CONTRIBUTIONS TO THE MOLECULAR BIOLOGY OF KELP

by

Michael Keith Liptack
B.Sc., University of Washington, 1991

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
in
the Department of Biological Sciences

SIMON FRASER UNIVERSITY
November 1999

ABSTRACT

Genetic relatedness between various kelp (Order Laminariales, Class Phaeophyceae, Division Heterokontophyta) taxa was investigated using DNA sequencing and PCR-typing. The rDNA ITS1 region of gametophytes generated by a naturally occurring apparent kelp hybrid of *Macrocystis* C. Agardh and *Pelagophycus* Areschoug were sequenced to determine parentage. All gametophytes examined had only *Macrocystis* rDNA suggesting either a non-hybrid, or more complicated hybridization than pure equal parental contribution occurred. Laboratory-generated intergeneric hybrids of *Alaria* Greville and *Lessoniopsis* Reinke were examined for parentage based on rDNA regions amplified using PCR. Both parental rDNA types were visible in one identified possible hybrid and non-hybrids were easily distinguished. Actin introns in both *Alaria* and *Nereocystis* Postels & Ruprecht were characterized and sequenced, representing the first actin intron sequences examined in the Heterokontophyta. The second actin intron from individuals of three *Alaria* species, spanning a geographic range of hundreds of kilometers, were sequenced to quantify variation and to examine individual relatedness for usage in studies of gene flow and population subdivision. Relatedness seemed to correlate with oceanographic distance but not with accepted species boundaries

Acknowledgments

I would like to thank the members of my committee for their support and reviews of this work. In particular, I would like to thank my senior supervisor, Louis D. Druehl, for his inspiration, support, and friendship, without which this thesis would have been impossible. Many other lab-mates shared in field collections, technical help, and general discusions on all manner of phycological topics, and to them I am eternally grateful: Ian Tan, Charlene Mayes, Handojo Kusumo, Andrew Swanson, and Ian MacKenzie. I would also like to especially thank Nikita Grigoriev, Barry Honda, and the late John D.G. Boom for all the intellectual encouragement, technical help, and friendship.

Much of this work was carried out while in residence at the Bamfield Marine Station. I would like to thank the Director, Andy Spencer, and the members of W.C.U.M.B.S. for material and scientific support during my studies at BMS. I also owe a huge debt to all the employees of the Bamfield Marine Station who helped me in innumerable ways: Nancy Christney, Heather Brooke, Cliff Haylock and the rest of the maintenance guys (I promise, no more Blazer repairs!!), Shirley Pakula, Linda Mather, Leslie Rimmer, Dave Hutchison, both Karls, and anyone else I might have forgotten.

Finally, I would like to thank my wife Elise for support, patience, and ever-present words of encouragement. I have no doubt, without her support I would not have made it. Thanks also to my daughters Hailey and Kyla for keeping me down to earth and my family for always believing in me.

Financial support for this study came from a Parks Canada grant through the Pacific Rim National Park as well as NSERC operating grants to Louis Druehl.

Table of Contents

List of Tables

List of Figures

General Introduction

This thesis explores aspects of kelp (Division Heterokontophyta, Class Phaeophyceae, Order Laminariales) relatedness using molecular genetics. Two general approaches have been taken to contribute to our understanding of kelp relationships: 1) molecular genetic analyses of putative kelp hybrids and 2) characterization and implementation of a novel genetic marker for resolving relationships of closely related kelp taxa.

This chapter provides a general overview of our present understanding of kelp evolutionary relationships.

Defining Kelp

Kelp are represented by 29 genera distributed among 5 families (Druehl et al. 1997). These genera are mostly found in cold waters of the Northern Hemisphere where they represent the bulk of multicellular photosynthetic biomass in shallow coastal waters. Kelp provide the architecture which houses and directly or indirectly sustains near-shore food webs (Duggins et al. 1989).

The conspicuous stage of the kelp life cycle is the sporophyte stage which may obtain lengths of up to 54 meters (Abbott and Hollenberg 1976). The sporophyte produces haploid meiospores through meiosis. These motile spores are released into the water, attach to solid sustrate, and then develop into dioecious

filamentous microscopic gametophytes. The gametophytes produce eggs or sperm. Upon maturity, the female produces a pheromone that attracts the sperm (Müller 1967) to the retained egg. The resulting zygote develops into the large conspicuous sporophyte.

Three families of kelp are commonly found in the northeast Pacific. Morphological features of the kelp sporophyte distinguish these families. The Alariaceae is characterized by the absence of branching and the presence of sporophylls. The Lessoniaceae is characterized by regular branching and in some cases the presence of sporophylls. Finally, the Laminareaceae is characterized by the lack of sporophylls and no branching.

Kelp Dispersal

The effectiveness of kelp dispersal partially dictates the degree to which kelp taxa, traditionally distinguished based on morphology, may be genetically related. Therefore, an understanding of dispersal is important to appreciating the process of gene flow and hybridization.

Because of both the heteromorphic alternation of generations and the dispersal of only the small non-sexual haploid meiospores, kelp do not easily fit present dispersal models based on wind-born pollen, seed banks, or motile sexual individuals. Only the haploid meiospore stage is known to disperse, while the haploid gametophytes and diploid sporophytes remain sessile. Both the two-dimensional distribution of their niche as well as the sessile nature of their sexual stages favor an

isolation by distance model of evolutionary processes, heavily dependent upon the dispersal of meiospores.

Because of the heteromorphic alternation of generations in kelp, reproductive cells and propagules experience a different set of evolutionary challenges when compared to the sporophyte stage. Nearly all studies in the Laminariales concentrate on the large visible sporophyte stage, such that the biology and ecology of microscopic meiospore and gametophyte stages are poorly known (Santelices 1990). Motile haploid meiospores that settle to form haploid gametophytes are released in massive quantities from each individual (Chapman 1984). Meiospores only seem to drift in the plankton for about 24 hours and recruitment is very spotty (Reed 1990). Although the 10 μm long kelp meiospores are chemotactic and flagellated, their motility is believed to be only effective at the microscopic level in recruitment site decisions (Amsler and Neushul 1989).

The most common hypothesis regarding long distance kelp meiospore dispersal effectiveness states that rare drifting fertile individuals inoculate new sites. An alternative explanation is that kelp forests produce many spores that can disperse over a wide range and therefore frequently re-colonize at extensive distances from parental stands. Determining the validity of either of these hypotheses is difficult because the studies undertaken so far on actual kelp recruitment have only shown recruitment to occur at distances of about five meters (e.g. Anderson and North 1966, Druehl 1981).

Kelp Hybridization

Hybridization, the crossing of two individuals from different populations, species, or genera has been an important tool in understanding evolution and taxonomy. In practice, hybridization is often used as a judge of relatedness (Lewis 1996b). Modern versions of the biological species concept (Dobzhansky 1937, Mayr 1963), or its various corollaries (Mallet 1995), often are at least partly defined by the ability to form hybrids. Certain groups of organisms, such as some tracheophytes, seem to easily form natural hybrids among related forms (Dressler 1981), whereas hybridization is much rarer in animals (Levin 1979). However, hybridization is a complex process, with a myriad of causes resulting in the lack of development, sterility, or death of the hybrid (Lewis 1996b).

Hybridization has been examined at several levels in the Laminariales. Intraspecific and interspecific hybrids seem quite common in some genera (Lewis 1996b). Intergeneric hybrids have also been identified in the field or produced artificially by crossing (Sanbonsuga and Neushul 1978, Coyer et al. 1992, Lewis and Neushul 1995). Even hybrids between members of different kelp families have been reported (Tokida et al. 1958, Cosson and Olivari 1982). The possibility that true hybridizations might occur between higher taxa is supported by the observation that members of the three common kelp families (Alariaceae, Lessoniaceae, and Laminareaceae) display sperm chemotaxis based on the same pheromone (Müller et al. 1985).

Recognition of hybridization in phenotypically plastic organisms such as kelp is challenging. Characters, previously used for taxonomic separation such as bullae, have been shown to have mendelian inheritence in laboratory crosses between species (Bolton et al. 1983). Maintaining pure stocks and verifying the parentage of laboratory crosses remains challenging. For example, laboratory rates of parthenogenesis in unisexual cultures have been reported as high as 25% (Chapman 1974).

Naturally occurring apparent hybrid individuals with morphologies that show evidence of both *Macrocystis* and *Pelagophycus* are occasionally found in southern California (Coyer et al. 1992). In this thesis, I examine the gametophytes from an apparent *Macrocystis* x *Pelagophycus* hybrid to determine whether rDNA sequencing could distinguish the presence of both parental types in the gametophytes.

Laboratory-generated hybrids, where male and female gametophytes are mixed *in vitro,* suffer from difficulties of distinguishing individuals generated through sexual union from uniparental individuals generated from apogamy, parthenogenesis, or androgenesis. In this thesis I also demonstrate a method using PCR to determine the parentage of individual blades obtained in attempted *in vitro* hybrid crosses.

Kelp Relatedness

Our understanding of evolutionary relationships among extant kelp taxa is hindered by the absence of a significant fossil record and limited molecular studies, particularly at the inter- and intra-species levels.

A major dichotomy exists between interpretations of relationships based on the traditional morphology of kelp and recent molecular evolutionary data (Druehl et al. 1997). Morphological studies have always proved difficult because of the wide morphological plasticity of apparently conspecific individual plants (Sundene 1958). The recent interest in molecular-based phylogenies is in part also fueled by a desire to avoid some of the problems of human subjective interpretations of phenotype.

Relationships at the Family level based on rDNA phylogenies in kelp show the greatest divergence from morphological phylogenies (Druehl et al. 1997). Many of the traditionally used Family-defining morphological features, such as sporophylls, splitting of the thallus, and midribs seem to create the most problems for reconciling the morphological and phylogenetic data (Fain et al. 1988, Druehl and Saunders 1992). Kelp families based on nuclear rDNA phylogenies (ITS1 and 5.8S) as well as chloroplast RFLPs seem to suggest three main groups (Druehl and Saunders 1992, Druehl et. al 1997). Group 1 contains *Alaria* Greville, *Lessoniopsis* Reinke and *Pterygophora* Ruprecht (members of the traditional Lessoniaceae and Alariaceae respectively) and is potentially distinguished morphologically by a single midrib and the presence of sporophylls. Group 2 contains *Costaria* Greville and *Dictyoneurum* Ruprecht (members of the traditional Laminariaceae and Lessoniaceae respectively), typified by a lack of sporophylls. Group 3 contains the rest of the kelp genera with many unresolved relationships and unclear morphological trends between the genera.

Members of the Laminariales are unique model organisms for studies of recent population-level evolutionary processes. Kelp exist in essentially a two-

dimensional ribbon of habitat bounded by the photic zone and high tide. In practical terms, kelp are found in temperate waters usually less than 30m deep.

Attempts to understand the genetic relatedness of individuals from a certain species or between species can be very difficult if there are essentially no well-established heritable traits. Even those traits that are visible pose conflicting challenges because of the need to examine the genotype, not the phenotype when determining relatedness or phylogeny. The kelp (Laminariales) are such taxa where poorly understood natural variation necessitates a different approach. Molecular genetic techniques, which have been used to study families, genera, and species in the brown algae (Tan and Druehl 1995, Druehl et al. 1997), are excellent tools for both phylogeny and relatedness. Various molecular techniques can examine the genotype directly and the mechanics of the origin of certain molecular variable traits are at least partially understood.

The ribosomal cistron, encompassing multiple repeats of three of the subunits that form ribosomes, has a rich legacy of use in studies at many levels of relatedness. Taxa from deep roots of the evolutionary tree to subspecies have been successfully characterized using rDNA. Each repeat contains regions that vary drastically in the average substitution rate and therefore can be applied to studies on different levels. Because of the multicopy nature of rDNA, concerted evolution tends to homogenize the copies, often numbering in the thousands, carried within individuals of a species. To examine relatedness at a population level or within a species, markers are needed

that minimize the masking effects of processes other than identity by descent. Selection, reversals, hitchhiking, drift and genetic bottlenecks all can hide the relatedness between panmictic individuals. For these reasons, genetic regions with phenotypic expression are not typically used in population level studies. Silent sites or synonymous sites are a notable exception to this rule though they utilize regions within individual codons that can vary without affecting the phenotype. Both cpDNA and mtDNA from chloroplasts and mitochondria respectively have been used for population-based studies, particularly in mammals and other metazoans. The use of organelle DNA in many of the less studied groups such as the brown algae is more difficult however because of a lack of understanding of their inheritance (Bisalputra et al. 1971). The tools left for study of relatedness between individuals are therefore narrowed to variable markers of unknown heritage or identity (AFLPs, RAPDs, etc.) and non-coding regions believed to be mostly inherited under the neutral model of evolution.

In this thesis I characterize the second intron of actin in a member of the Heterokontophyta for the first time, opening oportunities to uncover individual relatedness among selected local patches of kelp through a comparison of actin intron sequence data. DNA sequences of actin introns in three recognized local species of *Alaria* as well as portions of actin introns from *Nereocystis* are demonstrated and analyzed. Relationships between individuals based on actin intron sequences are also discussed.

Chapter I

ITS1 nrDNA Sequences of *Macrocystis pyrifera*, *Pelagophycus porra*, and a *Macrocystis* x *Pelagophycus* Hybrid.

Introduction

Macrocystis x Pelagophycus hybrids

Individuals plants with a morphology between *Macrocystis* and *Pelagophycus* are probably the best known examples of naturally occuring putative intergeneric hybrids in the Laminariales. Large sporophyte hybrids of *Macrocystis* and *Pelagophycus* have been grown in the lab (Sanbonsuga and Neushul 1978), and putative hybrids occur in the field (Neushul 1971, Coyer and Zaugg-Haglund 1982) with some regularity between beds of *Macrocystis pyrifera* and *Pelagophycus porra* off Santa Catalina Island, California (Figure 1). Both the laboratory-generated and naturally occurring *Macrocystis* x *Pelagophycus* hybrids have morphologies containing taxonomically distinctive elements from each of the parental taxa including: split stipes (a *Pelagophycus* character), small distal floats on each blade (a *Macrocystis* character), and rugous blades (a *Macrocystis* character).

A fertile individual *Macrocystis* x *Pelagophycus* hybrid, from which spores were released and gametophytes produced, was collected in the field by Coyer et al.

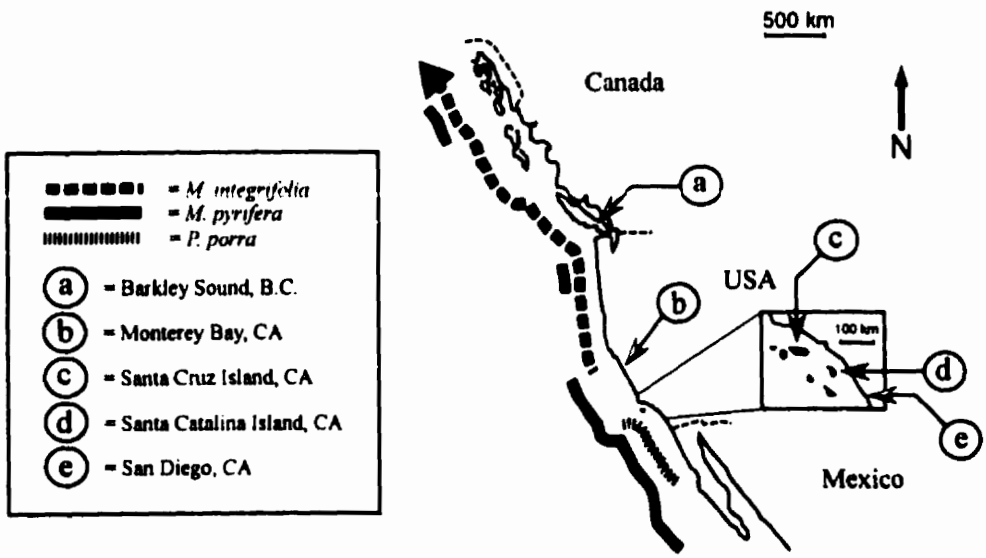


Figure 1. Species ranges and locations of samples. Only northeast Pacific species ranges excluding the Aleutians and south Alaska are shown (Womersley 1954, Druehl 1970, Abbott and Hollenberg 1976, Lewis and Neushul 1994, O'Clair et al. 1996).

(1992). Lewis and Neushul (1995) postulated that the fertile hybrid was an allopolyploid based on backcrosses and chromosome counts. Although the precise haploid chromosome number in these taxa is somewhat unclear (for review see Lewis 1996), the hybrid *Macrocystis* x *Pelagophycus* sporophytes seemed to have twice the haploid chromosome number (n=28-32) of the parental type's complement of about 16 haploid chromosomes (Lewis and Neushul 1994).

ITS1 nrDNA

In an effort to identify parental genomes in offspring and parents, a sequence-based analysis using nrDNA was undertaken. Specifically, the ITS1 region was examined because of the available data set encompassing most kelp genera (Saunders 1991) and because ITS1 regions have proved to be good markers in hybridization studies (Baldwin et al. 1995, Buckler and Holtsford 1996).

The ITS1 region appears as part of a tandemly repeated unit approximately proportional to genome size, of usually several hundred copies in most eukaryotes (Li 1983). The multicopy nature of nrDNA along with the homogenizing effects of concerted evolution allow for easier amplification of a large number of templates for direct sequencing (Hillis and Dixon 1991).

Materials and Methods

Culture Maintenance and Sources

The *Macrocystis* x *Pelagophycus* hybrid gametophytes and *Macrocystis pyrifera* gametophytes were provided by Dr. Sandy Benson at Neushul Mariculture Inc. (P.O. Box 1416, Goleta, Ca 93116). The *Pelagophycus porra* genomic sporophyte DNA as well as a preliminary ITS1 sequencing gel photo was provided by Dr. Gary Saunders (Department of Biology, University of New Brunswick, Fredricton, NB, Canada). Gametophyte strains used in the study (Table 1) were kept under 150 $\mu mol \cdot m^{-2} \cdot s^{-1}$ cool-white fluorescent light (16:8 LD photoperiod) at 13°C in 1X f/2 media (Lüning and Dring 1975). GeO_2 was added to a concentration of 500 $\mu g \cdot L^{-1}$ in cultures showing signs of diatom contamination (Chapman 1973).

DNA Isolation

To isolate DNA, a modification of the standard Chelex 100™ (BioRad #143-1441) extraction procedure was used (Walsh et al. 1991). Gametophyte tissue was briefly resuspended in 1 mL of sterile distilled water in a 1.8 mL microcentrifuge tube and collected at 16,000 g in a microcentrifuge for 30 seconds. If necessary, tissue was removed to bring the volume of the pellet to approximately 20 µL. 180 µL of a fresh 5% Chelex 100™ solution (w/v in sterile ddH_2O) was added immediately after vigorously resuspending the Chelex™ beads. The mixture was briefly vortexed and then incubated at 56°C for 90 minutes with occasional gentle mixing approximately

Table 1. Species, strain, general morphology, isolation locale, and culture source of gametophytes and sporophytes from which sequences were either generated or obtained . See map in Fig. 1 for locations.

Species	Strain	Morphology	Isolation locale
Pelagophycus porra	-	sporophyte	San Jose Is., CA[a]
Pelagophycus porra	Pp-DF#1	female gametophyte	Santa Cruz Is., CA[a]
Pelagophycus porra	Pp-DM#8	male gametophyte	Santa Cruz Is., CA[a]
Macrocystis pyrifera	Mp-AF#2	female gametophyte	Santa Catalina Is., CA
Macrocystis pyrifera	Mp-AM#3	male gametophyte	Santa Catalina Is., CA
Macrocystis x *Pelagophycus*	M/P-F#2	hybrid female gametophyte	Santa Catalina Is., CA
Macrocystis x *Pelagophycus*	M/P-M#1	hybrid male gametophyte	Santa Catalina Is., CA
Macrocystis integrifolia	1BS	sporophyte	Barkley Sound, B.C.
Macrocystis integrifolia	2BS	sporophyte	Barkley Sound, B.C.
Macrocystis integrifolia	3BS	sporophyte	Barkley Sound, B.C.
Macrocystis integrifolia	1MB	sporophyte	Monterey Bay, CA
Macrocystis integrifolia	2MB	sporophyte	Monterey Bay, CA
Macrocystis integrifolia	(Saunders 1991)	sporophyte	Barkley Sound, B.C.

[a] = collected from drift

every 15 minutes. After incubation, the mixture was vortexed for 10 seconds, and placed in a boiling waterbath for 8 minutes. The mixture was spun for 2 minutes in a microcentrifuge at 16,000 g. Finally, the supernatant was carefully moved to a new microcentrifuge tube to avoid transferring any Chelex™ beads and stored at –20°C.

PCR Amplification

BC1 (5'-GAT TCC GGA CTG TGG CTC GCG-3') (Saunders 1991) was used as the forward primer and BC2 (5'-CGA GTG GTG TCA ACA GAC ACT CC-3') (Saunders 1991) was used as the reverse primer in a Polymerase Chain Reaction (PCR)(Kleppe et al. 1971, Saiki et al. 1988) to generate templates for both single-and double-stranded sequencing. 25 µL reactions were used with 8.5 µL sterile H_2O, 2.5 µL 10x PCR buffer (500mM KCl, 100 mM Tris · Cl pH 8.3, and 0.1% gelatin), 2.0 µL 25 mM $MgCl_2$, 1.25 µL (200 pM) of each primer, 4.0 µL of dNTPs stock (1.25 mM each), 5.0 µL of template , and 0.5 µL of *Taq* polymerase (Perkin Elmer-Cetus). A modified 'hot start' (Erlich et al. 1991) procedure was used wherein the polymerase was added by pausing the thermocycler at the initial 70°C step. PCR conditions were an initial melting cycle (95°C for 5 minutes, 70°C for 1 minute, then 72°C for 30 seconds) followed by 27 'touchdown' (Don et al. 1991) amplification cycles (93°C for 45 seconds, 68°C for 1 minute with a 0.5°C decrease each cycle, then 72°C for 45 seconds), and finally a longer extension step (72°C for 5 minutes). Reaction products were stored at -20°C upon completion.

PCR reactions were divided into two fractions and run in duplicate sets on alternate lanes of 15 cm long 0.9% agarose/TAE (40 mM Tris-Acetate, 1 mM EDTA pH 8.0) gels (Sambrook et al. 1989). A haphazard cut was made between the two replicate sides of the gel and one half was post-stained for 15 minutes in a 250 ng · mL^{-1} Ethidium Bromide/TAE bath, and then destained in ddH_20 for 10 minutes. Appropriate bands identified under 260 nm UV illumination were marked with notches so that the corresponding bands in the other half of the original gel could be located. Bands from the non-stained side of the gel were carefully excised and then purified with the Sephaglas™ Band Prep kit (Pharmacia) using the manufacturer's protocol. To make sure the entire band was excised, the remaining portions of the gel were post-stained as above and viewed under UV illumination.

Four separate PCR reactions were done for each primer pair and for each individual. The four reactions were used for asymmetric PCR-based sequencing (Nichols and Raben 1994) using two separate sequencing kits (see below), and two rounds of DMSO-based double-stranded sequencing (Tan and Druehl 1994), respectively.

DNA Sequencing

Double-stranded templates were sequenced directly (Fuller 1992) using a DMSO-based method (Tan and Druehl 1994) for the denaturing of the template and the Sequenase™ version 2.0 (USB/Amersham) sequencing kit protocols for all subsequent steps. 7.5 µL of template, 0.5 µL (2.5 pM) of primer, and 1 µL of DMSO

were combined and brought to 95°C for 4 minutes. The reaction was then immediately quenched in a dry ice/ethanol bath for 5 minutes or put in a -80° C freezer for 20 minutes. The tube was quickly thawed at 37°C and briefly spun in a microcentrifuge to collect the contents at the bottom of the tube. 2 μL of Sequenase™ Reaction buffer was then added and the reaction was left at 20°C for 10 minutes. The extension and labeling steps were as described in the Sequenase™ 2.0 kit with the exception that 1 μL of ^{35}S dATP was used for the labeling step instead of 0.5 μL as described in the kit.

Single stranded templates for direct sequencing (Fuller 1992) were prepared using asymmetric PCR (Nichols and Raben 1994). Reactions consisted of 11.5 μL sterile ddH_20, 2.5 μL of 10x PCR buffer (500 mM KCl, 100 mM Tris · Cl pH 8.3, and 0.1% gelatin), 2.0 μL 25 mM $MgCl_2$, 4.0 μL primer (20 mM stock), 4.0μL dNTPs stock (1.25 mM each), 0.5μL of the original PCR as template, and 0.5 μL *Taq* polymerase (Perkin Elmer-Cetus). A modified "hot start" procedure (Erlich et al. 1991) was used as above with the following conditions: an initial melting cycle (95°C for 5 minutes, 70°C for 1 minute, then 72°C for 30 seconds) followed by 27 "touchdown" (Don et al. 1991) amplification cycles (93°C for 45 seconds, 68°C for 1 minute with a 0.5°C decrease each cycle, then 72°C for 45 seconds), and finally a long extension step (72°C for 5 minutes).

Reactions were checked on 15 cm 0.9% agarose/TAE gels post-stained as above to determine the best single-stranded product yield. Templates were purified

from the reactions using three separate 200 μL ddH_20 rinses in an Ultrafree-MC™ 30,000 MW cutoff spin-filter (Millepore) until only 7 μL remained.

Single-stranded sequencing reactions were done using the purified templates with both a Sequenase™ 2.0 kit (USB/Amersham) and a T7™ kit (Pharmacia) using the protocols for single-stranded templates provided with the kits. In the case of the Sequenase™ kit, the initial primer annealing was done at 74°C and 2 pM of primer was used. In the case of the T7™ kit, 1.8 pM of primer was used.

Sequencing reactions were run on glycerol-tolerant 6% polyacrylamide (19:1 acrylamide:bis-acrylamide), 8.3 M Urea, 1 X TTE (89 mM Tris · Cl, 29 mM Taurine, 500 μM 2Na · EDTA pH 8.0) gels in 0.8x TTE buffer. Gels were then soaked for 25 minutes in 10% glacial acetic acid/10% methanol, dried onto Whatman 3MM paper in a vacuum gel drier at 80°C, and exposed to PDB-1™ X-ray film (Kodak) for three to five days.

In all, a total of three separate runs of *P. porra* gametophytes and four separate runs for each of the two *M. pyrifera* and the two *Macrocystis* x *Pelagophycus* hybrid gametophytes were done in each direction to determine sequence identity. NC-IUB base nomenclature was used in all cases (NC-IUB 1984). Resulting sequences were deposited in GENBANK.

Sequence analysis

Consensus sequences and manual alignments were done by eye using the sequence editor ESEE (Cabot and Beckenbach 1989). Sequences generated from five

Macrocystis integrifolia Bory individuals (Mackenzie 1997) along with one individual from Saunders (1991) were used to construct the composite sequence of *Macrocystis integrifolia* for the alignment. Each sequence was compared to all other known Laminariales sequences from this region (Saunders 1991), and aligned to the sequence of *M. pyrifera* or the hybrid. Phylogenetic trees and bootstrapped data sets were generated with PHYLIP 3.5c. (Felsenstein 1993) using outgroup sequences from Saunders (1991).

Results

ITS1 Sequences

In total, nuclear ribosomal DNA (nrDNA) aligned sequences were generated for 64 basepairs from the 3' end of the 18S subunit, 291 basepairs of the ITS1, and 106 basepairs of the 5' end of the 5.8S subunit (Figure 2). Two *Macrocystis* x *Pelagophycus* hybrid gametophytes as well as two *Pelagophycus porra* gametophytes (GENBANK #U65084), one *Pelagophycus porra* sporophyte (GENBANK #U65084), and two *Macrocystis pyrifera* gametophytes (GENBANK #U65083) were sequenced. In all cases, the sequences generated from each gametophyte or sporophyte of the same species were identical.

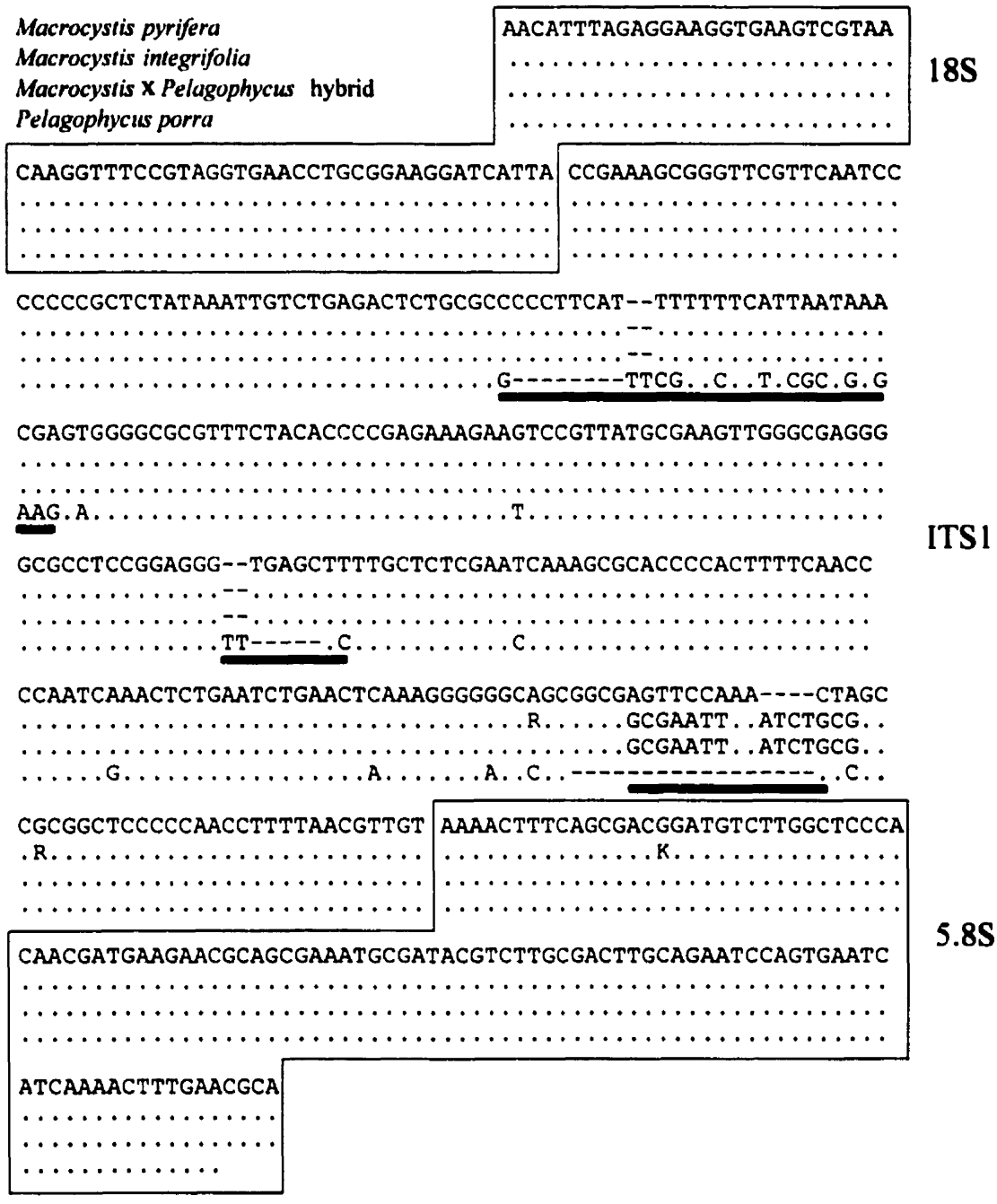


Figure 2. Nuclear ribosomal DNA (nrDNA) sequence alignment including regions of the 18S, ITS1, and 5.8S regions (see text). Sequences are aligned to *Macrocystis pyrifera*. Dots (.) indicate identity with the *Macrocystis pyrifera* sequence above, dashes indicate the location of insertions or deletions compared to *Macrocystis pyrifera*. Boxed areas delimit the 18S and 5.8S subunits and the unboxed region is the ITS1. The underlined regions were either too variable or too ambiguous to align using Saunders (1991) and were not included in phylogenetic tree generation.

Sequence Alignments

The generated sequences were aligned to most of the known kelp ITS1 sequences from Saunders (1991) including: *Alaria marginata* Postels & Ruprecht, *Lessoniopsis littoralis* (Tilden) Reinke, *Pterygophora californica* Ruprecht, *Egregia menziesii* (Turner) Areschoug, *Eisenia arborea* Areschoug, *Postelsia palmaeformis* Ruprecht, *Nereocystis luetkeana* (Mertens) Postels & Ruprecht, *Dictyoneurum californicum* Ruprecht, *Dictyoneuropsis reticulata* (Saunders) Smith, *Lessonia nigrescens* Bory, and *Costaria costata* (C. Agardh) Saunders. Regions of generated sequence with potential ambiguous alignment and regions of poor alignment in Saunders (1991) were removed from the data set for tree generation (Figure 2). The *M. integrifolia* ITS1 sequences generated by both Saunders (1991) and Mackenzie (1997) differed at only three sites (Figure 3), two G/A transitions and a single transversion in the 5.8S. Both variants of these two sites were seen in the Monterey Bay as well as in the Barkley Sound populations of *M. integrifolia.*

Phylogenetic analysis

For maximum parsimony trees both multiple deletion events and deletions beyond one base long were removed. For neighbor joining trees *Postelsia palmaeformis* and *Egregia menziesii* were included as outgroups and encompassed a total of 513 basepairs. These two species formed a clade with *Macrocystis integrifolia* based on the ITS1 assessment in Saunders (1991). Maximum parsimony trees were generated from a dataset of 413 basepairs and included three informative

Figure 3. Alignment of 3' end of 18S, ITS1, and 5' end of 5.8S nrDNA from a number of Laminariales genera. Dots (.) indicate identity with the *Macrocystis pyrifera* sequence. Dashes (-) are spaces inserted for alignment purposes. All sequences except *Macrocystis pyrifera*, *Macrocystis* x *Pelagophycus* hybrid and *Pelagophycus porra* are from Saunders (1991). In addition, the alignment itself is based heavily on that in Saunders (1991).

```
Macrocystis pyrifera                  AACATTTAGAGGAAGGTGAAGTCGTAACAAGGTTTCCGTAG
Macrocystis integrifolia              ....................................
Macrocystis x Pelagophycus hybrid     ....................................
Pelagophycus porra                    ....................................
Alaria marginata                      ....................................
Lessoniopsis littoralis               ....................................
Pterygophora californica              ....................................
Egregia menziesii                     ....................................
Eisenia arborea                       ....................................
Postelsia palmaeformis                ....................................
Nereocystis luetkeana                 ....................................
Dictyoneurum californicum             ....................................
Dictyoneuropsis reticulata            ....................................
Lessonia nigrens                      ....................................
Costaria costata                      ....................................
```

```
GTGAACCTGCGGAAGGATCATTACCGAAA------------GCGGGTTCGTTCAATCCCCCCCGCTCTAT
............................------------............................
............................------------............................
............................------------............................
..........................CA-----------C..A..AAAT---GGA----..T.-....A
..........................CA--------AAAA..A..AAAC---.GA----..T.-....A
........................GCAATCATCCCCCAA..A..AACT---.GA----..T.-....A
............................-------------..............A.............
............................------------..............A.............
............................------------..............C.............C
............................------------..............................
..........................CA------------.....AGGT....TATA--......T..C
..........................CA------------.....AGGT....TATA--......T..C
............................------------..............A.............C
..........................CA------------.....AGGT....TATA--......T..C
```

```
AAATTGTCTGAGACT-CTGC-GCCCCCTT-CAT-------TTTTTTCATTAATAAACGAGTGGGGCGCGTT
..............-....-........-...-------.....................................
..............-....-........-...-------.....................................
..............-....-..G--------------TTCG..C..T.CGC.G.GAAG.A..........
...A---..C.T...G....C.TTTGA.-----------------------------AC....GTC.
...A---G.-.T...G....CAGGT....AA-----------------------------.GAC....GTC.
...A---..-.T...G.C..T.T.CCTACGGTCGAGACTC....GAGC.GCCGGGAGTT..AC....GTC.
..............CG....C..TGGA-------------------------------------...CGG
..............-GTC..C.TTTGTAAC.TC--ACCTTA...AATTAATTA..C....G.A......CG
..............-....-...GAT..CAC---------------------------.A.A........
..............-....-....GAC.CAC-----------------------------G.A....GC..
.........CG.----.T..C.T.ATGAAA-------------------------------.A......A..G
.........CG.-----..R.C.T.ATGAAA-------------------------------.A......A..G
..........T...-GTC..C.TG--GAAA.TCCC---------------TTGG.GGCGAGC.A...ATCG
.........CG.----.CA.C...ATGAAA---------------------------------.A......A..G
```

```
TCTACACCCCGAGAAAG-----AAGTCCGTTATGCGAAGTTGGGCGAGGGGCGCCTCCGGAGGG-------
.................-----.....................................-------
.................-----.....................................-------
.................------..T.....................................TT-----
.G..A............------..AAT.....................................TTTTGTT
.G..A............------..AAT.....................................-------
.G..A............GA---..AA......................................GGGTTTT
C..........C.....------...TG..................T....A........A.TGTCGTG
....T............AATCTTTT.TT..................................A.AGTTGTT
.................------.AG.....................................-------
.................------GA.T...................................A.-------
GTC.T......C.....------..A.T...................................TTTTGGT
GTC.T......C.....------..A.T...................................TTTTGGT
....A............----TG.AG.....................T................TCTTTTC
GTC.T......T....T-----G.A.T.........................T..........TCTTCGC
```

```
-----------------------------------------------------------------------TGAG
-----------------------------------------------------------------------....
-----------------------------------------------------------------------....
------------------------------------------------------------------------
TA----------------------------------------------------------------------
------------------------------------------------------------------------
ACTACTACTACTGTTGTAG-----------------------------------------------------T.A
GTTGTCGTCGTCGTCGTCTTGGACGGCGGCGGCGGC-----------------------------G.C.
T-----------------------------------------------------------------C.
-----------------------------------------------------------------------CA
-----------------------------------------------------------------CTCA.T
CATTTCCTTGGTGACTTGGA-------------------------------------------CT..GA
CATTTCCTTGGTGACTTGGA-------------------------------------------CT..GA
TTTTCCCTTTCCCCTTCTTGTCTCTTTTCTTCCCTTTTCGGGGGATTTACGGGACGGGACGGGAAGGGTG.
------------------------------------------------------------------------
```

```
CTTTTGCTCTCG-AATCAAAGCGCACCCCAC-TTTTCAACCCCAA---TCAAACTCTGAATCTGAACTCAA
...........-................-...........---.....................
...........-................-...........---.....................
-.C........-..C.............-...........---..G.................A..
T.GAAA.C....-...............---.........----CT....................
T.AGGT......-...............---.........----.T....................
AG.C.T.A....-...............--..........----CT....................
GCGGC.......-..AA...........CA-.........TTTTTC...C.................
.AC.CT......-..C............ACA.........G.---.G....................
.GCCA.--....C..C............A...........----C.......C.............
.AAGC.......-..C............A-..........----.T.................A..
..GC........-..C............A-..........----.T....................
..GC........-..C............A-..........----.T....................
AGGAA.......-..C............-...........----.T....................
--.AGA......-..C............ACA.........----CT....................
```


AGGGGGGCAGCGGCGAGTTCCAAA-------CTAGCCGCGGCTCCCCCAACCTTTTAACGTTGTAAAACTT
.....GCGAATT..ATCT---GCG.....
.....GCGAATT..ATCT---GCG.....
.....A..C..----------..C.....
.....CTC-T..----------A.CG.G.....T.....-.....
.....CTCCT..CTTTTTTCGAAAGC.CG.....-.....
.....----------..TCACTG...T.....
.....G.GCCGCTC----------GT..GCGC..CT.....
.....CCTACTACTCC------CTCACA.GG.G..G.....A.....
.....C...CTC----------.....-.....
.....nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn.....
.....AGG.----------T.TTTTT.CTTC.....
.....AGG.----------T.TTTTT.CTTC.....
.....CT.----------T.....-.....
....A.....CTC----------G.C.....C.....-.....
TCAGCGACGGATGTCTTGGCTCCCACAACGATGAAGAACGCAGCGAAATGCGATACGTCTTGCGACTTGCA
GAATCCAGTGAATCATCAAAACTTTGAACGCA
Macrocystis pyrifera
Macrocystis integrifolia
Macrocystis x Pelagophycus hybrid
Pelagophycus porra
Alaria marginata
Lessoniopsis littoralis
Pterygophora californica
Egregia menziesii
Eisenia arborea
Postelsia palmaeformis
Nereocystis luetkeana
Dictyoneurum californicum
Dictyoneuropsis reticulata
Lessonia nigrens
Costaria costata

sites (Figure 4). Neighbor joining and parsimony phylogenetic trees were generated using PHYLIP 3.5c with 500 bootstrap replicated data sets, Jukes-Cantor corrections (Jukes and Cantor 1969) where applicable, and either random input order or ten-factor jumbling (Figure 5). Both neighbor-joining and maximum parsimony had the *Macrocystis* x *Pelagophycus* hybrid grouped with *M. integrifolia* in about 90% of the bootstrap replicates.

Discussion

The exact sequence identity between *Macrocystis integrifolia* sequences from Barkley Sound, B.C., Canada and Monterey Bay, CA, USA to the *Macrocystis* x *Pelagophycus* hybrid gametophytes was unexpected. In the northern hemisphere (Figure 1), *M. integrifolia* has a reported range from the Aleutian Islands to Monterey, California whereas *M. pyrifera* is found from Monterey, California to Magdalena Bay, Baja California, Mexico with possible populations reported in Alaska and Washington (Womersley 1954, Lewis and Neushul 1994, O'Clair et al. 1996). The *Macrocystis* x *Pelagophycus* hybrid was found nearly 475km south of the range of *M. integrifolia* (Figure 1).

ITS1 sequences from *M. pyrifera* found near the hybrid site were different from the hybrid at 13 of 461 nucleotides. The two identical *Pelagophycus porra* sequences of individuals from approximately 100km northwest and 100 km southeast

Figure 4. Variable sites and percent identity matrix for the ITS1 sequences from the alignment shown in Figure 2. a) All ITS1 variable sites are shown and used to calculate percent identities b) Ambiguous regions as denoted in Figure 2 were removed before calculating the percent similarities.

M/P hybrid CCCTTCAT--TTTATAAAACGATG--TGAGCTTACGAGGCGGCGAATAAATCTGCG
M. integrifolia .....
M. pyrifera .....AGTTCA..----CTA
P. porra G-------TTCGCTCGCGGAAGATTT-----CCGAAC----------CTC

	M/P hybrid	*Macrocystis integrifolia*	*Macrocystis pyrifera*
M/P hybrid	-		
Macrocystis integrifolia	100	-	
Macrocystis pyrifera	95.5	95.5	-
Pelagophycus porra	80.8	80.8	81.4

a)

Macrocystis/Pelagophycus hybrid CC-TTTATAAAACGATG-TTTACGAATGCG
Macrocystis integrifolia .....
Macrocystis pyrifera .....-CTA
Pelagophycus porra G-TCGCTCGCGGAAGATT-CCGAAC--CTC

	M/P hybrid	*Macrocystis integrifolia*	*Macrocystis pyrifera*
M/P hybrid	-		
Macrocystis integrifolia	100	-	
Macrocystis pyrifera	98.6	98.6	-
Pelagophycus porra	89.7	89.7	90.7

b)

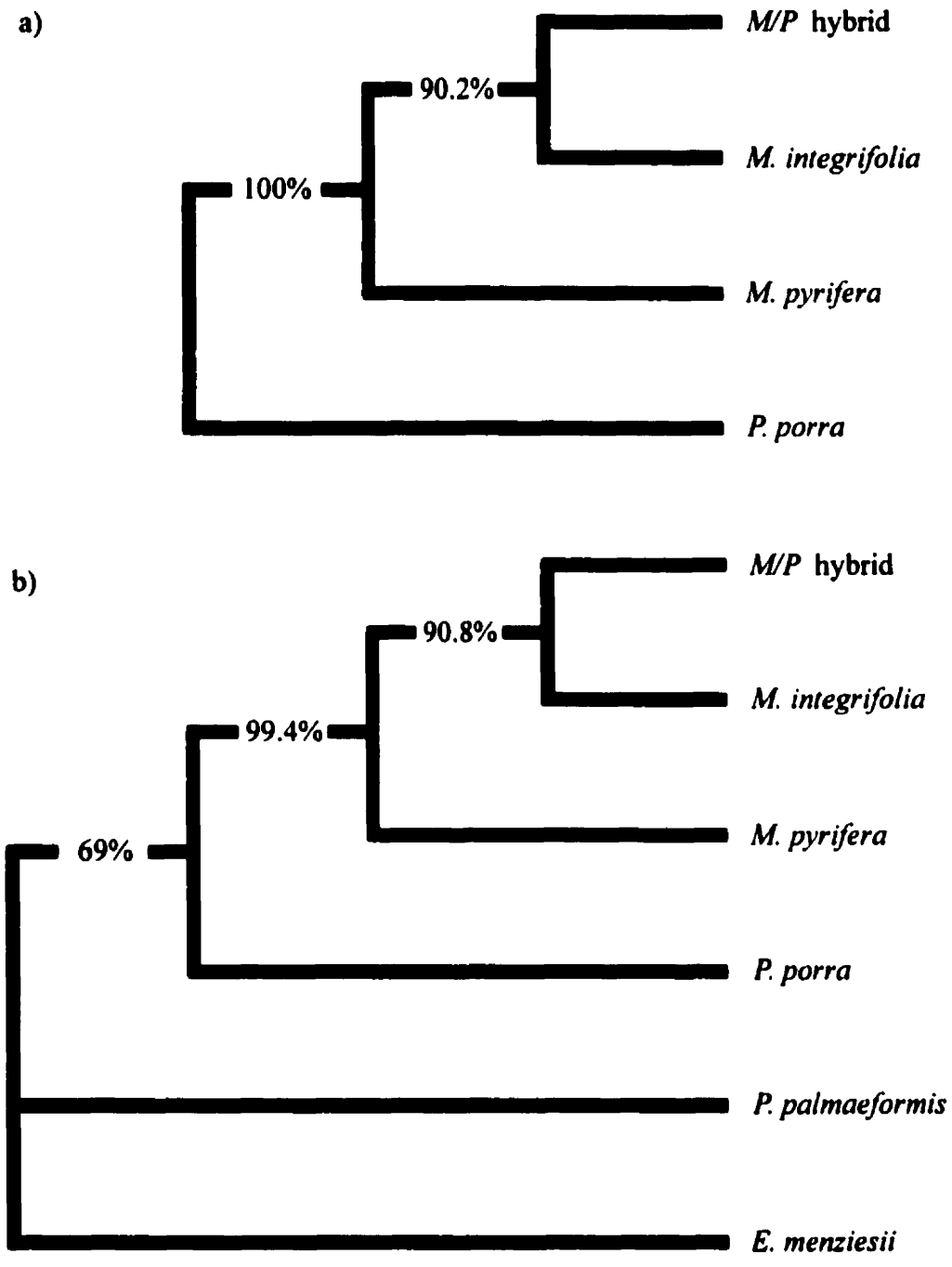


Figure 5. 50% Majority-Rule Consensus trees generated using PHYLIP 3.5c and the alignment from Figure 2. The percentages at the node indicate the frequency out of 500 replicate trees of the grouping of the taxa to the right of that node (bootstrapping). Nodes with percentages less than 50% were collapsed.
a) Maximum Parsimony-based consensus tree using 10-fold jumbling.
b) Neighbor Joining consensus tree using random input order.

of the hybrid (Table 1 and Figure 1) were different from the hybrid at 58 of 461 nucleotides (Figure 4).

Phylogenetic trees generated using taxa from the same clade as *M. integrifolia* in Saunders (1991) showed a close affinity between the hybrid and the two *Macrocystis* species (Figure 5), as can be seen in the 100% bootstrap values. Only three trees of the 500 bootstrapped Maximum Parsimony replicates failed to resolve the node separating the hybrid/*Macrocystis* spp. complex and *P. porra*. These results are somewhat self-evident given the fact that the hybrid sequence was identical to the *M. integrifolia* consensus sequence and only differed from the *M. pyrifera* sequence by three informative sites of the 413 examined (Figure 4). Bootstrapping, because it uses sampling with replacement of the dataset, can generate less than 100% node values in consensus trees when the sequences are identical but the number of informative sites is low (Felsenstein 1985). For this reason, the bootstrap node values did not reach 100% even though the hybrid and *M. integrifolia* sequences were identical.

In an examination of possible hybridizations, three possible outcomes would be evident by ITS1 sequencing. First, recombination could occur in the germ line leading to chimeric ITS1 sequences that contain region(s) of each parental type. However, the *Macrocystis* x *Pelagophycus* hybrid gametophytes appear to be quite different in their ITS1 from *P. porra*, and none of the differences between the hybrids and *M. pyrifera* could be accounted for by sequences in *P. porra*. Second, no evidence was ever seen, such as faint background bands or multiple sequence patterns

for the presence of both parental ITS1 types in either of the two hybrid gametophytes sequenced. Direct sequencing was purposely used so that multiple analogous sequences could be amplified and seen if present. Third, as was the case, only one parental ITS1 type could be found in the offspring.

Other workers have shown that the male hybrid *Macrocystis* x *Pelagophycus* gametophyte is a probable polyploid (2n) with about 30 chromosomes (Lewis and Neushul 1995). If the gametophytes are diploid and contain both parental genomes as hypothesized, then at least the *P. porra* nrDNA loci are missing from both the male and female hybrid gametophytes, and the gametophytes are therefore not complete allopolyploids.

A number of possibilities exist to explain the *Macrocystis* x *Pelagophycus* hybrid results. Meiotic nondisjunction or a lack of recombination in the F_1 germ line may have led to gametophytes with only one type of parental nrDNA. Meiotic nondisjunction could unevenly divide the parental chromosomes such that some or all of the gametophytes got only and entirely one parental type of ITS1 nrDNA-containing chromosome(s) in a way analogous to somatic cell hybridizations in tissue culture. For an as yet undetermined reason, in somatic cell hybridizations of mouse and human cells, human chromosomes are preferentially lost in later mitotic divisions, often leaving mostly mouse chromosomes and only one or two human chromosomes (Harris 1995). Since we do not yet know the chromosomal location(s) of the nrDNA repeats in any of the Laminariales, the number or identity of individual chromosomes from each of the putative parental types in *Macrocystis* x *Pelagophycus*

hybrids has yet to be determined. Possibly, if the nrDNA loci were all on one (or even less likely if more) chromosomes, and recombination between nrDNA loci from different parental chromosomes did not occur sometime during mitosis in the germ line, then the gametophytes would have one or the other parental nrDNA only. Then presumably the two gametophytes sequenced by chance each only contained *M. integrifolia* nrDNA.

Concerted evolution may have essentially removed any record of *P. porra* parental ITS1 (Campbell et al. 1997). Concerted evolution tends to homogenize multicopy genes like nrDNA given enough time (Dover 1982). Such a situation could occur if the hybridization event that led eventually to the gametophytes we tested happened many generations ago, or possibly if the number of mitoses in the germ line cells themselves was quite large. Analyzing the sequence of many more hybrid gametophytes from a number of hybrid sporophytes would be necessary to detect whether concerted evolution has homogenized the nrDNA in favor of the *M. integrifolia* parental type. Interestingly, polyploidy as proposed for the hybrids that we examined (Lewis and Neushul 1995), can inhibit concerted evolution by preventing sexual recombination (Campbell et al. 1997).

Finally, the hybrid may not actually be a hybrid of *Macrocystis* and *Pelagophycus* but instead just an apogamously produced sporophyte from a *Macrocystis* parent. *Macrocystis integrifolia* gametophytes have been reported to have either 16 or 28-32 chromosomes (Lewis 1996); The *Macrocystis* x *Pelagophycus* hybrid male gametophyte has 30-32 chromosomes (Lewis and Neushul

1995). Previously, other *Macrocystis* x *Pelagophycus* hybrids were produced in the lab (Sanbonsuga and Neushul 1978), but only female gametophytes were tested for apogamy. Other studies have tried crosses with these same hybrid gametophytes and gametophytes from *M. pyrifera* and *P. porra* (Lewis and Neushul 1995). Successful crosses were only seen between male *Macrocystis* x *Pelagophycus* hybrid gametophytes and *P. porra* females and not *M. pyrifera* females when judged by a lack of blades in single sex controls. This still leaves open the possibility that the original *Macrocystis* x *Pelagophycus* hybrids, which gave rise to the gametophytes we tested, were actually just apogamously produced blades of *M. integrifolia* similar to those seen in *Alaria crassifolia* Kjellman by Nakahara and Nakamura (1973). The male apogamously produced sporophytes of *A. crassifolia* lacked the characteristic mid-rib. Perhaps the morphological distinctiveness of the putative *Macrocystis* x *Pelagophycus* hybrid had a similar origin. The *Macrocystis* parental stock could have had either 16 chromosomes and then undergone autodiploidization (Müller 1967) or originally had about 30 chromosomes as reported by some authors for *M. integrifolia* (Yabu and Sanbonsuga 1987, Lewis and Neushul 1994).

Summary

This study showed that DNA sequence data could be used as a powerful tool for examining the genetic identity of naturally occurring hybrid kelp. The gametophytes produced by the F1 hybrid described in Coyer et. al (1992), were found

to not contain ITS1 rDNA from *Pelagophycus*. Instead, the gametophytes were found to be identical to *M. integrifolia*.

More data, gathered through sequencing of other sporophyte and gametophyte individual *Macrocystis* x *Pelagophycus* hybrids as well as other genic regions, are needed to solidify these results. Karyogamic or *in situ* hybridizations may also help explain the apparent uniparental inheritance seen in the hybrid ITS1 rDNA.

Chapter II

ITS1 nrDNA Fragments as Molecular Evidence for an Interfamilial Laminarialean Hybrid Cross Between *Alaria marginata* Postels & Ruprecht and *Lessoniopsis littoralis* (Tilden) Reinke.

Introduction

Laminarialean Hybrids

Many of the small blades observed in hybrid cross cultures of phaeophytes are impossible to classify as either the result of hybridization, parthenogeneis, androgenesis, apogamy, or some developmentally arrested partial hybridization. Confusion as to the nature of the resulting crosses arises because a gametophyte or sporophyte morphology does not necessarily correspond to a given ploidy level (Nakahara and Nakamura 1973). Processes such as autodiploidization (a spontaneous increase in chromosome number), apospory (the generation of gametophytes without the production of spores), apogamy (the generation of sporophytes without the production of gametes), and parthenogenesis/androgenesis (development of sporophytes from unfused gametes) can give rise to diploid gametophytes, haploid sporophytes, and other unusual ploidy levels (for review see Lewis 1996). What is needed to distinguish true and putative hybrids is a test of the genetic make-up and parental identities of the tissue in question.

ITS1 nrDNA

The Internal Transcribed Spacer 1 (ITS1) region of the nrDNA cistron was chosen for this study for a number of reasons. As mentioned earlier, the nrDNA cistron in most eukaryotes occurs as tandemly repeated units of usually several hundred copies approximately proportional to genome size (Li 1983). A number of studies in agriculturaly important species have had success using ITS1 regions to study hybridization (Buckler and Holtsford 1996). At present, the only published sequences which show variation amongst the recognized kelp taxa and encompass the various possible crosses within the Laminariales are from the ITS1 (Saunders 1991).

Materials and Methods

Hybrid and Test Crosses

Cultures used in this study (Table 2) were from our own collection. Gametophyte cultures were maintained in 1X f/2 medium under 20 μmol m^{-2} s^{-1}red light (16:8 LD photoperiod) at 13°C (Lüning and Dring 1975).

All materials used in manipulating and storing gametophytes were initially soaked in 500 mM HCl overnight, rinsed thoroughly in distilled water, and then autoclaved for 30 minutes at 1.1 kg $\cdot cm^{-2}$ and 121°C. Three independently isolated male and female gametophytes of *A. marginata* and *L. littoralis* from previously

Table 2. Strain, species, and general morphology of plants from which DNA was extracted. Both species were initially isolated from Barkley Sound, B.C., Canada on the 18th of January, 1990 and maintained in culture at the Bamfield Marine Station by L. Druehl.

Strain(s) used	Species	Morphology
Am1,1 female	*Alaria marginata*	female gametophyte
Am1,2 female	*Alaria marginata*	female gametophyte
Am1,3 female	*Alaria marginata*	female gametophyte
Ll1,1 female	*Lessoniopsis littoralis*	female gametophyte
Ll1,2 female	*Lessoniopsis littoralis*	female gametophyte
Ll1,3 female	*Lessoniopsis littoralis*	female gametophyte
Am1,1 male	*Alaria marginata*	male gametophyte
Am1,2 male	*Alaria marginata*	male gametophyte
Am1,3 male	*Alaria marginata*	male gametophyte
Ll1,1 male	*Lessoniopsis littoralis*	male gametophyte
Ll1,2 male	*Lessoniopsis littoralis*	male gametophyte
Ll1,3 male	*Lessoniopsis littoralis*	male gametophyte
Am1,1 female & Am1,1 male	*Alaria marginata*	sporophyte
Am1,2 female & Am1,2 male	*Alaria marginata*	sporophyte
Am1,3 female & Am1,3 male	*Alaria marginata*	sporophyte
Ll1,1 female & Ll1,1 male	*Lessoniopsis littoralis*	sporophyte
Ll1,2 female & Ll1,2 male	*Lessoniopsis littoralis*	sporophyte
Ll1,3 female & Ll1,3 male	*Lessoniopsis littoralis*	sporophyte
Am1,1 female & Ll1,1 male	*A. marginata* & *L. littoralis*	sporophyte
Am1,2 female & Ll1,2 male	*A. marginata* & *L. littoralis*	sporophyte
Am1,3 female & Ll1,3 male	*A. marginata* & *L. littoralis*	sporophyte
Am1,1 male & Ll1,1 female	*A. marginata* & *L. littoralis*	sporophyte
Am1,2 male & Ll1,2 female	*A. marginata* & *L. littoralis*	sporophyte
Am1,3 male & Ll1,3 female	*A. marginata* & *L. littoralis*	sporophyte

generated single-sex, unialgal cultures were ground in cooled depression grinding plates with a glass stirring rod until no filaments were visible to the naked eye. The grinding plates were pre-cooled to -20°C before use, then allowed to warm until the external frost melted to maintain a temperature as close to 4-10°C as possible during grinding. Only approximately 95% of the gametophyte was used for grinding, the other 5% was allowed to vegetatively propagate under red light conditions. Subsequently 200 µL of f/2 medium was added to each depression and mixed with the ground-up gametophyte. About half of the resulting dilution was added to 2 mLs of f/2 medium in a 35 mm x 10 mm sterile petri dish (Falcon #1008). The same procedure was repeated for the individual of the opposite sex used in the cross, and added to the same petri dish. Sporophyte cultures and crosses were kept under 150 µmol $m^{-2} \cdot s^{-1}$ cool-white fluorescent light (16:8 LD photoperiod) at 13°C for about six weeks. GeO_2 was added to a concentration of 500 µg/L in cultures showing signs of diatom contamination (Chapman 1973).

Cultures were checked under 100x magnification on an inverted microscope for the presence of blades and for any signs of diatom contamination on a weekly basis. For DNA extraction, individual blades were carefully dissected away from the gametophyte and placed in separate 1.8 mL microcentrifuge tubes.

DNA Extraction

The gametophyte DNA extraction method of Mayes et al. (1992) was utilized with the following modifications. Proteinase K digestions were done in 300 µL of

Proteinase K buffer (50 mM EDTA, 100 mM Tris pH 8.5, 200 mM NaCl, and 1% SDS) and at 20°C for 2 hours to reduce possible melting of intrinsic polysaccharides. Organic extractions were done first with a half volume of phenol (TE-saturated and pH 7.6 Tris-equibilarated)(Sambrook et al. 1989), without centrifugation or removal of the aqueous phase, and then with another half volume of Chloroform/IAA (24:1) followed by subsequent vortexing and centrifugation. One final extraction was done with Chloroform/IAA (24:1) as described in Mayes et al. (1992).

PCR Amplification

Two separate unique reverse primers were constructed from sequence in Saunders (1991) for the ITS1 region of the ribosomal cistron to determine the identity of *A. marginata*-based (AM1: 5'-GAGCCGCGCCCGGTAAAG-3') and *L. littoralis*-based (Ll1: 5'-GCGCTTTGATTCGAGAGACC-3') tissues using PCR (Kleppe et al. 1971, Saiki et al. 1988). A primer (either P1: 5'-TAATCTGTTGAACGTGCATCG-3' or BC1: 5'-GATTCCGGACTGTGGCTCGCG-3'), common to all known kelp in the 18S subunit (Saunders 1991), was used as the forward PCR primer. The expected band sizes were 512 or 336 basepairs for *A. marginata* and 438 or 265 basepairs for *L. littoralis* depending upon which reverse primer was used. 25 µL reactions were used with 10.25 µL sterile ddH_2O, 2.5 µL 10x PCR buffer (500 mM KCl, 100 mM Tris-Cl pH 8.3, and 0.1% gelatin), 2.5 µL 25 mM $MgCl_2$, 1.25 µL P1 or BC1 primer (20 µM stock), 1.25 µL of each downstream primer from 20 µM stocks (Ll1 or AM1), 5µL dNTPs (1.25 mM each stock), 2.0 µL template, and 0.25 µL *Taq*

polymerase (Perkin Elmer-Cetus). A modified 'hot start' procedure (Erlich et al. 1991) was used wherein the polymerase was added after pausing the thermocycler during the 67°C step of the melting cycle. PCR conditions were one initial melting cycle (95°C for 10 minutes, 67°C for 30 seconds, and then 72°C for 45 seconds), followed by thirty amplification cycles (94°C for 30 seconds, 66°C for 30 seconds, and then 72°C for 45 seconds), and finally a long extension step (72°C for 5 minutes). All PCRs from a hybrid cross were done concurrently using a master mix to obviate yield variations based on differing component ratios. PCR products were run on 15 cm long 0.9% agarose/TAE gels (Sambrook et al. 1989), post-stained for 15 minutes in a 250 ng · mL^{-1} Ethidium Bromide/TAE bath, and then destained in ddH_2O for 10 minutes.

PCR reactions from both male and female gametophyte tissue, sporophyte-like blades in single sex cultures (if seen), possible hybrid blades, and self crosses (where successful) were used with a forward primer and each species-specific reverse primer separately in such a fashion that the identity of the cultures was unknown to the person performing the tests. In addition, two pseudo-hybrids corresponding to the two observed hybrids were constructed by extracting both parental types from unialgal cultures in one tube. These pseudo-hybrids were subjected to the same PCR conditions as the possible hybrid blades.

PCR reactions were also done on blades from self crosses of the two taxa and a hybrid cross (cultures 45, 48 and 56) using three primers (Am1, Ll1, P1) in the

same reaction. All PCRs were performed at the same time in a parallel fashion using the same master mix (cocktail) and conditions as above.

Subsequent PCRs of blades seen in actual crosses were done once the predictive power of the tests were shown using controls of known samples. PCRs of blades in the actual hybrid crosses were run with an upstream primer and each species-specific primer pair separately.

Results

Hybrid Crosses

A total of seven reciprocal gametophyte crosses were attempted for *A. marginata* and *L. littoralis*, including seven *A. marginata* and six *L. littoralis* self crossess. Four of the crosses were done by Darcy Lightle and Louis Druehl. In addition, single male and female unisexual controls were established for each crossing attempt.

Three of the seven rounds of reciprocal crosses involved strains Am1,1 and Ll1,1, while two rounds involved Am1,2 and Ll1,2, and one involved Am1,3 and Ll1,3 (Table 2). Sporophyte morphologies developed in six of the fourteen attempted

hybrid crosses. Sporophyte morphologies were also seen in seven single female cultures and three male cultures of the twenty-eight single-sex cultures as well as five of the thirteen self crosses (Table 3).

PCR Tests

PCR reactions in which either species-specific primer along with an upstream primer were done using templates from both *A. marginata* and *L. littoralis*. Band sizes were as predicted in all cases and reflected the proper species in those cases where the identity was known. Faint unknown bands were only seen when three primers (P1, Am1, and Ll1) were used together (Figure 6).

Alaria marginata (strain Am1,1) genomic DNA from putative sporophyte blades, female apogamic or parthenogenic blades as well as individual male and female gametophytes in PCR reactions gave bands of the expected size (512 basepairs) only in reactions using the *A. marginata* -specific primer (Am1) and P1, but not when using the *L. littoralis*-specific primer (Ll1) and P1 (Figure 7).

Lessoniopsis littoralis male gametophytes, female gametophytes, and apogamic blades from crosses involving strain Ll1,1 were used in separate PCR reactions with either species-specific primer (Am1 or Ll1) and P1 (Figure 8). Both the male and female gametophytes gave bands of the expected size (438 base pairs). The apogamic male blade gave no results. The culture of *L. littoralis* contained only a few small diatom-engulfed blades and showed a faint band on the PCR of the

Table 3. *Alaria marginata* and *Lessoniopsis littoralis* gametophyte crossing attempts and resulting sporophyte morphologies.

Culture(s) used	Crosses Attempted	Crosses with sporophyte morphology
A. marginata female	7	2
A. marginata male	7	0
L. littoralis female	7	5
L. littoralis male	7	3
A. marginata female + male	7	2
L. littoralis female + male	6	3
A. marginata female & *L. littoralis* male	7	2
L. littoralis female & *A. marginata* male	7	4

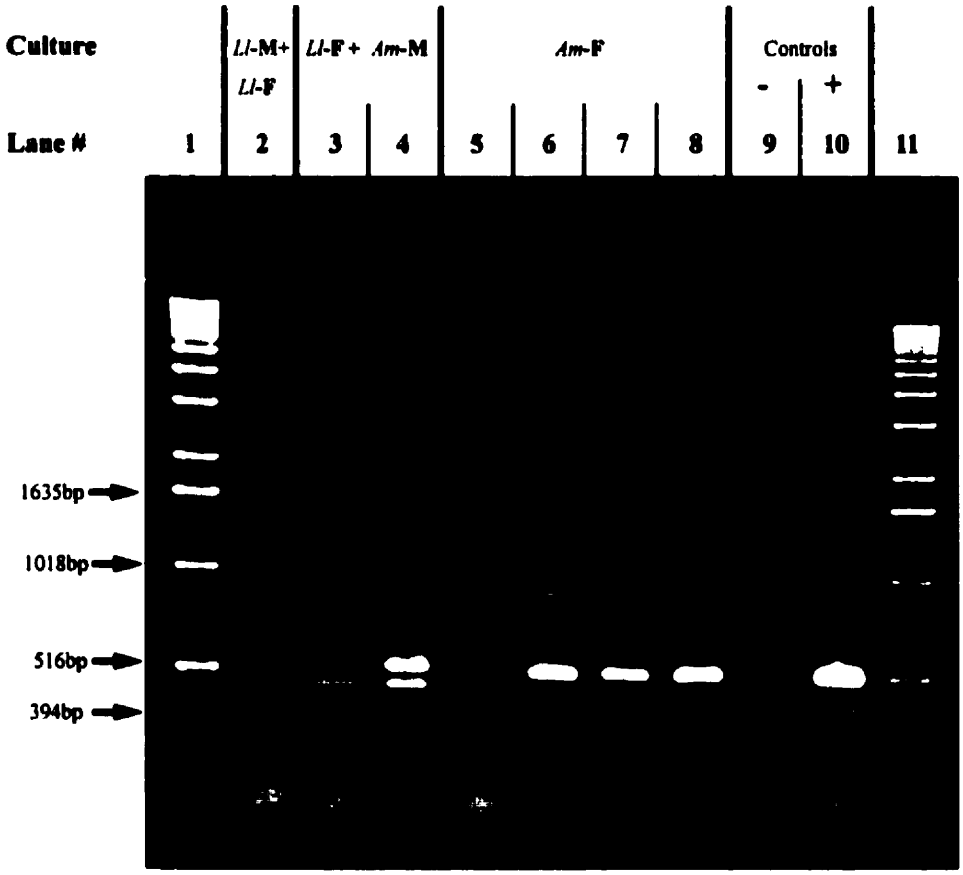


Figure 6. Agarose gel from PCRs of blades seen in cultures. Expected band sizes are 512bp for the *Alaria marginata* fragment and 438bp for the *Lessoniopsis littoralis* fragment. Both lanes 1 and 11 are 500 ng of 1 Kb DNA Marker. Lane 2 is a blade from a culture of both male and female *Lessoniopsis littoralis* gametophytes. Lanes 3 and 4 are separate DNA extractions from the same blade seen in a culture containing female *Lessoniopsis littoralis* and male *Alaria marginata* gametophytes. Lane 3 only faintly shows the *Alaria marginata* sized band, but does show the *Lessoniopsis littoralis*-sized band. Lane 4 has both appropriately-sized bands. Lanes 5 through 8 are separate blades seen in a culture of *Alaria marginata* female gametophytes, and all contain the correct-sized bands. Lane 9 is a no-template negative control and Lane 10 is a positive control using *Alaria marginata* meiospore DNA.

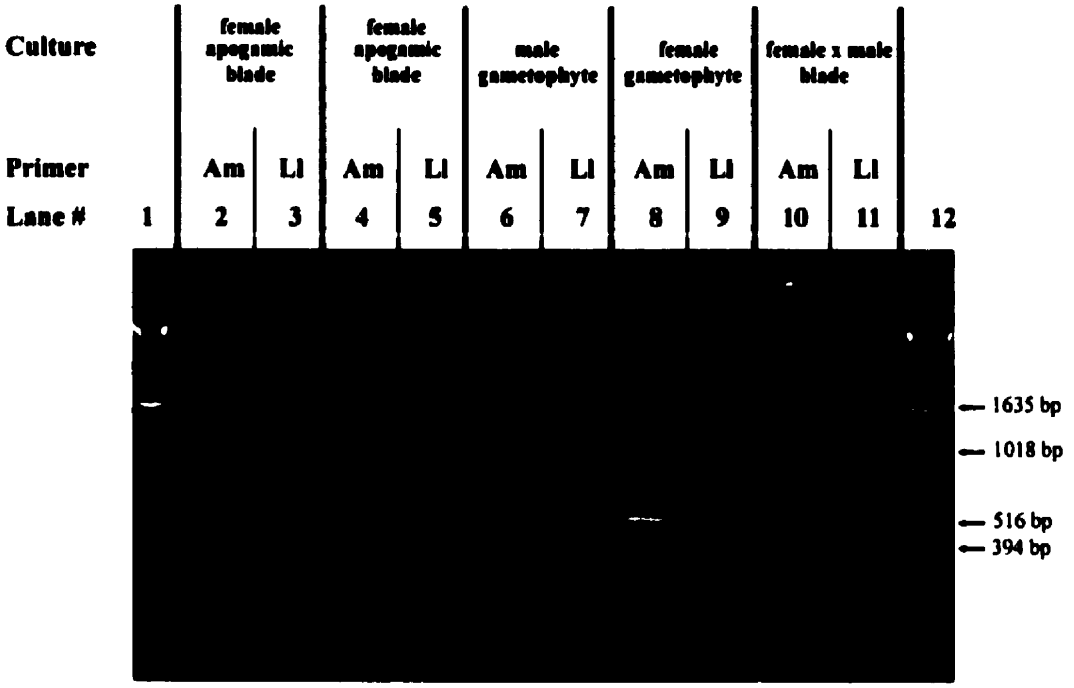


Figure 7. Agarose gel from PCRs of *Alaria marginata* tissues. The outer two lanes (1 and 12) are 500ng of 1Kb DNA Marker. The lanes are in pairs in which one of the pair is from a PCR with the *Alaria* primer (Am) and the universal primer P1 only, and the other of the pair is from a PCR with the *Lessoniopsis* primer (Ll) and the universal primer P1 only. Lanes 2 and 3 are from a blade seen in a female gametophyte culture. Lanes 4 and 5 are from a different blade seen in the same female gametophyte culture. Lanes 6 and 7 are from a male gametophyte, and lanes 8 and 9 are from a different female gametophyte. The final pair of lanes (10 and 11) are from a blade seen in a self-cross. In all cases only the Alaria primer (Am) when paired with the universal primer P1 produced a band. All bands were of the 512bp predicted size.

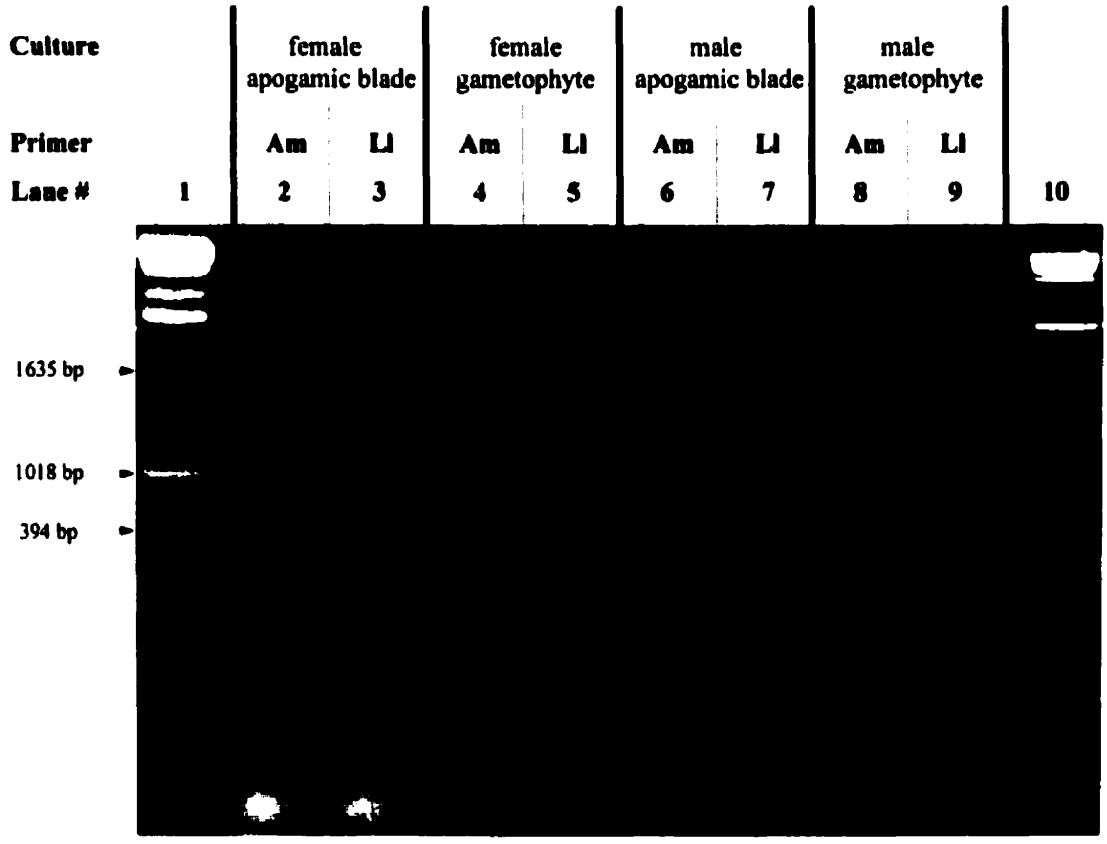


Figure 8. Agarose gel from PCRs of *Lessoniopsis littoralis* tissues. The outer two lanes (1 and 10) are 500ng of 1Kb DNA Marker (see Figure 1 for sizes). The lanes are in pairs in which one of the pair is from a PCR with the *Alaria* primer (Am) and the universal primer P1 only, and the other of the pair is from a PCR with the *Lessoniopsis* primer (Ll) and the universal primer P1 only. The first pair (lanes 2 and 3) are from a blade seen in a female gametophyte culture overgrown with diatoms and show a number of faint bands, barely visible in this figure, with the *Lessoniopsis* primer of which one is the correct size (438bp). Lanes 4 and 5 are from a female gametophyte and gave a band of the expected size in Lane 5. Lanes 6 and 7 are from the only small blade seen in a male gametophyte culture and showed no results. Lanes 8 and 9 are from a male gametophyte and showed a band of the correct size in Lane 9. The diffuse bands at the bottom of the figure are primer-oligomer PCR products. In all cases, only the *Lessoniopsis* primer (Ll) produced a band when paired with the universal primer P1. All bands were of the expected size.

correct size along with some faint other-sized bands (though no bands at the *A. marginata* size of 512 basepairs).

Genomic DNA from two possible putative hybrid blades resulting from a *L. littoralis* female x *A. marginata* male cross from strains Am1,2 and Ll1,2 and an *A. marginata* female x *L. littoralis* male cross from strains Am1,3 and Ll1,3 were individually subjected to PCR with each separate primer pair (Am1 & P1 or Ll1 & P1). One of the possible hybrid blades (an approximately 20 cell blade from a *L. littoralis* male x *A. marginata* female cross) gave no results, the other small blade (approximately 60 cells from a *L. littoralis* female x *A. marginata* male cross) resulted in only the appearance of the *L. littoralis* band. In addition the pseudo-crosses, performed by taking individual gametophytes from the parental type strains and coextracting them in the same tube followed by PCRs as above, produced both bands as expected (Figure 9).

PCR, using all three primers (P1, Ll1, and Am1) concurrently, was also attempted on templates from two possible hybrid blades and both parental gametophytes (Figure 6). Bands in the parental gametophyte PCRs were of the expected size but sometimes showed faint evidence for the incorrect band. Using all three primers only resulted in bands for six out of seven replicates of PCRs from two putative hybrids and often the *L. littoralis* band was brighter than the *A. marginata* band.

Blades seen in a hybrid cross of *A. marginata* strain Am1,1 with *L. littoralis* strain Ll1,1 were also amplified utilizing each species-specific primer (Am1 or Ll1)

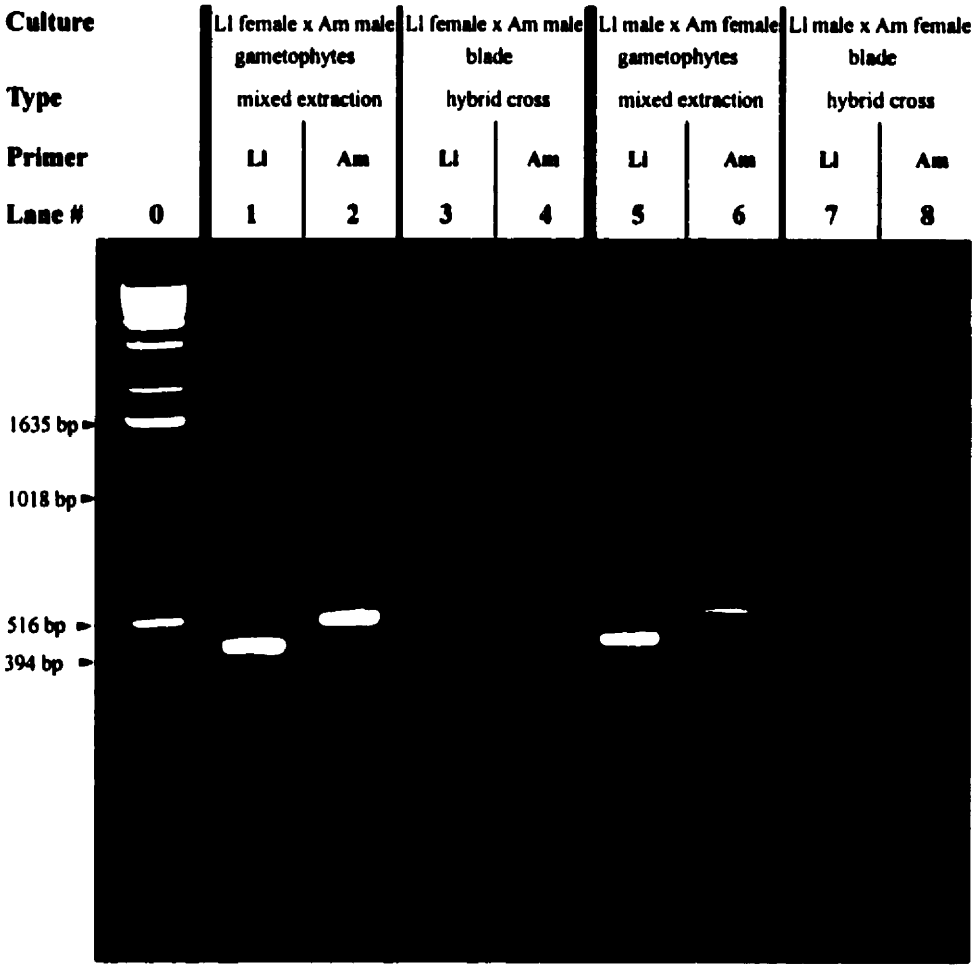


Figure 9. Agarose gel from PCRs of gametophyte-crosses and pseudo-crosses. Except for lane 0 which is 500ng of 1Kb DNA Marker (see Figure 1 for sizes), the lanes are in pairs in which one of the pair is from a PCR with the *Lessoniopsis* primer (Ll) and the universal primer P1, and the other is from a PCR with the *Alaria* primer (Am) and the universal primer P1. The first pair of lanes (1 & 2) are from a pseudo-cross done by extracting both a *Lessoniopsis littoralis* female and an *Alaria marginata* male together, and show the expected band sizes (438 and 512 bp) in each PCR. The next two lanes (3 & 4) are from a 60-cell blade seen in a hybrid-cross of the gametophytes from lanes 1 and 2 (a *Lessoniopsis littoralis* female and an *Alaria marginata* male). Lanes 5 and 6 are from a pseudo-cross like lanes 1 and 2, but instead using a *Lessoniopsis littoralis* male and an *Alaria marginata* female together, and show the expected band sizes (438 and 512 bp) in each PCR. The corresponding hybrid-cross yielded a 20-cell blade which gave no results with PCR (lanes 7 & 8).

individually along with the upstream primer BC1 (Figure 10). Bands were of the expected size in all of the apogamic blades seen in the single-sex controls. A putative hybrid blade seen in a culture containing *L. littoralis* female and *A. marginata* male only showed evidence for the *A. marginata* male parental genotype whereas another putative blade seen in a culture of *A. marginata* female and *L.littoralis* male showed bands from both parental types. No evidence of any incorrect bands was observed.

Discussion

Crosses

Every culture situation, except *A. marginata* males, produced plants having sporophyte morphology in at least one of the crosses attempted. The production of blades from unisexual female cultures is common (Lewis 1996). However, there are only a few reports of blades arising in unisexual male cultures (Nakahara and Nakamura 1973, Nakahara 1984, Lewis 1996). At least *in vitro*, *Alaria marginata* and *Lessoniopsis littoralis* apogamic or parthenogenic/androgenic blades are common.

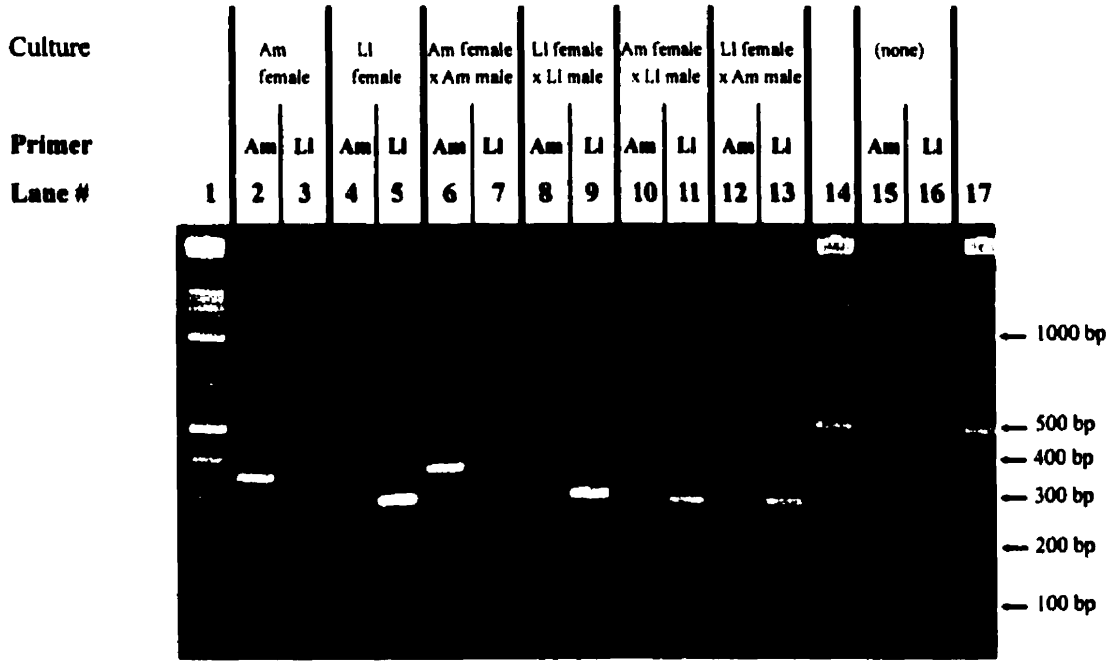


Figure 10. Agarose gel of PCRs from blades in a hybrid-cross using universal primer BC1 and either the Am or LI primer in adjacent pairs of lanes. Expected band sizes are 336 and 265 bp respectively. Lanes 1, 14, and 17 ar 100 basepair molecular weight markers. Lanes 15 and 16 are negative controls containing either primer Am or LI and BC1, but no template DNA. Lanes 2 and 3 are from a blade seen in an *Alaria marginata* female-only gametophyte culture. Lanes 4 and 5 are from a blade seen in a *Lessoniopsis littoralis* female-only gametophyte culture. Lanes 6 and 7 are from a blade seen in an *Alaria marginata* self-cross. Lanes 8 and 9 are from a blade seen in a *Lessoniopsis littoralis* self-cross. Lanes 10 and 11 are from a blade seen in a hybrid-cross of *Alaria marginata* female x *Lessoniopsis littoralis* male. Lanes 12 and 13 are from a blade seen in a hybrid-cross of *Lessoniopsis littoralis* female x *Alaria marginata* male, and cannot be distinguished from a *Lessoniopsis* self cross.

PCR Test

The species-specific primers were created based on the known ITS1 sequences from each genus (one species each in this case) and designed to anneal at different distances from the common primer. By choosing the species-specific primers to be approximately similar in melting temperature (T_m), a common primer of approximately the same T_m could be designed. The differences in the band sizes were purposely kept at or above about 50 basepairs (about 10% of the overall length) since smaller differences are difficult to see on agarose gels (Sambrook et al. 1989).

The results using species-specific primers imply that using each primer pair separately is ideal. The PCRs done using all three primers often were missing bands or had occasional faint spurious and questionable bands. No such false bands were seen when only two primers were used. Conceivably, bands did not appear or yields varied if all three primers were used because of slight competitve rate advantages magnified during the geometric growth phase of PCR. If one PCR primer pair was slightly more advantageous at the temperature used, or if the product is either shorter or forms less secondary structure, then rates of product production could vary. Although twice as many PCR reactions are needed, both of these problems are avoided by using each primer pair separately.

Interfamilial Hybrids

The parentage study indicates that most of the blades seen in cultures of gametophytes from *A. marginata* and *L. littoralis* were not true hybrids containing

both parental genomes. These blades were either partial hybrids composed of portions of each parent's chromosomes, or apogamic/parthenogenetic/androgenetic blades. However, one blade was seen that contained nrDNA ITS1 representatives of each genome. Although growth to maturity of this blade was not attempted, the method did allow for the determination of hybrid parentage in a blade as small as 60 cells.

This study is the first to directly attribute the genetic origin of a putative hybrid to its supposed parents. Earlier investigations have relied on indirect evidence of chromosome number and/or sporophyte morphology to define successful hybridizations. One assumption is that a doubling of chromosomes in blades arising from attempted crosses is indicative of successful sexual fusion. However, there is a possibility that autodiploidization may have occurred (Müller 1967). Increases in ploidy levels have been documented through several generations of parthenogenetic *Laminaria japonica* (Lewis et al. 1993).

Several studies have employed the general morphology of young blades as an indicator of their origin. Abnormal blades (mis-shaped, often consisting of irregular cell sizes and shapes, and cells having many nuclei) are often assumed to result from some non-sexual process, whereas blades having a normal morphology are usually considered the products of sexual fusion (Nakahara and Nakamura 1973). The assumption that abnormal blades reflect unisexual or mis-matched bisexual genomic contributions seems reasonable. However, the assumption that the origin of normal blades is restricted to successful sexual fusion is not so clear (Sundene 1958, Tom

Diek 1992). An apogamous haploid blade may undergo autodiploidization, resulting in a homozygous diploid sporophyte (Nakahara and Nakamura 1973).

One investigated blade arising from an *A. marginata* female x *L. littoralis* male cross shared both ITS1 nrDNA parental genomes, suggesting a successful sexual fusion took place between representatives of two laminarialean families (Alariaceae and Lessoniaceae). Other, putative interfamilial crosses within the Laminariales have been reported (Tokida et al. 1958, Cosson and Olivari 1982). Interfamilial fertility may indicate that laminarialean families, as defined on the basis of morphology, are more closely related than their taxonomic hierarchcial position would indicate (Druehl and Saunders 1992, Saunders 1991).

The findings reported here should be viewed as preliminary examinations of hybridizations between separate families of the Laminariales. The developed procedures would be useful for ploidy determinations as well as other genetic or hybrid studies because of the need for only a small amount of tissue. Once the inherent technical difficulties of rearing sporophytes were overcome, full-grown hybrids with parentage confirmed using the techniques described herein could be generated for further investigations.

Chapter III

Actin Introns in *Alaria* and *Nereocystis*

Introduction

In any study attempting to tease apart the events that led to the present genetic relatedness between conspecific individuals, or even to discover any inherent population structure, the challenge lies in picking traits or markers that are as unbiased as possible. In essence what is needed is a marker that will not be affected by processes such as recombination, selection, or other non-neutral events. A number of genetic markers are available, each with it's own difficulties, such as: mtDNA, cpDNA, introns, microsatellites and AFLPs, silent-site variation, and non-coding regions both upstream and downstream of known genes. Multicopy genetic regions have also been extensively studied, but they suffer from difficulties in determining whether the data generated represents the relationships of gene copies or individuals. Concerted evolution tends to mitigate these problems between species or higher taxa (Li and Graur 1991). This study aims to identify and characterize kelp actin intron regions from *Alaria* spp. and *Nereocystis luetkeana* for use in subsequent genetic analyses. Actin introns were chosen because of the availability of coding sequence, as so far partial actin coding regions of *Costaria costata* (C. Ag.) Saunders are the

only nuclear and non-ribosomal regions sequenced in any member of the Laminariales (Saunders 1991).

Alaria and Nereocystis

Alaria marginata Postels et Ruprecht, *Alaria tenuifolia* Setchell, and *Alaria nana* Schrader are the three recognized species of *Alaria* (in the family Alariaceae) found from mid-coastal British Columbia to northern California (Widdowson 1971). The three local species of *Alaria* are superficially similar in appearance with separate sporophylls and an elongated blade with midrib. *Alaria nana* is found on wave exposed sites, in the midtidal zone, and has irregular, highly elongated sporophylls with a length to width ratio greater than five; *Alaria tenuifolia* is found in protected areas, in the lower tidal or high subtidal zones, and is typified by a stipe length in excess of 15 cm with irregularly shaped sporophylls; *Alaria marginata* is found in both moderate exposed sites as well as sheltered sites in the lower intertidal zone, and has a stipe that is less than 15 cm in length and sporophylls with a length to width ratio less than five (Widdowson 1971). However, authors have noted the occurrence of individuals morphologically intermediate between these three recognized *Alaria* species (Widdowson 1971, Mróz 1989, Abbott and Hollenberg 1976).

Nereocystis luetkeana (Mertens) Postels et Ruprecht, in the family Lessoniaceae, is widely distributed subtidally down to 30m deep from the Aleutian Islands of Alaska to just south of Monterey, California (Scagel et al. 1989). *Nereocystis luetkeana* is the only species of the genus *Nereocystis*. It has a distinctive

morphology with a highly elongated stipe ending in a bulbous gas float from which many blades eminate.

Both *Nereocystis* as well as all three *Alaria* species are commonly found in the same general habitats (Waaland 1977). Only *Nereocystis* contains a gas-filled float that could theoretically disperse propagules over a wider area than *Alaria* which has no means of floatation. Since the distribution of *Alaria* and *Nereocystis* occurs along a narrow strip of the seashore, these kelp provide a unique opportunity to study dispersal (gene flow) unhindered by a third dimension over reasonable distances. This two dimensional nature allows for closer adherence to commonly invoked gene flow models such as the stepping-stone model (Hartl and Clark 1989) and combined with an appropriate genetic marker could allow for ecologically meaningful predictions of dispersal.

Actin introns have proved quite successful in quantifying gene flow and population subdivision in Pacific Humpback Whales (Palumbi and Baker 1994). My study attempted to also use actin introns to elucidate gene flow and population subdivision in the kelp *Alaria* spp. and the fairly closely related *Nereocystis luetkeana*.

Actin

Actin is a highly conserved gene found in most eukaryotes (Hightower and Meagher 1986). The highly conserved nature of actin is evident in the 65% amino acid identity between some of the most divergent eukaryotes (Kaine and Spear 1982,

Wildeman 1988). Actin is believed to be involved in cell motility, cytoplasmic streaming, muscle cell contraction, and cell surface organization (Novick and Botstein 1985, Wildeman 1988). At present (March 1999) 606 of the 1.9 million sequences in GENBANK are actin. Most eukaryotes have multicopy actin genes with the exception of various parasitic ciliates, some green algae, brown algae, certain oomycetes (water molds), many yeast, and most fungi (Bhattacharya and Stickel 1994).

The DNA sequence for most of the coding regions of the kelp *Costaria costata* actin gene, including the location of two intron splice sites, was determined by Bhattacharya et al. (1991). Restriction enzyme analysis and Southern blotting indicated that only one form of actin existed in *Costaria* and it occurred as a single-copy gene.

Introns

Introns are commonly found in many eukaryote genes. Two different types of introns have been identified, self-splicing and spliceosome-mediated introns (Cech 1986). Most mRNA coding for proteins uses snRNA spliceosomes to catalyze intron removal. Some self-splicing introns (called Group II) are similar to introns using spliceosomes. Self-splicing introns can be found in mitochondria, chloroplasts, and *Tetrahymena* rDNA (Costa et al. 1997).

Actin introns

Many actin genes have Group II introns, some of which occur at similar sites in different taxa. The actin genes of most multicellular eukaryotes have at least one, and usually two to six introns (Wildeman 1988). Some ciliates, oomycetes, and slime molds have no actin introns (Kaine and Spear 1982, Bhattacharya et al. 1991).

Materials and Methods

Sources and Meiospore Release

Individual plants were collected from a number of locations along the west coast of North America (Figure 11). The *Alaria* sporophyte samples were collected and DNA was extracted using a CsCl method by Mróz (1989). *Nereocystis* soral samples were collected and meiospores were released for subsequent DNA extraction.

DNA Extraction

DNA isolation from either meiospores or sporophyte tissues was attempted using a number of separate methods including Mayes et al. (1992), a novel method using adsorption to diatomaceous earth, a CTAB method, and a Chelex™ 100 based method.

The diatomaceous earth method (attempted on sample 3 from Figure 11) was loosely based on Myakishev et al. (1995), Carter and Milton (1993), and Jin Ngee

#	Designation	Location	Approximate position	Collector
1	NL01JB	Jones Bay, B.C.	44° 30'N 124° 05'W	-
2	NL01VR	Village Reef, B.C	48° 50'N 125° 17'W	-
3	NLCI04	Clarke Island, B.C	48° 50'N 125° 18'W	-
4	AM06SR	Seal Rock, OR, USA	44° 30'N 124° 05'W	L. Mróz
5	AM03KB	Kelsey Bay, B.C.	50° 24'N 125° 58'W	L. Mróz
6	AM03RB	Rosario Beach, WA, USA	48° 25'N 122° 40'W	L. Mróz
7	AM01WB	West Beach, WA, USA	48° 20'N 122° 40'W	L. Mróz
8	AT01OP	Orange Point, B.C.	50° 04'N 125° 17'W	L. Mróz
9	AM03SK	Sitka, AK, USA	57° 05'N 135° 15'W	L. Mróz
10	AM03JR	Jordan River, B.C.	48° 25'N 124° 03'W	L. Mróz
11	AN10BO	Botany Beach, B.C.	48° 32'N 124° 27'W	L. Mróz
12	AM01WS	Whiffen Spit, B.C.	48° 22'N 123° 47'W	L. Mróz
13	AM04CP1	Cattle Point, WA, USA	48° 45'N 122° 45'W	L. Mróz

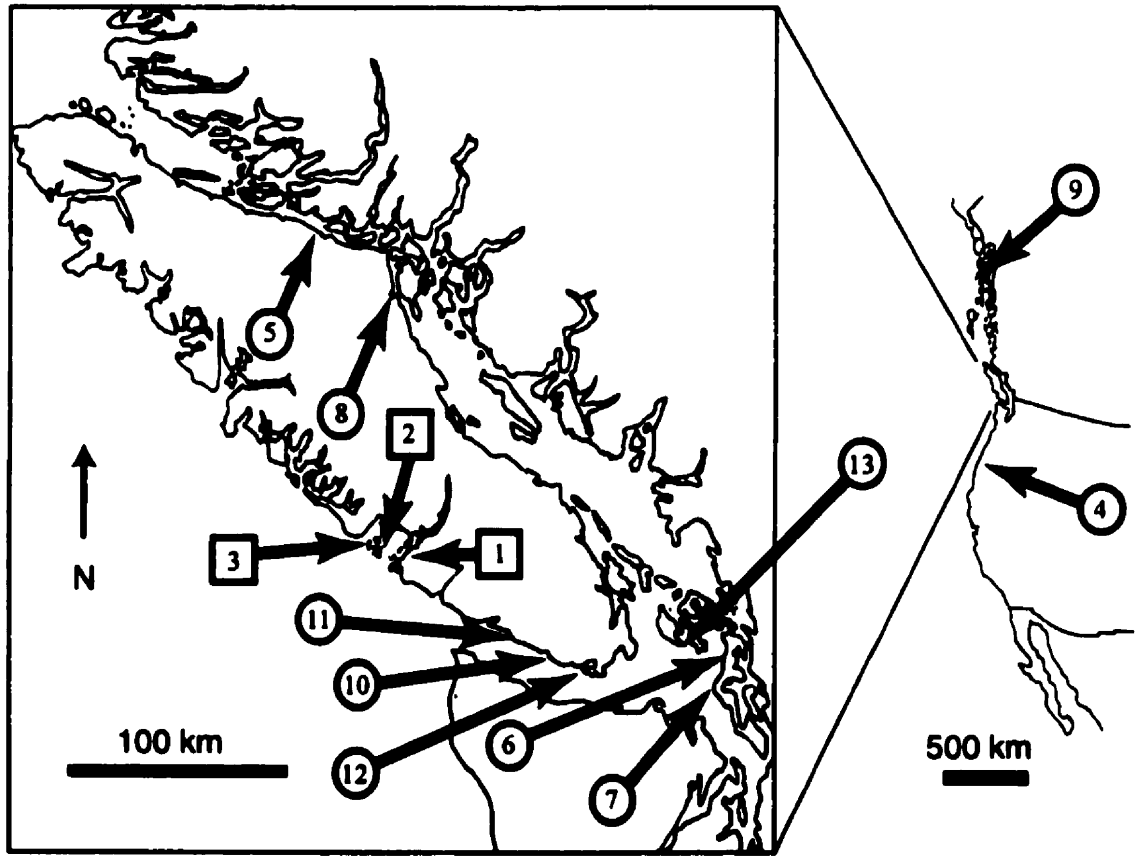


Figure 11. Designation, location, approximate position, and collector (if different from the author) for *Alaria* sp. and *Nereocystis leutkeana* individuals on map shown. Designations for *Alaria* individuals start with 'A' while followed by the species designation, the individual number, and then an abbreviation for the location. *Nereocystis* individuals are designated in a similar manner except that they begin with 'NL'. Designations for individuals collected by L. Mróz correspond to the designations assigned in Mróz (1989).

Chia (*pers. comm.*). A high-salt chelating solution, BS1 (50 mM Tris · HCl pH 7.6, 50 mM 2Na·EDTA pH 8.0, 100 mg · mL^{-1} RNase A), was designed to bind excess divalent cations from marine tissues. In a 1.5 mL microcentrifuge tube, approximately 30 μL of pelleted meiospores were resuspended in 400 μL of BS1 by vigorously vortexing for 30 seconds. The top of the tube was pierced with a needle, and the tube was placed in a boiling water bath for 5 minutes. If sporophyte tissues were used, a similar amount of thinly sliced sporophyte tissue was ground with 400 μL of BS3 minus diatoms (see below) in a tissue homogenizer (Polytron). Cellular debris was pelleted in a microcentrifuge for 15 minutes at 13,000 g. The supernatant was added to 700 μL of BS2 (6M guanidine hydrochloride, 50 mM Tris · Cl pH 8.0, 20 mM 2Na·EDTA pH 8.0, and 40 mg·mL^{-1} acid-washed, calcined diatomaceous earth [Sigma]) in a new tube. In addition, 8 μL of Nonidet P40 (Sigma) and 5 μL 1M DTT were added. The tube was incubated at 20°C for about five minutes with occasional shaking. The tube was centrifuged at 5,000 g for five minutes and the supernatant discarded. 1 mL of BS3 (50% EtOH, 5 mM 2Na·EDTA pH 8.0, 20 mM Tris HCL pH 7.6, and 200 mM NaCl) was added and gently mixed for 30 seconds or until no clumps were visible. The tube was again centrifuged at 5,000 g for 5 minutes and the supernatant was removed. The pellet was again resuspended in 1 mL of BS3 and loaded on a mini-column (Promega). With a syringe, the solution was pushed through the mini-column followed by an additional mL of BS3. 200 μL of 95%EtOH were pushed through the column followed by a 13,000 g centrifuge spin for 2

minutes. 50 µL of 80°C 10mM Tris · Cl pH 8.0 was added to the column and allowed to sit for 1 minute. The mini-column was centrifuged at 13,000 g for 2 minutes and the flow-through kept. 5 µL of 1xTE + RNase A (10 $\mu g \cdot mL^{-1}$) was added and the tube was placed at -20°C until needed.

A Chelex™ 100 based method (see Chapter I) was utilized to extract DNA from samples 1, 2 and 3 in Figure 11.

A CTAB-based method was also attempted on samples 1, 2 and 3, though no viable sequence was ever generated. About 20 mg of nitrogen-ground dry sporophyte tissue was placed in a tube with 300 µL Proteinase K buffer (Sambrook et al. 1989) and 5 mg Proteinase K at 65°C for 1 hour. The solution was placed in a boiling water bath for 45 seconds and allowed to cool at room temperature. 10 µL of RNase A solution (50 $mg \cdot mL^{-1}$) was added and the solution was put in a 42°C bath for 30 minutes. 300 µL of CTAB solution (5% (w/v) CTAB and 0.5 M NaCL) was added and the tube was incubated for 3 minutes at 65°C. The tube was centrifuged for 5 minutes at 13,000 g and the supernatant was removed. 300 µL of 1.2M NaCL, or enough to keep the concentration of NaCL above 0.8 M total (Del Sal et al. 1989), was added and the solution was gently mixed. 1 mL of 95% EtOH was added and the tube was centrifuged for 10 minutes at 13,000 g. The supernatant was removed and the pellet was rinsed with 70% EtOH followed by a 1 minute spin at 13,000 g. The EtOH was removed and the pellet was allowed to air dry for 10 minutes. 300 µL of 65°C 10 mM Tris · Cl pH 8.0 was added and the pellet gently resuspended. A

sequential organic extraction of the aqueous phase with equal volumes of Tris · Cl pH 8.0-saturated phenol, then 25:24:1 phenol / chloroform / isoamyl alcohol, and then chloroform (24:1 chloroform / isoamyl alcohol) was performed. 120 µL of 5M NH_4OAc and 1 mL of -20°C 95% EtOH was added and the solution was left at -20°C for 20 minutes. The solution was centrifuged at 13,000 g for 20 minutes at 4°C. The supernatant was removed and the pellet was rinsed with 70% EtOH followed by a 1 minute spin at 13,000 g. The EtOH was removed and the pellet was allowed to air dry for 10 minutes. 50 µL of H_2O was added and the solution kept at -20°C until needed.

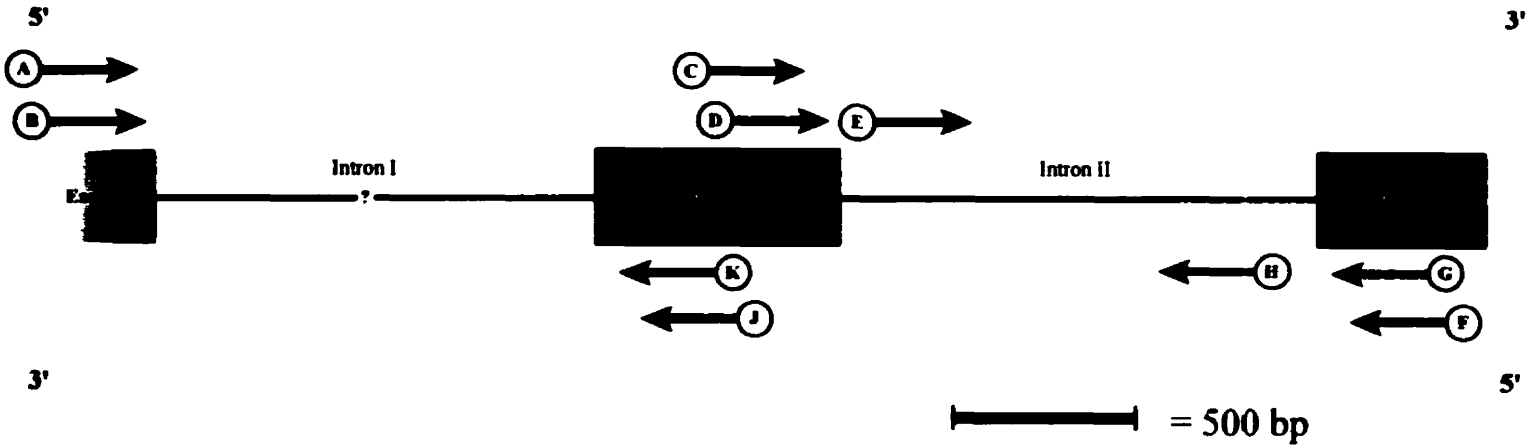
PCR Optimization

Primers were designed to anneal 20 to 100 bases from either end of both actin introns in *Costaria costata* (Bhattacharya et al. 1991). The primers were also designed as nested pairs and checked for complementarity or hairpin formations Figure 12) using OLIGO 5.0 (National Biosciences Inc.). Once intron sequences were generated, primers were designed 122 and 284 bases from the exon/intron boundary on the 5' and 3' end of the intron respectively.

PCR annealing temperatures for each primer pair were initially set based on a formula in Wu et al. (1991). A modified 'hot start' (Erlich et al. 1991) procedure was used wherein the polymerase was added by pausing the thermocycler at the initial 70°step. An initial five or 10 minute incubation at 95-98°C was performed to insure proper denaturation.

Figure 12. Schematic of actin based on sequences from Bhattacharya et al. (1991) and this study showing primer identity and locations. The question mark in the first intron as well as the shading in the first exon indicates uncertainty about length. The tips of the arrows indicate approximate positions of the primer binding sites.

Actin



Primer	Sequence	Distance from exon/intron boundary
A	5' - CGGTTCCGGTATGTGCAAGGC - 3'	- 76 bp
B	5' - GCGCTGTCTTCCCATCGATTG - 3'	- 30 bp
C	5' - GAGAAGTCGTACGAGCTCCCC - 3'	- 49 bp
D	5' - ATCGGAAACGAGCGCTTCCG - 3'	- 11 bp
E	5' - CCGGACAGAAGTACACGGCCC - 3'	+ 122 bp
F	5' - GGATGTCGACGTCGCACTTCATG - 3'	+ 69 bp
G	5' - GTGCAGTCGTGGATGCCCG - 3'	+ 38 bp
H	5' - GAATCCTCAAACAGTTCATACTGC - 3'	- 284 bp
J	5' - CGCGGAGCTCGTTGTAGAAGG - 3'	- 133 bp
K	5' - CCGTGCTCGATCGGGTACTTC - 3'	+ 69 bp

Optimization of the PCR conditions was performed. Mg^{2+} concentrations were varied from 1 to 4 mM. For primer concentration, 50 to 500 pM of each primer was tried. Annealing temperatures between 42°C and 68°C were tried as well as enzyme concentrations from 0.1 unit to 5 units per reaction. In each specific case, template concentration was also optimized. To test Mg^{2+}, primer concentration, as well as template amount and enzyme concentration, a 'touchdown' PCR was performed (Don et al. 1991). Initial PCR conditions were: a melting cycle (95°C for 5 minutes, 70°C for 1 minute, 72°C for 30 seconds) followed by 30 'touchdown' amplification cycles (93°C for 45 seconds, 68°C for 1 minute with a 0.5°C decrease each cycle, 72°C for 45 seconds), and finally a longer extension step (72°C for 5 minutes).

Actual working conditions for PCR used to generate bands in most cases did not use 'touchdown' PCR. Typical 25 μL PCR reactions for all primers included 8.5 μL ddH_20, 2.5 μL supplied 10x enzyme buffer, 2 μL $MgCl_2$ (2 mM final concentration), 5 total μL of primer (50 pM each), 4 μL dNTPs mix (200 nM each final concentration), 2.5 μL of template (amount added varied), and 0.5 μL of enzyme mix (1:15 Pfu to Taq). Typical reaction conditions were 98°C for 5 minutes followed by 75°C for 1 minute to allow enzyme addition, thirty cycles of 62°C for 1 minute, 72°C for 1 minute, and 94°C for 45 seconds. A 60°C step for 1 minute followed by a 72°C step for 5 minutes were performed.

PCR products were examined on 0.8 -1 % agarose gels post-stained with 0.01% EtBR and ddH_2O and visualized with 260 nm UV light.

Excision and Cleaning of PCR products

PCR reactions to be excised were divided into two fractions and run in duplicate sets of alternate lanes on 15 cm long 0.9% agarose/TAE gels as described in Chapter 1 (Sambrook et al. 1989). Instead of purification with the Sephaglas™ Band Prep kit , some bands were concentrated by placing the band in a 0.6 mL centrifuge tube, piercing the bottom of the tube with a 22-gauge needle, placing the tube in a 1.5 mL tube, and centrifuging at 5,000 g for a minute or so. To make sure the entire band was excised, the remaining portions of the gel were post-stained as above and viewed under UV illumination.

Single Stranded PCR Products

Single-stranded sequencing templates were generated using the Asymmetric PCR technique described in Nichols and Raben (1994) except that only 300 pM of primer was used for the initial PCR. Otherwise, the initial PCR was identical to the general conditions described above. Single-stranded PCRs from reactions carrying various amounts of the first PCR as template were run on 1% agarose gels to check for the best yield. Asymmetric PCRs were initially organically extracted as described above and heated at 65°C in a sand bath to remove chloroform. Finally, the single-stranded products were cleaned by passing the products over a 30,000 MW cut-off Ultrafree-MC™ spin filter (Millipore #UFC3 LTK00) followed by three rinses with 18 MΩ ddH_2O. The reactions were left at -20°C until ready for sequencing.

Cloning of PCR Products

One of two methods was used to clone PCR products. Some products were blunt-cloned into the EcoRV site of pBluescript (Stratagene). The rest of the products were cloned into a proprietary vector containing a 'killer' gene in the MCS as well as a kanamycin-resistance gene using the protocols supplied in the ZeroBlunt™ PCR Cloning Kit (Invitrogen #K2700-20). Positive clones on the kanamycin plates were cut with EcoRI to test for the correct insert and stored as glycerol stocks (Sambrook et al. 1989) at -80°C until needed for sequencing. Sequencing-ready minipreps were prepared from 5 mL overnight cultures picked from glycerol-stock streaked plates using a QIAprep™ Spin Miniprep Kit (Qiagen).

Cloning of fragments into pBluescript was performed two different ways. Initially a modification of TA cloning was used (Marchuk et al. 1991). A blunt-cut pBluescript vector with dTTP overhangs was constructed by cutting 10 μg of the vector in a thermocycler with EcoRV in a 40 μL reaction for 2 hours at 37°C followed by 70°C for 15 minutes. Total volume was adjusted to 100 μL with ddH_2O. A sequential organic extraction of the aqueous phase with equal volumes of Tris · Cl pH 8.0-saturated phenol, then 25:24:1 phenol / chloroform / isoamyl alcohol, and then chloroform (24:1 chloroform / isoamyl alcohol) was performed. 9 μL of 3M NaOAc was added followed by 1.8 mL of 95% EtOH and left at -20°C overnight. The tube was centrifuged at 13,000 g for 5 minutes and the supernatant was removed. A final 70% EtOH wash was performed followed by 13,000 g centrifugation as above after

which the pellet was allowed to air-dry for 10 minutes. The cut vector was resuspended in ddH_20 and brought up to 100 μL in a 0.6 mL tube which contained a final concentration of 1x PCR buffer (see Chapter I), 2 mM dTTP, 1.5 mM $MgCl_2$, and 5 units of Taq polymerase. The reaction was placed in a thermocycler for 3 hours at 70°C followed by a sequential organic extraction as above. An equal volume of 2M NH_4OAc was added followed by two volumes of –20°C isopropanol. The tube was spun in a 4°C microcentrifuge at 13,000 g for 20 minutes. The supernatant was removed and 100 μL of 70% EtOH was added followed by a 13,000 g 5 minute spin at 4°C. The supernatant was again removed and the pellet was allowed to air-dry for about 2 hours until completely dry. The pellet was stored at –20°C until needed at which time it was resuspended in 160 μL of 10 mM Tris · Cl pH 8.0 and checked on an agarose gel against λ Hind III marker band intensity to estimate concentration (usually about 50 ng · μL^{-1}). 20 μL TA ligations were set up containing 50 ng T vector, 2 μL 10x Ligation buffer (60 mM Tris-HCl pH 7.5, 60 mM $MgCl_2$, 50 mM NaCl, 1 mg · mL^{-1} BSA, 70 mM β-mercaptoethanol, 1 mM ATP, 20 mM dithiothreitol, 10 mM spermidine) 50 units of ligase, and about 500 ng of insert (about 10 to 1 ratio of insert : vector) for 16 hours at 14°C. Ligations were kept at –20°C until ready for transformation.

Blunt cloning into pBluescript was also done using a slight modification of a method wherein polishing of the fragment, cutting of the vector, and ligation of the two together all takes place in one tube (Chuang et al. 1995). EcoRV was used to

blunt the vector instead of SmaI. Five extra units of EcoRV were added and the tube placed at 37°C after being left overnight at 14°C. Five more units of ligase were added and the reactions were placed at 20°C for an hour. The reactions were left at –20°C until ready for transformation.

For the pBluescript-based constructs, 200 µL of XL-1 Blue (Stratagene) competent cells were used for transformation of about 10 µL of each ligation following the manufacturers protocol. Positives were picked based on *lac* Z α complementation (blue/white selection), copied onto X-gal plates, and grown as 5 mL overnight cultures. Boil minipreps were performed on 3 mL of pelleted overnight culture (Berghammer and Auer 1993), digested with appropriate restriction endonucleases to release the insert, and run on an agarose gel. Confirmed positives were stored as glycerol stocks (Sambrook et al. 1989) at -80°C until needed for sequencing. Positives with the correct insert were grown again as 5 mL overnight cultures, picked from newly-streaked plates of the glycerol stocks, and cleaned for both automated and manual sequencing using a QIAprep™ Spin Miniprep Kit (Qiagen).

DNA Sequencing

Manual sequencing was performed on clones, direct PCR products, and single-stranded Asymmetric PCR products, using either ^{35}S-dideoxy sequencing or ^{33}P-cycle sequencing. Direct sequencing of double and single-stranded PCR products was carried out using a DMSO-based modification of the ^{35}S Sequenase™ 2.0

(Amersham) protocol (Tan and Druehl 1994). Clones were sequenced using either the same DMSO-based modification of the ^{35}S Sequenase™ 2.0 (USB/Amersham) protocol (Tan and Druehl 1994) or the protocol supplied with the ^{33}P ThermoSequenase™ cycle sequencing kit (Amersham #US 79750)).

The automated thermal cycle sequencing of various clones were performed by the staff at University Core DNA Services, University of Calgary, Calgary, Alberta using an ABI Prism 377 (Applied Biosystems Inc.) and the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq® DNA Polymerase, FS.

Manual sequencing reactions were separated on polyacrylamide gels as described in Chapter 1.

Sequence analysis

All sequences generated by automated thermal cycle sequencing were manually checked for common base-call errors (Huntley 1996). Automated sequencing runs were also usually done twice to reduce reading errors. For each sequencing primer, both manual ^{35}S dideoxy and manual ^{33}P cycle sequencing generated sequences were repeated at least one additional time and read three times. In a number of cases, both separate clones as well as separate PCR bands from the original template were sequenced. NC-IUB base nomenclature was used in all cases (NC-IUB 1984).

Sequences were aligned by eye using the ESEE sequence editor (Cabot and Beckenbach 1989). Exon regions from the cDNA sequence in Bhattacharya et al. (1991) were used to orient the sequences. Sequence length and base composition was determined using the COMPARE.exe program written by Prof. Andy Beckenbach (Simon Fraser University).

Exon/Intron boundaries were re-calculated using a number of methods. Splice-junctions were deduced using a comparison to *Arabidopsis* splice junctions from Brown et al. (1996) as well as the NetPlantGene server (http://genome.cbs.dtu.dk, Hebsgaard et al. 1996). Two neural network-based methods for predicting human and *Drosophila* splice junctions were also tried: Splice-site-predictor (http://www-hgc.lbl.gov, Reese and Eeckman 1996), and GeneFinder (http://dot.imgen.bcm.tmc.edu:9331/gene-finder/, Solovyev et al. 1994). Finally, the "GT-AG rule" (Mount 1982, Breathnach and Chambon 1981), and inspection of the alignments of the *Alaria* and *Nereocystis* individuals, were used with the predicted splice-junctions to correlate the results.

Sequence identity was checked by performing Basic Local Alignment Search Tool (BLAST) searches (Altshul et al. 1990) with Open Reading Frames (ORFs) generated using translations from all six reading frames against a translated non-redundant (nr) database (tBLASTn) at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov). ORFs larger than 30 residues and with smallest sum probabilities smaller than approximately 0.05 were kept for comparisons.

Results

DNA Extraction

Alaria sequence was only generated using the CsCL prepared samples from Mróz (1989). The Diatomaceous Earth and the Chelex extraction method proved fruitful in generating *Alaria* genomic DNA preparations, but sequence results were poor. Only Chelex-based extractions were used for sequencing the *Nereocystis* samples (# 1-3 in Figure 11).

PCR methods

PCR utilizing hot-start and annealing temperatures near 60°C were found ideal for most primer pairs. The most reliable results were obtained when using primer pairs C/F and A/K (Figure 12). When using template DNA produced by any method except CsCL-purification, yields were both low and sporadic. PCRs using CsCl-purified templates originally produced by Mróz (1989) gave consistently better yields if the total template amount per reaction was kept below 0.4 ng (Figure 13).

Resulting band sizes were typically about 1.3 kbp for Intron II primers C/F and 1.4 for primers A/K. The Intron I-containing fragment was estimated to be approximately 1.4 kbp using primers A/J.

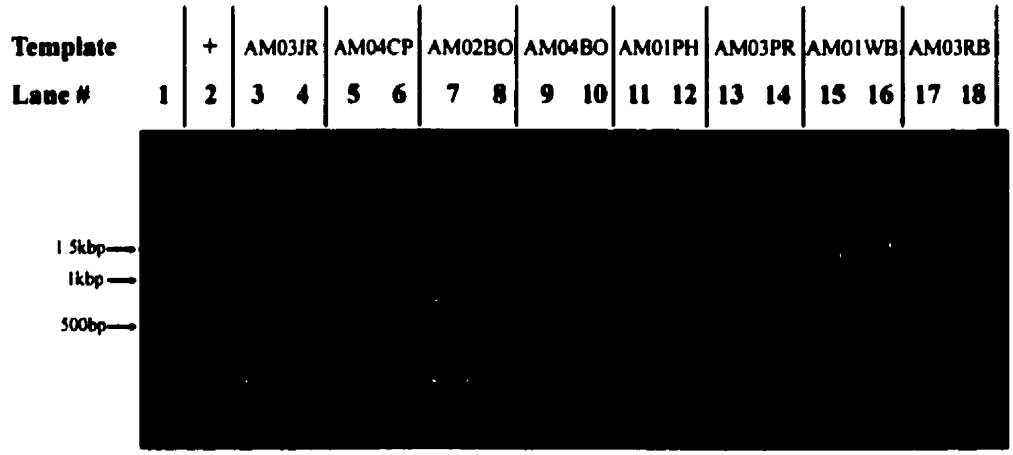


Figure 13. Agarose gel of PCR fragments from *Alaria marginata* actin introns I and II. Lane 1 contains 400ng of 100bp Molecular Weight Marker. Odd-numbered lanes are PCRs using primers C & F (see Figure 12), and should have an intron II band of approximately 1.3 kbp (except lane 1 which contains 400ng of 100bp Molecular Weight Marker). Even-numbered lanes are PCRs using primers A & K, and should have an intron I band of approximately 1.4 kbp. Lanes 3 - 12 have approximately 40 ng of template, while lanes 13 - 18 have 0.4 ng of template. Lane 2 is a positive control containing 0.4 ng of AM02JR template and is otherwise the same as the other even-numbered lanes. White arrows denote the two expected band sizes as generated in AM01WB.

Reactions using primer D in combination with either primer F, G, or H, often resulted in strong multiple banding patterns containing two or three additional bands consistently at about 1.9, 0.9, and 0.4 kbp (Figure 14). The five extra bands generated from AM06SR and AM04CP (see Figure 11 for designations) with primer pairs including primer D were excised and cloned using the pZero™ kit. Automated sequencing of the clones showed that the PCR products all had primer D sequence on both ends. tBlastn searches of all six reading frames of the resulting products showed no similarity greater than the 0.005 smallest-sum probability level. The only consistent similarities seen between the different primer D clones were very weak scores to a number of thermophilic bacteria sequences, suggesting contamination from the polymerase.

Cloning

As described above, three different cloning methods were attempted: a modified TA cloning method, TIG cloning, and Invitrogen's pZero™ kit. Of the three, TA cloning was found the least useful. After numerous attempts using TA cloning, only a few Intron II regions were successfully cloned. Three clones from AM06SR and two clones from NLCI04 were generated.

TIG cloning was much faster and resulted in more positive clones. Two Intron I regions were successfully cloned including two clones of AM03JR and one clone of AN10BO. In addition, Intron II positive clones included three clones from AM03RB, eight clones of AM03JR, and six clones from AN10BO.

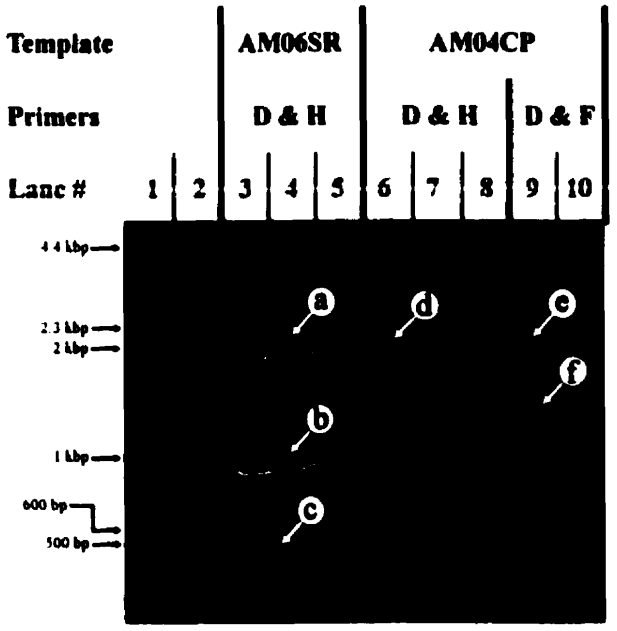


Figure 14. Agarose gel from PCRs of *Alaria marginata* actin intron II. Lanes 1 and 2 are 500ng of λ DNA/Hind III and 100bp molecular weight markers respectively. Lanes 5 and 8 were not loaded to allow for easier band excision. Lanes 3 and 4 are both AM06SR with primers D and G (see Figure 12 for primer designations) and differ by an order of magnitude in template concentration. Lanes 6 and 7 are the same as 3 and 4 except that AM04CP is the template. Lanes 9 and 10 again are AM04CP and only differ by an order of magnitude in template concentration, but were generated with primers D and F. Circled areas ⓐ through ⓔ (but not ⓕ which was the correct size) were excised for cloning (approximate sizes: a = 1.8 kbp, b = 900 bp, c = 450 bp, d = 1.8kbp, e = 1.8 kbp, f = 1.2 kbp).

Although pZero™ cloning tended to yield fewer numbers of positive clones per cloned fragment when compared to TIG cloning, a wider range of fragments were ultimately successful. For Intron I, one clone of AT01OP, one clone of AM01WS, re and two clones of AM01WB were generated. Successful Intron II clones included three from AM01WB, two from AM03KB, and one each from AT01OP and AM03SK.

Sequencing

Initial results utilizing direct ^{35}S sequencing of PCR products generated from Intron II gave poor results as compared to direct sequencing of rDNA PCR products (see Chapter 1). Sequence was only readable for approximately 125 bp and typically started almost 75 bp from the primer. Two separate PCR products from the second Intron of AM06SR were sequenced from both ends using internally nested primers (Figure 12). Six separate PCR products from NLCI04 were directly sequenced from both ends, as were three PCR products from AM03SK. Two PCRs of NL01VR and one of NL01JB were done from the 3' end of Intron II. All direct sequencing products were compared to cloned products, where applicable, and were consistent in all cases.

Automated cycle sequencing of cloned products using the T3 or T7 sites was attempted for a number of clones. In all cases, reads were better from T7 and complete coverage (overlap) of sequences was not achieved. For AM06SR Intron II, the three positive clones (5a2, 5b2, and 5c2) were sequenced from each end yielding

800, 1050, and 190 bp of sequence. 900 bp from two clones from intron II NLCI04 was completed. 860 bp of AM03SK Intron II was completed as well as about 400 bp of the Intron I fragment. The AM03SK Intron I fragment was the only Intron I fragment done by automated cycle sequencing. Only approximately 400 bp from clone 27.6 of AN10BO Intron II was readable on two separate reactions. 755 bp of AM03RB, 900 bp of AM03JR clone 16.10, and 870 bp of AN10BO clone 28.9 were also done.

Manual ^{33}P cycle sequencing was used on most of the Intron I (the AM06SR and AN10BO Intron I clones were only done by automated sequencing), and all of the Intron II clones examined. Manual cycle sequencing of individual clones was combined with sequences generated by direct ^{35}S sequencing and automated sequencing of clones. In all cases where multiple sequence methods were used to generate sequences from the same clone, the results were identical. In addition only one base difference (in AM03SK and resolved with further direct sequencing) was noted between direct sequences of PCR products and sequences of clones.

Sequence Composition and Length

In total, 392 bases from six separate individuals were completed for the Intron I region (Figure 15). Twenty four individual clones were sequenced in the Intron II region (Appendix I); Twenty one were *Alaria* clones encompassing eight individuals and three single-clone *Nereocystis* individuals. The *Alaria* Intron II fragment was an average of 1286 bp long (range: 1255 - 1370 bp).

Figure 15. Alignment of *Alaria* DNA sequences of the cloned PCR fragment including Intron I. Individual identities are as in Figure 11. The vertical slash ('pipe') symbol indicates the start of the *Costaria costata* sequence from Bhattacharya et al. (1991) and also indicates the predicted splice site. A 'dot' (.) indicates identity with the AN10BO sequence shown, while a 'dash' (-) indicates an insertion or deletion added to allow alignment. The gap in the sequence represents the unknown portion of the intron.

```
AN10BO          5'-GGCTTTGCCGGTGACGACGCGCCGCGCGCGGTGTTCCCTTCCA
AT01OP             ........................................
AM03JR             ........................................
AM01WS             ........................................
AM01WB             ........................................
AM06SR             ........................................
Costaria costata

TCGTAGGGCGCCCCAAGCACCCCGGAATCATGGTGAGTCCTTTTCAAACAAAACGAACCC
............................................................
............................................................
............................................................
.....................................................G......
.............................................A..............

TCGTTGCGGCCCACGCAAGAAGAATTGACACCTGCCTCGAGGATAAGAGGCGATCAGCAC
............................................................
............................................................
............................................................
...............-............................................
............................................................

TCTCCCTCGCCCGGCGCAA        TTTTTCTCAAATGTTTTCTCCCTTGCTTTCTGGT
...................        ...............................
...................        ...............................
...................        ...............................
...................        ..C............................
...................        ................G..............

CTTGTTCTTGCTCTGGTTTAATATGAGAACGAACACCCATTTATTCCCTAATGCGCAACC
.........................T..................................
.........................T..................................
.........................T..................................
............................................................
.........................T..................................

CTCCCACCACGATTCTACTTAATTTGCTTTCGATGTTTTCAGGTCGGCATGGACCAAAAG
............................................................
.............................................G..............
............................................................
..............C.............................................G...
............................................................G...
                                             |..............G...

GACGCTTACGTGGGCGATGAAGCCCAGTCCAAGCGAGGCGTCCTCACCC-3'AN10BO
.............................................    AT01OP
.............................................    AM03JR
.............................................    AM01WS
.............................................    AM01WB
.............................................    AM06SR
.............................................    Costaria costata
```

Protein Sequence and Splice Sites

Alaria and *Nereocystis* sequences from either end of the Intron II-containing fragment were translated, and aligned with the published *Costaria costata* sequence from Bhattacharya et al. (1991) (Figure 16).

Splice sites as determined by the "GT-AG" rule (Breathnach and Chambon 1981, Mount 1982), *Arabidopsis* consensus Intron sequences (Brown 1996), as well as both neural network methods (Reese and Eckman 1996, Solovyev 1994) agreed with a donor (5') splice site six basepairs upstream of that reported in Bhattacharya et al. (1991) (Figure 17). The existence of a stop codon in all of the *Alaria* individuals just before the splice site for *Costaria* from Bhattacharya et al (1991) also argues for a different splice site in the case of *Alaria*.

NetPlantGene (Hebsgaard et al. 1996) gave other sites towards the middle of the fragment higher scores than the sites predicted above for Intron II. The higher-scoring sites were, however, polymorphic amongst the *Alaria* individuals.

To determine the amino acid sequence and splice sites for Intron I, the *Achlya bisexualis* actin sequence (Bhattacharya et al. 1991) as well as the translated *Alaria* and 3' *Costaria* sequences were aligned (Figure 18). All of the splice site determination methods mentioned above either gave no prediction or agreed with the splice site for Intron I being identical to that in Bhattacharya et al. (1991).

Figure 16. Alignment of regions surrounding the second actin intron. Sequences are from clones 5a2 (AM06SR, *Alaria*), 5b and 5e (NLCI04, *Nereocystis*), and Bhattacharya et al. (1991). Protein regions are translated from the above DNA sequences in the case of *Alaria* or *Nereocystis*, and Bhattacharya et al. (1991) for *Costaria*. A 'dot' (.) indicates identity with the *Alaria* sequence shown, a 'dash' (-) indicates an insertion or deletion added to allow alignment, while the 'hats' (^^^) indicate the position of the intron based on Bhattacharya et al. (1991). Amino acids are shown using the single-letter code with a star (*) denoting stop codons.

```
Alaria                GAGAAGTCGTACGAGCTCCCCGACGGAAACGTCATCGTCATCG
Costaria              .....................................T....
Nereocystis           .....................................T....
Alaria - p.            E  K  S  Y  E  L  P  D  G  N  V  I  V  I
Costaria - p.          E  K  S  Y  E  L  P  D  G  N  V  I  V  I
Nereocystis - p.       E  K  S  Y  E  L  P  D  G  N  V  I  V  I
```

```
GAAACGAGCGCTTCCGTTGCCCCGAGGTATGATACAAAATG------------TAC
.............................C.T^^^^^^^^^^^^^^^^^^^^^^^^
..G...........................TAAATATGATCGTTAATCATTACCA
G  N  E  R  F  R  C  P  E  V  *
G  N  E  R  F  R  C  P  E  V  R
G  D  E  R  F  R  C  P  E  V  C  K  Y  D  R  *
```

```
    CAATTGTGCCTCCGTTCGTCCTCTC---TTTTTTTTCAACGTGTTTTGTTCAATCC
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
    GTGTCTGATCTTGTCTCTCTCCCCCCCGTCTCTTTGCTGCAACGTGTTTCGACAAT
                                    F F/S S S/T C  F  V  Q  S

          *  S  C  L  S  P  P  R  L  F  A  A  T  L  F  D  N
```

```
TTGCAGGTGCTGTTCCAGCCCTCGTTCATCGGAATGGAGTCCTCGGGCATCCACGACTGC
^^^^^^^^^^^.................................................
CC.......T...................................................
 L  Q  V  L  F  Q  P  S  F  I  G  M  E  S  S  G  I  H  D  C
             F  Q  P  S  F  I  G  M  E  S  S  G  I  H  D  C
 P  Q  V  L  F  Q  P  S  F  I  G  M  E  S  S  G  I  H  D  C
```

```
ACGTTCAAGACGATCATGAAGTGCGACGTCGACATCC   Alaria
.....................................   Costaria
.....................................   Nereocystis
 T  F  K  T  I  M  K  C  D  V  D  I     Alaria - p.
 T  F  K  T  I  M  K  C  D  V  D  I     Costaria - p.
 T  F  K  T  I  M  K  C  D  V  D  I     Nereocystis - p.
```

Figure 17. Alignment of regions surrounding the second actin intron. Sequences are from clones 5a2 (AM06SR, *Alaria*), 5b and 5e (NLCI04, *Nereocystis*), and Bhattacharya et al (1991). Protein regions are translated from the above DNA sequences in the case of *Alaria* or *Nereocystis*, and Bhattacharya et al. (1991) for *Costaria*. The 'pipe' or 'vertical slash' symbol (|) indicates the boundary of the proposed splice site in the case of *Alaria* and *Nereocystis*. A 'dot' (.) indicates identity with the *Alaria* sequence shown, a 'dash' (-) indicates an insertion or deletion added to allow alignment, while the 'hats' (^^^) indicate the position of the *Costaria* intron based on Bhattacharya et al. (1991). Amino acids are shown using the single-letter code with a star (*) denoting stop codons.

```
Alaria              GAGAAGTCGTACGAGCTCCCCGACGGAAACGTCATCGTCATCG
Costaria            .........................................
Nereocystis         .........................................
Alaria - p.          E  K  S  Y  E  L  P  D  G  N  V  I  V  I
Costaria - p.        E  K  S  Y  E  L  P  D  G  N  V  I  V  I
Nereocystis - p.     E  K  S  Y  E  L  P  D  G  N  V  I  V  I
```

```
GAAACGAGCGCTTCCGTTGCCCCGAG|GTATGATACAAAATG------------TAC
.......................... ...C.T^^^^^^^^^^^^^^^^^^^^^^^^
..G.....................|.....TAAATATGATCGTTAATCATTACCA
G  D  E  R  F  R  C  P  E | V  *
G  D  E  R  F  R  C  P  E | V  R
G  D  E  R  F  R  C  P  E | V  C  K  Y  D  R  *
```

```
     CAATTGTGCCTCCGTTCGTCCTCTC---TTTTTTTTCAACGTGTTTTGTTCAATC
     ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
     GTGTCTGATCTTGTCTCTCTCCCCCCCGTCTCTTTGCTGCAACGTGTTTCGACAA
                                    F F/S S S/T C  F  V  Q  S

           *  S  C  L  S  P  P  R  L  F  A  A  T  L  F  D  N
```

```
CTTGCAG|GTGCTGTTCCAGCCCTCGTTCATCGGAATGGAGTCCTCGGGCATCCACGACT
^^^^^^^ ^^^^^^.............................................
TCC....|...T.............................................
  L  Q | V  L  F  Q  P  S  F  I  G  M  E  S  S  G  I  H  D
               F  Q  P  S  F  I  G  M  E  S  S  G  I  H  D
  P  Q | V  L  F  Q  P  S  F  I  G  M  E  S  S  G  I  H  D
```

```
GCACGTTCAAGACGATCATGAAGTGCGACGTCGACATCC   Alaria
.....................................     Costaria
.....................................     Nereocystis
C  T  F  K  T  I  M  K  C  D  V  D  I     Alaria - p.
C  T  F  K  T  I  M  K  C  D  V  D  I     Costaria - p.
C  T  F  K  T  I  M  K  C  D  V  D  I     Nereocystis - p.
```

Figure 18. Alignment of regions surrounding the first actin intron. Sequences identities are as described in Figure 11. Blank regions are unknown. Protein regions are translated from the above DNA sequences in the case of *Alaria* and Bhattacharya et al. (1991) for *Costaria* and *Achlya*. The 'pipe' or 'vertical slash' symbol (|) indicates the boundary of the proposed splice site for *Alaria* and the published splice site for *Costaria* and *Achlya* (Bhattacharya et al. 1991). A 'dot' (.) indicates identity with the *Alaria* sequence shown. Amino acids are shown using the single-letter code with a star (*) denoting stop codons.

```
AN10BO          5'-GGCTTTGCCGGTGACGACGCGCCGCGCGCGGTGTTCCCTTCCA
AT01OP             ........................................
AM03JR             ........................................
AM01WS             ........................................
AM01WB             ........................................
AM06SR             ........................................
Costaria costata
Alaria - p.         G  F  A  G  D  D  A  P  R  A  V  F  P  S
Costaria - p.
Achlya - p.         G  F  A  G  D  D  A  P  R  A  V  F  P  S
```

```
TCGTAGGGCGCCCCAAGCACCCCGGAATCATG|GTGAGTCCTTTTCAAACAAAACGAACC
................................|........................
................................|........................
................................|........................
................................|..................G.....
................................|.............A..........

I  V  G  R  P  K  H  P  G  I  M  | V  S  P  F Q/K T K/R R  T

I  V  G  R  P  K  H  P  G  I  M  |
```

```
CTCGTTGCGGCCCACGCAAGAAGAATTGACACCTGCCTCGAGGATAAGAGGCGATCAGCA
............................................................
............................................................
............................................................
............................................................
............................................................

 L  V  A  A  H  A  R  R  I  D  T  C  L  E  D  K  R  R  S  A
```

```
CTCTCCCTCGCCCGGCGCAA
....................
....................
....................
....................
....................

 L  S  L  A  R  R
```

```
TTTTTCTCAAATGTTTTCTCCCTTGCTTTCTGGTCTTGTTCTTGCTCTGGTTT
.....................................................
.....................................................
.....................................................
..C..................................................
................G....................................

  F  S  Q  M  F  S  P  L  L  S  G  L  V  L  A  L  V
```

```
AATATGAGAACGAACACCCATTTATTCCCTAATGCGCAACCCTCCCACCACGATTCTACT
......T..............................................................
......T..............................................................
......T..............................................................
.....................................................................C...
......T..............................................................

*  Y E/V N  E  H  P  F  I  P  *  C  A  T  L  P  P  R  F H/Y
```

```
TAATTTGCTTTCGATGTTTTCAG|GTCGGCATGGACCAAAAGGACGCTTACGTGGGCGAT
.......................|.....................................
.......................|......G..............................
.......................|.....................................
.......................|.............G.....................
.......................|.............G.....................
                       |.............G.....................C
L  I  C  F  R  C  F  Q |  V  G  M  D  Q  K  D  A  Y  V  G  D
                       |  V  G  M  D  Q  K  D  A  Y  V  G  D
                       |  V  G  M  D  Q  K  D  A  Y  V  G  D
```

```
GAAGCCCAGTCCAAGCGAGGCGTCCTCACCC-3'   AN10BO
.............................        AT01OP
.............................        AM03JR
.............................        AM01WS
.............................        AM01WB
.............................        AM06SR
.............................        Costaria costata
 E  A  Q  S  K  R  G  V  L  T        Alaria - p.
 E  A  Q  S  K  R  G  V  L  T        Costaria - p.
 E  A  Q  S  K  R  G  V  L  T        Achlya - p.
```

BLAST Searches

tBLASTn results for predicted amino acid regions of the exon regions near Intron II gave high scores for a number of actins from GENBANK including exact identity to the only two other brown algal actins known (Figure 19).

Intron I tBLASTn results from all ORFs gave no results smaller than the 0.05 smallest-sum probability cut-off.

Intron II tBLASTn searches of all ORFs gave only three results smaller than the 0.05 smallest-sum probability. The best match (smallest probability) was to a rice (*Oryza sativa* L.) receptor kinase-like protein (GENBANK #U72724) at the 0.0076 smallest-sum probability level, still well above the 2.2×10^{-32} to 1.2×10^{-25} of the actin coding region probabilities tested.

Discussion

PCR and Cloning

Amplification of genomic DNA from *Alaria* and *Nereocystis* individuals was difficult. Although significant levels of high molecular weight genomic DNA was visible on EtBr-stained agarose gels, successful PCR from these extracts was difficult. Many methods were tried, most with sporadic or inconsistent results. Optimization of PCR conditions was never precisely achieved. Others have reported strong inhibition of *Taq* polymerase by a number of phaeophytes as well as other

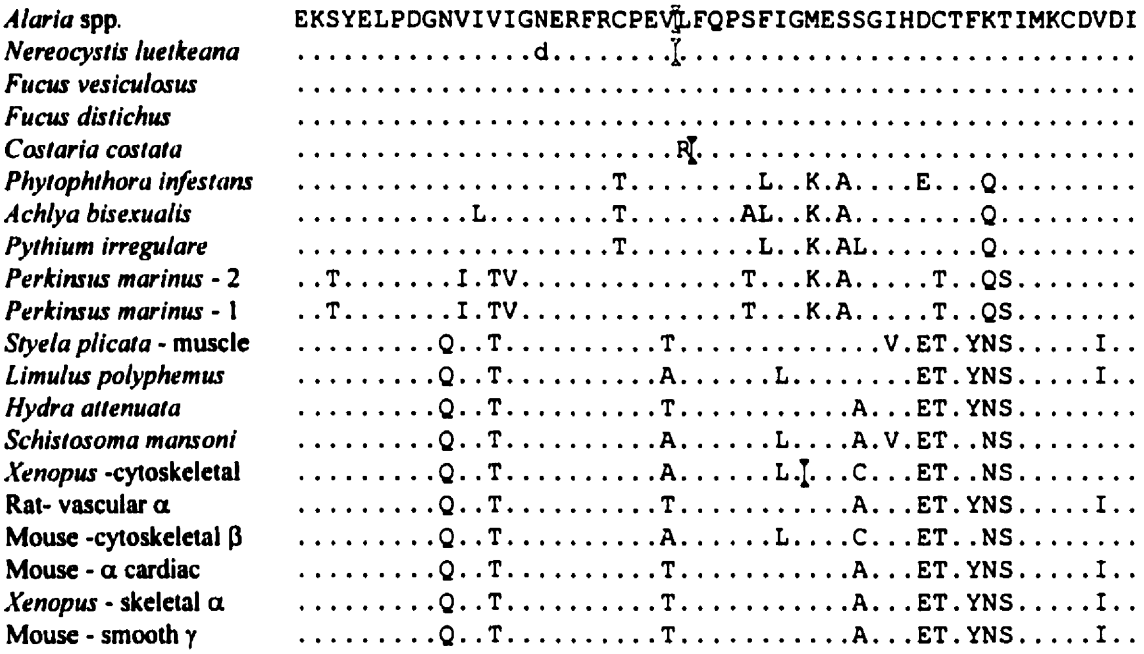


Figure 19. Protein alignment of actin fragments. Both *Alaria* and *Nereocystis* are shown as well as the closest 18 sequences based on tBlastn searches. A dot (.) indicates identity with the *Alaria* sequence. Lowercase letters represent amino acids for which the difference is based on only one DNA sequence of one clone. (I) represents published intron sites. (I) represents splice site determined as described in the text. All other sequences either did not have introns, were not described in the original paper, or were in other regions of the gene.

marine macrophytes including *Alaria nana* (Jin et. al 1997). In the end, CsCl purified DNA from an earlier survey (Mróz 1989) of the same species was the most fruitful in generating genomic DNA amplifications that yielded good sequencing results.

If the actin gene of kelp in general is truly single-copy, as reported for *Costaria costata* in Bhattacharya et al. (1991), then the difficulty PCRing the fragments may also reflect the low number of actin copies per given amount of genomic DNA. If this were true, increasing both the primer and template concentrations might have improved the yields seen, but no improvement was seen.

Alternately, the PCR reaction itself may have been inhibited by some factor that was removed in CsCl purification but not in any of the other methods tried. Inhibitors of PCR have been reported from studies using a number of different extraction techniques (Johnson et al. 1995, Wiedbrauk et al. 1995, Kreader 1996). Wide ranges of concentrations of all components were tried in an attempt to mitigate potential inhibitors, with no success. The possible inhibitor may have affected the PCR at all template concentrations and component combinations attempted.

Possibly, the low yields were due to the inefficient or low sensitivity of the particular primer pairs used. For unknown reasons in certain instances, individual primer pairs have been shown to vary widely in their ability to PCR the same fragment, even though all of the primers were exact matches (He et al. 1994).

Predicted lengths for the introns based on *Costaria* clone sizes were about 1.3 kbp each for a total of 2.6 kbp in length (Bhattacharya et al. 1991). Average Intron II

sequence lengths reported here were just under 1.3 kb and Intron I clones were about the same size based on agarose gels.

Splice Sites

Splice sites for the first intron of *Alaria* actin were identical to that reported for *Costaria costata* based on all of the prediction methods for splice sites attempted.

Proposed splice sites for the second intron of actin in *Alaria* and *Nereocystis* based on a number of prediction methods were different from that reported in *Costaria costata* (Bhattacharya et al. 1991). This upstream splice site would give a change from arginine to leucine between *Alaria/Nereocystis* and *Costaria*. The nearby amino acid code (with | denoting the splice site) after splicing would be PE|VLFQ for *Alaria* or *Nereocystis* and PE|VRFQ for *Costaria*.

The site that was predicted in Bhattacharya et al. (1991) would give two different codons at the site ('stop', arginine, and cysteine) between *Alaria*, *Nereocystis*, and *Costaria*. The nearby amino acid code would then be PEV*|FQ for *Alaria*, PEVC|FQ for *Nereocystis*, and PEVR|FQ for *Costaria*. Although it is not clear from Bhattacharya et al. (1991) how the splice sites were determined from *Costaria*, the study itself used cDNA clones and not genomic DNA to determine the coding sequence. An error in the cDNA generation step, caused by the low fidelity of Reverse Transcriptase (Sambrook et al. 1989), may have caused the sequence difference seen. Alternately, the taxa may splice their actin at slightly different sites.

The possibility that the actin sequences seen in all of the *Alaria* individuals examined were truncated forms of actin or pseudogenes could not be ruled out.

Comparisons of Exon Regions

DNA sequence generated for the regions on the 5' end of the proposed first intron from *Costaria costata* gave identical translated amino acid sequence to that reported for the water mold *Achlya bisexualis* (Bhattacharya et al. 1991). All *Alaria* individuals examined had identical sequences in the same region. The 3' end had two synonymous silent-site changes, one change in three of the individuals, and one change only in one individual. A non-synonymous change was seen in one individual giving a methionine to valine change. Only one clone was sequenced for each individual, however, so the methionine to valine change seen in one individual may represent an error from *Taq* amplification.

A predicted lariat branch point consensus sequence (YNYTRAY) was seen 51 to 54 bases upstream of the 3' splice site in the second intron. The site was present in both *Nereocystis* and *Alaria*, although AM06SR was missing the 5' pyrimidine. The first intron in all *Alaria* individuals sequenced has an exact lariat branch point consensus sequence, 50 bp from the 3' splice site.

Given the new proposed splice site, the exon sequence bracketing the second intron for *Alaria* and *Nereocystis* was nearly identical to that reported for *Costaria*. Outside of the splice site variation, only one non-synonymous change was noted in the one *Nereocystis* exon region examined. In *Nereocystis*, an aspartic acid replaced

an asparagine (two similar amino acids, with only aspartic acid negatively charged) in *Costaria* or *Alaria*.

Comparison of Actin Intron II Regions

Intron regions were appreciably different between the *Nereocystis* individuals and the *Alaria* individuals. Within five basepairs of the 5' splice site, the sequences of the two different genera are impossible to visually align. Only the last 25 basepairs of the intron near the 3' splice site could be aligned. Within the eight *Alaria* individuals examined encompassing members of three local species, the entire intron could be aligned if spaces were utilized.

Clone-to-clone variation was low when comparing the four individuals from which multiple clones were sequenced. The point mutant variation among clones per thousand basepairs sequenced was: 0.27 basepairs for AM06SR, 0.62 basepairs for AM03RB, 0.78 basepairs for AM03JR, and 0.89 basepairs for AN10BO. Because only two clones were generated for AM01WB, the 12 differences between them could not be resolved. The error rate for AM01WB point mutants was therefore 4.88 basepairs per thousand sequenced. Only two indels were seen between clones of the same individual. A single extra 'A' was found in one of the five clones from AN10BO. Four bases were missing from one of the clones of AM03RB, making that clone identical to the two AN10BO clones at the same position.

The point mutation differences between clones are of similar magnitude to published Taq error rates (0.285 per thousand basepairs, Tindall and Kunkel 1988).

Therefore the differences between clones, with the exception of the indels and high rates in AM01WB, are believed to be mainly PCR errors, and not necessarily allelic differences within the individuals.

The data generated herein represents the first non-ribosomal and non-coding nuclear sequences generated in any member of the Laminariales. As such this data is not as constrained by questions of selection, uniparental inheritence, or incomplete concerted evolution as other studies in the Laminariales. The sequences contained recognizable splice sites and other intron-specific features as well as measurable variation between individuals. Given enough individuals, these regions identified herein could provide a wealth of population level inferences for *Alaria* specifically, and most likely for other kelp as well.

Chapter IV

Actin Introns as Markers for Phylogeography in *Alaria*.

Introduction

A number of studies have shown that individual kelp plants have extremely limited dispersal. For example, settled male and female gametophytes, arising from meiospores released by sporophytes, must be close enough to allow fertilization of a non-motile egg by a swimming male sperm (Norton 1992, Santelices 1990). However, a spacing of approximately 1 mm was the maximum distance gametophytes could be separated in petri dish cultures of *Macrocystis* and still allow fertilization (Reed 1990). This spacing was also confirmed when the effective range of kelp pheromones was found to be 1 mm (Maier and Müller 1990). Similarly, Sundene (1962) only found new *Alaria esculenta* sporophytes within 10 m of his transplants to an area where *Alaria* is not usually found. Anderson and North (1966) reported finding new *Macrocystis* sporophytes mostly within 5 m of an isolated individual plant, while *Postelsia* seems to only disperse at distances of 1.5 – 3 m (Dayton 1973). Kelp apparently cannot reattach once adrift, so all sporophytes at a given location must have arisen from gametophytes present at that location (Santelices 1990).

How then do kelp invade new habitats or re-colonize after disasters if their dispersal is so limited? Interestingly, juvenile kelp have been found up to 5 km from the nearest adult stand (see Druehl 1981) and spores of many types of ephemeral

marine macrophytes have been recorded 30 – 35 km from the nearest source off North Carolina (Amsler and Searles 1980).

In southern California, heavy storms in the early 1980s denuded many kelp beds, but vigorous recruitment was observed subsequent to these storms (Ebeling et al. 1985). One hypothesis put forward to explain the rapid return of the kelp is that sporogenic drift material can densely inoculate the substrate as they float past the bottom (Anderson and North 1966). Although competent material has been observed leaving behind a patch of recruits (Fager 1971, Dayton et al. 1984), and holding fertile sori on the substrata will produce a large patch of gametophytes (Dayton et al. 1984), other studies dispute the hypothesis that fertile drifting individuals facilitate long-range dispersal (Reed and Ebeling 1991).

Reed et al. (1988) found that dispersal of *Pterygophora californica* over long distances was not as patchy as would be expected from drift-based dispersal only. This was shown to correlate well with long-range dispersal caused by storms keeping kelp meiospores suspended in the water column longer than usual. In addition, Reed et al. (1988) along with others (Anderson and North 1966, Palmer and Strathmann 1981, Reed and Ebeling 1991, Underwood and Denley 1984) have suggested that larger numbers of adults may help to insure sufficient densities of male and female gametophytes at distant sites by synchronizing dispersal or producing huge numbers of spores.

Because of the heteromorphic alternation of drastically different sized generations (the sporophyte and gametophyte) found in kelp, field observations of

many stages in their life histories are incomplete. This has often meant that studies of dispersal and recruitment have involved recording the appearance of juvenile sporophytes, and extrapolating back to determine the mechanisms facilitating their arrival (Anderson and North 1966, Fager 1971, Dayton 1973, Amsler and Searles 1980)

Long-range dispersal has been examined in barnacles which, like kelp, posses a planktonic stage (Palmer and Strathmann 1981, Underwood and Denley 1984). However, care must be taken in drawing parallels between barnacles and kelp since young barnacles can choose settlement sites, a trait which has not been demonstrated in kelp. Kelp and barnacles are also planktonic at different stages in their life cycle. Kelp release planktonic meiospores which each produce either a male or female filamentous gametophyte which subsequently produce sperm and eggs. In barnacles, fertilized zygotes are planktonic. The alternation of generations in kelp requires that meiospores settle in a density high enough to allow the gametes to find each other for fertilization (Maier and Müller 1990). The need for gametophytes to be near each other means that kelp gametophytes not only have to reach a new location, but they must also arrive in sufficient numbers to insure the eventual production of sporophytes on the substratum. In barnacles, small-scale hydrodynamic forces were found to be the most important influence on where planktonic stages settled (Underwood and Denley 1984). The wide dispersal observed for barnacles has been postulated to act like a form of evolutionary insurance that may mediate local extinction events (Palmer and Strathmann 1981).

Direct observation of migrations in kelp have been incomplete because the original source of any settled spores is not known. The most practical method for determining gene flow, and indirectly the source of individual patches and dispersal distances, is to look for alleles shared only by members of potential subpopulations (Slatkin 1985a, Slatkin and Barton 1989) or generate F-statistics based on DNA sequence data (Lynch and Crease 1990, Hudson et al. 1992). Both methods allow an analysis of the subdivision of heterozygosity (Wright 1951) and can therefore be used to describe the movement of kelp meiospores. Genetic similarity coefficients have also proved useful in population genetic structure (Kusumo 1998). Hypotheses can then be drawn that mimic the stepping-stone models (Slatkin 1985a) used for gene flow analyses. Kelp seem to be an ideal fit for stepping stone models due to their restriction to the narrow strip of the photic zone along the coast.

In analyzing natural populations and their subdivisions, often the goal is to determine if gene flow is high enough to prevent local ecotypes from forming through drift and/or selection. Theoretical studies have shown that only one individual entering the local population from the larger overall population per generation is enough gene flow to prevent genetic differentiation (Maruyama 1972). Selection, however, may be able to overwhelm gene flow if the ratio of mutation to selection is greater than one (Slatkin 1987). The result is that the importance of gene flow evolutionarily depends on the product of the population size (N) and the migration rate (m) (Slatkin 1985a). If this product (Nm) is greater than one, gene flow is

overcoming drift and selection in the subpopulation, thereby essentially preventing ecotype formation (Slatkin 1985b).

Comprehensive analyses of gene flow, however, demand large numbers of individuals from many apparent populations. Perhaps 20 individuals equally drawn from various hypothesized population subdivisions are needed to measure gene flow (Weir 1990). When sample sizes are smaller, other methods such as minimum spanning trees can be employed to discern phylogeography. Although direct measurement of gene flow is not possible using minimum spanning trees, traits which are not necessarily inherited in a tree-like manner can be utilized. In sexually reproducing taxa, traits or alleles can be acquired by recombination as well as sexual reproduction itself, both of which are not restricted to identity by descent.

To prevent spurious results arising from selection, neutral mutations (those whose effects on fitness are null) are generally needed for population studies (Kimura 1968). Various genic or chromosomal regions including 3' flanking regions, fourfold degenerate sites, introns and pseudogenes all have low selection rates and therefore serve as promising regions of neutral mutation (Li and Graur, 1991).

Presumably, by choosing untranslated regions, most non-neutral selection can be avoided. Mitigating the effects of recombination which causes structure not directly inherited by descent is more problematic. Ancient recombination is nearly impossible to detect but less important since gene flow can only be detected over the last few hundred generations (Slatkin 1985b). More recent recombination can

potentially be detected and is more likely to mask or confuse determinations of identity by decent (Aquadro et al. 1986, Templeton et al. 1992).

Because of both the supposed lack of selection and the accesibility of flanking gene regions that evolve more slowly, introns are believed to be excellent tools for genetic studies of populations and population substructure (Li and Graur 1991).

Actin introns have been used previously for gene flow determinations in Humpback Whales (Palumbi and Baker 1994), wherein the results were compared to earlier mtDNA studies. Combining both the mtDNA and actin intron results made it possible to differentiate gene flow rates of the male and female whales. The purpose of my study was to see if kelp actin introns can resolve relationships and suggest gene flow in *Alaria* spp. Although not enough individuals were sequenced to properly measure gene flow using Fst statistics as was the case with the whale study, insights into the relationships between distant stands of *Alaria* resulted. This study also aimed to shed light on the confusing relationships between the three common local *Alaria* species. Previous work by Mróz (1989) has shown that the three morphologically defined species actually appear to represent a RFLP (Restriction Fragment Length Polymorphisms) as well as a morphological cline. My study adds further evidence towards clarifying the relationships between the three morphotypes presently divided as species.

Determinations of gene flow or phylogeography are important not only for quantifying actual dispersal distance for a species but also as a baseline for investigations of other *in situ* ecological processes. On a regional scale, questions of

why kelp are or are not found at a certain locale hinges upon a determination of the capacity for dispersal. Therefore, if the dispersal distance of a species is known, the potential dispersal can serve as the null hypothesis for many ecological investigations.

A number of investigations have addressed the patchy distribution of kelp. Many authors have looked at the various causes for patchy kelp distribution (Dayton et al. 1984, Johnson and Mann 1988, Munda 1992). Currents, temperature, and other physical factors are invoked to explain the distributions, but without any quantification of dispersal potential, conclusions are difficult. Druehl (1967) was able to show that two similar *Laminaria* species have differing environmental tolerances and may face dispersal restrictions. Being able to quantify the actual potential for dispersal of the two *Laminaria* species would allow an integration of local oceanographic data and ecological forces to better understand their present distribution.

Materials and Methods

Samples, representing all three of the recognized local *Alaria* species (*Alaria marginata* Postels et Ruprecht, *Alaria tenuifolia* Setchell, and *Alaria nana* Schrader) as well as three *Nereocystis luetkeana* (Mertens) Postels et Ruprecht individuals were collected from a total of 13 sites (Figure 11).

DNA was extracted and sequences were generated as described in Chapter III. All sequences were aligned and manipulated using the computer program ESEE (Cabot and Beckenbach 1989).

A consensus sequence, based on multiple clones, was developed for each individual if multiple clones were available. The most common base at each site was deemed the consensus base for that site.

A number of potentially non-neutral sites and stretches of sequence were removed from the data set. Coding regions, as determined in Chapter III, were removed. Except when determining possible recombination events, inserts and deletions were removed from the data set and not scored. Sequence sites were also removed from the data set if only two clones of any individual were sequenced and the identity of the bases at that site in the two clones varied. Regions corresponding to the lariat consensus sequence and all sequences downstream within the intron as described in Chapter III were also removed from the analysis. In addition, a short 8 basepair region at the 5' end of the intron that was identical in *Nereocystis luetkeana* was removed.

Conversion to FASTA or interleaved format for PHYLIP was performed using a series of computer programs written and distributed by Andy Beckenbach (Simon Fraser University). SAV2FAS.EXE was used to convert ESEE documents to FASTA format. FA2FEL.EXE was used to convert from FASTA to PHYLIP format. To generate base content and pairwise comparisons of the sequences, COMPARE.EXE was used.

Most of the phylogenetic analyses were done using the programs provided in Joe Felsenstein's PHYLIP 3.573c package (Felsenstein 1993). Programs used were DNAPARS.EXE 3.572c , DNADIST.EXE 3.573c, DNAML.EXE 3.573c, NEIGHBOR.EXE 3.5, and CONSENSE.EXE 3.573c. For bootstrapping, SEQBOOT.EXE 3.5c was used with 1000 replicates for parsimony or neighbor-joining trees, and 100 replicates for maximum-likelihood trees.

Possible recombination sites were determined using a number of methods. Visual inspection of a figure showing sequential phylogenetically informative site identities was used to scan for possible large-scale recombination events. An analysis of potential convergent homoplasies within the parsimony tree was done by eye to meet the recombination site criteria of Templeton et al. (1992) and Aquadro et al. (1986). Several population genetic parameters as well as possible evidence for recombination were determined using Jody Hey's SITES package of programs (Hey and Wakeley 1997; http://heylab.rutgers.edu). Tajima-D statistics were also confirmed using an ANSI C++ program compiled for 32-bit x86 processors with

Visual C++ 6.0™ as well as for Silicon Graphics CC compiler for IRIX 6.2 (Appendix III).

Minimum Spanning Trees were generated by hand for the eight individuals using the tree generated by DNAPARS.EXE and the distance matrices generated by DNADIST.EXE. Kimura 2-parameter (Kimura 1980), maximum likelihood from DNAML.EXE, and Jukes Cantor (Jukes and Cantor 1969) corrections were each used to generate separate distance matrices. Arbitrarily, all vertices within three point mutations or two indels (insert or deletion events as compared to the rest of the trees) of the minimum tree were included.

Results

DNA Sequence Alignments

Sequences generated as described in Chapter III, were aligned and a number of regions of probable non-neutral and problematic sequences were removed from the alignments. Two alignments were generated from the reduced dataset. One alignment (Appendix IV) includes the inserts and deletions (indels) added to allow sequence alignment. The other alignment (Figure 20) has the insert and deletion regions as well as any ambiguously aligned regions removed.

The alignment that incorporated indels (Appendix IV) was used for three separate analyses of possible recombination within the intron itself. Polymorphic sites in which at least two individuals shared an identity that was different from the

most common allele (the shared derived character state) were numbered consecutively. Possible homoplasies were determined based on maximum parsimony results and noted on the alignment.

Potential Recombination

Three separate analyses were performed to discern any recombination events between individuals or their recent ancestors within the actin intron examined. In an attempt to detect large scale events, a figure representing the shared derived character state or ancestral state of all 45 phylogenetically informative sites presented sequentially, was constructed (Figure 21). The figure was adjusted to put taxa that shared the most sites based on distance matrixes, closest together wherever possible. Recombination events would appear as runs of shared derived characters from individuals not closely related in the distance matrix. Although individual shared derived characters did not always follow the distance matrix (see for example KB vs. OP), no obvious runs of multiple anomalous sites were apparent.

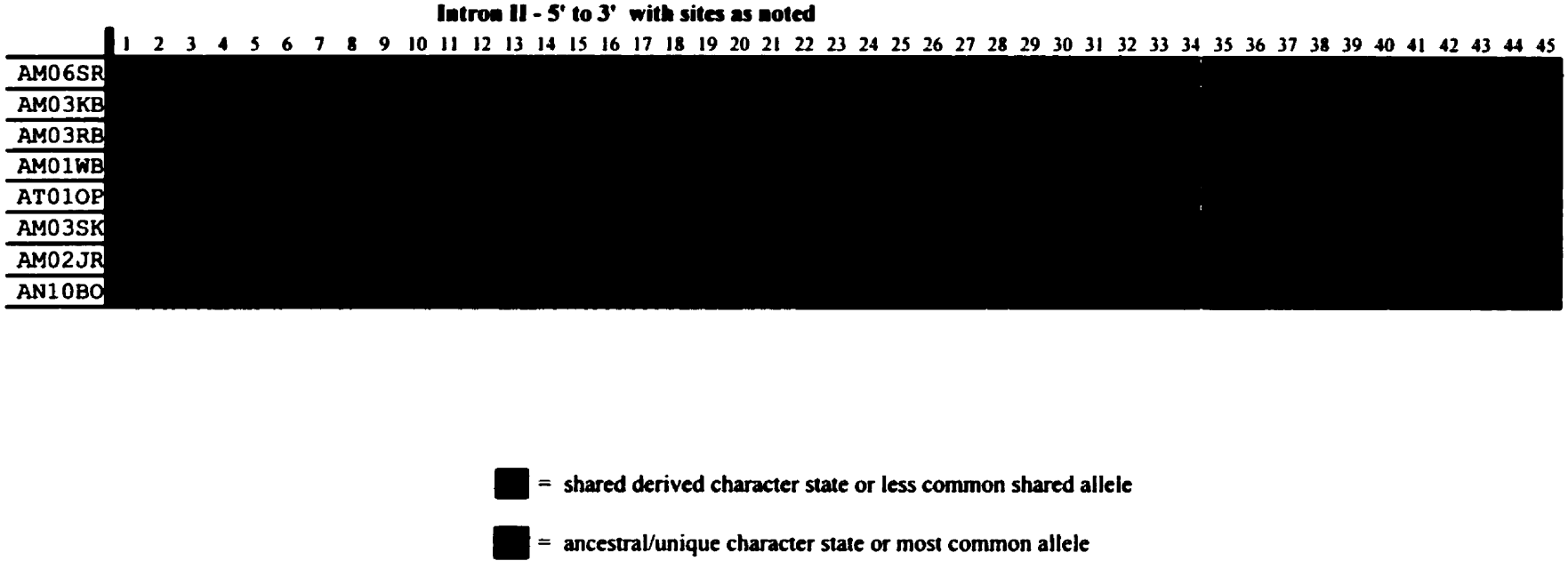
Figure 20. Alignment of *Alaria* actin intron II regions based on the splice sites from Bhattacharya et al (1991). Sequences are identical to Appendix IV except that inserts and deletions were removed. A 'dot' (.) indicates identity with the AM06SR sequence shown while lower case letters designate sites for which data from only one clone was determined.

```
AM06SR  ATACAAATGACTGTGACAAAGCGCTCTTGTGATGGTGTAAGTAACAGATAAC
AM03KB  ............GT....................A..T............
AM03RB  ....................................................
AM01WB  ....................................................
AT01OP  ....................................................
AM03SK  ..t.........gt............t.........................
AM03JR  ..T.........GT............T.........................
AN10BO  ..T.........GT............T.........................

ACTGTCCGAAAATGTAAGCACCGTTCTGTACGTGAGTATACAGGCGGTGCTAATTTTCCG
.........................T.................A.......T........
.........................T..................................
.T.......................T..................................
.........................t..................................
....................................................................
...................T...A..T..................................
...................T...A..T..................................

GACAGAAGTACACGGCCCAGAGCACTCTTGTTATACCTGGTATAGTCTTATTTTCAATTG
.....TG...T.....A............A....TAT.AT..CTACAG..GCAG......
.............................T...C.GTA.T....................
.............................T...G.GTA.T....................
.............................t...c.gta.t....................
.........................ct........gt....c.......g..a.......
.........................CT........GT....C.......G..AA......
.........................CT........GT....C.......G..AA......

TATATACTGGTATAAAAGTGGCGATGAAATAAATAGCCTGCTTCCGCTAGCGTACACCTG
.....GGC........G....G.......G.T............................
.............................G.............................C....
.............................G.............................C....
.............................g.............................c....
.............................G..............................
.............................G..............................
.............................G..............................

GATACCTGGTCCGGTAGCGACCTATGGTACCGGTGACGCACGTAGCATTAACGGATTTTT
.............................A..............................
.......................G......A..............................
.......................G......A..............................
.......................g......a..............................
.............................a..............................
.............................A...........G..................

GGGGGGTTTCGGGCTGCGCTGTTTCATTTTGCTCCTTCTCTCTGGAGACTCGGAGACTGG
....................................................................
....................................................................
....................................................................
....................................................................
.....................................c..............................
................G....................C..............................
................G....................C..............................
```

```
AGACCTCTCCGTCTACTTGGTACTAATTCTCTCTTTGGTCTGGTACGAGCGTTTTGCGTT
.............................................................
.............................................................
.............................................................
.............................................................
.............................................................
.............................................C...............
.............................................C...............

CCTCACGGCAGTTGCACCAGCTGCTAAAATTGCGGGTCTGTTCCGGGCGGGGGGTCTTGA
.............................................................
.............................................................
.............................................................
.............................................................
.............................................................
.T...G.....................................T.............
.T...G.....................................T.............

CTTCTGTGTGTGTGCCCTGACCGAAGATGGACGGCGGCGGTGTTTTTGTGTAATACATTT
..c..........................................c......c
.............................................T.......
.............................................T.......
.............................................t.......
.....................................................cc
.........................................G.......CC
.........................................G.......CC

TATAAACTCGTTTGGCTGATGGTGTACTTGCGCGCCGGCGATTTTTTTTGTTGGTACCTG
.....................................................g.....
.....................................................G.T.....
.....................................................
.....................................................
...c.................................................g.....
...C.................................................G.....
...C.................................................G.....

GTATCAAGAGTCCCCCCTTCAACTGCACACCCCCGGTTTTAAACACGAAGGGTAACTTGT
..............g..............................................
.....................................................C.....
.....................................................C.....
.....................................................c.....
..............g..............................................
...........T..G..............................................
...........T..G..............................................

ATTAACAGCAGTATCGCAACGCACAAAGTAATTTTGTGTCTCTTAACAGTTATCATGACA
.................G.AA....T.........C.....C.G....CA........G
...................A.....................................
...................A.....................................
...................a.....................g...............
...................a.....................................g.....
...................A.....................................G.....
...................A.....................................G.....
```

```
TGTTTTCAATGATAAGGTTCCCAAGTTAGCCTCCAGGTATATGGATGTTGGGATCTTTTG
A....C...G....T...AT.....................T......A..A.....
.............T.............................................
.............T.............................................
.............t.............................................
.............t.............................................a.
.............................................................A.
.............................................................A.

GCCCGCTTCCTTTGTGATAACGTTCGTTGGACGTACCGTATATTCACCTCGTCTTCAGTA
........A.........GC....T.................G..TG.T..........
.................C.........................................
.................C.........................................
.................c.........................................
.................g.............g...........................
.....................................G.......................
.....................................G.......................

CGTCTC  AM06SR
......  AM03KB
......  AM03RB
......  AM01WB
......  AT01OP
......  AM03SK
......  AM03JR
......  AN10BO
```

Figure 21. Representation of variable sites across actin intron II of *Alaria* for visual determination of possible recombination between individuals . Numbers denote sites indicated in Appendix IV. Black boxes indicate potential shared, derived character states while gray boxes indicate potential ancestral or unique character states.


Intron II - 5' to 3' with sites as noted
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
AM06SR
AM03KB
AM03RB
AM01WB
AT01OP
AM03SK
AM02JR
AN10BO
= shared derived character state or less common shared allele
= ancestral/unique character state or most common allele

110

An analysis of homoplasies using the criteria of Templeton et al. (1992) was performed. A Maximum Parsimony tree was generated using the alignment that included indels. The tree was used to identify potential homoplasies. In total, eight possible homoplasies were identified (Appendix IV). Under the criteria of Templeton et al. (1992), if two homoplasies sequentially show the same taxa grouped together or if any one homoplasy is an indel, then those included regions may be recombinantly related. No sequentially similar homoplasies were seen. However, a possible single indel homoplasy one base long was seen (site 35, Appendix IV).

Tajima-D statistics (*D*) were generated both with and without taxon KB. KB was excluded from some of the analyses because of the large divergence between KB and all other taxa. A Tajima-D statistic of -1.0580 was seen with taxa KB included in the analysis. When taxa KB was removed, the Tajima-D statistic was 0.5698. Both of these values indicate that the null hypothesis of neutral mutations cannot be rejected at the 90% or greater level (Tajima 1989, Simonsen et al. 1995).

Tree Generation

The alignment, with indels as well as potential non-neutral regions removed (Figure 20), was used to generate unrooted Maximum Parsimony, Neighbor Joining, and Maximum Likelihood trees (Figures 22-29). Minimum spanning trees however, also included indels. In addition, because of the large divergence between taxon KB (Kelsey Bay, B.C.) and the other taxa, replicate alignments were made with KB removed. Comparisons between trees generated, with and without KB included,

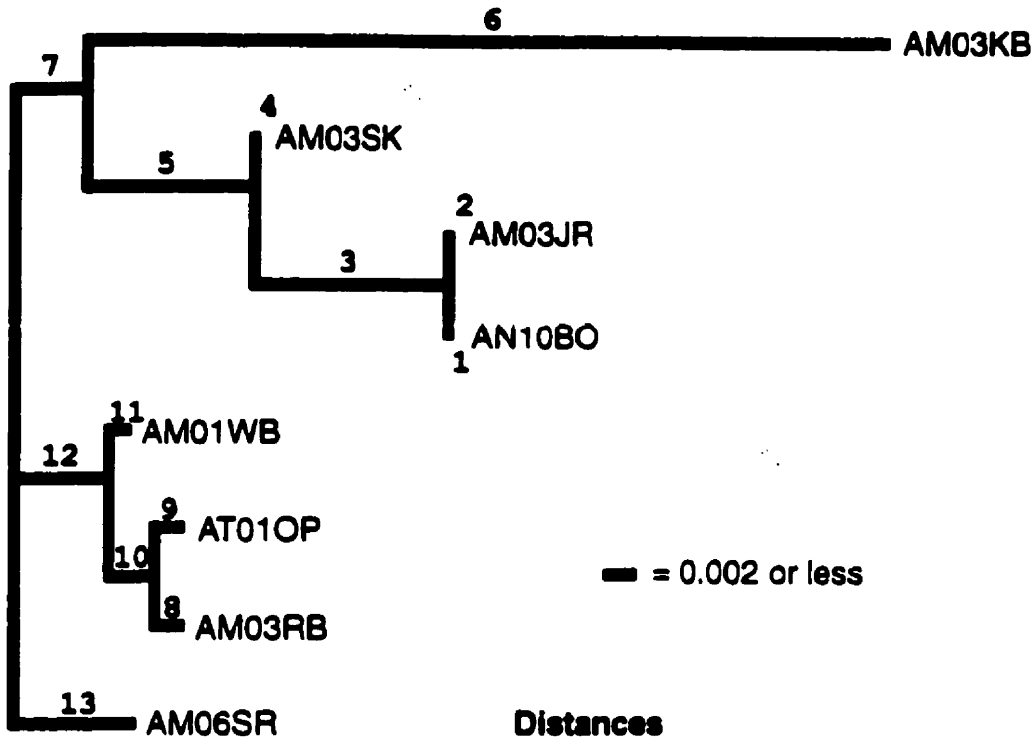
A)

	AM06SR	AM03RB	AM01WB	AT01OP	AM03SK	AM03JR	AN10BO
AM06SR		0.0230	0.0218	0.0218	0.0317	0.0455	0.0455
AM03RB			0.0048	0.0036	0.0342	0.0493	0.0493
AM01WB				0.0036	0.0329	0.0480	0.0480
AT01OP					0.0329	0.0480	0.0480
AM03SK						0.0181	0.0181
AM03JR							0.0000
AN10BO							

B)

	AM06SR	AM03KB	AM03RB	AM01WB	AT01OP	AM03SK	AM03JR	AN10BO
AM06SR		0.0859	0.0230	0.0218	0.0218	0.0317	0.0455	0.0455
AM03KB			0.0845	0.0832	0.0805	0.0805	0.0994	0.0994
AM03RB				0.0048	0.0036	0.0342	0.0493	0.0493
AM01WB					0.0036	0.0329	0.0480	0.0480
AT01OP						0.0329	0.0480	0.0480
AM03SK							0.0181	0.0181
AM03JR								0.0000
AN10BO								

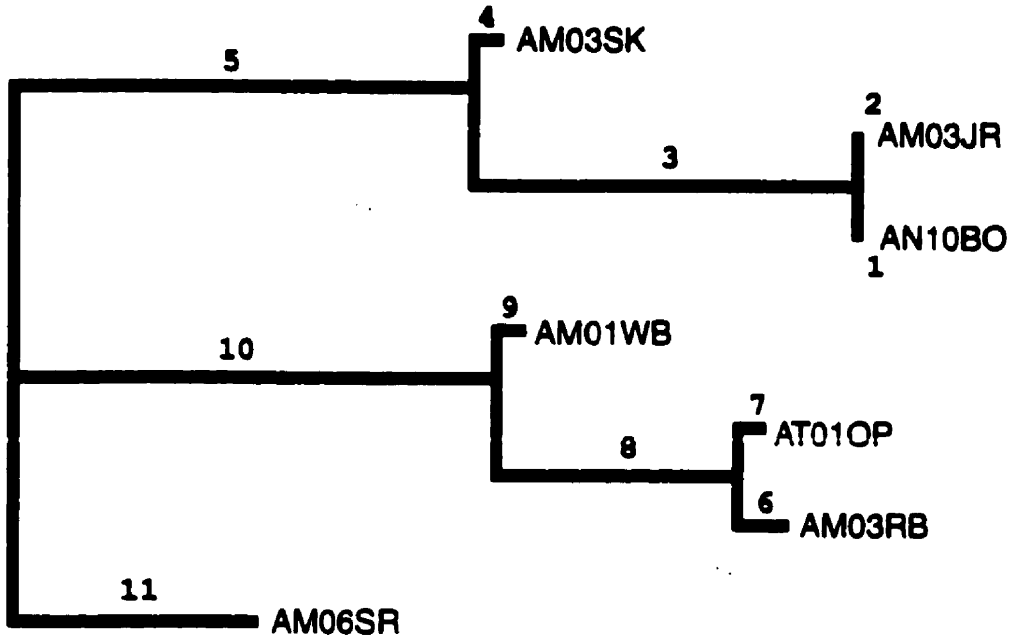
Figure 22. Jukes-Cantor corrected pairwise distance matrices generated using PHYLIP 3.573c DNADIST.EXE. Only upper-triangular portion of each matrix is shown with A) the alignment from Figure 20 minus AM03KB as input, or B) the entire alignment from Figure 20 as input .



Distances

1.	0.00000
2.	0.00000
3.	0.01685
4.	0.00125
5.	0.01432
6.	0.06658
7.	0.00595
8.	0.00258
9.	0.00102
10.	0.00044
11.	0.00196
12.	0.00873
13.	0.01132

Figure 23. Neigbor Joining phylogram from the Jukes-Cantor corrected distance matrix in Figure 22 using Neigbor.exe from PHYLIP 3.573c. Distances are the lengths of the segments noted on the tree. Note that the length of segment 10 is exagerated in the tree above for clarity of the branching structure, the other segments are accurate to approximately 0.001.



Distances

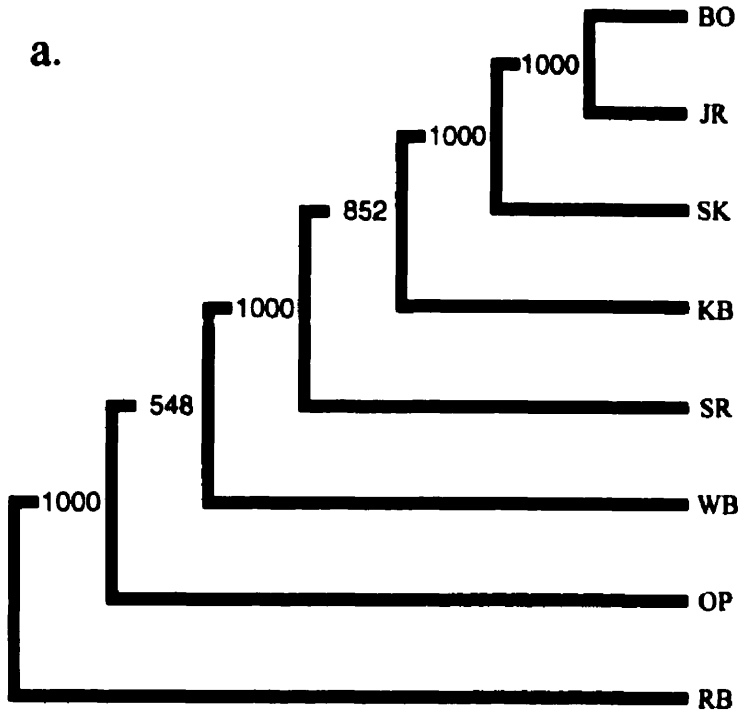
= 0.002

1.	0.00006
2.	0.00012
3.	0.01592
4.	0.00224
5.	0.01968
6.	0.00240
7.	0.00117
8.	0.00012
9.	0.00239
10.	0.01096
11.	0.00973

Figure 24. Neigbor Joining phylogram from the Jukes-Cantor corrected distance matrix in Figure 22 minus AM03KB using Neigbor.exe from PHYLIP 3.573c. Distances are the lengths of the segments noted on the tree. Segment length accuracy is approximately 0.001.

Figure 25. Neighbor joining trees using a Jukes-Cantor corrected distance matrix generated using PHLIP 3.57c. Sequences were from the Figure 20 alignment in which indels were removed. Tree "a." has all eight individuals while tree "b." does not include AM03KB. Numbers to the right of the node represent the bootstrap values for the node from 1000 replicates. SK = AM06SK, BO = AN10BO, JR = AM02JR, SR = AM06SR, RB = AM03RB, KB = AM03KB, OP = AT01OP, and WB = AM01WB.

a.


BO
1000
JR
1000
852
SK
1000
KB
548
SR
1000
WB
OP
RB

b.

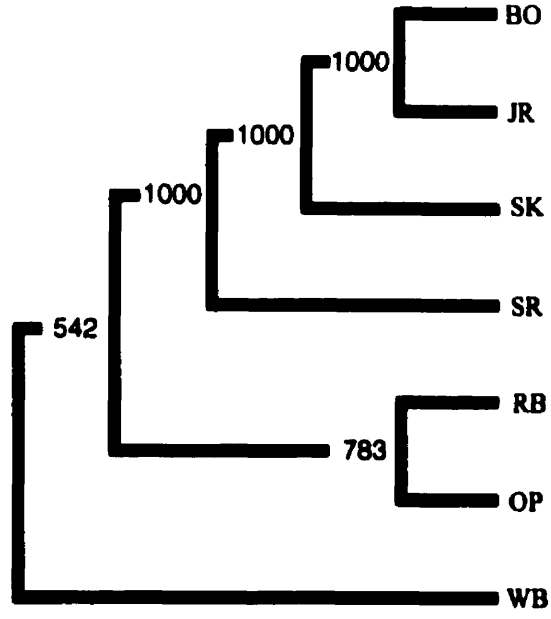
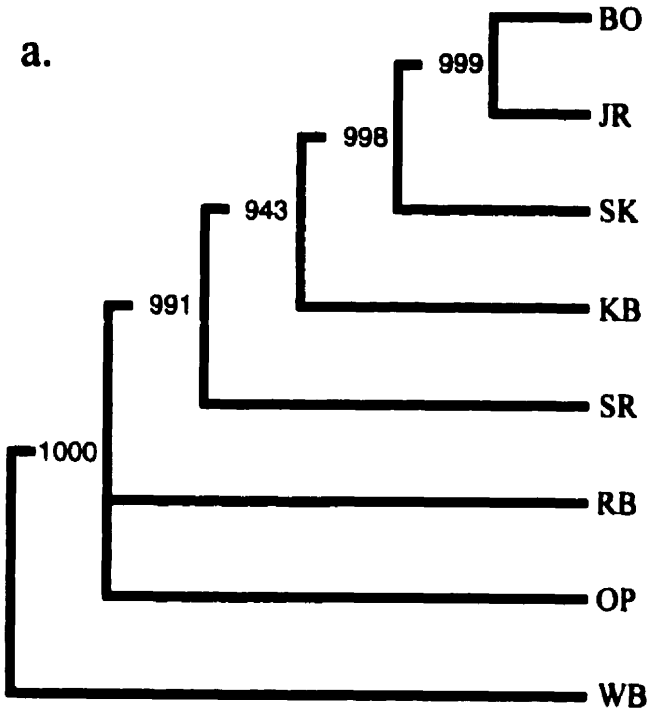
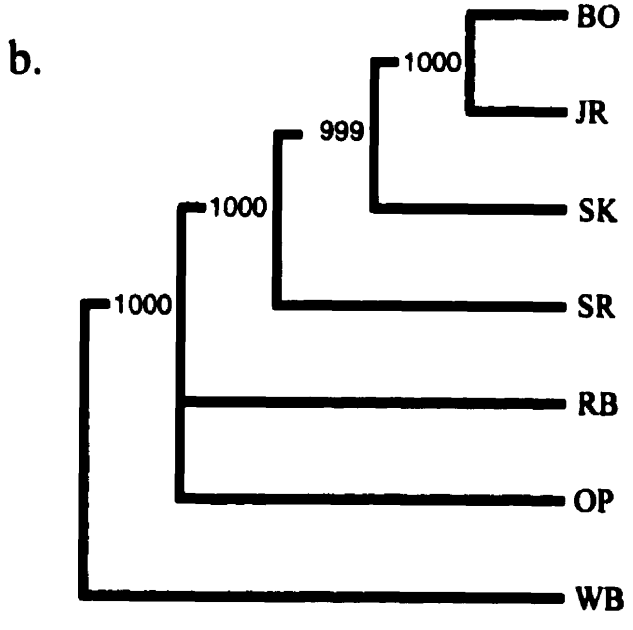
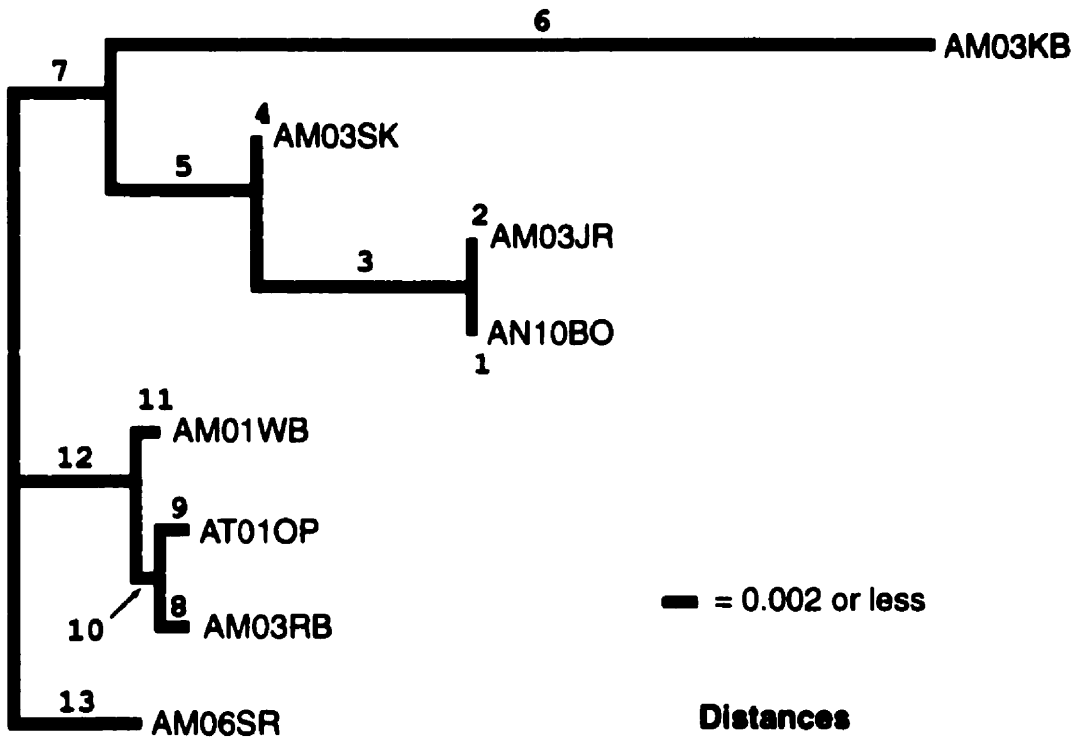

BO
1000
JR
1000
1000
SK
SR
542
RB
783
OP
WB

Figure 26. Maximum Parsimony trees using using PHLIP 3.57c. Sequences were from the Figure 20 alignment in which indels were removed. Tree "a." has all eight individuals while tree "b." does not include AM03KB. Numbers to the right of the node represent the bootstrap values for the node from 1000 replicates. SK = AM06SK, BO = AN10BO, JR = AM02JR, SR = AM06SR, RB = AM03RB, KB = AM03KB, OP = AT01OP, and WB = AM01WB.


a.
BO
999
JR
998
943
SK
991
KB
SR
1000
RB
OP
WB

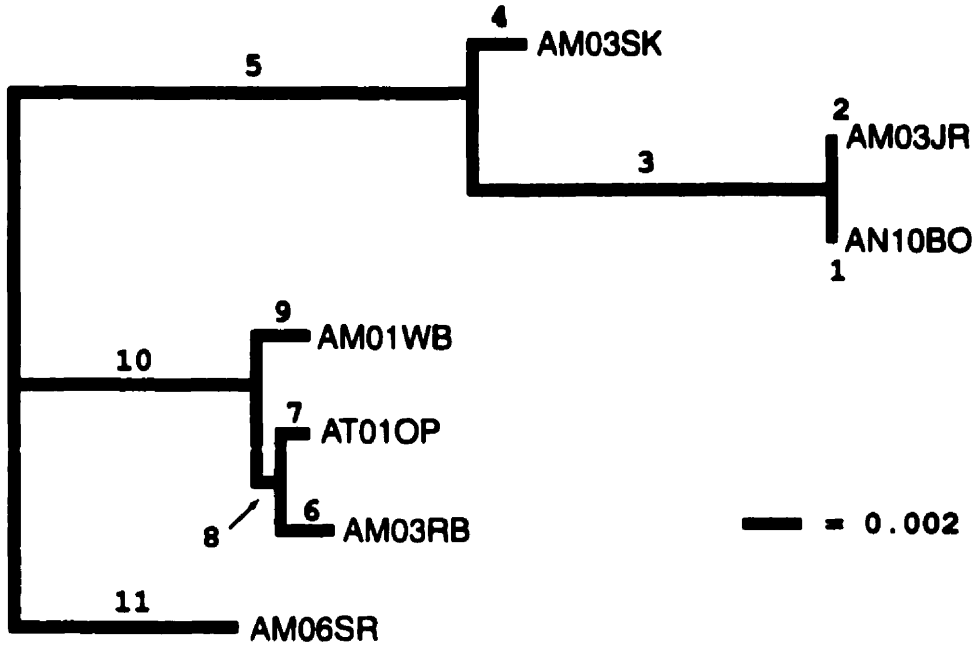

b.
BO
1000
JR
999
1000
SK
1000
SR
RB
OP
WB



Distances

1.	0.00012
2.	0.00012
3.	0.01822
4.	0.00006
5.	0.01270
6.	0.06827
7.	0.00893
8.	0.00893
9.	0.00121
10.	0.00006
11.	0.00239
12.	0.01066
13.	0.01051

Figure 27. Maximun Liklihood phylogram generated from the alignment in Figure 20 using dnaml.exe from PHYLIP 3.573c with an equal probability model. Distances are the lengths of the segments noted on the tree, except segment 10 which is exagerated for clarity. Segment length accuracy is otherwise approximately 0.001.



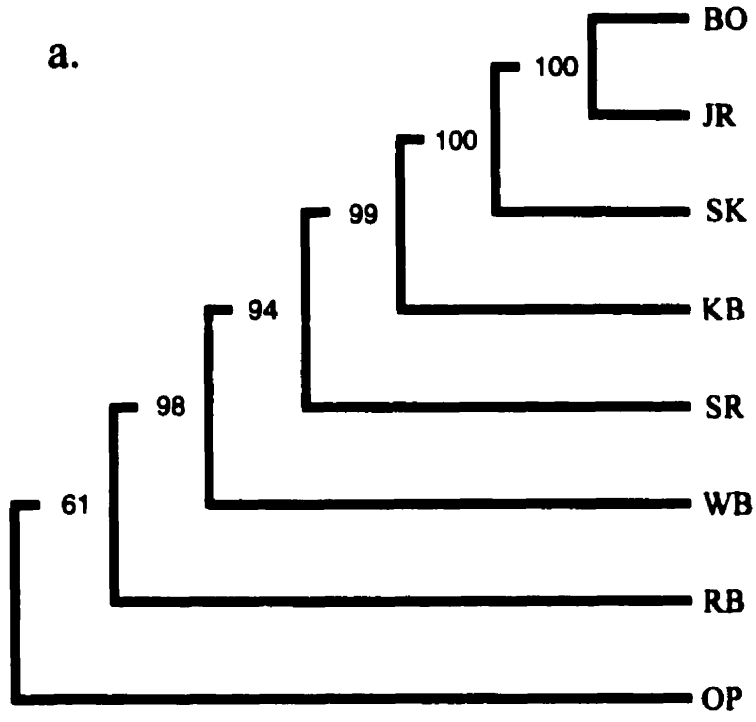
Distances

1.	0.00006
2.	0.00012
3.	0.01592
4.	0.00224
5.	0.01968
6.	0.00240
7.	0.00117
8.	0.00012
9.	0.00239
10.	0.01096
11.	0.00973

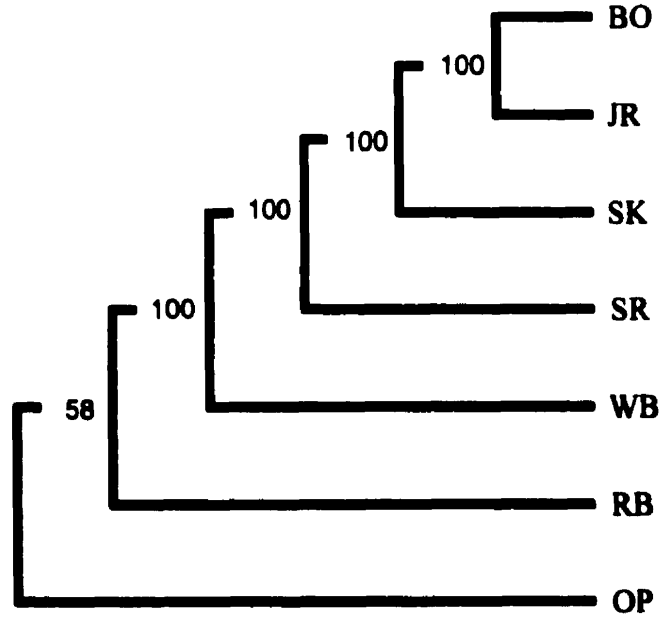
Figure 28. Maximun Liklihood phylogram generated from the alignment in Figure 20 minus AM03KB using dnaml.exe from PHYLIP 3.573c with an equal probability model. Distances are the lengths of the segments noted on the tree, except segment 8 which is exagerated for clarity. Segment length accuracy is otherwise approximately 0.001.

Figure 29. Maximum Likelihood trees using an equal probability evolutionary model generated using PHLIP 3.57c. Sequences were from the Figure 20 alignment in which indels were removed. Tree "a." has all eight individuals while tree "b." does not include AM03KB. Numbers to the right of the node represent the bootstrap values for the node from 100 replicates. SK = AM06SK, BO = AN10BO, JR = AM02JR, SR = AM06SR, RB = AM03RB, KB = AM03KB, OP = AT01OP, and WB = AM01WB.

a.


BO
JR
SK
KB
SR
WB
RB
OP
100
100
99
94
98
61

b.


BO
JR
SK
SR
WB
RB
OP
100
100
100
100
58

could indicate whether KB's large divergence of more than 69 point mutations and 19 indels was skewing the results.

Neighbor Joining Trees

Neighbor Joining trees were generated with 1000 bootstrap replicates from alignments both with and without taxon KB. Only one node was not well resolved and another only occurred 85% of 1000 bootstrap replicates in the Neighbor Joining tree that included KB (Figure 25). The weakly resolved node (548 of 1000 replicates) was between OP and WB. The next weakest bootstrap score of 852 of 1000 replicates occurred on the node separating KB and SR. The Neighbor Joining tree without KB, had two very weak (542 and 783 of 1000 replicates) branches, with the rest of the branches supported in 100% of the replicates (Figure 22). Collapsing both weakly supported branches gives a trichotomy of individuals RB, OP, and WB.

Phylograms from a Jukes-Cantor corrected pairwise distance matrix (Figure 22) were also generated using Neighbor Joining (Figures 23 and 24).

Maximum Parsimony Trees

Maximum Parsimony trees were generated with 1000 bootstrap replicates from alignments both with and without taxa KB. In the tree that included taxa KB, only one node (between KB and SR) was not supported at least 99% of the time (Figure 26a). The node between KB and SR was still quite strong however, occurring 94% of the time. The Maximum Parsimony tree generated from an alignment without

KB had one undetermined node (neither possible node seen more frequently than in 50% of the trees) between RB and OP (Figure 26). The other nodes all occurred either 999 or 1000 times in 1000 replicates.

Maximum Likelihood

Phylograms utilizing the resulting segment lengths from a Maximum Likelihood tree were generated using the alignment in Figure 20 both with and without AM03KB (Figures 27 and 28).

One hundred bootstrap replicates of both alignments with and without KB were analyzed using Maximum Likelihood. The Maximum likelihood tree generated with KB had two nodes not supported in at least 98% of the resulting trees (Figure 29). The node between SR and WB was only supported in 94% of the trees, and the node between OP and RB was only supported 61 times out of 100 replicates. The Maximum Likelihood tree generated without KB had nodes occurring 100% of the time for all nodes except the node between RB and OP, which was seen 58% of the time (Figure 29b).

Minimum Spanning Trees

A distance matrix was used as an initial guide to identify highly similar taxa for minimum spanning trees. Manual pairwise comparisons were done to determine the number of changes necessary to convert one taxon into any of the others. The resulting data were used to generate a minimum spanning tree (Figure 30). If the

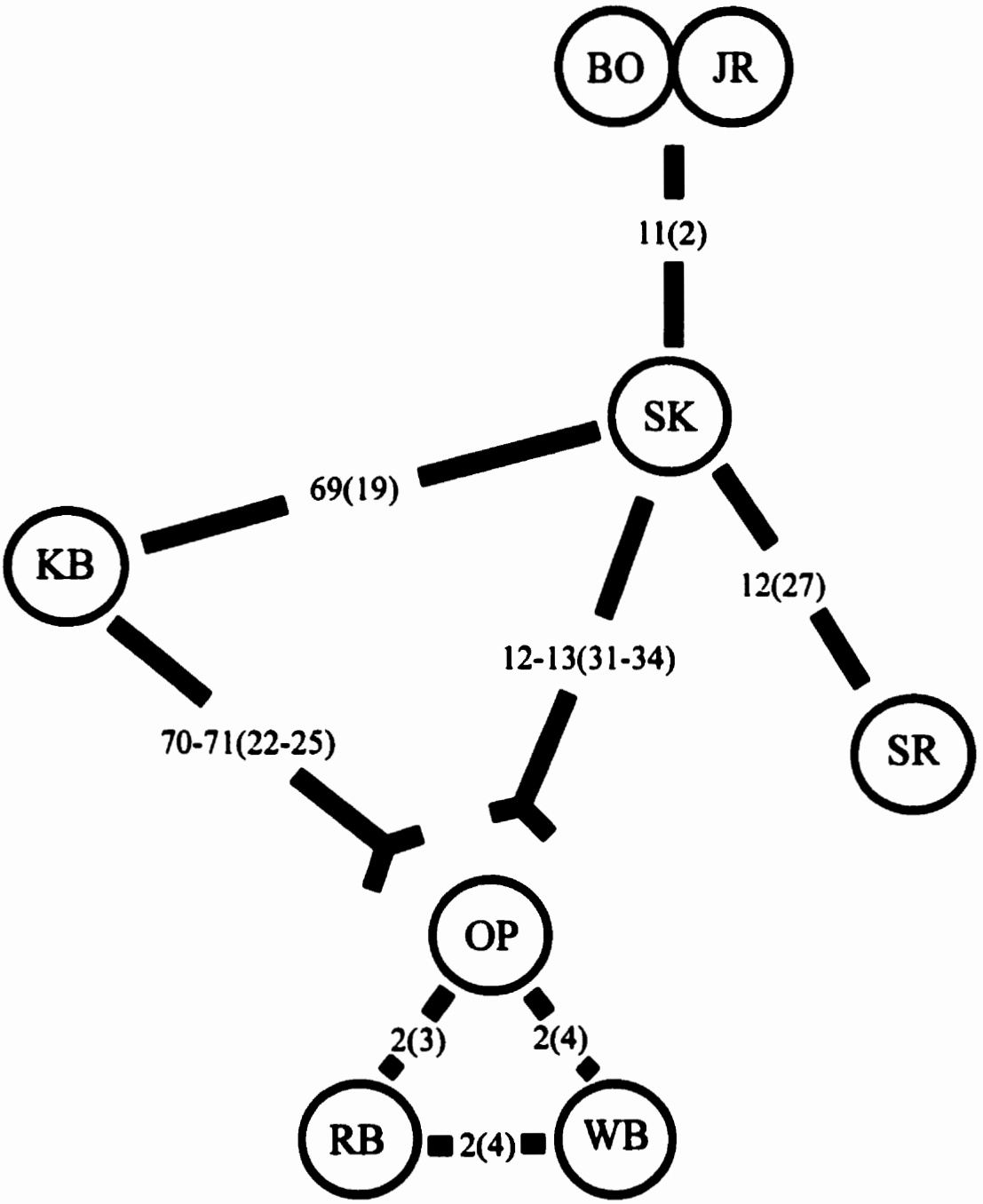


Figure 30. Minimum Spanning Tree from alignment in Appendix IV. Numbers indicate the smallest estimated number of changes necessary to convert one taxa into the other taxa. Numbers in parenthesis denote indels. Thick lines indicate shortest tree +/- five point mutations. Lines which branch at the end include all of the three closely related taxa: OP, RB, and WB. Taxa designations are as denoted in Figure 22.

most likely branch differed from any of the other possible branches by only five or less steps, both were shown. Five was chosen as the cutoff point because five is one step more than the maximum point mutation or indel differences between the three nearly identical individuals OP, RB, and WB. The next largest difference was thirteen steps. No other topologies were within ten steps of the one shown.

The minimum spanning tree resulted in two closely related clusters, two other nearby individuals, and one distant individual. The BO and JR taxa were identical in this analysis, and the three taxa OP, RB and WB differed from each other by only two point mutations or three to four indels. Looking at point mutations alone, the two clusters were nearly equidistant from SK, which was also equidistant from SR, forming a star-like topology. Finally, the KB individual was more than five times as far away from the nearest individual (69 point mutations or 19 indels).

Discussion

Alignments

Although data were generated for the individual plants encompassing an average of 1286 bp, some of the sequence, particularly those regions that are believed to be expressed, were not used. *Nereocystis luetkeana*, a kelp from a different family (Lessoniaceae), was used in the alignment to determine common non-neutral regions shared across the Laminariales. *Nereocystis* second intron sequences were so

divergent as to be un-alignable from a point 5 bp downstream of the 5' splice junction to the lariat structure near the 3' splice junction. Sequences denoting coding regions, lariat-structures of the splicing mechanism, and one additional short region were removed from the analysis. The deleted region encompassed five bases adjacent to the splice junction that were identical in all of the individual *Alaria* plants examined as well as in the *Nereocystis luetkeana* individuals. In addition, most of the analyses were done on sequences in which homologous indel regions were deleted.

The coding regions were removed from the analysis because of the possibility of selection on actin proteins. Variation in the coding region was quite low, only two point mutants were seen, as described in Chapter III. Sequences corresponding to the lariat-structure, which is required for proper splicing, were also removed (Brody and Abelson 1985). In addition, since the location of the lariat structure varies somewhat, and signals for it's positioning are not well understood, areas from the lariat structure to the 3' end of the intron were removed from the analysis. Indels (inserts or deletions) were removed because of theoretical difficulties in reconciling point mutation and indel mutation rates or processes (Swofford and Olsen 1990). Indel regions, however, do provide strong markers for determining recombination due to their uniqueness. Indels were therefore not removed from analysis of potential recombination.

Recombination

Recent recombination has a strong likelihood of giving false relationships if present (Aquadro et al. 1986). Sequences may be related by recombined regions in diploid (or polyploid) species from the same panmictic population. Recombination with non-functional pseudogenes or multigene families is also possible (Fitch et al. 1990). Such recombined regions could both mask true genetic relationships and erroneously imply others. In addition, most methods of reconstructing phylogenies rely upon the sequences having a single evolutionary history for their entire length (Jacobsen et al. 1997). Areas with apparent recombination since the divergence of the taxa are therefore suspect and should be removed.

A powerful method for determining possible recombination makes use of homoplasies. If two or more individuals share identical states, not because the state is a shared ancestral state, but instead due to other unknown processes, they are deemed homoplasies. For this analysis, a conservative assumption was made that the process creating misleading relationships (homoplasies) was recombination (Aquadro et al. 1986). Homoplasies were identified as, and can be functionally defined as, sites or regions that are identical between two or more individuals not grouped closely together in the phylogeny generated by all of the other sites. Homoplasies can be caused by parallelism, convergence, or reversals in the sequence examined (Hillis and Moritz 1990).

In determining recombination in homologous DNA sequence regions, opinions differ as to how much weight to place on each homoplasy. Aquadro et al. (1986) believe recombination should only be inferred if the recombination resolves

two or more homoplasies, or if one of the homoplasies can be regarded as completely parsimonious. Others have commented that such a conservative designation of recombination is too harsh, and may overestimate or falsely assign regions to recombination (Templeton et al 1992).

DNA sequence homoplasies are quite common if sequences are very divergent (Templeton 1983). Homoplasic point mutations have a relatively large chance of arising randomly given the smaller number of possible states for each site. Point mutations are therefore more difficult to assign as truly homoplasic. Because of the decreased likelihood of chance arisal as length increases, indels fit the criteria of being regions that are nearly completely parsimonious and therefore good markers for possible recombination.

In the data set encompassing the second actin intron, little evidence for recombination based on homoplasies was seen. Eight possibly homoplasic point mutations were identified. Only one indel was also identified as a possible homoplasy, but the indel is only one base long. Therefore, even under the stringent criteria set forth by Aquadro (1986), the evidence does not imply any recent recombination in the second intron amongst the individuals examined.

Recombination could also be detected by visual inspection of sequential phylogenetically informative sites in a way analogous to Jacobsen et al. (1997). This method is, however, far from an ideal way to show the lack of recombination because two outcomes of the analysis are possible. The lack of any visible switch in similarity of the sequence in question from one individual to another, could be due either to an

incomplete sample of all genotypes or an actual lack of recombination. Instead using phylogenetically informative sites could be fruitful if and only if sequence stretches showed strong evidence for recombination. In essence, the method cannot prove that recombination has occurred, and instead only hopes to detect some proportion of the recombination that has occurred. As such, a sequential inspection of all of the informative sites in this study showed no evidence for recombination.

Given a data set encompassing lengthy sequences of data from the same small number of individuals, limited statistical tests of neutrality (or recent recombination) are possible. Computer simulations have as well as qualitative studies, have shown that Tajima's D-statistic to be one of the most effective tools in discovering possible selection events (Simonsen et al. 1995). By considering both the number of polymorphic sites as well as the average number of pairwise differences between individuals, a D-statistic can be generated and compared to the confidence limits for the rejection of the null hypothesis of only neutral variation (Tajima 1989). The confidence limits of the statistic are based on an infinite sites model (non-overlapping point-mutation events). Further work using numerical simulations has shown the D-statistic to be the best alternative method utilizing presently attainable amounts of molecular polymorphism data for detecting selective sweeps, population bottlenecks, and population subdivision (Simonsen et al. 1995). The D-statistic, while better than other methods, can still only detect relatively recent bottlenecks or selective sweeps and long-standing population subdivision. Most estimates suffer from these problems because of the limitations set by the infinite alleles model and the difficulties inherent

in determining such parameters as the effective population size (N_e) and the mutation rate.

Tajima -D statistics generated from the data in Figure 20 (indels removed) were unable to reject the null hypothesis of neutral mutations. A D-statistic was also generated that did not include taxon KB because of KB's large contribution to the average number of pairwise differences. Even without KB, the null hypothesis could not be rejected. The D-statistic results give no evidence for recent recombination, population bottlenecks, selective sweeps or long standing population subdivision, and suggest that the sequences examined contain only neutral changes.

The dataset examined in this study showed little evidence for recombination based on homoplasies or informative sites. In addition, the null hypothesis of neutral mutation could not be rejected using the most powerful known statistical test for the lack of neutrality, the Tajima-D statistic. Therefore, the data show no evidence for recent recombination among the actin introns examined in this study.

Phylogenetic Hypotheses and Trees

Phylogenetic hypotheses and their resulting trees are powerful tools in reconstructing evolutionary relationships among taxa (Hillis and Moritz 1990). Their usefulness in population-based analyses is tempered by the assumptions inherent in phylogenetic inferences. One frequently overlooked basic assumption is that the individuals represent independent terminal branches (Davis and Nixon 1992). Genetic mixing of contemporaneous branch terminii creates possible non-tree-like

relationships and therefore violates one of the most basic assumptions of phylogenetics. If individual alleles are examined, one obvious potential manifestation of mixing between taxa is recombination between the alleles.

The algorithms used to reconstruct phylogenies also can affect the resulting tree topology, and each has its own assumptions and caveats as discussed below. In an attempt to mitigate the weaknesses of each method, three of the major methods (Parsimony, Neighbot-Joining, and Maximum Likelihood) were used in this study.

Parsimony, although powerful at finding most-likely topologies, can suffer from "Type II" errors. These errors are best understood using an analogy of Sewell Wright's two-dimensional surface with a number of peaks whose height is proportional to the "closeness" of the topology to the "real" tree (Swofford and Olsen 1990). Each peak represents a nearby maximum, and the parsimony algorithms try to "climb" to a peak. Difficulties arise in trying to determine whether the peak in question is the highest one in the two-dimensional space since the algorithms cannot "see" nearby peaks. Bootstrapping using multiple data sets generated by resampling the original data can help alleviate these problems (Felsenstein 1988).

Neighbor-joining based-trees bring different assumptions and problems. Whereas parsimony based methods utilize the state of each character when comparing the two, Neighbor-Joining utilizes the relative differences, codified in a pairwise distance matrix to group those taxa with the least divergence. Because the data are transformed, certain information is lost (Penny 1982). Often, bootstrapped Neighbor-Joining trees are less-able than parsimony-based methods to differentiate branches

(Hillis and Dixon 1990). Distance based methods, however, do not suffer from an *a priori* reliance on the least number of steps (minimum evolution) to create the tree, which basically defines parsimony (Hillis and Dixon 1990).

Maximum Likelihood trees rely upon sampling the data set, generating a tree and then attempting to improve the tree based on how well the tree fits the given evolutionary model. In a way, Maximum Likelihood is like Parsimony. The difference lies in the fact that Maximum Likelihood looks at the product of the log likelihood of all the ancestral states based on the evolutionary model to evaluate the tree, while Parsimony examines the sum of the changes to the ancestral state to arrive at the tree (Hillis and Dixon 1990). Because of the huge computational load, Maximum Likelihood based methods are often confined to smaller data sets.

An examination of the results generated, using all three phylogenetic methods employed, gave a consistent non-rooted tree (Figs. 23, 24, 26, 27, and 28). This tree had five main branches with the following relationships. One branch contains the Jordan River (JR) and Botany Beach (BO) individuals, with the Sitka (SK) individual's branch grouping the next closest. The Kelsey Bay (KB) individual represents the next branch, and then the Seal Rock (SR) individual. Finally a branch encompassing the Rosario Beach (RB), West Beach (WB), and Orange Point (OP) individuals was seen. These branches map nicely to the local geography, with branches from: 1) Puget Sound / Georgia Straight (RB, WB, OP), 2) Oregon (SR), 3) Johnstone Strait / Inside Passage (KB), 4) outer coastal southeast Alaska (SK), and finally 5) southwest Vancouver Island (JR, BO).

The trees generated using all three methodologies were essentially identical. Although not a true statistical test, the agreement among the three methods strengthens the support for the generated tree. The individual relationships also seem to be correlated over distances of 100km or less, with geographically-close individuals in the same clade. Relationships over larger distances were less clear since only two distant individuals were examined.

Minimum Spanning Trees

To alleviate some of the difficulties inherent in an assumption of a single tree-based phylogeny and to examine phylogeography using the small number of individuals analyzed, an entirely different net-based topology was hypothesized. Such network or star phylogenies allow for the taxa to have more than one relationship to each other, as can occur when the individuals in question do not have definite barriers to genetic exchange. Such relationships mimic the well-studied mathematical phenomenon known as Minimum Spanning Trees. If the number of units compared (in this case *Alaria* individuals) is large, calculating all possible relationships is daunting. In the case of this data set, however, a number of the individuals are either identical or nearly so (zero to two total point mutation differences), thereby reducing the number of possibilities and allowing for a manual generation of the minimum spanning tree.

The Minimum Spanning Tree generated from the data set has five major groups (Figure 30). The Kelsey Bay (KB) individual is very distant from all of the

other taxa, with either the Puget Sound / Georgia Strait (OP, WB, RB) or Sitka Alaska (SK) groups the closest, and only differing by less than 3% of the total point mutations (from 69 to 71 point mutations). All of the other three groups are almost equally distant to Sitka (SK).

The Minimum Spanning Trees seem to support a conclusion of divergence based on geography. Individuals from the same body of water, such as Puget Sound / Georgia Strait (OP, RB, and WB) or West Coast of Vancouver Island (BO and JR), are closer to each other than they are to individuals from the other water bodies.

Relationships between Individuals

Because the number individuals examined was less than the idealized situation of perhaps 20 individuals from each population and at each population sub-level, direct measures of gene flow utilizing F statistics were not possible. Instead, drawing upon the recombination results, the phylogenetic trees, as well as the Minimum Spanning Tree, a number of conclusions can be reached that may explain the data set generated. Each of the areas investigated carry their own assumptions and caveats. The question is whether the three areas can be combined to make a summary conclusion by, in essence, mitigating some of the inherent flaws of each.

By combining the three phylogenetic trees, an overall picture of the tree-based relationships can be drawn. These phylogenetic tree-based results however hinge upon a lack of recent gene flow between termini. Because kelp posses a diploid genome in the sporophyte stage, such gene flow between termini seems most likely to

manifest itself either in heterozygous alleles in an individual, or in recombination-based changes after syngamy. No such heterozygotes (with the possible exception of AM01WB) or recombinations were seen.

One cannot, however, discount the possibility of missing alleles, since the data were generated from diploid individuals. All efforts were made (such as cloning and direct sequencing of genomic DNA) to uncover any heterozygous individuals in this allele. Though never detected in this study, the possibility that alleles were missed means that some relationships may remain undetected.

In addition, except for possibly the Minimum Spanning Tree, the trees generated all represent likely best trees instead of actual best trees. Although the methods used to generate the trees are well documented and commonly utilized they can, under certain circumstances, miss the "best" tree. Boostrapping (in this case sometimes with as many as 1000 replicates) can lessen the likelihood of missing the "best" tree. In the end all phylogenetic trees are inferences, albeit strong ones if done correctly.

Given all of the above caveats, an overall picture of the relationships between the individuals studied can be drawn. These relationships can then be mapped on both the traditional species boundaries as well as the geographic locales of the individuals.

Although the taxa do group in a manner consistent with geography, a mere distance based grouping does not hold. Instead, in all cases except one the taxa seem

to group closely if separated by perhaps less than 100 km. At distances greater than 100 km, the boundaries appear to be hydrodynamic or perhaps historical in origin.

Temperature/salinity curves may explain some of the geographic boundaries in the data (Druehl 1981). In general, the temperature and salinity regimes in the Vancouver Island area vary seasonally. The outer coast has relatively lower salinity due to rainfall in the winter when temperature is low, and both higher temperature and salinity in the summer. Inner coastal regions of Vancouver Island are affected by the Fraser River runoff and so have higher salinity and lower temperature in the winter with both lower salinity and higher temperature during the summer runoff (Thomson 1981). Gross-level tidal currents in the protected areas descibed here flow from the mouth of the Strait of Juan de Fuca to the northern-most end of Georgia Strait (Thompson 1981). The temperature/salinity effect may help explain why the southwest Vancouver Island group (BO and JR) and Puget Sound / Georgia Strait group (OP, WB, RB), which are as close as 100 km apart and share the same tidal current regime, are so far apart genetically. In addition, less obviously, the large difference between the Orange Point (OP) and Kelsey Bay (KB) individuals, that are less than 100 km apart, may reflect a distinct temperature/salinity barrier. Orange Point's closeness to the mouth of the Fraser in comparison to Kelsey Bay's location in Johnstone Strait, may result in vastly different temperature/salinity regimes. The Fraser River runoff has a strong surface water influence in separating the central and northern ends (Orange Point and then Kelsey Bay in Johnstone Strait) from the southern regions of Georgia Strait and the Juan de Fuca Strait (Thomson 1981).

Perhaps the differences between the Orange Point and Kelsey Bay individuals reflect a genetic barrier based on preferred conditions for reproduction. Druehl (1981) explained the lack of *Macrocystis* individuals at Entrance Is. near the Fraser River, and presence in Nootka Sound on the outer coast of Vancouver Island, as a temperature/salinity effect. Other relationships when viewed from a temperature/salinity or oceanographic standpoint are less clear.

The eight individuals included in the study actually include three species: one from *A. nana*, one from *A. tenuifolia*, and six from *A. marginata*. If the inferred phylogeny were to represent the species boundaries, then each species should be monophyletic. Instead, the two non-*A. marginata* individuals (OP and BO), group more closely to certain *A. marginata* individuals (JR, or RB and WB) than the other *A. marginata* individuals (SK, KB, SR) do to the same *A. marginata* individuals (JR, or RB and WB). This finding supports earlier assertions, based on rDNA RFLPs and morphology, that the species boundaries are suspect (Mróz, 1989).

In addition to the oceanographic hypothesis above, the barrier between Orange Point and Kelsey Bay reflects a traditional taxonomic species boundary as well (Widdowson, 1971). The Kelsey Bay individual would be classified based on the commonly accepted taxonomy (Gabrielson et al. 1989) as an *Alaria marginata*, while the Orange Point individual would be an *Alaria tenuifolia*. The one *A. nana* individual shows less agreement with the accepted taxonomy since the nearest *A. marginata* individual was exactly genetically identical in this region. This identity, along with the polyphyletic nature of the *A. marginata* individuals as a whole in this

study, suggest that the three species of *Alaria* examined here need revision. As mentioned earlier, rDNA RFLPs also question the accepted taxonomic groupings of the three *Alaria* species (Mróz 1989). The RFLP study found "hybrid zones" with contrasting taxonomic and genetic affinities in these same geographic regions.

Distinctions between the Kelsey Bay individual and all other *Alaria* individuals in this study may be somewhat suspect. The Kelsey Bay individual is very divergent, in both point mutations as well as indels. The Kelsey Bay *Alaria* individual however is not as distant from the other *Alaria* as from *Nereocystis,* which is nearly impossible to align to any *Alaria* except in the lariat-structure and coding regions. Since most of the other *Alaria* individuals differ amongst themselves by less than a third of the differences between Kelsey Bay and it's closest individual, any relationships to the Kelsey Bay individual are suspect. Perhaps further investigations into the individuals in and around Kelsey Bay or other individuals from Johnstone Strait might help clarify this situation.

This study has demonstrated the usefulness of actin introns for population-level analyses in kelp. Actin introns satisfy many of the requirements for neutrality and do not appear to have any recent recombination in three local *Alaria* spp. The moderate number of differences between individuals of less than 5% should prove fruitful in further large-scale examinations of gene flow. Introns of other nuclear encoded genes may prove to have similar abilities to discern population substructure and could independently corroborate these results.

General Conclusion

The results presented here can be seen as a starting point for future research into population and intraspecific genetic studies in the Laminariales. Techniques were developed that shed light on areas with no previous data points. Of particular importance is my characterization of actin introns in a heterokont. Availability of such a powerful marker for population studies in the Heterokontophyta, and the Laminariales in particular, will allow population geneticists to employ neutral-based measures such as F statistics for gene flow analyses and population structure (Hudson et al. 1992, Lessa 1992). Although taxa-wide conclusions regarding gene flow are not possible with the number of individuals examined herein, these studies have shown where to concentrate further research and also refined techniques. In addition, the hybridization studies, involving both wild-collected and laboratory-generated hybrids, were able to determine parentage of small individual plants.

Hybrids of the Laminariales

The molecular rDNA sequence data generated from the *Macrocystis* x *Pelagophycus* hybrid was most unexpected. The gametophytes from the wild-collected hybrid were identical to *Macrocystis integrifolia*, and not *Macrocystis pyrifera* or *Pelagophycus* as expected. Although the relationships between the extant *Macrocystis* species, and perhaps even their distinction, has recently been questioned

(Mackenzie 1997), the lack of identity with the *Pelagophycus* parent was unexpected. A number of proposed causes for the observed identities were suggested including meiotic non-disjunction, loss through normal mendelian inheritance, or that the gametophytes examined were not true hybrids.

Nuclear DNA sequencing of the multicopy ribosomal cistron proved to be an excellent method for examining the parentage of wild-collected tissues. The results presented herein suggest that a simple biparental inheritence of the nuclear genome may not have occurred.

PCR-based typing of the results of laboratory generated crosses also proved to be an effective technique. Parentage was as expected in the controls and self crosses performed in the two species utilized, *Lessoniopsis littoralis* and *Alaria marginata*. A true hybrid individual was also found that contained both parental genomes. These results offered the first proven existence of intergeneric and interfamilial hybrids in the Laminariales.

Hybridizations and implications for families of the Laminariales

Hybrids between members of separate families of the Laminariales would seem at first to be surprising. However, if the "Groups" of genera proposed in Druehl et al. (1997) are considered, a hybrid between *Alaria* and *Lessoniopsis* would be less unexpected. Both *Lessoniopsis* and *Alaria* are members of Group1 within the Laminariales. Both have less than 2 % divergence in the 3' 18S - ITS1- 5.8S region,

as compared to more than 5% divergence between different Groups (Saunders and Druehl 1993).

Population-level differences in Alaria spp.

Determinations of gene flow in kelp could help to elucidate a number of long-standing debates and paradoxes in the study of the Laminariales. Kelp have only been shown to disperse about 5m (Druehl 1981), yet they manage to colonize rocky headlands separated by many inhospitable kilometers of sandy coastline. In addition, kelp are regularly seen colonizing offshore oil platforms as well as buoys (Davis et. al 1982).

Two general theories have been proposed to explain long-distance dispersal: occasional inoculation by drifting sporophytes (Anderson and North 1966), and episodic storm-induced spore clouds (Reed and Ebeling 1991). Although my study did not examine enough individuals to measure gene flow and thereby favor one dispersal hypothesis over the other, the techniques executed were able to distinguish individuals from different locales. In addition, the relatedness between individual *Alaria* thalli examined was approximately proportional to distance, especially when local hydrodynamics were included in the analyses.

Future Directions

Clearly, further work into the identification of individual kelp chromosomes and their segregation in *Macrocystis* x *Pelagophycus* hybrids is necessary to fully

understand the identity and nature of *Macrocystis* x *Pelagophycus* hybrid sporophytes and gametophytes. Through flow cytometry, *in situ* hybridization, etc. the number of chromosomes and the locations of the nrDNA could be determined, thereby allowing for a better understanding of the cytological events in both hybridization among the Laminariales and the kelp life cycle in general.

Rare or single-instance samples inherently suffer from questions of possible handling or culture errors. Analyses of differently obtained individuals of apparent hybrid origin between *Macrocystis* and *Pelagophycus* would be necessary to rule out errors such as mislabeled or cross-contaminated cultures.

PCR-based typing of laboratory crosses between kelp genera offers a potentially fruitful way to examine hybridizations in light of recent proposed revisions within the taxonomy of the Laminariales (Druehl et. al 1997). Interesting crosses could be performed between Groups as defined by Druehl et. al (1997), as well as within Groups. Since hybridization is a classical delimiter of species boundaries (Mallet 1995), a survey of hybridizational potential between kelp genera may be helpful in understanding the generic relationships.

The use of nuclear-encoded actin introns as examples of neutral regions appears to be a fruitful approach for future studies of population subdivision, and intraspecific genetic exchange. The inclusion of other introns or non-coding regions would strengthen the resulting hypotheses. Techniques involving anonymous or unknown genic regions, such as AFLP, could be utilized. AFLP data have been used

to differentiate kelp from adjacent nearby kelp stands of different wave exposure (Kusumo 1998).

The methods described in my study could be expanded to include more local individuals to see if higher-order structures, such as the recognized species boundaries, are reflected in the relationships between individuals. This study has suggested that local *Alaria* species boundaries are suspect, and more data points would probably add weight to such arguments.

Although unconventional, the apparent variation seen in this study using these methods could shed light on local oceanographic patterns. Kelp meiospores are restricted to perhaps 24 hours of time in the plankton (Reed 1990), restricting propagule migration to less than a hundred kilometers in the best circumstances. Given that resolution of nuclear gene flow measurements typically spans on the order of a few hundred generations (Slatkin 1985b, Slatkin 1987), an average of recent hydrodynamic currents in nearshore areas could be constructed from the relatedness of various patches. This method would assume that the extant populations examined are not recent recolonizations and therefore near equilibrium. Gene flow rates would then map to amounts of surface flow between regions.

References

Abbott, I. A., and G. J. Hollenberg. 1976. *Marine Algae of California*, Stanford University Press, Stanford, CA, 827 pp.

Amsler, C.D. and M. Neushul. 1989. Chemotactic effects of nutrients on spores of the kelps *Macrocystis pyrifera* and *Pterygophora californica*. *Mar. Biol.* **102**:557-564.

Amsler, C.D. and R.B. Searles. 1980. Vertical distribution of seaweed spores in a water column offshore of North Carolina. *J. Phycol.* **16**:617-619.

Anderson, E.K. and W.J. North. 1966. *In situ* studies of spore production and dispersal in the giant kelp *Macrocystis pyrifera*. *Proc. Int. Seaweed Symp.* **5**:73-86.

Aquadro, C. F., Desse, S. F., Bland, M. M., Langley, C. H. and C. C. Laurie-Ahlberg. 1986. Molecular population genetics of the alcohol dehydrogenase gene region of *Drosophila melanogaster*. *Genetics* **114**:1165-1190.

Baldwin, B. G., Sanderson, M. J., Porter, J. M., Wojciechowski, M. F., Campbell, C. S., and M. J. Donoghue. The ITS1 region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Ann. Mo. Bot. Gard.* **82**:247-277.

Berghammer, H. and B. Auer. 1993. "Easypreps": fast and easy plasmid minipreparation for analysis of recombinant clones in *E. coli*. *BioTechniques* **14**:519-520.

Bhattacharya, D., Stickel, S.K., and M.L. Sogin. 1991. Molecular phylogenetic analysis of actin genic regions from *Achlya bisexualis* (Oomycota) and *Costaria costata* (Chromophyta). *J. Mol. Evol.* **33**:525-536.

Bisalputra, T., Shields, C.M., and J.W. Markham. 1971. In situ observations of the fine structure of *Laminaria* gametophytes and embryos in culture. *J. Microscopie* **10**:83-98.

Bolton, J.J., Germann, I., and K. Lüning. 1983. Hybridization between Atlantic and Pacific representatives of the Simplices section of *Laminaria* (Phaeophyta), *Phycologia* **22**:133-140.

Breathnach, R. and P. Chambon. 1981. Organization and expression of eucaryotic split genes coding for proteins. *Ann. Rev. Biochem.* **50**:49-83.

Brody, E. and J. Abelson. 1985. The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. *Science* **228**:963-967.

Brown, J. W. S. 1996. *Arabidopsis* consensus intron sequences. *Plant Mol. Biol.* **32**:531-535.

Buckler, E. S., and T. P. Holtsford. 1996. *Zea* systematics: ribosomal ITS evidence. *Mol. Biol. Evol.* **13**:612-622.

Campbell, C. S., Wojciechowski, M. F., Baldwin, B. G., Alice, L. A., and M. J. Donoghue. 1997. Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* agamic complex (Rosaceae). *Mol. Biol. Evol.* **14**:81-90.

Cabot, E. L., and A. T. Beckenbach. 1989. Simultaneous editing of multiple nucleic acids and protein sequences with ESEE. *Comput. Applic. Biosci.* **5**:233-234.

Carter, M. J., and I. D. Milton. 1993. An inexpensive and simple method for DNA purifications on silica particles. *Nuc. Acids Res.* **21**:1044.

Cech, T. R. 1986. The generality of self-splicing RNA: relationship to nuclear mRNA splicing. *Cell* **44**:207-210.

Chapman, A. R. O. 1973. Methods for macroscopic algae. *In* Stein, J.R. [Ed.] *Handbook of Phycological Methods: Culture Methods and Growth Measurements*. Cambridge University Press, Cambridge, pp.87-104.

Chapman, A.R.O. 1974. The genetic basis of morphological differentiation in some *Laminaria* populations. *Mar. Biol.* **24**:85-91.

Chapman, A.R.O. 1984. Reproduction, recruitment and mortality in two species of *Laminaria* in southwest Nova Scotia. *J. Exp. Mar. Biol. Ecol.* **78**:99-109.

Chuang, S.-E., Wang, K.-C., and A.-L. Cheng. 1995. Single-step direct cloning of PCR products. *Trends Genet.* **11**:7-8.

Cosson, J and R. Olivari. 1982. Premiers résultats concernant les possibilités d'hybridization interspécifiques et intergénériques chez les Laminariales des côtes de la Manche. *Compte Rendu Hebdomadaire des Séances de l'Académie des Sciences, Paris* **295**:381-384

Costa, M., Fontaine, J. M., Loiseaux-de Goer, S., and F. Michel. 1997. A group II self-splicing intron from the brown alga *Pylaiella littoralis* is active at unusually low magnesium concentrations and forms populations of molecules with a uniform conformation. *J. Mol. Biol.* **274**:353-364.

Coyer, J. A., Engle, J. M. and R. C. Zimmerman. 1992. Discovery of a fertile *Pelagophycus* x *Macrocystis* (Phaeophyta) putative hybrid and subsequent production of F_2 sporophytes in the laboratory. *J. Phycol.* **28**:127-130.

Coyer, J. A., and A. Zaugg-Haglund. 1982. A demographic study of the elk kelp, *Pelagophycus porra* (Laminariales, Lessoniaceae), with notes on *Pelagophycus* x *Macrocystis* hybrids. *Phycologia* **21**:399-407.

Davis, J. I., and K. C. Nixon. 1992. Populations, genetic variation, and the delimitation of phylogenetic species. *Syst. Biolo.* **41**:421-435.

Davis, N., Van Blaricom, G.R., and P. K. Dayton. 1982. Man-made structures on marine sediments: Effects on adjacent benthic communities. *Mar. Biol.* **70**:295-303.

Dayton, P.K. 1973. Dispersion, dispersal, and persistence of the annual intertidal

alga, *Postelsia palmaeformis* Ruprecht. *Ecol.* **54**:433-438.

Dayton, P.K., Currie, V., Gerrodette, T., Keller, B. Rosenthal, R., and D. Ven Tresca. 1984. Patch dynamics and stability of some southern California kelp communities. *Ecol. Monogr.* **54**:253-289.

Del Sal, G., Manfioletti, G., and C. Schneider. 1989. The CTAB-DNA precipitation method: a common mini-scale preparation of template DNA from phagemids, phages, or plasmids suitable for sequencing. *Biotechniques* 7:514-520.

Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K., and J.S. Mattick. 1991. 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res.* **19**:4008.

Dover, G. A. 1982. Molecular drive: a cohesive mode of species evolution. *Nature* **299**:111-116.

Dressler, R. L. 1981. *The orchids: natural history and classification*, Harvard University Press, Cambridge, MA. 332pp.

Druehl, L. D. 1967. Distribution of two species of *Laminaria* as related to some environmental factors. *J. Phycol.* **3**:103-108.

Druehl, L. D. 1970. The pattern of Laminariales distribution in the northeast Pacific. *Phycologia* **9**:237-247.

Druehl, L.D. 1981. Geographical distribution. *In* (C.S. Lobban and M.J. Wynne, eds.) *The Biology of Seaweeds*. Univ. of Ca. Press, Berkeley, California. pp. 306-325.

Druehl, L. D., and G. W. Saunders. 1992. Molecular explorations in Kelp evolution. *Prog. Phycol. Res.* **8**:47-83.

Druehl, L. D., Mayes, C., Tan, I. H., & Saunders, G. W. (1997). Molecular and morphological phylogenies of kelp and associated brown algae. *Pl. Syst. Evol. [Suppl.]*, **11**:221-235.

Duggins, D.O., Simenstad, C.A. and J.A. Estes. 1989. Magnification of secondary production by kelp detritus in coastal marine ecosystems. *Science* **245**:170-173.

Ebeling, A.W., Laur, D.R., and R.J. Rowley. 1985. Severe storm disturbances and the reversal of community structure in a southern California kelp forest. *Mar.Biol.* **84**:287-294.

Fager, E.W. 1971. Pattern in the development of a marine community. *Limnol. Oceanogr.* **16**:241-253.

Erlich, H. A., Gelfand, D., and J. J. Sninsky. 1991. Recent advances in the Polymerase Chain Reaction. *Science* (Wash. D.C.) **252**:1643-1650.

Fain, S., Druehl, L. D., and D. L. Baillie. 1988. Repeat and single copy sequences are differently conserved in the evolution of kelp chloroplast DNA. *J. Phycol.* **24**:292-302.

Felsenstein, J. 1995. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**:783-791.

Felsenstein, J. 1988. Phylogenies from molecular sequences: Inference and reliability. *Annu. Rev. Genet.* **22**:521-565.

Felsenstein, J. 1993. *PHYLIP (Phylogeny Inference Package) version 3.5c.* Distributed by the author. Department of Genetics, University of Washington, Seattle, WA, USA.

Fitch, D. H. A., Mainone, C., Goodman, M., and J. L. Slightom. 1990. Molecular history of gene conversions in the primate fetal β-globin genes. *J. Biol. Chem.* **265**:781-793.

Fuller, C. W. 1992. Modified T7 DNA polymerase for DNA sequencing. *Methods in Enzymology*. **216**:329-354.

Harris, H. 1995. *The Cells of the Body: A History of Somatic Cell Genetics*, Cold Spring Harbor Press, New York, pp.131-134.

He, Q., Marjamaki, M., Soini, H., Mertsola, J., and M. K. Viljanen. 1994. Primers are decisive for sensitivity of PCR. BioTechniques **17**:82-87.

Hebsgaard, S. M., Korning, P.G., Tolstrup, N., Engelbrecht, J., Rouze, P., and S. Brunak. 1996. Splice site prediction in *Arabidopsis thaliana* DNA by combining local and global sequence information. *Nuc. Acids Res.* **24**:3439-3452.

Hightower, R.C., and R.B. Meagher. 1986. The molecular evolution of actin. *Genetics* **114**:315-332.

Hillis, D., M. and M. T. Dixon. 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. *The Quarterly Review of Biology* **66**:411-453.

Hillis, D. M. and C. Moritz. 1990. *Molecular Systematics*, Sinauer Associates, Sunderland, Massachusetts, 588 pp.

Hudson, R.R., Slatkin, M., and W.P. Maddison. 1992. Estimation of levels of gene flow from DNA sequence data. *Genetics* **132**:583-589.

Huntley, J. F. 1996. Peak patterns seen using AmpliTaq® DNA Polymerase, FS. In *Applied Biosystems*, DNA Sequencing Facility, Iowa State University, Ames, Iowa. 3 pp.

Jacobsen, I. B., Wilson, S. R., and S. Easteal. 1997. The partition matrix: exploring variable phylogenetic signals along nucleotide sequence alignments. *Mol. Biol. Evol.* **14**:474-484.

Jin, H-J, Kim, J-H, Sohn, C.H., DeWreede, R.E., Choi, T-J, Towers, G.H.N., Hudson, J. B., and Y-K Hong. 1997. Ihibition of *Taq* polymerase by seaweed extracts from British Columbia, Canada and Korea. *J. Appl. Phycol.* **9**:383-388.

Johnson, S. R., Martin, D. H., Cammarata, C., and S. A. Morse. 1995. Alterations in sample preparation increase sensitivity of PCR assay for diagnosis of chanchroid. *J. Clin. Microbiol.* **33**:1036-1038.

Johnson, C.R., and K.H. Mann. 1988. Density, patterns of adaptation, and stability of Nova Scotian kelp beds. *Ecol. Mono.* **58**:129-154.

Jukes, T. H. and C. R. Cantor. 1969. Evolution of protein molecules. *In* Munro, H. N. [Ed.] *Mammalian Protein Metabolism*. Academic Press, New York, USA. pp. 21-132.

Kaine, B. P. and B.B. Spear. 1982. Nucleotide sequence of a macronuclear gene for actin in *Oxytricha fallax*. *Nature* **295**:430-432.

Kimura, M. 1968. Evolutionary rate at the molecular level. *Nature* **217**:624-626.

Kimura, M. 1980. A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**:111-120.

Kleppe, K., Ohtsuka, E., Kleppe, R., Molineux, I., and H. G. Khorana. 1971. Repair replication of short synthetic DNAs as catalyzed by DNA polymerases. *J. Mol. Biol.* **56**:341-361.

Kusumo, H. T. 1998. *A Parallel Assessment of Morphological and Genetic Diversity*

In the Kelp Alaria marginata. Ph.D. thesis, Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada, 96 pp

Lessa, E.P. 1992. Rapid surveying of DNA sequence variation in natural populations. *Mol. Biol.Evol.* **9**:323-330.

Lewis, R.J. 1996. Chromosomes of the brown algae. *Phycologia* **35**:19-40.

Lewis, R. J., Jiang, B. Y., Neushul, M., & Fei, X. G. 1993. Haploid parthenogenetic sporophytes of *Laminaria japonica* (Phaeophyceae). *J. Phycol.* **29**:363-369.

Lewis, R.J. and M. Neushul. 1994. Northern and Southern Hemisphere hybrids of *Macrocystis* (Phaeophyceae). *J. Phycol.* **30**:346-353.

Lewis, R.J. and M. Neushul. 1995. Intergeneric hybridization among five genera of the family Lessoniaceae (Phaeophyceae) and evidence for polyploidy in a fertile *Pelagophycus* x *Macrocystis* hybrid. *J. Phycol.* **31**:1012-1017.

Li, W., -H. 1983. Evolution of duplicate genes and pseudogenes. *In* Nei, M & Koehn, R.K. [Eds.] *Evolution of Genes and Proteins*. Sinauer Assoc., Sunderland, MA, pp. 14-37.

Li, W., -H. and D. Graur. 1991. *Fundamentals of Molecular Evolution.* Sinauer Associates, Sunderland, Massachusetts, 284 pp.

Lüning, K., and M. J. Dring. 1975. Reproduction, growth and photosynthesis of gametophytes of *Laminaria saccharina* grown in blue and red light. *Mar. Biol. (Berl.)* **29**:195-200.

Lynch, M. and T.J. Crease. 1990. The analysis of population survey data on DNA sequence variation. *Mol. Biol. Evol.* 7:377-394.

Mackenzie, I. 1997. *Genetic and morphological variation in* Macrocystis *C. Agardh (Laminariales, Phaeophyceae) off the North American coast.* M.Sc. thesis, Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada, 78 pp.

Maier, I. and D.G. Müller. 1990. Chemotaxis in *Laminaria digitata* (Phaeophyceae). *J. Exp. Botany* **41**:869-876.

Mallet, J. 1995. A species definition for the Modern Synthesis. TREE **10**:294-299.

Marchuk, D., Drumm, M., Saulino, A., and F.S. Collins. 1991. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR

products. *Nuc. Acids Res.* **19**:1154.

Maruyama, T. 1972. Distribution of gene frequencies in a geographically structured finite population. I. Distribution of neutral genes and of genes with small effect. *Ann. Hum. Gent.* **35**:411-423.

Mayes, C., Saunders, G. W., Tan, I. H. and L.D. Druehl. 1992. DNA extraction methods for kelp (Laminariales) tissue. *J. Phycol.* **28**: 712-716.

Mount, S.M. 1982. A catalog of splice junction sequences. *Nuc. Acids Res.* **10**:459-472.

Mróz., L. W. 1989. *Ribosomal DNA assessment of genetic relationships among Alaria Greville (Phaeophyta) stands from Vancouver Island and vicinity.* M.Sc. thesis, Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada, 70 pp.

Müller, D. G. 1967. Generationswechsel, Kernphasenwechsel und Sexualität der Braunalge *Ectocarpus sliculosus* im Kulturversuch. *Planta* **75**:39-54.

Munda, I. M. 1992. Gradient in seaweed vegetation patterns along the North Icelandic coast, related to hydrographic conditions. *Hydrobiologia* **242**: 133-

147.

Myakishev, M. V., Kapanadze, G. I., Shaikhayev, G. O., Georgiev, G. P., and D. R. Beritashvili. 1995. *Extraction of DNA from the whole blood by silica gel.* Institute of Gene Biology, Moscow, Russia, http://www.ras.ru/biogen/blood.html

Nakahara, H. 1984. Alternation of generations of some brown algae in unialgal and axenic cultures. *Scientific Papers of the Institute of Algological Research, Faculty of Science, Hokkaido University* 7:77-292.

Nakahara, H. and Y. Nakamura. 1973. Parthenogenesis, Apogamy and Apospory in *Alaria crassifolia* (Laminariales). *Mar. Biol. (Berl.)* **18**:327-332.

NC-IUB (Nomenclature Commitee of the International Union of Biochemistry). 1984. Nomenclature for incompletely specified bases in nucleic acid sequences. *Eur. J. Biochem.* **150**:1-5

Neapolitan, R. and K. Naimipour. 1996. *Foundations of Algorithms*. Jones and Bartlett, Sudbury, MA. 523 pp.

Nichols, R. C., and N. Raben. 1994. Hints for direct sequencing of PCR-generated single-stranded DNA. *BioTechniques* **17**:412-414.

Neushul, M. 1971. *Macrocystis* and related kelps. *Nova Hedwigia* **32**:223-228.

Norton, T.A. 1992. Dispersal by Macroalgae. *Br. phycol. J.* **27**:293-301.

Novick, P. and D. Botstein. 1985. Phenotypic analysis of temperature-sensitive yeast actin mutants. *Cell* **40**:405-416.

O'Clair, R. M., Lindstrom, S. C., and I. R. Brodo. 1996. *Southeast Alaska's Rocky Shores*. Plant Press, Auke Bay, Alaska. pp. 50-52.

Palmer, A.R., and R.R. Strathmann. 1981. Scale of dispersal in varying environments and its implications for life histories of marine invertebrates. *Oecologia* **48**:308-318.

Palumbi, S. R., and C. S. Baker. 1994. Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. *Mol. Biol. & Evol.* **11**:426-435.

Penny, D. 1982. Towards a basis for classification: The incompleteness of distance measures, incompatibility analysis and phenetic classification. *J. Theoret. Biol.* **96**:129-142.

Polanshek, A. R., and J. A. West. 1975. Culture and Hybridization studies on *Petrocelis* (Rhodophyta) from Alaska and California. *J. Phycol.* **11**:434-439.

Reed, D.C. 1990. The effects of variable settlement and early competition on patterns of kelp recruitment. *Ecol.* **71**:776-787.

Reed, D.C. and A.W. Ebeling. 1991. Dispersal potential and the degree of openness in kelp populations. *Proc. 72nd mtg. West. Soc. Nat., Santa Barbara, CA, December 1991* (abstract)

Reed, D.C., Laur, D.R., and A.W. Ebeling. 1988. Variation in algal dispersal and recruitment: the importance of episodic events. *Ecol. Monogr.* **58**:321-335.

Reese, M. G., and F. H. Eckman. 1996. *Splice sites: A detailed neural network*

study. Poster at 1996 Genome Mapping & Sequencing Meeting, Cold Spring Harbor, New York (http://www-hgc.lbl.gov/inf/splice-abstract.html, Accessed 11/1999).

Rueness, J. 1973. Speciation in *Polysiphonia* (Rhodophyta, Ceramiales) in view of hybridization experiments: *P. hemisphaerica* and *P. boldii*. *Phycologia* **12**:107-109.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, T. T., Mullis, K. B. and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science (Wash. D.C.)* **239**:487-491.

Sambrook, J., Fritsch, E. F. and T. Maniatis. 1989. *Molecular Cloning: a Laboratory Manual, 2nd ed.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 400 pp.

Sanbonsuga, Y. and M. Neushul. 1978. Hybridization of *Macrocystis* with other float bearing kelps. *J. Phycol* **14**:214-24.

Santelices, B. 1990. Patterns of reproduction, dispersal and recruitment in seaweeds. *Oceanogr. Mar. Biol. Annu. Rev.* **28**:177-276.

Saunders, G. W. 1991. *Molecular evolution and phylogeny of the Laminariales: analysis of the nuclear encoded ribosomal cistron.* Ph.D. thesis, Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada, 130 pp.

Saunders, G. W. 1993. Gel purification of red algal genomic DNA: an inexpensive and rapid method for the isolation of polymerase chain reaction-friendly DNA. *J. Phycol* **29**:251-254.

Saunders, G. W., and L. D. Druehl. 1993. Revision of the kelp family Alariaceae and the taxonomic affinities of *Lessoniopsis* Reinke (Laminariales, Phaeophyta). *Hydrobiologia* **260/261**: 689-697.

Scagel, R. F. 1963. Distribution of attached marine algae in relation to oceanographic conditions in the northeast Pacific. *In* Dunbar, M. J. [Ed.] *Marine Distributions.* University of Toronto Press, Roy. Soc. Canada Spec. Publ. No. 5. Toronto, Canada. pp. 37-50.

Scagel, R. F., Gabrielson, P. W., Garbary, D. J., Golden, L., Hawkes, M. W., Lindstrom, S. C., Oliviera, J. C., and T. B. Widdowson. 1989. *A Synopsis of the Benthic Marine Algae of British Columbia, Southeast Alaska, Washington,*

and Oregon. University of British Columbia, Department of Botany. 532 pp.

Simonsen, K. L., Churchill, G. A., and C. F. Aquadro. 1995. Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics* **141**:413-429.

Slatkin, M. 1985a. Rare alleles as indicators of gene flow. *Evolution* **39**:53-65.

Slatkin, M. 1985b. Gene flow in natural populations. *Ann. Rev. Ecol. Syst.* **16**:393-430.

Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. *Science* **236**:787-792.

Slatkin, M., and N.H. Barton. 1989. A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* **43**:1349-1368.

Solovyev, V. V., Salamov, A. A., and C. B. Lawrence. 1994. Predicting internal exons by oligonucleotide composition and discriminant analysis of spliceable open reading frames. *Nuc. Acids Res.* **22**:5156-5163.

Sundene, O. 1958. Interfertility between forms of *Laminaria digitata*. *Nytt Mag. Bot.* **6**:121-128.

Swofford, D. L., and D. J. Olsen. 1990. Phylogeny reconstruction. In Hillis, D. M. and C. Moritz [Eds.] *Molecular Systematics*. Sinauer Associates, Sunderland, Massachusetts, pp. 411-500.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**:585-595.

Tan, I. H., and L. D. Druehl. 1994. A molecular analyses of *Analipus* and *Ralfsia* (Phaeophyceae) suggests the order Ectocarpales is polyphyletic. *J. Phycol.* **30**:721-729.

Templeton, A. R., Crandall, K. A, and C. F. Sing. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data III. Cladogram Estimation. *Genetics* **132**:619-633.

Tindall, K. R., and T. A. Kunkel. 1988. Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* **27**:6008-6013.

Tokida, J., Ohmi, H., and M. Imashima. 1958. A chimaera of *Alaria* and *Laminaria* found in nature. *Nature (Lond.)* **181**: 924-925.

Tom Diek (Bartsch), I. 1992. North Pacific and North Atlantic digitata *Laminaria* species (Phaeophyta): hybridization experiments and temperature responses. *Phycologia* **31**:147-163.

Underwood, A.J., and E.J. Denley. 1984. Paradigms, explanations, and generalizations in models for the structure of intertidal communities on rocky shores. *In* Strong, D.R., Simberloff, D., Abele, L.G.,and A.B. Thistle, [Eds.].*Ecological Communities: Conceptual Issues and the Evidence*. Princeton University Press, Princeton NJ. pp 151-180.

Waaland, J. R., 1977. *Common Seaweeds of the Pacific Coast.* Pacific Search Press, Seattle WA.

Walsh, P. S., Metzger, D. A., and R. Higuchi. 1991. Chelex™ 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* **10**:506-513.

Weir, B. 1990. Intraspecific differentiation. *In* Hillis, D. M. and C. Moritz [Eds.] *Molecular Systematics*. Sinauer Associates, Sunderland, Massachusetts, pp 373-410.

West, J. A., Polanshek, A. R., and D. E. Shevlin. 1978. Field and culture studies on *Gigartina agardhii* (Rhodophyta). *J. Phycol.* **14**:416-426.

Widdowson, T.A. 1971. A taxonomic revision of the genus *Alaria* Greville. *Syesis* **4**:11-49.

Wiedbrauk, D. L., Werner, J. C., and A. M. Drevon. 1995. Inhibition of PCR by aqueous and vitreous fluids. *J. Clin. Microbiol.* **33**:2643-2646.

Wildeman, A. G. 1988. A putative ancestral actin gene present in a thermophilic eukaryote: novel combination of intron positions. *Nuc. Acids Res.* **16**:2553-2564.

Womersley, H. B. S. 1954. The species of *Macrocystis* with special reference to those on southern Australian coasts. *Univ. Calif. Publ. Bot.* **27**:109-132.

Wright, S. 1951. The genetical structure of populations. *Ann. Eugen.* **15**:323-354.

Wu, D.Y., Ugozzoli, L., Pal, B.K., Qian, J., and R.B. Wallace. 1991. The effects of temperature and oligonucleotide primer length on the specificity and efficiency of amplification by the Polymerase Chain Reaction. *DNA & Cell Bio.* **10**:233-238.

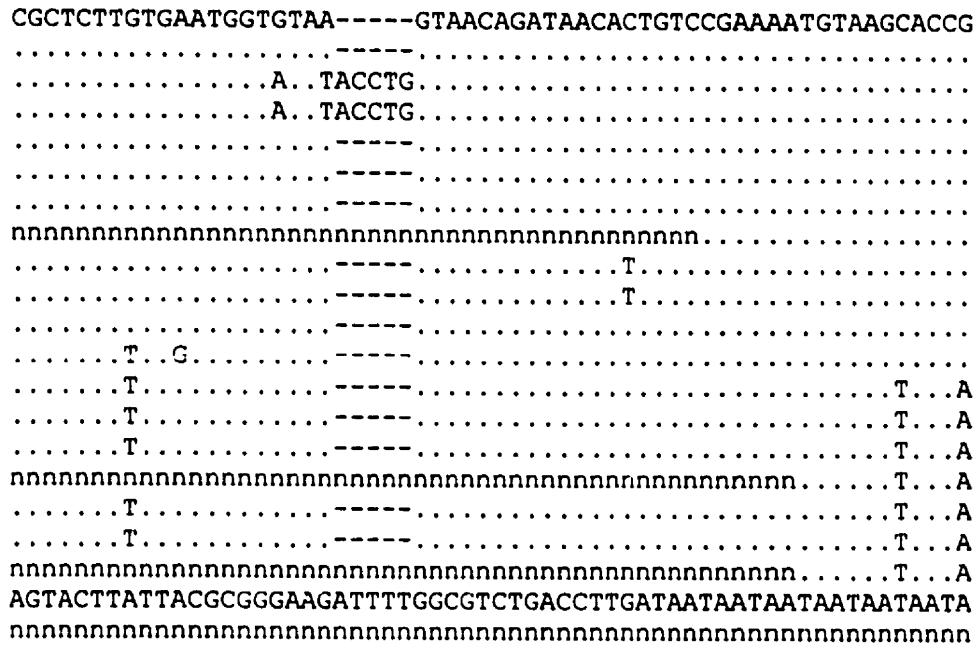
Yabu, H. and Y. Sanbonsuga. 1987. Chromosome count in *Macrocystis integrifolia* Bory. *Bull. Fac. Fish., Hokkaido Univ.* **38**:339-342.

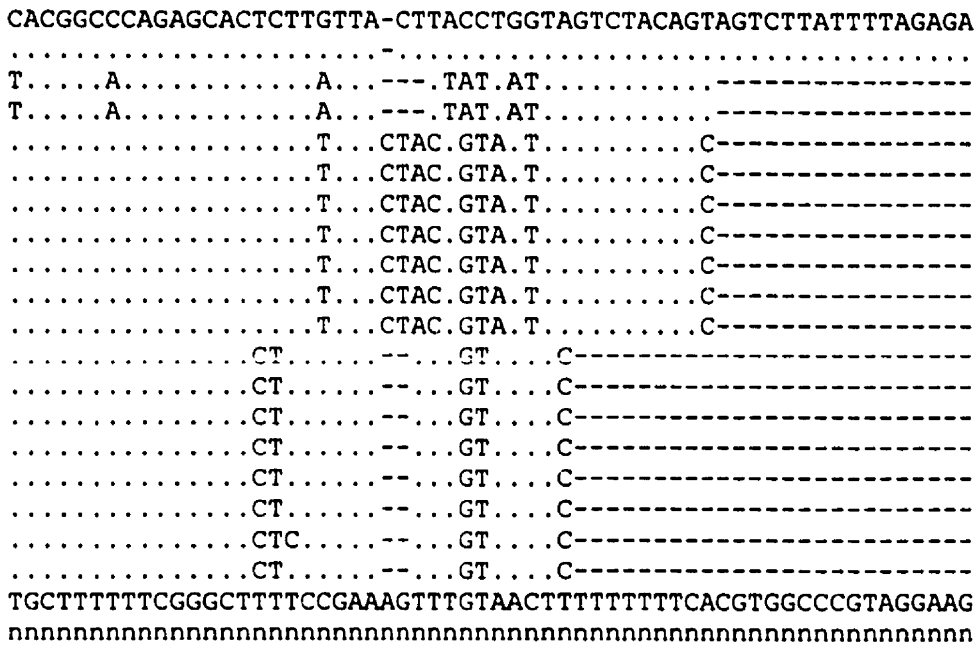
Appendix I

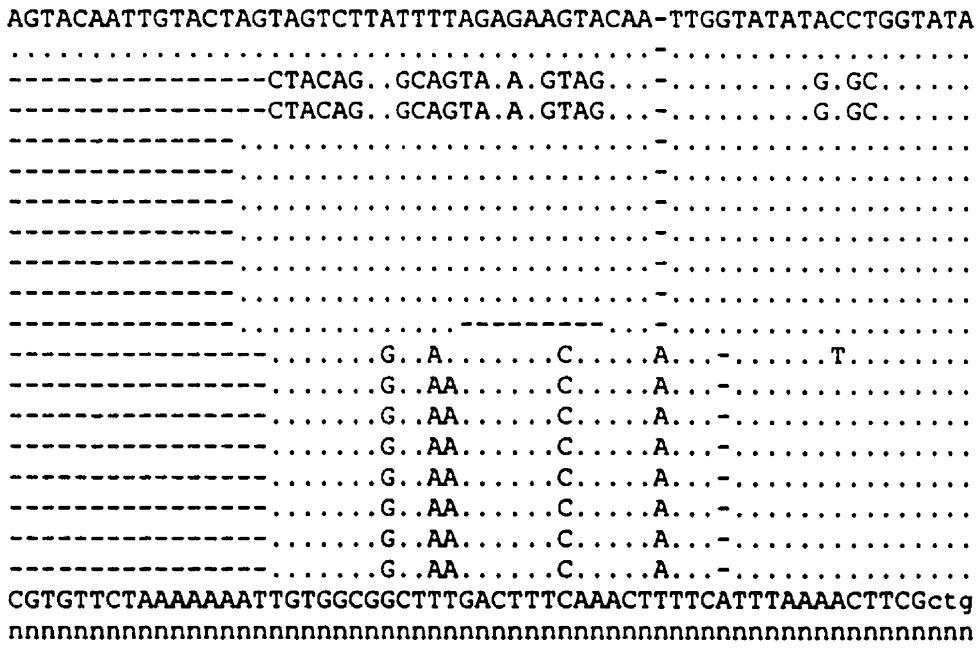
Alignment of sequences from clones of the actin Intron II-containing region of both *Alaria* and *Nereocystis*. Designations are clone numbers except for the *Nereocystis* sequences, which are indicated by individual designation. Clone numbers listed together (separated by slashes) indicate that the two clones were identical. In addition, NLCI04 represents two identical clones (5b and 5e). Clone identities are as follows: 5a2, 5b2, and 5c2 = AM06SR; 40.1 and 40.5 = AM03KB; 2.1, 2.2, 2.3, and 2.4 = AM03RB; 4.1 and 4.2 = AM01WB; 36.5 = AT01OP, 6 = AM03SK; 16.2, 16.5, and 16.10 = AM03JR; 28.4, 28.5, 28.8, 28.9, and 28.10 = AN10BO; 5b and 5e = NLCI04. A 'dot' (.) indicates identity with the AM06SR 5a2 sequence shown, a 'dash' (-) indicates an insertion or deletion added to allow alignment, while an 'n' indicates a base not determined in this study.

```
5a2          GAGAAGTCGTACGAGCTCCCCGACGGAAACGTCATCGTCATCGGAAACG
5b2/5c2      ................................................
40.1         nnnnnnnnnnnnnnnnnnn.............................
40.5         ................................................
2.1          ................................................
2.2          nnnnnnnnn.......................................
2.3          ................................................
2.4          nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
4.1          ................................................
4.2          ................................................
36.5         ................................................
6            ................................................
16.2         ............................................G....
16.5/16.10   ............................................G....
28.4         ................................................
28.5         nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
28.8         ................................................
28.9         ................................................
28.10        nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
NLCI04       ............................................G...
NLVR/NLJB    nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
```

```
AGCGCTTCCGTTGCCCCGAGGTATGATACAAAATG----------TACTGTGACAAAG
.....................................----------.............
.....................................----------.....GT......
.....................................----------.....GT......
.....................................ATAATAAAAATG.............
.....................................ATAATAAAAATG.............
.....................................ATAATAAAAATG.............
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
.....................................-.....ATAATAAAAATG.....
.....................................ATAATAAAAATGC............
.....................................ATAATAAAAATG.............
.............................T.......----------.....GT......
.............................T.......----------.....GT......
.............................T.......----------.....GT......
.............................T.......----------.....GT......
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
.............................T.......----------.....GT......
.............................T.......----------.....GT......
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
.....................GTATGTAAATATGATCGTTAATCATTACCAGGGTATTAAT
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
```


CGCTCTTGTGAATGGTGTAA-----GTAACAGATAACACTGTCCGAAAATGTAAGCACCG
AGTACTTATTACGCGGGAAGATTTTGGCGTCTGACCTTGATAATAATAATAATAATAATA
ATGATAATGATACATACCAGGTACCTGGTACCGATGTTGGCCCAACATCGGCCCCAGATT


CACGGCCCAGAGCACTCTTGTTA-CTTACCTGGTAGTCTACAGTAGTCTTATTTTAGAGA
TGCTTTTTTCGGGCTTTTCCGAAAGTTTGTAACTTTTTTTTTCACGTGGCCCGTAGGAAG
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn


AGTACAATTGTACTAGTAGTCTTATTTTAGAGAAGTACAA-TTGGTATATACCTGGTATA
CGTGTTCTAAAAAAAATTGTGGCGGCTTTGACTTTCAAACTTTTCATTTAAAACTTCGctg
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn

```
------------------------------------------------------------
------------------------------------------------------------
TGAATACCGTATTGTTATTACCAGCATAATTGATGTGTAAAATATATAGATGAACTGTAT
TGAATACCGTATTGTTATTACCAGCATAATTGATGTGTAAAATATATAGATGAACTGTAT
------------------------------------------------------------
------------------------------------------------------------
------------------------------------------------------------
------------------------------------------------------------
------------------------------------------------------------
------------------------------------------------------------
------------------------------------------------------------
------------------------------------------------------------
------------------------------------------------------------
------------------------------------------------------------
------------------------------------------------------------
------------------------------------------------------------
------------------------------------------------------------
tcccattggcatagtctcatggcaatatttggaaaagtgaagaatccgggnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
```

```
--------AAAGTGGCGATGAAATAAATAGCC-TGCTTCCGC----------------
--------.......................-.........-----------------
TCAAGATA..G....G.......G.T......-.........GCCTTTTTTTCGCGCCCG
TCAAGATA..G....G.......G.T......-.........GCCTTTTTTTCGCGCCCG
--------..............G........-.........CCCTTTTTT-CGCGCGGG
--------G.............G........-.........CCCTTTTTT-CGCGCGGG
--------..............G........-.........CCCTTTTTT-CGCGCGGG
--------..............G........-.........CCCTTTTTT-CGCGCGGG
--------..............G........-.........CCCTTTTTT-CGCGCGGG
--------..............G........-.........CCCTTTTTT-CGCGCGGG
--------..............G........-.........CCCTTTTTT-CGCGCGGG
--------..............G........-.........ACCTTTTTT-CGCGCGGG
--------..............G........-.........ACCTTTTTC-CGCGCGGG
--------..............G........-.........ACCTTTTTC-CGCGCGGG
--------..............G........-.........ACCTTTTTT-CGCGCGGG
--------..............G........-.........ACCTTTTTT-CGCGCGGG
--------..............G........-.........ACCTTTTTT-CGCGCGGG
--------..............G........A.........ACCTTTTTT-CGCGCGGG
--------..............G........-.........ACCTTTTTT-CGCGCGGG
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
```

```
----------------TAGCGTA----------CTTACCTGGATACCTGGTACCAGGTAG
----------------.......----------.........................
-TGGGG--TCTCTGTC..nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
-TGGGG--TCTCTGTC.......TAACCGTATA.CC.....................
TTGGG---TCTCCGTC.......--------TA...C.....................
TTGGG---TCTCCGTC.......--------TA...C.....................
TTGGG---TCTCCGTC.......--------TA...C.........T...........
TTGGG---TCTCCGTC.......--------TA...C.....................
TTGGG---TCTCCGTC.......--------TA...C................G.....
TTGGG---TCTCCGTC.......--------TA...C.....................
TTGGG---TCTCCGTC.......--------TA...C.....................
-TGGGGGGTCTCCGTC.......--------TA..................G.......
-TGGGGGGTCTCCGTC.......----------.........................
-TGGGGGGTCTCCGTC.......----------.........................
-TGGGGGGTCTCCGTC.......----------.........................
-TGGGGGGTCTCCGTC.......----------.........................
-TGGGGGGTCTCCGTC.......----------.........................
-TGGGGGGTCTCCGTC.......----------.........................
-TGGGGGGTCTCCGTC.......----------.........................
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
```

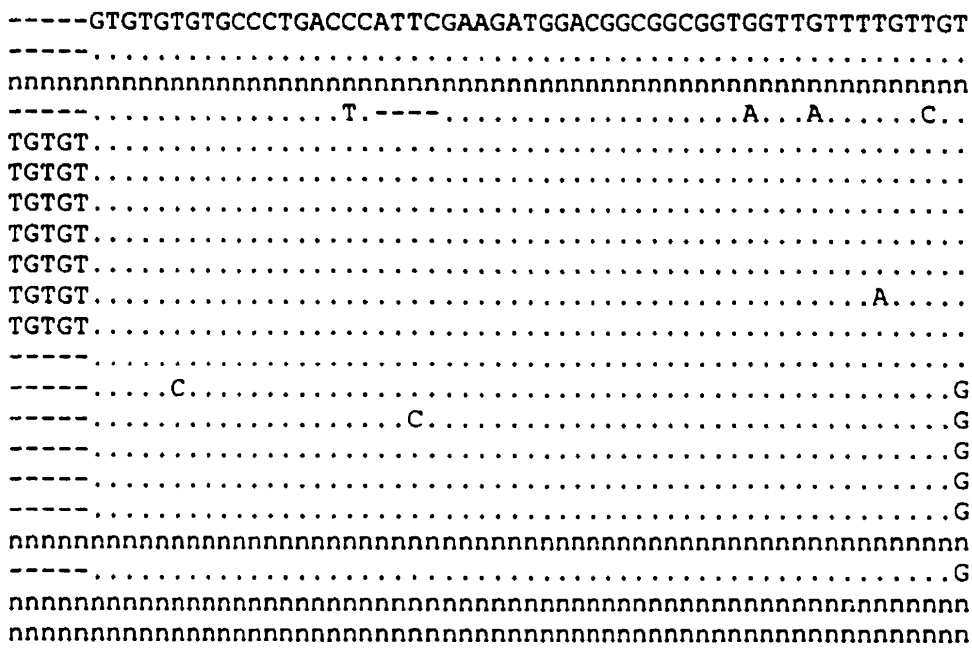
```
CGAGCCGTATGGTACCGGTGACC-GCACGTTAGCATTAACGGATTTTTGGGGGGTTTTCG
.......................-....................................
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
...-..A........A.....T.-......A.....................C.....
........G......A.......C.....................................
........G......A.......C.....................................
........G......A.......C.....................................
........G......A.......C.....................................
........G......A.......C.....................................
........G......A.......C.....................................
........G......A.......C.....................................
..............A.......-.....................................
..............A.......-......G..............................
..............A.......-......G..............................
..............A.......-......G..............................
..............A.......-......G..............................
..............A.......-......G..............................
..............A.......-......G..............................
..............A.......-......G..............................
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
```

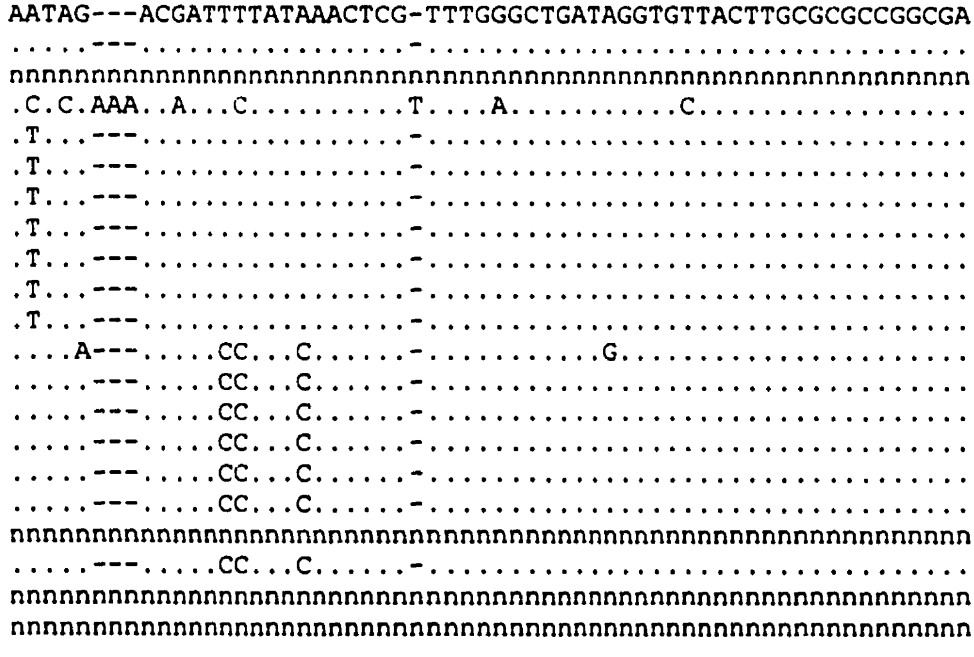
```
GGCATGCGCTGTTTCATTTTGCTCCTCGTACTCTCTGGAGACTCCCGGAGACTGGAGACC
.....-.....
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
.....T..G.....-.....
.....--.....
.....--.....
.....--.....
.....--.....
...G.....--.....
.....A.....--.....
.....--.....
.....C.....-.....
......G.....C.....-.....
......G.....C.....-.....
......G.....C.....-.....
......G.....C.....-.....G....
......G.....C.....-.....
......G.....nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
......G.....C.....-.....
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
```

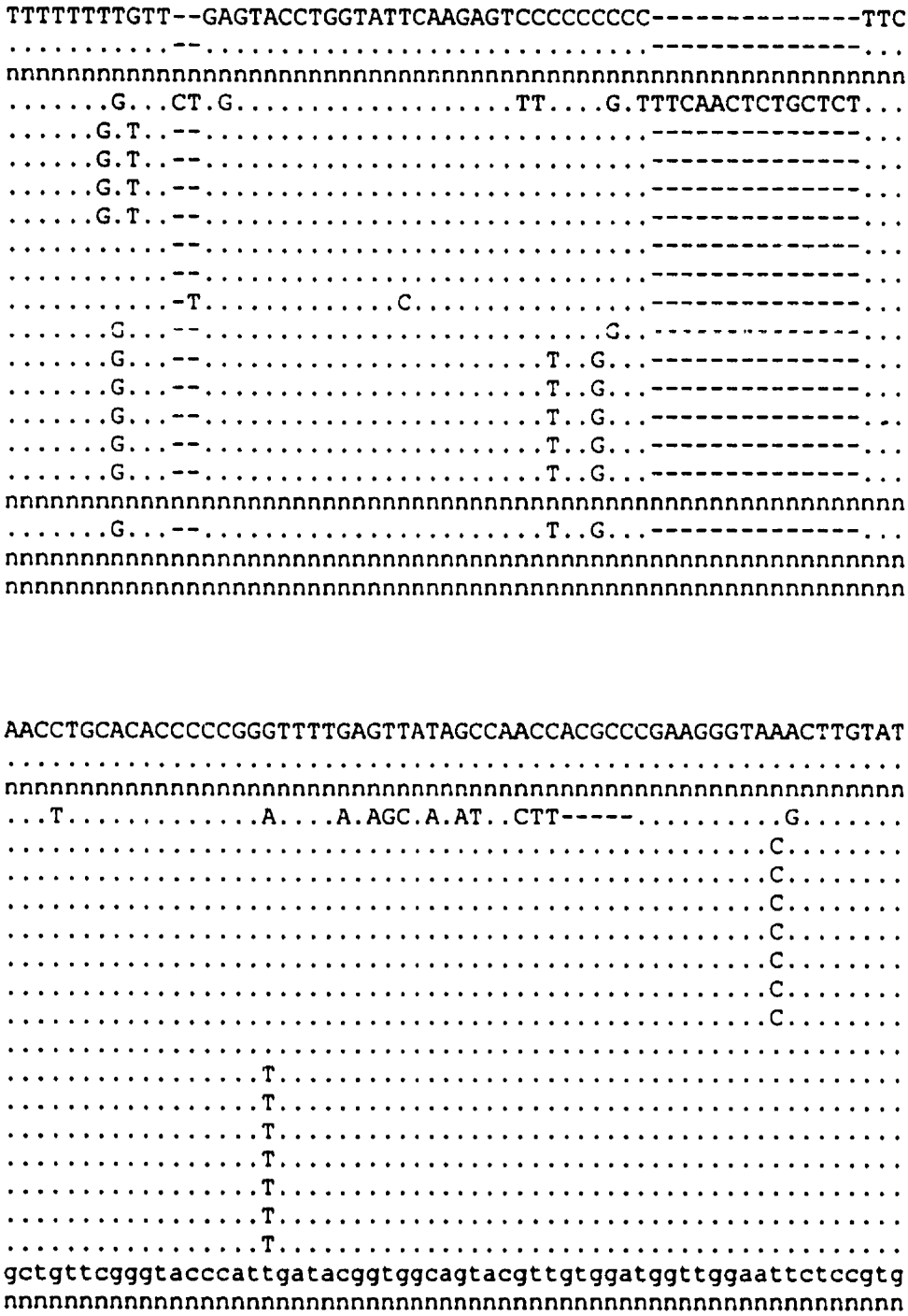
```
TCTCC-GTCTACTT----------TTTGGAGGTGGTACTAATTTCTCTCTTTTTGGTC
.....-.....----------.....
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
.....-.....-------ATCTGTT---ATGA.....C.....
.....C.....----------.....-.....
.....C.....----------.....-.....
.....C.....----------.....-.....
.....-.....----------.....-.....
.....-.....----------.....-.....
.....-.....----------.....-.....
.....-.....----------.....-.....
.....-.....TT-------ATCTG.....C.....-.....
.....-.....ACCTGGTATCTG.....
.....-.....ACCTGGTATCTG.....
.....-.....ACCTGGTATCTG.....
.....-.....ACCTGGTATCTG.....
.....-.....ACCTGGTATCTG.....
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
.....-.....ACCTGGTATCTG.....
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
```

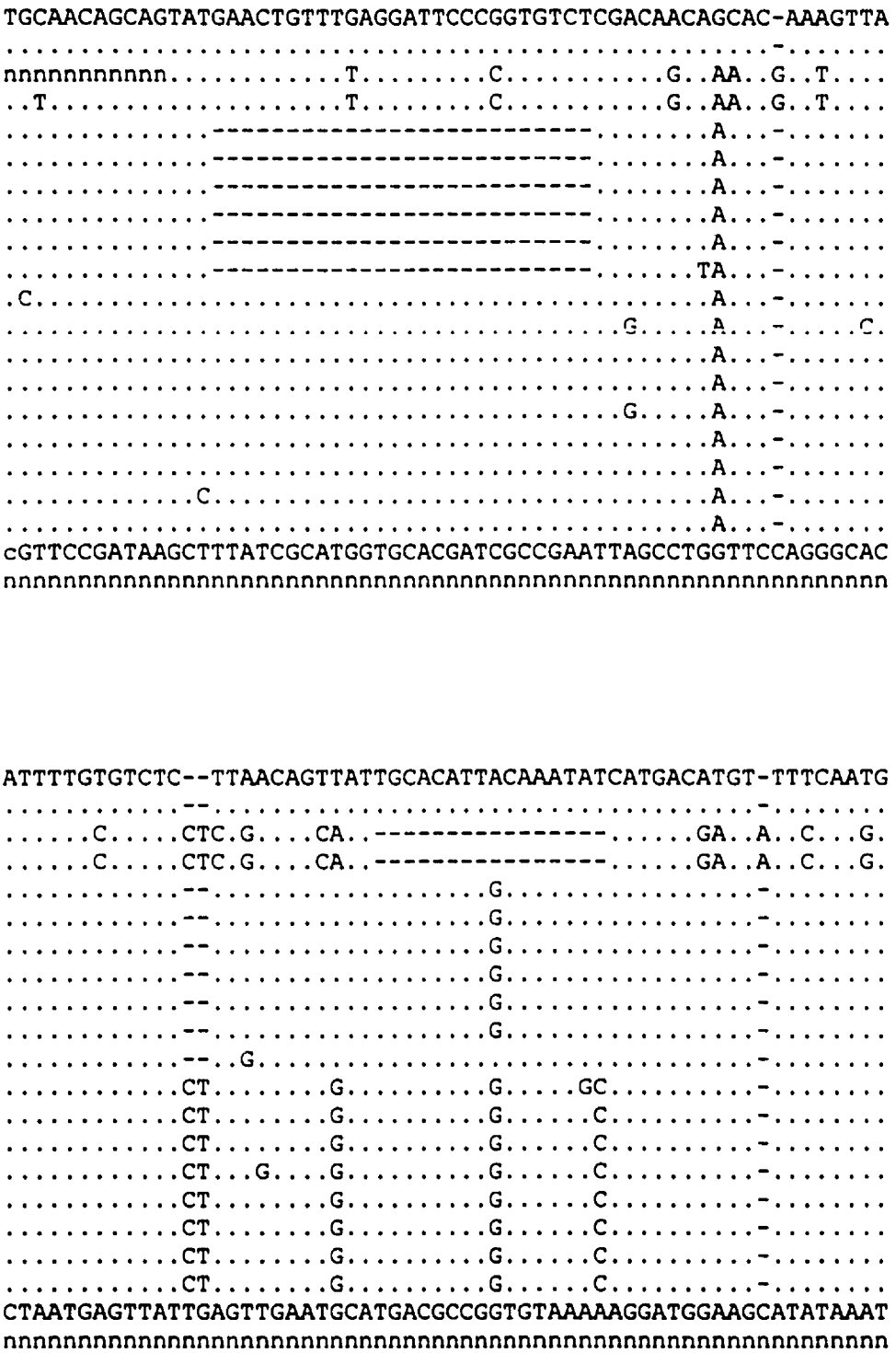
```
TGGTACGAGCGTTTTTGCGTTTCCTCACGGCAGTTGCACCAGCTGCTAAAATCTGGCGGG
.............................................................
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
..............C.....................................T..T....
.............................................................
.............................................................
.............................................................
.............................................................
.....................A.......................................
.............................................................
.............................................................
.............................................................
....C..................T...G...................................
....C..................T...G...................................
....C..................T...G...................................
....C..................T...G...................................
....C..................T...G...................................
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
....C..................T...G...................................
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
```

```
T-CTGTGTTCCGGGCGGGGGGTCTTGACTTCT----------------------------
.-...........................----------------------------
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
.CT........................C..GTGTCTTGTTT-----------------
.C..........................GTGTCTTGT----GTGTGTGTGTGCGTG
.C..........................GTGTCTTGT----GTGTGTGTGTGCGTG
.C..........................GTGTCTTGT----GTGTGTGTGTGCGTG
.C..........................GTGTCTTGT--------GTGTGTGCGTG
.C.....-....................GTGTCTTGT--------GTGTGTGCGTG
.C.....-....................GTGTCTTGT--------GTGTGTGCGTG
.C..........................GTGTCTTGTGTGTGTGTGTGTGCGTG
.C..........................GTGTCTTGT-------------------
.C...........T..............GTGTCTTGT-------------------
.C...........T..............GTGTCTTGT-------------------
.C...........T..............GTGTCTTGT-------------------
.C...........T..............GTGTCTTGT-------------------
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
.C...........T..............GTGTCTTGT-------------------
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
```






TTTTTTTTGTT--GAGTACCTGGTATTCAAGAGTCCCCCCCCC----------TTC
..........--..........----------...
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
.......G...CT.G.....TT....G.TTTCAACTCTGCTCT...
......G.T..--..........----------...
......G.T..--..........----------...
......G.T..--..........----------...
......G.T..--..........----------...
..........--..........----------...
..........--..........----------...
..........-T.....C.....----------...
.......G...--..........G..----------...
.......G...--.....T..G...----------...
.......G...--.....T..G...----------...
.......G...--.....T..G...----------...
.......G...--.....T..G...----------...
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
.......G...--.....T..G...----------...
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
AACCTGCACACCCCCGGGTTTTGAGTTATAGCCAACCACGCCCGAAGGGTAAACTTGTAT
.....
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
...T.....A....A.AGC.A.AT..CTT-----.....G.....
.....C.....
.....C.....
.....C.....
.....C.....
.....C.....
.....C.....
.....C.....
.....
.....T.....
.....T.....
.....T.....
.....T.....
.....T.....
.....T.....
.....T.....
gctgttcgggtacccattgatacggtggcagtacgttgtggatggttggaattctccgtg
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn


TGCAACAGCAGTATGAACTGTTTGAGGATTCCCGGTGTCTCGACAACAGCAC-AAAGTTA
cGTTCCGATAAGCTTTATCGCATGGTGCACGATCGCCGAATTAGCCTGGTTCCAGGGCAC
ATTTTGTGTCTC--TTAACAGTTATTGCACATTACAAATATCATGACATGT-TTTCAATG
CTAATGAGTTATTGAGTTGAATGCATGACGCCGGTGTAAAAAGGATGGAAGCATATAAAT

```
ATAAGGTTCCCAAGTTGTTAGCCTCGG---------CCAGGTATATGGATGTTGGGATCT
.........................---------.....................
...T...AT.......--.......--AACCAGGTA...........T......A..A.
...T...AT.......--.......--AACCAGGTA...........T......A..A.
...T.....................---------.....................
...T.....................---------.....................
...T.....................---------.....................
...T.....................---------.....................
...T...........C.........---------.....................
...T.....................---------.....................
...T.....................---------.....................
...T.....................CCAGGTATA.....................
..GT.....................---CTAGGTATA...................
..GT.....................---CTAGGTATA...................
..GT.....................---CTAGGTATG...................
..GT.....................---CTAGGTATA...................
..GT.....................---CTAGGTATA...................
..GT.....................---CTAGGTATA...................
..GT.....................---CTAGGTATA...................
GCCTATCAGAGACCTGGCATCTGTATTTTGTTAAGTGTCTGATGTTGTTTCTCTCGCCCG
nnnnnnnnnAGACCTGGCATCTGTATTTTGTTAAGTGTCTGATGTTGTTTCTCTCGCCCG
```

```
TTTGGCCCGCTTCACTTTTGTGATAACGTTCGTTGGACGTACGCGTATATTCACC--TCG
.....................................................--...
............A...........GC....T...................G..TGAG.T.
............A...........GC....T...................G..TGAG.T.
.......................C.............................--...
.......................C.............................--...
.......................C.............................--...
.......................C.............................--...
...........T...........C.............................--...
.......................C.............................--...
..............C.......C.................A...........--...
..A.....................G..............G.............--...
..A...C.................................G.............--...
..A.....................................G.............--...
..A.....................................G.............--...
..A.....................................G.............--...
..A.....................................G.............--...
..A.....................G..............G.............--...
..A.....................................G.............--...
TTTCTTTTTCACGACGTGATTCGACAATCCTCAGGAGCCTATTTATATGCCAGGAGACCT
TTTCTTTTTCACGACGTGATTCGACAATCCTCAGGAGCCTATTTATATGCCAGGAGACCT
```

```
TCTTCAGTACGT------CTC--TTTGCCAATTGTGCCTCCGTTCGTCCTCTC---TTTT
............------...--.................................---....
............TTGTCT...CAG.GT.T...ATG.TG...A.......C...TTT....
............TTGTCT...CAG.GT.T...ATG.TG...A.......C...TTT....
............------...GA...A.T.............................---....
............------...GA...A.T.............................---....
............------...GA...A.T.....C.......................---....
............------...GA...A.T.......nnnnnnnnnnnnnnnnnnnnnnnn
............------...GA...A.T.............................---....
............------...GA...A.T.............................---....
............------...--...A.T.............................---....
............------...--.....T.......T.....................---....
............------...--.....T.......T.....................A---....
............------...--.....T.......T.....................A---....
............------...--.....T.......T.....................A---....
............------...--.....T.......T.nnnnnnnnnnnnnnnnnnnnnn
............------...--.....T.......T.....................A---....
............------...--.....T.......T.....................A---....
............------...--.....T.......T.nnnnnnnnnnnnnnnnnnnnnn
GGTATTTATGTTTTTTTTCGAGTGTCTGATCTTGTCTCTCTCCCCCCCG....----...
GGTATTTATGTTTTnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
```

```
TTTTCAACGTGTTTTGTTCAATCCTTGCAGGTGCTGTTCCAGCCCTCGTTCATCGGAATG
............................................................
...C.....................TCA.............................
...C.....................TCA.............................
......T.....................................................
......T.....................................................
......T.....................................................
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
......T.....................................................
......T.....................................................nnnnnnnnn
......T.....................................................
............................................................
............................................................
............................................................
.c..........................................................
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
............................................................
............................................................
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
GC.G.......-...CGA......--........T..........................
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
```

```
GAGTCCTCGGGCATCCACGACTGCACGTTCAAGACGATCATGAAGTGCGACGTCGACATC
.............................................................
.............................................................
.............................................................
.........................................................nnnnnnn
.................................G...........................nnn
.............................................................
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
..........T..................................................
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
.............................................................
.............................................................G..
.............................................................
.............................................................
.............................................................
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
........................A....................................
.............................................................
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
.............................................................
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
```

```
C  5a2
.  5b2/5c2
.  40.1
.  40.5
n  2.1
n  2.2
.  2.3
n  2.4
.  4.1
n  4.2
.  36.5
.  6
.  16.2
.  16.5/16.10
.  28.4
n  28.5
.  28.8
.  28.9
n  28.10
.  NLCI04
n  NLVR/NLJB
```

Appendix II

Sequence alignment of *Nereocystis* individuals at the second actin intron. Sequences are from AM06SR (*Alaria*), two clones of NLCI04, (*Nereocystis*), one each of NL01VR and NL01JB (both *Nereocystis*), as well as *Costaria costata* from Bhattacharya et al (1991). See Figure 11 for locations of each individual. A 'dot' (.) indicates identity with the *Alaria* sequence shown, a 'dash' (-) indicates an insertion or deletion added to allow alignment, and a "star" (*) indicates intron regions of *Costaria costata* based on Bhattacharya et al. (1991).

```
Alaria (AM06SR)        GAGAAGTCGTACGAGCTCCCCGACGGAAACGTCATCGT
Costaria costata       .....................................
Nereocystis (NLCI04)   .....................................
Nereocystis (NL01VR)
Nereocystis (NL01JB)

CATCGGAAACGAGCGCTTCCGTTGCCCCGAGGTATGATACAAAATG------------TA
T.............................C.T************************
.......G.........................TAAATATGATCGTTAATCATTACC


CTGTGACAAAGCGCTCTTGTGAATGGTGTAA-----GTAACAGATAACACTGTCCGAAAA
************************************************************
AGGGTATTAATAGTACTTATTACGCGGGAAGATTTTGGCGTCTGACCTTGATAATAATAA


TGTAAGCACCGTTCTGTAC-----------TGTGAGTATACAGGCGGTGCTAATTTTTCC
************************************************************
TAATAATAATAATGATAATGATACATACCAGGTACCTGGTACCGATGTTGGCCCAACATC


GGACAGAAGTACACGGCCCAGAGCACTCTTGTTA-CTTACCTGGTAGTCTACAGTAGTCT
************************************************************
GGCCCCAGATTTGCTTTTTTCGGGCTTTTCCGAAAGTTTGTAACTTTTTTTTTCACGTGG


TATTTTAGAGAAGTACAATTGTACTAGTAGTCTTATTTTAGAGAAGTACAA-TTGGTATA
************************************************************
CCCGTAGGAAGCGTGTTCTAAAAAAATTGTGGCGGCTTTGACTTTCAAACTTTTCATTTA


TACCTGGTATA-----------------------------------------------
************************************************************
AAACTTCGctgtcccattggcatagtctcatggcaatatttggaaaagtgaagaatccgg


------------------AAAGTGGCGATGAAATAAATAGCCTGCTTCCGC--------
************************************************************
g
```

```
--------------------------TAGCGTA----------CTTACCTGGATACCTGG
**************************************************************



TACCAGGTAGCGAGCCGTATGGTACCGGTGACC-GCACGTTAGCATTAACGGATTTTTGG
**************************************************************



GGGGTTTTCGGGCATGCGCTGTTTCATTTTGCTCCTCGTACTCTCTGGAGACTCCGGAGA
**************************************************************



CTGGAGACCTCTCCGTCTAC--------------TTTGGAGGTGGTACTAATTTCTCTC-
**************************************************************



TTTTGGTCTGGTACGAGCGTTTTTGCGTTTCCTCACGGCAGTTGCACCAGCTGCTAAAAT
**************************************************************



CTGGCGGGT-CTGTGTTCCGGGCGGGGGGTCTTGACTTCT--------------------
**************************************************************



-------------GTGTGTGTGCCCTGACCCATTCGAAGATGGACGGCGGCGGTGGTTGT
**************************************************************
```

```
TTTGTTGTAATAG---ACGATTTTATAAACTCG-TTTGGGCTGATAGGTGTTACTTGCGC
************************************************************


GCCGGCGATTTTTTTTGTT--GAGTACCTGGTATTCAAGAGTCCCCCCCCC---------
************************************************************


-----TTCAACCTGCACACCCCCGGGTTTTGAGTTATAGCCAACCACGCCCGAAGGGTAA
************************************************************
  gcagctgctgttcgggtacccattgatacggtggcagtacgttgtggatggttggaat


ACTTGTATTGCAACAGCAGTATGAACTGTTTGAGGATTCCCGGTGTCTCGACAACAGCAC
************************************************************
tctccgtgcGTTCCGATAAGCTTTATCGCATGGTGCACGATCGCCGAATTAGCCTGGTTC


-AAAGTTAATTTTGTGTCTC--TTAACAGTTATTGCACATTACAAATATCATGACATGT-
************************************************************
CAGGGCACCTAATGAGTTATTGAGTTGAATGCATGACGCCGGTGTAAAAAGGATGGAAGC


TTTCAATGATAAGGTTCCCAAGTTGTTAGCCTCGG---------CCAGGTATATGGATGT
************************************************************
ATATAAATGCCTATCAGAGACCTGGCATCTGTATTTTGTTAAGTGTCTGATGTTGTTTCT
                  ..........................................
    wwwwtatcarara..........................................


TGGGATCTTTTGGCCCGCTTCACTTTTGTGATAACGTTCGTTGGACGTACGCGTATATTC
************************************************************
CTCGCCCGTTTCTTTTTCACGACGTGATTCGACAATCCTCAGGAGCCTATTTATATGCCA
............................................................
............................................................
```

```
ACC--TCGTCTTCAGTACGT------CTC--TTTGCCAATTGTGCCTCCGTTCGTCCTCT
**************************************************************
GGAGACCTGGTATTTATGTTTTTTTTCGAGTGTCTGATCTTGTCTCTCTCCCCCC.G...
....... ..............
.......................

C---TTTTTTTTCAACGTGTTTTGTTCAATCCTTGCAGGTGCTGTTCCAGCCCTCGTTCA
*******************************************.............
.----...GC.G.......-...CGA......--.......T.................



TCGGAATGGAGTCCTCGGGCATCCACGACTGCACGTTCAAGACGATCATGAAGTGCGACG
............................................................
............................................................
```

```
TCGACATCC  Alaria (AM06SR)
.........  Costaria costata
.........  Nereocystis (NLCI04)
           Nereocystis (NL01VR)
           Nereocystis (NL01JB)
```

Appendix III

ANSI C++ source code created for the console program Tajima-D.exe. The source will compile for Win32 as well as various UNIX flavors. No graphical interface is implied, so the executable will run in command prompt mode. No dynamic memory is allocated, and pointers are not employed such that the program could be compiled in Visual Basic or other language with only semantic modifications.

```
//  tajima.cpp
//
// A Test for Neutrality
//
//      by M. Liptack (mliptack@sfu.ca)
//      copyright 1997-1999
//      version 0.7
//
//
// A program for finding Tajima-D statistics:
//     Tajima, F.  1989.  Statistical method for testing
//      the neutral mutation hypothesis by DNA
//      polymorphism.  Genetics 123:585-595.
//
//  Also tests whether they are statistically
//    significant.  It requires aligned data without
//    gaps (or at least gaps reduced to a single "-"
//    character).  Input must be "G" or "A" or "T"
//    or "C"). Data other than DNA sequences can be
//    used if converted to 'G','A','T','C','.', or '-'.
//    Note that indels > 1 basepair must be either
//    collapsed to 1 or removed.
//
//    Max input = 16 sequences (memory dependent),
//      2000 bases long (this could be changed
//      by adjust the code in only a few places below)


//      NOTE:
//      Variables from Tajima are similar in that
//      they phonetically sound the same ( ie S is bigS, etc.)
//



#include <iostream.h>
#include <iomanip.h>
#include <math.h>
#include <stdlib.h>
#include <conio.h>



int i, j, k, m, p, q, r = (0);  //for loop integers only, declared here in case this
                                  //is eventually compiled in VB

int nucs, enn, bigS, kill, paircomp, nucdiffs, sequencenumberA, sequencenumberB;
               // math, logic, and basic stats of sequences

int dontuse, bigSflag; // flags.....

char working[16];//used to hold bases at one aligned position
//array decl. may need to be smaller!!
char sequence[16][2000] = {0};//array decl. may need to be smaller!!

char firstallele; //used in logic loop to compare bases at one site

char end; //detect end designation by user

float kayhat, hi, xji, dee, beel, bee2, cee2, ceel, eel, ee2, variancehatd, ayone,
       aytwo, confidencelim90, confidencelim95, confidencelim99, sqrtvard, TajimaD;
                 // math variables for computing the statistic


 int main()
{
```

```
//
// calc. total number of sequences and nucleotides, allows error check too!!


        cout << endl << endl << endl;
        cout << "*****************************************" << endl;
        cout << "*********   Tajima-D statistics     *********" << endl;
        cout << "*****************************************" << endl << endl;

        cout << "NOTE: input limits are 3-16 sequences,";
        cout << " 1-2000 nucleotides long.." << endl << endl << endl;

        try  //error handler
        {


// inputting sequence parameters
        nucs = 0;
        enn = 0;
        do
                {
                        cout << "How many sequences?" << endl;
                        cin >> enn;
                        if (enn < 3)
                                {
                                        cout << "you need at least 3 seqs...try again:";
                                        cin >> enn;
                                }
                        if (enn > 16)
                                {
                                        cout << "too many...  retry:";
                                        cin >> enn;
                                }
                        i += 1;
                        if (i == 10) throw "number of sequences is out of range";
                } while((enn > 16) && (enn < 3));

        i = 0;

        do
                {
                        cout << "How many aligned nucleotides in longest seq?";
                        cin >> nucs;
                        if (nucs < 2)
                                {
                                        cout << "you need at least 2 nucs...try again:";
                                        cin >> nucs;
                                        }
                        if (nucs > 2000)
                                  {
                                        cout << "too many...  retry:";
                                        cin >> nucs;
                                  }
                        cout << endl << "set to: " << nucs << endl;
                        i += 1;
                        if (i == 10) throw "number of nucleotides is out of range";
                } while ((nucs < 1) && (nucs > 2001));

        i = 0;

        cout << endl << endl << endl;
```

```
//
// data input, 2D array, output to screen to check the data
//     allows editing if incorrect
//when reach bottom of column: row enn (bottom) = zeros, end of each row = /0
//


        sequencenumberA = 1;

        while (sequencenumberA > 0)
               {
                        cout << "Enter the sequence # to input or edit (enter 0
                        when you've entered them all): ";
                cin >> sequencenumberA;

                cout << endl;
                kill = 0;

                if (sequencenumberA = 0) continue;

                while (sequencenumberA > enn)
                {
                        cout << "That's bigger than the number of ";
                        cout << "sequences you originaly specified...  ";
                        cout << " Try again:";
                        cin >> sequencenumberA;
                        kill += 1;
                if (kill > 9) throw "the number of the sequence requested is
                more than the total number available ";
                }
                sequencenumberB = sequencenumberA - 1;
                cout << "begin sequence " << sequencenumberA << " here:";

                //input data
                cin.getline(sequence[sequencenumberA], 2000, '\n');



                cout << endl<< "Here's everything so far...." << endl;
                for (i = 0; i <= sequencenumberB; i++)
                {
                        cout << "Sequence " << i + 1 << " :";
                        for (j = 0; j < nucs; j++)
                        {
                                cout << sequence[i][j];
                        }
                        cout << endl;
                }
                cout << endl << endl << endl;
        }


//computing the test statistic...Part I- the easy stuff, finding ee1 and ee2..
        ayone = 0; aytwo = 0;  //prevents divide by zero
        if (enn <= 1)
                throw "Cannot divide by zero";
        bee1 = (enn+1)/(3*(enn-1));
        bee2 = (2*(enn*enn+enn+3))/(9*enn*(enn-1));
        for (i = 1; i < enn; i++)
                {                       // calculates the sum terms a1 and a2
                ayone += 1/i;
                aytwo += 1/(i*i);
                }
        cee1 = bee1 - (1/ayone);
        cee2 = (bee2) - ((enn+2)/ayone*enn) + (aytwo/(ayone*ayone));
        ee1 = (cee1)/(ayone);
        ee2 = (cee2)/((ayone*ayone)+aytwo);
```

```
    }  //end of try

       // divided by zero and out-of-range error handler

       catch (const char errorMessage[])
       {
              cout << "ERROR: " << errorMessage << endl;
              return 1;

       }



// add term to check if the position has a zero, because if it does we can't use it


//computing the test statistic...Part II-the more difficult ind.terms & S
       bigS=0; bigSflag = 0, nucdiffs = 0;
       for (i = 0; i < 16; i++)  //fills working array with ^^^^^
              working[i] = '^';
       for (j = 0; j < nucs; j++) //shifts "window" to view a position at a time
              {
              for (k = 0; k < enn; k++) //input to working and change '.'
                     {
                     working[k] = sequence[k][j];
                     if (working[k] == '.')
                            working[k] = working[0];
                     }
              for (m = 1; m < enn; m++) //is it a segregating site?
                     {
                     if (working[m] != working[0])
                            bigSflag++;
                     }
              if (bigSflag > 0)
                     bigS++;
              bigSflag = 0;

              //calculate total # of diffs. aka nucdiffs
              for (p = 0; p < enn-1; p++)
                     {
                     for (q = p + 1; q < enn; q++)
                            {
                            if (working[q] == working[p])
                                   nucdiffs++;
                            }
                     }
              for (r = 0; r < 16; r++)
                     {//make sure working is empty
                     working[r] = '^';}
                     }


//calculating khat which is # of diffs/#pairwise compairisons
       paircomp =((enn * enn) - enn)/2;
       kayhat = nucdiffs / paircomp;



//more calculations that actually give D
       variancehatd = (ee1*bigS) + (ee2*bigS*(bigS-1));
       sqrtvard = sqrt(variancehatd);
       dee = kayhat - (bigS/ayone);

       TajimaD = dee/sqrtvard;
```

```
//add 95 and 99% limits (90% is below), then put at top.....


        float teststat90Tajima[18] = {0,0,0,-0.876, -1.255, -1.405,
                                -1.498, -1.522, -1.553, -1.559,
                                -1.572, -1.573, -1.580, -1.580,
                                -1.584, -1.583, 0, 0};

        confidencelim90 = teststat90Tajima[enn];
        cout << "Tajima-D is " << TajimaD << endl;
        cout << " 90% Confidence limit is: " << confidencelim90 << endl;
        cout << " 95% Confidence limit is: " << confidencelim95 << endl;
        cout << " 99% Confidence limit is: " << confidencelim99 << endl;
        cout << " ...If Tajima-D is less than above, the null hypothesis ";
        cout << "that the sequence is neutral, could not be disproved...";
        cout << endl << endl;

        end = 'x';
        do          //wait for user to read data, maybe add
                    //output to program named "TajimaD.txt"
            {
               cout << "hit q to end ";
               cin >> end;
               } while(end != 'q');


        return 0;
}
```

Appendix IV

Alignment of *Alaria* actin intron II regions. Sequences were based on consensus sequences for each individual in Appendix I. Protein coding regions, lariat consensus sequences, regions from individuals with only two clones that varied, and 5 bases identical at the 5' end of the intron to the homologous *Nereocystis* sequence were removed. A 'dot' (.) indicates identity with the AM06SR sequence shown, a 'dash' (-) indicates an insertion or deletion added to allow alignment, while lower case letters designate sites for which data from only one clone was determined. Numbered regions correspond to sites used for visual determination of possible recombination in Figure 21.

```
          1          _____2_____      34              5
AM06SR  ATACAAATG-----------ACTGTGACAAAGCGCTCTTGTGATGGTGTAA
AM03KB  .........-----------....GT.....................A..T
AM03RB  .........ATAATAAAAATG..............................
AM01WB  .........ATAATAAAAATG..............................
AT01OP  .........ataataaaaatg..............................
AM03SK  ..t......-----------....gt.............t...........
AM03JR  ..T......-----------....GT.............T...........
AN10BO  ..T......-----------....GT.............T...........


                                          6   7  8
-----GTAACAGATAACACTGTCCGAAAATGTAAGCACCGTTCTGTAC-----------T
ACCTG.....................................T.....ATTTAATAC---
-----.....................................T.....---------TG.
-----............T........................T.....---------TG.
-----.....................................t.....---------tg.
-----.....................................---------tg.
-----.............................T...A..T.....---------TG.
-----.............................T...A..T.....---------TG.


                                                        910
GTGAGTATACAGGCGGTGCTAATTTTCCGGACAGAAGTACACGGCCCAGAGCACTCTTGT
............A.......T............TG...T.....A............A.
.....................................................T.
.....................................................T.
.....................................................t.
.....................................................ct....
.....................................................CT....
.....................................................CT....


TA-CTTACCTGGTAGTCTACAGTAGTCTTATTTTAGAGAAGTACAATTGTACTAGTAGTC
..---.TAT.AT..........-------------------------------CTACA
..CTAC.GTA.T..........C-------------------------------.......
..CTAC.GTA.T..........C-------------------------------.......
..ctac.gta.t..........c-------------------------------.......
..--...gt....c----------------------------------------.....
..--...GT....C----------------------------------------.....
..--...GT....C----------------------------------------.....


  14 15             16  17
TTATTTTAGAGAAGTACAA-TTGGTATATACTGGTATA---------------------
G..GCAGTA.A.GTAG...-.........GGC......TGAATACCGTATTGTTATTACC
...................-.................----------------------
...................-.................----------------------
.......---------...-.................----------------------
..g..a.......c.....a...-.............----------------------
..G..AA......C.....A...-.............----------------------
..G..AA......C.....A...-.............----------------------
```

```
----------------------------------------------AAAGTGGCGATGAA
AGCATAATTGATGTGTAAAATATATAGATGAACTGTATTCAAGATA..G....G......
----------------------------------------------..............
----------------------------------------------..............
----------------------------------------------..............
----------------------------------------------..............
----------------------------------------------..............
----------------------------------------------..............
```

```
ATAAATAGCCTGCTTCCGC-----------------------------------TAGCGTA
.G.T..............GCCTTTTTTTCGCGCCCG-TGGGG--TCTCTGTC.......
.G................GCCTTTTTT-CGCGCCCG-TGGGG--TCTCTGTC.......
.G................GCCTTTTTT-CGCGCCCG-TGGGG--TCTCTGTC.......
.g................ccctttttt-cgcgcccgttggg---tctccgtc.......
.g................acctttttt-cgcgcccg-tgggggtctccgtc.......
.G................ACCTTTTTC-CGCGCCCG-TGGGGGGTCTCTGTC.......
.G................ACCTTTTTT-CGCGCCCG-TGGGGGGTCTCTGTC.......
```

```
           18                                  19             20     21
----------CACCTGGATACCTGGTCCGGTAGCGACCTATGGTACCGGTGAC-GCACGT
taaccgtata.....................................a......-......
--------TA.C.........................G......A......C......
--------TA.C.........................G......A......C......
--------ta.c.........................g......a......C......
--------ta.....................................a......c......
----------.....................................A......-.....G
----------.....................................A......-.....G
```

```
                                    22                   23
AGCATTAACGGATTTTTGGGGGGTTTCGGGCTGCGCTGTTTCATTTTGCTCCTTCTCTCT
.....................................................................
.....................................................................
.....................................................................
.....................................................................
.....................................................c.....c......
.................................G...................C......
.................................G...................C......
```

```
        24
GGAGACTCCGGAGACTGGAGACCTCTCC-GTCTACTT------------GGTACTAATTC
............................-........-------atctg...........
........-...................C........------------............
........-...................-........------------............
........-...................-........------------............
............................-........-------atctg...........
............................-........ACCTGGTATCTG...........
............................-........ACCTGGTATCTG...........
```

```
    25          26                27  28
TCTCTTTTGGTCTGGTACGAGCGTTTTGCGTTCCTCACGGCAGTTGCACCAGCTGCTAAA
...........................................................
....-......................................................
....-......................................................
....-......................................................
....-......................................................
.................C.................T...G...................
.................C.................T...G...................
```

```
                        29
ATTGCGGGT-TGTGTTCCGGGCGGGGGGTCTTGACTTCT---------------------
.........c.............................c..gtgtcttgt------------
.........C.............................GTGTCTTGTGTGTGTGTGC
.........C....-........................GTGTCTTGT----GTGTGTGC
.........c.............................gtgtcttgtgtgtgtgtgc
.........c.............................gtgtcttgt------------
.........C..........T..................GTGTCTTGT------------
.........C..........T..................GTGTCTTGT------------
```

```
                                                          30
--------GTGTGTGTGCCCTGACCGAAGATGGACGGCGGCGGTGTTTTTGTGTAATACA
--------.....................................................c....
GTGTGTGT.....................................................T....
GTGTGTGT.....................................................T....
gtgtgtgt.....................................................t....
--------.....................................................
--------.....................................................G......
--------.....................................................G......
```

```
3132  33                                                    34  35
TTTTATAAACTCGTTTGGCTGATGGTGTACTTGCGCGCCGGCGATTTTTTTTGTT-GGTA
..c.....................................................g...t....
.....................................................G.T..-....
.....................................................-....
.....................................................t....
.cc...c..............................................g...-....
.CC...C..............................................G...-....
.CC...C..............................................G...-....
```

```
                 363738                                          39
CCTGGTATCAAGAGTCCCCCCTTCAACTGCACACCCCCGGTTTTAAACACGAAGGGTAAC
..................g.....................................
.....................................................C.
.....................................................C.
.....................................................c.
..................g.....................................
................T..G.....................................
................T..G.....................................
```

```
TTGTATTAACAGCAGTATGAACTGTTTGAGGATTCCCGGTGTCTCGCAACGCAC-AAAGT
.........................T.........C..........G.AA..G..T..
.................-------------------------.....A...-.....
.................-------------------------.....A...-.....
.....................................................a...-.....
.....................................................a...-.....
.....................................................A...-.....
.....................................................A...-.....
```

```
             40  41
AATTTTGTGTCTC--TTAACAGTTATTGCACATTACAAATTCATGACATGT-TTTCAATG
.......C.....CTC.G....CA..---------------.....GA..A..C...G.
.............--...................G.............-........
.............--...................G.............-........
.............--..g................................-........
.............ct........g..........g.....c..........-........
.............CT........G..........G.....C..........-........
.............CT........G..........G.....C..........-........
```

```
  42
ATAAGGTTCCCAAGTGTTAGCCTCGG---------CCAGGTATATGGATGTTGGGATCTT
...T...AT......--.......--AACCAGGTA............T......A..A..
...T.....................---------.............................
...T.....................---------.............................
...t.....................---------.............................
...t.....................ctaggtata.............................
..GT.....................---CTAGGTATA...........................
..GT.....................---CTAGGTATA...........................


 43                                          44
TTGGCCCGCTTCCTTTGTGATAACGTTCGTTGGACGTACCGTATATTCACC--TCGTCTT
...........A.........GC....T..................G..TGAG.T.....
.....................C.............................--.......
.....................C.............................--.......
.....................c.............................--.......
.a...................g.............g.............--.......
.A.....................................G.............--.......
.A.....................................G.............--.......

                    45
CAGTACGT------CTC--  AM06SR
........TTGTCT...CA  AM03KB
........------...GA  AM03RB
........------...GA  AM01WB
........------...--  AT01OP
........------...--  AM03SK
........------...--  AM03JR
........------...--  AN10BO
```