGAACTTCCAGCCGTCCCTACAATTCTATGGGGACCAGCTCTCTCCCC
CTATAAATTCTATGGGCCCTACAATTCCAGCTCTTCGATAGCTTGCTAGAAT
AGGAGAATGTGTGTGGGGATGGGTGGGGG

4A8 Consensus (936 bases)

GATCCTGCACGTAATCCATCCATACTGCTAAAAATAGCAGGGGTCCCACGCC
GAGTGCTGAATACTGCTAAAAATAGCAACTCAGGCTAGAAGTGTCGGAACGA
GCCCCGAACGAAAGCCAACTGATCACAACCTGCCAACTGAAGTTGAGCCAAG
CTCCATCAAGACCCTCTATCCAATCTATCAGCCTACGAGGGTTCGCATAATAG
GCTAATCCTGTGCTCGGAAAGATGTCAACTAGAAGGGGACATGACAGTTCAT
TTGGGTGCCTAGTCAAATCAATTTACTAAAATGATTTAAAAAGCAAAAAATA
AATAAAAAAGTGCCAACTTGGATGCTGTAGGGACGGGTTCATCGTGGGTTT
GTCTGTATCTGCCCCAAGTTCGTCTGGGTTTGCCCCAGGCAATTGAATGGCCT
TTTGACGCCCTCGGCCCGGACCCACAGGCAAACCAAACTAGTACCCAGATCT
AGTTCAAACCGGATCAGGATCCAATCGATCACCTATAGGATCCAGTAGCATA

4E2 Consensus (808 bases)

5H3 Consensus (545 bases)

7F11 Consensus (1012 bases)

81

TAGAATTTAGCAAACATAAACTATGAACATTTTCATCTCATTAGAAAAGAAG
 GGCCCATTTCTACCATCACAAAATTGTACTAGTATCAAGTTAAAATGTACAAT
 GCAACATAGCTTACCACTATAAATACAACCTACACTACTTGAAGAGCAACTA
 GAGAATCTACATTTTCAACTTGAATCTAATGGGTCGCACTTACATCCTAAGCT
 CAAAGTGATAAAATTCTATAGGAGAAAATTGTTTGGTAGAATTCCCTAAAGA
 TGTGTTGTCATGGATTTCTATGATTGCAAAATAGCCACAAAATGGATGTGTTG
 AAAATTCCTTATAAACTTTCAAGCAAATGTAATCGACACATATAAAATTCGTC
 ACAAAAACATTTTCTTATCTTCCTTTTAATTGTGCCAAAATAGGAGCTCTATA
 ATAGGGTATTGCATCAATCACACATAATAGAAGGGGATTTTGTAGATTTTA
 TAGATGCAATTTTAACTATTAATCTCCTTCTCTCAATATCTATCTCTCCCTC
 TACCTCTTGGTCACTTTGTCTCTCTTCATCTATNTCTCTCTCTCTCAATAT
 TTGTGGCTCTCTCTATCTCCCTCTTCTCTCTTTGATATCCCTCTCTTTCTC
 AAGTCCTCTCTCTTTGTTCTTTATA.....CTCTCTCTCTCTCTCTCTCTCTN
 CTC
 TC
 CTGCTCCCTCTTTCTATCCCTGCTGCTGCTCCCTCTTTTTCTATTGCAAAC
 ATAACATTAATAATTGTTGGTAATGGAGCTTGATTATCTTCTTAATANCAATGT
 TTTGCTTTCAATCATGTTTGG

8E8 Consensus (798 bases)

GATCATAGTCTAATTACCTTTAATTAACATAATCTTGAAATCATAATTTAATT
 ACTTTAATTAATCCTAAAATCATGAATTAATCATTATTACATAATCGAATATA
 ATACATATATATGTATCTATATATATATGTATAACCATTATCAATTAGCTCT
 AGGCTTACAATACTCACCTTACACGGGCATGCCTATTATCTATACACACAC
 ACACACACACACACAATTAATTATATTATACTTTGTTGTACCCAAACCAA
 ATATTTTTGCTTCATCAAGGCATTTTGCATAAATTAGTATGAACCTTGTTAAA
 ATGATGTAAAATACAGTTTTTGGGCAATATCAAGCATGGAGTTTTGTTAGTGT
 TTTGTCAGTTAAAGGTAGTATCCATTGATCAAAACCTGAATCAGAGACTCAG
 AGCGAAGCATTTACAGTGGAGCATTTCACTTCGACACTTATACTTTTAAATG
 GTCATTTCAAGTGACCAAAATGATAATTTCAATGGTCAATTTTTGAAGATTTCT
 CAAGGAAAAGCAAATTTGCGCCCCCTCTAGCCCAAGGGCGCCCTTCTTTTTGA
 AGGCCTAAATTTTGTGATAAATGCTCATGAATACAGTCAATCTCTTTAGAATT
 AGTCTGCCTATCCGATGAGATATCCTGCAAGGCTATCCAAGACTCGTAACTCC
 GGTAACCTGGCAACAAAAGTTGAAGATTCCAAAGTTGCAGTTTGCAGAAGCTT
 GTCATTTTTAGCTTTTTCGCCCCCGTAGCTCAAGGGCGCATTCTTTAAAGGA
 GGGTTAGTC

9A3 Consensus (1157 bases)

TTGGGAAATGGGCTTGCTAGCATGTATGCAAGCTATAGATTCTTAGAAAATG
 TACTCCTTGTGAATAAGAAGATAGCTAAATTGCGATTCAACAACAATTACATT
 GATTCACATTTATTTAAAAGAATGATTGATTGTGTGTGTGTGTGTGTGTGTG
TGTGTTTGTA
TGTGTCTG

**CACACGCGCACGCGTTCACCACAACGCACACGCGCACGCGCACGCGTTC
ACCACAACGCACACGCACACGCGCACGCGTTCACCACAACGCACACGCA
CACGCGCACGCGTTCACCCAACGCACACGCGCACATACTGAACCGTT**

10H12 Consensus (642 bases)

**TGTATGCAGAAAAGCAGCAGATTGGAAATAAAGAATAATTATATATATAGA
TAAATAAGAATACCAATCAAATATAAATAAAGGAATATATATATATATAT
AGTTAGAAGACAGAATTCATATAAAATAAGTATGGGAAGATTGAATAGGGAT
AATAGATTCTTATTTGTTTGATTACGCTAACAATATCTTCTGAATGTAAAGG
ACAGATCGGACAACAAAGCAATTGGAGCAGATCAAACATAAGTATCTGCACC
CAAGGGCAGAATCAATAGTTATTACAATTACTGCTATGAATATAGCCATCTTA
TGTGTTAATAGATATTGTGGTATGCAGAGATAGTTGTTTTATGCTTATTTATC
AGATAAATGTAGATATATAATTGTACTAATTAAACACTGATATAATATCATAT
TATGTCTTCGGGTAGCATAACATTTACTAGTATTATGAATAGTGATATGATTTT
GATTTGGCAATAATTGGTCACAAAGGGACGAGGCAGGAATGCAAGGAGGGG
GGACTGTGAATGATTCCTCACTAGGAATGGCACCTCATGATGATGGTTAAGG
ACCATCGGTAAAGGACCACATGAGGCCAAGGAGGATTAGGTCCTATCCTTGG
GCCTAGGGTTGAGGATC**

11D11 Consensus (625 bases)

**CATGATCTGTTTATAAAGGAAGTGATGTCATAGATGCATCAAGTTCAATTATG
CATCTTTTAAACAATGAGGAAGCATTTCAAGTTGTTGTTAAGCCCCTCTCACT
GATAACCCTATTGGATAGTCACTCTATCACATAGGAACTTGAGTTGGTAGGTC
ACCCACTATATGTGACAGGTGGTTGTCCAATAAAGATAAAATGAGCAAGCCC
ATGTGTATTTATTACACATAATAAAGAAGATAAATGGTTGAATCAAGGTC
ATGATACAAGTACTAGGAGAGAGATAGAGGGTAGAGGGTTTAGAGTTTA
AGGGTTAAGGGTTTAAGGGTTAGGGTTAAGGGTAGAGGGGAGAGAGATA
GAG.....AG
AGAGAGAGAGAGGGTAGAGGGTAGAGGGTTTAAGGGTTTGGAGTTTAA
GGGTAAAGGGTTTAGGGTTTAGGGTTATGGGTAGAGGGGAGAGAGATA
GAGGGATTCAGTCAAAGACAAGACCCCAACCTTGGGATTTTTCCAATACCAT
AAGACAGGAAAGGACATATTCAAAGATGCAACATACTGTGTTGATATGTCC
TTCCTCTCTTAAGAA**

11G8 Consensus (1504 bases)

**TATAATTCACATAAATTATTTACCTAACCATAAATTATTAATGCCTCTATAAT
CAATCATGGGAAGAAGAGACAAAAGTACAAGAGAGGTCTGGAGAGATTCCA
TGCTAAGTAGCACCACCCTAATGGATGCCCATGTGTCTCTCATTAGTAGGCC
AAGGGAGTTGCCTATCTTCCCTTGGGCTTAGAAGTAGGAAAGACTTGAGGTT
TCGTCTCCAACCTTCCCAATGATTGATTGCTGGGGGAATAATAACTTTGAATT
AGTATTGTTAGTTACTCATATATGTGTGTGTGTGTGTGTGTGTGAATAA
TGTAATGCATGTGCAAACAATATACATGACCAATTAATTATTTTATGAATCAT
TGGCTAACATCATTGCAAGGTAAATTGGAGTTTATTAGAATTTGTAATTTAA**

TGAAATATATCTCGTACTTTCAAATTGTGATTTTCGGGCAAAAATCTAAGCAT
 ACCAAATAGGGGACATTACAATTGGTATCAAAGTTTGTTCCTACATGACATA
 CCTACTTTAATCATCATTCAATTAGAATTAACATTCAACCTTCTTAAATAGG
 CCAATAATGCATTTTTGTATGCATGTGCCTTCCACGTGTACACTTTGTGATTAT
 GTATATTCATGTGAATACTTTGCGACTTATGTGTATTTATGTATACTTTGTGCC
 TACATGTATTCATGTGTACACTTTGTGACTTGCATGTATTCATGTGTATACTTT
 GTGATTACA.....TTGACTCATATATTCAGTGTGCTTTGTGACTTACATGTATTC
 AAGTGTATACTTTGTGACTTACATACATGTATTCATGTGAATACTTTGTGACT
 TACATGTACTCATGTGTATACTTCACGACTTATGTGTTTCATGTGACATACAT
 ATATTCAAGTGTATACTTTGTGACTTACATATATTCATGTGTATACCTTGCGA
 ATTACTTTGTGACTAATGTATATGAATTTGTGTGCATGCTTTGTGCCTTATGTG
 TATCTATGTGTTTGCTTCATATTGGATGTGTCTTCAAGAAGATTACTCCATGC
 GTGTTTATAATTACTTTATGCTTAAAAATTATGTAAGTGAACATTA AAAATGT
 TGTCTATCAAAATGCTTCTTGGGTAAAGAAGCACTAACAAATTATCTTATTTA
 AAATTGAATAATTATTAATTATCCAAATGCAATAACAAAGAGTTGTAACATT
 GAAGTGAAAGGTGTGACTCCCTCATACTTGTGATATAAAAGGGAGAGCACTT
 CATTTGAGAAACACATCCATTACATCATTGAAGAATTCAATTGAGAATAGAA
 CCTTAAGCAATTTTACAAATGAATTAGTTTAGGTCTCACGAATAGTGATGAT
 GTGAGGAATGTGTGAATCGTAATACGAAAGATTGTGACTCTTCAAACGAAAG
 GTCGCGACTCTTCAAAGGCCAAAAGTTGTGGCCCTTCAAATGAGAAAGACTG
 TGA CTCTTTTATAACAAAGGGTGATC

13G7 Consensus (964 bases)

AGCAGGACGGCCATTGACACACGAGACACCATCGTGGAGGGAGGGAGAGAG
 GCCGAAGGGGAGGAAGGGGTCAAGTATCGAGAGTTAGTGGAGTGCAAGGGT
 AGGCGGGTGGGAGGAGGAGGAGGCAGATGCTGGGTGTGTGAGAGCGATGGT
 GAGTGAGAGTGTGCGAGGAGCGAGCGAGTGGGCTGGGCTGGTGCTGGAGTG
 GTCAGTGGTATGAGTGTGCGCCGGTGGTTGCAGTGATGGCGTGTGGGTGCTG
 AGTGTGACAGTGGGACACGACGACACGAGTCGTTGGTGCGGTTACGAGTGAT
 GCGTGGGGTGGCGGAAGGTGTGCGTGCAAGTGCAGCTGCAGCCGTCGCGTGATG
 GGAGCACGGGCGAGTGAACGATCGATGAGCGATGAGCGATGCGTGTGTGCG
 TGTGAGCGTGTGGCTGTCTATCGGCCCCGTCTCTGTATGAGAGTGCAGCTGCTC
 GTGTCGTGCCGTGCTCCTTCCATCGACAGCAGCGGGCGTGTGGAGTGTACGG
 CTTGTATGCGCTGTGTGAGCTGCTGCTGCTGCTAAGAGCGGACAGAGT
 CTGTGCGTGTGGCGCCGTGATATGCAGCGCCCGAGTCGCAAACAAGCAACGC
 AGCGGTGATGTGATGTGATGTGATGTGATGACAGCGACGCTGGTGCTGCC
 ACTCCACTGCATTGCATTGTTGTCACCACTCCTGCTGCTGCTGCTGCTGCTG
 CCTGCTGCCTGCTGCAAGCTGCGATGCGATGCTGCGACAAGTTGCGGTCCGC
 TGGTTGAATCACTCTCTGCCCTCGTCACTGTGCGGCGGTTTGTGTGCTTGGTCT
 GTGCTGCTGTGTGGGTGTTGGCTGCTGCCGAGTGCGTGGCTGACCGGCGGTGT
 ACGCTGCGTTCAACAGCGGCCTTGTCAATATCCGCTCGTTGGAGAATAGGG
 AACGGACGGGCACAGTGCAGTGATTCACGATC

15A4 Forward (604 bases)

CCGTTGAGAACTCACTGCGTCTGGTGCGGGTTTTTGTTCGGTTTCTGCTC
TCCCCGAGTGTTTTTCTCTCATTTTATGATTTTTGATGCGCTTCTTTCTTAGTC
AGAGGATAAATGTCGTCACGAGTTGTTCTGTT**TGTGTGTGTGTGTGTGTG**
TGTGTGTGTGTGTGTGTGTGTGTGNNNGGAGAAGTTTTTATGTGTATCA
TAATCAATCCCTTTTCTCTCTCTTACAAGCCTTAAAGAATGCCCTTTCTCATTT
CTAGGGGTGGCCGGGCTGGGGGGAACAAGGAACGACTGGATAGATGGTTGT
TGAGAGGCCTACTATCTATAGGGCAGGCCGGATAGGAGTTTTCTATTACTGG
AAGGAAAGGTTGGAGTTCAAGGGATAGATGAAAGAGAAGGAAAGAGATTAG
TTCTGGTAGATTCGACCGATTCTCCGCAGCAAGATTTGTGTTGGTGTTACAAG
GAGATTCCTATTTCTCGAACCGACCCGGCCCCCACTATATAAAGGGCCCTCA
ACTATCCCCTACTATTCTATGGGCAATATTACCCCCGGGTGGGGGGGGCTCTC
CGGAACCTACCCTTAGAATAGTAGGGG

15A4 Reverse (766 bases)

GATCCAGCGTCGTATGATTAGTCCTGGACACAACGAAACACACAGCACAGTT
ACTCAACAGAAGCGCACAGAAAGGCAACACTAAGCCATGTCAATGAGCAAA
GTTAACAAACAAGTGTTTACTTGAATCACACGGTCCAGGTGCATGTTGGTCTC
GTAACCAGCTATTGATATTGTTTGCTGCACACAAGACAAAAAGAAACAATAT
CACATTTCAAAAACGGATTGGCACTCGGAACCTGAAAAAATGACTGTAAGAAA
ACAAGGAAATCAAATAAGAAACAGGCGTACCTGAGGAATCGATCGTTTGA
GACAAGATACGCAGGATGACAGGAGGCTGAATCGAACTGTGTTGTGTTGACA
ATGCCAGACGAGACTGCACCCCAACAACAATGTCCATACTTAGTTACAATTT
CTTACGATGCTAAATTGCGATTAAGCAAGAAAGCTTGAGGTGAAACGGACCC
AACGAGCAAACGGACAGATATATACATACACAACAGAGAAAAAAAAAAAAAG
GGTGAGCAAAGAGAAACAGAGAGCATAAAAGCAAAGCAGAGCCACACCGG
AACGAAGTGAGGCAGGATGCGTGTGTTTGCAGCCGCTGTCAGCGGAGCCGC
GTCGATGGTGTACCGCAGCGCCGACACCGGAGGATGCACCGACGACTGAAG
ACCGACGCTGCCGCTGCCAACCGTGTAAGTCGGCCAGCCACTGCGAGGGACAC
AACACGGACACGCGCACACACAACAAGCAAACCTAGAACTTCAAT

15B3 Forward (731 bases)

GATCGAAATTTAGTCTCGCGTACCCGGCCGCGGCTCCGGCTCTCCCACCCCC
CCCCCTTCTCTGCGGTACGACGTGCTCGGGGGGAAGCAGTGTGAGAAGGGTG
ATGAGAGGGAAGGAGAAGCTCGCTGGGGGAGATGTATCCACACGCGTGGA
AGTCGGGTGGTAAGGGAGCACACGTACACACACCCGTATAAATATACACAAA
TTCTTACAGAGGAATTTAGAGAGGATGTAACCGTTGCATACAGGACGAGTGT
ATCAGGGGACTAGGCGCGGAGGACGAGGTGGGAGAGGAAGCGGGGAGGCAT
GGTGTGGGGAAATGAAGAAGAAGAAGATGAGGAGGAGGAAGAAGAAGAA
GGAGGAGAAGGAGAAGGAGAAGGATGAGAAGGAGAAGCTAAGAACTC
GACAACTCCGACCTAGAGGGAGGACGAAGAAAAGGAAAAGGAAAAGG
AAAAGGAAAAGGAAAAGGAAAAGGAAAAGGAAAAGGAAAAGGAAAAG
AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA

TAGGTAGAGGGCCAACTAAGTTAGAGAACTATATGTTCAAATGAATGTAACA
AAAGGTGGTGTAGCTATATATAGAAAAAGCTAACCTAGGAACTTCACCAACA
GGATGTTTATTTTTGGCTGGATGGTAATGAATTTTTTTGAGACCAAATGGAGC
AAGGTGGGAAAATTGTACAGAGG

20B3 Consensus (608 bases)

CTTGATGATGATTAGTGGTCCACATTGTGGTAAGAATTTTTTTATTGTATTATT
ATATTATTGATTTATTAATGTTTTCAACTTGTAATTTTCCCTTTTTGTGTT
AGCATTCACTTGTTGCAAAACAATTA AAAA ACTATTGTGTGACTACTAAGGATT
TTGAGCATCACACTGTCATTTGAGTCTTGTTCTTTTTCAAAAAAATTGTTTGTA
GATACATTTGCTCCTTTATCTCACCTATTGTGTGTGTCATTTCGATATGTTAATA
CAA ACTCTCCTAGTCTTAGAAACATTCAATAGTATGCTTAAACATATGAAACA
AACAAGTCATAGAAACGACCCTAATTTGGCATTATATTA AAAA ACATATCATG
CACATTGTGATGCATCAATAAAACATTATGAACACACGCGACACACACACAC
ACACACACATGGATGCACACACACCCACGCGACGCGAGGCACATTCCCTCTC
CTCCATTCAATCCCCCTTTCTTCTAATCCCCCACTTCCTCCTTTGAAGCCATT
AAAACACACACGCGATGCACACACACACACACACTCTTTTAACGTTTAAAGTC
ACCCTAATTGCATAAAGAGTGAGAGATC

20C5 Consensus (530 bases)

GATCAATTTATATATACCTATATATATATACTAACGAACCCAAAACGAAC
CAAACCCACCCCCCAAACATTTAGCCGAACCGGCGAACCAAACCCCAAAC
CCGAACATA
TATATATACTAACAAACCCAAAACGAACCGAACCCACCCCAAAAAATTTA
GCCAAACCGGCGAACCGAACCCCGAACCGGAACCGGAACCGGTAACCTTAG
TTCCTACTATTTTTTATCATGTTTTGCAGTTCAAAGCCCATAGA ACTCTAGTTT
TCACCTGGATGCATCAGAAGTCTATTACTAAGCTTAATGTTTTGCAGTTCAA
GCCCATAGA ACTTTAGTTTTACCTGGATGGATCAGAAGTCTATTACTAAGTT
TAGAAAAAGGGAAAGCACTTCACTTACTGCGAGCATCCCAAAGGCGCACAGT
CTTATCTCCAGAGGCAGTAGCAACCAACTCAGAATGCTTTGGATCCCAGCAC
AATTGATC

20E5 Consensus (563 bases)

GATCGGTGAAAACCTTAAAAACACCATTTTTCAAATAAACCTCTTTTTTGGCTG
CCGCTGTTGTCACTCTTTGCTTTAACAGGGTAGTCTTCTCTTATGGGTCTTGAA
TTGCTATCCCTTGTTGAGCTTGTTGATATTATCTCCATTAATATAAATCAAATA
TTCTTTTATCCATTATTTCAAGCAGGGAAAAGCTCCTCTATGGATGCGGTTTG
GGAATTGTGGACACAAACCCCGAGCCCGCCCCCACATTCTCTTCCCCCTCC
CCTTCCGCCACCCGACACACACACACACACACACACATATAATCAAT
CTGTCATGGTATGTGTAAATATTTTTCCAATTATCTTATTGGTTTCTTATAGAG
TAAAATTTGTTCAATTGTCAGAATGCTTGCAGACAAGCTTAACATGAGCTATGA
AGAGGCAGAACGATGGATTGTAAACCTTGTAAGAACGCAAAGCTGGATGCT

AAGATTGATTGACATCAGGGACTGTTGTCATGGGAACAAACCACACAAGCA
TGTAAGTGGGCATACTGACCTCTTAATCGAGATC

21E3 Consensus (1008 bases)

TTGTTGGTGTGGATGCTGTTGCTGGTGCTGGTGCTGATGCTGGTGCTGATG
CTGGTGCTGATGCTGAGCGAACTGCGATGCAGGTACCGATTGGTTGTACGT
TGCGGACGTGTAGACGGGCTGATGGATGTCTGCAACGCACTACCAATCAGT
ACTAAGCTGCCGGGATGAGTCGAGTTCTCGACATGATGTGCTTGTGCTTGGCG
GGGAGTCGGAGAGAGGCAAGGGGAGGTGCAGGAAGGGGAGTGTACATGTAG
GTGTGTCGGAGGGGAGGTGTCGTGGGAGGA.....GAGGTTGTGTGCCCCCGCT
CCACCGTCTGCTGCCACAATCCGCAGTCTCTTTGCTTTTAGAGTTGCCGCTAG
CATCATACCATAACATTAGTGCACGATGCTTGGAGGCCACATCTCAGGCCACA
TAGCACGGTGAGTGCGGTTTCAGGCACCTTCAGGGCACTGCAAACCGGGGGAT
ACGGATCCACCGGGTGAGAAATCCCAAGTATACTCATTCTAGGGCTATGAAA
CTTCTAGTTTGTCTGTGGGGTGAGAAAGCCCAAAGGGGAGATACATATGCTA
GTTGATATGAACTCACAACACAAAGGTGACTATGGGGAAAAGTTGTGCCAC
TAGATGGGATTCCTGCTCAACCATATCAAAGTTATACCTCCTAATAACTTTAA
CAAGGAACCACAAGGAAAATTTTCAGTTCGTGCTCCATGAGACAAGAAAAGC
TCTGCATGCTTATATCATGAACAAAAAGGAAGGTACCATAGGCATGCAAGTT
GCATTAATATTATGCAAGGTGGAAAAAATGCAAAATGAAATGATGTAAAATG
TAAAAGATGTTCAAAGGAGTCTTGTTAAGTTCATGCAAGGTTAGTCACTTAG
AGAACATTGTGTCTAATGTCTCCAAGAGAGGAGTAATCAAAGACTAGAATAT
TCACAACCTTAGGGGATGCAACGCATTGCATCAAGCCACCAAGAGGAATGTAA
TATGCTAAGATAGATC

21E5 Consensus (1383 bases)

GATCTCGTAGAATGCCAGACTACTACGATTTCTGAATCATCAGAGTGGACCAC
GCCACCGGACAGCGGCCCATTTCTATCCAATAGAAACCACTCCTCCCAATCT
CTCACGCACATTAGTGAGAGGAGCGAGAGGGGGAGATTCATTACGTTCTCT
AAGAGATTCTGGAGAGACCGCATCGATAGATGGAATCTGCCCCCCCCCACC
ACCCTCTATAGTTGTTGCAAGCAATTAGAAGGGATAGCCAGCGGAGCTATGA
GAAGAATTCCCTTTTCGTGTCACACGCACACGAAACGGTCTCATCATGATCGTG
ATGCGGCTTTGCCCCCGCTACTTACTGTGCCAGTCGTGCTAACGGCCTTCCTC
GTTCCCCCGCTGCCGAGGCTGGCCAAAGCGGTGCGGGCGGCCCGGGTTAGAT
TCTAGAATGATGAAGCATCCCTGCGCGCTAGCAGCTTTAAGTGTGACTTATTC
CCGATAGCGATAGGTGCTATATTCTTCCCAAGCAGCGCCGGGGGTGAACAAA
CACAAGCCGATCCGTTAGGCGCTGGGTTTCGTAGAACACCGGTAGGGGGGGGA
ACACAACGAAAGTGCCACAACAACCACTACCCTCTTGAATTCATTCCCTCCTC
ACCGACTACTTGATCTTCTCGCGGCGGGGGCGGAGACTAGTTCGACTTCTGGT
CGGTGTGTCCTGTGTTGCTCGGCTCCACCACCCACATGATATTATCACCGCCT
CCCACCCAGGGGGGGGGTGGGGATATTAAGGTGCCTAAGTGGNAGGGCCCCG
GCCAATTCTATGGTAGGTGAGGTAGATGCCGCCCTACTATTCTAAGGGGGGT
GGGTGGATATAGTACTCGTTCGAGCCCAGCTACCTCAACACCCGACCCTCCC

ACCAACCTTCGATACGGGGTATCATCGATCGATCGAGGGAGGGATCAAGTTA
TGTGGTGAACCTACTAAGGTTAACTCAGGTATACTGCTACTCGAGTATACTCC
ATTGAAGCTACTCGAGCTCCATTGGTTAACTCGAGGAGTATACTTACTTACT
TACTTACTTACTTACTTACTTACTCCATTGAACTGGAAGAGAGAGAGCTCA
AAGCTGAACACCGAACCACAGCTTTGACACATTGGACACTGGACATGCTGCT
CCGTTTCATTGGAATGCCAGCTCCGCTAGGAGGGACACTAGTTAATTCCCCCCC
CACCCGGACCAATGTGCAAACCTTAAGACTCACCGGCTATAGATTGGAGGGG
CGCCTCCCTCGGCCGGTGTGATGGCTGCTCTCCCATTTCCAACCATTGCAGTG
ATTGGCGAGTGGAATCACAGGTAGTTTAGGTATACCCGACCGCAGCCCAAG
ATATACCCGTGACGCATTGGGATC

23E12 Consensus (686 bases)

GATCAAGTGCCACAGGGTGGAGTATGGGTGGAAGCAGCAGCAGCAGCAGC
AGCAGTAGCAGTAGCAGTAGAAGAAGAAGATGGAAAAGGAGGCTGGTCA
AAGAAAAGAGAAAATCGGAAATTCAAATTCGACGAAAATTTTCTGTTTTT
TTTTTTGGTTTTTTTTCTCTTTGGAATTGTACTGTACACAGAGAGTGCAAAAG
GGAATGCAGTACGGAGCTCAAGACGGTGTCTCTGGATCCATCCCCCTCAGG
TCCCAACGCCATGCCGTGTGCAGCGAGACCACATCATGGCCGGGGGAAGCAA
ACAGCTATAAAATCAAATACACGTGCCATATCCCCATGTTCTGCTTGTATTCA
GATGGGGCTACTCGCCCTTGAGTTCGCTCGCCGCACCGCTCTCTTCGGCTTCC
ACTGCGGCTGCCTCCTCCTTGGGTGCCGTGGTCACCTCGGCCGTACTGCTCTC
CTCTTTCTTCTCTTCCGCAACAACCTTCTTCTCTCCTCGGTGGTCACCTCGCT
GCTCCTGGGAGCAGCAGCCGCCGCCGCTCCGCTGCAGGCGGTGGTTCCGTC
GATGGGCTGGCCACAGCACCGGCTGCCTTGGCTTTCCGCCCATGAACAGACC
CACTTCGGTCCTCTGTACTTTCTGCCACAAGCGCTGGGTCCTTGACAAACACA
GGAAGATC

23F9 Consensus (574 bases)

ACCCAGCCGCTCAAATTTCCAAAAGAGCATTGAAGCCCTTTTGAAGTAGGAC
AGCCTGCTAACCACATTCTCTCTATTTCAAACAAATGTTCAAAGATGCAATTT
ATGTTGGTCCATCATCTAAACCCTGGAAAACATAAACTACACACTAGCCTCCT
TGATAAAGAGCACTCTAAGTTGAATATGATGATGGAGGACATGAGGCAGAAT
TGATTGAGGACAAGCTGCTATATTATCATGGATGGGTGGATTGATATTAGAC
ATTGGTAATTCATGAACATGATAGTCACATGAGGGCTCTTATATTTTATAGAG
CCATTTATTGACATATCTCTCTCTTTCCCTCTCTCTGTCTATAATCTATTAA
TAAATTTTGAaaaaATCACACACACACACACACACACACACACACACACA
CACACACACACACACACATATATATATATATGTATTTTGTATGTTGCCGT
CCCCTAAAAATTGCCCCTAGGAAACGTCCAGAAATGTGTCCCTAGCTGTCC
CCCCGTCTCGTCCCCGAAACGCCGTGGAACATAGCCTGTAGGTGAATTCG
G

[illegible]

GATCCCCCTTGAACAGCAAAGCCACTACTACTGTCACTCAATTCAACAGGAGG
AAACTCTCTGGAAGGACGTATGTTGAATCGCGGAAGACTCCCCGACCTTTCTC
TTCCTAAACTATTTCACTTATATATATATATATATATATATATATATATATG
ATTTTCCCCAACTTCAATCTACACTTCAGTGTTGCGTTGCTTGAGAAAAGGCA
GGCCAATTAGCATGGCATTAGAAATGAGGGAAGCAACTGGTTCAGATCTTCC
CTCATCGGTTTCTTGCTCCGACTTATGTAGCGGTGAAAGCGGGTCAAGCCCCT
GGATTCTTTCAAGTGCTTGACTCCCTACTCCACAGACTTGTAGCCCGACATGAA
ACGCTGATGTTTTCTTCCGGTAAACAGATTTCAAGAGAAAATCTGTGTGTTTT
CTCCGCCATGCAAGTGCTGCCATAAGCTTACAAAAGTACTTGGAGAGGTTG
TTCAAGTATGTTTCGCTGCAGTCCTTCCGCATTTGTTGTGGCCTATGTTTATATT
GATC

ACGGAGGAACATGTGACAAACACACACACACACACACATCTAATCTAAAG
ATAAAGAAGAGGCCTCCTCCTAACACTGATCGACCAAGAGTTCATCCACTTC
CCACTCTAGATCATCGTAGGGAGGAGGATGAGACTATAGTGATGGGATGTTA
AGAGGAGGAGCATTCGACCCTCCTGCAACAAGGGCTAAGGAAGAGGGATCT
GTAAGATTCTCTCCACCGAGCTCTTATGGAGCTAAGGTCAGCAATTTTGTGAA
GATCGCTGACAAGGCCGAAACGTCATCTAGAGAGTTAAACATGCACTTTACA
AGAGTGTGATGTTTTACCTCAACAGATGGGGAAGGTTGCACTCGAGGTTTAT
CGCTAGGCTCACCAAATGATTTAGGAGTGGTGACAAACTCAAGGGAAGCTAG
GACTGGTTATGTCACTAAGACAGAACAAATCATCACGGGGCCACCTTTCAA
GAATTTGGGCTCTTCTTACGTCCTCCATTGCTAAAAGGACCACACTTCACCTT

TTTCTTTGTTGGAGGAGGTAAACTCCCCACAGACAAGGCAAGAGGTGCGAAG
GAAGAGGAGGAAAGAACTGCAGTAGATACGACCGAGGTGTTTCATCATATCC
ACGAGCTTTAGGAAAGGCTAGACATCCAAAAAGGACATCAATAGAGGTGGC
TTCACCAAAGGAGTGATGAACACCTACAACCTTTTGGGGTCATATGCATTAC
GCTCATAGAAAAAGTTGGAAGAGATACTGAGTTGGTTTGAACGGAAGCAGA
GGGAAGAGAGGAGGACGCGACAAGAGGCCCTGATATAGGGGTGGTCATCTC
CTCTTCCTTCCTCTGTAAGGCGGCCTACTCAACCACGAGAGCCTTTTTTCAGC
ACTTACGCTCATGCACAAGGTAGTTTCTGCAACTCTTTAACTAGAGGACAA
AGGCTACAATGGAATGAAAGGCTCCTGCTCAGCCATGAGAGTGTCCACTACA
GTCATTTCTAGTGGAGGAACAACTCCTAAATGGTAGTAGGAGGGACAGGGT
GAGAAGCAGCAAGGGGAGTCGGCTCAAGTGCCTCGACAGTGGACTTCTTCCC
ATTCACTACTTTAGGAGTAGGAGTGGTAGGAGGCCACGGAGGTTGAAACAAA
GAGGACTCTCCATGGGATCATTCTTCCTATTATACCAAACGGGCCTAGGCAT
AGGAAAAACAAGAAGCTCAAACGTGGCACGCTTGGGCTTTTGCCAAGCAGGC
ACCATAGGATTCTCTCACAAATGCGAAATGGAATCTGATAGTGCCTCAACGG
TGTGAGAGGAAGCATCACATTGAACAATAGAAGAGGATGAAGGAGGATC

27H5 Consensus (539 bases)

CACGATCCTTTAAACATTTAATACATTGTGTATAAGCCTTTAATCAAAACAGG
CCATAGGCCAACAAATGACTGGACAGTTCAAACATTAACTAGGAGTTAAAA
ATATCTTTTTTTTGTAGCTATAACTAGTGATATATATATAGATAGATAGAT
AGATAGATATATAAGTAAAAAACTTGATTACAAGTTTAAAAGAGGCTAGT
ATTGTATTATTGTTCTACTAACGAATTCTAATTATTATCTTGGAATTAGTCCTT
TAGGCTAAGTTGCCGAATCGGGTACGGGCACGGGTACGGGTACGTGGGTACG
GTACGTGGGTACGGCAAAAAAATTTGGGGGCTTGGGTACGGAATCGGGTACG
TCCGTACAAAAATATAATAAAATATTTAAATATATATATATATATATCAACA
TAGAAACAAAATAAGTTTAGAATGTCATATACCAATACATATGATAAACAAAT
ACCATGACAAATTTCAATTTTTGAAACATAATATACTAATTTAAATTCGGTTT
TCAATATCAAAAGGATC

28E5 Consensus (672 bases)

TAAGTTTGTTCATTAGAAATGAATTAAGGGAGGGTGCGGTTGCTGTGCTGA
GACAAAAGTTTATGTTTAACTATTGAATGAGCCTACAACCTGAAGTAAATCG
CTTATAGCAGGAAAGCAAATTCATGCCAGTTTGAGGTAAGAATTTAAAATAG
AGGTTCAAGTTTAAAAGATTAAAATGAATTCATGATAAGAGCAAACAAATA
GGTAGAAGGGAGGGGGGTGCACTGCTCTCTAAGGGCGCAACCCTGTGAGGA
GGGTGCTCCACAACCAGCTAATGTGCAACTTCAGGTTGCAAATGTTTAAATATA
TTTTTTCTTTTATTTGTTTATATAAAATTCAAATCAATAACCTAAAACAACATA
ATTCTAATCTCTCTTTGGTTTCAAATAAAAATTACAGTGCAAAGACATTTAT
GTGCATGTATATATATATATATGTGTGTGTGTGTGTGTGTGTGTGTGTGT
GTGTGTGTGTGTTATTTTATATAATTGTATATATTTATGTAGAATTAGAAAATA
AGAGGTTTGTATATCTTTTCTTTAGGAAATACAAATTCTAATTGGGTTTTCTAC

39G9 Forward (630 bases)

39G9 Reverse (720 bases)

94

42B11 Consensus (419 bases)

**TCACAATTGGCTCTCAATCATTCGCTCGCTCACACACACACACACACAC
ACACACACACGAACAATCATCGCACCCCGTTCACTCGCAGACATTATCCACG
CCACGTCCTGGGCGAAACAAAAGCGCGCCAAATACCCCTCTTCTCCACCCAG
CGCTCGCTAGCAAGAGCATGCCCTTCGGTTGAACAACGAAAACCATGGTCAA
GGTCAGCATAAACGCCTCACCCCTTGCGCGCACACAGAACACCATACTGACAA
GCGCGACAGTTTCGCCGATCCATGAAAAGCAGCGATAAGCACGACTCGGCGA
ATCGACCGGTGCAGATATCCATCCCACACAAGGACGCGACCGCAATCGAGCC
CCCAAGGAAGGTCATCAAAGCGCTCTACGATTACAACGCGCCCATAGATCCC
CGGG**

APPENDIX I

PCR AMPLIFICATION OF GENOMIC DNA

Applied Biosystems, 47 Wiggins Avenue, Bedford, MA 01730 USA

(Modified by D. Bruno 2002)

Prepare the cycling reaction by adding the following reagents in the order they appear.

<u>Reagent</u>	<u>Quantity (Per reaction)</u>
Autoclaved sterile water	27.25 μ l
10X PCR II buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl)	5 μ l
dNTPs (100 mM each)	0.5 μ l (250 μ M each final conc.)
Forward primer	5 μ l (50 pmol)
Reverse primer	5 μ l (50 pmol)
AmpliTaq® Gold DNA polymerase	0.25 μ l (2.5 units/100 μ l)
MgCl ₂ (2.5 mM solution)	3 μ l (1.5 mM final conc.)
Genomic DNA	4 μ l (5-20 ng)
	<u>50 μl</u>

Thermalcycle in a GeneAmp® 9700 PCR instrument with the following parameters.

<u>Temp. (°C)</u>	<u>Time</u>	
95	10 min	} Cycle 30 times
95	30 sec	
60	1 min	
72	2 min	
4	Hold	

Run 25 μ l of the reaction on an agarose gel to confirm the amplification products then run on polyacrylamide gel to resolve fragments.

APPENDIX J

EXAMPLE OF FALSE POSITIVES FROM LIBRARY SCREENING

Listed below are two clones that were false positives. The repetitiveness and clustering of nucleotides can be seen.

2B2 Consensus (1070 bases)

GAATGTGGGACGTGGAGGGAGGAGGAGGACGACTGTCGGCTGTGACGAGAA
GAGGCTGTGTGAAGGAGACGTGGAGCTGGGGGGTGGTCTGAGTGTGGCTGA
GGACGTACTGAGTGGCGTCATACGAGTTGGGGGCGCCTGGCACCTCCATGTC
CTCAGCTGTTCGGTGTGGTGTACCTTCTGCTGCGTGGTGTGTTGCTGCTGCTGT
AGCCTGTTCTGCTCTCTGCGTAGTCTCTTCCCTTCTGCTGCTGCTGCAGCTGT
TGCTCCTCCTCTTGTGCGCTCGCATCTCAGCTCGTCTCCGCTCCTCAGTCACG
TCCACCACATCCGGACTGTCCCCTCTCGTCCGCCACTGGCCTCATCCGGACC
GCCGATCGGAATCACACGTCGTCGTTACCCTCACTCCTCTGCCTCTGTGCAC
TTCCGGACGGCAGCGCCTGCTGTTCGGTCGCTCGGCCAGTCAGCCGCTCCGTC
GTGCCGTCGTGCCTGTGCAGCACTGTGCTCTCCTCAGTGCTCTGACGGCCGTT
GCGGAAGGTGGTGGTGTGGTGGTCAACACACGCTGACCCTGCTCGTCGAAC
GAGGTGCGGCTCGACACGCTCACCGCATTTCGCACCACCTGTCAGCGAGGGCC
GTCCCATCCCCATCCCCAGCCCTCCCATGCCCATCCCCATGCCGCCCGCAGCC
AGCAGGCCCAATGGTCCAGCGCTGAACAGCGAAGGCCAGCGCCAGCATG
CCAAAGCCGAACGGCATAGGCGCAATGCCAAAGCCGCCCGCGTGAGGCTCG
TCGAACGGCGAAGCGAAGGGGTGCTTGTGCGCCGAAGAAGTCGCGGAAGATG
TCGTGGGCGAGCGGAAGGTGAACACGGGCTGTCGGAAGGCGAAGTGGCGTG
GCGGGCCGAAGGCCATGAAGGATGTGCTGCGAAGCCTGCACTGGCGCCGAA
GGGTGCAGTGGAGTGTGGCGGCTGCTGTGGGAGTGAGGTGGCGAGGGAGGT
GCGTCGAGGGCGTCATCGTACTCGCGCGTTTGTCTGTGTGCGGAGAGGACGG
AGAATGCTTCACTGTT

39G9 Forward (550 bases)

CAGAAGTTAGCTTAAAGGAGTTGGAGAAGGAACTGCAGAGGGATTCTATT
TTTGAAAAAGGAAGAAGGCTGTTGTGTTTGGATTTCGTCATCAAACAAAGG
ACTCCAGGTTCTTTTCGATTAATTTATTCTTTTGCTTTTGATAGTTCTGT
AAGGAAAAATGAAAATCATGTCTACCATATTCAAATCAATATCTATGTT
TTGAAAAAGGTTTTCCTGTTGAAAACCTTGTCTTGTGTTAAACCTTTAAAT
CTTCTGTAGAAAATTAGTGAGAAACCTTCTGATTTCTTCAATAAATCTA
GCTATGACATTTTTTATTTGAAAAGAAAATTGATTTTTTGTGGAAATTAA
AATTGTTTACCCTATTTTATTTTTTTTGTGGAGAAAAATTGGACTATGG
TTTCATTGTGTTTGTATTATTTATTTGTACTTTGTGAGCAAATATATAT
ATATACCTATATATATAAATATGTTTAGCATATATATGTATATATACATA
TACATATTCATATTTTATAATTATGTCTCTTTTGAAAAGAAAAATATAAT