

**POPULATION DIVERSITY AND MOLECULAR EVOLUTION  
OF SELECTED EASTERN PACIFIC SEA CUCUMBERS  
(CLASS: HOLOMURROIDEA) BASED ON MITOCHONDRIAL DNA**

by

Allan David Arndt

B.Sc., University of Calgary, 1986; M.Sc., University of Calgary, 1990

**THESIS SUBMITTED IN PARTIAL FULFILLMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY**

in

the Department of Biological Sciences

© Allan D. Arndt 1996

**SIMON FRASER UNIVERSITY**

October, 1996

All rights reserved. This work may not be reproduced,  
in whole or in part, by photocopy or other means,  
without permission of the author.



National Library  
of Canada

Bibliothèque nationale  
du Canada

Acquisitions and  
Bibliographic Services Branch

Direction des acquisitions et  
des services bibliographiques

395 Wellington Street  
Ottawa, Ontario  
K1A 0N4

395, rue Wellington  
Ottawa (Ontario)  
K1A 0N4

Your file / votre référence

Our file / notre référence

**The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.**

**L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.**

**The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.**

**L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.**

ISBN 0-612-16777-1

APPROVAL

NAME: Allan David Arndt

DEGREE: DOCTOR OF PHILOSOPHY

TITLE OF THESIS:

Population Diversity and Molecular Evolution of Selected Eastern Pacific Sea Cucumbers  
(Class: Holothuroidea) Based on Mitochondrial DNA.

Examining Committee:

Chair: Dr. Z. K. Punja, Associate Professor

---

Dr. M. J. Smith, Professor, Senior Supervisor  
Department of Biological Sciences, SFU

---

Dr. A. Beckenbach, Professor  
Department of Biological Sciences, SFU

---

Dr. L. Druehl, Professor  
Department of Biological Sciences, SFU

---

Dr. F. Breen, Associate Professor  
Department of Biological Sciences, SFU  
Public Examiner

---

Dr. A. Meyer, Assoc. Professor  
Department of Ecology and Evolution  
University of New York  
External Examiner

Date Approved Oct. 26, 1996

**PARTIAL COPYRIGHT LICENSE**

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

**Title of Thesis/Project/Extended Essay**

Population Diversity and Molecular Evolution of Selected Eastern

Pacific Sea Cucumbers (Class: Holothuroidea) Based on

Mitochondrial DNA

**Author:** \_\_\_\_\_

(signature)

Allen Acast

(name)

30 Oct 1996

(date)

## Abstract

Sea cucumbers (Class: Holothuroidea) are one of six extant classes of the exclusively marine phylum Echinodermata. Unfortunately, these organisms are poorly represented in the fossil record and consequently, the taxonomy and systematics of this class remains problematic. As a preliminary investigation, a molecular phylogeny was constructed focusing on members of the family Cucumariidae from the eastern Pacific. Fourteen hundred nucleotides spanning portions of two mitochondrial genes, encoding the large ribosomal RNA subunit and cytochrome oxidase subunit one, were sequenced from 15 species, representing six families from two orders. A number of important phylogenetic issues were resolved, particularly the relationships among brooding species.

Following this phylogenetic assessment, an analysis of population structure in two species of sea cucumber as a function of their mode of development was undertaken, based on four hundred nucleotides of mitochondrial sequence with presumed homology to the replication control region. *Cucumaria pseudocurata* exhibits a brooding form of development while *Cucumaria miniata* passes through a pelagic nonfeeding larval stage lasting up to 14 days. These two species occupy extensive, overlapping ranges along the west coast of North America and populations were sampled between Pescadero Point, California and Juneau, Alaska. Significant geographic structuring of populations was evident in *C. pseudocurata* indicating limited dispersal. Despite the general trend of strong geographic structuring of populations, there is evidence of significant gene flow between adjacent sites separated by as little as four kilometres of ocean or as much as one hundred kilometres along the coast. The most significant genetic disjunction observed

separated Alaskan and Queen Charlotte Island (Haida Gwaii) samples from Vancouver Island and all sample sites further south. The northern population samples were found to contain equal or greater genetic variability compared to many southern sites. The most plausible explanation of this genetic disjunction is therefore survival in northern glacial refugia in addition to the persistence of populations south of the furthest extent of the Cordilleran ice sheet. Presumably, these surviving populations expanded to their present apparently continuous distribution of these externally indistinguishable types following the retreat of the ice sheet. The degree of sequence divergence suggests that this separation existed well prior to the last glaciation and may have been maintained for several million years.

Conversely *C. miniata* population samples exhibited a panmictic structure over their entire range indicating high levels of gene flow despite a relatively short dispersal phase. These results suggest that mode of development and life history traits have had a profound influence on gene flow.

Finally, mtDNA sequence from *C. miniata* and *C. pseudocurata* has revealed a novel tRNA gene arrangement compared to other echinoderms. Analysis of mtDNA indicates this novel arrangement has arisen by a nontandem duplication of an ancestral tRNA gene cluster including putative control elements. Gene sequence analysis indicated potential utility of the tRNAs as a phylogenetic marker at deeper levels than other mitochondrial genes which, in conjunction with tRNA gene order rearrangements, make this a region of promising utility in studies of molecular evolution.

**Dedication**

*Blessed are the poor in spirit, Blessed are the meek  
For theirs shall be the Kingdom that the power mongers seek  
Blessed are the dead for love, and those who cry for Peace  
And those who love the gift of Earth...  
May their gene pool increase*

Bruce Cockburn, 1990

Respectfully dedicated to the memory of John D.G. Boom,  
whose generosity, talent and wit are deeply missed.

### **Acknowledgments**

I would like to thank the members of my committee for their support and critical reviews of this work. In particular, I would like to thank the my senior supervisor, Michael J. Smith, for his enthusiastic support and friendship, without which this thesis would have been considerably less enjoyable and perhaps not possible.

I would like to thank the following people for assistance with field collections:

Robert Arndt, Veronica Franklin, Cameron Osborne, Dr. John Pearse, University of California, Santa Cruz, and Dr. Tom Shirley, University of Southeast Alaska, Auke Bay. In addition, I would like to give special thanks to Phil Lambert, invertebrate curator of the Royal B.C. Museum, Victoria, for numerous discussions, collaboration, and for introducing me to the wonderful world of sea cucumbers on the rocks of Sooke Regional Park.

I would also like to thank present and past members of the SFU Biosciences department for providing an excellent atmosphere for research and for helping a biochemist learn echinoderm biology. In particular, Karen Beckenbach, John Boom, Dr. J. R. Brown, Dr. D. ÓFoighil, Dr. M. Marra, Sheldon McKay, Cam Muir, Andrea Scouras, and Theresa Burg provided invaluable discussions concerning this work.

Finally, I wish to convey profound gratitude and thanks to Nancy Armitage for her encouragement and support, her artistic talents, and her keen eye and sharp pencil.

Financial support was provided by SFU graduate fellowships, an SFU President's PhD Stipend, SFU teaching assistantships and an NSERC grant to Dr. M.J. Smith.



**Table of Contents**

<b>Approval</b>	ii
<b>Abstract</b>	iii
<b>Dedication</b>	v
<b>Acknowledgments</b>	vi
<b>Table of Contents</b>	vii
<b>List of Tables</b>	ix
<b>List of Figures</b>	xi
<b>General Introduction</b>	I

**Chapter I. Phylogenetic Analysis of Selected Eastern Pacific Sea Cucumbers**

**Focusing on the Cucumariidae**

<b>Introduction</b>	4
<b>Materials and Methods</b>	9
<b>Results</b>	22
<b>Discussion</b>	47

**Chapter II. Genetic Diversity and Population Structure of Sea Cucumber Species  
with Differing Modes of Development**

<b>Introduction</b>	<b>53</b>
<b>Materials and Methods</b>	<b>61</b>
<b>Results</b>	<b>71</b>
<b>Discussion</b>	<b>119</b>

**Chapter III. Sea Cucumber Mitochondrial tRNAs: Gene Order, Secondary  
Structures and Patterns of Substitution**

<b>Introduction</b>	<b>136</b>
<b>Materials and Methods</b>	<b>139</b>
<b>Results</b>	<b>142</b>
<b>Discussion</b>	<b>181</b>
<b>Future Directions</b>	<b>194</b>
<b>References</b>	<b>197</b>
<b>Appendix A. Distance Matrix Among <i>C. pseudocurata</i> Haplotypes</b>	<b>220</b>
<b>Appendix B. Distance Matrix Among <i>C. miniata</i> Haplotypes</b>	<b>230</b>
<b>Appendix C. tRNA Structures</b>	<b>238</b>

**List of Tables**

Table 1.	Classification and location of species in study	10
Table 2.	Oligonucleotide primers	14
Table 3.	Nucleotide comparison of sea cucumber COI gene fragments	27
Table 4.	Overall codon usage for the COI gene fragments	30
Table 5.	Tamura-Nei distance matrix for the COI gene fragments	34
Table 6.	Transition/transversion ratios of pairwise species comparisons of the COI gene fragments	35
Table 7.	Nucleotide base composition of the aligned portions of the sea cucumber lrRNA gene fragments	37
Table 8.	PCR amplification primers	66
Table 9.	Frequency and distribution of <i>C. pseudocurata</i> mtDNA haplotypes	82
Table 10.	Frequency and distribution of shared <i>C. miniata</i> mtDNA haplotypes	100
Table 11A.	Mean intra-deme distance for <i>C. pseudocurata</i> samples	107
Table 11B.	Mean inter-deme nucleotide distances among <i>C. pseudocurata</i> samples	107
Table 12.	Sea cucumber population structure based on analysis of variance using the computer program, AMOVA	109

Table 13.	Pairwise $\Phi_{ST}$ estimates among <i>C. pseudocurata</i> population samples	111
Table 14A.	Mean intra-deme distance for <i>C. miniata</i> samples	113
Table 14B.	Mean inter-deme nucleotide distances for <i>C. miniata</i> samples	113
Table 15.	Nucleotide identity among sea cucumber tRNA genes	175
Table 16.	Substitutions in tRNA genes between <i>C. miniata</i> and <i>C. pseudocurata</i>	176
Table 17A.	Substitutions in tRNA gene stems between <i>C. miniata</i> and <i>P.californicus</i>	177
Table 17B.	Substitutions in tRNA gene loops between <i>C. miniata</i> and <i>P.californicus</i>	178
Table 18.	Maintenance of base pairing in tRNA stems	180

## List of Figures

Figure 1.	Aligned nucleotide sequence of the 3'-terminal 1rRNA gene fragment from 14 species of sea cucumber	17
Figure 2.	Inferred amino acid sequence for the sea cucumber COI gene fragments	24
Figure 3.	Alignment of the variable nucleotide sites of the COI gene fragment from ten Cucumariidae species.	31
Figure 4.	Phylogenetic trees for the family Cucumariidae based on the nucleotide sequence of the amplified COI gene fragments.	40
Figure 5.	Comparison of calcareous ossicle morphology of brooding sea cucumber species.	42
Figure 6.	Maximum likelihood phylogeny based on the nucleotide sequence of the sea cucumber 3'-terminal 1rRNA gene fragment	45
Figure 7.	Population sample site map.	62
Figure 8.	Alignment of mtDNA region used for population study	74
Figure 9.	Variability of sites in region sequenced for population study	77
Figure 10.	Variable nucleotides detected among <i>C. pseudocurata</i> haplotypes	79
Figure 11.	Neighbor-joining tree for <i>C. pseudocurata</i> haplotypes	86
Figure 12.	Minimum spanning network among northern <i>C. pseudocurata</i> haplotypes.	91

Figure 13.	Minimum spanning network among southern <i>C. pseudocurata</i> haplotypes.	93
Figure 14.	Detected variable nucleotides among <i>C. miniata</i> haplotypes.	97
Figure 15.	Neighbor-joining tree for <i>C. miniata</i> haplotypes.	101
Figure 16.	Minimum spanning network among <i>C. miniata</i> haplotypes.	104
Figure 17.	Comparison of ossicles from northern and southern populations of <i>C. pseudocurata</i> .	114
Figure 18.	Variable sites detected in the threonine tRNA gene	117
Figure 19.	Arrangement of clustered tRNA genes in sea cucumber mtDNA.	143
Figure 20.	<i>C. miniata</i> mtDNA sequence between the srRNA and ND1 genes.	146
Figure 21.	<i>C. pseudocurata</i> mtDNA sequence between the srRNA and methionine tRNA genes.	149
Figure 22.	PCR amplification products using pCmhd23.	152
Figure 23.	DNA sequence from the mitochondrial restriction fragment clone, pCmhd23.	154
Figure 24.	PCR amplification products of <i>C. miniata</i> mtDNA between the COI and ATP6 genes.	156
Figure 25.	<i>C. miniata</i> mtDNA sequence between the COI and ND4l genes.	158
Figure 26.	Arrangement of tRNA genes in <i>C. miniata</i> mtDNA.	162
Figure 27.	Alignment of <i>C. miniata</i> sequence corresponding to the two copies of the putative control region duplication	165

<b>Figure 28.</b>	<b>Possible stem loop structures for <i>C. miniata</i>.</b>	<b>168</b>
<b>Figure 29.</b>	<b>Putative stop element sequences.</b>	<b>171</b>
<b>Figure 30.</b>	<b>DNA distance matrix using tRNA gene sequence.</b>	<b>182</b>
<b>Figure 31.</b>	<b>Phylogenetic trees based on the tRNA data set.</b>	<b>184</b>

## General Introduction

### *Molecular Markers*

Over the past three decades, the introduction of molecular techniques has provided an enormous increase in the scope of opportunities in many areas of biological research. Through the use of such techniques, specific heritable traits can be identified and used in quantitative analyses at a level of sensitivity and throughput previously unattainable. One of the earliest applications of such an approach was based on immunological identification. Zuckerkandl and Pauling (1965) demonstrated that amino acid substitutions accumulated as a linear function of time over a wide range of taxa. This pattern of change is consistent with neutral mutation theory (Kimura, 1983) and indicates that at least some positions are not subject to selective constraints. This feature of selective neutrality in properly chosen molecular markers precludes problems associated with markers subject to sexual selection or phenotypic plasticity. One of the most important recent technical advances has been the development of polymerase chain reaction (Mullis et al., 1986; Saiki et al., 1988) (PCR) DNA amplification techniques. PCR allows DNA to be recovered in sufficient quantities to obtain specific gene sequence from extremely small amounts of tissue from living or even preserved organisms. Such direct analysis of the genetic material can provide a wealth of information for phylogenetic analysis as well as in assessing more recent events such as differentiation within populations.



Mitochondrial DNA (mtDNA) was chosen as a marker for several reasons. The mitochondrial genome of most organisms is a compact, circular DNA molecule containing a limited set of genes coding for 13 protein subunits, 22 tRNAs, and both the small (srRNA) and large (lrRNA) ribosomal RNAs. The mitochondrial genome is present at high copy number per cell in most tissues (Brown and Wilson, 1986; Wolstenholme, 1989). The attributes of high copy number and circular nature allow relatively straightforward mtDNA purification from very little material and has been successfully performed on many organisms. Consequently, complete genomic sequences from a diverse range of taxa are available (reviewed by Avise, 1994) including several echinoderms (Jacobs et al., 1989; Cantatore et al., 1989). This information can be utilized to identify highly conserved regions of the mtDNA genome through sequence comparisons of homologous regions. Kocher et al., (1989) first suggested that oligonucleotide primers designed to anneal to such highly conserved regions would have utility across a broad range of taxa and would allow amplification of DNA from many organisms. Such conserved primers offer a considerable advantage by eliminating the necessity of costly and time consuming cloning and sequencing of DNA in order to examine organisms not previously analyzed in this manner.

Among the 37 mitochondrial genes, there exists considerable heterogeneity in rates of accumulation of substitutions. A particular gene of interest can often be chosen which is thought to be evolving at an appropriate rate. Regions such as the replication control region and third positions of protein coding genes accumulate substitutions at higher rates, and are useful for the comparison of closely related species or even populations, while the

ribosomal genes, tRNA genes, and second positions of protein coding genes are appropriate for analyzing more ancient relationships. Furthermore, mtDNA accumulates mutations at a rate some 5-10 times higher than single copy nuclear genes in primates (Brown et al., 1979), however, approximately equal rates were observed in sea urchin (Vawter and Brown, 1986). This may be attributable to the fact that replication of mtDNA is dependent on DNA polymerase  $\gamma$ , an enzyme that apparently lacks a repair mechanism (Tomkinson and Linn, 1986) in most organisms. The mitochondrial genome is usually strictly maternally inherited with no evidence of intermolecular recombination, and can thus be treated as a haploid genome which provides a maternal genealogy. The rapid accumulation of mutations and uni-parental inheritance provide a sensitive marker for recent evolutionary events (reviewed in Moritz et al., 1987). It is therefore not surprising that mtDNA has become one of the most commonly used molecules for phylogenetic and population genetic analyses (Avise, 1994).

The investigations described here utilized mitochondrial DNA sequence analysis to address several problems. First, to provide an initial phylogenetic analysis of selected sea cucumber species and compare these results to more traditional morphological characters; second, to examine population structure or micro-evolutionary changes within two closely related species with differing modes of development; and finally, a third section examined observed tRNA gene rearrangements within two *Cucumaria* species and considers the implications concerning intra-molecular recombination and the use of mitochondrial gene order for more ancient phylogenetic analysis.

# Chapter I

## Phylogenetic Analysis of Selected Eastern Pacific Sea Cucumbers Focusing on the Cucumariidae

### Introduction

#### *Sea Cucumber Systematics*

The sea cucumbers, Class Holothuroidea, constitute one of the six extant classes of the exclusively marine Phylum Echinodermata. The approximately 1,200 species of sea cucumbers are partitioned into six orders. The earliest fossil evidence of the Holothuroidea, in the Upper Silurian, consists of calcareous deposits from the body wall, (Gilliland, 1992, Pawson, 1966). Unfortunately, the fossil record for the Holothuroidea is poor, providing little evidence in support of intraclass phylogeny. The first holothuroids were probably covered by calcareous plates, a condition maintained in only a few genera of the order Dendrochirotida. In most extant species, the calcareous accretions have been reduced to microscopic ossicles in the skin. Sea cucumber orders are distinguished on the basis of gross morphological characters such as the morphology of the tentacles, the presence or absence of tentacular retractor muscles, the presence or absence of an internal respiratory tree, and the presence and distribution of podia on the trunk. At the familial level, taxonomic distinctions are dependent on a finer level of characters such as the

morphology of dermal ossicles, the form of the esophageal calcareous ring, the distribution and morphology of the tube feet, etc. At the species level, the differences in a defining morphological character can be quite subtle, obscuring taxonomic distinctions.

An interesting feature of this group is the wide range of reproductive strategies that have been adopted. Included in this study are species with feeding or nonfeeding planktonic larval forms as well as several instances of benthic maternal brooding of embryos (see Table 1). In order to understand how such differences may have evolved, a firmly established phylogeny is required.

### *Molecular Phylogenetics*

In addition to the enormous amount of sequence data being generated, an equal effort has been put into improving analytical methods of handling such data. Three different approaches to phylogenetic analysis have been developed.

Maximum parsimony treats nucleotides as differing character states and reconstructs relationships based on the minimum number of character changes required to explain the observed data. The cornerstone of this concept is rooted in cladistics with the most important or informative characters demonstrating a shared derived state (reviewed in Hillis and Moritz, 1990, Avise, 1994). Such synapomorphic changes cluster taxa together and indicate a departure from the presumed ancestral state.

Distance based methods calculate the dissimilarity between sequences and usually report this value as the number of changes per nucleotide position. A model of the pattern of nucleotide substitution is assumed (see below) in order to calculate a corrected

distance matrix for all possible pairwise comparisons among the taxa examined. Various algorithms are then employed to combine these distances into an estimate of phylogenetic structure. One of the most commonly employed tree building algorithms based on such a distance matrix is the neighbor-joining method (Saitou and Nei, 1987). In this method, a pair of taxa are joined together so as to minimize the squared distances to all other taxa. Taxa are then added to this pair one at a time in a similar manner to define distances among all internal nodes and produce the final phylogenetic tree.

The third method is referred to as maximum likelihood (Felsenstein, 1981). In this method, a likelihood function is determined for how well the observed data fits a given set of relationships based on the nucleotide changes observed and seeks to maximize this likelihood value.

All the methods described above have certain advantages and drawbacks and the performance of these methods has been subjected to a multitude of tests based on theoretical and natural data sets and has been reviewed by numerous authors (Li and Grauer, 1991; Hillis and Moritz, 1990; Avise, 1994). Both parsimony and likelihood methods analyze a multitude of possible trees and choose the one which optimizes the defining criteria. In the case of maximum likelihood, alternative possibilities can be directly compared in order to test if the resultant tree is significantly more likely than others (Kashino and Hasegawa, 1989) based on the calculated log likelihood and confidence limits of a given tree. Parsimony determines if there is a unique most parsimonious tree or will indicate if more than one tree of equal length was found. This indicates uncertainty in one or more branch points and a consensus tree can be

constructed which reduces such uncertainties to polytomies. Distance-based methods are generally much faster computationally but produce only a single tree.

One commonly employed significance test for these methods is a bootstrap procedure (Felsenstein, 1985). A subsample of the data is selected by sampling nucleotide sites, with replacement, such that a particular site(s) may be selected more than once. A new tree is then constructed with this subsample of the data. This procedure is repeated numerous times (hundreds or thousands) and the frequency that nodes connecting particular taxa in the original tree occur in the resampled trees is recorded. This test is a measure of the consistency of the data across the entire region sequenced.

Among the three common approaches described, maximum likelihood has the fewest inherent assumptions concerning the model of substitution and appears to perform consistently among the best (Huelsenbeck and Hillis, 1990). Yet this method is computationally the most demanding and reconstruction with more than 15 taxa requires considerable computer time. One of the key assumptions in parsimony analysis is equal rates of variation among lineages. This condition is often violated in typical phylogenetic studies for a variety of reasons such as generation time, population structure or possibly even metabolic rate (Felsenstein, 1985; Palumbi, 1992). Under such conditions an incorrect phylogeny may be recovered. Assumptions are also made in distance methods where a model of nucleotide substitution is required to calculate the distance matrix. The simplest model assumes an equal probability of substitution among all four nucleotides (Jukes and Cantor, 1969). Kimura (1980) derived the two parameter model to calculate

transitions and transversions separately. This is an important concept since it appears that transitions tend to occur more frequently yet the proportion of transitions often decreases with time. This can be explained by the fact that a transversion will remove any trace of a previous transition at the same site while a transition at a site where a transversion has previously occurred will appear to be due to a single transversion (DeSalle et al., 1987). Thus over time, as the likelihood of multiple mutations at a single site becomes significant, transversions will mask a proportion of transitions and thus underestimate the actual number of mutations that has occurred. Tamura and Nei (1993) noticed that the pattern of substitutions in human and chimpanzee control region sequence data followed a gamma distribution and derived a method where all types of substitutions are considered separately. This method offers an advantage by being less sensitive than other distance methods to nucleotide composition. Basic parsimony analysis is also sensitive to base composition and recently another approach has been developed which calculates differences among taxa as the log of determinants of a matrix (Lockhart et al., 1994; Steel, 1993). Again this method, referred to as logDet, has been demonstrated to be insensitive to nucleotide composition bias among taxa and recovered the correct tree when other methods failed (Jermiin et al., 1994a; Jermiin et al., 1994b; Lockhart et al., 1994). Methods of phylogenetic inference based on molecular data are relatively new and improvements will no doubt continue to appear in the literature. For the time being, the most satisfactory approach to dealing with such uncertainties among phylogenetic reconstruction methods is to demonstrate agreement or congruence among various methods.

A phylogenetic analysis was undertaken utilizing partial DNA sequences of two mitochondrial genes; 870 nucleotides of the cytochrome oxidase subunit 1 (COI) and 350 nucleotides of the large ribosomal RNA subunit (1rRNA) were obtained from 16 species of sea cucumber from the eastern Pacific (Table 1). Included are most Pacific members of the genus, *Cucumaria*, with representative species from five other families in two orders, Dendrochirotida and Aspidochirotida. This molecular analysis primarily focuses on recent evolution within the Cucumariidae and examines a number of specific taxonomic and phylogenetic questions concerning these Pacific sea cucumbers. Members of other families were obtained opportunistically in order to a) provide a range of potential outgroups, b) examine the suitability of COI and 1rRNA as suprafamilial indicator sequences, and c) provide a framework for future molecular phylogenetic studies.

### **Materials and Methods**

#### ***Collection and Identification***

Specimens were collected subtidally by SCUBA diving or dredge, or intertidally by hand. Most species were collected along the west coast of North America from southeast Alaska to Baja California, with the exception of *Pseudocnus californicus*, which was collected from the Gulf of California (Table 1). Upon collection, specimens were cleaned in sea water, immediately frozen on dry ice, and subsequently stored at  $-70^{\circ}\text{C}$ .



Table 1. Classification and location of species in study

Order	Family	Genus, species	Development	Collection Site	PCR Fragment (nt)
Dendrochirotrida					
	Cucumariidae				
		<i>Cucumaria curata</i> Cowles, 1907	DD, EB	Pacific Grove, CA	1369
		<i>C. lubrica</i> Clark, 1901	DD, EB	Victoria, BC	1370
		<i>C. piperata</i> (Stimpson, 1864)	P, LL	Barkley Sound, BC	1364
		<i>C. pallida</i> Kirkendale and Lambert, 1995	P, LL	Victoria, BC	1364
		<i>C. miniata</i> (Brandt, 1835)	P, LL	Barkley Sound, BC	1367
		<i>C. pseudocurata</i> (Deichmann, 1938)	DD, EB	Barkley Sound, BC	1368
		<i>C. vegae</i> Theel, 1886	DD, EB	Juneau, AK	1368
		<i>Pseudocnus astigmatus</i> (Wells, 1924)	DD, EB	Victoria, BC	1370
		<i>P. californicus</i> (Semper, 1868)	NK	Gulf of California, Mex	1370
	Psolidae				
		<i>Psolus chitonoides</i> Clark, 1901	P, LL	Barkley Sound, BC	1406
	Phyllophoridae				
		<i>Pentamera lissoplaca</i> (Clark, 1924)	P, LL	Barkley Sound, BC	1417
	Sclerodactylidae				
		<i>Eupentacta quinquesemita</i> (Selenka, 1867)	P, LL	Barkley Sound, BC	1433

Order	Family	species	Development	Collection Site	PCR Fragment (nt)
Aspidochirotida					
	Synallactidae				
		<i>Pseudostichopus mollis</i> Theel, 1886	NK	West coast, Queen	1376
	Stichopodidae				
		<i>Parastichopus parvimensis</i> (Clark, 1913)	P, PL	Ensenada, Baja Cal., Mex.	1516
		<i>P. californicus</i> (Simpson, 1857)	P, PL	Barkley Sound, BC	1515

\* DD=Direct Development; EB=External Brooder; P=Pelagic; LL=Lecithotrophic larvae;

PL=Planktotrophic Larvae; NK=Not Known

Identifications were based on original descriptions or examination of the type specimens. Informative characters at this level are primarily ossicle structure or form. Ossicles were permanently mounted on glass slides (Lambert, 1985). The depth of occurrence for *Cucumaria lubrica* Clark, 1901 was not indicated in the original description, thus two putative *C. lubrica* were collected. Intertidal specimens are labelled *C. lubrica-i*, and subtidal specimens are referred to as *C. lubrica-s*.

#### *DNA Extraction*

Typically 200 milligrams (mg) of oral tentacle were ground in liquid N<sub>2</sub> in the presence of 600 uL of a proteinase solution (50 mM Tris- HCl, pH 7.5, 50 mM EDTA, pH 8.0, 0.4% SDS, 0.5 mg/ml Proteinase K). The ground samples were immediately placed at 65°C for a minimum of 2 hours until tissue was completely digested. The digested samples were repeatedly extracted with an equal volume of phenol:chloroform:isoamyl alcohol, 25:24:1, until the interface was clear. The resultant aqueous phase was adjusted to 0.5 M NaCl and to 1% cetyl trimethyl ammonium bromide (CTAB), and incubated for a further twenty minutes at 65°C. After two final extractions, total DNA was precipitated by adding an equal volume of isopropanol followed by sedimentation at 13000 RPM for 20 minutes in a microcentrifuge. The DNA pellet was rinsed with 500 uL of 70% ethanol, air-dried and resuspended in 200 uL of TE buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA).

### *Amplification Primers*

Amplification of specific mitochondrial DNA fragments was accomplished by designing conserved primers (Kocher et al., 1989) based on published sea urchin and sea star sequence data (Jacobs et al. 1988, Smith et al. 1993). All primers used in this study were synthesized by the Institute of Molecular Biology and Biochemistry (IMBB), at Simon Fraser University (Table 2). The first set of primers, which span the 3' end of the *12S*rRNA gene and the 5' third of the COI gene, were previously used to demonstrate a shared gene order between the echinoid and holothuroid classes (Smith et al., 1993). The second, overlapping primer set amplified the middle region of the COI gene, resulting in a data set of 350 nucleotides of the *12S*rRNA gene and 870 nucleotides of the COI gene.

### *PCR Amplification*

One hundred nanograms of template DNA and 25 picomoles of each primer were placed in a reaction mixture using either Amplitaq (Perkin-Elmer, Norwalk, CT), or Ultratherm (Eclipse Biologicals through Biocan Scientific, Mississauga, ON) thermal-stable DNA polymerase according to the manufacturers' specifications with a covering layer of mineral oil and held on ice until the thermal cycler had reached denaturation temperature. The amplification profile was as follows: An initial cycle of 95°C for 60 seconds(s), 50°C, 30s, 72°C, 60s, followed by 30 cycles of 95°C, 30s; 50°C, 30s; 72°C, 60s, followed by a final cycle with a 10 minute extension time.

Table 2. Oligonucleotide primers

Primer	Sequence 5'→3'	Mitochondrial Map Position <sup>1</sup>
16Sb	GACGAGAAGACCCTGTGGAGC	5297-5317
COIb	GGGTAGATAGTTCAGCCAGTTCC	6173-6150
COIef	ATAATGATAGGAGG[A/G]TTTGG	6000-6019
COIer	GCTCGTGT[A/G]TCTAC[A/G]TCCAT	6692-6673
16Scuc	TGACAAIIAGGATTGCGACC	5568-5587
16Sr	ACTTAGATAGAACTGACCTG	5710-5687

<sup>1</sup> Map position refers to the complete mitochondrial genome sequence of the sea urchin, *S. purpuratus*, (Jacobs, et al., 1988).

### *Product Purification*

Amplified products were electrophoresed on 1% agarose gels in a Tris-acetate-EDTA buffer (1X TAE ) containing 50 ng/ml ethidium bromide. Product bands, visualized at 365 nm wavelength, were excised with a razor blade. These agarose plugs were placed in micro-centrifuge spin columns. A minimal amount of siliconized glass wool was packed into the tip of the lower half of a 1 ml pipette tip. A 0.5 ml microcentrifuge tube with a hole punctured in the bottom served to hold the tip and was placed in a 1.5 ml tube to collect flow-through containing DNA. The tubes were centrifuged for 15 minutes at 5000 RPM and the column was discarded, leaving the DNA in the 1.5 ml tube. Product concentrations were estimated by agarose electrophoresis in gels containing DNA standards of known mass. If necessary, PCR products were precipitated and resuspended to achieve optimal concentration for DNA sequencing.

### *Cloning*

Direct sequencing of the 16S-COI PCR products was not always possible. In such cases the PCR products were cloned into the Sma I restriction enzyme site of pUC18 according to the T/A cloning protocol (Marchuk et al. 1991). A minimum of 3 clones were sequenced in order to eliminate any nucleotide sequence errors due to amplification or cloning (Saiki et al., 1988; Tindall and Kunkell, 1988; Kheovang and Thilly, 1989).

### *Sequencing*

DNA sequencing was based on the chain termination method (Sanger, 1977, Hatori and Sakaki, 1986). For manual sequencing, the Sequenase 2.0 System (US Biochemicals) was used as described previously (Smith et al., 1993), except that reactions were electrophoresed on a 4.5% polyacrylamide gels. Automated thermal cycle sequencing of the amplified fragments was utilized for most samples, according to manufacturer's instructions (Applied Biosystems Inc., ABI).

### *Sequence Alignment*

All sequences were manually entered and aligned using the Eyeball Sequence Editor program (ESEE, Cabot and Beckenbach, 1989). Initially the lrDNA sequences were aligned using Clustal V (Higgins et al., 1992) with a gap penalty of 10 and subsequently visually inspected to improve alignment. Two regions of ambiguous alignment were excluded from phylogenetic analysis (Figure 1).

COI coding sequence was translated using the echinoderm mitochondrial codon table (Himeno et al. 1987). A series of programs written by A. Beckenbach (pers. commun.) were used to transfer sequence files between program formats and to analyze substitution rates at each codon position.

### *Phylogenetic Analysis*

Initial inspection of nucleotide frequency and substitution patterns revealed significant departures from equal nucleotide usage. Recently concerns have been raised

Figure 1. Aligned nucleotide sequence of the 3' terminal lrRNA gene fragment from 14 species of sea cucumber. Capital letters in bold type indicate sites used for phylogenetic analysis, except for the 3' ATG of each sequence which represents the initiation codon of the COI gene; lower case letters refer to sites of ambiguous alignment which were excluded from the analysis. Original alignment was produced by CLUSTAL V (Higgins et al., 1993) with a transition/ transversion ratio of 2.0, final alignment was done manually (ESEE, Cabot and Beckenbach, 1989). Dots indicate nucleotide identity. *C.cura*, *Cucumaria curata*; *C.lubr-s*, *C. lubrica*-subtidal specimen; *Po. ast*, *Pseudocnus astigmatus*; *C.lubr-i*, *C. lubrica*-intertidal specimen; *C.pipe*, *C. piperata*; *C.mini*, *C. miniata*; *C.pseu*, *C. pseudocurata*; *Po.cal*, *Pseudocnus californicus*; *E. quin*, *Eupentacta quinquesemita*; *Pe.lis*, *Pentamera lissoplaca*; *P.chit*, *Psolus chitonoides*; *Ps. mol*, *Pseudostichopus mollis*; *Pa.cal*, *Parastichopus californicus*; *Pa.par*, *Parastichopus parvimensis*.



C.cura TGTGGAGCTTCAACT--TCCTT---aaaga-ata--gacaag-tttttgaaagaaaaatcCTCTTTAAAGAAGTTTTTGGTTGGGGCAA  
C.lubr-s .....a.a.a.....  
Po.ast .....t.t.atcc.a.....tcct...AA.T.....  
C.lubr-i .....t.t.atcc.a.....tcct...AA.T.....  
C.pipe .....T-----cctt..aacc.....ttat...AA.....  
C.mini .....c--t-.gcccc.a.....tcc.....A.....  
C.pseu .....c--t-.gacc.c.a.....tc.....A.....  
Po.cal .....AG.....-c.a.a.accacc.c.....c.c.tT...CCC.C.....T..  
E.quin .....ttag.....a-ctt.....t.....tc.tT...-A.T.....G.....T..  
Pe.lis .....A.C.....-a-ctt.....t.....tc.....-A.T.....GA.....T..  
P.chlt .....-C.TA---.ac--tttaa.cc.at.a.....-at.T...T...  
Ps.mol .....A.....-CTAAtttag.tta.-c.ac---tata.c.a.....-tt.tT...A--T..A.....T..  
Pa.cal .....A.G.AAC...gaaatgtc.tc.cctttt.ttc.....a.gg-t.aaT.T.AG.T--A...C.....  
Pa.par .....A.G.AAC...gaaatgtc.tc.cctttt.ttc.....a.gg-t.aaT.T.AG.T--A...C.....

C.cura CCACGGAG-TACAAAATACCTCCAGAAATTAACCGA tttt-acatcaataaaa---ttaaa-caa--AAAGAACCAG-AACCCCT  
C.lubr-s .....T...A.....  
Po.ast .....T.C.AG.....A.....t-.....cg.....  
C.lubr-i .....T.C.A.....A.....t-.....c.....  
C.pipe .....T...AC.....C.....c.t-a.....C-----T.....  
C.mini .....A...-A.....C.....t-.....C-----  
C.pseu .....A...TA.....T...G.....t-.....C-----  
Po.cal .....-G.T...A.T.....C.....A..aa..a.....g.....-C.....  
E.quin .....-A.A.TCA.T...T.T.TCA...AA.....a.t.ttc.....t.at..tt...TT.ATT..  
Pe.lis .....-A.A.CTAGT...T.T.CA...AA.....a.t.ttc.....t.a..tc...TT.ATT..  
P.chlt .....AA.T.G.AT.....CC...AA.....at.a...tt.....c.c.tg..-t.....ATA..  
Ps.mol .....-A.A..TA.....T...-AA...-act.a...ttt...c.c.ttt.....-c.....A-ATA..  
Pa.cal .....-A.A..G.AT...TT.....GGAA.....a...tt.c.....a...ttt.GC..C...TTTATT..  
Pa.par .....-A.A..G.AT...TT.....GGAA.....a...tt.c.....a...ttt.GC..C...TCTA.T..

C.cura GGCNAAACAGAAAAGTTACCGAGGGATAACAGCGTAATCTCCTTTAAGAGTTACATTTGACAAGGAGGATTGGACCCTCGATTGGAT  
C.lubr-s .....  
Po.ast .....T.....

C.lubr-i .T.....  
C.pipe .T.....  
C.mini .T.....  
C.pseu .T.....  
Po.cal .T.....C.....  
E.quin .T.....C.A.....CA.....  
Pe.lis .T.....C.A.....CA.....  
P.chit .A.....T.....A.A.....  
Ps.mol .T.....T.C.....CC.T.....G.A.....  
Pa.cal .A.....G.....T.....T.C.....CC.T.....G.A.....  
Pa.par .A.....G.....T.....T.....C.....CC.T.....G.A.....

*Deltocephalus*

3' End 360

C.cura TGGGGACACCTTAGGGTGCAGCAGCT-CTAATGGTTGGACTGTTCGTCCTAATAAACCCCTACATGATC-TGAGTTCAGACCCGAGAGT  
C.lubr-s .....  
Po.ast .C.....T.....T.....A.....  
C.lubr-i .C.....T.....T.....A.....  
C.pipe .A.....A.....-.....C.....-.....  
C.mini .C.....-.....C.....-.....  
C.pseu .C.....-.....C.....-.....  
Po.cal .A.....A.....-.....TA.....-.....  
E.quin .AT.....AA.A.....A.....-T.TTA.....T.T.C.....  
Pe.lis .AT.....AA.A.....A.....-T.TTA.....T.T.C.....  
P.chit .A.C.....AA.T.....T.....AT.....TC.....A.....  
Ps.mol .TA.....A.A.....T.A.....TT.....A.....TA.....  
Pa.cal .TA.....AA.....T.....CT.T.A.....C.....TT.....TA.....  
Pa.par .TA.....AA.....T.....CT.T.A.....T.....TT.....TA.....

*D. yakuba* 3' End

450

C.cura CAGGTCAGCTTCTATCTTCTT---AAAATTTTTTCTAGTACGAAAGGACCCGAAAAAATACTTCTATAAATTTTATCAAGGATAAatt  
C.lubr-s .....C.C.....C.C.....G.A.....G.A.....att  
Po.ast .....C.C.....C.C.....G.A.....G.A.....ttt  
C.lubr-i .....C.C.....C.C.....G.A.....G.A.....ttt  
C.pipe .....C.T.....C.....C.G.....C.G.....ctt  
C.mini ---C.T.....C.....G.C.....C.G.....cct  
C.pseu ---C.T.....C.....G.C.....G.C.....T.....tat  
Po.cal ---A.TCCA.C.C.C.....G.AG.....T.A.....CC.....A.....C.aaa

E.quin .....T.....CCC.C..A-.....C.....A.....TAT.....AA.....T.....A.Cata  
 Pe.lis .....T.....ACAT...AG.....C.....A.....TAT.....AA.....T.....A.Catt  
 P.chit .....TT-.T.TAACC.CC.C.....GG..T.A.....A.....T.....Ttta  
 Ps.mol .....A---.TG.A.....C.....G.....TC.....G.....C.....GA...tta  
 Pa.cal .....-----..TAACCC.....GGG.GG.C.T.A..G.CCAACG.T..A.G.....aaa  
 Pa.par .....-----..TAACCC.....GGG.GG.C.T.A..G.CCAACG.T.....G.....aaa

473

C.cura ttacaaataaaaaaaccccccaata**ATG**  
 C.lubr-s ttacaaataaaaaaaccccccaata**ATG**  
 Po.ast ttaaaaataaaaaaaccccccaata**ATG**  
 C.lubr-i ttaaaaataaaaaaaccccccaata**ATG**  
 C.pipe caaccaaacacaaaacacttcaata**ATG**  
 C.mini caactaaactaacaccacaatccccata**ATG**  
 C.pseu taccctaaacacactaaagtttataaca**ATG**  
 Po.cal ccacttaaaacaaaacaaaatccaata**ATG**  
 E.quin aacatacacaaccataaacctccaactaaacacacataatttaacaacataatttcaacataacttataaaa**ATG**  
 Pe.lis aatcataatcacaaattttaccacaaatcatacacacacaaacttttaacataatttataaaa**ATG**  
 P.chit gatcaacaaccataatttataaaaaacaaacccccaataataactataataaaaaat**ATG**  
 Ps.mol tttttcaaaaaaatttttaaaaaa**ATG**  
 Pa.cal aacaagaaaaaaaagttaaaaaaaaaccccttagtttaactagggaaacatcgaaaagggacacactgcaactcacttaaacctgcact  
 Pa.par aaaaacaataaaaaaaagttaaaaaaataaaccttagttaaatccagggaacgtcttaaaaattaacacactagcgactcacttaaacctac

611

Pa.cal gaaaaaaaagtttttggattttacacaaggggaccccccttatagttttccctagaaaattttggatttttaagaaaac**ATG**  
 Pa.par gctaaaaagaaaagtttttggattttacacaaggggaccccccttatggtttttctagaattttggatttttaagaaaac**ATG**

that many methods of phylogenetic reconstruction may be unsuitable when directed mutation pressure results in nucleotide composition bias (Jermiin et al., 1994a, 1994b; Lockhart et al., 1994). Therefore, several methods of phylogenetic reconstruction were utilized. The distance method of Tamura and Nei (1993), which takes into account differing nucleotide frequencies and substitution rates, was used with the neighbor-joining tree building method (Saitou and Nei, 1987) in the computer program package Molecular Evolutionary Genetic Analysis (MEGA) (Kumar et al., 1993). The maximum likelihood method (Felsenstein, 1981) using DNAML in Phylip 3.5 (Felsenstein, 1993), the maximum parsimony method using PAUP (Swofford, 1993), and a new method based on the determinants of a divergence matrix (Steel, 1993) (logDet) were also utilized. In cases involving irregular nucleotide composition, logDet has been shown to recover the correct tree when other methods have failed to do so (Lockhart et al., 1994). The logDet transformation was utilized in Splitstree (Huson and Wetzell, 1994), a computer program which allows differing portions of a tree to be represented by separate graphs and is thereby less constraining in cases of unequal rates of substitution or unique nucleotide composition among lineages.

## Results

### *PCR products*

The amplified mitochondrial DNA products from all the sea cucumbers contained 3' terminal regions of the lrRNA gene and the 5' half of the COI genes. The overall length of amplified products, as verified by DNA sequence analyses, varied between 1364 nucleotides (nt) (*Cucumaria piperata*, *C. pallida*) and 1515 nt (*Parastichopus parvimensis*) (Table 1). Except for a 6 nt (two codon) insertion in the COI sequence of *Eupentacta* and *Pentamera*, and minor insertions/deletions in the lrRNA gene sequences, this length variation mapped to a highly variable region directly upstream of the COI initiation codon and represents either terminal lrRNA gene sequence or unassigned sequence, UAS, (Figure 1). All of the large ribosomal genes contained nucleotide blocks that are identical to DNA segments close to the 3' end of the lrRNA gene of *Drosophila D. yakuba* 3'-terminal block represents the last highly conserved block among the sea cucumber sequences. Nonetheless, the exact 3' termini of individual sea cucumber mitochondrial lrRNA genes could not be assigned with certainty. The complete DNA sequences of the amplified fragments have been submitted to the Genbank data base (Accession Nos. U31901, U32210-U32221, U32198 and U32199).

### *Sequence Analysis of the PCR Fragments*

#### *Cytochrome oxidase I sequence.*

The identity of the COI gene sequences was determined by similarity of the peptide sequence to the COI gene of the sea urchin, *Strongylocentrotus purpuratus* (Jacobs et al. 1988). The resulting amino acid sequences of the sea cucumber COI fragment are aligned in Figure 2. All COI genes are assigned ATG as the initiation codon. Both *Eupentacta quinquesemita* and *Pentamera lissoplaca* COI genes contain a two amino acid insertion following the initiation codon (Figure 2). The amino acid sequences of both species of *Parastichopus* were identical and the species pair *C. pseudocurata/C. vegae* share identity with *C. miniata*. The number of amino acid differences within the Cucumariidae ranges from 1 to 8 between species with the maximal difference (8) between *Pseudocmus californicus* and the members of the genus *Cucumaria*. There are only two comparisons between families and orders show differences ranging from 11 to 31 amino acids.

A detailed analysis of base composition was undertaken in order to assess patterns of change and examine the validity of assumptions used in phylogenetic reconstruction. The GC content of the sea cucumber COI gene fragments ranged from a low of 34.7% in *Pseudostichopus mollis* to a high of 43.5% in *C. miniata*, with an average of 42%. As reported for numerous organisms, including a number of other echinoderms (Jacobs et al., 1988, Asakawa et al., 1991, Cantatore et al. 1989), the proportion of guanine (G) in the third codon position was low; 6-12% (Table 3), with a corresponding

Figure 2. Inferred amino acid sequences for the sea cucumber CO1 gene fragments. Species designations are as given in the legend to Figure 1 with the addition of *Cucumaria pallida*, abbreviated as C.pall. The inferred amino acid sequence of the *C. lubrica*-s CO1 gene fragment is identical to *C. curata*, and those of *C. pseudocurata* and *C. vegae* (abbreviated C.vega elsewhere) are identical to *C. miniata* and are not shown. E.quin and Pe.lis have a putative 2 amino acid insertion (IN) at codon positions 2-3. Dots indicate amino acid identity.

C.cura M--LNRWLFSTNKHKDIGTLYLLFGAWAGMVTAMSVIIRTELAQPGSLLNDDQIYNTIVTAHALVMIFFMVPIMIGGFGNWLIPLMIGA  
 Po.ast .....  
 C.lubr-i .....  
 C.pipe .....  
 C.pall .....  
 C.mini .....  
 Po.cal .....  
 E.quin .....  
 Pe.lis .....  
 P.chit .....  
 Ps.mol .....  
 Pa.cal .....

C.cura PDMAFPRMNNMSFWLIPPSFILLASAGVENGAGTGWTIYPPPLSSNIAHAGGSVDLAI FSLHLAGASSILASINFITTIINMRSPGVTFD  
 Po.ast .....  
 C.lubr-i .....  
 C.pipe .....  
 C.pall .....  
 C.mini .....  
 Po.cal .....  
 E.quin .....  
 Pe.lis .....  
 P.chit .....  
 Ps.mol .....  
 Pa.cal .....



270

C.cura	RLPLFWSVFITAFLLLSLPLVIAGAITMLLTDRNLNTSFPDPAGGDPILFQHLFWFFGHPEVYILILPGFGMISHVIAHYSGKQEPFG
Pe.ast	.....I.....
C.lubr-i	.....D.....I.....
C.pipe	.....I.....
C.pall	.....I.....
C.mini	.....I.....
Pe.cal	.....I..T.....
E.quin	.....I..T.....V..I.....S.....
Pe.lis	.....I..T.....V..I.....S.....
P.chit	.....I.....I..T.....N.....
Ps.mol	.....A.....I..T.....I.....N.....
Pa.cal	.....I.....I..T.....

293

C.cura	YLGMYAMVAIGILGFLWAHHM
Pe.ast	.....
C.lubr-i	.....T.....
C.pipe	.....
C.pall	.....
C.mini	.....
Pe.cal	.....
E.quin	.....
Pe.lis	.....
P.chit	.....I.....
Ps.mol	.....I..V.....
Pa.cal	.....

Table 3. Nucleotide composition of sea cucumber COI gene fragments

	Position 3			4-Fold Degenerate				Total				
	A3	T3	C3	G3	A4	T4	C4	G4	A	T	C	G
<i>Cucumara curata</i>	47.8	18.3	27.0	6.9	57.7	12.6	27.9	1.8	30.9	27.0	24.6	17.5
<i>C. lubrica-s</i>	47.1	17.3	28.4	7.3	55.9	12.6	29.1	2.3	30.7	26.6	25.0	17.6
<i>Pseudoecmus astigmatus</i>	45.7	19.4	28.0	6.9	57.7	15.3	25.2	1.8	30.4	27.3	24.8	17.4
<i>C. lubrica-i</i>	45.7	19.4	28.0	6.9	56.0	16.0	26.4	1.6	30.7	27.3	24.7	17.3
<i>C. piperata</i>	45.7	17.0	29.1	8.3	58.3	12.6	26.0	3.1	30.4	26.4	25.3	17.9
<i>C. pallida</i>	47.8	17.3	25.6	9.3	61.3	11.7	20.7	6.3	31.4	26.4	24.1	18.1
<i>C. miniata</i>	44.6	17.3	28.4	9.7	55.0	17.8	20.9	6.2	30.2	26.3	25.3	18.2
<i>C. pseudocurata</i>	42.6	24.6	21.1	11.8	55.2	19.2	15.2	10.4	29.5	29.3	22.3	18.9
<i>C. vegae</i>	43.3	24.2	21.1	11.4	56.8	18.0	16.2	9.0	29.8	29.1	22.4	18.8
<i>Pseudoecmus californicus</i>	56.7	9.3	27.7	6.2	70.6	8.7	20.6	0.0	34.1	23.8	24.8	17.3
<i>Eupentata quinquesmita</i>	47.8	24.7	20.6	6.9	54.0	27.0	18.2	0.8	31.7	29.3	21.9	17.1
<i>Pentamera fissoplaca</i>	43.6	24.7	23.0	8.6	50.0	25.8	20.2	4.0	30.5	29.3	22.8	17.4
<i>Psolus chitonoides</i>	55.4	17.6	19.4	7.6	63.1	13.1	20.5	3.3	34.6	26.9	21.3	17.2
<i>Pseudostichopus molitis</i>	50.9	33.6	9.0	6.6	58.9	33.1	8.1	0.0	32.2	33.1	18.0	16.7
<i>Parastichopus californicus</i>	36.7	29.8	22.5	11.1	35.4	31.5	24.6	8.5	27.1	30.9	22.7	19.3
<i>Parastichopus parvimensis</i>	34.9	31.1	21.5	12.5	35.4	30.1	24.8	9.7	26.5	31.0	22.7	19.7

increase in adenine (A) to an average of 47%. It is of interest that the echinoderm mtDNA genetic code utilizes only ATG for methionine and AAG for lysine. This bias against G therefore appears even stronger when one examines four-fold degenerate sites only (Table 3). This is most evident for the two species *Pseudocmus californicus* and *Pseudostichopus mollis* where G was not recorded at any four-fold degenerate sites. For *Pseudocmus californicus*, a bias against thymine (T) at degenerate sites was also apparent where the frequency of T was 8.7%, resulting in 70.6% of such sites being occupied by A. *Pseudostichopus mollis* shows a similar bias against cytosine (C) where a frequency of 8.1% was observed. The bias against G is reflected in observed pairwise substitution comparisons, where the vast majority of transversions do not involve G. Nearly all observed pairwise transversions were A↔T or A↔C while T↔C changes are predominant among transitions. In only two pairwise comparisons does the A↔G transition change exceed T↔C changes and both involve closely related species. The first case is the comparison between *Pseudocmus astigmatus* versus *C. lubrica*-i in which the COI comparison shows only two nucleotide changes, one a transition and the second a transversion. The second case is the pairwise comparison between *C. pseudocurata* and *C. vegae* where there are 13 A↔G and seven T↔C transitions and a single A↔T transversion.

The effect of the bias against G is further reflected in the codon usage in the COI sequence. The summed codon usage of 16 sea cucumber COI gene fragments is shown in

Table 4. There are 8 codons for serine (AGN and TCN). In the 16 sea cucumber COI fragments there are 299 serines. Of these there is a 2:1 usage of TCN vs. AGN codons (201:98). In third positions of serine codons there is a 30:1 ratio favoring third position A over G (only 6 of the 299 serine codons end in G) while T and C were approximately three fold less common than A. Six leucine codons are utilized a total of 592 times among gene fragments from the 16 species where A is 4-5 times more common than C or T and 11 times more common than G. In the 4 member codon families, the third position is preferentially occupied by A (920 codons) over C (361 codons) over T (353 codons) over G (79 codons). Only in the case of glycine codons is the frequency of G in the third position not the lowest while A is still by far the most common (316/461).

The aligned variable sites in the Cucumariidae COI gene fragment are shown in Figure 3. In that alignment we have included two putative *C. lubrica* as discussed, and two species of *Pseudocmus*, the recently reclassified *P. astigmatus* (Lambert, 1990) and were variable. Pairwise comparisons of Tamura-Nei distance estimates are given in Table 5.

Transition/transversion ratios ranged from a high of 20.0 for the 21 nucleotide changes between the closely related *C. pseudocurata*/*C. vegae* pair, to 0.64 for the 146 nt changes between *Eupentacta* and *C. miniata*. The *C. lubrica*-s/*C. curata* were the most similar with only 7 changes in 867 nucleotides examined. Despite this high similarity, the transitions to transversion ratio for this pair was only 1.3. In Table 6, the pairwise species comparisons in which the transition to transversion ratio exceeded 2.0 are listed.

Table 4. Overall codon usage for the 870 nucleotide CO1 gene fragments from 16 sea cucumber species.

F TTT 185	S TCT 37	Y TAT 45	C TGT 0
F TTC 180	S TCC 48	Y TAC 67	C TGC 0
L TTA 135	S TCA 111	# TAA 0	W TGA 121
L TTG 4	S TCG 5	# TAG 0	W TGG 7
L CTT 149	P CCT 83	H CAT 32	R CGT 4
L CTC 79	P CCC 71	H CAC 96	R CGC 18
L CTA 215	P CCA 115	Q CAA 62	R CGA 68
L CTG 10	P CCG 3	Q CAG 1	R CGG 6
I ATT 136	T ACT 50	N AAT 18	S AGT 9
I ATC 86	T ACC 45	N AAC 57	S AGC 17
I ATA 289	T ACA 149	N AAA 144	S AGA 71
M ATG* 241	T ACG 6	K AAG 32	S AGG 1
V GTT 81	A GCT 67	D GAT 36	G GGT 68
V GTC 39	A GCC 158	D GAC 109	G GGC 30
V GTA 105	A GCA 167	E GAA 61	G GGA 316
V GTG 11	A GCG 6	E GAG 3	G GGG 47

\* *excluding initiation codons*

Figure 3. Alignment of the variable nucleotide sites of the CO1 gene fragment from ten Cucumariidae species. The nucleotide position of the variable sites is indicated over the sequence. Dots indicate nucleotide identity. Species designations as given in the legend to Figure 1.







Table 5. Tamura-Nei distance matrix for the CO1 gene fragments

	1	2	3	4	5	6	7	8	9	10	11
1) <i>Cucumaria curata</i>		0.8	8.3	8.5	12.4	13.4	13.5	15.1	14.7	16.4	22.0
2) <i>C. lubrica-s</i>	0.3		8.4	8.7	12.7	13.6	13.9	15.1	15.0	16.8	22.1
3) <i>Pseudocnus astigmatus</i>	1.1	1.1		0.2	12.5	14.5	14.0	15.3	15.5	16.7	22.1
4) <i>C. lubrica-i</i>	1.1	1.1	0.2		12.7	14.7	14.2	15.5	5.8	17.0	22.4
5) <i>C. piperata</i>	1.3	1.4	1.4	1.4		13.0	12.2	13.7	13.1	15.1	22.0
6) <i>C. pallida</i>	1.4	1.4	1.4	1.4	1.3		6.5	10.7	10.7	15.3	19.6
7) <i>C. miniata</i>	1.4	1.4	1.4	1.4	1.3	0.9		11.1	10.7	15.1	19.3
8) <i>C. pseudocurata</i>	1.5	1.5	1.5	1.5	1.4	1.2	1.3		2.5	17.5	20.1
9) <i>C. vegae</i>	1.4	1.5	1.5	1.5	1.3	1.2	1.2	0.6		17.5	21.3
10) <i>Pseudocnus californicus</i>	1.6	1.6	1.6	1.6	1.5	1.5	1.5	1.6	1.6		19.9
11) <i>Enpentacta quinquesemita</i>	1.8	1.8	1.8	1.8	1.8	1.6	1.6	1.7	1.7	1.7	

Distances in the upper-right matrix

Standard Errors in lower-left matrix

All values in per cent

Table 6. Pairwise species comparisons of CO1 gene fragments with transition/transversion ratios greater than 2.00

Species	S/V	D (# changes/870)
<i>C. curata/P. astigmatus</i>	2.14	66
<i>C. curata/C. lubrica-i</i>	2.09	68
<i>C. lubrica-s/Pseudocmus astigmatus</i>	2.05	67
<i>C. lubrica-s/C. lubrica-i</i>	2.00	69
<i>C. pallida/C. pseudocurata</i>	2.40	85
<i>C. pallida/C. vegae</i>	2.27	85
<i>C. miniata/C. pseudocurata</i>	4.44	87
<i>C. miniata/C. vegae</i>	3.94	84
<i>C. pseudocurata/C. vegae</i>	20.00	21
<u><i>Parastichopus californicus/P. parvimensis</i></u>	3.22	38

### *Large Ribosomal RNA Gene Sequence*

The nucleotide composition of the lrRNA gene fragment for fourteen sea cucumbers representing two orders and six families are shown in Table 7. The complete sequence of the fragments amplified are shown in Figure 1, where the regions aligned for phylogenetic analyses are indicated. Pairwise comparisons of aligned sequences show that transition to transversion ratios are less than 2 for all comparisons except for the following instances: between the *Parastichopus* species (4 transitions and no transversions), between *Parastichopus californicus* or *Parastichopus parvimensis* and *Psolus* (transition to transversion ratios of 2.3 and 2.2 respectively), and between *Pseudostichopus mollis* and *C. miniata* (ratio = 2.2). In the pairwise comparisons A↔T transversions always exceed A↔C changes except in the *C. miniata/C. lubrica* comparison where there is a single A↔T and 3 A↔C transversion events.

### *Patterns of Nucleotide Substitution*

For the COI gene, the observed directional mutation pressure results in a restricted substitution pattern where C↔T transitions and transversions involving A predominate. A similar pattern is seen in the lrRNA gene where A↔G transitions are only half as common as C↔T events and transversions involving A occur 12 times more frequently than those involving G.

Table 7. Nucleotide base composition of the aligned portions of the sea cucumber lrRNA gene fragments

	A	T	C	G	Length
<i>Cucumaria curata</i>	32.2	25.7	20.6	21.5	354
<i>C. lubrica-s</i>	32.5	25.7	20.3	21.5	354
<i>Pseudocmus astigmatus</i>	32.4	25.6	20.8	21.1	355
<i>C. lubrica-i</i>	32.7	25.6	20.8	20.8	355
<i>C. piperata</i>	32.1	26.1	20.3	21.5	349
<i>C. miniata</i>	31.2	24.9	22.4	21.5	353
<i>C. pseudocurata</i>	31.1	26.3	20.9	21.8	354
<i>Pseudocmus californicus</i>	32.1	23.9	22.0	22.0	355
<i>Eupentacta quinquesemita</i>	33.3	27.7	19.0	19.9	357
<i>Pentamera lissoplaca</i>	34.4	26.8	17.9	20.9	358
<i>Psolus chitonoides</i>	34.0	25.6	20.5	19.9	356
<i>Pseudostichopus mollis</i>	34.7	26.3	17.8	21.2	354
<i>Parastichopus californicus</i>	30.5	24.6	20.7	24.1	357
<i>Parastichopus parvimensis</i>	30.3	24.1	21.3	24.4	357
Average	32.4	25.6	20.4	21.6	

All values in per cent (%) except total fragment length (in nucleotides)

*Molecular Phylogeny and Systematics of the Cucumariidae*

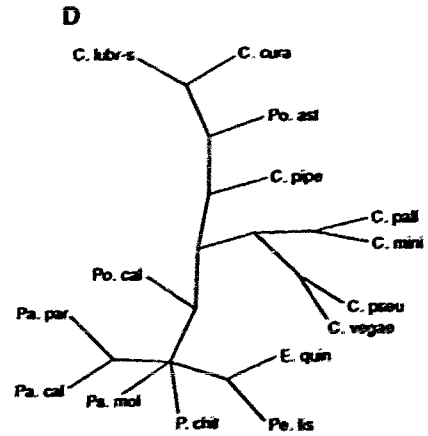
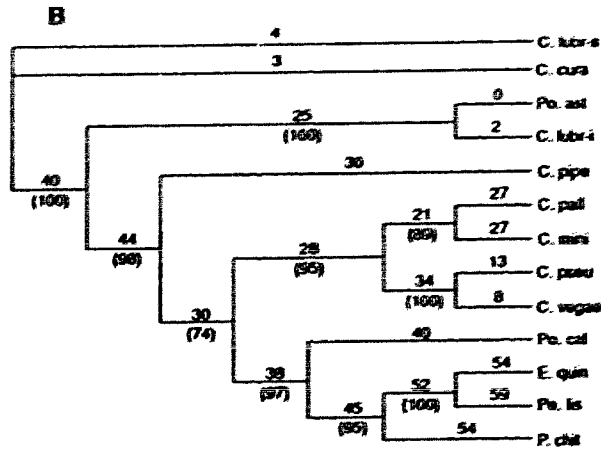
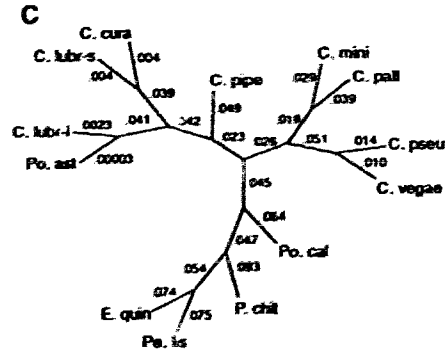
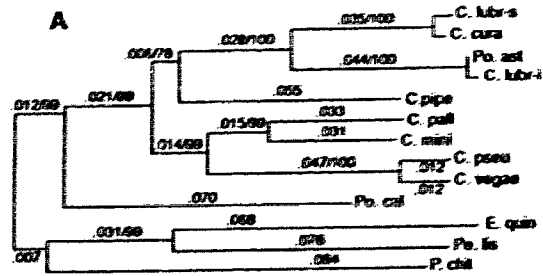
Prior to the molecular phylogenetic analysis of the family Cucumariidae using the COI nucleotide sequence, the intraspecific DNA polymorphism in that gene was assessed. DNA fragments containing COI gene sequence were amplified from 23 individual *C. pseudocurata* and six *C. miniata*. A 630 nt fragment of the gene from *C. pseudocurata*, mapping between codons 231 and 443 of the *S. purpuratus* gene (Jacobs *et al.*, 1988), was sequenced as was a 657 nt fragment from *C. miniata* (*S. purpuratus* codons 225-446). The DNA sequence of the 23 *C. pseudocurata* COI fragments fell into 9 haplotypes, while the *C. miniata* sequence fell into 4 haplotypes. The Tamura-Nei distance estimates between *C. pseudocurata* haplotypes ranged from 0.16% to 1.13% while that between *C. miniata* haplotypes ranged from 0.15% to 0.61%. Distance estimates between the two species varied from 10.8-12.2%, while a value of 11.2% was obtained from the fragments used for phylogenetic analysis, indicating that the rate of divergence does not vary significantly between regions of the COI gene fragments used. In addition, two *E. quinquesemita* specimens were collected and compared; one from northern Baja California, Mexico and one from Barkley Sound, Canada. These geographic isolates were of particular interest since the former possesses red-coloured tubefeet as opposed to the cream white tubefeet of northern specimens. The two specimens differed by 0.3% (two transitions: T↔C and A↔G; and one A↔C transversion) in the 870 nucleotides of COI sequence, a level of difference in the range noted between *C. miniata* individuals.

The phylogenetic analysis of the Cucumariidae utilized three species; *Eupentacta quinquevittata*, *Pentamera lissoplaca*, and *Psolus chitonoides*, belonging to three different families in the same order as the Cucumariidae, the Dendrochirotrida, as outgroups. The close relationships between the *C. curata*/*C. lubrica-s* pair and the *C. lubrica-i*/*P. astigmatus* pair can be seen in the trees in Figure 4. Also evident in these trees is the fact that the two *Pseudocmus* species clearly do not form a monophyletic group. All four methods used consistently place *P. astigmatus* in a paraphyletic arrangement within the *Cucumaria* (Figure 4A-D).

An examination of the type specimen of *C. lubrica* Clark, 1901 (MCZ #202) indicated that the intertidal *C. lubrica* most closely matched the original description (Clark, 1901) with lobed 4-holed button ossicles and some pine cone-shaped ossicles (Figure 5 D,E). Clark stated only that *C. lubrica* was collected in Puget Sound, with no indication of depth. These types of ossicles are also found in *Pseudocmus astigmatus* (Figure 5C). These two species exhibit COI gene nucleotide identity of 99.8%, well within the range of estimates of intraspecific variation for *C. pseudocurata* (0.16-1.13%). Similarly, *C. curata* from Pacific Grove exhibits 99.2% COI sequence identity with the seen in interpopulation comparisons of *C. pseudocurata*. Both *C. curata* and *C. lubrica-s* possess the same kind of smooth 4-holed button ossicles (Figure 5 A,B).

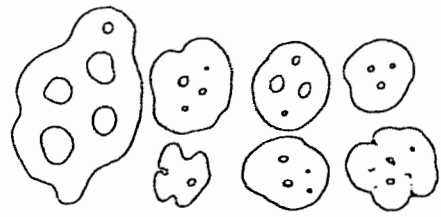
Due to the G bias discussed above, it is possible that 3rd codon positions are saturated for some comparisons even within the genus *Cucumaria* and simple parsimony analyses would not reflect multiple changes at nucleotide sites. Analysis of 1st and 2nd codon positions only, resulted in a branch and bound parsimony tree with one minor

Figure 4. Phylogenetic trees for the family Cucumariidae based on the nucleotide sequence of the amplified COI gene fragments. **A**, Neighbor-joining tree using Tamura-Nei distance matrix; both branch lengths and bootstrap confidence intervals from 500 replicates are indicated. **B**, Parsimony tree using branch and bound method with bootstrap values from 200 replicates. Branch lengths are indicated as number of changes, bootstrap values are given as percentages after the slash. **C**, Maximum likelihood estimate using DNAML, a total of 3157 trees were examined. **D**, LogDet tree using Splitstree (Huson and Wetzel, 1994) tree construction method. Dots indicate nucleotide identity. *C.cura*, *Cucumaria curata*, *C. lubr-s*, *C. lubrica*-subtidal specimen, *Po ast*, *Pseudocmus astigmatus*, *C. lubr-i*, *C. lubrica*-intertidal specimen, *C pipe*, *C. piperata*, *C. mini*, *C. miniata*, *C. pseu*, *C. pseudocurata*, *Po cal*, *Pseudocmus californicus*, *E. quin*, *Eupentacta quinqueemita*, *Pe lis*, *Pentamera lissoplaca*, *P. chit*, *Psolus chitonoides*, *Ps mol*, *Pseudostichopus mollis*; *Pa cal*, *Parastichopus californicus*, *Pa par*, *Parastichopus parvimensis*. Three species from the order, Aspidochirotida: *Pa cal*, *Parastichopus californicus* *Pa par*, *Parastichopus parvimensis* and *Ps mol*, *Pseudostichopus mollis* are also included to indicate uncertainty of branching order beyond the Family level. In both **C** and **D** all branch lengths are drawn equal.

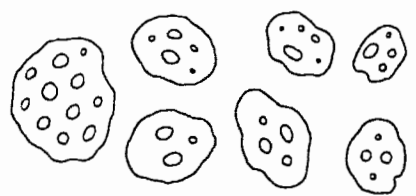




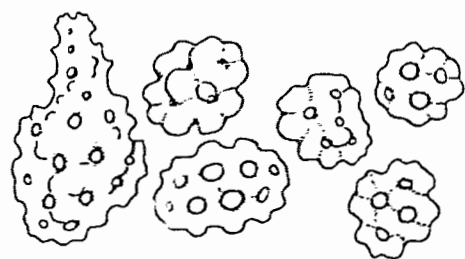
**Figure 5. Comparison of calcareous ossicle morphology of brooding sea cucumber species. All ossicles were obtained from the mid-dorsal region. Scale bar equals 100  $\mu$ m. A, *C. lubrica-s*, Victoria, B.C.; B, *C. curata*, Pacific Grove, CA; C, *P. astigmatus*, Sooke, B.C.; D, *C. lubrica-i*, Victoria, B.C.; E, *C. lubrica*, cotype, Puget Sound, WA.**



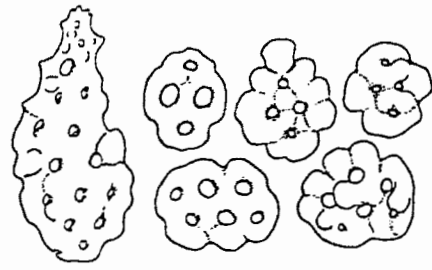
**A**



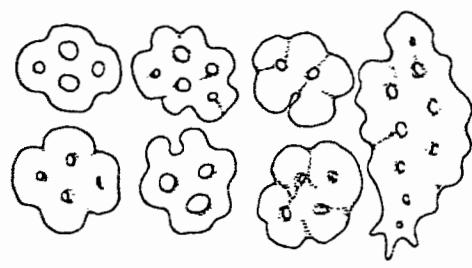
**B**



**C**



**D**



**E**

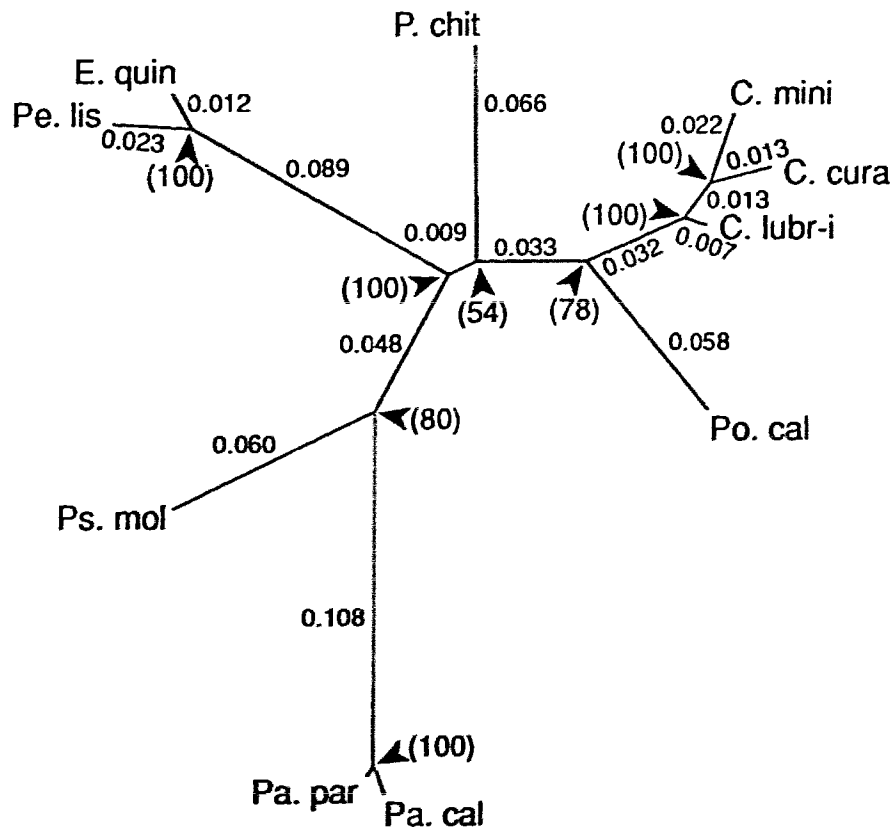
difference; placement of *C. piperata* next to the *C. miniata/C. pallida* pair rather than next to the *C. curata/C. lubrica* clade (not shown).

### *Phylogeny of Holothuroids at the Family Level*

Our preliminary analyses of pairwise comparisons of the COI gene fragment at the nucleotide level demonstrated that at the family level and above, most variable sites were saturated for change and, consequently, were of limited utility at such levels. Two distinct approaches were used to ascertain relationships among the sea cucumber families. The 18S rRNA gene sequences were aligned and those positions for which unambiguous alignment could be determined were used in maximum likelihood and bootstrapped parsimony estimates of sea cucumber phylogeny. For the maximum likelihood estimates, the Kishino-Hasegawa statistical test (1989) of user-defined trees was utilized to estimate the likelihood of alternative tree topologies.

Both maximum likelihood and bootstrap parsimony approaches mentioned above yielded congruent results. The best-fit DNA maximum likelihood estimate from 3037 trees examined for the aligned 18S rDNA sequences is shown in Figure 6. In that Figure are shown the confidence limits for the nodes from 200 bootstrap parsimony replicates. Several alternative tree topologies were assessed with the user tree option. The movement of *Psolus chitonoides* to the edge leading to the *Pentamera/Eupentacta* clade does not result in a statistically significant log likelihood difference. This demonstrates the ambiguity in the placement of the Psolidae. Of particular interest is the relationship between *P. lissoplaca* and *E. quinquesemita*, which are currently assigned to different

Figure 6. Maximum Likelihood phylogeny based on the nucleotide sequence of the sea cucumber 3'-terminal 18S rRNA gene fragment. 3037 trees were examined. Decimal values indicate branch lengths while those in brackets with arrowheads refer to percent bootstrap confidence levels from 200 replicates. Species designations as in the legend to Figure 1.



families. The distance estimate of 4.11% is very similar to the intrafamilial comparison of *C. miniata* and *C. lubrica* (3.5%) and no other interfamilial comparisons have a value less than 10%.

## Discussion

### *Phylogenetic Analyses*

DNA sequences of the PCR amplification products corresponding to portions of the COI and lrRNA genes revealed the presence of length variation among the taxa examined. This variation was localized to the 3' end of the lrRNA gene and represents either unique 3' terminal sequence in the lrRNA or variable length unassigned sequence. Direct RNA sequence analysis to confirm the precise 3' termini could provide insight into these two possibilities.

Both *Eupentacta quinquesemita* and *Pentamera lissoplaca* exhibited an insertion of six nucleotides (ATCAAC) corresponding to a two amino acid insertion (IN) between codons two and three of the COI gene. The initial four codons of all other species are MNLN, with a consensus sequence of ATGAACTAAAC. The insertion may represent duplication of either the first two, or the third and fourth, codons with subsequent nucleotide divergence. Either of these possibilities would require two transversions to produce ATCAAC seen in both *E. quinquesemita* and *P. lissoplaca*.

Similar patterns of nucleotide frequency and substitutions were observed in both the 18S rRNA and COI gene sequences. Although G accounted for 21.6 % and 16.9 % of sites, respectively, C↔T transitions were far more common than A↔G transitions and transversions involving G were less than one tenth as common as those involving A. Particularly illustrative is the extreme bias against G at four-fold degenerate sites of the COI gene fragment, where G accounts for 4.3 % of such sites. It therefore appears that directional mutational pressure against G has been exerted on both protein coding and RNA genes and may be a strand-specific effect as suggested (Osawa et al., 1992; Asakawa et al., 1991; W. K. Thomas, pers. commun.). Since Neighbor-Joining with the Tamura-Nei distance method (Tamura and Nei, 1993), maximum likelihood (Felsenstein, 1993) and Splitstree (Huson and Wetzel, 1994) using a logDet transformation (Lockart et al., 1994; Steel, 1993) produced congruent phylogenetic trees, it was concluded that the analysis has adequately compensated for mutational bias. Further analysis, particularly of nuclear loci would clearly strengthen the phylogenetic conclusions presented.

Phylogenetic analysis of the Cucumariidae clarified several controversial points concerning the taxonomy of this group. COI gene sequence analysis demonstrated that *P. astigmatus* and the intertidal *C. lubrica* specimen exhibit a level of similarity (99.8%) well within the estimates of intra-specific variation for another brooding species, *C. pseudocurata* ( $d = 0.0-1.13\%$ ). Examination of ossicle morphology also indicated that these two specimens were indistinguishable (Figure 5 C,D) and most closely matched the original description of *C. lubrica* Clark, 1901 (Figure 5E). Thus both morphological and

mtDNA evidence provide strong support that these two species are synonymous. This species synonymy could also resolve the apparent paraphyly in the genus *Cucumaria* (Figures 4). The recent revision of *Cucumaria fisheri astigmatus* to *Pseudocnus* (Lambert, 1990) may therefore have been incorrect. The genus *Pseudocnus* is described as having an outer layer of pine cone-shaped ossicles supported by a layer of plates (Panning 1962). *P. astigmatus* has pine cone ossicles only in clusters around the podia. Molecular evidence suggests that the mere presence of pine cone ossicles was not sufficient to warrant the placement in the *Pseudocnus* genus. Such an example demonstrates the difficulty of determining true synapomorphic traits that signify close relationships.

The evidence from morphology and mtDNA strongly suggest that the black subtidal species that is usually referred to as *C. lubrica* may well be conspecific with the earlier described *C. curata*. Although *C. curata* has only been unequivocally reported intertidally from the Monterey CA area, further characterization is required to determine whether this variation in habitat utilization represents a local adaptation or is more widespread. A detailed survey of both subtidal and intertidal populations along the coast of North America could clarify our understanding of how such differing populations or species arose.

This analysis also demonstrated that *C. curata* and *C. pseudocurata* display nucleotide distances sufficient to support distinct species status (COI: d=15%), as argued by Lambert (1985) based on ossicle morphology, and do not represent geographical variants of a single species as Rutherford (1977a) concluded. In fact, the mtDNA data



indicate that *C. pseudocurata* is much more closely related to the more northerly distributed species *C. vegae*, with a distance estimate ( $d=2.5\%$ ) approximately twice that of intraspecific values ( $d=0.16-1.13\%$ ). A discussion regarding the separate status and geographic partitioning of these two closely related species is presented in Chapter Two. If these congeners in fact represent geographic variants of a single species, *C. vegae* Theel, 1886, is the senior synonym.

The results also support the recent species designation of *C. pallida* (Kirkendale and Lambert, 1995) as distinct from *C. miniata*, which it closely resembles. The comparison of *C. pallida* to *C. miniata* resulted in a distance estimate of 5.1%, an order of magnitude greater than intraspecific *C. miniata* estimates ( $d=0.1-0.6\%$ ).

As discussed above, despite the similar size, reproductive biology, habitat, and overall appearance of *C. curata* and *C. pseudocurata*, both mtDNA and ossicle form suggest these species are quite distinct and, in fact, indicates a considerable degree of convergence. Assuming the planktonic larvae of *C. piperata*, *C. miniata* and *C. pallida* represent a more primitive condition, it appears that brooding has arisen in at least two separate lineages within the Cucumariidae (Figure 4B).

Analysis of unambiguously aligned lrRNA gene sequence demonstrated a suitable level of variation for phylogenetic analysis, both among genera and among families. Notwithstanding the difficulties of alignment of the ribosomal gene sequences, the lrRNA gene comparisons appear to be more useful at these levels than those using the COI gene due to the lower level of observed divergence. The resultant trees are largely supportive of currently accepted classification. Based on fossil evidence from the late Cretaceous,

Gilliland (1992) has proposed a recent shared ancestry for the Sclerodactylidae, the Phyllophoridae and the Cucumariidae. The lrRNA gene sequence data support a close relationship between the Sclerodactylidae (*E. quinquesemita*) and the Phyllophoridae (*P. lissoplaca*) (Figure 6) and in fact calls into question the separation of these two species into separate families. A primary discriminating character for these families is the form of the calcareous esophageal ring. Elements of the ring are formed from a mosaic of minute pieces in the Phyllophoridae while they are not formed from minute pieces in the Sclerodactylidae. Examination of additional species is necessary to determine whether such differences warrant separate family status. Our results clearly indicate an earlier separation of the Sclerodactylidae/Phyllophoridae and the Cucumariidae lineages. However, the suggested monophyly of these families is not strongly supported by our data. This is due to the ambiguous placement of the Psolidae (Figure 6) resulting in an unresolved branching order among these families. This may be somewhat surprising owing to the presumed appearance of psolid fossils in the Triassic (see Gilliland 1992), well before any evidence of the other groups. Our analysis may be reconciled with fossil evidence in one of several ways: First, the Sclerodactylidae/Phyllophoridae, Cucumariidae and Psolidae lineages may be of approximately equal age but the macroscopic calcareous plates of the Psolidae may have had a much greater probability of fossilization and hence an apparent earlier origin. Second, the macroscopic plates of the Psolidae may be a secondary acquisition, with the origin of this family being later than the first appearance of such presumed ancestral plates. Finally, bias may have been introduced by the limited number of species and sequences sampled. Examination of

other species in these families, as well as the inclusion of additional, more slowly evolving loci, are needed to resolve these questions.

The present study has demonstrated the utility of molecular analysis in light of morphological evidence, particularly in systems such as the Holothuroidea where simple form and poor fossilization have prevented complete elucidation of evolutionary history. Phylogenetic analysis of the COI and 18S rRNA genes has provided information at both the genus and family levels which will hopefully serve as a foundation for further molecular analysis of this class.

## CHAPTER II

### **Genetic Diversity and Population Structure of Sea Cucumber Species with Differing Modes of Development**

#### **Introduction**

Minor genetic differences arise by mutation within and among populations and Darwin (1859) was one of the first to note the importance of intrapopulation diversity in the process of speciation. However, these new forms or alleles are initially rare and are therefore usually lost in the stochastic sampling of gametes over time. Occasionally, such forms increase in frequency in populations of finite size, either by chance or due to some selective advantage of the new allele and, on a local scale, may replace the ancestral form, thus establishing a unique local population. Conversely, migration or dispersal of individuals or gametes with subsequent successful participation in reproduction, referred to as gene flow, will oppose such local differentiation through the random mixing of alleles from differing local populations. The level of gene flow is thus responsible for the scale at which populations randomly mix. Barriers to gene flow will result in isolation of local populations and lead to geographic differences or structuring of populations. Slatkin (1987) has described gene flow as being both a creative and constraining force in the generation of novel genetic forms or evolution. The most obvious consequence of

significant gene flow is in uniting local populations. Under such conditions, immigrants will provide a continual source of ancestral or parental alleles since these are by far the most common. This stabilizing inter-mixing will likely maintain similar allele frequencies in local populations, overshadowing rare mutational events and reducing or eliminating local effects of genetic drift. Conversely, prolonged disruption or barrier to gene flow may result in local differentiation. Gene flow acts as a creative force when novel advantageous alleles are established in a population and subsequently spread to nearby locales. The spread of such alleles may be due to rare movement among demes, or as a result of the removal of previous barriers. This may occur, for example, if a rise in ocean level flooded a narrow isthmus and connected two previously isolated marine populations. Gene flow may also act as a creative force when one or a few individuals colonize new or unoccupied habitat. Under favourable circumstances, such a new population may expand, allowing substantial divergence to accumulate rapidly through drift without a restriction on total population size and has been referred to as the founder effect (reviewed in Slatkin, 1985). The shifting balance theory (reviewed in Hartl and Clark, 1989) describes how isolation could result in novel combinations of genes resulting in higher adaptive fitness. Subsequently, gene flow from such local populations could spread these novel combinations to other populations.

If two populations in Hardy-Weinberg equilibrium (reviewed in Hartl and Clark, 1989) having differing allele frequencies are combined, the frequency of heterozygotes will always be lower than that observed after one generation of random mating restores Hardy-Weinberg equilibrium. Known as the Wahlund principle, this effect can be utilized

to detect any departure from random mating, which can be expressed as an inbreeding coefficient or fixation index,  $F$ . Wright (1951) incorporated these ideas into a hierarchical approach to assess departures from random panmictic mating. The first level assesses the average heterozygosity of an individual relative to its subpopulation ( $F_{IS}$ ), the second assesses the heterozygosity of an individual relative to the total population ( $F_{IT}$ ), the third the heterozygosity of a subpopulation relative to the total population ( $F_{ST}$ ). These levels can be related by the expression  $[F_{ST} = (F_{IT} - F_{IS}) / (1 - F_{IS})]$ . Commonly referred to as  $F$  statistics, these parameters are a standard means of assessing the geographic structure of populations.

In order to understand how these processes are operating in natural populations, a model of the pattern of migration is required. The simplest model of migration is the Fisher-Wright island model. A large population is assumed to be split into many discrete subpopulations with equal probability of migration between all subsets with the additional assumptions of no mutation, no natural selection, and subpopulations sufficiently numerous to negate, overall, the effects of genetic drift. Wright (1951) further considered such a model in the context of finite population size allowing for genetic drift by allowing  $2N$  individuals to produce infinite gametes leading to  $2N$  individuals in the next generation and obtained an estimate of the fixation index  $F$ , in terms of the migration rate,  $m$ , and the effective population size,  $N$ , where  $F = 1 / (4Nm + 1)$ . In this way,  $F$  statistics can be used to obtain estimates of gene flow ( $Nm$ ) on any given scale. Wright (1943) also considered the case of organisms in which dispersal is considerably restricted. In this case it is highly unlikely that the assumption of equally probable random mating

among all individuals is satisfied. If mating occurs only between nearby groups or individuals, the probability of differentiation in the entire population increases and Wright (1943) referred to this concept as isolation by distance. In fact, parapatric speciation can develop without the presence of barriers to gene flow, as long as gene flow occurs only between adjacent populations, (Endler, 1973). The stepping stone model (Kimura and Weiss, 1964) was developed for a series of discrete populations with exchange only between adjacent populations and predicts spatial differentiation will arise. It therefore follows that species with limited dispersal capabilities will eventually develop differences among local populations

### *Dispersal in Marine Invertebrates*

Development in many marine species includes a pelagic feeding larval stage followed by settlement and metamorphosis into the adult form (Strathmann, 1987). Fecundity in these organisms is commonly millions of eggs per female and therefore this pelagic form represents a potent means of dispersal for many marine organisms (Hedgecock, 1986), although to what extent this potential is realized remains controversial (Avise, 1994).

In many cases, species have presumably undergone some change, possibly affecting early development or settlement patterns, resulting in the loss of feeding functions or even loss of the entire pelagic phase. This loss would result in decreased dispersal and allow greater differentiation among populations due to local genetic drift and natural selection. Changes such as the loss of feeding in larvae are usually paralleled by an

increase in egg size which would be a necessary replacement to the feeding mechanisms of nutrient acquisition, and may also result in greatly reduced times between fertilization and adult metamorphosis, (Strathmann, 1987). A common assumption is that as the length of the pelagic larval phase decreases, the dispersal potential of the organism decreases, and thus the degree of local differentiation increases (Crisp, 1978). Berger (1973) assessed dispersal in three gastropod species from the genus, *Littorina* and found that two species which lack pelagic larval forms showed higher levels of population differentiation than a congener possessing a pelagic larval stage. Janson (1987) also examined genetic differentiation using allozymes in populations of two *Littorina* species along the Swedish west coast and found a linear relationship between geographic and genetic distance in *L. saxatilis* which lacks a pelagic form, compared to no such correlation in *L. littorea* which possesses a larval form. In a similar study, Campbell (1978) found strong differentiation among 12 populations of the gastropod (*Nucella*) *lamellosa* along the coast of North America, a species which also lacks a pelagic larval phase. However, no clear correlation between divergence and geographic distance or with environmental factors was detected. Campbell (1978) suggested that random genetic drift among local populations appears to explain the genetic structure in this species. Therefore, sedentary species which lack pelagic larvae almost always show significant differentiation on a small geographic scale (Burton, 1983; Jablonski, 1986).

**In contrast to this pattern of local differentiation in species with benthic development and limited motility, species which possess a pelagic larval phase, and thus a high dispersal potential, show a great variability in the degree of genetic differentiation**



(Burton, 1983; Avise, 1994). Allozyme data (reviewed in Hedgecock, 1986) for certain Atlantic species exhibited significant differentiation among local populations, despite evidence of pelagic dispersal. *Homarus americanus*, the American lobster, possesses a pelagic larvae lasting two to eight weeks yet shows genetic differences between the Gulf of St. Lawrence and the Atlantic ocean (Tracy et al., 1975; reviewed in Avise, 1994). On the other hand, southwestern Pacific populations of the red rock lobster, *Jasus edwardsii*, (Ovenden et al., 1992) or eastern Pacific populations of the sea urchin *Strongylocentrotus purpuratus* (Palumbi and Wilson, 1990), exhibit a near absence of genetic variation over a range of thousands of kilometres. Thus the underlying causes of differentiation and speciation in marine species with high dispersal potential are only beginning to emerge and include selection due to local adaptation, changes in physical barriers, as well as the evolution of gamete interactions (Palumbi, 1993). These results also reflect the strong influences that life history and habitat characteristics have on dispersal capabilities (reviewed in Avise, 1994).

### *Life History of Study Species*

*Cucumaria miniata* is a benthic filter feeding species occupying crevasses or burrowing under rocks for protection. Often the only visible or accessible part of the organism is the oral tentacles and upon settlement, is essentially a sessile organism. This species is common in shallow waters with current flow but is known to occur to at least 200 m in depth. Spawning occurs in spring and eggs averaging 400  $\mu\text{m}$  in diameter are

positively buoyant. After fertilization, nonfeeding lecithotrophic larvae develop in 2 days with the entire pelagic phase lasting up to 14 days.

*Cucumaria pseudocurata* is strictly an intertidal organism, occupying a zone from one to several feet above the mean low tide. This species is found on the exposed outer coast and occupies a series of microhabitats created by other organisms on rocky shores. The most common of these is among the byssal attachment threads of the mussel *Mytilus californianus*, where the sea cucumbers often form dense aggregations against the rock. Other associations seen are dense mats packed around the base of surf grass *Phyllospadix scouleri* or similarly with coralline algae, which offer less obvious protection from predators, but could be important as a wave break. It is difficult to assess movement of adult individuals in their natural setting but specimens held on water tables would seek out some type of protection or cover and were capable of travelling more than a metre in a 10 hour period (pers. obs.). How far adults might move in a lifetime in their natural habitat is unclear but, notwithstanding dispersal by storms, likely not more than a few metres. Spawning in this species occurs between December and February. *C. pseudocurata* eggs are 1 mm in diameter. Fertilization is external yet the female does not release her eggs but instead broods the young embryos underneath her body against the substrate for approximately one month. Brood size varies between 10 and > 100, according to female body mass (Rutherford, 1977b). The embryos then hatch and juveniles crawl away. Due to the lack of pelagic dispersal and the limited potential movement of adults, populations of *C. pseudocurata* would be expected to exhibit some

geographic structure since any break in the rocky shore habitat could present a substantial barrier to migration and gene flow.

The phylogenetic analysis presented in Chapter I demonstrated a close relationship between *C. pseudocurata* and its' northern sister species *C. vegae*. The distance estimate (2.5%) based on the COI gene fragment sequence was only twice that of the intraspecific estimates (up to 1.3%). Thus both the phylogenetic tree topology and the shortness of branch lengths, relative to intraspecific estimates, raise some question as to the separate species status of these two organisms.

*C. vegae* and *C. pseudocurata* have been distinguished on the basis of geographical, ecological, and morphological differences. The range occupied by *C. vegae* extends northward of *C. pseudocurata*, with an uncertain degree of overlap in central British Columbia, through to the Aleutian islands, (P. Lambert, pers. commun.). *C. vegae* is found in more protected waters than is *C. pseudocurata*, and is found, predominantly, in dense groupings in cracks and depressions of intertidal rock surfaces, without the typical faunal associations of *C. pseudocurata*. Calcareous ossicles from dorsal skin samples also differ among the two species. Those of *C. vegae* are spectacle shaped rods whereas *C. pseudocurata* ossicles are larger rounded plates with numerous holes (Lambert, 1985). Furthermore, ossicles from the Queen Charlotte Islands/Haida Gwaii and the central B.C. coast have been described as intermediate between *C. vegae* and *C. pseudocurata* (Lambert, pers. commun.).

Molecular genetic analysis provides a method to improve our understanding of how micro-evolutionary processes actually occur in nature. The present study was an

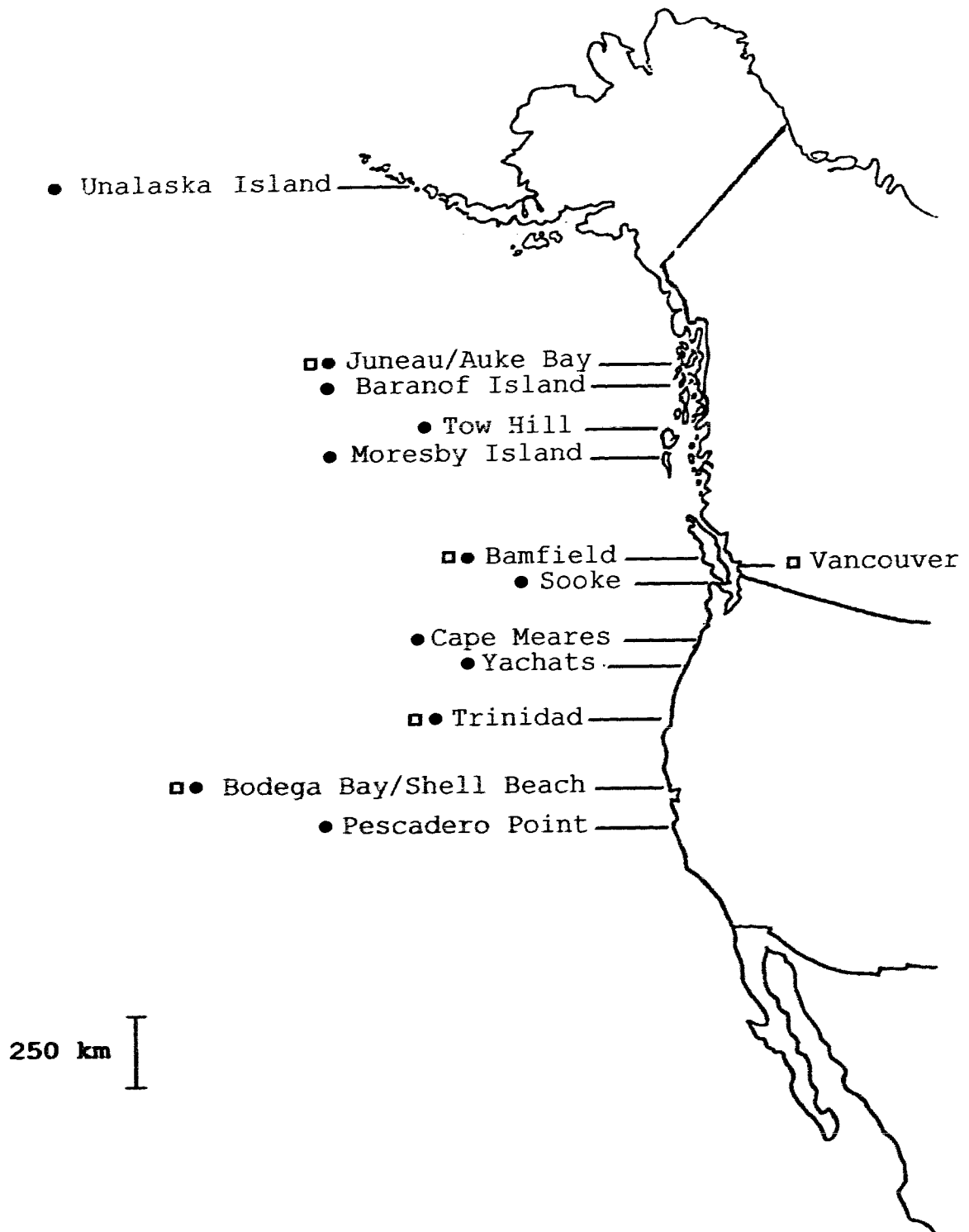
attempt to analyze the genetic structure of sea cucumber populations, as a function of mode of development and life history, along the west coast of North America and to examine the geographic structure and genetic distinction between the externally indistinguishable brooding species, *C. vegae* and *C. pseudocurata*.

### **Materials and Methods**

#### ***Sampling of Populations***

Populations were sampled at several locations along the west coast of North America (Figure 7). In the case of the brooding species, all population samples were obtained from exposed outer coast locations and are therefore referred to as *C. pseudocurata*. Specimens obtained from protected sites in Alaska were confirmed as *C. vegae* specimens by P. Lambert (pers. commun.) based on ossicle morphology and were included in the phylogenetic analysis. To sample *C. pseudocurata* populations, a transect was established passing through presumed suitable habitat between one and five feet above the low tide reference (Minister of Supply and Services, Government of Canada, 1995). Five individuals were taken at random from a 100 cm<sup>2</sup> quadrat at 10 m intervals along the transect up to a total of 30 times, if possible. It is worth noting that *C. pseudocurata* occupies a somewhat protected habitat created by other species which must often be removed in order to gain access to the sea cucumbers. The above sampling strategy was utilized in order to minimize the impact of sampling on the surrounding

**Figure 7. Population sample site map. Sample locations for *C. miniata* are indicated by open boxes (□). Sample locations for *C. pseudocurata* are indicated by closed circles (●).**



community. Once obtained, specimens were placed in containers of sea water for transport to the laboratory. On occasions when it was not possible to bring live animals back to the lab, specimens were placed immediately on dry ice at the field site. All specimens obtained from a particular location were subsequently placed in individual plastic tubes in a common storage box. Individual specimens were drawn at random from this pool for sequence analysis.

*C. miniata* populations were sampled at Auke Bay, Alaska, Diana Island, B.C., and Trinidad Head, California, with two smaller samples obtained at Vancouver B.C. and Bodega Bay, California (Figure 7). A transect was established by compass heading, and a piece of tissue was removed from an individual every 10 m, if present, until a minimum of 30 specimens had been obtained. At Kirby Point on Diana Island, the transect line was laid intertidally between the zero and four foot level with reference to published tide tables (Minister of Supply and Services, Government of Canada, 1995). This site is a cobble beach and sea cucumbers were obtained in the intertidal zone by overturning rocks. A small piece of body wall including a few tube feet was then removed from the organism using a scalpel and the sample was placed on dry ice as soon as possible. In order to prevent cross contamination of samples, scalpel blades were used only once. At the other collection sites, samples were obtained subtidally by SCUBA. In these cases, a pair of needle nosed pliers was used to remove a piece of oral tentacle from a feeding individual. Pliers were then rinsed extensively and visually inspected to ensure no tissue was still present prior to taking the next sample. Such a procedure was necessary since this species often settles in the cracks of large slabs of rock or burrows underneath large

boulders, with the oral tentacles being the only exposed or visible portion of the organism. Populations were also sampled at Vancouver, B.C., and Bodega Bay, California, but only six and eight specimens, respectively, were recovered.

### *Primers and PCR Amplification*

In order to effectively examine genetic diversity at the population level, it was first necessary to identify a region of high nucleotide variability in the mtDNA genome of these sea cucumbers. A strategy was developed based on published sea urchin mtDNA sequence (Jacobs et al. 1988) since previous work had demonstrated a conserved gene order between urchins and sea cucumbers (Smith et al., 1993). Studies with vertebrate species have demonstrated that the regions flanking the control region exhibit the highest levels of substitution observed and are suitable for population level studies (for example, Cann et al., 1984; J. Brown et al., 1993; Bowers et al., 1994; Stanley et al., 1996; T. Burg, pers. commun.). Although much reduced in size in sea urchins compared to vertebrates, the control region appeared to be a region with the potential for variation. In urchins this small domain is also flanked by a cluster of tRNA genes. Since gene boundaries may also provide opportunities for high rates of substitution, a characterization of this region in sea cucumbers was undertaken. Oligonucleotide primers were designed to anneal at a conserved region at the 3' end of the small ribosomal RNA (srRNA) gene and at the anticodon loop of several tRNA genes (Table 8). The combination of the srRNA gene



## TRANSFER RNA GENE ORDER PRIMERS

PRIMER	SEQUENCE (5'-3')	COMMENTS
12Sb	GATCGAGGTGCAGCTAATAAG	708-728 of <i>S. purpuratus</i> <sup>1</sup>
12SCUC	ATGTACCGGAAGGTGCCTC	3' End of srRNA Gene
UAS1	GATCCTTGAGTACCATCC	360-377: CmsrRNA-ND1 <sup>2</sup>
ALAc <sup>3</sup>	ATATTTTGGAAATCACAG	847-830: CmsrRNA-ND1
CYSc	TGCAATTAATATGTGAA	958-940: CmsrRNA-ND1
METBc	CATTCTTGGGATATGAGCC	1095-1077: CmsrRNA-ND1
ND1c	TGCATATAACCCAUYACCT	2268-2250 of <i>S. purpuratus</i>
TRP	GTAACTAAACTGAAAGCCTTCAAAG	Faces ATPase6 Gene
A6c	TGTGAGGTTGGCTGCTAGTCG	From 3' End of pCmhd23 <sup>4</sup>
VALc	AAGCGACTCTTTTACACAG	Faces CO1
COI3'	GGCTACACTCCAAGGGTC	6755-6773 of <i>S. purpuratus</i>
COI3'B	GTCAAAGGTACAATTCTTC	7010-7029 of <i>S. purpuratus</i>
CM21R	CGGGCTTAACCACTTTTC	253-270 of pCMHD23
CM21RCc	GAACATCACTCCTACC	75-58 of pCMHD23
CMARG	CCAACCCATAATCCATC	3' of tRNA ARG
CM21RC2c	GGTTTTCCATAGAGTTC	From CM21RCc Sequence
CM21RC3c	GGGGTTATCTTTACTC	From CM21RCc Sequence

1. Numbers refer to *S. purpuratus* mitochondrial genome position (Jacobs et al., 1989).
2. Numbers refer to *C. miniata* srRNA-ND1 gene fragment (Figure 20).
3. A lower case c is used to denote primers that anneal to the sense strand and therefore prime synthesis of the complement strand.
4. Numbers refer to *C. miniata* HindIII restriction fragment clone pCmhd23 (Figure 23).

primer, 12Scuc, and one of these tRNA gene primers, Metb, were used to PCR amplify a 1000 bp fragment that included the putative control region. The thermal cycle profile was 95<sup>o</sup>C, for 30 sec.; 50<sup>o</sup>C, 30 sec.; 72<sup>o</sup>C, 60 sec., for 25 cycles followed by a final cycle with a 5 minute extension step. For the purposes of comparison of detectable variation, conserved COI primers (Palumbi et al., 1991) were also used to amplify and sequence approximately 600 bp fragment from this gene using the same thermal cycle profile.

#### *Fragment Purification and DNA Sequencing*

PCR amplified DNA fragments were separated by agarose gel electrophoresis and purified using spin columns as in Chapter I. Sequencing for population level analysis utilized the 12Scuc and UAS1 (Table 8). Both these primers are oriented to sequence the sense strand. Automated thermal cycle sequencing was utilized for most samples according to the manufacturer's instructions (ABI), as in Chapter I. Other primers listed in Table 8 were used to determine complete sequence of the PCR fragment (discussed in Chapter III).

#### *Statistical Analysis*

Initial investigations of genetic structure were performed using the phylogenetic reconstruction methods described in Chapter I. Sequences were aligned using the program ESEE (Cabot and Beckenbach, 1989). The program MEGA, (Kumar et al., 1993) was

used to calculate a pairwise distance matrix among all haplotypes. Neighbor-joining trees (Saitou and Nei, 1987) based on this distance matrix were also constructed using MEGA. An alternative approach to phylogenetic reconstruction for the analysis of population genetics is that of a minimum spanning tree (MST). Rather than considering taxa as terminal nonrecombining units, an MST is constructed by computing the minimum length network required to connect all taxa but allows taxa to be assigned as internal nodes in the tree and also allows multifurcating events. This approach may more accurately represent relationships among populations (Excoffier et al., 1992; Stanley et al., 1996), since a common mtDNA haplotype may be the progenitor of several new variants with all such haplotypes present in the population at the time of sampling. Phylogenetic reconstruction assumes all taxa examined are terminal nodes, often assumed to arise through a bifurcating process from common ancestral types, and therefore does not allow taxa to be assigned to internal nodes. Minimum spanning trees were produced using the program NTSYS (Rohlf, 1990) with a pairwise distance matrix as input.

Using nucleotide sequence data it is possible to determine the precise nucleotide substitutions or changes identifying unique haplotypes. Such information can be used to estimate the overall patterns of nucleotide change occurring within populations and can subsequently be used to assess models of substitution compared to the observed patterns. Observed changes were superimposed onto the neighbor joining tree to obtain estimates of the number and types of changes. Changes which could be mapped to particular branch points were counted once. Other changes were found in several sequences in separated portions of the tree and were counted at each separate occurrence

*Population Structure and Gene Flow*

Fixation indices (F statistics) were originally derived by Wright (1951) in order to partition the genetic variation observed in populations into various components. These F statistics are applicable to mtDNA under the assumption of uniparental, nonrecombining inheritance. Hudson et al. (1992) derived an expression to estimate F<sub>ST</sub> values based on sequence data:

$$F_{ST} = 1 - H_w / H_b$$

where  $H_w$  is the nucleotide variation among all individuals, calculated pairwise, within a deme, and  $H_b$  is the average variation between demes. The computer program Heap Big, (Steve Palumbi, pers. commun., 1995) calculates this  $F_{ST}$  value and compares it to values obtained by randomly assigning sequences to demes using a Monte Carlo simulation in order to assign a P value for the observed  $F_{ST}$ . It has been shown (Hudson et al., 1992) that  $F_{ST}$  can be used to estimate gene flow and, in the case of mtDNA,  $N_e m = 1/2[1/F_{ST} - 1]$ , where  $N_e$  is the effective number of females and  $m$  is the migration rate.

A second approach derived by Excoffier et al., (1992) was also used to assess the detected variance within the data set using the authors' computer program, the analysis of molecular variance (AMOVA). Distance values between all haplotypes are computed and used to construct a Euclidian squared distance matrix. Thus, the distance values between haplotypes are treated as a distribution of estimates of genetic diversity which can be subjected to traditional variance analysis. This approach therefore does not depend on any assumptions concerning the model of evolutionary changes, and may

represent the most conservative approach to analyzing population structure. AMOVA will partition the variance of such a sampling procedure into three components analogous to Wright's F statistics (1951): the correlation of variance among a set of defined groups compared to the total population ( $\phi_{CT}$ ), the correlation among subpopulations compared to the defined group ( $\phi_{SC}$ ), and the correlation within subpopulations compared to the total population ( $\phi_{ST}$ ). The program will also calculate pairwise  $\phi_{ST}$  estimates among all populations and perform a randomization procedure to assess the significance of population structure.

Since one of the assumptions of these models is equal populations and sample sizes, analysis was restricted to the eight largest samples, which were of approximately equal size.

### *Ossicle Morphology*

In order to confirm species identity of *C. pseudocurata* versus *C. vegae*, ossicles were examined from representative specimens from each sample site. A small skin sample was removed from the mid-dorsal region and dissolved in 10% bleach for 10 minutes and rinsed with water as described by Lambert (1985). Ossicles were then viewed under a microscope (1000x) and drawn.

## Results

### *Site Description*

As mentioned, it was often necessary to remove other co-occurring fauna since *Cucumaria pseudocurata* is opportunistic in occupying protected microhabitats created by numerous rock adhering organisms. One such species is the mollusc, *Mytilus californianus*, beneath which, against the rock was the most commonly observed habitat of *C. pseudocurata*, and is the most well known habitat of this organism (Ricketts and Calvin, 1968). *C. pseudocurata* was also found crowded tightly around the base of surf grass *Phyllospadix scouleri*, particularly on extended intertidal benches of rock such as those at First Beach, Vancouver Island, British Columbia, and near Sitka, Alaska, or similarly associated with coralline algae of the genus *Corallina*. The least common association seen was at Tow Hill, B.C., where *C. pseudocurata* was found in crevices created between barnacles on the rock. Both *Semibalanus cariosus*, the thatched barnacle, and *Pollicipes polymerus*, referred to as either the leaf or goose barnacle, appeared to provide suitable habitat.

Population density of *C. pseudocurata* varied considerably among sites. *C. pseudocurata* is typically found in dense aggregations (Rutherford, 1973) and densities greater than 1000/m<sup>2</sup> were observed at Diana Island, First Beach, Cape Meares, and Yachats. At the Moresby Island, Pescadero Point, Trinidad Head and Shell Beach sites, organisms were still abundant but densities were an order of magnitude lower and no extensive mats were observed as was the case for the sites mentioned above. At Tow

Hill, and Baranof Island, densities were lower still. At these two sites, only 17 and 23 specimens, respectively, were collected. At Tow Hill in particular, there is only a small area of suitable rock habitat surrounded by extensive sand beaches and the total population size is likely extremely small in comparison to sites such as Diana Island and Yachats with extensive stretches of rocky shoreline. Population density was much lower at Tow Hill than at other sites (pers. obs.) and considerable effort was put into obtaining the 17 specimens collected. At the Baranof Island sample site, small offshore islets provide a natural wave break and, although not examined directly, are likely to constitute a much more extensive population site than that directly sampled on Baranof Island. DNA sequence was obtained from all specimens from the Tow Hill and Baranof Island sites. These population samples were thus not subjected to the randomization described in the Materials and Methods section for all other sites. This involved intensive sampling from numerous transect points followed by combining all specimens into a single large pool from which individuals were drawn at random for sequence analysis. Therefore, the smaller population samples may be subject to greater sampling error.

### *Preliminary Analysis*

An initial analysis of genetic diversity in *C. miniata* and *C. pseudocurata* populations was undertaken using a portion of the COI gene as described in Chapter I. This initial assessment was undertaken to provide an estimate of intraspecific variation prior to phylogenetic analysis. Nine haplotypes were detected among 23 *C. pseudocurata* specimens and distance estimates ranged between 0.16% and 1.13%.

Among nine *C. miniata* specimens, six haplotypes were detected and varied between 0.15% and 0.61%. Distance estimates between the two species ranged between 10.8% and 12.2% for this region of the mitochondrial genome.

### *MtDNA Amplification*

An amplification product of approximately 1200 bp was obtained from both species using the srRNA (12Sb) and the methionine tRNA (METb) oligonucleotide primer pair. DNA sequence analysis demonstrated conserved nucleotide blocks corresponding to the 3' end of the srRNA gene. Based on this conserved sequence, another oligonucleotide primer, 12Scuc, was designed for sequence analysis beyond the end of the srRNA gene (Table 8). Sequencing and subsequent analysis of the amplified fragment provided some unexpected results regarding tRNA gene arrangement and will be discussed in Chapter III. Manual sequence alignment (Figure 8), using the computer program ESEE (Cabot and Beckenbach, 1989), indicated this region to be highly divergent between the two species with a distance estimate of 24%, excluding an apparent 20 nucleotide insertion/deletion event (indel). This region accumulated changes at more than double the rate of the COI gene, based on both inter- and intraspecific comparisons, and appeared to be a suitable marker for population level analysis.

### *Sequence Analysis: C. pseudocurata*

Initially, sequence was obtained from 180 *C. pseudocurata* specimens from the eight major locations. In order to reduce distances between locations and gain a more



**Figure 8.** Alignment of mtDNA region utilized for population study. Sequences were aligned using Clustal IV and manually adjusted to maximize identity. The region used in the population study indicated in capitals, flanking regions in lower case. Asterisk below aligned sequences indicates nucleotide identity. The 12Scuc primer sequence is shown underlined and labelled above. The threonine gene sequence is indicated in bold with dashes enclosed in brackets above.

[ --12Sscu Primer-- ]

C.m. tacacatcgcccgctcactctcctcaaaagagagaaaagtcgtaacatagtagatataccggaaaggtgcctcraaaaaa 80  
C.p. tacacatcgcccgctcactctcctcaaaagagagaaaagtcgtaacatagtagccgtaccggaaaggtgcctcraaaaaa  
\*\*\*\*\*

[ ----- ]

C.m. TATTAAAAATTAAACACAGAAACTTCCAAAGTCTCAGTTCAGGTTTACACCTATAATAATT-CGCTATAAGTAGGATACAA 160  
C.p. TATTAAAAATTAAATATTGAGGCTTCCCAAGCCTCAAAAATGGGCTGT-CCCGTAATAATTATGCTATAAGTAGGAAACAA  
\*\*\*\*\*

-----THR tRNA----- ]

C.m. ACTTAGTCTTGTAAACTAAGTTAGAAAGTTAAACTCTTTTCATATAGCTAACAAATAAAGCCCTACCCAGCCCTCCGCTC 240  
C.p. ACTTAGTCTTGTAAACTAAGTTAGAAAGTTAAACTCTTTTCGATAGCTAACAAATAAAGTCAAATAAAGCTTACCATTTT  
\*\*\*\*\*

C.m. AACCTAAACTGAGTAAAAAACCCTTCAAACTTAACATATCCCAAAACCTTAATCATCCCTTCCCCACTTAAGAAAAGAT 320  
C.p. AGACAAAAATAAACATAAAAAGCCTTAAA-TTTATCTTTTACAAATAACAGATCATTCATTTTATCGCTTAAAGAAAAT  
\* \* \* \* \*

C.m. TAAAAACTT-----GGGCTTCCCCCCTA--TATAACCCCCCTTCCATGATCCTTGAGTACCATCC 400  
C.p. TTAGAACTTTAAAAATAAAACTCAGAAATTTGGGCCCTCCCCCTAGCCCTTTACCCCCCTTCCAAGATCCTTGAGTACCATCC  
\* \* \* \* \*

C.m. GGGCTTAACCACTTTCAAATAGTCCTTTTCTTACTTTTGGGGCTCAATAATCATTTATTCCCTCTCCTCGTGATTCCA 480  
C.p. AGGCTTATCCCTTTACCTATAGTCCTTTTCTTATACTTTGGGGGGTTA-CATCATTTATCCCTATCTTCCGGGATTTCCA  
\*\*\*\*\*

C.m. TATTAACCTTGctattattcttactctctgtctcttaccctgcacataacctatataattcacttytgagaaagagag 560  
C.p. C-TTAACTTGtcttcttccatgtctctctgtctcttaccctgcacacgtacctatataattcactt-tggagaaagtagg  
\*\*\*\*\* \* \* \* \* \*

Capitals indicate region used in study. Asterisks indicate conserved positions.

complete understanding of the patterns of divergence or genetic exchange, sequences from intermediate locations between the eight major sample sites were included. Numbers of sequences from these locations were considerably lower than the original seven sites. In total, 411 nucleotides of mtDNA sequence was obtained from 223 *C. pseudocurata* specimens from 11 locations listed in Figure 7. DNA sequences were entered and aligned in ESEE. This region of the genome is AT rich (68.6%) and exhibits some bias against G which accounted for only 11.5% of sites. Figure 9 shows the variability across the entire region in nonoverlapping blocks of ten nucleotides. Numbers of substitutions detected in a ten nucleotide block varied between zero and six. Several small blocks appeared more variable than others but generally substitutions were distributed throughout the region. In total 72 variable sites were detected (Figure 10), of which 51 were parsimoniously informative. A total of 63 unique mtDNA haplotypes were identified among the 223 *C. pseudocurata* specimens sequenced. The reduction in detected haplotypes compared to the number of sequences examined was due to a single or very few haplotypes predominating at each sample site, with the remainder of the population being represented by numerous considerably more rare haplotypes. A complete list of haplotype frequencies is shown in Table 9. At six of the eight major sites, a single haplotype accounts for greater than 50% of the specimens examined. The two exceptions are Baranof Island AK., and Yachats OR. At both these sites, three common haplotypes were found in approximately equal frequencies (Table 9).

**Figure 9. Variability of sites in region sequenced for population study. The number of variable sites in non-overlapping blocks of 10 nucleotides are shown across the entire region used for both *C. pseudocurata* and *C. miniata*. Nucleotides 330-350 are not applicable (NA) for *C. miniata* due to an insertion/deletion in the aligned sequences (see Figure 8). The threonine gene sequence corresponds to positions 63-128.**

Location	<i>C. pseudocurata</i>	<i>C. miniata</i>
1-10	2	1
11-20	2	4
21-30	3	1
31-40	4	1
41-50	6	4
51-60	1	2
61-70	2	0
71-80	1	2
81-90	0	0
91-100	1	0
101-110	0	0
111-120	2	0
121-130	1	2
131-140	1	3
141-150	2	3
151-160	3	4
161-170	3	2
171-180	0	2
181-190	3	2
191-200	4	1
201-210	0	2
211-220	0	0
221-230	4	1
231-240	1	3
241-250	1	2
251-260	2	3
261-270	5	0
271-280	1	1
281-290	2	2
291-300	0	1
301-310	2	2
311-320	0	0
321-330	1	0
331-340	1	NA
341-350	0	NA
351-360	1	3
361-370	3	0
371-380	1	2
381-390	0	2
391-400	0	1
401-410	1	2
411-420	0	2

**Figure 10. Variable nucleotides detected among *C.pseudocurata* haplotypes. Sequence shown is shared by three *Cucumaria vegae* specimens from Auke Bay, AK (Cva), other labels as in Table 9. A period indicates nucleotide identity. Insertion/deletion differences are only shown for multi-state sites, based on output options from the computer program MEGA (Kumar et al., 1993).**

111111111111111111111111222222222222222233333333334  
1112223333444444556670122344555666888999922223455666667890023566770  
7069378478913457832760902817156235356357824891735134566904563938023

Cva GTGTGTTTACTAGTTGTTGGAAATTAAGTCCCAATACTCTTCTTACAGTAATTCACCTAAGCAGGTTT  
Cvu .....A.....C.....C.....  
BIa ..A.C.....TT..CC.....T.....G.....TC...T...C.  
BIb ..A.C.....TT..CC.....T.....G.....TC...TT..C.  
BIc ..A.C.....CC.....T.....GG.....T...C.  
BIe ..A.C.....G.C.A.....TT..CC.....T.....G.....TC...T...C.  
BIf ..A.C.C.....CC.....TT..CC.....T.....G.....TTC...T...C.  
BIg ..A.C.....TT..CC.....T.....G.....TC...GT...C.  
BIi ..A.C.....CC.....T.....GG.....TC...TT..C.  
BIj ..A.C.....CC.....T.....G.....TC...TT..C.  
BIk ..A.C.....G.C.....TT..CC.....T.....G.....TC...TT..C.  
BIh ..A.C.....G.C..G.....TT..CC.....T.....G.....TC...TT..C.  
THc ..A.C.....GG.....C-.....TT..C.  
THe ..A.C.....G.....C.....-.....TT..C.  
THa ..A.C.....G.....C.....TT..C.  
THf ..A.C.....G.....C.....-.....TT..C.  
THg ..A.C.....C.....TT..C.  
THh ..A.C.....G.....C.....T.....TT..C.  
THd ..A.C.....G.....C.....A.....TT..C.  
THb ..A.C.....G.....C.....G.....TT..C.  
MIc ...C.....C.....C-..GCT..CT.....T...C.  
MIb ...C.....C.....C.....C.....G.....T...C.  
MIa ...C.....C.....C-..G...CT.....T...C.  
MId ...C.....C.....C-..G...CT.....T...C.  
MIe ...C.....C.....C-..G...CT.....T...C.  
MIk ...C.....C.....C.....C.....G...G...T...C.  
SKa A.AGCCC...AG..C...A.....--...GCG..T...C..TG...C.....TT..C.  
SKb A.AGCCC...AG..C...A.....--...GCG..T...C..TG...C.....TT..C.  
SKd A.AGCCC...AG..C...A.....--...GGCG..T...C..TGA.G.C.....TT..C.  
SKe A.AGCCC...AG..C...A.....--...GCG..T...C..TG..G.C.....TT..C.  
BFa A.AGCCC...AAG..C...A.....G.....GCG..T...C..TG...C.....C.  
BFb A.AGCCC...AAG..C...A.....G.....GCG..T...C..TG...C.....CC.  
BFc A.AGCCC...AG..C...A.....--...GCG..T...C..TG...C.....T..CC.  
BFd A.AGCCC...AG..C...A.....--...GCG..T...C..TG...C.....T...C.  
BFe A.AGCCC...AG..C...A.....--...GCG..T...C..TG...C.....T...C.  
BFf A.AGCCC...AG..C...A.....--...GCG..T...C..TG...C.....T...C.  
BFg A.AGCCC...AG..C...A.....--...GCG..T...C..TG...C.....T...C.  
BFh A.AGCCC...AG..C...A.....--...GCG..T...C..TG...G.C.....T...C.  
BFi A.AGCCC...AG..C...A.....G.....GCG..T...C..TG...CC.....C.  
BFj A.AGCCC..CAAG..C...A.....--...GCG..T...C..TG...C.....T...C.  
BFk A.AGCCC..CAAG..C...A.....G.....GGCG..T...C..TG...G.C.....C.  
BFl A.AGCCC..AAG..C...A.....G.....GCG..T...C..TG...C.....C.  
CMA A.AGCCC..AAG..C...A.G.G.....GCG..T...C..TG...C.....T...C.  
CMB A.AGCCC...GA.C...AG...G.A.T...GCG..T...C..TG...C.....A...T...C.  
CMC A.AGCCC...GA.C...AG...G.A.T...GCG..CT...C..TG...C.....T...C.  
CMd A.AGCCC...GA.C...A.....ACT..G.G.T...C..TG...C.....T...C.  
CME A.AGCCC...GA.C...AG...G.A.T...GCG..T...C..TG...C.....T...C.  
YHb A.AGCCC...GA.C...AG...G.A.T...GCG.T...C..TG...C.....T...C.  
YHc A.AGCCC...G.C...AG...G.A.T...GCG.T...C..TG...C.....T...C.  
YHe A.AGCCC...G.CC...AG...G.A.T...GCG.T...C..TG...C.....T...C.  
YHa A.AGCCC...GA.C...AG...G.A.T...GCG.T...C..TG...C.....T...C.  
YHd A.AGCCC...G.C...AG...G.A.T...GCG.T...C..TG...C.....T...C.  
SBa A.AGCCC...GA.C...A.....A.T...G.G.T...C..TG...C.....T...C.  
SBb A.AGCCC...GA.C...A.....A.T...G.G.T...C..TG...C.....G...T...C.  
TRa A.AGCCC..AAG..C...AG...G.....GCG..T...C..TG...C.....C.

11111111111111111111111111222222222222222233333333334  
11122233334444444556670122344555666888999922223455666667890023566770  
7069378478913457832760902817156235356357824891735134566904563938023  
Cva GTGTGTTTACTAGTTGTGGAAATTAAGTCCCAATACTCTTCTTACAGTAATTCACCTAAGCAGGTTT  
PP1 ACAGCCC...AG..C...A.....G.....GCG..T...C..TG..C..C.....G....T.C.  
PP2 A..AGCCC...AG..C...A.....G.....GCG..T...C..TG..CG..C.....G....T.C.  
PP3 ACAGCCC...AG..C...A.....G.....GCG..T...C..TG....C.....T.C.  
PP7 ACAGCCC...AG..C...A.....G.....GCG..T...C..TG..G.C.....T.C.  
PP14 ACAGCCC...AG..C...A.....G.....GCG..T...C..TG....C.....C.  
PP15 A..AGCCC...AG..C...A.....G.....GCG..T...C..TG..G.C.....G....T.C.



Table 9. Frequency and distribution of *C. pseudocurata* mtDNA haplotypes

<u>Location</u>	<u>Latitude</u>	<u>Sample Size</u>	<u>Haplotype</u>	<u>Frequency</u>
<b>Baranof Island, AK.</b>	57 <sup>0</sup> 10' N	23	B1a	6
			B1b	5
			B1c	4
			B1d	2
			B1e	1
			B1f	1
			B1g	1
			B1h	1
			B1i	1
			B1j	1
<b>Tow Hill, Graham Island, B.C.</b>	54 <sup>0</sup> 10' N	17	THa	8
			THb	2
			THc	2
			THd	1
			THe	1
			THf	1
			THg	1
			THh	1
<b>Moresby Island, B.C.</b>	53 <sup>0</sup> 00' N	20	M1a	10
			M1b	5
			M1c	2
			M1d	1
			M1e	1
			M1f	1
<b>Diana Island, B.C.</b>	48 <sup>0</sup> 50' N	26	BFa	14
			BFb	2
			BFc	2
			BFd	2
			BFe	1
			BFf	1
			BFg	1
			BFh	1
			BFi	1
			BFj	1

Table 9. Haplotype frequency distribution: *C. pseudocurata* (cont'd)

Location	Latitude	Sample Size	Haplotype	Frequency
First Beach, Central Vancouver Island, B.C.	48 <sup>0</sup> 48' N	23	BFa	14
			BFb	4
			BFj	3
			BFk	1
			BFl	1
Sooke, Southern Vancouver Island, B.C.	48 <sup>0</sup> 20' N	22	SKa	11
			SKb	9
			SKd	1
			SKe	1
Cape Meares, OR.	45 <sup>0</sup> 30' N	20	CMa	7
			CMb	1
			CMc	2
			CMd	1
			CMe	1
			CMf	1
			YHa*	4
			Bfa*	2
Yachats, OR.	44 <sup>0</sup> 40' N	27	YHa	8
			YHb	7
			YHc	7
			YHd	3
			THE	2
Trinidad, CA.	41 <sup>0</sup> 05' N	13	TRa	13
Shell Beach, CA.	38 <sup>0</sup> 30' N	7	SBa	5
			SBb	1
			SBc	1

Table 9. Haplotype frequency distribution: *C. pseudocurata* (cont'd)

Pescadero Point, CA.	36° 35' N	25	PPa	19
			PPb	2
			PPc	1
			PPd	1
			PPe	1
			PPf	1
<hr/>				
Total for all 11 Sites			60	223

\* The haplotypes Bfa and YHa were also shared by specimens from Cape Meares OR.

*Phylogenetic Analysis: C. pseudocurata*

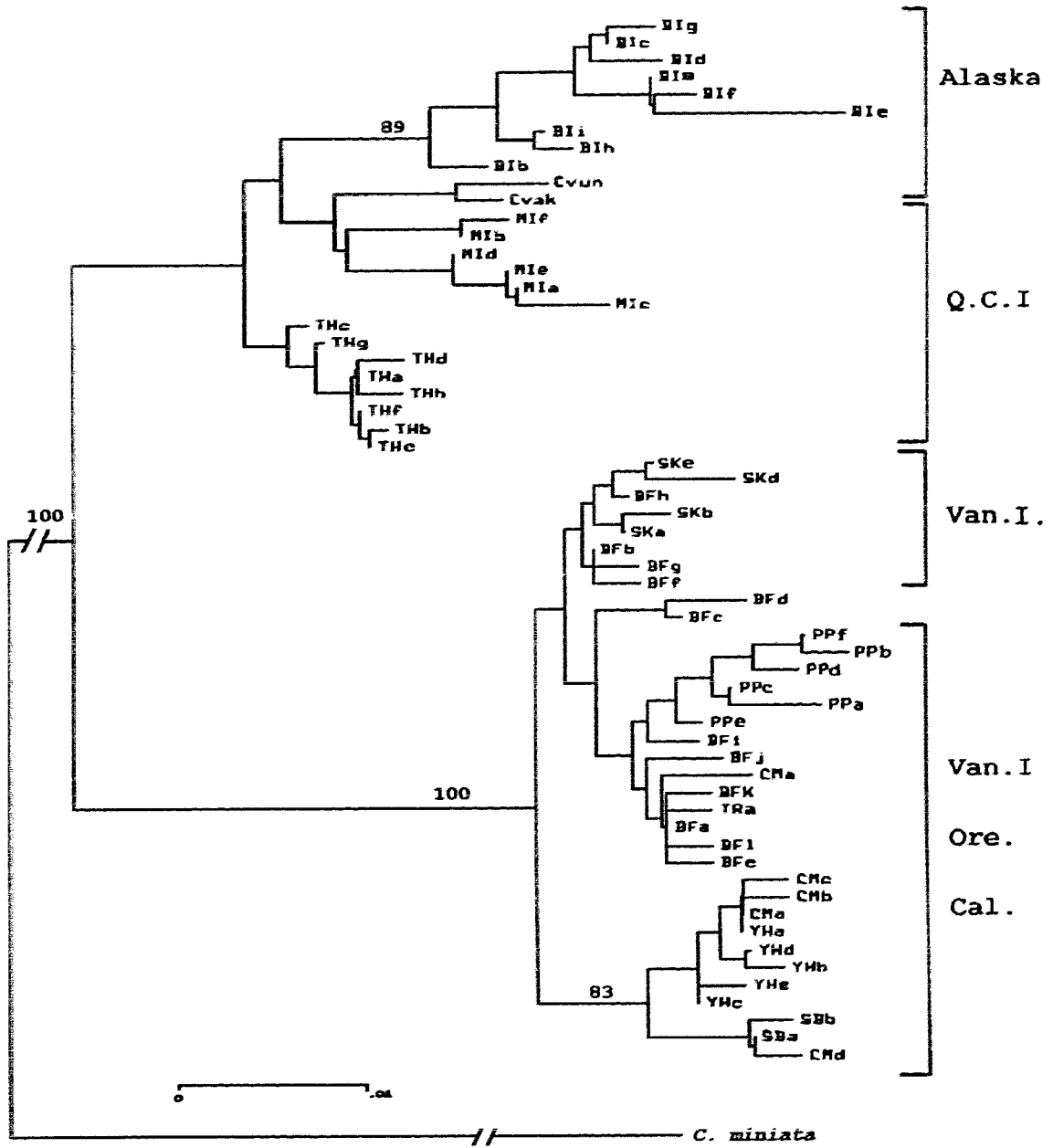
Phylogenetic analysis was performed on the haplotype data sets using the maximum likelihood (Felsenstein, 1981), neighbor-joining (Saitou and Nei, 1987), and maximum parsimony (Swofford, 1992) methods. All three methods produced nearly congruent trees and therefore only the neighbor joining tree is shown (Figure 11).

Pairwise distances were calculated using the Jukes-Cantor model of nucleotide substitution with the computer program MEGA (Kumar et al., 1993) (Appendix A). The mean number of nucleotide differences between haplotypes was 15, corresponding to a distance of 3.8% and ranged between 1 and 28. In most cases, individuals from the same sampling site are clustered together indicating some level of geographic structuring of the populations. The exceptions to this are Cape Meares OR., Bamfield B.C., and Tow Hill B.C. The most significant branch point of the trees separates the three northern sites from a set of more southern sites, with the break point occurring between the Queen Charlotte Islands/Haida Gwaii and Vancouver Island. The mean distance estimate between individuals across this discontinuity was 6.1%.

*Distribution of C. pseudocurata Haplotypes*

On a finer scale, relationships between population sample sites were less clear. In the south three major clades were detected. The first of these three clades contained all haplotypes from Yachats OR., most of the Cape Meares OR., haplotypes, and the only two haplotypes detected at Shell Beach CA. The second clade included haplotypes from Sooke and Bamfield on Vancouver Island B.C. The third clade included haplotypes from

**Figure 11. Neighbor-joining tree for *C. pseudocurata* haplotypes. The tree was based on a distance matrix calculated using the method of Tamura and Nei (1989) (see Appendix A) and was constructed using the computer program MEGA (Kumar et al., 1993). OTU labels indicate sample location (see Figure 7 and Table 9).**



*C. miniata*

several locations along the coast including Bamfield B.C., Cape Meares OR., Trinidad CA., and Pescadero Point CA. The Pescadero Point haplotypes formed a monophyletic group nested within this diverse clade. Sequence data was obtained for 20 individuals from Cape Meares and, except for the two Bamfield area sites, included the only two instances observed of haplotypes shared between locations. Two specimens shared the haplotype designated BFa and four shared the haplotype YHa. Another interesting feature of this site was that specimens were placed into two very distinct clusters in the tree (Figure 11), 9/20 individuals identical or most closely related to BFa and 11/20 identical or closely related to YHa. Therefore, this northern Oregon location has haplotypes similar to the adjacent southern site at Yachats OR., as well as haplotypes similar to the more broadly distributed group found all along the coast. An intermediate site examined was Trinidad Head, CA. Sequences from thirteen randomly chosen specimens, from a total collection of over 100, revealed a single haplotype, differing from haplotype BFa by only 0.5%, again placed within the broadly distributed group of lineages (Figure 11). This site was the only case observed where a single haplotype was apparently fixed in the population. The nearest sample sites in a geographic sense are Yachats OR., to the north, and Shell Beach CA., to the south, both of which fall into the Oregon/California group. The other intermediate site was Shell Beach, CA., between Trinidad Head and Pescadero Point. Sequence was obtained from only eight individuals and only two haplotypes were identified. These haplotypes are nearly identical to a two specimens from Cape Meares and appear closely related to haplotypes found in central Oregon.

Persistence of haplotypes along the coast can be described as follows; one lineage is predominant throughout Oregon, a second lineage is broadly distributed and includes specimens from British Columbia, northern Oregon and was the only haplotype detected in northern California, while a third major lineage is also present on Vancouver Island

Among the three major northern population samples, specimens were clustered into monophyletic groups reflecting sample location site. However, as in the South, there were exceptions to a direct correlation between geographic and genetic distance, particularly when larger geographic distances are involved. The first such instance is the smaller mean genetic distance observed between specimens obtained near Auke Bay, Alaska, and either Moresby Island haplotypes (1.9%) or Tow Hill haplotypes (1.8%) compared to the mean distance between Auke Bay and Baranof Island haplotypes (3.3%). This is despite the fact that Baranof Island is located between Auke Bay and Tow Hill with Moresby Island being even further to the south. It is also worth noting that the Auke Bay location is a protected, calm area unlike the exposed outer coast where *C. pseudocurata* is typically found. In fact, specimens analyzed from this site have been identified as *C. vegae* (P. Lambert) which raises the question of species designations. DNA sequence was obtained from a single specimen of *C. vegae* from Unalaska Island, Alaska, and its' haplotype falls within the Moresby Island cluster in the tree (Figure 11). The species designation of *C. vegae*, Theel 1898 precedes that of *C. pseudocurata* (Deichmann 1938). Interestingly, the two demes with the highest mean intrademe distance were from northern locations (Baranof Island and Moresby Island) and that the



overall mean distance among northern samples was almost twice that of the southern samples.

*Minimum Spanning Trees: C. pseudocurata*

A minimum spanning tree (MST) for *C. pseudocurata* is presented in two parts in Figures 12, and 13. The north/south split seen in the neighbor-joining tree is also evident in the MST, which requires 17 changes to connect haplotypes THg and SKa. It is interesting that these locations are not the geographically closest sample sites between the northern and southern groups, in fact the Bamfield and Moresby Island sites make up the closest geographic connection. As in the neighbor-joining tree, haplotypes from each site tended to be clustered together in the MST. The most notable exceptions to this pattern were once again the Bamfield and Cape Meares samples. The frequency at which haplotypes were detected in the samples is proportional to the area of the circle (the square root of the absolute frequency was used as the radius of the circle). The most common haplotypes detected in population samples were often the centre or hub of radiations connecting to many other rarer haplotypes (Figures 12 and 13). One such example is the haplotype BFa, found at a frequency of 14 with seven unique types most similar to BFa of which two connect to other clusters in the network. Similarly, BIa, THa, MIa, BFb, SKa, YHc, and SBa are all represented by multiple specimens and are all hubs of genetic radiations within the MST (Figures 12 and 13). NTSYS does not incorporate any resampling process to test the significance of nodes in alternative networks and does not incorporate the possibility of reticulating relationships.

Figure 12. Minimum spanning network among northern *C. pseudocurata* haplotypes. Network was constructed using the computer program NTSYS (Rohlf, 1990), based on a distance matrix calculated by the method of Tamura and Nei (1989) (see appendix A). All distances greater than a single change are indicated along branch lengths. The number of times a haplotype was detected is represented as a square root of the radius of the circle. This results in a given haplotype circle with an area proportional to the haplotype frequency. Absolute frequencies are also indicated in brackets within circles. Haplotype labels indicate sample location (see Table 9). A minimum of 17 changes is required to connect Northern and Southern (Figure 13) samples and is depicted as a truncated connection to SKa. Two specimens positively identified as *C. vegae* (P. Lambert, pers. commun.) are included and are labelled CVa (Auke Bay, AK) and CVu (Unalaska Island, AK).

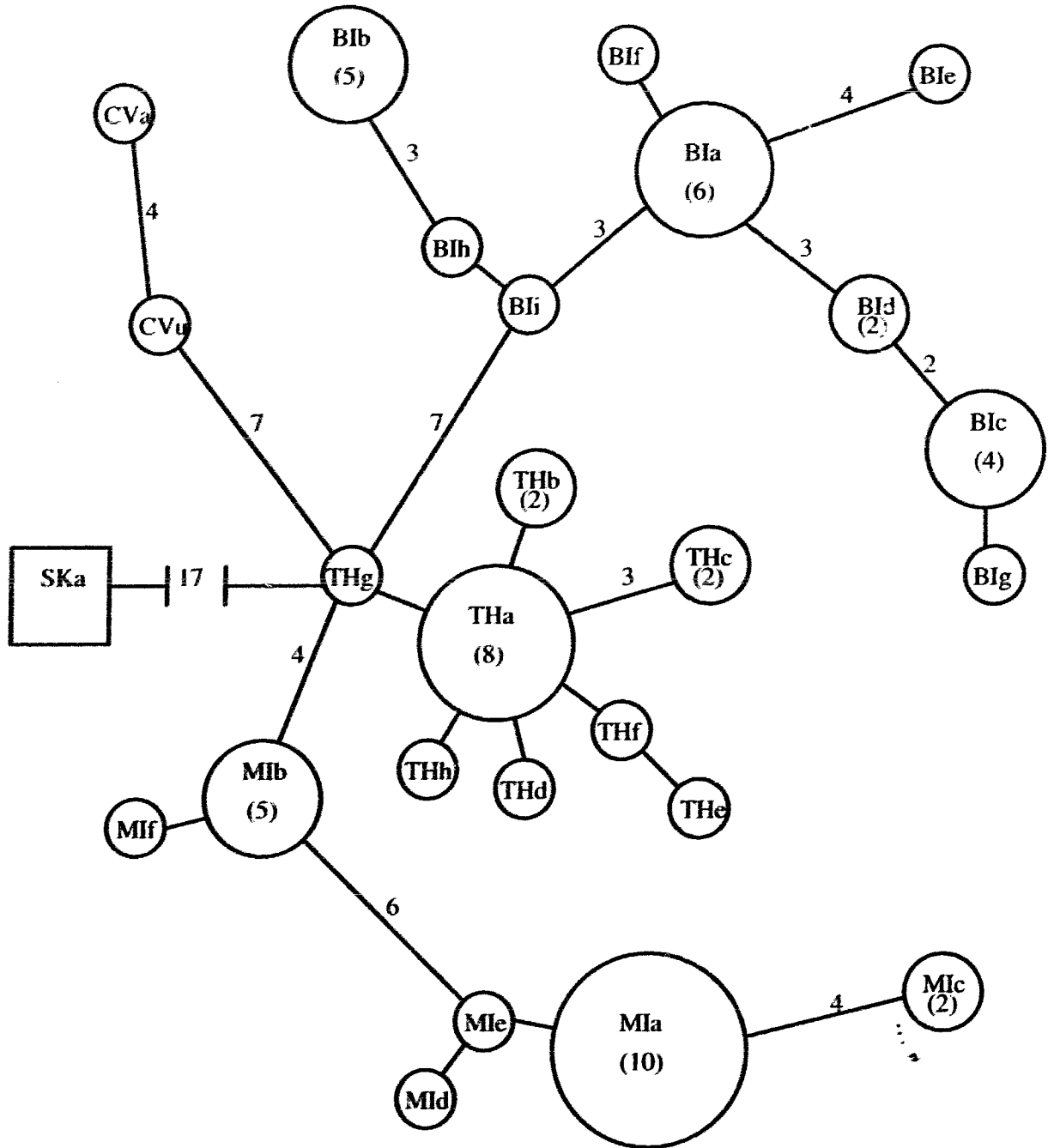
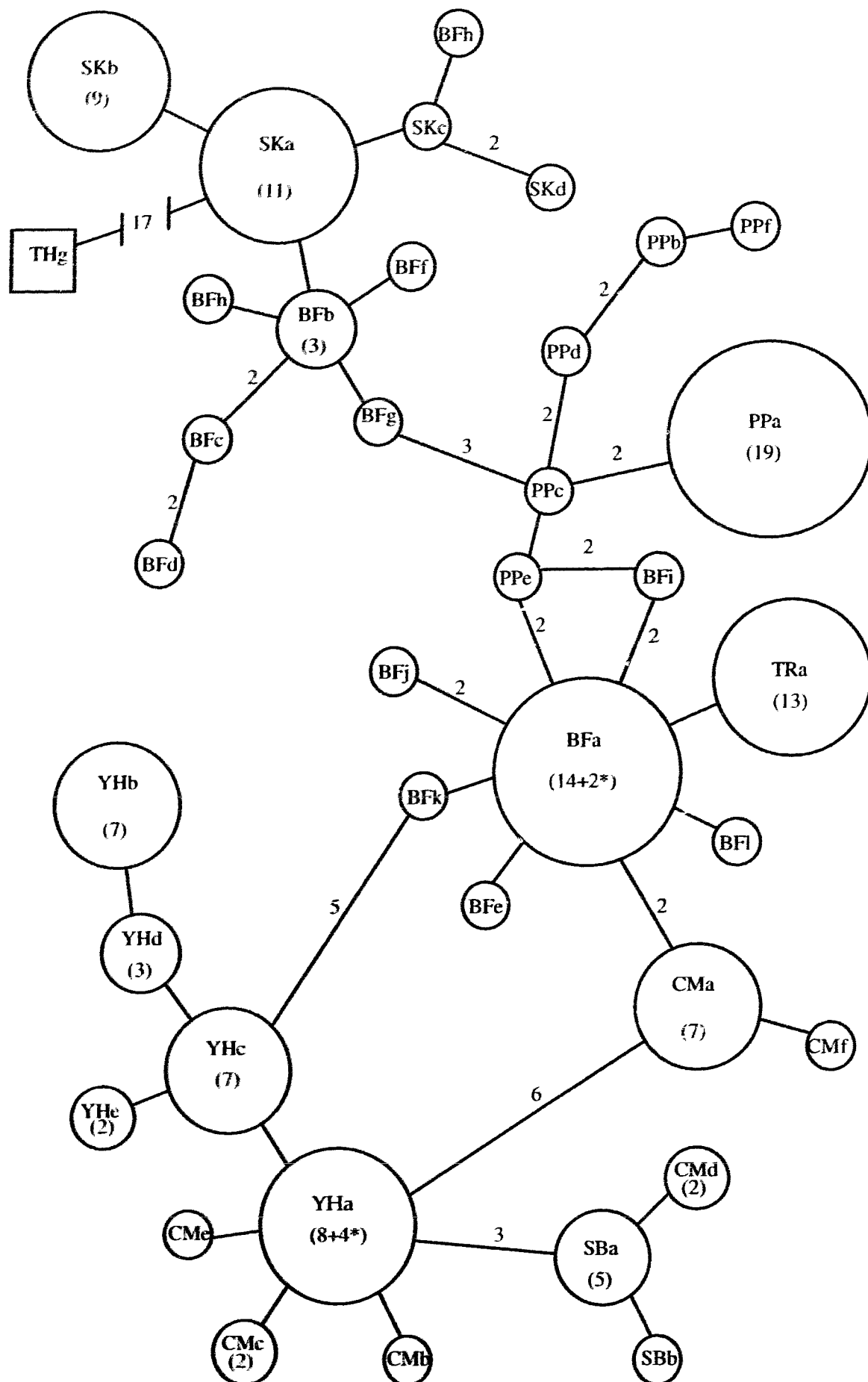


Figure 13. Minimum spanning network among southern *C. pseudocurata* haplotypes.

The network was constructed using the computer program NTSYS (Rohlf, 1990), based on a distance matrix calculated by the method of Tamura and Nei (1989) (see appendix A). All distances greater than a single change are indicated along branch lengths. The number of times a haplotype was detected is represented as a square root of the radius of the circle, resulting in a given haplotype circle having an area proportional to the haplotype frequency. Absolute haplotype frequencies are also indicated in brackets. Haplotype labels indicate sample location (see Table 9). The haplotypes BFa and YHa were also shared by 2 and 4 individuals respectively from Cape Meares, OR and are indicated by an \*. These were the only observed instances of haplotypes shared between sample sites. A minimum of 17 changes is required to connect Southern and Northern (Figure 12) samples and is depicted as a truncated connection to THg.



Therefore, the MST shown in Figures 12 and 13 are representations of the minimum continually connected networks and not meant to display the only possibilities

However, upon visual examination of the MST, two ambiguities were detected. The first is indicated as a triangle connecting haplotypes BFa, BFi and PPe (Figure 13). The second indicates that although CMA is most closely related to BFa, a shorter path to other Cape Meares haplotypes can be drawn through YHa than through BFk and YHc as determined by the program.

#### *Substitution Patterns in C. pseudocurata*

A total of 127 changes were required to account for all 63 haplotypes and included 87 transitions, 25 transversions, 15 indels. C-T and G-A transitions were equally common (43 versus 44 respectively). These values give a transition/transversion ratio of 3.5 and indels account for 12% of all substitutions. Mapping changes onto the MSTs required 122 changes, five substitutions fewer than the neighbor-joining tree. This resulted in 80 transitions with 42 A-G and 38 C-T changes. Twentynine transversions and 13 indels were also required. This results in an observed transition/transversion ratio of 2.9 and indels accounting for just over 10% of the observed changes. Despite their different approaches, both the neighbor-joining and MST trees result in very similar estimates of the patterns of substitutions.

### *Sequence Analysis of C. miniata Haplotypes*

The 12Scuc and Metb primers were used to amplify and sequence 391 nucleotides of mtDNA from 84 *C. miniata* specimens from 5 locations shown in Figure 7. Overall this region was AT rich (64%) with a slight bias against G which accounted for 9.8% of nucleotide sites. Of the 391 nucleotides, 63 were found to be variable and are shown in Figure 14, of which 33 were parsimonious sites. Figure 9 shows the variability across all sites. The number of substitutions in a ten nucleotide block ranged from zero to five. Among the 84 specimens sequenced, 68 unique haplotypes were detected. Therefore, only shared haplotypes are presented in Table 10. In contrast to *C. pseudocurata*, *C. miniata* shared haplotypes were often detected in distantly separated locations (Table 10) and at lower frequencies of occurrence. For example, the most common haplotype detected, Ja, was shared by five individuals corresponding to 6% of the entire population sampled; three from Juneau, AK., one from Diana Island, B.C., and one from Bodega Bay, CA.

### *Phylogenetic Analysis: C miniata*

The distance matrix shown in Appendix B was used to construct a neighbor-joining tree (Figure 15). Maximum likelihood trees were also generated (not shown) and were largely in agreement with the neighbor-joining tree. However a few taxa were placed in significantly different portions of the tree. Bootstrap resampling was employed to test the consistency of the data and provide an estimate of the confidence limits of the

Figure 14. Detected variable nucleotides among *C. miniata* haplotypes. Variable sites are presented according to position in region used for population analysis (see Figure 8).

Periods indicate nucleotide identity shared with haplotype T1. Only variable sites are shown.



11111111111111111111111122222222222222222233333333333333  
11112444445577233334445555667788000233444555789900333556678899  
45689023684878303682682379142378012756029457850134279242421912  
T1 AACCCTCATTTAATATTATACAGCCCTCGATCTATCAGGTCGCTTTGAGCCCTACCTTTTCT  
T2 .....G.....T.....  
T3 .....CCT.....T.....C...  
T4 .....T.....  
T5 .....T.....  
T6 .....C.....C.....  
T7 .....C.....T.....C.....A.T.....C...  
T8 .....C.....T.....C.....T.....T.....  
T9 .....C.....C.....CC...  
T10 .....G.....C.....G.....C...  
T11 .....A.....T.....  
T12 .....G..G.....C.....C...  
T13 G.....C.....T..G..C.....C...  
T14 .....G..C.....C.....  
T15 .....G..C.....C.....C...  
T16 .....G..C.....C.....  
T17 .....C.CG.....A...C.....  
T18 .T.....C.....  
T19 .....G.....A...C.....G.....  
T20 .....C.....  
J5 .....C...A.....  
J7 .....C.....  
J8 .....C.....C.....C.....G.....  
J4 .....CA.....  
J6 .T..... -- CCT.....TC.....C...  
J3 .T.....T.....C...  
J1 .....C.....C.....  
J2 .....C.....C.....C.....  
J9 .....C.....C.....CC...  
J10 .....CC...-A..C.....T.....  
J11 .....CC.....T.....  
J12 .T.....C.....-A..C.....CCT.....  
J13 .....C.....C...-A..C.....T.....  
J14 .....C.C.....C.....  
J15 .....C.....C.....T.....  
J16 .....C...-...C.....A.....  
J17 .....G.....A...C.....  
J18 .....A.....  
J19 .....C.C.....G.....  
D1 .....C.....C.....C.....  
D2 .....T.....C.....A.....C.....  
D3 .....T..A.....  
D4 .....T.....  
D5 .....A.....A.....  
D6 .....C..T.....CC.....T.....  
D7 .....C.....C.....GA.....  
D8 .....G..C.....A.....  
D9 .G.....C.....A.....  
D10 .T.....CCT.....C.....  
D11 .....C.....  
D12 .....C.....A.....C.....G.....  
D13 .....C.....

```

11111111111111111111222222222222222233333333333333
11112444445577233334445555667788000233444555789900333556678899
45689023684878303682682379142378012756029457850134279242421912
T1 AACCCTCATTTAATATTATACAGCCCTCGATCTATCAGGTCGCTTTGAGCCCTACCTTTTCT
D14 .G.....G.....CC.....TC
D15 .....G.....
D16 ....C.....G...C...T.....C.....
D17 .....C.....A.....C.....
D18 .....G.....C...A.....G.....
V1 .....C.....C.....
V2 .....C.....C.....G.....
V3 .....C.....C.....C.....T.....
V4 .....C.....C.....C.....T.....
V5 .....C.....C.....T.....
V6 .....CG.....
BB1 .....C.G.....A...C.....
BB2 ....T.....C.....C.C.....T.....
BB3 .....C.....C.....C.....T.....
BB4 .....G.....C.C.....C.....
BB5 .....C.....C.....C.....C.....

```

Table 10

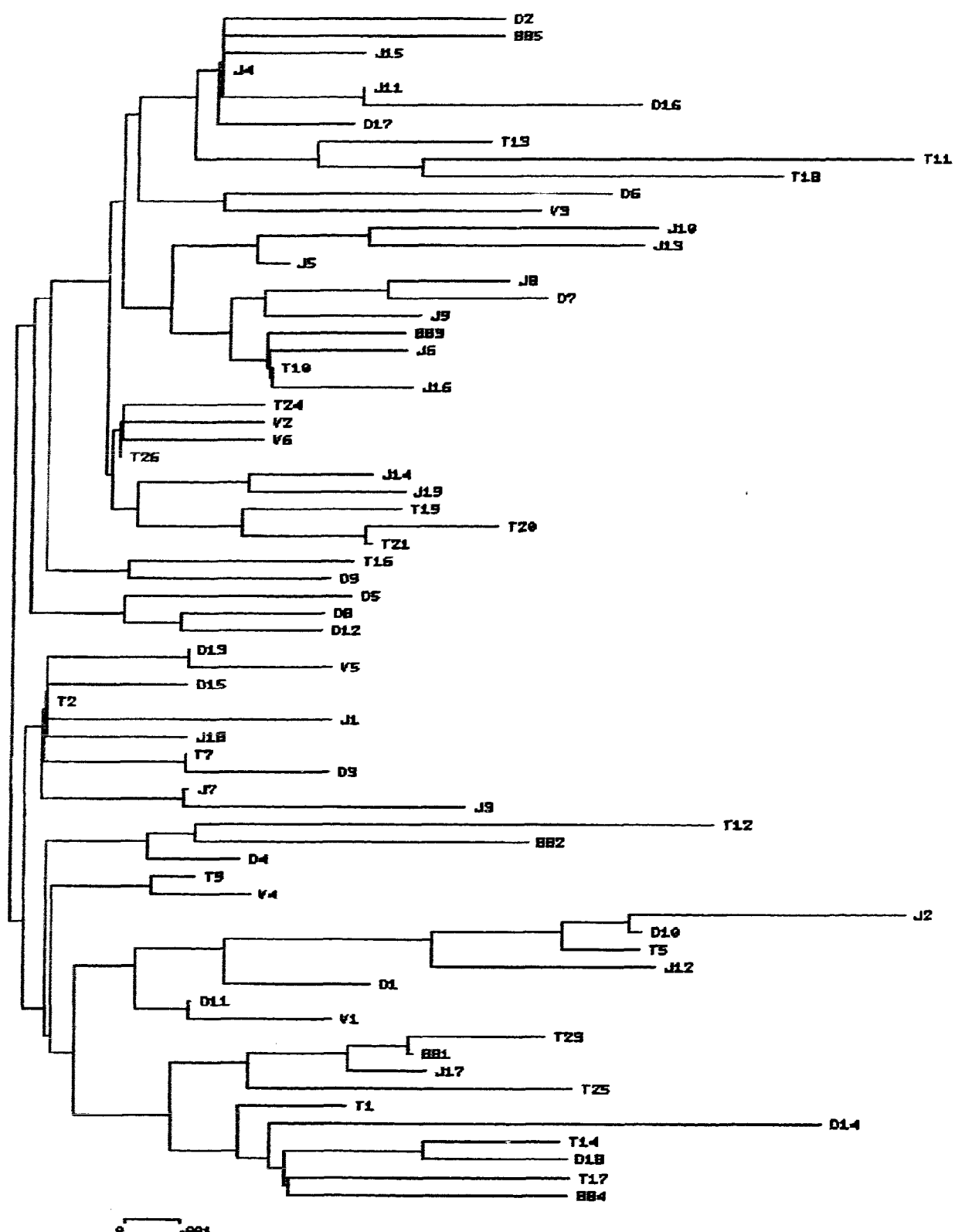
Frequency and Distribution of Shared *C. miniata* mtDNA Haplotypes<sup>1</sup>

Location	Latitude	Sample Size	Haplotype <sup>2</sup>					
			B1	D1	J1	J2	T1	V1
Juneau, AK.	58 <sup>0</sup> 10' N	24		1	3	1	2	
Vancouver, B.C.	49 <sup>0</sup> 15' N	7						1
Diana Island, B.C.	48 <sup>0</sup> 50' N	23		2	1	1	2	
Trinidad, CA.	41 <sup>0</sup> 05' N	23					4	
Bodega Bay, CA.	38 <sup>0</sup> 25' N	7	2		1			1
Totals			2	3	5	2	8	2

<sup>1</sup> Only haplotypes occurring more than one time are recorded. A total of 68 haplotypes were identified from 84 specimens.

<sup>2</sup> Haplotypes are labelled according to population from which they were first identified.

Figure 15. Neighbor-joining tree for *C. miniata* haplotypes. The tree was based on a distance matrix calculated using the method of Tamura and Nei (1989) (see Appendix B) and was constructed using the computer program MEGA (Kumar et al., 1993). OTU labels indicate sample location: Juneau, AK (J); Diana Island, B.C. (D); Vancouver, B.C. (V); Trinidad, CA (T); and Bodega Bay, CA (BB) (see Figure 7).



resulting tree, however, values for nearly all nodes were very low (not shown) and likely reflect the minimal number of characters separating branches.

#### *Minimum Spanning Tree: C. miniata*

The MST shown in Figure 16 resembles a star-like pattern of relatedness. A striking radiation of many haplotypes arising from one of a few central hubs is seen. Only five haplotypes were represented by more than one individual and as in the case of *C. pseudocurata*, most of these more common haplotypes were central hubs in the MST. In contrast to *C. pseudocurata*, the *C. miniata* mtDNA haplotypes showed no evidence of clustering based on geographic location.

#### *Patterns of Substitution: C. miniata*

Mapping substitutions onto the neighbor-joining and MST trees resulted in similar patterns of nucleotide substitutions. For the neighbor-joining tree, a total of 122 substitutions were required including 66 C-T and 32 A-G transitions, 11 transversions of which only 2 involved G, nine deletions and three insertions. Thus a transition/transversion ratio of 9.0 was observed and indels accounted for just over 10% of substitutions. Similar patterns were obtained using the MST tree where a total of 146 substitutions were required. These included 79 C-T and 35 A-G transitions, 19 transversions with only one involving G, 13 indels were detected. This results in a transition/transversion ratio of 8.1.

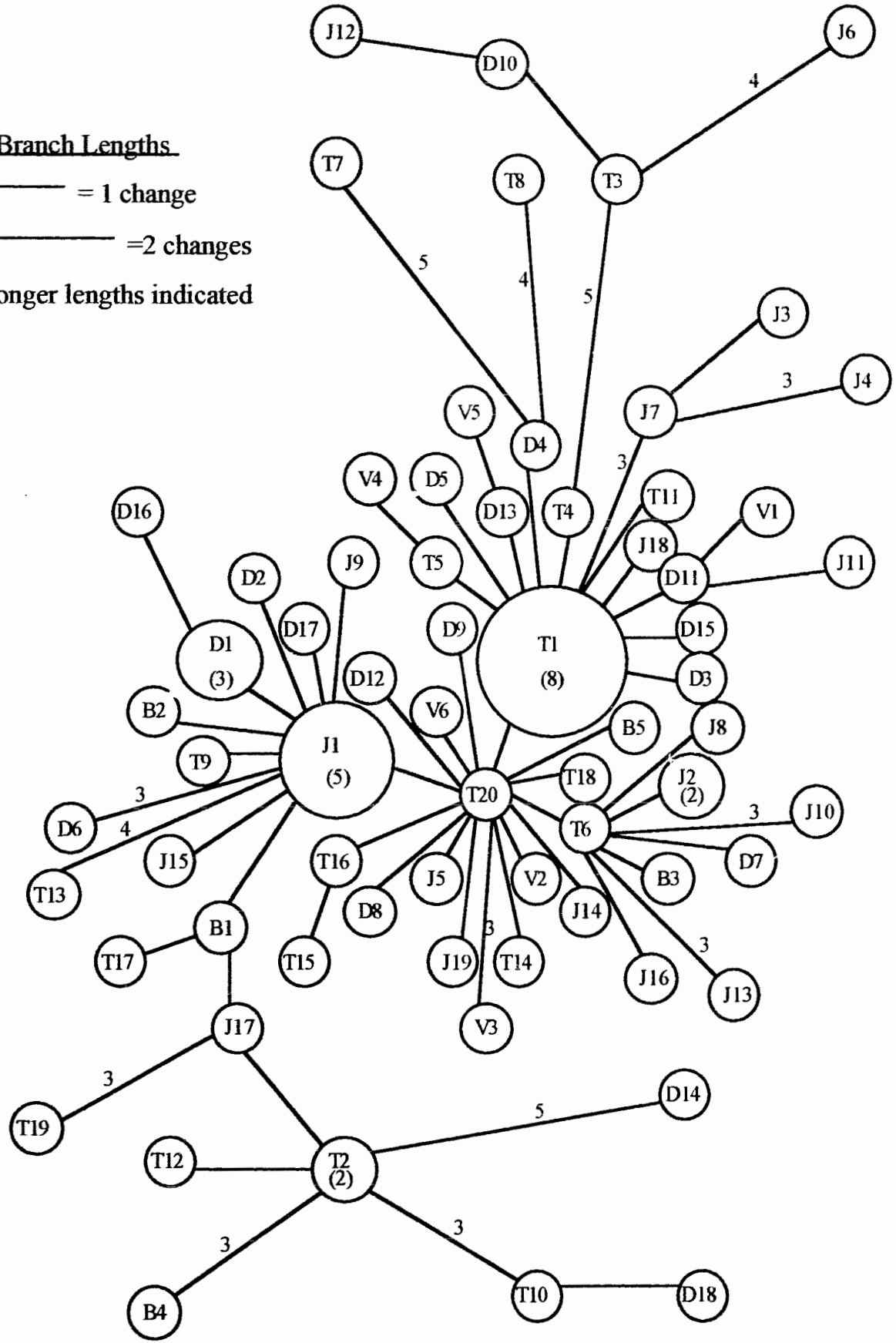
Figure 16. Minimum spanning network among *C. miniata* haplotypes. Network was constructed using the computer program NTSYS (Rohlf, 1990), based on a distance matrix calculated by the method of Tamura and Nei (1989) (see appendix B). All distances greater than two changes are indicated along branch lengths. The number of times a haplotype was detected is represented as a square root of the radius of the circle resulting in a given haplotype circle having an area proportional to the haplotype frequency. Haplotype labels indicate sample location: Juneau, AK. (J); Diana Island, B.C. (D); Vancouver, B.C. (V); Trinidad, CA. (T); and Bodega Bay, CA. (B).

Branch Lengths

— = 1 change

— = 2 changes

Longer lengths indicated





*Population Structure: C. pseudocurata*

Average nucleotide distances within and between demes for *C. pseudocurata* are presented in Table 11. Within demes, this distance ranged from a low of  $0.19 \pm 0.17\%$  at Tow Hill, B.C., to a high of  $0.75 \pm 0.60\%$  at Moresby Island, B.C. with an overall average of  $0.48 \pm 0.30\%$ . Variance estimates within demes were quite large and standard deviations were as large as the mean distance values (Table 11). Mean interdeme distances were typically an order of magnitude larger than intrademe values (Table 11). The mean interdeme distance among the three northern sample sites was  $2.43 \pm 0.46\%$  and  $1.44 \pm 0.35\%$  among the four southern sites. The overall interdeme distance among all seven sites was  $3.98 \pm 0.36\%$ . This value reflects the degree of the north/south discontinuity with the average distance between individuals across this break equal to  $6.1 \pm 2.4\%$ .

An  $F_{ST}$  estimate among all eight samples (0.88) was obtained according to the formula of Hudson et al., (1992). In order to test whether this high value was due primarily to the north/south discontinuity,  $F_{ST}$  was recalculated on the northern or southern groups alone and resulted in  $F_{ST}$  estimates of 0.77 and 0.76, respectively. Bootstrap analysis indicated zero out of 500 random samples resulted in a higher  $F_{ST}$  value, indicating a high level of significance in the geographic structure of populations.



Two population samples were obtained near Bamfield, B.C., in an attempt to examine the potential geographic size of populations. First Beach lies southwest of Bamfield on Vancouver Island and is approximately five kilometres distant from the Diana Island site included in the analysis above. Furthermore, these two sites are not part of a continual stretch of suitable habitat as they are separated by Trevor Channel, some three kilometres wide. Such separation would presumably form a formidable barrier to a brooding organism of the littoral zone. Despite this, haplotype frequency distributions at the two sites are nearly identical (Table 9). The two Bamfield area sample sites were also analyzed and resulted in an estimate of  $F_{ST} = 0.009$  and a Monte Carlo simulation indicated these two population samples were not significantly different.

Using Hudson et al.'s (1992) estimator for gene flow [ $\langle Nm \rangle = 1/2(1/F_{ST} - 1)$ ] yields a value of 0.07 migrants/generation for all eight populations, a value of 0.15 migrants/generation among either North or South populations and a surprisingly high value of 54 migrants/generation between the two Bamfield area sites.

Results of the variance analysis (Excoffier et al. 1992) using the computer program AMOVA are summarized in Table 12. Overall, 96.6% of the variance was due to differences among populations and only 3.4% due to differences within populations, resulting a value of  $\phi_{ST} = 0.97$  and an estimate of gene flow of  $N_e m = 0.017$ . Based on the genetic split observed in the phylogenetic analysis, population samples were assigned to one of two groups, designated North and South, separating Alaska and the Queen Charlotte Islands/Haida Gwaii from Vancouver Island, Oregon and California samples.

Table 12. Sea cucumber population structure based on analysis of variance using the computer program, AMOVA

Population Groupings	Variance Component <sup>1</sup>	% Total Variance	$\phi$ -Statistics	P (more extreme value)
<i>C. pseudocurata</i>				
Total (8 populations)	AP	96.6	$\phi_{ST} = 0.966$	<0.002
	WP	3.4		
North/South <sup>2</sup>				
(3 northern populations, 5 southern populations)	AG	88.6	$\phi_{CT} = 0.886$	<0.002
	AP/WG	9.5	$\phi_{SC} = 0.831$	<0.002
	WP	1.9	$\phi_{ST} = 0.981$	<0.002
Barkley Sound (2 nearest populations)	AP	6.6	$\phi_{ST} = 0.066$	<0.074
	WP	93.4		
<hr/>				
<i>C. miniata</i>				
Total (3 populations)	AP	2.5	$\phi_{ST} = 0.025$	<0.10
	WP	97.5		

1. Variance is broken into components of among groups (AG), among populations within groups (AP), and within populations and are given the  $\phi$ -statistic symbols  $\phi_{CT}$ ,  $\phi_{SC}$ , and  $\phi_{ST}$ , respectively, according to Excoffier et al., (1992). P values greater than 0.05 indicate the populations are not statistically different.

2. The North/South split of groups was assigned to minimize among population variance and was based on the split shown in the neighbor-joining tree (Figure 11). The three northern populations include Alaska & the Queen Charlotte Islands/Haida Gwaii, while the five southern populations include Vancouver Island, Oregon and California.

Splitting populations into these two groups resulted in 88.6% of the variance due to differences between these two groups, 9.5% due to differences among populations within groups and only 1.9% due to differences within populations. These values correspond to a  $\phi_{ST}$  and  $\phi_{SC}$  estimates of 0.98 and 0.83, respectively. The  $\phi_{SC}$  value is most relevant to examining gene flow among populations, since little is expected between the two groups, and results in an estimate of  $N_e m = 0.10$ . Analysis of the two Bamfield area samples indicated a very different trend, with 93.4% of the variance within populations and only 6.6% due to differences between these populations.

Pairwise  $\phi_{ST}$  estimates for all population samples were also calculated (Table 13). Again the two Bamfield samples were shown to not differ significantly ( $p$  value, 0.27) with a  $\phi_{ST}$  estimate of 0.066 corresponding to  $N_e m = 7.1$ . The calculation of  $\phi_{ST}$  is sensitive to any differences among populations and close relationships between two populations will not be detected if included in a larger set of separated populations. To illustrate this point, the two Bamfield samples were included with the Sooke sample to examine genetic structure on Vancouver Island. In this case,  $\phi_{ST} = 0.78$  and  $N_e m = 0.14$ , a value suggesting significant separation of these sample sites, despite the fact that the two Bamfield samples are indistinguishable as previously noted. Therefore, the pairwise  $\phi_{ST}$  calculations produced by AMOVA are very useful for detecting potential gene flow between any sites and are shown in Table 13. Except for the nearby Bamfield sites, the only pair of sites exhibiting genetic similarity were the two Oregon samples, ( $\phi_{ST} = 0.37$ ),

Table 13. Pairwise  $\phi_{ST}$  estimates among *C. pseudocurata* population samples

Deme	1	2	3	4	5	6	7	8	9
1		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.907		0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.865	0.788		0.0	0.0	0.0	0.0	0.0	0.0
4	0.983	0.994	0.978		0.0	0.0	0.0	0.0	0.0
5	0.979	0.985	0.977	0.704		0.0	0.0	0.0	<b>0.27</b>
6	0.963	0.957	0.953	0.502	0.503		0.0	0.0	0.0
7	0.987	0.997	0.986	0.955	0.909	0.370		0.0	0.0
8	0.986	0.993	0.985	0.893	0.747	0.767	0.967		0.0
9	0.982	0.992	0.980	0.844	<b>0.024</b>	0.468	0.939	0.814	0.0

Below diagonal:  $\phi_{ST}$  between pairs of populations

Above diagonal: Probability of Random distance ( $\phi_{ST}$ ) > Observed distance

Number of iterations : 500

Sample Sites:

1. Baranof Island, AK; 2. Tow Hill, B.C.; 3. Moresby Island, B.C.; 4. Sooke; B.C.; 5. Diana Island, B.C.; 6. Cape Meares, OR; 7. Yachats, OR; 8. Pescadero Point, CA; 9. First Beach, B.C.

resulting in a gene flow estimate of  $N_e m = 0.85$ . However, the bootstrap analysis indicated this value was not significantly different from random (Table 13).

#### *Population Structure: C. miniata*

The intra- and interdeme distances among all five *C. miniata* sample sites are presented in Table 14. The mean distance within demes was  $1.15 \pm 0.45\%$  and the mean distance between demes was  $1.20 \pm 0.48\%$ . Estimates of geographic structuring of populations were obtained initially for the three large *C. miniata* population samples with  $F_{ST} = 0.008$ . Gene flow was estimated at  $N_e m = 61$ .

Variance analysis of *C. miniata* population samples indicated that 95.2% of variance due to differences within populations and only 4.8% among populations. These results yield a  $\phi_{ST}$  estimate of 0.05 and a value of  $N_e m = 10$  is obtained, indicating very high levels of gene flow. This value is in considerably lower than the value obtained by the method of Hudson et al. (1992) of  $N_e m = 61$ .

#### *Ossicle Morphology of C. pseudocurata*

Ossicles from the dorsal skin were prepared from representative specimens from the major clades identified in the neighbor-joining tree (Figure 11) and are presented in Figure 17. Specimens from Vancouver Island, B.C., Yachats, OR., and Pescadero Point, CA., all possess large perforated plates typical of *C. pseudocurata* (Lambert, 1985). In contrast, ossicles from specimens from Baranof Island, AK., Auke Bay, AK., and

Table 14A. Mean intrademe distance for *C. miniata* samples

Deme	Location	Mean intrademe distance	Variance	Standard Deviation
1	Juneau	0.0123	$3.1 \times 10^{-5}$	0.0056
2	Diana I., B.C.	0.0116	$3.0 \times 10^{-5}$	0.0055
3	Burrard Inlet, B.C.	0.0107	$2.4 \times 10^{-5}$	0.0049
4	Trinidad Head, CA.	0.0119	$5.4 \times 10^{-6}$	0.0023
5	Bodega Bay, CA.	0.0110	$1.6 \times 10^{-5}$	0.0040

Table 14B. Mean interdeme nucleotide distances for *C. miniata* samples

	1	2	3	4	5
1		0.0122	0.0122	0.0116	0.0126
2	0.0053		0.0117	0.0112	0.0124
3	0.0051	0.0044		0.0113	0.0124
4	0.0054	0.0053	0.0042		0.0125
5	0.0046	0.0053	0.0043	0.0044	

Distance estimates in upper diagonal; Error estimates in lower diagonal

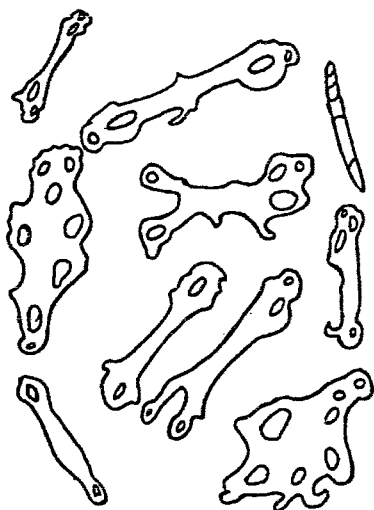
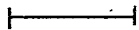
$F_{ST}$ : using all five population samples = 0.041

using samples 1,2, and 4 (equal sizes,  $>10$ ) = 0.0081

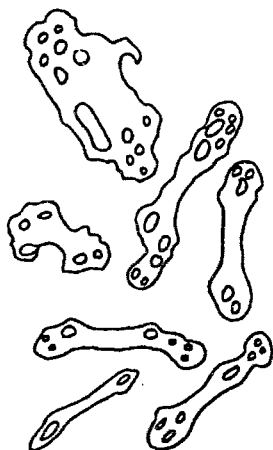


Figure 17. Comparison of ossicles from northern and southern populations. Skin samples were taken from central dorsal region, digested in 10% bleach, rinsed and mounted on slides and viewed under 100x magnification. Scale bar shown indicates 0.10 mm.

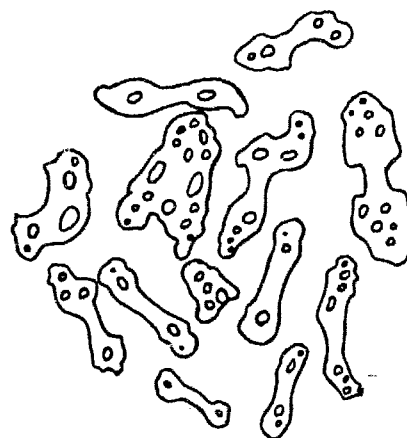
0.10 mm



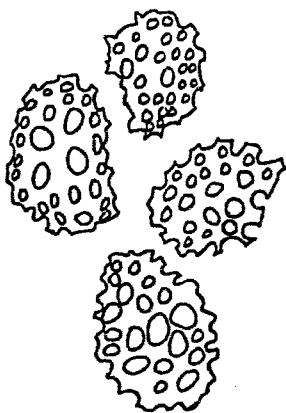
*C. Vegae*  
Auke Bay 2  
Northern A



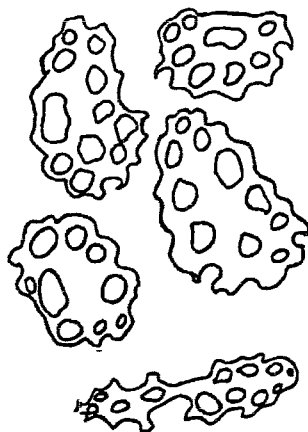
*C. pseudocurata*  
Baranof Island 1  
Northern B



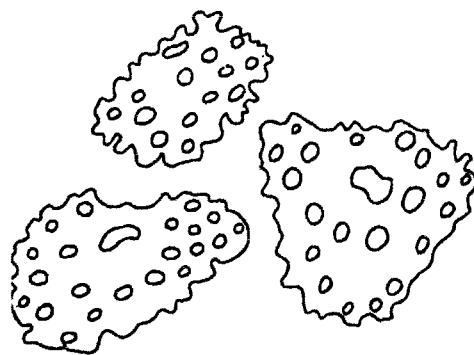
*C. pseudocurata*  
Moresby Island 6  
Northern C



*C. pseudocurata*  
Diana Island 3  
Southern A



*C. pseudocurata*  
Yachats 1  
Southern B



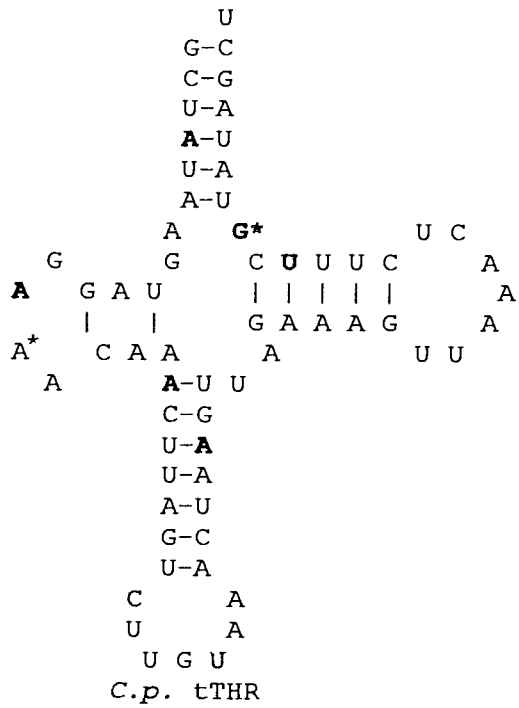
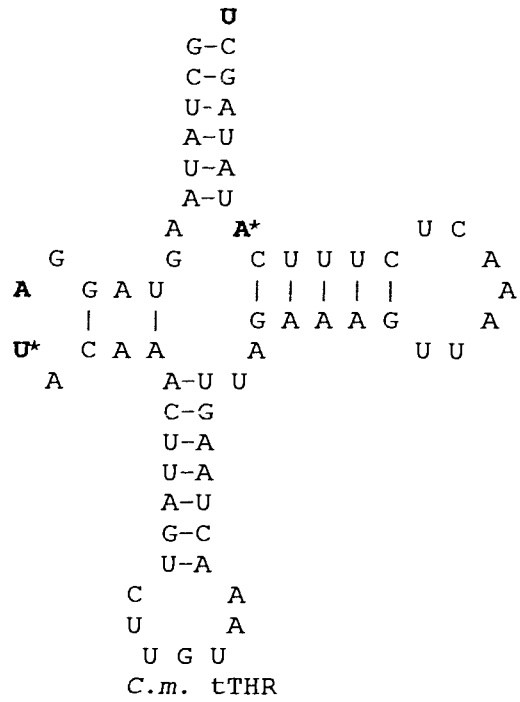
*C. pseudocurata*  
Pescadero Point 2  
Southern C

Moresby Island, B.C., are much more similar to the rod or dumbbell shaped ossicles characteristic of *C. vegae* (Lambert, pers. commun.).

#### *Substitutions in the Threonine tRNA Gene*

The threonine tRNA gene lies within the region sequenced for the population study. Substitutions detected within this gene were examined with respect to the secondary structure of the tRNA molecule. Figure 18 shows variable positions found within and between species. As indicated, the consensus sequence of the two species differ by only a single nucleotide transition in one of the loops. In the case of *C. miniata* four transitions were detected with none involving stem positions. Six changes were detected among *C. pseudocurata* haplotypes and included five transitions and one deletion. Three of these changes occurred at stem positions involved in base pairing. The first instance was the deletion of A19 (Figure 18) which could potentially pair with U39 in the anticodon stem. This change is likely minor, since it occurs at the top of the stem which would still involve five base pairs. The second change is also minor, an A to G transition which alters the U21/A37 pair in the anticodon stem to a U21/G37 pair, an acceptable pairing which would not significantly affect the stem stability. In contrast to these minor changes, a single haplotype from Baranof Island exhibited a U to C transition at position 58 which lies within the T stem in the consensus sequence (Figure 18). This results in a A-C pair in the stem and, although seen in other tRNAs, would appear to destabilize this stem to some degree.

Figure 18. Variable sites detected in the threonine tRNA gene. Character in bold indicate variable positions detected within species. The nucleotides marked with asterisks represents the only two differences detected between consensus sequences for *C. miniata* and *C. pseudocurata* threonine gene sequences.



## Discussion

### *Patterns of Nucleotide Frequency and Substitution*

Similar patterns of substitutions were observed in the two species. By assigning changes to particular branches of either minimum spanning networks or neighbor-joining phylogenetic trees, it was possible to determine the direction of mutation in most cases and the two approaches yielded nearly identical patterns within species. Consequently, only the patterns obtained based on the MSTs are discussed. Transitions were more common than transversions by a factor of 3.5:1 in the case of *C. pseudocurata*, and 8:1 in the case of *C. miniata*. The transition/transversion ratio is considerably lower than that reported for many organisms, particularly vertebrates where a ratio of 10:1 has been reported (Brown et al., 1979). As described in Chapter I for the COI and IrRNA genes, this region of the genome exhibited significantly fewer substitution events involving G in the case of *C. miniata*. C-T transitions were twice as common as A-G transitions, and transversions involving A were between five and nine times as common as those involving G among *C. miniata* haplotypes. Overall this region was AT rich (63.4%). In contrast, this pattern is not maintained in *C. pseudocurata* where A-G transitions were as common as C-T transitions and transversions involving G were more common than those involving A.

Indels were roughly half as common as transversions in *C. pseudocurata* and as common as transversions in *C. miniata*, thus constituting a significant class of mutation events in these non-protein coding regions. In both species, the most common size class

of such events was a single nucleotide. Discarding such information remains an unfortunate prerequisite for distance and maximum likelihood analysis.

### *Substitutions Within the Threonine tRNA Gene*

Among the 68 *C. miniata* haplotypes, four transitions were detected within the threonine tRNA gene. Two of these occurred in the Dhu loop, one occurred at the point of transition between the T $\phi$ C stem and the acceptor stem, and the fourth occurred at the 3' terminus. All these positions are likely to be under less structural constraint than stem positions and it is therefore not unexpected to detect differences at a population level.

Among the 63 *C. pseudocurata* haplotypes, six substitutions were detected within the threonine tRNA gene and, similar to the pattern observed for *C. miniata*, included five transitions and one deletion. Two A to G transitions at positions 4 and 37 altered base pairing in stems from A-U to G-U. A to G transitions were also detected at positions 13 and 60 that do not involve stem positions. In fact these two substitutions are at the homologous positions to identical A to G transitions described above in *C. miniata* suggesting the highly variable nature of the sites. The deletion of A19, which is depicted as pairing with U39 in Figure 17, was also detected in one *C. pseudocurata* haplotype. This change is likely not structurally important as the anticodon stem remains stable with six nucleotide pairs and the unpaired U would simply increase the variable loop by one nucleotide. One substitution of a more surprising nature was also detected at position 58 in the T $\phi$ C stem (Figure 18). These changes result in C-A mismatch pairing and barring

the utilization of some modified C or A nucleotide, would likely destabilize the stem structure to some extent. Alternatively, this tRNA may be subject to some form of post transcriptional editing to restore base pairing. It is worth noting that the Dhu stem, as depicted, is also somewhat unstable and the presented structure should be treated as preliminary. Since sequencing of the mitochondrial genome has not been completed for these species, it remains possible that this region does not constitute the functional threonine tRNA gene, although the near identity between species suggests some type of functional constraint exists for this region. It seems that several positions within the threonine tRNA gene are under relatively relaxed constraints and may accumulate substitutions even within populations. Recently, Kumazawa and Nishida (1993) suggested the utility of stem portions of tRNA genes for assessing deep phylogenies. The present work suggests caution at such levels due to the great variability in rate of variation among sites and therefore the potential for convergence in nucleotide data.

#### *Population Structure in C. pseudocurata*

The results of the population analyses indicate that the two species differ dramatically in terms of the degree of geographic structuring of genetic variation. The observed variance in the mtDNA sequences of *C. pseudocurata* population samples was almost completely among populations (96.6 %) (Table 13) rather than within populations (3.4%). Even when the sample sites are divided into northern and southern groups, the nested AMOVA analysis (Table 13) indicated that only 1.9% of the variance was due to



differences within populations, while 9.5% was partitioned among populations within groups and 88.6% due to differences among groups.

Estimates of gene flow among populations within groups were low using either the method of Hudson et al., (1992) ( $N_e m = 0.15$ ) or Excoffier et al., (1992), ( $N_e m = 0.10$ ). However, these two methods yielded very different values for the two Bamfield populations ( $N_e m = 54$  or  $7.0$ , respectively). The analysis of variance approach (Excoffier et al., 1992) partitions all observed variance into hierarchical levels, thereby utilizing all information in contrast to that of Hudson et al., (1992) which relies on mean estimates of distances within and between population. Since the variance approach uses squared distance values, any measurable difference among populations will be magnified and thus reduce an estimate of gene flow. The method of Excoffier et al., (1992) is therefore a more conservative approach to estimating gene flow, particularly in cases of high levels of gene flow. In any case, since gene flow values of greater than one are presumed sufficient to prevent population differentiation (Slatkin, 1987), these results suggest a connection of the two Bamfield area populations of *C. pseudocurata*. Alternatively, the assumption of populations being in equilibrium may not hold and insufficient time may have passed for differentiation to arise. As pointed out in Cockerham and Weir, (1993) such estimates of gene flow assume the observed population differentiation is due to the level of gene flow and in the absence of direct observation of successful migration, should be treated cautiously.

*C. pseudocurata* occupies a narrow zone of varied but specialized habitat. Over much of its range, there is a linear character to this habitat and is therefore closely

approximated by the stepping stone model (Kimura and Weiss, 1964). The expansive sand beaches along the coast of northern Washington and southern Oregon effectively separate the southern group into three main areas: British Columbia, Washington and Oregon, and California. Several extensive beaches also occur on the west coast of Vancouver Island, particularly those of Pacific Rim National Park including Long Beach and further south, Carmanah Beach, which may account for the separation seen between the Bamfield and Sooke sample sites. Areas on the Island north of Long Beach were not sampled but are likely to be distinct from those of the central and southern regions for the reasons outlined above.

Areas in the north such as Barkley Sound, within which the Bamfield samples were obtained, provide a much broader habitat with numerous islands and a predominance of rocky shores resulting in a significant departure from a simple linear distribution. Such areas in fact characterize the north coast of British Columbia and southeast Alaska and may provide a considerably increased opportunity for successful migration. The two Bamfield population samples were not significantly different from each other according to AMOVA ( $p > 0.07$ ) despite their physical separation of 5 km of ocean. Significant gene flow among these sites would presumably occur through a rafting mechanism as suggested by Highsmith (1989). The exchange of a few females could occur due to movements caused by storms. One such possibility would be with clumps of the surf grass *P. scouleri* that are often seen washed ashore following storms (pers. obs.) and with which *C. pseudocurata* is often associated. Alternatively, it is also possible that there is negligible exchange between these sites and the similarity is due to a common ancestral

source population associated with lower sea levels of the last glaciation and these populations have not undergone significant divergence in the post glacial period. Sample sites generally formed monophyletic groups and appear to be evolving independently. On the scale of hundreds of kilometres which typically separates sample sites in this study, there is little support for a stepwise dispersal pattern. Limited or rare genetic exchange is also suggested between the Cape Meares and Yachats in Oregon. Thus, a limited stepwise or linear dispersal pattern clearly seems to occur over distances of a few kilometres but less than 100 km. This upper limit may reflect the occurrence of uninhabitable zones along the coastline, resulting in barriers to dispersal and local differentiation. The detection of haplotypes assigned to two different clusters in the trees for both the Bamfield and Pescadero Point samples is consistent with long term persistence of two lineages along the south coast.

The distinction between the northern and southern population samples of *C. pseudocurata* was highly significant, with 88.6% of the observed variance occurring between the two groups (Table 14). The mean genetic distance between groups was 6.0% and the smallest value recorded was 3.8%. In contrast, the mean distance within groups was 2.4% in the northern group and 1.4% in the southern group, with the maximum observed distance within either group being 4.3%. The groups are significantly different (Table 13) although there is some overlap between the minimum between group and the maximum within group distances. One possible explanation of this separation is that the minimum geographic distance between northern and southern sample sites (750 km) was considerably larger than sequential distances within either group (typically 200-300 km).

This was primarily due to the difficulty and expense of sampling in the intermediate region along the central coast of British Columbia. However, several examples within each group indicate that a direct correlation between geographic and genetic distance is not well supported, particularly when large distances are involved. Haplotypes of some individuals from the Bamfield area were nearly identical to those from Pescadero Point, CA., several thousand kilometres away rather than samples geographically much closer. Similarly, the haplotype of a single specimen from Unalaska Island, AK., was most similar to those from Moresby Island, B.C., again on the order of thousands of kilometres distance and despite an intervening sample site on Baranof Island, AK. Such occurrences may be the result of convergence rather than identity by descent and are merely meant to indicate the significance of the north/south disjunction.

A second point concerning this genetic disjunction is the influence of the last glaciation. At its maximum extent, approximately 15,000 years before present, the Cordilleran ice sheet covered much of the coast of British Columbia. However, several coastal refugia likely existed along the Queen Charlotte Islands/Haida Gwaii and Vancouver Island and a larger Beringial refugia is thought to have existed in northeast Alaska along the Bering Sea (Clague, 1989). Substantial survival in these refugia seems probable given the high genetic diversity present in northern samples. In fact Baranof Island exhibited the highest mean intrapopulation distance of all samples. As the ice sheet retreated, colonization would then have spread south and east from these northern refugia as well as north from southern populations below the furthest extent of ice.

The present disjunction may be explained by the reproductive patterns and predominant coastal currents. *C. pseudocurata* spawns between December and January and juveniles hatch in early spring. As these young continue to grow, considerable spatial pressure may occur. Rutherford (1973) found a strong negative correlation between size class and frequency, with mortality particularly high in the first year of life (>96%). The author also reported a strong positive correlation between biomass density of the colony and mortality. It is possible that migration is therefore largely due to juveniles and would correspond to summertime space selection pressure. At this time of year, the northern Pacific current splits north and south off the coast of the Queen Charlotte Islands/Haida Gwaii into the Alaska and California Currents, respectively (Thomson, 1981). This current split coincides with the mtDNA genetic disjunction observed, separating Alaska and the Queen Charlotte Islands/Haida Gwaii from Vancouver Island and regions to the south. In contrast, during the winter months, the Davidson Current predominates off the west coast of Vancouver Island and regions to the south and moves in a northerly direction to join the Alaska Current (Thomson, 1981). If dispersal were to occur predominantly in the winter months due to the Davidson Current, it is improbable that such a disjunction as that detected would arise. Although a correlation between summer current patterns and genetic variance exist, alternative hypotheses could explain this phenomenon without suggesting that these current patterns regulate gene flow, presumably through their influence on rafting, although the possibility exists. In fact, predominant currents inside the continental shelf, which may have a far greater influence

on the short term dispersal of this species, are dependent on highly variable prevailing wind patterns (Thomson, 1981).

Alternatively, the observed disjunction may be a result of expansion following the last period of glaciation occurring from both northern and southern refugia. This expansion may only recently have connected these two population groups and genetic exchange may still be possible yet sufficient time may not have passed to allow detection of such exchange. This is certainly possible given that the proposed point of this disjunction was not sampled. It should be pointed out that these two concepts are not mutually exclusive. Higher genetic diversity in the northern samples compared to southern samples suggests substantial refugial survival while the disjunction is consistent with long term separation of these lineages. Furthermore, the lack of intermolecular recombination and maternal inheritance of mitochondrial DNA may preclude the earliest detection of genetic exchange following expansion. The use of nuclear markers such as microsatellites or allozymes would allow the detection of hybrids as heterozygotes possessing alleles previously restricted to either northern or southern populations.

It appears that, although limited gene flow does occur along the coast and may involve rafting mechanisms, particularly in area such as Barkley Sound, the predominant pattern is one of local populations evolving independently. The highly structured nature of these populations is likely due to isolation by distance as well as the presence of significant barriers to gene flow such as areas of sand beach or predominant current patterns. In the southern regions of this species' range, habitat is much more linear in

nature and sand coastline is more common than in the north and would suggest an even stronger pattern of local isolation.

Despite the uncertain influence of coastal currents on gene flow in *C. pseudocurata* populations, examples of predominant current patterns resulting in such genetic disjunctions are well known. Avise (1994) has reviewed evidence of a major genetic disjunction between Atlantic and Gulf of Mexico populations based on mtDNA sequence from numerous taxa including the horseshoe crab, *Limulus polyphemus*, the American oyster, *Crassostrea virginica*, the black sea bass, *Centropristis striata*, diamondback terrapin, *Malaclemys terrapin* and the seaside sparrow, *Ammodramus maritimus*. This break corresponds to the offshore currents at Cape Canaveral Florida effectively separating these two regions. Avise (1994) suggested that ancestral populations were most likely separated during the Pleistocene that, due to ecological and/or dispersal influences, have remained distinct. In contrast, Buroker (1983) reported uniform allozyme allele frequencies in American oyster populations from this region. Subsequently, Avise and Karl (1992) found that nuclear restriction fragment length polymorphisms agreed with the mtDNA data in detecting this genetic break, but that allozyme markers did not show this split. The authors suggested that one explanation of this discrepancy may be that allozyme loci are under balancing selective constraints and thus are not free to vary sufficiently to detect such a split.

On the Pacific coast, Burton (1983) has described local isolation among populations of the tide pool dwelling copepod, *Tigriopus californicus*, separated by as little as 500m despite a life history including pelagic larvae. Berger (1973) examined

population structure among three species of gastropod from the genus *Littorina* with differing lengths of larval stage and found that two species without pelagic larvae exhibited higher population differentiation than a congener with such a larval stage. Campbell (1978) examined the genetic divergence among populations of the marine gastropod (*Nucella*) *lamellosa* from San Francisco to Glacier Bay using allozymes and found a wide range of variation with little geographic correlation. This organism lacks a pelagic larval stage, is confined to rocky shores and often returns to the same breeding congregation which result in a low level of migration between demes. In a similar study, Hellberg (1993) examined allozyme variation in populations of the solitary coral, *Balanophyllia elegans*. Larval dispersal in this species has been estimated to be on the order of centimetres (Gerrodette, 1981) and, upon settlement, adults are sessile, resulting in an extremely low dispersal potential. Hellberg (1993) compared the pattern of genetic variation to expectations of the stepping stone model of Kimura and Weiss (1964). Overall, a high degree of genetic differentiation was detected with significantly less variation observed in more northerly samples. It is interesting that neither of these two studies detected the north/south genetic disjunction described for *C. pseudocurata* and may reflect the nature of the genetic marker utilized as discussed above. Hellberg (1993) included only one site which would undeniably belong to the northern region of the present study. It may therefore be difficult to discriminate between local differentiation and a larger phylogeographic split from such data. Alternatively, these studies may indicate that the sea cucumber populations examined have a different biogeographic history from those of the gastropods of the genus *Nucella* and the coral, *Balanophyllia*



*elegans*. It is possible that these groups were not present in coastal refugia and therefore were reestablished from southern populations. In the case of *Balanophyllia elegans*, this is consistent with the reduced genetic diversity observed in northern populations as discussed by Hellberg (1993).

Evidence from a mitochondrial DNA analysis of populations of *Oncorhynchus tshawytscha*, chinook salmon (Wilson et al., 1987), also indicate that British Columbia and Alaskan populations are derived from the Beringia refuge. Both mitochondrial DNA and allozyme data for sockeye salmon, *Oncorhynchus nerka*, also indicate post-glacial dispersal from both Beringian and Columbian refugia (Bickham et al., 1995). In this study, fish from the Skeena and Fraser river drainages were genetically distinct and in fact fish from the Skeena drainage were more closely related to Alaskan and Russian fish than those of the Fraser. The Skeena drains into the Pacific approximately across from the Queen Charlotte Islands/Haida Gwaii and this evidence is therefore consistent with the genetic break observed for sea cucumber populations north of Vancouver Island. A finer scale analysis of *O. nerka* populations also suggests a third possible refugia on unglaciated portions of Vancouver Island and/or the Queen Charlotte Islands/Haida Gwaii (Wood et al., 1995). Furthermore, a recent study of harbor seal populations also demonstrated a genetic distinction between northern and southern groups with a similar break point to this study (T. Burg, pers. commun.). These results further support growing evidence of substantial refugial survival along the coast resulting, at least for the present time, in at least two distinct biogeographic regions along the coast of British Columbia and Alaska.

*The Species Spectre: C. pseudocurata versus C. vegae*

Since this study was based on a single locus, mitochondrial DNA, it is not possible to unequivocally determine whether *C. vegae*/*C. pseudocurata* represent a single species or separate taxa having arisen through vicariance or allopatry as described above based on this study alone. Avise (1994) suggests that concordance among numerous loci is required to demonstrate separate species status. *C. pseudocurata* and *C. vegae* are externally indistinguishable and characterization is based on calcareous ossicle morphology. Ossicles from body wall of *C. vegae* are typically rods with some perforations at the ends (Theel, 1886) and have been described as spectacle rods (Lambert, 1985). The body wall ossicles of *C. pseudocurata* are much larger plates with numerous perforations (Deichmann, 1938). Representative specimens from the three main southern clades all possessed such large plates (Figure 17), while ossicles from three northern specimens possessed spectacle rods characteristic of *C. vegae* (Figure 17). Lambert (1985) developed a morphometric analysis of ossicle size and shape to examine geographic variation in ossicle morphology of these two species over a range similar to the present study. Despite the distinction associated with species descriptions, a linear relationship between geographic location and ossicle shape was detected (Lambert, 1985). It is interesting to note that the single most convincing intermediate form was from the northern tip of Vancouver Island. Unfortunately, the next adjacent site to the north was near Prince Rupert, B.C. and falls clearly within the northern group. Lambert (pers. commun.) describes the area of northern B.C. and southeast Alaska as being intermediate in nature with some spectacle rods and some smaller perforated plates. Such an

intermediate type of ossicle was noted in the Moresby Island specimen. The analysis of ossicle morphology, a character presumably produced under the influence of nuclear heritable traits, parallels the mtDNA evidence of a significant genetic split north of Vancouver Island. The description of intermediate forms from northern Vancouver Island suggests the possibility of a continuum of variation and further investigation with genetic markers is warranted. *C. vegae* is the senior synonym and a demonstration of hybridization would suggest a single species having gone through isolation due to glaciation with recontact along the central B.C. coast. At present it can be stated that *C. vegae* occurs from the Commander and Aleutian Islands down to southeast Alaska and the Queen Charlotte Islands/Haida Gwaii while *C. pseudocurata* occurs from Vancouver Island south to central California.

#### *Population Structure in C. miniata*

In contrast to *C. pseudocurata*, *C. miniata* does not exhibit any significant geographic separation over the range between Alaska and California. Analysis of indicated that 97.5% of the variance was due to that within populations and results in an estimate of gene flow of  $N_e m = 20$ . As in the case of *C. pseudocurata*, the method of Hudson et al., (1992) produced a much higher estimate of  $N_e m = 61$ . Despite this discrepancy, both approaches clearly indicate substantial gene flow among *C. miniata* populations, similar to the island model of migration.

At first, this may appear somewhat surprising considering that the nonfeeding larvae of this species spend less than two weeks in the water column and are essentially

sessile once settlement occurs. Furthermore, the doliolaria larvae of *C. miniata* has been described as a sluggish swimmer (Strathmann, 1987) and would depend on local currents for movement. McMillan, Raff and Palumbi (1993) utilized mtDNA restriction length polymorphisms to examine the structure of sea urchin populations of two species (Genus; *Heliocidaris*) in the south Pacific. The authors reported that *H. tuberculata*, a species with a feeding pelagic larval phase lasting several weeks, exhibited no significant geographic structuring of populations and appeared panmictic over a range of several thousand kilometres, while *H. erythrogramma*, possessing a nonfeeding pelagic larvae lasting only a few days, showed significant structure suggesting little gene flow. In terms of mode of development, *C. miniata* clearly is more similar to the latter, sharing a nonfeeding form of larvae which admittedly lasts significantly longer than *H. erythrogramma* (10-14 versus 4-6 days respectively). Yet the lack of geographic population structure and apparent high dispersal potential of *C. miniata* follow the pattern seen in *H. tuberculata*.

Several aspects of life history and habitat preference shed some light on this situation. It is possible that the greater length of duration of the larval stage of *C. miniata* compared to *H. erythrogramma* is sufficient to account for the differences in dispersal. This organism is a filter feeder and is highly abundant in semi-protected areas with moderate to high current flow. Stewart and Levitan (1993) monitored a natural spawning event near the Bamfield Marine Station and found greater than 90% fertilization rates which they attributed to synchronous spawning and high densities. These factors suggest a high potential for successful dispersal and therefore gene flow. Furthermore, *C. miniata*

is found from the intertidal shallows to over 200m in depth which provides a wide zone of suitable habitat along the coast which could increase the survivorship of settling larvae and may therefore reduce barriers to dispersal. It is interesting that this species does not appear to be affected by the splitting of the Alaska and California currents. *C. miniata* spawns in early summer when these two currents oppose each other, yet no geographic separation was detected. Coastal currents are complex and poorly understood so the data may simply indicate that *C. miniata* occupies waters where currents are more complex than this predominant movement. Another possibility is that the larvae may be capable of considerable oceanic movement due to currents and is consistent with the abundance of this organism in areas of high current flow. It therefore seems most consistent to suggest that *C. miniata* is capable of considerable dispersal and is able to circumvent the predominant coastal currents, forming a panmictic population along the western coast of North America. This species therefore appears to most closely fit an island model of dispersal with a high probability of mixing between many demes.

#### *Age Estimates Within and Between Species*

Since the fossil record for holothurians is poor, it was not possible to provide calibration of the molecular clock within this class. Nonetheless, the rate of overall mtDNA sequence divergence for sea urchins has been estimated at 1.0-1.5% per million years ( $\text{my}^{-1}$ ) (Vawter and Brown, 1986; Palumbi and Wilson, 1990). Given that the region used for the present study is accumulating substitutions at twice the rate of the COI gene, a rate of 2.0-3.0%  $\text{my}^{-1}$  was used.

Overall mean distance estimates between lineages varied considerably for the two species examined. Among *C. pseudocurata* haplotypes the mean distance was 3.8%. Separating these into northern and southern groups yields mean distance estimates of 2.4% and 1.4%, respectively with a mean of 6.1% between the groups. Assuming equal rates among lineages, southern *C. pseudocurata* lineages last shared a common ancestor 0.5-0.7 million years ago (mya), northern *C. pseudocurata* lineages 0.8-1.2 mya and the two groups last shared a common ancestor between 2.0 and 3.1 mya, indicating that these groups have remained separate since well before the last glacial retreat began. The mean distance among *C. miniata* haplotypes was 1.2% and *C. miniata* haplotypes last shared a common ancestor 0.4-0.6 mya. It appears that *C. pseudocurata* populations are evolving independently and the set of populations may accumulate change more rapidly than the panmictic *C. miniata* population. Unequal rates of fixation of changes may exist for organisms with such differing dispersal capabilities. Assuming a common rate of substitution may therefore result in either an underestimate of the age of *C. miniata* lineages or an overestimate of the age of *C. pseudocurata* lineages. This provides an important cautionary note when reconstructing the phylogenetic history among even closely related taxa and underlines the need to use algorithms that account for varying rates of substitution among lineages.

## CHAPTER III

### Sea Cucumber Mitochondrial tRNAs:

#### Gene Order, Secondary Structures and Patterns of Substitution

##### Introduction

Phylogenetic analysis based on mtDNA is a powerful technique, particularly for closely related species, due to the high rate at which substitutions accumulate in this molecule. However, this rapid accumulation of substitutions results in a saturation of changes, restricting the level at which informative data can be extracted. The tRNA genes appear to be among the most highly conserved of mitochondrial genes and have been suggested as potentially informative at deeper phylogenetic levels (Kumazawa and Nishida, 1993). It has also been suggested that mitochondrial gene order may serve an informative role at these deeper levels (W. Brown, 1985; Moritz et al., 1987). This is due to the fact that, although the mitochondrial genome has a highly conserved set of genes, the order or arrangement of these genes is not universal.

A comparison of mitochondrial gene order of different phyla reveals substantial rearrangement. For example, *Drosophila* differs from vertebrates by two major inversion events; the first including the genes for ND4l, ND4 and ND5, the second including ND1, srRNA and lrRNA in addition to numerous tRNA gene rearrangements (Clary and

Wolstenholme, 1985; Moritz et al., 1987; Wolstenholme, 1992). In comparing vertebrate and sea urchin mitochondrial genomes, Jacobs et al., (1989) noted that two transposition events are required to relate the positions of the lrRNA and the ND4l genes. However, tRNA genes appear to be significantly more mobile with 15 of the 22 tRNA genes located in differing positions for vertebrates and sea urchins (Jacobs et al., 1989). Even within the vertebrates, minor rearrangements have been detected involving tRNA genes. Desjardins and Morais (1991) reported the relocation of two tRNA genes in mtDNA from gallinaceous birds. A small cluster of tRNAs have also been relocated in marsupial mtDNA (Pääbo et al., 1991). These examples serve to illustrate that the tRNA genes appear to be among the most mobile elements of the mitochondrial genome. In fact, it has been suggested that tRNAs might facilitate genome rearrangements based on their structural similarity to replication control elements (Cantatore et al., 1987; Moritz et al., 1987; Jacobs et al., 1988).

During the evolution of echinoderms, rearrangements have occurred on a larger scale, which is not surprising given that the six extant classes had likely diverged by the early Cambrian, and represent considerably more ancient branches than seen among many of the vertebrate taxa examined to date. The first difference noted among echinoderm classes was the inversion of a 4.6 kbp fragment in asteroid relative to the echinoids (Jacobs et al., 1989; Smith et al., 1989). This region includes the lrRNA, ND1, and ND2 genes as well as a conserved cluster of tRNA genes. Two asteroid species, *Asterina pectinifera* and *Pisaster ochraceus*, belonging to the orders Spinulosida and Forcipulata, respectively, have been separated for at least 200 million years display the inversion



Based on these observations, Smith et al., (1993) examined the gene order in two of the remaining classes, the ophiuroids and holothuroids. The holothuroids *Cucumaria miniata* and *Parastichopus californicus* exhibited the sea urchin pattern while the ophiuroid *Ophiopholus aculeata* gene order is the same as the sea stars, with respect to the inversion event. Further gene rearrangements were subsequently detected in the brittle star, *O. aculeata* including a significantly modified tRNA gene cluster (Smith et al., 1993). Several independent rearrangement events appear necessary to relate this novel pattern to the conserved cluster seen in asteroids and echinoids.

The initial mtDNA sequence analysis of *C. pseudocurata* and *C. miniata* indicated a substantially different gene order in the tRNA cluster of these two species compared to the conserved tRNA cluster seen in urchins and sea stars. An effort to describe these differences and provide a potential explanation of how these differences arose was undertaken. This system provides an opportunity to examine potential mechanisms of gene rearrangement in a system presumably lacking intermolecular recombination. Furthermore, once described, this derived gene order could potentially be utilized as a character for phylogenetic analysis at deeper levels such as between families.

## Materials and Methods

### *Cloning*

For the purposes of cloning, mtDNA was purified according to the procedure of Lansman et al. (1981) with some modification. *C. miniata* tissue was homogenized using a Teflon dounce homogenizer in 20 ml of 210 mM mannitol, 70 mM sucrose, 50 mM Tris-HCl, pH7.5 (MSB) plus 3 mM calcium chloride. This homogenate was centrifuged at 1000xg for 20 min. to separate cytoplasm from nuclei. The supernatant was recentrifuged at 20000xg for 20 min. The resulting pellet containing mitochondria was rinsed with 20 ml of 100 mM NaCl, 50 mM Tris-HCl pH 8.0, 100 mM Na<sub>2</sub>EDTA (STE buffer) and recentrifuged. This pellet was resuspended in STE buffer and sodium dodecyl sulphate (SDS) was added to a final concentration of 1% to lyse the mitochondria. The crude mtDNA fraction was brought to 1M CsCl and 0.3 mg/ml ethidium bromide and centrifuged at 100,000xg for 72 hours to separate mtDNA from genomic DNA and RNA. The tubes were carefully placed in UV light and the mtDNA band extracted by puncturing the tube with a needle and syringe. The mtDNA was precipitated with ethanol and resuspended in 10 mM Tris-HCl, pH 7.5, 1 mM Na<sub>2</sub>EDTA (TE).

Aliquots of mtDNA were digested with various restriction enzymes and electrophoresed on 1% agarose gels to determine suitability for cloning. The enzyme HindIII was chosen as it appeared to cleave the mtDNA and digested aliquots were ligated into the HindIII restriction site of the plasmid pUC18. Ligation reactions were used to transform competent *Escherichia coli* cells, strain DH5 $\alpha$ , which were plated on

LB plates containing 100 µg/ml ampicillin to select for cells containing plasmids. Plates also contained 40 µg/ml 5-bromo-4-chloro-3-indolyl-β-D-galacto-pyranoside (x-gal). X-gal is included to distinguish colonies containing plasmids with mtDNA inserts from plasmids without inserts. Colonies with plasmids containing mtDNA inserted into the β galactosidase gene appeared white due to loss of enzyme function while those without mtDNA inserts appeared blue due to the metabolism of x-gal. IPTG was also included to induce expression of the β galactosidase gene. Plasmid was subsequently purified from several white colonies according to the alkaline lysis method (Birnboim, 1983), digested with HindIII, and separated by electrophoresis on 1% agarose gels in 40 mM Tris-Acetate, 2mM EDTA (1X TAE) to determine the presence and size of mtDNA inserts.

Sequencing and analysis of the isolated clones was accomplished by following standard procedures outlined in Chapter I using the M13 forward and reverse primers which have complementary sequences to sites flanking the cloning site in the pUC18 plasmid.

### *PCR Amplification*

In order to determine gene order, several primer designed to anneal at anticodon loops of several tRNA genes or conserved regions of the COI, ND1, and ATP6 genes (Table 8, Chapter II) were used in a variety of combinations. Subsequently, a tryptophan tRNA gene primer was used in combination with standard M13 forward and reverse

plasmid primers in an attempt to locate the tRNA genes which had presumably been relocated and were no longer present between the srRNA and ND1 genes.

### *Secondary Structure of tRNAs*

tRNA genes were identified by sequence identity of the anticodon loop and by the potential of the putative transcribed DNA sequence to be folded into the classic cloverleaf tRNA secondary structure. Attempts to use computer programs to generate tRNA secondary were largely unsuccessful due to the apparently less constrained forms of the sea cucumber mitochondrial tRNAs. The secondary structures were therefore produced manually by comparison to other mitochondrial tRNA structures, particularly those of the sea urchins *S. purpuratus* (Jacobs et al., 1988) and *P. lividus* (Cantatore et al., 1989).

### *Phylogenetic Analysis Using tRNA Gene Sequence*

The tRNA genes appear to be among the most highly conserved of the mitochondrial genes particularly with respect to secondary structure (reviewed in Wolstenhome, 1992). Kumazawa and Nishida (1993) utilized tRNA gene stem positions for phylogenetic analysis of vertebrate groups and suggested this approach may be useful for deeper levels than is possible with other mitochondrial genes. The authors also pointed out that stem comparisons avoid the problem of alignment ambiguities often encountered in comparisons of loop positions. Gene sequence was obtained for a set of ten tRNA genes from *C. miniata* and *P. californicus*. Sequences were aligned according to secondary structure. A preliminary phylogeny was constructed for these two sea

cucumber species as well as the sea urchins, *S. purpuratus* and *P. lividus*, two sea stars, *Asterina pectinifera* and *Pisaster ochraceus*, with *Xenopus laevis* as an outgroup.

Comparisons were based on all changes observed or on transversions alone, and both the entire tRNA gene sequence as well as only stem positions were analyzed by the maximum likelihood method (DNAML; PHYLIP 3.5, Felsenstein, 1993).

## Results

### *C. miniata* mtDNA Clones

Four *C. miniata* mtDNA HindIII restriction fragment clones with insert lengths of 1.5, 2.3, 2.6, and 3.9 kilobase pairs (kbp) were isolated and designated pCmhd15, pCmhd23, pCmhd26, and pCmhd39, respectively. Standard plasmid M13 forward and reverse primers were used to sequence the ends of these clones. Sequences were compared to an echinoderm mtDNA library in order to identify any gene fragments present.

A cluster of 13 tRNA genes is located between the srRNA gene and the NADH dehydrogenase subunit one (ND1) gene in sea urchins (Jacobs et al., 1988; Cantatore et al., 1989). This cluster is preceded by the threonine and glutamate tRNA genes as well as the replication control region (Figure 19). Sequence analysis of a portion of this region from one sea cucumber species, *P. californicus*, indicated that the portion of the gene cluster between the srRNA and the methionine tRNA genes was also conserved in

Figure 19. Arrangement of clustered tRNA genes in sea cucumber mtDNA. The sea urchin *S. purpuratus* tRNA gene cluster (Jacobs et al., 1988) is shown in the top panel. The tRNA genes are labelled as their corresponding single letter amino acid. Sequences for *P. californicus* and *C. pseudocurata* were obtained from a PCR fragment spanning the region between the srRNA and methionine tRNA genes. Consequently, the remainder of the cluster is not shown for these species. *C. miniata* sequence was obtained from a PCR product spanning the region between the srRNA and ND1 genes, encompassing the entire cluster. Hatched regions indicate short regions of unassigned sequence, presumably representing degenerated copies of the tRNA genes. The stippled area represents the putative control region and is listed as approximate for *C. pseudocurata* since it was not possible to distinguish between the end of the control region and any remaining portion of the degenerated proline tRNA.

*Strongylocentrotus purpuratus*

srRNA	E	T	120	P	Q	N	Lc	A	W	C	V	M	D	Y	G	Lu	NDI
-------	---	---	-----	---	---	---	----	---	---	---	---	---	---	---	---	----	-----

*Parastichopus californicus*

srRNA	E	T	475	P	Q	N	Lc	A	W	C	V
-------	---	---	-----	---	---	---	----	---	---	---	---

*Cucumaria pseudocurata*

srRNA	T	~405	Q	A	C
-------	---	------	---	---	---

*Cucumaria miniata*

srRNA	T	397	Q	A	C	M	D	Y	G	Lu	NDI
-------	---	-----	---	---	---	---	---	---	---	----	-----

members of this class of echinoderms (Smith et al., 1993) (Figure 19). DNA sequence was also obtained for this region from *C. miniata* (Figure 20) and from *C. pseudocurata* (Figure 21). However, DNA sequence analysis of this same region demonstrated that five of the tRNA genes, those for glutamate, proline, asparagine, leucine<sub>CUN</sub>, tryptophan, and valine, were no longer present in the cluster between the srRNA and ND1 genes (Figure 19). This results in a modified cluster that includes a short region of unassigned sequence (UAS) between the srRNA and threonine tRNA genes, a putative control region followed by the glutamine tRNA gene, another short UAS, the alanine tRNA gene, another short UAS, the cysteine tRNA gene, and another short UAS. In order to determine the extent of rearrangements throughout the cluster, a PCR product of approximately 1400 nucleotides (nt) in length was amplified from *C. miniata* using the 12Scuc and ND1 primers. Sequence analysis indicated that the five tRNA genes adjacent to ND1 in sea urchin mtDNA are also present in *C. miniata* mtDNA in the identical order and orientation (Figures 19 and 20). The relocation event(s) has therefore been restricted to the region between the srRNA gene and the methionine tRNA gene with only the five previously mentioned tRNA genes having been relocated.

In an attempt to determine the events underlying this rearrangement, an oligonucleotide primer was designed from the tryptophan tRNA gene sequence of the sea urchin, *S. purpuratus*. This tRNA appeared to be one of the most highly conserved of the relocated tRNA genes based on alignments of several sea urchin and sea star species (not shown). This primer was used in combination with forward or reverse universal plasmid primers in an attempt to identify and amplify a fragment containing this tRNA



Figure 20. *C. miniata* Mitochondrial DNA Sequence Between the srRNA and ND1 Genes. Identified genes are in bold type, with name and boundaries indicated above. In the case of ND1 gene, the initial methionine codon is underlined and the putative amino acid sequence is also indicated above. Intervening or unassigned sequence is in regular type. The tRNA genes are identified by the corresponding three letter amino acid label with the anticodons underlined.

```

-----3' End srRNA Gene-----
TACACATCGCCCGTCACTCTCATCAAAAAGAGAGAAAAGTCGTAACATAGTAGATGTACCGGAAGGTGCCCTCTAGAAAAT      80
-----]-----
TATTAAAATTAAAACACAGAAAACCTCCAAAAGTCTCAGTTCAGGTTACACCTATAATAATTCGCTATAAGTAGGATACAAA    160
----- THR -----]
CTTAGTCTTGTAACTAAGTTAGAAAAGTTAAACTCTTTCATATAGCTAACAAAATAAAAACCCCTACCAGCCCTCCGGCTCA    240
ACCTAAACTGAGTAAAAAAAACCTTCAAAAACCTAACATATCCCAAAAACCTTAAATCATCCCTTCCCCCCTTAAGAAAAGATT    320
AAAAACTTGGGCTTCCCCCTATATAAACCCTTCCATGATCCTTGAGTACCATCCGGGCTTAACCCACTTTTCAATAGT      400
CTTTTCTTACCTTTGGGGCTCAATATCATTTATTCCTCTCCTCGTGATTTCCATATTAACCTTGCTATATTTCITTT      480
ACTCTCTGCTCCTTACCCTCGCACATACCTATATAATTCACCTTGTGGAGAAAAGAGAGGGTAGACCGGGGGGGGGATGTA    560
TCGTTTATATAAATAATATGTTATTTATATAATAAACAAGAAACTAGTTTATTAATAAATAGCTATCTTGAGGGCAC      640
-----]----- GLN -----]
AGACAAAATATCTTTTACCCCTTGAAGAAGATGAGAAAACGACCTCATATCAAGGGTATCAAAAATCCCTCATTTTTCATT    720
-----]-----
AAAAATACTTTTC AACCTACAGAAAACCCCTTTGTGGATAAAAACCTTACTATATGGAAACATAACAACCTAAAAACCCCTAAT    800
-----]----- ALA -----]
ATTATTTCTGTATAAACA TAGACTCCAAAACGTGATTTCCAAAATAATCTTTGGATCATTTGCATGCAAAGCAAAATATTTT    880
-----]-----]----- CYS -----]
TATAAACTAAAATCACCCAAAATAAACA AAAAACAATACTGCAGGTTCTAAGGTGATTTACACATATTTAATTGCAAAAT    960
-----]-----]
TAAAAGTTACTAAAAAAAATTTAGTTAGAACTTCATAAGCTTCCCCCGAAACTAACCCACAATAAAAATCTTCATTTCTTATTT    1040
-----]----- MET -----]
ATATAAACACCGTAAAGTAAAGCTAACCTTAAGCTTTTAGGCTCATATCCCAAGAAATGGGAGATAAAAAC TCCCCTTTACTATA    1120

```

```

[-----ASP-----]
CCAAATTATTCGAGATCTACAGGGTCTAACCTGGGACTTTGGTCTGACAGTCCCAACGTTTACTAAACTAAAACCTCC 1200
[-----TYR-----]
CGACAAGATGGCTGAAAAAGCAAAGTTTGTA.AAACCTTTATGTAGGCACCCAGCCCTACTCTTTCGTCACGTTATTAGT 1280
[-----GLY-----]
ATAGAAAATATCTAGCTTCCAACTAGAAAGACCTTGCCCACTCACCCCAAGATTACGTTAAACACACACCTTAAATAGCAAAGCG 1360
[-----ND1-----]
[-----LEU-uur-----]
GTAAATGCAGAGGACTTAAAGATCCCCCTACTAAAGGTTTCGATTCTTTTTTAAGGTTAATGAGAACAGCAGCATTACTATCC 1440
[-----M S T A L F T I-----]
Q A L I Y I I P I L L S V A
AAGCACTAATATATAATACCTATTTTACTATCCGTCGCC 1481

```

Figure 21. *C. pseudocurata* nucleotide sequence between the srRNA Gene and the Methionine tRNA Gene. Identified genes are indicated in bold, unassigned regions in regular type. Labels above sequence indicate genes and their boundaries with tRNA genes identified by their corresponding amino acid with anticodons underlined. Genes transcribed from the opposite strand are indicated by labels ending with a lower case c. The precise end of the srRNA gene was not determined precisely and is indicated with a ? based on the last highly conserved region compared to *C. miniata*.

```

-----3' END srRNA Gene-----?
GGTACCCCTACACATCGCCCGTCACTCGTCAAAGAAGAGAAAAGTCGTAACATAGTAGGGGTACCCGGAAGTGCCCTCT 80
AGAAAATATATAAAATATAATTTGAGGCTTTC AAGCCCTCAAAAATGGGTCTGTCCCGTAATAATATGCTATAAAGTAGG 160
-----
      [-----]
AACAAAACCTTAGTCTTGTAAACTTAAGTTAGAAAAGTTAAAACTCTTTCGTATAGCTAACAAAGACAAGTCAAAATAACCTTAC 240
CATTTTAGACAAAATAAACATAAAAAGCCTTTAAATTTATCTTTTACAAAATACAGATCATTCAATTTATCGCTTAAAG 320
AAAAATTTAGAACCTTAAAAATAAAACTCAGAAATTTGGGCCCTCCCCCTTAGCCCTTACCCCCCTTCCAAGATCCTTGAGTAC 400
CATCCAGGCTTATCCCTTTTACCCTATAGTCTTTTTTTTTTATACCTTTGGGGGGTTACATCAATTTATCCCTATCTTCGGGATT 480
TCCACTTAACTTGTTGTTCTTTCCTCATGTCTCTGTCTTCCCTTGGCCCTCACACGTACCTATAATATTACCTTTGGAGAAAGTAG 560
GGGTAGACCCGGGGGGGGTGTATGTTATAATAATAAGAGCGACTCATTAACAATCCTTAAGAAATAGGGGCTAACGA 640
-----
      [-----]
CCCTTATTCCTACTTTGAAAAAATGAGAAGCGGACCCTCATATCAAGGGTATCCAAAATCCCTCATTTTTCATTAATAATAC 720
-----]
TTTTTCAATTAGATACTTAAGACACTTTTGTGGTACAGATTTACTACACAAAATACAAAATAGAATAATTAACACTCCCCCT 800
      [-----]
CATACCAGACTATAACTTAAACCTGTGATTTCCAAAATATCTTTGGATCTTTTGCATGCAAAAACAATAATTTTGTAAAC 880
-----]
      [-----]
TAAAAATCACTAATGAAAATACCCCAAATTTTAGGTTCTAAGGTGTAACCCACATATTTAATTGCAAAATAAAAAGTTACTAA 960
-----]
ACAACATTTAGTTAGAACTTCAAAAGTTTAAACCCAAAACTAATCCTACATAAAGATAGACATTTATAATAAATACACGTAA 1040

```

AGTAA 1045

gene from one of four HindIII restriction fragment clones. This experiment was based on the fact that the four HindIII clones represent more than 10 kbp of the *C. miniata* mitochondrial genome and would therefore have a high probability of containing the relocated tRNAs. One such attempt yielded a 1.8 kbp PCR product from the clone pCmhd23 (Figure 22) using the tryptophan and M13-forward primers. Analysis of DNA sequence obtained with the forward primer indicated homology to the ATP6 gene from sea urchin (Figure 23). Attempts to identify genes from sequence obtained with the reverse primer demonstrated the presence of a putative glutamate tRNA gene (Figure 23).

An ATP6 oligonucleotide primer (Table 8) was designed and used in combination with a COI primer in an attempt to PCR amplify the entire region surrounding the tryptophan tRNA gene from total mtDNA. A PCR product 4.8 kbp in length was obtained (Figure 24). Several oligonucleotide primers were designed for various tRNAs (Table 8) in order to obtain DNA sequence for this region (Figure 25). From this sequence information, it was possible to identify the five relocated tRNA genes between the COI and ND4l genes (Figures 25, 26). The glutamate tRNA gene is adjacent to a second copy of the putative control region, followed by the proline tRNA gene, a short UAS, the asparagine and leucine<sub>CUN</sub> tRNA genes, another UAS, the tryptophan tRNA gene, another UAS, the valine tRNA gene, another UAS and finally the 5' end of the ND4l gene.

All 13 tRNA genes from the two separated clusters in *C. miniata* mtDNA have retained the same transcriptional orientation relative to the original cluster in sea urchins. The glutamate and threonine tRNA genes have also maintained the same transcriptional

**Figure 22. PCR amplifications using pCmhd23. A standard alkaline lysis preparation of the plasmid pCmhd23 was diluted one in one hundred for use as a PCR template.**

**Reaction conditions were: 95<sup>0</sup> C, 30 s; 50<sup>0</sup> C, 30 s; 72<sup>0</sup> C, 60s for 25 cycles. Lanes are identified according to the primer combination used. Lane 1, M13 forward-Tryptophan tRNA; Lane 2, M13 forward-M13 reverse; Lane 3, One kb ladder; Lane 4, M13 reverse-Tryptophan tRNA.**

1 2 3 4

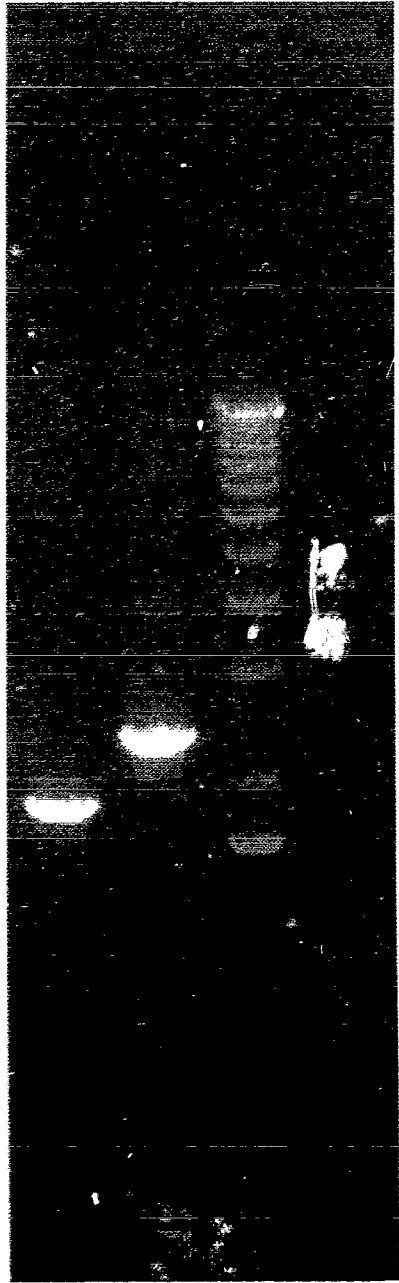




Figure 23. DNA sequence from the mitochondrial restriction-fragment clone, pCmhd23. Sequence from pCmhd23 was obtained using standard M13 forward and reverse primers. Sequence using the reverse primer is shown aligned with the glutamate tRNA gene sequence from the sea urchin, *S. purpuratus*. (Jacobs et al., 1988). The tRNA gene sequences are indicated in bold type with the anticodons underlined. Sequence obtained with the forward primer is aligned with ATPase 6 gene sequence from *S. purpuratus* (Jacobs et al., 1988). The putative amino acid sequences for this gene are also shown.

pCMHD23RN ACTGCTAAAGTTTGA-AAAAC-GAAGGCTTTTCAAGCCTTAGACCCAGGTTAAACCCCT 60  
 S.p. tGLU GCCCCTATAGTTGAAACACACAACCAAGAGCTTTTCACGCTCTTAAGTTTGAGTTAAAAATCT 60

pCMHD23RN GGTAGGAGTGATGTTACCCCAAGCTAATAAAAATAAAGCCCTACCAGCCCTCGCTTAACC 120  
 S.p. tGLU CAATAGGAGCT

pCMHD23RN TAAACTGAGTAAAAAACCTTCAAAAACCTTAACAATCCTCCAAAACCTTAAATCATCCCTTCC 180  
 pCMHD23RN CCCACTTAAGAAAGATTAAAACCTTGGGCTTCCCCCTGTATAAACCCCTTCCAATGATC 240  
 pCMHD23RN CTTGAGTACCATCCGGGCTTAACCACCTTTCAATAGTCTTTTTTCTTTACITTTGGGGC 300

pCMHD23FN CTAATCCCATTTATGGTCTGAATAGAAAACCCCTAAGTCTTTTTTGGCTCAACCCCTTAGCTCTA 60  
 L I P F M V W I E T L S L F A Q P L A L

SPATP6N CTAATACCATTAATGGTCTGAATAGAAAACACTAAGTCTATTTGGCAACCCCATAGCCCTA 60  
 L I P L M V W I E T L S L F A Q P I A L

pCMHD23FN GGTTAGGACTAGCAGCCAAACCTCACAGCAGGACACCTCCTCTATTCCTATTTCAAGA 120  
 G L R L A A N L T A G H L L L F L F S R

SPATP6N GGCCTACGACTTGCTGCTAACCTCACAGCTGGCCACTTGCTAATTTTCTTCTATCGACA 120  
 G L R L A A N L T A G H L L I F L L S T

pCMHD23FN GCTGCCTGAGTCCCTAAGAAAACCTTATACTAAGAAGAATAACAATAGTATCTTAATT 180  
 A A W V L R N N L I L R R I T I V I L I

SPATP6N GCTATATGGTACTTTCTCCAGTTTGATGATTAGGGTCCCAATCCTTATTTTATT 180  
 A I W L L S S S L M I S V P I L I F I

pCMHD23FN TTAATATTCATTCTAGAAAATAGGAGTAGCCTGCATA 216  
 L I F I L E I G V A C I

SPATP6N TTATTATTGCTAGAAAATAGGAGTAGCCTGCATC 216  
 L L F V L E I G V A C I

Figure 24. PCR amplification products of *C. miniata* mtDNA between COI and ATP6 genes. Products are identified by the primer pair used in the amplification. Lane1, COI3b-ATP6; Lane2, COI3'-ATP6; Lane 3, COII-ATP6; Lane 4, One kb ladder.

1 2 3 4

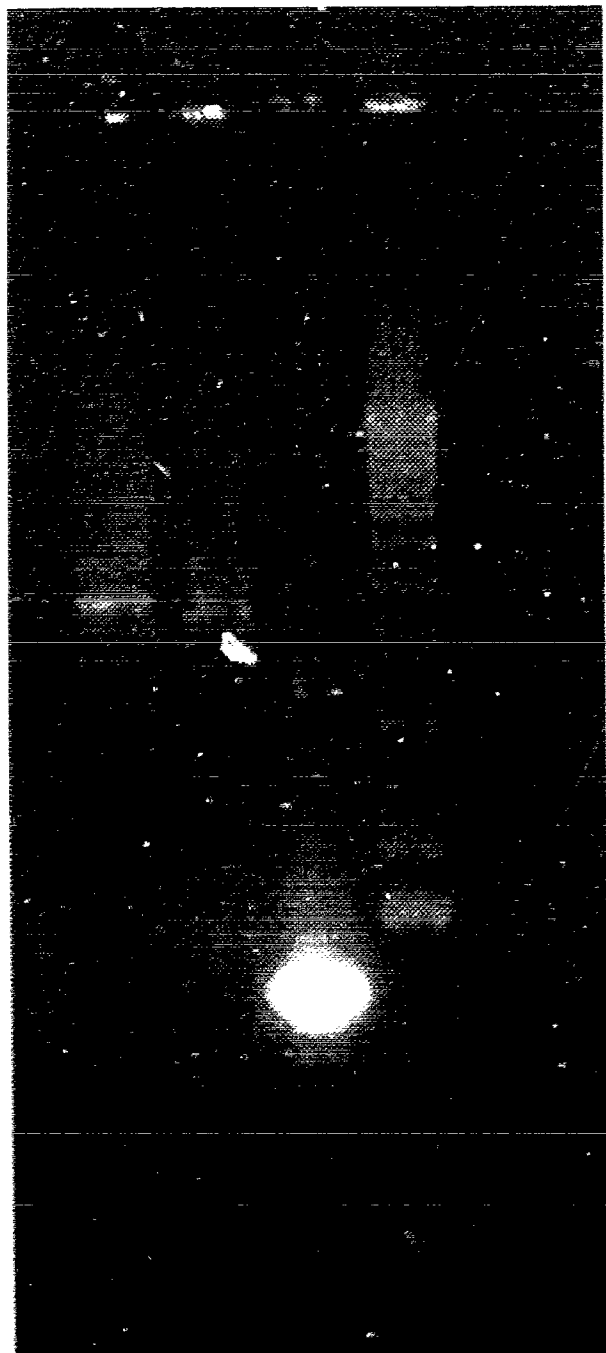


Figure 25. *C. miniata* mitochondrial DNA sequence between the COI and ND4l genes.

Identified genes presented in bold type with direction of transcription (arrows) and gene boundaries (square brackets) indicated above. Putative amino acid sequence for the COI and ND4L genes also indicated above DNA sequence. The tRNA genes are labelled by their corresponding three letter amino acid name, with anticodons underlined. Unassigned sequence in regular type. Copies of a 286 nt repeat identified next to the arginine tRNA gene are underlined. Complete sequence for this intervening region was not obtained due to technical difficulties (see text). Therefore, numbering of downstream sequence and the presence of four copies of the repeat were estimated from the size of PCR products spanning this region (see figure 24).



TAAAAAATAAAAATTATTTTATAAAGGAGGGGACCCCTTATAATTTTCCAAAAGATAAACTTAACCCCTACCC 1360  
 [-----]  
CCTACCACTAAACTATTAATAAAGCTTTACCCACTCTTTCAACACAAAGAACTGCTAAAGTTTAGAAAAACGAAGGCTT 1440  
GLU → [-----]  
TTCAAGCCCTTAGACCCAGGTTAAACCCCTGGTAGGAGTGATGTTCAACCCAAAGCTAATAAAATAAAGCCCTTACCAGCCCT 1520  
 CGCTTAACCTAAACTGAGTAAAAAAACCTTCAAAACTTAACATCTCCCAAACCTTAAATCATCCCTTCCCCCCACTTAAGA 1600  
 AAGATTAAAAACTTGGGCTTCCCCCTGTATAACCCCTTCCATGATCCTTGAGTACCATCCGGGCTTAACCACTTTTC 1680  
 AATAGTCTTTTTTCTTTACTTTTGGGGCTCCAATAATCAATTTATTTCCCTCCTCGTGATTTCCATATTAACCTTGCTAT 1760  
 TATTTCTTACTCTCTGCTCCTTACCCTCGCACATTCCTATATATTCACCTTTTGGAGACAGAGGGTAGACCTGGGGGG 1840  
 GGGTGATGGTTTTATAATAATAATAATTTATTTATAATTAATAATAACAGAAACTAGTTTATTAAAAATAGTTATTTTG 1920  
 [-----] [-----] --PRO-->  
GGGATAACAGATACaGACCTAGCCCTGTGTTCTGAAAATAAAACAGATCTTTAAAATGGGTTGAAGCTGAAACAATAGCAT 2000  
 TTGGCCGTTAACCAAAAGGATGAAGGTTAGTCCCTTTCAACCCAGCAAGTTTAAAATAGCAAAAATAGTAATGCATTAGATT 2080  
LEU-cun → [-----] [-----] --ASN--> [-----] [-----] --TRP-->  
TAGGTTCTATATTAAGGTTGCAAAATCCCTTTTAAACTTTTGGCCACACTAAAATTTTAAATTTAAACnATATAGAAAC 2160  
 [-----] [-----] --TRP-->  
 TTTCCCCCCCCCCCCAAACC GAAGATTTAAGTTAAATAAAACTTTGAGCCTTCAAGCTCACCCACACAAAATCAAATTTTGT 2240  
 [-----] [-----]  
AATCTTTGGACTATCCAAAACCTTCAACCAACTACAGCACAAACCCACTAAATACTAGTACAAAATCTTCAAAAACAGTTCCG 2320  
 [-----] [-----] --←-VAL-->  
 AGTTGCACGAACCTTAACCTCTGTGTAAGAGTGGCTTACTCCCTAGCTATATTCTGTATTTATAATAATACCCCAAGACA 2400

[ ND4L-→-----  
 M N I L T S I L 2480

AACCCCTATTAAAAAAGAGGTGCTGTTAAACTTATACCACAAAATACGAAAAATGAACATACTAACCAAGAATTTTAT

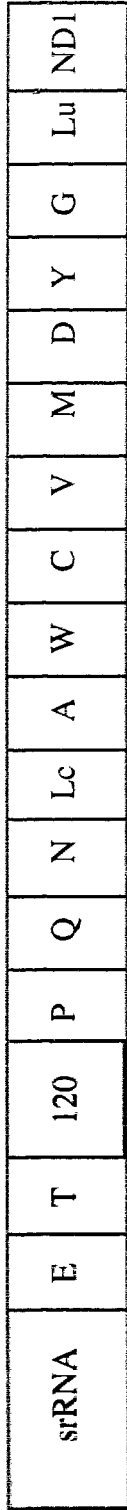
-----  
 F S L F F L G V L G V I N R L H L L S L L L C L E L 2560  
 TCTCCCTTTCTCCCTCGGGGTTTAGGAGTGGTTATAAAAACGGTTACACCTTCTCTCCCTTTTACTTTGTTAGAGCTA

-----  
 L L I S L F L N  
 CTTTAAATCTCCCTTTTCCATAAAA 2584

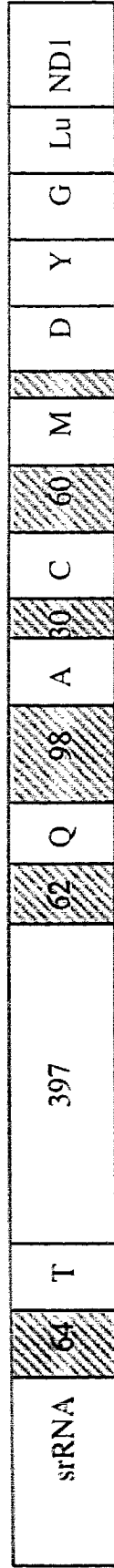


Figure 26. Arrangement of tRNA genes in *C. miniata* mtDNA. The sea urchin *S. purpuratus* tRNA gene cluster (Jacobs et al., 1988) is shown in the top panel. The tRNA genes are labelled as their corresponding single letter amino acid. *C. miniata* sequence was obtained from two distinct regions: The middle panel shows gene order spanning the region between the srRNA and ND1 genes; the lowest panel shows gene order between the COI and ND4L genes. Hatched regions indicate short regions of unassigned sequence, presumably representing degenerated copies of the tRNA genes. The stippled area represents the putative control region.

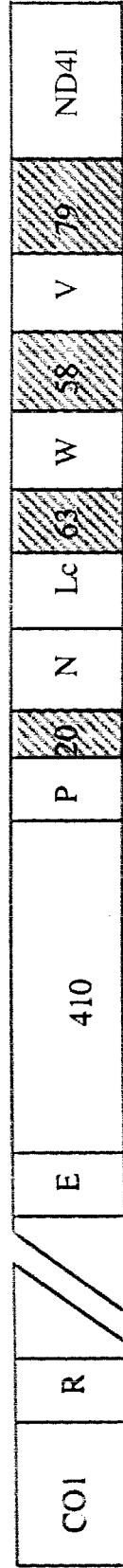
*Strongylocentrotus purpuratus*



*Cucumaria miniata*



*Cucumaria miniata*



orientation. Furthermore, the short UAS elements are found at complementary positions in the two clusters. Alignment of the clusters demonstrates that the location of a tRNA gene in one cluster, corresponds to a UAS element in the other cluster. Combining the two clusters produces a single functional copy of each tRNA gene in the same order and orientation as seen in the sea urchin cluster (Figure 26). It therefore appears likely that this gene arrangement has arisen through a single nontandem duplication. Attempts to align tRNA genes with putative degenerated copies showed no evidence of homology except in the case of the proline tRNA (not shown). The putative proline tRNA copy 5' of the glutamine tRNA gene has retained 70% identity to the proline tRNA gene sequence. Complete sequence was not obtained for the entire region between the arginine and glutamate tRNA genes due to the presence of a number of copies of a 286 nt repeat (Figure 25). The presence of this repeated element resulted in multiple binding sites for the internal primers cm21rc2c and cm21rc3c (Table 8) and consequently, rendered direct sequencing impossible. Based on the length of PCR products, and assuming that only copies of the repeat were present in this region, it would appear that there are four tandem copies of this element, although cloning and sequencing of this region would be required to confirm this.

In contrast to the degeneration of the putative duplicate tRNA genes, the two 400 nt segments 3' of either the threonine or glutamate tRNA genes have retained a high degree of sequence similarity to each other (Figure 27) and have not been reduced to a single copy. These two copies differ by only 3.0% if analysis is restricted to the first 420 nucleotides, suggesting some functional constraint. The first 240 nt of the copy

**Figure 27. Alignment of *C. miniata* sequence corresponding to the two copies of the putative control region duplication. Top sequence is from the PCR product spanning the srRNA and ND1 genes, bottom sequence is from the PCR product spanning COI and ND4L genes. Sequences begin just after the threonine tRNA gene for the srRNA-ND1 fragment and 16 nucleotides after the glutamate tRNA gene in the case of the COI-ND4L fragment with numbering referring to the position in Figures 20 and 25, respectively. Vertical bars denote nucleotide identity between the two regions. The proline tRNA gene in the COI-ND4L sequence is shown in bold with the anticodon underlined. The maintenance of the 3' portion of the proline tRNA gene sequence between the srRNA and ND1 genes represents the only recognizably conserved portion of any of the duplicated tRNA genes.**



between the srRNA and ND1 genes of the sequence shown in Figure 27 was also included in the population study. No decrease in the number of sites free to vary was observed in this region compared to the 160 nt 5' of this region corresponding to the threonine tRNA gene and the putative degenerated glutamate tRNA gene (Figure 9). Overall the region used for population analysis accumulated substitutions three times faster than the COI gene. Clearly some sites are free to vary in this region despite the apparent functional constraints.

In the sea urchin, *S. purpuratus*, the region between the threonine and proline tRNA genes is thought to encode the origin of replication. Jacobs et al. (1988) have presented a stem loop structure thought capable of initiation of replication. An attempt was made to manually align a stem-loop structure involving the polypurine (G) tract in *C. miniata* and *C. pseudocurata*. A possibly analogous structure, of similar dimensions to the urchin stem-loop, was detected in the duplicated region of *C. miniata* mtDNA 5' of both the glutamine and proline tRNA genes, having a 21 base pair stem and a loop of 37 nucleotides (Figure 28). This stem structure does involve a string of purine nucleotides as does the *S. purpuratus* stem, however there are four mismatch pairs in the *C. miniata* structure, compared to one for *S. purpuratus*. The *C. miniata* stem is composed of eight G-T, three A-T, and six C-G pairs, compared to seven G-T, one A-T and 13 C-G pairs for *S. purpuratus* indicating the considerably more stable nature of the sea urchin structure. A short sequence element was also identified 3' of the stem-loop structure for *C. miniata* which is quite similar to the stop sequence elements described for the mouse (Bibb et al., 1981. as shown in Jacobs et al., 1988), and the putative analogous elements

**Figure 28. Possible Stem Loop Structures for *C. miniata*. Numbers refer to sequence position in Figure 20. Nucleotides in brackets are the only substitutions observed in the duplication of this region between the arginine tRNA and ND4L genes (Figure 24).**





from sea urchin (Jacobs et al., 1988) are aligned in figure 29. A possible element was also detected in the *C. pseudocurata* sequence although similarity to the urchin and *C. miniata* elements is not as convincing.

An alternative stem loop structure for *C. miniata* was also detected using the computer program, RNAFOLD (Bairoch, 1993), 3' of the polypurine tract (Figure 28), which has only a single mismatch out of 14 base pairs. This structure was the most thermodynamically favorable stem detected, with a free energy value of  $\Delta G = -10.5$  kcal/mol. The putative stop element described above is contained in this alternate stem-loop. Only less convincing stem-loop structures could be formed from the *C. pseudocurata* sequence and are not presented.

### *tRNA Structures*

The putative secondary structures of the *C. miniata*, *C. pseudocurata*, and *P. californicus* tRNAs are shown in Appendix C, including the *C. miniata* arginine tRNA, which is immediately 3' of the COI gene (Figure 25). Also included in Appendix C is a reference tRNA structure indicating stem, loop and connector regions and abbreviations. Canonical pairing of seven, four, five/six, and five bases for the amino acid acceptor (AA), Dhu, anticodon (AC), and T $\phi$ C (hereafter referred to as T) stems, respectively, were observed in most structures. Only six of sixteen *C. miniata* tRNAs, and one of ten *P. californicus* tRNAs, had AC stems of six base pairs in length. Mismatch pairing of nucleotides was also observed in several of the cloverleaf stems in the structures

**Figure 29. Putative stop element sequences. Alignment of termination associated sequence (TAS) elements involved in arrest of replication at D loop (from Jacobs et al., 1988).**

---

Mouse	AATAGTTTAAATGT (Bibb et al., 1981)
	**** * ** *
<i>S.purpuratus</i>	AATAATATATAAT (Jacobs et al., 1988)
	***** ** *
<i>C.miniata</i>	AATAATATATTGT
	***** * *
<i>C.pseudocurata</i>	AATAAGAGCGACT

---

presented, however, numerous examples of similar mismatches have been described in other animal mitochondrial tRNA genes and in particular in the two sea urchins examined to date (Cantatore et al., 1988; Jacobs et al., 1988). Such mismatches shorten the T stem from five to four nucleotides in the proline, and tryptophan tRNA genes of *C. miniata* and in the cysteine tRNA of *P. californicus*. The Dhu stem has been reduced to three base pairs in the threonine, leucine(CUN), and valine tRNAs of *P. californicus* and in the leucine(CUN), valine, leucine(UUR), and tyrosine tRNAs of *C. miniata*. It was not possible to construct a convincing D stem for the threonine tRNA of either *C. miniata* or *C. pseudocurata* (Appendix C). Mismatches were also found at internal positions of stems. An A-A mismatch occurs at the base of the AC stem in the glutamine and asparagine tRNAs of *C. miniata*. Other unusual pairs seen in *C. miniata* tRNAs include a G-G mismatch in the AA stem of the glutamate tRNA, and A-C mismatches at the third or fourth base pair of the AC stem in the methionine and leucine(UUR) tRNAs, respectively. Among *P. californicus* tRNAs, an A-C mismatch occurs at the fourth pair of the glutamate AC stem, a U-C pair occurs at the third position of the proline AC stem, and U-C and U-U pairs occur at the second and third positions of the alanine AA stem (Appendix C).

### *Substitution Patterns in tRNA genes*

DNA sequence was aligned for a common set of ten tRNA genes from *C. miniata* and *P. californicus*, as well as the five tRNAs identified from *C. pseudocurata*, in order to gain a better understanding of substitution patterns within mitochondrial tRNA genes.

Pairwise nucleotide identity values were calculated for each tRNA (Table 15). The comparison of *C. pseudocurata* and *C. minata* demonstrated that these tRNA genes are very highly conserved at this level, with an average distance estimate of 4.0% for the five tRNAs examined. This is considerably lower than the COI gene sequence distance data from Chapter I, where a distance estimate of 11.1% was obtained, but is higher than the 2.0% observed in the lrRNA gene sequence data.

If one examines substitutions with respect to secondary structure, only five out of sixteen changes occurred at stem positions (Table 16) all of which are transitions and all but one involve a G-U pair in one of the species that maintains base pairing in the stem. The sole exception to this is an A-G transition at the base of the AA stem. This nucleotide is involved in an apparent mismatch pair with A in both species (Appendix C). Among the eleven changes observed in loop positions, four changes were transitions, three were transversions and four were insertion/deletion events (indels) (Table 16) with half of the observed changes occurring in the D and T loops.

The ten *C. minata* and *P. californicus* tRNAs used vary in sequence identity from 73.4%, for cysteine, to 88.6% for both asparagine and leucine(CUN) (Table 15). A distance estimate of 20.3% was calculated for the entire tRNA data set. This value is very similar to the estimates obtained with the COI or lrRNA gene sequence data of 22.3 and 24.0%, respectively. Pairwise substitutions according to position in the putative secondary structure were also determined (Table 17A,B). The number of substitutions at loop positions (66/264) indicated a higher proportion of variable sites compared to stem positions (76/424). Following the method of Wheeler and Honeycutt (1988) as outlined

Table 15. Nucleotide identity among sea cucumber tRNA genes (in percent)

tRNA	<i>C. miniata</i> <i>C. pseudocurata</i>	<i>C. miniata</i> <i>P. californicus</i>
Glutamate		78.2
Threonine	98.6	79.7
Proline		70.8
Glutamine	96.3	85.3
Asparagine		88.6
Leucine(cun)		88.6
Alanine	96.3	81.0
Tryptophan		87.1
Cysteine	92.7	73.4
Valine		88.0

Table 16. Substitutions in tRNA genes: *C. miniata* versus *C. pseudocurata*

<b>tRNA</b>	<b>AA stem</b>	<b>AC stem</b>	<b>D loop</b>	<b>Connector 2</b>	<b>T loop</b>
<b>T</b>	A <-> G				
<b>Q</b>	C <-> U				- <-> C U <-> C
<b>A</b>		C <-> U	U <-> C	U <-> A	
<b>C</b>			U <-> A U <-> C U <-> C	A <-> -	A <-> C - <-> C
<b>V</b>	G <-> A U <-> C				

Table 17A. Substitutions in tRNA Genes: *C. miniata* versus *P. californicus*

tRNA	Stem <sup>1</sup>								Total Compensatory			
	Ti/Tv <sup>3</sup>		D		AC		T		Ratio <sup>2</sup>			
	S <sup>4</sup>	D	S	D	S	D	S	D	S	D		
<b>E</b>	1	6				2		2	11	0.67		0.75
<b>T</b>	1		1			4		4	10	0.50	1.00	4.0
<b>P</b>		2			1	8		2	15	0.33	1.00	6.0
<b>Q</b>	1				1			4	6	0.50	1.00	2.5
<b>N</b>	2								2	na	1.00	na
<b>Lc</b>					1			1	2	0.50	na	1.0
<b>A</b>	2	2	1		1			1	9	0.60	1.00	na
<b>W</b>		2		2					4	na	1.00	na
<b>C</b>		4				2		4	10	na	1.00	na
<b>V</b>	2	2			1			2	7	1.00	1.00	
<b>Total</b>	9	18	2	2	5	16	4	20	76	0.70	0.96	7.2

<sup>1</sup> Stem labels refer to structural positions as shown in example tRNA in Appendix C.

<sup>2</sup> Compensatory substitutions(changes that maintain base pairing, including G-U)/total substitutions.

<sup>3</sup> Transition/transversion ratio.

<sup>4</sup> Stem changes are divided into two classes: Either a single(S) nucleotide or both(D) nucleotides of a stem pair were observed to have changed following Dixon and Hillis (1993).

<sup>5</sup> na indicates all observed changes maintain base pairing.



Table 17B. Substitutions in tRNA Genes: *C. miniata* versus *P. californicus*

tRNA	Loop <sup>1</sup>					Indels <sup>2</sup>	Total	Ti/Tv <sup>3</sup>
	C1	D	AC	C2	T			
<b>E</b>		1+1*		1*	2	2	6	1.0
<b>T</b>		1+5*				5	6	na <sup>4</sup>
<b>P</b>		1*	1	2	4	1	8	0.29
<b>Q</b>					2+2*	2	4	na
<b>N</b>		1+6*			4+1*	7	12	na
<b>Lc</b>		2	1	3	1*	1	7	1.0
<b>A</b>		2*	2		2	2	6	3.0
<b>W</b>		1		2+2*	2	2	7	0.67
<b>C</b>		1+2*		1*	2+1*	4	7	0.5
<b>V</b>	1*	1*		1		2	3	na <sup>5</sup>
<b>Total</b>	<b>1*</b>	<b>7+18*</b>	<b>4</b>	<b>8+4*</b>	<b>18+4*</b>	<b>28</b>	<b>66</b>	<b>1.2</b>

<sup>1</sup> Loop labels refer to structural positions as shown in example tRNA in Appendix C.

<sup>2</sup> Insertion/deletion events are indicated by \*.

<sup>3</sup> Transition/transversion ratio.

<sup>4</sup> Indicates transition/transversion ratio of infinity; no transversions observed.

<sup>5</sup> Transition/transversion ratio of zero; no transitions observed.

by Dixon and Hillis (1993), substitutions at stem positions were classified as compensatory mutations if an observed single substitution maintained allowable tRNA base pairing (including G-U), or if both nucleotides of a given base pair had been substituted and base pairing was maintained. A matrix can be constructed to demonstrate that single substitutions of this type occur randomly at a frequency of 0.125 while double events maintaining base pairing occur randomly at a frequency of 0.256 (Hillis and Dixon, 1993). Based on the observed substitutions, single and double compensatory events would have been expected to occur at random 2.5 and 14 times, respectively, compared to observed values of 14 and 54, respectively. Both these values are significantly different from expected random occurrence according to a Chi-square test ( $p < 0.01$ ). Put another way, 70% of single substitutions and 96% of double substitutions were compensatory.

The transition/transversion ratio of substitutions was 7.2 for stem positions, 1.8 for loop positions, and 4.7 for the complete gene sequence. Insertion/deletion events were only detected at loop positions where they constituted 42% of observed changes (Table 17B). As in the case of the *C. miniata*/*C. pseudocurata* comparison, the highest number of substitutions were recorded in the D and T loops (Table 17B). A substantial number of changes were recorded in what has been referred to as connector region two (Steinberg and Cedergren, 1994), or alternatively as the variable loop (Wolstenholme, 1992). One change was also identified in connector region one (Steinberg and Cedergren, 1994), which is usually two nucleotides in length. Expressing the number of substitutions as a fraction of the nucleotide positions examined results in values of 25/62 for the D loop, 12/38 for connector two, and 22/67 for the T loop. Table 18 lists all observed single

Table 18. Maintenance of base pairing in tRNA stems

Taxa	tRNA	Stem	Substitutions	Paired Nucleotide
<i>C. miniata</i> / <i>C. pseudocurata</i>	A	AC	C <-> U	G
<i>C. miniata</i> / <i>P. californicus</i>	A	AC	A <-> G	U
<i>C. miniata</i> / <i>C. pseudocurata</i>	Q	AA	C <-> U	G
<i>C. miniata</i> / <i>C. pseudocurata</i>	V	AA	G <-> A U <-> C	U G
<i>C. miniata</i> / <i>P. californicus</i>	V	AA AC	G <-> A U <-> C A <-> G	U G U

substitutions recorded and demonstrates that a G-U pair is always present in one of the species examined at such sites.

The tRNA gene sequence data was used for phylogenetic analysis of six echinoderm species, primarily to serve as an examination of the utility of this data set at deeper phylogenetic levels, as suggested by Kumazawa and Nishida (1993). Trees were constructed using both the complete gene sequence and stems plus anticodon loop sequence. Both neighbor-joining distance and maximum likelihood methods were used to construct trees based on both transitions and transversions or transversions alone. A distance matrix based on all positions (Figure 30) indicated that this data set was highly variant. Trees were generated using stem positions plus AC loop or the entire gene sequence and based on all changes, or transversions only, group the two members of each class together and support the presumed relationships among classes (Figure 31).

## **Discussion**

### ***tRNA Gene Arrangement***

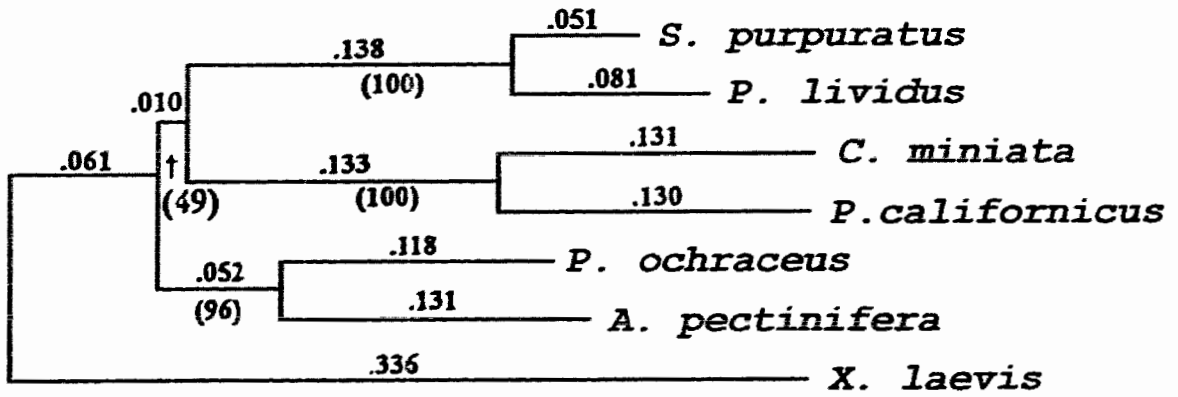
The tRNA genes appear to be among the most mobile elements of the mitochondrial genome (reviewed in Moritz et al., 1987; Pääbo et al., 1991; Smith et al., 1993). Movement of these genes appears to be have occurred by intramolecular transposition events or duplications (Cantatore et al., 1987; Moritz et al., 1987) involving

**Figure 30. DNA distance matrix using tRNA gene sequence. The method of Tamura and Nei (1989) was used to calculate the distance matrix based on 789 nucleotides from ten aligned tRNA gene sequences. Gap Sites were removed in the pairwise comparisons.**

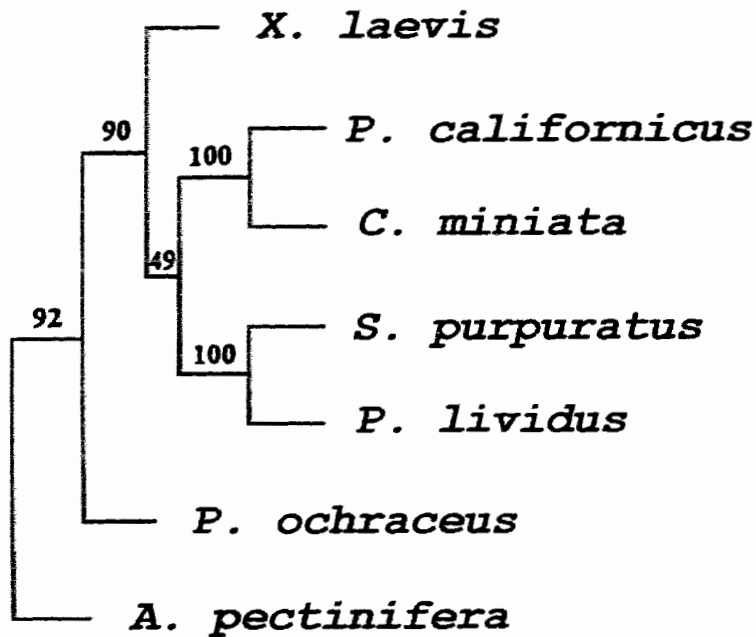


**Figure 31. Phylogenetic trees based on the tRNA data set. A: Neighbor-joining tree based on the Tamura-Nei distance matrix shown in Figure 30 using the entire gene sequences, excluding pairwise gaps. Branch lengths indicated above, bootstrap confidence estimates below in brackets. B: Maximum likelihood tree using only stem positions with bootstrap confidence values.**

A.



B.





one or more genes. In fact, several examples of mtDNA duplications have been reported, but most of these appear to be the result of tandem duplication events. Moritz and Brown (1987) characterized duplications in mtDNAs from seven lizard species. All duplications were tandem, direct, and either included, or were adjacent to, the control region. The authors also noted a nonrandom distribution of the duplication boundaries with most appearing to align with tRNA genes. However, the most parsimonious explanation of the novel mitochondrial gene order of *C. miniata* appears to be an ancestral nontandem duplication event involving the tRNA cluster region with subsequent drift randomly reducing the two cluster copies to a single set of functional tRNA genes. This is consistent with the short segments of noncoding DNA present between several of the tRNA genes in the mtDNA of the *Cucumaria* species, despite the fact that such regions are very rare in mitochondrial DNA (Moritz et al., 1987).

Of all the putative duplicate tRNA gene copies, only that region corresponding to the degenerated proline tRNA gene, between the srRNA and ND1 genes, retains any recognizable similarity to the functional copy, identified between the arginine tRNA and the ND4l genes. These two regions exhibit 70% nucleotide identity and in fact exhibit 100% identity over the first 30 nucleotides. This would seem to suggest that the duplication event is not of recent origin and that considerable time has since past, allowing most extra gene copies to become reduced to unrecognizable noncoding regions. This is also consistent with the fact that both *Cucumaria* species examined shared this duplication and also points to the potential utility of this gene arrangement as a phylogenetic character at deeper levels within the Holothuroidea.

The duplicated region between the threonine or glutamate tRNA genes and the remainder of the clustered tRNA genes has not been reduced to a single copy, and exhibits only 3% divergence. By comparison, this level is even lower than the 4% divergence observed between *C. miniata* and *C. pseudocurata* for the gene sequence of five tRNAs. It is possible that some type of homogenization process is maintaining the two copies nearly identical or, alternatively, that considerable functional constraints are imposed on this region. If this region has retained functional homology with that of sea urchins, it represents the putative replication control region (Jacobs et al., 1988). The authors proposed a stem-loop structure for mtDNA from *S. purpuratus* involving a polypurine string as an analogous structure to the cloverleaf involved in replication control in vertebrates (Clayton, 1982; G. Brown et al., 1986). These elements are suggested to provide recognition or regulatory sites for RNA/DNA polymerases. Adjacent to the proline tRNA gene copies is a region containing a string of 12 G nucleotides and an AT rich region. A potential stem-loop structure can be formed involving the poly G tract of *C. miniata*, and is followed by a potential TAS element. The *C. miniata* stem structure has considerably fewer G-C pairs and three more mismatches than the *S. purpuratus* structure, and appears considerably less stable. The similarities of this region at least suggest the possible involvement of this segment in the control of replication but more direct evidence of involvement in replication is required to confirm such a functional role. It is curious that no convincing stem-loop structure could be identified for *C. pseudocurata* and that the TAS element does not share higher identity with that of the putative element from *C. miniata*, given that the TAS is presumed to be a highly

conserved' element. However, since mtDNA sequence was not obtained for the duplicate region from *C. pseudocurata* between the COI and ND4I genes, it is possible that only a single functional replication control region has been retained in this species. Obtaining mtDNA sequence for that portion of the mitochondrial genome from *C. pseudocurata* could provide important insight into the functional requirements of mitochondrial replication and transcription in *Cucumaria*.

The tandem duplications observed in lizard mitochondrial DNA (Moritz and Brown, 1987) were also associated with the control region and tRNA genes were also commonly present at the boundaries of these regions and may therefore have arisen by similar mechanisms. However, the rearrangement in *C. miniata* mtDNA is distinct, since the duplicated regions are in the same strand polarity but are separated by several genes (ND1, ND2, COI, and the arginine tRNA), encompassing 3 kbp of DNA, and would appear to have required a duplicative transposition event. It is not known how such a duplication would affect replication but may simply result in alternate possible sites of initiation. In fact, this may provide a replicative advantage compared to a molecule with only a single origin.

A second function of this region may be as a transcriptional control element for the adjacent tRNA genes. Such a role for this noncoding element could also explain the maintenance of these two nearly identical copies since initiation or control of transcription may be required from both sites in order to transcribe all essential tRNAs. Were this region not required for some such function, it seems highly unlikely that both copies would have been maintained given the observed drift in the duplicated tRNA genes

themselves. Whether or not this region has such a functional role would require characterization and localization of the elements necessary for initiation of transcription in *C. miniata* and *C. pseudocurata* mitochondrial DNA.

### *tRNA Structure*

Among the structures presented for the tRNAs, mismatches were detected but numerous cloverleaf structures proposed for sea urchin tRNAs (Jacobs et al., 1988; Cantatore et al., 1989) or for *Xenopus laevis* (Roe et al., 1985) involve similar mismatches. Analysis of mature tRNAs is required to confirm whether these mismatch pairs exist in the mature structure. Base modifications are known to occur which stabilize tRNA structure (Derrick and Horowitz, 1993; Arnez and Steitz, 1994) and may allow unusual base pairs to form. One such example is the dimethylation of G at certain positions which alters the preferred base pairing from G-C to G-A (see Steinberg and Cedergren, 1995). The suggested recognition sequence for the dimethyl transferase is two sets of G-C pairs with one of these sets immediately 5' of the G nucleotide to be modified, however, no such recognition element was detected for any of the sea cucumber mismatches.

Recently (Steinberg and Cedergren, 1993) an analysis of atypical mitochondrial tRNAs lead to the proposal of a simple set of rules for these unusual structures. Given an AC stem of N base pairs, connector one (see Appendix C) was a minimum of 8-N nucleotides and connector 2 was a minimum of 9-N nucleotides. According to this rule, tRNAs with a five base pair AC stem would require connector regions of three and four

nucleotides, respectively. However, a five base pair stem with a two nucleotide connector one appears to be the most common arrangement in the sea cucumber tRNAs (Appendix C) with 7/15 *C. miniata* and 7/10 *P. californicus* tRNAs having this arrangement. The *C. miniata* glutamine tRNA has an AC stem of six and only a one nucleotide connector one and the valine tRNA has a five base pair stem with only a single nucleotide connector one.

One of the most uncertain structures presented is the threonine tRNA. The D stem for both *Cucumaria* species is depicted as two base pairs with an intervening mismatch, and the AA stem has a purine-purine mismatch as the basal pair. The AC stem is somewhat ambiguous as it can be folded into five, six, or seven base pairs. Steinberg and Cedergren (1994) suggested that a seven base pair AC stem and a one base pair D stem for the serine(GCU) tRNA from *S. purpuratus* could adopt a conventional three dimensional structure with smaller connector regions. It is therefore possible that the extra length of the AC stem compensates for the poor D stem but connector one remains two nucleotides and this does not resolve the question of the mismatch in the AA stem. The structures for *P. californicus* and two urchin species have only three nucleotide base pairs in the D stem while the reported gene sequence for *Pisaster ochraceus* threonine tRNA also involves two base pairs and an intervening mismatch. Despite these structural problems, the threonine gene sequence from the two *Cucumaria* species are nearly identical, with only a single observed substitution, which coincidentally, corresponds to the 3' position of the AA stem mismatch. This level of substitution constraint suggests that this is the functional threonine gene, but without

complete mitochondrial genome sequence from at least one of these species, the possibility of an alternative functional threonine gene copy with a more conventional structure cannot be ruled out.

*Substitution Patterns in tRNA Genes: C. miniata vs. C. pseudocurata*

The comparison gene sequence from five tRNAs revealed several interesting features. Although only 15 changes were observed, strikingly different patterns of substitutions were detected in stem versus loop positions. As expected, observed changes at stem positions maintained base pairing through the occurrence of a G-U pair in one of the species (Table 18). All changes detected at stem positions were transitions which is consistent with the constraint of maintenance of base pairing. For example, a G-C pair may be changed to a G-U pair through a C→U transition. Substitutions were twice as common at loop positions and transversions and indels were equally common. The five tRNA genes identified in both *Cucumaria* species differed by 4.0%, indicating that these genes are accumulating substitutions at a considerably lower rate than the COI gene (11.3%). An expanded set of tRNA genes would therefore allow analysis at phylogenetic levels that the COI gene is no longer useful, as was suggested by Kumazawa and Nishida (1993), such as among sea cucumber families. An analysis of the evolution of developmental patterns in sea stars (Hart, Byrne, and Smith, unpublished), successfully utilized a combination of COI and tRNA gene sequence, where it was not possible to completely resolve all relationships using either data set alone.

*Substitution Patterns in tRNA Genes: C. miniata vs. P. californicus*

Considerably more changes were observed between the gene sequences of ten tRNAs from *C. miniata* and *P. californicus*, although similar patterns were observed. Transitions dominated at stem positions ( $T_i/T_v = 7.2$ ) and 14/20 single substitutions and 54/56 double substitution events were classified as compensatory changes. A recent study by Ortí et al., (1996) described similar patterns of compensatory mutations in the ribosomal RNA genes of Piranhas mitochondria and suggested a weighting factor of 0.5 for such positions. It therefore appears that a weighting factor of close to 0.5 seems appropriate for tRNA stem positions also. The presence of a G-U pair in all cases where stem changes were observed between *C. miniata* and *C. pseudocurata* suggests that this base pair allows an intermediate through which A-U pairs are converted to G-C pairs, or vice versa. Such a mechanism would allow the possibility of convergence at these positions over long periods of time and may limit the utility of such gene sequence

Observed changes between *P. californicus* and *C. miniata* at loop positions included a much higher proportion of transversions ( $T_i/T_v = 1.2$ ), than at stem positions ( $T_i/T_v = 7.2$ ). Indels were also common, accounting for 28/66 changes. The fact that no indels were detected at stem positions is not surprising, given their role in secondary structure, however, an inherent bias exists since alignment is based largely on stem structure. In regions with alternative pairing possibilities, indels could occur but would most probably be scored as events in the adjacent loop. In any case, observed changes

support the notion that patterns of substitution are heavily influenced by secondary structural constraints.

### *Phylogenetic Utility of tRNA Gene Sequence*

A simple phylogenetic analysis was undertaken to test the utility of tRNA gene sequence at deeper levels. Both distance and maximum likelihood methods consistently placed the two members of each echinoderm class included together. However, distance estimates beyond intraclass comparisons were large, and support for grouping sea urchins with sea cucumbers, as indicated by adult morphological characters and mitochondrial gene order (Smith et al., 1993), was only poorly supported. Utilizing transversions alone did not improved performance of tree-building methods compared to both transitions and transversions. Analysis of stem positions plus the anticodon loop did not appear to recover the correct tree more often than did the inclusion of all positions. It therefore appears that tRNA gene sequence will be of considerable utility for relationships within echinoderm classes but perhaps not for relationships as ancient as several hundred million years, as suggested by Kumazawa and Nishida (1993). The exclusion of loop positions from analysis is not necessarily warranted, although substitutions accumulated more rapidly than at stem positions and the high degree of compensatory mutations indicates that a weighting factor near 0.5 is appropriate for stem positions.



## **Future Directions**

### ***Population Structure***

This study has demonstrated the presence of a significant mtDNA genetic discontinuity along the central coast of British Columbia. Further sampling of northern Vancouver Island and the mainland coast would allow for a more precise localization of this genetic break. One question not answered by this study is whether these northern and southern groups represent different species or simply long separated variants which have subsequently expanded their ranges following the retreat of the Cordilleran ice sheet and have reestablished contact. Since mtDNA is maternally inherited without intermolecular recombination, it is difficult to detect recent hybridization with this marker. The inclusion of some type of nuclear marker would greatly improve our understanding of this system since such markers could discriminate between the competing hypotheses of hybridization versus speciation between the northern and southern groups. Allozymes, microsatellites, or analysis of anonymous nuclear loci polymorphisms are three possible approaches. The detection of heterozygotes would confirm that the northern and southern groups are capable of interbreeding despite separation during the last glaciation.

The presence of this discontinuity also raises the possibility that other organisms with similar life histories also exhibit this discontinuity. Analysis of population structure of other marine invertebrates with limited dispersal in these regions could be combined with sea cucumber data to examine patterns of dispersal at a fine scale along the British

Columbia coast. Concordance among species with limited dispersal capabilities would then reinforce the concept of two different biogeographic regions along the British Columbia coast. Numerous other organisms appear to share this north/south genetic split including chinook (Wilson et al., 1987) and sockeye (Wood et al., 1995) salmon, harbor seals (T. Burg, pers. commun.), and threespine stickleback (O'Reilly et al., 1993). These results have important implications for conservation biology, since considerably less habitat has been protected on the central and northern coast of British Columbia than in the south. A growing body of evidence indicates that considerable diversity has been maintained in the northern regions despite past glaciation and efforts must continue to preserve and improve our understanding of the origins of biodiversity in this region.

#### *Duplication of tRNA Genes*

The apparent duplication of tRNA genes identified in *C. miniata* raises some interesting questions. One of the most intriguing is why the putative control region has been maintained in both copies despite the fact that most of the tRNA genes have been reduced to a single functional copy. Possible explanations, as outlined in above, are, first, that a second functional replication origin would not necessarily be deleterious and may in fact provide a selective advantage for the mtDNA molecule and second, that a second transcriptional regulator is required as a result of the tRNA gene duplication. One experimental approach to test these ideas would be to develop an in vitro transcription assay for the mitochondrial system, or to use a bacterial expression vector to test the

potential of this region to initiate transcription. Deletion analysis could also be undertaken to determine essential elements for either transcription or replication.

Another point of interest is that the duplication event appears to be shared by several members of the *Cucumaria*. This duplication event may thus provide a highly informative character for phylogenetic analysis. COI gene sequence comparisons indicated that variable sites were saturated for changes beyond the familial level and thus contain little phylogenetic information. LrRNA gene sequence was more informative at higher taxonomic levels but gene order characters may be informative at still higher levels, as has been suggested (Cantatore et al., 1987, Moritz et al., 1987). Furthermore, Hart, Byrne, and Smith (unpublished) utilized tRNA gene sequence for phylogenetic analysis of developmental patterns in sea stars when COI gene sequence appeared to be highly variable and consequently largely uninformative. Therefore, localization and sequencing of tRNA genes could greatly improve our understanding of the radiation of sea cucumber families. Phylogenetic information at such levels may lie not only within stem regions but at loop positions as well. With proper consideration given to weighting and class of rate variation, tRNA genes comprise an informative yet under-utilized portion of the mitochondrial genome.

## References

- Alberts, B., D. Bray, J. Lewis, M. Raff, K. Roberts, and J.D. Watson. 1983. *Molecular Biology of the Cell*. Garland Publishing, New York.
- Arnez, J.G., and T.A. Steitz. 1994. Crystal structure of unmodified tRNA (Gln) complexes with glutamyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure. *Biochemistry*. **33**:7560-7567.
- Asakawa, S., Y. Kumasawa, T. Araki, H. Himeno, K. Miura, and K. Watanabe. 1991. Strand-specific nucleotide composition bias in echinoderm and vertebrate mitochondrial genomes. *J. Mol. Evol.* **32**:511-520.
- Avise, J.C., J. Arnold, R.M. Ball, E. Bermingham, T. Lamb, J.E. Neigel, C.A. Reeb, and N.C. Saunders. 1987. Intraspecific Phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Ann. Rev. Ecol. Syst.* **18**:489-522.
- Avise, J.C. 1994. *Molecular Markers, Natural History and Evolution*. Chapman and Hall, New York.

- Bairoch, A. 1993. PC GENE, release 6.80. University of Geneva, Intelligenetics Inc., Switzerland.
- Berger, E. 1973. Gene-enzyme variation in three sympatric species of *Littorina*. *Biol. Bull.* **145**:83-90.
- Bibb, M.J., R.A. Van Etten, C.T. Wright, M.W. Walberg, and D.A. Clayton. 1981. Sequence and gene organization of mouse mitochondrial DNA. *Cell* **26**:167-186.
- Bickham, J.W., C.C. Wood, and J.C. Patton. 1995. Biogeographic implication of cytochrome b sequences and allozymes in sockeye (*Oncorhynchus nerka*). *J. Hered.* **86**:140-144.
- Birnboim, H.C. 1983. A rapid alkaline lysis extraction method for the isolation of plasmid DNA. *Meth. Enzymol.* **100**:243-255.
- Bowers, N., J.R. Stauffer, and T.D. Kocher. 1994. Intra- and interspecific mitochondrial DNA sequence variation within two species of rock-dwelling cichlids (Teleostei: Cichlidae) from lake Malawi, Africa. *Mol. Phyl. Evol.* **3**:75-82.
- Brandt, J.F. 1835. Prodrum descriptionis animalium ab H. Mertensio in orbis terrarum circumnavigatione observatorum. Sumptibus Academiae, St. Petersburg.

- Brown, G.G., G. Gadaleta, G. Pepe, C. Saccone, and E. Sbisà. 1986. Structural conservation and variation in the D-loop-containing region of vertebrate mitochondrial DNA. *J. Mol. Biol.* **192**:503-511.
- Brown, J.R., A.T. Beckenbach, and M.J. Smith. 1993. Intraspecific DNA sequence variation of the mitochondrial control region of white sturgeon (*Acipenser transmontanus*). *Mol. Biol. Evol.* **10**:326-341.
- Brown, W.M., M. George, Jr., and A.C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* **76**:1967-1971.
- Brown, W.M. 1985. Evolution of the animal mitochondrial DNA genome. *In* Molecular Evolutionary Genetics, ed. R. MacIntyre. Plenum Press, New York.
- Burton, R.S. 1983. Protein polymorphisms and genetic differentiation of marine invertebrate populations. *Mar. Biol. Letters* **4**:193-206.
- Campbell, C.A. 1978. Genetic divergence between population of *Thais lamellosa* (Gmelin). *In: Marine Organisms: Genetics, Ecology, and Evolution*, pp. 157-170. (Eds.) B. Battaglia, and J.A. Beardmore. Plenum Press, New York.

- Cabot, E.L., and A.T. Beckenbach. 1989. Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. *Comput. Appl. Biosci.* **5**:233-234.
- Cann, R.L., W.M. Brown, and A.C. Wilson. 1984. Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. *Genetics.* **106**:479-499.
- Cantatore, P., M.N. Gadaleta, M. Roberti, C. Saccone, and A.C. Wilson. 1987. Duplication and remoulding of tRNA genes during the evolutionary rearrangement of mitochondrial genomes. *Nature* **329**:853-855.
- Cantatore, P., M. Roberti, G. Renaldi, M. Gadaleta, and C. Saccone. 1989. The complete nucleotide sequence, gene organization, and genetic code of the mitochondrial genome of *Paracentrotus lividus*. *J. Biol. Chem.* **264**:10965-10975.
- Clague, J. 1989. The Cordilleran ice sheet. *In* Quaternary Geology of Canada and Greenland. ed. R. J. Fulton, Geological survey of Canada, Ottawa.
- Clark, H.L. 1901. No.15: Echinoderms from Puget Sound: Observations made on the echinoderms collected by the parties from Columbia University, in Puget Sound in 1896 and 1897. *Proc. Boston Soc.* **29**:323-337.

- Clark, H.L. 1924. Some holothurians from British Columbia. *Can. Field Nat.* **38**:54-57.
- Clary, D.O., and D.R. Wolstenholme. 1985. The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization and genetic code. *J. Mol. Evol.* **22**:252-271.
- Clayton, D.A. 1982. Replication of the animal mitochondrial DNA. *Cell* **28**:693-705.
- Cockerham, C.C., and B.S.Weir. 1993. Estimation of gene flow from F-statistics. *Evolution* **47**:855-863.
- Cowles, R.P. 1907. *Cucumaria curata* sp.nov. Johns Hopkins Univ. Circ. new series **3**:8-9.
- Crisp, D.J. 1978. Genetic consequences of different reproductive strategies in marine invertebrates. Pp. 257-273 *in*: Marine Organisms: Genetics, Ecology, and Evolution, B. Battaglia and H.A. Beardmore (eds), Plenum Press, New York.
- Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favored Races in the Struggle for Life.* John Murray, London.



- Deichmann, E. 1938. New holothurians from the western coast of North America and some remarks on the Genus *Caudina*. *Proc. New Eng. Zool. Club* **16**:103-115.
- Derrick, W.B., and J. Horowitz. 1993. Probing structural differences between native and in vitro transcribed *Escherichia coli* valine transfer RNA: evidence for stable base modification-dependent conformers. *Nucl. Acids Res.* **21**:4948-4953.
- DeSalle, R., T. Freedman, E.M. Prager, and A.C. Wilson. 1987. Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. *J. Mol. Evol.* **26**:157-164.
- Dixon, M.T., and D.M. Hillis. 1993. Ribosomal RNA secondary structure: compensatory mutations and implications for phylogenetic analysis. *Mol. Biol. Evol.* **10**:256-267.
- Endler, J.A. 1973. Gene flow and population differentiation. *Science* **179**:243-250.
- Excoffier, L., P.E. Smouse, and J.M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**:479-491.

- Fang, Q., W.C. Black IV, H.D. Blocker, and R.F. Whitcomb. 1993. A phylogeny of new world *Deltocephalus*-like Leafhopper genera based on mitochondrial 16S ribosomal DNA sequences. *Mol. Phyl. Evol.* 2:119-131.
- Felsenstein, J. 1981. Evolutionary trees from DNA sequence: A maximum likelihood approach. *J. Molec. Evol.* 17:368-376.
- Felsenstein, J. 1982. Numerical methods for inferring evolutionary trees. *Quart. Rev. Biol.* 57:379-404.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783-791.
- Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) Version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle.
- Gerrodette, T. 1981. Dispersal of the solitary coral *Balanopyhllia elegans* by demersal planular larvae. *Ecology* 62:611-619.
- Gilliland, P.M. 1992. The skeletal morphology, systematics and evolutionary history of Holothurians. *Spec. Papers in Palaeontol.* 47:1-147.

Hartl, D.L., and A.G. Clark. 1989. Principles of Population Genetics, 2nd Ed. Sinauer, Sunderland, MA.

Hattori, M., and Y. Sakaki. 1986. Dideoxy sequencing method using denatured plasmid templates. *Anal. Biochem.* **152**:232-238.

Hedgecock, D. 1986. Is gene flow from pelagic larval dispersal important in the adaptation and evolution of marine invertebrates? *Bull. Mar. Sci.* **39**:550-564.

Higgins, E.G., A.J. Bleasby, and R. Fuchs. 1992. CLUSTALV: improved software for multiple sequence alignment. *Comput. Appl. Biosc.* **8**:189-191.

Highsmith, R.C. 1985. Floating and algal rafting as potential dispersal mechanisms in brooding invertebrates. *Mar. Ecol.* **25**:169-179.

Hillis, D.M., and C. Moritz. 1990. Molecular Systematics. Sinauer, Sunderland, MA.

Himeno, H., H. Masaki, T. Kawai, T. Ohta, I. Kumagai, I. Miura and K. Watanabe. 1987. Unusual genetic codes and novel gene structure for tRNA<sub>AGY</sub><sup>Ser</sup> in starfish mtDNA. *Gene* **56**:219-230.

- Hudson, R.R., Slatkin, M., and W.P. Maddison. 1992. Estimation of levels of gene flow from DNA sequence data. *Genetics* **132**:583-589.
- Huelsenbeck, J.P., and D.M. Hillis. 1993. Success of phylogenetic methods in the four-taxon case. *Syst. Biol.* **42**:247-264.
- Huson, D., and R. Wetzell. 1994. Splitstree. FSP Mathematisierung, Universitat Bielefeld. Bielefeld, Germany.
- Jablonski, D. 1986. Larval ecology and macroevolution in marine invertebrates. *Bull. Mar. Sci.* **39**:565-587.
- Jacobs, H.T., D.J. Elliott, V.B. Math, and A. Farquharson. 1988. Nucleotide sequence and gene organization of sea urchin mitochondrial DNA. *J. Mol. Biol.* **202**:185-217.
- Jacobs, H., S. Asakawa, T. Araki, K. Mikura, M.J. Smith, and K. Watanabe. 1989. Conserved tRNA gene cluster in starfish mitochondrial DNA. *Curr. Genet.* **15**:193-206.

- Janson, K. 1987. Allozyme and shell variation in two marine snails (*Littorina*, Prosobranchia) with different dispersal abilities. *Biol. J. Linnean Soc.* 30:245-256.
- Jermiin, L.S., D. Graur, R.M. Lowe, and R.H. Crozier. 1994a. Analysis of directional mutation pressure and nucleotide content in mitochondrial cytochrome b genes. *J. Mol. Evol.* 39:160-173.
- Jermiin, L.S., D. Graur, and R.H. Crozier. 1994b. Strand-specific directional mutation pressure in metazoan mitochondrial DNA. *Mol. Biol. Evol.* 12:558-563.
- Jukes, T.H., and C.R. Cantor. 1969. Evolution of protein molecules. Pp. 21-132. *in* Mamalian Protein Metabolism, H.N. Munro (ed.), Academic Press, New York.
- Karl, S.A., E.W. Bowen, and J.C. Avise. 1992. Global population genetic structure and male-mediated gene flow in the green turtle (*Chelonia mydas*): RFLP analyses of anonymous nuclear loci. *Genetics* 131:163-173.
- Kishino, H. and M. Hasegawa. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* 29:170-179.

- Keohavong, P., and W.G. Thilly. 1989. Fidelity of DNA polymerase in DNA amplification. *Proc. Natl. Acad. Sci. USA.* 86:9253-9257.
- Kirkendale, L. and P. Lambert. 1995. *Cucumaria pallida*, a new species of sea cucumber from the northeast Pacific ocean (Echinodermata: Holothuroidea). *Can. J. Zoo.* 73:542-551.
- Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16:111-120.
- Kimura, M. 1983. *The Neutral Theory of Molecular Evolution.* Cambridge Univ. Press, Cambridge, England.
- Kimura, M., and G.H. Weis. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* 49:561-576.
- Kishino, H., and M. Hasegawa. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* 29:170-179.

- Kocher, T.D., W.K. Thomas, A. Meyer, S.V. Edwards, S. Paabo, F.X. Villablanca, and A.C. Wilson. 1989. Dynamics of mitochondrial DNA evolution in animals: Amplification primers and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA* **86**:6196-6200.
- Kumar, S., K. Tamura, and M. Nei. 1993. MEGA: Molecular evolutionary genetics analysis, version 1.0. The Pennsylvania State University, University Park, PA 16802.
- Kumazawa, Y., and M. Nishida. 1993. Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. *J. Mol. Evol.* **37**:380-398.
- Lambert, P. 1985. Geographic variation of calcareous ossicles and the identification of three species of sea cucumbers (Echinodermata: Holothuroidea) from the eastern Pacific Ocean. *In* Proceedings of the Fifth International Echinoderm Conference, Galway. B. F. Keegan and B. D. S. O'Connor, *Eds.* A.A. Balkema, Rotterdam. pp. 437-443.
- Lambert, P. 1990. A new combination and synonymy for two subspecies of *Cucumaria fisheri* Wells (Echinodermata: Holothuroidea). *Proc. Biol. Soc. Wash.* **103**:913-921.

- Lansman, R.A., R.O. Shade, J.F. Shapira, and J.C. Avise. 1981. The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. *J. Mol. Evol.* **17**:214-226.
- Li, W.-H. and D. Graur. 1991. *Fundamentals of Molecular Evolution*. Sinauer, Sunderland, MA.
- Liu, L.L., D.W. Foltz, and W.B. Stickle. 1991. Genetic population structure of the southern oyster drill *Stramonita (=Thais) haemostoma*. *Marine Biology* **111**:71-79.
- Lockhart, P.J., M.A. Steel, M.D. Hendy, and D. Penny. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. *Mol. Biol. Evol.* **11**(4):605-612.
- Marchuk, D., M. Drumm, A. Saulino and F.S. Collins. 1991. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucl. Acids Res.* **19**:1154.
- Martin, A.P., and S.R. Palumbi. 1993. Body size, metabolic rate, generation time and the molecular clock. *Proc. Natl. Acad. Sci.* **90**:4087-4091.



Mayhook, A.G., A-M. Rinaldi, and H.T. Jacobs. 1992. Replication origins and pause sites in sea urchin mitochondrial DNA. *Proc. R. Soc. Lond. B.* **248**:85-94.

McMillan, W.O., R.A. Raff, and S.R. Palumbi. 1992. Population genetic consequences of developmental evolution in sea urchins (genus *Heliocidaris*). *Evolution* **46**:1299-1312.

Minister of Supply and Services Canada. 1995. Canadian Tide and Current Tables 1996 (vols. 5,6). Fisheries and oceans communication directorate, Ottawa.

Moritz, C., and W.M. Brown. 1987. Tandem duplications in animal mitochondrial DNAs: Variation in incidence and gene content among lizards. *Proc. Natl. Acad. Sci. USA.* **84**:7183-7187.

Moritz,C., T.E. Dowling, and W.M. Brown. 1987. Evolution of animal mitochondrial DNA; relevance for population biology and systematics. *Ann. Rev. Ecol. Syst.* **18**:269-292.

Mullis, K.F. Faloona, S. Scharf, R. Saiki, G. Horn, and H. Erlich. 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* **51**:263-273.

- O'Reilly, P., T.E. Reimchen, R. Beech, and C. Strobeck. Mitochondrial DNA in *Gasterosteus* and pleistocene glacial refugium on the Queen Charlotte Islands, British Columbia. *Evolution* **47**:678-684.
- Ortí, G., P. Petry, J. I. R. Porto, M. Jégu, and A. Meyer. 1996. Patterns of nucleotide change in mitochondrial ribosomal RNA genes and the phylogeny of Piranhas. *J. Mol. Evol.* **42**:169-182.
- Osawa, S., T.H. Jukes, K. Watanabe, and A. Muto. 1992. Recent evidence for evolution of the genetic code. *Microbiological Reviews* **56**:229-264.
- Ovenden, J.R., B.J. Brasher, and R.W.G. White. 1992. Mitochondrial DNA analyses of the red rock lobster *Jasus edwardsii* supports an apparent absence of population subdivision throughout Australasia. *Mar. Biol.* **12**:319-326.
- Pääbo, S., W.K. Thomas, K.M. Whitfield, Y. Kumazawa, and A.C. Wilson. 1991. Rearrangements of Mitochondrial Transfer RNA Genes in Marsupials. *J. Mol. Evol.* **33**:426-430.
- Palumbi, S.R., and A.C. Wilson. 1990. Mitochondrial DNA diversity in the sea urchins *Strongylocentrotus Purpuratus* and *S. Droebachiensis*. *Evolution* **44**(2):403-415.

- Palumbi, S.R., A. Martin, S. Romano, W.O. McMillan, L. Stice, and G. Grabowski. 1991. The simple fool's guide to PCR, version 2. University of Hawaii Zoology Department, Honolulu, Hawaii.
- Palumbi, S.R. 1992. Marine speciation on a small planet. *Trends Ecol. Evol.* 7:114-118.
- Panning, A., von. 1962. Bemerkungen über die holothurien-familie Cucumariidae (Ordnung Denrochirota) 3. Teil die gattung *Pseudocnus* Panning 1949. *Mitt Hamburg Zool. Mus. Inst.* 60:57-60.
- Pawson, D.L. 1966. Phylogeny and evolution of holothuroids. *In: Treatise on Invertebrate Paleontology. Vol. Part U Echinodermata 3. Edited by R.C. Moore.* University of Kansas Press, Lawrence. pp. 641-646.
- Ricketts, E. F., and J. Calvin. 1968. *Between Pacific Tides.* 4th ed. revised by J.W. Hedgpeth. Stanford University Press, Stanford, CA.
- Roe, B.A., D.-P. Ma, R.K. Wilson, and J. F.-H. Wong. 1985. The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome. *J. Biol. Chem.* 260:9759-9774.

- Rohlf, F. J. 1990. NTSYS. Numerical taxonomy and multivariate analysis system, version 1.60. Exeter Publishing, Setauket, New York.
- Rutherford, J.C. 1973. Reproduction, growth and mortality of the Holothurian *Cucumaria Pseudocurata*. *Mar. Biol.* **22**:167-176.
- Rutherford, J.C. 1977a. Geographical variation in morphological and electrophoretic characters in the holothurian *C. curata*. *Mar. Biol.* **43**:165-174.
- Rutherford, J.C. 1977b. Variation in egg numbers between populations and between years in the Holothurian *Cucumaria Curata*. *Mar. Biol.* **43**:175-180.
- Saiki, R.K., D.H. Gelfand, S. Stoeffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis, and H.A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**:487-491.
- Saitou, N., and M. Nei. 1987. The neighbor joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406-425.
- Sanger, F.S., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74**:5463-5467.

- Seutin, G., B. Franz Lang, D.P. Mindell, and R. Morais. 1994. Evolution of the WANCY region in amniote mitochondrial DNA. *Mol. Biol. Evol.* **11**(3):329-340.
- Selenka, E. 1867. Beiträge zur anatomie und systematik der holothurien. *Z. Wiss. Zool.* **17**:291-374.
- Semper, C. 1868. Reisen im archipel der Phillippinen zweiter theil wissenschaftliche resultate erster band holothurien. Weisbaden.
- Sewell, M.A., and D.R. Levitan. 1992. Fertilization success during natural spawning of the dendrochiro sea cucumber *Cucumaria miniata*. *Bul. Mar. Sci.* **51**:161-166.
- Slatkin, M. 1985. 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.
- Smith, M.J., D.K. Banfield, K. Doteval, S. Gorski, and D.J. Kowbel. 1989. Gene arrangement in sea star mitochondrial DNA demonstrates a major inversion event during echinoderm evolution. *Gene* **76**:181-185.

- Smith, M.J., A. Arndt, S. Gorski, and E. Fajber. 1993. The phylogeny of echinoderm classes based on mitochondrial gene rearrangements. *J. Mol. Evol.* **36**:545-554.
- Stanley, H.F., S. Casey, J. M. Carnaham, S. Goodman, J. Harwood, and R.K. Wayne. 1996. World wide patterns of mitochondrial DNA differentiation in the harbour seal (*Phoca vitulina*). *Mol. Biol. Evol.* **13**:368-382.
- Steel, M. 1993. Recovering a tree from the leaf colourations it generates under a Markov model. Research rep. 103, Mathematics Department, University of Christchurch, N.Z.
- Steinberg, S., and R. Cedergren. 1994. Structural compensation in atypical mitochondrial tRNAs. *Nature Struct. Biol.* **1**:507-511.
- Steinberg, S., and R. Cedergren. 1995. A correlation between N2-dimethylguanosine presence and alternate tRNA conformers. *RNA*. **1**:886-891.
- Stimpson, W. 1857. Crustacea and Echinodermata of the Pacific shores of North America. Boston *J. Nat. Hist.* **6**:524-531.

- Stimpson, W. 1864. Descriptions of new species of marine invertebrata from Puget Sound, collected by the Naturalists of the North-west Boundary Commission, A.H. Campbell, Esq., Commissioner. *Proc. Acad. Nat. Sci. Philadelphia* **16**:153-161.
- Strathmann, M.F. 1987. Reproduction and development of marine invertebrates of the northern Pacific coast. University of Washington Press, Seattle, Washington.
- Sueoka, N. 1962. On the genetic basis of variation and heterogeneity of DNA base composition. *Proc. Nat. Acad. Sci. USA*. 2653-2657.
- Swofford, D. L. 1993. PAUP: Phylogenetic analysis using parsimony, version 3.1.1. Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois.
- Takahata, N., and S.R. Palurabi. 1985. Extranuclear differentiation and gene flow in the finite island model. *Genetics* **109**:441-457.
- Tamura, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**:512-526.

- Theel, H. 1886. Report on the Holothurioidea dredged by the HMS Challenger during the years 1873-1876 Part II. Report of the Scientific Results of the Voyage of H.M.S. Challenger 1873-76 **14**:1-290.
- Tindall, K.R., and T. Kunkel. 1988. Fidelity of DNA synthesis by *Thermus aquaticus* DNA polymerase. *Biochemistry* **27**:6008-6013.
- Thomson, R.E., 1981. Oceanography of the British Columbia coast. Department of Fisheries and Oceans, Government of Canada. Ottawa, ON.
- Tomkinson, A.E., and S. Linn. 1986. Purification and properties of a strand-specific endonucleases from mouse cell mitochondria. *Nucl. Acids Res.* **14**:9579-9593.
- Tracey, M.L., K. Nelson, D. Hedgecock, R.A. Shleser, and M.L. Pressick. 1975. Biochemical genetics of lobsters: genetic variation and the structure of America lobster (*Homarus americanus*). *J. Fish. Res. Bd. Can.* **32**:2091-2101.
- Van Syoc, R.J. 1994. Genetic divergence between subpopulations of eastern Pacific goose barnacle *Pollicipes elegans*: mitochondrial cytochrome *c* subunit 1 nucleotide sequences. *Mol. Mar. Biol. Biotech.* **3**:338-346.



- Vawter, L., and W.M. Brown. 1986. Nuclear and mitochondrial DNA comparisons reveal extreme rate variation in the molecular clock. *Science* **234**:194-
- Wells, H. 1924. IV. New species of *Cucumaria* from Monterey Bay, California. *Ann. Mag. Nat. Hist. Ser.9* **14**:113-121.
- Wheeler, W.C., and R.L.Honeycutt. 1988. Paired sequence difference in ribosomal RNAs: evolutionary and phylogenetic implications. *Mol. Biol. Evol.* **5**:90-96.
- Wilson, G. M., W.K. Thomas, and A.T. Beckenbach. 1987. Mitochondrial DNA Analysis of Pacific northwest populations of *Oncorhynchus tshawytscha*. *Can. J. Fish Aquat. Sci.* **44**:1301-1305.
- Wolstenholme, D.R. 1992. Animal mitochondrial DNA: structure and evolution. *In* Mitochondrial genomes. D.R. Wolstenholme and K.W. Jeon, (eds). *Int. Rev. Cytol.* **141**:173-216.

Wood, C.C., B.E. Riddell, D.T. Rutherford, and R.E. Withler. 1995. Biochemical genetic survey of sockeye salmon (*Oncorhynchus nerka*) in Canada. *Can. J. Fish. Aquat. Sci.* **51** (suppl):114-131.

Wright, S. 1943. Isolation by distance. *Genetics* **28**:114-138.

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

Zuckerkindl, E., and L. Pauling. Evolutionary divergence and convergence in proteins. 1965. *In* *Evolving Genes and Proteins*. V. Bryson and H.J. Vogel, (eds.), Academic Press, New York.

Appendix A: Distance Matrix for *Cnemidaria pseudocurata* Haplotypes (Jukes-Cantor)

## OTU Labels:

1, cvak; 2, cvun; 3, BIa, 4, BIb, 5, BIc, 6, BIe; 7, BIf, 8, BIh, 9, BIi, 10, BIc;  
 11, BIg; 12, THc; 13, THe; 14, THa, 15, THf, 16, THg; 17, THh, 18, THd; 19, THb; 20, Mic;  
 21, MIb; 22, MIa; 23, MIc; 24, MIe; 25, MIf, 26, SKa; 27, SKb; 28, SKd; 29, SKe; 30, Bfa;  
 31, BFe; 32, BFc; 33, BFg; 34, BFb; 35, BFh; 36, BFj; 37, BFc; 38, BFd; 39, BFj; 40, BFk;  
 41, BFi; 42, CMa; 43, CMB; 44, CMc; 45, CMd; 46, CMe; 47, YHb; 48, YHe; 49, YHa;  
 50, Yha;  
 51, YHd; 52, SBa; 53, SBb; 54, TRa; 55, PP1; 56, PP2; 57, PP3; 58, PP7; 59, PP14; 60, PP15;

Distances in the upper-right matrix

Standard Errors in lower-left matrix

\* indicates an invalid distance value

OTUs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		0.0074	0.0299	0.0223	0.0376	0.0402	0.0325	0.0299	0.0274	0.0376	0.0402	0.0199	0.0174	0.0173	0.0174
2	0.0043		0.0325	0.0249	0.0402	0.0428	0.0350	0.0325	0.0299	0.0402	0.0428	0.0224	0.0199	0.0198	0.0199
3	0.0087	0.0091		0.0123	0.0074	0.0098	0.0024	0.0098	0.0074	0.0074	0.0098	0.0275	0.0275	0.0274	0.0275
4	0.0075	0.0079	0.0055		0.0198	0.0223	0.0148	0.0074	0.0098	0.0198	0.0223	0.0199	0.0174	0.0198	0.0174
5	0.0098	0.0101	0.0043	0.0070		0.0173	0.0098	0.0173	0.0148	0.0049	0.0074	0.0352	0.0352	0.0350	0.0351
6	0.0101	0.0104	0.0049	0.0075	0.0066		0.0123	0.0198	0.0173	0.0173	0.0198	0.0378	0.0378	0.0376	0.0377
7	0.0091	0.0094	0.0024	0.0061	0.0049	0.0055		0.0123	0.0098	0.0098	0.0123	0.0301	0.0301	0.0299	0.0300
8	0.0087	0.0091	0.0049	0.0043	0.0066	0.0070	0.0055		0.0024	0.0123	0.0148	0.0224	0.0199	0.0223	0.0199
9	0.0083	0.0087	0.0043	0.0049	0.0061	0.0066	0.0049	0.0024		0.0098	0.0123	0.0199	0.0199	0.0198	0.0199
10	0.0098	0.0101	0.0043	0.0070	0.0035	0.0066	0.0049	0.0055	0.0049		0.0024	0.0301	0.0301	0.0299	0.0300
11	0.0101	0.0104	0.0049	0.0075	0.0043	0.0070	0.0055	0.0061	0.0055	0.0024		0.0326	0.0326	0.0325	0.0326
12	0.0071	0.0075	0.0083	0.0071	0.0095	0.0098	0.0087	0.0075	0.0071	0.0087	0.0091		0.0025	0.0025	0.0025
13	0.0066	0.0071	0.0083	0.0066	0.0095	0.0098	0.0087	0.0071	0.0071	0.0087	0.0091	0.0025		0.0000	0.0000
14	0.0066	0.0070	0.0083	0.0070	0.0094	0.0098	0.0087	0.0075	0.0070	0.0087	0.0091	0.0025	0.0000		0.0000
15	0.0066	0.0070	0.0083	0.0066	0.0094	0.0098	0.0087	0.0070	0.0070	0.0087	0.0091	0.0025	0.0000	0.0000	

16	0.0061	0.0066	0.0079	0.0066	0.0091	0.0094	0.0083	0.0070	0.0066	0.0083	0.0087	0.0035	0.0025	0.0024	0.0025
17	0.0070	0.0075	0.0079	0.0066	0.0091	0.0094	0.0083	0.0070	0.0066	0.0083	0.0087	0.0035	0.0025	0.0024	0.0025
18	0.0070	0.0075	0.0087	0.0075	0.0098	0.0101	0.0091	0.0079	0.0075	0.0091	0.0094	0.0035	0.0025	0.0024	0.0025
19	0.0070	0.0075	0.0087	0.0066	0.0098	0.0101	0.0091	0.0070	0.0075	0.0091	0.0094	0.0035	0.0000	0.0024	0.0000
20	0.0079	0.0079	0.0091	0.0079	0.0101	0.0105	0.0094	0.0083	0.0079	0.0094	0.0098	0.0079	0.0075	0.0075	0.0075
21	0.0061	0.0066	0.0087	0.0075	0.0098	0.0101	0.0091	0.0079	0.0075	0.0091	0.0094	0.0061	0.0056	0.0055	0.0055
22	0.0071	0.0071	0.0083	0.0071	0.0095	0.0098	0.0087	0.0075	0.0071	0.0087	0.0091	0.0071	0.0066	0.0066	0.0066
23	0.0066	0.0066	0.0080	0.0066	0.0091	0.0095	0.0084	0.0071	0.0066	0.0084	0.0088	0.0061	0.0056	0.0056	0.0056
24	0.0071	0.0071	0.0084	0.0071	0.0095	0.0098	0.0088	0.0075	0.0071	0.0088	0.0091	0.0066	0.0061	0.0061	0.0061
25	0.0066	0.0070	0.0091	0.0079	0.0101	0.0104	0.0094	0.0083	0.0079	0.0094	0.0098	0.0066	0.0061	0.0061	0.0061
26	0.0120	0.0123	0.0129	0.0120	0.0126	0.0140	0.0132	0.0123	0.0120	0.0120	0.0123	0.0100	0.0100	0.0100	0.0100
27	0.0116	0.0120	0.0132	0.0123	0.0129	0.0143	0.0135	0.0126	0.0123	0.0123	0.0126	0.0104	0.0104	0.0103	0.0103
28	0.0129	0.0132	0.0137	0.0123	0.0135	0.0149	0.0140	0.0126	0.0129	0.0129	0.0132	0.0111	0.0107	0.0110	0.0107
29	0.0123	0.0126	0.0132	0.0116	0.0129	0.0143	0.0135	0.0120	0.0123	0.0123	0.0126	0.0104	0.0100	0.0103	0.0100
30	0.0118	0.0121	0.0127	0.0118	0.0124	0.0138	0.0130	0.0127	0.0124	0.0124	0.0127	0.0109	0.0115	0.0115	0.0115
31	0.0121	0.0124	0.0130	0.0121	0.0127	0.0141	0.0132	0.0130	0.0127	0.0127	0.0130	0.0112	0.0118	0.0118	0.0118
32	0.0120	0.0123	0.0129	0.0120	0.0126	0.0140	0.0132	0.0129	0.0126	0.0126	0.0129	0.0107	0.0107	0.0107	0.0107
33	0.0120	0.0123	0.0123	0.0120	0.0120	0.0135	0.0126	0.0123	0.0120	0.0120	0.0123	0.0107	0.0107	0.0107	0.0107
34	0.0116	0.0120	0.0126	0.0116	0.0123	0.0137	0.0129	0.0126	0.0123	0.0123	0.0126	0.0104	0.0104	0.0103	0.0103
35	0.0120	0.0123	0.0129	0.0113	0.0126	0.0140	0.0132	0.0123	0.0126	0.0126	0.0129	0.0107	0.0104	0.0107	0.0103
36	0.0118	0.0121	0.0127	0.0118	0.0124	0.0138	0.0130	0.0127	0.0124	0.0124	0.0127	0.0109	0.0115	0.0115	0.0115
37	0.0123	0.0126	0.0132	0.0123	0.0129	0.0143	0.0135	0.0132	0.0129	0.0129	0.0132	0.0111	0.0111	0.0110	0.0110
38	0.0129	0.0132	0.0137	0.0123	0.0135	0.0149	0.0140	0.0132	0.0135	0.0135	0.0137	0.0117	0.0114	0.0116	0.0114
39	0.0124	0.0127	0.0132	0.0118	0.0130	0.0144	0.0135	0.0127	0.0130	0.0130	0.0132	0.0115	0.0118	0.0121	0.0118
40	0.0121	0.0124	0.0130	0.0121	0.0127	0.0141	0.0132	0.0130	0.0127	0.0127	0.0130	0.0112	0.0118	0.0118	0.0118
41	0.0121	0.0124	0.0130	0.0121	0.0127	0.0141	0.0132	0.0130	0.0127	0.0127	0.0130	0.0112	0.0118	0.0118	0.0118
42	0.0124	0.0127	0.0132	0.0124	0.0130	0.0144	0.0135	0.0127	0.0124	0.0124	0.0127	0.0109	0.0115	0.0115	0.0115
43	0.0130	0.0132	0.0138	0.0130	0.0135	0.0149	0.0141	0.0132	0.0130	0.0130	0.0132	0.0115	0.0121	0.0121	0.0121



10	0.0274	0.0274	0.0325	0.0325	0.0351	0.0325	0.0301	0.0276	0.0302	0.0350	0.0543	0.0570	0.0624	0.0570	0.0588
11	0.0299	0.0299	0.0350	0.0377	0.0377	0.0350	0.0326	0.0302	0.0327	0.0376	0.0570	0.0597	0.0651	0.0597	0.0614
12	0.0049	0.0049	0.0049	0.0250	0.0250	0.0149	0.0200	0.0150	0.0175	0.0174	0.0386	0.0412	0.0465	0.0412	0.0458
13	0.0025	0.0025	0.0000	0.0225	0.0225	0.0124	0.0175	0.0125	0.0150	0.0149	0.0386	0.0412	0.0439	0.0386	0.0511
14	0.0024	0.0024	0.0024	0.0224	0.0224	0.0123	0.0174	0.0124	0.0149	0.0148	0.0384	0.0410	0.0463	0.0410	0.0508
15	0.0025	0.0025	0.0000	0.0224	0.0224	0.0124	0.0174	0.0124	0.0150	0.0149	0.0385	0.0411	0.0438	0.0385	0.0510
16	0.0049	0.0049	0.0049	0.0199	0.0199	0.0098	0.0149	0.0099	0.0124	0.0123	0.0384	0.0410	0.0463	0.0410	0.0482
17	0.0035	0.0049	0.0049	0.0199	0.0199	0.0148	0.0149	0.0099	0.0124	0.0173	0.0410	0.0436	0.0489	0.0436	0.0535
18	0.0035	0.0035	0.0049	0.0249	0.0249	0.0148	0.0199	0.0149	0.0174	0.0173	0.0410	0.0436	0.0436	0.0436	0.0535
19	0.0035	0.0035	0.0035	0.0249	0.0249	0.0148	0.0199	0.0149	0.0174	0.0173	0.0410	0.0436	0.0436	0.0384	0.0535
20	0.0070	0.0070	0.0079	0.0079	0.0079	0.0149	0.0049	0.0074	0.0049	0.0174	0.0571	0.0544	0.0653	0.0598	0.0616
21	0.0049	0.0061	0.0061	0.0061	0.0061	0.0149	0.0149	0.0124	0.0149	0.0024	0.0489	0.0463	0.0570	0.0516	0.0535
22	0.0061	0.0061	0.0071	0.0035	0.0035	0.0061	0.0000	0.0025	0.0000	0.0174	0.0519	0.0492	0.0600	0.0546	0.0564
23	0.0050	0.0050	0.0061	0.0061	0.0043	0.0056	0.0025	0.0025	0.0025	0.0149	0.0466	0.0440	0.0547	0.0493	0.0539
24	0.0056	0.0056	0.0066	0.0066	0.0035	0.0061	0.0000	0.0025	0.0025	0.0174	0.0493	0.0466	0.0574	0.0520	0.0565
25	0.0055	0.0066	0.0066	0.0066	0.0066	0.0024	0.0066	0.0061	0.0066	0.0116	0.0516	0.0489	0.0597	0.0543	0.0561
26	0.0100	0.0103	0.0103	0.0103	0.0123	0.0113	0.0117	0.0111	0.0114	0.0116	0.0025	0.0025	0.0075	0.0025	0.0100
27	0.0103	0.0107	0.0107	0.0107	0.0120	0.0110	0.0114	0.0107	0.0111	0.0113	0.0025	0.0100	0.0050	0.0126	0.0126
28	0.0110	0.0113	0.0107	0.0107	0.0132	0.0123	0.0126	0.0121	0.0124	0.0126	0.0043	0.0050	0.0050	0.0177	0.0177
29	0.0103	0.0107	0.0107	0.0100	0.0126	0.0116	0.0120	0.0114	0.0117	0.0120	0.0025	0.0035	0.0035	0.0126	0.0126
30	0.0112	0.0118	0.0118	0.0118	0.0127	0.0118	0.0121	0.0119	0.0122	0.0121	0.0050	0.0056	0.0067	0.0056	0.0056
31	0.0115	0.0121	0.0121	0.0130	0.0121	0.0121	0.0124	0.0122	0.0125	0.0124	0.0056	0.0062	0.0072	0.0062	0.0025
32	0.0107	0.0110	0.0110	0.0129	0.0129	0.0120	0.0123	0.0117	0.0121	0.0123	0.0035	0.0043	0.0056	0.0043	0.0050
33	0.0107	0.0110	0.0110	0.0129	0.0129	0.0120	0.0123	0.0117	0.0121	0.0123	0.0035	0.0043	0.0056	0.0043	0.0050
34	0.0103	0.0107	0.0107	0.0107	0.0126	0.0116	0.0120	0.0114	0.0117	0.0120	0.0025	0.0035	0.0050	0.0035	0.0043
35	0.0107	0.0110	0.0110	0.0129	0.0129	0.0120	0.0123	0.0117	0.0121	0.0123	0.0035	0.0043	0.0056	0.0025	0.0050
36	0.0112	0.0118	0.0118	0.0127	0.0127	0.0118	0.0121	0.0119	0.0122	0.0121	0.0050	0.0056	0.0067	0.0056	0.0035
37	0.0110	0.0113	0.0113	0.0132	0.0132	0.0123	0.0126	0.0121	0.0124	0.0126	0.0043	0.0050	0.0062	0.0050	0.0043

38	0.0116	0.0120	0.0120	0.0113	0.0138	0.0129	0.0132	0.0127	0.0130	0.0132	0.0056	0.0062	0.0050	0.0050	0.0056
39	0.0118	0.0124	0.0124	0.0118	0.0133	0.0124	0.0127	0.0125	0.0128	0.0127	0.0062	0.0067	0.0056	0.0056	0.0035
40	0.0115	0.0121	0.0121	0.0121	0.0130	0.0121	0.0124	0.0122	0.0125	0.0124	0.0056	0.0062	0.0072	0.0062	0.0025
41	0.0115	0.0121	0.0121	0.0121	0.0130	0.0121	0.0124	0.0122	0.0125	0.0124	0.0056	0.0062	0.0072	0.0062	0.0025
42	0.0112	0.0118	0.0118	0.0118	0.0127	0.0118	0.0121	0.0119	0.0122	0.0121	0.0050	0.0056	0.0067	0.0056	0.0035
43	0.0118	0.0124	0.0124	0.0124	0.0133	0.0124	0.0127	0.0125	0.0128	0.0127	0.0067	0.0072	0.0081	0.0072	0.0071
44	0.0118	0.0124	0.0124	0.0124	0.0133	0.0124	0.0127	0.0125	0.0128	0.0127	0.0067	0.0072	0.0081	0.0072	0.0071
45	0.0115	0.0121	0.0121	0.0121	0.0130	0.0121	0.0124	0.0122	0.0125	0.0124	0.0062	0.0067	0.0076	0.0067	0.0075
46	0.0115	0.0121	0.0121	0.0121	0.0130	0.0121	0.0124	0.0122	0.0125	0.0124	0.0062	0.0067	0.0076	0.0067	0.0066
47	0.0118	0.0124	0.0124	0.0124	0.0133	0.0124	0.0127	0.0125	0.0128	0.0127	0.0067	0.0072	0.0081	0.0072	0.0071
48	0.0112	0.0118	0.0118	0.0118	0.0127	0.0118	0.0121	0.0119	0.0122	0.0121	0.0056	0.0062	0.0072	0.0062	0.0061
49	0.0115	0.0121	0.0121	0.0121	0.0130	0.0121	0.0124	0.0122	0.0125	0.0124	0.0062	0.0067	0.0076	0.0067	0.0066
50	0.0115	0.0121	0.0121	0.0121	0.0130	0.0121	0.0124	0.0122	0.0125	0.0124	0.0062	0.0067	0.0076	0.0067	0.0066
51	0.0115	0.0121	0.0121	0.0121	0.0130	0.0121	0.0124	0.0122	0.0125	0.0124	0.0062	0.0067	0.0076	0.0067	0.0066
52	0.0112	0.0118	0.0118	0.0118	0.0127	0.0118	0.0121	0.0119	0.0122	0.0121	0.0056	0.0062	0.0072	0.0062	0.0071
53	0.0115	0.0121	0.0121	0.0121	0.0130	0.0121	0.0124	0.0122	0.0125	0.0124	0.0062	0.0067	0.0076	0.0067	0.0075
54	0.0115	0.0121	0.0121	0.0121	0.0130	0.0121	0.0124	0.0122	0.0125	0.0124	0.0056	0.0062	0.0072	0.0062	0.0025
55	0.0115	0.0121	0.0121	0.0121	0.0136	0.0127	0.0130	0.0125	0.0128	0.0130	0.0056	0.0062	0.0072	0.0062	0.0056
56	0.0118	0.0124	0.0124	0.0118	0.0138	0.0130	0.0133	0.0128	0.0131	0.0132	0.0062	0.0067	0.0056	0.0056	0.0061
57	0.0108	0.0115	0.0115	0.0115	0.0130	0.0121	0.0124	0.0119	0.0122	0.0124	0.0043	0.0050	0.0062	0.0050	0.0043
58	0.0115	0.0121	0.0121	0.0115	0.0136	0.0127	0.0130	0.0125	0.0128	0.0130	0.0056	0.0062	0.0050	0.0050	0.0056
59	0.0112	0.0118	0.0118	0.0118	0.0127	0.0118	0.0121	0.0119	0.0122	0.0121	0.0050	0.0056	0.0067	0.0056	0.0035
60	0.0115	0.0121	0.0121	0.0115	0.0136	0.0127	0.0130	0.0125	0.0128	0.0130	0.0056	0.0062	0.0050	0.0050	0.0056
OTUs	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
1	0.0561	0.0543	0.0543	0.0516	0.0543	0.0535	0.0570	0.0624	0.0588	0.0561	0.0561	0.0588	0.0641	0.0641	0.0561
2	0.0588	0.0570	0.0570	0.0543	0.0570	0.0561	0.0597	0.0651	0.0614	0.0588	0.0588	0.0614	0.0668	0.0668	0.0588
3	0.0641	0.0624	0.0570	0.0597	0.0624	0.0614	0.0651	0.0705	0.0668	0.0641	0.0641	0.0668	0.0722	0.0722	0.0695

4	0.0561	0.0543	0.0543	0.0516	0.0489	0.0535	0.0570	0.0570	0.0561	0.0588	0.0641	0.0641	0.0614	0.0588	0.0641	0.0641	0.0641	0.0614	0.0614
5	0.0614	0.0597	0.0543	0.0570	0.0597	0.0588	0.0624	0.0624	0.0614	0.0678	0.0641	0.0614	0.0614	0.0641	0.0641	0.0695	0.0695	0.0668	0.0668
6	0.0749	0.0733	0.0678	0.0705	0.0733	0.0722	0.0760	0.0760	0.0776	0.0816	0.0749	0.0749	0.0749	0.0776	0.0831	0.0831	0.0831	0.0803	0.0803
7	0.0668	0.0651	0.0597	0.0624	0.0651	0.0641	0.0678	0.0678	0.0695	0.0733	0.0668	0.0668	0.0668	0.0695	0.0749	0.0749	0.0749	0.0722	0.0722
8	0.0641	0.0624	0.0570	0.0597	0.0570	0.0614	0.0651	0.0651	0.0614	0.0651	0.0641	0.0641	0.0641	0.0614	0.0668	0.0668	0.0668	0.0641	0.0641
9	0.0614	0.0597	0.0543	0.0570	0.0597	0.0588	0.0624	0.0624	0.0641	0.0678	0.0641	0.0614	0.0614	0.0588	0.0641	0.0641	0.0641	0.0614	0.0614
10	0.0614	0.0597	0.0543	0.0570	0.0597	0.0588	0.0624	0.0624	0.0641	0.0678	0.0641	0.0614	0.0614	0.0588	0.0641	0.0641	0.0641	0.0614	0.0614
11	0.0641	0.0624	0.0570	0.0597	0.0624	0.0614	0.0651	0.0651	0.0668	0.0705	0.0641	0.0641	0.0641	0.0614	0.0668	0.0668	0.0668	0.0641	0.0641
12	0.0484	0.0439	0.0439	0.0412	0.0439	0.0458	0.0465	0.0465	0.0511	0.0519	0.0484	0.0484	0.0484	0.0458	0.0511	0.0511	0.0511	0.0537	0.0537
13	0.0537	0.0439	0.0439	0.0412	0.0412	0.0511	0.0465	0.0465	0.0537	0.0492	0.0537	0.0537	0.0537	0.0511	0.0564	0.0564	0.0564	0.0537	0.0537
14	0.0535	0.0436	0.0436	0.0410	0.0436	0.0508	0.0463	0.0463	0.0561	0.0516	0.0535	0.0535	0.0535	0.0508	0.0561	0.0561	0.0561	0.0535	0.0535
15	0.0536	0.0438	0.0438	0.0411	0.0411	0.0510	0.0464	0.0464	0.0536	0.0491	0.0536	0.0536	0.0536	0.0510	0.0562	0.0562	0.0562	0.0536	0.0536
16	0.0508	0.0436	0.0436	0.0410	0.0436	0.0482	0.0463	0.0463	0.0535	0.0516	0.0535	0.0535	0.0535	0.0482	0.0535	0.0535	0.0535	0.0508	0.0508
17	0.0561	0.0463	0.0463	0.0436	0.0463	0.0535	0.0489	0.0489	0.0588	0.0543	0.0561	0.0561	0.0561	0.0535	0.0588	0.0588	0.0588	0.0561	0.0561
18	0.0561	0.0463	0.0463	0.0436	0.0463	0.0535	0.0489	0.0489	0.0588	0.0543	0.0561	0.0561	0.0561	0.0535	0.0588	0.0588	0.0588	0.0561	0.0561
19	0.0561	0.0463	0.0463	0.0436	0.0436	0.0535	0.0489	0.0489	0.0535	0.0489	0.0561	0.0561	0.0561	0.0535	0.0588	0.0588	0.0588	0.0561	0.0561
20	0.0643	0.0625	0.0625	0.0598	0.0625	0.0616	0.0653	0.0653	0.0669	0.0707	0.0643	0.0643	0.0643	0.0616	0.0669	0.0669	0.0669	0.0643	0.0643
21	0.0561	0.0543	0.0543	0.0516	0.0543	0.0535	0.0570	0.0570	0.0588	0.0624	0.0561	0.0561	0.0561	0.0535	0.0588	0.0588	0.0588	0.0561	0.0561
22	0.0591	0.0573	0.0573	0.0546	0.0573	0.0564	0.0600	0.0600	0.0617	0.0654	0.0591	0.0591	0.0591	0.0564	0.0617	0.0617	0.0617	0.0591	0.0591
23	0.0565	0.0520	0.0520	0.0493	0.0520	0.0539	0.0547	0.0547	0.0592	0.0601	0.0565	0.0565	0.0565	0.0539	0.0592	0.0592	0.0592	0.0565	0.0565
24	0.0592	0.0547	0.0547	0.0520	0.0547	0.0565	0.0574	0.0574	0.0619	0.0629	0.0592	0.0592	0.0592	0.0565	0.0619	0.0619	0.0619	0.0592	0.0592
25	0.0588	0.0570	0.0570	0.0543	0.0570	0.0561	0.0597	0.0597	0.0614	0.0651	0.0588	0.0588	0.0588	0.0561	0.0614	0.0614	0.0614	0.0588	0.0588
26	0.0126	0.0050	0.0050	0.0025	0.0050	0.0100	0.0075	0.0075	0.0126	0.0126	0.0151	0.0126	0.0126	0.0100	0.0177	0.0177	0.0177	0.0151	0.0151
27	0.0151	0.0075	0.0075	0.0050	0.0075	0.0126	0.0100	0.0100	0.0177	0.0151	0.0177	0.0151	0.0151	0.0126	0.0202	0.0202	0.0202	0.0177	0.0177
28	0.0202	0.0126	0.0126	0.0100	0.0126	0.0177	0.0151	0.0151	0.0202	0.0202	0.0202	0.0202	0.0202	0.0177	0.0254	0.0254	0.0254	0.0202	0.0202
29	0.0151	0.0075	0.0075	0.0050	0.0075	0.0126	0.0100	0.0100	0.0126	0.0100	0.0151	0.0151	0.0151	0.0126	0.0202	0.0202	0.0202	0.0177	0.0177
30	0.0025	0.0100	0.0100	0.0075	0.0100	0.0049	0.0075	0.0075	0.0049	0.0126	0.0049	0.0025	0.0025	0.0049	0.0199	0.0199	0.0199	0.0025	0.0025
31		0.0126	0.0126	0.0100	0.0126	0.0074	0.0100	0.0100	0.0074	0.0151	0.0074	0.0049	0.0049	0.0074	0.0224	0.0224	0.0224	0.0049	0.0049



32	0.0056		0.0050	0.0025	0.0050	0.0100	0.0075	0.0126	0.0151	0.0126	0.0126	0.0126	0.0151	0.0228	0.0228	0.0228	0.0202
33	0.0056	0.0035		0.0025	0.0050	0.0100	0.0075	0.0126	0.0151	0.0126	0.0126	0.0126	0.0151	0.0228	0.0228	0.0228	0.0202
34	0.0050	0.0025	0.0025		0.0025	0.0075	0.0050	0.0100	0.0126	0.0100	0.0100	0.0100	0.0126	0.0202	0.0202	0.0202	0.0177
35	0.0056	0.0035	0.0035	0.0025		0.0100	0.0075	0.0075	0.0100	0.0126	0.0126	0.0126	0.0151	0.0228	0.0228	0.0228	0.0202
36	0.0043	0.0050	0.0050	0.0043	0.0050		0.0126	0.0177	0.0099	0.0074	0.0074	0.0074	0.0099	0.0199	0.0199	0.0224	
37	0.0050	0.0043	0.0043	0.0035	0.0043	0.0056		0.0050	0.0126	0.0100	0.0100	0.0100	0.0126	0.0254	0.0254	0.0228	
38	0.0062	0.0056	0.0056	0.0050	0.0043	0.0067	0.0035		0.0075	0.0151	0.0151	0.0151	0.0177	0.0305	0.0305	0.0279	
39	0.0043	0.0062	0.0062	0.0056	0.0050	0.0050	0.0056	0.0043		0.0074	0.0074	0.0074	0.0099	0.0250	0.0250	0.0275	
40	0.0035	0.0056	0.0056	0.0050	0.0056	0.0043	0.0050	0.0062	0.0043		0.0049	0.0049	0.0074	0.0174	0.0174	0.0199	
41	0.0035	0.0056	0.0056	0.0050	0.0056	0.0043	0.0043	0.0062	0.0043	0.0035			0.0074	0.0224	0.0224	0.0250	
42	0.0043	0.0062	0.0062	0.0056	0.0062	0.0050	0.0056	0.0067	0.0050	0.0043	0.0043			0.0199	0.0199	0.0224	
43	0.0075	0.0076	0.0076	0.0072	0.0076	0.0071	0.0081	0.0089	0.0079	0.0066	0.0075	0.0075	0.0071			0.0049	0.0124
44	0.0075	0.0076	0.0076	0.0072	0.0076	0.0071	0.0081	0.0089	0.0079	0.0066	0.0075	0.0075	0.0071	0.0035			0.0124
45	0.0079	0.0072	0.0072	0.0067	0.0072	0.0075	0.0076	0.0085	0.0083	0.0071	0.0079	0.0079	0.0075	0.0056			
46	0.0071	0.0072	0.0072	0.0067	0.0072	0.0066	0.0076	0.0085	0.0075	0.0061	0.0071	0.0071	0.0066	0.0025	0.0025	0.0050	
47	0.0075	0.0076	0.0076	0.0072	0.0076	0.0071	0.0081	0.0089	0.0079	0.0066	0.0075	0.0075	0.0071	0.0035	0.0035	0.0056	
48	0.0066	0.0067	0.0067	0.0062	0.0067	0.0061	0.0072	0.0081	0.0071	0.0056	0.0066	0.0066	0.0061	0.0035	0.0035	0.0056	
49	0.0071	0.0072	0.0072	0.0067	0.0072	0.0066	0.0076	0.0085	0.0075	0.0061	0.0071	0.0071	0.0066	0.0043	0.0043	0.0061	
50	0.0071	0.0072	0.0072	0.0067	0.0072	0.0066	0.0076	0.0085	0.0075	0.0061	0.0071	0.0071	0.0066	0.0025	0.0025	0.0050	
51	0.0071	0.0072	0.0072	0.0067	0.0072	0.0066	0.0076	0.0085	0.0075	0.0061	0.0071	0.0071	0.0066	0.0043	0.0043	0.0061	
52	0.0075	0.0067	0.0067	0.0062	0.0067	0.0071	0.0072	0.0081	0.0079	0.0066	0.0075	0.0075	0.0071	0.0050	0.0050	0.0025	
53	0.0079	0.0072	0.0072	0.0067	0.0072	0.0075	0.0076	0.0085	0.0083	0.0071	0.0079	0.0079	0.0075	0.0056	0.0056	0.0035	
54	0.0035	0.0056	0.0056	0.0050	0.0056	0.0043	0.0050	0.0062	0.0043	0.0035	0.0035	0.0035	0.0043	0.0066	0.0066	0.0079	
55	0.0061	0.0056	0.0056	0.0050	0.0056	0.0056	0.0062	0.0072	0.0066	0.0061	0.0061	0.0061	0.0066	0.0083	0.0083	0.0087	
56	0.0066	0.0062	0.0062	0.0056	0.0056	0.0061	0.0067	0.0056	0.0050	0.0066	0.0066	0.0066	0.0071	0.0087	0.0087	0.0091	
57	0.0050	0.0043	0.0043	0.0035	0.0043	0.0043	0.0050	0.0062	0.0056	0.0050	0.0050	0.0050	0.0056	0.0075	0.0075	0.0079	
58	0.0061	0.0056	0.0056	0.0050	0.0043	0.0056	0.0062	0.0050	0.0043	0.0061	0.0061	0.0061	0.0066	0.0083	0.0083	0.0087	
59	0.0043	0.0050	0.0050	0.0043	0.0050	0.0035	0.0056	0.0067	0.0050	0.0043	0.0043	0.0043	0.0050	0.0071	0.0071	0.0075	



26	0.0151	0.0177	0.0126	0.0151	0.0151	0.0126	0.0151	0.0126	0.0126	0.0151	0.0075	0.0126	0.0100	0.0126
27	0.0177	0.0202	0.0151	0.0177	0.0177	0.0151	0.0177	0.0151	0.0151	0.0177	0.0100	0.0151	0.0126	0.0151
28	0.0228	0.0254	0.0202	0.0228	0.0228	0.0202	0.0228	0.0202	0.0202	0.0228	0.0151	0.0100	0.0177	0.0100
29	0.0177	0.0202	0.0151	0.0177	0.0177	0.0151	0.0177	0.0151	0.0151	0.0177	0.0100	0.0100	0.0126	0.0100
30	0.0174	0.0199	0.0149	0.0174	0.0174	0.0199	0.0174	0.0025	0.0124	0.0149	0.0074	0.0124	0.0049	0.0124
31	0.0199	0.0224	0.0174	0.0199	0.0199	0.0224	0.0199	0.0049	0.0149	0.0174	0.0099	0.0149	0.0074	0.0149
32	0.0202	0.0228	0.0177	0.0202	0.0202	0.0177	0.0202	0.0126	0.0126	0.0151	0.0075	0.0126	0.0100	0.0126
33	0.0202	0.0228	0.0177	0.0202	0.0202	0.0177	0.0202	0.0126	0.0126	0.0151	0.0075	0.0126	0.0100	0.0126
34	0.0177	0.0202	0.0151	0.0177	0.0177	0.0151	0.0177	0.0100	0.0100	0.0126	0.0050	0.0100	0.0075	0.0100
35	0.0202	0.0228	0.0177	0.0202	0.0202	0.0177	0.0202	0.0126	0.0126	0.0100	0.0075	0.0075	0.0100	0.0075
36	0.0174	0.0199	0.0149	0.0174	0.0174	0.0199	0.0174	0.0074	0.0124	0.0149	0.0074	0.0124	0.0049	0.0124
37	0.0228	0.0254	0.0202	0.0228	0.0228	0.0202	0.0228	0.0100	0.0151	0.0177	0.0100	0.0151	0.0126	0.0151
38	0.0279	0.0305	0.0254	0.0279	0.0279	0.0254	0.0279	0.0151	0.0202	0.0126	0.0151	0.0100	0.0177	0.0100
39	0.0224	0.0250	0.0199	0.0224	0.0224	0.0250	0.0224	0.0074	0.0174	0.0099	0.0124	0.0074	0.0099	0.0074
40	0.0149	0.0174	0.0124	0.0149	0.0149	0.0174	0.0149	0.0049	0.0149	0.0174	0.0099	0.0149	0.0074	0.0149
41	0.0199	0.0224	0.0174	0.0199	0.0199	0.0224	0.0199	0.0049	0.0149	0.0174	0.0099	0.0149	0.0074	0.0149
42	0.0174	0.0199	0.0149	0.0174	0.0174	0.0199	0.0174	0.0074	0.0174	0.0199	0.0124	0.0174	0.0099	0.0174
43	0.0025	0.0049	0.0049	0.0074	0.0025	0.0074	0.0099	0.0124	0.0275	0.0301	0.0224	0.0275	0.0199	0.0275
44	0.0025	0.0049	0.0049	0.0074	0.0025	0.0074	0.0099	0.0124	0.0275	0.0301	0.0224	0.0275	0.0199	0.0275
45	0.0099	0.0124	0.0124	0.0149	0.0099	0.0149	0.0025	0.0049	0.0250	0.0301	0.0250	0.0301	0.0224	0.0301
46		0.0025	0.0025	0.0049	0.0000	0.0049	0.0074	0.0099	0.0149	0.0250	0.0199	0.0250	0.0174	0.0250
47	0.0025		0.0049	0.0074	0.0025	0.0099	0.0025	0.0099	0.0124	0.0275	0.0224	0.0275	0.0199	0.0275
48	0.0025	0.0035		0.0025	0.0025	0.0025	0.0099	0.0124	0.0224	0.0250	0.0174	0.0224	0.0149	0.0224
49	0.0035	0.0043	0.0025		0.0049	0.0124	0.0149	0.0149	0.0250	0.0275	0.0199	0.0250	0.0174	0.0250
50	0.0000	0.0025	0.0025	0.0035		0.0049	0.0074	0.0099	0.0149	0.0250	0.0199	0.0250	0.0174	0.0250
51	0.0035	0.0025	0.0025	0.0035	0.0035		0.0124	0.0149	0.0250	0.0275	0.0199	0.0250	0.0174	0.0250
52	0.0043	0.0050	0.0050	0.0056	0.0043	0.0056		0.0025	0.0224	0.0275	0.0301	0.0224	0.0199	0.0275
53	0.0050	0.0056	0.0056	0.0061	0.0050	0.0061	0.0025		0.0250	0.0275	0.0250	0.0301	0.0224	0.0250

54	0.0061	0.0066	0.0056	0.0061	0.0061	0.0061	0.0075	0.0079	0.0149	0.0174	0.0099	0.0149	0.0074	0.0149
55	0.0079	0.0083	0.0075	0.0079	0.0079	0.0079	0.0083	0.0079	0.0061	0.0074	0.0049	0.0099	0.0074	0.0099
56	0.0083	0.0087	0.0079	0.0083	0.0083	0.0083	0.0087	0.0083	0.0066	0.0043	0.0124	0.0074	0.0149	0.0025
57	0.0071	0.0075	0.0066	0.0071	0.0071	0.0071	0.0075	0.0079	0.0050	0.0035	0.0056	0.0049	0.0025	0.0099
58	0.0079	0.0083	0.0075	0.0079	0.0079	0.0079	0.0083	0.0087	0.0061	0.0050	0.0043	0.0035	0.0074	0.0049
59	0.0066	0.0071	0.0061	0.0066	0.0066	0.0066	0.0071	0.0075	0.0043	0.0043	0.0061	0.0025	0.0043	0.0124
60	0.0079	0.0083	0.0075	0.0079	0.0079	0.0079	0.0083	0.0079	0.0061	0.0050	0.0025	0.0050	0.0035	0.0056

Appendix B  
 JUKES-CANTOR DISTANCE MATRIX AMONG C. miniata HAPLOTYPES

OTU Labels

1=T1 2=T2 3=T3 4=T4 5=T5 6=T6 7=T7 8=T8 9=T9 10=T10 11=T11 12=T12 13=T13 14=T14 15=T15 16=T16 17=T17 18=T18  
 19=T19 20=T20 21=J5 22=J7 23=J8 24=J1 25=J2 26=J3 27=J4 28=J9 29=J10 30=J11 31=J13 32=K1 33=J12 34=J14 35=J15 36=J16 37=J17  
 38=J18 39=J19 40=J20 41=K2 42=K3 43=K4 44=K5 45=K6 46=K7 47=K8 48=K9 49=K10 50=K11 51=K12 52=K13 53=K14 54=K15  
 55=K16 56=K17 57=K18 58=V1 59=V2 60=V3 61=V4 62=V5 63=V6 64=BB1 65=BB2 66=BB3 67=BB4 68=BB5

Distances in the upper-right matrix  
 Standard Errors in lower-left matrix  
 \* indicates an invalid distance value

OTUs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1		0.005	0.016	0.008	0.008	0.010	0.018	0.016	0.013	0.008	0.010	0.008	0.018	0.013	0.016	0.013	0.010	0.010	0.013
2	0.004		0.013	0.003	0.003	0.005	0.016	0.013	0.010	0.010	0.005	0.010	0.016	0.008	0.010	0.008	0.013	0.005	0.010
3	0.006	0.006		0.010	0.016	0.018	0.024	0.021	0.018	0.013	0.013	0.018	0.024	0.021	0.018	0.016	0.021	0.018	0.018
4	0.004	0.003	0.005		0.005	0.008	0.018	0.016	0.013	0.013	0.008	0.013	0.018	0.010	0.013	0.010	0.016	0.008	0.013
5	0.004	0.003	0.006	0.004		0.008	0.018	0.016	0.013	0.013	0.008	0.013	0.018	0.010	0.013	0.010	0.016	0.008	0.013
6	0.005	0.004	0.007	0.004	0.004		0.021	0.018	0.010	0.016	0.010	0.016	0.016	0.008	0.010	0.008	0.013	0.005	0.016
7	0.007	0.006	0.008	0.007	0.007	0.007		0.018	0.016	0.021	0.016	0.016	0.016	0.024	0.021	0.024	0.021	0.021	0.021
8	0.006	0.006	0.007	0.006	0.006	0.006	0.006		0.018	0.018	0.018	0.018	0.024	0.021	0.024	0.021	0.021	0.018	0.018
9	0.006	0.005	0.007	0.006	0.006	0.006	0.006	0.007		0.016	0.016	0.010	0.010	0.013	0.010	0.013	0.013	0.010	0.016
10	0.004	0.005	0.006	0.006	0.006	0.006	0.007	0.007	0.006		0.016	0.010	0.021	0.018	0.021	0.018	0.013	0.016	0.016
11	0.005	0.004	0.006	0.004	0.004	0.005	0.006	0.007	0.006	0.006		0.016	0.021	0.013	0.016	0.013	0.018	0.010	0.016
12	0.004	0.005	0.007	0.006	0.006	0.006	0.006	0.007	0.005	0.005	0.006		0.016	0.018	0.016	0.018	0.013	0.016	0.016
13	0.007	0.006	0.008	0.007	0.007	0.006	0.006	0.008	0.005	0.007	0.007	0.006		0.018	0.016	0.018	0.018	0.016	0.021
14	0.006	0.004	0.007	0.005	0.005	0.004	0.008	0.007	0.006	0.007	0.006	0.006	0.007	0.004	0.008	0.005	0.016	0.008	0.013
15	0.006	0.005	0.007	0.006	0.006	0.005	0.007	0.008	0.005	0.007	0.006	0.006	0.006	0.004	0.008	0.003	0.018	0.010	0.016
16	0.006	0.004	0.006	0.005	0.006	0.006	0.006	0.007	0.006	0.006	0.006	0.007	0.007	0.004	0.003	0.006	0.016	0.008	0.013
17	0.005	0.006	0.007	0.006	0.006	0.006	0.008	0.007	0.006	0.006	0.007	0.006	0.006	0.006	0.007	0.006	0.006	0.013	0.016
18	0.005	0.004	0.007	0.004	0.004	0.004	0.007	0.007	0.005	0.006	0.005	0.006	0.006	0.004	0.005	0.004	0.006	0.006	0.006
19	0.006	0.005	0.007	0.006	0.006	0.006	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.003	0.006
20	0.004	0.003	0.006	0.004	0.004	0.003	0.007	0.006	0.004	0.006	0.004	0.006	0.006	0.004	0.004	0.004	0.006	0.004	0.006
21	0.005	0.004	0.007	0.004	0.004	0.004	0.007	0.007	0.005	0.006	0.005	0.006	0.006	0.004	0.005	0.004	0.006	0.004	0.006
22	0.005	0.003	0.005	0.004	0.004	0.005	0.007	0.006	0.006	0.005	0.005	0.006	0.007	0.005	0.006	0.005	0.006	0.005	0.006
23	0.006	0.005	0.007	0.006	0.006	0.004	0.008	0.008	0.006	0.006	0.006	0.007	0.007	0.006	0.006	0.004	0.007	0.005	0.007
24	0.005	0.004	0.007	0.004	0.004	0.005	0.007	0.007	0.006	0.006	0.005	0.006	0.007	0.006	0.006	0.006	0.007	0.005	0.006
25	0.007	0.007	0.004	0.006	0.007	0.008	0.009	0.008	0.008	0.006	0.007	0.007	0.009	0.008	0.008	0.007	0.007	0.008	0.008
26	0.006	0.005	0.006	0.005	0.005	0.006	0.008	0.007	0.007	0.006	0.006	0.007	0.008	0.006	0.007	0.006	0.007	0.006	0.007
27	0.004	0.004	0.006	0.004	0.004	0.004	0.006	0.006	0.004	0.005	0.005	0.005	0.005	0.004	0.005	0.004	0.004	0.004	0.005
28	0.006	0.004	0.007	0.005	0.005	0.003	0.008	0.007	0.006	0.007	0.006	0.007	0.007	0.005	0.006	0.005	0.006	0.004	0.007
29	0.006	0.005	0.006	0.006	0.006	0.004	0.007	0.007	0.005	0.006	0.006	0.006	0.006	0.006	0.005	0.004	0.006	0.005	0.006



OTUs	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
1	0.008	0.010	0.008	0.016	0.010	0.018	0.013	0.008	0.013	0.013	0.018	0.010	0.010	0.016	0.018	0.013	0.010	0.013	0.005
2	0.003	0.005	0.003	0.010	0.005	0.018	0.008	0.005	0.008	0.010	0.013	0.008	0.008	0.013	0.013	0.008	0.008	0.008	0.008
3	0.016	0.018	0.010	0.018	0.018	0.005	0.016	0.013	0.021	0.013	0.026	0.010	0.016	0.010	0.026	0.016	0.016	0.021	0.016
4	0.005	0.008	0.005	0.013	0.008	0.016	0.010	0.008	0.010	0.013	0.016	0.010	0.010	0.010	0.016	0.010	0.010	0.010	0.010
5	0.005	0.008	0.005	0.013	0.008	0.021	0.010	0.008	0.010	0.013	0.016	0.010	0.010	0.010	0.016	0.010	0.010	0.010	0.010
6	0.003	0.005	0.008	0.005	0.010	0.024	0.013	0.005	0.003	0.005	0.008	0.013	0.008	0.018	0.008	0.008	0.008	0.003	0.013
7	0.018	0.021	0.018	0.026	0.021	0.029	0.024	0.016	0.024	0.021	0.029	0.018	0.018	0.024	0.018	0.024	0.018	0.024	0.018
8	0.016	0.018	0.016	0.024	0.018	0.027	0.021	0.013	0.021	0.018	0.026	0.016	0.016	0.021	0.021	0.021	0.016	0.021	0.016
9	0.008	0.010	0.013	0.016	0.016	0.024	0.018	0.005	0.013	0.010	0.018	0.013	0.008	0.018	0.018	0.013	0.008	0.013	0.013
10	0.013	0.016	0.008	0.016	0.016	0.016	0.013	0.010	0.018	0.016	0.024	0.013	0.013	0.018	0.024	0.018	0.013	0.018	0.008
11	0.008	0.010	0.008	0.016	0.010	0.018	0.013	0.010	0.013	0.016	0.018	0.013	0.013	0.018	0.018	0.013	0.013	0.013	0.013
12	0.013	0.016	0.013	0.021	0.016	0.021	0.018	0.010	0.018	0.016	0.024	0.013	0.013	0.018	0.024	0.018	0.013	0.018	0.008
13	0.013	0.016	0.018	0.021	0.021	0.029	0.024	0.010	0.018	0.016	0.024	0.018	0.013	0.024	0.024	0.018	0.013	0.018	0.018
14	0.005	0.008	0.010	0.013	0.013	0.027	0.016	0.008	0.010	0.013	0.016	0.016	0.010	0.021	0.016	0.010	0.010	0.010	0.016
15	0.008	0.010	0.013	0.010	0.016	0.024	0.018	0.010	0.013	0.010	0.018	0.013	0.013	0.018	0.018	0.008	0.013	0.013	0.018
16	0.005	0.008	0.010	0.008	0.013	0.021	0.016	0.008	0.010	0.008	0.016	0.010	0.010	0.016	0.016	0.005	0.010	0.010	0.016
17	0.010	0.013	0.016	0.018	0.018	0.021	0.021	0.008	0.016	0.013	0.021	0.016	0.010	0.021	0.021	0.010	0.010	0.016	0.005
18	0.003	0.005	0.008	0.010	0.010	0.024	0.013	0.005	0.008	0.010	0.013	0.013	0.008	0.018	0.013	0.008	0.008	0.008	0.013
19	0.013	0.016	0.013	0.021	0.016	0.024	0.018	0.010	0.018	0.016	0.024	0.013	0.013	0.018	0.024	0.018	0.013	0.018	0.008
20	0.003	0.005	0.005	0.008	0.008	0.021	0.010	0.003	0.005	0.008	0.010	0.010	0.005	0.016	0.010	0.005	0.005	0.005	0.010
21	0.003	0.005	0.008	0.010	0.010	0.024	0.013	0.005	0.008	0.010	0.008	0.013	0.008	0.018	0.008	0.008	0.008	0.008	0.013
22	0.004	0.005	0.006	0.013	0.008	0.016	0.005	0.008	0.010	0.013	0.016	0.010	0.010	0.016	0.016	0.010	0.010	0.010	0.010
23	0.004	0.005	0.006	0.006	0.016	0.024	0.018	0.010	0.008	0.005	0.013	0.013	0.013	0.018	0.013	0.008	0.013	0.008	0.018
24	0.004	0.005	0.005	0.006	0.006	0.024	0.013	0.010	0.013	0.016	0.018	0.013	0.013	0.018	0.018	0.013	0.013	0.013	0.013
25	0.007	0.008	0.006	0.008	0.008	0.024	0.016	0.018	0.027	0.018	0.032	0.016	0.021	0.011	0.032	0.018	0.021	0.027	0.019
26	0.005	0.006	0.004	0.007	0.006	0.006	0.006	0.013	0.016	0.018	0.021	0.010	0.016	0.016	0.021	0.016	0.016	0.016	0.016
27	0.003	0.004	0.005	0.005	0.005	0.007	0.006	0.004	0.008	0.005	0.013	0.008	0.003	0.013	0.013	0.008	0.003	0.008	0.008
28	0.004	0.004	0.005	0.004	0.006	0.008	0.006	0.004	0.004	0.008	0.010	0.016	0.010	0.021	0.010	0.010	0.010	0.005	0.016
29	0.004	0.005	0.006	0.004	0.006	0.007	0.007	0.004	0.004	0.006	0.007	0.008	0.008	0.013	0.013	0.008	0.008	0.008	0.013
30	0.005	0.004	0.006	0.006	0.007	0.009	0.007	0.006	0.005	0.006	0.007	0.007	0.016	0.026	0.010	0.016	0.016	0.010	0.021
31	0.005	0.006	0.005	0.006	0.006	0.006	0.005	0.004	0.005	0.004	0.006	0.005	0.010	0.010	0.021	0.010	0.010	0.010	0.010
32	0.004	0.004	0.005	0.006	0.006	0.007	0.006	0.003	0.005	0.004	0.006	0.005	0.010	0.016	0.016	0.010	0.005	0.010	0.010
33	0.006	0.007	0.006	0.007	0.007	0.005	0.006	0.006	0.007	0.006	0.008	0.005	0.006	0.016	0.026	0.016	0.016	0.021	0.016
34	0.005	0.004	0.006	0.006	0.007	0.009	0.007	0.006	0.005	0.006	0.008	0.005	0.006	0.006	0.006	0.016	0.016	0.010	0.016
35	0.004	0.004	0.005	0.004	0.006	0.007	0.006	0.004	0.005	0.004	0.006	0.005	0.004	0.006	0.006	0.016	0.016	0.021	0.016
36	0.004	0.004	0.005	0.006	0.006	0.008	0.006	0.003	0.005	0.004	0.006	0.005	0.004	0.006	0.006	0.016	0.016	0.010	0.016
37	0.004	0.004	0.005	0.004	0.006	0.008	0.006	0.004	0.004	0.004	0.005	0.006	0.005	0.007	0.005	0.005	0.005	0.010	0.016
38	0.005	0.006	0.005	0.007	0.006	0.007	0.006	0.004	0.006	0.006	0.007	0.005	0.005	0.006	0.006	0.006	0.006	0.006	0.005
39	0.004	0.004	0.004	0.006	0.004	0.007	0.005	0.004	0.005	0.006	0.006	0.005	0.005	0.006	0.006	0.005	0.005	0.005	0.005
40	0.004	0.004	0.005	0.004	0.006	0.008	0.006	0.004	0.005	0.006	0.006	0.006	0.005	0.007	0.006	0.004	0.005	0.005	0.006





12	0.013	0.018	0.016	0.013	0.016	0.018	0.021	0.018	0.018	0.013	0.018	0.013	0.016	0.013	0.018	0.013	0.018	0.013	0.010
13	0.018	0.018	0.016	0.021	0.018	0.018	0.021	0.018	0.018	0.018	0.018	0.018	0.026	0.018	0.018	0.013	0.018	0.013	0.021
14	0.010	0.010	0.013	0.010	0.013	0.010	0.013	0.010	0.010	0.010	0.010	0.021	0.024	0.010	0.016	0.010	0.016	0.010	0.018
15	0.013	0.013	0.016	0.013	0.016	0.013	0.016	0.013	0.013	0.013	0.013	0.018	0.021	0.013	0.013	0.013	0.018	0.013	0.021
16	0.010	0.010	0.013	0.010	0.013	0.010	0.013	0.010	0.010	0.010	0.010	0.016	0.018	0.010	0.010	0.010	0.016	0.010	0.018
17	0.016	0.010	0.013	0.016	0.018	0.016	0.018	0.016	0.016	0.016	0.016	0.021	0.018	0.016	0.016	0.016	0.016	0.010	0.013
18	0.008	0.008	0.010	0.008	0.010	0.008	0.010	0.008	0.008	0.008	0.008	0.018	0.021	0.008	0.008	0.013	0.008	0.008	0.016
19	0.013	0.018	0.016	0.013	0.016	0.018	0.021	0.018	0.018	0.013	0.013	0.018	0.021	0.013	0.013	0.018	0.013	0.013	0.016
20	0.005	0.005	0.008	0.008	0.005	0.008	0.008	0.005	0.005	0.005	0.005	0.016	0.018	0.005	0.005	0.005	0.010	0.005	0.013
21	0.008	0.008	0.010	0.008	0.010	0.008	0.010	0.008	0.008	0.008	0.008	0.018	0.021	0.008	0.008	0.013	0.008	0.008	0.016
22	0.005	0.010	0.013	0.008	0.005	0.008	0.013	0.010	0.010	0.010	0.005	0.010	0.018	0.005	0.010	0.005	0.016	0.010	0.013
23	0.013	0.008	0.016	0.013	0.016	0.013	0.016	0.013	0.013	0.013	0.008	0.013	0.021	0.013	0.013	0.013	0.018	0.010	0.016
24	0.008	0.013	0.016	0.010	0.010	0.018	0.016	0.013	0.013	0.013	0.008	0.018	0.021	0.008	0.013	0.008	0.018	0.013	0.016
25	0.021	0.024	0.024	0.018	0.024	0.027	0.029	0.027	0.027	0.027	0.016	0.005	0.021	0.021	0.027	0.021	0.027	0.021	0.021
26	0.010	0.016	0.018	0.013	0.010	0.013	0.021	0.018	0.016	0.016	0.010	0.010	0.024	0.010	0.016	0.024	0.021	0.016	0.018
27	0.008	0.008	0.005	0.010	0.008	0.010	0.008	0.010	0.008	0.008	0.013	0.013	0.016	0.008	0.008	0.016	0.008	0.003	0.010
28	0.010	0.010	0.013	0.010	0.013	0.010	0.013	0.010	0.010	0.010	0.010	0.021	0.024	0.010	0.010	0.024	0.010	0.016	0.018
29	0.013	0.013	0.010	0.016	0.013	0.013	0.013	0.013	0.013	0.013	0.008	0.013	0.013	0.013	0.013	0.016	0.013	0.005	0.016
30	0.016	0.016	0.018	0.016	0.018	0.016	0.013	0.016	0.016	0.016	0.016	0.026	0.029	0.016	0.016	0.029	0.018	0.013	0.024
31	0.010	0.016	0.013	0.010	0.013	0.010	0.013	0.010	0.016	0.016	0.005	0.010	0.013	0.010	0.010	0.013	0.010	0.010	0.013
32	0.010	0.010	0.008	0.013	0.010	0.013	0.010	0.013	0.010	0.010	0.010	0.016	0.018	0.010	0.010	0.018	0.010	0.005	0.013
33	0.016	0.021	0.018	0.013	0.016	0.013	0.024	0.021	0.021	0.021	0.021	0.005	0.018	0.016	0.018	0.016	0.021	0.016	0.018
34	0.016	0.016	0.018	0.010	0.018	0.010	0.018	0.021	0.013	0.016	0.016	0.026	0.029	0.016	0.016	0.029	0.018	0.013	0.024
35	0.010	0.005	0.013	0.010	0.013	0.010	0.013	0.010	0.010	0.010	0.005	0.010	0.013	0.010	0.010	0.018	0.010	0.005	0.013
36	0.010	0.010	0.008	0.013	0.010	0.010	0.013	0.010	0.010	0.010	0.010	0.016	0.018	0.010	0.010	0.018	0.010	0.005	0.013
37	0.010	0.010	0.013	0.010	0.013	0.010	0.013	0.010	0.010	0.010	0.010	0.021	0.024	0.010	0.010	0.024	0.010	0.008	0.013
38	0.010	0.016	0.013	0.010	0.013	0.010	0.013	0.016	0.016	0.016	0.010	0.016	0.013	0.010	0.010	0.013	0.010	0.016	0.008
39		0.010	0.013	0.008	0.005	0.008	0.016	0.008	0.010	0.010	0.005	0.016	0.018	0.005	0.010	0.018	0.005	0.010	0.013
40	0.005		0.013	0.013	0.010	0.013	0.016	0.008	0.010	0.010	0.010	0.021	0.024	0.010	0.010	0.024	0.010	0.010	0.013
41	0.006	0.006	0.006	0.013	0.016	0.013	0.016	0.013	0.013	0.013	0.013	0.018	0.021	0.013	0.013	0.021	0.013	0.008	0.016
42	0.004	0.006	0.006	0.008	0.010	0.013	0.016	0.013	0.013	0.013	0.008	0.013	0.018	0.013	0.013	0.021	0.008	0.013	0.016
43	0.004	0.005	0.006	0.008	0.008	0.013	0.016	0.013	0.013	0.013	0.005	0.010	0.013	0.005	0.010	0.018	0.005	0.010	0.013
44	0.004	0.006	0.006	0.004	0.005	0.004	0.016	0.018	0.008	0.013	0.008	0.018	0.021	0.008	0.008	0.021	0.008	0.013	0.016
45	0.006	0.006	0.006	0.007	0.007	0.006	0.018	0.016	0.016	0.016	0.016	0.021	0.024	0.016	0.016	0.024	0.016	0.010	0.018
46	0.004	0.004	0.006	0.006	0.006	0.006	0.007	0.013	0.013	0.013	0.013	0.021	0.026	0.013	0.013	0.026	0.013	0.010	0.016
47	0.005	0.005	0.006	0.006	0.006	0.006	0.006	0.010	0.010	0.010	0.005	0.021	0.024	0.010	0.010	0.024	0.010	0.010	0.018
48	0.005	0.005	0.006	0.006	0.006	0.006	0.006	0.005	0.005	0.005	0.005	0.021	0.018	0.010	0.010	0.018	0.010	0.010	0.018
49	0.006	0.007	0.007	0.006	0.006	0.006	0.007	0.007	0.007	0.007	0.010	0.021	0.018	0.010	0.010	0.018	0.010	0.016	0.018
50	0.004	0.005	0.006	0.004	0.004	0.004	0.006	0.005	0.005	0.005	0.010	0.021	0.018	0.010	0.010	0.018	0.010	0.016	0.018
51	0.005	0.005	0.006	0.005	0.005	0.004	0.006	0.005	0.004	0.005	0.005	0.007	0.005	0.010	0.005	0.013	0.005	0.010	0.013
52	0.004	0.005	0.006	0.004	0.004	0.004	0.006	0.005	0.005	0.005	0.005	0.006	0.007	0.005	0.005	0.024	0.010	0.016	0.018



24	0.010	0.010	0.016	0.010	0.010	0.010	0.010	0.010	0.010	0.016	0.018	0.013	0.016	0.016	0.016
25	0.018	0.024	0.024	0.018	0.024	0.024	0.024	0.024	0.024	0.021	0.027	0.021	0.021	0.021	0.024
26	0.013	0.013	0.018	0.013	0.013	0.013	0.013	0.013	0.013	0.018	0.021	0.016	0.018	0.018	0.018
27	0.010	0.005	0.010	0.005	0.010	0.005	0.010	0.005	0.010	0.005	0.008	0.008	0.010	0.005	0.005
28	0.013	0.008	0.013	0.013	0.013	0.008	0.013	0.008	0.013	0.013	0.016	0.005	0.018	0.013	0.013
29	0.010	0.010	0.010	0.010	0.016	0.010	0.016	0.010	0.010	0.010	0.013	0.008	0.016	0.010	0.010
30	0.018	0.013	0.013	0.018	0.018	0.013	0.018	0.013	0.018	0.018	0.021	0.010	0.024	0.018	0.018
31	0.008	0.013	0.013	0.008	0.013	0.013	0.013	0.013	0.013	0.013	0.016	0.016	0.013	0.013	0.013
32	0.013	0.008	0.013	0.008	0.013	0.008	0.013	0.008	0.008	0.008	0.010	0.010	0.013	0.008	0.008
33	0.013	0.018	0.018	0.013	0.018	0.018	0.018	0.018	0.018	0.018	0.021	0.021	0.018	0.018	0.018
34	0.018	0.013	0.018	0.018	0.018	0.018	0.018	0.013	0.018	0.018	0.016	0.010	0.024	0.018	0.018
35	0.008	0.008	0.008	0.013	0.013	0.008	0.013	0.008	0.013	0.013	0.016	0.010	0.018	0.013	0.013
36	0.013	0.008	0.013	0.008	0.013	0.008	0.013	0.008	0.008	0.008	0.010	0.010	0.013	0.008	0.008
37	0.013	0.008	0.013	0.013	0.013	0.008	0.013	0.008	0.013	0.013	0.016	0.005	0.018	0.013	0.013
38	0.013	0.013	0.018	0.008	0.013	0.013	0.013	0.013	0.013	0.003	0.016	0.016	0.008	0.013	0.013
39	0.008	0.008	0.013	0.008	0.008	0.008	0.008	0.008	0.008	0.013	0.016	0.010	0.013	0.013	0.013
40	0.013	0.008	0.013	0.013	0.013	0.013	0.013	0.008	0.013	0.013	0.016	0.010	0.018	0.013	0.013
41	0.016	0.010	0.016	0.010	0.016	0.010	0.016	0.010	0.010	0.010	0.013	0.013	0.016	0.010	0.010
42	0.010	0.010	0.016	0.010	0.010	0.010	0.010	0.010	0.010	0.016	0.018	0.013	0.016	0.016	0.016
43	0.008	0.008	0.013	0.008	0.008	0.008	0.008	0.008	0.013	0.013	0.010	0.010	0.013	0.013	0.013
44	0.010	0.010	0.016	0.010	0.010	0.010	0.010	0.010	0.010	0.016	0.018	0.013	0.016	0.016	0.016
45	0.018	0.013	0.013	0.013	0.018	0.013	0.018	0.013	0.013	0.013	0.016	0.016	0.018	0.013	0.013
46	0.016	0.010	0.016	0.016	0.016	0.016	0.016	0.010	0.016	0.016	0.018	0.008	0.021	0.016	0.016
47	0.010	0.008	0.013	0.013	0.013	0.013	0.013	0.008	0.013	0.013	0.016	0.010	0.018	0.013	0.013
48	0.013	0.008	0.013	0.013	0.013	0.013	0.013	0.008	0.013	0.013	0.016	0.010	0.018	0.013	0.013
49	0.013	0.018	0.018	0.013	0.018	0.013	0.018	0.018	0.018	0.018	0.021	0.021	0.018	0.018	0.018
50	0.003	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.013	0.016	0.010	0.013	0.013	0.013
51	0.013	0.008	0.013	0.013	0.013	0.013	0.013	0.008	0.013	0.013	0.016	0.010	0.018	0.013	0.013
52	0.008	0.008	0.013	0.008	0.003	0.008	0.003	0.008	0.008	0.013	0.016	0.010	0.013	0.013	0.013
53	0.016	0.021	0.021	0.016	0.021	0.021	0.021	0.021	0.021	0.016	0.024	0.024	0.016	0.021	0.021
54	0.008	0.008	0.013	0.008	0.008	0.008	0.008	0.008	0.008	0.013	0.016	0.010	0.013	0.013	0.013
55	0.018	0.013	0.018	0.013	0.018	0.013	0.018	0.013	0.013	0.013	0.016	0.016	0.018	0.013	0.013
56	0.013	0.008	0.013	0.008	0.013	0.008	0.013	0.008	0.008	0.008	0.010	0.008	0.013	0.008	0.008
57	0.016	0.016	0.021	0.010	0.016	0.016	0.016	0.016	0.016	0.010	0.018	0.018	0.010	0.016	0.016
58	0.005	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.018	0.013	0.016	0.016	0.016
59	0.005	0.005	0.010	0.010	0.010	0.005	0.010	0.005	0.010	0.010	0.013	0.008	0.016	0.010	0.010
60	0.005	0.005	0.006	0.016	0.016	0.010	0.016	0.010	0.016	0.016	0.018	0.013	0.021	0.016	0.016
61	0.005	0.005	0.006	0.006	0.010	0.010	0.010	0.010	0.010	0.010	0.013	0.013	0.010	0.010	0.010
62	0.005	0.005	0.006	0.005	0.005	0.010	0.010	0.010	0.010	0.016	0.018	0.013	0.016	0.016	0.016
63	0.005	0.004	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.010	0.013	0.008	0.016	0.016	0.016
64	0.006	0.005	0.006	0.005	0.006	0.005	0.006	0.005	0.005	0.013	0.013	0.013	0.010	0.010	0.010

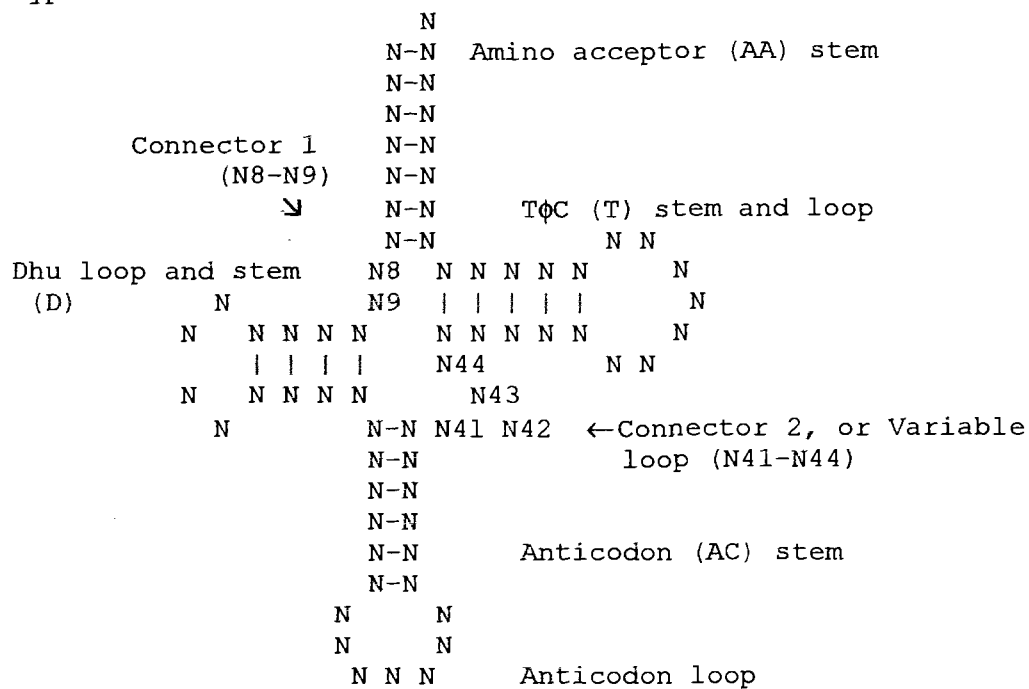
65	0.007	0.006	0.007	0.006	0.007	0.006	0.006	0.006	0.016	0.013	0.013
66	0.006	0.004	0.006	0.006	0.006	0.004	0.006	0.006		0.018	0.013
67	0.006	0.006	0.007	0.005	0.006	0.006	0.005	0.006	0.007		0.010
68	0.006	0.005	0.006	0.005	0.006	0.005	0.005	0.006	0.006	0.006	0.005

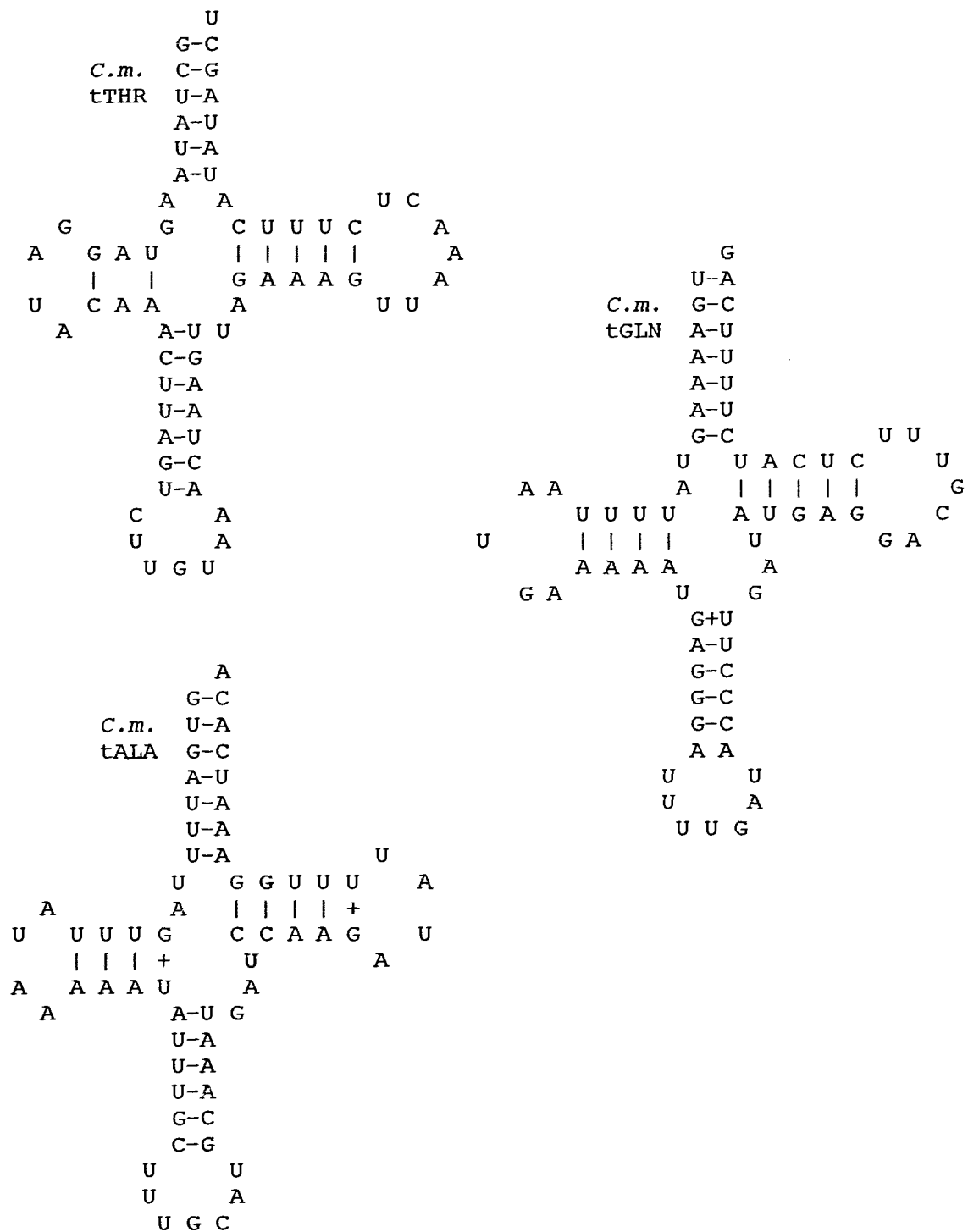
## Appendix C: tRNA Structures

Species abbreviations: *Cucumaria miniata*, *C. m.**C. pseudocurata*, *C. p.**Parastichopus californicus*, *P. c.*

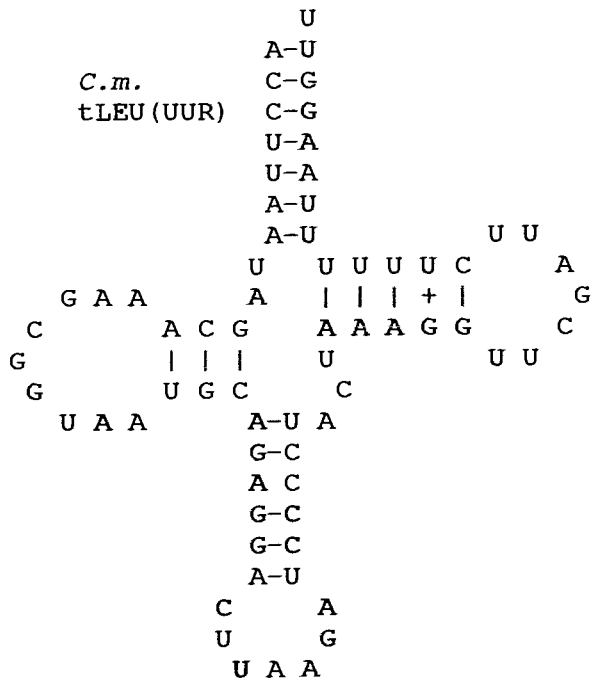
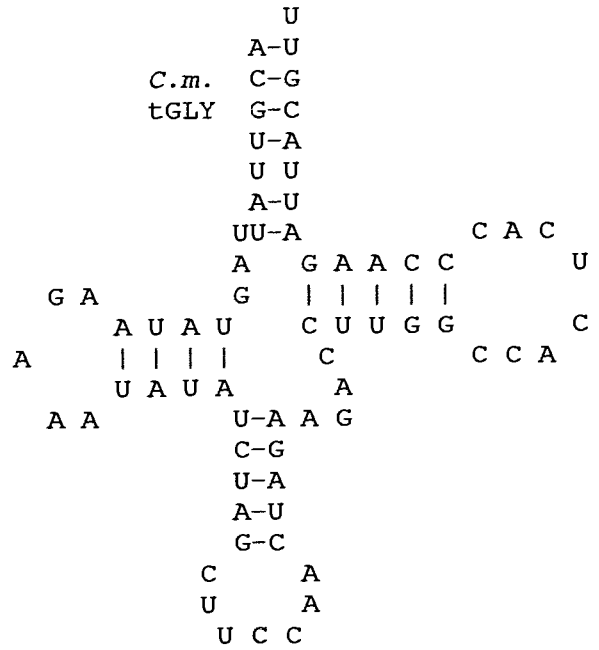
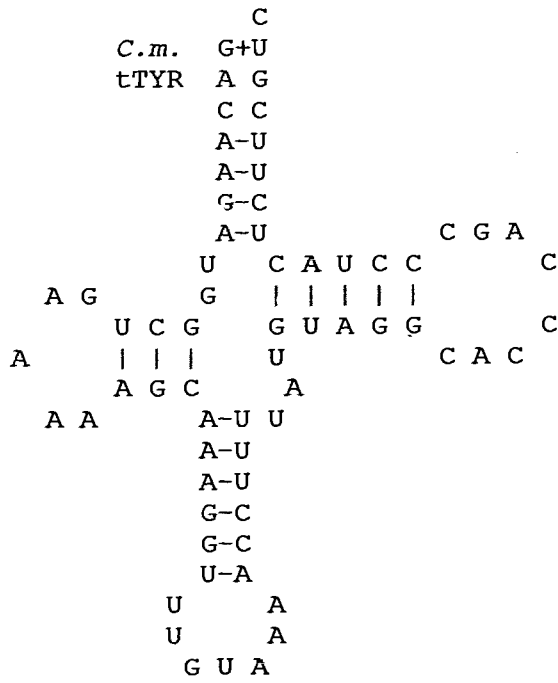
tRNAs labelled by corresponding three-letter amino acid symbol.

## Typical structure of mitochondrial tRNAs

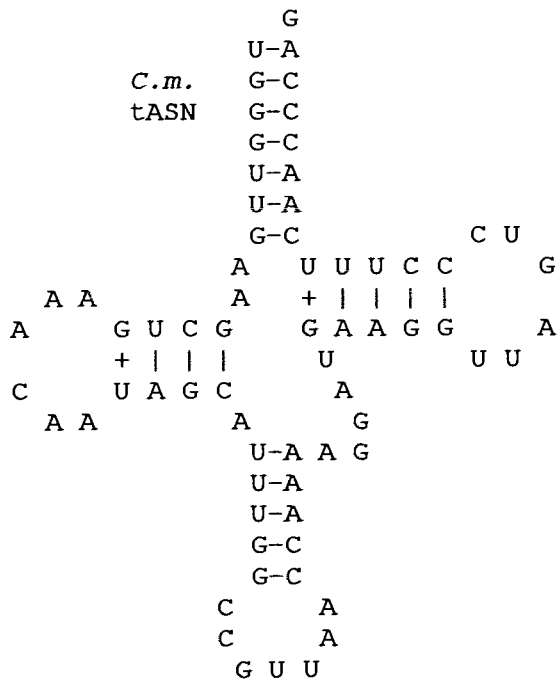
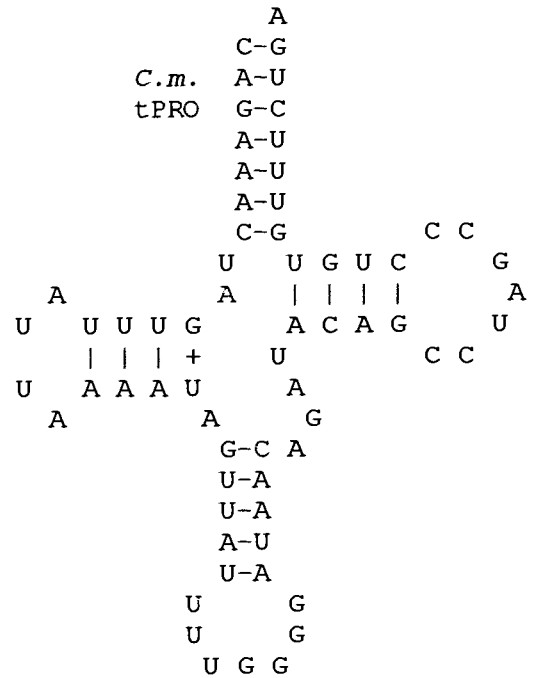
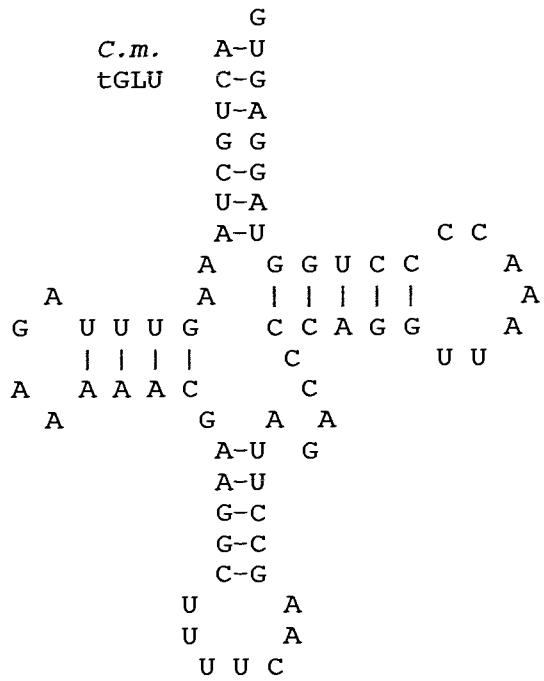


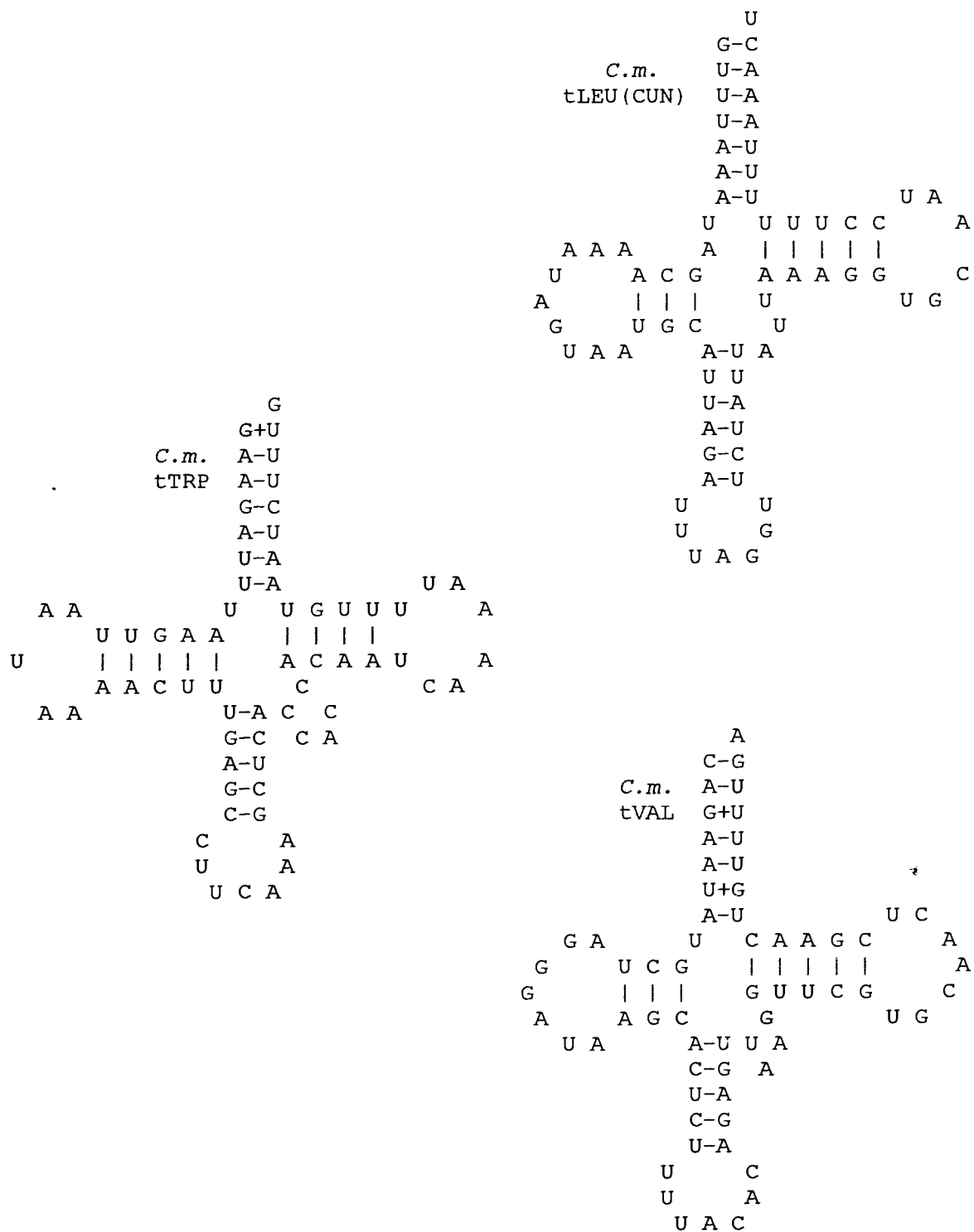


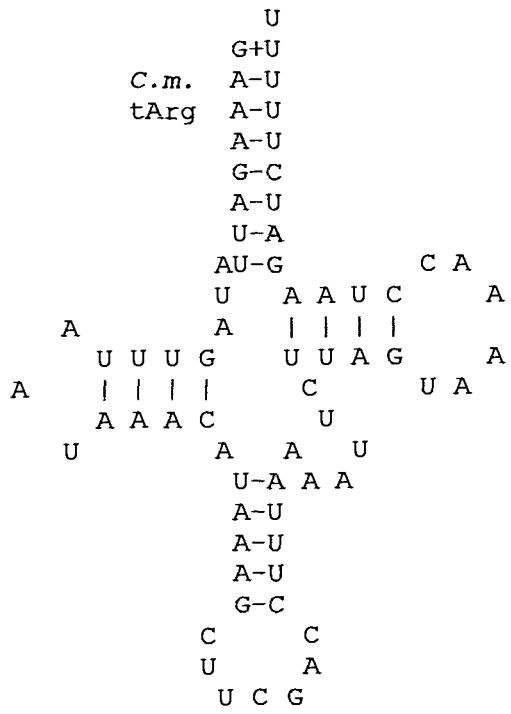












```

      U
      G-C
C.p.  C-G
tTHR  U-A
      A-U
      U-A
      A-U
      A G      U C
      G G C U U U C A
A G A U | | | | A
      | | G A A A G A
A C A A A U U
      A-U U
      C-G
      U-A
      U-A
      A-U
      G-C
      U-A
      C A
      U A
      U G U
  
```

```

      A
      U-A
      G-C
C.p.  A-U
tGLN  A-U
      A-U
      A-U
      G+U
      U U C
      A A U U A C U C G
      U U U U A | | | |
      | | | | A U G A G C
      A A A A U U G U C
      G A G+U G A
      A-U
      G-C
      G-C
      G-C
      A A
      U U
      U A
      U U G
  
```

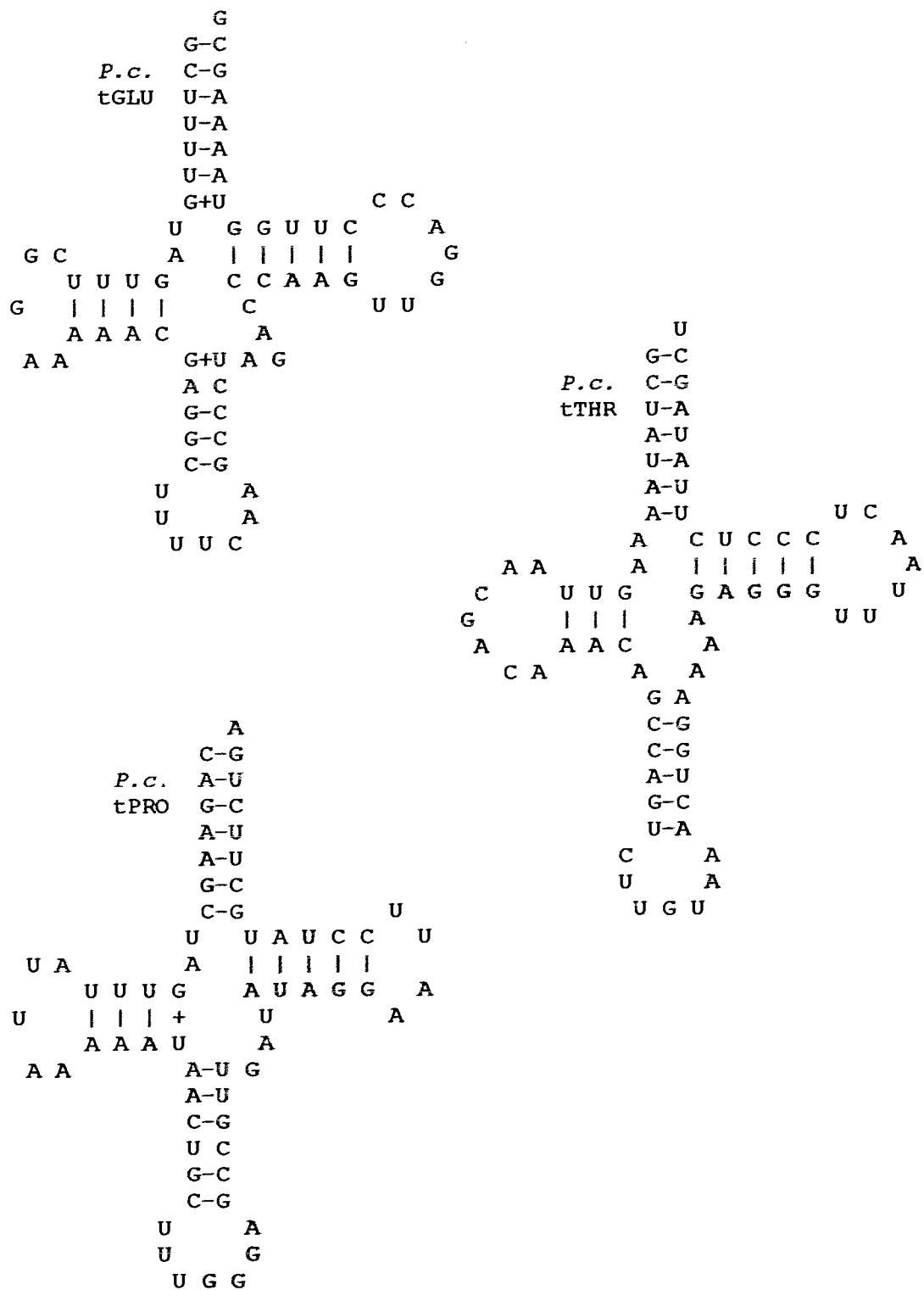
```

      A
      G-C
C.p.  U-A
tALA  G-C
      A-U
      U-A
      U-A
      U-A
      U A
      U G G U U U
      A | | | | U
C U U U G C C A A A
      | | | + U G A
A A A A U A
      A G
      U-A A
      U-A
      U-A
      G-C
      U+G
      U U
      U A
      U G C
  
```

```

          U
          G-C
    C.p.  U-A
    tCYS  U-A
          C-G
          U-A
          A-U
          A-U
          A C
    A A      G  G A U U U
          U G U G  | | | | |
    C      | | | +  C U A A A
          A C A U  A      C A
    C C      A      U
          U-A G U
          U-A
          U-A
          A-U
          A-U
          U      A
          U      A
          G C A

```



```

                G
                U-A
                P.c. G-C
                tGLN A-U
                G+U
                A-U
                A-U
                G-C
                CU
    AU          U  CATTTC  U
    A          UUUA  | | | | |  G
    U          | | | |  GUAAG  C
    G          AAAU  U          UU
    AA          A          A

```

```

                A-UAG
                G-C
                G-C
                G-C
                A-U
                C      U
                U      A
                UUG

```

```

                G
                U-A
                G-C
                P.c. A-U
                tASN G-C
                U-A
                U-A
                G-C
                UC
    CA          A  UUUCC  A
    G  A          A  | | | | |  U
    A          GUCG  GAAGG  G
    A          | | | |  U          UU
    A  A          UAGC  A
    A  A          A      G
    CG          U-AAG
                U-A
                U-A
                G-C
                G-C
                C      A
                C      A
                CUU

```

```

                U
                G-C
                U-A
                P.c. U-A
                tLEU U-A
                (CUN) A-U
                A-U
                A-U
                UA
                U  UUUCC  A
                A  | + | | |  A
    UGA          UUUG  AGAGG  C
    U          | | + |  C          UG
    A          AUGC  C
    G          A      A
    UAA          A      A
                A-UG
                U-A
                A-U
                G-C
                A-U
                U      A
                U      G
                UAG

```

```

      A
P.c. G-C
tALA U-A
      G-C
      G-C
      U U
      U C
      U-A
      AA
      U G A U U U
      A | | | | |
U A U U U G C U A A A U
      | | | | |
U A A A C A A G
      A A
      U-A
      U+G
      U-A
      G-C
      C-G
      C A
      U A
      U G C

```

```

      G
      G+U
P.c. A-U
tTRP A-U
      G-C
      A-U
      C-G
      U-A
      U A
      U U G U U U A
      A | | | | | G
      U U G G A C A A A A
      | | | | | U U
      A A C C A
      U A U-A A A
      G-C
      A-U
      G-C
      C-G
      C A
      U A
      U C C

```

```

      A
      G A
P.c. G+U
tCYS U-A
      C-G
      U-A
      U-A
      A-U
      A A A
      A A G A C A
      G | | | | |
U U G U G U C U G A
      | | | | |
U A C A U U G U U
      U
      G G
      U-A A
      U-A
      U-A
      G-C
      A-U
      U A
      U A
      G C A

```



```

          A
    P.c.  C-G
    tVAL  A-U
          A-U
          G-C
          A-U
          C-G
          A-U
          U C
          U C A A A C   A
    G G A   A | | | |   A
    G       U C G   G U U U G   C
          | | |   G       U G
    A       A G C   A
    G U A   U   G A
          A-U
          C-G
          U-A
          C-G
          U+G
          U   C
          U   A
          U A C
  
```