GLUTAMINE SYNTHETASE AS A BIOLOGICAL MARKER FOR FISH PHYLOGENETICS: SOME NEW INSIGHTS

by

TAMMY LABERGE MACDONALD

B.Sc., Simon Fraser University, 1991

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES
(Department of Zoology)

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

JANUARY 2003

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Zoology

The University of British Columbia
Vancouver, Canada

Date January 9th, 2002.

DE-6 (2/88)

Abstract

Glutamine synthetase is a key enzyme for nitrogen metabolism. It occurs in all organisms and is one of the oldest functioning genes. Many vertebrates have only one functional copy of this gene, while many plants have been shown to be multicopy for this gene. Pseudogenes for glutamine synthetase have also been reported in mammals. Until recently only a single copy of glutamine synthetase had been described in fish. However, six copies of this gene are expressed in rainbow trout *Oncorhynchus mykiss* and two copies of this gene are expressed in the gulf toadfish *Opsanus beta*.

We investigated a variety of intertidal fishes from British Columbia, Canada using PCR amplification of genomic DNA product and reverse transcriptase PCR to explore the diversity of glutamine synthetase in fish. We recovered two isoforms of glutamine synthetase in fourteen out of twenty-one fish. We describe the partial sequences for the two copies of this gene that differed in nucleotide composition by 8 to 22 percent.

Phylogenetic analysis was performed using the different glutamine synthetase isoforms to generate trees for intertidal fishes collected in this study. Fish from the following orders were represented in this study: Myxiniformes, Lepisosteiformes, Salmoniformes, Gasterosteiformes, Syngnathiformes, Scorpaeniformes, Perciformes and Pleuronectiformes. Most species adhered to the traditional taxonomic classification although some representative fish did not.

Table Of Contents

List Of Tables

List Of Figures

Acknowledgements

There are always so many people to thank when writing a thesis. I would like to start with my labmates: Shannon, Steve, Amanda, Ally and Lance. Someone once told me that labmates are special people because of what you have shared and that you will always have a tie to them. I have to agree. Shannon, you gave me my first fishing lesson and made coming to school fun. Steve, you always kept things interesting and made me realize life has a lot of aspects yet to explore. Amanda, you are always full of new ideas and were able to teach me not only about science, but about different ways to look at things. Lance, you were helpful, not only in the lab, but also in discussing future possibilities. Ally you helped me believe I could finish. You gave me confidence when I needed it and you make the best cookies in town!

I have also had the pleasure of having two supervisors so far in my life, Andy Beckenbach and Martin Adamson. Andy, I can't adequately describe the gratitude I have for you. You took a chance on me and let me try. When I first came to your lab I knew nothing about molecular biology. You let me learn and were very patient with me. You gave me the freedom to decide what I liked and never pressured me. I learned so much in the years I spent in your lab. It gave me the foundation so I could move on. I would like to say thank you for this, but it really doesn't seem like enough. Martin, you also took a chance on me. You also gave me freedom, freedom to choose a project that wasn't your main area of interest. Martin, I must say it has always been fun in your lab. I do not think I will find another supervisor that will go for coffee everyday with his students. I learned that education is a much larger lesson in life in your lab.

I have been fortunate to have a few mentors along the way and I would like to thank them now. Karen Beckenbach, Don Nelson and Murray Gilbert, thank you for teaching me

molecular techniques. I would never have been able to do this project without your guidance. I would also like to thank Don McPhail for his thoughtful discussions relating to my fish and the trees that resulted from my data. It is always great to be able to go to an expert for advice. I must also thank Bob Devlin and Rick Taylor who were on my committee. Thank you for your input into my thesis and its direction. Bob, you were able to make me feel much better about my project when I had some doubts.

There are many family and friends I would like to thank for their support. I would like to thank my Mom and Dad, my sister Penny and my brother Todd for putting up with me for all these years. I would also like to thank Garvin and Shirley MacDonald for their prayers and well wishes for me and for my thesis. Thank you to the friends from school that kept me sane: Manu and Phillip Gardner, Beth Zimmer, Jean Paul Danko, Scott Morgan, Karen Needham, Durrell Kapan and Dawn Cooper. Karmavores, you are a big part of my life, thanks for playing with me. It is so nice to have such a great loving and supportive group of friends.

I have one person left to thank. Glenn, thank you for always believing in me. You are my biggest fan. You allow me to be myself and support me in all that I do. I am always grateful for all that you do for me. Thank you and I love you.

Introduction

Glutamine synthetase (GS) is considered to be one of the oldest functioning genes (Kumada et al. 1993) and may have been present during the origin of life (Kumada et al. 1993; Tateno 1994). It is involved in nitrogen metabolism of all living organisms (Pesole et al. 1991; Kumada et al. 1993; Tateno 1994) where it converts glutamate to glutamine (Meister 1985; Eisenberg et al. 2000); in fish, GS removes toxic ammonia during this conversion (Meister 1985; Mommsen and Walsh 1992; Eisenberg et al. 2000) mediating the reaction is shown in Equation 1.

$$\text{Glutamate} + NH_3 + \text{ATP} \xleftrightarrow{Mg^{+2}} \text{Glutamine} + \text{ADP} + P_i$$

Equation 1.

For teleosts this reaction occurs just prior to the Ornithine-Urea Cycle (O-UC) (For review of O-UC see Mommsen and Walsh 1992).

Glutamine synthetase has been used to study phylogenetic relationships in all major groups of prokaryotes and eukaryotes (Pesole et al. 1991; Kumada et al. 1993; Tateno 1994; Pesole et al. 1995; Saccone et al. 1995). Glutamine synthetase evolves very slowly and is therefore used to look at older relationships of organisms (Tateno 1994). Pesole et al. (1991) report that GS evolves in a clock-like manner and GS follows the neutral evolution model (Kumada et al. 1993; Tateno 1994).

Of the two types of GS known; glutamine synthetase I occurs only in prokaryotes and glutamine synthetase II occurs mainly in eukaryotes, although some prokaryotes have been found with GSII (Hill et al. 1989; Goodman and Woods 1993; Kumada et al. 1993; Eisenberg et al. 2000). Glutamine synthetase I molecule is a dodecamer while the glutamine

synthetase II molecule is proposed to have eight subuits (Eisenberg et al. 2000). The existence of two types of GS implies gene duplication prior to the Prokaryote-Eukaryote split (Pesole et al. 1991; Kumada et al. 1993; Pesole et al. 1995; Saccone et al. 1995). Active sites of GSI and GSII are invariant indicating that their function is similar (Eisenberg et al. 2000).

Glutamine synthetase is a multimeric enzyme, and occurs in multigene families in plants (Cullimore et al. 1984; Tingey et al. 1987; Li et al. 1993; Temple et al. 1995). Some researchers found GS to be single-copy in a few vertebrates (Kuo and Darnell 1989; Pu and Young 1989; Campbell and Smith 1992; Laud and Campbell 1994). Multiple copies of GS were recently found in fish (Murray 2002; Walsh et al. 2002). This study reports on the sequence structure of a portion of the GS gene in a variety fish; it presents evidence of gene amplification and assesses phylogenetic relationships of these fish using GS as a biological marker.

Genomic DNA is examined for the presence GS sequence and revealed more than one GS-like sequence with different introns. Genomic DNA is comprised of functional genes with introns and pseudogenes and therefore complementary DNA (cDNA) is examined to determine which sequence was the functional GS gene since cDNA expresses only functional genes.

Twenty-one fish are examined in this study representing eight orders: Myxiniformes, Lepisosteiformes, Salmoniformes, Gasterosteiformes, Syngnathiformes, Scorpaeniformes, Perciformes and Pleuronectiformes. Myxiniformes is the most primitive of all of these orders containing the hagfish (Nelson 1994; Helfman et al. 1997) and the Lepisosteiformes containing gars is also fairly primitive (Nelson 1994; Helfman et al. 1997). The Salmoniformes are within the superorder Protacanthopterygii and are a sistergroup to the

superorder Acanthopterygii (Nelson 1994). Seventeen of 21 fish used in this study are within the last five orders and all fall within the superorder Acanthopterygii in the Series Percomorpha. The relationship of fishes within Series Percomorpha is an area of fish taxonomy that is still in flux (Johnson and Patterson 1993; Nelson 1994). While the composition of the species within Acanthopterygii is agreed upon, the taxonomy within this superorder is still unresolved (Johnson and Patterson 1993; Nelson 1994).

There are two main topics addressed in this thesis:

1. Is glutamine synthetase a multicopy gene in fishes? Is glutamine synthetase an appropriate marker for phylogenetic analysis?

2. Do the fish used in this study follow the traditional classification system for these types of fishes? Do the phylogenies produced in this study help resolve the superorder Acanthopterygii?

Materials and Methods:

Sample Collection and Preservation

Fish were sampled opportunistically by pole seining, rock tipping and dip netting and were identified using keys in Pacific Fishes of Canada (Hart 1988) and Fishes of the World (Nelson 1994). Tidepool sculpin (*Oligocottus maculosus* Girard 1856) were collected from Popham Island, B.C. in February 2000. Tubesnout (*Aulorhynchus flavidus* Gill, 1861), bay pipefish (*Syngnathus leptorhynchus* Girard, 1854), cabezon (*Scorpaenichthys marmoratus* Girard,1854), striped seaperch (*Embiotoca lateralis* Agassiz, 1854), shiner perch (*Cymatogaster aggregata* Gibbons, 1854), high cockscomb (*Anoplarchus purpurescens* Gill, 1861), penpoint gunnel (*Apodichthys flavidus* Girard, 1854), crescent gunnel (*Pholis laeta* (Cope, 1873)), speckled sanddab (*Citharichthys stigmaeus* Jordan and Gilbert, 1882), buttersole (*Isopsetta isolepis* (Lockington, 1880)), and starry flounder (*Platichthys stellatus* (Pallas, 1788)) were collected from the waters surrounding Stanley Park, Vancouver, B.C. in August 2000. Fish were euthanized with tricane methane sulfonate in seawater before being cut open from pectoral girdle to anus and placed in 95% ethanol. Three spined stickleback (*Gasterosteus aculeatus* Linnaeus, 1758), white spotted greenling (*Hexagrammos stelleri* Tilesius, 1810), buffalo sculpin (*Enophrys bison* (Girard, 1854)), Pacific sanddab (*Citharichthys sordidus* (Girard, 1854)), *A. flavidus*, *S. leptorhynchus*, *E. lateralis*, *C. aggregata*, *A. flavidus*, *P. laeta*, *P. ornata*, *C. stigmaeus* and *I. isolepis* were sampled while fishing in August 2001 in the waters around Stanley Park, Vancouver, B.C. These fish were euthanized, their carcasses were cut open and tissue samples were collected and immersed in liquid nitrogen and then stored at –80.0 °C. Danny Kent from the Vancouver Aquarium provided additional samples of *S. marmoratus*, and *P. stellatus* collected from the waters around Stanley Park, Vancouver, B.C. These samples were frozen in liquid nitrogen and

stored at –80.0 °C. Chum salmon (*Oncorhynchus keta* (Walbaum,1792)) and coho salmon (*Oncorhynchus kitsutch* (Walbaum, 1792)) were provided by Dr. Robert Devlin of DFO West Vancouver Labs. These samples were stored at –20.0 °C. Three spined stickleback (*Gasterosteus aculeatus* Linnaeus, 1758) and a fin clip of mossy sculpin (*Clinocottus embryum* (Jordan and Starks, 1895)) were provided by Patrick Tamkee from UBC. These samples were caught off Wizard Islet near Bamfield , B.C. and were preserved in 95% ethanol. Pacific hagfish (*Eptatretus stoutii* (Lockington, 1878)) samples were provided by Doug Fudge from UBC and were frozen at –20.0 °C and –80.0 °C. These samples were caught in the waters of Barclay Sound near Bamfield, B.C. Alligator gar (*Astractosteus spatula* Lacepede, 1803)) was provided by Dr. Robert Blake from UBC and was caught in the Gulf of Mexico. This sample was stored at –20.0 °C.

Primer Design

DNA sequence data for glutamine synthetase genes were retrieved from Genbank for all vertebrate specimens sequenced to August 1999 (Table 1). Initial alignment of Genbank sequences was performed using Clustal W (Thompson et al. 1994), and later adjusted by eye using ESEE 3.2S sequencing editor (Cabot and Beckenbach 1989)(Appendix 1). Several primers for glutamine synthetase were designed using OLIGO 4.04 (Rychlik and Rhoads 1989) based on regions of high conservation in the aligned sequences with the objective of providing specificity of amplification for GS sequences. Primers were designed to avoid primer-dimers and hairpins. Primers were designed to have a G-C content between 40 and 65% and have similar annealing temperatures. Primers were also designed to have the last six bases of the three prime end match at least seven animals in the aligned vertebrate sequences. Primers were constructed by the Nucleic Acid Protein Services unit of the University of British Columbia. Primers were initially tested only on *Oncorhynchus keta, Oncorhynchus kitsutch*, and *Oligocottus maculosus*. Primers that amplified glutamine synthetase in any of these fish and were used in the present study are

listed in Table 2. The alignment created during the primer design phase was also used to assess nucleotide composition of glutamine synthetase for all vertebrates in Genbank prior to August 1999 (Appendix 2).

Genomic DNA Extractions

Muscle, skin or liver tissue was used for DNA extractions. Using a UV sterilized scalpel, 30 to 40 mg of tissue was cut away from the fish carcass and placed in a 1.7 ml microcentrifuge tube. Ethanol-preserved specimens were soaked in 0.5 ml proteinase K buffer (0.05M EDTA, 0.5% SDS, 0,01 M Tris, 2.0 M NaCl) for 5 minutes to remove traces of ethanol. The buffer was replaced with 0.36 ml fresh proteinase K buffer and 0.04 ml proteinase K enzyme (28 mg/ml). Samples were incubated at 65.0 °C until the tissue was digested (usually within 8 to 12 hours). Digested samples were extracted twice with phenol:sevag, and once with sevag (24 Chloroform: 1 isoamyl alcohol). DNA was precipitated in cold 95% ethanol and left overnight at –20.0 °C. DNA was pelleted by centrifugation, washed twice with cold 70% ethanol and left to air dry overnight. DNA pellets were resuspended in 0.1 ml water and stored at –20.0 °C.

RNA Extractions

RNA extractions were performed on frozen tissue only. RNA was extracted using Qiagen RNeasy Midi extraction kit (Mississauga, Ontario). The protocol for isolation of total RNA from heart, muscle, and skin tissue with the following modifications to the protocol: tissue was ground in liquid nitrogen with a mortar and pestle before homogenizing it in homogenization buffer in a 10 ml Kontes tissue grinder; samples were centrifuged at 4500 rpm for twice the recommended time outlined in the protocol. The RNA product was taken up in 50 – 100 ul of RNase-free water, precipitated in 3 volumes cold 95% ethanol and left overnight at –20.0 °C. RNA was pelleted by centrifugation, washed twice with cold 80%

ethanol and dried by heating in a 37.0 °C heating block for 15 minutes. RNA pellets were resuspended in 0.1 ml water and stored at –20.0 °C.

Reverse Transcriptase Reactions

Single stranded cDNA was generated from RNA by reverse transcription (RT). RNA was prepared for RT by combining 50 pmol Oligo $d(T)_{12\text{-}18}$ (Amersham Pharmacia Biotech) with 1-2 ug of RNA, and denatured by heating to 95.0 °C for two minutes then put on ice. RT was performed in 25 ul reactions containing the denatured RNA, 1x PCR buffer (20 mM Tris-HCl (pH 8.4) and 50 mM KCl) (Invitrogen, Carlsbad, California), 0.01 M DTT, 2.5 mM $MgCl_2$, 0.4 mM each dNTP, 15 units RNAguard (Amersham Pharmacia Biotech), 200 units Superscript II RT (Invitrogen) and ddH_2O. The RT reaction was placed into a thermocycler and incubated at 42.0 °C for 50 minutes, 65.0 °C for 15 minutes and then cooled to 4.0 °C. The cDNA product was stored at –20.0 'C.

PCR and Sequencing

One to five microlitres of genomic DNA or cDNA were used for PCR. PCR reactions were performed in 25.0 ul volumes each containing 1x PCR buffer, 0.2 mM each dNTP, 1.6 mM $MgCl_2$, 1.25 units Taq polymerase (Invitrogen), 0.38 mM each primer, and water. The primers used for glutamine synthetase amplification are outlined above. PCR was carried out on a Perkin Elmer Geneamp® PCR system 2400. PCR conditions for amplification were as follows: One denaturation cycle of 95.0 °C for 3 minutes followed by three initial amplification cycles of 95.0 °C for 90 seconds, 48.0 °C for 45 seconds, 70.0 °C for 2:00 minutes, then thirty two regular amplification cycles of 95.0 °C for 1:00 minute, 50.0 °C for 30 seconds, 70.0 °C for 2:00 minutes followed by a 72.0 °C extension for 5:00 minutes, and finally a 4.0 °C soak file was activated. Five microlitres of PCR product was

electrophoresed on a 1.0% agarose gel stained with ethidium bromide (5.7×10^{-4} mg/ml) to confirm presence of fragment.

PCR product was gel purified using a QIAquick gel extraction kit (QIAGEN). The concentration of the PCR product was determined and 30-90 ng of the purified dsDNA PCR product was used for sequencing reactions. The PCR product was sequenced from both the 5' and 3' direction. Automated sequencing reactions used AmpliTaq FS DyeDeoxy Terminator Cycle Sequence chemistry (Applied Biosystems (ABI)). Excess terminators were removed by running the sequencing reaction product through Centri-Sep Spin columns (emp Biotech GmbH). Sequencing reactions were sent to the UBC NAPS sequencing facility where they were run on an ABI Model 373 Stretch DNA sequencer or an ABI Prism 377 DNA Sequencer. Sequence printouts were visually inspected for any anomalies.

Sequence Analysis and Phylogenetics

Sequences were manually aligned with ESEE Version 3.2S (Cabot and Beckenbach 1989) or IMSEA (Beckenbach unpublished) sequencing editors and compared to published glutamine synthetase sequence of gulf toadfish (*Opsanus beta*). Sequences were also compared to GS sequences present in Genbank using BLAST (Altschul et al. 1997). Introns were identified for genomic DNA by determining intron spice sites in the sequence using the methods of S. Mount (1982). Only sequence data from coding regions of DNA was used for analysis. Sequence data was analyzed for a 432 bp fragment of glutamine synthetase. Base composition, parsimony informative sites by codon position and pairwise distances were determined using IMSEA (Beckenbach unpublished). Phylogenetic trees were generated using parsimony, distance and likelihood methods of the PAUP* Version 4.0b 10 (Swofford 2002) and of MEGA version 2.1 (Kumar et al. 2001). Trees were

created using a heuristic search with random addition (50 replicates for parsimony and 10 replicates for maximum likelihood), and TBR branch swapping algorithm for PAUP* or CNI for MEGA. Neighbor joining trees were generated using Kimura-2-parameter distance. All trees were bootstrapped (n=100). Trees were run unweighted and weighted (2:4:1 by codon position). Pacific hagfish was used as an outgroup. Trees for individual isoforms of glutamine synthetase were also assessed. Trees were compared with a morphological tree based on orders of fishes (Nelson 1994; Helfman et al. 1997). Common names, family names and orders of fish used in this study are listed in Table 3.

Results

DNA Amplification

Primer sets involving eleven primers successfully amplified glutamine synthetase-like product in *Oncorhynchus keta, Oncorhynchus kitsutch*, and *Oligocottus maculosus*. These primers were then used on the remainder of the fish. Amplification was not successful for all primers on all fish, so multiple primers were used on some fish. Fragment size varied due to amplification of different types of DNA product, complementary DNA (cDNA) and genomic DNA (Table 4). Complementary DNA produced a smaller fragment size and did not vary because cDNA does not contain introns. Genomic DNA produced products of varying lengths due to variation in intron lengths between species. Five intron sites were identified between positions 230/231, 392/393, 539/540, 667/668, and 867/868 of the coding sequence (Figure 1). Introns had an average length of 108bp, 88bp, 90bp, 102bp, and 122bp respectively. No intron site data was generated for fish whose GS product was generated from cDNA only.

Five primers were largely successful and therefore used extensively. Primer GS-237 or primer GS-232 used with primer GSR-911 produced ~ 700 bp product when amplifying cDNA, and ~1050bp to ~1150 bp product when amplifying genomic DNA. Primer GS-448 used with primer GSR-977 amplified ~550 bp product with cDNA and ~800 bp to 900 bp product with genomic DNA.

Sequence from the region of overlap amplified from the two primer sets (corresponding to positions 467 through 899 of published GS sequence *Xenopus laevis* Genbank accession number D50062) in the cDNA of the above fish produced two different glutamine synthetase products for shiner perch, coho salmon, white spotted greenling, penpoint gunnel, Pacific sanddab, three spined stickleback, tubesnout, buttersole, cabezon and

crescent gunnel (Figure 1 and Appendix 2). The average difference between these two isoforms from within the same fish was ~18 % with the largest difference being 22% and occurring in both speckled sanddab and Pacific sanddab and the smallest difference of 8% occurring in shiner perch (Figure 2).

Two different products from cDNA indicate that glutamine synthetase has more than one transcript for GS which is therefore not a single copy gene but a multicopy gene in these animals. A second copy of glutamine synthetase product was also observed in the coding sequence of the genomic product for speckled sanddab, starry flounder, high cockscomb, and mossy sculpin. These sequences also produced different introns between the two different GS products (Figure 1). Only one copy of glutamine synthetase was observed for hagfish, alligator gar, striped seaperch, buffalo sculpin, tidepool sculpin, bay pipefish and chum salmon.

Isoform designation

Neighbor joining compares the distances or raw sequence similarity between sequences and was performed for the multiple cDNA products and genomic products (after removal of introns). This produced a tree (Figure 3) with a distinct clade (bootstrap value 99) for one isoform in 12 of the 14 fishes compared: cabezon, white spotted greenling, mossy sculpin, tubesnout, three spine stickleback, high cockscomb, penpoint gunnel, crescent gunnel, starry flounder, buttersole, Pacific sanddab and speckled sanddab. Also grouped within this clade were both isoforms from the shiner perch. The distances between the fish within this clade was ≤ 15 % (Figure 2). Therefore for this paper, an isoform of glutamine synthetase is identified as the sequence from either cDNA or genomic DNA (less introns) amplification which is resolved into a distinct clade by neighbor joining, whose overall similarity to other sequences within the clade is 15 % or less. A recent differentiation of an isoform may

occur for a fish within a clade, but is labeled by numbers after the isoform designation indicating more than one copy of the isoform from a particular fish originates somewhere within that clade.

The clade with 13 of the fish represented is herein referred to as the A clade, with the isoforms found in it designated as A for each species and with two distinct shiner perch isoforms referred to as A1 and A2, where the average distances within the A clade was smaller for A1 than A2 of the shiner perch isoforms. Figure 3 also shows a separation of the second isoforms of the above listed fish in Clade A (except shiner perch), indicating that the second sequence products isolated from these fish did not all represent the same glutamine synthetase isoform. Eight fish were grouped into the same clade for the second isoform (bootstrap 99): white spotted greenling, three spine stickleback, tubesnout, mossy sculpin, cabezon, high cockscomb, penpoint gunnel and crescent gunnel. The pairwise distances for this clade were $\leq 8\%$ (Figure 2). This clade is therefore labeled as B and all the fish within it have an isoform designation of B. Four fish were grouped into another clade for their second isoform (bootstrap 79): Pacific sanddab, speckled sanddab, starry flounder and buttersole. The pairwise distances within this clade was $\leq 15\ \%$ (Figure 2). This clade is therefore referred to as clade C and all the fish within it have an isoform designation of C. Also observed within this tree (Figure 3) was the separation of the coho genes into their own clade (bootstrap 72), with the pairwise distance between the two isoforms being only 13 % (Figure 3) and were therefore labeled D1 and D2.

Overall Analysis GS products

All amplification products from all fish used in this study, including those that only produced one gene product, were analyzed. A region corresponding to positions 467 to 899 of the published sequence of *Xenopus laevis* (Genbank accession number D50062) was

used for analysis. Base composition for fish sequence is reported in Table 5. and is similar with that reported for GS of other vertebrates (Appendix 3). Isoforms did not vary in overall base composition.

Amino acid translation shows sites that are conserved in GS sequences of all other organisms (Eisenberg et al. 2000) are also conserved in the fish used in this study. One exception however, occurred at position 135, where all organisms code for alanine, whereas fish with the B isoform coded for isoleucine. Active sites for this region of GS (Eisenberg et al. 2000) were also maintained (Figure 4).

Amino acid composition also loosely supports the existence of multiple genes of GS. At position 86 of Figure 4, fish with isoform C code for isoleucine while the amino acid for this position varies for the other isoforms from either valine or alanine. Fish with the D isoform code for serine at position 33 whereas other isoforms code for alanine at this position. Also seen in the D isoform a methionine at position 55 but the B isoform codes for aspartic acid here. At position 102 isoforms A, C and D code for alanine whereas isoform B codes for valine; at position 135 isoform A codes for alanine whereas isoform B codes for isoleucine.

Phylogenetic Analysis GS products

Phylogenetic analysis using neighbor joining, regardless of software used, weighted or unweighted, supported trees with similar topology (Figure 5). In each case isoforms A, B, C and D occurred in separate clades. Shiner perch had two sequences which clustered within the A clade. Both coho salmon isoforms clustered within the D clade. The pipefish GS isoform appears to have arisen as a sistergroup to gene A for GS although this branch

was not well supported but was consistently outside of the gene A cluster. The alligator gar isoform always arose on its own branch.

Analysis of sequence data using IMSEA (Beckenbach unpublished) revealed 239 fixed sites in the nucleotide data. There were 193 parsimony informative sites, 41 occurring at first codon positions, 22 occurring at second codon positions and 130 occurring at third codon positions. Parsimony analysis both weighted and unweighted produced trees with similar topologies separating the GS isoforms (Figure 6) but not completely identical topologies as with the neighbor joining tree above. Using weighted parsimony analysis, the C isoforms did not separated into a single clade with two branches, but were part of a ladder from which the B clade branched off (Appendix 4). Unweighted parsimony analysis resulted in the separation of the two isoforms of coho salmon onto different branches but the branch with coho D1 has low bootstrap support (under 50% not shown) (Figure 6). Both shiner perch isoforms clustered within the A clade. The pipefish isoform showed the same pattern as seen in the neighbor joining tree (Figure 5) branching as a sistergroup to the A clade.

Maximum likelihood analysis also separated isoforms A and B both with weighted and unweighted analysis but isoform C did not form its own clade in either likelihood analysis. Isoform D formed its own clade in the weighted analysis (Figure 7), but not for the unweighted analysis (data not shown). Again both shiner perch genes clustered within the A cluster, however the pipefish isoform was also within this cluster although the branch was weakly supported (under 50% not shown). Weighted and unweighted maximum likelihood trees had similar topologies, however the weighted tree had better resolution and higher bootstrap support (Figure 7).

Isoform analysis and phylogenetic analysis of fishes

Isoform A analysis produced trees with similar topologies regardless of method of analysis or weighting method (Figures 8, 9 and 10). The only difference between isoform A trees was the resolution level. Isoform B analysis also gave similar tree topologies regardless of method used for analysis or weighting method (Figures 11, 12 and 13). Isoform C trees were all the same (Data not shown), with one clade of the order Pleuronectiformes with two branches, each with 100% bootstrap support; one for the family Paralichthyidae which includes the Pacific sanddab and the speckled sanddab and one for the family Pleuronectidae which includes the starry flounder and the buttersole.

Generally fish clustered within their families. Order separation was not evident except in Scorpaeniformes. Within the order Scorpaeniformes, white spotted greenling (family Hexagammidae) never arose within the Cottidae clade (sculpins and cabezon) but always arose near this familial group. In all trees Pacific sanddab and speckled sanddab were grouped within their family Pleuronectidae, however the A isoform trees did not join the order Pleuronectiformes together with its two represented families Pleuronectidae and Paralichthyidae. In all trees was a clade of the order Gasteriformes, which paired as sister taxa tubesnout with three spine stickleback.

Common to both isoform A and isoform B trees was a clade of the family Pholidae (penpoint gunnel and crescent gunnel) as sister group to the family Stichaeidae (high cockscomb). These fish are within the order Perciformes but never clustered with Embiotocidae, the other family within this order (shiner perch and striped seaperch). Shiner perch isoform A1 always paired with striped seaperch isoform A in all A isoform trees but not with the Pholidae/Stichaeidae clade. Shiner perch A2 isoform always appeared as a

sistergroup to the butterole isoform A which is in a completely different order than the shiner perch.

Since more species produced isoform A and isoform B products, these isoforms were used for phylogenetic analysis of the taxa sampled for this study with the exception of isoform C's grouping of the order Pleuronectiformes. The 432 bp fragment of isoform A was amplified in all fish except alligator gar. For isoform A there was 272 conserved sites and 160 variable sites. Pairwise differences revealed that penpoint gunnel and crescent gunnel were the most genetically similar for isoform A (0.5% difference), while shiner perch A2 and Pacific hagfish were the least genetically similar (27.1 % difference) and speckled sanddab and mossy sculpin were the least similar of all the A isoforms (15.0 % difference)(Figure 14). Weighted analyses gave better resolution and higher bootstrap values than unweighted analyses for isoform A. Isoform B was only amplified in eight taxa: white spotted greenling, penpoint gunnel, three spine stickleback, tubesnout, mossy sculpin, high cockscomb, cabezon and cresent gunnel. Again weighted analyses gave better resolution and higher bootstrap values than unweighted analyses. For isoform B there was 370 fixed sites and 62 variable sites. Pairwise differences again showed that penpoint gunnel and crescent gunnel were the most genetically similar for isoform B (1.0 % difference) and tubesnout was most different from Pacific hagfish (27.6% difference for these distances) while within the B isoforms, cabezon and three spine stickleback were the least genetically similar (7.7 % difference)(Figure 14).

Discussion

Glutamine synthetase gene duplication(s)

DNA and cDNA roducts isolated in this study are indeed glutamine synthetase gene(s) since the functional sites as given in Eisenberg (2000) are conserved (Figure 4) and the products are similar to GS products in Genbank. Glutamine synthetase can no longer be considered a single copy gene in eukaryotes. Multiple isoforms of GS in cDNA indicates that this gene is a multicopy gene perhaps even part of a gene family. Multiple copies of glutamine synthetase are also found in kidney bean, peas, alfalfa, and corn (Cullimore et al. 1984; Tingey et al. 1987; Li et al. 1993; Temple et al. 1995). Different genes in plants are expressed in different plant tissues. Multiple copies of GS have also been found in fish not used in this study and showed differential tissue expression (Murray 2002; Walsh et al. 2002). In other teleosts three copies of GS have been isolated in diploid fish and six GS isoforms were isolated in fish with a polyploid ancestry (2002 Busby, Ellen, University of Victoria, pers. comm.). Pseudogenes have also been observed in human and mouse (Chakrabarti et al. 1995) which may imply multiple copies of glutamine synthetase were present and were subsequently lost by mutation. Glutamine synthetase has recently been shown to have differential expression in different developmental stages in two eukaryotes *Zea mays* (Li et al. 1993) and *Oncorhynchus mykiss* (2002 Wright, P. A., University of Guelph, pers. comm.), however, no developmental information was obtained for this study.

More than two different isoforms of GS are represented in this study. The neighbor-joining analysis of all isoforms reflects this (Figure 5). Multiple clades are formed with a higher degree of divergence than the within clade divergence. Most fish fall have an isoform that falls within the A clade; exceptions are bay pipefish, coho salmon, chum salmon and alligator gar. These fish likely have the A isoform but it was not amplified by the methods utilized in this study. Lack of PCR product of the A isoform only implies that

the primers utilized in this study were unable to amplify the A gene for the tissue used for extraction.

Bay pipefish did not group within any clade and instead formed its own branch. Both coho salmon isoforms were isolated in a clade, distinct from the A, B or C clade. Chum salmon grouped with coho salmon.

Alligator gar formed its own branch. This might have been predicted because it is a more primitive fish than the other fish used in this study and could be used as an outgroup with the Pacific hagfish, but this branch may represent a paralogous gene since gars are polyploid (Schultz 1980).

The separation of B and C isoforms into separate clades indicates that the second isoform isolated from most fish in this study did not represent the same gene. Isoforms A and B were isolated for white spotted greenling, penpoint gunnel, crescent gunnel, high cockscomb, mossy sculpin, cabezon, three spine stickleback and tubesnout. Isoforms A and C were isolated for the Pacific sanddab, speckled sanddab, starry flounder, and buttersole.

At the beginning of this study, differential expression of GS was not known, and therefore care was not taken to isolate specific organs. In some cases all of the organs within the gut cavity were combined. The GS sequence determined in bay pipefish was isolated from RNA and therefore sensitive to tissue specificity. The visceral tissue of this animal is surrounded by a large block of muscle tissue, and it is likely that the RNA extracted from this animal came predominantly from muscle tissue and secondarily from visceral tissue. Most other fish had large enough visceral cavities, that the predominant tissue isolated was liver or intestinal tissue for RNA extractions. Tissue from shiner perch was extracted multiple times due to amplification problems and isoform A2 was predominantly from gill

tissue, not liver, intestinal or muscle tissue. Isoform A2 from shiner perch represents a paralogous gene within shiner perch which has only differentiated a small amount from the A isoform. Differential expression of GS genes may explain why sequences in these fish were different from the other dominant isoforms and showed up in unexpected locations for the overall phylogenetic analyses.

Glutamine synthetase enzyme is thought to have undergone a duplication event. Bacterial GS I is differs from eukaryotic GS II (Kumada et al. 1993; Tateno 1994; Pesole et al. 1995) and another type of GSIII also occur in *Bacteroides fragilis* (Hill et al. 1989) and *Butyrivibrio fibrisolvens* (Goodman and Woods 1993). This duplication and subsequent divergence occurred prior to the divergence of prokaryotes and eukaryotes (Pesole et al. 1991; Kumada et al. 1993; Tateno 1994; Pesole et al. 1995; Saccone et al. 1995).

Gene duplications arise by multiple methods and can produce multilple sized products. Duplications can arise within a gene or spanning a complete gene by unequal crossing over during recombination resulting in tandem repeats on a chromosome (Ohno 1970; Li and Graur 1991; Twyman 1998; Freeman and Herron 2001). Evidence for tandem duplication occurs in the vertebrate lineage of the globin gene family (Proudfoot and Maniatis 1980; Freeman and Herron 2001). Duplication can also occur on a larger scale where regional portions of chromosomes or entire chromosomes (aneuploidy) are duplicated (Li and Graur 1991) (Ohno 1970; Allendorf and Thorgaard 1984; Twyman 1998). Finally, gene duplications can arise from polyploidy (Ohno 1970; Allendorf and Thorgaard 1984; Li and Graur 1991; Twyman 1998; Freeman and Herron 2001). Polyploidization results from a multiplication of an organism's entire genome. Many eukaryotes are descendants of lineages that have undergone polyploidization (For review see Otto and Whitton, 2000). Multiple rounds of genome duplication occurred in fish and this could explain the multiple

copies of genes or even gene families found in fish (Ohno 1970; Holland et al. 1994; Wittbrodt et al. 1998; Meyer and Schartl 1999; Taylor et al. 2001).

Only three fish used in this study are known polyploids: alligator gar (Schultz 1980), coho salmon (Schultz 1980; Allendorf and Thorgaard 1984), and chum salmon (Schultz 1980; Allendorf and Thorgaard 1984). The latter two are partial tetraploids – only a portion of their genome remains polyploid. Chromosome numbers in polyploid fish are generally higher than non-polyploid fish (Schultz 1980; Allendorf and Thorgaard 1984). There is no indication that any of the other fish used in this study are polyploid. Their chromosome numbers are similar to the average chromosome number for non-polyploid fish (Froese and Pauly 2000). In alligator gar, coho salmon and chum salmon, the GS isoform(s) expressed did not group with the GS isoforms expressed in the majority of the rest of the fish in this study. It is likely that the isoforms expressed in the polyploid fish represented different paralogous genes not found in the non-polyploid fish and is therefore reflected in the phylogenetic trees when all isoforms were analyzed (Figures 5, 6 and 7) by appearing as separate branches. Isoforms orthologous to either the A gene or the B gene in the polyploid fish were not amplified likely due to the differential expression of the GS isoforms. The method of gene duplication for glutamine synthetase cannot be determined in this study.

Phylogenetic analysis of GS products

With the exception of Pacific hagfish, alligator gar, coho salmon and chum salmon, fish used in this study are within the superorder Acanthoptyergii. This superorder is not phylogenetically resolved by morphology alone. There are two competing hypotheses for the resolution of this group that differ mostly in the composition of the Series Percomorpha (Figure 15). Johnson and Patterson (1993) combine the five orders Synbranchiformes, Elassomatidae, Gastereosteiformes, Mugiloidei and Atherinomorpha into a monophyletic

group within the Percomorpha series and refer to this group as Smegmamorpha. Dactylopteriformes, Scorpaeniformes, Perciformes, Pleuronectiformes and Tetradontiformes orders remain as unresolved in the Johnson and Patterson (1993) classification system. Nelson (1994) however, does not recognize the Smegmamorpha and defines the Percomorpha differently, although Nelson's overall classification of the superorder Acanthoptyerigii comprises the same orders and families as that of Johnson and Patterson (Nelson 1994). Fish used in this study, except those listed above, fall into the Percomorpha.

Phylogenies generated by the data for GS isoform A does not support either Johnson and Patterson's classification or Nelson's classification system. For both GS isoforms A and GS isoforms B, Gasteriformes (three spined stickleback and tubesnouts) always appear as sister taxa to the branch containing Pholidae (penpoint gunnel and crescent gunnel) and Stichaeidae (high cockscomb) which are members of the order Perciformes (Figures 5, 6, 7, 8, 9, 10, 11, 12 and 13). Stichaidae always form as a sister group to Pholididae.

The Perciformes also includes the family Embiotocidae (shiner perch and striped seaperch) which do not group with the other Perciform families the Pholids or Stichaeids (Figures 5, 6, 7, 8, 9 and 10) anywhere in this study. This is not completely surprising as the Perciform order is considered polyphyletic (Johnson 1993; Johnson and Patterson 1993; Nelson 1994) and has no synapomorphy to support it (Nelson 1994). Nelson expects this order to undergo a re-classification in the near future (Nelson 1994); therefore Pholidae and Stichaeidae may eventually be re-classified outside of the Perfiformes branch and closer to the Gasteriformes.

Glutamine synthetase isoform A did not result in a monophyletic grouping of the Order Pleuronectiformes (Pacific sanddab, speckled sanddab, starry flounder and buttersole) as

was predicted by the morphological analysis of Chapleau (1993). Species within the two families Paralichthyidae (Pacific sandddab and speckled sanddab) and Pleuronectidae (starry flounder and buttersole) did group together for both GS isoform A and GS isoform C. GS isoform C grouped the flatfish together, but the only Pleuronectiformes produced isoform C. Therefore grouping Pleuronectiformes together may be artificial due to the presence of a common isoform and not actually a phylogenetic resolution of the order. The familial relationships of the Pleuronectiformes are still not determined and are likely to change in the future (Johnson 1993; Johnson and Patterson 1993). Isoform A analysis clearly does not support monophyly of this order perhaps reflecting the gene phylogeny and not the species phylogeny.

Scorpaeniformes (cabezon, buffalo sculpin, tidepool sculpin, mossy sculpin and white spotted greenling) grouped together for GS isoform A on one branch (Figures 8, 9 and 10) but did not form one branch for GS isoform B (Figures 11, 12 and 13). Bootstrap support for GS isoforms B was also lower than that for GS isoforms A. The lack of resolution for GS isoform B may just reflect the overall smaller sequence divergence in GS isoform B than in GS isoform A (Figure 14) and therefore less informative sites available to group the Scorpaeniformes into one order on one branch. The phylogeny of the Scorpaeniform fish generated in this study for GS isoform A agrees with classification of Scorpaeniformes presented by Imamura and Shinohara (1998) and the classification of Nelson 1994.

Glutamine synthetase isoform A separated the orders Scorpaeniformes and Gasteriformes into their own clades, but did not support monophyly of Perciformes or Pleuronectiformes (Figure 8, 9 and 10). Glutamine synthetase isoform B separated the Perciformes and Gasteriformes, but did not support monophyly of Scorpaeniformes. These differences may reflect both the differences in type and number of species available for analysis of both

isoforms for this study. These differences my also be caused by different phylogenetic patterns for the genes studied.

Conclusions and future work

Glutamine synthetase is a multicopy gene or part of a gene family in fish. Not all the fish in this study produced two copies of glutamine synthetase but all of these fish likely have at least three copies of GS. This study did not assess the question of multiple copies of GS systematically. Differential expression of GS within the tissues was not addressed. Tissues sampled for this study were mostly from visceral tissue and not from heart, brain, gill (except shiner perch) or skin. In future studies each organ within a fish should be sampled separately for RNA in order to determine the exact number of GS genes and where in the fish these new genes are active. It is also important to sample fish at different developmental stages to determine if there is any developmental pattern of gene activity.

The primers used in this study were based on vertebrate sequence alignments from Genbank (Appendix 1) and therefore may have a bias favoring one GS gene so the primers may not be suitable to amplify all three (or even two) GS genes. If all copies within a fish are determined from multiple fish, it may be possible to design new primers that target specific isoforms of GS. The number of GS isoforms may also be identified using cDNA libraries, but all tissues would have to be represented and each species of fish addressed separately.

Glutamine synthetase may no longer be useful for phylogenetic analysis of fishes unless all isoforms of GS are isolated for each fish and compared for analysis. Most of the fish used in this project were frozen and therefore effort should be made to isolate the remaining isoforms from the different tissues of the fish.

As for the phylogeny of the fishes used in this study, there is still much work to be done. Although this study mostly agreed with classical fish taxonomy, there were exceptions. I would not suggest re-classifying the Pholidae (penpoint and crescent gunnels) and Stichaeidae (high cockscomb) out of the order Perciformes and into their own order based on the evidence of one gene, but the relationship of these species should be re-examined and more genes should be analyzed. More molecular work should be done to try and aid in the classification of these fish. Future studies should be carefully planned to address some of the classification questions still unresolved and should include additional representative species from each order.

Species Name	Common Name	Genbank Accession Number
Cricetulus griseus	Chinese hamster	AF150961
Opsanus beta	Gulftoadfish	AF118103
Heterodontus francisci	Hornshark	AF118104
Gillichthys mirabilis	Long-jawed mudsucker	AF266200
Scyliorhinus torazame	Cloudy catshark	AF306642
Danio rerio	Zebrafish	AW076779
Ictalurus punctatus	Channel catfish	BE469571
Xenopus laevis	African clawed frog	D50062
Bos taurus	Cow	J03604
Gallus gallus	Chicken	M29076
Rattus norvegicus	Norway Rat	M29579
Mus musculus	House Mouse	M60803
Squalus acanthias	Spiny dogfish	U04617
Cricetulus longicaudatus	Long-tailed hamster	X03495
Homo sapiens	Human	X59834
Sus scrofa	Pig	Z29636

Table 1. Glutamine synthetase sequences available in Genbank for all vertebrates prior to August 1999.

Primer Name	Direction	Position*	Sequence (5' to 3')
GS-101	Forward	101 - 119	GTGAAGAAGCAGTACATGG
GS-232	Forward	232 - 249	TCTACCTGAATGGAACTT
GS-237	Forward	237 - 254	CAGAATGGAACTTTGATGG
GS-300	Forward	300 - 318	TCGTTCCTGCTGCCATGTT
GS-448	Forward	448 - 465	CCCTTGGTTTGGAATGGA
GS-537	Forward	537 - 554	AAGGTCCCTATTACTGTG
GSR-548	Reverse	548 - 565	TGCTCCAAATCCACAGTA
GSR-802	Reverse	802 - 819	CACCCAGCACCATTCCAG
GSR-911	Reverse	911 - 928	GTAGGCAAGGATGTGGTA
GSR-977	Reverse	977 - 994	GTTGATGTTGGAGGTTTC
GSR-1069	Reverse	1069 - 1086	CGGCGGTCTTCAAAGTAG

Table 2. Primers used in the amplification and sequencing of glutamine synthetase product. * indicates sequence position in *X. laevis* of Appendix 1.

Common Name	Latin Name	Family	Subfamily	Order
Pacific hagfish	*Eptatretus stoutii* (Lockington, 1878)	Myxinidae	Eptatretinae	Myxiniformes
Alligator gar	*Atractosteus spatula* (Lacepede, 1803)	Lepisosteidae	N/A	Lepisosteiformes
Chum salmon	*Oncorhynchus keta* (Walbaum, 1792)	Salmonidae	Salmoninae	Salmoniformes
Coho salmon	*Oncorhynchus kisutch* (Walbaum, 1792)	Salmonidae	Salmoninae	Salmoniformes
Three spined stickleback	*Gasterosteus aculeatus* Linnaeus, 1758	Gasterosteidae	Gasterosteinae	Gasterosteiformes
Tubesnout	*Aulorhynchus flavidus* Gill, 1861	Aulorhynchidae	N/A	Gasterosteiformes
Bay pipefish	*Syngnathus leptorhynchus* Girard, 1854	Syngnathidae	Syngnathinae	Syngnathiformes
White spotted greenling	*Hexagrammos stelleri* Tilesius, 1810	Hexagrammidae	Hexagramminae	Scorpaeniformes
Cabezon	*Scorpaenichthys marmoratus* Girard, 1854	Cottidae	N/A	Scorpaeniformes
Buffalo sculpin	*Enophrys bison* (Girard, 1854)	Cottidae	N/A	Scorpaeniformes
Tidepool sculpin	*Oligocottus maculosus* Girard, 1856	Cottidae	N/A	Scorpaeniformes
Calico sculpin a.k.a. Mossy Sculpin	*Clinocottus embryum* (Jordan & Starks, 1895)	Cottidae	N/A	Scorpaeniformes

Table 3. Names of animals used in this study including common names, species names, family names, subfamily names and order names.

Common Name	Latin Name	Family	Subfamily	Order
Striped seaperch	*Embiotoca lateralis* Agassiz, 1854	Embiotocidae	N/A	Perciformes
Shiner perch	*Cymatogaster aggregata* Gibbons, 1854	Embiotocidae	N/A	Perciformes
High cockscomb	*Anoplarchus purpurescens* Gill, 1861	Stichaeidae	N/A	Perciformes
Penpoint gunnel	*Apodichthys flavidus* Girard, 1854	Pholidae	N/A	Perciformes
Crescent gunnel	*Pholis laeta* (Cope, 1873)	Pholidae	N/A	Perciformes
Pacific sanddab	*Citharichthys sordidus* (Girard, 1854)	Paralichthyidae (Bothidae)	N/A	Pleuronectiformes
Speckled sanddab	*Citharichthys stigmaeus* Jordan & Gilbert, 1882	Paralichthyidae (Bothidae)	N/A	Pleuronectiformes
Buttersole	*Isopsetta isolepis* (Lockington, 1880)	Pleuronectidae	Pleuronectinae	Pleuronectiformes
Starry flounder	*Platichthys stellatus* (Pallas, 1788)	Pleuronectidae	Pleuronectinae	Pleuronectiformes

Table 3. continued

Primer Pair	Approximate Size of Genomic DNA Fragment	Approximate size of cDNA Fragment
GS-101 with GSR-548	660 bp	470 bp
GS-101 with GSR-1069	1500 bp	990 bp
GS-232 with GSR-911	1100 bp	700 bp
GS-232 with GSR-977	1160 bp	760 bp
GS-232 with GSR-1069	1250 bp	850 bp
GS-237 with GSR-911	1100 bp	700 bp
GS-237 with GSR-977	1160 bp	760 bp
GS-237 with GSR-1069	1250 bp	850 bp
GS-300 with GSR-802	800 bp	520 bp
GS-448 with GSR-977	860 bp	550 bp
GS-448 with GSR-1069	920 bp	640 bp
GS-537 with GSR-977	680 bp	460 bp
GS-537 with GSR-1069	780 bp	780 bp

Table 4. Primers used to amplify PCR product for glutamine synthetase. Numbers associated with primers give positional information of the location of the primers relative to the published glutamine synthetase sequence for *Xenopus laevis* (Genbank accession number D50062). Approximate fragment size indicates the size of the fragment produced when the specific primer pair indicated is used for amplification.

bp = base pairs.

Name	Length	G	A	T	C	?
Cabezon_A	432	30.32	23.38	21.76	24.54	0
White_spotted_greenling_A	432	30.56	22.45	22.92	24.07	0
Mossy_sculpin_A	432	30.09	23.15	23.61	23.15	0
Tubesnout_A	432	31.02	22.92	21.99	24.07	0
Three_spine_stickleback_A	432	32.18	21.76	20.83	25.23	0
High_cockscomb_A	432	30.56	23.38	21.76	24.31	0
Penpoint_Gunnel_A	432	30.56	23.15	22.92	23.38	0
Crescent_gunnel_A	432	30.32	23.38	22.69	23.61	0
Starry_flounder_A	432	28.47	24.07	23.61	23.84	0
Buttersole_A	432	28.70	24.54	24.07	22.69	0
Pacific_sanddab_A	432	29.17	24.07	19.91	26.85	0
Speckled_sanddab_A	432	29.40	23.84	20.37	26.39	0
Shiner_perch_A1	432	28.94	24.54	22.92	23.61	0
Shiner_perch_A2	432	29.40	23.38	24.07	23.15	0
Buffalo_sculpin_A	432	29.63	23.84	22.69	23.84	0
Striped_seaperch_A	432	28.94	24.54	23.15	23.38	0
Tidepool_sculpin_A	432	30.56	22.92	22.92	23.61	0
White_spotted_greenling_B	432	30.79	22.45	21.76	25.00	0
Three_spine_stickleback_B	432	31.71	21.53	19.68	27.08	0
Tubesnout_B	432	31.48	22.69	20.14	25.69	0
Mossy_sculpin_B	432	31.59	21.89	20.65	25.87	30
Cabezon_B	432	30.98	22.20	21.46	25.37	22
High_cockscomb_B	432	32.41	21.30	20.37	25.93	0
Penpoint_gunnel_B	432	32.18	21.06	20.60	26.16	0
Crescent_gunnel_B	432	31.94	21.06	20.37	26.62	0
Starry_flounder_C	432	31.48	23.15	22.45	22.92	0
Buttersole_C	432	31.46	23.17	21.95	23.41	22
Speckled_sanddab_C	432	30.79	21.99	21.53	25.69	0
Pacific_sanddab_C	432	30.79	21.99	21.76	25.46	0
Coho_salmon_D1	432	30.56	23.15	22.45	23.84	0
Coho_salmon_D2	432	30.09	23.38	22.92	23.61	0
Chum_salmon_D	432	30.09	22.69	22.92	24.31	0
Bay_pipefish_E	432	30.79	22.92	17.82	28.47	0
Alligator_gar_F	432	30.56	22.92	19.68	26.85	0
Hagfish	432	29.86	21.06	27.55	21.53	0

Table 5. Base composition percentage statistics for amplified glutamine synthetase fragments. Fragments correspond to 432 bp fragment positioned from 467 to 899 of *Xenopus laevis* published glutamine synthetase sequence (Genbank accession number D50062). Letters after fish name indicate putative isoform designation.

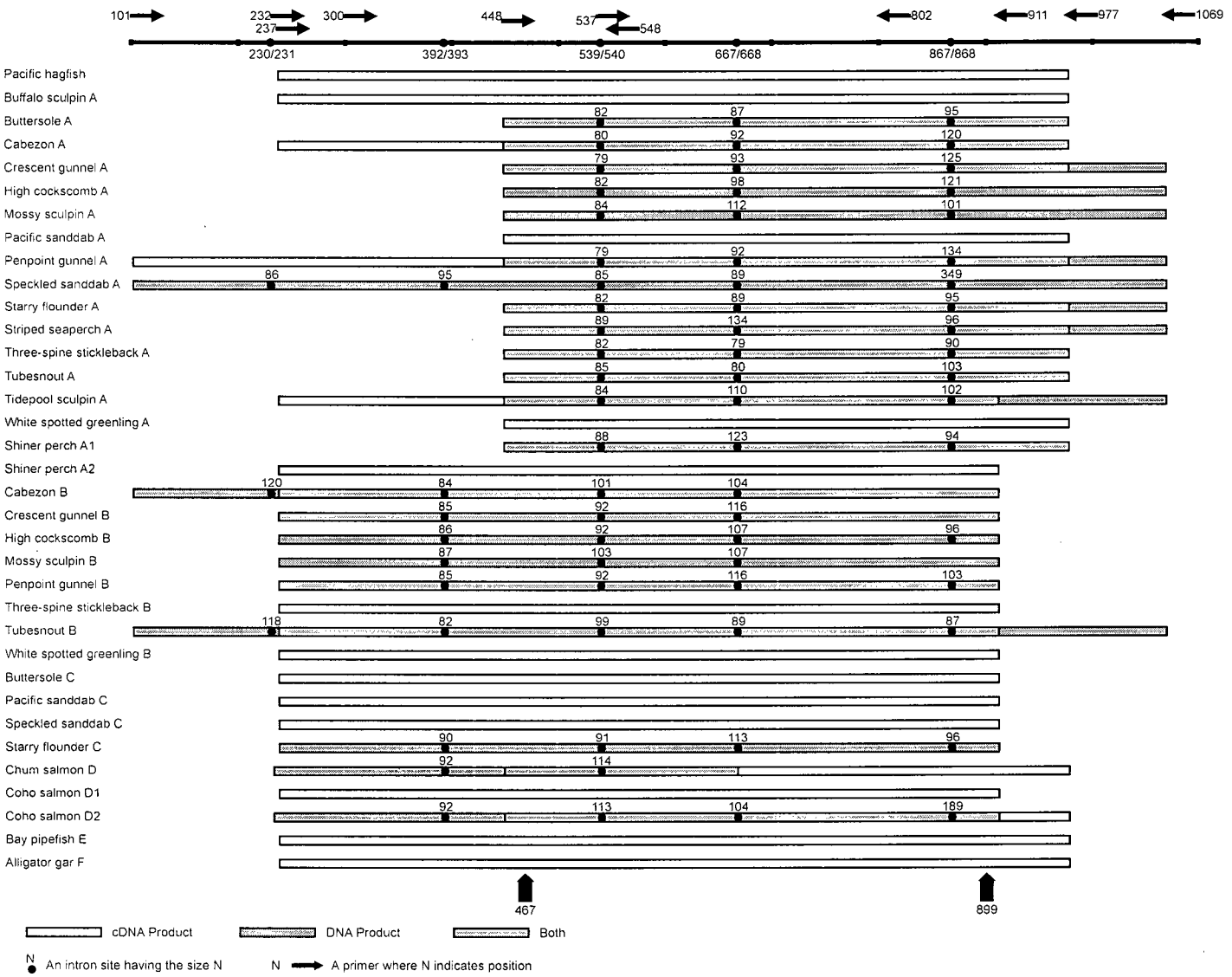


Figure 1: Summary of DNA fragments of glutamine synthetase amplified by PCR for fishes used in this study. Top line gives positional information relative to published sequence of *Xenopus laevis* glutamine synthetase gene (Genbank Accession number D50062) from position 100bp to 1100bp. Wide arrows at 467 and 899 indicate region of sequence used for phylogenetic analysis performed in this study.

Gene Copy	Sm A	Hs A	Ce A	Auf A	Gaa A	Ap A	Apf A	Pl A	Ca A1	Ps A	Ii A	Ca A2	Cso A	Cst A	Hs B	Gaa B	Auf B	Ce B	Sm B	Ap B	Apf B	Pl B	Ps C	Ii C	Cst C	Cso C	Oki D1	Oki D2
Sm A	-	**4**	**7**	**8**	**9**	**8**	**9**	**9**	**10**	**11**	**11**	**12**	**13**	**13**														
HsA	4	-	**7**	**8**	**10**	**7**	**8**	**8**	**10**	**10**	**10**	**10**	**12**	**12**														
Ce A	7	7	-	**10**	**12**	**9**	**10**	**10**	**12**	**11**	**11**	**12**	**15**	**15**														
Auf A	8	8	10	-	**7**	**9**	**9**	**9**	**13**	**10**	**11**	**13**	**13**	**12**														
Gaa A	9	10	12	7	-	**10**	**10**	**10**	**13**	**11**	**12**	**13**	**14**	**13**														
Ap A	8	7	9	9	10	-	**3**	**3**	**9**	**10**	**10**	**10**	**13**	**12**														
Apf A	9	8	10	9	10	3	-	**1**	**10**	**10**	**10**	**11**	**13**	**12**														
Pl A	9	8	10	9	10	3	1	-	**10**	**10**	**10**	**11**	**13**	**12**														
Ca A1	10	10	12	13	13	9	10	10	-	**10**	**10**	**8**	**13**	**13**														
Ps A	11	10	11	10	11	10	10	10	10	-	**3**	**5**	**14**	**14**														
Ii A	11	10	11	11	12	10	10	10	10	3	-	**3**	**14**	**14**														
Ca A2	12	10	12	13	13	10	11	11	8	5	3	-	**14**	**13**														
Cso A	13	12	15	13	14	13	13	13	13	14	14	14	-	**1**														
Cst A	13	12	15	12	13	12	12	12	13	14	14	13	1	-														
Hs B	16	16	18	19	17	15	16	16	17	16	15	15	18	17	-	**7**	**6**	**5**	**4**	**5**	**6**	**7**						
Gaa B	18	19	19	20	18	16	19	19	19	19	18	18	17	17	7	-	**4**	**8**	**8**	**6**	**6**	**7**						
Auf B	18	19	20	20	18	17	19	19	18	18	18	18	19	18	6	4	-	**6**	**7**	**6**	**6**	**7**						
Ce B	19	19	20	21	18	17	18	18	18	17	18	17	19	18	5	8	6	-	**3**	**4**	**5**	**5**						
Sm B	18	19	20	21	19	18	20	20	20	18	18	18	20	19	4	8	7	3	-	**4**	**4**	**5**						
Ap B	18	18	19	20	18	17	18	18	18	17	18	17	20	19	5	6	6	4	4	-	**1**	**2**						
Apf B	18	18	18	20	18	17	18	18	17	17	18	17	19	19	6	6	6	5	4	1	-	**1**						
Pl B	18	19	19	21	18	17	18	18	18	18	18	17	20	20	7	7	7	5	5	2	1	-						
Ps C	21	20	22	22	21	19	20	20	18	19	20	19	22	21	13	15	13	10	11	11	12	12	-	**2**	**15**	**15**		
Ii C	22	21	23	23	22	20	21	21	19	20	21	20	22	21	14	16	14	10	12	13	13	13	2	-	**15**	**15**		
Cst C	21	21	22	22	21	20	21	21	21	22	23	23	22	22	15	16	16	15	15	15	15	16	15	15	-	**1**		
Cso C	21	22	22	22	22	21	21	21	21	22	23	23	22	22	15	17	17	14	15	15	15	16	15	15	1	-		
Oki D1	18	18	18	19	19	18	17	18	16	17	17	17	17	17	17	18	19	18	19	18	18	18	19	19	19	19	-	**13**
Oki D2	18	18	18	18	19	16	17	17	15	14	15	15	18	18	14	16	15	14	16	16	15	16	15	15	18	18	13	-

Figure 2 Pairwise comparisons for % differences in GS isoforms for fish used in this study that possessed two GS isoforms. Species are designated by initials for species name and isoform designation is indicated by A, B, C or D. Ap – high cockscomb, Apf – penpoint gunnel, Auf – tubesnout, Ca – shiner perch, Ce – mossy sculpin, Cso – Pacific sanddab, Cst – speckled sanddab, Gaa – three spine stickleback, Hs – white spotted greenling, Ii – buttersole, Oki – coho salmon, Pl – crescent gunnel, Ps – starry flounder, Sm – cabezon. For clarity each isoform grouping is shown above the diagonal in bold.

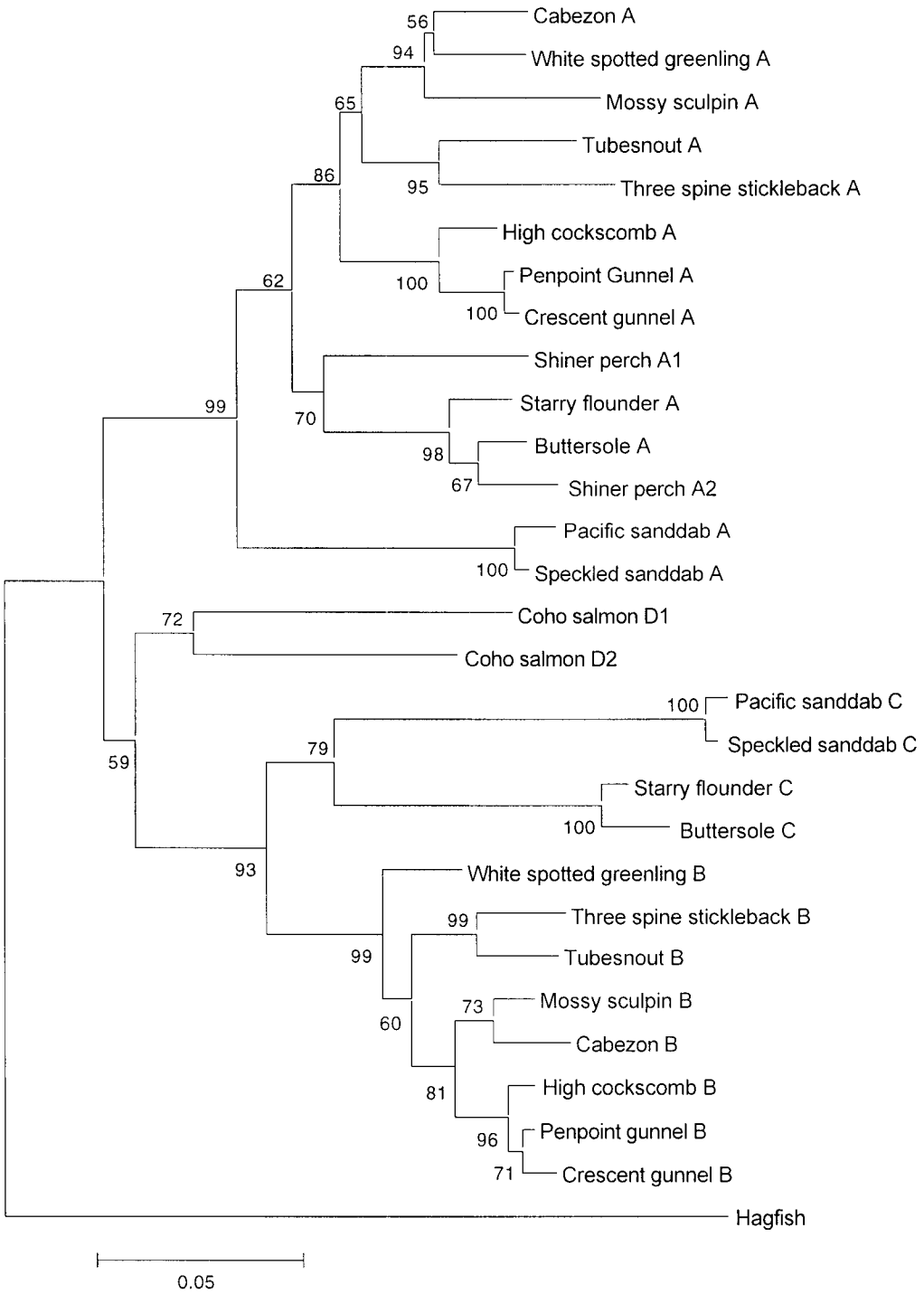


Figure 3 Neighbor joining tree constructed for all fish that produced two isoforms of glutamine synthetase for this study. Tree is based on Kimura-2-parameter distance constructed with unweighted 432 bp fragment of glutamine synthetase product, using CNI branch swapping algorithm in MEGA version 2.1. Tree was bootstrapped 100 times.

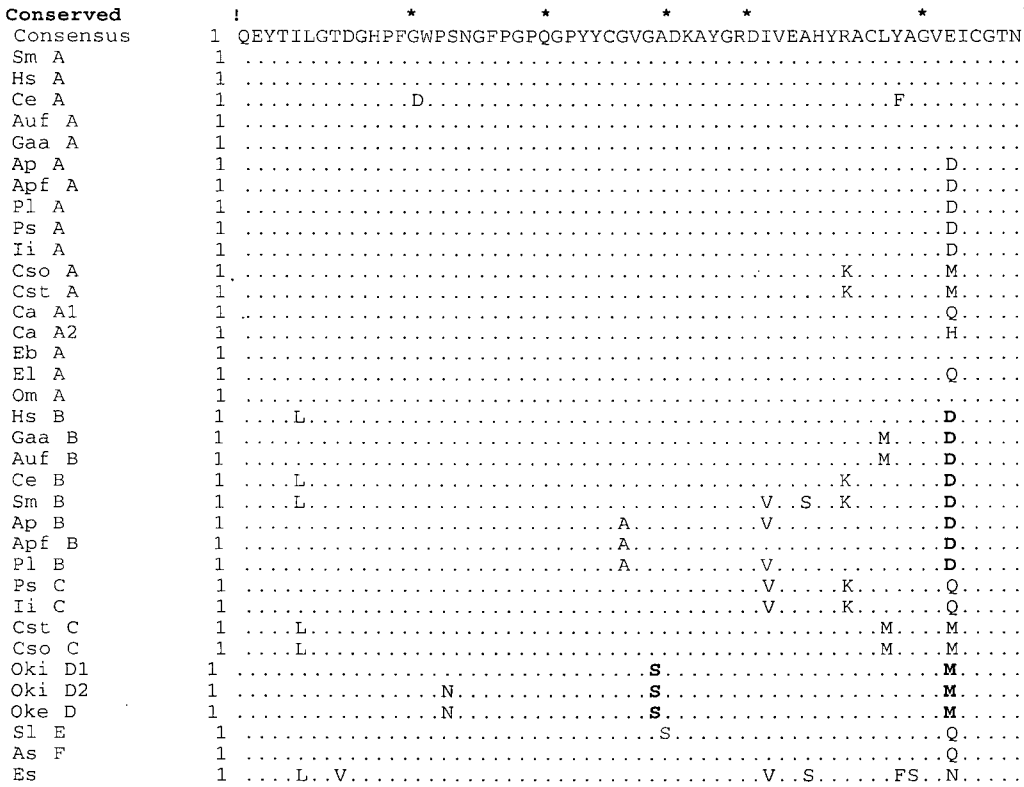


Figure 4. Amino acid translation of glutamine synthetase products for all isoforms of GS. * - indicate site conserved in all organisms for GS (Eisenberg et al. 2000). ! – indicate active site of GS conserved in all organisms for GS (Eisenberg et al. 2000). Numbers correspond to amino acid number 135 to 279 of the published *Xenopus laevis* sequence (Genbank accession number D50062). Amino acid translation sites that support isoform designations shown at positions 102 and 135 for isoform A; Sites 55, 102, and 135 for isoform B; Sites 86 and 102 for isoform C; Sites 33, 55, and 102 for isoform D.

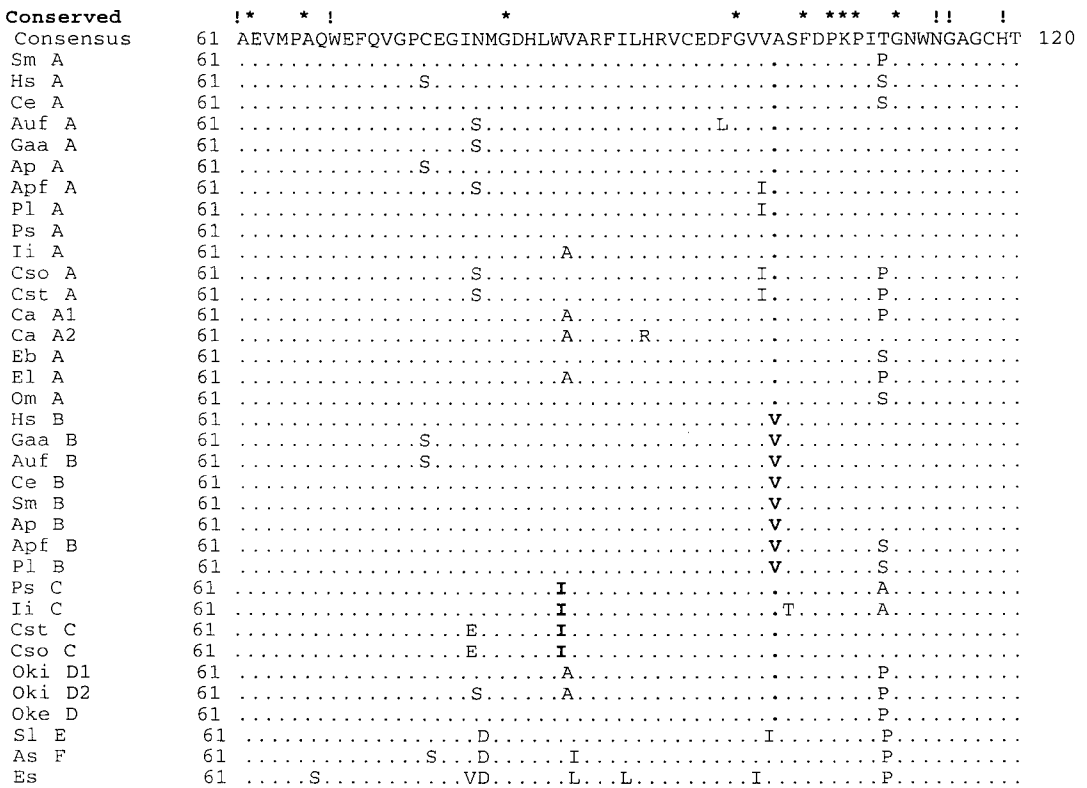


Figure 4. continued.

```
Conserved          *           *
 Consensus    121 NFSTKEMREDGGLKAIEESIEKLG 144
Sm A          121 ...............D......
Hs A          121 ...............D......
Ce A          121 ...............D......
Auf A         121 ...............D......
Gaa A         121 ...............D......
Ap A          121 ........................
Apf A         121 .V......................
Pl A          121 .V......................
Ps A          121 ........................
Ii A          121 .........E..............
Cso A         121 ...............D......
Cst A         121 ...............D......
Ca A1         121 ...............D......
Ca A2         121 .........E..........M..
Eb A          121 ...............D......
El A          121 ...............D......
Om A          121 ...............D......
Hs B          121 .........E....I..D.....E
Gaa B         121 .........E....I......R.P
Auf B         121 .........E....I......R.A
Ce B          121 ..............
Sm B          121 .......P......I.
Ap B          121 ..............I......R.A
Apf B         121 ..............I......R.A
Pl B          121 ..T...........I......R.A
Ps C          121 ..............V......R.A
Ii C          121 ..............L
Cst C         121 ....Y.........V......R.A
Cso C         121 ....Y.........V......R.A
Oki D1        121 ..............G..D......
Oki D2        121 .........E..........R..
Oke D         121 .........E..........R..
Sl E          121 ...............D......
As F          121 .........EN...Y......R.S
Es            121 ...SLA..QA...QH..YA....A
```

Figure 4. continued.

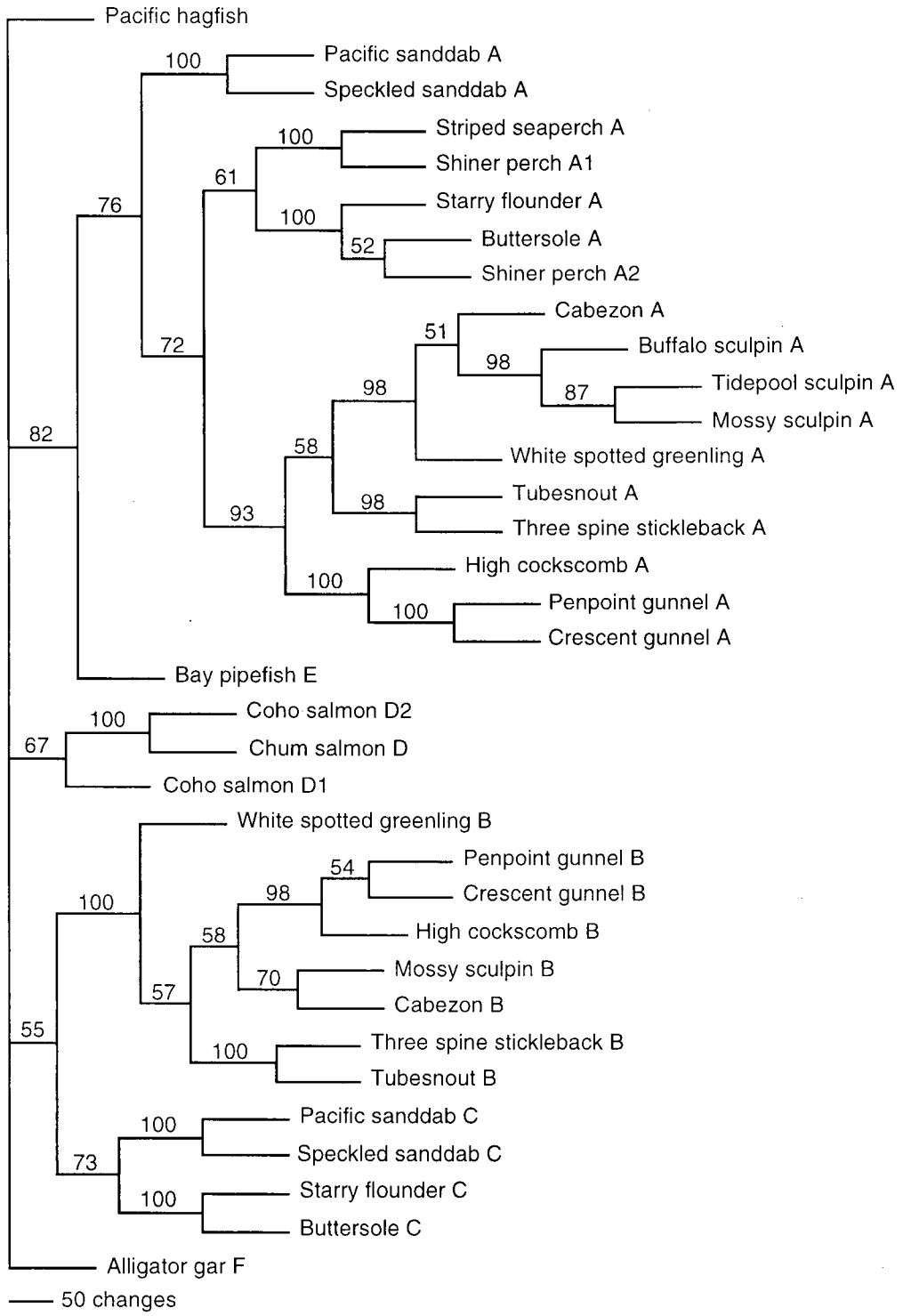


Figure 5. Neighbor joining tree based on Kimura-2-parameter distances constructed from 432 bp fragment of glutamine synthetase from all isoforms amplified in all fish used in this study. I used TBR branch swapping algorithm and tree was bootstrapped 100 times.

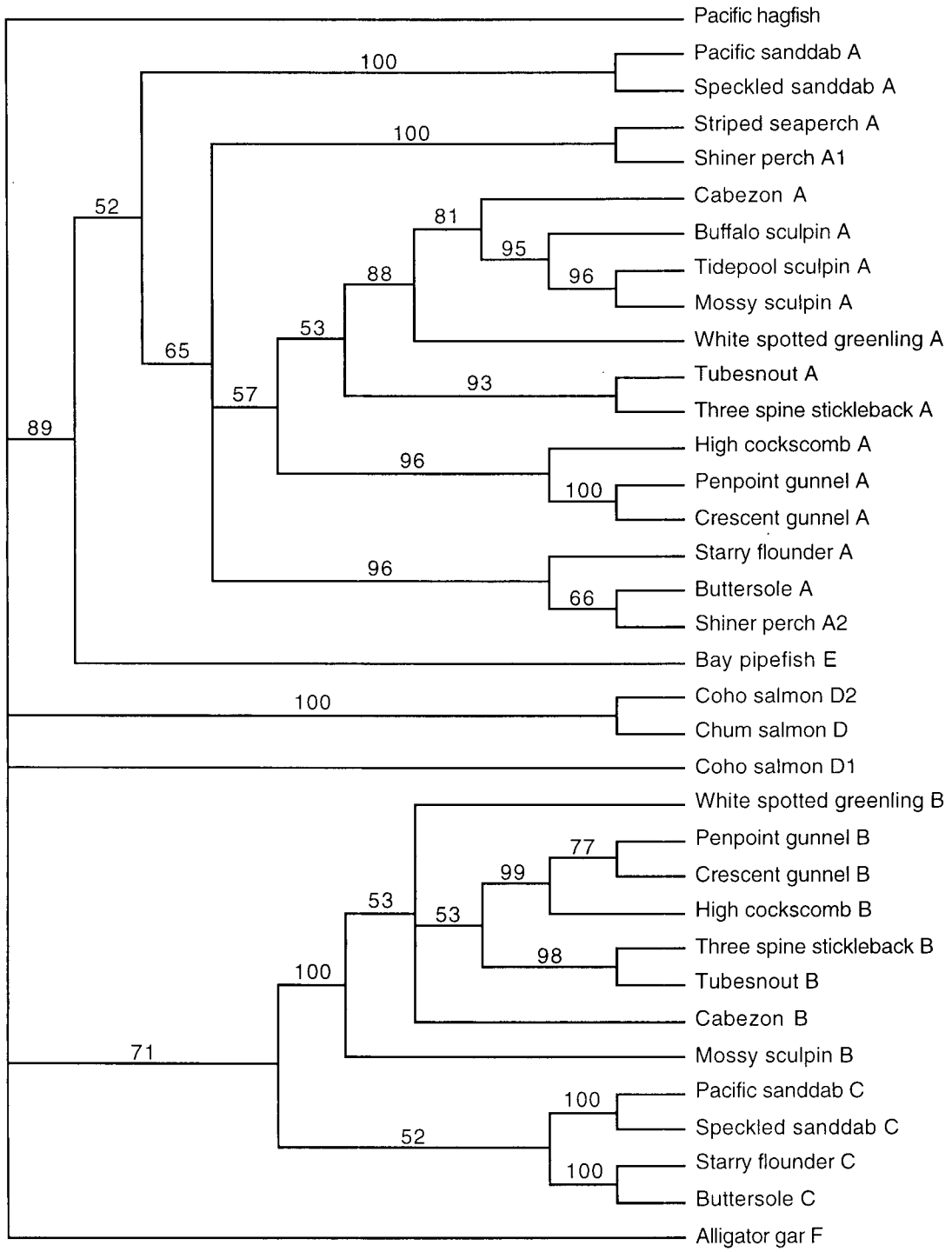


Figure 6. Maximum parsimony tree constructed from 432 bp fragment of glutamine synthetase from all isoforms amplified in all fish used in this study. Parsimony criterion was set to random addition, 50 replicates, using TBR branch swapping algorithm. Tree was bootstrapped 100 times.

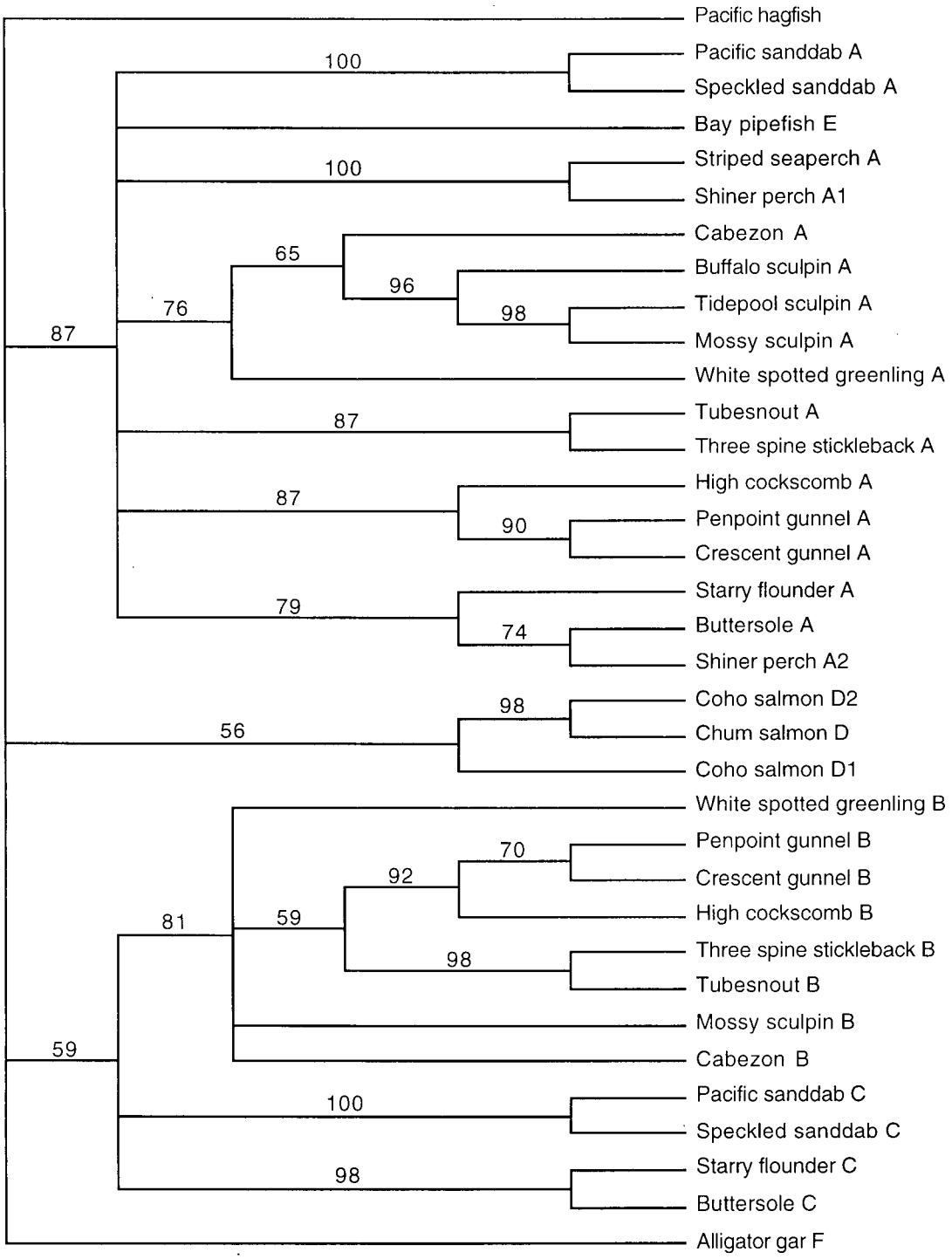


Figure 7. Maximum likelihood tree constructed from 432 bp fragments of all isoforms of glutamine synthetase for all fish used in this study. Likelihood criterion was set to random addition, 10 replicates, TBR branch swapping algorithm. I used HKY85 for the Likelihood model. Sequence data was weighted 2:4:1 by codon position and tree was bootstrapped 100 times.

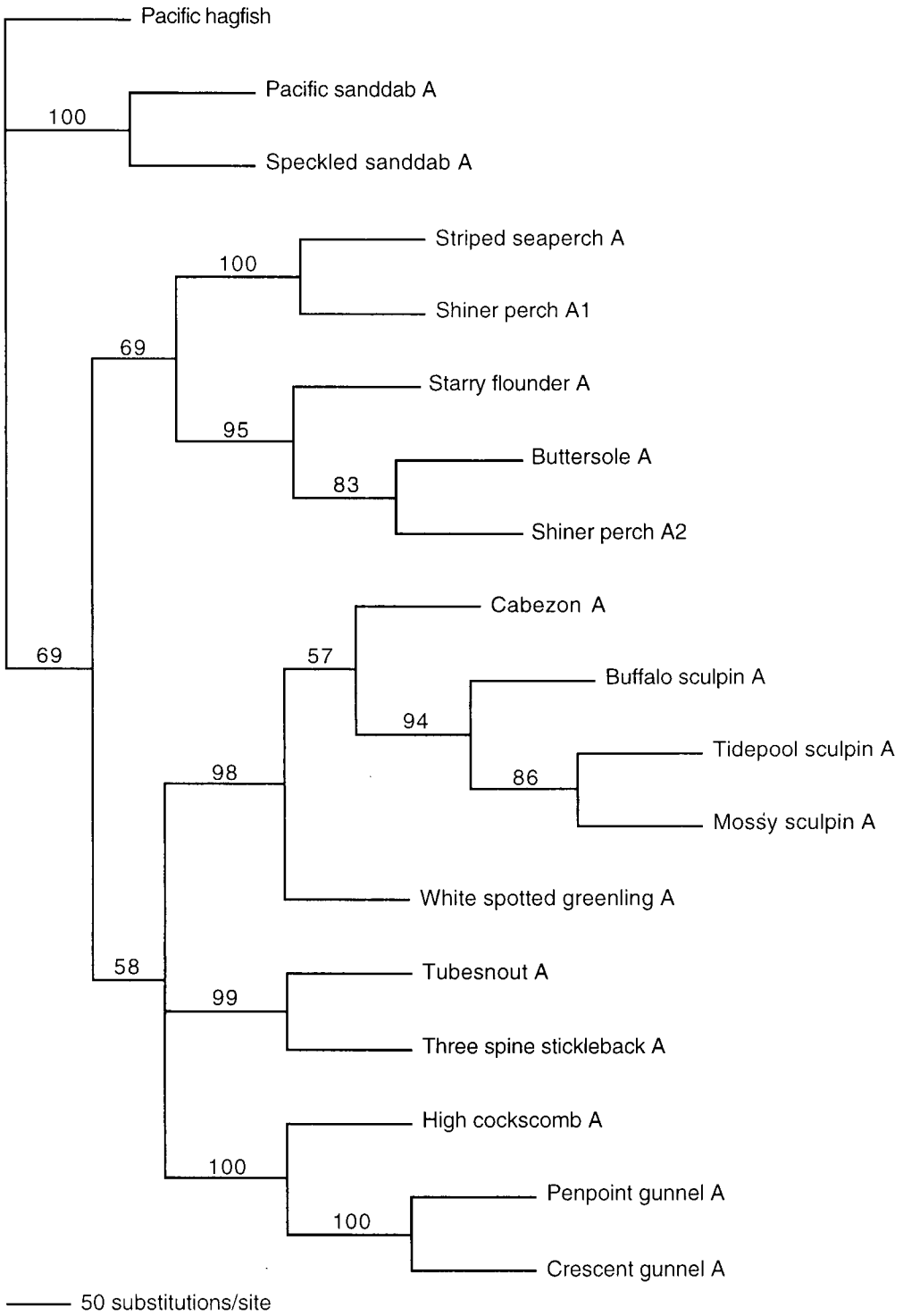


Figure 8. Neighbor joining analysis of isoforms A only of glutamine synthetase using Kimura-2- parameter distance. Sequence data was weighted 2:4:1 by codon position and tree was bootstrapped 100 times.

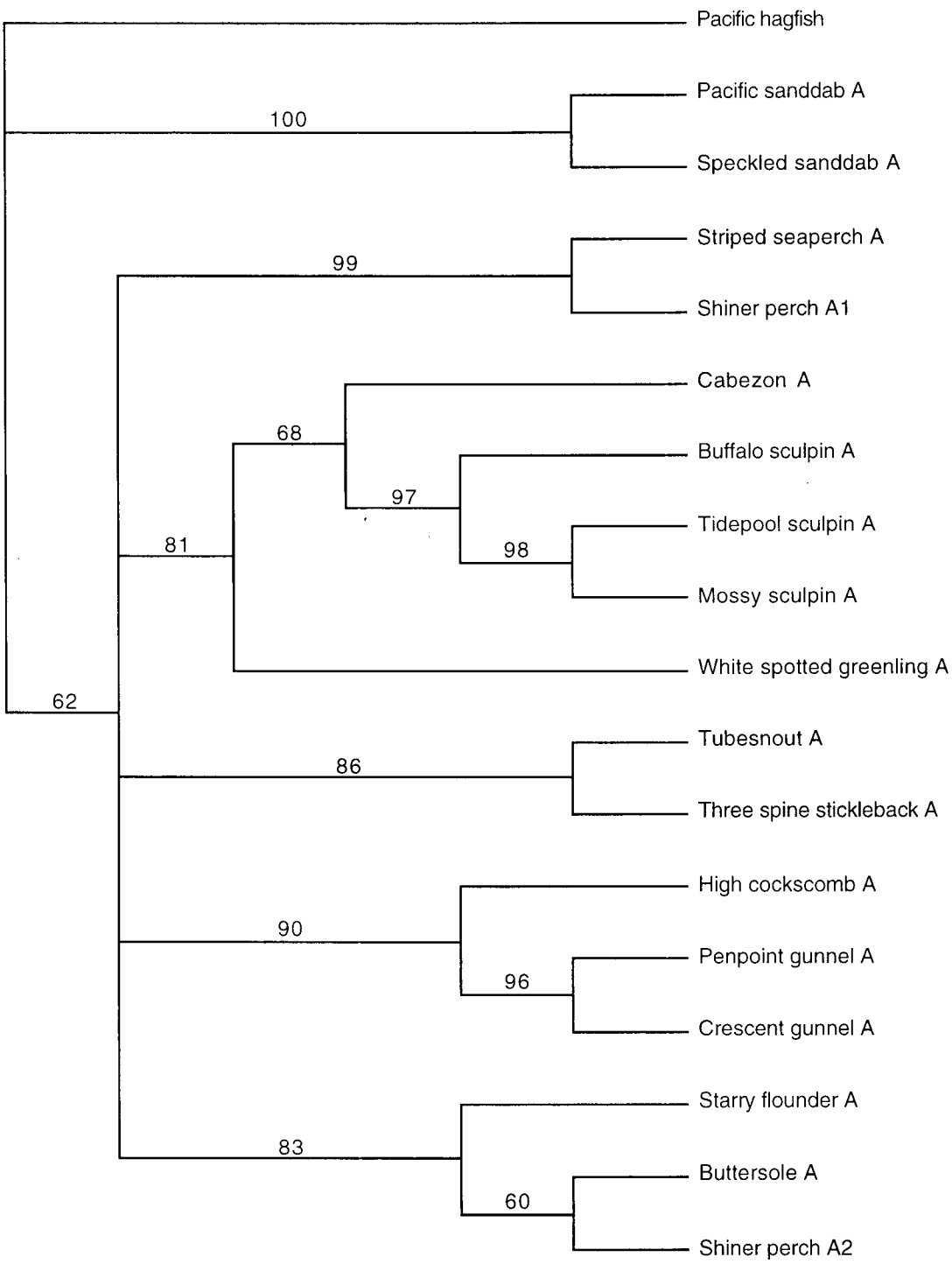


Figure 9. Maximum parsimony tree constructed from A isoforms only of glutamine synthetase. Parsimony criterion was set to random addition, 50 replicates, TBR branch swapping algorithm. Sequence data was weighted 2:4:1 by codon position and tree was bootstrapped 100 times.

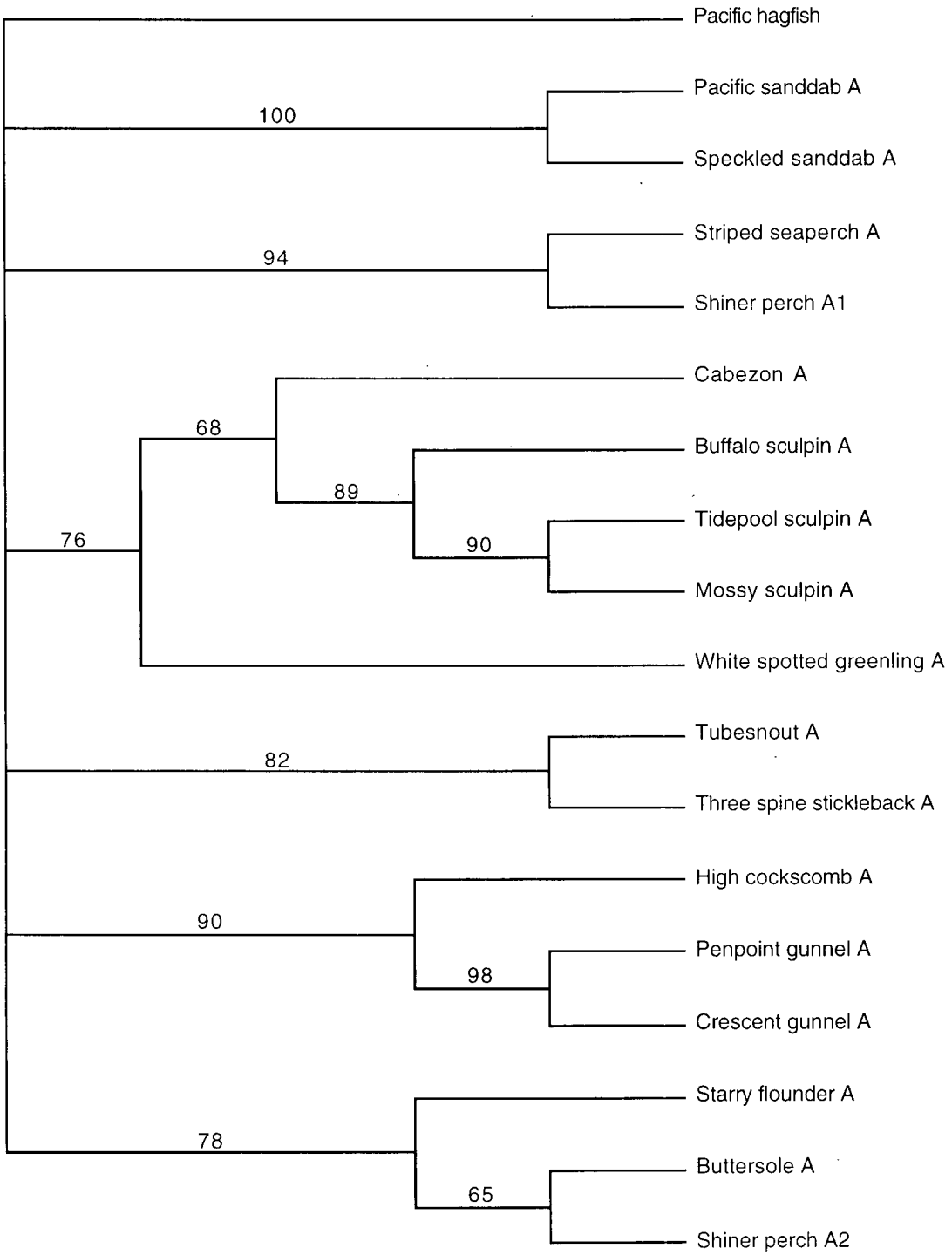


Figure 10. Maximum likelihood tree constructed A isoforms only of glutamine synthetase. Likelihood criterion set to random addition, 10 replicates, TBR branch swapping algorithm. I used HKY85 for the Likelihood model. Sequence data was weighted 2:4:1 by codon position and tree was bootstrapped 100 times.

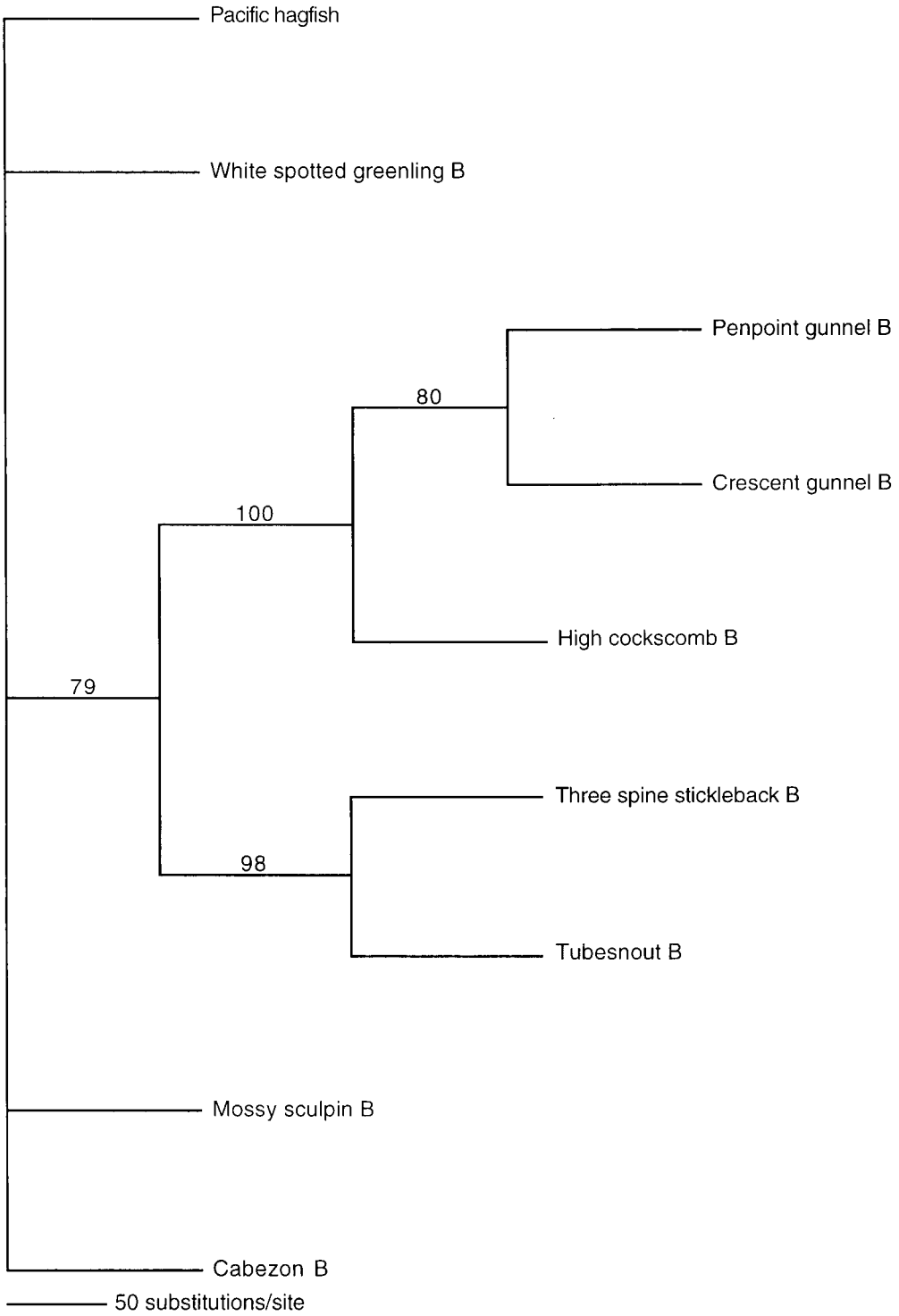


Figure 11. Neighbor joining tree constructed from B isoforms only of glutamine synthetase based on Kimura-2-parameter distances. I used TBR branch swapping algorithm. Sequence data was weighted 2:4:1 by codon position and tree was bootstrapped 100 times.

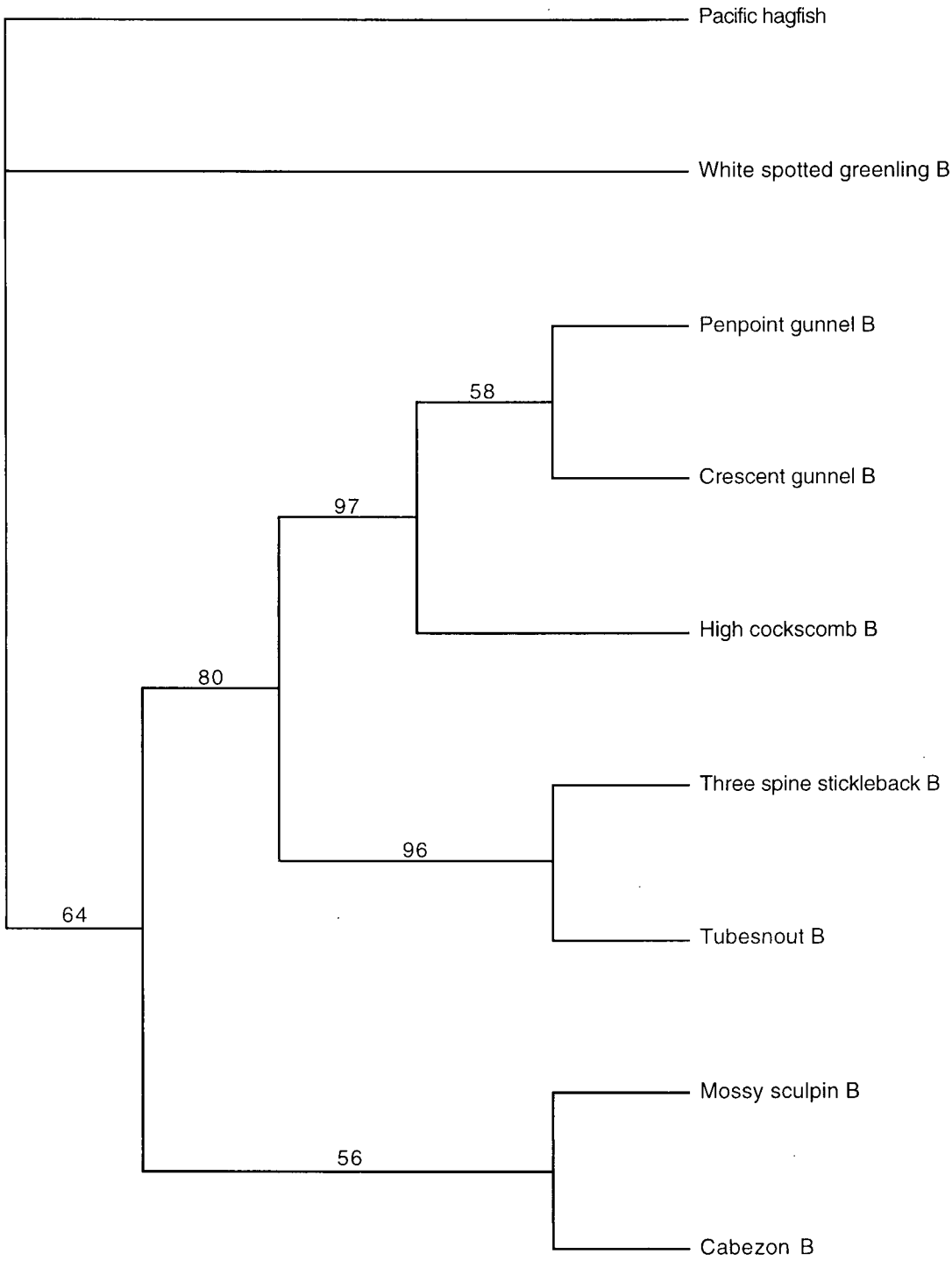


Figure 12. Maximum parsimony tree constructed from B isoforms only of glutamine synthetase. Parsimony criterion was set to random addition, 50 replicates, TBR branch swapping algorithm. Sequence data was weighted 2:4:1 by codon position and tree was bootstrapped 100 times.

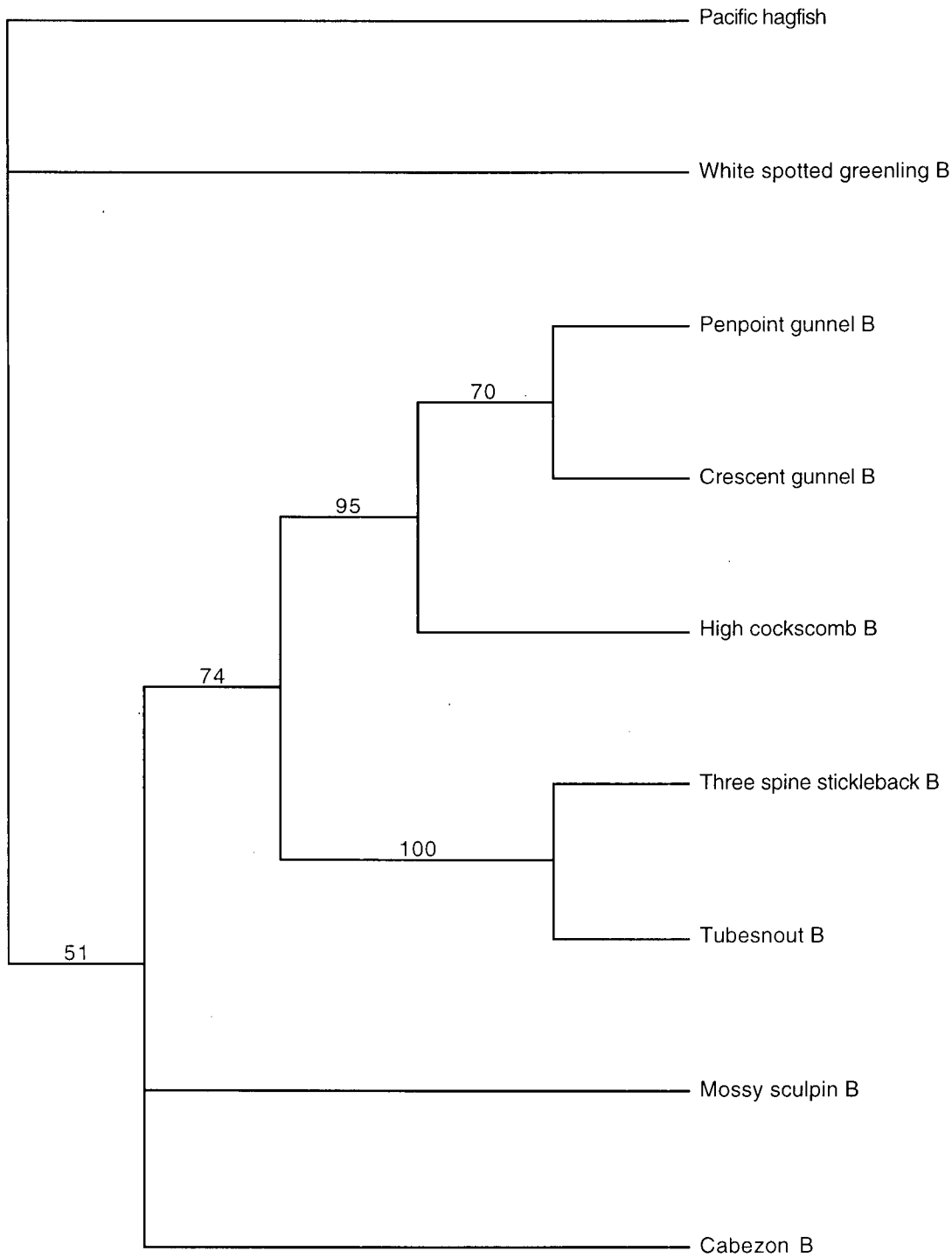


Figure 13. Maximum likelihood tree constructed B isoforms only of glutamine synthetase. Likelihood criterion was set to random addition, 10 replicates, TBR branch swapping algorithm. I used HKY85 for the Likelihood model. Sequence data was weighted 2:4:1 by codon position and tree was bootstrapped 100 times.

Fish	El	Ca 1	Ca 2	Cso	Cst	Eb	Om	Ii	Ps	Hs	Ce	Sm	Gaa	Auf	Apf	Pl	Ap	Es
El	-																	
Ca 1	1.2	-																
Ca 2	9.0	7.9	-															
Cso	12.7	13.4	13.9	-														
Cst	12.5	13.2	13.4	1.2	-													
Eb	10.6	10.9	12.5	13.7	13.4	-												
Om	11.1	11.3	12.7	14.6	14.4	2.3	-											
Ii	10.4	9.7	3.0	14.1	13.7	11.1	11.3	-										
Ps	10.4	9.7	4.4	13.9	13.4	11.1	11.3	2.5	-									
Hs	10.6	10.4	11.1	12.7	12.5	5.1	5.3	10.2	10.0	-	5.0	4.2	7.0	6.2	5.5	6.5	5.2	26.1
Ce	11.6	11.8	12.7	15.0	14.6	3.7	2.1	11.3	11.3	6.9	-	2.7	7.5	6.2	4.5	5.0	4.2	25.4
Sm	10.0	10.2	12.0	12.5	12.7	4.4	4.9	11.3	11.6	4.9	6.0	-	7.7	6.5	4.2	4.7	3.5	25.6
Gaa	12.3	12.5	13.4	13.9	13.4	10.4	10.4	12.0	11.3	9.7	11.3	8.8	-	4.0	6.2	7.2	6.2	26.6
Auf	12.0	12.3	12.7	13.0	12.7	8.6	8.6	10.6	10.4	8.1	9.3	7.9	6.7	-	6.0	7.0	5.7	27.6
Apf	9.5	9.7	11.6	13.2	12.3	9.3	9.3	10.4	9.7	8.1	9.0	8.8	9.7	8.6	-	1.0	1.2	26.9
Pl	9.5	9.7	11.6	13.2	12.3	9.3	9.3	10.4	9.7	8.1	9.0	8.8	10.2	8.6	0.5	-	1.7	26.6
Ap	8.8	9.0	10.6	13.2	12.5	7.9	7.9	9.7	9.7	7.2	8.3	7.9	9.7	9.0	3.0	3.0	-	26.9
ES	23.8	24.5	27.1	25.7	25.5	25.7	26.6	26.4	25.2	26.2	26.4	25.9	25.2	26.0	25.7	25.7	26.2	-

Figure 14. Pairwise comparisons of % differences in GS sequences of isoforms A (below diagonal) and isoforms B (above diagonal). Isoform B was amplified in fewer fishes than isoform A and therefore only requires a small portion of the table compared to isoform A. Apf – penpoint gunnel, Ap – high cockscomb, Auf – tubesnout, Ca – shiner perch, Ce – mossy sculpin, Cso – Pacific sanddab, Cst – speckled sanddab, Eb – buffalo sculpin, El – striped seaperch, Es – Pacific hagfish, Gaa – three spine stickleback, Hs – white spotted greenling, Ii – buttersole, Om – tidepool sculpin, Pl – crescent gunnel, Ps – starry flounder, Sm – cabezon.

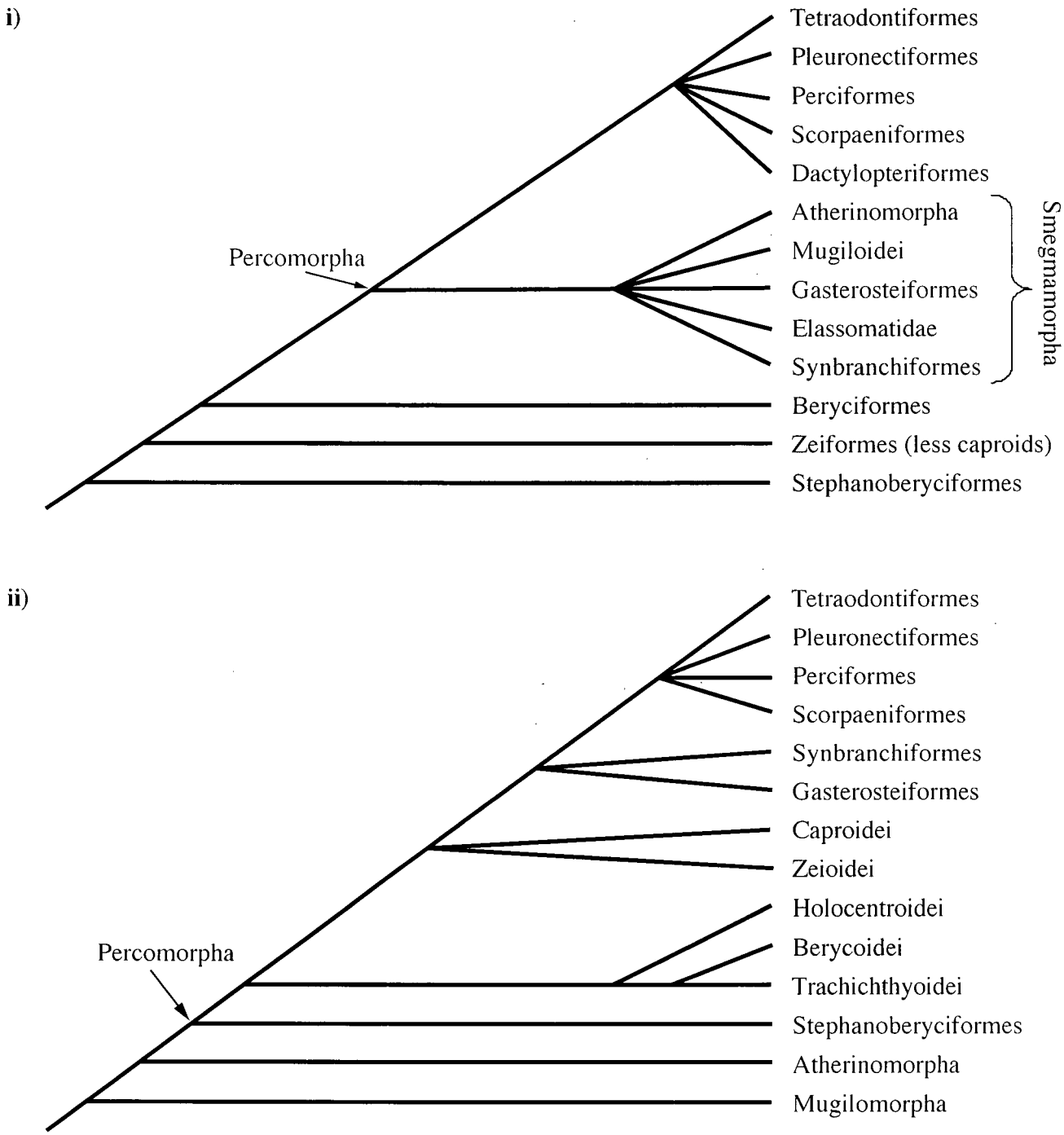


Figure 15. The two competing hypotheses for the resolution of the Superorder Acanthoptyerygii. i) Phylogenetic relationships of the Acanthoptyerygii as presented by Johnson and Patterson (1993). ii) Phylogenetic relationships of the Acanthoptyerygii as presented by Nelson (1994). A small arrow on each cladogram shows where the author(s) believe the series Percomorpha begins.

References

Allendorf, F. W. and G. H. Thorgaard (1984). Tetraploidy and the Evolution of Salmonid Fishes. Evolutionary genetics of fishes. B. J. Turner. New York, Plenum Press: 1-53.

Altschul, S. F., T. L. Madden, et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25**(17): 3389-402.

Cabot, E. L. and A. T. Beckenbach (1989). Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. Comput Appl Biosci **5**(3): 233-4.

Campbell, J. W. and D. D. Smith, Jr. (1992). Metabolic compartmentation of vertebrate glutamine synthetase: putative mitochondrial targeting signal in avian liver glutamine synthetase. Mol Biol Evol **9**(5): 787-805.

Chakrabarti, R., J. B. McCracken, Jr., et al. (1995). Detection of a functional promoter/enhancer in an intron-less human gene encoding a glutamine synthetase-like enzyme. Gene **153**(2): 163-99.

Chapleau, F. (1993). Pleuronectiform relationships: A cladistic reassessment. Bulletin of Marine Science **52**(1): 526-540.

Cullimore, J. V., C. Gebhardt, et al. (1984). Glutamine synthetase of Phaseolus vulgaris L.: organ-specific expression of a multigene family. J Mol Appl Genet **2**(6): 589-99.

Eisenberg, D., H. S. Gill, et al. (2000). Structure-function relationships of glutamine synthetases. Biochim Biophys Acta **1477**(1-2): 122-45.

Freeman, S. and J. C. Herron (2001). Evolutionary Analysis. Upper Saddle River, New Jersey, Prentice-Hall Inc.

Froese, R. and D. Pauly (2000). FishBase 2000: concepts, design and data sources. Los Baños, Laguna, Philippines. 344 p., ICLARM. **2000:** 344.

Goodman, H. J. and D. R. Woods (1993). Cloning and nucleotide sequence of the Butyrivibrio fibrisolvens gene encoding a type III glutamine synthetase. J Gen Microbiol **139**(Pt 7): 1487-93.

Hart, J. L. (1988). Pacific fishes of Canada. Ottawa, Canada, Canadian Government Publishing Centre.

Helfman, G. S., B. B. Collette, et al. (1997). The diversity of fishes. Malden, Mass., Blackwell Science.

Hill, R. T., J. R. Parker, et al. (1989). Molecular analysis of a novel glutamine synthetase of the anaerobe Bacteroides fragilis. J Gen Microbiol **135**(Pt 12): 3271-9.

Holland, P. W., J. Garcia-Fernandez, et al. (1994). Gene duplications and the origins of vertebrate development. Dev Suppl: 125-33.

Imamura, H. and G. Shinohara (1998). Scorpaeniform fish phylogeny: An overview. Bulletin of the National Science Museum, Tokyo, Series A (Zoology). **24**(3): 85-212.

Johnson, G. D. (1993). Percomorph phylogeny: Progress and Problems. Bulletin of Marine Science **52**((1)): 3-28.

Johnson, G. D. and C. Patterson (1993). Percomorph phylogeny: A survey of Acanthomorphs and a new proposal. Bulletin of Marine Science **52**((1)): 554-626.

Kumada, Y., D. R. Benson, et al. (1993). Evolution of the glutamine synthetase gene, one of the oldest existing and functioning genes. Proc Natl Acad Sci U S A **90**(7): 3009-13.

Kumar, S., K. Tamura, et al. (2001). MEGA2: molecular evolutionary genetics analysis software. Bioinformatics **17**(12): 1244-5.

Kuo, C. F. and J. E. J. Darnell (1989). Mouse glutamine synthetase is encoded by a single gene that can be expressed in a localized fashion. Journal of Molecular Biology **208**: 45-56.

Laud, P. and J. W. Campbell (1994). Genetic basis for tissue isozymes of glutamine synthetase in elasmobranchs. Journal of Molecular Evolution **39**: 93-100.
Li, M. G., R. Villemur, et al. (1993). Differential expression of six glutamine synthetase genes in Zea mays. Plant Mol Biol **23**(2): 401-7.
Li, W.-H. and D. Graur (1991). Fundamentals of molecular evolution. Sunderland, Mass., Sinauer Associates.
Meister, A., Ed. (1985). Glutamate, Glutamine, Glutathione, and Related Compounds. Methods in Enzymology. Orlando, Florida, Academic Press, Inc.
Meyer, A. and M. Schartl (1999). Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. Curr Opin Cell Biol **11**(6): 699-704.
Mommsen, T. P. and P. J. Walsh (1992). Biochemical and environmental perspectives on nitrogen metabolism in fishes. Experientia **48**: 583-593.
Mount, S. M. (1982). A catalogue of splice junction sequences. Nucleic Acids Res **10**(2): 459-72.
Murray, B. W., Busby,E., Mommsen,T. and Wright,P.A. (2002). Evolution of glutamine synthetase in vertebrates: Multiple glutamine synthetase genes expressed in rainbow (Oncorhynchus mykiss).Journal of Experimental Biology **In Review**
Nelson, J. S. (1994). Fishes of the world. New York, J. Wiley.
Ohno, S. (1970). Evolution by gene duplication. Berlin ; New York, Springer-Verlag.
Otto, S. P. and J. Whitton (2000). Polyploid incidence and evolution. Annual Review of Genetics **34**: 401-437.
Pesole, G., M. P. Bozzetti, et al. (1991). Glutamine synthetase gene evolution: a good molecular clock. Proc Natl Acad Sci U S A **88**(2): 522-6.
Pesole, G., C. Gissi, et al. (1995). Glutamine synthetase gene evolution in bacteria. Mol Biol Evol **12**(2): 189-97.
Proudfoot, N. J. and T. Maniatis (1980). The structure of a human alpha-globin pseudogene and its relationship to alpha-globin gene duplication. Cell **21**(2): 537-44.
Pu, H. F. and A. P. Young (1989). The structure of the chicken glutamine synthetase-encoding gene. Gene **81**(1): 169-75.
Rychlik, W. and R. E. Rhoads (1989). A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA. Nucleic Acids Res **17**(21): 8543-51.
Saccone, C., C. Gissi, et al. (1995). Molecular classification of living organisms. Journal of Molecular Evolution **40**: 273-279.
Schultz, R. J. (1980). Role of polyploidy in the evolution of fishes. Polyploidy, Biological Relevance. W. H. Lewis. New York, Plenum Press. **13:** 313-340.
Swofford, D. L. (2002). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Sunderland, Massachusetts., Sinauer Associates.
Tateno, Y. (1994). Evolution of glutamine synthetase genes is in accordance with the neutral theory of molecular evolution. Jpn J Genet **69**(5): 489-502.
Taylor, J. S., Y. Van de Peer, et al. (2001). Comparative genomics provides evidence for an ancient genome duplication event in fish. Philos Trans R Soc Lond B Biol Sci **356**(1414): 1661-79.
Temple, S. J., J. Heard, et al. (1995). Characterization of a nodule-enhanced glutamine synthetase from alfalfa: nucleotide sequence, in situ localization, and transcript analysis. Mol Plant Microbe Interact **8**(2): 218-27.
Thompson, J. D., D. G. Higgins, et al. (1994). CLUSTAL W: improving the sensitivity of progressivemultiple sequence alignment through sequence weighting,position-specific gap penalties and weight matrix choice. Nucleic Acids Research **22**: 4673-4680.

Tingey, S. V., E. L. Walker, et al. (1987). Glutamine synthetase genes of pea encode distinct polypeptides which are differentially expressed in leaves, roots and nodules. Embo J **6**(1): 1-9.

Twyman, R. M. (1998). Advanced Molecular Biology: A Concise Reference. New York, Springer-Verlag New York Inc.

Walsh, P. J., G. D. Mayer, et al. (2002). A Second Glutamine Synthetase Gene with Expression in the Gills of the Gulf Toadfish (Opsanus beta). Journal of Experimental Biology **In Review**.

Wittbrodt, J., A. Meyer, et al. (1998). More genes in fish? BioEssays **20**: 511-515.

```
GLFTDFISH      1 AGTCAGGGTTGGAAGACTCTTTCTACAACTTCCTCGTTTGTCCAGGTGTGTTACAGCAGC  60
XENOPUS        1 AACAAACTTGCTGAGGAAACGAAGATCGGGGCTTCTTCCCTTGCTCATAAGATCAGCTGC  60
DOGFISH        1 GTTTCTGGTAAAAAAATGCGGCAATATCACGCCGACGATTTGGCGGAATCAACATACTTA  60
CATSHARK       0
HORNSHARK      0
CHICKEN        1 GGAGCCCCGCGCCGAGCCCCTGCCCGCAGCCCAGCCCAGGACAGCCCTCGCCAGCTCCGC  60
CATFISH        0
MOUSE          1 TCGTTCTCGTGACCTGTTCACCCATCCATCATCCAGCTGGCCACTGTTCTGAACACCTTC  60
PIG            1 CCGATTCTCGCTCTCGCGGCCTGCCCGCCCTGCCTCCTGCTCGCCGCCCAGAACACCGTC  60
COW            0
RAT            0                     GTCCACCCATCCATCATCCTGCCGGCCACCGCTCTGAACACCTTC  45
LT-HAMSTER     0    GCTCGTGGCCCTGTCCACCCCGTCCATCATCCCGCCGGCCACCGCTCAGAGCACCTTC  58
C_HAMSTER      0
HUMAN          0                                GCTTTACCCGCCCGCCTGCTCGGCGACCAGAACACCTTC  39
ZEBRAFISH      1 AGCCAGTGTTATAGTGACTATTTGTACTTTTTGATAGGTTTGATATAGCATCACTAGCAC  60
MUDSUCKER      0                                      TTCGTCCATCCAGGTGTGTTATAGTAGCGGT  31

GLFTDFISH     61 AAAAATGGCCACATCAGCCAGCGCCAGCTTGAGTAAAGCTGTCAAGAAGCATTACATGGG  120
XENOPUS       61 CGCCATGTCAGTCTCCCACAGTTCGAGACTCAACAAGGGGGTGAGGGAACAGTACATGAA  120
DOGFISH       61 CAAGATGGCCACGTCAGCCAGCGCCAATCTCAGCAAAATCGTCAAGAAGAATTACATGGA  120
CATSHARK      0
HORNSHARK     0
CHICKEN       61 AGCCATGGCCACCTCGGCGAGCTCCCACCTGAGCAAAGCCATCAAGCACATGTACATGAA  120
CATFISH       0
MOUSE         61 CACCATGGCTACCTCAGCAAGTTCCCACGTGAACAAAGGCATCAAGCAAATGTACATGTC  120
PIG           61 CACCATGGCCACCTCAGCGAGCTCCCACTTGAACAAAGGCATCAAGCAGGTGTACATGTC  120
COW           0
RAT           46 CACCATGGCCACCTCAGCAAGTTCCCACTTGAACAAAGGCATCAAGCAGATGTACATGAA  105
LT-HAMSTER    59 CACCATGGCCACCTCAGCAAGTTCCCACTTGAACAAAAACATCAAGCAAATGTACTTGTG  118
C_HAMSTER      0      ATGGCCACCTCAGCAAGTTCCCACTTGAACAAAGGCATCAAGCAAATGTACATGTC  56
HUMAN         40 CACCATGACCACCTCAGCAAGTTCCCACTTAAATAAAGGCATCAAGCAGGTGTACATGTC  99
ZEBRAFISH     61 CACCATGGCCACTTCAGCCAGTTCCAAATTAAGCAAAGCTATGAAACAGCAGTACATGGA  120
MUDSUCKER     32 GAGGATGGCCACGTCCTCCAGCTCCCATTTGAGCAAAGCTGTGAAGCAGCAGTACATGGA  91

GLFTDFISH    121 ACTCCCTCAGGGGGATAAAGTCCAAGCTATGTACATTTGGATTGATGGAACAGGGGAGGG  180
XENOPUS      121 ACTGCCCCAAGGAGAAAAGGTCCAGGTCACCTACGTGTGGATCGACGGCACCGGGGAAGG  180
DOGFISH      121 GCTGCCCCAAGATGGCAAGGTGCAAGCGATGTACATCTGGATAGACGGCACAGGGGAGGC  180
CATSHARK       0
HORNSHARK      0       CCCCAAGATGGCAAGGTGCAAGCTATGTATATCTGGATCGATGGCACAGGAGAGGC  56
CHICKEN      121 GCTGCCGCAGGGTGAGAAGGTCCAAGCCATGTACATCTGGATCGACGGGACTGGGGAGCA  180
CATFISH        0
MOUSE        121 CCTGCCCCAGGGTGAGAAAATCCAAGCCATGTATATCTGGGTTGATGGTACCGGAGAACC  180
PIG          121 CCTGCCCCAGGGCGAGAAAGTCCAAGCTATGTACATCTGGATTGACGGTACGGGAGAGGG  180
COW            0
RAT          106 CCTGCCCCAGGGCGAGAAGATCCAACTCATGTATATCTGGGTTGATGGTACCGGGGAAGG  165
LT-HAMSTER   119 CCTGCCCCAGGGTGAGAAAGTCCAAGCCATGTATATCTGGGTTGATGGTACTGGAGAAGG  178
C_HAMSTER     57 CCTGCCCCAGGGTGAGAAAGTCCAAGCCATGTATATCTGGGTTGATGGTACCGGAGAAGG  116
HUMAN        100 CCTGCCTCAGGGTGAGAAAGTCCAGGCCATGTATATCTGGATCGATGGTACTGGAGAAGG  159
ZEBRAFISH    121 TCTCCCTCAGGGAGAAAAAGTTCAGGTCATGTACATCTGGATTGTNGGATCCGTAGAGGG  180
MUDSUCKER     92 GCTGCCTCAGGGAGATCTGGTGCAGGCTATGTACATCTGGATCGACGGCACTGGAGAGGG  151
```

Appendix 1. Sequence alignment of all vertebrate sequences present in Genbank prior to 1999. Glftdfish = *Opsanus beta*, Xenopus = *Xenopus laevis*, dogfish = *Squalus acanthias*, catshark = *Scyliorhinus torazame*, hornshark = *Heterodontus francisci*, chicken = *Gallus gallus*, catfish = *Ictalurus punctatus*, mouse = *Mus musculus*, pig = *Sus scrofa*, cow = *Bos taurus*, rat = *Rattus norvegicus*, Lt-hamster = *Cricetulus longicaudatus*, C-hamster = *Cricetulus griseus*, human = *Homo sapiens*, zebrafish = *Danio rerio*, mudsucker = *Gillichthys mirabilis*. Position 1 corresponds to position 1 of *Xenopus laevis* published glutamine synthetase sequence (Genbank accession number D50062).

```
GLFTDFISH    181 ACTCAGATGTAAAACCAGA---ACGCTGGATTCTGAACCCAAAAGCATTGAAGATCTTCC 237
 XENOPUS     181 AGTGAGGTGCAAAACCAGG---ACTCTGGATCAGGAACCCAAAACCATAGATGAAATCCC 237
 DOGFISH     181 CGTCCGCTGCAAGACCAGA---ACCTTGGACAATGAGCCCAAGAGCATTGCCGAACTCCC 237
CATSHARK      0
 HORNSHARK    57 AGTCCGCTGTAAAACCAAA---ACCTTGGACAAGGAGCCCAAGAACATTACTGACCTCCC 113
 CHICKEN     181 CCTCCGCTGCAAAACCCGC---ACTCTGGACCACGAACCCAAGAGCCTGGAAGATCTCCC 237
CATFISH       0
 MOUSE       181 ACTGCGCTGCAAGACCTGTCGTACCCTGGACTGTGAGCCCAAGTGTGTGGAAGAGTTACC 240
 PIG         181 ACTGCGCTGCAAGACCCGG---ACCCTGGATTCTGAGCCCAAGTGTATAGAAGAGTTGCC 237
COW           0
 RAT         166 GCTACGCTGCAAGACCCGT---ACTCTGGACTGTGACCCCAAGTGTGTAGAAGAGTTACC 222
 LT-HAMSTER  179 ACTGCGCTGCAAAACCCGC---ACCCTGGACTGTGAGCCCAAGTGTGTAGAAGAGTTACC 235
 C_HAMSTER   117 ACTGCGCTGCAAAACCCGC---ACCCTGGACTGTGAGCCCAAGTGTGTAGAAGAGTTACC 173
 HUMAN       160 ACTGCGCTGCAAGACCCGG---ACCCTGGACAGTGAGCCCAAGTGTGTGGAAGAGTTGCC 216
 ZEBRAFISH   181 ATTGAGATGCAAAACCAGG---ACTCTAGACTCTGAACCTAAATCTGTTGAAGAACTTNC 237
 MUDSUCKER   152 GCTGCGCTGCAAAACCAGG---ACACTAGACTCTGAACCCAAAAGCATTGAAGATCTGCC 208

 GLFTDFISH   238 GGAATGGAACTTTGACGGTTCCAGCACGTACCAGGCTGAGGGCTCCAACAGCGACATGTA 297
 XENOPUS     238 TGAATGGAACTTCGATGGATCCAGTACTCACCAAGCAGAAGGCTCAAACAGTGACATGTA 297
 DOGFISH     238 AGAATGGAACTTCGATGGCTCAAGTACGTATCAGTCAGAGGGGTCCAACAGCGACATGTA 297
CATSHARK      0
 HORNSHARK   114 AGAATGGAACTTTGATGGCTCAAGTACATATCAGTCAGAGGGGTCCAACAGCGACATGTA 173
 CHICKEN     238 CGAGTGGAACTTTGATGGCTCCAGCACCTTCCAAGCCGAAGGCTCCAACAGCGACATGTA 297
CATFISH       0
 MOUSE       241 TGAGTGGAACTTTGATGGCTCCAGTACCTTTCAGTCTGAAGGCTCCAACAGCAACATGTA 300
 PIG         238 CGAGTGGAATTTCGATGGCTCTAGTACTTTTCAGTCTGAAGGCTCCAACAGTGACATGTA 297
COW           0
 RAT         223 CGAGTGGAACTTTGATGGTTCTAGTACGTTTCAGTCTGAAGGCTCCAACAGCGACATGTA 282
 LT-HAMSTER  236 TGAGTGGAATTTTGATGGCTCTAGTACCTTTCAGTCTGAGGGCTCCAACAGTGACATGTA 295
 C_HAMSTER   174 TGAGTGGAATTTTGATGGCTCTAGTACCTTTCAGTCTGAGAGCTCCAACAGTGACATGTA 233
 HUMAN       217 TGAGTGGAATTTCGATGGCTCTAGTACTTTACAGTCTGAGGGTTCCAACAGTGACATGTA 276
 ZEBRAFISH   238 TGAGTGGAACTTTGATGGTTCCAGCACATATCAGGCTGAGGGGTCCAACAGTGACATGTA 297
 MUDSUCKER   209 AGAATGGAACTTTGATGGCTCCAGCACATATCAAGCAGAAGGTTCCAATAGTGACATGTA 268

 GLFTDFISH   298 CTTGGTTCCCGCTGCCATGTTCCGTGATCCCTTTCGCGAAGATCCCAACAAGCTTGTCCT 357
 XENOPUS     298 TCTCATCCCAGTCCAGATGTTCAGAGACCCATTCTGCCTGGACCCCAATAAACTGGTTAT 357
 DOGFISH     298 CCTGGTTCCATCTGCCATGTTCCGGGATCCTTTCCGTAGGGATCCAAACAAGCTCGTCCT 357
CATSHARK      0
 HORNSHARK   174 CCTCATCCCATCTGCCATGTTCCGGGATCCTTTCCGTAAGGATCCAAACAAGCTCATCCT 233
 CHICKEN     298 CCTGCGACCTGCTGCCATGTTCCGGGACCCTTTTCGCAAGGATCCCAACAAATTAGTTCT 357
CATFISH       0
 MOUSE       301 TCTCCATCCTGTTGCCATGTTTAGAGACCCCTTCCGC---------AACAAGCTGGTGCT 351
 PIG         298 TCTTGTCCCTGCTGCCATGTTTCGGGACCCTTTCCGCAAGGACCCCAACAAGCTGGTGTT 357
COW           0
 RAT         283 CCTCCATCCTGTGGCCATGTTTCGAGACCCCTTCCGCAGAGACCCCAACAAGCTGGTGTT 342
 LT-HAMSTER  296 TCTCAGCCCTGTTGCCATGTTTCGGGACCCCTTCCGCAGAGATCCCAACAAGCTGGTGTT 355
 C_HAMSTER   234 TCTCAGCCCTGTTGCCATGTTTCGGGACCCCTTCCGCAAAGAGCCCAACAAGCTGGTGTT 293
 HUMAN       277 TCTCGTGCCTGCTGCCATGTTTCGGGACCCCTTCCGTAAGGACCCTAACAAGCTGGTGTT 336
 ZEBRAFISH   298 TTTGTTCCCTCAAGCCATGTTCAGAGACCCTTTCAGGAAAGACCCCAACAAACTGGTTCT 357
 MUDSUCKER   269 TCTGGTCCCTGCTGCCATGTTCCGTGACCTTTCCGCAAGACCCAACNAACTGGTCCTGTG 328
```

Appendix 1. continued

```
GLFTDFISH    358 TTGTGAAGTGCTGAAGTACAACCGCAAACCATCAGAATCCAATCTTCGGTTGAACTGTAA 417
XENOPUS      358 GTGTGAAGTCTTGAAATACAACCGCAAGTCTGCAGAGACCAACCTGAGACACACATGCAA 417
DOGFISH      358 CTGTGAGGTCCTCAAGTATAACAGGAAGCCAGCAGAATCTAATCTTAGACACTCATGCCA 417
CATSHARK       0
HORNSHARK    234 CTGTGAAGTCTTCAAGTACAACAGAAAGCCAGCAGAAACTAATCTTAGAAACTCATGCCA 293
CHICKEN      358 CTGTGAGGTCTTCAAATACAACCGCCAGTCTGCAGACACAAATCTTCGGCACACCTGTAG 417
CATFISH        0
MOUSE        352 ATGTGAAGTTTTCAAGTATAACCGGAAGCCTGCAGAGACCAACTTGAGGCACATCTGTAA 411
PIG          358 CTGTGAGGTCTTCAAGTACAACCGAAAGCCTGCAGAGACCAACTTAAGGCACACCTGTAA 417
COW            0
RAT          343 CTGCGAAGTATTCAAGTATAACCGGAAGCCCGCAGAGACCAACCTGAGGCACAGCTGTAA 402
LT-HAMSTER   356 CTGTGAAGTTTTCAAGTACAACCGGAAGCCTGCAGAGACCAATTTAAGGCACTCGTGTAA 415
C_HAMSTER    294 CTGTGAAGTCTTCAAGTACAACCAGAAGCCTGCAGAGACCAATTTAAGACACACGTGTAA 353
HUMAN        337 ATGTGAAGTTTTCAAGTACAATCGAAGGCCTGCAGAGACCAATTTGAGGCACACCTGTAA 396
ZEBRAFISH    358 GTGCGATGTTCTGAAATACAACCATAAACCTGCAGAAACCAATCTTCGTCAGTCCTGTAA 417
MUDSUCKER    329 TGAAGTGCTCAAGTTCCACCGCCAGCCTGCAGAAACCAACCTGAAGATTACATGTT      384

GLFTDFISH    418 CAAGGTGATGAACATGGTCAAGGACCAGCATCCTTGGTTTGGCATGGAGCAAGAGTACAC 477
XENOPUS      418 GAAGATCATGGAGATGGTGAATGACCACCGCCCGTGGTTTGGAATGGAGCAGGAATACAC 477
DOGFISH      418 GAAAATCATGTCCATGATCGCAAATGAATATCCATGGTTTGGAATGGAACAAGAGTACAC 477
CATSHARK       0
HORNSHARK    294 GAAAGTCATGTCCATGGTCGCAGGTGAACACCCATGGTTTGGAATGGAACAGGAATACAC 353
CHICKEN      418 GCGGATTATGGATATGGTGTCCAACCAGCACCCCTGGTTTGGGATGGAGCAGGAGTACAC 477
CATFISH        0                                          GTTTGGCATGGAGCAGGAGTACAC  24
MOUSE        412 ACGGATAATGGACATGGTGAGCAACCAGCACCCCTGGTTTGGAATGGAGCAGGAATATAC 471
PIG          418 ACGGATAATGGACATGGTGAGCAACCAGCACCCCTGGTTTGGAATGGAGCAGGAATATAC 477
COW            0
RAT          403 GCGTATAATGGACATGGTGAGCAGCCAGCACCCCTGGTTTGGAATGGAACAGGAGTATAC 462
LT-HAMSTER   416 ACGGATAATGGACATGGTGAGCAACCAGCACCCCTGGTTTGGAATGGAACAGGAGTATAC 475
C_HAMSTER    354 ACGGATAATGGACATGGTGAGCAACCAGCACCCCTGGTTTGGAATGGAACAGGAGTATAC 413
HUMAN        397 ACGGATAATGGACATGGTGAGCAACCAGCACCCCTGGTTTGGCATGGAGCAGGAGTATAC 456
ZEBRAFISH    418 GAAGATTATGGATATGGTCCAGAACCAGCATCCTTGGTTTGGAATGGAACAGGAGTACAC 477
MUDSUCKER

GLFTDFISH    478 CATTCTTGGCACAGATGGACATCCTTTCGGCTGGCCATCTAATGGATTTCCCGGACCACA 537
XENOPUS      478 CTTGCTGGGCATTAATGGGCACCCGTATGGCTGGCCAGAAAATGGTTTCCCAGGGCCACA 537
DOGFISH      478 TTTGCTGGGAACGGACGGTCATCCCTTTGGATGGCCTTCCAATTGCTTTCCTGGACCACA 537
CATSHARK       0                                                    GGACCGCA   8
HORNSHARK    354 TCTTCTGGGAACAGATGGACATCCCTTTGGATGGCCTTCCAATGGGTTTCCTGGACCACA 413
CHICKEN      478 CCTTCTGGGAACAGATGGTCATCCGTTTGGCTGGCCTTCCAATTGCTTCCCTGGACCCCA 537
CATFISH       25 CATCCTGGGAACGGACGGTCACCCGTTCGGCTGGCCTTCCAACGGCTTTCCCGGTCCTCA  84
MOUSE        472 TCTCATGGGAACAGACGGCCACCCGTTTGGTTGGCCTTTCAATGGCTTCCCTGGACCCCA 531
PIG          478 TCTCATGGGCACAGATGGACACCCCTTTGGTTGGCCTTCCAATGGCTTCCCTGGGCCCCA 537
COW            0
RAT          463 TCTCATGGGAACAGACGGCCACCCTTTCGGCTGGCCTTCTAATGGCTTCCCTGGACCCCA 522
LT-HAMSTER   476 TCTGATGGGAACAGATGGGCACCCTTTTGGTTGGCCTTCCAATGGCTTTCCTGGGCCCCA 535
C_HAMSTER    414 TCTCTTGGGAACAGATGGGCACCCTTTTGGTTGGCCTTCCGATGGCTTCCCTGGGCCCCA 473
HUMAN        457 CCTCATGGGGACAGATGGGCACCCCTTTGGTTGGCCTTCCAACGGCTTCCCAGGGCCCCA 516
ZEBRAFISH    478 TCTTCTCGGCACAGATGGTCATCCTTTCGGTTGGCCCTCCAATGGCTTCCCTGGACCTCA 537
MUDSUCKER
```

Appendix 1. continued

```
GLFTDFISH     538 AGGTCCATATTACTGTGGTGTGGGAGCAGACAAGGCCTACGGCAGAGACATAGTGGAGGC 597
XENOPUS       538 AGGTCCCTATTACTGCGGCGTTGGAGCGGACAAGGTGTATGGCCGGGATGTGGTAGAGTC 597
DOGFISH       538 AGGGCCCTATTACTGTGGAGTTGGTGCAGACAAAGCTTACGGCAGAGATATTGTCGAGGC 597
CATSHARK        9 GGGACCCTATTACTGTGGCGTTGGTGCAGATAAAGCTTATGGTCGGGATATTGTGGAGGC 68
HORNSHARK     414 AGGGCCCTATTACTGTGGAGTTGGTGCAGACAAAGCTTACGGTAGAGATATTGTGGAAGC 473
CHICKEN       538 AGGTCCGTACTACTGCGGTGTAGGAGCTGACAAAGCCTATGGCAGAGACATTGTGGAGGC 597
CATFISH        85 GGGGCCTTACTACTGTGGAGTCGGAGCGGACAAGGCGTACGGCAGGGATATTGTGGAAGC 144
MOUSE         532 AGGCCCATATTACTGCGGTGTGGGAGCAGACAAAGCCTATGGCAGGGACATCGTGGAGGC 591
PIG           538 AGGTCCGTACTATTGTGGTGTTGGAGCAGACAAAGCCTATGGCAGGGACATTGTGGAGGC 597
COW             0
RAT           523 AGGACCCTATTACTGCGGTGTGGGAGCTGACAAGGCTTATGGCCGAGATATCGTGGAGGC 582
LT-HAMSTER    536 AGGTCCGTATTACTGTGGTGTGGGCGCAGACAAAGCCTATGGCAGGGATATCGTGGAGGC 595
C_HAMSTER     474 AGGTCTGTATTACTGTGGTGTGGGCGCAGACAAAGCCTATCGCAGGGATATCATGGAGGC 533
HUMAN         517 GGGTCCATATTACTGTGGTGTGGGAGCAGACAGAGCCTATGGCAGGGACATCGTGGAGGC 576
ZEBRAFISH     538 AGGTCCATATTACTGTGGTGTTGGAGCTGATAANGCCTATGGACGAGATGTTGTAGAAGC 597
MUDSUCKER

GLFTDFISH     598 CCATTACAGAGCCTGTCTCTATGCTGGAGTCCAGATTTGTGGCACAAATGCAGAAGTAAT 657
XENOPUS       598 GCATTATAAGGCCTGTCTGTACGCTGGCATTAAAATCTGTGGCACCAACGCAGAAGTCAT 657
DOGFISH       598 TCACTACCGGGCGTGTCTGTATGCTGGAATTGAACTCAGTGGAACCAATGCTGAAGTTAT 657
CATSHARK       69 TCACTACCGAGCATGTCTATATGCTGGGATTCACTTGTCTGGTACCAATGCTGAAGTGAT 128
HORNSHARK     474 TCACTACCGGGCTTGTCTGTATGCTGGAATCCATCTCTCTGGCACCAATGCTGAAGTGAT 533
CHICKEN       598 CCACTACCGAGCGTGCCTGTATGCTGGTGTGAAAATTGGAGGAACCAACGCAGAAGTGAT 657
CATFISH       145 CCACTACAGAGCGTGTCTGTACGCCGGCGTGAATATCTGCGGCACGAACGCTGAGGTCAT 204
MOUSE         592 TCACTACCGGGCCTGCTTGTATGCCGGAGTCAAGATCACGGGGACAAATGCGGAGGTTAT 651
PIG           598 TCACTACCGGGCCTGCTTGTATGCCGGCATCAAGATTGGGGGCACCAATGCCGAGGTCAT 657
COW             0
RAT           583 TCACTACCGGGCCTGCTTGTATGCTGGAATCAAGATCACAGGGACAAATGCCGAGGTTAT 642
LT-HAMSTER    596 TCACTACCGCGCCTGCTTGTATGCTGGGGTCAAGATTACAGGAACAAATGCTGAGGTCAT 655
C_HAMSTER     534 TCACTACCGTGCCTGCTTGTATGCTGGGGTCAAGATTACAGGAACATATGCTGAGGTCAA 593
HUMAN         577 CCATTACCGGGCCTGCTTGTATGCTGGAGTCAAGATTGCGGGGACTAATGCCGAGGTCAT 636
ZEBRAFISH     598 ACATTATAGAGCCTGTCTGTATGCTGGGGTAAAATCTGTGGCACCAATGCTGAGTCATGC 657
MUDSUCKER

GLFTDFISH     658 GCCTGCACAGTGGGAGTTTCAGGTAGGACCTTGTGAGGGTATCAACATGGGCGATCATTT 717
XENOPUS       658 GCCCTCGCAGTGGGAGTTCCAAGTGGGTCCGTGCGAAGGTATCGACATGGGGGACCACCT 717
DOGFISH       658 GGCTGCTCAGTGGGAATACCAAGTTGGACCTTGTGAAGGTATCCAGATGGGTGACCACTT 717
CATSHARK      129 GGCTTCTCAGTGGGAGTACCAGGTTGGACCTTGCGAGGGCATCCATATGGGTGACCACTT 188
HORNSHARK     534 GGCTTCTCAGTGGGAGTACCAAGTTGGACCTTGTGAAGGTATCAAGGTGGGTGACCACTT 593
CHICKEN       658 GCCAGCCCAGTGGGAGTTCCAGGTGGGACCGTGCGAAGGGATTGAGATGGGGGATCACCT 717
CATFISH       205 GCCAGCTCAGTGGGAGTTCCAGGTGGGGCCGTGCGAGGGTATCGAGATGGGAGATCACCT 264
MOUSE         652 GCCTGCCCAGTGGGAATTCCAGATAGGACCCTGTGAGGGGATCCAGATGGGAGATCATCT 711
PIG           658 GCCCGCCCAGTGGGAATTCCAGATCGGACCCTGTGAAGGAATCGACATGGGAGATCACCT 717
COW             0
RAT           643 GCCTGCCCAGTGGGAATTCCAGATAGGACCCTGCGAAGGGATCCGCATGGGAGATCATCT 702
LT-HAMSTER    656 GCCTGCCCAGTGGGAATTCCAAATAGGACCCTGTGAAGGAATCCGCATGGGAGATCATCT 715
C_HAMSTER     594 GCATGCCCAGTGGGAATTCCAAATAGGACCCTGTGAAGGAATCCGCATGGGAGATCATCT 653
HUMAN         637 GCCTGCCCAGTGGGAATTTCAGATTGGACCTTGTGAAGGAATCAGCATGGGAGATCATCT 696
ZEBRAFISH     658 CTGCACAGTGG                                                  668
MUDSUCKER
```

Appendix 1. continued

```
GLFTDFISH     718 CTGGGCGGCACGTTTCATCCTGCACCGTGTCTGTGAGGATTTGGGCGTGGTCGCTTCATT 777
XENOPUS       718 GTGGATGGCCAGGTTCATCCTTCATCGGGTCTGTGAAGACTTTGGGGTGGTGGCGACTCT 777
DOGFISH       718 GTGGATTTCCAGGTTTATTCTGCACAGGGTGTGCGAGGACTTCGGTATCATTGCTAGCTT 777
CATSHARK      189 ATGGATGTCGAGGTTTATTCTGCACCGCGTGTGTGAGGACTTTGGGATCATCGCTAGCTT 248
HORNSHARK     594 GTGGATTTCAAGGTTTATTCTGCACAGGGTGTGCGAGGACTTTGGTATCATTGCTAGCTT 653
CHICKEN       718 CTGGATAGCACGTTTCATCCTCCACCGGGTGTGCGAAGACTTTGGTGTCATTGTGTCCTT 777
CATFISH       265 GTGGGTGGCTCGTTTCATCCTGCACAGGGTGTGTGAAGACTTCGGCATCGTCGCCTCGTT 324
MOUSE         712 TTGGATAGCCTGTTTTATCTTGCATCGGGTATGCGAAGACTTTGGGGTGATAGCAACCTT 771
PIG           718 CTGGGTGGCCCGATTCATCTTGCATCGTGTGTGCGAAGACTTCGGAGTGATCGCCACCTT 777
COW             0
RAT           703 CTGGGTAGCCCGTTTTATCTTGCATCGGGTATGCGAAGACTTTGGGGTGATAGCAACCTT 762
LT-HAMSTER    716 CTGGGTGGCCCGTTTCATCTTGCATCGAGTATGTGAAGACTTTGGGGTAATAGCAACCTT 775
C_HAMSTER     654 CTGGGTGGCCCGTTTCATCTTGCATCGAGTATGTAAAGACTTTGGAGTAATAGCAACCTT 713
HUMAN         697 CTGGGTGGCCCGTTTCATCTTGCATCGTGTGTGTGAAGACTTTGGAGTGATAGCAACCTT 756
ZEBRAFISH
MUDSUCKER

GLFTDFISH     778 TGACCCTAAGCCCATCCCCGGAAACTGGAACGGTGCTGGCTGCCATACAAACTTCAGCAC 837
XENOPUS       778 GGACCCCAAACCCATGACCGGAAACTGGAACGGAGCCGGGTGCCACACCAACTACAGCAC 837
DOGFISH       778 TGACCCTAAGCCCATTCCTGGCAACTGGAATGGTGCTGGGTGCCACACTAACTTTAGCAC 837
CATSHARK      249 TGACCCGAAGCCTATTCCTGGGAACTGGAACGGTGCTGGATGTCATACCAACTTTAGCAC 308
HORNSHARK     654 TGACCCGAAGCCCATTCCTGGCAACTGGAATGGGGCAGGGTGCCACACCAACTTTAGCAC 713
CHICKEN       778 CGATCCCAAACCCATCCCTGGGAACTGGAACGGTGCTGGCTGTCACACCAACTTCAGCAC 837
CATFISH       325 CGACCCCAAACCCATCCCTGGGAACTGGAACGGCGCGGGATGTCACACCAACTTCAGCAC 384
MOUSE         772 TGACCCCAAGCCCATTCCAGGGAACTGGGATGGTGCAGGCTGCCATACCAACTTCAGCAC 831
PIG           778 TGATCCTAAGCCCATTCCTGGGAACTGGAATGGTGCCGGCTGCCACACCAACTTTAGCAC 837
COW             0
RAT           763 TGACCCCAAGCCCATTCCAGGGAACTGGAATGGGGCAGGCTGCCACACCAACTTTAGCAC 822
LT-HAMSTER    776 TGACCCCAAGCCCATTCCTGGGAACTGGAATGGTGCAGGCTGCCATACCAACTTTAGCAC 835
C_HAMSTER     714 TGACTCCAAGCCCATTCCTGGGAACTGGAATGGTGCAGGCTGCCATACCAACTTTAGTAC 773
HUMAN         757 TGATCCTAAGCCCATTCCTGGGAACTGGAATGGTGCAGGCTGCCATACCAACTTCAGCAC 816
ZEBRAFISH
MUDSUCKER

GLFTDFISH     838 GAAAGAGATGAGGGAAGACGGCGGATTAAAAGCCATTGAAGATGCGATTGAGAAGCTCGG 897
XENOPUS       838 GGAGAGCATGAGGGTGGAAGGAGGACTCAAACACATTGAAGATGCCATAGAGAAGCTGGG 897
DOGFISH       838 CAAAGCCATGCGGGATGATGGAGGGTTGAAGTACATTGAAGACTCAATTGAAAAACTGGG 897
CATSHARK      309 AAAATCTATGCGGGATGAGGGCGGTTTGAAATTCATTGAAGAGTGTATTGAAAAACTGGG 368
HORNSHARK     714 CAAATCCATGCGGGAAGAGGGAGGGCTGAAGTACATTGAAGACTCCATTGAAAAACTGGG 773
CHICKEN       838 CAAGAACATGAGGGAAGATGGAGGTCTCAAGCACATCGAGGAGGCCATCGAGAAGCTGAG 897
CATFISH       385 TAAAGAGACGCGGGAAGAAGGCGGGCTCAAATGCATTGAGGAATGTATCGAGAAACTGGC 444
MOUSE         832 CAAGGCCATGCGGGAGGAGAATGGTCTGAAGTGCATTGAGGAGGCCATTGACAAACTGAG 891
PIG           838 CAAGGCCATGCGAGAGGAGAATGGTCTGAAGTACATCGAGGAGGCCATCGAGAAGCTAAG 897
COW             0                                 GTCTGAAGTACATTGAGGAGGCCATTGAGAAGCTAAG  37
RAT           823 CAAGGCCATGCGGGAGGAGAATGGTCTGAGGTGCATTGAGGAGGCCATTGATAAACTGAG 882
LT-HAMSTER    836 CAAGGCCATGCGGGAGGAGAATGGTCTGAAGCACATCGAGGAGGCCATCGAGAAACTAAG 895
C_HAMSTER     774 CAAGACCATGCGGGAGGAGAATGGTCTGAAGCACATCAAGGAGGCCATTGAGAAACTAAG 833
HUMAN         817 CAAGGCCATGCGGGAGGAGAATGGTCTGAAGTACATCGAGGAGGCCATTGAGAAACTAAG 876
ZEBRAFISH
MUDSUCKER
```

Appendix 1. continued

```
GLFTDFISH    898 GAAGAGGCACCACTACCACATTCGTGCCTATGACCCCAAAGGGGGGCTGGACAACGCCCG 957
XENOPUS      898 GAAGAGACACGACTACCACATCTGCGTCTACGACCCGCGGGGAGGGAAAGACAACTCCCG 957
DOGFISH      898 CAAGAGGCATCAGTACCACATTCGTGCCTATGATCCTAAAGGAGGGTTGGACAATGCTAG 957
CATSHARK     369 CAAGAGGCACCAATACCACATTCGTGCCTATGATCCTAAA                      408
HORNSHARK    774 CAAGAGGCATCAGTACCACATTCGTGCCTATGACCCTAAAGGAGGGTTGGACAATGCTAG 833
CHICKEN      898 CAAGCGCCACCAGTACCACATCCGTGCCTACGACCCCAAAGGAGGGCTGGACAACGCCCG 957
CATFISH      445 GAAGAGACACAACTACCACATCCGTGCCTACGATCCTAAAGGAGGCCTGGACAACGCTCG 504
MOUSE        892 CAAGAGGCACCAGTACCACATCCACACCTACGATCCCAAGGGGGGCCTGGACAACTCCCG 951
PIG          898 CAAGCGGCACCAGTACCACATCCGAGCCTACGATCCCAAGGGGGGCCTGGACAACACACG 957
COW           38 CAAGCGCCACCAGTACCACATCCGAGCCTACGATCCCAAGGGGGGCCTGGACAACGCCCG 97
RAT          883 CAAGAGGCACCAGTACCACATCCGTGCCTACGACCCCAAGGGGGGCCTGGACAACGCCCG 942
LT-HAMSTER   896 CAAGCGGCACCGGTACCACATTCGAGCCTACGATCCCAAGGGGGGCCTGGACAATGCCCG 955
C_HAMSTER    834 CAAGCGGCACCGGTACCATATTCGAGCCTACGATCCCAAGGGGGGGCTGGACAATGCCCG 893
HUMAN        877 CAAGCGGCACCAGTACCACATCCGTGCCTATGATCCCAAGGGAGGCCTGGACAATGCCCG 936
ZEBRAFISH
MUDSUCKER

GLFTDFISH    958 CCGTCTCACCGGCCACCACGAAACCTCAAACATCCACGAGTTCTCTGCAGGTGTGGCCAA 1017
XENOPUS      958 GAGACTCACCGGCCAACACGAGACGTCGAGTATTCACGAGTTCTCGGCCGGCGTGGCCAA 1017
DOGFISH      958 AGCTTTGACAGGCCACCATGAAACCTCAAATATCAATGAGTTCTCAGCTGGTGTTGCCAA 1017
CATSHARK
HORNSHARK    834 GCGTTTGACAGGCCACCATGAAACCTCAAATATCAATGAGTTCTCAGCTGGCGTTGCCAA 893
CHICKEN      958 GCGCCTGACGGGCTTCCACGAGACGTCCAGCATCCACGAGTTCTCCGCCGGCGTGGCCAA 1017
CATFISH      505 CCGCCTGACTGGCCACCACGAGACCTCCAACATCCACGAGTTCTCTGCCGGCGTCC      560
MOUSE        952 GCGTCTGACTGGATTCCACGAAACCTCCAACATCAACGACTTTTCTGCCAGTGTTGCCAA 1011
PIG          958 GCGCCTAACTGGATTCCATGAAACCTCCAACATCAACGACTTTTCTGCCGGCGTGGCCAA 1017
COW           98 GCGCCTAACTGGGTTCCACGAAACCTCCAACATCAACGACTTCTCTGCCGGCGTGGCCAA 157
RAT          943 CCGTCTGACTGGATTCCACGAAACCTCCAACATCAACGACTTTTCCGCTGGCGTTGCCAA 1002
LT-HAMSTER   956 TGGTCTGACTGGGTTCCACGAAACGTCCAACATCAACGACTTTTCTGCTGGTGTCGCCAA 1015
C_HAMSTER    894 TCGTCTGACTGGGTTCCACAAAACGTCCAACATCAACGACTTTTCAGCTGGCGTCGCCGA 953
HUMAN        937 ACGTCTAACTGGATTCCATGAAACCTCCAACATCAACGACTTTTCTGCTGGTGTAGCCAA 996
ZEBRAFISH
MUDSUCKER

GLFTDFISH   1018 CCGCGGCGCCAGCATTCGCATTCCCCGTAGTGTCGGCCAGGAGAAGAAGGGCTACTTTGA 1077
XENOPUS     1018 CCGGGGCGCCAGTATCCGCATCCCGCGTCAGGTGGGCCAGGAAGGCTACGGCTACTTTGA 1077
DOGFISH     1018 TAGAGGAGCCAGCATCCGAATCCCTCGATCCGTTGGCCAGGACAAGAAAGGCTACTTTGA 1077
CATSHARK
HORNSHARK    894 TAGAGGAGCTAGCATCCGAATCCCTCGATCTGTTGGCCAGGACAAGAAAGGCTACTTTGA 953
CHICKEN     1018 CCGCGGCGCCAGCATCCGCATCCCACGCAACGTGGGCCATGAGAAGAAAGGCTACTTCGA 1077
CATFISH
MOUSE       1012 CCGCAGTGCCAGTATCCGCATTCCCTGGACTGTCGGCCAGGAGAAGAAGGGCTACTTTGA 1071
PIG         1018 CCGTGGCGCTAGCATCCGCATTCCCCGGACTGGGGGCCAGGAGAAGAAGGGTTACTTCGA 1077
COW          158 CCGTGGTGCTAGCATCCGCATCCCCCGGACTGTTGGCCAGGAGAAGAAGGGCTACTTCGA 217
RAT         1003 CCGCAGCGCCAGTATCCGCATTCCCCGGATTGTCGGCCAGGAGAAGAAGGGTTACTTTGA 1062
LT-HAMSTER  1016 TCGCAGTGCCAGCATCCGCATTCCCCGGACTGTCGGCCAGGAGAAGAAAGGTTACTTTGA 1075
C_HAMSTER    954 TCGCAGTGCCAGCATCCGCATTCCCCGGACTGTCGGCCAGGAGAAGAAAGGTTACTTTGA 1013
HUMAN        997 TCGTAGCGCCAGACTACGCATTCCCCGGACTGTTGGCCAGGAGAAGAAGGGTTACTTTGA 1056
ZEBRAFISH
MUDSUCKER
```

Appendix 1. continued

```
GLFTDFISH   1078 GGACCGCCGACCGTCTGCCAACTGTGACCCGTACGGCGTAACGGAGGCCCTGATCCGCAC 1137
XENOPUS     1078 AGACCGACGGCCGGCAGCCAACTGCGACCCCTACGCAGTAACCGAGGCGCTGGTCAGGAC 1137
DOGFISH     1078 AGACCGCCGTCCATCTGCTAATTGTGACCCTTATGCAGTCACAGAAGCATTGGTCCGCAC 1137
CATSHARK
HORNSHARK    954 AGACCGCCGTCCCTCTGCTAATTGTGACCCTTATGCAGTCACAGAAGCATTGGTCCGCAC 1013
CHICKEN     1078 GGACCGCGGGCCTTCAGCCAACTGCGATCCCTACGCCGTGACGGAGGCCCTGGTCCGTAC 1137
CATFISH
MOUSE       1072 AGACCGTCGGCCTTCTGCCAATTGTGACCCCTATGCGGTGACAGAAGCCATCGTCCGCAC 1131
PIG         1078 AGACCGTCGCCCTTCTGCCAACTGTGACCCCTTTGCGGTGACAGAAGCTCTCATCCGCAC 1137
COW          218 AGACCGTCGCCCATCTGCCAACTGTGACCCCTTCGCCGTGACCGAAGCCCTCATCCGCAC 277
RAT         1063 AGACCGTCGGCCTTCTGCCAATTGCGACCCCTATGCGGTGACGGAAGCCATCGTCCGCAC 1122
LT-HAMSTER  1076 AGACCGCCGCCCCTCTGCCAATTGTGACCCCTTTGCAGTGACAGAAGCCATCGTCCGCAC 1135
C_HAMSTER   1014 AGCCCGCTGCCCCTCTGCCAATTGTGACCCCTTTGCAGTGACAGAAGCCATCGTCCGCAC 1073
HUMAN       1057 AGATCGTCGCCCCTCTGCCAACTGCGAGCCCTTTTCGGTGACAGAAGCCCTCATCCGCAC 1116
ZEBRAFISH
MUDSUCKER

GLFTDFISH   1138 GTGTTTGCTGAGCGAGGAAGGAGATGAACCTTTAGCTTACTGAATCCCACTCCCCTCCTG 1197
XENOPUS     1138 CACCATCCTGAACGAAACCGGCAGCGAGACCAAAGACTATAAGAACGGAGCTGGATTCTC 1197
DOGFISH     1138 ATGCCTATTGGATGAGTCTGGGGACAAGCCTATTGAGTACAACAAAAATTAAGCAAAATA 1197
CATSHARK
HORNSHARK   1014 ATGCCTATTGGATGAGTCTGGGGACAAGCCT                               1044
CHICKEN     1138 GTGTCTCCTCAACGAAACCGGGGACGAGCCTTTTGAGTACAAGAACTAAGTGGACTCGTG 1197
CATFISH
MOUSE       1132 GTGTCTCCTCAACGAAACAGGCGACGAACCCTTCCAATACAAGAACTAAGCAGACTAGAC 1191
PIG         1138 GTGTCTCCTCAACGAAACTGGCGACGAGCCCTTCCAGTACAAAAACTAAGTGGACTAGAC 1197
COW          278 ATGTCTTCTGAATGAAACTGGCGACGAGCCCTTCCAGTACAAGAACTAAGTGGACTAGAC 337
RAT         1123 GTGTCTCCTCAACGAAACTGGCGACGAGCCCTTCCAATACAAGAACTAAGCGGACTCGAC 1182
LT-HAMSTER  1136 ATGCCTTCTCAATGAGACTGGCGACGAGCCCTTCCAATACAAAAACTAATTAGACTTTGA 1195
C_HAMSTER   1074 ATGCCTTCTCAATGAGACTGGCGACCAGCCCTTCCAATACAAAAACTAA              1122
HUMAN       1117 GTGTCTTCTCAATGAAACCGGCGATGAGCCCTTCCAGTACAAAAATTAAGTGGACTAGAC 1176
ZEBRAFISH
MUDSUCKER

GLFTDFISH   1198 ACATTCTTTTCTTTAAACTAGTACATTGTTTCTGTTCTCCTACTGAGATGATTTAACCTG 1257
XENOPUS     1198 CCGGGCAATCGGTATGGCATCTCCCCGAGACGCCGCTGTGTTTTAACCCGTTAGTCTCCC 1257
DOGFISH     1198 ATGCACTAATGGACCTGGCATTTGTAGCAGTGATAGCTGTTGAAATGTGGGACCTTTGGG 1257
CATSHARK
HORNSHARK
CHICKEN     1198 CCCACAGACACCGCCTTCCCCCTCCCCCCACCCCCCCCGTGCTCCCCGTACCCCTAAACT 1257
CATFISH
MOUSE       1192 TTCCAGTGATCCCTCTCCCAGCTCTTCCCTCTCCCAGTTGTCCCCACTGTAACTCAAAGG 1251
PIG         1198 GGGCAGCCATCAAAACCCCTCCAATTCTACACCGCCCCCCCCCCCCTCGCCCTCTCAACT 1257
COW          338 TTGCAGCCCTCGAAACCCCTCTTAATTCTACATCTTACTCCCACTCTCG             386
RAT         1183 TTCCAGTGATCTTGAGCCCTTCCTAGTTCACCCCACTCCCAACTGTTCCCTCTCCCACTG 1242
LT-HAMSTER  1196 GTGATCTTGAGCCTTTCCTAGTTCATCCCACCCCGCCCCAGCTGTCTCATTGTAACTCAA 1255
C_HAMSTER
HUMAN       1177 CTCCAGCTGTTGAGCCCCTCCTAGTTCTTCATCCCTGACTCCAACTCTTCCCCCTCTCCC 1236
ZEBRAFISH
MUDSUCKER
```

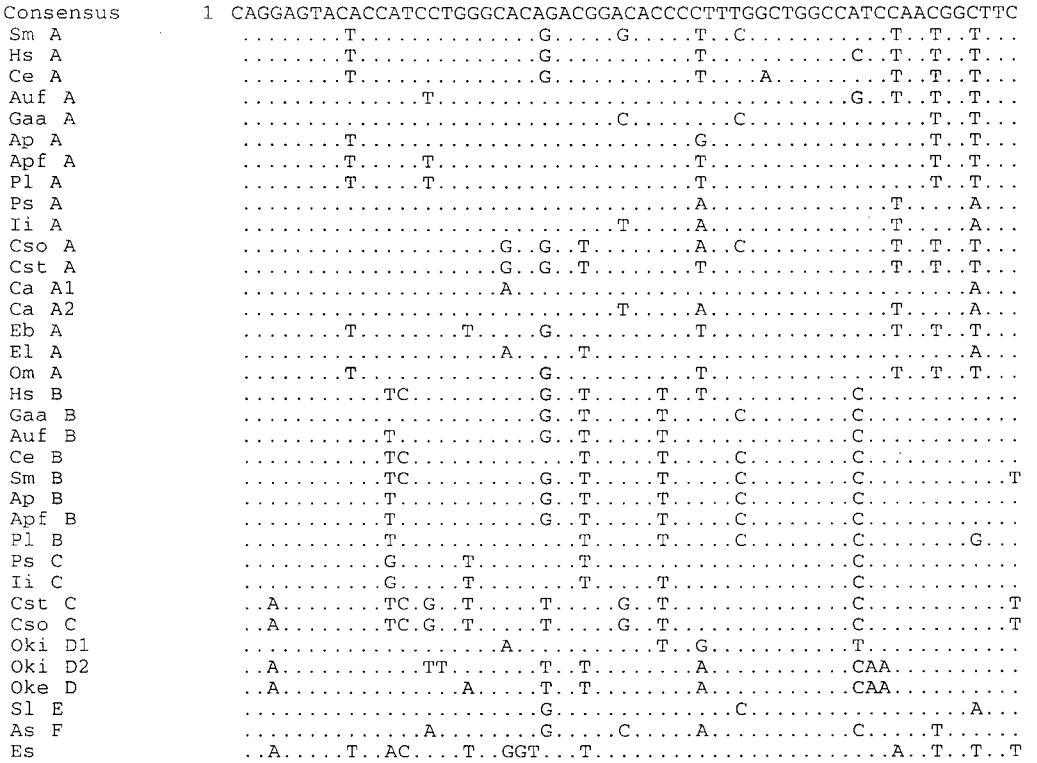
Appendix 1. continued

```
GLFTDFISH   1258 CATTTTAATGGTTTAAAAGTTGGCTGGTCAACTTAAAACAAGGCGGTCTTGTCCTTGGTA 1317
XENOPUS     1258 GACACTACTTGAATTCTTGTGAAGTAAAAATTCCTTTGGAAAGGAGGGGCATTCCTAGAA 1317
DOGFISH     1258 TCTCTACTCTACTCCTATACTGTACAGGTGCTAAAGGGGGAGGGTCAGAAGGGTTTTATT 1317
CATSHARK
HORNSHARK
CHICKEN     1258 TCCCTTCTAGTTGTAATCCTGAGGGTACAAGATAACACCTTCGTGTCTCAGTAACTCTTG 1317
CATFISH
MOUSE       1252 ATGGAATACCAAGGTCTTTTTATTCTTCGTGCCAAAAAAAAGAGAAATTTTTAACTTACT 1311
PIG         1258 GTCGTAATAGCTGTAACTCAAAGGGCGGAATAGCAAGGTCTTTTTTATTCCTCAAAAAAA 1317
COW
RAT         1243 GTCCCCACTGTAACTCAAAAGGATGGAATATCAAGGTCTTTTTATTCCTTGCGCCCAGTT 1302
LT-HAMSTER  1256 AGGATGGAATATCAAGGTCTTTTTATTCCTCGTGCCCAGTTAATCTTGCTTTTATTGGTC 1315
C_HAMSTER
HUMAN       1237 AGTTGTCCCGATTGTAACTCAAAGGGTGGAATATCAAGGTCGTTTTTTTCATTCCATGTG 1296
ZEBRAFISH
MUDSUCKER

GLFTDFISH   1318 GTGGTGAGCTGGTAATAGCAGGGTATGTTCCGCTTGCCTTCTGACGGGACTGGCCTTTTG 1377
XENOPUS     1318 CCAGAGACCGTAACATGGTTTCTCCCGTGTTATCTGCCGAGATGGAGGGCCAATTTGGGC 1377
DOGFISH     1318 TGTATTTCAGAACCTAATTTCTTCTGTTGTTATCTGGAAGGTGAGGAATGAGGCTTGCGA 1377
CATSHARK
HORNSHARK
CHICKEN     1318 TTGTTTTGAGGTGGGGGAGGAGGGCAGGTTTAGTTTTATTAATGTCTGTTTGTCATTGAC 1377
CATFISH
MOUSE       1312 GTGAATCTAATCATATTCATTTTTTTCCATTTTTATATTATCCATGAACAACTTTTAGTG 1371
PIG
COW
RAT         1303 AATTTTTGCCTTTATTGGTCAGAATAGAGGGGTCAGGTTCTTAATCTCTACACACCCAAC 1362
LT-HAMSTER  1316 AGAATAGAGGAGTCAAGTTCTT                                        1337
C_HAMSTER
HUMAN       1297 CCCAGTTAATCTTGCTTTCTTTTGTTTGGCTGGGATAGAGGGGTCAAGTTATTAATTTCT 1356
ZEBRAFISH
MUDSUCKER

GLFTDFISH   1378 TAACACTGTATATAGTCTGTGTCCTGGCATTTCCTTTTTTGTTTTTTTTGGTCTGTCGGT 1437
XENOPUS     1378 ATGGGGGGGGGGGGGGTTGCACAGATTTTATAGGAATAAGAAACAAAGCAAGTGACCTGCT 1437
DOGFISH     1378 TAGGACAACAAAACTGTTCTCTTATTTAGAACAGTTAATAACTCTTCAAGTTGACTGGTC 1437
CATSHARK
HORNSHARK
CHICKEN     1378 TCTTCAAAAGGCGAGAGGAGGAGGGGGGTGGGGGGGGGGGTAGATGATTTTTAACCACTGT 1437
CATFISH
MOUSE       1372 ATCTTTGTCTACTACATTTTATTATGTTTTTGATTACATATTATTCTAAAAACAGCACCA 1431
PIG
COW
RAT         1363 CCCTTCTTTCCTAGCTAGCTTTCCAGTGGGGAACGGGAGGGGGTGGGGAAGGGTAACCCA 1422
LT-HAMSTER
C_HAMSTER
HUMAN       1357 TCACACCTACCCTCCTTTTTTTCCCTATCACTGAAGCTTTTTAGTGCATTAGTGGGGAGG 1416
ZEBRAFISH
MUDSUCKER
```

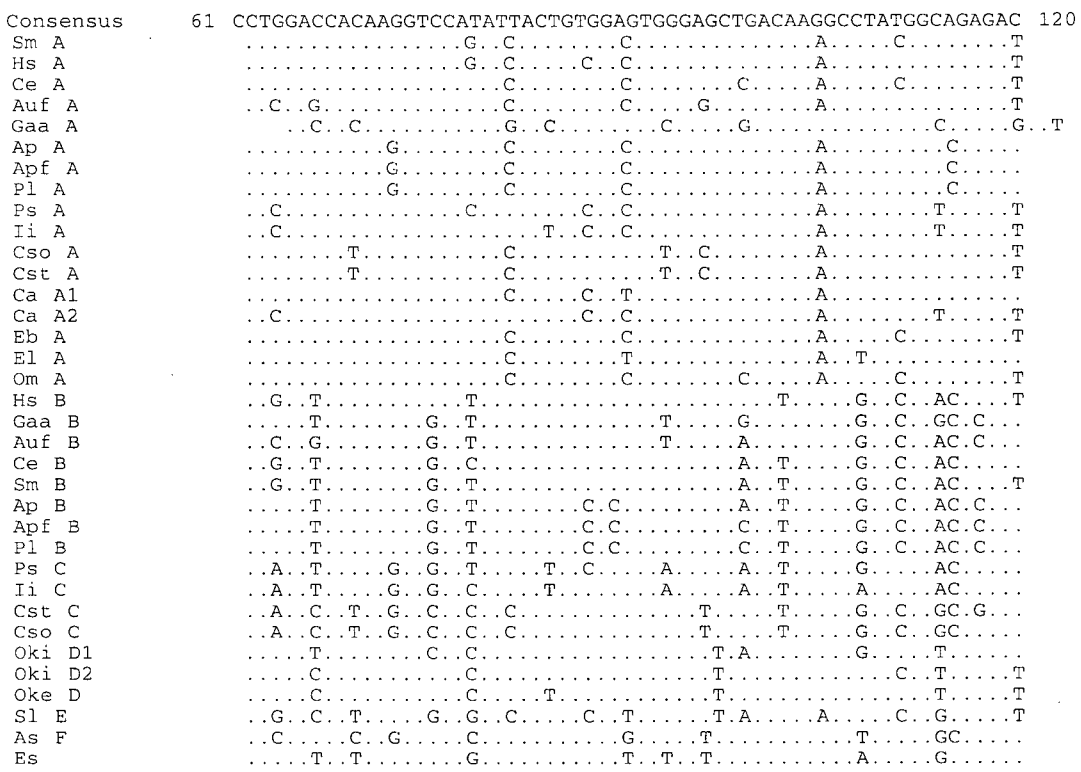
Appendix 1. continued

Name	Length	G	A	T	C
Gulftoadfish	432	29.17	25.23	22.69	22.92
African clawed frog	432	32.41	24.54	19.44	23.61
Spiny dogfish	432	28.47	24.31	25.93	21.30
Cloudy catshark	369	30.35	23.04	27.37	19.24
Hornshark	432	28.70	24.54	25.23	21.53
Chicken	432	30.32	23.84	20.60	25.23
Channel catfish	432	32.87	21.76	18.98	26.39
House Mouse	432	30.56	24.07	21.53	23.84
Pig	432	29.63	22.92	21.30	26.16
Norway Rat	432	30.09	23.61	21.53	24.77
Long-tailed hamster	432	29.63	24.31	22.45	23.61
Chinese hamster	432	28.24	25.23	23.61	22.92
Human	432	31.02	22.92	22.69	23.38
Zebrafish	202	27.36	21.89	27.36	23.38

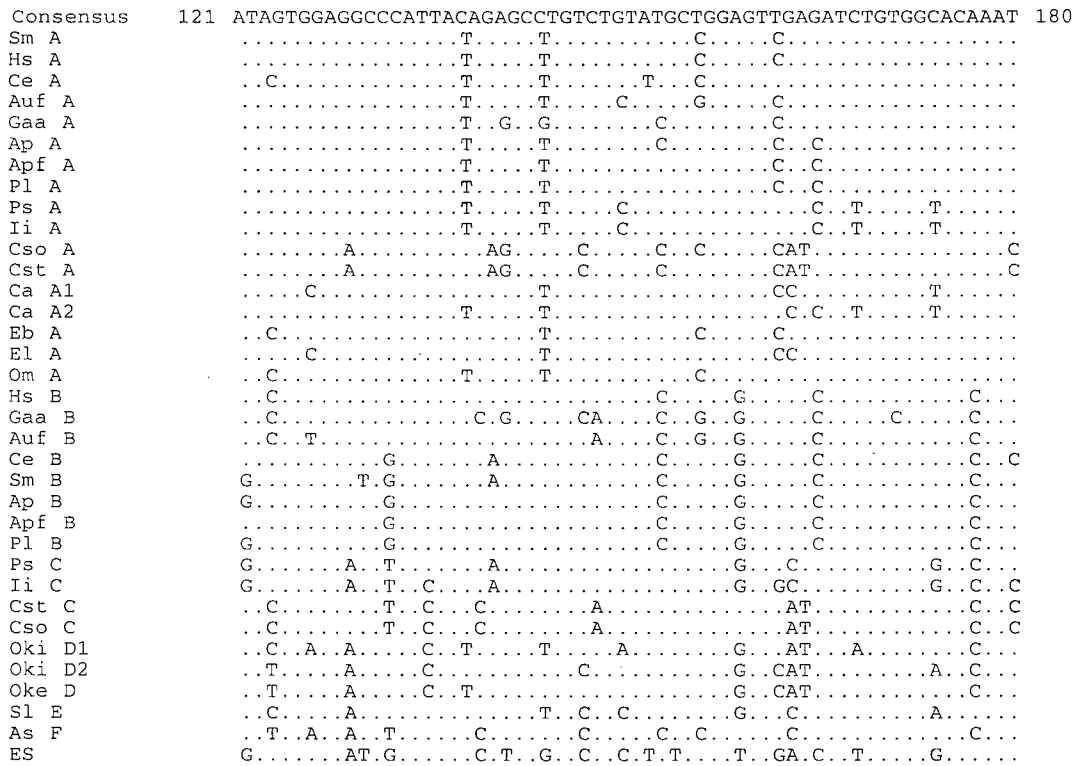
Appendix 2. Base composition percentage statistics for vertebrate sequences available in Genbank prior to August 1999. The fragment length of 432 bp spanned positions 467 to 899 of *Xenopus laevis* sequence in Appendix 1. Gulftoadfish (*Opsanus beta),* African clawed frog (*Xenopus laevis),* spiny dogfish (*Squalus acanthias),* cloudy catshark (*Scyliorhinus torazame),* hornshark *(Heterodontus francisci),* chicken *(Gallus gallus),* channel catfish (*Ictalurus punctatus),* house mouse (*Mus musculus),* pig *(Sus scrofa*), Norway rat (*Rattus norvegicus),* long-tailed hamster (*Cricetulus longicaudatus),* Chinese hamster (Cricetulus griseus), human *(Homo sapiens),* zebrafish *(Danio rerio).* Note: the zebrafish fragment and the cloudy catshark fragment only spanned a portion of the 432bp fragment length.



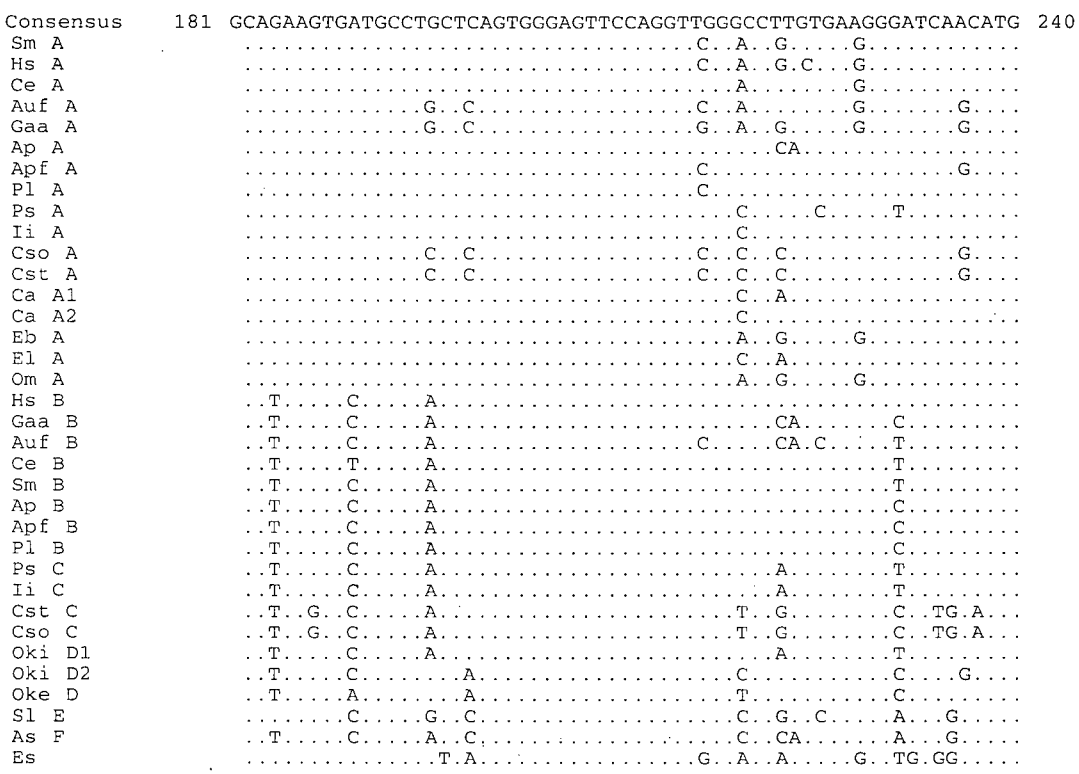
Appendix 3. Sequence alignment for amplified glutamine synthtase product for fish used in this study. The 432 bp region of glutamine synthetase amplified corresponds to the region of 467 to 899 of the published GS sequence for *Xenopus laevis* (Genbank accession number D50062). A dot represents an identical nucleotide base to the base given in the concensus sequence. Numbers indicate positional information relative to position 467 (with 1 being 467 and 432 being position 899) of the GS gene for *Xenopus laevis*. Species are designated by initials for species name and isoform designation is indicated by A, B, C, D, E or F. Ap – high cockscomb, Apf – penpoint gunnel, As – alligator gar, Auf – tubesnout, Ca – shiner perch, Ce – mossy sculpin, Cso – Pacific sanddab, Cst – speckled sanddab, Eb – buffalo sculpin, El – striped seaperch, Es – Pacific hagfish, Gaa – three spine stickleback, Hs – white spotted greenling, Ii – buttersole, Oke – chum Salmon, Oki – coho salmon, Om – tidepool sculpin, Pl – crescent gunnel, Ps – starry flounder, Sl – bay pipefish, Sm – cabezon. Note: Es does not have an isoform designation. ? – indicate base at that position is unknown for that fish.



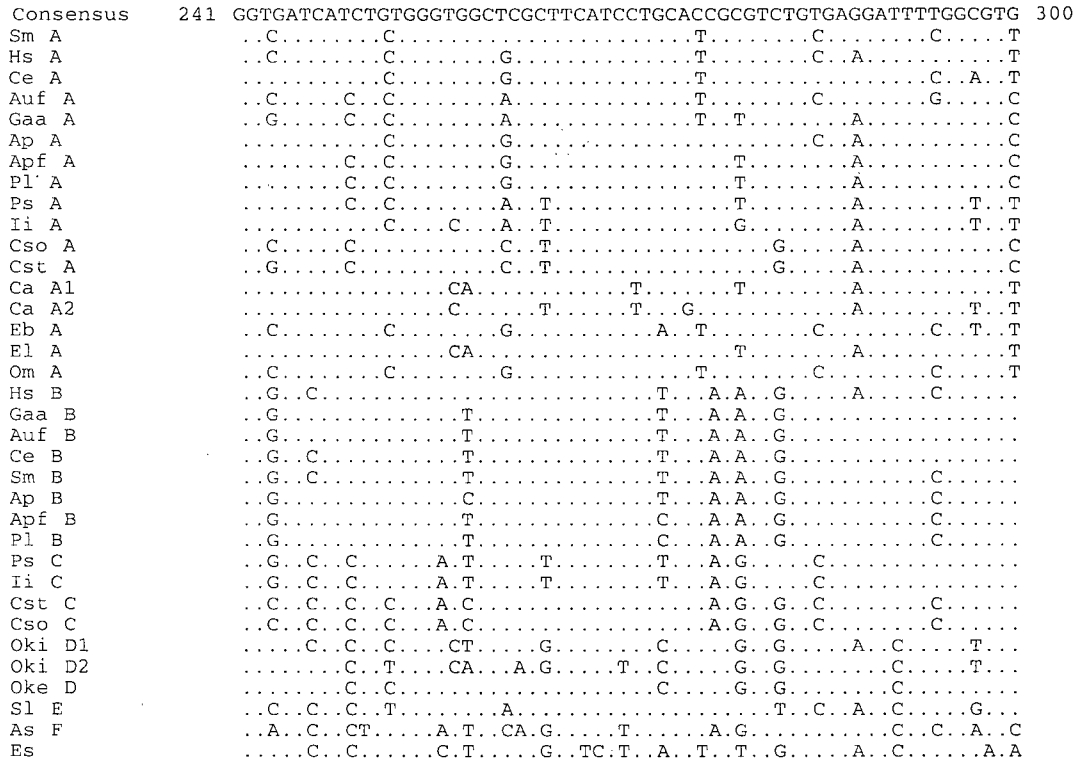
Appendix 3.continued.



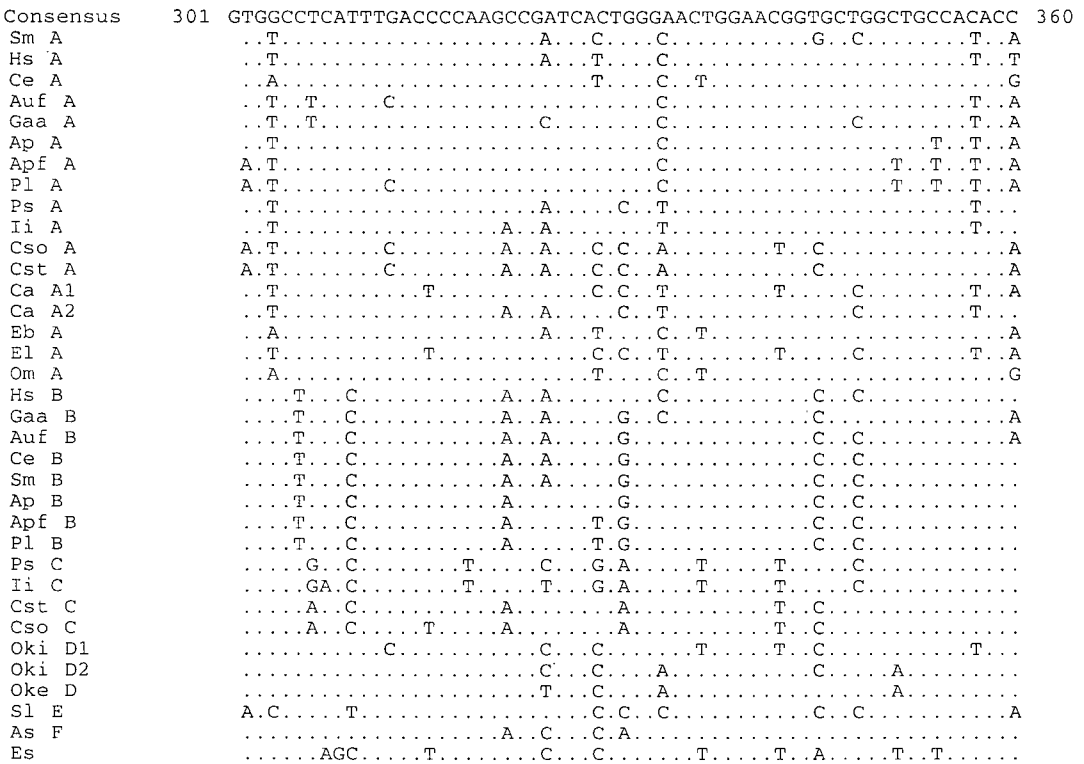
Appendix 3. continued.



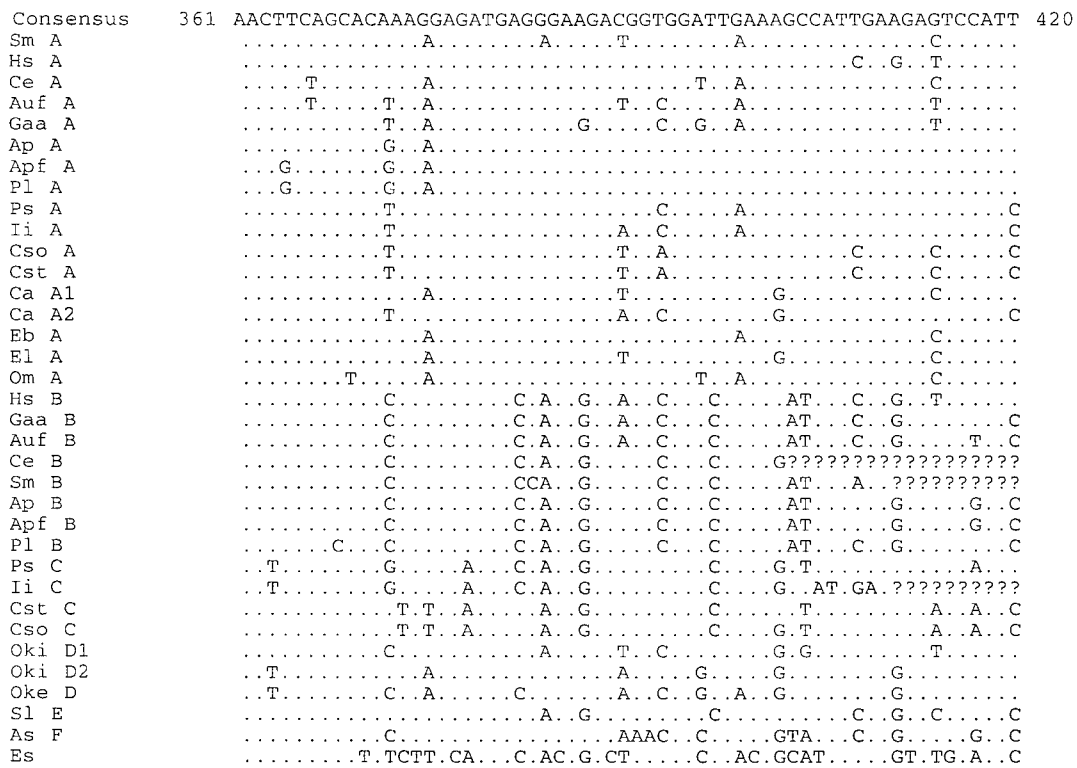
Appendix 3. continued.



Appendix 3. continued.



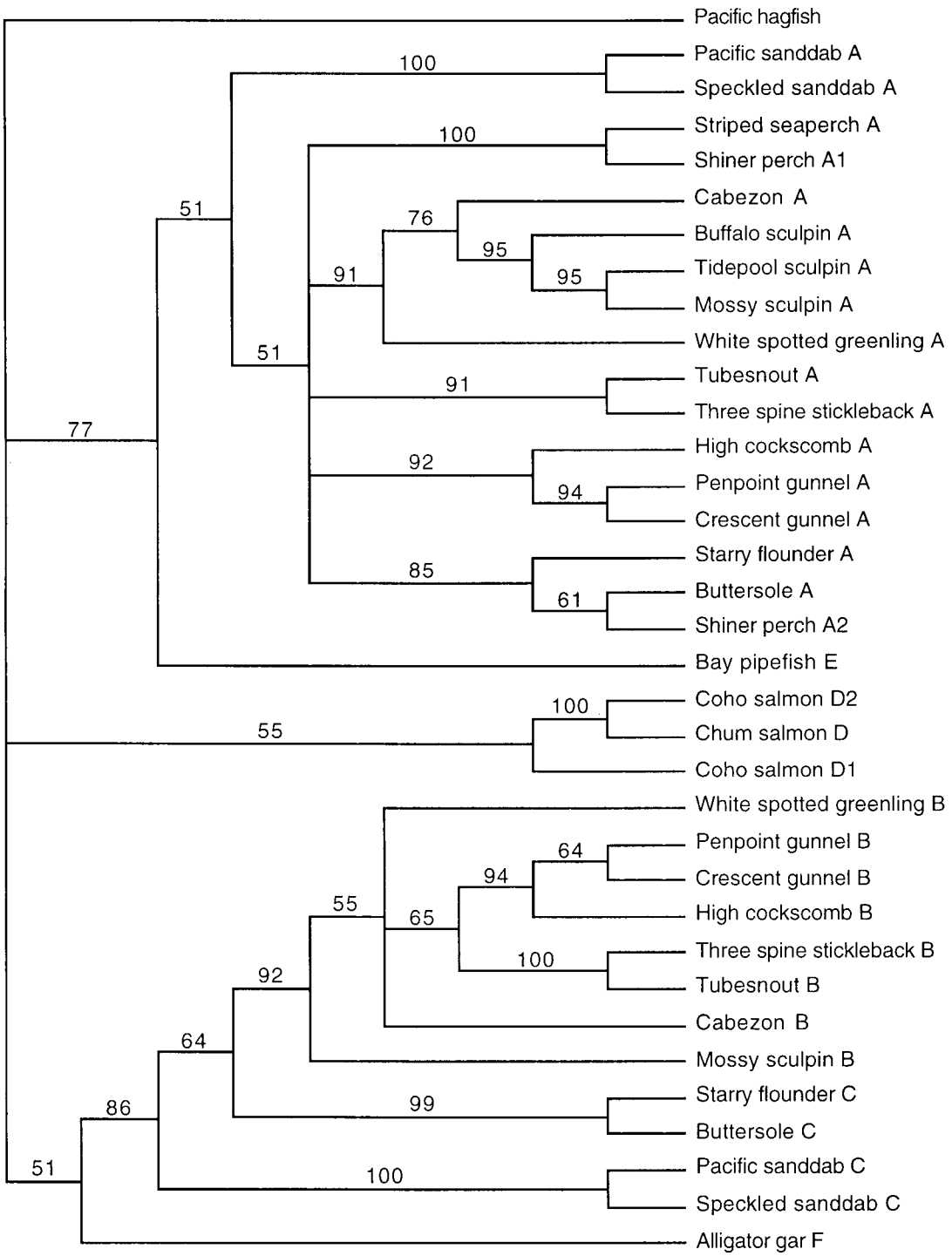
Appendix 3. continued.



Appendix 3. continued.

```
Consensus    421 GAGAAGCTGGGG 432
Sm A             ........C...
Hs A             ........T...
Ce A             ........C...
Auf A            ........T...
Gaa A            ........T...
Ap A             ........C...
Apf A            ........C...
Pl A             ........C...
Ps A             ........T..C
Ii A             ........T..C
Cso A            ........C..C
Cst A            ........C..C
Ca A1            ........T..A
Ca A2            ....T...T..C
Eb A             ........T...
El A             ........T..A
Om A             ........T...
Hs B             ..........AA
Gaa B            ....G....CCC
Auf B            ....GA....C.
Ce B             ????????????
Sm B             ????????????
Ap B             ....G.....C.
Apf B            ....G.....C.
Pl B             ....G.....C.
Ps C             ..A.G.....CA
Ii C             ????????????
Cst C            ..A.G.....CA
Cso C            ..A.G.....CA
Oki D1           ......T....A
Oki D2           ..A.G.......
Oke D            ....G.......
Sl E             ..........C
As F             ....G....A.C
Es               ..........CA
```

Appendix 3. continued.



Appendix 4. Maximum parsimony tree constructed from 432 bp fragments of all isoforms of glutamine synthetase for all fish used in this study. Parsimony criterion was set to random addition, 50 replicates, TBR branch swapping algorithm. Sequence data was weighted 2:4:1 by codon position and tree was bootstrapped 100 times.