

**Halibut mitogenomics: A study of the complete
mitochondrial genome sequences of Atlantic-, Pacific- and
Greenland halibut.**

**A master thesis in molecular biology
Master of Aquaculture
By
Kenneth Andre Mjelle**

**Department of Fisheries and Natural Sciences
Bodø Regional University**

August 2006

Acknowledgement

This work was done at the Department of Fisheries and Natural Sciences, Bodø Regional University, from August 05 to May 06.

I am grateful to my supervisor Steinar Johansen for establishing the group in molecular biology at Bodø Regional University. Both he and supervisor Truls Moum have provided valuable advice during the whole process of work with this thesis.

Thanks to Tor Erik Jørgensen for striving to keep the lab a nice and tidy place to work, and for the initial work performed on Atlantic halibut. I am also thankful to Bård Ove Karlsen for teaching me basal lab techniques.

Arild Eeg at Risør aquarium, Ruth Withler at Fisheries and Oceans Canada - Pacific Region and Michael E. Cobb kindly provided samples of several species in this study.

Ragna Breines, I still owe you a trip to Stetind or that bag of jelly-bears for helping me with figures. The choice is yours.

Kristin Pedersen, thanks for your immense patience and tolerance the last year.

Bodø, August 2006

Kenneth Andre Mjelle

Content

Summary.....	5
Introduction	6
The Pleuronectidae family.....	6
Population studies of Atlantic halibut	9
Organelles: Mitochondria and plastids	10
The organization of organelle genomes.....	10
Vertebrate mitochondrial genomes.....	11
Transmembrane helices in mitochondrial proteins.....	13
Organization of the fish mitochondrial DNA	14
The genetic code for vertebrate mitochondrial DNA.....	14
Codon usage bias	16
Molecular markers	16
Phylogenetic analysis	17
Aims of the study.....	21
Materials	22
Methods	25
DNA extraction.....	25
Polymerase chain reaction (PCR).....	25
PCR primer pairs	26
Agarose gel electrophoresis.....	30
Extraction of DNA from gel	31
Cloning of the fragments into the TOPO vector.....	31
Screening for recombinants	32
Sequencing.....	33
Bioinformatics	33
Results	35
Genome content.....	35
Protein coding genes.....	39
Ribosomal RNA genes	39
Transfer RNA genes	40
Non coding regions.....	45
Base composition.....	52
Codon usage bias	52
Intraspecific variation in nucleotide sequences	57
Nucleotide diversity compared in Atlantic-, Pacific- and Greenland halibut.....	63
Intraspecific variation in amino acid sequences	65
Interspecific variation in amino acid sequences	66
Non-conservative amino acid substitutions.....	67
Transmembrane helices	68
Phylogenetic investigation of mtDNA sequences from Pleuronectiform species.....	68
Estimation of divergence times	73
Authenticity of the Mitochondrial Sequences	73
Discussion.....	74
Genetic variation.....	74
Codon usage bias	80
Nucleotide composition	80
Phylogenetic investigation of mtDNA sequences	81
Species divergence times	82
Species identification based on molecular markers.....	83
References	85

Appendix	92
----------------	----

Summary

The complete mitochondrial DNA sequence was determined in four individuals of Atlantic halibut (*Hippoglossus hippoglossus*), Pacific halibut (*Hippoglossus stenolepis*) and Greenland halibut (*Reinhardtius hippoglossoides*) respectively. Thirteen protein-coding genes, twenty-two tRNA genes, two rRNA genes and a large non-coding control region were identified, and the conserved vertebrate gene order was confirmed. Extensive length variation of the mtDNA genome was observed, due to variations in copy number of a 61 bp heteroplasmic repeated motif in the control region.

Furthermore, 800 bp from the mtDNA genes ND2, COI and control region respectively was sequenced in 30 individuals from the Atlantic halibut broodstock at Mørkvedbukta Research Station. In addition, approximately 13420 bp from Common sole (*Solea vulgaris*), 15012 bp from European plaice (*Pleuronectes platessa*) and 7678 bp from Turbot (*Scophthalmus maximus*) were determined.

The overall genetic variation based on the number of nucleotide substitutions was greatest in Greenland halibut, followed by Pacific halibut and Atlantic halibut. The control region and some of the ND genes had the highest nucleotide diversity, while the rRNA genes and ATP8 (Pacific and Greenland halibut) were most conserved. Four amino acid substitutions between the three halibut species in ND5, with a negative mutation matrix score suggest that these substitutions could have an impact on functional and structural properties of the gene.

Phylogenetic investigations based on the complete mtDNA genome revealed that Atlantic halibut and Pacific halibut are closely related species, potentially separated at the sub-species level. Of all other available mtDNA genomes Greenland halibut is the closest relative to the *Hippoglossus* genus. Furthermore, Atlantic- and Pacific halibut was estimated to have separated 2 Ma, while divergence between Greenland halibut and the *Hippoglossus* genus took place approximately 6 million years ago.

Introduction

The Pleuronectidae family

The order of Pleuronectiformes contains 11 families and more than 700 species. Three families containing 16 species are usually represented along the Norwegian coast: Pleuronectidae, Soleidae and Scophthalmidae. The Pleuronectidae family consists of 39 genera and a total of 93 species world wide (www.fishbase.org). The three species being the main object in this thesis; Atlantic halibut (*Hippoglossus hippoglossus*), Pacific halibut (*Hippoglossus stenolepis*) and Greenland halibut (*Reinhardtius hippoglossoides*) all belong to the Pleuronectidae family.

Atlantic halibut (*Hippoglossus hippoglossus*)

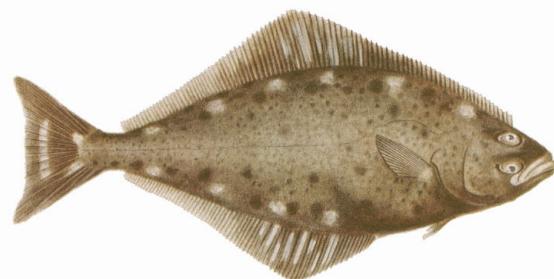
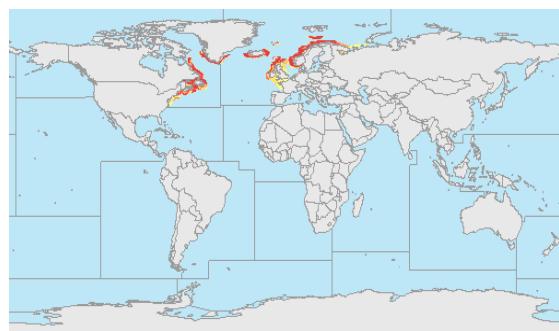


Figure 1: Distribution of the Atlantic halibut in the Atlantic ocean (left). A specimen of Atlantic halibut (right) (www.fishbase.org and www.fishlarvae.com)

Physical description

The Atlantic halibut is the largest fish naturally belonging to the Northern Atlantic Ocean. The largest known specimen was captured near Iceland. It was 3,65 m long and weighed 266 kg (Bethon 2005). It has a brown-greyish eyed side and white blind side. The lateral line is arched around the pectoral fin. The caudal fin is cut straight or slightly concave. Furthermore it is a narrower fish, relatively, than most flatfishes (Fig 1, right).

Distribution

The Atlantic halibut is found in the Northern Atlantic Ocean, in the eastern Atlantic from the coast of Labrador and Newfoundland to Cape Cod and far north on the west coast of

Greenland. Further along the coast of Iceland, from the Biscaya bay to Novaja Zemlja, including the islands of Svalbard (Bethon 2005) (Fig 1, left).

Biology

This fish prefers water at 0-10 °C. It is found on different kind of bottom substrate, and is able to hide in the substrate, with only it's eyes visible. The Atlantic halibut is also an excellent swimmer hunting at all depths, and able to migrate long distances. Important prey includes other fish, mollusca, arthropoda and other benthos. Both males and females mature sexually at the age of 12-13 years. Many details regarding spawning are still unknown. Along the Norwegian coast spawning has been registered at depths of 300-700 m, and as far north as Hammerfest. Other important spawning grounds are found near the Faeroe islands, at the Atlantic ridge, the Danish strait, the banks of Newfoundland and the David's strait. Spawning takes place in late winter. Eggs are dispatched and fertilized close to the bottom, but slowly rise in the water column. Eggs hatch after 9-16 days. The larvae go through metamorphosis when 3-3,5 cm long; the left eye moves over to the right side and pigmentation is increased on the right side. The fry gradually prefer deeper waters as they grow. (Bethon 2005).

Pacific halibut (*Hippoglossus stenolepis*)

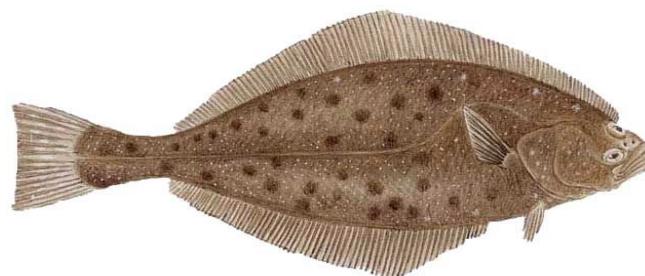


Figure 2: Distribution of Pacific halibut in the Pacific ocean (left). A specimen of Pacific halibut (right).
www.fishbase.org www.landbigfish.com

Physical description

The pacific halibut can reach a length of 267 cm and weigh up to 363 kg. Maximum observed age is 42 years (www.fishbase.org). It is coloured only on the right (top) side. The colour varies from mottled greenish to dark brown, depending on camouflage requirements. (Fig 2,

right) The underside is white. It has a concave tail, and its lateral line makes a large loop around the pectoral fin (Rodger 2005)

Distribution

Pacific halibut can be found throughout most of the marine waters of the North Pacific. It is distributed along the continental shelf of the North Pacific from Southern California to Nome, Alaska and along the coasts of Japan and Russia (www.fishbase.org).

Biology

Pacific halibut move from deep water at the edge of the continental shelf to shallower banks during the summer and return to deep water in the winter to spawn. This migration may involve hundreds of km. It is found on various types of bottoms. Young are found near shore, moving out to deeper waters as they grow older. Older individuals typically move from deeper water along the edge of the continental shelf where they spend the winter, to shallow coastal water (27-274 m) for the summer. Pacific halibut feed on fishes, crabs, clams, squids and other invertebrates (www.fishbase.org). Males mature sexually at the age of 8 years, while females mature after 12 years. Eggs are released pelagic at depths of 900-1200 m. They hatch after 16 days. When larvae are about 2 cm long the left eye begins to travel to the right side of the head, and they become bottom dwellers (Rodger 2005).

Greenland halibut (*Reinhardtius hippoglossoides*)

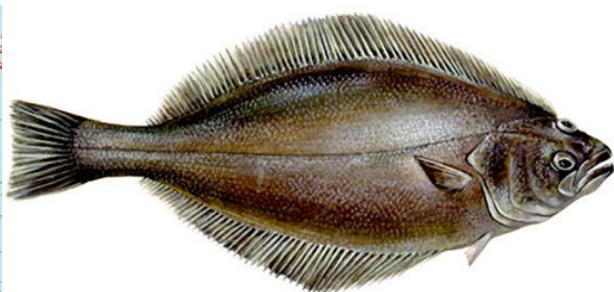
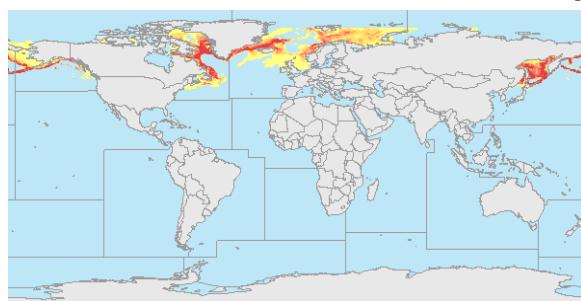


Figure 3: Distribution of Greenland halibut (left). A specimen of Greenland halibut (right). (www.fishbase.org and www.galizacig.com)

Physical description

The Greenland halibut can grow up to 1,2 m long and weigh 45 kg (Bethon 2005). Its left side is almost as dark as its right side. The lateral line does not make an arch around the pectoral

fin, but is almost straight from head to caudal fin. Its left eye does not move completely over to the right side during metamorphosis (Rodger 2005). This feature makes it easy to separate Greenland halibut from Atlantic- and Pacific halibut.

Distribution

Both the Atlantic and the Pacific Ocean are home to Greenland halibut. In the Atlantic Ocean the species is found off the North American and Greenland coasts, in the Iceland / Faroes area, and in the eastern Norwegian and Barents Seas (Fig 3, left) (Andriyashev 1954). There is a subspecies in the Pacific Ocean, from the Bering Sea to northern Japan and southern California; *R.h . matsuurae*. They are separated at the subspecies level (Fairbarn 1981).

Biology

The Greenland halibut is a eurytopic species, adapted to life in a wide range of environmental conditions. It is found at a depth of 20 m to 1200 m, in temperatures between –1 and 10 °C (Fedorov 1971). It is most abundant where bottom temperatures range from –1 to 3 °C (Templeman 1973). This flatfish is benthopelagic; it spends time both close to the bottom and swimming in the pelagic. It can adopt a vertical swimming position when it moves freely in the pelagic (De Groot 1970). Sexual maturation for males and females takes place at the age of 5 and 10 years, respectively. The main spawning grounds in the Norwegian / Barents Sea are at depths between 600 and 900 m along the continental slope between 70°N and 75°N. The Davis Strait near Greenland is also an important spawning area. Eggs and larvae are pelagic. Eggs are 3,8-4,3 mm in diameter (Godø and Haug 1989). Metamorphosis is unusually slow, and takes place when the fry is 6-9 cm long. Greenland halibut in the Norwegian and Barents Sea and halibut close to Iceland and Greenland are considered to belong to different populations (Bethon 2005).

Population studies of Atlantic halibut

Joensen (1954) and Mccracken (1958) where the first who suggested that Atlantic halibut may consist of more than a single population. Mork and Haug (1983) conducted a study seeking to find genetic differentiation in halibut from three different localities along the Norwegian coast using tissue enzymes as a marker. Their study revealed no significant differences. Grant et al. (1984) also concluded with some stock differentiation. Then Haug and Fevolden (1986) screened 43 loci in halibut from Greenland and three different localities in Norwegian waters. A possible heterogeneity was reported between the southernmost and the two northernmost

localities. Foss et al. (1998) found halibut from Greenland, Faroes and Iceland to be similar in terms of allele frequencies. Whereas halibut from Honningsvåg (northern Norway) deviated significantly. Reid et al. (2005) investigated population structure in Atlantic halibut using 18 microsatellite markers (see “molecular markers” below). The analysis revealed no significant differentiation between samples, although uncertainties surrounding Atlantic halibut reproductive behaviour made it difficult to ascertain that only a single breeding population had been sampled at each location.

This study is by no means a population study. A population study would have to include more individuals from a wider geographic area. But mitochondrial DNA is a marker well suited for population studies, and information found in this study will be useful in future investigation of population structure.

Organelles: Mitochondria and plastids

Organelles are structures in the cytoplasm with a specific function surrounded by a membrane. Mitochondria and plastids are two of many organelles found in the eukaryotic cell. Mitochondria are found in both plant and animal cells, whereas plastids are found in plants and algae. Plastids have the ability to differentiate and redifferentiate. The most known form is probably chloroplasts (Mathews et al. 2000).

Mitochondria are surrounded by a smooth outer membrane and a membrane on the inside that is folded into cristae. The cristae contain a set of enzymes for oxidation of fatty acids and oxidative phosphorylation of ADP into ATP. Their task is to transform energy in organic molecules from digested food items into energy in the form of ATP, the most common energy source for energy-requiring reactions in the cell (Pollard and Earnshaw 2004).

The organization of organelle genomes

Mitochondria and chloroplasts both possess DNA genomes that code for RNA species and some of the proteins involved in the functions of the organelle. Different kinds of RNA are also known to be imported into mitochondria (Entelis et al. 2001). Acceptance of the fact that mitochondria contain their own genome followed the demonstration that circular DNA was isolated from mitochondrial fractions of cells (Van Bruggen et al. 1966). Short thereafter

circular DNA molecules were isolated from mitochondria from other vertebrates and some invertebrates.

In some lower eukaryotes the mitochondrial DNA (mtDNA) is linear, but more usually it is a single circular molecule of DNA. So far only a few linear mitochondrial DNA have been isolated. One of the exceptions is the mtDNA from the cnidarian *Hydra attenuata*, which consists of two unique 8-kb linear DNA molecules (Warrior and Gall 1985). Chloroplast DNA falls in the range of 120-200 kb, whereas mtDNA varies enormously in size. In animals it is relatively small, usually less than 20 kb. But in plants mtDNA can be as big as 2000 kb (Primrose and Twyman 2003).

Mitochondria are essential components of all eukaryotic cells. Despite the great diversity of genome organization, they have the same function in every organism: Cellular respiration and oxidative phosphorylation. Many of the genes for mitochondrial proteins and RNAs are found in the nucleus. The mechanisms behind the transport of proteins into mitochondria are still under debate. An enzyme called translocase is known to recognise and translocate proteins into mitochondria. There are different translocases in the outer and inner membrane. So far six different translocases have been identified (Mokranjac and Neupert 2005). Some organisms use the standard genetic code to translate nuclear mRNAs, and a different code for their mitochondrial mRNAs (Primrose and Twyman 2003).

Vertebrate mitochondrial genomes

MtDNA is a much used molecular marker because it is present in much higher copy number than nuclear DNA (Michaels et al. 1982). However, with todays Polymerase Chain Reaction (PCR) obtaining enough material for analysis is usually not a problem, even for nuclear genes. A typical somatic cell contains 500–1000 mitochondria, each with a few DNA molecules. Whereas an oocyte has a much higher concentration of mitochondria. (May-Panloup et al. 2005) measured an average of 256000 mitochondrial genomes in human oocytes. As mentioned above, mtDNA is a relatively small genome. It is mainly maternally inherited (Hutchison et al. 1974). MtDNA is generally considered to lack recombination. However, recent work in several fields has questioned the validity of this characteristic. Studies of natural populations of several animal species have provided convincing evidence for the presence of mtDNA recombination in at least some taxonomic groups (Burzynski et al. 2003; Hoarau 2002; Lunt and Hyman 1997). Although there is no doubt that certain animal

mitochondrial genomes do recombine, we remain basically ignorant as to how common mtDNA recombination might be in the animal kingdom. The mitochondrial genome has a higher mutation rate than the nuclear genome, perhaps due to repair inefficiency (Parsons et al. 1997). This elevated rate of evolution is 5-10 times faster than single copy nuclear genes (Brown et al. 1979)

An mtDNA molecule is regarded as a haplotype and is treated as a single locus. Since mtDNA is maternally inherited, most copies are identical (i.e. monoclonal) in an individual. However, it is possible that more than one mtDNA type is present in an individual. This phenomenon is called heteroplasmy and it is now thought that all individuals are heteroplasmic, at least at some level. Differences can be found among single nucleotides or in the number of copies of a repeated motif. Individuals may have more than one mtDNA type in a single tissue or specific differences may be located in specific tissues. Heteroplasmy in one tissue sample and homoplasmic in another tissue sample is also a possibility. Length heteroplasmy in a homopolymeric stretch (i.e., C stretches), is more common than point heteroplasmy. See table 31 for an overview of heteroplasmy in fish.

Vertebrate mitochondrial genomes normally carry genes encoding 13 proteins: Seven NADH dehydrogenase subunits, three cytochrome oxidase subunits, ATPase 6, ATPase 8 and cytochrome B. The genome also consists of 22 transfer RNA genes (tRNAs) and two ribosomal RNA genes; 12 S and 16 S, necessary for translation of the proteins encoded by mtDNA. The proteins are involved in electron transport and ATP synthesis (Billington 2003). In addition a non-coding control region, known as displacement loop (D-loop), is present. The d-loop is only a short region of the whole control region. The non-coding control region is responsible for size variation in the mitochondrial genome due to differences in length and copy number of short repeats in various groups (Brown et al. 1996; Gissi et al. 1998; Ursing and Arnason 1998).

One strand is known as the light (L) strand and the other one as the heavy (H) strand. The two strands are called heavy and light because there is an imbalance in nucleotide composition in the two strands. The heavy strand has a higher content of purines (adenine and guanine), while the light strand is richer in pyrimidines (tymine and cytosine). Purines are heavier than pyrimidines. The majority of the genes are transcribed from the H-strand. The replication origin for the heavy-strand (O_H) is located in the control region, whereas the replication origin of the light strand (O_L) is located in the WANCY cluster between the tRNAs for asparagine and cysteine

(Fig 4). The WANCY cluster is a short region containing five tRNA genes: Trp (W), Ala (A), Asn (N), Cys (C) and Tyr (Y).

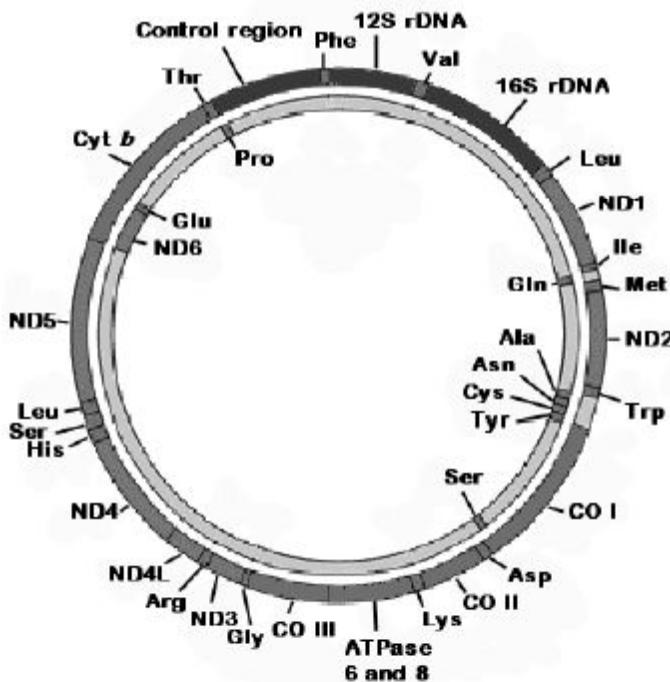


Figure 4: Schematic representation of the circular molecule of the “conserved” vertebrate mitochondrial genome organisation. Genes outside and inside the circle are transcribed from the H and L strands, respectively. Protein coding genes are presented as follows: Cyt B – cytochrome b; CO I, CO II, CO III – subunits of the cytochrome oxidase; ND1-6 – subunits 1 to 6 of the NADH reductase. tRNAs are presented by their three letter amino acid abbreviations. Figure from (Pereira 2000)

A great number of phylogenetic studies using mitochondrial gene sequences have been reported. For example, the control region is often used in population studies due to the high variability in its nucleotide sequence. While protein-coding genes, such as cytochrome b are generally used for phylogenetic analysis of taxa above the species level (Pereira 2000).

Transmembrane helices in mitochondrial proteins

Embedded within the inner membrane of mitochondria are the proton carriers, primarily cytochromes, which constitute the respiratory chain. There are totally eight of these carrier complexes. They are assembled in the form of multiprotein complexes. ND1-ND6 and ND4L are 7 of a total of 25 subunits of the NADH dehydrogenase (Complex I). NADH pumps protons into the intermembrane space of the mitochondrion, thus passing energy on to

complex III. Complex II transfers energy from FADH₂ to complex III. Cyt-b is one of 9-10 polypeptides in complex III. COI-COIII along with 10 other polypeptides make up cytochrome oxidase (complex IV) that pumps protons into the intermembrane region of mitochondria. ATPase6 and ATPase8 are subunits of FoH⁺-ATPase, together with 10-12 other polypeptides (Complex V) which catalyses ATP synthesis from ADP (Mathews et al. 2000). These proteins will be further characterized by determining the number of transmembrane helices in protein coding genes in Atlantic-, Pacific- and Greenland halibut.

Organization of the fish mitochondrial DNA

Until about 1995 one believed that gene order was conserved among all fish mtDNA. The reason for this is that the first complete mitochondrial genome sequences found in vertebrate taxa had no variation in the position of the genes along the molecule. This had been observed in taxa as diverse as humans and other mammals (Anderson et al. 1981; Anderson et al. 1982; Árnason and Gullberg 1993; Gadaleta et al. 1989), *Xenopus* (Roe et al. 1985), and some fish species (Johansen et al. 1990; Tzeng et al. 1992; Zardoya et al. 1995). But in 1995 sea lamprey was reported to have a translocation of tRNA Gln – Cytb and in the location of the non coding regions (Lee and Kocher 1995). The same rearrangement was also found in river lamprey (Delarbre et al. 2000). Miya and Nishida (1999) reported an example of tRNA gene rearrangement in bony fishes between the ND6 and the control region (ND6 - Cyt b - tRNAGlu - tRNAPro - tRNAThr - control region). This changed the “conserved” status of mtDNA in fishes. Until this study three Pleuronectiform species were available from Gene Bank. No deviation from the “conserved vertebrate gene order” were found in these species.

As reviewed by Pereira (2000) rearrangements associated with tRNA genes have been reported in birds, lizards, crocodilians, marsupial mammals, snakes and tuatara. Tandem duplication and multiple deletion events associated with tRNA genes appear to be the most probable mechanisms for new gene rearrangements seems. Placental mammals, turtles, fishes, some lizards and the frog *Xenopus* have the most conserved mitochondrial gene order (Pereira 2000).

The genetic code for vertebrate mitochondrial DNA

As seen in Table 1 it is possible to have a change in the DNA sequence without any corresponding change in the gene product; the polypeptide. The reason for this is that all

amino acids are encoded by more than one codon (Table 1). The 3rd position (synonymous position) within a codon is often a fourfold degenerate site. So a substitution in this position will have no effect on the amino acid level. A change in one of the nucleotides in the 1st codon position will in most cases change the amino acid. The exception is two tRNA genes that read different codons. tRNA^{Leu(UUR)} reads the codons TTA and TTG. While tRNA^{Leu(CUN)} reads the codons CTT, CTC, CTA and CTG. tRNA^{Ser(UCN)} reads TCA, TCG, TCC and TCT, and tRNA^{Ser(AGY)} reads AGC or AGT.

Table 1: The vertebrate mitochondrial genetic code. (<https://bioinformatics.org/JaMBW/2/3/TranslationTables.html>)

Amino acid	Codon						
Phe (F)	TTT	Ser (S)	TCT	Tyr (Y)	TAT	Cys (C)	TGT
Phe (F)	TTC	Ser (S)	TCC	Tyr (Y)	TAC	Cys (C)	TGC
Leu (L)	TTA	Ser (S)	TCA	STOP	TAA	Trp (W)	TGA
Leu (L)	TTG	Ser (S)	TCG	STOP	TAG	Trp (W)	TGG
Leu (L)	CTT	Pro (P)	CCT	His (H)	CAT	Arg (R)	CGT
Leu (L)	CTC	Pro (P)	CCC	His (H)	CAC	Arg (R)	CGC
Leu (L)	CTA	Pro (P)	CCA	Gln (Q)	CAA	Arg (R)	CGA
Leu (L)	CTG	Pro (P)	CCG	Gln (Q)	CAG	Arg (R)	CGG
Ile (I)	ATT	Thr (T)	ACT	Asn (N)	AAT	Ser (S)	AGT
Ile (I)	ATC	Thr (T)	ACC	Asn (N)	AAC	Ser (S)	AGC
Met (M)	ATA	Thr (T)	ACA	Lys (K)	AAA	STOP	AGA
Met (M)	ATG	Thr (T)	ACG	Lys (K)	AAG	STOP	AGG
Val (V)	GTT	Ala (A)	GCT	Asp (D)	GAT	Gly (G)	GGT
Val (V)	GTC	Ala (A)	GCC	Asp (D)	GAC	Gly (G)	GGC
Val (V)	GTA	Ala (A)	GCA	Glu (E)	GAA	Gly (G)	GGA
Val (V)	GTG	Ala (A)	GCG	Glu (E)	GAG	Gly (G)	GGG

There are four distinct differences between the mitochondrial genetic code and the standard genetic code (Table 2)

Table 2: Differences between the mitochondrial genetic code and the standard genetic code

	Mt genetic code	Standard genetic code
Codon	Amino acid	Amino acid
AGA	STOP	Arg (R)
AGG	STOP	Arg (R)
AUA	Met (M)	Ile (I)
UGA	Trp (W)	STOP

Codon usage bias

Synonymous codons (i.e. codons coding for the same amino acid) are often not used with equal frequency. This phenomenon is known as codon bias. There has been a debate concerning the functional role of codon bias, namely, does codon bias maximize translational efficiency (elongation rates) or translational accuracy (reducing misincorporation of incorrect amino acids). Different kind of organisms apparently have evolved unique genomic signatures (Bernardi and Bernardi 1986). It has been proposed that each species is subjected to specific genomic pressures on base composition, in turn resulting in a distinctive bias in codon choice (Grantham et al. 1980). In principle, biases in nucleotide composition and codon usage can result from natural selection and/or differential mutational pressure. In many organisms (*E. coli*, *D. melanogaster* and *C. elegans*) codons that use abundant tRNAs are selectively favoured (Gouy and Gautier 1982; Powell and Moriyama 1997; Stenico et al. 1994). Codon selection of this type most likely functions to optimise translational speed and/or translational accuracy, although additional factors, such as transcription efficiency and mRNA secondary structure could also exert selection pressures (Xia 1996; Zama 1990). It was found for many *E. coli* genes that the degree of bias in codon choice is directly related to the expression level of a particular gene (Klump and Maeder 1991). In mammals codon bias seems to be closely related to mutation pressure (Sharp et al. 1993). Furthermore, in a recent study in humans Kotlar and Lavner (2006) concluded that selection acts to enhance translation efficiency in highly expressed genes by preferring some codons, and acts to reduce translation rate in lowly expressed genes by preferring a different set of codons. Codon usage bias will be thoroughly investigated in the three halibut species.

Molecular markers

Repeated sequences of short motifs, usually 1-13 bp long, are known as microsatellites. Such sequences are probably created when DNA helicase slips back a few bases, so that the same bases are inserted twice in the new strand. The number of copies of the repeated motif changes fast. This is the reason why microsatellites are used for population studies. They can reveal information about the history and geographical dispersal of populations. The Polymerase Chain Reaction is usually used to amplify microsatellites (Jarne and Lagoda 1996; Queller et al. 1993).

RFLPs (Restriction fragment length polymorphism) are fragments of restricted DNA (i.e. DNA cut at specific recognition sites with a restriction enzyme) and then separated according

to size with gel electrophoresis. A specific DNA sequence will yield a specific pattern after gel electrophoresis. Different sequences may lack or have present different endonuclease recognition sites, thus creating different patterns (Baxevanis and Francis Ouellette 2001).

The methods mentioned above are based on differences in DNA sequence. In this method the products of the DNA sequence are used, namely proteins. Allozymes are different alleles of proteins. They are separated by electrophoresis because of differences in charge. Isozymes have been used for the same purpose. Isoenzymes have differences in the amino acid sequence, but target the same chemical reaction. Only a fraction of the genetic variation that is present can be detected by the use of allozymes. The reason being that synonymous substitutions will not cause changes in the amino acid sequence that makes up the protein. The lack of resolution is this methods main drawback (Higgs and Attwood 2005; Kephart 1990; May 1992).

Single Nucleotide Polymorphisms (SNP) are variations in one nucleotide position. This molecular marker is one of the new tools used for example in population studies. Both mitochondrial and nuclear genes are screened for SNPs (Vignal et al. 2002). Introns are DNA regions that have no known function. They are mainly found in the nuclear genome. Introns evolve faster than the rest of the gene and intron-SNPs are therefore interesting as a marker in population studies, e.g. Li et al. (2006).

Sequencing of DNA and even complete mitochondrial genomes has in the recent years become a routine, e.g. Miya et al. (2001). DNA sequences have the potential of being a powerful molecular marker. Analysis using DNA sequences often provide higher resolution and accuracy than the methods mentioned above. This is the reason why DNA sequencing was chosen in this study

Phylogenetic analysis

A phylogenetic analysis tries to explain evolutionary relationships. In molecular phylogenetics, likenesses and differences, usually in DNA or amino acid sequences, are used to create a phylogenetic tree. In such a tree a clade is a monophyletic taxon that include the most recent common ancestor and all of the descendants of that most recent ancestor. A taxon is simply a named group of organisms. It is not necessarily a clade. The length of each branch

correspond to the number of changes that has occurred (Baxevanis and Francis Ouellette 2001).

An alignment were homologous nucleotides are compared is the basis from which a phylogenetic tree is created. A tree based only on observations within a single gene should be referred to as a gene tree, not a species tree. A gene tree represents the evolutionary history of the gene, but the history of the whole species might be different. To created a trustworthy species tree it is required to use data from multiple genes (Krane and Raymer 2003).

In an unrooted tree you only get information about the relationship between the nodes. While in a rooted tree evolution can be followed from the most recent common ancestor of all taxa included in the tree. Trees are usually rooted by choosing one or more uncontroversial species as an outgroup (Krane and Raymer 2003).

Molecular data can be used in two ways to generate phylogenetic trees: Character based methods and distance based methods. Distance methods summaries the differences between the sequences and use the amount of change between sequences to derive a tree. A distance method is only able to tell the true phylogenetic story if all genetic divergence events were accurately recorded in the sequence (Swofford et al. 1996). The biggest advantage is that distance methods require far less computer power (Baxevanis and Francis Ouellette 2001).

Neighbor Joining starts with a tree were all species come a single central node. Neighbors are then calculated so that the least total branch length is found. This method is far from optimal, but usually creates close to the optimal tree (Saitou and Nei 1987).

The minimum-evolution (ME) method searches for the optimal tree by finding the one with the smallest sum of branch lengths as estimated under the least-squares criterion, were negative branch lengths is not allowed (Rzhetsky and Nei 1992).

The character based methods use character data at all steps in the analysis. Maximum parsimony is based on the principle that the simplest method is the best. A MP tree is the one that requires the fewest changes to explain the inferred topology (Swofford et al. 1996). To do this, all trees are given a length, equal to the minimum number of transitions which can explain the observed distribution of character states across taxa assuming the tree. The tree with the shortest length is the maximum parsimony tree.

Long branch attraction, as reviewed by Bergsten (2005), is a problem that arises when sequences that evolve rapidly and are not closely related, cluster together. When the nucleotide substitution rate is high there is a possibility that the same nucleotide will occur at the same position, simply by chance. This is possible because there are only four different nucleotides to choose between. The problem seems to be greatest when using the MP method. By employing the maximum likelihood method this problem is reduced. It is also possible to avoid the problem by adding taxa that are related to those taxa with long branches.

Bootstrapping is a way of determining the confidence of a taxon. A part of the original sequence from each species is drawn and a new, pseudoreplicate tree is made from this data set. The process is repeated multiple times, in this study 2000 times. This is a way of testing the reliability of the tree topology (Hall 2004). The method was invented by Efron (1979). It has been suggested that a bootstrap of more than 70 % correspond to a probability of greater than 95 % that the true phylogeny has been found (Hillis and Bull 1993). Though, under less favourable conditions, a bootstrap greater than 50 % will be an overestimate of the accuracy (Hillis and Bull 1993). So conditions of the analysis must be considered.

As a general rule, if a data set yields similar trees when analysed by the fundamentally different distance matrix and parsimony methods, that tree can be considered fairly reliable (Krane and Raymer 2003).

The patterns of substitution in protein coding genes are well known. The knowledge of how mutations accumulate makes protein coding genes good candidates for phylogenetic studies in fish. Mutations in third (and rarely in first) positions of codons that do not result in amino acid substitutions occur at higher rates than substitutions causing amino acid replacements.

Transitions (purine→ purine or pyrimidine → pyrimidine) in third codon position are most frequently observed, followed by transversions (purine ↔ pyrimidine) in third codon positions and silent transitions in some first codon positions (Meyer 1994). Among distantly related species, transitions in third positions will probably experience substitutional saturation, and will be unreliable for revealing evolutionary descent (Bakke and Johansen 2005). Transversions are less common and have a higher chance of being reliable indicators of descent. Second positions of codon are most conserved and contain phylogenetic information among distantly related species.

It is difficult to resolve phylogenetic relationships among the Pleuronectiformes due to their reduced morphology (Berendzen and Dimmick 2002). Traditionally eye position has been an important character in classification of flatfishes. Flounders and soles have both been divided into right- and left-hand taxa (Norman 1934). Chapleau (1993) concluded that sidedness had been derived multiple times within flatfishes and that eye position was not a good indicator of relationships within the group.

Berendzen and Dimmick (2002) used 12S and 16S mitochondrial ribosomal DNA to elucidate relationships within the Pleuronectiformes. Of the taxa included in the present thesis Soleidae and Scophthalmidae was recognized as the most basal groups, while Pleuronectidae was the most derived group. Pacific halibut was the only Pleuronectidae species from the present thesis that was also included in the study by Berendzen and Dimmick (2002). Atlantic halibut and Pacific halibut are recognized as closely related species, but this knowledge is not based on DNA sequence information (Nelson 1994).

Molecular clocks

The molecular clock hypothesis arised some four decades ago, when it was discovered that divergence in nucleotide and amino acid sequence often increase linearly as time goes by Zuckerkandl and Pauling (1962). The divergence rate vary from gene to gene, and synonymous positions evolve faster than nonsynonymous ones. Fossils records are used to estimate divergence rates and thus calibrate the molecular clock. But fossils tend to underestimate divergence times (Hedges 2002). Geological events like the rise of the Isthmus of Panama have also been used as calibration points e.g. Domingues et al. (2005). Ideally multiple calibration points should bee used. The rise of the Isthmus of Panama was chosen to calibrate the molecular clock that was used to estimate time of divergence for Atlantic-, Pacific- and Greenland halibut.

Even if homologous genes are considered, divergence rate may differ from taxa to taxa. Therefore relative rate test have been developed to detect differences in evolutionary rate among lineages, e.g. Tajima (1993). Still, divergence times are only rough estimates that depend highly on the accuracy of the calibration point and conservation of divergence rates among taxa.

Aims of the study

- Determine the complete mitochondrial DNA sequence from 4 individuals of Atlantic halibut, Pacific halibut and Greenland halibut, respectively.
- Determine part of the mtDNA sequence from Common sole, Turbot and European plaice for use as a molecular marker sequence in phylogenetic analysis.
- Investigate intraspecific and interspecific genetic variation in Atlantic halibut, Pacific halibut and Greenland halibut, both at the nucleotide- and amino acid levels.
- Contribute to phylogenetic relationship determination among Atlantic halibut, Pacific halibut, Greenland halibut, Common sole, Turbot and European plaice.
- Investigate the possibility of creating a test for easy and fast species identification based on the mtDNA molecular marker.

Materials

Table 3 displays which species were included in this study.

Table 3: English and Latin names of species from which mtDNA was completely or partially determined.

Common name	Scientific name
Atlantic halibut	<i>Hippoglossus hippoglossus</i>
Pacific halibut	<i>Hippoglossus stenolepis</i>
Greenland halibut	<i>Reinhardtius hippoglossoides</i>
Common sole	<i>Solea vulgaris</i>
Turbot	<i>Scophthalmus maximus</i>
European plaice	<i>Pleuronectes platessa</i>

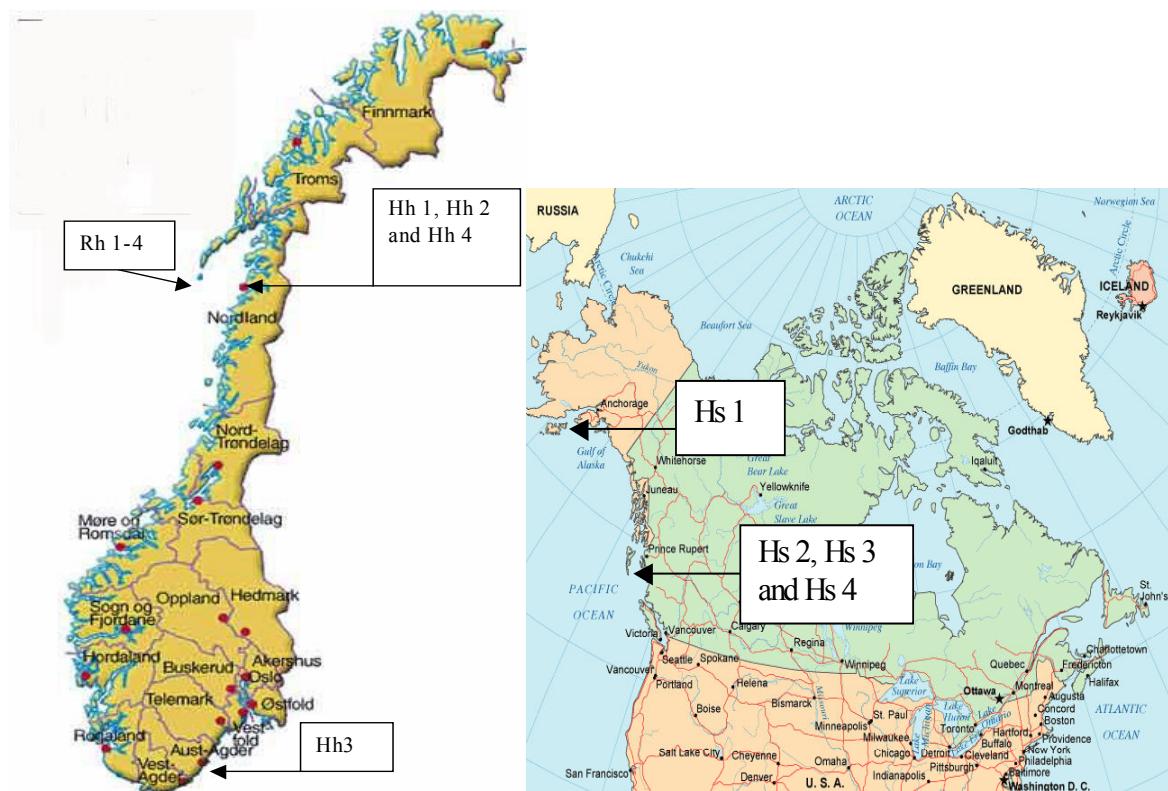


Figure 5: Location where the individuals of Atlantic halibut (Hh), Pacific halibut (Hs) and Greenland halibut (Rh) were caught. (www.tagkompaniet.se and http://encarta.msn.com/map_701511416/Canada.html)

Atlantic-, Greenland- and Pacific halibut

Figure 5 displays the location where the different individuals were caught. Atlantic halibut number one and two were born in captivity at Mørkvedbukta Research Station . Their parents were wild caught at an unknown location, but probably not very far from Bodø (N67 16 20,

E14 22 04). Atlantic halibut number three was wild caught in the waters close to Bodø. A sample from Atlantic halibut number four were obtained from Risør aquarium in southern Norway (N58 42 34, E9 13 54). All four individuals of Greenland halibut were caught off the coast of the island Røst (N67 31 39, E12 06 39). Pacific halibut number one were caught outside Cook Inlet, about 37 km off Homer (N 59 35 41 W152 06 18), Alaska. Pacific halibut 2-4 were caught in Hecate Strait (N52 53 40, W130 38 10) at the west coast of Canada.

30 individuals of Atlantic halibut

Approximately 800 bp of the ND2, COI and control region respectively was sequenced in 30 individuals from the Atlantic halibut broodstock at Mørkvedbukta Research Station. The first eight individuals originates from Tysfjord (N68 04 11, E16 13 66). They were bought from local halibut farmers, and the relationship among them is not known. Individuals 9-23 are born at Mørkvedbukta Research Station in 2001. They are possibly siblings or half-siblings, but this information is uncertain. The last six individuals were wild-caught in 2005 at an unknown location, but probably close to Bodø (N6716201, E1422043)

Common sole, Turbot and European plaice

Samples of Common sole and Turbot were obtained from Risør aquarium (N58 42 34, E9 13 54). The European plaice was caught close to Bodø (N6716201, E1422043). Approximately 13420 bp from Common sole, 15012 bp from European plaice and 7678 bp from Turbot were determined

Table 4 presents the species, from which mtDNA sequences were used in the different alignments and in the phylogenetic studies.

Table 4: Species from which mtDNA was used in phylogenetic studies.

Common name	Latin name	Accession number
Atlantic halibut	<i>Hippoglossus hippoglossus</i> 1	This study
Atlantic halibut	<i>Hippoglossus hippoglossus</i> 2	This study
Atlantic halibut	<i>Hippoglossus hippoglossus</i> 3	This study
Atlantic halibut	<i>Hippoglossus hippoglossus</i> 4	This study
Pacific halibut	<i>Hippoglossus stenolepis</i> 1	This study
Pacific halibut	<i>Hippoglossus stenolepis</i> 2	This study
Pacific halibut	<i>Hippoglossus stenolepis</i> 3	This study
Pacific halibut	<i>Hippoglossus stenolepis</i> 4	This study
Greenland halibut	<i>Reinhardtius hippoglossoides</i> 1	This study
Greenland halibut	<i>Reinhardtius hippoglossoides</i> 2	This study
Greenland halibut	<i>Reinhardtius hippoglossoides</i> 3	This study
Greenland halibut	<i>Reinhardtius hippoglossoides</i> 4	This study
European plaice	<i>Pleuronectes platessa</i>	This study
Turbot	<i>Scophthalmus maximus</i>	This study
Common sole	<i>Solea vulgaris</i>	This study
Stone flounder	<i>Platichthys bicoloratus</i>	AP002951
Bastard halibut	<i>Paralichthys olivaceus</i>	AB028664
Spotted halibut	<i>Verasper variagatus</i>	DQ403797
Cod	<i>Gadus morhua</i>	X99772
Haddock	<i>Melanogrammus aeglefinus</i>	DQ020497
Alaska pollock	<i>Theragra chalcogramma</i>	AB094061
Atlantic salmon	<i>Salmo salar</i>	U12143
Rainbow trout	<i>Oncorhynchus mykiss</i>	L29771
Arctic charr	<i>Salvelinus alpinus</i>	AF154851
American gizzard shad	<i>Dorosoma cepedianum</i>	NC_008107
Naked shellear	<i>Cromeria nilotica</i>	NC_007881
Korean rockfish	<i>Lycodes toyamensis</i>	NC_004409
	<i>Sebastes schlegeli</i>	NC_005450

Methods

Various kits have been used in this study. The protocol provided by the manufacturer was followed most of the time. Any deviation from the manufacturers protocol is mentioned below. Recipes for buffers, medium and agar plates are found in the manufacturers protocol provided with each kit and also available as pdf-files from each manufacturers home page.

DNA extraction

Total DNA was extracted from red muscle with a kit from Roche (cat.# 11814770001). The manufacturer's instructions were followed. DNA was eluted in double sterilized water or TE-buffer, and stored at -20° celsius.. Concentration of DNA after elution was determined using OD₂₆₀ readings from a Eppendorf biophotometer (cat # 952000004). DNA purity was determined using the optical density ratio between 260 and 280 nm measurements. This is based on the fact that OD at 260 nm is twice that at 280 nm if the solution contains pure DNA. Clean DNA has a OD-260/OD-280 between 1.8 and 2.0.

Polymerase chain reaction (PCR)

There are three basic steps in PCR. First, the target genetic material must be denatured—that is, the strands of its double helix must be unwound and separated—by heating to 90–96°C. The second step is hybridisation or annealing, in which the primers bind to their complementary bases on the now single-stranded DNA. The third is DNA synthesis by a polymerase. Starting from the primer, the polymerase reads the template strand in the 3'-5' direction and match it with complementary nucleotides very quickly. The result is two new double helices in place of the first, each composed of one of the original strands plus its newly assembled complementary strand. This cycle is repeated typically 30–35 times. PCR is performed in a thermal cycler which is programmed to change the temperature between the respective steps in each cycle.

PCR conditions used:

Component	Amount
10x PCR buffer	5 ul
DNTP	8 ul
L-primer (10 uM)	1 ul
H-primer (10 uM)	1 ul
LA-taq polymerase (Takara)	0,5 ul
Template	100-200 ng
DEPC water	X ul
Total volume	50 ul

The template was total DNA.

Cycle parameters where depending on the length of the fragment to be amplified:

Fragments of 4-5 kb		
Initial denaturation	94 °C	3 min
Denaturation	94 °C	1 min
Annealing	48 °C	1 min
Extension	72 °C	4 min
Number of cycles	15	
Denaturation	94 °C	1 min
Annealing	53 °C	1 min
Extension	72 °C	4 min
Number of cycles	15	
Final extension	72 °C	10 min

The Heteroplasmic Tandem Repeat (HTR) region is located in the control region and is approximately 1 kb long.

HTR region, 1kb		
Initial denaturation	94 °C	3 min
Denaturation	94 °C	1 min
Annealing	53 °C	1 min
Extension	72 °C	1 min
Number of cycles	30	
Final extension	72 °C	10 min

PCR primer pairs

The primer combinations used to amplify mtDNA from the four halibut individuals are presented in table 5. L indicates the light strand and H the heavy strand. Specific primers were designed based on an alignment of mtDNA from *Platichthys bicoloratus* and *Paralichthys olivaceus*. These are species that belong to the Pleuronectiformes. Primer

positions were based on the *Platichthys bicoloratus* and *Paralichthys olivaceus* mtDNA. The entire mitochondrial genome in Atlantic-, Pacific- and Greenland halibut was amplified in fragments of 3,5-5 kb, which partially overlapped each other.

Table 5: Primer pairs used to amplify mtDNA in Atlantic halibut, Pacific halibut and Greenland halibut.

Primer pair	Fragment	Sequence 5'-3'
L466	1	GTAGCTCTACTCATCCTGAA
H3977		AAGTGGTGTAGAGGAAGCAC
L3850	2	CGCTGGTCTCCCACCACAGCT
H7740		AGTACAAAATTGGATGGATGG
L7104	3	ATCAATTGGTTCAAGCCAA
H10010		GACCTCCTTGCAATTCTCG
L9615	4	TACTGATGAGGATCTTAATC
H13882		GGTAGGTTAGAGGATGCAATGG
L13125	5	CAACCCCAACTAGCATTCTTC
H590		ATATGGAGTGTATAGTGCAA

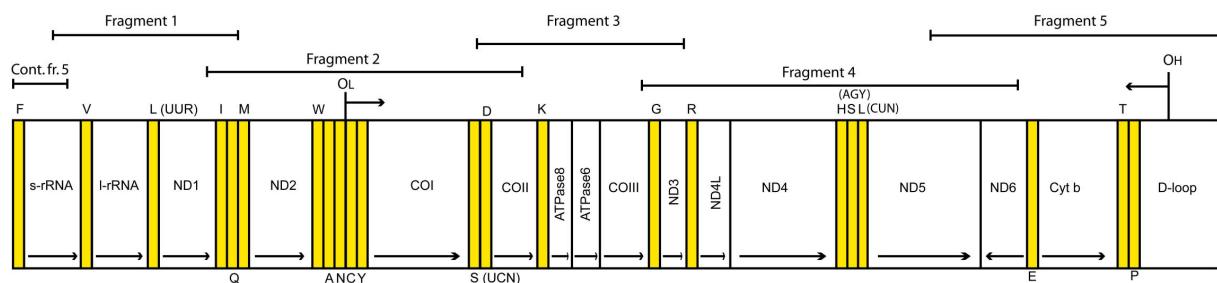


Figure 6: A schematic view of the mtDNA. Regions amplified in Atlantic halibut, Pacific halibut and Greenland halibut are presented above the mtDNA. The tRNA genes are coloured yellow and are indicated by one letter amino acid code. All tRNAs on the upper line are coded on the light strand. Arrows within genes indicates from which strand the genes are coded. Arrows pointing towards right indicated light strand, while genes encoded by heavy strand points towards left.

In some cases the primer combination presented in table 5 gave no PCR product in Common sole, Turbot and European plaice. Alternative primer combinations are presented in table 6.

Table 6: Alternative primer combinations used to amplify mtDNA from Common sole, Turbot, American plaice and European plaice, in case no PCR was produced with the primer pairs in table 3.

Primer pair	Fragment	Sequence 5'-3'
L1249	1	CGCAAGGGAAAGCTGAAA
H3058		CTACCCCTAACCTCCAAAGC
L2642	2	GTCCTACGTGATCTGAGTTC
H6200		AAAGAATCAGAATAGGTGTT
L5572	3	TCGAGCAGAGCTAAGTCAAC
H8589		CTACCCCTAACCTCCAAAGC
L8329	4	CTACCCCTAACCTCCAAAGC
H10433		CTACCCCTAACCTCCAAAGC
L9615	5	TACTGATGAGGATCTTAATC
H11931		CTAAGACCAACGGATGAGCT
L11812	6	AAACACTAGATTGTGATTCT
H13882		GGTAGGTTAGAGGATGCAATGG
L13125	7	CAACCCCAACTAGCATTTCTTC
H16200		GAGAACCCCTTACCGCTGGAGTGAAC
L15662	8	CTACCCCTAACCTCCAAAGC
H1838		GTCCGTTCCGATTACAC

To amplify the HTR region of Atlantic halibut and Pacific halibut primers L 16276 and H 17250 was used (Table 7). For Greenland halibut primer combination L 16604 – H 17300 was used.

Table 7: Primer pair used to amplify the HTR region in Atlantic halibut, Pacific halibut and Greenland halibut.

Primer pair	Sequence 5'-3'
L 16276	CGGAGACGTTAAAGGGT
H 17250	CTTGAAATAATCTAACAGTA
L 16604	ACATAAAGGGATATCATGTGC
H 17300	GTTGGGATGATGCAAGAAAT

Table 8 shows sequencing primers used in this study.

Table 8: Sequencing primers used in this study

Primer	Sequence 5'-3'
H30	CCCATCTAACATCTTCA
L466	GTAGCTCTACTCATCCTGAA
H590	ATATGGAGTGTATAGTC
L1249	CGCAAGGGAAAGCTGAAA
L1630	GATCATATTCAAATAAGGACA
L2164	TAAATGAAGACCTGTATGAA
L2642	GTCCTACGTGATCTGAGTTC
L2935	TTGCTAGCCGTAGCATTCT
H3977	AAGTGGTGTAGAGGAAGCAC
L4100	ACTGGCTCCTCGCTTGAATA
H4720	CATTGACGGTGGTCGCTTTGT
L4750	CCCTCGCCATCTCGTGAACA
L5601	AGTCTGCTTATTGGGCAGA
H6200	AAAGAACAGAACATAGGTGTT
L7104	ATCAATTGGTTCAAGCCAA
H7173	ACAATTCTGCCTTGACAAGG
H7740	AGTACAAAATTGGATGGATGG
L8678	CTCATTGCAACAGCTGCCAA
L8820	AAGCCTCTACCTACAAGAAA
H8828	GGGTGTGCTTGATGGCCAT
L9615	TACTGATGAGGATCTTAATC
H9700	GCTGTAGCTCAGGTGAACGTC
H10010	GACCTCCTTGCATTCAATTG
L10149	TAACAGGCCTGGCATTCCAC
L11350	ATTGCACACGGCTTACTTC
H11555	CACC GGTTAATGCCAGAGTT
L11812	AAACACTAGATTGTGATTCT
H11931	CTAAGACCAACGGATGAGCT
L12716	CATGGCTTGAATAGCAACAAAC
H12909	AGTAGGGCAGATA CCGGTGTAG
L13125	CAACCCCAACTAGCATTCTTC
H13882	GGTAGGTTAGAGGATGCAATGG
L14236	ACACCATTACCGACCAACTACC
L14331	CCACCGTTGTTATTCAACT
H14390	TTGTAGTTGAATAACAAACGG
L15662	CTACCCCTAACTCCCAAAGC
H16200	GAGAACCCCTTACCCGCTGGAGTGAAC

PCR primers and nested sequencing primers used to amplify part of ND2, COI and the control region in mtDNA from 30 individuals of Atlantic halibut are presented in table 9.

Table 9: Amplified fragments and sequencing primers used in SNP analysis part of ND2, COI and the control region in 30 individuals of Atlantic halibut.

Primer	Fragment	Sequence
L4143	1	ACTGGCTCCTCGCTTGAATA
H5087		GAGAAGAAGGGTGGTTGCTG
L4287	Sequencing primer	CAAGCACTACCAACGCTTGA
L5931	2	CAATCTTCTCACTTCACCTTGC
H6601		ACTACATAGTATGTGTCATGC
H6581	Sequencing primer	CAGAACAACTCTGAGAGAGG
L15662	3	CTACCCCTAACTCCCAAAGC
H16385		GGATGACAGAAAGAAATGAGC
L15711	Sequencing primer	CATAAATGTACAATGAAGG

Agarose gel electrophoresis

Gel electrophoresis is a technique for separating charged molecules with different sizes. In an agarose gel, the negatively charged DNA fragments move toward the positive electrode at a rate inversely proportional to their length. After the electric field is applied for a certain period, DNA fragments of different size will be separated. The DNA is visualised in the gel by addition of ethidium bromide. This binds strongly to DNA by intercalating between the bases and is fluorescent meaning that it absorbs invisible UV light and transmits the energy as visible light.

1. Melt 0,7% - 2,5% agarose in 0,5 x TBE electrophoresis buffer in a microwave oven and cool to 65°C.
2. Add ethidium bromide (EtBr) to a final concentration of 0,5 µg/ml and pour the gel into an appropriate mold with a slot former in place.
3. When the gel is fully set, place it in an electrophoresis chamber.
4. Add 1 µl loading buffer (Takara) to 5 µl of sample, before loading the samples. Pour electrophoresis buffer (0,5 x TBE) over the gel to cover it. 6 µl 1 kb DNA plus ladder (Invitrogen™) is run alongside the samples for estimation of DNA size in each band and estimation of the DNA quantity in each band.
5. Run the gel at 4-5 V/cm (measured as the distance between the cathode and the anode) until the bromophenol blue colour band has migrated 2/3 of the gel.
6. After separation, place the gel on top of a UV-transilluminator (365 nm) and photograph the gel.

10xTBE buffer stock

900 mM Tris Borate
20 mM EDTA pH 8.0

Extraction of DNA from gel

DNA bands were excised from an agarose gel and extracted from the gel using E.Z.N.A.[®] Gel Extraction Kit (cat. # D2500-02). This removes primers and remaining nucleotides (dNTP). The manufacturers protocol for gel extraction was followed. DNA was eluted in 40 µl of elution buffer. Concentration and purity of DNA after elution was determined using OD₂₆₀ and OD₂₈₀ readings from a biophotometer. The PCR product was now ready for cloning or sequencing.

Cloning of the fragments into the TOPO vector

Due to heteroplasmy, PCR fragments containing the control region were cloned into a plasmid vector (pCR[®]4-TOPO[®]) using the TOPO TA Cloning kit for Sequencing from Invitrogen (cat.# K4575-01). The critical steps in the procedure are shown in figure 7 below.

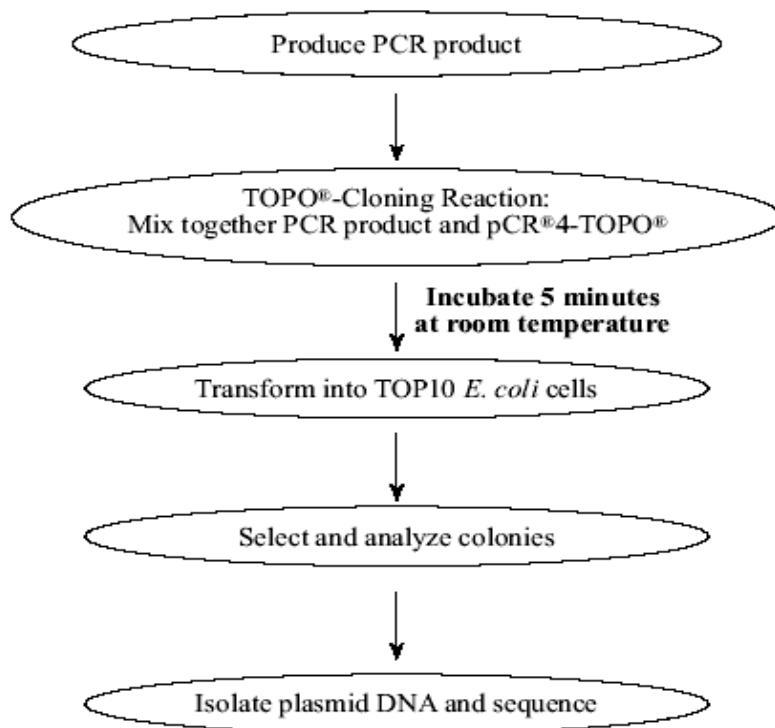


Figure 7: The flow chart above outlines the experimental steps necessary to clone PCR products. (Instruction manual TOPO TA cloning kit, Invitrogen)

The protocol for chemically competent cells was followed. LB agar plates and LB medium were made from recipies in the manufacturers protocol. Bacteria where incubated at 37°C overnight on LB agar plates, employing kanamycin (50 µg/ml final concentration) as the selective agent and x-gal for blue-white screening (40 µl per plate, 20 mg/ml stock solution).

Assumed positive clones (white) were used to inoculate 10ml of LB medium containing 50 µg/ml kanamycin. 10 colonies from each plate were picked. The inoculated medium was incubated overnight in a shaking incubator at 37°C.

Plasmid DNA was isolated using Wizard® Plus SV Minipreps from Promega (cat. # A1460). The centrifugation protocol was followed.

Screening for recombinants

To determine whether or not the colonies contained the insert of interest colony PCR or restriction analysis were performed.

Colony PCR

1. Pick a well isolated colony and transfer half of it to 50µl of sterile water. The other half is transferred to a 10µl disposable pipette tip and kept in a sterile 1,5 ml eppendorf tube at 4 c, awaiting the result of colony PCR.
2. Boil the 50 µl of water containing the bacteria for 10 minutes. This causes the bacteria to lyse and release the plasmid.
3. Centrifuge at 16000 x g for 5 minutes.
4. Use 5 µl of the supernatant in a 50 µl PCR with nested primers, or the same primers as in the original PCR if no suitable primers are available.
5. The product is run on a gel. A band at the expected length indicates a positive colony.

Restriction analysis

Restriction enzymes cut double stranded DNA at specific recognition sequences. The plasmid used for cloning is designed with recognition sites for several different restriction enzymes at each side of a potential DNA insert. After plasmids had been isolated they were digested with an enzyme called EcoR I for 2 hours at 37 C and the reaction was run on a gel. EcoR I cuts after the first nucleotide in the following sequence: GAATTC. A band corresponding to the length of the plasmid will always occur. The insert of interest can occur as one band or several shorter bands, that together add up to the expected length.

A digestion with EcoR I consisted of:

Plasmid approx. 100ng/ul	2 ul
EcoR I 10 U/ul	0,5 ul
Buffer 10x	2 ul
<u>Water</u>	<u>15,5 ul</u>
Total volume	20 ul

Sequencing

Sequencing was performed on PCR fragments and cloned inserts using a BigDyeTerminator Ver3.1 kit (Applied Biosystems). Primers used were the same as those used for PCR including inner primers, all presented in tables 5, 6, 7, 8 and 9.

Component	Amount
PCR product	30-300 ng
Big Dye	1 ul
Sequencing buffer 5x	2 ul
Primer 3,5 uM	1 ul
Water to adjust volume)	x ul
Total volume	10 ul

150–300 ng of plasmid or 60-80 ng of a PCR product of 4 kb was used as template for sequencing with the following cycling parameters:

Initial denaturation	96°C	1 min
Denaturation	96°C	30 sec
Annealing	50°C	15 sec
Elongation	60°C	4 min
Cycles	35	

After sequencing-reactions samples were precipitated using ethanol / EDTA and sent to UNN in Tromsø for separation analysis using the Applied Biosystems 3130xl Genetic Analyzer.

Procedure for Ethanol/EDTA precipitation:

1. 1 µl 125 mM EDTA, 1 µl 3 M Natrium acetat, pH 7.5, 27 µl 96% ethanol and 10 µl sequencing reaction were added to a 1.5 ml eppendorf tube.
2. The mixture was gently mixed and incubated at room temperature for 15 minutes.
3. The tube was centrifuged at room temperature for 15 minutes at 13000 rpm.
4. The supernatant was removed.
5. 150 µl 70% ethanol was added.
6. The tube was centrifuged at room temperature for 5 minutes at 13000 rpm.
7. All the ethanol was removed.
8. The sample was air dried.

Bioinformatics

Computer analysis of DNA sequences, such as editing and assembling was performed using the Lasergene DNA Star package. EditSeq was used to edit the sequences. The first and last part of a sequence is often of poor quality and was therefore removed. Undetermined bases

found internally in the sequence were changed into the correct bases if this could be determined by manual reading of the chromatogram files. EditSeq was also used for translation of protein coding DNA into amino acids.

After editing the sequences, they were assembled in SeqMan. When assembling the first Atlantic halibut mtDNA sequence, the entire mitochondrial sequence from stone flounder was used as a backbone to help assemble the sequences correctly. The complete mtDNA sequence from Atlantic halibut was then used as a backbone when assembling subsequent individuals of Atlantic halibut, Pacific halibut and Greenland halibut.

A variety of programs was used for further analysis: For statistical analysis of the mtDNA sequences DNAsP was used (Rozas et al. 2003). The putative open reading frames (ORF) were identified Gene finder program available in NCBI using the mitochondrial genetic code. Similarity searches were done using BLAST (Altschul et al. 1990). Both programs are available at the NCBI web site www.ncbi.nlm.nih.gov/gorf/gorf.htm and www.ncbi.nlm.nih.gov/blast/. 21 out of 22 tRNA were identified using the computer software tRNAscan-SE 1.21 (Lowe and Eddy 1997) found at www.genetics.wustl.edu/eddy/tRNAscan-SE, with the default parameters for mitochondrial DNA. tRNA-Ser was identified by sequence homology with Stone flounder.

Phylogenetic analysis were performed using Clustal X (Thompson et al. 1997) and MEGA 3.1 (Kumar et al. 2004) as follows: An alignment of sequences from the species to be included was created by Clustal X A phylogenetic tree based on the alignment were produced by the use of Neighbor Joining (NJ), Minimal Evolution (ME) and maximum parsimony (MP) methods. A bootstrap with 2000 replicates was performed to determine the reliability of the different parts of the inferred trees. In trees including only Pleuronectiform species, *Lycodes toyamensis* and *Sebastes schlegeli* were used as an outgroup to root the tree. A tree including Salmonidae and Gadidae species were rooted by *Dorosoma cepedianum* and *Cromeria nilotica*. These species were chosen because they belong to the sister taxa of the least common taxonomic group of the species in question. Two species was used as an outgroup to make sure that the outgroup really was an outgroup relative to the species in focus.

Results

Genome content

37 genes (two rRNAs, 13 proteins and 22 tRNAs) were identified in Atlantic halibut, Pacific halibut and Greenland halibut. This is the same as is found in other vertebrates, as reviewed by Boore (1999). A large non-coding control region was identified by its content of conserved sequence motifs. The majority of the genes were encoded on the H-strand, except those for ND6 and 8 tRNA genes. In some of the tRNAs and protein coding genes, base-sharing and reading frame overlaps were observed. All genes in the three species were of identical size, except for the 16S rRNA gene. Here the gene in Greenland halibut was one bp longer than the one found in Atlantic halibut and Pacific halibut. In addition to the origin of replication of the light strand (Ori L, 38 bp), a few non-coding insertion sequences were found in Atlantic-, Pacific- and Greenland halibut. The largest one was located between tRNA-Ser^{UCN} and tRNA-Asp, and was 14 bp. Extensive length variation was observed in the control region, due to variation in copy number of a 61 bp repeat motif.

Localization of protein coding genes, tRNA and rRNA, as well as other features in the mitochondrial genome from Atlantic halibut, Pacific halibut and Greenland halibut are presented in Tables 10, 11 and 12, respectively

The mtDNA genome in Atlantic halibut 1, 2, 3 and 4 was 17546 bp, 17619 bp, 17973 bp and 17729 bp.

Table 10: Localization of genes and features in the mitochondrial genome of Atlantic halibut (*Hippoglossus hippoglossus* 1). tRNA genes denoted by their three letter abbreviations

Gene	From	To	Size (bp)	Codon (Init./Term.)		Anticodon (tRNA)
Phe	1	68	68			GAA
12S rRNA	69	1017	949			
Val	1018	1091	74			TAC
16S rRNA	1092	2806	1715			
Leu ^{UUR}	2807	2880	74			UUA
ND1	2881	3855	975	ATG	TAG	
Insertion	3855	3860	6			
Ile	3861	3931	71			GAU
Gln	4001	3931 (L)71				UUG
Met	4001	4069	69			CAU
ND2	4070	5115	1046	ATG	TA-	
Trp	5116	5187	72			
Insertion	5188		1			
Ala	5257	5189 (L)69				UGC
Insertion	5258		1			
Asn	5331	5259 (L)73				GUU
Ori L	5332	5369	38			
Cys	5434	5370 (L)65				GCA
Tyr	5502	5435 (L)68				GUU
Insertion	5503		1			
COI	5504	7069	1566	GTG	AGA	
Ser ^{UCN}	7135	7065 (L)71				UGA
Insertion	7136	7149	14			
Asp	7150	7220	71			GUC
Insertion	7221	7226	6			
COII	7227	7917	691	ATG	T--	
Lys	7918	7990	73			UUU
Insertion	7991		1			
ATP8	7992	8159	168	ATG	TAA	
ATP6	8150	8832	683	ATG	TA-	
COIII	8833	9617	785	ATG	TA-	
Gly	9618	9689	72			UCC
ND3	9690	10038	349	ATG	T--	
Arg	10039	10107	69			UCG
ND4L	10108	10404	297	ATG	TAA	
ND4	10398	11778	1381	ATG	T--	
His	11779	11848	70			GUG
Ser ^{AGY}	11849	11915	67			
Insertion	11916	11918	3			
Leu ^{CUN}	11919	11991	73			UAG
ND5	11992	13830	1839	ATG	TAA	
ND6	14348	13827 (L)	522	ATG	TAG	
Glu	14417	14349 (L)	69			UUC
Insertion	14418	14421	4			
Cyt-b	14422	15562	1141	ATG	T--	
Thr	15563	15635	73			UGU
Pro	15705	15635 (L)	71			UGG
Control region	15706	17546	1841			

The mitochondrial genome of Pacific halibut 1, 2, 3 and 4 was 17841 bp, 17841 bp, 17963 bp and 17902 bp.

Table 11: Localization of genes and features in the mitochondrial genome of Pacific halibut (*Hippoglossus stenolepis* 1). tRNA genes denoted by their three letter abbreviations

Gene	From	To	Size (bp)	Codon (Init./Term.)		Anticodon (tRNA)
Phe	1	68	68			GAA
12S rRNA	69	1017	949			
Val	1018	1091	74			UAC
16S rRNA	1092	2806	1715			
Leu ^{UUR}	2807	2880	74			UAA
ND1	2881	3855	975	ATG	TAG	
Insertion	3856	3860	5			
Ile	3861	3931	71			GAU
Gln	4001	3931 (L)71				UUG
Met	4001	4069	69			CAU
ND2	4070	5115	1046	ATG	TA-	
Trp	5116	5187	72			
Insertion	5188		1			
Ala	5257	5189 (L)69				UGC
Insertion	5258		1			
Asn	5331	5259 (L)73				GUU
Ori-L	5332	5369	38			
Cys	5434	5370 (L)65				GCA
Tyr	5502	5435 (L)68				GUA
Insertion	5503		1			
COI	5504	7069	1566	GTG	AGA	
Ser ^{UCN}	7135	7065 (L)71				UGA
Insertion	7136	7149	14			
Asp	7150	7220	71			GUC
Insertion	7221	7226	6			
COII	7227	7917	691	ATG	T--	
Lys	7918	7990	73			UUU
Insertion	7991		1			
ATP8	7992	8159	168	ATG	TAA	
ATP6	8150	8832	683	ATG	TA-	
COIII	8833	9617	785	ATG	TA-	
Gly	9618	9689	72			UCC
ND3	9690	10038	349	ATG	T--	
Arg	10039	10107	69			UCG
ND4L	10108	10404	297	ATG	TAA	
ND4	10398	11778	1381	ATG	T--	
His	11779	11848	70			GUG
Ser ^{AGY}	11849	11915	67			
Insertion	11916	11918	3			
Leu ^{CUN}	11919	11991	73			UAG
Insertion	11992		1			
ND5	11993	13831	1839	ATG	TAA	
ND6	13828	14349 (L)	522	ATG	TAG	
Glu	14418	14350 (L)	69			UUC
Insertion	14419	14422	4			
Cyt-b	14423	15563	1141	ATG	T--	
Thr	15564	15636	73			UGU
Pro	15706	15636 (L)	71			UGG
Control region	15707	17841	2135			

The mitochondrial genome of Greenland halibut 1, 2, 3 and 4 was 18017 bp, 18139 bp, 17895 bp and 18078 bp.

Table 12: Localization of genes and features in the mitochondrial genome of Greenland halibut (*Reinhardtius hippoglossoides* 1). tRNA genes denoted by their three letter abbreviations

Gene	From	To	Size (bp)	Codon (Init./Term.)		Anticodon (tRNA)
Phe	1	68	68			GAA
12S rRNA	69	1017	949			
Val	1018	1091	74			UAC
16S rRNA	1092	2805	1714			
Leu ^{UUR}	2806	2879	74			UAA
ND1	2880	3854	975	ATG	TAA	
Insertion	3855	3859	5			
Ile	3860	3930	71			GAU
Gln	4000	3930 (L)	71			UUG
Met	4000	4068	69			CAU
ND2	4069	5114	1046	ATG	TA-	
Trp	5115	5186	72			
Insertion	5187		1			
Ala	5256	5188 (L)	69			UGC
Insertion	5257		1			
Asn	5330	5258 (L)	73			GUU
Ori-L	5331	5368	38			
Cys	5433	5369 (L)	65			GCA
Tyr	5501	5434 (L)	68			GUA
Insertion	5502		1			
COI	5503	7068	1566	GTG	AGA	
Ser ^{UCN}	7134	7064 (L)	71			UGA
Insertion	7135	7148	14			
Asp	7149	7219	71			GUC
Insertion	7220	7225	6			
COII	7226	7916	691	ATG	T--	
Lys	7917	7989	73			UUU
Insertion	7990		1			
ATP8	7991	8158	168	ATG	TAA	
ATP6	8149	8831	683	ATG	TA-	
COIII	8832	9616	785	ATG	TA-	
Gly	9617	9688	72			UCC
ND3	9689	10037	349	ATG	T--	
Arg	10038	10106	69			UCG
ND4L	10107	10403	297	ATG	TAA	
ND4	10397	11777	1381	ATG	T--	
His	11778	11847	70			GUG
Ser ^{AGY}	11848	11915	67			
Insertion	11916	11918	3			
Leu ^{CUN}	11919	11991	73			UAG
ND5	11992	13830	1839	ATG	TAA	
ND6	13827	14348 (L)	522	ATG	TAG	
Glu	14417	14349 (L)	69			UUC
Insertion	14418	14421	4			
Cyt-b	14422	15562	1141	ATG	T--	
Thr	15563	15635	73			UGU
Pro	15705	15635 (L)	71			UGG
Control region	15706	18017	2312			

Protein coding genes

Thirteen protein coding genes were identified by sequence alignment with mtDNA from bastard halibut (*Paralichthys olivaceus*) and stone flounder (*Platichthys bicoloratus*) as well as putative open reading frames (www.ncbi.nlm.nih.gov/projects/gorf/).

Among the protein coding genes, cytochrome oxidase subunit 1 (COI) begins with the GTG codon, while the others begin with the ATG codon, the standard initiation codon. Seven of 13 protein coding genes end with a truncated termination codon, such as T or TA, judging from the 5'end of downstream tRNA genes. Complete termination codons are proposed to be created after post-transcriptional addition of adenosine at the RNA level. COI ends with AGA, and ND6 ends with TAG.

ATP8 and ATP6 share 10 bases, ND4L and ND4 share 7 bases and ND5 and ND6 share 4 bases, but ND5 and ND6 are coded on opposite strands.

Ribosomal RNA genes

The two rRNA genes are located between the tRNA^{Phe} and tRNA^{Leu(UUR)}, separated by tRNA^{Val}. The length of the 12S rRNA gene was 949 bp long in Atlantic halibut, Pacific halibut and Greenland halibut. 16S rRNA gene was 1714 bp long in Atlantic halibut and Pacific halibut, and 1715 bp long in Greenland halibut.

Transfer RNA genes

Twenty-two tRNA genes between 65 and 74 bp in size were identified in Atlantic halibut, Pacific halibut and Greenland halibut (figure 9). tRNA genes were derived from their gene structures and identified and located by anticodon recognition, their proposed secondary structure, and homology with already published mtDNAs within the Pleuronectiformes. The classic secondary structure of tRNAs is shown in figure 8. All but one (tRNA-Ser (AGY)) folded into the cloverleaf secondary structure with tRNAscan-SE 1.21 (Lowe and Eddy 1997). Two tRNA genes specify leucine, Leu(UUR) and Leu(CUN), and two tRNA genes specify serine, Ser(UCN) and Ser(AGY). Nucleotide substitutions between the three halibut species were found. Mostly substitutions were located in the loop structures the tRNAs. But in some cases substitutions in the stem structures were observed. Atypical (non-Watson and Crick) base pairings (G-U and G-A) and mismatches were identified in all tRNA genes. Four tRNA genes were perfectly conserved among the 12 individuals; Ser (UCN), Arg, Leu (CUN) and Glu.

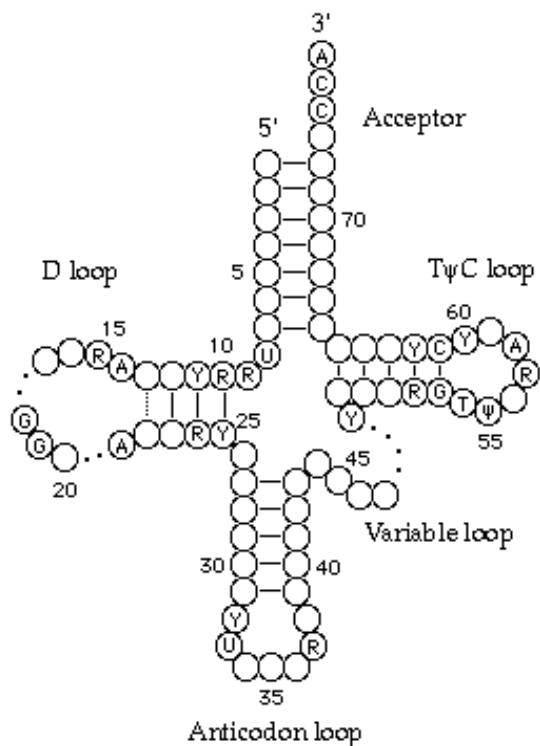
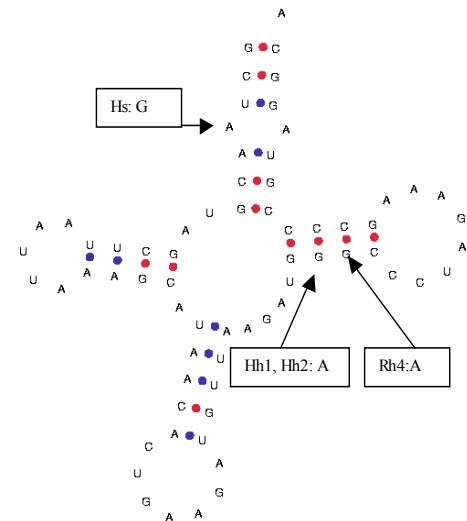
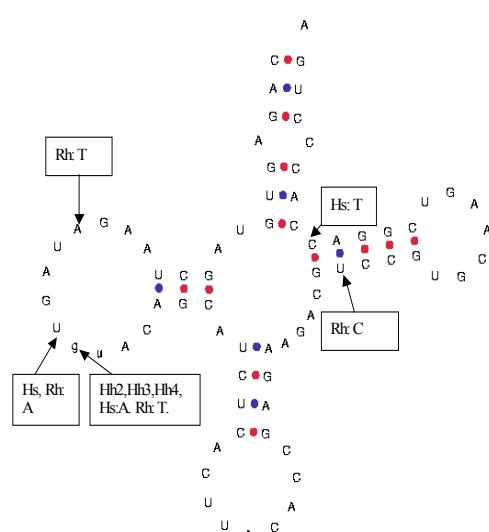


Figure 8: Secondary structure of a typical transfer RNA. Figure from www.genetics.wustl.edu/eddy/tRNAscan-SE

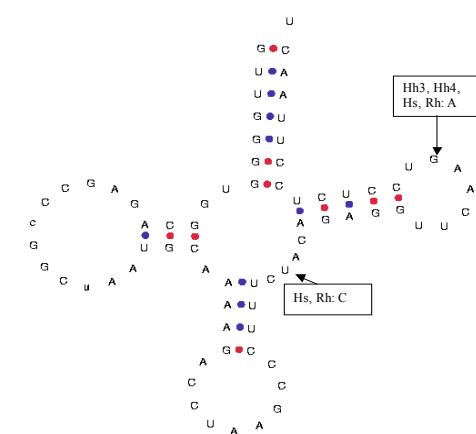
tRNA-Phe



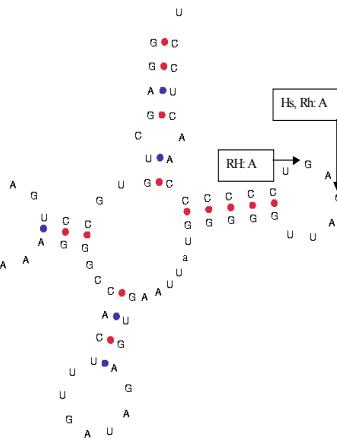
tRNA-Val



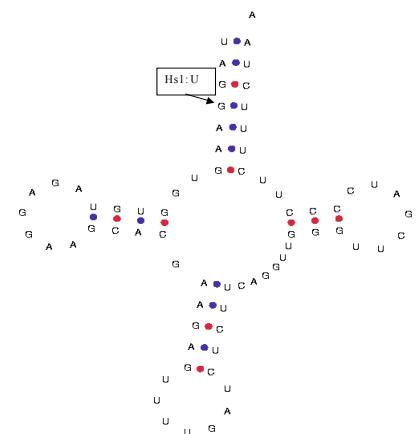
tRNA-Leu (UUR)



tRNA-Ile



tRNA-Gln



tRNA-Met

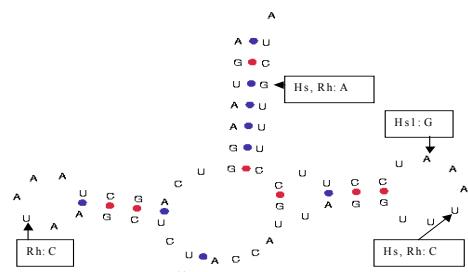
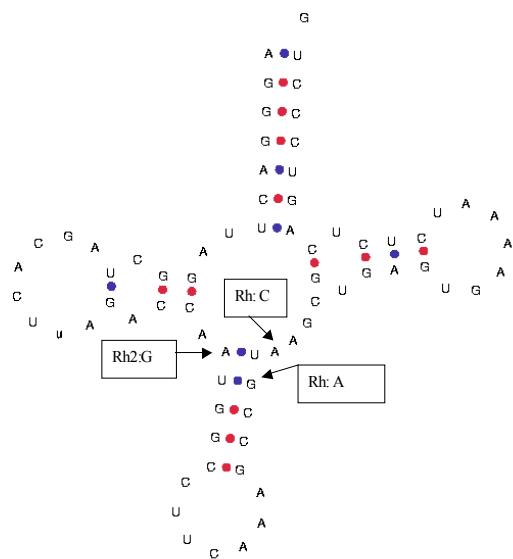
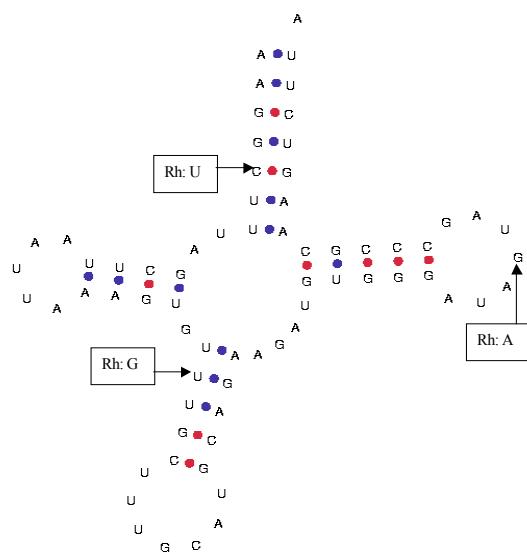


Figure 9: Proposed secondary structure of mitochondrial tRNA genes in Atlantic halibut 1, with nucleotide substitutions found in individuals of Atlantic halibut (Hh), Pacific halibut (Hs) and Greenland halibut (Rh). The amino acid recognition site in the 3' end is not shown. Figures created by tRNA-SE 2.1 (Lowe and Eddy 1997).

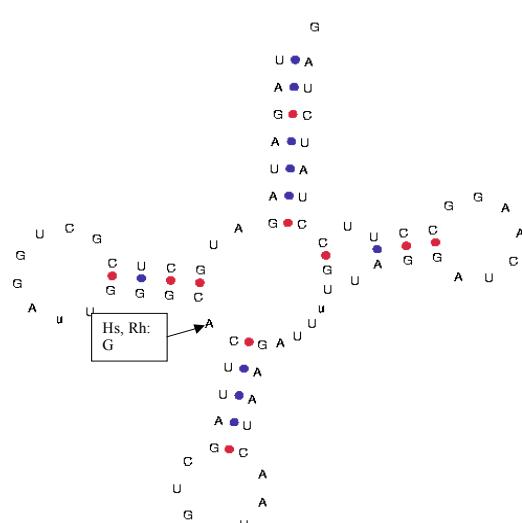
tRNA-Trp



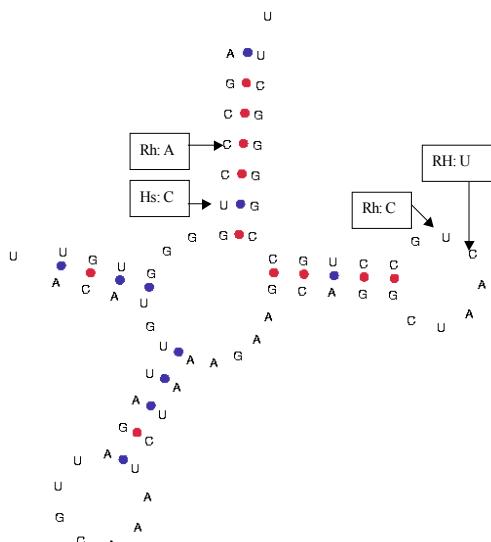
tRNA-Ala



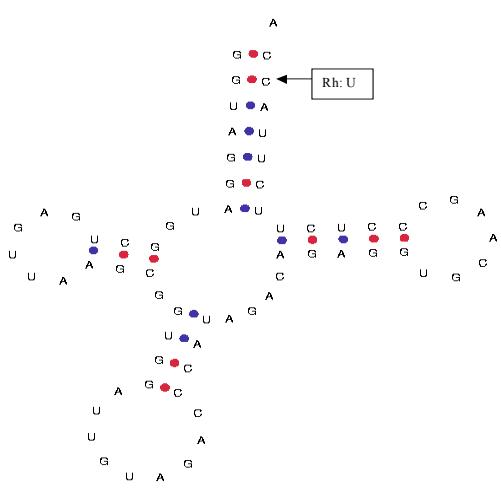
tRNA-Asn



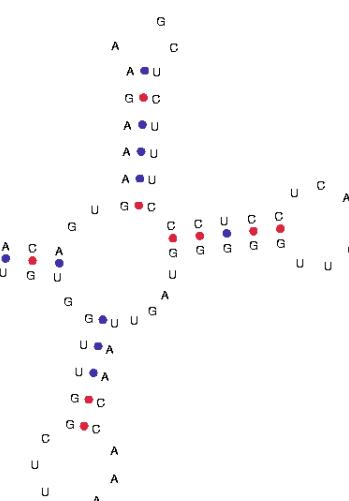
tRNA-Cys



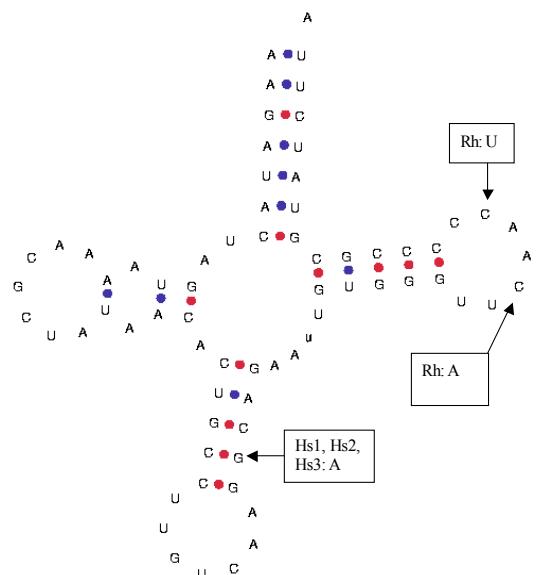
tRNA-Tyr



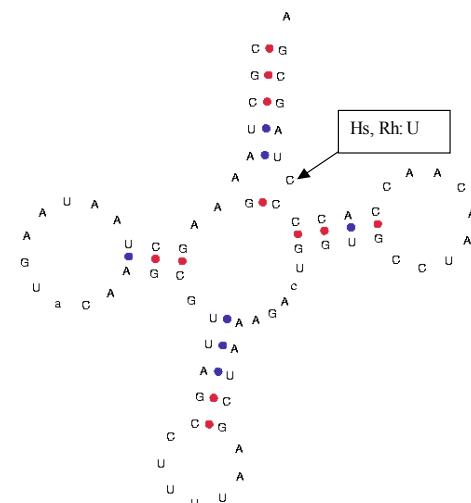
tRNA-Ser (UCN)



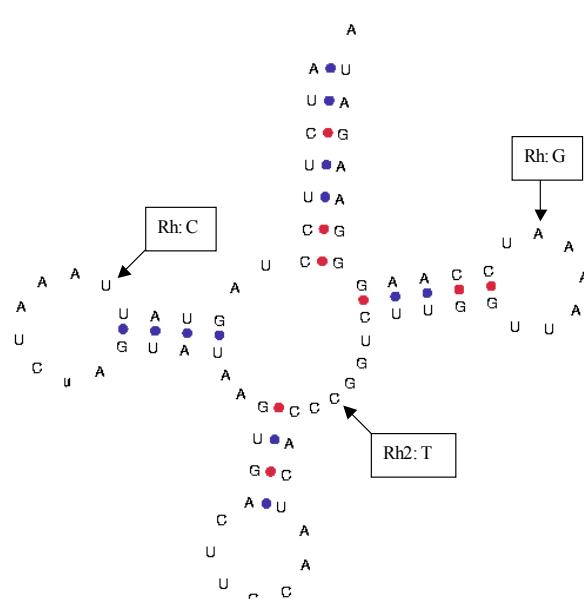
tRNA-Asp



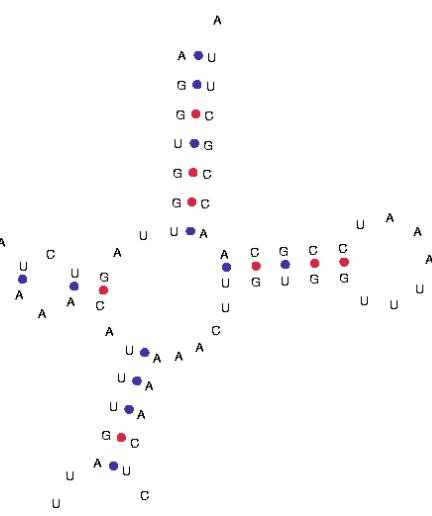
tRNA-Lys



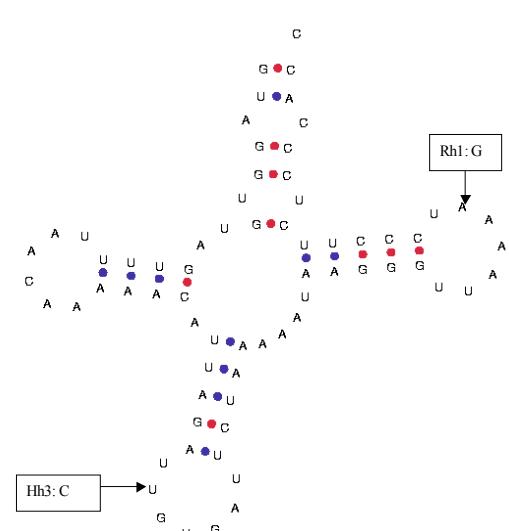
tRNA-Gly



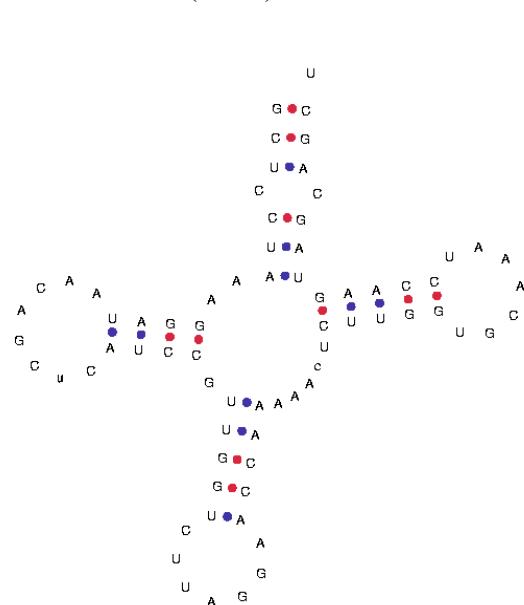
tRNA-Arg



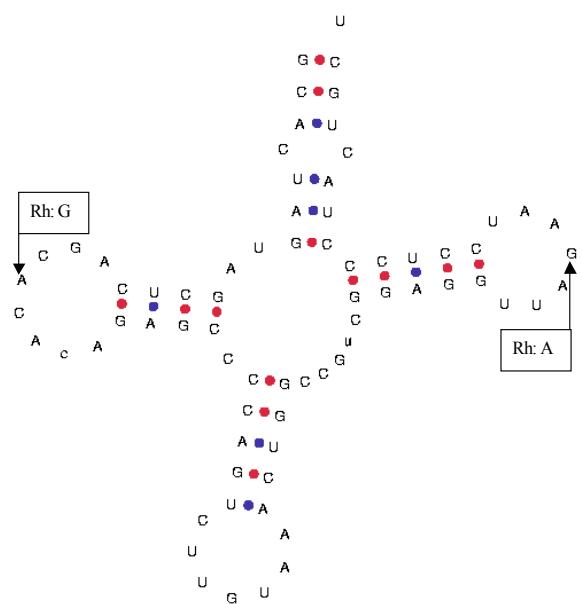
tRNA-His



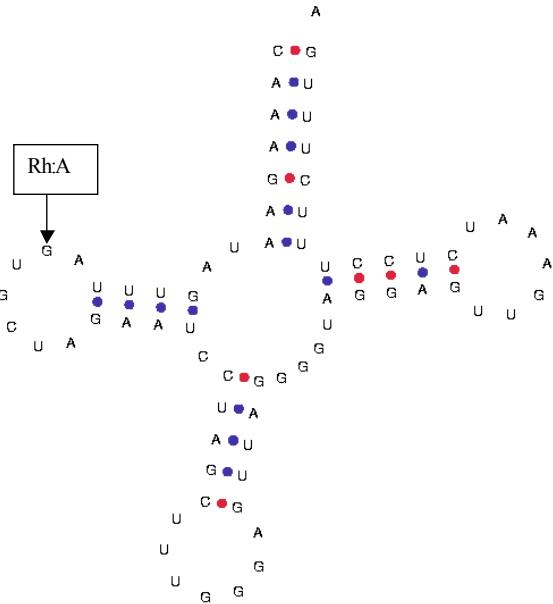
tRNA-Leu (CUN)



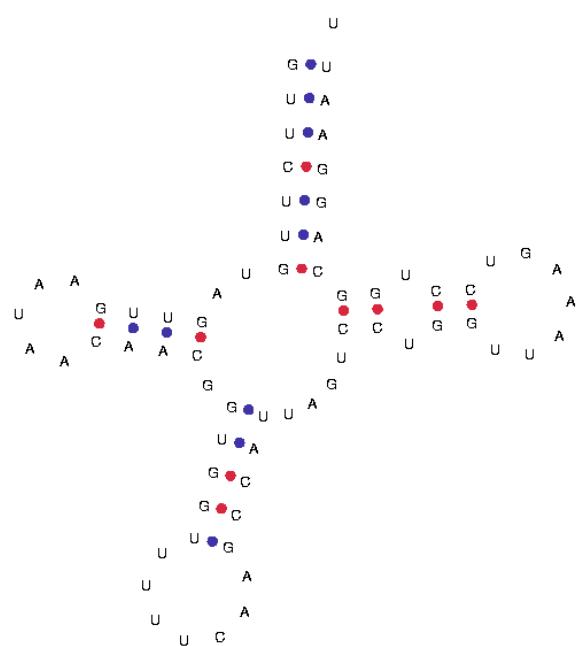
tRNA-Thr



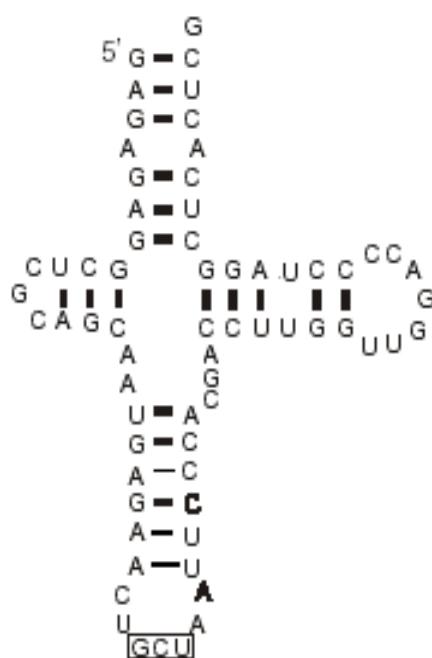
tRNA-Pro



tRNA-Glu



tRNA-Ser (AGY)



Non coding regions

Replication origin of light strand (Ori-L)

A non-coding sequence associated with the putative L-strand replication origin is 38 bp long and located between tRNA-Asn and tRNA-Cys. This region, together with a part of tRNA-Cys is able to form a stable stem-loop structure with 12 paired bases in the stem (Fig 10). The conserved motif (5'-GCCGG-3') was found in the stem of the tRNA-Cys.

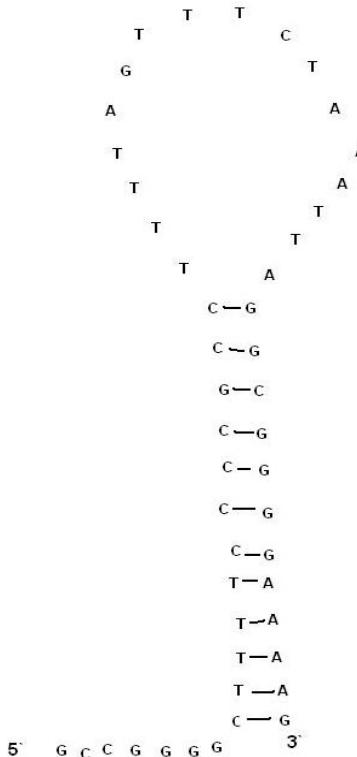


Figure 10: Proposed secondary structure of the replication origin of light strand in Atlantic halibut, with the conserved motif 5'-GCCGG-3'. Nucleotide sequence of the H-strand is shown.

Control region

The control region is the largest non coding region. It is flanked by the tRNA^{Pro} and tRNA^{Phe}-genes. Three domains are found in the control region; one hyper variable domain known as the ETAS region (Extended Termination Associated Sequences), one central conserved domain (CC region) and one conserved sequence block domain (CSB-region) (Fig 11). Several conserved motifs were identified by comparison with other fish control regions (Breines 2005; Inoue et al. 2001; Kim and Lee 2004; Lee et al. 1995; Nesbø et al. 1998; Saitoh et al. 2000) (Fig 11 and 12).

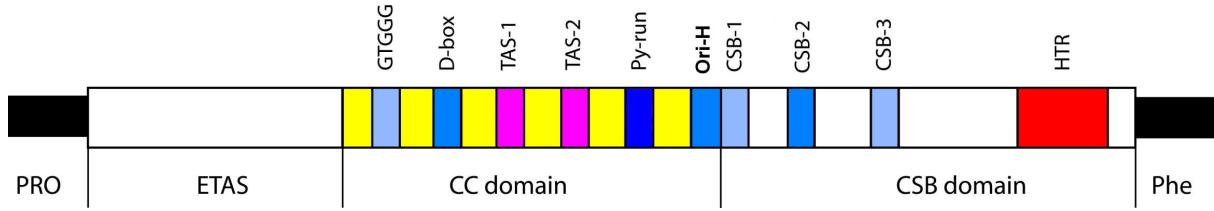


Figure 11: Schematic view of control region consisting of the three domains; ETAS, CC and CSB. The region is localized between tRNA-Pro and tRNA-Phe. A number of conserved sequence motifs were identified

<pre>Hippoglossus_hippoglossus_1 Hippoglossus_stenolepis_1 Reinhardtius_ippoglossoides_1 Verasper_variegatus Limanda_ferruginea Hippoglossoides_plateossoides Paralichtys_olivaceus Solea_vulgaris</pre> <pre>Hippoglossus_hippoglossus_1 Hippoglossus_stenolepis_1 Reinhardtius_ippoglossoides_1 Verasper_variegatus Limanda_ferruginea Hippoglossoides_plateossoides Paralichtys_olivaceus Solea_vulgaris</pre>	<pre>-----TGGTACA-TAAATGTACAATGAAGGGATTTCA-TATACATG -----TGGTACA-TAAATGTACAATGAAGGGATTTCA-TATACATG -----CGGCATAATGTATGTAAAGCAAAGGATTTA-TATACATG TTGCATACGCACTGTT-CATACGATAACGCAATTGTTA-TGTGCATA -----GACACAATACATGTACA--TGAAGGTTTCA-TGTACATG -----ATGAATGTACAATAAAGGATTTAATGTACATC -----GGAACATATGTTTATGAAAATT---AATATACATA -----TTCTGTTTATG---TATGAAAAT---AATA--CATA * * * * * * * * * * * * * * * * ***</pre> <pre>TATGTAATAACACCATAATTATAGTAACCATTATGTGATGTACTAG TATGTAATAACACCATAATTATAGTAACCATTATGTGATGTACTAG TATGTAATAACACCATAATTATAGTAACCATTATGTGATGTACTAA TTTGTAAATAACACCATAATTATAGTAACCATTATATAATGAACTAG TATGTAATAACACCATAATTATAGTAACCATTATGTGATGTACTAG TATGTAATTACACCATAATTATAGTAACCATTAAGTCCGATG-TACAA TATGTTATTATCACCATTAATCTATATTAAACCATTATATAATGCTTCC **</pre> <pre>GACATACATGTATAATCACCAATCTCGTAAATACAGCACTCATTCAAA GACATACATGTATAATCACCACTCTCGTAAATAGAACACTCATTCAAC GACATACATGTATAATCACCACTAGTAATATAGCACTCATTCAAC GACATTCATGTATTAACTAATCTAGTAATATAGCACTCATTCAAC GACATTCATGTATAATACCTAATCTAGTAATTAAGCACTCATTCAAC GACATACATGTATAATACCTAATCTAGTAATAGCACTCATCACAC GACACAAATGGATG-TGAACAAAATGGTGTCAAACATTCAATACCAG TACATTACTGATTAATCACCATCTCGAAATAAACCACTTCACTTACAC *</pre> <pre>CATTTTAACTAAGATAG-ACTAAAACC-TGAATAATCACTAATCTAAA CATTTTAACTAAGACGA-ACTAAAACC-TGAGTGATCACTAATCTAAA CATTTTAACTAAGAAATTACTAGAACCTGACTTCACATTACGTTACA CATTTTAACTAAGTAA-ACTAAAACC-TAATTAACTAATTTAAA CAATTAACTTAACAGATAG-ACTAGAACCC-TAGACAAGCACTAACCTTACA CA-TTTTAACTAAGAGTT-ACTAGAACCC-CAACTTACTAATATAACA C-TATATAACTAAATATACAAAAACCAAA-ACCTATAA--GGTATA-- T-TACTATACTAAG-ACGGACCCAAACCAACTTGTAA--AATTCAACT *</pre> <pre>TATGTGA-AAGTCCAGGACCAGTCGA---AATCTAACAGCGAACACAACA TAAGTGA-AAGTCCAGGACCAGTCGA---GATCTAACAGCGAACACAACA TATGTGA-AATTCCAGGACCAGCGA---CATTAAAGACCGAACACAACA TATGTAACCTCTAGGACCAATCGA---AATTAAAGACCGAACACGACA TTAGTGA-AATTCCAGGACCAGTCGA---AACTTAAGACCGAACACAACA TATGTGG-AAGTCAAGGAGTAGTCGA---AACTTAAGACCGAACACAACA -CAATAAGAATTAAAGACTAGTCGA---AACTTTACACCGAACACAACC TCAATAA-GAATTCACTGCTAATCGATCCGACAAGACTCCACATAAAAC *</pre>
--	--

Figure 12: Alignment of ETAS and CC region of the control region in 8 Pleuronectiform species. Nine conserved domains were identified. CSB; conserved sequence blocks, TAS; termination associated sequence (putative position), D-box, replication origin H-strand (O_H) and Py-run; repeats of Pyrimidine bases. Nucleotide sequence of the L-strand is shown.

Hippoglossus_ippoglossus_1
Hippoglossus_stenolepis_1
Reinhardtius_ippoglossoides_1
Verasper_variegatus
Limanda_ferruginea
Hippoglossoides_platessoides
Paralichtys_olivaceus
Solea_vulgaris

CTCATCAGTCGAGT-TATACCAAGACTCAAATCTGCCGACCCAA-AA
CTCATCAGTCGAGT-TATACCAAGACTCAAATCTGCCGACCTCAA-AA
CTCACCGGTCAAGT-TATACCAAGACTCAAATTCACTCACTAAA-AA
CTCATCAGTTAAGT-TATACCAAGACTCAAATCTGCCATCATAA-AT
CTCATCAGTCGAGT-TATACCAAGACTCAAATCTC-TTCACGCCAA-AA
CTCATCAGTCGAGT-TATACCAAGACTCAAATCCCTAACATCAA-AA
TTCATATGTCAAGT-TATACCAAGACTCAAACCTCTGTCGATCCAA-AT
TG-ACACGGTAAATATACATTGATCCAACATCCTATAAAATTAACTAC
* *

Hippoglossus_ippoglossus_1
Hippoglossus_stenolepis_1
Reinhardtius_ippoglossoides_1
Verasper_variegatus
Limanda_ferruginea
Hippoglossoides_platessoides
Paralichtys_olivaceus
Solea_vulgaris

TCTATGTAGTAAGAGCCTACCAACC-GGTGATTCTTAATGATAACTCT
TCTATGTAGTAAGAGCCTACCAACC-GGTGATTCTTAATGATAACTCT
ACCAACCCAATAAGAGCCTACCAACC-GGTGATCCCTTAATGATAACTCT
TCTATGTAGTAAGAGCCTACCAACC-GGTGATTTCTGAATGATAACTCT
TTCGATGTAGTAAGAGTCTACCAACC-GGTGATTTCTAAATGATAACGGC
CCCTATGTAGTAAGAGCCTACCAACC-GGTGATCCCTTAATGATAACGGT
TCCCAGTCAGTAAGAGCCTACCATCA-GTTGATTACTTAATGCCAACGGT
AAATGCGCAGTAAGAGCCTACCAACAAGCTCATATCTTAATGATAACGGT
* *

GTGGG-box

TATTGAGGGTGGGGACAAAATTGTGGGGGTTTCACTCGGTGAATTATT
TATTGAGGGTGGGGACAAAATTGTGGGGGTTTCACTCGGTGAATTATT
TATTGAGGGTGGGGACAAAATTGTGGGGGTTTCACTCGGTGAATTATT
TATTGATGGTCAGGGACAGAAATCGTGGGGGTTTCACTCAGTGAACATA
TATTGAAGGTGAGGGACAAAATTGTGGGGGTTTCACTCGGTGAATTATT
TATTGAAGGTGAGGGACAAAATTGTGGGGGTTTCACTCGGTGAATTATT
TATTGAAGGTGAGGGACAAAATTGTGGGGGTTTCACTCGGTGAACATA
TATTGATGGTCAGGGACAGATACGTGGGGGTTTCACTTAGTGAATTATT
***** *

D-box

CCTGGCATTGGTTCC-TACTTCAGGGCCATTGATTGATGTTATCCCTCA
CCTGGCATTGGTTCC-TACTTCAGGGCCATTGATTGATGTTATCCCTCA
CCTGGCATTGGTTCC-TACTTCAGGGCCATTGATTGATGTTATCCCTCA
CCTGGCATTGGTTCC-TACTTCAGGGCATATTATCGATATTATCCCTCA
CCTGGCATTGGTTCC-TACTTCAGGGCATATTACCGAT-TTATCCCCCA
CCTGGCATTGGTTCC-TACTTCAGGGCACGGATCGAT-TTATCCCCCA
CCTGGCATTGGTTCC-TACTTCAGGGCATGACCTGATATTATCCCTCA
CCTGGCATTGGCTCACATCTCAAGGGCATATTATATCCTTTCTCA
***** *

TAS-1 (both putative) TAS-2

CACTTCATCGACACTTACATAAGTTAATGTTGATAATACATACGACTCG
CACTTCATCGACACTTACATAAGTTAATGTTGATAATACATACGACTCG
CACTTCATCGACACTTACATAAGTTAATGTTGATAATACATACGACTCG
CACTTCATCGACGCTTGACATAAGTTAATGTCGATAATACATACGACTCG
CACTTCATCGACACTTACATAAGTTAATGTTGATAATACATACGACTCG
CACTTCATCGACACTTACATAAGTTAATGTTGATAATACATACGACTCG
CACTTCATCGACACTTACATAAGTTAATGTTGGTATACATACGACTCG
CAGATTTACTGGCCCTGACATTGACTGATGCTTCGAC-TATACGACTCG
* *

Py-run

TTACCCACCAAGCCGGCGTTCACTCCAGCGGGTAAGGGGTTCTCTTTTT
TTACCCAGCAAGCCGGCGTTCACTCCAGCGGGTAAGGGGTTCTCTTTTT
TTACCCACCAAGCCGGCGTTCACTCCAGCGGGTAAGGGGTTCTCTTTTT
TTACCCAGCAAGCCGGCGTTCACTCCAGCGGGTAAGGGGTTCTCTTTTT
TTACCCAGCAAGCCGGCGTTCACTCCAGCGGGTAAGGGGTTCTCTTTTT
TTACCCAGCAAGCCGGCGTTCTCTCCAGCGAGCAAGGGGTTCTCTTTTT
***** *

Py-run

-TTTTTCCTTCACTTGA~~CATTCAAAAGTGC~~CATCCAGCCAACGGAGACG
-TTTTTTCCTTCACTTGA~~CATTCAAAAGTGC~~CATCCAGCCAACAGAGACG
-TTTTTTCCTTCACTTGA~~CATTCAAAAGTGC~~CATCCAGCCAACAGAGACG
-TTTTTTCCTTCACTTGA~~CATTCAAAAGTGC~~CATCCAGCCAACAGAGACG
GTTTTTCCTTCACTTGA~~CATTCAAAAGTGC~~CATCCAGCCAACGGAAACG
-TTTTTTCCTTCAATCG-CATTCAAAAGGCATCCAGCCAAATAAAACG
-TTTTTTCCTTCACTTGCAGAGTGCATCTAGCCGATAGAAATA
-TTTTTTCCTTCACTTGCAGAGTGCAGCG---CGGTAATAGTT
***** *

Ori-H

```

Hippoglossus_hippoglossus_1
Hippoglossus_stenolepis_1
Reinhardtius_ippoglossoides_1
Verasper_variegatus
Limanda_ferruginea
Hippoglossoides_platessoides
Paralichtys_olivaceus
Solea_vulgaris

```

```

Hippoglossus_hippoglossus_1
Hippoglossus_stenolepis_1
Reinhardtius_ippoglossoides_1
Verasper_variegatus
Limanda_ferruginea
Hippoglossoides_platessoides
Paralichtys_olivaceus
Solea_vulgaris

```

```

Hippoglossus_hippoglossus_1
Hippoglossus_stenolepis_1
Reinhardtius_ippoglossoides_1
Verasper_variegatus
Limanda_ferruginea
Hippoglossoides_platessoides
Paralichtys_olivaceus
Solea_vulgaris

```

```

Hippoglossus_hippoglossus_1
Hippoglossus_stenolepis_1
Reinhardtius_ippoglossoides_1
Verasper_variegatus
Limanda_ferruginea
Hippoglossoides_platessoides
Paralichtys_olivaceus
Solea_vulgaris

```

```

Hippoglossus_hippoglossus_1
Hippoglossus_stenolepis_1
Reinhardtius_ippoglossoides_1
Verasper_variegatus
Limanda_ferruginea
Hippoglossoides_platessoides
Paralichtys_olivaceus
Solea_vulgaris

```

```

TTAAAGGGTGAGCACTTCTGCACGCCGTACATAGTATCCGTG-
TTAAAGGGGAGCATTCTGCACGCCGTACATAGTATCCATGT-
TTAAAGGGTGAGCACTTCTGCACCGCGCGTACATAGTATCCATGT-
TTAAAGGGTGAGCACTTCTGCACCGCGCGTACATAGTATCCATGT-
TCAAAGGGGAGCATTCTGCACCGCGTACATAGTATCCATGT-
TCAAAGGGGAGCATTCTGCACCGCGTACATAGTATCCATGT-
TCAAGGTG-GAAC-A-TTCCCTGCGCAA-GTAATAGTCATGT-
GACAAGGGG-GTACA-TTCTGCTCCAA-GGAAATAGTGAATGTT
* * * * * * * * * * * * * * * * * * * * * *

```

CSB1

```

-AATGAGTCTTATTAGAAGGATAACATTAAGTTACGTCATAA
-AGGGAGTCTTATTAAAAAATAACTAAGTATCGCATAA
-TATAAGTCTTATTAGAAGGATAACTAAGGATACTGCATAA
-AACAAGTCTTATTAGAAGGATAACTAAGGATACTGCATAA
-AATAAGACTTTTAGAAGGATAACTAAGGATCAGTGCATAA
-AATAAGACTTTTAGAAGGATAACTAAGGATCAGTGCATAA
AAAATGATATGATA-GAAGAATTTCAT-AACTGATATCATGAGCATAAA
* * * * * * * * * * * * * * * * * * * * * * * * * * *

```

```

GATTTGCTCATTTCTCTGCATCC-----CCAGGATAC
GATTTGCTCATTTCTCTATCATCC-----CCAGGATAC
AATGTGCTCCCCCTCTATCATCC-----CCAGGATAC
AATATGCTCATTTATCTGCTTCC-----CCAGGATAC
GATGTGCTTGTATTCTATCTTCT-----CCAGGATAC
GATGTGCTTGTATTCTATCTTCC-----CCTTGATTG
GGTGGTCTTGTAAATCAAAGATCC-----CTTAGATCAC
TAATTATTTTCCCCTAAATAACCTGATATACGCCCTCTAACAAAC
* * * * *

```

CSB-2

```

CCCCTT-----TTT-GCGCGC-AAAACCCCC-TACCCCCTAAAC
CCCCTT-----TTTTGCGCGC-AAAACCCCC-TACCCCCTAAAC
CCCCTT-----TTT-CGCCT-AAAACCCCC-TACCCCCTAAAC
CCCCTT-----TTT-GCGCGT-AAAACCCCC-TACCCCCTAAAC
CCCCTT-----TTT-GCGCGC-AAAACCCCC-TACCCCCTAAAC
CCCCTT-----TTT-ACGCGCGT-AAAACCCCC-TACCCCCTAAAC
CCCC-----TTTTGCGCGC-AAAACCCCC-CCACCCCCCAAAC
TCTCGTCAAAAGGTTTTGCGT-TAAACCCC-TACCCCCTTTTC
* * * * * * * * * * * * * * * * * * * * * *

```

CSB-3

```

CCCTGAAGTTGCTAAGACCCCTGAAAACCCCCCGGAAACACGGACAC
CCCTGAAGTTGCTAAGACCCCTGAAAACCCCCCGGAAACACGGACAC
TCCTGAAGTTGCTAAGACCCCTGAAAACCCCCCGGAAACACGGACAC
TCCTGAAGTTGCTAACACTCCTGAAAACCCCCCGGAAACACGGACAC
CCCTGAAGTTGCTAACAC-CCTGAAAACCCCCCGGAAACACGGACAC
CCCTGAAGTTGCTAACAC-CCTGAAAACCCCCCGGAAACACGGACAC
CCCTGAAGTTGCTAACAC-CCTGAAAACCCCCCGGAAACACGGACAC
TCCTAAGGTTATCTATACCTCTGAAAACCCCCCGGAAACACGGACAC
TCGAAAGGTTCTGTTACTTCTGAAAACCCCCCGGAAACACGGACAC
* * * * * * * * * * * * * * * * * * * * * *

```

Characterisation of mtDNA heteroplasmy in the control region

Heteroplasmy is the presence of a mixture of more than one type of an organelle genome within a cell or individual. Since every eukaryotic cell contains many hundreds of mitochondria with hundreds of copies of mtDNA, it is possible and indeed very common for mutations to affect only some of the copies, while the remaining ones are unaffected (Grzybowski 2000). In this case the objective was to determine whether or not there was heteroplasmy in the control region in a single individual, and determine the most common number of copies of the repeated element. The most likely mechanism behind the creation of repeated motifs is slipped-strand mispairing (Levinson and Gutman 1987)

Table 13 shows the dominating copy number of a heteroplasmic 61 bp motif found in the control region of Atlantic-, Pacific- and Greenland halibut.

Table 13: Most common copy number of a 61 bp repeat in the HTR region of the control region in Atlantic halibut (Hh), Pacific halibut (Hs) and Greenland halibut (Rh).

Individual	Most common copy number of 61 bp repeat
Hh1	13
Hh2	12
Hh3	19
Hh4	15
Hs1	17-18
Hs2	17
Hs3	Apprx. 19
Hs4	18-19
Rh1	20-21
Rh2	22-23
Rh3	18
Rh4	21

The repeated motif in the control region of Atlantic halibut, Pacific halibut and Greenland halibut was identified with the following program: <http://tandem.bu.edu/trf/trf.html> (Benson 1999). The motif was quite conserved among the three halibut species (Fig 13). The G + C content in the motif was lower than the average for the rest of the mtDNA: 41 % in Atlantic halibut, 41 % in Pacific halibut and 36 % in Greenland halibut.

Hh CAACCCACAAATACCCCTGGCTTATCGTAACCCCACCAGTTGTTTAATAATACCACTTT

Hs CAACCCACAAATACCCCTGGCTTATCACAACCCCACCAGTTGTTTAATAATACCACTTT

Rh CAACCCCTAAAATACCCCTGACCAAAGATAAGCCCACCAGTTATTTAATAATGGTACTTT

Figure 13: Sequence of the HTR motif in the control region of Atlantic halibut (Hh), Pacific halibut (Hs) and Greenland halibut (Rh). Nucleotide sequence from L-strand is shown.

Deducing the number of copies

A common strategy to determine copy number of heteroplasmic repeated motifs is to slice out the band of interest and clone or direct sequence it. In Atlantic halibut, Pacific halibut and Greenland halibut the repeated motif is 61 bp long. Copy number is as high as 20, giving a repeated motif region of about 1200 bp. Even with a sequencing primer located perfectly outside the repeated motif region it would be very difficult to read through this region and determine the number of repeats this way.

L16276

AGCCAA **CGGAGACGTTAAAGGGT** GAGC ACTTTCTTCACGCCGTACATAGTATCCGTGAATGAGTCTTATTAGAAGGATAACATTAAG
TGTATCATGTCATAAGATTGCTATTCTCTGCATCCCCAGGATACCCCTTTTGCCTGCAAAACCCCCCTACCCCCCTAAACCCC
TGAAGTTGCTAAGACCCCTGAAAACCCCCCGGAAACAGGACAAACCTCTGGTAGCTCAGAAAGGG
CAATTACTTT
CAACCACTTT 11 bp motif
CAACCACTTT
CAACCCACAATACCCCTGGCTTATCATAACCCACCAGTTGTTTAATAATACCACTTT 61 bp motif
CAACCCACAATACCCCTGGCTTATCGTAACCCACCAGTTGTTTAATAATACCACTTT
CAACCCACAATACCTAAACCCACAAACGCCGTATCACCTCTCATGATCAAAACT **ACTGTTTAGATTATTCAAG** TATTCCCT

H17250

Figure 14: Part of the sequence of the HTR region of Atlantic halibut 1. Location of HTR primers in red and truncate copy of the 61 bp repeat in blue. Nucleotide sequence of the L-strand is shown.

To confirm a possible heteroplasmy detected in the control region during cloning, primers on each side of the heteroplasmic region were designed (Table 7). MtDNA was not fully assembled when primers were designed, so the primer numbers do not precisely correspond to the actual position in mtDNA. For Atlantic halibut the primers are located 280 bp before and 80 bp after the repeated motifs (Fig 14). This corresponds to and PCR product of 375 bp if no repeated motif is included, only a truncated motif of 15 bp. To determine the sequence of the four shortest bands in figure 13 these were carefully cut out of the gel and cloned. The shortest band contained no repeated motif, only one truncated motif. The rest of the clones contained the truncated motif and one or more copies of the 61 bp repeated motif. Thus it is possible to count the number of bands and determine which number of repeats is most common in each individual. The same approach was used for Pacific halibut. In figure 15 there are some strong, shorter bands. These could be mistaken as the most common copy number in each individual. But they are probably not, because all sequenced clones from the control region contained a larger number of repeated motifs than what is observed among the quite strong, shorter bands. A closer look at the bands for Hh2 reveals that they are slightly longer than the rest of the Atlantic halibut individuals (Fig 15). The explanation is that Hh2 has one extra copy of a 11 bp repeat (Fig 14)

For Greenland halibut primer combination L 16472 – H17300 was used to produce a PCR product that was 154 bp long when containing no repeated motifs and only a truncate motif.

The band was difficult to visualize under the 200 bp marker (Fig 15). But with proper settings of contrast and light it can be seen. There was uncertainty as how to interpret the “extra” band around 500 bp. It was chosen to regard it as a PCR artefact.

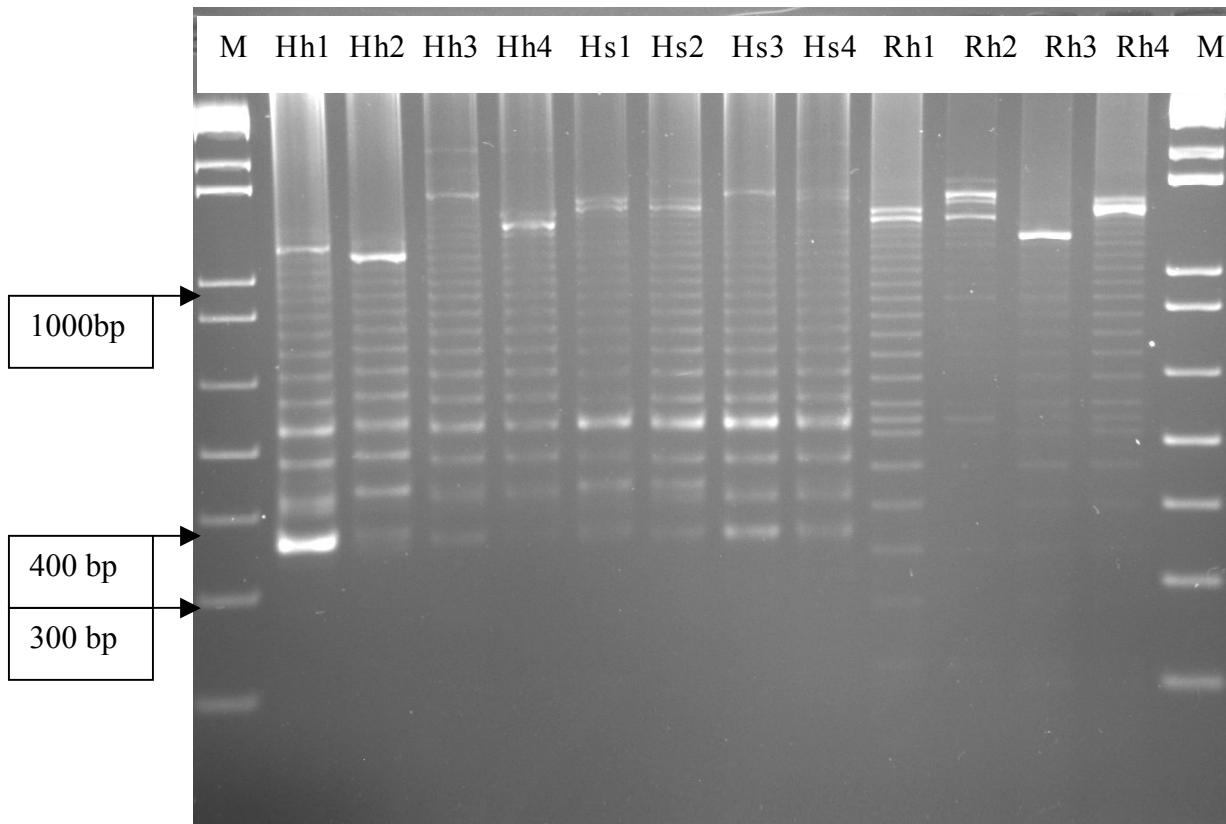


Figure 15: Gel picture of PCR containing the Heteroplasmic Tandem Repeat region in Atlantic halibut (Hh 1-4), Pacific halibut (Hs 1-4) and Greenland halibut (Rh 1-4). Most common copy number of a 61 bp repeat in each individual is seen as the strong bands above the 1000 bp marker. M: 1 kb plus DNA ladder from Invitrogen.

Secondary structure of the repeated motif

Various approaches have been devised for predicting secondary structures of RNA molecules. Many of these search to minimize the free energy of the folded macromolecule, thus searching for the most stable structure (Krane and Raymer 2003). Various number of copies of the HTR motif (Fig 14) was subjected to Zucker's M-fold program (Zuker et al. 1999). One copy of the repeated motif in Atlantic halibut was folded with a free energy of -2,6 kcal/mol, while 13 copies had a free energy of -55 kcal/mol. The same values for Pacific halibut and Greenland halibut was -3,8 kcal/mol and -55 kcal/mol (17 copies), -3,4 kcal/mol and -155kcal/mol (20 copies) respectively (structures not shown). A lower free energy indicates a more stable secondary structure.

Base composition

The G + C content of the mtDNA in Atlantic halibut, Pacific halibut and Greenland halibut varied between 45,1 % and 46,1 % (Table 14). In protein coding genes the G + C content in the 1st codon position was 52,2 % to 53,6 %. While 2nd and 3rd codon position contained 41,5 % to 43,8 % and 43,3 % to 45,7 % G +C respectively. G + C content in tRNA was 47,2% to 47,6%. Also rRNA showed at slightly higher G + C content than total mtDNA with a variation between 47,0 % and 48,9%.

Table 14: Base composition of the H-strand of the Atlantic halibut, Pacific halibut and Greenland halibut mitochondrial genome. Numbers in parenthesis are data for ND-6

Region	G+C content Hh	G+C content Hs	G+C content Rh
Codon pos.			
1 st	53,4 (49,7)	52,2 (50,9)	53,6 (51,7)
2 nd	41,7 (45,7)	43,8 (45,7)	41,5 (44,8)
3 rd	45,7 (39,3)	43,4 (40,4)	43,3 (43,1)
Total protein genes	47,0 (45,4)	46,8 (46,0)	46,2 (46,6)
tRNA	47,6	47,5	47,2
rRNA	47,0	48,4	48,9
Total mtDNA	46,1	45,7	45,1

Codon usage bias

Atlantic halibut

Every organism seems to prefer the use of some equivalent triplet codons over others (Krane and Raymer 2003). CUC, CUA and AUU are the most abundant codons in the protein coding genes of Atlantic halibut (Fig 16). The two codons encode the amino acid leucine, and AUU specify isoleucine. AGG, AGA, and UAA are the least common codons. This should come as no surprise as they all are termination codons.

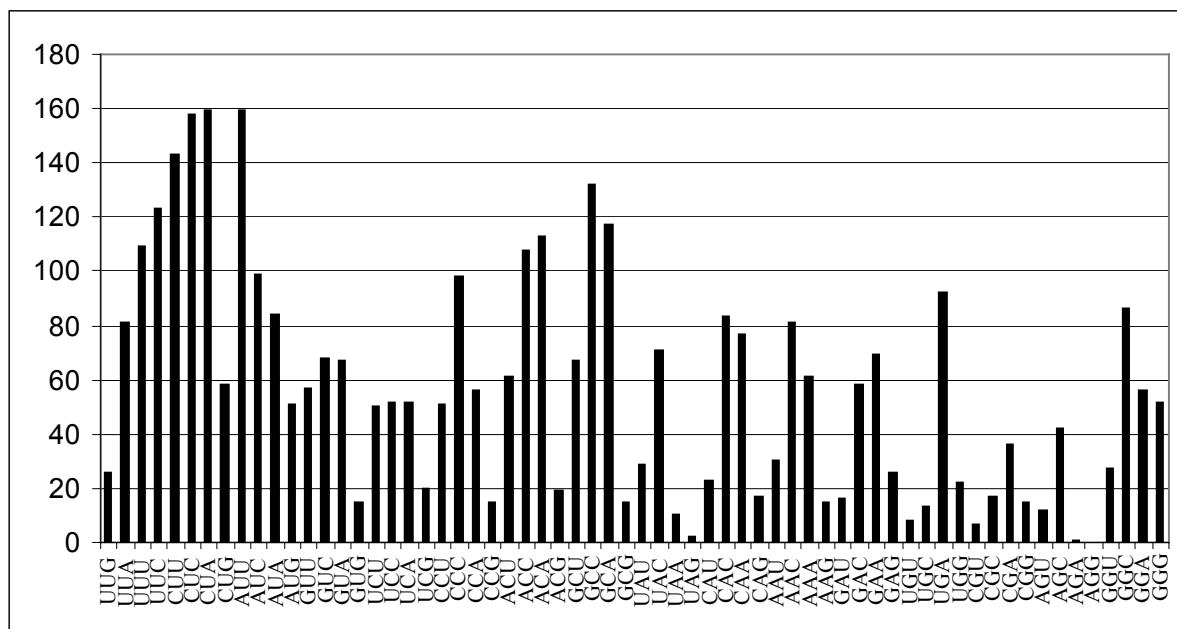


Figure 16: Codon usage in Atlantic halibut (Hh1) mtDNA. Codons along the abscissa and absolute number of each codon in mtDNA protein coding genes along the ordinate axis

In leucine, valine, serine, proline, threonine and alanine there seems to be a non-preference for codons ending with guanine (Table 15). For two-fold degenerate amino acids (tyrosine, histidine, asparagine, lysine, aspartic acid, glutamic acid, cysteine, tryptophan and serine) one codon is clearly dominating over the other. In these amino acids cytosine or adenine is the nucleotide of choice in third codon position.

Table 15: Codon preference calculated for each amino acid in mtDNA proteins of Atlantic halibut

Amino acid	Codon	%									
Phe (F)	TTT	47	Ser (S)	TCT	29	Tyr (Y)	TAT	29	Cys (C)	TGT	38
Phe (F)	TTC	53	Ser (S)	TCC	29	Tyr (Y)	TAC	71	Cys (C)	TGC	62
Leu (L)	TTA	13	Ser (S)	TCA	30	STOP	TAA	83	Trp (W)	TGA	81
Leu (L)	TTG	4	Ser (S)	TCG	12	STOP	TAG	17	Trp (W)	TGG	19
Leu (L)	CTT	23	Pro (P)	CCT	23	His (H)	CAT	21	Arg (R)	CGT	9
Leu (L)	CTC	25	Pro (P)	CCC	45	His (H)	CAC	79	Arg (R)	CGC	23
Leu (L)	CTA	26	Pro (P)	CCA	25	Gln (Q)	CAA	82	Arg (R)	CGA	48
Leu (L)	CTG	9	Pro (P)	CCG	7	Gln (Q)	CAG	18	Arg (R)	CGG	20
Ile (I)	ATT	62	Thr (T)	ACT	20	Asn (N)	AAT	27	Ser (S)	AGT	22
Ile (I)	ATC	38	Thr (T)	ACC	36	Asn (N)	AAC	73	Ser (S)	AGC	78
Met (M)	ATA	62	Thr (T)	ACA	38	Lys (K)	AAA	80	STOP	AGA	100
Met (M)	ATG	38	Thr (T)	ACG	6	Lys (K)	AAG	20	STOP	AGG	0
Val (V)	GTT	28	Ala (A)	GCT	21	Asp (D)	GAT	22	Gly (G)	GGT	12
Val (V)	GTC	33	Ala (A)	GCC	40	Asp (D)	GAC	78	Gly (G)	GGC	39
Val (V)	GTA	32	Ala (A)	GCA	35	Glu (E)	GAA	72	Gly (G)	GGA	25
Val (V)	GTG	7	Ala (A)	GCG	4	Glu (E)	GAG	28	Gly (G)	GGG	24

Pacific halibut

In Pacific halibut (Fig 17) two codons encoding leucine (CUU, CUC) and one encoding isoleucine (AUU) appear with the highest frequencies. The termination codons AGG, AGA, TAG and TAA have the lowest frequencies.

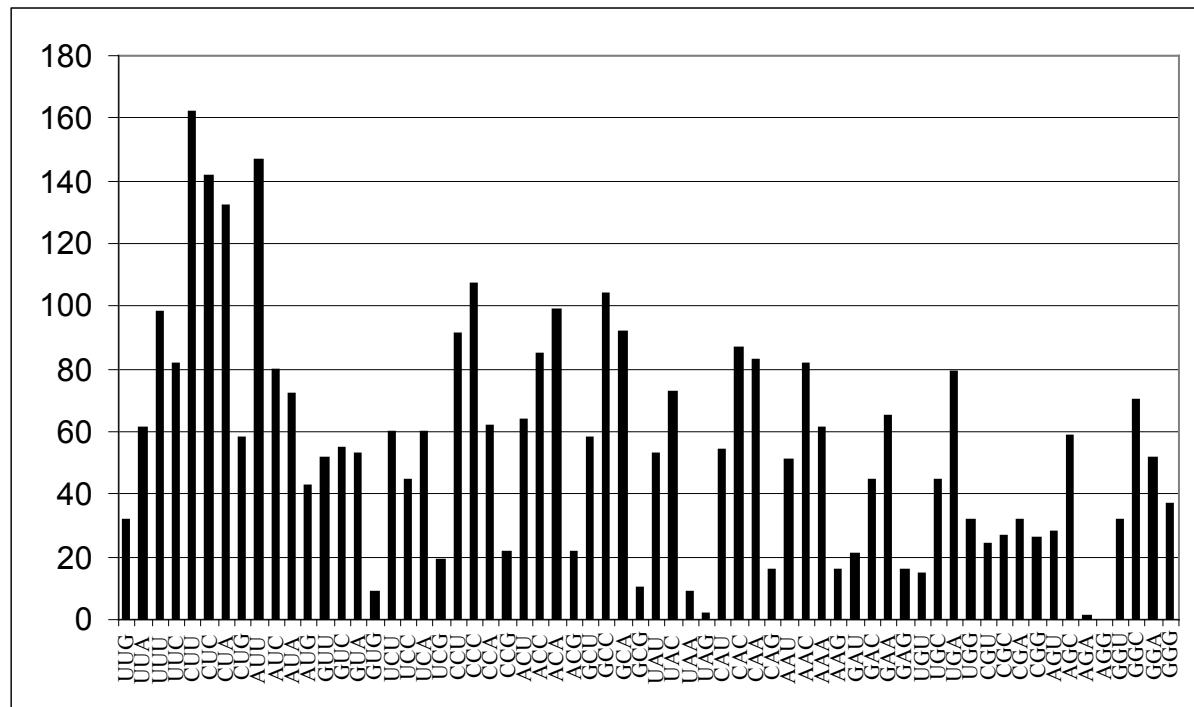


Figure 17: Codon usage in Pacific halibut (Hs1) mtDNA. Codons along the abscissa

and absolute number of each codon in mtDNA protein coding genes along the ordinate axis.

Table 16: Codon preference calculated for each amino acid in mtDNA proteins of Pacific halibut.

Amino acid	Codon	%									
Phe (F)	TTT	54	Ser (S)	TCT	33	Tyr (Y)	TAT	42	Cys (C)	TGT	25
Phe (F)	TTC	46	Ser (S)	TCC	24	Tyr (Y)	TAC	58	Cys (C)	TGC	75
Leu (L)	TTA	10	Ser (S)	TCA	33	STOP	TAA	81	Trp (W)	TGA	71
Leu (L)	TTG	5	Ser (S)	TCG	10	STOP	TAG	19	Trp (W)	TGG	29
Leu (L)	CTT	28	Pro (P)	CCT	33	His (H)	CAT	62	Arg (R)	CGT	22
Leu (L)	CTC	24	Pro (P)	CCC	37	His (H)	CAC	38	Arg (R)	CGC	25
Leu (L)	CTA	23	Pro (P)	CCA	22	Gln (Q)	CAA	84	Arg (R)	CGA	29
Leu (L)	CTG	10	Pro (P)	CCG	8	Gln (Q)	CAG	16	Arg (R)	CGG	24
Ile (I)	ATT	65	Thr (T)	ACT	24	Asn (N)	AAT	38	Ser (S)	AGT	32
Ile (I)	ATC	35	Thr (T)	ACC	31	Asn (N)	AAC	62	Ser (S)	AGC	68
Met (M)	ATA	63	Thr (T)	ACA	37	Lys (K)	AAA	79	STOP	AGA	100
Met (M)	ATG	37	Thr (T)	ACG	8	Lys (K)	AAG	21	STOP	AGG	0
Val (V)	GTT	31	Ala (A)	GCT	22	Asp (D)	GAT	32	Gly (G)	GGT	17
Val (V)	GTC	33	Ala (A)	GCC	39	Asp (D)	GAC	68	Gly (G)	GGC	37
Val (V)	GTA	31	Ala (A)	GCA	35	Glu (E)	GAA	80	Gly (G)	GGG	27
Val (V)	GTG	5	Ala (A)	GCG	4	Glu (E)	GAG	20	Gly (G)	GGG	19

In third codon position of leucine, valine, serine, proline, threonine and alanine there is a strong preference for tyamine, cytosine or adenine over guanine (Table 16). For arginine and glycine there is a relatively even distribution of the respective four possible triplet codons in the protein coding genes.

Greenland halibut

Three codons encoding leucine; CUC, CUA and CUU are most abundant in Greenland halibut (Fig 18). Termination codons AGG, AGA and UAA are the least common codons.

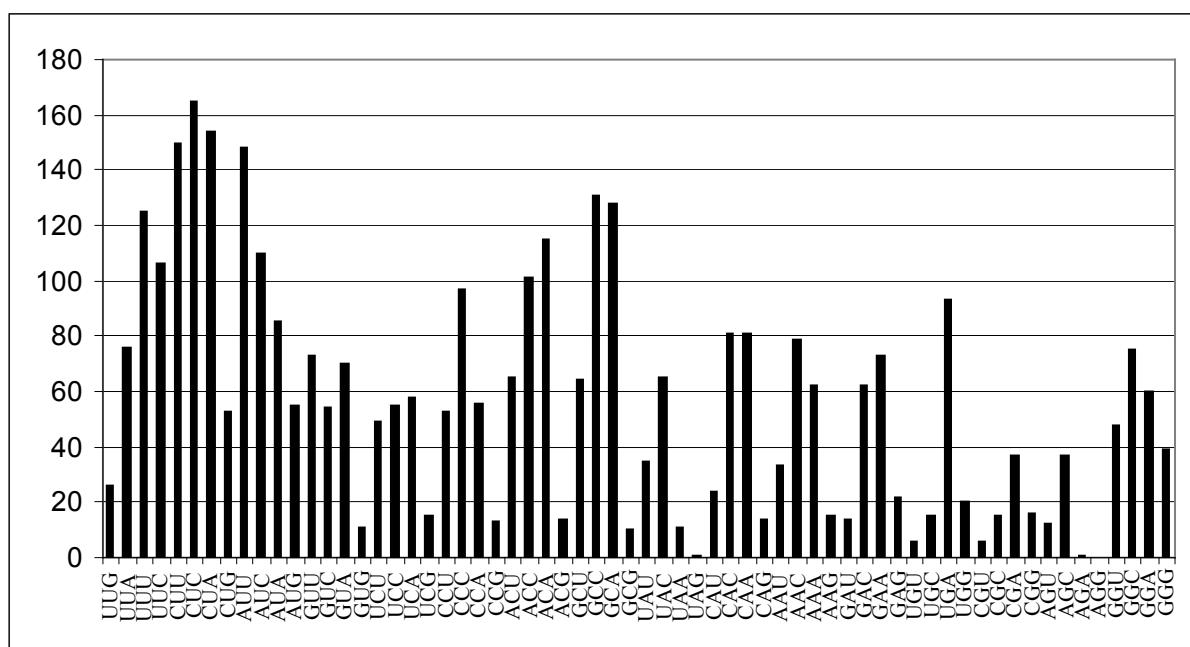


Figure 18: Codon usage in Greenland halibut (Rh1) mtDNA. Codons along the abscissa and absolute number of each codon in mtDNA protein coding genes along the ordinate axis.

Some amino acids show a relatively even distribution of triplet codons, though there seems to be a non-preference for any triplet codons ending with guanine (Table 17). In other amino acids the distribution is skewed. For example, the amino acid arginine is specified by CGA 50 % of the time. Especially amino acids with only two alternative codons have a very uneven distribution of codons, for example aspartic acid, histidine and glutamine.

Table 17: Codon preference calculated for each amino acid in mtDNA proteins of Greenland halibut.

Amino acid	Codon	%	Amino acid	Codon	%	Amino acid	Codon	Amino acid	Codon	
Phe (F)	TTT	54	Ser (S)	TCT	28	Tyr (Y)	TAT	Cys (C)	TGT	28
Phe (F)	TTC	46	Ser (S)	TCC	31	Tyr (Y)	TAC	Cys (C)	TGC	72
Leu (L)	TTA	12	Ser (S)	TCA	33	STOP	TAA	Trp (W)	TGA	69
Leu (L)	TTG	4	Ser (S)	TCG	8	STOP	TAG	Trp (W)	TGG	16
Leu (L)	CTT	24	Pro (P)	CCT	24	His (H)	CAT	Arg (R)	CGT	8
Leu (L)	CTC	26	Pro (P)	CCC	44	His (H)	CAC	Arg (R)	CGC	20
Leu (L)	CTA	25	Pro (P)	CCA	26	Gln (Q)	CAA	Arg (R)	CGA	50
Leu (L)	CTG	9	Pro (P)	CCG	6	Gln (Q)	CAG	Arg (R)	CGG	22
Ile (I)	ATT	57	Thr (T)	ACT	22	Asn (N)	AAT	Ser (S)	AGT	24
Ile (I)	ATC	43	Thr (T)	ACC	34	Asn (N)	AAC	Ser (S)	AGC	76
Met (M)	ATA	60	Thr (T)	ACA	39	Lys (K)	AAA	STOP	AGA	100
Met (M)	ATG	40	Thr (T)	ACG	5	Lys (K)	AAG	STOP	AGG	0
Val (V)	GTT	24	Ala (A)	GCT	19	Asp (D)	GAT	Gly (G)	GGT	22
Val (V)	GTC	36	Ala (A)	GCC	39	Asp (D)	GAC	Gly (G)	GGC	34
Val (V)	GTA	39	Ala (A)	GCA	38	Glu (E)	GAA	Gly (G)	GGA	27
Val (V)	GTG	6	Ala (A)	GCG	4	Glu (E)	GAG	Gly (G)	GGG	17

In figure 19 the frequency of every codon is compared in Atlantic halibut, Pacific halibut and Greenland halibut. The over-all codon frequency patterns are highly conserved among the halibut species. However, for some codons Pacific halibut appears to differ from Atlantic halibut and Greenland halibut. Compared to Atlantic halibut and Greenland halibut codons UUU, UUC (both phenylalanine), CUA (leucine), AUG (methionine), ACA (threonine) and GCC (alanine) have lower frequencies in Pacific halibut. CGU and CGG (both arginine) have higher frequencies.

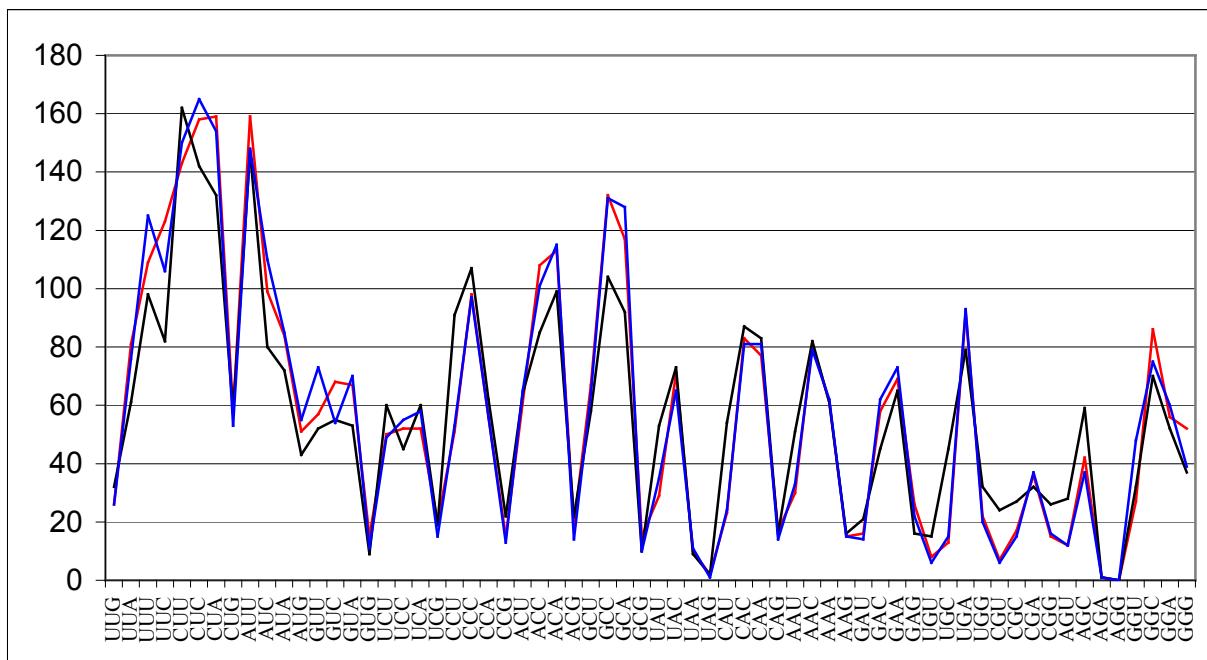


Figure 19: Comparison of codon preference in mtDNA proteins of Atlantic halibut (Hh1, red), Pacific halibut (Hs1, black) and Greenland halibut (Rh1, blue). Absolute abundance of codons along the ordinate-axis.

Intraspecific variation in nucleotide sequences

Atlantic halibut

A comparison of the complete mtDNA from four individuals of Atlantic halibut revealed 91 variable nucleotide positions, also known as single nucleotide polymorphisms (SNP) (Table 18). There were 79 singleton variable positions and 12 parsimony informative sites. In protein coding genes 83 % of the variability was in the 3rd codon position, 3 % in 2nd position and 14 % in 1st codon position. The transition to transversion ratio was 3,8. One insertion was observed in the control region of Hh2.

Table 18: Single nucleotide polymorphisms in mtDNA from four individuals of Atlantic halibut, except for HTR region. Nucleotide positions are based on Hh1.

Region	Phe	12S rRNA				Val	16S rRNA				Leu	ND1		
Position	46	312	340	734	1038	1372	1909	1999	2866	3117	3237	3315		
Codon pos											3	3	3	
Hh1	G	C	A	A	G	A	A	G	G	G	A	G		
Hh2	G	A	A	C	A	G	G	G	G	G	G	A		
Hh3	A	C	G	C	A	A	A	A	A	A	A	A		
Hh4	A	C	A	C	A	A	A	G	A	A	A	A		

Region	ND1					ND2								
Position	3507	3585	3669	3720	3849	4318	4384	4462	4501	4507	4522	4642		
Codon pos	3	3	3	3	3	3	3	3	3	3	3	3		
Hh1	A	A	T	G	G	G	A	G	G	G	G	G		
Hh2	G	A	C	A	A	A	G	A	A	C	G	G		
Hh3	A	A	C	A	G	G	A	A	A	A	G	C		
Hh4	A	G	C	A	G	G	A	A	A	A	G	C		

Region	CO1						CO2			ATP8		ATP6
Position	5633	5971	6220	6244	6337	6379	6442	7444	7616	8046	8072	8234
Codon pos	1	3	3	3	3	3	3	2	3	1	3	1
Hh1	C	A	C	C	G	A	G	T	C	A	A	A
Hh2	C	A	C	T	A	G	G	T	T	A	A	G
Hh3	T	C	T	C	G	A	A	C	C	G	G	A
Hh4	C	C	T	C	G	A	A	T	C	A	A	A

Region	ATP6				CO3			ND3			ND4L		ND4	
Position	8290	8662	8680	8815	9405	9418	9821	9992	10175	10313	10599	10748	11129	
Codon pos	3	3	3	3	3	1	3	3	3	2	1	3	3	
Hh1	G	T	A	C	A	A	G	C	G	G	G	G	A	
Hh2	A	C	G	T	A	A	G	C	A	G	G	G	A	
Hh3	G	T	A	C	G	G	A	T	G	G	C	A	G	
Hh4	G	T	A	C	G	A	A	T	G	C	G	G	G	

Region	His				ND5							ND6		
Position	11778	11809	12079	12363	12654	12984	13206	13642	13684	13893	13914	14061		
Codon pos			1	3	3	3	3	1	1	3	3	3		
Hh1	T	T	A	G	A	A	A	C	T	T	T	G		
Hh2	T	T	G	A	G	G	G	T	C	C	T	A		
Hh3	C	C	A	A	A	A	G	C	T	T	C	G		
Hh4	T	T	A	A	A	A	G	C	T	T	C	G		

Region	ND6						Cyt-b					Control region		
Position	14115	14283	14538	14901	14910	15051	15453	15492	15801	15828	15867	15879		
Codon pos	3	3	3	3	3	3	3	3						
Hh1	C	T	C	A	A	G	A	C	C	C	A	T		
Hh2	T	C	A	G	G	A	G	T	T	A	G	C		
Hh3	C	T	C	A	A	G	A	T	C	C	A	T		
Hh4	C	T	C	A	A	G	A	T	C	C	A	T		

Region	Control region											
Position	15981	16260	16277	16327	16352	16400	16402	16404	16405	16408	16410	16411
Hh1	C	A	G	C	G	G	-	A	T	C	A	G
Hh2	A	G	A	T	A	A	A	C	A	T	T	A
Hh3	C	G	A	T	A	G	-	A	T	C	A	G
Hh4	C	G	A	T	A	G	-	A	T	C	A	G

Region	Control region					
Position	16510	16515	17363	17436	17540	17557
Hh1	T	C	G	A	G	C
Hh2	G	G	A	G	T	G
Hh3	T	C	G	A	G	C
Hh4	T	C	G	A	G	C

Nucleotide diversity is a concept in molecular genetics that is used to measure the degree of polymorphism within a population. It was first introduced by Nei and Li (1979). It is defined as the average number of nucleotide differences per site between any two DNA sequences chosen randomly from the sample population. The highest nucleotide diversity was found in the control region at 1,05%, followed by ATP8 0,60% and ND6 0,51. The lowest nucleotide diversity was found in 16S rRNA with 0,087%. Both COII and COIII had a nucleotide diversity of 0,15% (Table 19).

Table 19: Nucleotide diversity in mtDNA from 4 individuals of Atlantic halibut

Region	Variable sites	Variable sites / sites in gene (%)	Nucleotide diversity (%)
12S rRNA	3	0,32	0,16
16S rRNA	3	0,17	0,087
ND1	8	0,82	0,43
ND2	7	0,67	0,35
COI	7	0,44	0,26
COII	2	0,29	0,15
ATP8	2	1,2	0,60
ATP6	5	0,73	0,37
COIII	2	0,25	0,15
ND3	2	0,57	0,38
ND4L	2	0,67	0,33
ND4	4	0,29	0,16
ND5	7	0,38	0,19
ND6	5	1,0	0,51
Cyt-b	6	0,52	0,26
Control region	22	2,0	1,1

Pacific halibut

The four individuals of Pacific halibut had a total of 99 variable nucleotide positions (Table 20). In protein coding genes 94 % of the variability was found in the 3rd codon position. 3 % of variability was found in both 1st and 2nd codon position. The transition to transversion ratio was 4,7.

Table 20: Single nucleotide polymorphisms in mtDNA except for HTR region in four individuals of Pacific halibut. Nucleotide positions are based on Hs 1.

Region	12SrRNA		16SrRNA			ND1				Gln	Met	ND2	
Position	953		1408	1518	1529	3330	3516	3564	3735	3998	4055	4150	4342
Codon pos						3	3	3	3			3	3
Hs1	A	G		A	A	C	A	G	A	A	G	T	C
Hs2	G	A		A	G	T	G	A	G	C	A	C	T
Hs3	A	A		G	A	T	A	G	A	C	A	C	T
Hs4	A	A		G	A	T	A	G	A	C	A	C	T

Region	ND2					COI							
Position	4507	4849	4957	5074	5800	5923	6133	6199	6331	6403	6475	6541	
Codon pos	3	3	3	3	3	3	3	3	3	3	3	3	3
Hs1	G	A	A	C	C	G	G	G	C	T	C	T	
Hs2	A	G	G	T	T	A	A	G	T	T	T	C	
Hs3	A	A	G	T	T	A	A	A	T	C	C	C	
Hs4	A	A	G	T	T	A	A	G	T	C	C	C	

Region	COI		Asp		COII		ATP6					
Position	6571	6763	7188	7223	7616	8212	8251	8284	8404	8617	8656	8662
Codon pos	3	3			3	3	3	3	3	3	3	3
Hs1	C	A	G	C	C	T	G	G	A	T	C	T
Hs2	T	A	A	C	T	A	A	A	G	T	C	C
Hs3	T	G	A	C	T	A	G	A	G	T	C	T
Hs4	T	A	A	A	T	A	G	A	G	C	T	T

Region	ATP6		COIII				ND3	ND4				ND4L
Position	8668	8853	9078	9165	9468	9594	9836	9926	10254	10263	10271	10935
Codon pos	3	3	3	3	3	3	3	3	3	3	2	1
Hs1	A	A	A	A	G	C	G	T	A	C	A	C
Hs2	G	C	T	G	G	C	A	A	C	T	G	G
Hs3	A	C	T	G	G	C	A	A	C	T	G	C
Hs4	A	A	T	G	A	T	A	A	C	T	G	C

Region	ND4L						ND5					
Position	10943	11033	11441	11474	11655	11681	11711	11997	12157	12310	12586	12595
Codon pos	3	3	3	3	1	3	3	2	3	3	3	3
Hs1	T	G	A	G	A	A	T	A	T	G	T	A
Hs2	C	A	G	A	C	G	T	A	T	A	A	A
Hs3	C	A	A	C	G	T	A	C	G	A	G	
Hs4	C	A	A	A	C	G	C	G	C	G	A	A

Region	ND5								ND6		Cyt-b	
Position	13015	13018	13066	13114	13378	13453	13459	13573	13654	13975	14041	14653
Codon pos	3	3	3	3	3	3	3	3	3	3	3	3
Hs1	T	A	T	C	G	C	T	C	T	T	T	A
Hs2	C	G	C	C	A	T	G	T	C	T	T	A
Hs3	C	G	T	C	A	T	T	C	C	C	C	G
Hs4	C	G	T	T	A	T	T	C	C	T	T	G

Region		Cyt-b						Control region					
Position		14719	14839	14971	15217	15274	15475	15876	15878	15900	15989	16255	16264
Codon pos		3	3	3	3	3	3						
Hs1		C	G	T	C	C	G	G	G	A	T	T	T
Hs2		T	A	C	C	T	A	A	A	G	C	T	T
Hs3		T	A	C	T	T	A	A	A	G	C	A	G
Hs4		T	A	C	C	T	A	A	A	G	C	T	T

Region		Control region											
Position		16279	16281	16314	16336	16349	16352	16383	16410	16432	16472	16518	17599
Hs1		G	G	G	G	A	A	G	A	G	G	G	G
Hs2		G	G	G	G	A	A	G	A	G	G	G	C
Hs3		A	A	A	T	A	A	A	C	A	C	A	C
Hs4		G	G	G	T	G	G	G	A	G	G	G	C

Region		Control region		
Position		17689	17758	17807
Hs1		G	A	A
Hs2		G	G	G
Hs3		G	G	G
Hs4		A	G	G

The three regions with the highest nucleotide diversity was the control region 0,97%, ATP6 0,58% and ND4L 0,51. ATP8, 12S rRNA and COII had the lowest nucleotide diversity with 0%, 0,053 % and 0,072% respectively 8 (Table 21).

Table 21: Nucleotide diversity in mtDNA from 4 individuals of Pacific halibut

Region	Variable sites	Variable sites / sites in gene (%)	Nucleotide diversity (%)
12S rRNA	1	0,11	0,053
16S rRNA	3	0,18	0,097
ND1	4	0,41	0,21
ND2	6	0,57	0,29
COI	10	0,64	0,33
COII	1	0,15	0,072
ATP8	0	0	0
ATP6	8	1,2	0,59
COIII	5	0,64	0,34
ND3	2	0,57	0,28
ND4L	3	1,0	0,51
ND4	8	0,58	0,29
ND5	14	0,76	0,39
ND6	2	0,38	0,19
Cyt-b	7	0,60	0,31
Control region	21	1,9	0,97

Greenland halibut

The number of single nucleotide polymorphisms found in four individuals of Greenland halibut was 110 (Table 22). Of the nucleotide variability found in protein coding genes 89 % was in 3rd codon position, 10 % was in 1st codon position and 1 % was in 2nd codon position. The ratio transitions to transversions were 5,5.

Table 22: Single nucleotide polymorphisms (SNP) in mtDNA except for HTR region in four individuals of Greenland halibut. Nucleotide positions are based on Rh1.

Region	Phe	16S rRNA						ND1								
Position	47	1262	1540	1815	2261	2937	2972	3146	3161	3197	3296	3299	3398	3503	3551	3605
Codon pos.						1	3	3	3	3	3	3	3	3	3	3
Rh1	A	A	A	A	A	C	G	A	G	G	A	A	A	A	A	C
Rh2	G	A	A	A	A	T	A	G	G	G	G	G	G	G	G	T
Rh3	G	A	A	T	G	T	G	A	G	C	G	G	G	G	G	C
Rh4	G	G	T	A	A	T	A	G	C	G	G	G	G	G	G	T
Region	ND2												Trp	Asn		
Position	4089	4194	4422	4440	4452	4539	4542	4560	4602	4792	4872	4884	5141	5336		
Codon pos.	3	3	3	3	3	3	3	3	3	1	3	3				
Rh1	T	G	A	T	A	T	A	G	G	C	G	G	A	T		
Rh2	C	A	G	C	T	C	G	A	C	C	A	A	G	C		
Rh3	C	A	G	C	T	C	G	A	C	T	G	A	A	T		
Rh4	C	A	G	C	T	C	G	A	C	C	A	A	A	A		
Region	COI				COII				ATP6	COIII	Gly		ND4L			
Position	6510	6555	6774	6849	7597	7615	7618	7669	7831	8295	9461	9660	10218	10247		
Codon pos.	3	3	3	3	3	3	3	3	3	3	3	3		1	3	
Hh1	A	C	G	A	C	T	G	C	A	A	C	C	T	G		
Hh2	G	G	A	T	T	T	A	T	G	G	T	T	C	A		
Hh3	G	G	A	T	T	C	A	T	G	G	T	C	C	A		
Hh4	G	G	A	T	T	T	A	T	G	G	T	C	C	A		
Region	ND4										HIS	ND5				
Position	10436	10612	10876	10928	11032	11119	11200	11287	11313	11650	11833	12174				
Codon pos.	1	3	3	1	3	3	3	3	2	3		3				
Hh1	A	A	T	T	G	A	A	G	C	G	G	C				
Hh2	A	G	C	C	A	A	G	A	C	A	A	T				
Hh3	A	A	C	C	A	A	G	A	C	A	A	T				
Hh4	T	G	C	C	A	G	G	A	T	A	A	T				
Region	ND5															
Position	12342	12348	12528	12534	12585	12960	13023	13107	13156	13191	13308	13347				
Codon pos.	3	3	3	3	3	3	3	3	1	3	3	3				
Hh1	G	G	C	G	G	T	G	A	A	T	T	C				
Hh2	A	A	T	A	A	C	A	G	G	T	T	T				
Hh3	A	A	T	A	A	C	A	G	A	T	C	T				
Hh4	A	A	T	A	A	C	A	G	A	C	T	T				
Region	ND5				ND6				Cyt-b							
Position	13425	13543	13794	13908	13947	13959	14067	14166	14586	14724	14838	14893				
Codon pos.	3	1	3	3	3	3	3	3	3	3	3	1				
Hh1	C	T	A	C	T	G	C	C	C	G	C	G				
Hh2	T	T	G	T	C	G	T	T	T	A	G	A				
Hh3	T	T	G	T	T	A	T	T	T	A	G	A				
Hh4	T	C	G	T	C	G	T	T	T	A	G	A				

Region		Cyt-b						Control region					
Position		15006	15159	15372	15435	15510	15516	15724	15842	15859	16257	16263	16266
Codon pos.		3	3	3	3	3	3						
Hh1		C	T	T	T	A	A	G	T	G	A	G	T
Hh2		T	A	C	T	G	G	A	C	A	A	A	G
Hh3		T	A	C	C	A	G	A	C	G	T	A	G
Hh4		T	A	C	T	G	G	A	C	A	A	A	T

Region		Control region											
Position		16281	16283	16286	16289	16296	16314	16316	16351	16354	16355	16377	16384
Codon pos.													
Hh1		G	G	G	T	T	T	A	G	G	G	T	G
Hh2		A	A	G	C	G	G	G	A	A	A	G	A
Hh3		A	A	T	C	G	G	G	A	G	G	T	A
Hh4		A	A	G	C	G	G	G	A	G	G	T	A

Region		Control region					
Position		16472	16518	17888	17912	17915	17942
Codon pos.							
Hh1		G	G	A	T	G	G
Hh2		A	A	G	C	A	A
Hh3		A	G	G	C	A	A
Hh4		A	A	G	C	A	A

The highest nucleotide diversity in Greenland halibut was found in the control region with 1,1%, ND1 0,62% and ND2 0,59%. Three genes showed no nucleotide diversity: 12S rRNA, ATP8 and ND3 (Table 23).

Table 23: Nucleotide diversity in mtDNA from 4 individuals of Greenland halibut

Region	Variable sites	Variable sites / sites in gene (%)	Nucleotide diversity (%)
12S rRNA	0	0	0
16S rRNA	4	0,23	0,12
ND1	11	1,1	0,62
ND2	12	1,1	0,59
COI	4	0,26	0,13
COII	5	0,72	0,36
ATP8	0	0	0
ATP6	1	0,15	0,073
COIII	1	0,13	0,064
ND3	0	0	0
ND4L	2	0,67	0,34
ND4	10	0,72	0,37
ND5	16	0,87	0,44
ND6	5	0,96	0,51
Cyt-b	10	0,88	0,45
Control region	24	2,2	1,1

Nucleotide diversity compared in Atlantic-, Pacific- and Greenland halibut

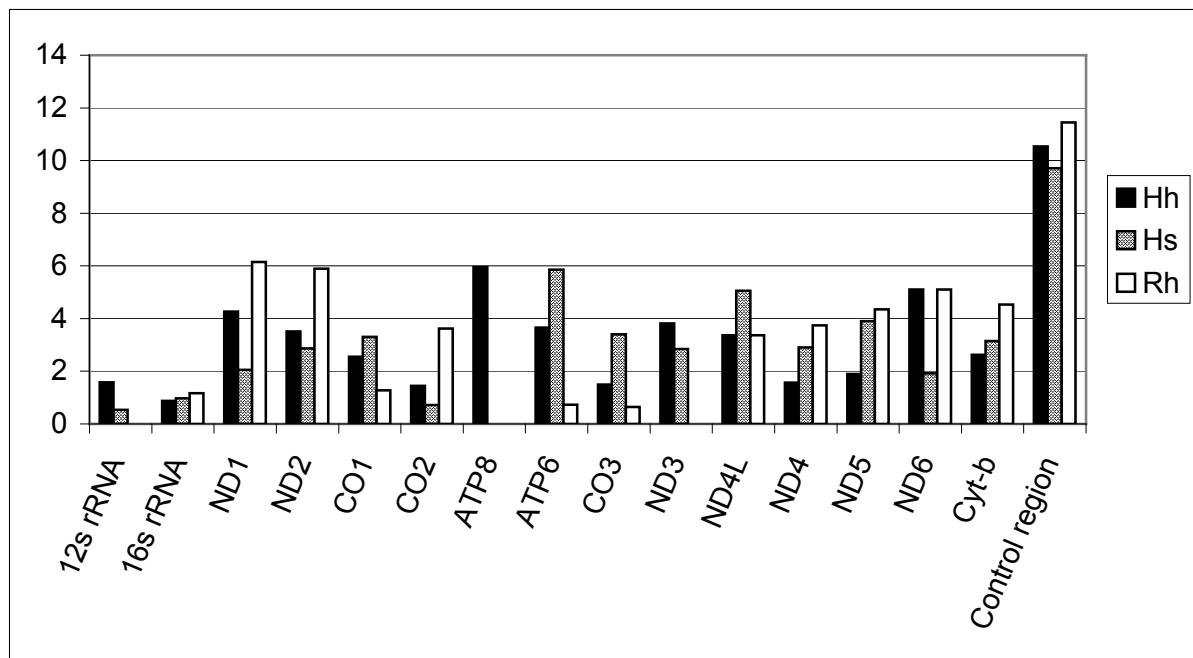


Figure 20: Comparison of nucleotide diversity in different regions of mtDNA in Atlantic halibut (blue), Pacific halibut (pink) and Greenland halibut (yellow). Respective regions along the abscissa and nucleotide diversity along the ordinate axis per thousand.

When comparing the nucleotide diversity in the same gene of Atlantic halibut, Pacific halibut and Greenland halibut similarities were discovered in some genes, while other genes had quite different nucleotide diversities (Fig 20). The two ribosomal RNA genes showed a very low nucleotide diversity in all species. ATP8 had no nucleotide substitutions in Pacific halibut and Greenland halibut, but a relatively large (0,6%) nucleotide diversity in Atlantic halibut. In ND3 Greenland had no nucleotide polymorphisms and both Atlantic halibut and Pacific halibut had a relatively equal higher nucleotide diversity. ND4L, ND4, ND5-6, Cytb and control region had relatively similar nucleotide diversities, when compared to the large differences in some of the other genes.

SNPs in 30 individuals of Atlantic halibut

When the complete mtDNA sequence from four individuals of Atlantic halibut had been determined, regions with high concentration of single nucleotide polymorphisms and interesting substitutions (i.e. transversions) were identified. 800 bp from ND2, COI and the

control region respectively from 30 individuals of the Atlantic halibut brood stock population at Mørkvedbukta Research Station was determined. For primers and fragments see table 9.

Half of the individuals had genotype f (table 24). This observation and additional information available suggest that they probably have the same mother.

Table 24: Single nucleotide polymorphisms (SNPs) in selected regions within the genes ND2, COI and the control region of 30 individuals of Atlantic halibut

Region	ND2								COI				Control region			Genotype
	Pos	4450	4501	4507	4718	4828	4849	4882	5971	6280	6442	15867	15921	16114		
Codon	3	3	3	1	3	3	3	3	3	3	3					
pos																
Hh1	A	A	A	G	T	G	A	A	A	G	A	T	A	a		
Hh2	A	A	A	G	T	G	A	A	A	G	A	C	A	b		
Hh3	G	A	A	G	T	G	A	A	C	G	A	T	A	c		
Hh4	G	G	G	G	T	G	A	A	A	G	A	C	A	d		
Hh5	G	G	G	G	T	G	A	A	A	G	A	C	A	d		
Hh6	G	A	A	G	T	G	A	A	C	G	A	C	A	c		
Hh7	G	A	A	G	T	G	A	A	C	G	A	C	A	c		
Hh8	G	A	A	A	T	G	A	A	A	G	G	C	A	e		
Hh9	G	G	G	T	G	A	A	A	A	G	A	C	A	f		
Hh10	G	G	G	G	T	G	A	A	A	G	A	C	A	f		
Hh11	G	G	G	G	T	G	A	A	A	G	A	C	A	f		
Hh12	G	G	G	G	T	G	A	A	A	G	A	C	A	f		
Hh13	G	G	G	G	T	G	A	A	A	G	A	C	A	f		
Hh14	G	G	G	G	T	G	A	A	A	G	A	C	A	f		
Hh15	G	G	G	G	T	G	A	A	A	G	A	C	A	f		
Hh16	G	G	G	G	T	G	A	A	A	G	A	C	A	f		
Hh17	G	G	G	G	T	G	A	A	A	G	A	C	A	f		
Hh18	G	G	G	G	T	G	A	A	A	G	A	C	A	f		
Hh19	G	G	G	G	T	G	A	A	A	G	A	C	A	f		
Hh20	G	G	G	G	T	G	A	A	A	G	A	C	A	f		
Hh21	G	G	G	G	T	G	A	A	A	G	A	C	A	f		
Hh22	G	G	G	G	T	G	A	A	A	G	A	C	A	f		
Hh23	G	G	G	G	T	G	A	A	A	G	A	C	A	f		
Hh24	G	A	A	G	T	A	G	C	A	A	G	C	A	g		
Hh25	G	A	A	A	T	G	A	A	A	G	G	C	A	e		
Hh26	G	A	A	G	C	G	A	C	A	A	G	C	A	h		
Hh27	A	A	A	G	T	G	A	A	A	G	A	C	A	b		
Hh28	G	A	A	G	C	G	A	C	A	A	A	C	C	i		
Hh29	G	A	A	G	T	G	A	A	A	G	A	C	A	j		
Hh30	G	A	A	G	T	G	A	C	A	A	A	C	A	k		

Intraspecific variation in amino acid sequences

A table of physiochemical distances, based on properties of the amino acid like polarity, molecular volume and chemical composition, was used to investigate amino acid substitutions (Grantham 1974). The smallest distance is 5, while 215 is the highest possible distance.

Atlantic halibut

A total of ten amino acid substitutions were observed between four individuals of Atlantic halibut (Table 25). The Pro→Ser substitution in COI is a change from a non-polar to a polar amino acid, both being neutral. The same is true for the Thr → Val substitution in COIII. In ND4L the Arg → Gln represents a change from a basic to a neutral, non-polar amino acid. The ND4 Cys → Arg substitution is a change from a neutral, polar to a basic amino acid. In ND5 the neutral, polar Thr is exchanged with the neutral, but non-polar Ala. The other amino acid substitutions were relatively innocent.

Table 25: Substitutions of amino acids in mtDNA in four individuals of Atlantic halibut

Region	COI	COII	ATP8	ATP6	COIII	ND4L	ND4L	ND4	ND4	ND5
Dist. ^A	74	98	29	29	69	43	112	32	180	58
Hh1	Pro	Leu	Ile	Ile	Thr	Arg	Cys	Val	Cys	Thr
Hh2	Pro	Leu	Ile	Val	Thr	Gln	Cys	Val	Cys	Ala
Hh3	Ser	Pro	Val	Ile	Val	Arg	Cys	Leu	Arg	Thr
Hh4	Pro	Leu	Ile	Ile	Thr	Arg	Ser	Val	Cys	Thr

A: Physiochemical distance (Grantham 1974)

Pacific halibut

The four amino acid substitutions in four individuals of Pacific halibut (Table 26) appeared to be relatively non-dramatic, as all changes involved amino acids belonging to the same group. The substitution in ND5 have a very high distance value, even though they are both polar, hydrophilic amino acids.

Table 26: Substitutions of amino acids in mtDNA in four individuals of Pacific halibut

Region	ND4L	ND4	ND4	ND5
Dist. ^A	46	32	5	194
Hs1	Asn	Leu	Ile	Tyr
Hs2	Ser	Val	Leu	Tyr
Hs3	Ser	Val	Leu	Tyr
Hs4	Ser	Val	Leu	Cys

A: Physiochemical distance (Grantham 1974)

Greenland halibut

Five substitutions in amino acids were registered in four individuals of Greenland halibut (Table 27). Of these both the Ser → Leu substitution in ND4 and the Thr → Ala substitution in ND5 is a change from a neutral and from a polar to a neutral, non-polar amino acid. In ND5 the neutral, polar Tyr is also exchanged with the basic His. Finally, in Cyt-b the neutral, polar Asn is replacing the acidic Asp.

Table 27: Substitutions of amino acids in mtDNA

in four individuals of Greenland halibut

Region	ND4	ND4	ND5	ND5	Cyt-b
Dist. ^A	58	145	58	83	23
Rh1	Thr	Ser	Thr	Tyr	Asp
Rh2	Thr	Ser	Ala	Tyr	Asn
Rh3	Thr	Ser	Thr	Tyr	Asn
Rh4	Ser	Leu	Thr	His	Asn

A: Physiochemical distance (Grantham 1974)

Interspecific variation in amino acid sequences

A comparison of the amino acid sequence form Atlantic halibut 1, Pacific halibut 1 and Greenland halibut 1 revealed 70 substitutions, but no indels of amino acids (Table 28). COI, ATP8 and ND4L were perfectly conserved. ND5, ND4, ND2 and Cyt-b showed the greatest amount of variability, with 29, 10, 9 and 8 substitutions, respectively. Each end of a polypeptide seemed to be more prone to accumulating substitutions. Specially the first and last hundred amino acids of ND5 were regions with a very high concentration of amino acid substitutions.

Table 28: Amino acid substitutions between Atlantic halibut 1 (Hh1), Pacific halibut 1 (Hs1) and Greenland halibut (Rh1)

ND1		ND2				COII		ATP6	COIII		ND3		ND4																			
Hh1	I	V	L	V	T	L	T	I	V	V	T	I	I	F	S	S	I	K	H	L	V	S	V	S	L	S	A	L	T	V		
Hs1	I	L	L	V	T	V	T	I	V	V	T	I	I	F	S	S	I	K	H	L	V	N	V	S	L	S	A	I	T	V		
Rh1	V	L	M	I	A	L	A	V	M	M	A	L	M	Y	V	T	V	M	V	I	I	S	I	N	S	G	T	I	A	I		
ND5																					ND6											
Hh1	Y	I	I	R	D	T	N	V	S	S	H	L	A	T	S	I	E	P	N	M	A	V	I	T	A	A	V	L	T	L	G	T
Hs1	Y	I	I	R	G	T	N	I	N	S	H	I	A	T	S	I	E	P	G	M	A	V	I	T	S	A	V	M	T	L	G	T
Rh1	H	V	M	Q	E	S	D	I	N	L	N	L	V	S	N	L	K	S	D	V	T	K	V	I	S	S	I	L	A	M	A	I
Cyt-b																																
Hh1	I	L	V	I	N	L	S	W																								
Hs1	I	L	V	I	N	L	S	W																								
Rh1	V	I	I	V	D	V	A	L																								

Non-conservative amino acid substitutions

All amino acid substitutions have a mutation score. A negative score denotes a non-conservative replacement for transmembrane proteins. Scores were determined by the amino acid mutation matrix for transmembrane proteins (Jones et al. 1994). Nine of the 70 substitutions between Atlantic halibut, Pacific halibut and Greenland halibut had a negative score (Table 29). Four were in ND5, two in COIII and one in ATP6, ND4 and Cyt-b respectively. In all mutations that had a negative score Atlantic halibut and Pacific halibut had the same amino acid, and Greenland halibut had a different one. The exception was one substitution in ND5 where all species had a different amino acid. The mutation between Atlantic halibut and Pacific halibut (N/G) had a score of -2. Amino acid substitutions with a negative mutation score were also observed between four individuals of Atlantic halibut and four individuals of Greenland halibut

Table 29: Amino acid substitutions showing a negative mutation matrix score

	ATP6	COIII	ND4	ND5	Cyt-b	COII	ND4	ND4
Hh1	S	K	H	L	S P V N*	W Hh1	L C Rh1	S
Hs1	S	K	H	L	S P V G	W Hh2	L C Rh2	S
Rh1	V	M	V	S	L S K D	L Hh3	P R Rh3	S
						Hh4	L C Rh4	L
Score	-1	-1	-4	-2	-2 -1 -4 -2	-2	-3	-4

*Score for mutation N-G is -2

Transmembrane helices

The number of helices was predicted by the help of the TMHMM 2.0 program (<http://www.cbs.dtu.dk/services/TMHMM/>). Results are presented in table 30. The same number of transmembrane helices was predicted in all individuals of Atlantic halibut, Pacific halibut and Greenland halibut. So the amino acid substitutions observed in some individuals did not have any effect on transmembrane helices.

Table 30: Number of predicted transmembrane helices in 13 mitochondrial proteins of Atlantic halibut (Hh), Pacific halibut (Hs) and Greenland halibut (Rh).

	No of transmembrane helices		
	Hh	Hs	Rh
ND1	8	8	8
ND2	9	9	9
COI	11	11	11
COII	2	2	2
ATP8	1	1	1
ATP6	6	6	6
COIII	6	6	6
ND3	3	3	3
ND4L	3	3	3
ND4	12	12	12
ND5	16	16	16
ND6	5	5	5
Cyt-b	9	9	9

Phylogenetic investigation of mtDNA sequences from Pleuronectiform species.

Three different trees were created, each by the help of the methods Neighbor Joining, Minimum Evolution and Maximum Parsimony. NJ and ME trees based on nucleotide sequence were inferred with the Kimura 2 Parameter distance model. This model take into consideration that transitions occur more often than transversion, which were the case here. The K2P model gives better distance estimates than the often used Jukes-Cantor method, because the JK method does not correct for a higher rate of transitional substitutions as compared to transversional substitutions (Higgs and Attwood 2005). P-distances were used in analysis of amino acid sequences. MP analysis were performed with the branch and bound method. The MP method works best when the amount of variation among sequences is similar. Distance methods like NJ and ME are less stringent about sequence similarity, while

Maximum Likelihood methods do not require a clear recognizable sequence similarity (Mount 2001).

When considering which model to chose one have to remember that the single most important step in tree construction is the sequence alingment. No amount of care, tweaking of parameters or choise of model can infer a correct tree if the alignment is poorly done. The key to determining evolutionary distances between sequences lies in the degree of divergence between sequences. All alignments were manually inspected and adjusted if necessary. Ambiguously aligned regions and gaps were excluded from the analyses. The number of bootstrap replicates was 2000. When using the complete mtDNA genome the highest possible resolution is achieved, but there is a possibility that 3rd codon position will experience substitutional saturation between taxa that diverged a long time ago. These positions will not be informative and only create “noise” in the analysis. To avoid this problem when inferring deep branch phylogenies one can use 1st and 2nd position of codons and stem regions of tRNA genes e.g. Miya et al. (2001)

The tree in figure 21 is based on the complete mtDNA sequence from all available Pleuronectiform species. By using the complete mitochondrial genome a high resolution and accuracy is acchieved.

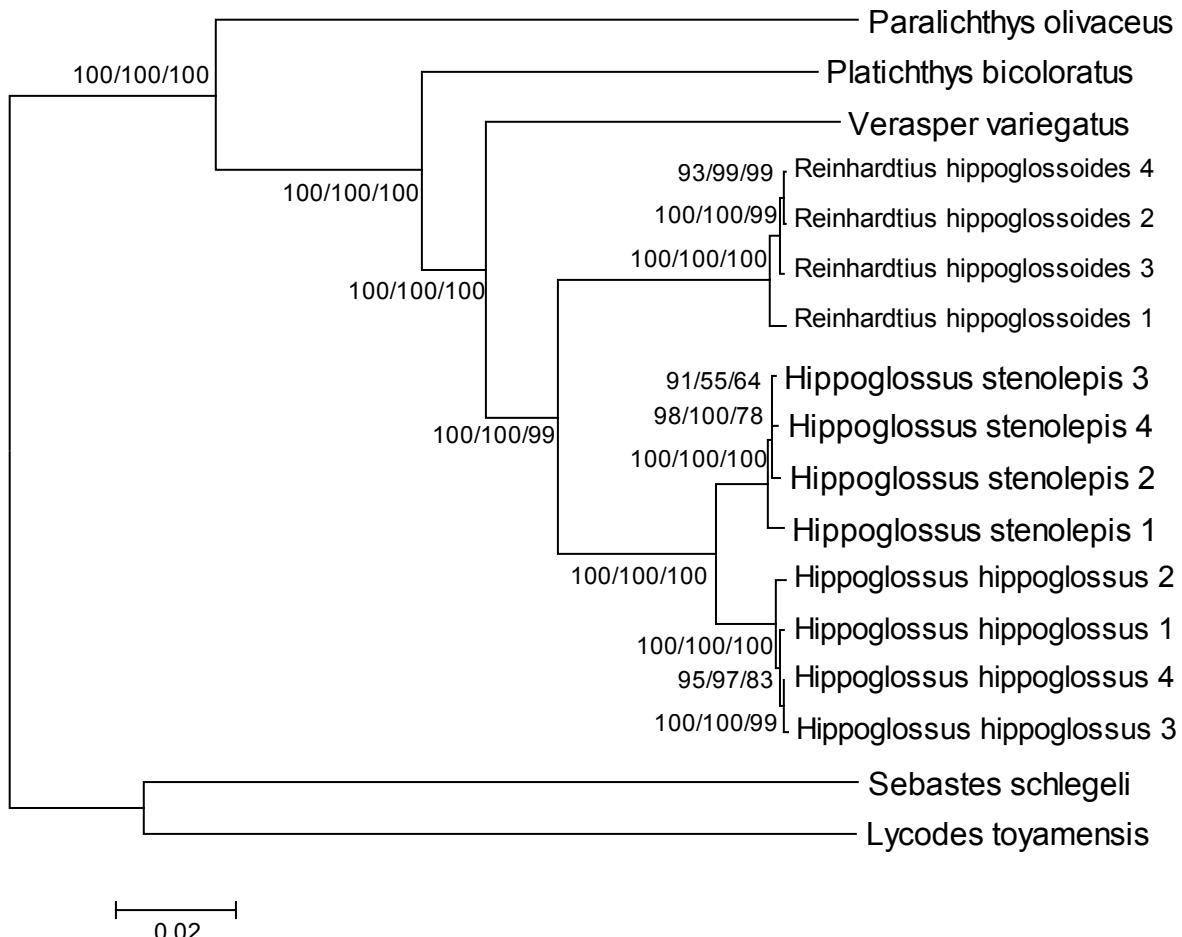


Figure 21: Minimum evolution tree based on complete mtDNA sequences, except control region. The numbers beside the nodes denote bootstrap percentages for 2000 replicates for the tree models Neighbor Joining, Minimum Evolution and Maximum Parsimony. Outgroup: *Sebastes schlegeli* and *Lycodes toyamensis*.

The complete mtDNA sequence was not obtained from Common sole (*Solea vulgaris*), European plaice (*Pleuronectes platessa*) and Turbot (*Scophthalmus maximus*). The available mtDNA sequences from these species were aligned together with the sequences used in figure 21. Regions that lacked sequence in one or more species were excluded. Then one was left with approximately 5000 homologous base pairs between position 30 and 9890, from which the tree presented in figure 22 was made. This includes ND1, COII, ATP8, ATP6, partial sequence from ND2 and ND3 and sequence from multiple tRNAs in this region. These protein coding genes and tRNAs were identified in the same way as in the halibut species, to make sure all sequences were homologous.

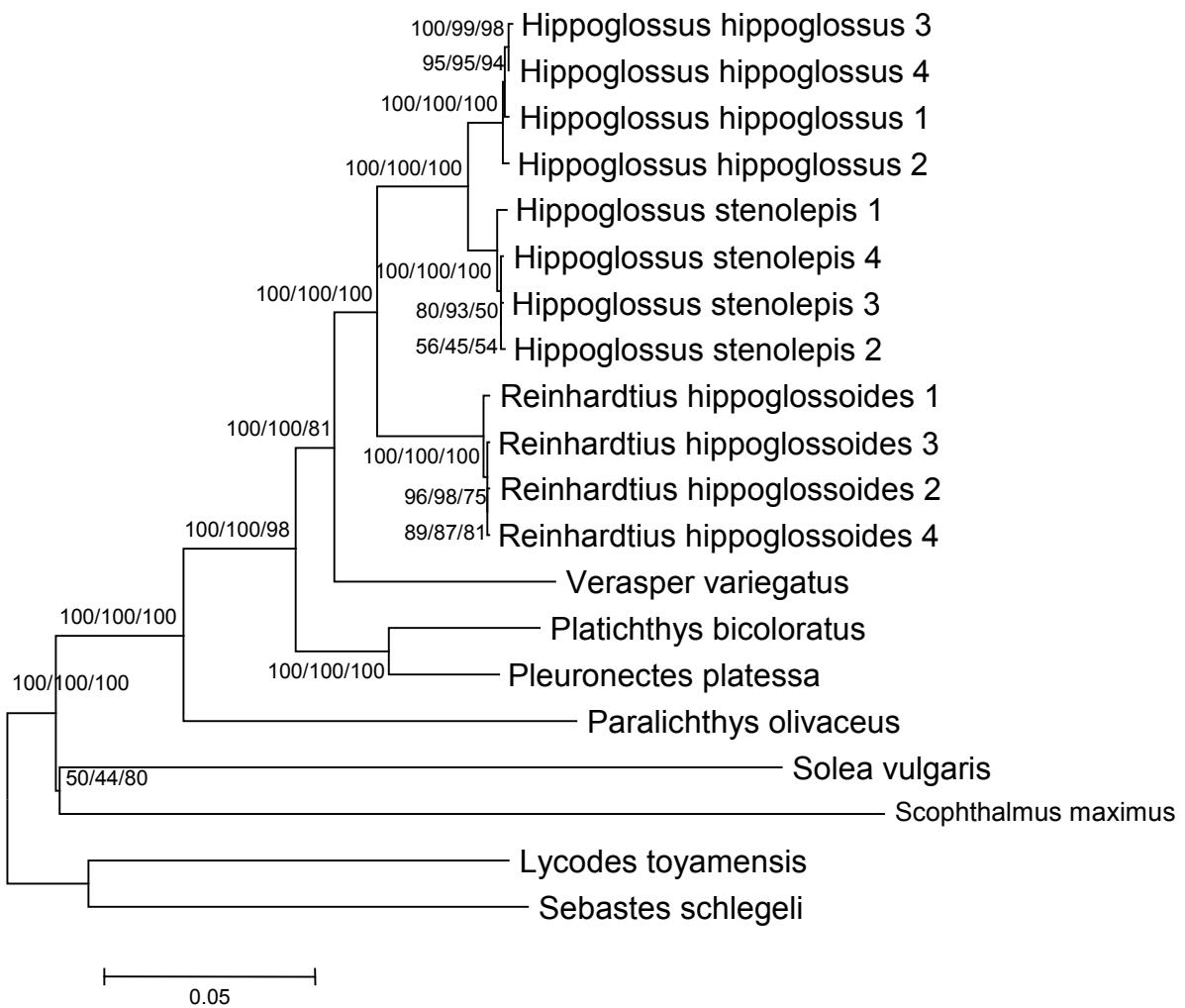


Figure 22: Minimum evolution tree based on alignments of approximately 5000 bp of 9 pleuronectiform species. The numbers beside the nodes denote bootstrap percentages for 2000 replicates for the tree models Neighbor Joining, Minimum Evolution and Maximum Parsimony. A dash instead of a bootstrap value indicates that the method did not support the clade in question. Outgroup: *Sebastes schlegeli* and *Lycodes toyamensis*.

The phylogenetic information in figure 21 and figure 22 can be summarized as follows: Soleidae (*Solea vulgaris*) and Scophthalmidae (*Scophthalmus maximus*) appeared to be sister taxa and also the most basal taxa. Paralichthyidae formed a sister taxa to Pleuronectidae. Within the Pleuronectidae, *Platichthys bicoloratus* and *Pleuronectes plateissa* clustered into one group and was most basal. *Verasper variegatus* appeared to be the closest taxon to the three halibut species. Furthermore, the *Reinhardtius* genus formed a sister taxa to the *Hippoglossus* genus. The *Hippoglossus* genus was the most derived clade. Finally, the interspecific genetic variation between *Hippoglossus hippoglossus* and *Hippoglossus stenolepis* was greater than the intraspecific variation within each species, meaning that they are probably separate species, but speciation took place relatively recently. Bootstrap support was generally high in all trees when considering phylogenetic relationship above species level. The tree based on

5000 bp did not separate all individuals well enough with all methods (Fig 22). Bootstrap values as low as 55 % and 64 % were observed between individuals of *Hippoglossus stenolepis* in the tree produced with the help of complete mtDNA sequences (Fig 21).

To investigate the phylogenetic relationship between Pleuronectiformes and the “well-known” taxa Gadidae and Salmonidae a concatamer of amino acid sequences from the thirteen mtDNA protein-coding genes were used to create the tree in figure 23. Amino acid sequences were chosen because the problems with saturation in third codon position is, to a high degree, excluded, as many amino acids are two- or four-fold degenerated (i.e. substitution can occur in third codon position without changing the amino acid). Phylogenetic investigation of Pleuronectiformes, Gadidae and Salmonidae based on amino acid sequences revealed that Salmonidae is the most basal taxon, while Gadidae is closer related to the Pleuronectiformes (Fig 23).

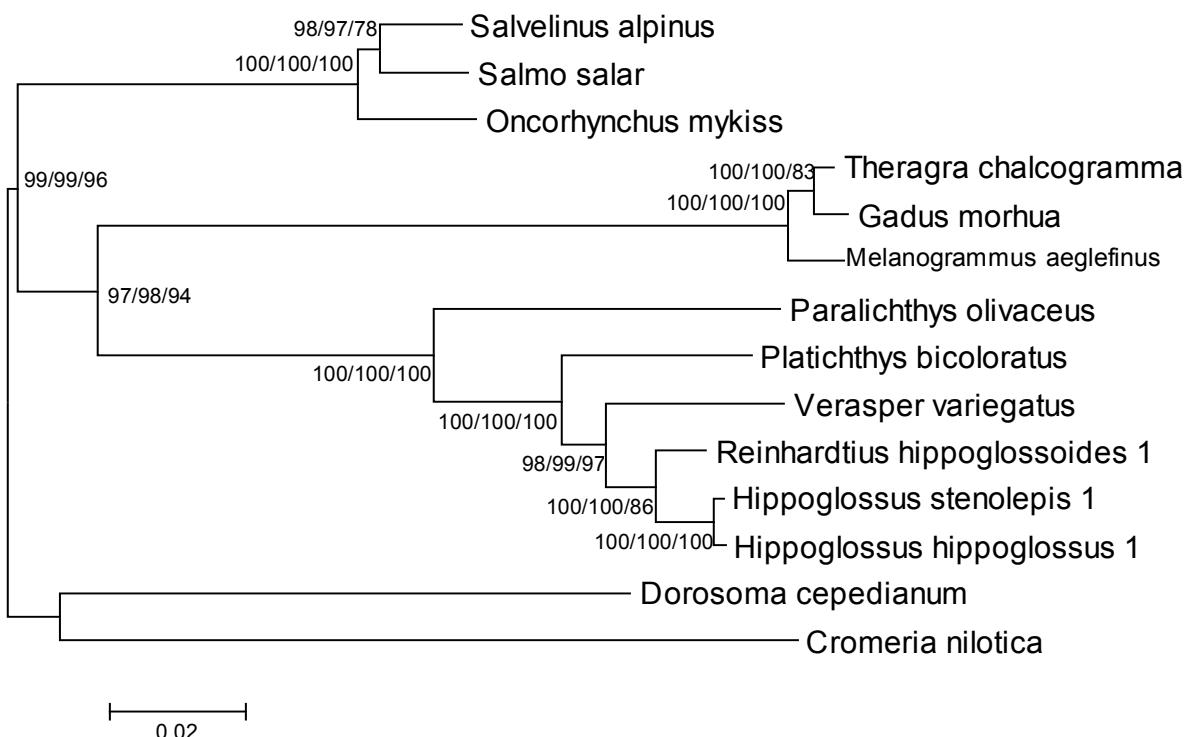


Figure 23: Minimum evolution tree of 6 pleuronectiform species, 3 Gadidae species and 3 Salmonidae species tree based on amino acid sequences from 13 mitochondrial protein coding genes. The numbers beside the nodes denote bootstrap percentages for 2000 replicates for the tree models Neighbour Joining, Minimum Evolution and Maximum Parsimony. Outgroup: *Dorosoma cerpedianum* and *Cromeria nilotica*.

Estimation of divergence times

Sequence divergence in COI between Atlantic halibut 1 and Pacific halibut 1 was 2,3% .

Divergence between Atlantic halibut 1 and Greenland halibut 1 was 7,0%, while divergence between Pacific halibut 1 and Greenland halibut 1 was 7,3%. According to Bermingham et al. (1997) sequence divergence in COI of recently separated teleost taxa is roughly 1,2% per million years. Based on these numbers the divergence time between Atlantic- and Pacific halibut was estimated to 2 million years, while divergence between Greenland halibut and the *Hippoglossus* genus took place approximately 6 million years ago (Ma). Tajimas test of relative rate was used to evaluate rate constancy among lineages (Tajima 1993).

Authenticity of the Mitochondrial Sequences

Nuclear mitochondrial pseudogenes (numt) are parts of the mtDNA molecule that have been copied and incorporated into the nuclear genome. As reviewed by Bensasson et al. (2001) numts have contaminated PCR-based mitochondrial studies of over 64 different animal species. When numt sequences are present, the nuclear and mitochondrial products are often co amplified by PCR, and sequences from these products may exhibit double peaks and ambiguities at some positions (Sorenson and Fleischer 1996). In this study PCR products were mostly rather large, 3kb - 4,5 kb, and all PCR products appeared as clean single bands of the expected molecular weight when controlled using agarose gel electrophoresis. Data files from sequencing reactions that exhibited multiple peaks were discarded. Unspecific primer binding is also a possible explanation for multiple peaks. Thus, it did not appear that co-amplification of both nuclear and mitochondrial copies of genes were a problem.

Discussion

Genetic variation

Complete mitochondrial genome

The number of single nucleotide polymorphisms (SNP) observed among Atlantic halibut individuals was 91. The same numbers for Pacific halibut and Greenland halibut was 99 and 110 respectively. Thus, no significant difference in genetic variation within the three halibut species could be observed. Very few population studies based on complete mtDNA sequences have been performed. In a study of complete mtDNA sequences from 10 individuals of Walley pollock Yanagimoto et al. (2004) observed 179 SNPs. The number of nucleotide substitutions found between Norwegian Coastal stock and Norwegian Shelf stock of cod was 75. Between Norwegian Coastal stock and Arcto-Norwegian stock of cod, 83 nucleotide substitutions were identified, while the number of nucleotide substitutions between Norwegian Shelf stock and Arcto-Norwegian stock was 68 (Breines 2005). These are comparisons between two individuals only, and more variable sites are expected if more individuals are analysed and compared. Brown et al. (2006) detected 156 nucleotide substitutions between five individuals of Rainbow trout. However, this number could be lower, because they questioned the accuracy of the sequencing work in one individual done by Zardoya et al. (1995)

Ribosomal RNA genes

The overall rRNA substitutions rates are about half of those compared to protein coding genes (Brown and Simpson 1982; Hixson and Brown 1986). The same appears true for Atlantic halibut, Pacific halibut and Greenland halibut. In a study of mtDNA from ten individuals of Walleye pollock (*Theragra chalcogramma*) Yanagimoto et al. (2004) found relatively low substitution rates in rRNA genes. Similar findings were discovered in cod (*Gadus morhua*) by Breines (2005) and in a comparison between saithe (*Pollachius virens*) and pollack (*Pollachius pollachius*) by Nymark (2006).

Protein coding genes

The patterns of substitution in protein coding genes are well known. The knowledge of how mutations accumulate makes protein coding genes good candidates for phylogenetic studies in fish.

The genes coding for the cytochrome oxidase subunits and cytochrome b are the most conserved genes (Meyer 1994). This appears to be true also for Atlantic halibut, Pacific halibut and Greenland halibut (Fig 20). These genes were found to be very conserved in saithe and pollack by Nymark (2006), but only relatively well conserved in cod (Breines 2005) and Walley pollock (Yanagimoto et al. 2004). A study that compared mtDNA from Arctic charr and Brook charr found 4.8 %, 3.5 % and 5.9 % sequence divergence COI, COII and COIII respectively (Doiron et al. 2002). The most variable are the ND and ATPase genes (Billington 2003). The ND genes have a nucleotide diversity of 0.2-0.6 % in Atlantic halibut 1-4, Pacific halibut 1-4 and Greenland halibut 1-4, respectively. Though ND3 in Greenland halibut 1-4 had no nucleotide substitutions. ATP8 had a high nucleotide diversity (0.6%) in Atlantic halibut 1-4, and ATP6 had the same nucleotide diversity in Pacific halibut 1-4 and 0.3% in Atlantic halibut 1-4. Otherwise the ATP genes had no or little nucleotide diversity. Brown et al. (2006) found the highest nucleotide variability in ND4 and ND5 in a study of intraspecific variability in five individuals of Rainbow trout. Saithe and pollack had a very low nucleotide variability in the ATP genes (Nymark 2006). While Walley pollock and cod had a moderately substitution rate in these genes (Breines 2005; Yanagimoto et al. 2004). Sequence divergence was high between Arctic charr and Brook charr in all ND genes, except ND4L. Sequence divergence in ATP6 was also high (Doiron et al. 2002). The total number of nucleotide substitutions in protein coding genes of Atlantic halibut, Pacific halibut and Greenland halibut was 59, 71 and 79 respectively. The same number in five individuals of Rainbow trout was 47 (Brown et al. 2006)

Some suggestions as to why some genes are well conserved, while others show a higher variability, are given below. Substitution rates are influenced by the rate of mutation and the probability of fixation of a mutation. Fixation depends on whether the mutation provides an advantage, is neutral or deleterious. Substitution rates in regions that are likely to affect the function of the gene product will be lower (Graur and Li 1999). Apparently, mitochondrial genes have different mutation rates and some genes evolve faster than others. The rate of evolution of a gene or mutation that is under selection will be very different from one that is

not. Selection resists substitutions at some sites, while allowing substitutions at other sites to become fixed. Different genes with different functions, or different parts of a gene with different functions will have different rates of evolution. Thus, different regions of DNA with different functional constraints will evolve at different rates. This is true for both nuclear and mitochondrial genes (Higgs and Attwood 2005). Genes evolve at different rates because, although mutations are usually random events, some proteins are much more tolerant to changes in their amino acid sequence than are other proteins. For this reason, the genes that encode these more tolerant, less constrained proteins evolve faster (Nei 1987). Furthermore, structure, function and expression level affect the rate of evolution. So does the position of a gene in a pathway or network. Genes more central to a network have a greater number of pleiotropic effects on other genes and biological processes. These genes evolve slower, because they are very constrained: a mutation that has a beneficial effect on one process is very likely to have deleterious effects on the other processes, and would be negatively selected (Graur and Li 1999).

The rate of mutations might not always be similar along the mitochondrial genome. Sites in the mtDNA control region of humans known to be hyper variable have been identified as mutational hotspots. Though, it is not known what features of these hyper variable sites and/or the mtDNA replication machinery cause them to be mutational hotspots (Stoneking 2000). Some of the mechanisms discussed here are likely to be the cause of the differentiated evolutionary rate observed in mtDNA genes in Atlantic halibut, Pacific halibut and Greenland halibut.

Variability in protein complexes

Eight protein complexes mediate the processes carried out by the electron transport chain. Each of these complexes plays a vital role in the health of the cell and any slight mutation in any one of the proteins that make up these complexes can lead to cell death or stress, which can both in turn lead to a number of diseases. As an example a mutation in complex I genes been associated with Parkinsons disease and aging (Smigrodzki et al. 2004).

Proteins with more interactors are known to evolve more slowly not because they are more important to the organism, but because a greater proportion of the protein is directly involved in its function (Fraser et al. 2002). The faster evolving ND genes belong to multiprotein complex I, consisting of 25 polypeptides. It seems that the ND genes do not have a central

role in the network of polypeptides in complex I. Otherwise the high occurrence of mutations in the ND genes would make it difficult to cooperate properly with the rest of the polypeptides in the complex. Apparently, there is no strong negative selection acting on mutations in ND genes.

On the other hand, the more conserved CO genes belonging to complex IV could hold a central position among the 13 polypeptides in this complex. It is also possible that there is a regional heterogeneity along the mtDNA molecule, when it comes to occurrence of mutations. Maybe the CO genes, along with ATP8 and ATP6 are located at a part of the molecule that is somehow more protected against mutations. ATP8 was very conserved in Pacific halibut and Greenland halibut, but ATP8 and ATP6 were very variable in Atlantic halibut and Greenland halibut respectively. This hypothesis seems unlikely as Cyt-b, a relatively conserved gene, is located in a different part of the mtDNA molecule, close to the control region, which has an elevated evolutionary rate. Regions that are frequently exposed as a single strand during transcription or replication (O_H and O_L) have also been reported to have a higher rate of mutations (Ramiro et al. 2003).

Control region

The control region is partially constrained in primary sequence and secondary structure in order to regulate replication and transcription of the mtDNA molecule (Clayton 1991). The control region is characterized by the displacement loop (d-loop), a stretch of DNA that is complementary to the light strand; the d-loop displaces the heavy strand. Of all mitochondrial defined regions and genes the control region has the highest substitution rate, at least when considering the two peripheral domains, while the central domain is remarkably conserved (Pesole et al. 1999). Nucleotide diversity was not calculated for the three domains in the control region separately. But one gets an impression of variability in the different domains by viewing the alignment in figure 12. The nucleotide diversity for the control region as a whole, except heteroplasmic tandem repeats, was approximately 1%, about 3 times higher than the average nucleotide diversity in protein coding genes. The intraspecific substitution rate in the control region of cod was found to be slightly higher than that of the rest of the protein coding genes (Breines 2005). Interspecific substitution rate in 10 individuals of Walley pollock was not found to be significantly higher than the protein coding genes (Yanagimoto et al. 2004). The same observation was done regarding the interspecific substitution rate in the control region of saithe and pollack (Nymark 2006).

Functional consequences of amino acid substitutions

Substitutions at the amino acid level will change the properties of the polypeptide. Some changes are small and would probably not have any adaptive value. Others that, for instance, could improve the function of an enzyme at certain temperatures would be of great adaptive value if the organism preferred this temperature (Doiron et al. 2002). A negative score suggests that this substitution serve a vital function in the structure and function of a polypeptide. Doiron et al. (2002) found a high concentration of substitutions with a negative score in the ND genes. They suggested that the transmembrane multiprotein complex I, which ND genes are a part of, could be a target where selection acted on the char's mitochondrial genome for specific environmental factors. For example temperature niches. Five out of nine amino acid substitutions with a negative mutation score between Atlantic halibut, Pacific halibut and Greenland halibut were located in the ND genes. This could be an adaptation to temperature niches. Greenland halibut prefer lower temperatures than both Atlantic halibut and Pacific halibut (see introduction), and the observed mutation was, with one exception, always observed between Atlantic halibut / Pacific halibut and Greenland halibut.

Transmembrane helices

The number of transmembrane helices in the three halibut species is compared to similar data from Saithe (*Pollachius virens*) and Pollack (*Pollachius pollachius*) in a study by Nymark (2006) in table 31. Saithe and Pollack have one more helic in ND1, COI and ND4. Saithe has two helices in ND4L, while all other species have 3.

Table 31: Comparison of transmembrane helices in mtDNA proteins in Atlantic halibut (Hh), Pacific halibut (Hs), Grenland halibut (Rh), Saithe (Pv) and Pollack (Pp)

No of transmembrane helices		
	Hh/Hs/Rh	Pv/Pp
ND1	8	9
ND2	9	9
COI	11	12
COII	2	2
ATP8	1	1
ATP6	6	6
COIII	6	6
ND3	3	3
ND4L	3	2/3
ND4	12	13
ND5	16	16
ND6	5	5
Cyt-b	9	9

Heteroplasmic tandem repeats in the control region

The control region is responsible for tremendously length variation in mtDNA due to tandem duplications (Brown 1985). Such heteroplasmic tandem repeats (HTR) have been reported in several teleost species (Table 32).

Table 32: Selected teleost species with HTR reported in the control region.

Organism	Repeat size (bp)	No. of repeats	Heteroplasmy	Reference
Atlantic halibut (<i>Hippoglossus hippoglossus</i>)	61	12-19	Yes	This study
Pacific halibut (<i>Hippoglossus stenolepis</i>)	61	17-19	Yes	This study
Greenland halibut (<i>Reinhardtius hippoglossoides</i>)	61	18-23	Yes	This study
American Plaice (<i>Hippoglossoides platessoides</i>)	61	2	Not determined	(Lee et al. 1995)
Winter flounder (<i>Pseudopleuronectes americanus</i>)	62	6-8	Not determined	(Lee et al. 1995)
Yellowtail flounder (<i>Limanda ferruginea</i>)	62	6-8	Not determined	(Lee et al. 1995)
European flounder (<i>Platichthys flesus</i>)	19	1-10	Yes	(Hoarau 2002)
Sea bass (<i>Dicentrarchus labrax</i>)	17	1-19	Yes	(Cesaroni et al. 1997)
Cod (<i>Gadus morhua</i>)	48	6-19	Yes	
Sturgeon (<i>Acipenser transmontanus</i>)	40	2-6	Yes	(Árnason and Rand 1992)
	82	1-6	Yes	(Brown et al. 1992)

Atlantic halibut, Pacific halibut and Greenland halibut are all heteroplasmic with respect to a 61 bp motif in the CSB domain of the control region. There were large differences in the sequence of the motif in the three species. It was possible to fold both one copy and the most common copy number of the motif into secondary structures. The G + C content was low in all motifs. This could be a way of resisting the formation of a stable secondary structure. The function of these repeated elements is not known. The most likely mechanism behind the creation of repeated motifs is slipped-strand mispairing (Levinson and Gutman 1987)

A number of conserved sequence motifs were identified in the control region of several pleuronectiform species (Fig 11 and 12). Some sequence blocks were more conserved than others, suggesting that these are vital in mitochondrial metabolism. Different approaches have been tried (Mignotte et al. 1987; Saccone et al. 1987), but the function of the central conserved region is not well understood. Studies of the control region in mammals (Hoelzel et al. 1991) and in *Xenopus* (Wong et al. 1983) have revealed that secondary structure of sequences found in the control region are not conserved among vertebrates. This could suggest that secondary structure of RNA does not play a major role in the function of these sequences.

Codon usage bias

In this study it was observed that Atlantic halibut, Pacific halibut and Greenland halibut in four-fold degenerate codons have a non-preference for codons with guanine in third codon position. For two-fold degenerate codons there was a clear preference for one codon over the other. Here adenine or cytosine was the preferred nucleotide in third codon position. Johansen et al. (1990) also found a clear tendency to exclude guanine in third codon position. This phenomenon is observed in many other species of teleosts, from which codon usage tables are available (<http://www.kazusa.or.jp/codon/Mitochondrion.html>). It seems likely that some of the mechanisms mentioned in the introduction, translational efficiency and / or translational accuracy, are the cause of the observed codon bias in the three halibut species. The pattern seems too clear to be random.

Nucleotide composition

G+C content was compared between pleuronectiform species (*Hippoglossus hippoglossus*, *Hippoglossus stenolepis*, *Reinhardtius hippoglossoides*, *Platichthys bicoloratus* and *Paralichthys olivaceus*), Gadidae species (*Theragra chalcogramma*, *Boreogadus saida* and *Gadus macrocephalus*) and Salmonidae species (*Salmo salar* and *Salmo trutta*) (<http://www.kazusa.or.jp/codon/Mitochondrion.html>). Pleuronectiform species had a G+C content in protein coding genes of approximately 46%, Gadidae 42% and Salmonidae 46%. G+C content in 1st position in pleuronectiform species was around 53%, Gadidae 52% and Salmonidae 55%. 2nd position had a G+C content of 42% in pleuronectiform, 39% in Gadidae and 42% in Salmonidae. Pleuronectiform species had a G+C content of 45% in 3rd codon position, Gadidae 35% and Salmonidae 41%. So there seems to be a selection against a high

G+C content in protein coding genes of Gadidae species, compared to pleuronectiform and Salmonidae species. The specific reason is not known, but it seems to be less beneficial for Gadidae species to have guanine in third codon position. A possible explanation could be related to transcriptional and translational efficiency and accuracy, as mentioned above.

Lin et al. (2006) reported a G+C content in mtDNA from Asian seabass (*Lates calcarifer*) of 46,1% as high compared to many other teleosts. They suggest that a high G+C content is associated with higher temperatures in tropical waters. However, non-tropical fishes like Atlantic salmon (*Salmo salar*) (Hurst et al. 1998), Arctic charr (*Salvelinus alpinus*) (Doiron et al. 2002), Atlantic halibut, Pacific halibut and especially Greenland halibut have G+C contents in the range of 45,1% - 46,2%. A possible explanation could be that these species once originated in warmer waters and then changed their temperature niches faster than the evolution of mtDNA, given that the assumptions by Lin et al. (2006) are correct.

SNP analysis of 30 Atlantic halibut individuals

Part of ND2, COI and the control region were target for an SNP analysis in 30 Atlantic halibut individuals from the broodstock in Mørkvedbukta Research Station (Table 24). The relationship between these individuals was not fully known beforehand. Eleven different genotypes were observed. What is interesting is that one genotype (f) appear 15 times. A possible explanation is that they are siblings or half-siblings, sharing the same mother. Additional information about the breeding of their possible parents suggests that this is true. By also using nuclear microsatellites it would be possible to finally determine the genetic relationships. Normally one would seek to have a wide variety of genetic material in a broodstock. The use of closely related individuals in a broodstock will eventually lead to inbreeding, i.e. increased homozygosity, which leads to reduced larval and juvenile survival and performance (Beaumont and Hoare 2003). For this reason it is important to know the history of each individual in a breeding programme, and preferably genotype every individual to avoid inbreeding.

Phylogenetic investigation of mtDNA sequences

The focus on the three halibut species through this study will continue in the phylogenetic investigation. Their position among other known complete Pleuronectiform mtDNA genomes will be considered. The Norwegian waters normally host three different families of

Pleuronectiform species: Soleidae, Scophthalmidae and Pleuronectidae (Bethon 2005). Partial mtDNA sequence from *Solea vulgaris* (Soleidae) and *Scophthalmus maximus* (Scophthalmidae) was also determined, and the relationship between these three families will be discussed.

Results from this study suggest that Soleidae and Scophthalmidae are the most basal families. (Berendzen and Dimmick 2002) were not able to fully determine the position of Scophthalmidae, though it appears to be more basal than Soleidae. Chapleau (1993) resolved Scophthalmidae as a polytomy containing Bothidae, Paralictyidae and Pleuronectidae. So the position of these two families remain uncertain.

Paralichthyidae has been recognized as polyphyletic (Chapleau 1993; Hensley and Ahlstrom 1984). One clade was determined to be closely related to Pleuronectidae (Berendzen and Dimmick 2002). Paralichthyidae represented by *Paralictys olivaceus* was recognized as the family closest to Pleuronectidae in this study.

Within the Pleuronectidae *Platichtys bicoloratus* and *Pleuronectes platessa* formed a group. This group is recognized as Pleuronectini (Haaramo 2004), a sister group to Hippoglossini. Hippoglossini includes *Verasper variegatus*, *Reinhardtius hippoglossoides*, *Hippoglossus stenolepis* and *Hippoglossus hippoglossus*. The last two species appear to be very close related. They could even be separated at the sub-species level. There are 358 SNPs between Hh1 and Hs1. 1100 SNPs was found between HH1 and Rh1 (control region was excluded due to gaps). Nymark (2006) detected 877 nucleotide substitutions between Saithe (*Pollachius virens*) and Pollack (*Pollachius pollachius*). The number of nucleotide substitutions between Arctic char (*Salvelinus alpinus*) and Brook trout (*Salvelinus fontalis*) was approximately 760 (Doiron et al. 2002).

Species divergence times

Atlantic- and Pacific halibut was estimated to have separated 2 Ma, while divergence between Greenland halibut and the *Hippoglossus* genus took place approximately 6 million years ago.

One should be very careful with drawing conclusions based on estimated divergence times, but some cautious suggestions as to what caused this speciation can be made. Important geological events in the area where these species live or their ancestors might have lived

include the generally warmer periods Miocene 23-5.3 Ma (Billups 2002) and Pliocene 5.3-1.8 Ma (Crowley 1991), the submergence of the Bering Strait 3.5 Ma (Vermeij 1991) and the rise of the Isthmus of Panama 3.1-3.5 Ma (Coates and Obando 1996). These events might have influenced speciation that led to the three halibut species.

Miocene was warm until 14 Ma when temperatures dropped. This happened again 8 Ma (Billups 2002). The first speciation was the separation of Greenland halibut from the ancestor of the *Hippoglossus* genus. It is difficult to tell why this happened. So possible explanations suggested here would only be speculations. One have to bear in mind that other species not included in this study could be involved in this process.

One possible mechanism behind the separation of the *Hippoglossus* genus is that the ancestor of this genus might have been able to move freely between the Atlantic- and the Pacific Ocean via the Arctic basin during the warm period of Pliocene. Speciation might have occurred when temperatures in the Arctic waters dropped too low during late Pliocene for the ancestor to inhabit these high latitudes.

Species identification based on molecular markers

Illegal fishing could potentially be a threat to many species, specially those with a slow growth rate and long generation time, such as Atlantic halibut, Pacific halibut and Greenland halibut. Therefore, accurate species identification is a necessary part of each fishery management program as well as any conservation approach (Ludwig 2005). Morphological species identification is very difficult and unreliable when the fish has been processed into filet and other products. PCR based restriction fragment length polymorphism (RFLP) has many advantages. After establishing the species-specific restriction patterns, this approach was used to differentiate 17 Acipenseriform species by Ludwig et al. (2002). Wolf et al. (1999) were able to distinguish between several species of game fish and even closely related deer species with PCR-RFLP. The same approach was used by Graves (2001) to separate several popular fish species. He digested amplification products from ND4 and 12S/16S rRNA region in mtDNA with restriction enzymes and identified species-specific patterns. This approach should prove useful in separating between the three halibut species in this study. The regions of choice should be the control region or one of the ND genes, due to the high degree of nucleotide variability found here.

An alignment of the complete mtDNA sequence from the four individuals of Atlantic halibut, Pacific halibut and Greenland halibut revealed 27 SNPs where each species had a different nucleotide (i.e. species specific SNP). Some regions had a higher concentration of these SNPs than others. Six SNPs were found between position 3090 and 3696. By sequencing this stretch, or a part of it, one can easily determine which species the mtDNA sample belongs to.

Real time PCR (RT-PCR) is a innovative technique in molecular genetics that makes it possible to quantitative measure polymorphic DNA regions and identify SNPs at the same time. After PCR amplification, SNPs can be detected by analysing a melting-curve: The temperature is raised from somewhere between 40°C and 75°C to a the temperature were a hybridisation probe melts off the DNA. No SNP means a perfect match between DNA and probe and is indicated by a higher melting temperature. If a mutation is present the melting temperature will be lower (Reuter et al. 2005). Then one would have to identify an SNP were species 1 have a specific nucleotide (e.g. C) and species 2 and 3 have a different nucleotide from species 1 (but both the same, e.g. G). Now you have excluded or confirmed species 1. If species 1 was excluded, the process is repeated, but now with a SNP were species 1 and 2 have the same nucleotide and species 3 have a different one. Then species 2 or 3 is also ruled out, and the correct species has been identified. Two sensor probes would have to be designed according to SNP 1 and 2 to separate between three species. More possible species requires more SNPs and probes. Multiple sets of SNPs can be added depending on the degree of certainty that is required in species identification. To make sure that the SNP really is diagnostic and no intraspecific variation is found at this position, a large number of samples from each species might be needed when establishing the test.

References

- (www.fishbase.org).
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." *J. Mol. Biol.*(215): 403-410.
- Anderson, S., A. T. Bankier, B. G. Barrell, M. H. L. de Bruijin, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreier, A. J. H. Smith, R. Standen and I. G. Young (1981). "Sequence and organization of the human mitochondrial genome." *Nature*(290): 457-465.
- Anderson, S., M. H. L. de Bruijin, A. R. Coulson, I. C. Eperon, F. Sanger and I. G. Young (1982). "Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome." *J Mol Biol*(156): 683-717.
- Andriyashev, A. P. (1954). "Fishes of the northern seas of U.S.S.R. (transl. from Russian)." (1964).
- Árnason, E. and D. M. Rand (1992). "Heteroplasmy of short tandem repeats in mitochondria of Atlantic cod, *Gadus morhua*." *Genetics*(132): 211-220.
- Árnason, U. and A. Gullberg (1993). "Comparison between the complete mtDNA sequences of the blue and the fin whale, two species that can hybridize in nature." *J. Mol. Evol.*(37): 312-322.
- Bakke, I. and S. D. Johansen (2005). "Molecular Phylogenetics of Gadidae and Related Gadiformes Based on Mitochondrial DNA Sequences." *Marine Biotechnology* 7(61-69).
- Baxevanis, A. D. and B. F. Francis Ouellette (2001). *Bioinformatics; a practical guide to the analysis of genes and proteins*, Wiley-interscience.
- Beaumont, A. R. and K. Hoare (2003). *Biotechnology and genetics in fisheries and aquaculture*, Blackwell publishing.
- Bensasson, D., D.-X. Zhangb, D. L. Hartla and G. M. Hewittc (2001). "Mitochondrial pseudogenes: evolution's misplaced witnesses." *Trends in Ecology & Evolution* 16(6): 314-321.
- Benson, G. (1999). "Tandem repeats finder: a program to analyze DNA sequences." *Nucleic Acids Research* 27(2): 573-580.
- Berendzen, P. B. and W. W. Dimmick (2002). "Phylogenetic relationships of Pleuronectiformes based on molecular evidence." *Copeia*(3): 642-652.
- Bergsten, J. (2005). "A review of long-branch attraction." *Cladistics*(21): 163-193.
- Bermingham, E., S. S. McCafferty and A. P. Martin (1997). Fish biogeography and molecular clocks: Perspectives from the Panamanian isthmus. *Molecular systematics of fishes*. T. D. Kocher and C. A. Stepien, Academic press.
- Bernardi, G. and G. Bernardi (1986). "Compositional constraints and genome evolution." *J Mol Evol*(24): 1-11.
- Bethon, P. (2005). *Aschehougs store fiskebok*.
- Billington, N. (2003). *Mitochondrial DNA in Population genetics: Principles and applications for Fisheries Scientists*, ed. Hallerman, Eric M. American Fisheries Society, Bethesda, Maryland.
- Billups, K. (2002). "Late Miocene through early Pliocene deep water circulation and climate change viewed from the sub-Antarctic South Atlantic." *Palaogeography palaeoclimatology palaeoecology* 185(3-4): 287-307.
- Boore, J. L. (1999). "Animal mitochondrial genomes." *Nucleic Acids Res* 27(8): 1767-1780.
- Breines, R. (2005). Genetic relationship of Gadidae species: A study of complete mtDNA from Atlantic cod, Polar cod and Arctic cod. *Department of Molecular Biotechnology, Institute of Medical Biology, University of Tromsø. Master of science*.

- Brown, G. G. and M. V. Simpson (1982). "Novel features of animals mtDNA evolution as shown by sequences of two rat cytochrome oxidase subunit II genes." Proc Natl Acad Sci USA(79): 3246-3250.
- Brown, J. R., A. T. Beckenbach and M. J. Smith (1992). "Mitochondrial DNA variation and heteroplasmy in populations of the white sturgeon (*Acipenser transmontanus*)."
Genetics(132): 221-228.
- Brown, J. R., K. Beckenback, A. T. Beckenback and M. J. Smith (1996). "Length variation, heteroplasmy and sequence divergence in the mitochondrial DNA of four species of sturgeon (*Acipenser*)."
Genetics(142): 525-535.
- Brown, K. H., R. E. Drew, L. A. Weber and G. H. Thorgaard (2006). "Intraspecific variation in the rainbow trout mitochondrial DNA genome."
Comparative Biochemistry and Physiology Part D: Genomics and Proteomics 1(2): 219-226.
- Brown, W. M. (1985). The mitochondrial genome of animals, in Molecular Evolutionary Genetics. (ed. R.J. Macintyre) pp. 95-130, Plenum Press.
- Brown, W. M., M. J. George and A. C. Wilson (1979). "Rapid evolution of mitochondrial DNA."
Proc Natl Acad Sci USA(76): 1967-1971.
- Burzynski, A., M. Zbawicka, D. O. F. Skibinski and R. Wenne (2003). "Evidence for recombination of mtDNA in the marine mussel *Mytilus trossulus* from the Baltic."
Mol. Biol. Evol.(20): 388-392.
- Cesaroni, D., F. Venanzetti, G. Allegrucci and V. Sbordoni (1997). "Mitochondrial length variation and heteroplasmy in natural populations of the European sea bass, *Dicentrarchus labrax*."
Mol Biol Evol(14): 560-568.
- Chapleau, F. (1993). "Pleuronectiform relationships; a cladistic reassessment."
Bull. Mar. Sci.(52): 516-540.
- Clayton, D. A. (1991). "Replication and transcription of vertebrate mitochondrial DNA."
Ann. Rev. Cell. Biol.(7): 453-478.
- Coates, A. G. and J. A. Obando (1996). The geologic evolution of the Central American Isthmus. Evolution and Environment in Tropical America. J. B. C. Jackson, A. F. Budd and A. G. Coates, University of Chicago Press: 21-56.
- Crowley, T. J. (1991). "Modeling Pliocene warmth."
Quaternary Sci. Rev(10): 275-282.
- De Groot, S. J. (1970). "Some notes on an ambivalent behaviour of the Greenland halibut *Reinhardtius hippoglossoides* (Walb.) Pisces: Pleuronectiformes."
J. Fish Bid.(2): 275-279.
- Delarbre, C., H. Escriva, C. Gallut, V. r. Barriel, P. Kourilsky, P. Janvier, V. Laudet and G. Gachelin (2000). "The Complete Nucleotide Sequence of the Mitochondrial DNA of the Agnathan *Lampetra fluviatilis*: Bearings on the Phylogeny of Cyclostomes."
Mol. Biol. Evol. 17(4): 519-529.
- Doiron, S., L. Bernatchez and P. U. Blier (2002). "A Comparative Mitogenomic Analysis of the Potential Adaptive Value of Arctic Charr mtDNA Introgression in Brook Charr Populations (*Salvelinus fontinalis* Mitchell)."
Mol. Biol. Evol. 19(11): 1902-1909.
- DOMINGUES, V. S., G. BUCCIARELLI, V. C. ALMADA and G. BERNARDI (2005). "Historical colonization and demography of the Mediterranean damselfish, *Chromis chromis*."
Molecular Ecology(14): 4051-5063.
- Efron, B. (1979). "Bootstrapping methods: Another look at the jackknife."
Ann. Stat.(7): 1-26.
- Entelis, N. S., O. A. Kolesnikova, S. Dogan, R. P. Martin and I. A. Tarassov (2001). "5 S rRNA and tRNA Import into Human Mitochondria."
J. Biol. Chem 276(49): 45642-45653.
- Fairbarn, D. J. (1981). "Biochemical genetic analysis of population differentiation in Greenland halibut (*Reinhardtius hippoglossoides*) from the northwest Atlantic, Gulf of St. Lawrence and Bering Sea."
Can. J. Fish. aquat. Sci.(38): 669-677.
- Fedorov, K. Y. (1971). "Zoogeographic characteristics of the Greenland halibut, *Reinhardtius hippoglossoides* (Walbaum))."
J. Ichthyol. 11(6): 971-976.

- FOSS, A., A. K. IMSLAND and G. NÆVDAL (1998). "Population genetic studies of the Atlantic halibut in the North Atlantic Ocean." *Journal of Fish Biology*(53): 901–905.
- Fraser, H., A. E. Hirsh, L. M. Steinmetz, C. Scharfe and M. W. Feldman (2002). "Evolutionary Rate in the Protein Interaction Network." *Science* **296**(5568): 750 - 752.
- Gadaleta, G., D. Pepe, G. De Candia, C. Quagliariello, E. Sbisà and C. Saccone (1989). "The complete nucleotide sequence of the *Rattus norvegicus* mitochondrial genome: cryptical signals revealed by comparative analysis between vertebrates." *J. Mol. Evol.*(28): 497-516.
- Gissi, C., A. Gullberg and A. Ulfur (1998). "The complete mitochondrial DNA sequence of the rabbit, *Oryctolagus cuniculus*." *Genomics*(50): 161-169.
- Godø, R. O. and T. Haug (1989). "A review of the natural history, fisheries and management of Greenland halibut (*Reinhardtius hippoglossoides*) In the eastern Norwegian and Barents Seas." *J. Cons. int. Explor. Mer.*(46): 62-75.
- Gouy, M. and C. Gautier (1982). "Codon usage in bacteria, correlation with gene expressivity." *Nucleic Acids Res*(10): 7055-7047.
- Grant, W. S., D. J. Teel, T. Kobayashi and C. Schmitt (1984). "Biochemical population genetics of Pacific halibut (*Hippoglossus stenolepis*) and comparison with Atlantic halibut (*H. hippoglossus*)."*Canadian Journal of Fisheries and Aquatic Sciences*(41): 1083–1088.
- Grantham, R. (1974). "Amino acid difference formula to help explain protein evolution." *Science*(185): 862-864.
- Grantham, R., C. Gautier, M. Guoy, R. Mercier and A. Pave (1980). "Codon catalogue usage and the genome hypothesis." *Nucleic Acids Res*(8): 49-62.
- Graur, D. and W.-H. Li (1999). *Fundamentals of Molecular Evolution*, Sinauer.
- Graves, J. (2001). *Fisheries Genetics Program*. Unpublished. The Virginia institute of Marine Science School of Marine Science College of William and Mary Gloucester Point, Virginia 23062.
- Grzybowski, T. (2000). "Extremely high levels human mitochondrial DNA heteroplasmy in single root hairs." *Electrophoresis*(21): 548-553.
- Hall, B. G. (2004). *Phylogenetic trees made easy. Second edition.*, Publishers Sunderland, Massachusetts U.S.A.
- Haug, T. and S. E. Fevolden (1986). "Morphology and biochemical genetics of Atlantic halibut, *Hippoglossus hippoglossus* (L.) from various spawning grounds." *Journal of Fish Biology*(28): 367–378.
- Hedges, S. B. (2002). "The origin and evolution of model organisms." *Nat. Rev. Genet*(3): 838-849.
- Hensley, D. A. and E. H. Ahlstrom (1984). Pleuronectiform relationships p. 670-687 in. *Ontogeny and systematics of fishes*, Am Soc Ichtyol Herpetol Spec.
- Higgs, P. G. and T. K. Attwood (2005). *Bioinformatics and molecular evolution*, Blackwell science.
- Hillis, D. M. and J. J. Bull (1993). "An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis." *Syst. Biol.*(42): 182-192.
- Hixson, J. E. and G. G. Brown (1986). "A comparison of the small ribosomal RNA genes from the mitochondrial DNA of the great apes and humans, sequence, evolution and phylogenetic implications." *Mol Biol Evol*(3): 1-18.
- Hoarau, G. (2002). "Heteroplasmy and evidence for recombination in the mitochondrial control region of the flatfish *Platichthys flesus*." *Mol. Biol. Evol*(19): 2261–2264.
- Hoelzel, A. R., J. M. Hancock and G. A. Dover (1991). "Evolution of the cetacean mitochondrial D-loop region." *Mol Biol Evol* **8**(3): 475-493.
- Hurst, C. D., S. E. Bartlett, I. J. Bruce and W. S. Davidson (1998). "The complete nucleotide sequence of the mitochondrial DNA of the Atlantic salmon, *Salmo salar*."

Unpublished, Norwood Laboratory, School of Chemical and Life Sciences, University of Greenwich, Wellington Street, Woolwich, London, SE18 6PF.

- Hutchison, C. A., J. E. Newbold, S. S. Potter and M. H. Edgell (1974). "Maternal inheritance of mammalian mitochondrial DNA." *Nature*(251): 536–538.
- Haaramo, M. (2004). "www.fmnh.helsinki.fi/users/Haaramo."
- Inoue, J. G., M. Miya, J. Aoyama, S. Ishikawa, K. Tsukamoto and M. Nishida (2001). "Complete mitochondrial sequence of the Japanese eel *Anguilla japonica*." *Fisheries science* **67**: 118-125.
- Jarne, P. and P. J. L. Lagoda (1996). "Microsatellites, from molecules to populations and back." *Trends in Ecology and Evolution*(11): 424-429.
- Joensen, J. S. (1954). "On the life story of halibut in Faroe waters." *Meddelelser fra Danmarks Fiskeri- og Havundersøgelser Ny serie 1*: 1-25.
- Johansen, S., P. H. Guddal and T. Johansen (1990). "Organization of the mitochondrial genome of Atlantic cod, *Gadus Morhua*." *Nucleic Acids Res.* **18**(3): 411-419.
- Jones, D. T., W. R. Taylor and J. M. Thornton (1994). "A mutation data matrix for transmembrane proteins." *FEBS Lett*(339): 269-275.
- Kephart, S. R. (1990). "Starch gel electrophoresis of plant isozymes: a comparative analysis of techniques." *American Journal of Botany*(77): 693-712.
- Kim, I. C. and J. S. Lee (2004). "The complete mitochondrial genome of the rockfish *Sebastodes schegeli* (Scorpaeniformes, Scorpaenidae)." *Mol. Cells* **17**(2): 322-328.
- Klump, H. and D. Maeder (1991). "The thermodynamic basis of the genetic code." *Pure & Appl. Chem*(63): 1357- 1366.
- Kotlar, D. and Y. Lavner (2006). "The action of selection on codon bias in the human genome is related to frequency, complexity, and chronology of amino acids." *BMC Genomics*(7): 67.
- Krane, D. E. and M. L. Raymer (2003). *Fundamental concepts of bioinformatics*, Benjamin Cummings.
- Kumar, J., K. Tamura and M. Nei (2004). "MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment." *Briefings in Bioinformatics* **5**(2): 150-163.
- Lee, W.-J. and T. D. Kocher (1995). " Complete sequence of a sea lamprey (*Petromyzon marinus*) mitochondrial genome: early establishment the vertebrate genome organization." *Genetics*(139): 873-887.
- Lee, W. J., J. Conroy, W. H. Howell and T. D. Kocher (1995). "Structure and evolution of teleost mitochondrial control regions." *J Mol Evol* **41**: 54-66.
- Levinson, G. and G. A. Gutman (1987). "Slipped-Strand Mispairing: A Major Mechanism for DNA Sequence Evolution." *Mol. Biol. Evol.* **4**(3): 203-221.
- Li, X.-L., Z.-L. Wu, Z.-Z. Liu, Y.-F. Gong, R.-Y. Zhou and G.-R. Zheng (2006). "SNP Identification and Analysis in Part of Intron 2 of Goat MSTN Gene and Variation within and among Species." *JOURNAL OF HEREDITY* **97**(3): 285-289.
- Lin, G., L. C. Lo, Z. Y. Zhu, F. Feng, R. Chou and G. H. Yue (2006). "The Complete Mitochondrial Genome Sequence and Characterization of Single-Nucleotide Polymorphisms in the Control Region of the Asian Seabass (*Lates calcarifer*)." *Mar. Biotechnol* **8**(71–79).
- Lowe, T. M. and S. R. Eddy (1997). "tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence." *Nucleic Acids Res*(25): 955-964.
- Ludwig, A. (2005). *Molecular species identification - state of the art and future prospect*, Unpublished. Institute for Zoo and Wildlife Research, Department of Evolutionary Genetics, PO Box 601103, 10252 Berlin, Germany.
- Ludwig, A., L. Debus and L. Jenneckens (2002). "A molecular approach for trading control of black caviar." *Int Rev Hydrobiol*(87): 661-674.

- Lunt, D. H. and B. C. Hyman (1997). "Animal mitochondrial DNA recombination." *Nature*(387): 247.
- Mathews, C. K., K. E. van Holde and K. G. Ahern (2000). *Biochemistry*, Benjamin Cummings.
- May-Panloup, P., M. F. Chrétien, C. Jacques, C. Vasseur, Y. Malthièry and P. Reynier (2005). "Low oocyte mitochondrial DNA content in ovarian insufficiency." *Human Reproduction* **20**(3): 593-597.
- May, B. (1992). *Starch gel electrophoresis of allozymes*. Hoelzel AR (ed.), Molecular Genetic Analysis of populations: a practical approach, Oxford University Press, Oxford.
- McCracken, F. D. (1958). "On the biology and fishery of the Canadian Atlantic halibut, *Hippoglossus hippoglossus* L." *Journal of the Fisheries Research Board of Canada* **15**: 1269-1311.
- Meyer, A. (1994). *DNA technology and phylogeny in fish*, in *Genetics and evolution of aquatic organisms*, ed. Beaumont, A.R. Chapman & Hall.
- Michaels, G. S., W. W. Hauswirth and P. J. Laipis (1982). "Mitochondrial DNA copy number in bovine oocytes and somatic cells." *Dev. Biol*(94): 246-251.
- Mignotte, B., D. Dunon-Bluteau, C. Reiss and J.-C. Mounolou (1987). "Sequence deduced physical properties in the D-loop region common in five vertebrate mitochondrial DNAs." *J Theor Biol*(124): 57-69.
- Miya, M., A. Kawaguchi and M. Nishida (2001). "Mitogenomic Exploration of Higher Teleostean Phylogenies: A Case Study for Moderate-Scale Evolutionary Genomics with 38 Newly Determined Complete Mitochondrial DNA Sequences." *Mol. Biol. Evol.* **18**(11): 1993-2009.
- Miya, R. and M. Nishida (1999). "Organization of the mitochondrial genome of a deep-sea fish, *Gonostoma gracile* (Teleostei: Stomiiformes): first example of transfer RNA gene rearrangements in bony fishes." *Mar. Biotechnol*(1): 416-426.
- Mokranjac, D. and W. Neupert (2005). "Protein import into mitochondria." *Biochemical Society Transactions* **33**(5): 1019-1023.
- Mork, J. and T. Haug (1983). "Genetic variation in halibut, *Hippoglossus hippoglossus* (L.), from Norwegian waters." *Hereditas*(98): 167-174.
- Mount, D. W. (2001). *Bioinformatics: sequence and genome analysis*, CSHL press.
- Nei, M. (1987). *Molecular Evolutionary Genetics*, Columbia Univ. Press, New York.
- Nei, M. and W. H. Li (1979). "Mathematical model for studying genetic variation in terms of restriction endonucleases." *Proc Natl Acad Sci USA* **76**: 5269-5273.
- Nelson, J. S. (1994). *Fishes of the world*. New York, John Wiley & Sons Inc.
- Nesbø, C. L., M. O. Arab and K. S. Jakobsen (1998). "Heteroplasmy, Length and Sequence Variation in the mtDNA Control Regions of Three Percid Fish Species (*Perca fluviatilis*, *Acerina cernua*, *Stizostedion lucioperca*)."*Genetics* **148**: 1907-1919.
- Norman, J. R. (1934). *A systematic monograph of the flatfishes (Heterostomata)*. Vol. 1. Psettodidae, Bothidae, Pleuronectidae. British museum (Natural history) London.
- Nymark, M. (2006). Gadidae mitogenomics: Molecular characterization of the complete mitochondrial genomes of Saithe (*Pollachius virens*) and Pollack (*Pollachius pollachius*). *Department of Molecular Biotechnology, Institute of Medical Biology*, University of Tromsø. **Master of science**.
- Parsons, T. J., D.S. Muniec, K. Sullivan, R. Allison-Greiner, M.R. Wilson, D.L. Berry, K.A. Holland, V.W. Weedn, P. Gill and M. M. Holland (1997). "A high observed substitution rate in the human mitochondrial DNA control region." *Nature Genet*(15): 363-368.
- Pereira, S. L. (2000). "Mitochondrial genome organization and vertebrate phylogenetics." *Genetics and molecular biology* **23**(4): 745-752.
- Pesole, G., C. Gissi, A. De Chirico and C. Saccone (1999). "Nucleotide substitution rate of mammalian mitochondrial genomes." *J. Mol. Evol.* **48**: 427-434.

- Pollard, T. D. and W. C. Earnshaw (2004). Cell biology, Saunders.
- Powell, J. R. and E. N. Moriyama (1997). "Evolution of codon usage bias in *Drosophila*." Proc Natl Acad Sci USA(94): 7784-7790.
- Primrose, S. B. and R. M. Twyman (2003). Principles of Genome Analysis and Genomics, Blackwell publishing.
- Queller, D. C., J. E. Strassmann and C. R. Hughes (1993). "Microsatellites and kinship." Trends in Ecology and Evolution(8): 285-288.
- Ramiro, A. R., P. Stavropoulos, M. Jankovic and M. Nussenzweig (2003). "Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand." Nature Immunology(4): 452 - 456.
- Reid, D. P., S. PONGSOMBOON, T. JACKSON, C. MCGOWAN, C. MURPHY, D. MARTIN-ROBICHAUD and M. REITH (2005). "Microsatellite analysis indicates an absence of population structure among *Hippoglossus hippoglossus* in the north-west Atlantic." Journal of Fish Biology(67): 570-576.
- Reuter, M., Y. Küpper, A. Schmitz, J. P. Breuer, U. Wend and J. Hennig (2005). "Detection of new single nucleotide polymorphisms by means of real time PCR." Journal of Genetics 84(3).
- Rodger, R. W. A. (2005). The Fisheries of North America: An Illustrated Guide to Commercial Species, Canadian Marine Publications.
- Roe, B. A., D.-P. Ma, R. K. Wilson and F.-H. Wong (1985). "The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome." J. Biol. Chem(260): 9759-9774.
- Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer and R. Rozas (2003). "DnaSP, DNA polymorphism analyses by the coalescent and other methods." Bioinformatics(19): 2496-2497.
- Rzhetsky, A. and M. Nei (1992). "A simple method for estimating and testing minimum evolution trees." Mol. Biol. Evol.(9): 945-967.
- Saccone, C., M. Attimonelli and E. Sbisa (1987). "Structural elements highly preserved during the evolution of the D-loop containing region in vertebrate mitochondrial DNA." J Mol Evol(33): 205-211.
- Saitoh, K., K. Hayashizaki, Y. Yokoyama, T. Asahida, H. Toyohara and Y. Yamashita (2000). "Complete nucleotide sequence of Japanese flounder (*Paralichthys olivaceus*) mitochondrial genome: Structural properties and cue for resolving teleostean relationships." JOURNAL OF HEREDITY 91(4): 271-278.
- Saitou, N. and M. Nei (1987). "The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees." Mol. Biol. Evol. 4(4): 406-425.
- Sharp, P. M., M. Stenico, J. F. Peden and A. T. Lloyd (1993). "Codon usage, mutational bias, translational selection, or both?" Biochem Soc Trans(21): 835-841.
- Smigrodzki, R., J. Parks and W. D. Parker (2004). "High frequency of mitochondrial complex I mutations in Parkinson's disease and aging." Neurobiol. Aging(25): 1273-1281.
- Sorenson, M. D. and R. C. Fleischer (1996). "Multiple independent transpositions of mitochondrial DNA control region sequences to the nucleus." Proc. Natl. Acad. Sci. USA(93): 15239-15243.
- Stenico, M., A. T. Lloyd and P. M. Sharp (1994). "Codon usage in *Caenorhabditis elegans*, delineation of translational selection and mutational biases." Nucleic Acids Res(22): 2437-2446.
- Stoneking, M. (2000). "Hypervariable Sites in the mtDNA Control Region Are Mutational Hotspots." Am. J. Hum. Genet.(67): 1029-1032.
- Swofford, D. L., G. J. Olsen, P. J. Waddel and D. M. Hillis (1996). Phylogenetic inference. In Molecular Systematics, Sunderland MA: Sinauer Associates.

- Tajima, F. (1993). "Simple methods for testing molecular clock hypothesis." *Genetics*(135): 599-607.
- Templeman, W. (1973). "Distribution and abundance of the Greenalnd halibut, *Reinhardtius hippoglossoides* (walbaum) in the northwest Atlantic." *ICNAF Res. Bull.* **1973**(10): 83-98.
- Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin and D. G. Higgins (1997). "The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools." *Nucleic Acids Res* **24**: 4876-4882.
- Tzeng, C.-S., C.-F. Hui, S.-C. Shen and P. C. Huang (1992). "The complete nucleotide sequence of the *Crossostoma lacustre* mitochondrial genome: conservation and variations among vertebrates." *Nucleic Acids Res*(20): 4853-4858.
- Ursing, B. M. and U. Arnason (1998). "The complete mitochondrial DNS sequense of the pig, *Sus scrofa*." *J. Mol. Evol.*(47): 302-306.
- Van Bruggen, E. F., J. Borst, P. Ruttenberg, G.J.C.M, M. Grubber and A. M. Kroon (1966). *Biochim. Biophys. Acta*(119): 437-439.
- Vermeij, G. J. (1991). "Anatomy of an invasion: the trans-Arctic interchange." *Palaebiology* **17**(3): 281-307.
- Vignal, A., D. Milan, M. Sanristobal and A. Eggen (2002). "A review on SNP and other types of molecular markers and their use in animal genetics." *Genet. Sel. Evol.*(34): 275-305.
- Warrior, R. and J. Gall (1985). *Arch. Sci. Geneva*(38): 439-445.
- Wolf, C., J. Rentsch and P. Hubner (1999). "PCR-RFLP analysis of mitochondrial DNA: a reliable method for species identification." *J Agric Food Chem.* **47**(4): 1350-1355.
- Wong, J. F., E. P. Ma, R. K. Wilson and B. A. Roe (1983). "DNA sequence of the *xenopus laevis* mitochondrial heavy and light strand replication origin and flanking tRNA genes." *Nucleic Acids Res* **11**(14): 4977-4995.
- Xia, X. (1996). "Maximizing transcription efficiency causes codon usage bias." *Genetics*(144): 1309-1320.
- Yanagimoto, T., T. Kitamura and T. Kobayashi (2004). "Complete nucleotide sequence and variation of mitochondrial DNA from 10 individuals of walleye pollock, *Theragra chalcogramma*." *Fisheries science*(70): 885-895.
- Zama, M. (1990). "Codon usage and secondary structure of mRNA." *Nucleic Acids Symp Ser*(22): 93-94.
- Zardoya, R., A. Garrido-Pertierra and J. M. Bautista (1995). "The complete nucleotide sequence of the mitochondrial DNA genome of the rainbow trout, *Oncorhynchus mykiss*." *J. Mol. Evol.*(41): 942-951.
- Zuckerkandl, E. and L. Pauling (1962). Molecular disease, evolution, and genetic heterogeneity. *Horizons in Biochemistry*. M. Marsha and B. Pullman, Academic Press: 189-225.
- Zuker, M., D. H. Mathews, D. H. Turner, J. Barciszewski and B. F. C. Clark (1999). Algorithms and Thermodynamics for RNA Secondary Structure Prediction: A Practical Guide in RNA Biochemistry and Biotechnology. N. A. Series, Kluwer Academic Publishers.

Appendix

A. Complete L-strand nucleotide sequence of the mitochondrial genome from Atlantic halibut 1-4, Pacific halibut 1-4 and Greenland halibut 1-4. The sequences do not yet have an accession number, but they will be submitted to Gene bank.

B. Concatamer of all available mtDNA sequence from European plaice, Turbot and Common sole. L-strand.

C. Part of the results from this thesis was used in a poster presentation:
Kenneth A. Mjelle, Tor E. Jørgensen, Bård O. Karlsen, Steinar D. Johansen and Truls Moum
2006. Genetic markers in Atlantic halibut: Tools for breeding success.

The poster was presented at the 2006 Winter Meeting of the Norwegian Biochemical Society

Hippoglossus hippoglossus_1 17546 bp

```
1 GCTAACGTAG CTTAATTAAA GCATAACACT GAAGATGTTA AGATGGGCC TAGAAAGCCC
61 CGGGAGCAC AAGGCTTGT CCTGACTTAA CTGTCGACTT TAACTAAACT TACACATGCA
121 AGTATCCGCC CCCCTGTGAG AATGCCACA ACTCCCTGCT TGGGAACTAG GAGCCGGTAT
181 CAGGCACAAG CCCAGCTAGC CCACGACGCC TTGCTTAGCC ACACCCTCAA GGGAACTCAG
241 CAGTGATAAA TATTAAGCCA TAAGTGAAAAA CTTGACTTAG TTAAGGTTAA GAGGGCCGGT
301 AAAACTCGTG CCAGGCCACCG CGGTTATACG AGAGGCCAA GTTGACAAAC AACGGCGTAA
361 AGAGTGGTTA GGGGATTTC TAAACTAGAG CGGAACGCTT TCAAAGCTGT TATACGCACC
421 CGAAAGTATG AAACCCAATT ACGAAAGTAG CTCTACTCAT CCTGAACCCA CGAAAGCTAA
481 GAAAACAAACT GGGATTAGAT ACCCCACTAT GCTTAGCCCT AAACATCGAT TGCACTATAC
541 ACTCCATATC CGCCCGGGAA TTATGAACAT CAGTTAAAAA CCCAAAGGAC TTGGCGGTG
601 TTAACATCCA CCTAGAGGAG CCTGTTCTAG AACCGATAAC CCCGTTAAA CCTCACCTTC
661 TCTTGTTTA TCCGCTATA TACCAACGTC GTCAGCTTAC CTCGTGAAGG CTTTACAGTA
721 AGCAAAATTG GCAAAGCCC AAACGTCAAG TCGAGGTGTA GTGAATGAGG GGGGAAGAAA
781 TGGGCTACAT TTGCTAAACA TAGCAACAC GAATGTTGCA TTGAAACATG TAACTGAAGG
841 AGGATTAGC AGTAAGCAGG AAATAGAGCG TCCCGCTGAA ACTGGCCCTA AAGCGCGCAC
901 ACACGGCCCG TCACCCCTCC CAAGCCCCCT AAACTAAACT ATTAAAACC CTACAACCCG
961 CGAAGGGGAG GAAACTCGTA ACATGTAAG TGACCGGAA GTGCGCTTIC GAAAATCAG
1021 AGTGTAGCTA AGATAGTGT CAGCATCTCA CTTACACCGA AAAGACGTCC GTGCAAGTCG
1081 GACCACCTGTG ACGCTTATTA GCTAGGCCAA CCCCTTAACA CAACAGACCC CCATTATATAA
1141 CCCCTAAAGC ACGAACACCC CACGTAGCTA AACCAATTCTC CCCCCTAAGT CCAGGCATA
1201 AAAAAGGAAA TTTTGGAGCA ATAGAAAAAG TACCGCAAGG GAAAGCTGAA AGAGAGATGA
1261 AAAAGCCAG TAAAGCTAA GAAAGCAGAG CTTAGTGCTC GTACCTTTG CATCATGATT
1321 TAGCTAGCAC TTTCAAGCAA AGAGAACCTA CAGTTTGAA CCCCAGAAACT GAGTGAGCTA
1381 CTCCAAGACA GCCTATTAT AGGGCGAACCC CGTCTCTGTG GCAAAGAGGT GGGAAAGAGCT
1441 TTGAGTAGAG GTGACAAACT TACCGAACTC AGTTATAGCT GGTTGCCTGT GAATTGAATA
1501 GAAGTCAGC CCCCTGGGT CTCCACTCAT GCACCTTGT ACCCTTCAG ATGCAATGAG
1561 AAACCGAGGG TGTTAGCTAA AGGGGGTACA GCCCTTTGTA TACAAGACAC AACTTTCTCA
1621 GCAGGATAAA GATCATATTC AAATAAGGAC AAATGTTTA GTGGGCCTAA AAGCAGCCAC
1681 CTAATAGAA AGCGTTAAAG CTCAAACATA AAGCCCTCCC GGTATACCGA TAACCCTATC
1741 TTATTCCCCC ACATTTAACCA GGCCCTCTA TGCAACACAT AGGAACGACT ATGCTAATAT
1801 GAGTAATAAG AGGGTATAAT CACCCCTCTCC TTGACCATGT GTAAATCGGA ACGGACCCCC
1861 CACCGAATCT TAACGGCCCC AATCAAAGAG GGTATTGGAA ACTACCCACAA ATTAGGCCA
1921 GAAAACATC CAATGCAAAC CCGTTAACCC CACACTGGTG TGCCCCAAAG GAAAGACCAA
1981 AGGGGGGAGA AGGAACCTGG CAAACATACC CCAAGCCTCG CCTGTTTACG AAAACATCG
2041 CCTCTTGAT AACCACAGTA TAAGAGGTCC CGCCTGCCA GTGACCATAT AGTTCAACCG
2101 CGCGGGTATT TTGACCGTGC AAAGGTAGCG TAATCACTTG TCTTTAAAT GAAGACCTGT
2161 ATGAATGGCA TAACGAGGGC TTAACTGTCT CCTCCCCCG GTCAATGAAA TTGATCTCCC
2221 CGTGCAGAAG CGGGGATTAA ACCATAAGAc GAGAAGGCCA TATGGAGCTT TAGACACACA
2281 GGTGGGCCAT GTCAAATACC CCTAGCTAAG GGCCTGAACCT AATGGAAATC TGCTTGATG
2341 TCTTCGGTTG GGGGACCTAAG GGGGAAATACA AAACCCCCCA GTGGAAAGGG AGCACACCCC
2401 TAAGTTACTT CTTCTCCCG AAGCCAGAGC AACAGCTCTA ACCAGCAGAA ATTCTGACCA
2461 AACTGATCCG GTAAAACCGA TCAACCGAACCC AAGTTACCCCT AGGGATAACA GCGCAATCCC
2521 CTTTTAGAGC CCATATCGAC AAGGGGTTT ACGACCTCGA TGTTGGATCA GGACATCCTA
2581 ATGGTGCAGC CGCTATTAAAG GTTCCGGTTG TTCAACGATT AAAGTCCTAC GTGATCTGAG
2641 TTCAGACCGG AGTAATCCAG GTCAGTTCT ATCTATGACA TGATCTTTTC TAGTACGAAA
2701 GGACCGAAAAA GAGGAGGCC CTGCTAAAG TACGCCCTCAC CCCCACCTAA TGAAAAAAATC
2761 TAAACTAGGC AAAAGGGCAT AACCCCTTTG CTGGAGATAAA CAGCAAGTTG GGGTGGCAGA
2821 GCCCGGCTAA TGCAAAAGC CTAAGCCCTT TCTACAGAGG TTCAAGTCCT CTCCTTAACT
2881 ATGATTCAA CTCTCATTAAC TCATATTATC AACCCACTAA CCTTTATTGT GCCTGTATTG
2941 CTAGCCGTAG CATTCTCAC CCTCCTGAA CGAAAAGTGC TAGGCTACAT ACAACTACGA
3001 AAAGGGCCAA ACATTGTTGG GCCTTACGGC CTCCCTCAAC CTATTGCTGA TGGCGTAAAA
3061 CTCTTTATTA AAGAACCCAT TCGACCTCA ACCGCATCAC CGCTCCTATT CCTCGTGGCC
```

3121 CCTATGCTCG CTCTCACACT AGCCCTAAC CTTTGAGGCC CCATACCTTT CCCGTACCC
 3181 GTCGTGGACC TTAACCTAGG TATTATTT ATTCTAGCTC TATCTAGCCT TGCAGTATAT
 3241 TCTATTCTGG GCTCCGGCTG AGCATCTAAT TCAAAATATG CCCTAGTAGG GGCACCG
 3301 GCTGTCGAC AAACGATTTC TTACGAAGTT AGCCTTGGAC TCATCTTACT AACATTATC
 3361 ATCTTCACCG GGGGCTTCAC ACTCCAAACC TTTAACACAG CCCAAGAGGC CATTGACTG
 3421 GTAGTGCCTG CTTGACCCTGG CTGACCATA TGATATATCT CCACCTTGC CGAAACAAAC
 3481 CGTGACCCCT TCGACTTAAC AGAAGGAGAA TCTGAACITG TCTCGGGCTT CAACGTAGAG
 3541 TACCGAGGGG GACCCCTTCG CTTGTTCTTC TTAGCCGAAT ATTCAAATAT TCTCCTAATA
 3601 AATACCCCTG CCGCTACACT ATTCTAGGG GCCTCCCACA TCCCATCTAT CCCGGAATTA
 3661 ACTAGTATTA ATATTATGAC TAAAGCAGC CTTCTATCAA TTGTCCTCCT GTGGGTCCGG
 3721 GCTTGTACCG CGCATTCCG TTACGACCAA CTTATGCACC TCATTTGAAA AAACTTCTT
 3781 CCACTGACTT TGGCACTAGT TATCTGACAC TTGGCGCTCC CTATTGCATT CGCTGGCTC
 3841 CCACCACAGC TATAGCGCA GGAGCTGTGC CTGAATTAA AGGGCCACTT TGATAGAGTG
 3901 AATTATGGGG GTTAGAGTCC CCCCAACTCC TTAGAAAGAA GGGGATCGAA CCCAACCTGA
 3961 AGAGATCAAAC ACTCTTCGTG CTTCTCTAC ACCACTCTCT AGTAAGGTCA GCTAAATAAG
 4021 CTCTGGGGC CATAACCCAA CCATGTAGGT TAAATCCTT CTTTGCTAA TGAACCCCTA
 4081 CATCTTGACC ACTCTCTAT TTGTTTGGG CTTGGAACA ACACTCACGT TTGCAAGCTC
 4141 ACACTGGCTC CTCGCTTGA TAGGACTTGA GATTAATACA CTAGCCATT TTCCCCCTAAT
 4201 AGCCCAACAC CACCACCCCC GGGCAGTCGA AGCTACTACT AAATACTTTC TCGCACAAGC
 4261 CACAGCAGCC GCCACCCCTC TGTTGCAAG CACTACCAAC GCTGACTCA CAGGCCAGTG
 4321 GGATATTCAA CAGATAACAC ATCCCCCTCC CACAACAATG ATTGTTATTG CCCTCGCGCT
 4381 AAAGGTCGGA CTGGCACCAA TGCACCTTG ACTCCCGAA GTACTCCAAG GGCTAGACCT
 4441 GACCACCGGG CTCATCCTC CAACCTGACA AAAACTGCC CCCTTGCCTC TCTTACTACA
 4501 GATTAGACACA ACCAACCTC CGCCTCTAAT CATTATTGGC TTACTATCCA CTCTCGTTGG
 4561 TGGCTGAGGT GGTCTAACCC AAACCCAGCT GCGCAAGGTG CTCGCCTACT CCTCAATTGC
 4621 CCACCTTGGC TGAATAATGC TGATCCTCA ATTTTCACCC CTCTTGCC TTCTCACCT
 4681 CCTCACGTAC TTTACTATGA CTTTCTCAGC ATTCCTAATT TTTAAAGTAA ACAAAAGCGAC
 4741 CACCGTCAAT GCCCTCGCA TCTCGTGAAC AAAAACTCTT GCCCTCACAG CCCTAGCACC
 4801 CCTGTTTTA CTTTCTCG GCGGCCCTCC TCCCTTACCG GGCTTCATGC CCAAGTGATT
 4861 CATTCTTCAG GAACGTACTA AACAAGACCT TCCGGCACTT GCCACCCCTCC CCGCTTTAAC
 4921 TGCCCTTGG AGTCTCTACT TCTACTTGCG CCTCTCGTAT GCAATAACCC TCACAATATT
 4981 CCCCaACAAC CTAATTGGTG TCACTCCCTG ACGATTCTAC TCCCCCCTAAT TTACACTTCC
 5041 CCTCGCCGTC TCCACTGCAAG CAACCAACCT TCTCTCCCA CTAGCCCCCG CTGCGTAGC
 5101 ACTACTGGTC ACCTAAGGGG CTTAGGCTAG CACTTAGACC AATGGCCTTC AAAGCCGTAA
 5161 GCGTGAGTGA AAATCTCTA GTCCCTGTTA AGACTTGCAG GCTACTATCC CACATCTCT
 5221 GCATGCAAAG CAAACACTT AATTAAGCTA AAGCCTTCT AGATAGGAAG GCCTTGATCC
 5281 TACAAAATCT TAGTTAACAG CTAAGTGCAGG AATCCAGCGA GCATCTATCT ACTTTTCCCC
 5341 GCCTAATTAA GAAACTAAAA GCGGGGGAAA AGCCCCGGCA GGAGCTTAGC CTGCTCTTT
 5401 AGATTGCAA TCTAACATGT AACACCCAG GGCTTGGTAA GAAGAGGGCT TGCACCTCTG
 5461 TCTATGGGTC TACAATCCAC CGCTTAACCTC AGCCATCCCTA CTCGTGGCAA TCACACGTTG
 5521 ATTTTCTCG ACCAATCACA AAGACATCGG CACCCCTCTAT CTCGTATTG GTGCGTAGC
 5581 CGGAATAGTG GGGACAGGCC TAAGTGTGCT TATTGGGCA GAACTAAGCC AACCCGGGGC
 5641 TCTCTGGG GACGACCAA TTTATAATGT GATCGTCACC GCACACGCC TTGTAATAAT
 5701 CTTTTTATA GTAATACCCA TTATGATTGG GGGGTTCGGA ACTGGCTTA TTCCACTAAT
 5761 AATTGGGGCC CCAGACATGG CGTCCCTCG AATGAAATAAT ATGAGTTTCT GACTTCTTCC
 5821 CCCCTCTTT CTCCCTCTT TAGCCTCTC AGGTGTTGAA GCGGGAGCG GTACCGGATG
 5881 AACCGTGTAC CCCCCACTAG CTGGCAATTAGC AGCCCACGCC GGGGCATCC TAGACGTGAC
 5941 AATCTTCTCA CTTCACCTG CAGGAATTTC ATCAATTCTG GGGCAATTAA ACTTTATTAC
 6001 TACCATCATT AACATGAAAC CCACACAGT CACTATGTAC CAAATCCCGT TATTGTTTG
 6061 AGCCGTTCTT ATTACAGCCG TACTCTCTCT TCTGTCCCTG CCGGTTTTAG CGCAGGGAT
 6121 TACAATGCTA CTAACAGACC GCAACCTTAA CACGACCTTC TTGACCCCTG CGGGAGGAGG
 6181 TGACCCATT CTCTACCAAC ACCTATTCTG ATTCTTGGC CACCCAGGAG TATACATCT
 6241 TATCTCTCA GGCTTGGAA TAATTCTCA CATTGGCAGA TACTATGCG GAAAGAAAAGA
 6301 ACCTTTGGC TACATGGGA TAGTCTGAGC TATAATGGCC ATTGGACTCC TGGGCTTCAT
 6361 TGTCTGGGC CATCACATAT TTACAGTCGG AATAGACGTA GATACACGAG CCTACTTTAC
 6421 CTCTGCCACA ATAATCATG CGATTCCAAC TGGCGTAAAA GTCTTAGCT GACTCGAAC
 6481 CCTCCATGGG GGAAGCATTA AATGAGAAAC GCCCCCTCTA TGAGCCCTCG GCTTTATTT
 6541 CCTCTTACA GTAGGGGTC TCACTGGCAT TGCTTAGCT AACTCCCTC TCGATATTGT
 6601 TCTCTGATGAC ACATACTATG TAGTCGCCA CTTCCCATAT GACTATCTA TGGGTGCTGT
 6661 ATTTGCAATC GTTGGCGCTC TCGTCATTG ATTCCCGTTA TTACAGGCT ATACCCCTCA
 6721 CTCCACATGA ACAAAAATCC ACTTCGGCTT GATGTTTATT GGGGTCAATC TAACATTCTT
 6781 CCCTCAACAT TTTCTGGGC TGGCTGGAT ACCCCGACGG TACTCAGACT ACCCAGACGC
 6841 ATACACCCCT TGAAACACTG TTTCATCAAT TGGGCCCTA ATGCCCCCTG TTGCTGTAAT
 6901 TTTATTCTTA TTCATTATTT GAGAAGCATT TACAGCCAA CGAGAAGTCG GAGCAGTAGA
 6961 ACTAACTGCA ACTAACATGT AATGACTTTA CGGCTGCCCT CCCCCCTACC ACACATTGAA
 7021 AGAGCCCGA TTGCTACAAAG TTGCTATAAA TTGCAACAA CTAACGAGAA AGGGAGGAGT
 7081 TGAACCCCCA TCAATTGGT TCAAGGACAC CACATAACCC TCTGTCACT TTCTTCACAA
 7141 CACACCAATA AGATACTAGT AAAACGCTAT AACACTGCC TGTCAAGGCA GAATTGTTGG
 7201 TCAACCCCCC GCGTATCTA AACGCAATGG CACATCCCTC ACAACTCGGA TTTCAAGACG
 7261 CAGCTTCAACC CTTAATAGAA GAACTACTTC ACTTCCACGA CCACGCCCTA ATAATTGAA
 7321 TTCTCATTAG CACAATGGTA CTTTATATAA TTGCGGCTAT AGTCACAGCC AAACAAACAG
 7381 ATAAActTGT ACTAGACTCT CAAGAAATTG AAATTATCTG AACAGTTCTC CCAGCTGTTA
 7441 TTCTTATCCT TATCGCCCTA CCATCCATCC GAATTGTTGA CTTAATAGAC GAAATTAAATG
 7501 ACCCCACCT TACAATCAA GCCATCGGG ACCAATGATA CTGAAGCTAC GAATACACAG
 7561 ACTACGAAGA CCTCGGTTT GACTCATACA TAATCCCCAC ACAAGACCTG ACCCCCGGAC
 7621 AATTCCGACT ACTAGAAGCA GACCAACGGAA TAGTAATCCC AGTACAATCT CCAATTCCG
 7681 TTTTAATCTC GGCTGAAGAT GTCTTACATT CCTGGCAGT CCCCTCTCTA GCGTAAAG
 7741 TTGACGCCGT ACCTGGACGA TTGAAACCAAG CAACCTTTAT TGTTAGCCGA CCAGGCCGTAT
 7801 TTTTCGGACA ATGTTCTGAA ATTTGCGGGG CTAACCATAG TTTTATACCC ATTGTTGTAG

7861 AAGCAGTCCC CCTTGAAACAC TTTGAGAACT GGTCTTCACT ATAATTGAA GACGCCTCGC
 7921 TAAGAAGCTA ATAAGTACAA GCGTTAGCCT TTTAAGCTAA AGACTGGTGC CTAACAACCA
 7981 CCCCTAGCGA CATGCCCAA CTGAACCCCG CACCCCTGATT TGCAATTAA GTTTCTCTT
 8041 GAATAATCTT TTAACTGTT ATTCCCCCA AAGTTTTAGC ACACACCTAT CCCAATGAAC
 8101 CCACCTCCA AAGCACACAA AAACCTAAAA CAGAGCCCTG AAACGTGACCA TGATACTAAG
 8161 CTTCTTGAC CAATTATAT CCCCTTCATA CCTCGGCATT CCACTAATTG CACTAGCGAT
 8221 TAGCTGCCT TGAATTCTGT TCCCAACCCC TCAAGCCGC TGATTAAACA ACCGCGTGCT
 8281 AACACTTCAG GGGTATTAA TTAGCGCCT CACCTCACA CCTCTTCTAC CCCTAAACCC
 8341 TGGGGGACAG AAGTGGGCAG TTCTATTCA CTCGTTAAC GTTTCCCTAT TTCTATCAA
 8401 CATGCTAGGA CTCCATCCAC ACACCTCAC ACCAACAAACC CAACTCTCCC TAAACATGGG
 8461 CCTCCAGTA CCCCTTGAT TGGCAACTGT CATTATCGGG ATACGAAATC AACCAACGCA
 8521 TGCCCTAGGC CACCTTC CAGAAGGCAC CCCCCACTGCC CTCATCCCTG TACTTATTAT
 8581 CATTGAAACA ATTAGCCTAT TCATTGACCC CTTAGCTCTC GGTGTCCGAC TAACAGCAAA
 8641 TCTCATAGCA GGCCACCTCC TTATTCAACT CATTGCAACA GTCGCCCTCG TCCTTCTCC
 8701 CCTCATACAT GTTGTGCTA TCTTAACAC GGTGGCTCTC TTCCCTCCTCA CTCTACTAGA
 8761 AGTAGCCGT TCCATGATTC AAGCTTGT CTTGTCTA CTCTCTAACCC TATACTACA
 8821 AGAGAACGTT TAATGGCCCA TCAAGCACAC CCATACCCACA TAGTCGACCC CAGCCCATGA
 8881 CCCCTAACAG GGGCTATTG TGCCCTATTG ATAACATCTG GCCTTGCTAT CTGATTCCAC
 8941 TTCCACTCCA CAACCTAAT AACTATTGGG ACAATCCTTC TCACCTTAAC AATCTCCAA
 9001 TGATGACGGG ACGTGTCAG AGAAGGCACA TTCAAGGAC ACCACACTCC CCCCCTCCTA
 9061 AAAGGCCTCC GATATGGAT AATTTTATT ATTACCTAG AAGTCCTATT CTTCTTAGGC
 9121 TTCTCTGAG CCTTTTACCA CTCAGCCTA GCACCTACCC CTGAGCTAGG CGGCTTCTGA
 9181 CCACCGCAG GCATTACTCC CTTAGACCCCA TTGAAGTCA CACTCTTAA CACAGCAGTC
 9241 CTGCTCGCTC CTGGCGTAA TGTACCTGA GCACATCACA GCATTATGGA GGGTAAACGA
 9301 AAACAAACTA TTCAATCTC TGCTCTACA ATCTTACTCG GGGGGTACTT CACTTCTCTT
 9361 CAAGGCCTAG AGTATCACGA AGCCCCCTTC ACCATTGCG AGCGAGTTTA CGGCGCTACA
 9421 TTCTTGTG TGACCGGCTT CCACGGGTT CCACGGGTTA CACGTCTAA TTGGCACATC ATTCTTAGCC
 9481 GTTGTCTAT TACGCCAGAT TCTTCACCAT TTACACATCAA ACCACCACCT TGGGTTGAA
 9541 GCAGCCGCAT GATACTGACA CTTCTGTAGAC GTCTGTCTGAC TTTCTCTCA TATCTCTATT
 9601 TACTGATGAG GATCTTAATC TTCTCTAGT TAAATCTAGT ATAATGTAATC TCCAATCACC
 9661 CGGTCTTGGT TAAAATCCAA GGGAAAGATA TGAGCCTCT CTAACCATC ATTTCGATTG
 9721 CCACCCCTCT CTCGACAGTA CTAGCCCTTG TATCCTTTG ATTACCCCAA ATTACACCAG
 9781 ACCACGAAAA GCTATCCCG TAGCAGTGG GCTTGTACCC GATAGGATCC GCCCGGCTGC
 9841 CTTTTCACT GCGGTTTTG CTCATGCCA CCCTCTTCT TCTCTTCGAT TTAGAAATTG
 9901 CTCTCTCCT CCCTCTTCA TGAGGAGACC AACTAGCATC GCCACTACTG ACGTTCACCT
 9961 GAGCTACAGC TGCTCTAGCC CTTCTAACCC TCAGCCTCAT TTACGAATGA ATGCAAGGAG
 10021 GTCTAGAATG GGCTGAATAG GTGGTAGTC TAAGAAAAAC ATTTGATTTG GGCTAAAAAA
 10081 CTTGTGGTTT AAATCCGAA CCGCTTAATG ACCCCCCACAC ACTTTGCTT CTCTCAGGCC
 10141 TTTCTCTGG GTTTAACAGG CCTGGCATTC CACCGGTTTC ACCTCTTTC CGCCCTATTG
 10201 TGCTCTGAAG GAATAATACT ATCCCTATT ATTGCTCTC CCCTATGAAC GCTCCAACCT
 10261 GACTCAACCA ACTTTTCAGC ATCTCCAAATA CTTCTGCTTG CATTTCAGC CTGCGAAGCA
 10321 AGGCCGGCC TCGCCCTATT AGTAGGCCACT GCTCGAACCC ACGGGACCGA CCGATTACAA
 10381 AGCTAAATC TTCTCAATG CTAAAAGTCC TAATCCCAAC ACTTATGCTA ATCCCAACAG
 10441 CCTGGTTACT CAAACCCAGC TGACTCTGG CCATGACTTT ATGTATAGC TTTTGCATCT
 10501 CCTGGTTAG CTTCTCATGA CTAAGAACCT CTCAGAAAC CGGCTGATCG TCACTTAGCC
 10561 TCTCTTATAGC TACCGACTC ATCTAACCC CCCTCCTCGT TCTTACATGC TGATTACTCC
 10621 CACTAATAAT TTTGGCAAGC CAAAAACACA CAGCCTTAGA ACCTCTTAGT CGCCAGCGCA
 10681 TGTACATCAC ACTTCTCGCC TCACTCCAGT TTTCTTAAT CCTAGCATTT AGGCCACCG
 10741 AACTGGTGT GTTTACGTA ATATTGAAAG CCACCCCTCAT CCCCCACGCTA ATCATTATCA
 10801 CCCGCTGAGG AAACCAAACA GAACGCTAA ACGCAGGAAC CTACTTTCTC TTCTATACAT
 10861 TAGCGGGCTC ACTCCCTCTT CTCGTTGCT TACTCTACT CAAAAACACA TCCGGCACCC
 10921 TGTGCTGTG AACCCCTCAC TATACCGACC CGCTCGCTCT CTCGCTTAT GCAGACAAAC
 10981 TATGATGAGC AGGCTGTCTT TTAGCATTC TGTTAAAAAT ACCCTCTAC GGGGTCCACC
 11041 TATGGCTCCC TAAAGCCCAT GTTGAAGGCC CAATTGCGAGG CTCATAATC CTTGCAGCGG
 11101 TTCTACTGAA GCTAGGGGAA TACGGCATAA TCCGCATAAT ACAATATTG GAACCTCTAA
 11161 CTAAGGAGTT AAGCTACCC TTCATTGCT TTGCACTTTG AGGGGTAATT ATAATCTGGG
 11221 CAATTGTCT ACGCCAGACA GACCTTAAGT CCCTAATTGC CTACTCCTCA GTCAAGCCACA
 11281 TAGGCTCTAGT GGCTGGAGGA ATTCTATCC AATCACCTTG GGGCTTAACA GGGGCTCTTA
 11341 CCCTCATATA TGCACACGGG CTTACTCTAT CAGCCCTCTTG CTGCTTAGCA AACACAAAC
 11401 ATGAAACGAAAC CCACAGTCGA ACGATAGTC TTGCGCGAGG ACTCTAAGTA GCCCTACCC
 11461 TAATAGCCAC TTGATGATT ATTCTCTAGT TGCCCAACCT AGCTCTGCCA CCACCTACCA
 11521 ACCTCATGGG GGAGCTGATA ATTATCACTT CTTATTCAA CTGGCTCTGA TGAACCTCTGG
 11581 CATTAAACGGG TGCAAGCACT CTAATCACAG CTGGTTACTC CTTTACATA TTCTTAATAA
 11641 CTCAGCGGG CCCTCTCCCA ACACATGTCA TTGCACTTG GAACATCTCAC ACCCGAGAAC
 11701 ACCTCTCTAT TGCACTTCAC CTCATCCCT TAATTCTACT CGTACTTAAG CCCGAGCTGA
 11761 TCTGAGGCTG AACTGCCTGT AGGTGTAGT TTAACAAAAA CATTAGATTG TGATTCTAAA
 11821 ATAAGGGTT AAAATCCCT CTCCTCCAGGA GAGAGGCTCG CAGCAATGAG AACTGCTAAT
 11881 TCCCACGACC TTGGTTGGAC CCCTAGGCTC ACTCGAAAGC TTCTAAAGGA TAACAGCTCA
 11941 TCCGTTGGTC TTAGGAACCA AAAACTCTTG GTGCAAATCC AAGTAGCAGC TATGTACCCC
 12001 ACCTCCCTAA TGATTCATC GAGCCTGATC ACAATCTTG CGTTACTAGC CTATCCTTTA
 12061 ATCACCAAGA TCCGCCCTAC ACCTCGGGAC CCTCAATGGG CTACCAACCCA TGTAAAACA
 12121 GCTGAAAAAA TAGCTTTCTT TGTTAGCTG CTGCCCTCTGA CCCTATTCTTA TAATGAGGGC
 12181 GCAGAAACAA TCGTTACTAA CTGAACCTGA ATAATACCA ACTCCTTCCTA TGTCAGCATT
 12241 AGCTTGAAAT TTGATTTCTA CTCAATTATT TTTACACCC TTGCCCTCTA CGTAACTTGA
 12301 TCCATCTGG AGTTCGCATC ATGGTATATG CACCGAGACC CACACATGAA CGCCTCTTT
 12361 AAGTACCTCC TGACCTTCT GATTGCCATA ATCGCTCTAG TAACTGCAA CAATATGTTT
 12421 CAACTATTAA TCGGTTGAGA GGGCGTGGGG ATCATATCGT TCTTACTTAT TGGATGGTGG
 12481 TTCGGACGGG CGGATGCTAA CACTGAGCG CTTCAAGCAG TACTTTACAA CCGAGTAGGG
 12541 GACATCGGGC TTATCTTCG CATTGGCTGA ATAGCAGCAA ACCTAAACTC CTGGGAGATG

12601 CAACAAATCT TTGCAACCAC CAAAGACATA GACTTAACCT ACCCTTCTCT CGGACTCATC
 12661 ATTGCAGCAA CTGGCAAATC GGCTCAATTG GGACTTCACC CTTGACTACC CTCCGCTATG
 12721 GAGGGTCCCA CACCGGTCTC TGCCCTACTT CATTTAGCA CTATGGTCGT TGCCGGCATT
 12781 TTTCTGCTAG TACGAATAAG CCCGCTACTG GAAAACAACC AAACCGCCCT CACAACCTGC
 12841 CTATGCCTCG GAGCTTAC AACACTATTG ACTGCAACCT GTGCCCTCAC GAAAAATGAT
 12901 ATTAACCAAA TCGTCGCATT CTCTACATCC AGCCAGCTTG GGCTTATAAT GGTAACCATT
 12961 GGTCTAAACC aGCCACAAT AGCATTCTT CACATCTGTA caCATGCCTT CTTCAAGGCC
 13021 ATACTATTTT TATGCTCCGG ATCTATTATC CACAGCCTAA ATGATGAACA AGATATCCGG
 13081 AAAATAGGGG GCATACACCA CCTTGACCTT TTACATCCTT CTGCTTAAC TATTGGCAGC
 13141 CTCGCCCTCA CAGGAACCCC CTTCTGGCT GGCTCTTCT CTAAGATGC CATCATCGAA
 13201 GCATTAACCA CATCCCACCT AAACGCCCTGA GCCCTAACCT TAACCCTCT AGCCACCTCT
 13261 TTCACAGCAA TCTACAGCTT TCGAGTAGTA TTTTTGTAC CCATGGGCCA CCCCCGATTT
 13321 AGTCCCCTCT CTCCGATCAA CGAAAACAAC CCAGCAGTGA TAAACCCACT TAAGCGACTA
 13381 GCATGGGGGA GTATCGTGC AGGCTTAAGT ATTACCTCTA ATATTACACC CTTGAAAACC
 13441 CCTGTCATAT CTATACCTCC CCTCCCTAAA CTAGCTGCTC TTGCACTCAC AATCATCGGC
 13501 CCTCTTGTG CTATAGAGCT CGGCATGTTA ACCAACAAAC ATATGAGCC AATACCCAAC
 13561 ACGCACATAC ACCACTCTC TGCCTTATTA GGCTCTTCC CGGGGTTAT ACACCGCTA
 13621 ACCCCCTGTAT TGCTCTCTT CCTAGGCAA AACATTGCTA ACCAACAGT AGATCAGACA
 13681 TGATTAGAAA AAACGGGCC CAAAGCATT GAAGCCGCTA ACCGACCTCT CGTCTCTCT
 13741 ACAAGTAACA TTCAACGGGG GGTAATCAAA ACATATCTCA CCCTCTTCCT CTTACACTA
 13801 GCCCTTATAA TCCCCCTATT CATCACCTAA ACCGCTCGCA ACGCCCTCG ACTCATACCC
 13861 CGAGTTAATC CAAGGACTAC AAATAAAGTT AAAAGCAATA CCCATGCACT AGTAATTAAAC
 13921 ACCCACCCCC CAGAAGAGTA TATTAAGGCT ACCCCTCCAA TATCACCTCG AAACACGGAA
 13981 AACTCCCTA ACTCATCTG TAGAACCCAA GAAGATTCT ACCACCCCCC TCAAAACAGC
 14041 CCGGAAACTA ATAATACTCC CCCTGTGAGT ACAACTCCGT ATACCATTAC CGACCAACTA
 14101 CCCCAGCTCT CTGGGTACGG CTCAGCAGCT AACGCTGCTG ATATATGAAA TACAACAAAC
 14161 ATCCCGCTA AGTAAATTAA GAACAAGACT AGTGACAAAA AGGGCCCTCC ATGCCGACT
 14221 AATATACCAC ACCCCATGCC AGCTACCACA ACCAAACCCA AAGCAGCAAA ATAAGGGGAC
 14281 GGATTTAGAAG CAACTGCAAC TAACCCCAAC ACAAGAGAAA ATAAAACCTAA ATATATAGCG
 14341 AAAGTCATAA TTCTGCGAG GACTTTAACG AGGACTAATG GCTTGAAAAA CCACCGTTGT
 14401 TATTCAACTA CAAGAACCTT AATGGCCAGT CTACGTAAAT CCCACCCCTCT TCTAAAATC
 14461 GCAACAGATG CTTTAGTCGA CCTCCCCGCC CCCTCTAATA TCTGGTTTG ATGGAACCTT
 14521 GGGTCTCTT TAGGACTCTG TTTAATTACC CAAATTGCGA CGGGCTTATT TCTAGCCATA
 14581 CACTACACAT CAGACATTG TACTGCCCT ACCTCCGTGG CCCACATCTG TCGAGACGTC
 14641 AACTACGGCT GACTTATCCG AAGCATTCT CATCACCGC CATCATTCTT TTTCATTG
 14701 CTCTACCTTC ATATTGCCCG AGGACTATAC TATGGCTCTT ACCTCTATAA AGAAACATGA
 14761 ACTGTTGGGG TTATCTCTT CCTTCTCGTA ATAATAACAG CTTCTGGTGG ATACGTCTC
 14821 CCTTGAGGAC AAATGCTGTT TTGAGGTGCA ACCGTTATTA CCAACCTCTT ATCTGCCGTC
 14881 CCCTACGTG GAAACACCC AGTCAATGA ATCTGAGGGG GTTTTCTGT AGACAATGCT
 14941 ACCCTCACCC GGTTCTTGC CTTCCACTT CTCTTCCCT TCATCATCG TGCCGCAACA
 15001 TTAATTCACTC TACTCTTCT TCACGAGACT GGCTCAAACA ACCCCACCGG GCTAAACTCA
 15061 GACTCCGACA AAGTTCCCT TCACCCCTAC TTCACTTACA AAGACCTCTT GGGCTTGCA
 15121 GTACTTCTTA CTGATTAGC ATCCCTAGCC CTATTTTCCC CAAATCTCTT AGGAGACCCA
 15181 GATAACTTTA CCCCTGCAA TCCGCTGTT ACTCCCCAC ACATTAAGCC AGAGTGTATAC
 15241 TTCTCTGTTG CCTATGCTT TCTACGCTC ATCCCAAACA AACTGGAGG AGTACTTGCC
 15301 CTTTTTATTCT CCATCTTAGT TCTTATGCTC GTCCCTTCTC TTCTACACCTC TAAACACGA
 15361 AGCTTAATAT TCCGCTCCCT AACACAATTG TTGTTTTGGT CTTTAGTAGC AGACGTTATG
 15421 ATTCTAACTT GAATTGGAGG GATGCCGTA GAACACCCCT TCGTCATCAT TGGACAAGTA
 15481 GCGTCTCTCC TCTACTTCTC CCTCTTCCCTA GTCTGTATCC CAACAGCAGG TTGACTAGAG
 15541 AATAAAAGTCC TTGGATGAAA ATGCACTAGT AGCTCAGCAC ACAGAGCCCC AGTCTTGAA
 15601 ACTGCCGTC GGAGGTTAGA ATCCCTCTA CTGCTCAAAG AAAGGAGATT TCAACTCTA
 15661 CCCTCTAACTC CCAAAGCTAG GATTCTAGCA CTAAACTATT TTTGTGGTA CATAAAATGTA
 15721 CAATGAAGGA TTTTCATATA CATGTATGTA ATAACACATT ATTTTATAG TAACCATT
 15781 ATGTGATGTA CTAGGACATA CATGTATAAT CACCAAATCT CGTAATACAG CACTCATCA
 15841 TCAACATTTT TAACTAAGAT AGACTAAAAC CTGAATAATC ACTAATCTTA AATATGTGAA
 15901 AGTCAGGAC CAGTCGAAAT CTAAGACCGA ACACAACACT CATCAGTCGA GTTATACCAA
 15961 GACTCAAAT CTCGCCGACC CCAAAATCTT ATGTTAGTAAG AGCCTACCAA CGGGTGATTC
 16021 CTTAATGATA ACTCTTATG AGGGTGAGGG ACAAAAATTG TGGGGGTTTC ACTCGGTGAA
 16081 TATTTCTGG CATTGGTTC CTACTCAGG GCCATTGATT GATGTTATCC CTCACACTT
 16141 CATGACACT TACATAAGT AATGTTGATA ATACATACAG CTCGTTACCC ACCAAGCCGG
 16201 GCGTTCACTC CAGCGGTAA GGGGTTCTCT TTTTTTTTT CTTTCACTT GACATTCAA
 16261 AGTGCATCCA GCCAACGGAG ACGTTAAAG GGTGAGCACT TTTCTTGAC GCGCGTACA
 16321 TAGTATCCGT GTAATGAGTC TTTATTAGAA GGATAACATT AAGTGTATC ATGTGCATAA
 16381 AGATTTGCTC ATTTCTCTG TCATCCCCAG GATAACCCCT TTTTGCGCG CAAAACCCCC
 16441 CTACCCCTTA AAACCCCTGA AGTTGCTAAG ACCCCTGAAA ACCCCCCGGG AACAGGACAA
 16501 ACCTCTGGTA GCTCAGAAAG GGCAATTACT TTTCAACCC TTTCAACCA CTTTCAACC
 16561 CACAAATACC CCTGGCTTAT CATAACCCCA CCAGTTGTTA TATAAATACC ACTTTCAAC
 16621 CCACAAATAC CCCTGgCTTA TCATAACCCCA ACCAGTTGTT TATAAATAC CACTTTCAA
 16681 CCCACAAATA CCCCTGGCTT ATCGTAACCC CACCAAGTTG TTTAATAATA CCACCTTCA
 16741 ACCCACAAAT ACCCCCTGGCT TATCGTAACC CCACCAAGTTG TTTAATAAT ACCACTTT
 16801 AACCCACAA TACCCCTGGC TTATCGTAAC CCCACCAAGTT GTTTTAATAA TACCACT
 16861 CAACCCACAA ATACCCCTGG CCTTATCGTA CCCACCAAGT GTTTTTAATAA ATACCACT
 16921 TCAACCCACA AATACCCCTG GCTTATCGTA ACCCCACAG TGTTTTAAT AATACCACT
 16981 TTCAACCCAC AAATACCCCT GGCTTATCGT AACCCCAACCA GTTGTGTTAA TAATACCACT
 17041 TTTCAACCA CAAATACCC TGGCTTATCG TAACCCCAAC AGTTGTTAA ATAATACCA
 17101 TTTTCACCC ACAAAATACCT CGTGGCTTAC GAAACCCAC CAGTTGTTT AATAATACCA
 17161 CTTTTCAACC CACAAATACC CCTGGCTTAT CGTAACCCCA CCAGTTGTTT TAATAATACC
 17221 ACTTTCAAC CCACAAATAC CCCTGGCTTA TCGTAACCC ACCAGTTGTT TTAATAATAC
 17281 CACTTTCAA CCCACAAATA CCTAAACCC ACAACACGGC TATCACCTC TTGATGATCA

17341 AAAAACTACT GTTTAGATTA TTTCAAGTAT TCCTTCACC ACCCTAACTA CGAAAGGCAT
 17401 AAAACAGAAT CCAAGAAAAC TGCTCCATCC AAAACTATCT TTTTAACCG ATAAGCCCTT
 17461 AAAGTCCCCCT GAGAACCGAGC CTTTTACCGC CCCAAAGGCT AAACACAAAC AACTACCAC
 17521 CTCTAGGTGT GTTTAGAAAA TTTACT

Hippoglossus hippoglossus 2 17619 bp

1 GCTAACGTAG CTTAATTAAA GCATAACACT GAAGATGTTA AGATGGGCC TAGAAAGCCC
 61 CGGGAGCACA AAGGCTGGT CCTGACTTTA CTGTCGACTT TAACCTAAACT TACACATGCA
 121 AGTATCCGCC CCCCTGTGAG AATGCCACA ACTCCCTGCT TGGGAACCTAG GAGCCGGTAT
 181 CAGGCCACAAG CCCAGCTAGC CCACGACCC TTGCTTAGCC ACACCCCTCAA GGGAACTCAG
 241 CAGTGTATAA TATTAAGCCA TAAGTAAAAA CTTGACTTAG TTAAGGTTAA GAGGGCCGGT
 301 AAAACTGTGAG CAAAGCCACCG CGGTTATACG AGAGGCCCAA GTTGACAAAC AACGGCGTAA
 361 AGAGTGGTTA GGGGATTAC TAAACTAGAG CCGAACGCTT TCAAAGCTGT TATACGCACC
 421 CGAAAGTATG AAACCCAAATT ACGAAAGTAG CTCTACTCAT CCTGAACCCA CGAAAGCTAA
 481 GAAACAAACT GGGATTAGAT ACCCCACTAT GCTTAGCCCT AAACATCGAT TGCACCTAC
 541 ACTCCATATC CGCCCGGGAA TTAtgAACt CAGTTAAAAA CCCAAAGGAC TTGGCGGtGC
 601 TTACATCCA CCTAGAGGAG CCTGTTCTAG AACCGATAAC CCCCGTTAAA CCTCACCTTC
 661 TCTTGTTTA TCCGCTATA TACCACCGT GCAGCTTAC CTTGTAAGG CTTTACAGTA
 721 AGCAAAATTG GCACAGCCCA AAACGTCAGG TCGAGGTGTA GTGAATGAGG GGGGAAGAAA
 781 TGGGCTACAT TTGCTAAACA TAGCAAACAC GAATGTTGCA TTGAAACATG TAACTGAAGG
 841 AGGATTAGC AGTAAGCAGG AAATAGAGCG TCCCGCTGA ACTGGCCCTA AAGCGCGCAC
 901 ACACCGCCCG TCACCCCTCC CAAGCCCCCT AAACCTAAACT ATTAAAACC CTACAACCCG
 961 CGAAGGGGAG GAAAGTCGTA ACATGGTAAG TGTACCGGAA GGTGCGCTTG GAAAATCAG
 1021 AGTGTAGCTA AGATACTATA CAGCATCTCA CTTACACCGA GAAGACGTCC GTGCAAGTCG
 1081 GACCACCCCTG ACGCCTATA GCTAGCCAA CCCCTTAACA CAACAGACCC CCATTTATAA
 1141 CCCCTAAAGC ACGAACACCC CACGTAGCTA AACCTATTCTC CCCCCTAAGT CCAGGCGATA
 1201 AAAAGGAAA TTTTGGAGCA ATAGAAAAAG TACCGCAAGG GAAAGCTGAA AGAGAGATGA
 1261 AAAAGCCAG TAAAGCTAA GAAAGCAGAG CTTAGTGCTC GTACCTTTTG CATCATGATT
 1321 TAGCTAGCAC TTTCAAGCAA AGAGAACCTA CAGTTGTA CCCCAGAAACT GGGTGAGCTA
 1381 CTCCAAGACA GCCTATTAT AGGGCGAACCG CGTCTCTGTG GAAAAGAGT GGGAAAGAGCT
 1441 TTGAGTAGAG GTGACAAACC TACCGAACTC AGTATAGCT GTTGCCTGT GAATTGAATA
 1501 GAAGTTCAGC CCCCTGGTT CTCCACTCAT GCACCTTGT AACCCCTCAG ATGCAATGAG
 1561 AAACCAAGGG TGTAGTCA AGGGGTACA GCCCCTTGTA TACAAGACAC AACTTCTCA
 1621 GCAGGATAAA GATCATATTC AAATAAGGAC AAATGTTTA GTGGGCCTAA AAGCAGCCAC
 1681 CTAAGATGAA AGCGTTAAAG CTCAAACATA AAGCCCTCCC GGTATACCGA TAACCCATAC
 1741 TTATCCCCC ACATTTAACAA GGCCTCTCTA TGACCCACAT AGGAACGACT ATGCTAATAT
 1801 GAGTAATAAG AGGGTATAAT CACCCCTCTCC TTGCACATGT GTAAATCGGA ACGGACCCCC
 1861 CACCGAATCT TAACGGCCC AATCAAGAG GGTATTGAA ACTACCACGA ATTTAGGCCA
 1921 GAAAACATC CAATGCAAAC CCGTTAACCC CACACTGTC TGCCCAAAG GAAAGACCAA
 1981 AGGGGGGAGA AGGAACCTGG CAAACATACC CCAAGCCTCG CCTGTTTAC AAAAACATCG
 2041 CCTCTTGATC AACACAGATA TAAGAGGTC CGCCTGCCA GTGACCATAT AGTTCAACCG
 2101 CGCGGTATT TTGACCGTGC AAAGGTAGCG TAACTACTTG TCTTTAAAT GAAGACCTGT
 2161 ATGAATGGCA TAACGAGGGC TTAACTGTCT CCTCCCCCG GTCAATGAAA TTGATCTCCC
 2221 CGTGCGAGAG CGGGGATTAA ACCATAAGAC GAGAAGACCC TATGGAGCTT TAGACACACA
 2281 GGTGGGCCAT GTCAAAATACC CCTAGCTAAG GGCTCTGAAC AAATGGAATC TGCCTTGATG
 2341 TCTCGGTTG GGGCAGCAT GGGGATAACA AAACCCCCAC GTGGAAAGGG AGCACACCCC
 2401 TAAGTTACTT CTTCTCCCCG AAGCCAGAGC AACAGCTCTA ACCAGCAGAA ATTCTGACCA
 2461 AACTGATCCG GTAAACACCG TCAACGAAAC AAGTTACCTC AGGGATAACA GCGCAATCCC
 2521 CTTTTAGAGC CCATATCGAC AAGGGGTTT ACGACCTCGA TGTTGGATCA GGACATCCTA
 2581 ATGGTGCAGC CGCTATTAAG GTTCTGTTG TTCAACGATT AAAGTCCTAC GTGATCTGAG
 2641 TTCAGACCGG AGTAATCCAG GTCAGTTCT ATCTATGACA TGATCTTTTC TAGTACGAA
 2701 GGACCGAAAA GAGGAGGCC CTGCTAAAG TACGCCCTCAC CCCCACCTAA TGAAAAATC
 2761 TAAACTAGGC AAAAGGGCAT AACCTTTTG CTGAGATAA CAGCAAGTTG GGGTGGCAGA
 2821 GCGCGGCTAA TGCAAAAGAC CTAAGCCCTT TCTACAGAGG TTCAAGTCCT CTCCTTAAC
 2881 ATGATTCAA CTCTCATTAC TCATATTATC AACCCACTAA CCTTATTG TGCTGTATTG
 2941 CTAGCGTAG CATTCTCAC CCTCTTGA CGAAAAGTC TAGGCTACAT ACAACTACGA
 3001 AAAGGCCAA ACATTGTTGG GCCTTACGGC CTCTTCAAC CTATTGCTGA TGGCGTAAAA
 3061 CTCTTATTA AAGAACCCAT TCGACCTCA ACCGCATCAC CGTCCTATT CCTCGTGGCC
 3121 CCTATGCTCG CTCTCACACT AGCCCTAACCT CTTGAGGCC CCATACCTT CCTCGTACCC
 3181 GTCTGGGACCT TAACTCTAGG TATTCTTTT ATTCTAGCTC TATCTAGCTT TGCACTGTAT
 3241 TCTATTCTGG GCTCGGCTG AGCATCTAAT TCAAAATATC CCTCTAGTAGG GGCACCTCGG
 3301 GCTGTCGAC AAAACATTTC TTACGAAGT AGCCTTGGAC TCATCTTACT AAACATTATC
 3361 ATCTTCACCG GGGGCTTCAC ACTCCAAACCC TTTAACACAG CCCAAGAGGC CATTGACTG
 3421 GTAGTGCCTG CTTGACCGCT GGCTGCCATA TGATATATCT CCACCTTGC CGAAACAAAC
 3481 CGTGCACCCCT TCGACTAAC AGAAGGGAA TCTGAACCTTG TCTCGGCTT CAACGTAGAG
 3541 TACGCAGGG GACCCCTCGC CTTGTTCTC TTAGCCGAAT ATTCAAATAT TCTCCTAATA
 3601 AATACCCCTGT CCGCTACACT ATTCTTCTGGG GCCTCCACCA TCCCACATCTAT CCCGGAATTA
 3661 ACTAGTATCA ATATTATGAC TAAAGCAGCC CTTCTATCAA TTGCTTCCCT GTGGGTCCGA
 3721 GCTTCGTCACC CGCGATTCCG TTACGACCAA CCTTATGACCC TCATTGAAA AAACCTTCTT
 3781 CCACTGACTT TGGCACTAGT TATCTGACAC TTGGCGCTCC CTATTGCTT CGCTGGTCTC
 3841 CCACCAAC TATAGGCCA GGAGCTGTC CTGAATTAA AGGGCCACTT TGATAGAGTG
 3901 ATTATGGGG GTTAGAGTCC CCCCAACTCC TTAGAAAGAA GGGGATCGAA CCCAACCTGA
 3961 AGAGATCAA ACTCTTCGTG CTTCCTCTAC ACCACTTCCT AGTAAGGTCA GCTAAATAAG
 4021 CTCTTGGGCC CATAACCCAA CCATGTAGGT TAAATCCCT CTTTGCTAA TGAACCCCTA
 4081 CATCTTGACCC ACTCTTCAT TTGGTTGGG CCTTGGAAAC AACTCACGT TTGCAAGCTC
 4141 ACATCTGGCTC CTCGCTTGA TAGGACTTGA GATAATACA CTAGCCATTA TTCCCTTAAT
 4201 AGCCCAACAC CACCAACCCCC GGGCAGTCGA AGCTACTACT AAATACCTTC TCGCACAAGC
 4261 CACAGCAGCC GCCACCCCTCC TGTTGCAAG CACTACCAAC GCTTGACTCA CAGGCCAATG

4321 GGATATTCAA CAGATAACAC ATCCCCCTCC CACAACAATG ATTGTTATTG CCCTCGCGCT
 4381 AAAAGTCGA CTGGCACCAA TGCACCTTG ACTCCCCGA GTACTCCAAG GGCTAGACCT
 4441 GACCACCGGG CTCATCCTCT CGACCTGACA AAAACTTGCA CCCTTGCCC TCTTACTACA
 4501 AATTCAAACAA ACCAACCTA CCCCTCTAAT CATTATTGGC TTACTATCCA CTCTCGTTGG
 4561 TGGCTGAGGT GGTCTCAACC AAACCCAGCT GCGCAAGGTC CTCGCCTACT CCTCAATTGC
 4621 CCACCTTGGC TGAATAATGC TGATCCTCA ATTTtCACCC CTCCCTGCC TTCTCACCC
 4681 CCTCACGTAC TTTACTATGA CTTTCTCAGC ATTCCTAATT TTTAAAGTAA ACAAAGCGAC
 4741 CACCGTCAAT GCCCTCGCA TCTCGTGAAC AAAAACTCTC GCCCTCACAG CCCTAGCACC
 4801 CCTGGTTTA CTTTCTCTCG GCGGCCCTCC TCCCCTTACC GCCTTCATGC CCAAGTGATT
 4861 CATTCTTCAG GAACGTACTA ACAAGACCT TCCGGCACTT GCCACCCCTCG CCGCTTAAAC
 4921 TGCCCTTGT AGTCTCTACT TCTACTTGGC CCTCTCGTAT GCAATAACCC TCACAATATT
 4981 CCCCAACAAAC CTAATTGGTG TCACTCCCTG ACGATTCTAC TCCCCCCTAAT TTACACTTCC
 5041 CCTCGCCGTC TCCACTGCAG CAACCACCT TCTCTCCCA CTAGCCCCCG CTGCCGTAGC
 5101 ACTACTGGTC ACCTAACGGGA CTTAGGCTAG CACTTAGACC ATGGCCCTTC AAAGCCGTAA
 5161 GCGTGAAGTGA AAATCTCTCA GTCCCTCTTA AGACTTGGCG GCTACTATCC CACATCTCT
 5221 GCATCGAAAG CAAACACTT AATTAACTA AAGCCTTCT AGATAGGAAG GCCTTGATCC
 5281 TACAAAATCT TAGTAAACAG CTAAGTGCCTT AATCCAGCGA GCATCTATCT ACTTTTCCCC
 5341 GCCTAATTAA GAAACTAAA GGCGGGAAA AGCCCCGGCA GGCGAGTTAGC CTGCTTCTT
 5401 AGATTTGCAA TCTAACATGT AACACCCAG GGCTTGGTAA GAAGAGGGCT TGCACCTCTG
 5461 TCTATGGGTC TACAATCCAC CGCTTAACCTC AGCCATCCTA CCTGTGGCAA TCACACGTTG
 5521 ATTTTCTCG ACCAATCACA AAGACATCGG CACCCCTCAT CTCGTATTG GTGCCTGAGC
 5581 CGGAATAGTG GGGACAGGCC TAAGTCTGCT TATCCTGGCA GAACTAAGCC AACCCGGGGC
 5641 TCTCCTGGGA GACGACCAAA TTTATAATGT GATCGTCACC GCACACGCC TTGTAATAAT
 5701 CTTTTTATA GTAATACCAAA TTATGATTG GGGGTTGGAA AACTGGCTTA TTCCACTAAT
 5761 AATTGGGCC CCAGACATGG CGTTCCTCG AATGAATAAT ATGAGTTCT GACTTCTTCC
 5821 CCCCTCCTT CTCCTCCTCT TAGCCTCTTC AGGTGTTGAA GCCGGAGCAG GTACCGGATG
 5881 AACCGTGTAC CCCCCACTAG CTGGCAATT TGGCCACGCC GGGGCATCC TAGACCTGAC
 5941 AATCTCTCA CTTCACCTG CAGGAATTTC ATCAATTCTG GGGGCAATTA ACTTTATTAC
 6001 TACCATCATT AACATGAAAC CCACACAGT CACTATGTAC CAAATCCCGT TATTGTTTG
 6061 AGCCGTTCTT ATTACAGCCG TACTTCTCT TCTGTCCTG CCGGTTTTAG CCGCAGGGAT
 6121 TACAATGCTA CTAACAGACC GCAACCTAA CACGCCCTTC TTTGACCCCTG CCGGAGGAGG
 6181 TGACCCCAT TCTCACCAAC ACCTATTCTG ATTCTTTGGC CACCCAGAGG TATACATTCT
 6241 TATTCTCCCA GGCTTGGAA TAATTCTCA CATGTTGCA TACTATGCAG GTAAGAAAGA
 6301 ACCTTTGGC TACATGGGA TAGTCTGAGC TATAATAGCC ATTGGACTCC TGGGCTTCAT
 6361 TGTCTGGCC CATCACATGT TTACAGTCGG AATAGACGTAA GATACACGAG CCTACTTAC
 6421 CTCTGCCACA ATAATCATG CGATTCCAAC TGGCGTAAAAA GTCTTTAGCT GACTCGAAC
 6481 CCTCCATGGG GGAAGCTTA AATGAGAAAC GCCCCTCTA TGAGCCCTCG GCTTTATTTT
 6541 CCTCTTACA GTAGCGGTCACTC TCACTGGCAT TGTCTTAGCT AACTCCTCTC TCGATATTGT
 6601 TCTGATGAC ACATACTATG TAGTCGCCCA CTTCCACTAT GTACTATCTA TGGGTGCTGT
 6661 ATTTGCAATC GTTGCCTCT TCGTCATTG ATTTCCTGTTA TTTACAGGCT ATACCCTCTA
 6721 CTCCACATGA ACAAAAATCC ACTTCGGCT GATGTTTATT GGGGTCAATC TAACATTCTT
 6781 CCCTCAACAT TTTCTGGGC TGGCTGGAT ACCCCGACGG TACTCAGACT ACCCAGACGC
 6841 ATACACCCCT TGAAACACTG TTTCATCAAT TGGGTCCCTA ATGTCCTCG TTGCTGTAAT
 6901 TTTATTCTTA TTCATTATTG GAGAAGCATT TACAGCCAA CGAGAAGTCG GAGCAGTAGA
 6961 ACTAAACTGCA ACTAACATG AATGACTTTA CGCGTGCCT CCCCCCTACC ACACATTGAA
 7021 AGAGCCCGA TTCTGACAAAG TTCTGATAAA TTGCAACAAA CTAACGAGAA AGGGAGGAGT
 7081 TGAACCCCCA TCAATTGGT TCAAGGCAAC CACATAACCG CTCTGTACT TTCTTCACAA
 7141 CACACCAATA AGATACTAGT AAAACGCTAT AACACTGCCT TGTCAAGGCA GAATTGTGGG
 7201 TTCAACCCCC GCGTATCTTA AACGCAATGG CACATCCCTC ACAACTCGGA TTCAAGACG
 7261 CAGCTTCACC CTTAATAGAA GAACTACTTC ACTTCCACGA CCACGCCCTTA ATAATTGTAA
 7321 TTCTCATTAG CACAACTGTA CTTTATTTA TTGCGGCTAT AGTCACAGCC AAACTAACAG
 7381 ATAAACTTGT ACTAGACTCT CAAGAAATTG AAATTATCTG AACAGTTCTC CCAGCTGTTA
 7441 TTCTTATCTCT TATGCCCTA CACATCCATC GAATTGTTG CTTAATAGAC GAAATTATG
 7501 ACCCCCCACCT TACAATCAA GCCATCGGCC ACCAATGATA CTGAAGCTAC GAATACACAG
 7561 ACTACGAAGA CCTCGGTTT GACTCATACA TAATCCCCAC ACAAGACCTG ACCCCTGGAC
 7621 AATTCCGACT ACTAGAACG GACCACCGAA TAGTAATCCC AGTAGAAATCT CCAATTCCG
 7681 TTTTAATCTC GGCTGAAGAT GTCTTACATT CCTGGGCAGT CCCCTCTCTA GCGTAAAAG
 7741 TTGACGCCGT ACCTGGACGA TTGAACCAAG CACCTTTAT GTTAGCCGA CCAGGCGTAT
 7801 TTTTCGGACA ATGTTCTGAA ATTGCGGGG CTAACCATAG TTTTATACCC ATTGTTGTAG
 7861 AAGCAGTCCC CCTTGAAACAC TTTGAGAACT GGTCTTCACT AATAATTGAA GACGCCCTCG
 7921 TAAGAAGCTA ATAAGTACAA GCGTTAGCTT TTTAAGCTAA AGACTGGTGC CTAACAACCA
 7981 CCCCTAGCGA CATGCCCTAA CTGAACCCCG CACCTGATT TGCAATTCTT GTTTCTCTT
 8041 GAATAATCTT TTTAACTGTT ATTCCCCCA AAGTTTAGC ACACACCTAT CCCAATGAAC
 8101 CCACCTCCCA AAGCACACAA AAACCTAAAAA CAGAGCCCTG AAACTGACCA TGATACTAAG
 8161 CCTCTTGTAC CAATTATAT CCCCCTCATC CCTCGGCATT CCACTAATTG CACTAGCGAT
 8221 TAGCCTGCCT TGAGTTCTGT TCCCAACCCC TCAAGCCGC TGATTAACAA ACCCGTGTCT
 8281 AACACTTCAGCA GGGTGAATTG TTAGCGCTT CACCTCACAA CCTCTCTAC CCCTAAACCC
 8341 TGGGGGCCAC AAGTGGGAG TTCTATTAC CTCGTTAATA GTTTTCCCTAT TTTCTATCAA
 8401 CATGCTAGGA CTCCCTCCAT ACACCTTCAC ACCAACAAAC CAACTCTCCC TTAACATGGG
 8461 CCTCGCAGTA CCCCTTGAT TGGCAACTGT CATTATCGGG ATACGAAATC AACCAACGCA
 8521 TGCCCTAGGC CACCTTCTTC CAGAAGGCAC CCCCACTGCC CTCATCCCTG TACTTATTAT
 8581 CATTGAAACA ATTAGCTTAT TCATTGACCC CTTAGCTCTC GGTGTCCGAC TAACAGCAAA
 8641 TCTCACAGCA GCCCACCTCC TCATTCAACT CATTGCAACG GCTGCCCTCG TCCTTCTTCC
 8701 CCTCATACCT GTTGTGCGTA TCTTAACAAAC GGTGGTCCTC TTCTCTCTA CTCTACTAGA
 8761 AGTAGCGTT GCCATGATC AAGCTTATGT CTTGTTCTA CTCCTAAGCC TATATCTACA
 8821 AGAGAACGTT TAATGCCCA TCAAGCACAC CCATACCCACA TAGTCGACCC CAGCCCATGA
 8881 CCCCTAACAG GGGTATTG TGCCCTATTG ATAACATCTG GCCTTGCTAT CTGATTCCAC
 8941 TTCCACTCCA CAACCTAAT AACTATTGGG ACAATCCTTC TCACCTTAAC AATCTTCAA
 9001 TGATGACGGG ACGTGTCAG AGAAGGCACA TTTCAAGGAC ACCACACTCC CCCCGTTCAA

9061 AAAGGCCTCC GATATGGGAT AATTCTTATT ATTACCTAG AAGTCCTATT CTTCTTAGGC
 9121 TTCTTCTGAG CCTTTTACCA CTCAAAGCCTA GCACCTACCC CTGAGCTAGG CGGCTTCCTGA
 9181 CCACCGAGCAG GCATTACTCC CTTAGACCCA TTGAAGTCC CACTTCTTAA CACAGCAGTC
 9241 CTGCTCGCCT CTGGCGTAAC TGTTACCTGA GCACATCACA GCATTATGGA GGGTAAACGA
 9301 AAACAAACTA TTCAATCTCT TGCTCTTACA ATCTTACTCG GGGGGTACTT CACTTCCTT
 9361 CAAGGCCTAG AGTATCACGA AGCCCCCTTC ACCATTGCAG ACGGAGTTA CGGCGCTACA
 9421 TTCTTGTGAG CCACCGCTT CCACGGATTA CACGTCTTAA TTGGCACATC ATTCTTAGCC
 9481 GTTTGTCTAT TACGCCAGAT TCTTCACCAT TTACATCAACC ACCACCACTT TGGGTTTGAA
 9541 GCAGCGCCTA GATACTGACA CTCGAGTAC GTCGCTGAC TTTCTCTA TATCTCTATT
 9601 TACTGATGAG GATCTTAATC TTCTAGTAA TAAATCTAGT ATAAGTGACT TCCAATCACC
 9661 CGGTCTTGGT TAAAATCCAA GGGAAAGATAA TGAGCCTTCT CTTAACCATC ATTTCGATTG
 9721 CCACCCCTCT CTCGACAGTA CTAGCCCTG TATCCCTTTG ATTACCCCAA ATTACACAG
 9781 ACCACGAAAA GCTATCGCCG TACGAGTGC GCTTGACCC GATAGGATCC GCCCGGCTGC
 9841 CTTTTTCACT GCGGTTTTC CTCATCGCCA TCCCTCTTCT TCTCTTCGAT TTAGAAATTG
 9901 CTCTCCTCCCT CCCTCTTCCA TGAGGAGAC AACTAGCATC GCCACTACTG ACGTTCACCT
 9961 GAGCTACAGC TGTCTTAGCC CTTCTAACCC TCGGCCCTCAT TTACGAATGA ATGCAAGGAG
 10021 GTCTAGAATG GGCTGAATG TGTTAGTAC TAAGAAAAAA ATTGTGATTC GGCTCAAAAAA
 10081 CTTGTGGTT AAATCCCAA CCGCTTAATG ACCCCCCACAC ACTTTGCCTT CTCCCTAGCC
 10141 TTTCTTCTGG GTTTAACAGG CTCGGCATTC CACCAAGTTT ACCTCCCTTC CGCCCTATTG
 10201 TGCCCTGAAG GAATAATACT ATCCCTATT ATTGCTCTCT CCCTATGAAC GCTCCAACCT
 10261 GACTCAACCA ACTTTTCAGC ATCTCCAATA CTTCTGCTT CATTTCAGC CTGCGAAGCA
 10321 AGCGCCGGCC TCGCCCTATT AGTAGGCACT GTCGAACCC ACGGGACCGA CCGATTACAA
 10381 AGCCTAAATC TTCTCCAATG CTAAAAGTCC TAATCCCAAC ACTTATGCTA ATCCCAACAG
 10441 CCTGGTTACT CAAACCCAGC TGACTCTGG CCATGACTTT AATGTATAGC TTTTGATCT
 10501 CTTTGTGTT CTTTGTGTT CTTTGTGTT CTTTGTGTT CTTTGTGTT CTTTGTGTT
 10561 TCTTTATAGC TACCGACTCC CTATCAACCC CCCTCCTCGT TCTTACATGC TGATTACTCC
 10621 CACTAATAAT TTTGGCAAGC CAAAACACA CAGCCTTACA ACCTCTTAGT CGCCAGCGCA
 10681 TGTACATCAC ACTTCTCGCC TCACTCCAGT TTTCTCTAAT CCTAGCATTT AGGCCACCG
 10741 AACTGTTGAT GTTTTACGTA ATATTGAAAG CCACCCATC CCCCACGCTA ATCATTATCA
 10801 CCCCGTGAGG AAACCAACA GAACGCTAA ACCGAGGAAC CTACTTTCTC TTCTATACAT
 10861 TAGCGGGCTC ACTCCCTCTT CTCGTTGTT TACTCTACT CAAAACACA TCCGGCACCC
 10921 TGTCGTGTT AACCCCTCAC TATACGGACC CGCTCGCTCT CTCGTTGTT GCAGACAAAC
 10981 TATGATGAGC AGGCTGTT TTAGCATTCC TGGTTAAAAT ACCCCCTCTAC GGGGTCCACC
 11041 TATGGCTCCC TAAAGCCAT GTTGAAGCCC CAATTGCAGG CTCATAATC CTTGCAGCGG
 11101 TTCTACTGAA GCTAGGGGA TACGGCATAA TCCGCATAAT ACAATATTG GAACCTCTAA
 11161 CTAAGGAGTT AAGCTACCC TTCATTGCT TTGCACTTTG AGGGGTAATT ATAACGGGT
 11221 CAATTGTCT ACGCAGACA GACCTTAAGT CCCTAATGCT CTAACCTCTCA GTCAGCCACA
 11281 TAGGCTTAGT GGCTGGACC ATTCTTATC AATCACCCCTG GGGCTTAACA GGGGCTCTTA
 11341 CCCTCATTAAT TGACACGGG CTTACTCATC CAGCCCTCTG CTGCTTAGCA AACACAAACT
 11401 ATGAACGAAC CCACAGTCGA ACGATAGTCC TGGCGCGAGG ACTCTAAGTA GCCCTACCC
 11461 TAATAGCCAC TTGATGATT ATTCTCTAGCT TGGCAACCT AGCTCTGCCA CCACTACCA
 11521 ACCTCATGGG GGAGCTGATA ATTATCACTT CTTATTCAA CTGGTCTGTA TGAACCTGG
 11581 CATTAAACGGG TGCAGGCACT CTAATCACAG CTGTTACTC CTTTACATA TTCTTAATAA
 11641 CTCAGCGGGG CCCTCTCCCA ACACATCTCA TTGCACTTGA GCCATCTCAC ACCCGAGAAC
 11701 ACCTCTTAT TGCACTTCAC CTCATCTCC TAATTCTACT CGTACTTAAG CCCGAGCTGA
 11761 TCTGAGGCTG AACTGCTGT AGGTGTAGT TTAAACAAAAA CTTAGATTG TGATTCTAAA
 11821 AATAAGGGTT AAAATCCCTT CTCCCAACCGA GAGAGGCTCG CAGCAATGAG AACTGCTAAT
 11881 TCCCACGACC TTGGTTGGAC CCCTAGGCTC ACTCGAAAGC TCCTAAAGGA TAACAGCTCA
 11941 TCCGTTGGTC TTAGgAACCA AAAACTCTTGTG CTCGAAATCC AAGTAGCAGC TATGTACCC
 12001 ACCTCCCTAA TGATTTCTAC GAGCCTGATC ACAATCTTGTG CGTTACTAGC CTATCCTTA
 12061 ATCACACCGA TCCGCCCTGC ACCTCGGGAC CCTCAATGGG CTACCACTCCA TGTCAAACAA
 12121 GCTGTAAAAA TAGCTTCTT TGTTAGTCTG CTGCTCTGA CCCTATTCTTA TAATGAGGGC
 12181 GCAGAAACAA TCGTTACTA CTGAACCTGA ATAATACACA ACTCCTCTAA TGTCACTGATT
 12241 AGCTTGAAT TTGATTCTA CTCATTTTACCA TTGCCCCCTCA CGTAACCTGTA
 12301 TCCATCCTGG AGTCGCATC ATGGTATATG CACCGAGACC CACACATGAA CCGCTTCTT
 12361 AAATACCTCC TGACGTTCT GATTGCCATA ATCGCTCTAG TAACTGCAA CAATATGTT
 12421 CAACTATTTA TCGGTTGAGA GGGCGTGGGG ATCATATCGT TCTTACTTAT TGGATGGTGG
 12481 TTCGACAGGG CGGATCTAA CACTCGACCG CTTCAAGCAG TACTTTACAA CCGAGTAGGG
 12541 GACATCGGGC TTATCTTCG CATGGCTTGA ATAGCAGCAA ACCTAAACTC CTGGGAGATG
 12601 CAACAAATCT TTGCAACAC CAAAGACATA GACTTAACCT ACCCTCTTCT CGGGCTCATC
 12661 ATTGCAAGCA CTGGCAACATC GGCTCAATTC GGACTTCACC CTTGACTACC CTCCGCTATG
 12721 GAGGGTCCCA CACCGGTCTC TGCCCTACTT CATCTCTAGCA CTATGGTCGT TGCCGGCATT
 12781 TTCTGTCTAG TACGAAATAAG CCCGCTACTG GAAAACAACC AAACCGCCCT CACAACCTGC
 12841 CTATGCTCG GAGCTTAC AACACTATTC ACTGCAACCT GTGCCCTCAC GCAAAATGAT
 12901 ATTAAAAAAA TCGTCGCTT CTCTCATCC AGCAGCTTG GGCTTATAAT GTAAACCAATT
 12961 GGTCTAAACCC AGCCACAATC AGCGTTCTT CACATCTGTA CACATGCCCT CTTCAAGGCC
 13021 ATACTATTTT TATGCTCCCG ATCTTATAC CACAGCTTAA ATGATGAACA AGATATCCGG
 13081 AAAATAGGGG GCATACACCA CCTTGACCTT TTACATCTC CTTGCCCTAAC TATTGGCAGC
 13141 CTCGCCCTCA CAGGAACCC CTTCTGGCT GGCTTCTTCT CAAAGATGC CATCATCGAA
 13201 GCATTGAACA CATCCCACCT AAACGCCCTGA GCCCTAACCT TAACCCCTCT AGCCACCTCT
 13261 TTCACAGCAA TCTACAGCTT TCGAGTAGTA TTTTTGTAC CCATGGGCCA CCCCCGATTT
 13321 AGTCCCCTCT CTCCGATCAA CGAAAACAAC CCAGCAGTGA TTAACCCACT TAAGCGACTA
 13381 GCATGGGGGA GTATCGTGC AGGCTACTA ATTACCTCTA ATATTACACC CTTGAAAACC
 13441 CCTGTCATAT CTATACCTCC CCTCCCTAAA CTAGCTGCTC TTGCACTGC AATCATCGGC
 13501 CTTCTGTGAG TCTATAGAGT CGCCATGTT ACCAACAAAC AATATGAGCC AATACCAAC
 13561 ACGCACATAC ACCACTCTC TGCCTTATTA GGCTTCTTCC CGGGGGTTAT ACACCGCTCA
 13621 ACCCCTGTAT TGCTCTTCTT CTTAGGCCAA AACATTGCTA ACCAACACAGT AGATCAGACA
 13681 TGACTAGAAA AAACGGGCC CAAAGCCATT GAAGCCGCTA ACCGACCTCT CGTCTCTCT
 13741 ACAAGTAACA TTCAACGGGG GGTAATCAAAC ATCATATCTCA CCCTCTTCTT CTTACACTA

13801 GCCCTTATAA TCCCCCTATT CATCACCTAA ACCGCTCGCA ACGCCCCCTCG ACTCATACCC
 13861 CGAGTTAACT CAAGGACTAC AAATAAAGTT AAGAGCAATA CCCATGCACT AGTAATTAAAC
 13921 AACCAACCCC CAGAAGAGTA TATTAAGGCT ACCCCTCCAa TATCACCTCG AAACACGGAA
 13981 AACTCCCCTA ACTCATCTGC TAGAACCCAA GAAGATTCAT ACCACCCCCC TCAAAACAGC
 14041 CCGGAAACTA ATAATACTCC TCCTGTGTAG ACAACTCCGT ATACCATTAC CGACCAACTA
 14101 CCCCAGCTCT CTGGATACGG CTCAGCAGCT AACGCTGCTG AATATGCAA TACAACAAAC
 14161 ATCCCGCTA AGTAATTAA GAACAAGACT AGTGACAAAA AGGGCCCTCC ATGCCGACT
 14221 ATATACCAAC ACCCCATGCC AGCTACCACA ACCAAACCCA AAGCAGCAA ATAAGGGGAC
 14281 GGGTTAGAAG CAACTGCAAC TAACCCCAAC ACAAGAGAAA ATAAAACATAA ATATATAGCG
 14341 AAAGTCATAA TTCCGTCCAG GACTTTAACC AGGACTAATG GCTGAAAAAA CCACCGTTGT
 14401 TATTCAACTA CAAGAACCT AATGGCCAGT CTACGTAAT CCCACCCCTCT TCTAAAATC
 14461 GCAAACGATG CTTTAGTCGA CCTCCCCGCC CCCTCTAATA TCTCGGTTTG ATGGAACCTT
 14521 GGGTCTCTT TAGGACTATG TTTAATTACC CAAATTGCGA CCGGCTTATT TCTAGCCATA
 14581 CACTACACAT CAGACATTGC TACTGCCCTC ACCCTCGTGG CCCACATCTG TCGAGACGTC
 14641 AACTACGGT GACTTATCCG AAGCATTCA GCAACCGGG CATCATTCTT TTTCATTGTC
 14701 CTCTACCTTC ATATTGGCG AGGACTATAC TATGGCTCTT ACCTCTATAA AGAAACATGA
 14761 ACTGTTGGGG TTATCTCTT CTTCTCGTA ATATAAACAG CCTCGTGG ATACGTCTC
 14821 CCTTGAGGAC AAATGTCGG TTGAGGTGCA ACCGTTATTA CCAACCTTCT ATCTGCCGTC
 14881 CCCTACGTG GAAACACCC GGTCAATGG ATCTGAGGGG GTTTTCTGT AGACAATGCT
 14941 ACCCTCACCC GGTTCTTGC CTTCACTT CTCTCCCTCT TCATCATCGC TGCGCAACA
 15001 TTAATTCACTC TACTCTTCT TCACGAGACT GGCTCAAACA ACCCCACCGG ACTAAACCTCA
 15061 GACTCCGACA AAGTCCCTT TCACCCCTAC TTCACTTACA AAGACCTCTT GGGCTTTGCA
 15121 GTACTCTTA CTGCATTAGC ATCCCTAGC CTATTTCCC CAAATCTCTT AGGAGACCCA
 15181 GATAACTTTA CCCCTGCAA TCCGCTGTG ACTCCCCCAC ACATTAAGCC AGAGTGATAC
 15241 TTCTGTGTTG CCTATGCTAT TCTACGCTC ATTCAAACAA AACTTGGAGG AGTACTTGCC
 15301 CTTTTATTCT CCATCTTAGT TCTTATGCTC GTCCCTTTC TTCATACCTC TAAACAAACGA
 15361 AGCCTAATAT TCCGCCCCCT AACACAATT TTGTTTGGT CTTAGTAGC AGACGTTATG
 15421 ATTCTAACTT GAATTGGAGG GATGCCGTA GAGCACCCCT TCGTCATCAT TGGACAAGTA
 15481 GCGCTCTCTC TTACTCTCTC CCTCTCTCTA GTCTGATCC CAACAGCAGG TTGACTAGAG
 15541 AZATAAGTCT TTGGATGAAA ATGCACTAGT AGCTCAGCAC ACAGAGCCCC AGTCTTGAA
 15601 ACTGGCCGTC GGAGGTTAGA ATCCTCCCTA CTGCTCAAAG AAAGGAGATT TCAACTCTA
 15661 CCCCTAACTC CCAAAGCTAG GATTCTAGCA CTAAACTATT CTTGTGGTA CATAAATGTA
 15721 CAATGAAGGA TTTTCATATA CATGTATGTA ATAACACCAT ATATTTATAG TAACCATT
 15781 ATGTGATGTA CTAGGACATA TATGTATAAT CACCAAATCT CGTAATAAAG CACTCATCA
 15841 TCAACATTTT TAACTAAGAT AGACTAGAAC CTGAATAACC ACTAATCTTA AATATGTGAA
 15901 AGTCCAGGAC CAGTCGAAT CTAAGACCGA ACACAACACT CATCAGTCGA GTTATACCAA
 15961 GACTCAAAT CTCGCCGACC ACAAAATCCT ATGTAGTAAG AGCCTACCAA CGGGTGATT
 16021 CTTAATGATA ACTCTTATTG AGGGTGGAGG ACACAAATTG TgGGGGTTTC ACTCGGTGAA
 16081 TTATTCTGG CATTGGCTC CTACTCAGG GCCATGATT GATGTTATCC CTCACACTT
 16141 CATCGACACT TACATAAGTT AATGTTGATA ATACATACGA CTCGTTACCC ACCAACCGG
 16201 GCGTCACTC CAGCGGGTAA GGGGTTCTCT TTTTTTTT CTTTCACTT GACATTTCAG
 16261 AGTGCATCCA GCCAACAGAG ACGTTAAAG GGTGAGCACT TTTCTTGAC GCGCGTACA
 16321 TAGTATTCTG GTAATGAGTC TTTATTAGAA GAATAACATT AAGTGTATC ATGTGCATAA
 16381 AGATTTGCTC ATTTCTCTA TACCCACTCT AGATACCCCC TTTTTGCGC GCAAACCCCC
 16441 CCTACCCCCC TAAACCCCTG AAGTTGCTAA GACCCCTGAA AACCCCCCGG AAACAGGACA
 16501 AACCTCTGGG AGCTGAGAAA GGGCAATTAC TTTCAACCA CTTTTCAACC ACTTTCAAC
 16561 CACTTTCAA CCCACAAAAT CCCcTGGCTT ATCGtAACCC CaCCaGTTGT TTTAATAATA
 16621 CCACTTTCA ACCCACAAT ACCCcTGACT TATCGTAACC CCACCAAGTT TTTAATAAT
 16681 ACCACTTTCA AACCCACAAT TACCCGTAC TTATCGTAAC CCCACCAGTT GTTTTAATAA
 16741 TACCACTTTT CAACCCACAAT ACCCCCTGA CTTATCGTAAC CCCCACCAGT TGTTTTAATA
 16801 ATACCACTTT TCAACCCACA AATACCCCTG ACTATCGTA ACCCCACCAG TTGTTTTAAT
 16861 ATACCACTT TTCAACCCAC AAATACCCCT GACTTATCGT AACCCCCACCA GTTGTGTTAA
 16921 TAATACCACT TTTCAACCC CAAATACCC TGACTTATCG TAACCCCCAC AGTTGTGTTA
 16981 ATAATACCACT TTTCAACCC ACAAAATCCC CTGACTTATC GTAACCCCCAC CAGTTGTTT
 17041 ATAATACCA CTTTCAACCC CACAAATACC CCTGACTTAT CGTAACCCCCAC CAGTTGTTT
 17101 TAATAATACC ACTTTCAAC CCACAAATAC CCCTGGCTTA TCGTAACCCCC ACCAGTTGTT
 17161 TTAATAATAC CACTTTCAA CCCACAAAATA CCCCTGGCTT ATCGTAACCC CACCAGTTGT
 17221 TTTAATAATA CCACCTTTCA ACCCACAAT ACCCCCTGGCT TATCGTAACC CCACCAAGTT
 17281 TTTAATAATA ACCACCTTTCA ACCCACAAT TACCCCTGGC TTATCGTAAC CCCACCAGTT
 17341 GTTTAATAAA TACCACTTT CAACCCACAAT ATACCTAAAC CCCACAAAGC CCATATCACC
 17401 TTCTTCATGA TCAAAAGACT ACTGTTTAGA TTATTTCAAG TATTCTTCTC ACCACCCCAA
 17461 CTACGGAAGG CATAAAACAG AATCCAAGAA AACTGCTCCA TCCAAAACATA TCTTTTAA
 17521 CGATAAGCC CTTAAGTCCC CCTTAGAAC AGCCTTTAC CGCCCCAAAG GCTAAACACA
 17581 AACAAAGTAC CACCTCTAGG TGTGTTAGA AAATTTACT

Hippoglossus hippoglossus 3 17973 bp

1 GCTAACGTAG CTTAATTAA GCATAACACT GAAGATGTTA AGATGAGCCC TAGAAAGCCC
 61 CGGGAGCACA AAGGCTGGT CCTGACTTTA CTGTCGACTT TAACCTAAACT TACACATGCA
 121 AGTATCCGCC CCCCTGTGAG AATGCCACA ACTCCCTGCT TGGGAACTAG GAGCCGGTAT
 181 CAGGCACAAG CCCAGCTAGC CCACGACGCC TTGCTTAGCC ACACCCCTAA GGGAACTCAG
 241 CAGTGATAAA TATTAAGCCA TAAGTAAAAA CTTGACTTAG TTAAGGTTAA GAGGGCCGGT
 301 AAAACTCGTG CCAGGCCACCG CGGTTATACG AGAGGCCAG GTTGACAAAC AACGGCGTAA
 361 AGAGTGGTTA GGGGATTAC TAAACTAGAG CCGAACGCTT TCAAAGCTGT TATACGCAAC
 421 CGAAAGTATG AAACCCAAATT ACGAAAGTAG CTCTACTCAT CCGAACCCCA CGAAAGCTAA
 481 GAAACAAACT GGGATTAGAT ACCCCCATAT GCTTAGCCCT AAACATCGAT TGCACATATAC
 541 ACTCCATATC CGCCCGGAA TTATGAACAT CAGTTAAAAA CCGAAAGGAC TTGGCGGTG
 601 TTAACATCCA CCTAGAGGAG CCTGTTCTAG AACCGATAAC CCCCGTTAAA CCTCACCTTC
 661 TCTTGTGTTA TCCGCCTATA TACCAACGTC GTCAGCTTAC CCTGTGAAGG CTTTACAGTA

721 AGCAAAATTG GCACAGCCC AAACGTCAGG TCGAGGTGA GTGAATGAGG GGGGAAGAAA
 781 TGGCTACAT TTGCTAAACA TAGCAAACAC GAATGTTGA TTGAAACATG TAACTGAAGG
 841 AGGATTAGC AGTAAGCAGG AAATAGAGCG TCCCGCTGAA ACTGGCCCTA AAGCGCGCAC
 901 ACACGGCCCG TCACCCCTCC CAAGCCCCCT AAACCTAAACT ATTAAAACC CTACAACCG
 961 CGAAGGGGAG GAAAGTCGA ACATGGTAAG TGTACCGGAA GGTGCGCTTG GAAAATCAG
 1021 AGTGTAGCTA AGATAGTATA CAGCATCTCA CTTACACCAGA GAAGACGTCC GTGCAAGTCG
 1081 GACCACCTG ACGCTTATA GCTAGGCCAA CCCCTTAACA CAACAGACCC CCATTTATAA
 1141 CCCCTAAAGC ACGAACACC CACGCTAGCA AACCATTCCTC CCCCCTAAAGT CCAGGCAGATA
 1201 AAAAAGGAAA TTTGGAGCA ATAGAAAAAG TACCGAAGG GAAAGCTGAA AGAGAGATGA
 1261 AAAAGCCAG TAAAGCTTA GAAAGCAGG CTTAGTGCT GTACCTTTG CATCATGATT
 1321 TAGCTAGCAC TTTCAAGCAA AGAGAACCTA CAGTTGTAA CCCCCTAAAGT GAGTGAGCTA
 1381 CTCCAAGACA GCCTATTAT AGGGCGAAC CGTCTCTGTG CAAAAGAGT GGGAAAGAGCT
 1441 TTGAGTAGAG GTGACAAACC TACCGAACTC AGTTATAGCT GGTTGCCTGT GAATTGAATA
 1501 GAAGTTCAGC CCCCTGGTT CTCCACTCAT GCACCTTGT ACCCCTTCAG ATGCAATGAG
 1561 AAACCCAGGG TGTTAGTCAA AGGGGTACA GCCCTTGTG TACAAGACAC AACTTTCTCA
 1621 CGAGGATAAA GATCATATTC AAATAAGGAC AAATGTTAA GTGGCCCTAA AAGCAGCCAC
 1681 CTAATAGAA AGCGTTAAAG CTCAAACATA AAGCCCTCCC GGTTACCGA TAACCCCTATC
 1741 TTATCCCCC ACATTTAAC GGGCTCTTA TGACCCACAT AGGAACGACT ATGCTAATAT
 1801 GAGTAATAAG AGGGTATAAT CACCCCTCTC TTGCACATGT GTAAATCGGA ACGGACCCCC
 1861 CACCGAATCT TAACGGCCCC AATCAAAGAG GGTATTGGAA ACTACCACAA ATTTAGGCCA
 1921 GAAAACATC CAATGCAAAC CCGTTAACCC CACACTGGTG TGCCCAAAG GAAAGACCAA
 1981 AGGGGGGAGA AGGAACCTAG CAAACATACC CCAAGCCTCG CTCGTTTACCC AAAAACATCG
 2041 CCTCTTGCA AACCACAGTA TAAGAGGTG CGCCTGCCA GTGACCATAT AGTTCAACGG
 2101 CCGCGGTATT TTGACCGTGC AAAGGTAGCG TAATCACTTG TCTTTAAAT GAAGACCTGT
 2161 ATGAATGGCA TAACGAGGGC TTAACGTCT CTTCCCCCG GTCAATGAAA TTGATCTCCC
 2221 CGTCAGAAG CGGGGATTAA ACCATAAGAC GAGAAGACCC TATGGAGCTT TAGACACACA
 2281 GGTGGGCAT GTCAAATACC CCTAGCTAAG GGCCTGAAC AAATGGAATC TGCCTTGATG
 2341 TCTTCGGTTG GGGCGACCAT GGGGAATACA AAACCCCCAC GTGGAAAGGG AGCACACCCC
 2401 TAAGTTACTT CTTCTCCCGC AAGCCAGAGC AACAGCTCTA ACCAGCAGAA ATTCTGACCA
 2461 AACTGATCCG GTAAACCGA TCAACGAAG AAGTTACCTT AGGGATAACA GCGCAATCCCC
 2521 CTTTTAGAGC CCATATCGAC AAGGGGTTT ACGACCTCGA TTGTTGGATCA GGACATCCTA
 2581 ATGGTGCAGC CGCTATTAAAG GGTCGGTTG TTCAACGAT AAAGTCCTAC GTGATCTGAG
 2641 TTCAGACCGG AGTAATCCAG GTCACTTTCT ATCTATGACA TGATCTTTT TAGTACGAAA
 2701 GGACGAAAAA GAGGAGGCC CTCGCTAAAG TACGCCCTAC CCCCACCTAA TGAAAAAATC
 2761 TAAACTAGGC AAAAGGCAT AACCCCTTTG CTGGAGATAA CAGCAAGTTT GGGTGGCAGA
 2821 GCCCCGCTAA TGCAAAAGAC CTAAGCCCTT TCTACAGAGG TTCAAATCCT CTCCTTAAC
 2881 ATGATTTCAA CTCTCATTAC TCATATTAC AACCCACTAA CCTTTATTG TGCTGTATTG
 2941 CTAGCGTAG CATTCTCAC CCTCCCTGAA CGAAAAGTGC TAGGCTACAT ACAACTACGA
 3001 AAAGGGCAA ACATTGTTGC GCCTTACGGC CTCTTCACT TATTGCTGA TGGCGTAAAA
 3061 CTCTTATTAA AAGAACCCAT TCGACCTCA ACCGCATCAC CCGCTCTATT CTCCTGAGCC
 3121 CCTATGCTCG CTCTCACACT AGCCCTAAC CTTGAGGCC CCATACCTT CCGTACCC
 3181 GTCGTGGACC TTAACCTAGG TATTTTATT ATTCTAGCTC TATCTAGCCT TGCAGTATAT
 3241 TCTATTCTGG GCTCCGGCTG AGCATCTAAT TCAAATATG CCCTAGTAGG GGCACGCG
 3301 GCTGTCGAC AAACAATTTC TTACGAGTT AGCCTTGGAC TCATCTTACT AAACATTATC
 3361 ATCTTCACCG GGGGCTTCAC ACTCCAAACCTT TTAAACACAG CCAAGAGGG CATTGACTG
 3421 GTAGTCCCG CTTGACCGCT TGCTGCCATA TGATATATCT CCACTCTTGC CGAAACAAAC
 3481 CGTGCACCCCT TCGACTTAAC AGAAGGAGAA TCTGAACCTTG TCTCGGGCTT CAACGTAGAG
 3541 TACGCAGGGG GACCCCTCGC CTTGTTCTTC TTAGCCGAAT ATTCAAATAT TCTCCTAATA
 3601 AATACCCCTGT CCGCTACACT ATTTTAGGG GCCTCCCACA TCCCATCTAT CCCGAATTA
 3661 ACTAGTATCA ATATTATGAC TAAAGCAGCC CTTCTATCAA TTGTCCTCCGT GTGGGTCCGA
 3721 GCTTCGTACC CGCGATTCCG TTACGACCAA CTTATGACCC TCATTTGAAA AAACTTCTT
 3781 CCACTGACTT TGGCACTAGT TATCTGACAC TTGGCGCTCC TATTGCTC CGCTGGTCTC
 3841 CCACACAAAC TATAGGCCA GGAGCTGTC CTGAATTAAAGGGCAACTT TGATAGAGTG
 3901 AATTATGGGG GTTAGAGTCC CCCCCACTTC TTAGAAAGAA GGGGATCGAA CCCAACCTGA
 3961 AGAGATCAAAC ACTCTTCGTG CTTCCCTCTAC ACCACTTCCT AGTAAGGTCA GCTAAATAAG
 4021 CTCTGGGCC CATACCCAA CCATGTAGGT TAAATCCTT CTTTGCTAA TGAACCCCTA
 4081 CATCTTGACCC ACTCTTCTAT TTGGTTGGG CCTTGGAAACA ACACTCACGT TTGCAAGCTC
 4141 ACACTGGCTC CTCGCTTGA TAGGACTTGA GATAATACA CTAGCCATTAA TTCCCTTAAT
 4201 AGCCACAC CACCAACCCCCC GGGCAGTCGA AGCTACTAAC AAATACTTTC TCGCACAAGC
 4261 CACAGCAGCC GCCACCCCTC TTGTTGCAAG CACTACCAAC GCTTGACTCA CAGGCCAGTG
 4321 GGATATTCAA CAGATAACAC ATCCCCCTCC CACAACAATG ATTGTTATTG CCCTCGCGCT
 4381 AAAGGTCGGA CTGGCACCAA TGCACCTTG ACTCCCCGAA GTACTCCAAG GGCTAGACCT
 4441 GACCACGGG CTCATCCTCT CAACCTGACA AAAACTTGCA CCCTTGCCC TCTTACTACA
 4501 AATTCAAACCA ACCAACCCCTA CGCCTCTAAAT CATTATTGGC TTACTATCCA CTCTCGTTGG
 4561 TGGCTGAGGT GGTCTCAACC AAACCCAGCT GCGCAAGGTC CTCGCCTACT CCTCAATTG
 4621 CCACCTTGGC TGAATAATGC TCATCTTCA ATTTCACCC CTCCCTGGCC TTCTCACCC
 4681 CCTCACGTAC TTACTATGAC CTTCTCAGC ATTCTTAAGTAA ACAAAAGCGAC
 4741 CACCGTCAAT GCCCTCGCA TCTCGTGAAC AAAACTCCT GCGCTCACAG CCCTAGCACC
 4801 CCTGGTTTA CTTTCTCTCG GCGGCCCTCC TCCCTTACCG GGCTTCATGC CCAAGTGTATT
 4861 CATTCTTCAG GAACTGACTA AACAAAGACCT TCCGGCACTT GCCACCCCTCG CCGCTTAAAC
 4921 TGCCCTTGT AGTCTACT TCTACTGCG CCTCTCGTAT GCAATAACCC TCACAATATT
 4981 CCCAACAAAC CTAATTGGTG TCAACTCCCTG ACGATTCTCA TCCCCCCAAAT TTACACTTCC
 5041 CCTCGCCGTC TCCACTGCAAG CAACCCACCT TCTCTCCCA CTAGCCCCCG CTGCGTAGC
 5101 ACTACTGGTC ACCTAAGGGG CTTAGGCTAG CACTTAGACCC AATGGCCTTC AAAGCCGTAA
 5161 GCGTGAGTGA AAATCTCTCA GTCCCTGTTA AGACTTGCGG GCTACTATCC CACATCTTCT
 5221 GCATGCAAAG CAAACACTT AATTAAGCTA AAGCCTTTCT AGATAGGAAG GCCTTGATCC
 5281 TACAAAATCT TAGTTAACAG CTAAGTCCCC AATCCAGCGA GCATCTATCT ACTTTCCCC
 5341 GCCTAATTAA GAAACTAAAAA GGCGGGGAAA AGCCCCGGCA GGCAGTTAGC CTGCTTCTT
 5401 AGATTTGCAA TCTAACATGT AACACCCAG GGCTTGGTAA GAAGAGGGCT TGCACCTCTG

5461 TCTATGGGTC TACAATCCAC CGCTTAACTC AGCCATCCTA CCTGTGGCAA TCACACGTTG
 5521 ATTTTCTCG ACCAATCACA AAGACATCGG CACCCCTCAT CTCGTATTG GTGCCTGAGC
 5581 CGGAATAGTG GGGACAGGCC TAAGTCTGCT TATTGGGCA GAACTAAGCC AATCCGGGC
 5641 TCTCTGGGA GACGACCAAA TTTATAATGT GATCGTCACC GCACACGCC TTGTAATAAT
 5701 CTTTTTATA GTAATACCA TTATGATTGG GGGGTCGGA AACTGGCTTA TTCCACTAAT
 5761 AAtGGGGCC CCAGACATGG CGTCCCTCG AATGAATAAT ATGAGTTTCT GACTTCTCC
 5821 CCCCTCCTT CTCCCTCT TAGCCCTTC AGGTGTTGAA GCCGGAGCAG GTACCGGATG
 5881 AACCGTGTAC CCCCACTAG CTGGCAATT AGCCCACGCC GGGGCATCCG TAGACCTGAC
 5941 ATCTCTCA CTTCACCTG CAGGAATTCTG CTCAATTCTG GGGGCAATTAA ACTTTATTAC
 6001 TACCATCATT AACATGAAAC CCACAAGT CACTATGTAC CAAATCCCGT TATTGTTG
 6061 AGCCGTTCTT ATTACAGCCG TACTTCTCT TCTGTCCCTG CCCGTTTTAG CGCAGGGAT
 6121 TACAATGCTA CTAACAGACC GCAACCTAA CACGACCTTC TTTGACCTCT CGGGAGGAGG
 6181 TGACCCATT CTCTACCAAC ACCTATTCTG ATTCTTTGGT CACCCAGAGG TATACATTCT
 6241 TATCTCTCCA GGCTTCGGA TAATTCTCA CATTGTTGCA TACTATGCAG GTAAGAAAGA
 6301 ACCTTTGGC TACATGGGA TAGTCTGAGC TATAATGGG ATTGGACTCC TGGGCTTCAT
 6361 TGTCTGGGGC CATCACATAT TTACAGTCGG AATAGACGTA GATACACGAG CCTACTTAC
 6421 CCTCTGGCAC AATACTATTG CAATTCCAAAC TGGCTAAAGA GCTTTAGCT GACTCGAAC
 6481 CCTCCATGGG GGAAGCATTA AATGAGAAC GCCCCTTCTA TGAGCCCTCG GCTTTATT
 6541 CCTCTTACA GTAGGCGTC TCACTGGCAT TGTCTTAGCT AACTCCCTC TCGATATTGT
 6601 TCTGCATGAC ACATACTATG TAGTCGCCA CTTCACATAT GTACTATCTA TGGGTGCTGT
 6661 ATTTGAATC GTTGCCTGC TCGTCCATTG ATTCCGTTA TTTACAGGCT ATACCCTTCA
 6721 CCTCACATGA ACAAAAATCC ACTTCGGCT GATGTTTATT GGGGTCAAATC TAACATTCTT
 6781 CCTCTCAACAT TTTCTGGGC TGGCTGGAT ACCCCGACGG TACTCAGACT ACCCAGACGC
 6841 ATACACCTTCA TGAAACACTG TTTCATCAAT TGGGTCCTA ATGTCCTCTG TTGCTGTAAT
 6901 TTTTATTCTA TTCATTATTG GAGAACATT TACAGCCAA CGAGAAGTCG GAGCAGTAGA
 6961 ACTAACTGCA ACTAACATTG AATGACTTTA CGGCTGCCCT CCCCCCTACC ACACATTGA
 7021 AGAGCCGCA TTCGTACAAG TTCGTATAAA TTCGAACAAA CTAACGAGAA AGGGAGGAGT
 7081 TGAACCCCCA TCAATTGGTT TcAAGGCAAC CACATAACCG CTCTGTCACT TTCTTCACAA
 7141 CACACCAATA AGATACTAGT AAAACGCTAT AACACTGCC TGTCAAGGCA GAATTGTTGG
 7201 TCAACCCCCC GCGTATCTA AACGCAATGG CACATCCCTC ACAACTCGGA TTTCAAGACG
 7261 CAGCTTCACT CTTAATAGAA GAACACTTC ACTTCCACGA CACGCCCTTA ATAATTGTA
 7321 TTCTCATTAG CACAATGGT CTTTATTTA TTGCGGCTAT AGTCACAGCC AAACAAACAG
 7381 ATAAACTTGT ACTAGACTCT CAAGAAATTG AAATTATCTG AACAGTTCTC CCAGCTGTTA
 7441 TTCTATCCT TATCGCCCTA CCATCCATCC GAATTGGTA CTTAATAGAC GAAATTAAATG
 7501 ACCCCCACCT TACAATCAA GCCATCGGCC ACCAATGATA CTGAAGCTAC GAATACACAG
 7561 ACTACGAAGA CCTCGGTTT GACTCATACA TAATCCCCAC ACAAGACCTG ACCCCCAGAC
 7621 ATTCCGACT ACTAGAACGA GACCACCGA TAGTAATCCC AGTAGAAATCT CCAATTGCG
 7681 TTTAATCTC GGCTGAAGAT GTCTTACATT CCTGGCAGT CCCCCTCTA GGCCTAAAG
 7741 TTGACCCCGT ACCTGGACCA TTGAACCAAG CAACCTTTAT TGTAGCCGA CCAGCGTAT
 7801 TTTTCGGACA ATGTTCTGAA ATTTGGGGG CTAACCATAG TTTTATACCC ATTGTTGTTAG
 7861 AAGCAGTCCC CCTTGAACAC TTTGAGAACT GGTCTTCACT ATAATTGAA GACGCCCTCG
 7921 TAAGAAGCTA ATAAGTACAA GCGTTAGCT TTTAAGCTAA AGACTGGTGC CTAACAACCA
 7981 CCCCTAGCGA CATGCCCTAA CTGAACCCCG CACCCCTGATT TGCAATTCTT GTTTCTCTT
 8041 GAATAGTCTT TTTAATCTGTT ATTCCCCCA AGGTTTACG ACACACCTAT CCCAATGAAC
 8101 CACACCTCCA AAGCACACAA AAACCTAAAA CAGAGCCCTG AAACATGACCA TGACTAAAG
 8161 CCTCTTGCAC CAATTATAT CCTCCCTCAT CTCGGCCTT CCACTAATTG CACTAGCGAT
 8221 TAGCCTGCCT TGAATTCTGT TCCCAACCC TCAAGCCCG TGATTAACACA ACCGCGTGCT
 8281 AACACTTCAG GGGTGATTG TTAGCCGCTT CACCTCACAA CTTCTTCTAC CCCTAAACCC
 8341 TGGGGCCAC AAGTGGCAG TTCTATTAC CTCGTTAATA GTTTCCCTAT TTTCTATCAA
 8401 CATGCTAGGA CTCCCTCCAT ACACCTTCAC ACCAACAAACC CAACTCTCCC TTAACATGGG
 8461 CCTCCGAGTA CCCCTTGAT TGGCACTGT CATATCGGG ATACGAAATC AACCAACCGA
 8521 TGCCTCTAGGC CACCTCTTC CAGAACGCAC CCCCCACTGCC CTCATCCCTG TACTTATTAT
 8581 CATTGAAACA ATTACGCTT CATTGACCC CTTAGCTCT GGTGTCGAC TAACAGCAA
 8641 TCTCACAGCA GGCCACCTC TTATTCAACT CATTGCAACA GTCGCCCTCG TCCTTCTTCC
 8701 CCTCATACCT GTTGTGCGTA TCTTAACAAAC GGTGGCTCTC TTCTCTCTA CTCTACTAGA
 8761 AGTAGCCGTT GCCATGATTC AAGCTTATGT CTTGTTCTA CTCCCTAAGCC TATACCTACA
 8821 AGAGAACGTT TAATGGCCCA TCAAGCACAC CCATACCACA TAGTCGACCC CAGCCCATGA
 8881 CCCCTAACAG GGGTATTG TGCCCTATTG ATAACATCTG GCCTGCTAT CTGATTCCAC
 8941 TTCCACTCCA CAACCTTAAT AACTATGGG ACAATCCTTC TCACCTTAAC AATCTTCAA
 9001 TGATGACGGG ACGTCGTAC AGAAGGCACA TTCAAGGAC ACCACACTCC CCCCCTTCAA
 9061 AAAGGCTCC GATATGGGAT AATTTTATT ATTACCTCAG AGTCCTATT CTTCTTAGGC
 9121 TTCTCTGAG CCTTTTACCA CTCAAGCTA GCACCTACCC CTGAGCTAGG CGGCTTCTG
 9181 CCACCAAGCAG GCATTACTCC CTTAGACCA TTTGAAGTCC CACTCTTAA CACAGCAGTC
 9241 CTGCTGCCT CTGGCGTAAC TGTTACCTGA GCACATCACA GCATTATGGA GGGTAAACGA
 9301 AAACAAACTA TTCAATCTCT TGCTCTTACA ATCTTACTCG GGGGGTACTT CACTTCTCTT
 9361 CAAGGCCTAG AGTATCAGCA AGCCCCCTT ACCATTGCGAC AGGGGGTTA CGGCGCTGCA
 9421 TTCTTGTGTC CCACGGCTT CACCGGATTA CACGCTTAAAT TTGGCACATC ATTCTTAGCC
 9481 GTTTGTCTAT TACGCCAGAT TCTTCACCAT TTCAACATCAA ACCACCACTT TGGGTTGAA
 9541 GCAGCCGCAT GATACTGACA CTTCTGAGAC GTCGCTGAC TTTCTCTA TATCTCTATT
 9601 TACTGATGAG GATCTTAATC TTCCTAGTAT TAAATCTAGT ATAAGTGACT TCCAATCACC
 9661 CGGTCTGGT TAAAATCCAA GGGAAAGATAA TGAGCCTCT CTTAACCATC ATTCGATTG
 9721 CCACCCCTCCT CTCGACAGTA CTAGCCCTG TATCCCTTTG ATTACCCCCAA ATTACACCAAG
 9781 ACCACGAAAA GCTATCGCGC TACGAGTCGG GCTTGACCC AATAGGATCC GCCCGGCTGC
 9841 CTTTTCTACT GCGGTTTTC CTCATGCCA TCCTCTTCT TCTCTTCGAT TTAGAAATTG
 9901 CTCTCTCTC CCTCTCTCA TGAGGAGACC AACTAGCATC GCCACTACTG ACGTTCACCT
 9961 GAGCTACAGC TGTCTAGCC CTTCTAACCC TTGGCCTCAT TTACGAATGA ATGCAAGGAG
 10021 GTCTAGAATG GGCTGAATAG GTGGTTAGTC TAAGAAAAAC ATTTGATTG GGCTCAAAAAA
 10081 CTTGTGGTTT AAATCCGAA CGCTTAATG ACCCCCCACAC ACTTTGCCTT CTCCCTAGCC
 10141 TTTCTCTGG GTTTAACAGG CCTGGCATTC CACCGGTTTC ACCTCCTTTC CGCCCTATTG

10201 TGCCTTGAAG GAATAATAC ATCCCTATTT ATTGCTCTCT CCCTATGAAC GCTCCAACCT
 10261 GACTCAACCA ACTTTTCAGC ATCTCCAATA CTTCTGCTTG CATTTCAGC CTGCGAAGCA
 10321 AGCGCCGGCC TCGCCCTATT AGTAGCCACT GCTCGAACCC ACGGGACCGA CCGATTACAA
 10381 AGCTAAATC TTCTCCAATG CTAAAAGTCC TAATCCAAC ACTTATGCTA ATCCCAACAG
 10441 CCTGGTTACT CAAACCAGC TGACTCTGGC CCATGACTTT AATGTATAGC TTTTGATCT
 10501 CCTTGGTTAG CCTTTCATGA CTAAAGAACC TCTCAGAAC CGGCTGATCG TCACCTAGCC
 10561 TCTTATAGC TACCGACTCC CTATCAACCC CCCCTCTCT TCTTACATGC TGATTACTCC
 10621 CACTAATAAT TTTGGCAAGC CAAAAACACA CAGCCTTACA ACCTCTTAGT CGCCAGCGCA
 10681 TGTACATCAC ACTTCTGCC TCACTCCAGT TTTCTTAAT CCTAGCATT AGCGCCACCG
 10741 AACTGGTAAT GTTTCAGTA ATATTGAGA CCACCCCTCAT CCCCACGCTA ATCATTATCA
 10801 CCCCTGAGG AAACCAACAA GAACGCTAA ACGCAGGAAC CTACTTTCTC TTCTATACAT
 10861 TAGCGGGCTC ACTCCCCTT CTCGTTGCTT TACTCCTACT CAAAACACA TCCGGCACCC
 10921 TGTCGCTGTT AACCCCTCAC TATACGGACC CGCTCGCTCT CTCGTCTTAT GCAGACAAAC
 10981 TATGATGAGC AGGCTGCTT TTAGCATTCC TGGTTAAAT ACCCCTCTAC GGGGTCCACC
 11041 TATGCTCCC TAAAGCCAT GTTGAAGCCC CAATTGCAAGG CTCAATAATC CTTGCAGCGG
 11101 TTCTACTGAA GCTAGGGGA TACGGCATGA TCCGCTATAAT ACAAAATATTG GAACCTCTAA
 11161 CTAAGGAGTT AAGCTACCCC TTCATTGTCT TTGCACTTTG AGGGGTAATT ATAACCTGGG
 11221 CAATTGCTC ACGCCAGACA GACCTTAAGT CCCTAATTGC CTACTCCTCA GTCAGCCACA
 11281 TAGGCCTAGT GGCTGGAGGA ATTCTTATCC AATCACCCCTG GGGCTTAACA GGGGCTCTA
 11341 CCCTCATATA TGCACACGGG CTTACTTCAT CAGCCCTCTT CTGCTTAGCA AACACAAACT
 11401 ATGAACGAAC CCACAGTCGA ACGATAGTCC TGGCGCGAGG ACTTCAAGTA GCCCTACCC
 11461 TAATAGCCAC TTGATGATT ATTCTACTGTT GGCCAAACCT AGCTCTGCCA CCACTACCC
 11521 ACCTCATGGG GGAGCTGATA ATTATCACTT CTTATTCAA CTGGTCTGTA TGAACCTCTGG
 11581 CATTAACCGG TGCAGGCACT CTAATCACAG CTGGTTACTC CTTTACATA TTCTTAATAA
 11641 CTCAGGGGG CCCTCTCCA ACACATGTC TTGCACTTG GCCATCTCAC ACCCGAGAAC
 11701 ACCTCTTAT TGCACTTCAC CTCATCCCCT TAATTCTACT CGTACTTAAG CCCGAGCTGA
 11761 TCTGAGGCTG AACTGCCGT AGGTGTAGTT TAAACAAAAA CATTAGATCG TGATTCTAA
 11821 ATAAGGGTT AAAATCCCTT CTCCCACCGA GAGAGGCTCG CAGCAATGAG AACTGCTAAT
 11881 TCCCACGACC TTGGTTGGAC CCCTAGGCTC ACTCGAAAGC TCCTAAAGGA TAACAGCTCA
 11941 TCCGTTGGTC TTAGGAACCA AAAACTTTG GTGCAAATC AAGTAGCAGC TATGTACCCC
 12001 ACCTCCCTAA TGATTCATC GAGCCTGATC ACAATCTTTG CGTTACTAGC CTATCCTTTA
 12061 ATCACACGA TCCGCCCCAT ACCTCCGGAC CCTCAATGGG CTACCACCCA TGTAAAACA
 12121 GCTGTAAAAA TAGCTTCTT TGTTAGCTG CTGCTCTGA CCCTATTTCT TAATGAGGGC
 12181 GCAGAAACAA TCGTTACTAA CTGAACCTGA ATAAATACCA ACTCCTTCAA TGTCAGCATT
 12241 AGCTGAAAT TTGATTCTA CTCAATTATT TTTACACCCA TTGCCCTCTA CGTAACCTGA
 12301 TCCATCCTGG AGTCGCATC ATGGTATATG CACCGACAGC CACACATGAA CCGCTTCTTT
 12361 AAATACCTCC TGACGTTCT GATTGCCATA ATCGCTCTAG TAACTGCAA CAATATGTTT
 12421 CAACTATTAA TCGGTTGAGA GGGCGTGGGG ATCATATCGT TCTTACTTAT TGGATGGTGG
 12481 TTCGGACGGG CGGATGCTA CACTGCAGCG CCTCAAGCAG TACTTTACAA CCGAGTAGGG
 12541 GACATCGGGC TTATCTCGC CATGGCTIGA ATAGCAGCAA ACCTAAACTC CTGGGAGATG
 12601 CAACAAATCT TTGCAACCAC CAAAGACATA GACTTAACCT ACCCTCTTCT CGGACTCATC
 12661 ATTGCAGCAA CTGGCAAAATC GGCTCAATTG GGACTTCACC TTGACTTACCC CTCCGCTATG
 12721 GAGGGTCCCA CACCGCTCTC TGCCCTACTT CATTCTAGCA CTATGGTCGT TGCCGGCATT
 12781 TTCTGCTAG TACGAAATAAG CCCGCTACTG GAAAACAAAC AAACCGCCCT CACAACCTGC
 12841 CTATGCCTCG GAGCTCTTAC AACACTATTG ACTGCAACCT CTGCCCCCTC GCAAAATGAT
 12901 ATTAAAAAAA TCGTCGCTT CTCTACATCC AGCCAGCTT GGCTTATAAT GGTAAACCAatt
 12961 GGTCTAAACC AGCCACAACT AGCATTCTT CACATCTGTA caCATGCCTT CTTCAAGGCC
 13021 ATACTATTTT TATGCTCCCG ATCTATTATC CACAGCTAA ATGATGAACA AGATATCCGG
 13081 AAAATAGGGG GCATACACCA CCTTGACACTT TTACATCTT CCTGCCCTAAC TATTGGCAGC
 13141 CTCGCCCTCA CAGGAACCCC CCTCTTGGCT GGCTTCTTCT CTAAAGATGC CATCATCGAA
 13201 CGATTGAACCA CATCCCACCT AAACGCCCTGA GCCCTAACCT TAACCCCTCT AGCCACCTCT
 13261 TTACAGCAA TCTACAGCTT TCGAGTAGT TTTTTGTAC CCATGGGCCA CCCCCGATT
 13321 AGTCCCTCT CTCCGATCAA CGAAAACAAAC CCAGCAGTGA TAAACCCACT TAAGCGACTA
 13381 GCATGGGGGA GTATCGTTGC AGGCTTACTA ATTACCTCTA ATATTACACC CTTGAAAACC
 13441 CCTGTATAT CTATACCTCC CCTCCTTAACT CTAGCTGCTC TTGCACTGC AATCATCGGC
 13501 CTTCTGTTG CTATAGAGCT CGCCATGTTA ACCAACAAAC AATATGAGCC AATACCAAC
 13561 ACGCACATAC ACCACTTCTC TGCCTTATTA GGCTTCTTCC CGGGGTTAT ACACCGCTTA
 13621 ACCCTGTAT TGCTCTTCTT CCTAGGCCAA AACATTGCTA ACCAACACAGT AGATCAGACA
 13681 TGATTAGAAA AAACGGGCC CAAAGCCTT GAAGGCCGCTA ACCGACCTCT CGTCTCTCT
 13741 ACAAGTAACA TTCAACGGGG GGTAACTAAAC ACATATCTCA CCTCTTCTCT CTTTACACTA
 13801 GCCCTTATAA TCCCCCTATT CATCACCTAA ACCGCTCGA ACGCCCCCTCG ACTCATACCC
 13861 CGAGTTAACT CAAGGACTAC AAATAAAGTT AAAAGCAATA CCCATGCAGT AGTGTAAAC
 13921 ACCACCCCC CAGAAGAGTA TATTAAGCT ACCCCCTCAA TATCACCTCG AAACACGGAA
 13981 AACTCCCTA ACTCATCTGC TAGAACCCAA GAAGATTCTAC ACCACCCCCC TCAAAACAGC
 14041 CCGGAAACTA ATAATACCTCC CCCTGTGTA ACAACTCCGT ATACCATTAC CGACCAACTA
 14101 CCCAGCTCT CTGGGTACGG CTCAGCAGCT AACGCTGCTG AATATGCAA TACAACAAAC
 14161 ATCCCGCTA AGTAATTAA GAACAGACT AGTGACAAA AGGGCCCTCC ATGCCCCGACT
 14221 AATATACCCAC ACCCCATGTC AGCTACCCAA ACCAAACCCA AAGCAGCAA ATAAGGGGAC
 14281 GGAGTTAGAAG CAACTGCAAC TAACCCAAAC ACAAGAGAAA ATAAAACAAAT ATATATAGCG
 14341 AAAGTCATAA TTCTGCCCCAG GACTTTAACCT AGGACTAATG GCTTGAAAAA CCACCGTTGT
 14401 TATTCAACTA CAAGAACCTT AATGGCCAGT CTACGTTAAAT CCCACCCCTCT TCTAAAATC
 14461 GCAACACGATG CTTTACTCGA CCTCCCCGCC CCCTCTAATA CCTCGGTTTG ATGGAACCTT
 14521 GGGTCTCTTT TAGGACTCTG TTTAATTACCA CAAATTGCGA CCGGCTTATT TCTAGCCATA
 14581 CACTACACAT CAGACATTGC TACTGCCCTC ACCTCCGTTG CCCACATCTG TCGAGACGTC
 14641 AACTACGGCT GACTTATCCG AAGCATTCTATGCCAACCGGG CATTCTTCTT TTTCATTGC
 14701 CTCTACCTTC ATATTGCCCC AGGACTATAC TATGGCTCTT ACCTCTATAA AGAAACATGA
 14761 ACTGTTGGGG TTATTCTCTT CCTCTCGTA ATAATAACAG CCTTCGTTGG ATACGTCTC
 14821 CCTTGAGGAC AAATGCGTT TTGAGGTGCA ACCGTTATTA CCAACCTTCT ATCTGCCGTC
 14881 CCCTACGTGCG GAAACACCTT AGTCCAATGA ATCTGAGGGG GTTTTCTGT AGACAATGCT

14941 ACCCTCACCC GGTTCTTTCG CTTCCACTTT CTCTTCCCCC TCATCATCGC TGCCGCAACA
 15001 TTAATTCACTC TACTCTTCTC TCACGAGACT GGCTCAAACA ACCCCCACCGG GCTAAACTCA
 15061 GACTCCGACA AAGTTCCCTT TCACCCCTAC TTCACTTACA AAGACCTCTT GGGCTTTGCA
 15121 GTACTTCTTA CTGATTAGC ATCCCTAGCC CTATTTTCCC CAAATCTCTT AGGAGACCCA
 15181 GATAACTTTA CCCCTGCAAA TCCGCTGTG ACTCCCCCAC ACATTAAGCC AGAGTGATAC
 15241 TTCTGTGTTG CCTATGCTAT TCTACGCTCC ATTCCAAACA AACTTGGAGG AGTACTTGCC
 15301 CTTTTTATTCT CCATCTTAGT TCTTATGCTC GTCCCTTTCC TTCATACCTC TAAACAAACGA
 15361 AGCCTAAATCT TCCGTCCTCT AACACAACTT TTGTTTGTG CTTTAGTAGC AGACGTTATG
 15421 ATTCTAACTT GAATGGAGG GATGCCGTA GAACACCCCT TCCTGATCAT TGGACAAGTA
 15481 GCGCTCTCC TTTACTCTC CCTCTCTTA GTCTGATCC CAACAGCAGG TTGACTAGAG
 15541 AATAAAAGTCC TTGGATGAAA ATGCACTAGT AGCTCAGCAC ACAGAGCCCC AGTCTTGAA
 15601 ACTGGCCGTC GGAGGTTAGA ATCCTCCCTA CTGCTCAAAG AAAGGAGATT TCAACTCCTA
 15661 CCCCTAACTC CCAAAGCTAG GATTCTAGCA CTAAACTATT CTTGTGGTA CATAAATGTA
 15721 CAATGAAGGA TTTTCATATA CATGTATGTA ATAACACCAT ATATTTATAG TAACCATT
 15781 ATGTGATGTA CTAGGACATA CATGTATAAT CACCAAATCT CGTAATACAG CACTCATCTA
 15841 TCAACATTTT TAACTAAGAT AGACTAAAC CTGAATAATC ACTAATCTTA AATATGTGAA
 15901 AGTCAGGAC CAGCTGAATC TAAAGACCGA ACACAAACACT CATCAGTCGA GTTATACCAA
 15961 GACTAAAAT CTCGCCGACC CCAAATCCTT ATGTTAGTAAG AGCCTACCAA CCGGTGATT
 16021 CTTAATGATA ACTCTTATG AGGGTGAGGG ACAAAAATTG TGGGGGTTTC ACTCGGTGAA
 16081 TTATTCCCTGG CATTGGTTC CTACTTCAGG GCCATTGATT GATGTTATCC CTCACACTT
 16141 CATCGACACT TACATAAGTT AATGTTGATA ATACATACGA CTCGTTACCC ACCAACCGG
 16201 GCGTTCACTC CAGCGGGTAA GGGGTTCTCT TTTTTTTTCT CTTTCACTT GACATTTCAG
 16261 AGTCATCCA GCCAACAGAG ACGTTAAAG GGTGAGCACT TTTCTTGAC GCGCCGTACA
 16321 TAGTATTCTG TGAATGAGTC TTATTTAGAA GAATAACATT AAGTGTATC ATGTGCATAA
 16381 AGATTTGCTC ATTTCTCTG TCATCCCGAG GATAACCCCT TTTTGCAGG CAAACCCCC
 16441 CTACCCCCCT AAACCCCTGA AGTTGCTAAG ACCCCCTGAAA ACCCCCCCGGA AACAGGACAA
 16501 ACCTCTGGTA GCTCAgAAAG GGCAATTACT TTCAACCAC TTTCAACCA CTTTTCAACC
 16561 CACAAATACC CCTGGTTAT CgTAACCCCA CCAGTTGTT TAATAATACC ACTTTTCAAC
 16621 CCACAAATAC CCCTGGCTA TCATAACCCC ACCAGTTGTT TTAATAATAC CACTTTCAAC
 16681 CCCACAAATA CCCCTGGCTT ATCATAACCC CACCAAGTTGTT TTTAATAATA CCACTTTCA
 16741 ACCCACAAAT ACCCCCTGGT TATCATAACC CCACCAAGTT TTTAATAAT ACCACTTT
 16801 AACCCACAA TACCCCTGGC TTATCATAAC CCCACCAAGT GTTTTAATA TACCACTTT
 16861 CAACCCACAA ATACCCCTGG CTTATCATAA CCCACCAAGT GTTTTAATA ATACCACTTT
 16921 TCAACCCACA AATACCCCTG GCTTATCATA ACCCCACCAAG TTGTTTTAAT AATACCAC
 16981 TTCAACCCAC AAATACCCCT GGCTTATCAT ACCCCACCA GTTGTGTTAA TAATACCACT
 17041 TTTCAACCCCA CAAATACCC TGGCTTATCA TAACCCACCA AGTTGTGTTAA ATAATACCA
 17101 TTTTCAACCC ACAAAATACCC CTGGCTTATC ATAACCCACCA CAGTTGTGTT AATAATACC
 17161 CTTTCAACCC CACAAATACCC CCTGGCTTATCATAACCCCA CCAGTTGTGTT TAATAATACC
 17221 ACTTTCAAC CCACAAATAC CCCTGGCTT CACAAACCCCA ACCAGTTGTT TTAATAATAC
 17281 CACTTTCAA CCCACAAATA CCCCTGGCTT ATCATAACCC CACCAAGTTGTT TTTAATAATA
 17341 CCACTTTCA ACCCCACAAAT ACCCCCTGGCT TATCATAACC CCACCAAGTTGTT TTTAATAATA
 17401 ACCACTTTTC AACCCACAA TACCCCTGGC TTATCATAAC CCCACCAAGTT GTTTTAATAA
 17461 TACCACTTTT CAACCCACAA ATACCCCTGG CTTATCATAA CCCACCAAGT TGTCTTAATA
 17521 ATACCACTTT TCAACCCACA AACACCCCTG GCTTATCATA ACCCCACCAAG TTGCTTAAT
 17581 ATACCACTT TTCAACCCAC AAATACCCCT GGCTTATCAT ACCCCACCA GTTGTCTTAA
 17641 TAATACCACT TTTCAACCC CAAATACCC CAGTTGTTAA CCCAAAGGAC TTGGCGGTG
 17701 ATAATACCCAC TTTTCAACCC ACAAAATACCT AAAACCCACA AACGCCGTAT CACCTTCTC
 17761 ATGATCAAAA AACTACTGTT TAGATTATTT CAAGTATTCC TTTCACCACCA CTAACATCGA
 17821 AAGGCATAAA ACAGAAATCCA AGAAAATGTC TCCATCCAAA ACTATCTTT TTAACCGATA
 17881 AGCCCTTAAG TCCCCCTGAG AACAGCCTT TTACCGCCCC AAAGGCTAAA CACAAACAAA
 17941 CTACCACCTC TAGGTGTGTT TAGAAAATTT ACT

Hippoglossus hippoglossus 4 17729 bp

1 GCTAACGTAG CTTAATTAAA GCATAACACT GAAGATGTTA AGATGAGCCC TAGAAAGCCC
 61 CGGGAGCACA AAGGCTGGT CCTGACTTTA CTGTCGACTT TAACCTAAACT TACACATGCA
 121 AGTATCCGCC CCCCTGTGAG AATGCCACA ACTCCCTGCT TGGGAACCTAG GAGCCGGTAT
 181 CAGGCACAAG CCCAGCTAGC CCACGACGCC TTGCTTAGCC ACACCCCTCAA GGGAACTCAG
 241 CAGTGATAAA TATTAAGCCA TAAGTAAAAA CTTGACTTAG TTAAGGTTAA GAGGGCCGGT
 301 AAAACTCGTG CCAGGCCACCG CGGTTATACG AGAGGCCAA GTTGACAAAC AACGGCGTAA
 361 AGAGTGGTTA GGGGATTTC TAAACTAGAG CCGAACGCTT CAAAGCTGT TATACGCA
 421 CGAAAGTATG AAACCCAAAT AGCAAGTAGT CTCTACTCAT CCTGAAACCCA CGAAAGCTAA
 481 GAAACAAACT GGGATTAGAT ACCCCCATAT GCTTAGCCCT AACACATCGAT TGCACTATAC
 541 ACTCCATATC CGCCCGGGAA TTATGAACAT CAGTTTAAAC CCCAAAGGAC TTGGCGGTG
 601 TTAACATCCA CCTAGAGGAG CCTGTTCTAG AACCGATAAC CCCCGTTAAA CCTCACCTC
 661 TCTTGTGTTA TCCGCCTATA TACCAACCGTC GTCACTTAC CCTGTGAAGG CTTTACAGTA
 721 AGCAAATTG GCACAGCCA AAACGTCAAG TCGAGGTGTA GTGAATGAGG GGGGAAGAAA
 781 TGGGCTACAT TTGCTAACAC TAGCAACAC GAATGTTGCA TTGAAACATG TAACTGAAGG
 841 AGGATTTAGC AGTAAGCAGG AAATAGAGCG TCCCGCTGAA ACTGGCCCTA AAGCGCGCAC
 901 ACACCGCCCG TCACCCCTCC CAAGCCCCCT AAACCTAAACT AATTAAAACC CTACAACCCG
 961 CGAAGGGAG GAAAGCTGA ACATGGTAAG TGTAACGGAA GGTGCGCTTG GAAAATCAG
 1021 AGTGTAGCTA AGATACTATA CAGCATCTCA CTTACACCGA GAAGACGTCC GTGCAAGTC
 1081 GACCACCTG ACGCCTATA GCTAGCCCAA CCCCTTAACA CAACAGACCC CCATTATAA
 1141 CCCCTAAAGC ACGAAACACC CACGTAGCTA ACCATTCTC CCCCCTAAGT CCAGGGCGATA
 1201 AAAAGGAAA TTTTGGAGCA ATAGAAAAG TACCGAAGG GAAAGCTGAA AGAGAGATGA
 1261 AAAAGCCAG TAAAGCTTAA GAAAGCAGAG CTTAGTGTCTC GTACCTTTIG CATCATGATT
 1321 TAGCTAGCAC TTTCAAGCAA AGAGAACTA CAGTTGTAAC CCCCAGAAACT GAGTGAGCTA
 1381 CTCCAAGACA GCCTATTAT AGGGCGAAC CGTCTCTGTG GCAAAAGAGT GGGAAAGAGCT
 1441 TTGAGTAGAG GTGACAAAC TACCGAACCTC AGTTATAGCT GGTTGCCTGT GAATTGAATA

1501 GAAGTTCAAG CCCCTGGGTT CTCCACTCAT GCACTTTGT AACCCTTCAG ATGCAATGAG
 1561 AAACCAGGGG TGTTAGTCAA AGGGGGTACA GCCCCTTGA TACAAGACAC AACTTCTCA
 1621 GCAGGATAAA GATCATATTC AAATAAGGC AAATGTTTA GTGGGCCTAA AAGCAGCCAC
 1681 CTAATAGAA AGCGTTAACG CTCAAACATA AAGCCCTCCC GGTATACCGA TAACCCTATC
 1741 TTATTCCCCC ACATTTAACG GGCCCTCCTA TGCACCAT AGGAACGACT ATGCTAATAT
 1801 GAGTAATAAG AGGGTATAAT CACCCTCTCC TTGCACATGT GTAAATCGGA ACGGACCCCC
 1861 CACCGAATCT TAACGGCCC AATCAAAGAG GGTATTGGAA ACTACCACAA ATTTAGGCCA
 1921 GAAAAACATC CAATGCAAAC CGGTTAACCC CACACTGGT TGCCCCAAAAG GAAAGACCAA
 1981 AGGGGGAGA AGGAACCTGG CAAACATCC CCAAGCCTCG CTTGTTTACG AAAAACATCG
 2041 CCTCTTGATC AACACAGTA TAAGAGGTC CGCCTGCCA GTGACCATAT AGTTCAACGG
 2101 CCGCGGTATT TTGACCGTGC AAAGGTAGCG TAATCACTTG TCTTTAAAT GAAGACCTGT
 2161 ATGAATGGCA TAACGAGGGC TTAACTGTCT CCTTCCCCG GTCAATGAAA TTGATCTCCC
 2221 CGTGAGAAG CGGGGATTAA ACCATAAGAC GAGAAGACCC TATGGAGCTT TAGACACACA
 2281 GGTGGGCAT GTCAAATACC CCTAGCTAAG GGCTCTGAAC AAATGGAATC TGCTTGATG
 2341 TCTTCGGTTG GGGGACCAT GGGGAATACA AAACCCACAC GTGGAAAGGG AGCACACCCC
 2401 TAAGTTACTT CTTCTCCCCG AAGCCAGAGC AACAGCTTA ACCAGCAGAA ATTCTGACCA
 2461 AACTGATCCG GTAAACCCGA TCAACGAAAC AAGTTACCC AGGGATAACA GCGCAATCCC
 2521 CTTTTAGAGC CCATATCGAC AAGGGGTTT ACGACCTCGA TGTTGGATCA GGACATCCTA
 2581 ATGGTGCAGC CGCTATTAAAG GGTCGTTTG TTCAACGATT AAAGTCCTAC GTGATCTGAG
 2641 TTCAGACCGG AGTAATCCAG GTCAGTTCT ATCTATGACA TGATCTTTT TAGTACGAAA
 2701 GGACCGAAAA GAGGAGGCC CTGCTAAAG TACGCCCTAC CCCCACCTAA TGAAAAAATC
 2761 TAAAATAGGC AAAAGGGCAT AACCTTTTG CTGAGATAA CAGCAAGTTG GGGTGGCAGA
 2821 GCGCGGCTAA TGCAAAAGAC CTAAGCCCTT TCTACAGAGG TCTAAATCCT CTCCTTAAC
 2881 ATGATTCAA CTCTCATTAC TCATATTATC AACCCACTAA CCTTTATTGT GCCTGTATTG
 2941 CTAGCGTAG CATTCTCAC CCTCCTGAA CGAAAAGTC TAGGCTACAT ACAACTACGA
 3001 AAAGGGCCAA ACATTGTTGG GCCTTACGGC CTCCCTCAAC CTATTGCTGA TGGCGTAAAA
 3061 CTCTTATTAA AAGAACCCAT TCGACCTCA ACCGCATCAC CCGCCTATT CCTCGTAGCC
 3121 CCTATGCTCG CTCTCACACT AGCCCTAAC CTTTGAGGCC CCATACCTTT CCCGTACCC
 3181 GTCGTGGACC TTAACCTAGG TATTTTATT ATTCTAGCTC TATCTAGCCT TGCAGTATAT
 3241 TCTATTCTGG GCTCGGCTG AGCATCTAAT TCAAAATATG CCCTAGTAGG GGCACCTGG
 3301 GCTGTCGCAC AAACAATTTC TTACGAAGT AGCCTTGGAC TCATCTTACT AAACATTATC
 3361 ATCTTCACCG GGGGCTTCAC ACTCCAAACCC TTTAACACAG CCGAAGAGGC CATTGACTG
 3421 GTAGTGCCTG CTTGACCGCT GGCTGCCATA TGATATATCT CCACTTTGC CGAAACAAAC
 3481 CGTGCACCCCT TCGACTTAAC AGAAGGAGAA TCTGAACTTG TCTCGGGCTT CAACGTAGAG
 3541 TACGCAGGGG GACCCCTCGC CTTGTTCTC TTAGCCGAAT ATTGAAATAT TCTCCTAATA
 3601 AATAACCTGT CCGCTACACT ATTTTTAGGG GCCTCCACA TCCCCTATCAT CCCGGAATTA
 3661 ACTAGTATAC ATATTATGAC TAAAGCAGCC CTTCTATCAA TTGTTCTCCT GTGGGTCCGA
 3721 GCTTCGTACC CGCGATTCCG TTACGACCA CCTATGACCA TTGTTGAAA AAACCTTCTT
 3781 CCACTGACTT TGGCACTAGT TATCTGACAC TTGGCGCTCC CTATTGCTT CGCTGGTCTC
 3841 CCACCAACAC TATAGGCCA GGAGCTGTG CTTGAAATTAA AGGGCCACTT TGATAGAGTG
 3901 ATTATGGGG GTTAGAGTCC CCCCCAACTCC TTAGAAAGAA GGGGATCGAA CCCAACCTGA
 3961 AGAgATCAAAC ACTCTTCGTG CTTCCCTAC ACCACTTCCT agTAAGGTCA GCTAAATAAG
 4021 CTCTGGGCC CATAACCCAA CCATGTAGGT TAAATCCCTT CCTTTGCTAA TGAACCCCTA
 4081 CATCTTGACC ACTCTTCTAT TTGGTTGGG CCTTGGAAACA ACACCTCACGT TTGCAAGCTC
 4141 ACACGGCTC CTCGCTTGA TAGGACTTGA GATTAATACA CTAGCCATTA TTCCCCCTAAT
 4201 AGCCAAACAC CACCAACCCC GGGCAGTCGA AGCTACTACT AAATACTTTC TCGCACAAGC
 4261 CACAGCAGCC GCCACCCCTC TGTTGCAAG CACTACCAAC GCTTGACTCA CAGGCCAGTG
 4321 GGATATTCAA CAGATAACAC ATCCCCCTCC CACAACATG ATTGTTATTG CCCTCGCGCT
 4381 AAAGGTCGA CTGGCACCAA TGCACCTTG ACTCCCGAA GTACTCCAAG GGCTAGACCT
 4441 GACCACGGG CTCATCCTCT CAACCTGACA AAAACTTGCA CCCTTGCCC TCTTACTACA
 4501 ATTCTAAACCA ACCAACCCCTA CGCCTCTAAT CATATTGGC TTACTATCCA CTCTCGTGG
 4561 TGGCTGAGGT GGTCTCAACC AAACCCAGT GCGCAAGGTC CTGCGCTACT CCTCAATTG
 4621 CCACCTTGGC TGAATAATGC TCATCTTCA ATTTCACCC CCTCTTGCCC TTCTCACCC
 4681 CCTCACGTAC TTTACTATGA CTTTCTCAGC ATTCTCTAATT TTTAAAGTAA ACAAAGCGAC
 4741 CACCGTCAAT GCCCTCGCCA TCTCGTGAAC AAAAACTCCT GCCCTCACAG CCCTAGCACC
 4801 CCTGGTTTA CTTTCTCTCG GCGGCCTTCC TCCCCCTTACG GGCTTCATGC CCAAGTGATT
 4861 CATTCTTCAG GAACTGACTA AACAAAGACCT TCCGGCACTT GCCACCCCTCG CCGCTTTAAC
 4921 TGCCCTTGT AGTCTCTACT TCTACTTGG CCTCTCGTAT GCAATAACCC TCACAATATT
 4981 CCCAACAAAC CTAAAGTGT TCACTCCCTG ACGATTCTAC TCCCCCAAT TTACACTTCC
 5041 CCTCGCCGTC TCCACTGAG CAACACCCCT TCTCTCCCA CTAGCCCCCG CTGCCGTAGC
 5101 ACTACTGGTC ACCTAAAGGG CTTAGGCTAG CACTTAGACCC ATGGCCTTC AAAGCCGTAA
 5161 GCGTGAGTGA AAATCTCTCA GTCCCTGTAA AGACTTGCAG GCTACTATCC CACATCTCT
 5221 GCATGCAAAG CAAACACTT AATTAAGCTA AAGCCTTCT AGATAGGAAG GCCTTGATCC
 5281 TACAAAATCT TAGTTAACAG CTAAGTGCCTC AATCCAGCGA GCATCTATCT ACTTTTCCCC
 5341 GCCTAAATTAA GAAACTAAAG GGCAGGGAAA AGCCCCGGCA GGCAGTTAGC CTGCTTCTT
 5401 AGATTGGCAA TCTAACATGT AACACCCAG GGCTTGGTAA GAAGAGGGCT TGACACCTCTG
 5461 TCTATGGTC TACAATCCAC CGCTTAACCTA AGCCATCTCA CCTGTGGCAA TCACACGTTG
 5521 ATTCTTCTCG ACCAACATCAC AAGACATCGG CACCCCTCATC CTCTGTATTG GTGCCTGAGC
 5581 CGGAATAGTG GGGCAGGGCC TAAGTCTGCT TATCTGGCA GAACTAAGCC AACCCGGGGC
 5641 TCTCCTGGGA GACGACCAA TTTATAATGT GATCGTCACC GCACACGCCT TTGTAATAAT
 5701 CTTTTTATA GTAATACCA TTATGATTGG GGGTTCGGA AACTGGCTTA TTCCACTAAT
 5761 ATTGGGGCC CCAGACATGG CGTTCCCTCG AATGAATAAT ATGAGTTTCT GACTTCTCC
 5821 CCCCTCCTT CTCCCTCTCT TAGCCTCTC AGGTGTTGAA GCCGGAGCAG GTACCGGATG
 5881 AACCGTGTAC CCCCCACTAG CTGGCAATTAG CACCCACGCC GGGGCATCCG TAGACCTGAC
 5941 AATCTCTCA CTTCACCTG CAGGAATTTC CTCAATTCTG GGGCAATTAA ACTTTATTAC
 6001 TACCATCATT AACATGAAAC CCACAAACAGT CACTATGTAC CAAATCCCGT TATTTGTTG
 6061 AGCCGTTCTT ATTACAGCCG TACTTCTCT TCTGTCCCTG CCCGTTTAG CGCAGGGAT
 6121 TACAATGCTA CTAACAGACC GCAACCTAA CACGACCTTC TTGACCCCTG CGGGAGGAGG
 6181 TGACCCATT CTCTACCAAC ACCTATTCTG ATTCTTTGGT CACCCAGAGG TATACATTCT

6241 TATCCTCCCA GGCTTCGGAA TAATTTCTCA CATTGTTGCA TACTATGCAG GTAAGAAAGA
 6301 ACCTTTGGC TACATGGGA TAGTCTGAGC TATAATGCC ATTGGACTCC TGGGCTTCAT
 6361 TGTCTGGCC CATCACATAT TTACAGTCGG AATAGACGTA GATACACGAG CCTACTTAC
 6421 CCTCTGCCACA ATAATCATG CAATTCACAC TGGCGTAAAA GTCTTAGCT GACTCGAAC
 6481 CCTCCATGGG GGAAGCATTA AATGAGAAC GCCCCCTCTA TGAGCCCTCG GCTTTATTT
 6541 CCTCTTACA GTAGGCGGTC TCACTGGCAT TGTCTTAGCT AACTCCTCTC TCGATATTGT
 6601 TCTGCATGAC ACATACTATG TAGTCGCCA CTTCCACTAT GTACTATCTA TGGGTGCTGT
 6661 ATTTGCAATC GTTGCGGCCT TCGTCATTG ATTCCGTTA TTTACAGGCT ATACCCTCA
 6721 CTCCACATGA ACAAAAATCC ACTTCCGGCCT GATGTTTATT GGGGTCATC TAACATTCTT
 6781 CCCTCAACAT TTTCTGGGAT TGGCTGGGAT ACCCCGACGG TACTCAGACT ACCCAGACGC
 6841 ATACACCCCTT TGAAACACTG TTTCATCAAT TGGGTCCCTA ATGTCCTCG TTGCTGTAAT
 6901 TTTATTCTTA TTCATTATTG GAGAACATT TACAGCCAA CGAGAACGTC GAGCAGTAGA
 6961 ACTAACTGCA ACTAACATG AATGACTTTA CGGCTGCCCT CCCCCCTTAC ACACATTGA
 7021 AGAGCCGCA TTCGTACAAG TTCGTATAAA TTCGAACAA CTAACGAGAA AGGGAGGAGT
 7081 TGAAACCCCA TCAATTGGTT TCAAGAACAC CACATAACCG CTCTGTCACT TTCTTCaCaA
 7141 CACACAAATA AGATACTAGT AAAACGCTA AACACTGCT TGTCAAGGcA GAATTGTTGGG
 7201 tTCAACCCCCC GCGTATCTTA AACGCAATGG CACATCCCTA ACAAATCGGA TTTCAAGACG
 7261 CAGCTTCACC CTTAATAGAA GAACACTTC ACTTCCACGA CCACGCCCTA ATAATTGTA
 7321 TTCTCATTAG CACAATGGTA CTTTATATTA TTGCGGCTAT AGTCACAGCC AAACAAACAG
 7381 ATAAACTTGT ACTAGACTCT CAAGAAATTG AAATTATCTG AACAGTTCTC CCAGCTGTTA
 7441 TTCTTATCCT TATCGCCCTA CCATCCATCC GAATTGTTGA CTTAATAGAC GAAATTATG
 7501 ACCCCCCACCT TACAATCAA GCCATCGGCC ACCAATGATA CTGAAGCTAC GAATACACAG
 7561 ACTACGAAGA CCTCGGTTT GACTCATACA TAATCCCCAC ACAAGACCTG ACCCCCAGAC
 7621 AATTCCGACT ACTAGAACCA GACCCACGG TAGAATCTC AGTAGAATCT CCAATTCCG
 7681 TTTTAATCTC GGCTGAAGAT GTCTTACATT CCTGGGCACT CCCCTCTCTA GCGTAAAG
 7741 TTGACGCCGT ACCTGGACGA TTGAACCAAG CAACCTTTAT TGTTAGCCGA CCAGCGTAT
 7801 TTTTCGGACA ATGTTCTGAA ATTTGCGGGG CTAACCATAG TTTTATACCC ATTGTTGTAG
 7861 AAGCAGTCCC CCTTGAACAC TTTGAGAACT GGTCTTCACT AATAATTGAA GACGCCCTCG
 7921 TAAGAAGCTA ATAAGTACAA GCGTAGCCT TTTAAGCTAA AGACTGGTGC CTAACAAACCA
 7981 CCCCTAGCGA CATGCTCAA CTGAACCCCG CACCCCTGATT TGCAATTTTA GTTTCTCTT
 8041 GAATAATCTT TTAACTGTT ATTCCCCC CAAAGTTTACG ACACACCTAT CCAATGAAC
 8101 CCACCTCCA AAGCACCAA AAACCTAAAA CAGAGCCCTG AAACGTGACCA TGATACTAAG
 8161 CTTCTTGAC CAATTATAT CCCCTCATA CCTCGGCATT CCACTAATTG CACTAGCGAT
 8221 TAGCCTGCCT TGAATTCTGT TCCCAACCCC TCAAGCCCGC TGATTAACAA ACCCGCTGCT
 8281 AACACTTCAG GGGTGATTG TTAGCCGTT CACCTCACAA CTTCTTCTAC CCCTAAACCC
 8341 TGGGGGCCAC AAGTGGCAG TTCTATTAC CTCGTTAATA GTTTCTCTAT TTTCTATCAA
 8401 CATGCTAGGA CTCCTTCCAT ACACCTTCAC ACCAACACCC CAACTCTCCC TTAACATGGG
 8461 CCTCGCAGTA CCCCTTGAT TGGCAACTGT CATTATCGGG ATACGAAATC AACCAACGCA
 8521 TGCCCTAGGC CACCTCTC CAGAAGGCC CCCCACTGCG CTCATCCCTG TACTTATTAT
 8581 CATTGAAACA ATTAGCTAT TCATTGACCC CTTAGCTCTC GGTGTCGAC TAACAGCAA
 8641 TCTCACAGCA GGCCACCTCC TTATTCACT CATTGCAACA GTCGCCTTCG TCCTCTTCC
 8701 CCTCATACCT GTTGTGCTA TCTTAACAAC GGTGGCTCTC TCCCTCCTCA CTCTACTAGA
 8761 AGTAGCCGTT GCCATGATTC AAGCTTATGT CTTTGTCTA CTCCTAAGCC TATACTTACA
 8821 AGAGAACGTT TAATGCCCA TCAAGCACAC CCATACACCA TAGTCGACCC CAGCCCATGA
 8881 CCCCTAAACAG GGGTATTG TGCCCTATTG ATAACATCTG GCCTTGCTAT CTGATTCCAC
 8941 TTCCACTCCA CAACCTTAAT AACATTGGG ACAATCCTTC TCACCTTAAC AATCTTCAA
 9001 TGATGACGGG ACGTCTACG AGAAGGCACA TTCAAGGAC ACCACACTCC CCCCCTCAA
 9061 AAAGGCCTCC GATATGGGAT AATTATTTT ATTACCTCAG AAGTCCTATT CTTCTTAGGC
 9121 TTCTCTGAG CCTTTTACCA CTCAAGCCTA GCACCTACCC CTGAGCTAGG CGGCTCTGA
 9181 CCACCAAGC GCAATTACTCC CTTAGACCA TTTGAAGTCC CACTTCTTAA CACAGCAGTC
 9241 CTGCTCGCCT CTGGCGTAAC TGTTACCTGA GCACATCACA GCATTATGGA GGGTAAACGA
 9301 AAACAAACTA TTCAATCTCT TGCTCTTACA ATCTTACTCG GGGGGTACTT CACTTCTCTT
 9361 CAAGGCCTAG AGTATCACGA AGCCCCCTTC ACCATTGCG ACCGGGGTTA CGCGCTACA
 9421 TTCTTGTGTC CCACGGCTT CCACGGATTA CACGTCTAA TTGGCACATC ATTCTTAGCC
 9481 GTTTGTCTAT TACGCCAGAT TCTTCACCAT TTACACATCAA ACCACCACTT TGGGTTGAA
 9541 GCAGCCGCAT GATACTGACA CTTCTGAGAC GTCTGCTGAC TTTCTCTCTA TATCTCTATT
 9601 TACTGATGAG GATCTTAATC TTCTAGTAT TAAATCTAGT ATAAGTGAATC CCAATCACC
 9661 CGGTCTGGT TAAAATCCAA GGGAAAGATAA TGAGCCTTCT CTTAACCATC ATTTCGATTG
 9721 CCACCCCTCC CTCGACAGTA CTAGCCCTTG TATCCTTTTG ATTACCCCAA ATTACACCAAG
 9781 ACCACGAAAAA GCTATCGCCG TACGAGTCG GCTTTGACCC AATAGGATCC GCCCGCTG
 9841 CTTTTCTACT GCGGTTTCTC CTCATCGCCA TCCCTTTCTC TCTCTTCGAT TTAGAAATTG
 9901 CTCTCCTCCCT CCCTCTTCA TGAGGAGACC AACTAGCATC GCCACTACTG ACGTTCACCT
 9961 GAGCTACAGC TGTCTAGCC CTTCTAACCC TTGGCCTCAT TTACGAATGA ATGCAAGGAG
 10021 GTCTAGAATG GGCTGAATAG GTGGTTAGTC TAAGAAAAAC ATTGATTTG GGCTCAAAAAA
 10081 CTTGTGGTTT AAATCCCAA CGCGCTTAATG ACCCCCCACAC ACTTTGCCTT CTCCCTAGCC
 10141 TTCTCTCTG GTTAAACAGG CCTGGCATTC CACCGGTTTC ACCTCTTTC CGCCCTATTG
 10201 TGCTTGAAG GAATAATACT ATCCCTATTG ATTGCTCTCT CCTCTATGAC GCTCCAACCT
 10261 GACTCAACCA ACTTTCTACG ATCTCCATAA CTTCTGCTT CATTTCAGC CTCCGAAGCA
 10321 AGCGCCGGCC TCGCCCTATT AGTAGGCCACT GCTCGAACCC ACCGGGACCGA CCGATTACAA
 10381 AGCCTAAATC TTCTCCAATG CTAAAAGTCC TAATCCCAAC ACTTATGCTA ATCCCAACAG
 10441 CCTGGTTACT CAAACCCAGC TGACTCTGGC CCATGACTTT AATGTATAGC TTTTGATCT
 10501 CCTTGTTAG CTTTCTCATGA CTAAAGAACCC TCTCAGAACCG CGGCTGATCG TCACTTAGCC
 10561 TCTTCTATAGC TACCGACTCC CTATCACCC CCCTCCTCGT TCTTACATGC TGATTACTCC
 10621 CACTAATAT TTTGGCAAGC CAAACACACCA CAGCTTAAAGA ACCTCTTAGT CGCCAGCG
 10681 TGTACATCAC ACTTCTCGCC TCACTCCAGT TTTCTCTAAT CCTAGCATT AGCGCCACCG
 10741 AACTGTTGAT GTTTTACGTA ATATTGAAAG CCACCCCTCAT CCCCCACGCTA ATCATTATCA
 10801 CCCGCTGAGG AAACCAAAACA GAACGTCTAA ACCGAGGAAC CTACTTTCTC TTCTATACAT
 10861 TAGCGGGCTC ACTCCCTCTT CTCGTTGCTT TACTCCTACT CAAAACACA TCCGGCACCC
 10921 TGTCGCTGTT AACCCCTCAC TATACGGACC CGCTCGCTCT CTCGCTTAT GCAGACAAAC

10981 TATGATGAGC AGGCTGTCTT TTAGCATTCC TGGTTAAAAT ACCCCTCTAC GGGGTCCACC
 11041 TATGGCTCCC TAAAGCCCAT GTTGAAGCCC CAATTGCAGG CTCATAAATC CTTGCAGCGG
 11101 TTCTACTGAA GCTAGGGGA TACGGCATGA TCCGCATAAT ACAATATTG GAACCTCTAA
 11161 CTAAGGAGTT AAGCTACCCC TTCATTGTCTT TTGCACTTTG AGGGGTAATT ATAACGGT
 11221 CAATTTGTCT ACGCCAGACA GACCTTAAGT CCCTAATTGC CTACTCCTCA GTCAGGCCA
 11281 TAGGCCTAGT GGCTGGAGGA ATTCTTATCC AATCACCTG GGGCTTAACA GGGGCTCTTA
 11341 CCCCTCATATAAT TGCACACGGG CTTACTTCAT CAGGCCCTTT CTGCTTAGCA AACACAAACT
 11401 ATGAACGAAC CCACAGTCGA ACGATAGTCC TGGCGCAGG ACTTCAAGTA GCCCTACCC
 11461 TAATAGGCCA TTGATGATT ATTCTAGT TGCCCAACATT AGCTCTGCCA CCACTACCCA
 11521 ACCTCATGGG GGAGCTGATA ATTATCACTT CTTATTCAA CTGGTCTGTA TGAACCTG
 11581 CATTAAACCGG TGCAGGCACT CTAATCACAG CTGGTTACTC CTTTACATA TTCTTAATAA
 11641 CTCAGGGGG CCCTCTCCCA ACACATGTCA TTGCACTTG GCCATCTCAC ACCCGAGAAC
 11701 ACCTCTTAT TGCACTTCAC CTCATCCCC TAATTCTACT CGTACTTAAG CCCGAGCTGA
 11761 TCTGAGGCTG AACTGCGTGT AGGTGTAGTT TAAACAAAAA CATTAGATTG TGATTCTAAA
 11821 AZATAAGGGT AAAATCCCTT CTCCACCGA GAGAGGCTCG CAGCAATGAG AACTGCTAAT
 11881 TCCCACGACC TTGGTTGGAC CCCTAGGCTC ACTCGAACAG TCCTAAAGGA TAACAGCTCA
 11941 TCCGTTGGTC TTAGGAACCA AAAACTTTG GTGCAATTC AAGTAGCAGC TATGTACCC
 12001 ACCTCCCTAA TGATTTCATC GAGCCGATC ACAATCTTGC CGTACTAGTC CTATCCTTA
 12061 ATCACCACGA TCCGCCCTAC ACCTCGGGAC CCTCAATGGG CTACCACCCA TGTAAAACA
 12121 GCTGTAAAAA TAGCTTCTT TGTTAGTCTG CTGCCTCTGA CCCTATTCT TAATGAGGGC
 12181 GCAGAAACAA TCGTTACTAA CTGAACCTGA ATAAATACCA ACTCCTTCAA TGTCAGCATT
 12241 AGCTTGAAT TTGATTCTA CTCATTATT TTTACACCA TTGCCCTCTA CGTAACTTGA
 12301 TCCATCCTGG AGTTCGATC ATGGTATATG CACCGAGCATCACATGAA CGCCTCTTT
 12361 AAATACCTGG TGACGTTTCT GATGGCCATC ATCGCTCTAG TAACTGCAA CAATATGTTT
 12421 CAACTATTAA TCGGTTGAGA GGGCGTGGGG ATCATATCGT TCTTACTTAT TGGATGGTGG
 12481 TTCGGACGGG CGGATGCTAA CACTGCAGCG CTTCAAGCAG TACTTTACAA CCGAGTAGGG
 12541 GACATCGGGC TTATCTTCG CATGGCTTGA ATAGCAGCAA ACCTAAACTC CTGGGAGATG
 12601 CAACAAATCT TTGCAACCAC CAAAGACATA GACTTAACCT ACCCTCTTCT CGGACTCATC
 12661 ATTGCGACAA CTGGCAAATC GGCTCAATTC GGACTTCAACC CTTGACTACC CTCCGCTATG
 12721 GAGGGTCCCA CACCGCTCTC TGCCCTACTT CATTCTAGCA CTATGGTCTG TGCCGGCATT
 12781 TTTCTGCTAG TACGATAAG CCCGCTCTG GAAAACAACC AAACCGCCCT CACAACCTGC
 12841 CTATGCTCG GAGCTTCTAC AACACTATTG ACTGCAACCT GTGCCCTCAC GCAAAATGAT
 12901 ATTAAAAAAA TCGTCGCTT CTCTACATCC AGCCAGCTTG GGCTTATAAT GGTAACCAT
 12961 GGTCTAAACC AGCCACAACT AGCATTCTT CACATCTGTA CACATGCCCTT CTTCAAGGCC
 13021 ATACTATTTT TATGCTCCGG ATCTATTATC CACAGCCTAA ATGATGAACA AGATATCCGG
 13081 AAAATAGGGG GCATACACCA CCTTGCACCT TTACATCTT CTTGCCTAAC TATTGGCAGC
 13141 CTCGCCCTCA CAGGAACCC CTTCTGGCT GGCTTCTCT CAAAGATGCA CATCATCGAA
 13201 GCATTGAAACA CATCCCACCT AAACGCCATGA GCCCTAACCT TAACCCTCCT AGCCACCTCT
 13261 TTCACAGCAA TCTACAGCTC TCGAGTAGTA TTTTTGTAC CCATGGGCCA CCCCCGATT
 13321 AGTCCCTCTC CTCCGATCAA CGAAAACAAC CCAGCAGTGA TTAACCCACT TAAGCAGACTA
 13381 GCATGGGGGA GTATCGTTGC AGGCTACTA ATTACCTCTA ATATTACACC CTTGAAAACC
 13441 CCTGTCTAT CTATACTCC CCTCCTTAACT CTAGCTGCTC TTGCAGTCAC AATCATCGGC
 13501 CTTCTTGTG CTATAGAGCT CGCCATGTT ACCAACAAAC AATATGAGCC AATACCCAAAC
 13561 ACGCACATAC ACCACTCTC TGCCCTTAA GGCTTCTCC CGGGGTTAT ACACCGCTA
 13621 ACCCTGTAT TGCTCTTCT CCTAGGCCAA AACATTGCTA ACCAAACAGT AGATCAGACA
 13681 TGATTAGAAA AAACGGGGCC CAAAGGCTT GAAGCCGCTC ACCGACCTCT CGTCTCTCT
 13741 ACAAGTAACA TTCAACGGGG GGTAATAAA ACATATCTCA CCCTCTTCTT CTTTACACTA
 13801 GCCCTTATAA TCCCCCTATT CATCACCTAA ACCGCTCGCA ACGCCCTCG ACTCATACCC
 13861 CGAGTTAACT CAAGGACTAC AAATAAGTT AAAAGCAATA CCCATGCACT AGTGATTAAC
 13921 AACCACCCCC CAGAAGAGTA TATTAAGCT ACCCCTCCAA TATCACCTCG AAACACGGAA
 13981 AACTCCCCCTA ACTCATCTGC TAGAACCCAA GAAGATTCT ACCACCCCCC TCAAAACAGC
 14041 CCGGAAACTA ATAATACTCC CCTCTGTAG ACAACTCCGT ATACCATTAAC CGACCAACTA
 14101 CCCAGCTCT CGGGTAGCC CTCAGCAGCT AACCGCTGCT ATATGCAA TACAACAAAC
 14161 ATCCGGCTA AGTAATTTAA GAACAAGACT AGTGACAAAAA AGGGCCCTCC ATGCCGACT
 14221 AATATACCAAC ACCCCATGCC AGCTACCACA ACCAACCCCCA AAGCAGCAA ATAAGGGGAC
 14281 GGATTAGAAG CAACTGCAAC TAACCCCAAC ACAAGAGAAA ATAAAACCTAA ATATATAGCG
 14341 AAAGTCATAA TTCTGCCAG GACTTTAACC AGGACTAATG GCTTGAAAAA CCACCGTTGT
 14401 TATTCAACTA CAAGAACCTT AATGGCCAGT CTACGTAAT CCCACCCCTCT TCTAAAATC
 14461 GCAAACGATG CTTCAGTCGA CCTCCCCGG CCCTCTAAAT TCTCGGTTTG ATGGAACCTT
 14521 GGGTCTCTT TAGGACTCTG TTAAATTACCA CAAATTGCGA CGGGCTTATT TCTAGCCATA
 14581 CACTACATC CAGACATTG TACTGCCCTC ACCCTCCGTGG CCCACATCTG TCGAGACGTC
 14641 AACTACGGCT GACTTATCCG AAGCATTCTA GCCAACGGCG CATCATTCTT TTTCATTG
 14701 CTCTACCTTC ATATTGGCCG AGGACTATAC TATGGCTCTT ACCTCTATAA AGAAACATGA
 14761 ACTGTTGGGG TTATTCTCT CTTCTCGTA ATAATAACAG CTTCTGTTGG ATACGTCTC
 14821 CCTTGAGGAC AAATGTCGTT TTGAGGGTCA ACCGTTATT CCAACCTCTTCT ATCTGCCGTC
 14881 CCTCTACGTCG GAAACACCT AGTCCAAATGAA ATCTGAGGGG GTTTTCTGT AGACAATGCT
 14941 ACCCTCACCC GGTCTTTGC CTTCACCTT CTCTCCCTT CTACATCGC TGCCGCAACA
 15001 TTAATTCTAC TACTCTTCT TCACGAGACT GGCTCAAACA ACCCCACCGG GCTAAACTCA
 15061 GACTCCGACA AAGTTCCCTT TCACCCCTAC TTCACTTACA AAGACCTCTT GGGCTTGC
 15121 GTACTTCTTA CTGCATTAGC ATCCCTAGCC CTATTTCTT CAAATCTCTT AGGAGACCA
 15181 GATAACTTTA CCCCTGCAAA TCCGCTGTT ACTCCCCCAC ACATTAAGCC AGAGTGATAC
 15241 TTCTCTGTTG CCTATGCTAT TCTACGCTCC ATTCCAAACA AACTTGGAGG AGTACTTG
 15301 CTTTATTCTC CCATCTTAGT TCTTATGCTC GTCCCTTCC TTCACTACCTC TAAACAAACGA
 15361 AGCCTAATAT TCCGTCCTT AACACAAATTG TTGTTTTGGT CTTTAGTAGC AGACGTTATG
 15421 ATTCTAACTT GAATTGGAGG GATGCCGTA GAACACCCCT TCGTCATCAT TGGACAAGTA
 15481 GCGCTCTCC TTTACTCTC CCTCTCTTA GTCTGATCC CAAACAGCAGG TTGACTAGAG
 15541 AATAAAAGTCC TTGGATGAAA ATGCACTAGT AGCTCAGCAC ACAGAGCCCC AGTCTGTAA
 15601 ACTGGCCGTC GGAGGTAGA ATCCTCCCTA CTGCTCAAAG AAAGGAGATT TCAACTCTA
 15661 CCCCTAACTC CCAAAGCTAG GATTCTAGCA CTAAACTATT CTTTGTGGTA CATAAATGTA

15721 CAATGAAGGA TTTTCATATA CATGTATGTA ATAACACCCT ATATTTATAG TAACCATTAA
 15781 ATGTGATGTA CTAGGACATA CATGTATAAT CACCAAATCT CGTAATACAG CACTCATC
 15841 TCAACATTAA TAACTAAGAT AGACTAAAC CTGAATAATC ACTAATCTT AATATGTGAA
 15901 AGTCCAGGAC CAGTCGAAT CTAAGACCGA ACACAACACT CATCAGTCGA GTTATACCAA
 15961 GACTCAAAT CTCGCCGACC CCAAAATCCT ATGTAGTAAG AGCCTACCA CCGGGTATTC
 16021 CTTAATGATA ACTCTTATTG AGGGTGAGGG ACAAAAATTG TGGGGGTTTC ACTCGGTGAA
 16081 TTATTCCTGG CATTGGTTC CTACTTCAAGG GCCATTGATT GATGTTATCC CTCACACTT
 16141 CATGCACACT TACATAAGTT AATGTGATGATA ATACATACGA CTCGTTACCC ACCAAGCCG
 16201 CGCTTCACTC CAGGGGTTAA GGGGGTCTCT TTTTTTTT CTTCTCACTT GACATTTCAA
 16261 AGGGCATCCA GCCAACAGAG AGCTTAAAG GGTGAGCCT TTTCTTGAC GCGCCGTACA
 16321 TAGTATTGCGT GTAATGAGTC TTTATTAGAA GAATAACATT AAGTGTATC ATGTGCATAA
 16381 AGATTTGCTC ATTTCTCTG TCATCCCCAG GATACCCCCT TTTTGCAGG CAAAACCCCC
 16441 CTACCCCCCT AAACCCCTGA AGTTGCTAAG ACCCCTGAAA ACCCCCCCGGA AACAGGACAA
 16501 ACCTCTGGTA GCTCAGAAAG GGCAATTACT TTCAACCCAC TTTTCaACCA CTTTTCAACC
 16561 CACAAATACC CCTGGCTTAT CATAACCCCCA CCAGTTGTTT TAATAATACC ACTTTTCAAC
 16621 CCACAAATAC CCCTGGCTTA TCATAACCCC ACCAGTTGTTT TAATAATAC CACTTTCAA
 16681 CCCACAAATA CCCCCTGGCT ATCTAAACCC CACCAAGTTGTT TAATAATAA CCACTTTCA
 16741 ACCCACAAAT ACCCCCTGGCT TATCATAACC CCACCAAGTTGTT TAATAATAA ACCACTTT
 16801 AACCCACAA TACCCCTGGC TTATCATAAC CCCACCAAGTT GTTTAATAA TACCACTTT
 16861 CAACCCACAA ATACCCCTGG CTTATCGTA CCCACCAAGTT GTTTTAATAA ATACCAC
 16921 TCAACCCACA AATACCCCTG GCTTATCATA ACCCCCACAG TTGTTTTAAT AATACCACTT
 16981 TTCAACCCAC AAATACCCCT GGCTTATCAT AACCCCAAGG GTTGTGTTAA TAATACCACT
 17041 TTCAACCCCA CAAATACCC TGGCTTATCA TAACCCCAAGG AGTTGTCTTA ATAATACCA
 17101 TTTCAACCC CAAATACCC TGGCTTATC ATAACCCCAAC CAGTTGTTT ATAATACCA
 17161 CTTTTCAACC CACAAATACC CCTGGCTTAT CGTAACCCCCA CCAGTTGTTT TAATAATACC
 17221 ACTTTCAAC CCACAAATAC CCCTGGCTTA TCGTAACCCCC ACCAGTTGTTT TAATAATAC
 17281 CACTTTCAA CCCACAGATA CCCCTGGCTT ATCGTAACCC CACCAAGTTGTT TAATAATAA
 17341 CCACTTTCA ACCCACAAAT ACCCCCTGGCT TATCATAACC CCACCAAGTTGTT TAATAATAA
 17401 ACCACTTTTC AACCCACAA TACCCCTGGC TTATCATAAC CCCACCAAGTT GTTTAATAA
 17461 TACCACTTTT CAACCCACAA ATACCTAAA CCCACAAAGG CGTATCACC TTCTTCATGA
 17521 TCAAAAAACT ACTGTTAGA TTATTCAGA TATTCCTTC ACACCCCTAA CTACGAAAGG
 17581 CATAAAACAG AATCCAAGAG AACTGCTCA TCCAAAACATA TCTTTTTAA CCGATAAGCC
 17641 CTTAAGTCCC CCTGAGAACCC AGCCTTTAC CGCCCCAAAG GCTAAACACA AACAAACTAC
 17701 CACCTCTAGG TGTGTTAGA AAATTTACT

Hippoglossus stenolepis 1 17841 bp

1 GCTGACGTAG CTTAATTAAA GCATAACACT GAAGATGTTA AGATGGGCC TAGAAAGCCC
 61 CGGGAGCACA AAGGCTTGGT CCTGACTTTA CTGTCGACTT TAACCTAAACT TACACATGCA
 121 AGTATCCGC CCCCCTGAG AATGCCACA ACTCCCTGCT TGGGAACCTAG GAGCCGGTAT
 181 CAGGCACAAG CCCAGCTAGC CCACGACGCC TTGCTTAGCC ACACCCCTCAA GGGAACTCAG
 241 CAGTGATAAA TATTAAGCCA TAAGTGAAAA CTTGACTTAG TTAAGGTTAA GAGGGCCGGT
 301 AAAACTCGTG CCAGCCACCG CGGTTATACG AGAGGCCAA GTTGACAAAC AACGGCGTAA
 361 AGAGTGGTTA GGGGATTAC TAAACTAGAG CGAACCGCTT CAAAGCTGT TATACGCA
 421 CGAAAGTATG AAACCCAAATT ACGAAAGTAG CTCTACTCAT CTCGAACCCA CGAAAGCTAA
 481 GAAACAAACT GGGATTAGAT ACCCCACTAT GCTTAGCCCT AAACATCGAT TGCACTATAC
 541 ACTCCATATC CGCCCGGGAA TTATGAACAT CAGTTAAAA CCCAAAGGAC TTGGCGGTGC
 601 TTAACATCCA CCTAGAGGAG CCTGTTCTAG AACCGATAAC CCCCGTTAA CCTCACCTTC
 661 TCTTGTTTA TCCGCCTATA TACCACCGTC GTCACTTAC CCTGTGAAGG CTTTACAGTA
 721 AGCAAAATTG GCACAGCCC AAACGTCAAGG TCGAGGTGA GTGAATGAGG AGGGAAGAAA
 781 TGGGCTACAT TTGCTAAACA TAGAAACAC GAATGTTGCA TGAAACATG TAACTGAAGG
 841 AGGATTAGC AGTAACCGAG AAATAGACGG TCCCGCTGA ACTGGCCCTA AAGCGCGCAC
 901 ACACCGCCCG TCACCCCTCC CAAGCCCCCT AAACAAACT ATTAAAACCC CTACAACCCG
 961 CGAAGGGGAG GAAAGTCGA ACATGGTAAG TGTACCGGG AGTGCGCTTG GAAAAATCAG
 1021 AGTGTAGCTA AGATAGATA CAGCATCTA CTTACACCGA GAAGACGTCC GTGCAAGTCG
 1081 GATCACCTG ACGCTTATA GCTAGCCAA CCCCTTAACA CAACAAACCC CCATTATAA
 1141 CCCCTAAAGC ACGAACACCC CACGTAGCTA AACCAATTCTC CCCCCTAAGT CCAGGCAGATA
 1201 AAAAAGGAAA TTTTGGAGCA ATAGAAAAAG TACCGCAAGG GAAAGCTGAA AGAGAGATGA
 1261 AAAAGCCAG TAAAGCTTA GAAAGCAGAG CTTAGTGCTC GTACCTTTTG CATCATGATT
 1321 TAGCTAGCAC TTTCAAGCAA AGAGAACCTA GAGTTTGAA CCCCCTAAACT GAGTGAGCTA
 1381 CTCCAAGACCA GCCTATTAT AGGGCGAGCC CGTCTCTGTG GCAAAAGAGT GGGAAAGAGCT
 1441 TTGACTTAGAG GTGACAAACCT TACCGAACTC AGTATAGCT GTTGCCTGT GAATTGAATA
 1501 GAAGTTCACTG CCCCCTGGATT CTCCACTCAT GCACCTTATT AACCCCTTCAG ATGCAATGAG
 1561 GAACCAGGGG TGTAGTCAA AGGGGGTACA GCCCCTTGTA TACAAGACAC AACTTTCTCA
 1621 GCAGGATAAA GATCATATTG AAATAAGGAC AGATGTTTA GTGGGCCTAA AAGCAGCCAC
 1681 CTTAATAGAA AGCGTTAAAG CTCAAACATA AAGCCCTCCC GTTATACCGA TAACCCCTATC
 1741 TTATCCCCC ACATTTAACCA GGCCTCTTA TGACCCACAT AGGAACGACT ATGCTAATAT
 1801 GAGTAATAAG AGGGTATAAT CACCCCTCTT TTGACCATGT GAAATCGGA ACGGACCCCC
 1861 CACCGGATCT TAACGGCCCA AATCAAAGAG GGTATTGGAA ACTACCAAA ATTAGGCCA
 1921 GAAAACATC CAAATGAAAC CCGTTAACCC CACACTGGTG TGCCCCAAAG GAAAGACCAA
 1981 AGGGGGGAGA AGGAACCTGG CAAACATACC CCAAGCCTCG CCTGTTTACCA AAAAACATCG
 2041 CCTCTTGCAT AACCACAGTA TAAGAGGTCC CGCCCTGCCA GTGACCATAT AGTCAACGG
 2101 CCGCGGTATT TTGACCGTGC AAAGGTAGCG TAATCACTTG TCTTTTAAAT GAAGACCTGT
 2161 ATGAATGGCA TAACGAGGGC TTAACCTGCT CTTCCCCCG GTCAATGAAA TTGATCTCCC
 2221 CGTGCAGAAG CGGGGATTAA ACCATAAGAC GAGAAGACCC TATGGAGCTT TAGACACACA
 2281 GGTGGCCCAT GTCAAATACC CCCAGCTAAG GGCTGACTAATGGAAATC TGCCTTGATG
 2341 TCTTCGGTTG GGGCGACCAT GGGGAATACA AAACCCCCAC GTGGAAAGGG AGCACACCCCC
 2401 TAAGTTACTT CTTCTCCCGC AAGCCAGAGC AACAGCTCTA ACCAGCAGAA ATTCTGACCA
 2461 AACTGATCCG GTAAAACCGA TCAACGAACC AAGTTACCT AGGGATAACA GCGCAATCCC

2521 CTTTTAGAGC CCATATCGAC AAGGGGTTT ACGACCTGA TGTTGGATCA GGACATCCTA
 2581 ATGGTGCAGC CGCTATTAAAG GTTCTGTTG TTCAACGATT AAAGTCCTAC GTGATCTGAG
 2641 TTCAGACCGG AGTAATCCAG GTCAGTTCT ATCTATGACA TGATCTTTc TAGTACGAAA
 2701 GGACCGAAAA GAAGAGCCC CTGCTAAAG TACGCCCTAC CCCCACCTAT TGAAAAAATC
 2761 TAAACTAGGC AAAAGGGCAT AACCCCTTTG CTGGAGATAA CAGCAAGTTG GGGTGGCAGA
 2821 GCCCGGCTAA TGCAAAAGAC CTAAGCCCTT TCCACAGAGG TTCAAATCCT CTCCTTAAC
 2881 ATGATTCAA CTCTCATTAC GCATATTATC AACCCACTAA CCTTATTGAT ACCTGTATTG
 2941 CTAGCCGTAG CATTCTTAC CCTCTTGAA CGAAAAGTGC TAGGCTACAT ACAACTCCGA
 3001 AAAGGCCAA ACATCGTAG GCCTTACGGC CTCCCTAAC CATTGCTGA TGGTGTAAAA
 3061 CTCTTATTAA AAGAACCCAT TCGACCTTCG ACCGCATCAC CGTCCTATT CCTCCTGGCC
 3121 CCTATGCTCG CTCTCACACT AGCCCTAAC CTTTGAGGCC CCATACCTT CCGTACCC
 3181 GTCGTGGACC TAAACCTAGG TATTTTATT ATTCTAGCCC TATCTAGCCT TGCAGTATAT
 3241 TCTATTCTGG GCTCCGGCTG AGCATCTAAT TCAAAATATG CCCTAGTAGG GGCACGTGCG
 3301 GCTGTGACAA AAACAATTTC TTACGAGTC AGCCTTGGAC TCATCTTGCT AAACATTATC
 3361 ATCTTACCG GAGGCTTCAC ACTTCAACAG TTTAACACAG CCCAAGAGGC CATTGACTA
 3421 GTAGTGGCCC CTTGACCATC GGCTGGCATA TGATATATC CCACTCTTGC CGAAACAAAC
 3481 CGTGCACCCCT TCGACTTAAC AGAAGGAGAG TCTGAACTTG TCTCGGGCTT CAACGTAGAG
 3541 TACGCAGGCG GACCCCTCGC CTTGTTCTC TTGGCCGAAT ATTCAAATAT TCTCCTAATA
 3601 AATACCTGT CCGTACACT ATTTTAGGG GCCTCCCACA TTCCATCTAT CCCAGAATTA
 3661 ACCAGTATCA ATATTATGAC TAAAGCAGCC CTTCTGTCAA TTGTCTTCCT ATGAGTCCGA
 3721 GCTTCGTACC CGCGATTCCG TTACGACCA CTTATGCACC TCATTTGAAA AAACTTCTT
 3781 CCACTGACTT TGGCACTAGT TATCTGACAC TTGGCGCTCC CTATTGCAATT CGCTGGCTC
 3841 CCAACCAACAA TATAGGCCA GGAGCTGTGC CTGAaTTAA AGGgCCACTT TGATAGAGTG
 3901 AATTATGGGG GTTAAAGTCC CCCCACATC TTAGAAAGAA GGGGATCGAA CCCAACCTGA
 3961 AGAGATCAA ACTCTTGTG CTTCTCTAC ACCACTTACT AGTAAGGTCA GCTAAATAAG
 4021 CTCTGGGCC CATAACCCAA CCATGTAGGT TCAAGTCCTT CTTTACTAA TGAACCCCTA
 4081 CATCTTGACC ACTCTTCTAT TTGGTTGGG CCTAGGAACA ACACTCACGT TTGCAAGCTC
 4141 ACACTGGCTT CTCGCTGAA TAGGACTTGA GATTAATACA CTAGCCATTA TTCCCCTAAT
 4201 AGCCCAACAT CACCAACCC GGGCAGTCGA AGCTACCACT AAGTACTTTT TCGCACAAGC
 4261 CACAGCAGCC GCCACCTCC TGTTTGCAG TACTACCAAC GCTTGACTCA CAGGCCAATG
 4321 GGACATTCAA CAGATAACAC ACCCCCCCCC AACAACAATC ATTGTTATTG CCCTCGCACT
 4381 AAAAGTCGGA CTGGCACCAA TGCACTTGTG ACTTCCCAGA GTACTCCAAG GACTAGACCT
 4441 AACCAACCGGA CTTATCCTCT CAACCTGACA AAAACTTGC CCCTTGCCCC TCTTACTACA
 4501 AATTCAAGACA ACCAACCTTA CGCCTCTAA CATTATTGGC TTACTATCCA CGCTCGTTGG
 4561 TGGCTGAGGT GGTCTAACCC AAACCCAGTT ACGCAAAGTC CTCGCCTACT CCTCAATTG
 4621 CCACCTTGGC TGAATAATGC TGATCGTTCA ATTTCACCC CCTCTTGCCCC TTCTCACCC
 4681 CCTCAGCTAC TTTACTATGA CTTTCTCAGC ATTCTCTAAATT TTTAAAGTAA ATAAGGCCAC
 4741 CACTGTCAAT GCCCTCGCA TCTCGTGAAC AAAGACTCTC GCCCCTCACAG CCCTAGCACC
 4801 CCTGGTTTA CTATCCCTCG GCGGCCCTCC CCCCTTCTC GGCTTCATAC CCAAGTGATT
 4861 TATCTTCAG GAACTGACTA AGCAAGACCT TCCAGCACTT GCCACCCCTCG CCGCTTAAAC
 4921 TGCCCTTGT AGTCTCTACT TCTACCTCG TCTCTCATAT GCAATAACCC TCACAATATT
 4981 CCCCAACAAAC CTCATTGGCG TCACTCCCTG ACGATTCTAC TCCCCCCAAAT TTACGCTTCC
 5041 CCTCGCGTC TCCACTGCAG CAACCAACCT TCTCCTCCCA CTAGCCCCCG CTGCGTAGC
 5101 ACTACTGGTC ACCTAAAGGGA CTTAGCTGAG CACTTAGACCA ATGGCCTTC AAAGCCGTA
 5161 GCGTGAATGAA AAATCTCTCA GTCCCTGTTA AGACTTGCAG GCTACTATCC CACATCTTCT
 5221 GCATGCAAAAG CAAACACTT AATTAAAGCTA AAGGCTTTCT AGATAGGAAG GCCTTGATCC
 5281 TACAAAATCT TAGTTAACAG CTAAGCCGGC AATCCAGCGA GCATCTATCT ACTTTTCCCC
 5341 GCCTAAATTAA GAAACTAAAA GGCGGGGAAA AGCCCCGGCA GGCAGTTAGC CTGCTTCTT
 5401 AGATTTGCAA TCTAACATGT AACACCCCG GGCTTGGTAA GAAGAGGGCT TGCACCTCTG
 5461 TCTATGGGTC TACAATCCAC CGCTTAACCTC AGCCATCCTA CCTGTGGCAA TCACACGTTG
 5521 ATTTTCTCG ACCAACATCACA AAGACATCGG CACCCCTAT CTCGTATTTG GTGCCTGAGC
 5581 CGGAATAGTG GGGCAGCGCA TAAGTCTGTT TATTGGGGCA GAAACTAAGCC AACCCGGGGC
 5641 TCTCTGGGA GACGCCAAA TTATATAAGT AATCGTCACC GCACACGCC TTGTAATAAT
 5701 CTTTTTATAA GTAATACCCA TTATGATGG GGGGTTCGGA AACTGGCTTA TTCCACTAAT
 5761 AATTGGGGCC CCAGATATGG CGTCCCTCG AATGAATAAC ATGAGTTCT GACTTCTTCC
 5821 CCCCTCCTT CTCCTCCTCT TAGCCTCTC AGGTGTTGAA GCCGGAGCAG GTACCGGATG
 5881 AACCGTATAC CCCCCACTAG CTGGCAATT AGCCCACGCC GGGGCATCCG TAGACCTGAC
 5941 AATCTCTCA CTTCACCTG CAGGAATTTC ATCAATTCTG GGGCAATTAA ACTTTTATTAC
 6001 TACTATCATT AACATGAAAC CCACAAACAGT CACTATGTAC CAAATCCCT TATTGTTTG
 6061 AGCCGTTCTT ATTACAGCCG TACTTCTCT TCTATCCCT CCGGTTTTAG CGCAGGAAT
 6121 TACAACTACTA CTGACGACCC GCAACCTAA CACGCCCTTC TTTGACCCCTG CGGGAGGAGG
 6181 TGACCCCATC CTCTACCCAGC ACCTATTCTG ATTCTTTGGT CACCCAGAGG TATACATTCT
 6241 TATTCTCCCG GGCTCGGAA TAATTCTCA CATTGTTGCA TACTATGCAG GTAAGAAAGA
 6301 ACCTTTGGC TACATGGGA TAGTCTGAGC CATAATGGCC ATTGGACTCC TGGGGTTCAT
 6361 TGCTGGGCC CATCACATAT TTACAGCTGG AATAGATGTA GATACACAGAG CCTACTTAC
 6421 CTCTGCCACA ATAATCTG CGATTCCAAC TGGCGTAAAG GTCTTTAGCT GACTCGAAC
 6481 CCTCCATGGA GGAACCATTA AATGAGACCG ACCCCCTCTA TGAGCCCTCG GCTTTATT
 6541 TCTCTTACA GTAGGGGTC TCACCTGGCAT CGCTCTTAGCT AACTCCTCTC TCGATATTGT
 6601 TCTGATGAC ACATACTATG TAGTCGCTCA CTTCCACTAT GTACTATCTA TGGGTGCTGT
 6661 ATTTGCAATC GTTGGCCCT TCGTCCACTG GTTCCATTAA TTTACAGGCT ATACCCCTCA
 6721 CTCCACATGA ACAAAATCC ACTTCCGCT GATATTATT GGAGTCAATC TAACATTCTT
 6781 CCCCAACAT TTTCTGGCC TGGCCGAAT ACCCCGACGG TACTCAGACT ACCCAGATGC
 6841 ATACACCCCT TGAAACACTG TTTCATCAAT CGGGTCCCTA ATGTCCTCTCG TTGCTGTAA
 6901 TTTATTTTA TTCAATTGAGAAGCATT TACAGCCAAA CGAGAAGTGC GAGCAGTAGA
 6961 ACTAACTGCA ACTAACTATG AATGACTTTA CGGCTGCCCT CGGCCCTACC ACACATTG
 7021 AGAGCCCGCA TTGCTACAAG TTGCTATAAA TTGCAACAAG CTAACGAGAA AGGGAGGAGT
 7081 TGAACCCCCA TCAATTGGTT TCAAGCCAAAC CACATAACCG CTCTGTCACT TTCTTCACAA
 7141 CACACCAATA AGATACTAGT AAAACGCTAT AACACTGCCT TGTCAAGGCA GAATTGTGGG
 7201 TTCAACCCCCC GCGTATCTTA AACGCAATGG CACATCCCTC ACAACTCGGA TTTCAAGACG

7261 CAGCTTCACC CTTAATAGAA GAACTACTTC ACTTCCACGA CCACGCCCTTA ATAATTGTAA
 7321 TTCTCATTAG CACAATAGTA CTTTATATTA TTGCGGCTAT AGTCACAGCC AAACTAACAG
 7381 ATAAACTTGT ACTAGACTCT CAAGAAATTG AAATTATCTG AACAGTTCTC CCAGCTGTTA
 7441 TTCTTATCCT TATCGCCCTA CCATCCATCC GAATTGGTGA CTTAATAGAC GAAATTAAATG
 7501 ACCCCCACCT GACAATTAAA GCCATTGGCC ACCAGTGATA CTGAAGCTAC GAATACACAG
 7561 ACTACGAAGA CCTCGGTTT GACTCATACA TAATCCCCAC ACAAGACCTG ACCCCCAGAC
 7621 AATTCCGACT ACTAGAACCA GACCACCGAA TAGTAATCCC GGTAGAATCC CCAATTCCGCG
 7681 TTTTAATCTC GGCTGAAGAT GTCTTACATT CCTGGGCAGT CCCCTCTCTA GGCGTAAAAG
 7741 TTGACGCCGT ACCTGGACGA TTGAAACAG CAACCTTTAT TGTTAGCCGA CCAGGCCGTAT
 7801 TTTTCGGACA ATGTTCTGAG ATTGCGGGG CTAAACATAG TTATTTACCC ATTGTTGTAG
 7861 AAGCAGTCCC CCTTGAAACAC TTTGAGAACT GGTCCTCACT ATAATTGAA GACGCCCTCGC
 7921 TAAGAAGCTA ATAAGTACAA GCGTTAGCCT TTTAAGCTAA AGACTGGTGC CTAACAAACCA
 7981 CCCTTAGCGA CATGCCCTAA CTGAACCCCG CACCCCTGATT TGCAATTAA GTTTCTCTT
 8041 GAATAATCTT TTAACTGTT ATTCCCCCA AAGTTTTAGC ACACACCTAT CCTAATGAAC
 8101 CAACCTCCCA AAGCACACAA AAACCTAAAA CAGAGCCCTG AACTGACCA TGATACTAAG
 8161 CCTCTTGAC CAATTTATAT CCCCTCTATA CCTCGGCATC CCACAAATTG CTCTAGCGAT
 8221 TAGCCTGCCT TGAATCTGTT GCCAACCCCC GCAAGCCCG TGATTAAACA ACCGCGTACT
 8281 AACGCTTCAA GGGTATTGTT TTAGCCGCTT CACCTCGAA CCTTCTCTAC CCCTAAACCC
 8341 CGGAGGCCAC AAGTGGCAG TTCTATTAC CTCGTTAATA GTTTCTCTAT TTTCTATCAA
 8401 CATACTAGGA CTCCCTCCAT ACACCTTCAC ACCAACGACC CAACTCTCCC TTAATATAGG
 8461 CCTCGCAGTA CCCCTTGAT TAGCAACTGT CATTATCGGG ATACGAAATC AACCAACGCA
 8521 TGCCCTAGGC CACCTTCTC CAGAAGGCAC CCCACTGCG CTCATCCCTG TACTTATTAT
 8581 CATTGAAACA ATTAGCCTAT TCATTGACCC CCTCGCTCTC GGTCGCGAC TAACAGCAAA
 8641 TCTCACAGCA GGCCACCTC TTATTCAC CATTGCAACA GCTGCCCTCG TCCTTCTCTC
 8701 CCTCATACCT GTTGTGCTA TCTTAAACAC GGTGGCTCTC TTCTCTCTCA CCCTACTAGA
 8761 AGTAGCCGTT GCCATGATTC AAGCTTATGT CTTGTTCTA CCTCTAAGCC TATATCTACA
 8821 AGAGAACGTT TAATGGCCA TCAAGCACAC CCATACCACA TAGTCGACCC CAGCCCATGA
 8881 CCCCTAACAG GGGCTATTGC TGCCCTATTG ATAACATCTG GCCTTGCTAT CTGATTCCAC
 8941 TTCCACTCCA CAACCTGTAT AACTATTGGG ACAATCCTTC TCACCTTAAC AATCTTCCAA
 9001 TGATGACGGG ACGTGCTAGC GGAAGGCACCA TTTCAAGGAC ACCAACATC CCCCCTCAA
 9061 AAaGGCCTCC GATATGGAAT AATTTTATT ATTACCTCG AAGTCCTATT CTTCTTAGGC
 9121 TTCTCTGAG CCTTTTACCA TTCAAGCTTA GCACCTACCC CTGACTTAGG CGGCTTCTGA
 9181 CAACGGCAG GCATTACTCC CTTAGACCCCTT TTGAAGTCC CACTTCTAA CACAGCAGTC
 9241 CTGCTTGCT CTGGCTAAC TGTTACCTGA GCACACCACA GCATTATAGA GGGTAAACGA
 9301 AAACAAACTA TTCAATCTCT TGCTCTTACA ATCTTACTCG GGGGGTACTT CACTTCTCTT
 9361 CAAGGCTAG AGTATCACGA AGCCCCCTTC ACCATTGAG ACGGAGTTTA CGGCGCTACA
 9421 TTCTTGTTG CCACCGGCTT CCACGGACTA CACGTCTAA TTGGCACGTC ATTCTTAGCC
 9481 GTTGTCTAT TACGCCAAT TCTTCACCAT TTCAACATTC ACCACCAACTT TGGTTGAA
 9541 GCAGCCGCTAT GATACGACA CCTCGTAGAC GTCGCTCTGAC TTTCTCTCTA TATCTCTATT
 9601 TACTGATGAG GATCTTAAATC TTCTAGTAT TAAATCTAGT ATAAGTGAATC TCCAATCACC
 9661 CGGTCTTGGT TAAAATCCAA GGGAAAGATAA TGAGCCTTCT TCTAACCATC ATTCGATTG
 9721 CAACCCCTCT CTCGACAGTA CTTGCCCTTG TGTCCTTTG ATTACCCCAA ATTACACCAG
 9781 ACCACGAAAA ACTATCACCT TACGAGTGG GCTTGACCC GATAGGATCC GCGGGCTGC
 9841 CTTTTCACT GCGGTTTTC CTCATCGCCA TCCCTTTCT TCTATTGAT TTAGAAATTG
 9901 CTCTCTCTC CCTCTCTTCA TGAGGTGACCA AACTAGCATC ACCACTACTG ACATTCCACCT
 9961 GAGCTACAGC TGTCTAGAC CTTCTAACCC TTGGCCTCAT TTACGAATGA ATGCAAGGGG
 10021 GTCTAGAATG GGCGGAATAG GTGGTTAGTC TAAGAAAAAC ATTGATTTG GGCTCAAAAAA
 10081 CTTGTGGTTT AAATCCGAA CCGCTTAATG ACCCCCCACAC ACTTTGCCTT CTCTCAGCC
 10141 TTCTCTCTGG GTTAAACAGG CCTGGCATTC CACCGGTTTC ACCTCCTTTC CGCTCTATTG
 10201 TGCTCTGAAG GAATAATACT ATCCCTATT ATTGCTCTTT CCCTATGAAC GCTACAACCT
 10261 GACTCAACCA ACTTTTCAGC ATCTCCAAATA CTTCTGCTTG CATTTCAGC CTGCGAAGCA
 10321 AGCGCCGGTC TCGCCCTATT AGTAGGCACT GCTCGAACCC ACGGGACCGA CCGATTACAA
 10381 AGCCTAAATC TTCTCAATG CTAAAATGTA TAATCCCAAC ACTTATGCTA ATCCCAACAG
 10441 CCTGGTTACT CAAACCAAC TGACTCTGAC CCATGACTTT ATGTATAGC TTTTGATCT
 10501 CCTGGTCAG CCTTTCTGTA CTAAAAAACCTC TCTCAGAAAC CGGCTGATCG TCACCTAGCC
 10561 TCTTATAGC TACCGACTCC CTATCGACCC CCCTCCTCGT TCTTACATGC TGATTACTCC
 10621 CACTAATAAT TTTGGCAAG CAGAAACACA CAGCCTTAGA ACCTCTTAGT CGCCAGCGCA
 10681 TGATCATCAC ACTTCTCGCC TCACTCCAGT TCTCCATAT CCTAGCATT AGCGCCACCG
 10741 ACTGGTAAT GTTCTACGTA ATATTGAAAG CCACCCCTAT CCCACGCTA ATCATTATCA
 10801 CCCGCTGAGG AAACCAAAACCA GAACGTCTAA ACGCAGGAAC CTACTTTCTC TTCTATACAT
 10861 TAGCGGGCTC ACTCCCCCTT CTCGTTGCTT TACTCCTACT CAAAACACA TCGGGCACCC
 10921 TATCGTTGTT AACCCCTCAC TATACGGACC CGCTCGCTCT CTCGTTTAT GCAGACAAAC
 10981 TATGATGAGC AGGCTGCTT TTAGCATTCC TGGTTAAAT ACCCCCTCTAC GGGGTCCACC
 11041 TATGGCTCCC TAAAGCCCAT GTTGAAGCTC CAATTGCAGG CTCAATAATC CTTGCAGCAG
 11101 TTCTACTGAA GCTAGGGGA TACGGCATGA TCCGCATGAT AACGATATTG GAACCTCTAA
 11161 CTAAGGAATT AAGCTACCC TTCTATTGTT TTGCACTTTG AGGGGTAATT ATAACCTGGCT
 11221 CAATTGCTCT ACGCCAGAGA GACCTTAAGT CCTAATTGCT CACTCATCA GTCAGCCACA
 11281 TGGGCTAGT GGCTGGGGG ATTCTTATTAC AATCACCTCTG AGGCTTAACA GGAGCCCTCA
 11341 CCCTCATATAAT TGCAACAGGT CTTACTTCTAT CAGCCCTCTT CTGCTTGGCA AACACAAACT
 11401 ATGAACGAAC CCACAGTCGA ACAATGGTCC TGGCGCGAGG ACTTCAAGTA GCCCTACAC
 11461 TAATAGCCAC TTGGTGGTTT ATTCTACTGTT TGGCAACCT AGCTCTGCCA CCACTACCA
 11521 ACCTCATGGG GGAGCTGATA ATTATCACTT CTTATTCAA CTGGTCTCTGG TGAACCTCTGG
 11581 CATTAACCGG TGCAGGCACT CTAATCACAG CTGGTTACTC CCTTTACATA TTCTTAATAA
 11641 CTCAACAGGGG CCCTATCCCA ACACATGTCA TTGCACTGTA ACCATCTCAC ACCCGAGAAC
 11701 ACCTCTTAT TGCACTTCAC CTTATCCCT TAATTCTGCT CGTACTTAAG CCGGAGCTGA
 11761 TCTGAGGCTG AACTGCTGT AGGTGAGTT TAAACAAAAA CATTAGATTG TGATTCTAAA
 11821 ATAAGGGTT AAAATCCCTT CTCCCACCGA GAGAGGCTCG CAGCAATGAG AACTGCTAAT
 11881 TCCCACGACC TTGGTGGAC CCCTAGGCTC ACTCGAAAAG CTCCCTAAAGG ATAACAGCTC
 11941 ATCCGTTGGT CTTAGGAACC AAAACTCTT GGTGCAAATC CAAGTAGCAG CTATGTACCC

12001 CACCTCCCTA ATGATTTCAT CGAGCCTGAT CACAATCTTT GCATTACTAG CCTATCCTTT
 12061 AATCACCAAG ATCCGTCCTA CACCTCGAGG CCCTCAATGA GCTACCACCC ATGTCAAAC
 12121 AGCTGTAAAA ATAGCTTCT TTGTTAGTCT GCTACCTCTG ACCCTATTTC TTAATGAGGG
 12181 CGCAGAAACA ATCGTTACTA ATTGAACCTG AATAAATACT AACTCCTTCA ATATCAACAT
 12241 TAGCTTGAAA TTTGATTCT ACTCAATTAT TTTACACCC ATTGCCCTCT ACGTAACCTG
 12301 GTCCATCCTG GAGTCGCAT CATGATACAT GCACGCAGAC CCACACATGA ACCGCTTCTT
 12361 TAAATACCTC CTGACGTTTC TGATTGCCAT AATCGTCCTA GTAACTGCAA ACAATATGTT
 12421 TCAACTATTAT ATCGGCTGGG AGGGAGTGGG GATCATATCG TTCTTACTTA TTGGATGGTG
 12481 GTTCGGACGG GCGGATGCTA ACACTGCAGC GCTTCAGCA GTAATTTACA ACCGAGTAGG
 12541 AGACATCGGA CTTATCTCG CCATGGCTT AATAGCAGCA AACCTTAAC CCGAGAGAT
 12601 ACAACAAATC TTTGCAACCA CCAAGGCAT AGACTTAACC TACCCCTCTTC TCAGGACTCAT
 12661 CATTGCAGCA ACTGGCAAAT CGGCTCAATT CGGACTTCAC CTTGACTAC CCTCCGCTAT
 12721 GGAGGGTCCC ACACCGGTCT CTGCCCTACT TCATTCTAGC ACTATGGTCTG TTGCCGGCAT
 12781 TTTCTGCTA GTACGAATAA GCCCCCTCTT GGAAAACAAAC CAAACCGCCC TCACAACCTG
 12841 CCTATGCCTC GGAGCTCTA CAACACTT CACTGCAACC TGTGCCCTCA CGCAAAATGA
 12901 TATTTAAAAA ATCGTCGCAT TCTCTACATC CAGGCCAGCTT GGGCTTATAA TGGTAACCAT
 12961 TGGTCTAAC ACACACAAAC TAGCATTCT TCACATCTGT ACACATGCCT TCTTTAAAGC
 13021 CATACTGTT TTATGCTCG GATCTATTAT CCACAGGCCAT AATGATGAAC AAGATATCCG
 13081 CAAAATAGGG GGCATACACC ACCTTGCAACC TTTCACATCC TCTTGCCTAA CTATTGGCAG
 13141 CCTCGCCCTC ACAGGAACCC CTTCTTGGC TGGCTTCTTC TCTAAAGATG CCATCATCGA
 13201 AGCATTGAAC ACATCCCAAC TAAACGCCCTG AGCCCTAACCT TAACCCTCC TAGCCACCTC
 13261 TTTCACAGCA ATCTACAGCT TTCGAGTAGT ATTTTTGTA CCCATGGGCC aCCCCCGATT
 13321 TAGTCCCCCTC TCTCCGATCA ACGAAAACAA CCCAGCAGT ATTAACCCAC TTAAGCGGCT
 13381 AGCATGGGA AGTATCGTT CAGGCTTAACCT AATATTACAC CTTAAAAC
 13441 CCCTGTCATA TCCATGCCCTC CTCTCTCTAA ACTAGCTGCT CTTGAGGTTA CAATCATCGG
 13501 CCTTCTGTT GCTATAGAGC TCGCCATATT AACAACAAG CAATATGAGC CAATACCCGG
 13561 CACGCACATA CACCACTTCT CTGCCCTATT AGGCTTCTTC CCCGGGGtCA TACACCGCT
 13621 AACCCCTGTA TTGCTTCTCT TCCTAGGCCA AAATATTGCT AACCAAACAG TAGATCAGAC
 13681 ATGATTAGAA AAAACAGGCC CCAAAGCCAT TGAATCCGCT AACCGACCTC TAGTGTCTC
 13741 TACAAGTAAT ATTCAACAGGG GGGTGTCAA AACATATCTT ACCCTCTTCC TCCTAACACT
 13801 AGCCCTATA ATCCCATAT TCATCACCTA AACCGCTCG AACGCCCTC GGCTCATACC
 13861 CCGAGTTAAC TCAAGGACTA CAAACAAAGT TAAGAGCAAT ACCCATGAC TAGTGATTAA
 13921 CAACCAACCC CCAGAAGAGT ATATTAAGGC TACCCCTCCG ATATCCCTC GAAATACGGA
 13981 AAACCCCCC AACTCATCTG CTAAGACCCA AGAAGATTCA TACCACCCCC CTCAAAACAA
 14041 TCCAGAAACT AATAATACTC CTCCGTATA GACAACCCCG TAGACCATT CCGACCAACT
 14101 ACCCCAGCTC TCTGGGTACG GCTCAGCAGC TAACGCTGCT GAATATGCAA ATACAACCTAA
 14161 CATCCCGCCT AAGTAATTAGA AGAACAGGAC TAGTGACAAA AAGGGCCCTC CATGCCAAC
 14221 TAATATACCA CACCCATGC CAGCTACCAC AACCAAAACCC AAAGCAGCAA ATAAGGGGA
 14281 CGGATTAGAA GCAACTGCAA CTAACCCCAA CACAAGAGAA ATAAAAACTA ATATATAGC
 14341 GAAAGTCATA ATTCTGCTCA GGACTTTAAC CAGGACTAAT GGCTTGAAA ACCACCGTTG
 14401 TTATTCACT ACAAGAACCC TAATGCCAG TCTACGTTAA TCCCACCCCTC TTCTAAAAAT
 14461 CGCAACAGAT GCTTGTGCT ACCTCCCAGC CCCCTCTAAAT ATCTCTGTT GATGAAACTT
 14521 TGGGCTCTT TTAGGACTCT GTTTAATTAC CCAAATTGCA ACCGGCTTAT TTCTAGGCAT
 14581 ACACTACACA TCAGACATG CTACTGCCCT CACCTCCGTA GCCCACATCT GTCGAGACGT
 14641 CAACTACGGC TGACTTATCC GAAGCATTCA TGCCAAACGGC GCATCATTCT TTTCATTTG
 14701 CCTCTACCTT CATATTGGCC GAGGCTCTATA CTATGGCTCT TACCTCTATA AAGAAACATG
 14761 AACTGTTGGG GTTATTCTCC TCCTCTCTG AATAATAACA GCCTTCGTTG GATACGCTCT
 14821 CCCTTGAGGA CAAATGCTGT TTTGAGGTGC AACCGTTATT ACCAACCTTC TATCTGCCGT
 14881 CCCCTACGTC GGAAACACCC TGGTCCAATG GATCTGAGGT GTTTTTCTG TAGACAATGC
 14941 TACCCCTACC CGGTTCTTG CCTTCCACTT TCTCTTCCCC TTCATCATCG CTGCCGCAAC
 15001 GTTAATTCTA CTACTCTTC TTCAAGAAC TGGCTCAAC AACCCAACCG GGCTAAACTC
 15061 AGACTCCGAC aAGCTTCCCTT TTCAACCCCTA CTCCTACTAC AAGGACCTCT TAGGCTTTC
 15121 AGTCTCTCTT ACTGCTTACG CATCCCTAGC CCTATTTC CCAAATCTCT TAGGAGACCC
 15181 AGATAACATT ACCCCCTGCAA ATCCGTTGT TACTCCCCCA CACATTAAGC CAGAGTGATA
 15241 CTTCTGTGTT GCCTATGCTA TTCTCCGCTC CATCCCAAAC AAACCTGGAG GAGTACTG
 15301 CCTTTTATTC TCAATCTTAG TTCTTATGCT CGTCCCCCTC TTTCATACCT CTAAACAAACG
 15361 AAGCCTGATA TTCCGTCCTC TAACACAATT TTTGTTCTGA TCTTGTAGTAG CAGACGTTAT
 15421 GATTCTAATC TGAATTGGAG GAATGCCGT AGAACACCCCC TTGTTTATCA TTGGGCAAGT
 15481 AGCGTCTCTC CTTTACTCT CCTCTCTCTG AGCTCTGAT CCAACAGCAG GTGACTAGA
 15541 GAATAAAAGTC CTTGGATGAA AATGCACTAG TAGCTCAGCA CACAGAGGCC CAGTCTTGTA
 15601 AACTGCCGT CGGAGGTTAG AATCTCCCT ACTGCTAAA GAAAGGAGAT TTCAACTCCT
 15661 ACCCCTAACT CCCAAAGCTA GGATTCTAGC ACTAAACTAT TCTTGTGTT ACATAAAATGT
 15721 ACAATGAAGG ATTTTCATAT ACATGTATGT AATAACACCA TATATTATA GAAACATT
 15781 TATGTAATGT ACTAGGACAT ACATGTATAA TCACCCACATC TCGTAATAGA AACTCATTC
 15841 ACCACCAATT TTAACTAAGA CGAACTAAA CCTGAGTGT CACTAATCTT AAATAAGTGA
 15901 AAGTCCAGGA CCAGTCGAGA TCTAAGACCG AACACAACAC TCATCAGTCC AGTTATACCA
 15961 AGACTCAAAA TCTCGCCGAC CTCAAATTCTG TATGAGTAA GAGCCTACCA ACCGGTATT
 16021 CCTTAATGAT AACTCTTCT GAGGGTGGAG GACAAAAATT GTGGGGGTTT CACTCGGTGA
 16081 ATTATTCCTG GCATTGGTT CCTACTTCAG GGCATTGAT TGATATTATC CCTCACACTT
 16141 TCATCGACAC TTACATAAGT TAATGTTGAT AATACATACG ACTCGTTACC CAGCAAGCCG
 16201 GGCCTTCACT CCAGCGGGTA AGGGGTTCTC TTTTTTTT TTCTTCACT TGACTTTCA
 16261 AAGTGCATCC AGCCAAACAGA GACGTTAAA GGGGGAGCAC TTTTCTTGCA CGCGCCGTAC
 16321 ATAGTATCCA TGTAGGGAGT CTTTATTAAA AAAATAACAT TAAGTGTAT CATGTGCTA
 16381 AAGATTGCT CATTCTCTTCT ATCATCCCA GGATACCCCC TTTTTTGCG CGAAAAACCC
 16441 CCCTACCCCC CTAAACCCCT GAAGTGTCTA AGACCCCTGA AAACCCCCCG GAAACAGGAC
 16501 AAACCTCTGG TAGCTCAGAA AAGGCAATT CTTTATACCT ACTTTCAAC CCACAAATAC
 16561 CCCTGGCTTA TCACAAACCC ACCAGTTGTT TTAATAATAC CACTTTCAA CCCACAAATA
 16621 CCCCTGGCTT ATCACAACCC CACCAAGTTG TTTAATAATA CCACCTTTCA ACCCACAAAT
 16681 ACCCCTGGCT TATCACAACCC CCACCAAGTTG TTTAATAAT ACCACTTTTC AACCCACAAA

16741 TACCCCTGGC TTATCaCAAC CCCaCcAGTT GtTTTAATAA TACCACTTT CAACCCACAA
 16801 ATACCCCTGG CTTATCACAA CCCCCACCAGT TGTTTTAATA ATACCACTTT TCAACCCACA
 16861 AATACCCCTG GCTTATCAC ACACCCACAG TTGTTTTAAT AATACCACTT TTCAACCCAC
 16921 AAATACCCCT GGCTTATCAC AACCCCACCA GTTGTTTTA TAATACCACT TTTCAACCCA
 16981 CAAATACCC TGGCTTATCA CAACCCCAC AGTGTTTTA ATAATACCAC TTTCAACCC
 17041 ACAAATACCC CTGGCTTATC ACAACCCCAC CAGTTGTTTT ATAATACCA CTTTCACCC
 17101 CACAAATACC CCTGGCTTAT CACAACCCCAC CCAGTTGTTT TAATAATACC ACTTTCAAC
 17161 CCACAAATAC CCCTGGCTT TCACAAACCC ACCAGTTGTT TTAATAATAC CACTTTCAA
 17221 CCCACAAATA CCCCTGGCTT ATCACAACCC CACCGTTGTT TTAATAATAA CCACCTTCA
 17281 ACCCACAAAT ACCCCCTGGCT TATCACAACCC CACCGAGTT TTTAATAAT ACCACTTTC
 17341 AACCCACAAA TACCCCTGGC TTATCACAAAC CCCACCAAGTT GTTTAATAA TACCACTTT
 17401 CAACCCACAA ATACCCCTGG CTTATCACAA CCCCCACCAAGT TGTTTTAATA ATACCACTTT
 17461 TCAACCCACA AATACCCCTG GCTTATCAC ACACCCACAG TTGTTTTAAT AATACCACTT
 17521 TTCAACCCAC AAATACCCCT GGCTTATCAC AACCCCACCA GTTGTTTTA TAATACCACT
 17581 TTCAACCCACA AAACATACCGC AAACCCACAA ACGCCCTGTC ACCTTCTTCA TGATCAAAAA
 17641 ACTGCTGTTT AGATTATTTC AAGTATTCT tTCACCACCC TAACTACGGA AGGCACAAAAA
 17701 CAAAAATCAA GAAAATGCT CCTATTAAAT ACCTTTTTT AACCAGATAGA CCCTTAAATC
 17761 CTCTTGAAAAA CCAGCCTTT AACGCCCAA AGGCTAAACA CAAACAAAGT ACCACCTCTA
 17821 GATGTGGCTA GAAAATTAC T

Hippoglossus stenolepis 2 17841 bp

1 GCTGACGTAG CTTAATTAAA GCATAAACACT GAAGATGTTA AGATGGGCC TAGAAAGGCC
 61 CGGGAGCACA AAGGCTTGGT CCTGACTTTA CTGTCGACTT TAACCAAAC TACACATGCA
 121 AGTATCCGCC CCCCTGTGAG AATGCCACA ACTCCCTGCT TGGGAACCTAG GAGCCGGTAT
 181 CAGGCACAG CCCAGCTAGC CCACGAGCC TTGCTTAGCC ACACCCCTCAA GGGAACTCAG
 241 CAGTGATAAA TATTAAGCCA TAAGTGAAAA CTTGACTTAG TTAAGGTTAA GAGGGCCGGT
 301 AAAACTCGTG CCAGGCCACCG CGGTTATACG AGAGGCCAA GTTGACAAAC AACGGCGTAA
 361 AGAGTGGTTA GGGGATTTC TAAACTAGAG CGAACAGCTT TCAAAGCTGT TATACGCACC
 421 CGAAAGTATG AAACCCAAATT ACGAAAGTAG CTCTACTCAT CTCGAACCCCA CGAAAGCTAA
 481 GAAACAAACT GGGATTAGAT ACCCCACTAT GCTTAGCCCT AAACATCGAT TGCACTATAC
 541 ACTCCATATC CGCCCGGGAA TTATGAACAT CAGTTAAAAA CCAAAGGAC TTGGCGGTGC
 601 TTAACATCCA CCTAGAGGAG CCTGTTCTAG AACCGATAAC CCCCGTTAAA CCTCACCTTC
 661 TCTTGTNTTA TCCGCTATA TACCAACCGTC GTCAAGCTTAC CTCGTGAAGG CTTTACAGTA
 721 AGCAAAATTG GCACAGCCC AAACGTCAAG TCGAGGTGA GTGAATGAGG AGGGAAAGAAA
 781 TGGGCTACAT TTGCTAAACA TAGCAACAC GAATGTTGCA TGAAACATG TAACTGAAGG
 841 AGGATTCTGC AGTAAGCAGG AAATAGAGCG TCCCGCTGAA ACTGGCCCTA AAGCGCGCAC
 901 ACACCGCCCG TCACCCCTCC CAAGGCCCTT AAACCAAAC TATTAAAACC CTGCAACCCG
 961 CGAAGGGGGAGA GAAAGTCGTA ACATGGTAAG TGACCGGAA GTGCGCTTGC GAAAATCAG
 1021 AGTGTAGCTA AGATAGAAAT CAGCATCTCA CCTACACCGA GAAGACGTCC GTGCAAGTCG
 1081 GATCACCTG ACGCCATTAA GCTAGGCCAA CCCCTTAACA CAACAAACCC CCATTATAAA
 1141 CCCCTAAAGC ACGAACACCC CACGTAGCTA AACCATTCCTC CCCCTTAAGT CCAGGCGATA
 1201 AAAAAGGAAA TTTTGGAGCA ATAGAAAAAG TACCGCAAGG GAAAGCTGAA AGAGAGATGA
 1261 AAAAGCCAG TAAAGTTAA GAAAGCAGAG CTTAGTGCTC GTACCTTTTG CATCATGATT
 1321 TAGCTAGCAC TTTCAAGCCAA AGAGAACCTA GAGTTTGAA CCCCAGAAACT GAGTGAGCTA
 1381 CTCCAAGACA GCCTATTAT AGGGCGAAC CGTCTCTGTG CAAAGAGGT GGGAAAGAGCT
 1441 TTGAGTAGAG GTGACAAAC TACCGAACTC AGTTATAGCT GTTGCCTGT GAATTGAATA
 1501 GAAGTTCAGC CCCCTGATT CTCCACTCGT GCACCTTATT AACCCCTTCAG ATGCAATGAG
 1561 GAACCAGGGG TGTTAGCTAA AGGGGGTACA GCCCCTTGA TACAAGACAC AACTTTCTCA
 1621 GCAGGATAAA GATCATATTCA AAATAAGGAC AGATGTTTA GTGGGCCTAA AAGCAGCCAC
 1681 CTTAATAGAA AGCGTTAAAG CTCAAACATA AAGGCCCTCC GTTATACCGA TAACCCCTATC
 1741 TTATCCCCC ACATTTAACCA GGCCCTCTA TGACCCACAT AGGAACGACT ATGCTAATAT
 1801 GAGTAATAAG AGGGTATAAT CACCCCTCTA TTGACACATGT GAAATCGGA ACGGACCCCC
 1861 CACCGAACATCT TAACGGCCC AATCAAAGAG GGTATTGGAA ACTACCAAA ATTAGGCCA
 1921 GAAAACATC CAATGCAAAAC CCGTTAACCC CACATGGTG TGCCCAAAG GAAAGACCAA
 1981 AGGGGGGAGA AGGAACCTGG CAAACATACC CCAAGCCTCG CCTGTTTACCA AAAACATCG
 2041 CCTCTTGCAT AACACAGTA TAAGAGGTCC CGCCTGCCA GTGACCATAT AGTCAACCG
 2101 CGCGGTATT TTGACCGTGC AAAGGTAGCG TAATCACTTG TCTTTAAAT GAAGACCTGT
 2161 ATGAATGGCA TAACGGAGGC TTAACTGCTC CCTTCCCCCG GTCAATGAAA TTGATCTCCC
 2221 CGTGCAGAAC CGGGGATTAA ACCATAAGAC GAGAAGCACCC TATGGAGCTT TAGACACACA
 2281 GGTGGCCCAT GTCAAATACC CCCAGCTAAG GGCTGTAACCT AAATGGAATC TGCCTTGATG
 2341 TCTTCCGGTTT GGGCGACCAT GGGGAATACA AAACCCCCCAC GTGAAAGGG ACCACACCC
 2401 TAAGTTACTT CTTCTCCCG AAGCCAGAGC AACAGCTCTA ACCAGCAGAA ATTCTGACCA
 2461 AACTGATCCG GTAAAACCGA TCAACCGAACC AAGTTACCCCT AGGGATAACA GCGCAATCCC
 2521 CTTTTAGAGC CCATATCGAC AAGGGGTTT ACGACCTCGA TGTTGGATCA GGACATCTA
 2581 ATGGTGCAGC CGCTATTAAAG GGTTCGTTT TGCAACGATT AAAGTCCTAC GTGATCTGAG
 2641 TTCAGACCGG AGTAATCCAG GTCAGTTCT ATCTATGACA TGATCTTTTC TAGTACGAA
 2701 GGACCGAAAAA GAAGAGGCC CTGCTAAAG TACGCCCTCAC CCCCACCTAT TGAAAAAAATC
 2761 TAAACTAGGC AAAACGGGCT AACCCTTTG CTGGAGATAA CAGCAAGTTT GGGTGGCAGA
 2821 GCCCCGGCTAA TGCAAAAGAC CTAAGCCCTT TCCACAGAGG TTCAAATCTC CTCCCTAACT
 2881 ATGATTCAA CTCTCATTAC GCATATTATC AACCCACTAA CCTTTATTGT ACCTGTATTG
 2941 CTAGCCGTAG CATTCTTAC CCTCCTGAA CGAAAAGTGC TAGGCTACAT ACAACTCCGA
 3001 AAAGGGCCAA ACATCGTAGG GCCTTACGGC CTCCCTCAAC CTATTGCTGA TGGTGTAAAA
 3061 CTCTTATTAA AAGAACCCAT TCGACCTTCG ACCGCATCAC CGTCTCTATT CCTCCTGCC
 3121 CCTATGCTCG CTCTCACACT AGCCCTAACCT CTTGAGGCC CCATACCTT CCCGTACCC
 3181 GTCGTGGACC TTAACCTAGG TATTATTATT ATTCTAGCCC TATCTAGCCT TGCAGTATAT
 3241 TCTATTCTGG GCTCCGGCTG AGCATCTAAAT CAAAATATG CCCTAGTAGG GGCACATGCC
 3301 GCTGTCGCAC AAACATTTC TTACGAAGTT AGCCTTGGAC TCATCTTGCT AAACATTATC
 3361 ATCTTACCG GAGGCTTCAC ACTTCAAACAG TTTAACACAG CCCAAGAGGC CATTGACTA

3421 GTAGTGCCTCG CTTGACCACT GGCTGCCATA TGATATATCT CCACTCTTC CGAAACAAAC
 3481 CGTGCACCCCT TCGACTTAAC AGAAGGAGAG TCTGAGCTTG TCTCAGGGCTT CAACGTAGAG
 3541 TAGCGAGCG GACCCTTCGC CTTATTCTTC TTGGCCGAAT ATTCAAATAT TCTCCTAATA
 3601 AATACCCCTGT CCGCTACACT ATTTTTAGGG GCCTCCCACA TTCCATCTAT CCCAGAATTA
 3661 ACCAGTATCA ATATTATGAC TAAAGCAGCC CTTCTGTCAA TTGTCTTCCT ATGAGTCCGA
 3721 GCTTCGTACC CGCGGTTCCG TTACGACCA CTTATGCACC TCATTTGAAA AAACTTCTT
 3781 CCACTGACTT TGGCACTAGT TATCTGACAC TTGGCGCTCC CTATTGCATT CGCTGGTCTC
 3841 CCACCCAAAC TATAGGCCA GGAGCTGTGC CTGAaTTTAA AGGGCCACTT TGATAGAGTG
 3901 AATTATGGG GTTAAAGTCC CCCCAACTC TTAGAAAGAa GGGGATCGAA CCCAACCTGA
 3961 AGAGATCAA ACTCTTCGTG CTTCTCTAC ACCACTTCCT AGTAAGGTCA GCTAAATAAG
 4021 CTCTGGGCC CATACCCCAA CCATGTAGGT TCAAATCCCTT CTTTACTAA TGAACCCCTA
 4081 CATCTTGACC ACTCTTCTAT TTGGTTGGG CCTAGGAACA ACACTCACGT TTGCAAGCTC
 4141 ACACCTGGTC CTCGCTGAA TAGGACTTGA GATAATACA CTAGCCATTAA TTCCCCTAAT
 4201 AGCCCAACAT CACCACCCCC GGGCAGTCGA AGCTACCACT AAGTACTTTTC TCGCACAAGC
 4261 CACAGCAGCC GCCACCCCTC TGTTTGCAG TACTACCAAC GCTTGACTCA CAGGCCAATG
 4321 GGACATTCAA CAGATAACAC ATCCCCCTCC AACAAACAATG ATTGTTATTG CCCTCGCACT
 4381 AAAAGTCGGA CTGGCACCA TGCACTCTTG ACTTCCCAGA GTACTCCAAG GACTAGACCT
 4441 AACCAACCGGA CTTATCCTC CAACCTGACA AAAACTTGCA CCCTTTGCC TCTTACTACA
 4501 AATTCAAACA ACCAACCTA CGCCTCTAAT CATTATTGGC TTACTATCCA CGCTCGTGG
 4561 TGGCTGAGGT GGTCTAACCC AAACCCAGTT ACGCAAAGTC CTCGCCTACT CCTCAATTG
 4621 CCACCTGGC TGAATAATGC TGATCGTTCA ATTTTCACCC CTCCCTGCC TTCTCACCC
 4681 CCTCACGTAC TTTACTATGA CTTTCTCAGC ATTCCTAATT TTTAAAGTAA ATAAAGCCAC
 4741 CACTGTCAT GCCCCTCGAA TCTCGTGAAC AAAGACTCTC GCCCCTCACAG CCCTAGCACC
 4801 CCTGGTTTA CTATCCCTCG CGGCCCTTACCC CCCCCCTTACCC GGCTTCATGC CCAAGTGATT
 4861 TatCTTCAG GAACTGACTA AGCAAGACCT TCCAGCACTT GCCACCCCTG CCGCTTTAAC
 4921 TGCCCTTGT AGTCTCTACT TCTACCTCCG TCTCTCGTAT GCAATAACCC TCACAATATT
 4981 CCCCAACAAAC CTCATTGGCG TCACTCCCTG ACGATTCTAC TCCCCCCTAAT TTACGCTTCC
 5041 CCTCGCCGTC TCCACTGCAG CAACCACCC TCTCTCTCCA CTAGCCCCCG CTGCCGTAGC
 5101 ACTACTGGT ACCTAAAGGA CTTAGGCTAG CACTTAGACCC AATGGCCTTAA AGAGCCGTAA
 5161 GCGTGAAGTGA AAATCTCTCA GTCCCTGTTA AGACTTGGG GCTACTATCC CACATCTTCT
 5221 GCATGCAAG CAAACACTT AATTAAGCTA AAGCCTTTCT AGATAGGAAG GCCTTGATCC
 5281 TACAAAATCT TAGTAAACAG CTAAGGCCA AATCCAGCGA GCATCTATCT ACTTTTCCCC
 5341 GCCTAATTAA GAAACTAAAGG GCGGGGGAA AGCCCCGGCA GGCAGTTAGC CTGCTTCTT
 5401 AGATTTGCAA TCTAACATGT AACACCCCGG GGCTTGGTAA GAAGAGGGCT TGCACCTCTG
 5461 TCTATGGTC TACAATCCAC CGCTTAACCTC AGCCATCCTA CCTGTGGCAA TCACACGTTG
 5521 ATTTTTCTCG ACCAACTACA AAGACATCGG CACCCCTCTA CTCGTATTG GTGCCGTGAGC
 5581 CGGAATAGTG GGGCAGGCC TAAGTCTGCT TATTCGGCA GAACTAAGCC AACCCGGGGC
 5641 TCTCCTGGGA GACGACCAA TTTATAATGT AATCGTCACC GCACACGCC TTGTAATAAT
 5701 CTTTTTATAA GTAAATACCA TTATGATTGG GGGTTCTGGA ACTGGCTTA TTCCACTAAT
 5761 ATTTGGGCC CCAGATATGG CGTCTCCCTG AATGAATAAT ATGAGTTCT GACTTCTTCC
 5821 CCCCTCCTT CTCCTCCTC TAGCCTCTC AGGTGTTGAA GCCGGAGCAG GTACCGGATG
 5881 AACCGTATAC CCCCCACTAG CTGGCAATT TAGCCCACGCC GGAGCATCCG TAGACCTGAC
 5941 AATCTCTCA CTTCACCTG CAGGAATTTC ATCAATTCTG GGGGCAATTAA ACTTTATTAC
 6001 TACTATCATT AACATGAAAC CCACACAGT CACTATGTC CAAATCCCCCT TATTGTTTTG
 6061 AGCCGTTCTT ATTACAGCCG TACTTCTTCT TCTATCCCTG CCGGTTTTAG CCAGGAAT
 6121 TACAATACTA CTAACAGACC GCAACCTAA CACGACCTTC TTGACCCCTG CCGGAGGAGG
 6181 TGACCCCATC CTCTACCCAGC ACCTATTCTG ATTCTTTGGT CACCCAGAGG TATACTTCT
 6241 TATTCTCCCG GGCTTGGAA TAATTCTCA CATTGTTGCA TACTATGCAG GTAAGAAAGA
 6301 ACCTTTGGC TACATGGGA TAGTCTGAGC TATAATGCC ATTGGACTCC TGGGGTTCAT
 6361 TGTCTGGCC CATCACATAT TTACAGTCGG AATAGATGTA GATAACAGAG CCTACTTTAC
 6421 CTCTGCCACA ATAATCATG CGATTCCAAC TGGCGTAAAAA GTCTTTAGCT GACTTGCAC
 6481 CCTCCATGGGA GGAAGCCTTA AATGAGAAAC ACCCCCTCTA TGAGCCCTCG GCTTTATTT
 6541 CCTCTTACA GTAGCGGCT TCACTGGCAT TGTCTTAGCT AACTCCTCTC TCGATATTG
 6601 TCTCGATGAC ACATACTATG TAGTCGCTCA CTTCACCTAT GTACTATCTA TGGGTGCTGT
 6661 ATTTGCAATC GTTGCCTCCT TCGTCCACTG GTTCCATTAA TTTACAGGCT ATACCCTCTA
 6721 CTCCACATGA ACAAAAATCC ACTTCGGCT GATATTATT GGAGTCAATC TAACATTCTT
 6781 CCCCCAACAT TTTCTGGCC TGGCGGAAAT ACCCCGACGG TACTCAGACT ACCCAGATGC
 6841 ATACACCCCTT TGAAACACTG TTTCAAT CCGGTCCCTA ATGTCCTCG TTGCTGTAAT
 6901 TTTATTTTTA TTCATTATT GAGAACCTT TACAGCCAAA CGAGAAGTCC GAGCAGTAGA
 6961 ACTAAACTGCA ACTAAACATT AATGACTTTA CGGCGTCCCT CGCCCTTACACACATTG
 7021 AGAGCCCGCA TTGCTACAG TTGCTATAAA TTGCAACAG CTAACGAGAA AGGGAGGAGT
 7081 TGAACCCCCA TCAATTGGTT TCAAGGCAAC CACATAACCG CTCTGTCACT TTCTTCACAA
 7141 CACACCAATA AGATACTAGT AAAACGCTAT AACACTGCC TGTCAAGACA GAATTGTTGG
 7201 TtCAACCCCC GCGTATCTTA AACGCAATGG CACATCCCTC ACAACTCGGA TTCAAGACG
 7261 CAGCTTCACC CTTAATAGAA GAAACTCTTC ACTTCACGA CCACGCCCTTA ATAATTGTA
 7321 TTCTCATTAG CACAATAGTA CTTTATATT TTGCGGCTAT AGTCACAGCC AAAACTAACAG
 7381 ATAAACTTGT ACTAGACTT CAGAACATTG AAATTATCTG AACAGTTCTC CCAGCTGTTA
 7441 TTCTTATCTT TATCGCCCTA CCATCCATCC GAATTTTGTA CTTAATAGAC GAAATTAAATG
 7501 ACCCCCCACCT GACAATTAA GCCATTGGCC ACCAGTGATA CTGAAGCTAC GAATACACAG
 7561 ACTACGAAGA CCTCGGTTT GACTCATACA TAATCCCCAC ACAAGACCTG ACCCCCTGGAC
 7621 AATTCCGACT ACTAGAACCA GACCACCGAA TAGTAATCCC GGTAGAATCC CCAATTCCG
 7681 TTTTAATCTC GGCTGAAGAT GTCTTACATT CCTGGGCAGT CCCCTCTCTA GGCCTAAAG
 7741 TTGACGCCGT ACCTGGACGA TTGAACCAAG CAACTTTAT TGTTAGCCGA CCAGGCCTAT
 7801 TTTTCGGACA ATGTTCTGAG ATTGCGGGG CTAAACCATAG TTTTATACCC ATTGTTGTAG
 7861 AAGCAGTCCC CCTTGAAACCT TTGAGAAACT GGTCTTCACT ATAATTGAA GACGCCCTCG
 7921 TAAGAAGCTA ATAAGTACAA CGCTTAGCCT TTTAAGCTAA AGACTGGTGC CTAACAACCA
 7981 CCCTTAGCGA CATGCCCTAA CTGAACCCCG CACCCCTGATT TGCAATTAA GTTTTCTTCTT
 8041 GAATAATCTT TTTAATGTT ATTCCCCCA AAGTTTAGC ACACACCTAT CCTAATGAAC
 8101 CCACCTCCCA AAGCACACAA AAACCTAAAA CAGGCCCTG AAACGTGACCA TGATACTAAG

8161 CTTCTTTGAC CAATTTATAT CCCCTTCATA CCTCGGCATC CCACTAATTG CACTAGCGAT
 8221 TAGCCTGCCCT TGAATTCTGT TCCCAACCCC ACAAGCCCG TGATTAACA ACCCGCTACT
 8281 AACACTCAA GGGTGATTAA TTAGCCGTT CACCTCGCAA CCTTCTTCTAC CCCTAAACCC
 8341 CGGAGGCCAC AAGTGGCAG TTCTATTCAAC CTCGTTAATA GTTTCTCTAT TTTCTATCAA
 8401 CATGCTAGGA CTCCCTCAT ACACCTTCAC ACCAACGACC CAACTCTCCC TTAATATAGG
 8461 CCTCGCAGTA CCCCTTGAT TAGCAACTGT CATATCGGG ATACGAAATC AACCAACGCA
 8521 TGCCCTAGGC CACCTTCCTC CAGAAGGCAC CCCCACTGCC CTCATCCCTG TACTTATTAT
 8581 CATTGAAACA ATTAGCCTAT TCATTGCGACT CTCGCTCTC GGTGTCGAC TAACAGCAAA
 8641 TCTCACAGCA GGCCACCTCC TTATTCAAGCT CATTGCAAAAC GGTGGCTCTC TCCTCTCTCA CCCTACTAGA
 8701 CCTCATACCT GTTGTGCTA CTTAAACAC GGTGGCTCTC TCCTCTCTCA CCCTACTAGA
 8761 AGTAGCCGTT GCCATGATTC AAGCTTATGT CTTGTGCTA CTCCTAAGGC TATATCTACA
 8821 AGAGAACGTT TAATGGCCA TCAAGCACAC CCCTACCAACA TAGTCGACCC CAGCCCATGA
 8881 CCCCTAACAG GGGTATTGC TGCCCTATTG ATAACATCTG GCCTTGCTAT CTGATTCCAC
 8941 TTCCACTCCA CAACCTGAT AACTATTGGG ACAATCCCTC TCACCTTAAC AATCTTCAA
 9001 TGATGACGGG ACGTGCTACG GGAAGGCACCA TTTCAAGGAC ACCACACTCC CCCCGTTCAA
 9061 AAAGGCCGAT GATAATGGTAT AATTATTATT ATTACCTCTAG AAGTCTTATT CTTCTTAGGC
 9121 TTCTCTGAG CTTTTTACCA TTCAAGCCTA GCACCTTACCG CTGAGCTAGG CGGCTTCTGA
 9181 CCACGGCAG GCATTACTCC CTTAGACCCCC TTTGAAGTCC CACTTCTTAA CACAGCAGTC
 9241 CTGCTTGCCT CTGGCGTAAC TGTTACCTGA GCACACCACA GCATTATAGA GGGTAAACGA
 9301 AAACAAACTA TTCAATCTCT TGCTCTTACA ATCTTACTCG GGGGGTACTT CACTTTCTT
 9361 CAAGGCCAG AGTATCACGA AGCCCCCTTC ACCATTGCAG ACGGAGTTA CGGCGCTACA
 9421 TTCTTGTGAG CCACCGCCTT CCACGGACTA CACGTCTTAA TTGGCACGTC ATTCTTAGCC
 9481 GTTGTCTAT TACGCCAAT TCTTCACCAT TTACATCAAC ACCACCACTT TGGGTTGAA
 9541 GCAGCCGCTAT GATACTGACA CTCGCTAGAC GTCTGCTGAC TTTCTCTTA TATCTCTATT
 9601 TACTGATGAG GATCTTAACT TTCTAGTAT TAAATCTAGT ATAAGTACT TCCAATCACC
 9661 CGGTCTTGGT TAAAATCCAA GGGAAAGATAA TGAGCCTCT TCTAACCATC aTTTCGATTG
 9721 CCACCCCTCCT CTCGACAGTA CTTGCCCTTG TGCCCTTTG ATTACCCCAA ATTACACCAG
 9781 ACCACGAAAA ACTATCACCT TACGAGTGGC GCTTGTACCC GATAGGATCC GCGCGACTGC
 9841 CTTTTTCACT GCGGTTTTC CTCATCGCCA TCCTCTTCTC TCTATTGAT TTAGAAATTG
 9901 CCTCTCTCTC CCTCTTCTCA TGAGGAGAC AACTAGCATC ACCACTACTG ACATTCAACCT
 9961 GAGCTACAGC TGTCTTAGC CTTCTAACCC TTGGCCTCT TTACGAATGA ATGCAAGGGG
 10021 GTCTAGAATG GGCGGAATAG GTGGTTAGTC TAAGAAAAAA ATTGTGATTC GGCTCAAAAAA
 10081 CTTGTGGTT AAATCCCAA CCGCTTAATG ACCCCCCACAC ACTTTGCCTT CTCCCTAGCC
 10141 TTTCTCTGG GTTTAACAGG CCTGGCATTC CACCGGTTTC ACCTCCTTTC CGCTCTATTG
 10201 TGCTTGAAG GAATAATACT ATCCCTATT ATTGCTCTT CCCTATGAAC GCTCCAACCTT
 10261 GATTCAACCA GCTTTTCAGC ATCTCCAATA CTTCTGCTTG CATTTCAGC CTGCGAAGCA
 10321 AGCGCCGGTC TCGCCCTATT AGTAGGCACT GCTCGAACCC ACGGGACCGA CCGATTACAA
 10381 AGCTTAAATC TTCTCCAATG CTAAAAGTC TAATCCCAAC ACTTATGCTA ATCCCAACAG
 10441 CCTGGTTACT CAAACCCAC TGACTCTGAC CCATGACTTT ATGTATAGC TTTTGATCT
 10501 CCTGGTCAG CCTTTCGTGA CTAAAAAACCC TCTCAGAAC CGGCTGATCG TCACTTAGCC
 10561 TCTTTATAGC TACCGACTCC CTATCGACCC CCCTCCTCGT TCTTACATGC TGATTACTCC
 10621 CACTATAAT TTTGGCAAGC CAGAAACACA CAGCCTTACA ACCTCTTAGT CGCCAGCGCA
 10681 TGTACATCAC ACTTCTCGCC TCACTCCAGT TCTTCTTAAT CCTAGCATT AGCGCCACCG
 10741 AACTGGTAAT GTTTTACGTA ATATTGAAAG CCACCCCTAT CCCCACGCTA ATCATTATCA
 10801 CCCGCTGAGG AAACCAAACA GAACGCTAA ACGCAGGAAC CTACTTTCTC TTCTATACAT
 10861 TAGCGGGCTC ACTCCCCCTT CTCGTTGTT TACTCCTACT CAAAACACA TCCGGCACCC
 10921 TATCTGTGTT AACCGTTCAC TACACGGACC CGCTCGCTCT CTCGCTTTAT GCAGACAAAC
 10981 TATGATGAGC AGGCTGTCTT TTAGCATTCC TGGTTAAAAT ACCCCCTCTAC GGAGTCCACC
 11041 TATGGCTCC TAAAGCCAT GTGAAGCTC CAATTGCAGG CTCAATAATC CCTGCAGCAG
 11101 TTCTACTGAA GCTAGGGGA TACGGCATGA TCCGCATGAT AACGATATTG GAACTCTAA
 11161 CTAAGGAATT AAGCTTACCC TTCATGTCT TTGCACTTTG AGGGGTAATT ATAAGTGGCT
 11221 CAATTGGTCT ACGCAGACA GACCTTAAGT CCTTAATTGC TACTCATCA GTCAAGCCACA
 11281 TGGGCTTAGT GGCTGGGGGA ATTCTTATTC AATCACCCCT AGGCTTAACA GGAGCCCTCA
 11341 CCCTCATATAAT TGCACACGGT CTTACTTCAT CAGGCCCTTCT CTGCTTGGCA AACACAAACT
 11401 ATGAACGAAC CCACAGTCGA ACAATGGTCC TGGCGCGAGG GCTTCAAGTA GCCCTACCC
 11461 TAATAGCCAC TTGATGGTT ATTCTCTAGCT TGGCCAACCT AGCTCTGCCA CCACTACCA
 11521 ACCTCATGGG GGAGCTGATA ATTATCACTT CTTTATTCAA CTGGTCTGG TGAACCTGG
 11581 CTTAACCGG TGCAGGCACT CTAATCACAG CTGTTACTC CTTTACATAA TTCTTAATAA
 11641 CTCAACGGGG CCCTCTCCCA ACACATGTCA TTGCACTTGA GCCATCTCAC ACCCGAGAAC
 11701 ACCTCTTAT TGCACCTCAC CTTATCCCT TAATTCTGCT CGTACTTAAG CCCGAGCTGA
 11761 TCTGAGGCTG AACTGCTGT AGGTGTAGTT TAAACAAAAA CATTAGATTG TGATTCTAAA
 11821 AATAAGGGTT AAAATCCCTT CTCCCCACCGA GAGAGGCTCG CAGCAATGAG AACTGCTAAT
 11881 TCCCACGACC TTGGTTGGAC CCCTAGCTC ACTCGAAAAG CTCTTAAAGG ATAACAGCTC
 11941 ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAAATC CAAGTAGCAG CTATGTACCC
 12001 CACCTCCCTA ATGATTCTAT CGAGCTGTAT CACCAATCTT GCATTACTAG CCTATCTT
 12061 AATCACCAAG ATCCGTCAGA CACCTCGAGG CCCTCAATGA GCTACCAACCC ATGTCAAAC
 12121 AGCTGTTAAAAT ATAGCTTCTT TTGTTAGTC GCTACTCTC ACCCTATTTC TTAATGAGGG
 12181 CGCAGAAACA ATCGTTACTA ATTGAACCTG AATAAAACT AACTCTTCA ATATCAACAT
 12241 TAGCTGAAA TTTGATTCTT ACTCAATTAT TTTACACCC ATTGCCCTCT ACGTAACCTG
 12301 GTCCATCTA GAGTTGCAT CATGATACAT GCACGCAGAC CCACACATGA ACCGCTTCTT
 12361 TAAATACCTC CTGACGTTT TGATTGCCAT AATCGTCCTA GTAACTGCAA ACAATATGTT
 12421 TCAACTATTT ATCGGCTGGG AGGGAGTGGG GATCATATCG TTCTTACTTA TTGGATGGTG
 12481 GTTGGACGG GCGGATGCTA ACACGCGAC GCTTCAAGCA GTAATTTACA ACCGAGTAGG
 12541 AGACATCGGA CTTATCTCG CCATGGCTTG AATAGCAGAC AACCTAAACT CCTGAGAGAT
 12601 ACAACAAATC TTGCAACCA CCAAGGACAT AGACTTAACC TACCCCTCTC TCAGACTCAT
 12661 CATTGCGACCA ACTGGCAAAT CGGCTCAATT CGGACTTCAC CTTGACTAC CCTCCGCTAT
 12721 GGAGGGTCCC ACACGGTCT CTGCCCTACT TCATTCTAGC ACTATGGTCC TTGCCGGCAT
 12781 TTTCTGTCA GTACGAATAA GCCCCCTCT GGAAAACAAC CAAACCGCCC TCACAACCTG
 12841 CCTATGCTCTC GGAGCTCTTA CAACACTATT CACTGCAACC TGTGCCCTCA CGCAAAATGA

12901 TATTAACAAA ATCGTCGAT TCTCTACATC CAGCCAGCTT GGGCTTATAA TGGTAACCAC
 12961 TGGTCTAAC ACACACAAAC TAGCATTTCT TCACATCTGT ACACATGCCT TCTTCAAGGC
 13021 CATACTGTT TTATGCTCG GATCTATTAT CCACAGCCTA AATGACGAAC AAGATATCCG
 13081 CAAATAGGG GGCATCACCC ACCTTGACCC TTTCACATCC TCTTGCCTAA CTATTGGCAG
 13141 CCTCGCCCTC ACAGGAACCC CCTTCTGGC TGGCTTCTTC TCTAAAGATG CCATCATCGA
 13201 AGCATTGAAC ACATCCCAC TAAACGCCCTG AGCCCTAACCT TAACCCTCC TAGGCCACCTC
 13261 TTTCACAGCA ATCTACAGCT TTCGAGTAGT ATTTTTGTA CCCATGGGCC ACCCCCCGATT
 13321 TAGTCCCCCTC TCTCCGATCA ACGAAACAA CCCAGCAGTA ATTAACCCAC TTAAGCGACT
 13381 AGCATGGGG AAGTATCGTG CAGGCTACT AATTACCTCT AATATTACAC CTTAAAAAAC
 13441 CCCTGTCTA TCTATGCCG CTCTCCCTAA ACTAGCTGCT TTGAGCTTA CAATCATCGG
 13501 CCTCTTGTT GCTATAGAGC TCGCCATATT AACCAACAAAG CAATATGAGC CAATACCCGG
 13561 CACGCACATA CATCACTTCT CTGCCTTATT AGGCTTCTTC CCCGGGGTCA TACACCGCT
 13621 AACCCCTGTA TTGCTTCT TCCTAGGCCA AAACATTGCT AACCAAACAG TAGATCAGAC
 13681 ATGATTAGAA AAAACAGGCC CCAAAGCCAT TGAATCCGCT ACCCGACCTC TAGTGTCTC
 13741 TACAAGTAAT ATTCAACGGG GGGTGTACAA AACATATCTT ACCCTCTTCC TCCTAACACT
 13801 AGCCCTTATA ATCCCCATAT TCATCACCTA AACCGCTCG AACGGCCCTC GGCTCATACC
 13861 CCGAGTTAAC TCAAGGACTA CAAACAAAGT TAAGAGCAAT ACCCATGAC TAGTGATTA
 13921 CAACCAACCC CCAGAAAGAGT ATATTAAGGC TACCCCTCCG ATATCCCCTC GAAATACGGA
 13981 AAACCTCCCC AACTCATCTG CTAAGACCCA AGAAGATTCA TACCACCCCC CTCAAAACAA
 14041 TCCAGAAACT AATAATACTC CTCCTGTATA GACAACCCCG TAGACCATTAA CCGACCAACT
 14101 ACCCCAGCTC TCTGGGTACG GCTCAGCAGC TAACGCTGCT GAATATGCAA ATACAACCTAA
 14161 CATCCCGCCT AAGTAAATTA AGAACAGGAC TAGTGACAAA AAGGGCCCTC CATGCCAAC
 14221 TAATATACCA CACCCCATGC CAGCTACCAC AACCAAACCC AAAGCAGCAA AATAAGGGGA
 14281 CGGATTAGAA GCAACTGCAA CTAACCCCAA CACAAGAGAA AATAAAACTA AATATATAGC
 14341 GAAAGTCATA ATTCTTGCA GGACTTTAAC CAGGACTAAAT GGCTTGAAAA ACCACCGTTG
 14401 TTATTCAACT ACAAGAACCC TAATGCCAG TCTACGTAAA TCCCACCCCTC TTCTAAAAAT
 14461 CGAACACGAT GCTTTAGTCG ACCTCCCAGC CCCCTCTAAAT ATCTCTGTT GATGAAACCTT
 14521 TGGGCTCTT TTAGGACTCT GTTTAATTAC CCAAATTGCA ACCGGCTTAT TTCTAGCCAT
 14581 ACACATACACA TCAGACATG CTACTGCCTT CACCTCCGTA GCCCCACATCT GTCGAGACGT
 14641 CAACTACGGC TGACTTATCC GAAGCATTCA TGCCAAACGGC GCATCATTCT TTTTCATTG
 14701 CCTCTACCTT CATATTGGTC GAGGTCATCA CTATGGCTATA TACCTCTATA AAGAAACATG
 14761 AACTGTTGGG GTTATTCTCG TCCTCTCGT AATAATAACA GCCTTCTGTT GATACGTCCT
 14821 CCCTTGAGGA CAAATGTCAT TTTGAGGTGC AACCGTTATT ACCAACCTTC TATCTGCCGT
 14881 CCCCTACGTC GGAAACACCC TGGTCCAATG GATCTGAGGT GTTTTTCTG TAGACAATGC
 14941 TACCCCTCACC CGGTTTTG CCTTCCACTT CCTCTTCCCC TTCATCATCG CTGCCGCAAC
 15001 GTTAATTCTAT CTACTCTTC TTCACCGAAC TGGCTCAAAC AACCCAACCG GGCTAAACCTC
 15061 AGACTCCGAC AAAGTTCCCT TTCACCCCTA CTTCACTTAC AAGGACCTCT TAGGCTTGC
 15121 AGTCCTTCTT ACTGCATTAG CATCCCTAGC CCTATTTC CCAAATCTCT TAGGAGACCC
 15181 AGATAACTTT ACCCCCTGCA ATCCGCTGT TACTCCCCCA CACATTAAGC CAGAGTGATA
 15241 CCTCTGTGTT GCCTATGCTA TTCTCCGCTC CATTCACAAAC AAACTTGGAG GAGTACTTGC
 15301 CCTTTTATTC TCAATCTTAG TTCTTATGCT CGTCCCTTC CTTCATACCT CTAAACAACG
 15361 AAGCCTGATA TTCCGTCCTC TAACACAATT TTTGTTCTGA TCTTTAGTAG CAGACGTAT
 15421 GATTCTAATCT TGAATTGGAG GAATGCCGT AGAACACCCCC TTGTTTATCA TTGGACAAGT
 15481 AGCGTCTCTC CTTTACTTCT CCCTCTTCT AGTCTGATC CCAACAGCAG GTGACTAGA
 15541 GAATAAAAGTC CTTGGATGAA AATGCACTAG TAGCTCAGCA CACAGAGCCC CAGTCTTGT
 15601 AACTGGCCGT CGGAGGTTAG AATCCTCCCT ACTGCTCAA GAAAGGGAGAT TTCAACTCCT
 15661 ACCCTTAACCT CCCAAAGCTA GGATTCTAGC ACTAAACTAT TCTTTGTGGT ACATAAAATGT
 15721 ACAATGAAGG ATTTTCATAT ACATGTATGT AATAACACCA TATATTATA GTAACCATT
 15781 TATGTAATGT ACTAGGACAT ACATGTATAA TCACCACATC TCGTAATAGA ACACTCATTC
 15841 ACCACCATT TTAACTAAGA CGAACTAAAA CCTGAATAAT CACTAATCTT AAATAAGTGG
 15901 AAGTCCAGGA CCAGTCGAGA TCTAACAGCG AACACAAACAC TCATCAGTCC AGTTATACCA
 15961 AGACTCAAAAC TCTCGCCGAC CTCAAATCTC TATGTTGAA GAGCCTACCA ACCGGTGATT
 16021 CCTTAATGAT AACTCTTATT GAGGGTGAG GACAAAAAA GTGGGGGTTT CACTCGGTGA
 16081 ATTATTCTG GCATTGGTT CCTACTTCAG GGCATTGAT TGATATTATC CCTCACACTT
 16141 TCATCGACAC TTACATAAGT TAATGTTGAT AATACATACG ACTCGTTACC CAGCAAGCCG
 16201 GGCCTTCACT CCAGCGGTTA AGGGGTTCTC TTTTTTTT CCCTTTCACT TGACTTTCA
 16261 AAGTGCATCC AGCCAAACAGA GACGTTAAA GGGGGAGCAC TTTCTTGCA CGGCCGTAC
 16321 ATAGTATCCA TGTAGGGAGT CTTTATTTAA AAAATAACAT TAAGTGTAT CATGTGCATA
 16381 AAGATTTGCT CATTCTCTT ATCATCCCCA GGATACCCCCC TTTTTTGCG CGAAAAAACCC
 16441 CCCTACCCCCC CTAAACCCCT GAAGTTGCTA AGACCCCTGA AAACCCCCCG GAAACAGGAC
 16501 AAACCTCTGG TAGCTCAGAA AAGGCAATTA CTTTATACCA ACTTTCAAC CCACAAATAC
 16561 CCCTGGCTTA TCACAAACCC ACCAGTTGTT TTAATAATAC CACTTTCAA CCCACAAATA
 16621 CCCCTGGCTT ATCACAACCC CACCAAGTTGTT TTAATAATAA CCACCTTTCA ACCCACAAAT
 16681 ACCCCCTGGCT TATCACAACCC CCACCAAGTTGTT TTAATAATA ACCACCTTTCA ACCCCACAAA
 16741 TACCCCTGGC TTATCACAC CCCCACAGT GTTTAATAA TACCACTTT CAACCCACAA
 16801 ATACCCCTGG CTTATCACAA CCCCACAGT GTTTAATAA ATACCACTTT TCAACCCACAA
 16861 ATACCCCTGG GCTTATCACAA ACCCCACAGT TTGTTTAAAT ATACCACTT TTCAACCCAC
 16921 AAATACCCCT GGCTTATCAC AACCCACCA GTGTTTTAAAT AATACCACTT TTCAACCCAC
 16981 CAAATACCCC TGGCTTATCA CAACCCACCA AGTGTGTTTA ATAATACCA TTTCAACCC
 17041 ACAAAATACCC CTGGCTTATCA CAACCCACCA CAGTTGTTTA ATAATACCA CTTTCAACCC
 17101 CACAAATACC CCTGGCTTATCA CAACACCCCA CCAGTTGTTTA ATAATACCA ACTTTCAAC
 17161 CCACAAATAC CCCTGGCTTA TCACAAACCC ACCAGTTGTT TTAATAATAC CACTTTCAA
 17221 CCCACAAATA CCCCTGGCTT ATCACAACCC CACCAAGTTGTT TTAATAATAA CCACCTTTCA
 17281 ACCCACAAAT ACCCCCTGGCT TATCACAACCC CCACCAAGTTGTT TTTAATAAT ACCACCTTT
 17341 AACCCACAA TACCCCTGGC TTATCACAA CCCACAGT GTTTAATAA TACCACTTT
 17401 CAACCCACAA ATACCCCTGG CTTATCACAA CCCACAGT GTTTTAAATA ATACCACTTT
 17461 TCAACCCACA AATACCCCTG GCTTATCACAA ACCCCACAGT TTGTTTTAAAT AATACCACTT
 17521 TTCAACCCAC AAATACCCCT GGCTTATCAC AACCCACCA GTGTTTTAAAT AATACCACT
 17581 TTCAACCCAC CAAATACCCC AAACCCACAA ACGCCCTGTC ACCTTCTTCA TGATCAAAAA

17641 ACTGCTGTTT AGATTATTC AAGTATTCC TTCACCACCC TAACTACGGGA AGGCACAAAAA
 17701 CAAATCCAA GAAAATGCT CCATTCAAAT ACCTTTTTT AACCGATAGA CCCTTAAGTC
 17761 CTCCCTGAAAAA CCAGCCTTT AACGCCCAA AGGCTAAACA CAAACAGAGT ACCACCTCTA
 17821 GATGTGGCTA GAAAATTAC T

Hippoglossus stenolepis 3 17963 bp

1 GCTGACGTAG CTTAATTAAA GCATAACACT GAAGATGTTA AGATGGGCC TAGAAAGCCC
 61 CGGAGCACAA AAGGCTTGT CCTGACTTTA CTGTCGACTT TAACCAAAC TACACATGCA
 121 AGTATCCGCC CCCCTGTGAG AATGCCACCA ACTCCCTGCT TGGGAACTAG GAGCCGGTAT
 181 CAGGCACAAG CCCAGCTAGC CCACAGGCC TTGCTTAGAC ACACCCCTCAA GGGAACTCAG
 241 CAGTGATAAA TATTAAGCCA TAAGTGAAAAA CTTGACTTAG TTAAGGTTAA GAGGGCCGGT
 301 AAAACTCGTG CCAGGCCACCG CGGTTATAGC AGAGGCCAA GTTGACAAAC AACGGCGTAA
 361 AGAGTGGTTA GGGGATTTC TAAACTAGAG CGAACAGCTT TCAAAGCTGT TATACGCACC
 421 CGAAAGTATG AAACCCAATT ACGAAAGTAG CTCTACTCAT CCTGAACCCA CGAAAGCTAA
 481 GAAAACAAACT GGGATTAGAT ACCCCACTAT GCTTAGGCCCT AACACATCGAT TGCACATAC
 541 ACTCCATATC CGCCCGGAA TTATGAACAT CAGTTTAAAC CCCCAGGAC TTGGCGGTGC
 601 TTAACATCCA CCTAGAGGAG CCTGTTCTAG AACCGATAAC CCCCCTTAAAC CTCACCTTC
 661 TCTTGTGTTA TCCGCTATA TACCAACCGT GCAGCTTAC CTCGTGAAGG CTTTACAGTA
 721 AGCAAAATTG GCACAGCCA AAACGTCAGG TCGAGGTGTA GTGAATGAGG AGGGAAGAAA
 781 TGGGCTACAT TTGCTAAACA TAGCAACAC GAATGTTGCA TTGAAACATG TAACTGAAGG
 841 AGGATTAGC AGTAAGCAGG AAATAGAGCG TCCCGCTGAA ACTGGGCCCTA AAGCGCGCAC
 901 ACACCGCCCG TCACCCCTCC CAAGCCCCCT AAACCAAAC TATTAAAACC CTACAACCCG
 961 CGAAGGGGAG GAAAGCTGTA ACATGTAAG TGACCGGAA GGTGCGCTTG GAAAATCAG
 1021 AGTGTAGCTA AGATAGAAA CAGCATCTA CTTACACCCGA AAAGACGTCC GTGCAAGTCG
 1081 GATCACCTG ACGCCTATTA GCTAGGCCAA CCCCTTAACCA CAACAAACCC CCATTATAA
 1141 CCCCTAAAGC ACGAACACCC CACGTAGCTA AACCAATTCTC CCCCCTAAGT CCAGGCATA
 1201 AAAAAGGAAA TTTTGGAGCA ATAGAAAAAG TACCGCAAGG GAAAGCTGAA AGAGAGATGA
 1261 AAAAGCCAG TAAAGCTAA GAAAGCAGAG CTTAGTGCTC GTACCTTTTG CATCATGATT
 1321 TAGCTAGCAC TTTAACGCAA AGAGAACCTA GAGTTTGAA CCCCAGAAACT GAGTGAGCTA
 1381 CTCCAAGACA GCCTATTAT AGGGCGAACCG CGTCTCTGTG GAAAAGAGT GGGAAAGAGCT
 1441 TTGAGTAGAG GTGACAAAC TACCGAACTC AGTTATAGCT GGTGCGCTGT GAATTGAATA
 1501 GAAGTCAGC CCCCTGGTT CTCCACTCAT GCACCTTATT AACCCCTTCAG ATGCAATGAG
 1561 GAACCAAGGG TGTTAGTCAA AGGGGGTACA GCCCCTTTGA TACAAGACAC AACTTCTCA
 1621 GCAGGATAAA GATCATATTCA AAATAAGGAC AGATGTTTA GTGGGCCTAA AAGCAGCCAC
 1681 CTTAATAGAA AGCGTTAAAG CTCAAACATA AAGCCCTCCC GTTATACCGA TAACCCTATC
 1741 TTATCCCCC ACATTTAACCA GGCCCTCTA TGACCCACAT AGGAACGACT ATGCTAATAT
 1801 GAGTATAAG AGGGTATAAT CACCCCTCTC TTGACACATGT GTAAATCGGA ACGGACCCCC
 1861 CACCGAATCT TAACGGCCC AATCAAAGAG GTTATTGGA ACTACCAACAA ATTAGGCCA
 1921 GAAAACATC CAATGCAAAC CCGTAAACCC CACACTGGT TGCCCAAAG GAAAGACCAA
 1981 AGGGGGGAGA AGGAACCTGG CAAACATACC CCAAGCCTCG CCTGTTTACCC AAAACATCG
 2041 CCTCTTGCAT AACACAGTA TAAGAGGTCC CGCCTGCCA GTGACCATAT AGTTCAACGG
 2101 CGCGGTATT TTGACCGTGC AAAGGTAGCG TAATCACTTG TCTTTAAAT GAAGACCTGT
 2161 ATGAATGGCA TAACGAGGGC TTAACGTCT CCTCCCCCG GTCAATGAA TTGATCTCCC
 2221 CGTCCAGAAG CGGGGATTAA ACCATAAGAC GAGAAGACCC TATGGAGCTT TAGACACACA
 2281 GGTGGCCCAT GTCAAATACC CCCAGCTAAG GGCCTGAAC AAATGGAATC TGCCCTGATG
 2341 TCTTCGGTTT GGGCGACCAT GGGGAATACA AAACCCCCCAC GTGGAAAGGG AGCACACCCCC
 2401 TAAGTTACTT CTTCTCCCG AAGGCCAGAG AACAGCTCTA ACCAGCAGAA ATTCTGACCA
 2461 AACTGATCCG GTAAAACCGA TCAACCGAACC AAGTTACCCCT AGGGATAACA GCGCAATCCC
 2521 CTTTTAGAGC CCATATCGAC AAGGGGTTT ACGACCTCGA TGTTGGATCA GGACATCTA
 2581 ATGGTGCAGC CGCTATTAAAG GGTCGTTTG TTCAACGATT AAAGTCCTAC GTGATCTGAG
 2641 TTCAAGACCGG AGTAATCCAG GTCAAGTTCT ATCTATGACA TGATCTTTTC TAGTACGAAA
 2701 GGACCGAAAAA GAAGAGGCCCTGCTAAAG TACGCCCTAC CCCCACCTAT TGAAAAAAATC
 2761 TAAACTAGGC AAAAGGCAT AACCCCTTTG CTGGAGATAA CAGCAAGTTTG GGGTGGCAGA
 2821 GCCCGGCTAA TGCAAAAGAC CTAAGCCCTT TCCACAGAGG TCAAATCCT CTCCTTAAC
 2881 ATGATTCAA CTCTATTAC GCATATTATC AACCCACTAA CCTTTATTGT ACCTGTATTG
 2941 CTAGCCGTAG CATTCTTAC CCTCCTGAA CGAAAAGTGC TAGGCTACAT ACAACTCCGA
 3001 AAAGGCCAA ACATCGTAGG GCCTTACGGC CTCCCTCAAC CTATTGCTGA TGGTGTAAAA
 3061 CTCCTTATTA AAGAACCCAT TCGACCTTCG ACCGCATCAC CGTCCTATT CCTCCTGCC
 3121 CCTATGCTCG CTCTCACACT AGCCCTAACCT CTTTGAGGCC CCATACCTT CCCGTACCC
 3181 GTCGTGGACC TTAACCTAGG TATTATTT ATTCTAGCCC TATCTAGCCT TGCACTATAT
 3241 TCTATTCTGG GCTCCGGCTG AGCATAAT TCAAATATG CCTCTAGTAGG GGCACCTGG
 3301 GCTGTCGAC AAACAAATTTC TTACGAAGTT AGCCTTGGAC TCATCTTGC AAACATTATC
 3361 ATCTTACCG GAGGCTTCAC ACTTCAAACCC TTTAACACAG CCAAGAGGC CATTGACTA
 3421 GTAGTGCCTG CTTGACCACT GGCTGCCATA TGATATATCT CCACTCTTGC CGAAACAAAC
 3481 CGTGCACCCCT TCGACTTAAC AGAAGGAGAG TCTGAACCTTG TCTCGGGCTT CAACGTAGAG
 3541 TACCGAGGCG GACCCCTCGC TTGTTCTTC TTGGCCGAAT ATTCAAATAT TCTCTAAATA
 3601 AATACCTGTC CCGTACACT ATTATTTAGGG GCCTCCCAA TTCCATCTAT CCCAGAAATTA
 3661 ACCAGTATCA ATATTATGAC TAAAGCGGG CTTCTGTGCT TGTCCTTCC ATGAGTCCGA
 3721 GCTTGTACCG CGCGATTCCG TTACGACCAAA CCTATGACACC TCATTTGAAA AAACCTTCTT
 3781 CCACTGACTT TGGCACTAGT TATCTGACAC TTGGCGCTCC CTATTGACATT CGCTGGCTC
 3841 CCACCAACAC TATAGCGCA GGAGCTGTGC CTGAATTAA AGGGCCACTT TGATAGAGTG
 3901 AATTATGGGG GTTAAAGTCC CCCCAACTCC TTAGAAAGAA GGGGATCGAA CCCAACCTGA
 3961 AGAGATCAAAC ACTCTTCTG CTTCCCTAC ACCACTCTCT AGTAAGGTCA GCTAAATAAG
 4021 CTCTGGGGC CATACCCCAA CCATGTAGGT TCAAATCCTT CTTTACTAA TGAACCCCTA
 4081 CATCTTGACCC ACTCTTCTAT TTGGTTGGG CCTAGGAACAC ACTCTCACGT TTGCAAGCTC
 4141 ACACCTGGCTC CTCGCTTGAA TAGGACTTGA GATTAATACA CTAGCCATTAA TTCCCCTAAT
 4201 AGCCCAACAT CACCACCCCCC GGGCAGTCGA AGTACCAACT AAGTACTTTTC TCGCACAAGC
 4261 CACAGCAGCC GCCACCCCTCC TTGTTGCGAG TACTACCAAC GCTTGACTCA CAGGCCAATG

4321 GGACATTCAA CAGATAAACAC ATCCCCCTCC AACAACAATG ATTGTTATTG CCCTCGCACT
 4381 AAAAGTCGGA CTGGCACCAA TGCACCTTG ACTTCCCAGA GTACTCCAAG GACTAGACCT
 4441 AACACCGGA CTTATCCTCT CAACCTGACA AAAACTTGCA CCCTTGCCC TCTTACTACA
 4501 AATTCAAACAA ACCAACCTCA CGCCTCTAAT CATTATTGGC TTACTATCCA CGCTCGTTGG
 4561 TGGCTGAGGT GGTCTCAACC AAACCCAGTT ACGCAAAGTC CTCGCCTACT CCTCAATTGC
 4621 CCACCTTGGC TGAATAATGC TGATCGTTCA ATTTTCACCC CTCCCTGCCC TTCTCACCC
 4681 CCTCACGTAC TTTACTATGA CTTTCTCAGC ATTCCTAATT TTTAAAGTAA ATAAAGCCAC
 4741 CACTGTCAAT GCCCTCGCAA TCTCGTGAAC AAAGACTCCT GCCCTCACAG CCCTAGCACC
 4801 CCTGGTTTA CTATCCCTCG GCGGCCCTCC CCCCTTACC GCCTTCATAC CCAAGTGATT
 4861 TATTCTTCAG GAACGTGACTA AGCAAGACCT TCCGACTTGC CCGACCCCTCG CCGCTTAAAC
 4921 TGCCCTTGT AGTCTCTACT TCTACCTCCG TCTCTCGTAT GCAATAACCC TCACAATATT
 4981 CCCAACAAAC CTCATTGGCG TCACTCCCTG ACGATTCTAC TCCCCCCTAAT TTACGCTTCC
 5041 CCTCGCCGTC TCCACTGCAG CAACCACCT TCTCTCCCA CTAGCCCCCG CTGCCGTAGC
 5101 ACTACTGGTC ACCTAACGGGA CTTAGGCTAG CACTTAGACC ATGGCCTTC AAAGCCGTAA
 5161 GCGTGAAGTGA AAATCTCTCA GTCCCTGTAA AGACTTGCAG GCTACTATCC CACATCTCT
 5221 GCATCGAAAG CAAACACTT AATTAACTA AAGCCTTCT AGATAGGAAG GCCTTGATCC
 5281 TACAAAATCT TAGTAACTA CTAAGGCC CAACTCAGCGA GCATCTATCT ACTTTTCCCC
 5341 GCCTAATTAA GAAACTAAA GGCGGGAAA AGCCCCGGCA GGCGAGTTAGC CTGCTTCTT
 5401 AGATTTGCAA TCTAACATGT AACACCCGG GGCTTGGTAA GAAGAGGGCT TGCACCTCTG
 5461 TCTATGGGTC TACAATCCAC CGCTTAACCTC AGCCATCCTA CCTGTGGCAA TCACACGTTG
 5521 ATTTTCTCG ACCAATCACA AAGACATCGG CACCCCTCAT CTCGTATTG GTGCCTGAGC
 5581 CGGAATAGTG GGGACAGGCC TAAGCTGCT TATCGGGCA GAACTAAGCC AACCCGGGGC
 5641 TCTCCTGGAA GACGACCAAA TTTATAATGT AATCGTCACC GCACACGCC TTGTAATAAT
 5701 CTTTTTATA GTAATACCAAA TTATGATTG GGGGTTCGGA AACTGGCTTA TTCCACTAAT
 5761 AATTGGGCC CCAGATATGG CGTCCCTCG AATGAATAAT ATGAGTTCT GACTTCTTCC
 5821 CCCCTCCTT CTCCTCCTCT TAGCCTCTTC AGGTGTTGAA GCCGGAGCAG GTACCGGATG
 5881 AACCGTATAC CCCCCACTAG CTGGCAATT TGGCCACGCC GGAGCCTCC TAGACCTGAC
 5941 AATCTCTCA CTTCACCTG CAGGAATTTC ATCAATTCTG GGGGCAATTA ACTTTATTAC
 6001 TACTATCATT AACATGAAAC CCACAAACAGT CACTATGTAC CAAATCCCC TATTTGTTG
 6061 AGCCGTTCTT ATTACAGCCG TACTTCTCT TCTATCCCTG CCCGTTTTAG CCCGAGGAAT
 6121 TACAATACTA CTAACAGACC GCAACCTAA CACGCCCTTC TTTGACCCCTG CCGGAGGAGG
 6181 TGACCCCATC CTCTAACAC ACCTATTCTG ATTCTTTGGT CACCCAGAGG TATACATTCT
 6241 TATTCTCCCG GGCTTGGAA TAATTCTCA CATGTTGCA TACTATGCAG GTAAGAAAGA
 6301 ACCTTTGGC TACATGGGA TAGTCTGAGC TATAATGGCC ATTGGACTCC TGGGTTCAT
 6361 TGTCTGGCC CATCACATAT TTACAGTCGG AATAGATGTA GACACACGAG CCTACTTAC
 6421 CTCTGCCACA ATAATCATG CGATTCCAAC TGGCGTAAAA GTCTTTAGCT GACTCGAAC
 6481 CCTCCATGGA GGAAGCTTA AATGAGAAAC ACCCCTCTA TGAGCCCTCG GCTTTATTTT
 6541 CCTCTTACA GTAGCGGTTC TCACTGGCAT TGTCTTAGCT AACTCCTCTC TCGATATTGT
 6601 TCTCGATGAC ACATACTAT TAGTCGCTCA CTTCCACTAT GTACTATCTA TGGGTGCTGT
 6661 ATTTGCAATC GTTGCCTCT TCGTCCACTG GTTCCATTAA TTTACAGGCT ATACCCTCTA
 6721 CTCCACATGA ACAAAAATCC ACTTCCGCT GATATTATT GGGGTCAATC TAACATTCTT
 6781 CCCCCAACAT TTTCTGGCC TGGCCGAAT ACCCCGACGG TACTCAGACT ACCCAGATGC
 6841 ATACACCCCT TGAAACACTG TTTCATCAAT CGGGTCCCTA ATGTCCTCG TTGCTGTAAT
 6901 TTTTATTATA TTCATTATTG GAGAAGCATT TACAGCCAA CGAGAAGTCG GAGCAGTAGA
 6961 ACTAAACTGCA ACTAACATG AATGACTTTA CGCGTGCCTC CGCCCTTAC ACACATTGTA
 7021 AGAGCCCGA TTCTGACAAG TTCTGATAAA TTGAGAACAG CTAACGAGAA AGGGAGGAGT
 7081 TGAACCCCCA TCAATTGGT TCAAGGAAAC CACATAACCG CTCTGTACT TTCTTCACAA
 7141 CACACCAATA AGATACTAGT AAAACGCTAT AACACTGCCT TGTCAAGACA GAATTGTGGG
 7201 TTCAACCCCC GCGTATCTA AACGCAATGG CACATCCCTC ACAACTCGGA TTCAAGACG
 7261 CAGCTTCACC CTTAATAGAA GAACTACTTC ACTTCCACGA CCACGCCCTA ATAATTGTA
 7321 TTCTCATTAG CACAATAGTA CTTTATATAA TTGCGGCTAT AGTCACAGCC AAACTAACAG
 7381 ATAAACTTGT ACTAGACTCT CAAGAAATTG AAATTATCTG AACAGTTCTC CCAGCTGTTA
 7441 TTCTTATCTC TATGCCCTA CCATCCATCC GAATTGGTAA CTAAATAGAC GAAATTATG
 7501 ACCCCCCACCT GACAATTAA GCCATTGGCC ACCAGTGTAA CTGAAGCTAC GAATACACAG
 7561 ACTACGAAGA CCTCGGTTT GACTCATACA TAATCCCCAC ACAAGACCTG ACCCCTGGAC
 7621 AATTCCGACT ACTAGAACG GACCACCGAA TAGTAATCCC GGTAGAATCC CCAATTCCG
 7681 TTTTAATCTC GGCTGAAGAT GTCTTACATT CCTGGGCAGT CCCCTCTCTA GCGTAAAAG
 7741 TTGACGGCGT ACCTGGACGA TTGAACCAAG CACCTCTTAT GTTAGCCGA CCAGGCGTAT
 7801 TTTTCGGACA ATGTTCTGAG ATTTGCGGGG CTAACCATAG TTTTATACCC ATTGTTGTAG
 7861 AAGCAGTCCC CCTCTAACAC TTTGAGAACT GGTCTTCACT AATAATTGAA GACGCCCTCG
 7921 TAAGAAGCTA ATAAGTACAA GCGTTAGCTT TTTAAGCTAA AGACTGGTGC CTAACACCA
 7981 CCCTTAGCGA CATGCCCTAA CTGAACCCCG CACCTGATT TGCAATTAA GTTTTCTCTT
 8041 GAATAATCTT TTTAACTGTT ATCCCCC AAGTTTAGC ACACACCTAT CCTAATGAAC
 8101 CCACCTCCCA AAGCACACAA AAACCTAAAA CAGAGCCCTG AAACGTGACCA TGATACTAAG
 8161 CCTCTTGTAC CAATTATAT CCTCCCTCATC CCTCGGCATC CCACTAATTG CACTAGCGAT
 8221 TAGCCTGCCT TGAATTCTGT TCCCAACCCC GCAAGCCGC TGATTAACCA ACCCGTACT
 8281 AACACTTCAA GGGTGAATTG TTAGCGCTT CACCTCGCAA CCTCTCTAC CCCTAAACCC
 8341 CGGAGGCCAC AAGTGGGAG TTCTATTAC CTCGTTAATA GTTTTCTCAT TTTCTATCAA
 8401 CATGCTAGGA CTCCCTTCCAT ACACCTTCAC ACCAACGACC CAACTCTCCC TTAATATAGG
 8461 CCTCGCAGTA CCCCTTGAT TAGCAACTGT CATTATCGGG ATACGAAATC AACCAACGCA
 8521 TGCCCTAGGC CACCTTCTTC CAGAAGGCAC CCCCACTGCC CTCATCCCTG TACTTATTAT
 8581 CATTGAAACA ATTAGCTTAT TCATTGACCC CTCGCTCTC GGTGTCCGAC TAACAGCAAA
 8641 TCTCACAGCA GGCCACCTCC TCATTCAACT CATTGCAACA GCTGCCTTCG TCCTTCTTCC
 8701 CCTCATACCT GTTGTGCTA TCTTAACAAAG GTGGTCCCTC TCCCTCCTCA CCCTACTAGA
 8761 AGTAGCGTT GCCATGATC AAGCTTATGT CTTGTTCTA CTCCTAAGCC TATATCTACA
 8821 AGAGAACGTT TAATGCCCA TCAAGCACAC CCCTACCCACA TAGTCGACCC CAGCCCATGA
 8881 CCCCTAACAG GGGTATTGC TGCCCTATTG ATAACATCTG GCCTTGCTAT CTGATTCCAC
 8941 TTCCACTCCA CAACCTGAT AACTATTGGG ACAATCCTTC TCACCTTAAC AATCTTCAA
 9001 TGATGACGGG ACGTGTCAGC GGAAGGCACA TTTCAAGGAC ACCACACTCC CCCCGTTCAA

9061 AAAGGCCTCC GATATGGTAT AATTATTT ATTACCTAG AAGTCCTATT CTTCTTAGGC
 9121 TTCTTCTGAG CCTTTTACCA TTCAAGCCTA GCACCTACCC CTGAGCTAGG CGGCTTCCTGA
 9181 CCACCGGCAG GCATTACTCC CTTAGACCCC TTGAAGTCC CACTTCTTAA CACAGCAGTC
 9241 CTGCTTGCCT CTGGCTAAC TGTTACCTGA GCACACCACA GCATTATAGA GGGTAAACGA
 9301 AAACAAACTA TTCAATCTCT TGCTCTTACA ATCTTACTCG GGGGGTACTT CACTTCCTT
 9361 CAAGGCCTAG AGTATCACGA AGCCCCCTTC ACCATTGCAG ACGGAGTTA CGGCCTAC
 9421 TTCTTGTGAG CCACCGCTT CCACGGACTA CACGTCTTAA TTGGCACGTC ATTCTTAGCC
 9481 GTTTGTCTAT TACGCCAAT TCTTCACCAT TTACATCAGA ACCACCACTT TGGGTTTGAA
 9541 CGAGCCGCTA GATACTGACA CTCGAGAC GTCGCTGAC TTTCTCTTA TATCTCTATT
 9601 TACTGATGAG GATCTTAATC TTCTAGT TAAATCTAGT ATAAGTGACT TCCAATCACC
 9661 CGGTCTTGGT TAAAATCCAA GGGAAAGATAA TGAGCCTTCT TCTAACCATC ATTCGATTG
 9721 CCACCCCTCT CTCGACAGTA CTTGCCCTTG TGCCCTTTG ATTACCCCAA ATTACACAG
 9781 ACCACGAAAA ACTATCACCT TACGAGTGC GCTTGACCC GATAGGATCC GCGCGACTGC
 9841 CTTTTTCACT GCGGTTTTTC CTCATCGCCA TCCCTCTTCT TCTATTGAT TTAGAAATTG
 9901 CTCTCCTCCCT CCCTCTTCCA TGAGGAGAC AACTAGCATC ACCACTACTG ACATTAC
 9961 GAGCTACAGC TGTCTTAGCC CTTCTAACCC TTGGCCTCAT TTACGAATGA ATGCAAGGGG
 10021 GTCTAGAATG GGCGGAATAG TGTTAGTC TAAGAAAAAA ATTGATTT GGCTCAAAAAA
 10081 CTTGTTGTTT AAATCCGAA CCGCTTAATG ACCCCCCACAC ACTTTGCCTT CTCCCTAGCC
 10141 TTTCTTCTGG GTTTAACAGG CTCGGCATTC CACCGGTTTC ACCTCCCTTC CGCTCTATTG
 10201 TGCCTTGAAG GAATAATACT ATCCCTATT ATTGCTCTT CCCTATGAAC GCTCCAAC
 10261 GATTCAACCA GCTTTTCAGC ATCTCCAATA CTTCTGCTT CATTTCAGC CTGCGAAGCA
 10321 AGCGCCGGTC TCCTCCATT AGTAGGCACT GCTCGAACCC ACGGGACCGA CCGATTACAA
 10381 AGCCTAAATC TTCTCCAATG CTAAAAGTCC TAATCCCAAC ACTTATGCTA ATCCCAACAG
 10441 CCTGGTTACT CAAACCCAA TGACTCTGAC CCATGACTTT AATGTATAGC TTTGATCT
 10501 CCTTGTGTCAG CCTTTCTGTA CTAAAAAAAC TCTCAGAAAC CGGCTGATCG TCACCTAGCC
 10561 TCTTTATAGC TACCGACTCC CTATCGACCC CCCTCCTCGT TCTTACATGC TGATTACTCC
 10621 CACTAATAAT TTTGGCAAGC CAGAAACACA CAGCCTTACA ACCTCTTAGT CGCCAGCGCA
 10681 TGTACATCAC ACTTCTCGCC TCACTCCAGT TCTTCTTAAT CCTAGCATTT AGGCCACCG
 10741 AACTGGTAAT GTTTTACGTA ATATTGAAAG CCACCCCTTAC CCCCACGCTA ATCATTATCA
 10801 CCCCGTGAGG AAACCAAAACAA GAACGCTAA ACCGAGGAAC CTACTTTCTC TTCTATACAT
 10861 TAGCGGGCTC ACTCCCCCTT CTCGTTGCT TACTCCTACT CAAAACACA TCCGGCACCC
 10921 TATCGTTGTT AACCTTCAC TACACGGACC CGCTCGCTCT CTCGCTTAT GCAGACAAAC
 10981 TATGATGAGC AGGCTGCTT TTAGCATTCC TGGTTAAAAT ACCCCTCTAC GGAGTCCACC
 11041 TATGGCTCCC TAAAGCCAT GTGAAGCTC CAATTGCaGG CTCATAATC CTTGCGAGCAG
 11101 TTCTACTGAA GCTAGGGGA TACGGCATGA TCCGCATGAT AACGATATTG GAACCTCTAA
 11161 CTAAGGAATT AAGCTACCCC TTCATTGCT TTGCACTTTG AGGGGTAATT ATAACTGCT
 11221 CAATTGTCAT ACGCCAGACA GACCTTAAGT CCTTAATTGC CTACTCATCA GTCAGCCACA
 11281 TGGGCTAGT GGCTGGGAA ATTCTTATT AATCACCCCTG AGGCTTAACA GGAGCCCTCA
 11341 CCCTCATTAAT TGACACGGG CTTACTCAT CAGGCCCTCT CGCTTGGCA AACACAAACT
 11401 ATGAACGAAC CCACAGTCGA ACAATGGTCC TGGCGCGAGG ACTCTCAAGTA GCCCTAC
 11461 TAATAGCCAC TTGATGGTT ATTTCTAGCT TGGCAACCT AGCTCTGCCA CCACTACCA
 11521 ACCTCATGGG GGAGCTGATA ATTATCACTT CTTATTCAA CTGGTCCCTGG TGAACCTG
 11581 CATTAAACGGG TGCAGGCACT CTAATCACAG CTGTTACTC CTTTACATAA TTCTTAATAA
 11641 CTCAACGGGG CCCTCTCCCA ACACATCTCA TTGCACTTGA GCCATCTCAC ACCCGAGAAC
 11701 ACCTCTTAT TGCACCTCAC CTTATCTCC TAATTCTGCT CGTACTTAAG CCCGAGCTGA
 11761 TCTGAGGCTG AACTGCTGT AGGTGTAGTT TTAACAAAAA CTTAGATTG TGATTCTAAA
 11821 AATAAGGGTT AAAATCCCTT CTCCCAACCGA GAGAGGCTCG CAGCAATGAG AACTGCTAAT
 11881 TCCCACGACC TTGGTTGGAC CCCTAGGCTC ACTCGAAAAG CTCTAAAGG ATAACAGCTC
 11941 ATCCGTTGGT CTTAGGAACC AAAAACCTTT GGTGCAAATC CAAGTAGCAG CTATGTACCC
 12001 CACCTCCCTA ATGATTTCAT CGAGCCTGAT CACAATCTT GCATTACTAG CCTATCCTT
 12061 AATCACCGAC ATCCGCTCA CACCTCGAGG CCCTCAATGA GCTACCACCC ATGTCAAAC
 12121 AGCTGTAAGG AATAGCTTCT TTGTTAGT GCTACCCCTG ACCCTTATTT TTAATGAGGG
 12181 CGCAGAAACAA ATCGTTACTA ATTGAACCTG AATAAAATC AACTCTTCA ATATCAACAT
 12241 TAGCTTGAAA TTTGATTCT ACTCAATTAT TTTACACCC ATTGCCCTCT ACGTAACCTG
 12301 GTCCATCCTG GAGTTGCGAT CATGATACAT GCACGCGAC CCACACATGA ACCGCTCTT
 12361 TAAATACCTC CTGACGTTTC TGATTGCCAT AATCGCTCTA GTAACGTCAA ACAATATGTT
 12421 TCAACTATTG ATCGGCTGGG AGGGAGTGGG GATCATATCG TTCTTACTTA TTGGATGGTG
 12481 GTTCGGACGG GCGGATGCTA ACACGCGAC GCTCAAGCA GTAATTACAA ACCGAGTAGG
 12541 AGACATCGGA CTTATCTCG CCATGGCTTG AATAGCAGCA AACCTAAACT CCTGGGAGAT
 12601 ACAACAAATC TTTGCAACCA CAAAGGACAT AGACTTAACC TACCCCTCTC TCGGACTCAT
 12661 CATTGCGAGCA ACTGGCAATT CGGCTCAATT CGGACTTCA CTTGACTAC CCTCCGCTAT
 12721 GGAGGGTCCC ACACCGGTCT CTGCCCTACT TCATTCTAGC ACTATGGTCC TTGCCGGCAT
 12781 TTCTCTGCTA GTACGAATAA GCCCCCTCT GGAAAACAAAC CAAACCGCCC TCACAACCTG
 12841 CCTATGCTC GGAGCTCTA CAACACTATT CACTGCAACC TGTGCCCTCA CGCAAAATGA
 12901 TATTAACAAAAA ATCGCTCAT TCTCTACATC CAGCCAGCTT GGGCTTATAA TGGTAACC
 12961 TGGCTTAACAC CAAACACAAAC TAGCATTCT TCACATCTGT ACACATGCCT TCTTCAGGC
 13021 CATACTGTTT TTATGCTCG GATCTTAT CCACAGCCTA ATGATGAAC AAGATATCCG
 13081 CAAACATAGGG GGCATACACC ACCTTGCACC TTTCACATC TCTTGCCTAA CTATTGGCAG
 13141 CCTCGCCCTC ACAGGAACCC CCTTCTTGGC TGGCTTCTTC TCTAAAGATG CCATCATCGA
 13201 AGCATTGAAC ACATCCCACC TAAACGCCCTG AGCCCTAACCT TAAACCCCTCC TAGCCAC
 13261 TTTCACAGCA ATCTACAGCT TTCGAGTAGT ATTCTTGTAA CCCATGGGCC ACCCCCGATT
 13321 TAGTCCCCCTC TCTCCGATCA ACGAAAAACAA CCCAGCAGTA ATTAACCCAC TTAAGCGACT
 13381 AGCATGGGGAA AGTATGTTG CAGGCTACT AATACCTCT AATATTACAC CCTTAAAAAAC
 13441 CCCTGTCATA TCTATGCTC CTCTCCCTAA ACTAGCTGCT CTTGCGAGTTA CAATCATCGG
 13501 CCTTCTTGTGTT GCTATAGAC TCGCCATATT AACCAACAAAG CAATATGAGC CAATACCG
 13561 CACCGACATA CACCACTCT CTGCCCTATT AGGCTTCTTC CCCGGGGTCA TACACCG
 13621 AACCCCTGTA TTGCTTCTC TCCTAGGCCA AAACATTGCT AACAAACAG TAGATCAGAC
 13681 ATGATTAGAA AAAACAGGCC CCAAAGCCAT TGAATCGCT AACCGACCTC TAGTGTCTC
 13741 TACAAGTAAT ATTCAACGGG GGGTGTCAA AACATATCTT ACCCTCTTCC TCCTAACACT

13801 AGCCCTTATA ATCCCCATAT TCATCACCTA AACCGCTCGC AACGCCCTC GGCTCATACC
 13861 CCGAGTTAAC TCAAGGACTA CAAACAAAGT TAAGAGCAAT ACCCATGCAC TAGTGATTA
 13921 CAACCACCCC CCAGAAGAGT ATATTAAGGC TACCCCTCCG ATATCCCCTC GAAACACGGA
 13981 AAACCTCCCC AACTCATCTG CTAAGACCCA AGAAGATTCA TACCACCCCC CTCAAAACAA
 14041 CCCAGAAACT AATAATACTC CTCCTGTATA GACAACCCCG TAGACCATTAG CCGACCAACT
 14101 ACCCCAGCTC TCTGGGTACG GCTCAGCAGC TAACGCTGCT GAATATGCAA ATACAACCTAA
 14161 CATCCGCCT AAGTAAATTA AGAACAGGAC TAGTGACAAA AGGGCCCTC CATGCCAAC
 14221 TAATATACCA CACCCCATGC CAGCTACCCAC AACCAAACCC AAAGCAGCAA ATAAGGGGA
 14281 CGGATTAGAA GCAACTGCAA CTAACCCCA CACAAGAGAA AATAAAACTA AATATATAGC
 14341 GAAAGTCATA ATTCTGCCA GGACTTAAAC CAGGACTAA GGCTTGAAAA ACCACCGTTG
 14401 TTATTCACCT ACAAGAACCC TAATGCCAG TCTACGTTAA TCCCACCCCTC TTCTAAAAAT
 14461 CGCAAACGAT GCTTTAGTCG ACCTCCCAGC CCCCTCTAAT ATCTCTGTT GATGAAACCTT
 14521 TGGGCTCTT TTAGGACTCT GTTTAATTAC CCAAATTGCA ACCGGCTTAT TTCTAGCCAT
 14581 ACACATACACA TCAGACATG CTACTGCCT CACCTCCGTA GCCCACATCT GTCGAGACGT
 14641 CAACTACGGC TGGCTTATCC GAAGCATTCA TGCAACGGC GCATCATTCT TTTTCATTG
 14701 CCTCTACCTT CATATTGGTC GAGGTCTATA CTATGGCTATA TACCTCTATA AAGAAACATG
 14761 AACTGTTGGG GTTATTCTCC TCCCTCTCGT AATAATAACA GCCTTCGTTG GATACGTCCT
 14821 CCCTTGAGGA CAAATGTCGT TTTGAGGTGC AACCGTTATT ACCAACCTTC TATCTGCCGT
 14881 CCCCTACGTC GGAAACACCC TGGTCCAATG GATCTGAGGT GGTTTTCTG TAGACAATGC
 14941 TACCCCTACC CGGTTCTTG CCTTCCACTT CCTCTTCCCC TTCATCATCG CTGCCGCAAC
 15001 GTTAATTCAT CTACTCTTC TTCACGAAAC TGGCTAAAC AACCCAACCG GGCTAAACCTC
 15061 AGACTCCGAC AAAGTCCCTT TTCCACCTA CTTCACTTAC AAGGACCTCT TAGGCTTTC
 15121 AGTCCTCTT ACTGCATTAG CATCCCTAGC CCTATTTCCT CCAAATCTCT TAGGAGACCC
 15181 AGATAACTTT ACCCCCTGCA ATCCGCTTGT TACTCCTCCA CACATTAAGC CAGAGTGATA
 15241 CTTCTGTGTT GCCTATGCA TTCTCCGTC CATTCCAAAC AAACTTGGAG GAGTACTTGC
 15301 CCTTTTATTC TCAATCTTAG TTCTTATGCT CGTCCCCCTC CTTCATACCT CTAAACAACG
 15361 AAGCCTGATA TTCCGTCCCC TAACACAATT TTTGTTCTGA TCTTTAGTAG CAGACGTAT
 15421 GATTCTAACT TGAATTGGAG GAATGCCGT AGAACACCCCC TTGTTTATCA TTGGACAAGT
 15481 AGCGTCTCTC CTTTACTTCT CCCTCTTCTC AGTCTGATC CCAACAGCAG GTTGAUTAGA
 15541 GAATAAGTC CTTGGATGAA AATGCACTAG TAGCTCAGGA CACAGAGCCC CAGTCTTGT
 15601 AACTGGCGT CGGAGGTAG AATCCTCCCT ACTGCTCAA GAAAGGGAGAT TTCAACTCCT
 15661 ACCCTTAACCT CCCAACGCTA GGATTCTAGC ACTAAACTAT TCTTGTGTT ACATAATGT
 15721 ACAATGAAGG ATTTTCATAT ACATGTATGT AATAACACCA TATATTATA GTAACCATT
 15781 TATGTAATGT ACTAGGACAT ACATGTATAA TCACCAACATC TCGTAATAGA AACTCATT
 15841 ACCACCATT TTAACTAAGA CGAACTAAAA CCTGAATAAT CACTAATCTT AAATAAGTGG
 15901 AAGTCCAGGA CCAGTCGAGA TCTAACGACCG AACACAACAC TCATCAGTCAG AGTTATACCA
 15961 AGACTCAAAA TCTCGCCGAC CTCAAATCC TATGTTGAA GAGCCTACCA ACCGGTGATT
 16021 CCTTAATGAT AACTCTTATT GAGGGTGAGG GACAAAATT GTGGGGGTTT CACTCGGTGA
 16081 ATTATTCTG GCATTTGGT CCTACTTCAG GGCCATTGAT TGATATTATC CCTCACACTT
 16141 TCATCGACAC TTACATAAGT TAATGTTGAT AATACATACG ACTCGTTACC CAGCAAGCCG
 16201 GGCCTTCACT CCAGCGGGTA AGGGGTTCTC TTTTTTTTT TCCTTTCACT TGACATTCA
 16261 AAGGGCATCC AGCCAACAA AACGTTAAA GGGGGAGCAC TTTCTTGCA CGCACCGTAC
 16321 ATAGTATCCA TGTAGTGTAGT CTTTATTAAA AAAATAACAT TAAGTGTAT CATGTGCATA
 16381 AAAATTGCT CATTCTCTT ATCATCCCCC GGATACCCCCC TTTTTTGCG CAAAAAACCC
 16441 CCCTACCCCC CTAAACCCCT GAAGTGTCTA ACACCCCTGA AAACCCCCCG GAAACAGGAC
 16501 AAACCTCTGG TAGCTCAAAA AAGGcAAATTA CTTTATACCT ACTTTTCAA CCACAAATAC
 16561 CCCTGGCTtA TCAaaCCCC aCCAGTTGTT TTAatAATAC CACTTTTCA CCCACAAATA
 16621 CCCCTGGCTT ATCACAAACCC CACCAAGTTGT TTTAATAATA CCACCTTTCA ACCCACAAAT
 16681 ACCCCTGGCT TATCACAAAC CCACCAAGTT TTTAATAAT ACCACTTTTC AACCCACAAA
 16741 TACCCCTGGC TTATCACAAAC CCCACCAAGTT GTTTAATAAT TACCACTTTT CAACCCACAA
 16801 ATACCCCTGG CTTATCACAA CCCCAACAGT TGTTTTAAT ATACCACTTT TCAACCCACA
 16861 ATACCCCTG GCTTATCACAA ACCCCACAG TGTTTTAAT AATACCACTT TTCAACCCAC
 16921 AAATACCCCT GGCTTATCACAA AACCCCCCA GTTGTGTTAA TAAATACCT TTTCAACCC
 16981 CAAATACCCC TGGCTTATCA CAACCCCCAC AGTTGTTTTA ATAATACCA TTTTCACCC
 17041 ACAAAATACCC CTGGCTTATC ACAACCCCCAC CAGTTGTTTT ATAATACCA CTTTTCAACC
 17101 CACAAATACC CCTGGCTTAT CACAACCCCCA CCAGTTGTTT TAATAATACC ACTTTCAAC
 17161 CCACAAATAC CCCTGGCTT TCACAAACCCCC ACCAGTTGTT TTAATAATAC CACTTTCA
 17221 CCCACAAATA CCCCTGGCTT ATCACAAACCC CACCAAGTTGT TTTAATAATA CCACCTTTCA
 17281 ACCCCAAATA ACCCCCTGGT TATCACAAAC CCACCAAGTT TTTAATAAT ACCACTTTTC
 17341 AACCCAAATA TACCCCTGGC TTATCACAAAC CCCACCAAGTT GTTTAATAAT TACCACTTT
 17401 CAACCCACAA ATACCCCTGG CTTATCACAA CCCCAACAGT TGTTTTAATA ATACCACTTT
 17461 TCAACCCACA AATACCCCTG GCTTATCACAA ACCCCACAG TGTTTTAAT AAtaccact
 17521 ttcaACCCAC AAATACCCCT GGCTTATCAT AACCCCCACCA GTTGTGTTAA TAATACCACT
 17581 TTTCAACCC CAAATACCC TGGCTTATCA CAACCCCCAC AGTTGTTTTA ATAATACCA
 17641 TTTTCACCC ACAAAATACCC CTGGCTTATC ACAACCCCCAC CAGTTGTTTT ATAATACCA
 17701 TTTTCACCC CACAAATACC CCAAACCCAC AAACGCCCTG TCACCTCTT CATGATCAA
 17761 AAACCTGTG TTAGATTATC TCAAGTATTCTT CTTTACCCAC CTTAACTACG GAAGGCACAA
 17821 AACAAATACCC AAGAAAATCG CTCCATTCAA ATACCTTTTT TTAACCGATA GACCCTTAAG
 17881 TCCTCTGAA AACCAACCTT TTAACGCCCA AAAGGCTAAA CACAAACAGA GTACCACCTC
 17941 TAGATGTGGC TAGAAAATT ACT

Hippoglossus stenolepis 4 17902 bp

1 GCTGACGTAG CTTAATTAAA GCATAACACT GAAGATGTTA AGATGGGCC TAGAAAGCCC
 61 CGGGAGCACA AAGGCTGGT CCTGACTTTA CTGTCGACTT TAATCTAACT TACACATGCA
 121 AGTATCGCC CCCCTGTAG AATGCCACA ACTCCCTGCT TGGGAACTAG GAGCCGGTAT
 181 CAGGCACAAG CCCAGCTAGC CCACGACGCC TTGCTTAGCC ACACCCCTAA GGGAACTCAG
 241 CAGTGTAAA TATTAAGCCA TAAGTAAAAA CTTGACTTAG TTAAGGTTAA GAGGGCCGGT
 301 AAAACTCGTG CCAGCCACCG CGGTTATACG AGAGGCCAA GTTGACAAAC AACGGCGTAA

361 AGAGTGGTTA GGGGATTAC TAAACTAGAG CCGAACGCTT TCAAAGCTGT TATACGCACC
 421 CGAAAGTATG AAACCCAATT ACGAAAGTAG CTCTACTCAT CCTGAACCCA CGAAAGCTAA
 481 GAAACAAACT GGGATTAGAT ACCCCACTAT GCTTAGCCCT AAACATCGAT TGCACATAC
 541 ACTCCATATC CGCCCGGGAA TTATGAACAT CAGTTAAAA CCCAAAGGAC TTGGCGGTGC
 601 TTAACATCCA CCTAGAGGAG CCTGTTCTAG AACCGATAAC CCCCCTTAAA CCTCACCTTC
 661 TCTTGTTTA TCCGCCTATA TACCACCGTC GTCAAGCTAC CCTGTGAAGG CTTTACAGTA
 721 AGCAAAATTG GCACAGCCC AAACGTCAAG TCGAGGTGTA GTGAATGAGG AGGGAAAGAAA
 781 TGGGCTACAT TTGCTAACAA TAGCAACAC GAATGTTGCA TTGAAACATG TAACTGAAGG
 841 AGGATTTAGC AGTAAGCAGG AAATAGAGC TCCCGCTGAA ACTGGCCCTA AAGCGCGCAC
 901 ACACCGCCCG TCACCCCTCC CAAGCCCCCT AAACATAAAC ATTAAAACC CTACAACCCG
 961 CGAAGGGGAG GAAAGTCGA ACATGTAAG TGTAACGGAA GGTGCGCTTG GAAAATCAG
 1021 AGTGTAGCTA AGATAGAATA CAGCATCTCA CTTACACCGA GAAGACGTCC GTGCAAGTCG
 1081 GATCACCTG ACGCCTATT GCTAGCCAA CCCCTTAACA CAACAAACCC CCATTTATAA
 1141 CCCCTAAAGC ACGAACACC CACGTAGCTA AACCAATTCTC CCCCCTAAGT CCAGGCGATA
 1201 AAAAGGAAA TTTTGGAGCA ATAGAAAAAG TACCGCAAGG GAAAGCTGAA AGAGAGATGA
 1261 AAAAGCCCG TAAAGCTTA GAAAGCAGAG CTTAGTGTCT GTACCTTTG CATCATGATT
 1321 TAGCTAGCAC TTTCAAGCAA AGAGAACCTA GAGTTGTAA CCCCGAAACT GAGTGAGCTA
 1381 CTCCAAGACA GCCTATTATG AGGGCGAAC CGTCTCTGTG GCAAAAGAGT GGGAAAGAGCT
 1441 TTGAGTAGAG GTGACAAACC TACCGAACTC AGTTATAGCT GGTTGCCTGT GAATTGAATA
 1501 GAAGTTCAGC CCCCTGGTT CTCCACTCAT GCACCTTATT AACCCCTTCAG ATGCAATGAG
 1561 GAACCAGGGG TGTTAGTCAA AGGGGGTACA GCCCCTTGA TACAAGACAC AACTTTCTCA
 1621 GCAGGATAAA GATCATATTC AAATAAGGAC AGATGTTTA GTGGGCCTAA AAGCAGGCCAC
 1681 CTTAATAGAA AGCGTTAACAG CTCAAACATA AAGCCCTCCC GTTATACCGA TAACCCATAC
 1741 TTATTCCCCC ACATTAACA GGCCTCTTA TGCAACACAT AGGAACGACT ATGCTAATAT
 1801 GAGTAATAAG AGGGTATAAT CACCCCTCTC TTGACATGT GTAAATCGGA ACAGGACCCCC
 1861 CACCGAATCT TAACGGCCCC AATCAAAGAG GGTATTGGAA ACTACCACAA ATTTAGGCCA
 1921 GAAAACATC CAATGAAAC CCGTTAACCC CACACTGGTG TGCCCCAAAG GAAAGACCAA
 1981 AGGGGGGAGA AGGAACCTGG CAAACATACC CCAAGCCTCG CCTGTTTACC AAAAACATCG
 2041 CCTCTTGCAT AACACACAGTA TAAGAGGTCC CGCCTGCCA GTGACCATAT AGTCAACCG
 2101 CCGGGTATT TGACCGTGC AAAGGTAGCG TAATCACTG TCTTTAAAT GAAGACCTGT
 2161 ATGAATGGCA TAACGAGGGC TTAACTGTCT CTTCCCCCG GCATATGAAA TTGATCTCCC
 2221 CGTGCAGAAG CGGGGATTAA ACCATAAGAC GAGAAGACCC TATGGAGCT TAGACACACA
 2281 GGTGCCCAT GTCaAATACC CCCAGCTAAC GGCTCTGAAC AAATGGAATC TGCTTGATG
 2341 TCTTCGGTTG GGGCGACCAT GGGGAATACA AAACCCCCAC GTGGAAAGGG AGCACACCCC
 2401 TAAGTTACTT CTTCTCCCG AAGCCAGAGC AACAGCTCTA ACCAGCAGAA ATTCTGACCA
 2461 AACTGATCCG GTAAAACCGA TCAACCGAAC AAGTTACCTT AGGGATAACA GCGCAATCCCC
 2521 TTCTTAGAGC CCATATCGAC AAGGGGTTT ACACCTCGA TGTTGGATCA GGACATCCTA
 2581 ATGGTGCAGC CGCTTAAAG GGTCCGTTG TTCAACAGT AAAGTCCTAC GTGATCTGAG
 2641 TTCAGACCGG AGTAATCCAG GTCAGTTCT ATCTATGACA TGATTTTTC TAGTACGAAA
 2701 GGACGGAAAA GAAGAGCCCC CTGCTAAAG TACGCCCTAC CCCCACCTAT TGAAAAAATC
 2761 TAAACTAGGC AAAAGGGCAT AACCCCTTTG CTGGAGATAA CAGCAAGTTT GGGTGGCAGA
 2821 GCCCGGCTAA TGCAAAAGAC CTAAGCCCTT TCCACAGAGG TTCAAATCCT CTCTTAACT
 2881 ATGATTTCAA CTCTCATTAC GCATATTATC AACCCACTAA CCTTTATTGT ACCTGTATTG
 2941 CTAGCCGTAG CATTCTTAC CCTCCCTGAA CGAAAAGTC TAGGCTACAT ACAACTCCGA
 3001 AAAGGGCCAA ACATCGTAG GCCTTACGGC CTCCCTAAC CTTATGCTGA TGTTGTAAAA
 3061 CTCTTATTAA AAGAACCCAT CTCGACCTCG ACCGCATCAC CCGTCCTATT CCTCCTGGCC
 3121 CCTATGCTCG CTCTCACACT AGCCCTAAC CTTTGAGGCC CCATACCTT CCCGTACCC
 3181 GTCGTGGACC TAAACCTAGG TATTTTATT ATTCTAGCCC TATCTAGCCT TGAGTATAT
 3241 TCTATTCTGG GCTCCGGCTG AGCATCTAAT TCAAATATG CCCTAGTAGG GGCACACTGG
 3301 GCTGTCGCAC AAACAATTTC TTACGAGTT AGCCTTGGAC TCATCTTGCT AAACATTATC
 3361 ATCTTACCG GAGGCTCAC ACTTCAACACC TTTAACACAG CCAAGAGGG CATTGACTA
 3421 GTAGTGCCTG CTTGACCCAT GGTCACCTA TGATATATCT CACTCTTG CGAACACAAAC
 3481 CGTGCACCCCT TCGACTTAC AGAAGGAGAG TCTGAACCTG TCTCGGGCTT CAACGTAGAG
 3541 TACGCAGGCG GACCCCTCGC CTTGTTCTC TTGGCCGAAT ATTCAAATAT TCTCCTAATA
 3601 AATACCTGT CCGCTACACT ATTTTAGGG GCCTCCACA TTCCATCTAT CCCAGAAATTA
 3661 ACCAGTATCA ATATTATGAC TAAAGCAGCC CTTCTGTCAA TTGTCTTCCT ATGAGTCCGA
 3721 GCTTCGTACC CGCGATTCCG TTACGACCA CTTATGCACC TCATTTGAAA AAACTTCTT
 3781 CCACTGACTT TGGCACTAGT TATCTGACAC TTGGCGCTCC CTATTGCTATT CGCTGGCTC
 3841 CCACCAACAC TATAGGCCA GGAGCTGTGC CTGAAATTAA AGGGCCACTT TGATAGAGTG
 3901 AATTATGGGG GTTAAAGTC CCCCCAACCC TTAGAAAGAA GGGGATCGAA CCCAACCTGA
 3961 AGAGATCAA ACTCTTCTG CTTCCCTAC ACCACCTTCT AGTAAAGGTCA GCTAAATAAG
 4021 CTCTGGGCC CATAACCCAA CCATGTAGGT TCAAATCCTT CCTTTACTAA TGAACCCCTA
 4081 CATCTTGACC ACTCTTCTAT TTGGTTGGG CCTAGGAACA ACACTCACGT TTGCAAGCTC
 4141 ACACTGGCTC CTCGCTGAA TAGGACTTGA GATTAATACA CTAGCCATTA TTCCCCTAAT
 4201 AGCCCAACAT CACCAACCCCC GGGCACTCGA AGCTACCAACT AAGTACTTTTC TCgCACAAGC
 4261 CACAGCAGCG GCCACCTCC TGTTCGAG TACTACCAAC GCTGACTCA CAGGCCAATG
 4321 GGACATTCAA CAGATAACAC ATCCCTCCCA AACAACAATC ATTGTTATTG CCCTCGCACT
 4381 AAAAGTCGGA CTGGCACCA TGCACTCTG ACTTCCCGAA GTACTCCAAG GACTAGACCT
 4441 AACCAACCGGA CTTATCCTCT CAACCTGACA AAAACTTGCA CCCTTGCCCG TCTTACTACA
 4501 AATTCAAACCA ACCAACCTTA CGCCTCTAA CATATTGGC TTACTATCCA CGCTCGTGG
 4561 TGGCTGAGGT GGTCTAACCC AAACCCAGTT ACGCAAAGTC CTCGCCTACT CCTCAATTG
 4621 CCACCTTGGC TGAATAATG TGATCGTCA ATTTCACCC CCTCTTGCCCC TTCTCACCC
 4681 CCTCACGTAC TTTACTATG CTTTCTCAGC ATTCCCTAATT TTAAAGTAA ATAAAGCCAC
 4741 CACTGTCAAT GCCCTCGAA TCTCGTGAAC AAAGACTCTT CCCCCTCACAG CCCTAGCACC
 4801 CCTGGTTTA CTATCCCTCG GCGGCCCTCC CCCCCCTTACCG GGCTTCATAC CCAAGtGATT
 4861 TATTCTTCAG GAACTGACTA AGCAAGACCT TCCAGCACTT GCCACCCCTCG CCGCTTTAAC
 4921 TGCCCTTGT AGTCTCTACT TCTACCTCG TCTCTCGTAT GCAATAACCC TCACAATATT
 4981 CCCCAACAAAC CTCATTGGCG TCACTCCCTG ACGATTCTAC TCCCCCAAT TTACGCTTCC
 5041 CCTCGCCGTC TCCACTGCAG CAACCAACCT TCTTCTCCCA CTAGCCCCCG CTGCGTAGC

5101 ACTACTGGTC ACCTAAGGGA CTTAGGCTAG CACCTAGAAC AATGGCCTTC AAAGCCGTAA
 5161 GCGTGAGTGA AAACTCTCA GTCCCTGTT AGACTTGGG GCTACTATCC CACATCTCT
 5221 GCATGCAAAG CAAACACTT AATTAAGCTA AAGCCTTTCT AGATAGGAAG GCCTTGATCC
 5281 TACAAAATCT TAGTTAACAG CTAAGGCC CAAATCCAGGA GCATCTATCT ACTTTTCCCC
 5341 GCCTAATTAA GAAACTAAAA GGCGGGGAA AGCCCCGGCA GGCAGTTAGC CTGCTTCTT
 5401 AGATTTGCAA TCTAACATGT AACACCCGG GGCTTGGTAA GAAGAGGGCT TGCACCTCTG
 5461 TCTATGGTC TACAATCCAC CGCTTAACCT AGCCATCCCA CTCGTGGCAA TCACACGTTG
 5521 ATTTTCTCG ACCAACATCA AAGACATCGG CACCCCTAT CTCGTATTTG GTGCCTGAGC
 5581 CGGAATAGTG GGGACAGGCC TAAGTCTGTT TATTCGGGCA GAACTAAGCC AACCCGGGGC
 5641 TCTCTGGGA GACGACCAA TTTATAATGT AATCGTCACC GCACACGCCT TTGTAATAAT
 5701 CTTTTTTATA GTAATACCA TTATGATGG GGGGTTCGGA AACTGGCTTA TTCCACTAAT
 5761 AATTGGGGCC CCAGATATGG CGTCCCCTCG AATGAATAAT ATGAGTTCT GACTTCTCC
 5821 CCCCTCCTT CTCCTCCTC TAGCCTCTC AGGTGTTGAA GCCGGAGCAG GTACCGGATG
 5881 AACCGTATAC CCCCCACTAG CTGGCAATTAGC AGCCCACGCC GGAGCATCCG TAGACCTGAC
 5941 AATCTCTCA CTTCACCTG CAGGAATTTC ATAATTCTG GGGGCAATTAA ACTTTTATTAC
 6001 TACTATCATT AACATGAAA CCACACAGT CACTATGTAC CAAATCCCT TATTTGTTT
 6061 AGCCGTTCTT ATTACAGCCG TACTTCTCT TCTATCCCT CCGGTTTTAG CCGCAGGAAT
 6121 TACAATACTA CTAACAGACC GCAACCTTAA CACGACCTTC TTTGACCCCTG CCGGAGGAGG
 6181 TGACCCCATC CTCTACCGC ACCTATTCTG ATTCTTTGGT CACCCAGAGG TATACATCT
 6241 TATTCTCCCG GGCTTCGGAA TAATTCTCA CATTTGTCGA TACTATGCAG GTAAGAAAGA
 6301 ACCTTTGGC TACATGGGA TAGTCTGAG TATAATGGCC ATTGGACTCC TGGGGTTCAT
 6361 TGCTTGGGC CATCACATAT TTACAGCTGG AATAGATGTA GACACACGAC CCTACTTTAC
 6421 CTCTGCCACA ATAATCTG CGATTCCAAC TGGCGTAAAGC GTCTTTAGCT GACTCGAAC
 6481 CCTCCATGGA GGAAGCATTA AATGAGAAAC ACCCCCTTGTG GAGGCGCTG GCTTTATTT
 6541 CCTCTTACA GTAGGCGTC TCACTGGCAT TGTCTTAGCT AACTCCTCTC TCGATATTGT
 6601 TCTGCATGAC ACATACTATG TAGTCGCTCA CTTCCACTAT GTACTATCTA TGGGTGCTGT
 6661 ATTTGCAATC GTTGCCTGC TCGTCCACTG GTTCCATTAA TTTACAGGCT ATACCTCTCA
 6721 CTCCACATGA ACAAAAATCC ACTTCGGCT GATATTATTG GGAGTCATC TAACATTCTT
 6781 CCCCCAACAT TTTCTGGGC TGGCGGAAT ACCCCGACGG TACTCAGACT ACCCAGATGC
 6841 ATACACCCCT TGAAACACTG TTCATCAAT CGGGTCCCTA ATGCCCCCTC TTGCTGTAAT
 6901 TTATTTTTA TTCATTATTG GAGAACATT TACAGCCAA CGAGAAGTGG GAGCAGTAGA
 6961 ACTAACTGCA ACTAACATG AATGACTTTA CGGCTGCCCC CCGCCCTTAC ACACATTGA
 7021 AGAGCCCGCA TTGCTACAAG TTGCTATAAA TTGCAACAAG CTAACGAGAA AGGGAGGAGT
 7081 TGAACCCCCA TCAATTGGTT tCAAGGCAAC CACATAACCG CTCTGTCACT TTCTTCACAA
 7141 CACACCAATA AGATACTAGT AAAACGCTAT AACACTGCCT TGTCAAGACA GAATTGTGGG
 7201 TTCAACCCCCC GCGTATCTA AAAGCAATGG CACATCCCTC ACAACTCGGA TTCAAGAGC
 7261 CAGCTTCACC CTTAATGAA GAACTACTTC ACTTCCACGA CCACGCCCTA ATAATTGAA
 7321 TTCTCATTAG CACAATAGTA CTTTATTTA TTGCGGCTAT AGTCACAGCC AAACAAACAG
 7381 ATAAACTTGT ACTAGACTT CAAGAAATTG AAATTATCTG AACAGTTCTC CCAGCTGTTA
 7441 TTCTTATCCT TATGCCCCCTA CCATCCATCC GAATTGTTGTA CTTAATAGAC GAAATTAAATG
 7501 ACCCCCCACCT GACAATTAAA GCCATTGGCC ACCAGTGATA CTGAAGCTAC GAATACACAG
 7561 ACTACGAAGA CCTCGGTTT GACTCATACA TAATCCCCAC ACAAGACCTG ACCCCTGGAC
 7621 AATTCCGACT ACTAGAACCA GACCAACCGA TAGTAATCCC GGTAGAACATCC CCAATTCCG
 7681 TTTTAACTCTC GGCTGAAGAT GTCTTACCT CCGGGCAGT CCCCCTCTCTA GGCCTAAAG
 7741 TTGACGGCGT ACCTGGACG TTGAAACAG CAACCTTTAT TGTAGCCGA CCAGGGCGTAT
 7801 TTTTCGGACA ATGTTCTGAG ATTTGCGGGG CTAACCATAG TTTTATACCC ATTGTTGTAG
 7861 AAGCAGTCCC CCTTGAAACAC TTGAGAACT GGCTTCACT ATAATTGAA GACGCCCTCG
 7921 TAAGAACGTA ATAAGTACAA GCGTTAGCCT TTAAAGCTAA AGACTGGTGC CTAACAAACCA
 7981 CCCTTAGCGA CATGCCCTAA CTGAACCCCG CACCCCTGATT TGCAATTAA GTTTCTCTT
 8041 GAATAATCTT TTAAACTGTT ATTCCCCCA AAGTTTTAGC ACACACCTAT CCTAATGAAC
 8101 CACCCCTCCA AAGCACACAA AAACCTAAAA CAGAGCCCTG AACTGACCA TGATACTAAG
 8161 CTTCTTGTAC CAATTATTC CCTCCCTCATC CCTCGGCATC CCACAAATTG CACTAGCGAT
 8221 TAGCTGCCT TGAAATCTG TCCCAACCCCG GCAAGCCCG TGATTAACAA ACCGGCTACT
 8281 AACACTTCAA GGGTATTAA TTAGCCGTT CACCTCGAA CCTCTTCTAC CCCTAAACCC
 8341 CGGAGGCCAC AAGTGGCAG TTCTATTAC CTCGTTAATA GTTTCTCTAT TTTCTATCAA
 8401 CATGCTAGGA CTCCCTCCAT ACACCTTCAC ACCAACGACC CAACTCTCCC TTAATATAGG
 8461 CCTCGCAGTA CCCCTTGAT TAGCAACTGT CATTATCGGG ATACGAAATC AACCAACGCA
 8521 TGCCCTAGGC CACCTCTC CAGAAGGCAC CCCACTGCC CTCATCCCTG TACTTATTAT
 8581 CATTGAAACA ATTACCTAT TCATTGACCC CCTCGCCCTC GGTGTCGGAC TAACAGCAA
 8641 TCTCACAGCA GGCCCATCTC TTATTCAACT CATTGCAACA GCTGCCCTCG TCCTTCTTCC
 8701 CCTCATACCT GTTGTGCTA TCTTAAACAC GGTGGCTCTC TTCCCTCCTCA CCCTACTAGA
 8761 AGTAGCCGTT GCCATGATTC AAGCTTATGT CTTGTTCTA CTCCTAAAGCC TATATCTACA
 8821 AGAGAACGTT TAATGGCCCA TCAAGCACAC CCATACCCACA TAGTCGACCC CAGCCCATGA
 8881 CCCCTAACAG GGGCTATTGC TGCCCTATTG ATAACATCTG GCCTTGCTAT CTGATTCCAC
 8941 TTCCACTCCA CAACCTCTGAT AACTATTGGG ACAATCCCTC TCACCTTAAC AATCTTCAA
 9001 TGATGACGGG ACGTCGTAC GGAAGGCACA TTCAAGGAC ACCACACTCC CCCCCTCAA
 9061 AAAGGCTCC GATATGGTAT AATTATTTT ATTACCTCTG AACGTCTTATT CTTCTTAGGC
 9121 TTCTCTGAG CCTTTTACCA TTCAAGCCTA GCACCTACCC CTGAGCTAGG CGGCTTCTGA
 9181 CCACCGCAG GCATTACTCC CTTAGACCCCTT TTGAGAGTC CACTTCTAA CACAGCAGTC
 9241 CTGCTTGCCT CTGGCGTAAC TGTTACCTGA GCACACCACA GCATTATAGA GGGTAAACGA
 9301 AAACAAACTA TTCAATCTCT TGCTCTTACA ATCTTACTCG GGGGGTACTT CACTTTCCCT
 9361 CAAGGCCTAG AGTATCACGA AGCCCCCTTC ACCATTGAGC AGGGAGTTTA CGGCGCTACA
 9421 TTCTTGTGGT CCACCGGCTT CCACGGACTA CACGTCTAA TTGGCACATC ATTCTTAGCC
 9481 GTTGTCTAT TACGCCAAAT TCTTCACCAT TTCAACATCAA ACCACACTT TGGGTTTGAA
 9541 GCAGCCGCT GATACTGACA CCTCGTAGAC GTCGTCTGAC TTTCTCTCA TATTCTATT
 9601 TACTGATGAG GATCTTAATC TTCTAGTAA TAAATCTAGT ATAAGTGAATC TCCAATCACC
 9661 CGGTCTGGT TAAAATCCAA GGGAAAGATAA TGAGCCTTCT TCTAACCATC ATTCGATTG
 9721 CCACCCCTCCT CTCGACAGTA CTTGCCCTTG TGTCCTTTG ATTACCCCAA ATTACACCAG
 9781 ACCACGAAAA ACTATCACCT TACGAGTGC GCTTGTACCC GATAGGATCC GCCCGACTGC

9841 CTTTTTCACT GCGGTTTTC CTCATGCCA CCCTCTTCT TCTATTGAT TTAGAAATTG
 9901 CTCTCCTCCT CCCCTTCCA TGAGGAGACC AACTAGCATC ACCACTACTG ACATTACACCT
 9961 GAGCTACAGC TGTCTAGCC CTTCTAACCC TTGGCCTCAT TTACGAATGA ATGCAAGGGG
 10021 GTCTAGAATG GGCGGAATAG GTGGTAGTC TAAGAAAAAC ATTTGATTT GGCTCAAAAA
 10081 CTTGTGGTT AAATCCGAA CCGCTTAATG ACCCCCACAC ACTTTGCCTT CTCCCTAGCC
 10141 TTTCTCTGG GTTAAACAGG CCTGGCATTC CACCGGTTTC ACCTCCTTTC CGCTCTATTG
 10201 TGCTTGAAAG GAATAATACT ATCCCTATT ATTGCTCTT CCCTATGAAC GCTCCAACCT
 10261 GATTCAACCA GCTTTTCAGC ATCTCCATA CTCTGCTTG CATTTCAGC CTGCGAAGCA
 10321 AGCGCCGGT TCGCCCTATT AGTAGGCACT GCTCGAACCC ACAGGGACCGA CCGATTACAA
 10381 AGCCTAAATC TTCTCAAATG CTAAAATG TAATCCCAAC ACTTATGCTA ATCCCAACAG
 10441 CCTGTTACT CAAACCCAC TGACTCTGAC CCATGACTTT ATGTATAGC TTTTGATCT
 10501 CCTGGTCAG CCTTCTGTGA CTAAAAAAC TCTCAGAAC CCGCTGATCG TCACCTAGCC
 10561 TCTTATAGC TACCGACTCC CTATCGACCC CCCTCCTCGT TCTTACATGC TGATTACTCC
 10621 CACTAATAAT TTTGGCAAGC CAGAAACACA CAGCCTTAGA ACCTCTTAGT CGCCAGCGCA
 10681 TGTCATCAC ACTTCTGCC TCACTCTCAGT TCTCCTAAT CCTAGCATT AGCGCCACCG
 10741 ACTGGTAAT GTTTACGT ATATTGAAAG CCACCTTAT CCCCACGCTA ATCATTATCA
 10801 CCCGCTGAGG AAACAAACAA GACGCTAAAC CACGAGGAAC CTACTTCTC TTCTATACAT
 10861 TAGCGGGCTC ACTCCCCCTT CTCGTTGTT TACTCCTACT CAAAACACA TCCGGCACCC
 10921 TATCGTTGTT AACCCCTCAC TACACGGACC CGCTCGCTCT CTCGCTTAT GCAGACAAAC
 10981 TATGATGAGC AGGCTGTCTT TTAGCATTCC TGGTTAAAT ACCCCTCTAC GGAGTCCACC
 11041 TATGGCTCCC TAAAGCCAT GTGGAAGCTC CAATTGCAGG CTCAATAATC CTTGCAGCAG
 11101 TCTACTGAA GCTAGGGGA TACCGCATGA TCCGCATGAT AACGATATTG GAACTCTAA
 11161 CTAAGGAATT AAGCTACCC TTCATTGTCT TTGCACTTGG AGGGGTAATT ATAACCTGGCT
 11221 CAATTGCTC ACGCCAGACA GACCTAAAGT CCTAATTGCT CTAACATCA GTCAGCCACA
 11281 TGGGCTAGT GGCTGGGG ATTCTTATT AATCACCCCTG AGGCTTAACA GGAGCCCTCA
 11341 CCCTCATATA TGACACCGG CTTACTTCAT CAGCCCTTCT CTGCTGGCA AACACAAACT
 11401 ATGAACGAAC CCACAGTCGA ACAATGGTCC TGGCGCGAGG ACTTCAAGTA GCCCTACAC
 11461 TAATAGCCAC TTGATGGTT ATTTCTAGCT TGGC AACCT AGCTCTGCCA CCACTACCA
 11521 ACCTCATGGG GGAGCTGATA ATTATACACTT CTTTATTCAA CTGGTCTG TGAACTCTGG
 11581 CATTAAACGG TGCAGGCACT CTAATCACAG CTGGTTACTC CTTTACATAA TTCTTAATAA
 11641 CTCAACGGGG CCCTCTCCA ACACATGTCA TTGCACTTGA GGCATCTCAC ACCCGAGAAC
 11701 ACCTCTTAT CGCACTTCAC CTTATCCCTT TAATTCTGCT CGTACTTAAG CCGAGCTGA
 11761 TCTGAGGCTG AACTGCGTGT AGGTGTAGTT TAAACAAAAA CATTAGATTG TGATTCTAAA
 11821 ATAAGGGTT AAAATCCCTT CTCCCACCGA GAGAGGCTCG CAGCAATGAG AACTGCTAAT
 11881 TCCCACGACC TTGGTTGGAC CCCTAGGCTC ACTCGAAAAG CTCTAAAGG ATAACAGCTC
 11941 ATCCGTTGGT CTTAGGAACC AAAAAACTTT GGTGCAAATC CAAGTAGCAG CTATGTGCC
 12001 CACCTCCCTA ATGATTCTAT CGAGCTGTAT CACAATCTT GCATTACTAG CCTATCCTT
 12061 AATCACCGC ATCCGTCTCA CACCTCGAGG CCCTCAATGA GCTACCACCC ATGTCAAAC
 12121 AGCTGTTAA ATAGCTTTT TTGTTAGTCT GCTACCCCTG ACCCTATTTC TTAATGAGGG
 12181 CGCAGAAACA ATCGTTACTA ATTGAACCTG AATAAAATACT AACTCTTCA ATATCAACAT
 12241 TAGCTGAAA TTGATTCT ACTCAATTAT TTTACACCC ATTGCCCTCT ACGTAACCTG
 12301 GTCCATCTG GAGTCGAT CATGATACAT GCACGAGAC CCACACATGA ACCGCTCTT
 12361 TAAATACCTC CTGACCTTC TGATTGCCAT AATCGTCTTA GTAAC TGCAA ACAATATGTT
 12421 TCAACTATT ATCGCTGGG AGGGAGTGG GATCATATCG TTCTTACTTA TTGGATGGTG
 12481 GTTCCGACGG GCGGATGCTA ACACCTGCAGC GCTCAAGCA GAAATTTACA ACCGAGTAGG
 12541 AGACATCGGA CTTATCTTG CCATGGCTTG AATAGCAGAC AACCTAAACT CCGAGAGAT
 12601 ACAACAAATC TTTGCAACCA CCAAGGACAT AGACTTAACC TACCTCTTC TCAGACTCAT
 12661 CATTGCaGCA ACTGGCAAAT CGGCTCAATT CGGACTTCAC CTTGACTAC CCTCCGCTAT
 12721 GGAGGGTCCC ACACGGTCT CTGCCCTACT TCATTCTAGC ACTATGGTCC TTGCCGGCAT
 12781 TTTCTGCTA GTACGAATAA GCCCCCTCTT GGAAAACAAC CAAACCGCCC TCACAACCTG
 12841 CCTATGCCCTC GGAGCTCTA CAACACTATT CACTGCAACC TGTGCCCTCA CGCAAAATGA
 12901 TATTAACCAAA ATCGTCGAT TCTCTACATC CAGCCAGCTT GGGCTTATAA TGGTAACCAT
 12961 TGGCTAAAC CAACACACAA TAGCTTCT TCACATCTG ACACATGCC TCTTCAAGGC
 13021 CATACTGTT TTATGCTCG GATCTATTAT CCACAGCCTA ATGATGAAAC AaGATATCCG
 13081 CAAAATAGGG GGCATACACC ACCTTGACCC TTTACATCC TCTTGCCTAA CTATTGGCAG
 13141 CCTCGCCCTC ACAGGAACCC CTTCTTGGC TGGCTCTTC TCTAAAGATG CCATCATCGA
 13201 AGCATTGAAAC ACATCCCAAC TAAACCGCTG AGCCCTAACCT TAAACCCCTC TAGCCACCTC
 13261 TTTCACAGCA ATCTACAGCT TCGAGTAGT ATTCTTGTAA CCTATGGGCC ACCCCCGATT
 13321 TAGCTCCCTC TCTCCGATCA ACGAAAACAA CCCAGCAGTA ATTAACCCAC TTAAGCGACT
 13381 AGCATGGGA AGTATCGTT CAGGCTACT AATTACCTCT AATTATCAC CTTAAAAAAC
 13441 CCCTGTCATA TCTATGCCCTC CTCTCTTAA ACTAGCTGCT CTTGCAGTTA CAATCATCGG
 13501 CCTTCTTGTG GCTATAGAGC TCGCCATATT AACCAACAAG CAATATGAGC CAATACCCGG
 13561 CACGCACATA CACCACTCT CTGCCATTAGGCTTCC CCGGGGTCA TACACCGCT
 13621 AACCCCTGTA TTGCTCTTCT TCCTAGGCCA AAACATTGCT AACCAAAACAG TAGATCAGAC
 13681 ATGATTAGAA AAAACAGGCC CCAAAGCCAT TGAATCCGCT AACCGACCTC TAGTGTCTC
 13741 TACAAGTAAT ATTCAACGGG GGGTGTCAA AACATATCTT ACCCTCTTCC TCCTAACACT
 13801 AGCCCTATA ATCCCATAT TCATACACTA AACCGCTCGC AACGGCCCTC GGCTCATACC
 13861 CCGAGTTAAC TCAAGGACTA CAAACAAAGT TAAGAGCAAT ACCCATGAC TAGTGTATTA
 13921 CAACCAACCC CCAGAAGAGT ATATTAAGGC TACCCCTCCG ATATCCCTC GAAATACGGA
 13981 AAACCTCCCC AACTCATCTG CTAAGACCA AGAAGATTCA TACCACCCCC CTCAAAACAA
 14041 TCCAGAAACT AATAATACTC CTCCTGTATA GACAACCCCG TAGACCATT CCGACCAACT
 14101 ACCCCAGCTC TCTGGTAGC GCTCAGCAGC TAACGCTGCT GAATATGAA ATACAACCTAA
 14161 CATCCCGCCT AAGTAAATT AGAACAGGAC TAGTGACAAA AAGGGCCCTC CATGCCAAC
 14221 TAATATACCA CACCCATGTC CAGCTACCG AACCAAACCC AAAGCAGCAA AATAAGGGGA
 14281 CGGATTAGAA GCAACTGCAA CTAAACCCAA CACAAGAGAA AATAAAACAA AATATATAGC
 14341 GAAAGTCATA ATTCTGCTTA GGACTTTAAC CAGGACTAAT GGCTTGAAA ACCACCGTTG
 14401 TTATTCAACT ACAAGAACCC TAATGCCAG TCTACGTTAA TCCCACTC TTCTAAAAAT
 14461 CGCAACAGAT GCTTTAGTCG ACCTCCAGC CCCCTCTAAT ATCTCTGTTT GATGAAACTT
 14521 TGGGCTCTT TTAGGACTCT GTTTAATTAC CCAAATTGCA ACCGGCTTAT TTCTAGGCCAT

14581 ACACATACACA TCAGACATTG CTACTGCCTT CACCTCCGT AGCCACATCT GTCGAGACGT
 14641 CAACTACGGC TGGCTTATCC GAAGCATTCA TGCCAACGGC GCATCATTCT TTTTCATTG
 14701 CCTCTACCTT CATATTGGTC GAGGTCTATA CTATGGCTCT TACCTCTATA AAGAAACATG
 14761 AACTGTTGGG GTTATTCTCC TCCTTCTCGT AATAATAACA GCCTTCGTG GATACGCTCT
 14821 CCCTTGAGGA CAAATGTCGT TTGAGGTGC AACCGTTATT ACCAACCTTC TATCTGCCGT
 14881 CCCCTACGTC GGAAACACCC TGGTCCAATG GATCTGAGGT GTTTTCTG TAGACAATGC
 14941 TACCCCTACC CGGTTCTTG CCTTCCACTT CCTCTTCCCC TTCATCATCG CTGCCGCAAC
 15001 GTTAATTCTAT CTACTCTTC TTCACCAAAC TGGCTCAAAC AACCCAAACCG GGCTAAACTC
 15061 AGACTCCGCA AAaGTTCCCTT TTCAACCCCTA CTCACCTAC AAGGACCTCT TAGGCTTGC
 15121 AGTCTCTTACT ACTGCATTAG CATCCCTAGC CCTATTCTC CCAATCTCT TAGGAGACCC
 15181 AGATAACTTT ACCCTGCAA ATCCGCTGT TACTCCCCA CACATTAAGC CAGAGTGATA
 15241 CTTCTGTGTT GCCTATGCTA TTCTCCGCTC CATTCACAAAC AAACCTGGAG GAGTACTTGC
 15301 CCTTTTATTC TCAATCTTAG TTCTTATGCT CGTCCCCTTC CTTCATACCT CTAAACAACG
 15361 AAGCTGTATA TTCCGTCCTC TAACACAAATT TTTGTTCTGA TCTTTAGTAG CAGACGTAT
 15421 GATTCTAATCTG TGAATTGGAG GAATGCCGT AGAACACCCC TTCGTTATCA TTGGACAAGT
 15481 AGCGTCTCTC CTTTACTCTC CCTCTCTCC AGTCTGATC CCAACAGCAG GTTGACTAGA
 15541 GAATAAAGTC CTTGGATGAA AATGCATAG TAGCTCAGCA CACAGAGCCC CAGTCTTGT
 15601 AACTGGCCGT CGGAGGTTAG AATCTCCCT ACTGCTCAA GAAAGGAGAT TTCAACTCCT
 15661 ACCCCTAACT CCCAAAGCTA GGATTCTAGC ACTAAACTAT TCTTTGTGGT ACATAAATGT
 15721 ACAATGAAGG ATTTTCATAT ACATGTATGT AATAACACCA TATATTATA GTAACCATT
 15781 TATGTAATGT ACTAGGACAT ACATGTATAA TCACCCACATC TCGTAATAGA AACTCATT
 15841 ACCACCATTT TTAACTAAGA CGAACAAAAA CCTGAATAAT CACTAATCTT AAATAAGTGG
 15901 AGTCCAGGA CCAGTCGAGA TCTAAGACCG AACACAAAC TCATCAGTCC AGTTATACCA
 15961 AGACTCAAAA TCTCGCCGAC CTCAAAATCC TATGTTAGTAA GAGCCTACCA ACCGGTGATT
 16021 CCTTAATGAT AACTCTTATT GAGGGTGGAG GACAAAAATT GTGGGGGTTT CACTCGGTGA
 16081 ATTATTCCTG GCATTTGGTT CCTACTTCAG GGCCATTGAT TGATATTATC CCTCACACTT
 16141 TCATCGACAC TTACATAAGT TAATGTTGAT AATACATACG ACTCGTTACC CAGCAAGCCG
 16201 GGC GTTCACT CCAGCGGGTA AGGGGTTCTC TTTTTTTTTT CCCTTTCACT TGACTTTCA
 16261 AAGTGCATCC AGCAACAGA GACGTTAAA GGGGGAGCAC TTTTCTTGCA CGCGCCGTAC
 16321 TACTATGCC TGTAGTGGAGT CTTTACTTAA AGAATAACAT TAAGTGTAT CATGTGCATA
 16381 AAGATTTGCT CATTCTCTT ATCATCCCCA GGATACCCCCC TTTTTTGCG CGAAAAAACCC
 16441 CCCTACCCCCC CTAAACCCCCC GAAGTGTCA AGACCCCTGA AAACCCCCCG GAAACAGGAC
 16501 AAACCTCTGG TAGCTCAGAA AAGGCAATTAA CTTTATAC ACCTTTCAAC CCACAAATAC
 16561 CCCTGGCTTA TCATAACCCC ACCAGTTGTT TAAATAATAC CACTTTCAA CCCACAAATA
 16621 CCCCTGGCTT ATCACAACCC CACCAAGTTGT TTTAATAATA CCACCTTTCA ACCCACAAAT
 16681 ACCCCTGGCT TATCACAACCC CCACCAAGTTGT TTTAATAAT ACCACCTTTCA ACCCCACAA
 16741 TACCCCTGGC TTATCACAAC CCCACCAAGT GTTTAATAAA TACCACTTT CAACCCACAA
 16801 ATACCCCTGG CTTATCACAAC CCCACCAAGT GTTTAATAAT ACCACACTTT TCAACCCAC
 16861 AATACCCCTGG GCTTATCACAAC CCCACCAAGT TTGTTTTAA TAAATACCACT TTCAACCCAC
 16921 AAATACCCCCT GGCTTATCAC AACCCACCA GTTGTAA TAAATACCACT TTCAACCCCA
 16981 CAAATACCCC TGGCTTATCA CAACCCACCA AGTGTAA ATAATACCACT TTCAACCCCA
 17041 ACAAAATACCC CTGGCTTATCA CAACCCACCA CAGTTGTAA ATAATACCA CTTTCAACCC
 17101 CACAAATACC CCTGGCTTATCA CAACCCACCA CCAGTTGTAA TAAATACCA ACCTTTCAAC
 17161 CCACAAATAC CCCTGGCTTA TCACACACCC ACCAGTTGT TTAATAATAC CACTTTCAA
 17221 CCCACAAATA CCCCTGGCTT ATCACAACCC CACCAAGTTGT TTAATAATAA CCACCTTTCA
 17281 ACCCCACAAAT ACCCCCTGGCT TATCACAACCC CCACCAAGTTGT TTTAATAAT ACCACCTTT
 17341 AACCCACAAAT TACCCCTGGC TTATCACAAC CCCACCAAGT GTTTAATAAA TACCACTTT
 17401 AACCCACAAAT TACCCCTGG CTTATCACAAC CCCACCAAGT GTTTAATAAA TACCACTTT
 17461 TCAACCCACA AaTACCCCTG GCTTATCATA ACCCCACCAAG TTGTTTTAA TAAATACCACT
 17521 TTCAACCCAC AAATACCCCCT GGCTTATCAT ACCCCACCAAG TTGTTTTAA TAAATACCACT
 17581 TTCAACCCCA CAAATACCCC TGGCTTATCA TAACCCACCA AGTGTAA ATAATACCACT
 17641 TTTCAACCCCA CAAATACCCC TAACCCACCA AACGCCCTGT CACCTTCTTC ATGATCAAAA
 17701 AACTGCTGTT TAGATTATT CAAGTATTCC TTTCACCCAG CTAACATACGA AAGGCACAAA
 17761 ACAAAATCCA AGAAAATGTC TCCATTAAAC TACCTTTTTT TAACCGATAG ACCCTTAAGT
 17821 CCTCCTGAAA ACCAGCCTT TAACGCCCA AAGGCTAACAC AACAAACAGAG TACCACTCT
 17881 AGATGTGGCT AGAAAATTAA CT

Reinhardtius hippoglossoides 1 18017 bp

1 GCTAACGTAG CTTAATTAAA GCATAACACT GAAGATGTTA AGATGGACCC TAGAAAGCCC
 61 CGGGAGCACA AAGGCTGGT CCTGACTTTA CTGTCGACTT TAACTAAACT TACACATGCA
 121 AGTATCGCC CCCCTGTGAG AATGCCACA ACTCCCTGT TGGGAACTAG GAGCTGGTAT
 181 CAGGCCACAAG CCCCGCTAGC CCACGACGCC TTGCTTAGCC ACACCCCTAA GGGAACTCAG
 241 CAGTGATAAA TATTAAGCCA TAAAGTAAAAA CTTGACTTAG TTAAGGTTAA GAGGGCCGGT
 301 AAAACTCGT CCAGCCACCG CGGTTATACG AGAGGCCAA GTTGACAAAC AACGGCGTAA
 361 AGAGTGGTTA GGGGATTAC TAAACTAGAG CGGAACGCTT TCAGAGCTGT TATACGCACC
 421 CGAAAGTATG AAACCCAAAT ACAGAACTAG CTCTACTCAT CCGAACCCCA CGAAAGCTAA
 481 GAAACAAACT GGGATTAGAC ACCCCACAT GCTTAGCCCT AACACATCGAT TGCACATATAC
 541 ACTCCATATC CGCCCGGGAA TTATGAACGT CAGTTAAAAA CCAAAGGAC TTGGCGGTG
 601 TTAACATCCA CCTAGAGGAG CCTGTTCTAG AACCGATAAC CCCCGTTAAA CCTCACCCCTC
 661 TCTTGTTTA TCCGCTATA TACCAACCGTC GTCAGCTTAC CCGTGAAGG CTTTACAGTA
 721 AGCAAAATTG GCACAGCCCA AAACGTCAAG TCGAGGTGA GTGAATGAGG GGGGAAGAAA
 781 TGGGCTACAT TTGCTAAACA TAGCAACAC GAATGTTGCA TTGAAACATG CAACTGAAGG
 841 AGGATTAGC AGTAAGCAGG AAATAGACCG TCCCGCTGAA ACTGGCCCTA AAGCGCGCAC
 901 ACACCGCCCG TCACCCCTCC CAAGGCCCTA GAACCTAAACT AATTTAAACCA CAAACACCCG
 961 CGAAGGGAG GAAAGTCGTAAG ACATGGTAAG TGTACCGGAA GTGCGCTTG GAAAATCAG
 1021 AGTGTAGCTA AGTTAGATTA CAGCATCTCA CTTACACCGA GAAGACGCC GTGCAAGTCG
 1081 GACCACCCCTG ACGCCTATTA GCTAGCCCA ACCCTTAACA CAACAAACCC CCATTTATAA
 1141 CCCCTAAAGC ACGAACACCC CACGTAGCTA AACCAATTCTC CCCCTAAGT CCAGGCAGATA

1201 AAAAAGGAAA TTTGGAGCAA TAGAAAAAGT ACCGCAAGGG AAAGCTGAAA GAGAGATGAA
 1261 AAAGCCCAGT AAAGCTTAAG AAAGCAGAGC TTAAaGCTCG TACCTTTGC ATCATGATT
 1321 AGCTAGCACT TTCAAGCAA GAGAACCTAA AGTTTGTAC CCCGAAACTG AGTGAGCTAC
 1381 TCCAAGACAG CCTATTATA GGGCAAACCC GTCTCTGTGG CAAAGAGTG GGAAGAGCTT
 1441 TGAGTAGAGG TGACAAACCT ACCGAACCTCA GTTATAGCTG GTGCCTGTG AATTGAATAG
 1501 AAGTTCAGCC CCCTGGGTT TCCACTCATG CACTTTATTA ACCCTTCAGA TGCAATGAGG
 1561 AACCAGGGT GTTAGTCAAA GGGGTACAG CCCCTTGAT ACAAGACACA ACTTTCTCAG
 1621 TAGGATAAAG ATCATAATAA AATAAGGACA GATGTTTTAG TGGGCTAAA AGCAGCCACC
 1681 TTAACAGAAA GCGTTAACG TCAAACATGA AGCCCTCCG TTATACCGAT AACCTATCT
 1741 TAATCCCCA CATTAAACAG GCCCTCTAT GCATCACATA GGAACGACTA TGCTAATATG
 1801 AGTAATAAAGA GGGTATAAAC ACCCTCTCT TGCACATGT TAAATCGAA CGGACCCCC
 1861 ACCGAATCTT AACGGCCCA ATCAAAGAGG GTATTGAAA CTACCACAAA TTCAGGCCAG
 1921 AAAAACATCC AATGATAACC CGTTAACCCC ACACTGGTGT GCCCAAAAGG AAAGACCGAA
 1981 GGGGGGAGAA GGAACCTGGC AAACATACCC CAAGCCTCGC CTGTTTACCA AAAACATCGC
 2041 CTCTTGCTA ACCACAGTAA AGAGGTCCG GCCTGCCAG TGACCATATA GTTCAACGGC
 2101 CGCGTATT TGACCGTCA AAGGTACCGT AATCACTTGT CTTTTAAATG AAGACCTGTA
 2161 TGAATGGCAT AACGGGGTAA TAACTGTC CTTCCCCCG TCATGAAAT TGATCTCCCC
 2221 GTGCAGAAGC GGGGATTAA CCATAAGACG AGAAGACCT ATGGAGCTTT AGACACACAG
 2281 GTGGACCATG TCAAATACCC CCAACTAAGG GCCTGAAC TA GTGGAACCT GCCTTGATGT
 2341 CTTCGGTTGG GGCGACCATG GGGAAACAA AACCCCCAG TGAAAGGGA GCACACCCCT
 2401 AAGTTACTTC TTCTCCGCA AGCCAGAGCA ACAGCTCTAA CCAGCAGAAA TTCTGACCAA
 2461 ACTGATCCGG TAAAACCGAT CAACGAACCA AGTACCTCA GGGATAACAG CGCAATCCCC
 2521 TTTTAGAGCC CATATCGACA AGGGGTTTA CGACCTCGAT GTGGGATCAG GACATCCTAA
 2581 TGGTGCAGCC GCTATTAAAGG GTCTGTTGT TCAACGATTA AAGTCCCTAG TGATCTGAGT
 2641 TCAGACCGGA GTAATCCAGG TCAGTTCTA TCTATGACAT GATCTTTCT AGTACGAAAG
 2701 GACCGAAAAG AAGGGGCCA TGCTAAAAGT ACGCCTCACC CCCACCTAAT GAAAAAAATCT
 2761 AAACTAGGCA AAAGGGCATA ACCCTTTGC TGGAGATAAC AGCAAGTTGG GGTGGCAGAG
 2821 CCCGGCTAAT GCAAAAGACC TAAGCCCTT CCACAGAGGT TCAAATCCTC TCCTTAACTA
 2881 TGATTTCAAC TCTCATTACC CATATTATA ACCCCCTAAC CTTTATTGTA CCTGTAAC
 2941 TAGCCGTAGC ATTCCCTCACC CTCCTGAAAC GGGAAAGTCT AGGCTACATA CAACTCCGAA
 3001 AAGGGCTAA CATCGAGGG CCTTACGGC TCCTTCACCC CATTGCTGAC GGCCTAAAC
 3061 TCTTTATTAA AGAACCCGTT CGGGCTTCCA CCGCATCCCC CGTTCTGTT CTCCTGGCCC
 3121 CTATACTCGC ACTCACACTG GCCCTAACCC TTGAGCCCC GATGCCCTTC CCGTACCC
 3181 TTGTGGACCT TAACCTGGGT ATCCTTTTA TTCTAGCACT ATCTAGCCTT GCAGTATACT
 3241 CTATTCTGG GTCTGGTGA GCATCCAATT CAAAATACGC TTTAGTAGGA GCACTACGG
 3301 CTGTTGCACA AACAAATTCT TACGAAGTTA GCCTAGGACT CATCCTCCCT AACATCATCA
 3361 TCTTACCGG AGGATTACA CTTCAACGCT TTAACACAGC CCAAGAGGCC ATCTGACTAG
 3421 TTGTACCCGC CTGACCCCTC GCTGCCATAT GATACATCTC CACTCTGCC GAAACAAACC
 3481 GTGCACCCCT CGACATAACA GAAGGGAAAT CTGAACTTGT TCTAGGCTTC AACGTAAGAT
 3541 ACGCAGGAGG ACCCTTCGCT TTGTTCTTT TAGCCGAATA CTCAAATATC CTCCTAAATA
 3601 ATACCTGTC CGAACATTA TTTTGGCG CCTCCACAT TCCCTCTATC CCAGAATTAA
 3661 CTAGCATTAA TATTATAACT AAAGCAGCCC TTCTCTCAAT TGTTTTCTG TGAGTCCGAG
 3721 CTTCGTACCC GCGATTCCGT TACGACCAAC TTATGCACCT CATTGAAAG AACTTTCTC
 3781 CACTGACTTT AGCACTAGTT ATTTGACACT TAGCGCTCCC TATTGCACTT GCTGGCTC
 3841 CACCACAACT ATAAACGAG GAGCTGTGCC TGAATTAAA GGGCCACTTT GATAGAGTGA
 3901 ATTATGGGG TAAATACCC CCCAACCTCT TAGAAAAGG GGGATCGAAC CCAACCTGAA
 3961 GAGATCAAAA CTCTCGTC TTCTCTACCA CCACTCTTA GTAAGGTAG CTAACAAAGC
 4021 TCTTGGGCC ATACCCCAAC CATGTAGTT CAAATCCTC CTTACTAAT GAAACCC
 4081 ATCTTGACTA CTCTTATATT TGGTTGGC CTAGGAACAA CACTTACGTT TGCAAGCTCA
 4141 CACTGACTCC TCGCTTGAAT AGGCCTCGAA ATTAATACAC TAGCCATTAT TCCGCTAATA
 4201 GCCAACATC ATCACCCCCG AGCAGTCGA GCTACTACTA AATACTTTCT CGCACAAGCC
 4261 ACAGCAGCCG CTACCCCTT GTTGTCAAGT ACCACCAACG CATGACTAAC AGGCCAATGA
 4321 GACATCAAC AAATAACACA CCCCCCTCCC ACACAAATAA TTGTTATTGCA CCTAGCACTA
 4381 AAAATTGGAC TAGCACCAAT AACTCTTGA CTTCAGAAAG TACTTCAGG ACTAGACCTT
 4441 ACCACCGGAC TAATCCTCTC GACTTGACAA AAACCTGCAC CTTTGCCT ATTACTACAA
 4501 ATTCAAGCAA CAAACCTAC GCCTCTAAC ATTATTGGTT TACTATCCAC CTCGTTGGG
 4561 GGATGAGGTG GTCTCAATCA AACTCAATT CGTAAAGTCC TGGCCTACTC CTCGATTGCC
 4621 CATCTGGCT GAATAATAC AATATTCAAG TTTCTACCC TCCTCGCCCT CTTGCC
 4681 CTCACGTATT TTACTATGAC TTTTCTAGCA TTCTCTAAC TAAAGTAAA CAAGGCTACC
 4741 ACTGTCATG CCCTCGCAAT CTCTACAA AAACTCCCTG CGCTCACAGC CCTAGCACCC
 4801 CTAGCTTAT TGTCTTGG TGGCTCTCTT CCCCCCTACCG GCTTCATACC CAAATGATT
 4861 ATTCTTCAAG AGCTGACTAA GCAGGACCTC CCAGCCCTTG CTACCTCGC CGCATTAACT
 4921 GCCCTCTAA GCCTTACTT CTACTGCGC CTCTCGTAGC CGATAACCC GACAATGTT
 4981 CCTAACAAAC TCGTCGGGT TACCCCTGA CGGTTCTATT CTCCCCAATT TAGCCTTCCC
 5041 CTCGCCATGT CCACTCGAGC AACCCCTCTT CTCTCTCCAG TAGCCCTGC TGCTGTAGCA
 5101 CTACTCATAG CTAAAGGAC TTAGGTAGC ACTTAGACCA ATGGCCTTCA AAGCCATCAG
 5161 CGTGAATGAA ATACTCTAG TCCCTGTAA GACTGCGGG TATTATCCC ACATCTCTG
 5221 CATGCAAAGC AGACATCTTA ATTAAGCTAA AACCTTTCTA GATAGGAAGG CTTGATCCT
 5281 ACAAAATCTT AGTTAACAGC TAAGCCCGCA ATCCAGCGAG CATCTATCTA CTTTCTCCG
 5341 CTTGATTTAC CGAGTAAAAA GCGGGAGAAA GCCCCGGCAG GCGATTAGCC TGCTCTTA
 5401 GATTGCAAT CTAACATGTA ACACCCAGA GCTTGATAAG AAGAGGGCTT GCACCTCTGT
 5461 CTATGGTCT ACAATCCACC GCTTAACCTCA GCCATCCATC CTGTGGCAAT CACACGTG
 5521 TTTTCTCGA CCAATCACAA AGACATCGGC ACCCTCTATC TCGTATTG TGCTGAGCC
 5581 GGTATAGTGG GAACAGGGCT AAGTCTGCTT ATTGGGGCAG AACTAAGCCA ACCTGGGGCT
 5641 CTCTGGAG ACGACCAAAAT TTATAACGTA ATCGTCACCG CACACGCCCT TGTAAATAATC
 5701 TTTTTATAG TAATACCCAT TATGATCGGG GGTTTGGAA ACTGGCTTAT TCCACTAATA
 5761 ATTGGAGGCC CAGATATGGC TTTCCCTGA ATAAATAACA TGAGTTCTG ACTTCTCCC
 5821 CCATCCTTC TTCTCTCTT AGCCTCTCA GGTGTTGAAG CTGGGGCAGG TACGGGGTGA
 5881 ACCGTTATC CACCACTAGC TGGTAATCTG GCCCACGCCG GAGCATCCGT TGACCTAACA

5941 ATCTTCTCAC TTCACCTTGC AGGGATTCG TCAATTCTGG GGGCAATTAA CTTTATTACT
 6001 ACCATCATCA ACATGAAACC CACAACAGTT ACTATGTACC AAATCCCATT ATTTGTTG
 6061 GCCGTTCTAA TTACCGCCGT ACTTCTTCTT CTGTCCTTC CGTCTTAGC CGCAGGGATT
 6121 ACAATGCTAC TAACAGACCG CAACCTAAC ACAACCTTTT TTGACCCCGC CGGAGGTGGT
 6181 GACCCCATCC TCTATCAACA CCTATTCTGA TTCTTGGCC ACCCAGAGGT ATATATTCTC
 6241 ATTCTTCCAG GCTTCGGAAT AATTCCCAC ATTGTTGCAT ACTATGCAGG TAAAAAAGAA
 6301 CCTTTGGTT ACATAGGAAT AGTCTGAGCT ATAATAGCCA TTGGACTCCCT GGGATTCA
 6361 GTATGGGCCA ATCACATGTT TACAGTCGGG ATAGACGTCG ACACACGAGC CTACTTTACA
 6421 TCTGCCACAA TAATCATTGC AATCCCAGC GGCGTAAAG TCTTAGCTG ACTCGCAACC
 6481 CTCCACGGGG GAAGCATCAA ATGAGAAAACA CCACCTCTT GGCCCGCTCG CTTTATTTC
 6541 CTCTTACAG TAGGCGCT TACTGGCATT GTCTTGGCTA ACTCCTCTCT TGACATTGTT
 6601 CTGCATGACA CATACTATGT AGTAGGCCAC TTCCACTATG TACTATCTAT GGGTGCTGTA
 6661 TTCGCAATCG TTGCGCCCT CGTCCACTGA TTCCCCTAT TTACAGGTTA TACCCCTCAC
 6721 TCCACATGAA CAAAATCCA CTTCGGCTG ATATTTATTG GGGTCAATCT AACGTTCTC
 6781 CCCAACACT TCCTGGCCT AGCCGGAATG CCCCGACGAT ACTCAGACTA CCCAGACGCA
 6841 TATACCTAT GAAACACTGT CTCATCAATC GGATCCCTAA TGTCGCTCGT CGCTGTAATC
 6901 TTATTTGT TCATTATTG AGAACGATT ACAGCCAAAG GGAAGTTGG GGCAGTAGAA
 6961 CTTACTGCAA CTAACATCGA ATGACTTTAC GGCTGCCCTC CACCCCTACCA CACATTGAA
 7021 GAGCCGCAT TCGTCCAAGT TCGAATAAAT TCGAACAAAC TAACGAGAAA GGGAGGAGTT
 7081 GAACCCCAT CAATTGGTT CaagCCAACC ACATAACCACG TCTGTCACTT TCTTCACAAC
 7141 ACACCAATAA GATACTAGTA AAACGCTATA ACACGCTT GTCAAGGCAG AATTGTGGT
 7201 TAAATCCCCG CGTATCTTA ACGAATGGC ACATCCCTCA CAACTCGGAT TTCAAGACGC
 7261 AGCTTCACCC TTAATAGAAG AACTACTTCA CTTCACGAT CACGCCCTAA TAATTGTAAT
 7321 TCTCATCAGC ACGATAGTAC TTTATATTG TGCGGCTATA GTTACAGCCA AACTAACAGA
 7381 TAAACTGTT CTAGATTCTC AAGAAATGAA ATTATCTGA ACAGTTCTCC CAGCTGTTAT
 7441 TCTTATTCTC ATGCCCTAC CATCTCTCCG AATTGTTGAC TTAATAGACG AAATTAATGA
 7501 TCCCCACTTA ACAATCAAAG CCATGGCCA CCAGTGGTAC TGAAGCTACG AATACACAGA
 7561 CTATGAAGAC CTCGGTTTG ACTCTTACAT AATCCCACAA CAAGACCTGA CCCCTGGCA
 7621 GTTCCGACTA CTTGAAGCAG ATCACCGAT AGTAATCCCG GTAGAATCTCC CTATTGCGT
 7681 TCTAATCTCA GCTGAAGATG TATTACATC CTGAGCAGTC CCCTCCCTGG GCGTAAAGT
 7741 TGACGGCGTA CCTGGACGAT TAAATCAAGC AACCTTTATT GTTAGCCGGC CAGCGTATT
 7801 TTATGTCAA TGCTGTAAA TTGCGGAGC AAACCATAGT TTTATACCCA TCGTTGAGA
 7861 AGCAGTCCCC CTTGAACACT TTGAAAATG GTCTTCACTA ATAATTGAAAG ACGCCTCGCT
 7921 AAGAAGCTAA TAAGTACAAG CGTTAGCCTT TTAAGCTAAA GACTGGTGCC TAACAACCAC
 7981 CCTTAGCGAC ATGCCCTAAC TGAACCCGC ACCCTGATT GCAATTCTTG TTTCTCTTG
 8041 AATAATCTT TTAACTGTTA TTCCCCCAA AGTTTAGCA CATACTTACCG CCAATGAACC
 8101 TACCTCCCAA AGCACACAAA AACCTAAAC AGAACCTGAA AACTGACCAT GATACTAACG
 8161 TTTTTGACC AATTATATC CCCAGTACAC CTGGTATTG CGCTAATTGCA ACTAGCAATT
 8221 ACCCTCCCT GAATTCTGGT CCCAACCCCT CAAGCCCGCT GATTAACCAA CGGTGTC
 8281 ACACCTCAAG GATGATTAT TAGCGCTTC ACCTCACAC TTCTTCTTCC CCTAAACCCC
 8341 GGGGGCACA AATGAGCAGT TTTATTAC TCATTAATGG TTTTCTTATT TTCTATTAA
 8401 ATACTAGGAC TCCTTCATA TACCTTCACA CCAACAAACCC AACTCTCCCT TAATATGGC
 8461 CTCGAGTTC CCCTCTGATT AGCAACTGTC ATTATTGGAA TGCAGAACATCA ACCTACACAT
 8521 GCCCTAGGTC ACCTCTTCC AGAACGCCAC CCTACTGTC TCATCCCTGT ACTCATCATC
 8581 ATTGAAACAA TTAGCCTATT TATTGGCTCCTCG GTGTCGACT AACAGCGAAT
 8641 CTCACAGCAG GCCACCTCTC TATTCAACCTC ATTGACACAG CTGCTTCTG CTTCTTCCG
 8701 CTCATACCTG TTGTTCTAT CTTAACACAG GTGGTCTCT TCCCTTCTAC TCTGCTAGAA
 8761 GTCGCGGTCG CCATGATCCA AGCCTATGTC TTGTTCTAC TCTTAAGCCT ATATCTACAA
 8821 GAGAACGTTT AATGGCCAT CAAGCACACC CATACCACAT AGTCGACCCCC AGTCCATGAC
 8881 CCCTCACGGG GGCTATTGCT GCCCTATTGA TAACATCTGG CCTTGCTATC TGATTTCACT
 8941 TCCATTCCCA AACCCTTATA ACTATTGGGA CAGTCCTCT CACCTTAACA ATCTTCCAAT
 9001 GATGACGGGA CGTCGTACGA GAAGGCACAT TTCAAGGACA CCACACTCCC CCTGTTCAA
 9061 TAGGCCCTCG ATACGAAATA ATTCTTACCA TTACCTCAGA AGTTCTGTT TTCTTAGGCT
 9121 TCTTCTGAGC CTTTTACAC TCAAGGCTAG CACCTACTCC AGAACTAGGA GGCTTCTGAC
 9181 CCCCAGCAGG CATTACCCCT TTAGACCCAT TTGAAAGTCCC GCTTCTTAAC ACAGCAGTCC
 9241 TCCTTGCCTC TGGTGTAACT GTTACCTGAG CACATCACAG CATTATAGAG GGTAAACGGA
 9301 AACAAACTAT TCAATCTT GCCCTACAA TCTTAACCGG GGGTTACTTC ACTTTCTCC
 9361 AAGGCCCTGA GTATGTCGA GCCCCCTCA CCATTGCGAGA CGGAGTTTAC GGTGCTACAT
 9421 TCTTCTGTC CACGGCTTC CACGGCTTC ACGCTTAAAT CGGCACCTCA TTCTTAGCCG
 9481 TTTGTCGTT ACGCCAAATC CTTCACCTT TTACATCAAAC CACCCACTTT GGGTTTGAAG
 9541 CAGCCCGATG ATACTGACAC TTGAGCAGC TCGTCTGACT TTTCTCTAT ATCTCTATT
 9601 ACTGATGAGG ATCTTAATCT TCCTAGTAC AAATCTAGTA TAAGTGAATT CCAATCACCC
 9661 GGTCTGGTT AAAGCCAAG GGAAGATAAT GAGCCTTCTT CTAACCATCA TTTCAATCGC
 9721 CACCCCTCTC TCAACAGTAC TTGCCATTGTT ATCTCTTCTA CTCCCCAAA TTACACCAGA
 9781 CCATGAAAG CTTACACCAT ACGAATGCCG CTTGACCCG ATAGGCTCTG CCCGACTTAC
 9841 TTTTCGCTG CGATTTTCTC TCATGCCCT CTTCTTCTC CTGTTGACT TAGAGATG
 9901 ACTCTCTC CCCCCCTCCCT GAGGAGACCA ATTACGATCC CCAACTCTA CATTCACTT
 9961 AGCTACAGCT GTCTAGCCG TTCTAACCTC TGGCCTCATT TACGAATGAA TGCAAGGAGG
 10021 CCTAAGGTGA GCTGAATAGG TGGTTAGCT AAGAAAAACA TTGATTTG CTCACCAAA
 10081 TTGTGGTTA AATCCGCAAC CGCTTAATGA CCCCCACACA CTTGCGCTTC TCCTCAGCCT
 10141 TTCTCTGGG TTTAACAGGC CTGGCATTTC ACCGGTTCCA CCTCCTTCC GCTTATTAT
 10201 GTCTTGAAGG AATAATATTA TCCCTATTAA TTGCCCTCTC CCTATGGACA CTCCAACCTTG
 10261 ACTCAACAAA CTTCAGCA TCTCCTATGC TTCTACTTGC ATTTTCAGCC TGCGAAGGAA
 10321 CGCCGGCCCT CGCCCTGCTA GTAGCCACTG CTCGAACCCCA CGGAACCGAC CGACTACAAA
 10381 GCCTAAATCT TCTCAATGC TAAAATCTT AATCCAACA CCTATGCTAA TCCCAACAGC
 10441 CTGGTTACTC AAACCCAGCT GACTCTGCC CATAACCTTG ATGTATAGCT TTGTCATCTC
 10501 CTTAATTAGC CTTTCATGAT TGAAGAACCT CTCAGAACCC GGCTGATCAT CACTTAACCT
 10561 CTTTATAGCC ACCGACTCCC TATCAACCCCC CCTCCTCGTT CTCACATGCT GACTGCTACC
 10621 ACTAATGATT TTGGCAAGCC AAAACACAC AGCCTCAGAG CCTCTGGTC GCCAGCGCAT

10681 GTACATTACA CTTCTCGCT CACTCCAGT CTTCTTGATT CTAGCATTAA GCGCCACTGA
 10741 GCTGGTAATG TTTTACGTA TATTGAAAGC CACCCCATC CCCACTCTGA TTATTATTC
 10801 CCGCTGAGGG AACCAAACAG AACGCCAAA TGCAGGGACC TACTTTCTT TTTACACATT
 10861 GGCAGGCTCA CTACCTCTC TCGTTGCTT ACTTCTTCTC CAGAACACAT CGGGCACCT
 10921 GTCGTTGTTA ACCCTACACT ATACAGACCC ACTCACTCTC TCATCTTATG CAGACAAATT
 10981 ATGATGAGCA GGCTGTCT TAGCATTCT CGTAAAATA CCTCTCTATG GGGTCCACCT
 11041 GTGACTCCCT AAAGCCCCG TTGAAGCCCC CATTGCAGGC TCAATGATCC TTGCAGGGT
 11101 TCTCTTAAAAA CTGGGGAGGAT ATGGCATGAT CCGCATAATA ACAATTAG AGCCTCTAAC
 11161 TAAGGAGCTA AGCTACCCCT TCATTGCTT TGCACTTGA GGGTAATCA TAACTGGCTC
 11221 AATTGCTCA CGACAGACAG ACCCTTAAGT CCTAATTGCC TACTCATCTG TAAGTCACAT
 11281 GGGCTGGTG GCGGGGGAA TTCTTATTCA GTCACCCCTGG GGCTTAACAG GGGCCCTAC
 11341 CCTAATAATT GCACACGGTC TTACCTCATC AGCCCTCTTC TGCCTAGCAA ACACAAACTA
 11401 TGAACGAACC CACAGTCGA CGATAGTACT AGCACGAGGG CTTCAAGTAG CTCTACCATT
 11461 GATGGCCACT TGATGATTAA TTTCTAGTT AGCCAACCTA GCCCTTCCAC CACTACCTAA
 11521 CCTCATAGGT GAACTAATAA TTATTACTTC TCTATTCAAC TGATCTGTT GAACCTAGC
 11581 ATTAACCGGG GCAGGCACCTC TAATTACAGC TGTTACTCC CTTTATATGT TCTTAATAAC
 11641 CCTACGGGG CACTCCCATC CACATCATC TGCGCTTGAG CACATCACA CCCGAGAACAA
 11701 CCTCTTATT GCACCTCATC TTATCCCTT GATTCTCCCTC GTGCTTAAGC CCGAGCTGAT
 11761 CTGAGGTTGG ACTGCCTGTA GGTGAGTTT TAACAAAAAC ATTAGATTGT GATTCTAAAAA
 11821 ATAAGGGTTA AAGTCCCTC TCCCACCCgag AGAGGCTCG AGCAATGGGA ACTGCTAATT
 11881 CCCATGACCT TGGTTGACC CCCAGGCTCA CTCGAGAACG TCCTAAAGGA TAACAGCTCA
 11941 TCCGTTGGTC TTAGGAACCA AAAACTCTTG GTGCAAATCC AAGTAGCAGC TATGCACCC
 12001 ACCTCCCTAA TGATTCGTC AAGCCTGATC ACAATCTTG CATTGCTAGC CTATCCCTA
 12061 GTCACCAAA TAGCACCAC GCCCCAAGAA CCTCAGTGGG CCTCTACCCA TGTTAAAACA
 12121 GCTGAAAAA TAGCTTCTT TGTTAGCTG TTACCCCTGA CCCTATTTCT TAACGAGGGT
 12181 GCGAGACAA TCGTTACTAA CTGAACCTGA ATAAATACTA ACTCCTTCGA CATCAACATT
 12241 AGCCTAAAT TTGACTCTA CTCATTATT TTTACCCCCA TTGCCCTCTA CGTAACCTGG
 12301 TCCATCCTAG AATTGCTATT ATGATACATG CATGCAGACC CGAACATGAA CCGCTTCTC
 12361 AAGTATCTAC TGACATTCT GATTGCCATA ATCGTTCTCG TAACCGCAAA CAATATGTTT
 12421 CAACTATTCA TCGGGTGGAGA GGGTGTGGGG ATCATATCTG TTCTCCTTAT CGGGTGGTGG
 12481 TTCCGACGGG CGGATGCTAA CACTGCAGCA CTTCAAGCAG TACTTTACAA CGGGTTGGG
 12541 GACATCGGTC TTATCTTCGC CATGGCTTGAA ATAGCAGCAA ACCTGAACTC CTGGGAAATG
 12601 CAACAAATCT TTGTAACCTC CAAAGACATA GACTTAACCTT ATCCTCTCTC CGGGCTCATC
 12661 ATCGCAGCAA CGGGCAAGTC GGCTCAATTG GGTCTTCACC CctGACTACC CTCCGCTATG
 12721 GAGGGTCTA CACCGGTCTC TGCCCTACTT CATCTAGCA CTATAGTCGT TGCCGGCATT
 12781 TTCTGCTGG TCCGAATGAG CCCCCCTCTG GAAAACAACC AAACCTGCCCT CACAACCTGC
 12841 CTGTCCTCG GGGCCCTTAC AACACTGTT ACTGCAACCT GTGCCCTCAC ACAAAATGAT
 12901 ATTAAAAAAA TCGTGCATT CTCCACATCA AGCCAGCTG GACTTATGAT GTAAACCAT
 12961 GGACTAAAC AACCACAACT AGCGTTCTT CACATTGCA CACAGCTTT CTTCAAGGCC
 13021 ATGCTATTT TATGCTCGG ATCTATTATT CATACTCTTA ATGACGAACA AGACATCCG
 13081 AAAATAGGAG GCATACACCA CTTGACACCT TTTACATCTT CTTGCCTAAC TATTGGCACT
 13141 CTTGCCCTTA CAGGGACCC CTTCTTAGCT GGCTTCTTCT CTAAAGATGC TATCATCGAA
 13201 GCACTAAACAA CATCCACCT AAACGCCCTGA GCCCTAACTT TAACCCCTCT AGCCACCTCT
 13261 TTACACAGCAA TTTACAGCTT TCGAGTAGTA TTTTTGTC GCATGGGTCA CCCCCGGTTT
 13321 ATACCCCTCT CCCCCATCAA CGAAAACAAC CCAGCAGTGA TTAACCCACT TAAGCGATTA
 13381 GCATGAGGAA GCATCGTTGC AGGCTTAACCTTAACTTCTA ACATCACACC CCTAAAAAAC
 13441 CCTGTTATAT CTATACCCCCC TCTCTTTAA CTAGCTGCTC TTGAGTAAC AATCCTGGC
 13501 CTACTGTTG CTATGAACT CGCCATGCTA ACTAATAAGC AATATAAAATC AATACCGAC
 13561 ACACACGTAC ACCACTCTC CACCTGTTA GGCTCTTCC CGGGAGTAAT GCACCGCTA
 13621 GCCCCCTAAAT TACTCTCTT CCTAGGCCA AACGTTGCCA ACCAAACAGT AGATCAAACC
 13681 TGGCTAGAAA AAATCGGCC CAAAGGCAATT GAGTCTCTA ACCGACCCCT GATCTCTCT
 13741 CAAAGTAACA TTCAGCGAGG AGTAATCAAG ACATATCTCA CCCTTTCTC CTAACACTA
 13801 GCCCCTTATAA TCCCACTT CATCGCTTAA ACTGCCCGCA AGCCCCCTCG ACTTATAACCC
 13861 CGAGTCAACT CGAGAACCCAA AACAAAGTT AAAAGCAATA CCCACGCGCT AATAATTAAT
 13921 AACCACCCCC CAGAAGAATA TATTAAAGCT ACACCTCCGA TATCCCCCCC AAACACGGAG
 13981 AACTCCGCCA ACTCATCTGC TAGGACCCAA GAGCACTCGT ATCACCCCCC TCAGAACAGC
 14041 CCAGAAACTA ACATTACCCC TCCTGTGTA ACTACCCCCAT AAACCATTAC CGACCAACTA
 14101 CCCCCAGCTCT CCGGATACGG CTCAGCAGCT AACGCTGCGG AGTATGCAAAT TACAACCTAGC
 14161 ATCCCGCCCA AGTAAATTAA GAACAAACAT AGTGACAAAAA AAGGCCCCCC ATGCCCAACT
 14221 AATATACAC ACCCCATCCAC AGCTACTACA ACCAAACCCAA AAGCAGCGAA ATAAGGAGAT
 14281 GGATTAGAAG CAACTGCAAC TAACCTAAC ACgAGAGAAA ATAAAACAA ATATATAGCA
 14341 AAAGTCATAA TTCTGCCAG GACTTTAACCT AGGACTAATG GCTTGAAAAA CCACCGTTG
 14401 TATTCAACTA CAAGAACCT AATGGCCAGT CTACGTAAT CCCACCCCTCT TCTAAAATC
 14461 GCAAACGATG CTTTAGTGA CCTCCAGCC CCCTCTAATA TCTCCGTTTG GTGAAACTTT
 14521 GGGCTCTTTC TAGGACTCTG TTTAGTAACC CAAATTGCTA CGGGCTTGT TCTAGCCATA
 14581 CACTACACAT CAGATATTGC TACTGCCCT ACCTCCGTTG CTACATCTG CGGGGACGTC
 14641 ACTACAGCTG GACTTATCCG AACGATTCTAC GCCAACGGCC CATCATTCTT TTTCATTG
 14701 ATCTACCTAC ACATCGGCCG AGGGCTCTAC TATGGCTCAT ACCTCTATAA AGAAACATGA
 14761 ACTATTGGAG TCGTCCTCTT CCTCTCTGTA ATAATAACGG CTTCTGCGG ATACGCTCTC
 14821 CCTTGAGGAC AAATGTCTT TTGAGGTGCA ACCGTTATTA CCAACCTCT ATCCGCCGTC
 14881 CCCTATGTT GCGACACCT AGTCCAATGG ATTGAGGTG GTTTTCTGT AGACAATGCT
 14941 ACCCTTACCC GGTTCTTGC CTTCCACTTC CTTTCCCTCT TCATCATCGC CGCCGCAACA
 15001 GTAATCCCATC TACTATTCTC TCACGAAACT GGCTCAAATA ATCCCAACCGG CTTAAACTCA
 15061 GACTCAGACA AGGTTCCCTT CCACCCCTAC TTCACTTACA AAGACCTCTT AGGTTTGCA
 15121 GTGCTCTCA CTGCATTAGC ATCCCTCGCT TTGTTTCTC CTAATCTTT AGGAGACCA
 15181 GATAACTTTA CCCCCGCAAA TCCACTGTT ACACCCCCAC ACATTAACCC AGAGTGTAC
 15241 TTCTCTTTG CTTACGCTAT TCTCCGGTCC ATCCCAAACA AACTTGGAGG CGTTCTGCC
 15301 CTTTTATTTT CGATCTTAGT TCTCATGCTT GTCCCTTCC TTCACACCTC TAAACAAACGC
 15361 AGCCTAATGT TTCGCCCTCT AACGCAATT TTGTTCTGGT CTCTGGTAGC AGACGTTATA

15421 ATTCTAACTT GAATTGGAGG AATGCCGT AAACACCCCT TCGTTATCAT TGGACAAGTA
 15481 GCATCACTCC TTTACTTCG CCTCTTCTA GTCTTAATCC CAACAGCAGG TTTACTAGAG
 15541 AATAAAGTCC TTGGATGAAA ATGCACTAGT AGCTCAGCG ACAGAGCCCC AGTCTTGAA
 15601 ACTGCCGTC GGAGGTTAAA ATCCTCCCTA CTGCTCAAAG AAAGGAGATT TCAACTCCTA
 15661 CCCCTAACTC CCAAAGCTAG GATTCTAGCA TTAAACTATT CTTGCGGCA TAATGTATGT
 15721 AAAGCAAAGG ATTTTATAT ACATGTATGT ATTAACACCA TATATTATA GTAACCATTT
 15781 TGTGTAATGT ACTAGGACAT ACATGTATAA TCACCTAAC TAGTAATATA GCACTCATTC
 15841 ATCACCATT TTAACTAAGA AATTACTAGA ACCCTGACTT CACATTCAAGG TTACATATGT
 15901 GAAATTCCG GACCAGCGA CATTAAAGCG CGAACACAAAC ACTCACCGGT CAAGTTATAC
 15961 CAAGACTCAA AATTTCCTACT AAATTTAAACCA CAAACCCAAT AAGAGCCTAC CAACCGGTGA
 16021 TCCCCTAAATG ATAACCTCTA TTGAGGGTGA GGGACAAAAA TTGTGGGGGT TTCACTCGGT
 16081 GAATTATTCC TGGCATTGG TTCCTACTTC AGGGCCATTG ATTGATATTAA TCCCTCACAC
 16141 TTTCATCGAC ACTTACATAA GTTAATGTT ATAATACATA CGACTCGTTA CCCACCAAGC
 16201 CGGGCGTTCA CTCCAGCGG TAAGGGTTT CTTTTTTTTT TTTCCTTTCA CTTGACATTT
 16261 CAGAGTGCAT CCAGCCAACA GAGACGTTA AAGGGTGACG ACTTTTCTT CACTCAGCGT
 16321 ACATAGTATC CATGTTATAA GTCTTTT CAGGATAAC TAAAGGGAT ATCATGTGCA
 16381 TAAGAATGTC CTCCCTTCTT CTATCATCCC CAGGATACCC CCTTTTTT CCGTAAACCC
 16441 CCCTACCCCC CTAAACTCTT GAAGTGTCA AGACCCCTGaaa aacCCCCCG GAAACAGGAC
 16501 AACACCTCTGG TAGCTTAGAA AAGGCAATTAA CTTTCAACC TACTTTCAA CCCTAAAATA
 16561 CCCCTGACCA AAGATAAGCC CACCAAGTT TTTAATAATG GTACTTTCA ACCCTAAAAT
 16621 ACCCCTGAcC AAAGAtAAGC CCACCAAGTT TTTAATAATG GGTACTTTTC AACCCtAAGA
 16681 TACCCCTGAC CAAAGATAAG CCCCaCAGTT ATTAAATAA TGGTaCTTTT CAACCCCTAA
 16741 ATACCCCTGA CCAAAGATAA GCCCACCAGT TATTAAATAA ATGGTACTTT TCAACCCCTAA
 16801 AATACCCCTG ACCAAAGATA AGCCCACAG TTATTTAAAT AATGGTACTT TTCAACCCCTAA
 16861 AAATACCCCT GACCAAGAT AAGCCACCA GTTATTTAA TAATGGTACT TTTCAACCCCT
 16921 AAAATACCCC TGACCAAGA TAAGCCACCA AGTTATTTA ATAATGGTAC TTTCAACCC
 16981 AAAATACCC CTGACCAAG ATAAGCCAC CAGTTATTT AATAATGGTAC TTTCAACCC
 17041 CTAAAATACC CCTGACCAAA GATAAGCCCA CCAGTTATTT TAATAATGGT ACTTTCAAC
 17101 CCTAAAATAC CCCTGACCAAA AGATAAGCCC ACCAGTTATTT TTAATAATGG TACTTTCAAC
 17161 CCCCTAAATA CCCCTGACCA AAGATAAGCC CACCAAGTT TTTAATAATG GTACTTTCAAC
 17221 ACCCTAAAT ACCCCCTGAC CAAAGATAAGG CCACCAAGTT TTTAATAATG GGTACTTTTC
 17281 AACCCCTAAA TACCCCTGAC CAAAGATAAGG CCCACCAAGTT TTTAATAATG GGTACTTT
 17341 CAACCCCTAAA ATACCCCTGA CAAAGATAAGG CCCACCAAGTT TTTAATAATG GGTACTTT
 17401 TCAACCCCTAA AATACCCCTG ACCAAAGATA AGCCCACAG TTATTTAAAT AATGGTACTT
 17461 TTCAACCCCTAA AATACCCCT GACCAAGAT AAGCCACCA GTTATTTAA TAATGGTACT
 17521 TTTCAACCCCT AAAATACCCC TGACCAAGA TAAGCCACCA AGTTATTTA ATAATGGTAC
 17581 TTTTCAACCCCT TAAATACCC CTGACCAAG ATAAGCCACAG TTATTTAA ATAATGGTAC
 17641 CTTTCAACCCCT TAAATACCC CCTGACCAAA GATAAGCCCA CCAGTTATTT TAATAATGGT
 17701 ACTTTCAACCCCT TAAATACCC CCTGACCAAG AGATAAGCCC ACCAGTTATTT TTAATAATGG
 17761 TACTTTCAA CCCTAAATG ACTGGGACCT GTAACGCCGT GTCATCTTCT TCATGATCAA
 17821 AAAActACTG TTAGATTAT TTCAAGTATT TCTGGCATCA TCCCAACTAC GAAGGGcACA
 17881 GAACAAAAAC CGAGAAGACT TTGCTCTCCA TTCAAAACT ATCTTTTTT AACCGACAAG
 17941 CCCCTAAACC AACCTTCTAT TATCCTAAAAA CTTGACAAGA CTAGGGAAAC ACCTTTGGT
 18001 GTAAACAGATA TTTAACT

Reinhardtius hippoglossoides 2 18139 bp

1 GCTAACGTAG CTTAATTAAA GCATAAACACT GAAGATGTTA AGATGGGCC TAGAAAGCCC
 61 CGGGAGCACA AAGGCTTGGT CCTGACTTTA CTGTCGACTT TAACCTAAACT TACACATGCA
 121 AGTATCCGCC CCCCTGTGAG AATGCCACA ACTCCCTGCT TGGGAACCTAG GAGCTGGTAT
 181 CAGGCCACAAG CCCCGTAGC CCACGACGCC TTGCTTAGCC ACACCCCTCAA GGGAACTCAG
 241 CAGTGATAAA TATTAAGGCCA TAAGTGAAAAA CTTGACTCTAG TTAAGGTTAA GAGGGCCGGT
 301 AAAACTCGTG CCAGCCACCG CGGTTATACG AGAGGCCAA GTTGACAAAC AACGGCGTAA
 361 AGAGTGGTTA GGGGATTATC TAAACTAGAG CCGAACGCTT TCAGAGCTGT TATACGCACC
 421 CGAAAGTATG AAACCAAAAT ACGAAAGTAG CTCTACTCAT CCGAACCCCA CGAAAGCTAA
 481 GAAACAAACT GGGATTAGAT ACCCCACTAT GCTAGCCCT AACACATCGAT TGCACATAC
 541 ACTCCATATC CGCCCGGGAA TTATGAACGT CAGTTAAAAA CCAAAGGAC TTGGCGGTGC
 601 TTAACATCCA CCTAGAGGAG CCTGTTCTAG AACCGATAAC CCCCCTTAA CCTCACCCCTC
 661 TCTTGTGTTA TCCGCTATA TACCAAGCTC GTCACTTAC CTCGTGAAGG CTTTACAGTA
 721 AGCAAAATTG GCACAGCCCA AAACGTCAGG TCGAGGTGA GTGAATGAGG GGGGAAGAAA
 781 TGGGCTACAT TTGCTAAACAC TAGCAAAACAC GAATGTTGCA TTGAAACATG CAACTGAAGG
 841 AGGATTAGCAG AGTAAGCAGG AAATAGAGCG TCCCGCTGAA ACTGGCCCTA AACCGCGCAC
 901 ACACGGCCCG TCACCCCTCC CAAGCCCCCT GAACTAAACT ATTAAAACC CAACAACCCG
 961 CGAAGGGGAG GAAAGTCGTA ACATGGTAAG TGTACCGGAA GGTGCGCTTG GAAAATCAG
 1021 AGTGTAGCTA AGTTAGATTA CAGCATCTCA CTTACACCAGA GAAGACGCC GTGCAAGTCG
 1081 GACCACCCCTG ACGCCTATTA GCTAGCCCAA CCCCTTAACA CAACAAACCC CCATTTATAA
 1141 CCCCTAAAGC ACGAAACACC CACGCTAGTA AACCATTCCTC CCCCCTAAGT CCAGGCGATA
 1201 AAAAGGAAA TTGAGGCAA TAGAAAAAGT ACCGCAAGGG AAAGCTGAAA GAGAGATGAA
 1261 AAAGCCCGATG AAAGCTTAAAG AAGCAGAGC TTAAAGCTCG TACCTTTGCA ATCATGATTT
 1321 AGCTAGCACT TTCAAGCAAAG GAGAACCTAA AGTTGTAAC CCGGAAACTG AGTGAAGCTAC
 1381 TCCAAGACAG CCTATTATA GGGCAACCC GTCTCTGTGG CAAAGAGTG GGAAGAGCTT
 1441 TGAGTAGAGG TGACAAACCT ACCGAACTCA GTTATAGCTG GTTGCCTGTG AATTGAATAG
 1501 AAGTTCAGCC CCCTGGGTT TCCACTCATG CACTTTATTA ACCCTTCAGA TGCAATGAGG
 1561 ACCAGGGGT GTTAGCTAAAG GGGGGTACAG CCCCTTTGAT ACAAGACACAA ACTTTCTCAG
 1621 TAGGATAAAAT ATCATAATAA AATAAGGACA GATGTTTTAG TGGGCTAAA AGCAGCCACC
 1681 TTAACAGAAA CGCTTAAAGC TCAACATGAG AGCCCTCCCG TTATACCGAT AACCTATCT
 1741 TAATCCCCCA CATTAAACAG GCCCTCTTAT GCATCACATA GGAACGACTA TGCTAATATG
 1801 AGTAATAAGA GGGTATAACC ACCCTCTCTTG TGCACATGTG TAAATCGGAA CGGACCCCC
 1861 ACCGAATCTT AACGGCCCCA ATCAAAGAGG GTATTGGAAA CTACCACAAA TTCAGGCCAG

1921 AAAAACATCC AATGATAAACC CGTTAACCCC ACACTGGTGT GCCCAAAGG AAAGACCGAA
 1981 GGGGGGAGAA GGAACCTCGG AAACATACCC CAAGCCTCGC CTGTTTACCA AAAACATCGC
 2041 CTCTTCATA ACCACAGTAT AAGAGGTCCC GCCTGCCAG TGACCATATA GTTCAACGGC
 2101 CGCGTATTG TGACCGTGC AAGGTAGCGT AATCACTTGT CTTTTAAATG AAGACCTGTA
 2161 TGAATGGCAT AACGAGGGCT TAACTGTCTC CTCCCCCGG TCAATGAAAT TGATCTCCC
 2221 GTGCAGAAGC GGGGATTAAA CCATAAGACG AGAAGACCT ATGGAGCTTT AGACACACAG
 2281 GTGGACCATG TCAAATACCC CCAACTAAGG GCCTGAACTA AGTGGAACCT GCCTTGATGT
 2341 CTTCGGTTGG GGCGACCATG GGGAAATACAA AACCCCCACG TGGAAAGGGG GCACACCCCT
 2401 AGTTACTTC TTCTCCCGCA AGCCAGAGCA ACAGCTCTAA CCAGCAGAAA TTCTGACCAA
 2461 ACTGATCCGG TAAAACCGAT CAACGAACCA AGTACCCCTA GGATAACAG CGCAATCCCC
 2521 TTTTAGAGCC CATATCGACA AGGGGGTTTA CGACCTCGAT GTTGGATCATG GACATCTAA
 2581 TGGTGCAGCC GCTATTAAAG GTTCGTTGT TCAACGATTA AAGTCCTACG TGATCTGAGT
 2641 TCAGACCGGA GTAATCCAGG TCAGTTCTA TCTATGACAT GATCTTTCT AGTACGAAAG
 2701 GACCAGAAAG AAGGGGCCA TGCTAAAAGT ACGCCTCACCC CCCACCTAA GAAAAAAATCT
 2761 AAACTAGGC AAGGGCATA ACCCTTTGC TGGAGAAC AGCAAGTTGG GGTGGCAGAG
 2821 CCCGGCTAAT GCAAAGACG TAAGCCCCCTT CCACAGAGGT TCAAATCCTC TCCTTAACTA
 2881 TGATTCAAC TCTCATTAC CATATTATTA ACCCCCCCTA CTTTATTGTA CCTGTATTGC
 2941 TAGCCGTAGC ATTCCCTCACC CTCCCTGAAC GAAAAGTCTT AGGCTACATA CAACTCCGAA
 3001 AAGGGCCTAA CATCGTAGGG CCTTACGGCC TCCTTCAACC CATTGCTGAC GGCCTAAAC
 3061 TCTTTATTAA AGAACCCGT CCGCCTCCA CGCGATCCCC CGTTCTGTT CTCCTGGCC
 3121 CTATACTCGC ACTCACACTG GCCCTGACCC TTGAGGCC GATGCCCTTC CGTACCCCTG
 3181 TTGTGGACCT TAACCTGGT ATCCCTTTA TTCTAGCACT ATCTAGCCTT GCAGTATAC
 3241 CTATTCTGG GTCTGGTGA GCATCCAATT CAAAATACGC TTAGTAGGA GCACTGCGGG
 3301 CTGTTGCACA AACAAATTCT TACGAAGTTA GCCTAGGACT CACCTCCCT AACATCATCA
 3361 TCTTTACCGG AGGATTACA CTTCAGACCT TTACACACGGC CAAAGAGGCC ATCTGACTAG
 3421 TTGTACCCGC CTGACCCCTC GCTGCCATAT GATACATCTC CACTCTTGCC GAAACAAACC
 3481 GTGCACCCCT CGACCTAACAA GAGGGGAAT CTGAACCTGT TTCAGGCTTC AACGTTAGAAT
 3541 ACGCAGGAGG GCCCTTCGCT TTGTTCTTT TAGCCGAATA CTCAAATATC CTCCTAATAA
 3601 ATACTCTGTC CGCAACATTA TTTTGGGGG CCTCCACAT TCCCTCTATC CCAGAATTAA
 3661 CTACCAATTAA TATTAAACT AAAGCAGCCC TTCTCTCAAT TGTTTTCTG TGAGTCCGAG
 3721 CTTCTGTACCC GCGATTCCGT TACGACCAAT TTATGACACT CATTGAAAG AACTTTCTTC
 3781 CACTGACTTT AGCACTAGT ATTGACACT TAGCGCTCCC TATTGCAATC GCTGGCTCC
 3841 CACCACAACT ATAAACCGAG GAGCTGTGCC TGAATTAAA GGGCCACTTT GATAGAGTGA
 3901 ATTATGGGGG TTAAAATCCC CCCAACCTCT TAGAAAGAAG GGGATCGAAC CCAACCTGAA
 3961 GAGATCAAA CTCTCGTGC TTCCCTTACA CCACCTCCTA GTAAGGTCAG CTAAACAAGC
 4021 TCTTGGGCC ATACCCCAAC CATGAGTT CAAATCCTTC CTTTACTAAAT GAAACCCCTAC
 4081 ATCTTGACCA CTCTTATATT TGTTGGCC CTAGGAACAA CACTTACGTT TGCAAGCTCA
 4141 CACTGACTCC TCGCTTGAAT AGGCCTCGAA ATTAATACAC TAGCCATTAT TCCACTAATA
 4201 GCCAACATC ATCACCCCG AGCAGTCGA GCTACTACTA ATACTTTCT CGCACAAAGCC
 4261 ACAGCAGCCG CTACCCCTT GTTGTCAAGT ACCACCAACG CATGACTAAC AGGCCAATGA
 4321 GACATCCAAC AAATAACACA CCCCCCTCCCC ACAACAATAA TTGTTATTGC CCTAGCACTA
 4381 AAAATTGGAC TAGCACCAAT ACACCTTGA CTTCCAGAAG TGCTTCAAGG ACTAGACCTC
 4441 ACCACCGGAC TTATCTCTC GACTTGACAA AAACCTTGAC CTTTGCCTT ATTACTACAA
 4501 ATTCAAGCAA CAAACCTAC GCCTCTAATC ATTATTGGCT TGCTTACCAAC CCTCGTTGGA
 4561 GGATGAGGTG GTCTCAATCA AACTCAATT CGTAAAGTCC TCGCTACTC CTCGATTGCC
 4621 CATCTGGCT GAATAATACT AATCATTCAAG TTTTCACCC CCCTCGCCCT CTTGCCCCTA
 4681 CTCACGTATT TTACTATGAC TTTTCAGCA TTCTTAATCT TAAAGTAAA CAAGGCTACC
 4741 ACTGTCAATG CCCTCGCAAT CTCATGAACA AAAACCCCTG CCCTCACAGC CCTAGCACCC
 4801 CTAGTCTTAT TGTCCTTGG TGGCCTCCCT CCCCTTACCG GCTTCATACC CAAATGATT
 4861 ATTCTTCAAG AACTGACTAA GCAAGACCTC CCAGCCCTTG CTACCCCTCG CGCATTAACT
 4921 GCCCTCTTAA GCCTTTACTT CTACTGCGC CTCTCGTACCG CGATAAACCTT GACAATGTT
 4981 CTCACAAAC TCGTGGCGT TACCCCTGA CGGTTCTATT CTCCCCAATT TACGCTTCCC
 5041 CTCGCCATGT CCACCGCAG AACCCACCTT CTCTCCCAAC TAGCCCCCTG TGCTGTAGCA
 5101 CTACTCATAG CTTAAGGGAC TTAGGCTAGC ACTTAGACCA GTGGCCTTC AAGCCATCAG
 5161 CGTGAGTGAA AATCTCTAG TCCCTGTTAA GACTTGCAGG CTATTATCCC ACATTTCTG
 5221 CATGCAAAGC AGACACTTA ATTAAGCTAA AACCTTTCTA GATAGGAAGG CTTGATCCT
 5281 ACAAAATCTT AGTTAACAGC TAAGGCCAATCAGCAG CATCTATCTA CTTTCCCCCG
 5341 CTTGATTAC CGAGTAAAAA GCGGGAGAAA GCCCGGAG GCGATTAGCC TGCTTCTTAA
 5401 GATTTGCAAT CTAACATGTA ACACCCCGA GCTGTAAAG AAGAGGGCTT GCACCTCTGT
 5461 CTATGGGTCT ACAATCCACC GCTTAATCA GCCATCCTAC CTGTGGCAAT CACACGGTGA
 5521 TTTTCTCGA CCAATCACAA AGACATCGGC ACCCTCTATC TCGTATTG TGCTGAGGCC
 5581 GGTATAGTGG GAACAGGCCT AAGTCTGCTT ATTCCGGCAG AACTAAGCCA ACCTGGGGCT
 5641 CTCTTGGAG ACGACAAAT TTATAACGTA ATCGTACCG CACACGCCCT TGTAATAATC
 5701 TTTTTATAG TAATACCCAT TATGATCGGG GTTTCGGAA ACTGGCTTAT TCCACTAATA
 5761 ATTGGAGGCC CAGATATGGC TTCCCTCGA ATAAATAACCA TGAGTTCTG ACTTCTTCCC
 5821 CCATCTTTC TTCTCTCTT AGCCTCTTA GGTGTTGAAG CTGGGGCAG TACGGGGTGA
 5881 ACCGTTTAC CACCACTAG TGGAATCTG GCCCAGCAGG GAGCATCCGT TGACCTAAC
 5941 ATCTCTCAG TTCACCTTCG AGGGATTCTG TCAATTCTGG GGGCAATTAA CTTTATTACT
 6001 ACCATCATCA ACATGAAACC CACAACAGTT ACTATGTACCA AAATCCCTT ATTGTTGAA
 6061 GCCGTTCTAA TTACCGCCGT ACTTCTCTT CTGCTCCCTC CCGCTTACG CGCAGGGATT
 6121 ACAATGCTAC TAACAGACCG CAACCTAAC ACAACCTTT TTGACCCCGC CGGAGGTGGT
 6181 GACCCCATCC TCTATCAAC CCTATTCTGA TTCTTTGGCC ACCCAGAGGT ATATATTCTC
 6241 ATTCTTCCAG GCTTCGGAAT AATTTCACAT ATTGTTGCAT ACTATGCGAG TAAAAAAAGAA
 6301 CCTTTGGTT ACATAGGAAT AGTCTGAGCT ATAATAGCCA TTGGACTCCTT GGGATTCTT
 6361 GTATGGGCC ATCACATGTT TACAGTGGGG ATAGACGTG ACACACGAGC CTACTTACA
 6421 TCTGCCACAA TAATCATTGC AATCCCGACC GGCCTAAAG TCTTAGCTG ACTCGCAACC
 6481 CTCCACGGGG GAAGCATCAA ATGAGAAACG CCACCTCTT GGGCCCTCGG CTTTATTTC
 6541 CTCTTACAG TAGGGGGTCT TACTGGCATT GTCTTGGCTA ACTCCTCTCT TGACATTGTT
 6601 CTGCATGACA CATACTATGT AGTAGCCAC TTCCACTATG TACTATCTAT GGGTGCTGTA

6661 TTGCGCAATCG TTGCCGCCTT CGTCCACTGA TTCCCACAT TTACAGGTTA TACCCCTCAC
 6721 TCCACATGAA CAAAATCCA CTTCGGCCTG ATATTATTC GGGTCAATCT AACATTCTC
 6781 CCCAACACT TCCTGGCCT AGCCGAATG CCCCGACGAT ACTCAGACTA CCCAGACGCA
 6841 TATAACCTTT GAAACACTGT CTCATCAATC GGATCCCTAA TGTCGCTCGT CGCTGTAATC
 6901 TTATTTTGT TCATTATTG AGAACGATTT ACAGCCAAAC GAGAACGTTGG GGCAGTAGAA
 6961 CTTACTGCAA CTAACATCGA ATGACTTTAC GGCTGCCCTC CACCCCTACCA CACATTGAA
 7021 GAGCCCCCAT TCGTCCAAGT TCGAATAAT TCGAACAAAC TAACGAGAAA GGGAGGAGTT
 7081 GAACCCCAT CAATTGGTT CAAGGCAACC ACATAACCGC TCTGTCACCTT TCTTCACAAC
 7141 ACaCCAATAA GATACTGTA AAACGCTATA ACACTGCCTT GTCAAGGAG AATTGTGGT
 7201 TAAATCCCG CGTATCTAA ACGAAATGGC ACATCCCTCA CAACTCGGAT TTCAAGACGC
 7261 AGCTTCACCC TTAATAGAAG AACTACTTCA CTTCCACGAT CACGCCCTAA TAATTGTAAT
 7321 TCTCATCAGC ACGATAGTAC TTTATATTAT TGCGGCTATA GTTACAGCCA AACTAACAGA
 7381 TAAACTGTT CTAGATTCTC AAGAAATTGA ATTATCTGA ACAGTTCTCC CAGCTGTTAT
 7441 TCTTATTCTC ATCGCCCTAC CATCTCTCG AATTGGTAC TTAATAGACG AAATTAATGA
 7501 TCCCCACTTA ACAATCAAAG CCATGGGCA CCAGTGGTAC TGAAGCTACG AATACACAGA
 7561 CTATGAGAC CTCGGTTTG ACTCTTACAT AATCCCTACA CAAAGACCTGA CCCCTGGACA
 7621 GTTCCGACTA CTTGAAGCAG ATCACCGGAT AGTAATCCCG GTAGAATCTC CTATTGCGT
 7681 TCTAACTCTA GCTGAAGATG TATTACATC CTGAGCAGTC CCCTCCCTGG GCGTAAAAGT
 7741 TGACGCCGTA CCTGGACGAT TAAATCAAGC AACCTTTATT GTTAGCCGGC CAGGCGTATT
 7801 TTATGGTCAA TGCTCTGAAA TTGCGGAGC GAACCATAGT TTTATACCA TCGTTGTTAGA
 7861 AGCAGTCCCC CTTGAACACT TTGAAAATG GTCTTCACTA ATAATTGAAG AGCCTCGCT
 7921 AAGAAAGCTAA TAAGTACAAG CGTTAGCCTT TTAAGCTAAA GACTGGTGCCTA TAACAACAC
 7981 CCTTAGCGAC ATGCCCTAAC TGAACCCCGC ACCCTGATTT GCAATTCTAG TTTTCTCTTG
 8041 AATAATCTTT TTAACTGTT TTACCCCCAA AGTTTAGCA CATACTTAC CCAATGAACC
 8101 TACCTCCCAA AGCACACAA AACCTAAAC AGAACCCCTGA AACTGACCAT GATACTAAGC
 8161 TTTTTGACC AATTATATAC CCCAGTATAC CTCGGTATTG CGCTAATTG ACTAGCAATT
 8221 ACCCTCCCCCT GAATTCTGTT CCCAACCCCT CAAGCCCGCT GATTAAACAA CCGTGTGCTA
 8281 ACACCTCAAG GATGGTTAT TAGCCGCTTC ACCTCACAAC TTCTTCTTCC CCTAAACCCC
 8341 GGGGGCCACAA AATGAGCAGT TTTATTTAC TCATTAATGG TTTTTCTTATT TTCTATTAAAC
 8401 ATACTAGGAC TCCTTCCATA TACCTTACCA CCAACAAACCC AACTCTCCCT TAATATGGC
 8461 CTCGCAGTTC CCCTCTGATT AGCAACTGTC ATTATGGAA TGCGAAATCA ACCTACACAT
 8521 GCCCTAGGTC ACCTCCCTTC AGAAGGCCAC CCTACTGCTC TCATCCCTGT ACTCATCATC
 8581 ATTGAAACAA TTAGCCTATT TATTGACCC CTCGCTCTCG GTGTTGACT AACAGCGAAT
 8641 CTCACAGCAG GCCACCTCT TATTCAACTC ATTGCAACAG CTGCCTTGT CCTTCTCCG
 8701 CTCATACCTG TTGTTGCTAT CTTAACACG GTGGCCTCT TCCTTCTTAC TCTGCTAGAA
 8761 GTCGGGTCG CCATGATCCA AGCCTATGTC TTGTTGCTAC TCTTAAGCCT ATATCTACAA
 8821 GAGAACGTTT AATGGCCCAT CAAGCACACC CATAACCAT AGTCGACCC AGTCCATGAC
 8881 CCCTCACGGG GGCTATTGCT GCCCTATTGA TAACATCTGG CCTGCTATC TGATTCACT
 8941 TCCATTCCAC AACCTTATA aCTATTGGG CAGTCCTTCAC CCTCTTAACA ATCTTCAAT
 9001 GATGACGGGA CGTCGTACGA GAAGGCACAT TTCAAGGACA CCACACTCCC CCTGTTCAAA
 9061 TAGGCCTCCG ATACGGAATA ATCTTATTAA TTACCTCAGA AGTTCTGTT TTCTTAGGCT
 9121 TCTTCTGAGC CTTTACAC TCAAGCCTAG CACCTACTCC AGAACTAGGA GGCTTCTGAC
 9181 CCCAGCAGG CATTACCCCT TTAGACCCAT TTGAAAGTCCC GCTTCTTAAC ACAGCAGTCC
 9241 TCCCTGCCTC TGGTGTAACT GTTACCTGAG CACATCACAG CATTATAGAG GGTAACACGGA
 9301 ACAAACAT TCAATCTT GCCCTACAA TCTTAACCGG GGGTTACTTC ACTTTCTCC
 9361 AAGGCCCTGA GTATGTCGAA GGGCCCTTC CCATTGCAAGA CGGAGTTTAC GGTGCTACAT
 9421 TCTTGTGTC CACCGCTTC CACGGCCCTC ACGTCTTAAT TGGCACCTCA TTCTTAGCCG
 9481 TTTGTCTGTT ACGCCAATC CTTCACCAATT TTACATCAAA CCACCACTTT GGGTTGAAG
 9541 CAGCCGATG ATACTGACAC TTCGTAGACG TCGTCTGACT TTTCCTCTAT ATCTCTATTT
 9601 ACTGATGAGG ATCTTAATCT TCCTAGTATC AAATCTAGTA TAAGTGAACCT CCAatCACCT
 9661 GGTCTGGTT AAAGTCAAG GGAAGATAAT GAGCCTCTT CTAACCATCA TTCAATCGC
 9721 CACCCCTCTC TCAACAGTAC TTGCCATTG ATCCCTTGA CTCCCCAAA TTACACCAGA
 9781 CCATGAAAG CCATCACCATC ACGAATGCCG CTTGACCCG ATAGGCTCTG CCCGACTACC
 9841 TTTTCGCTG CGATTTTTC TCATCGCCAT CCTCTTTCTC CTGTTGACT TAGAGATTG
 9901 ACTCCTCCTC CCCCTCCCT GAGGAGACCA ATTACGATCG CCACTACTCA CATTCACTTG
 9961 AGCTACAGCT GTCCTAGCGC TTCTAACCT TGGCCTCATT TACGAATGAA TGCAAGGAGG
 10021 CCTAGAGTGA GCTGAATAGG TGGTTAGTCT AAGAAAAACA TTGATTTCG GCTAAAAAAC
 10081 TTGTTGTTA AATCCCAAC CGCTTAATGA CCCCCACACA CCTTGCTTCC TCCTCAGCCT
 10141 TTCTTCTGGG TTTAACAGGC CTGGCATTTA ACCGGTTCC CTCCTTTCC GCTTTATTAT
 10201 GTCTTGAAGG AATAATAC TCCCTATTAA TTGCCCCTCTC CTTATGAACA CTCCAACCTG
 10261 ACTCAACAA CTTTCTAGCA TCTCTTATGC TTCTACTTGC ATTTTCAGCC TGCAGGCAA
 10321 GCGCCGGCCT CGCCCTGCTA GTAGCCACTG CTCGAACCCA CGGAACCGAC CGACTACAA
 10381 GCCTAAATCT TCTCCAATGC TAAAATCCT AATTCACAA CCTATGCTAA TCCCAACAGC
 10441 CTGGTTACTC AAACCCAGCT GACTCTGGCC CATAACCTTG ATGTATAGCT TTTGCATCTC
 10501 CTTAAATTAGC CTTTCTATGC TGAAGAACCT CTCAGAACCC GGCTGATCAT CACTTAACCT
 10561 CTTTATAGCC ACCGACTCCC TATCAACCCCT CTCCTCTGTT CTCACATGCT GGCTGCTACC
 10621 ACTAATGATT TTGGCAAGGG AAAAACACAC AGCCTCAGAG CCTCTGGTC GCCAGCCGAT
 10681 GTACATTACA CTTCTGCTC CACTCCAGTT CTTCTGATT CTAGCATTAA GCGCCACTGA
 10741 GCTGTTAATG TTTTACGTA TATTGAAAGC CACCCCTCATC CCCACTCTGA TTATTATTAC
 10801 CGCTGAGGG AACCAAACAG AACGCCCTAA TGCAAGGACC TACTTTCTTT TTTACACATT
 10861 GGCAGGCTCA CTACCCCTTC TCGTTGCTT ACTCTTCTC CAGAACACAT CGGGCACCC
 10921 GTCGGTGCCTA ACCCTACACT ATACAGACCC ACTCACTCTC TCATCTTATG CAGACAAATT
 10981 ATGATGAGCA GGCTGTCTC TAGCATCCT CGTAAAATA CCTCTCTATG GAGTCCACCT
 11041 GTGACTCCCT AAAGCCCACG TTGAAGCCCC CATGCGAGGC TCAATGATCC TTGCAAGGGT
 11101 TCTCTAAAAA CTGGGAGGAT ATGGCATGAT CCGCATAATA ACAATATTAG AGCCTCTAAC
 11161 TAAGGAGCTA AGCTACCCCT TCATTGCTT TGCACTTTGG GGGGTAATCA TAACTGGCTC
 11221 AATTGTCATA CGACAGACAG ACCTTAAGTC CCTAATTGCC TACTCATCTG TAAGTCACAT
 11281 GGGCTAGTG GCGGGGGAA TTCTTATTCA GTCACCCCTGG GGCTTAACAG GGGCCCTTAC
 11341 CCTAATAATT GCACACGGTC TTACCTCATC AGCCCTCTTC TGCCTAGCAA ACACAAACTA

11401 TGAACGAACC CACAGTCGAA CGATAGTACT AGCACGAGGG CTTCAAGTAG CTCTACCAT
 11461 GATGGCCACT TGATGATT A TTTCTAGCTT AGCCAACCTA CCCCTTCCAC CACTACCTAA
 11521 CCTCATAGGT GAACTAATAA TTATTACTTC TCTATTCAAC TGATCCTGGT GAACCTAGC
 11581 ATTAACCGGG GCAGGCACTC TAATTACAGC TGGTTACTCC CTTTATATGT TCTTAATAAC
 11641 TCAACGGGGA CCTCTCCCAG CACATATCAT TGCGCTTGAG CCATCTCACA CCCGAGAAC
 11701 CCTCCTTATT GCACTTCATC TTATCCCCTT GATTCCTCCTC GTGCTTAAGC CCGAGCTGAT
 11761 CTGAGGTTGG ACTGCCTGTA GGTGTAGTT TAACAAAAAC ATTAGATTGT GATTCTAAA
 11821 ATAAGGGTTA AAATCCCTC TCCCACCGAG AGAGGCTCG AGCAATGGGA ACTGCTAATT
 11881 CCCATGACCT TGGTGGACC CCCAGCTC CTCGAGAAGC TCCTAAAGGA TAACAGCTCA
 11941 TCCGTGGTC TTAGGAACCA AAAACTTGTG TGCGAAATCC AAGTAGCAGC TATGCACCC
 12001 ACCTCCCTAA TGATTTCTG AAGCGTATC ACAATCTTG CATTGCTAGC CTATCCCC
 12061 GTCACCACAA TACGACCCAC GCCCCAAGAA CCTCAGTGGG CTTCTACCCA TGTTAAA
 12121 GCTGTAAAAA TAGCTTCTT TGTTAGTCTG TTACCCCTGA CCCTATTTCT TAATGAGGG
 12181 GCGGAGACAA TCGTTACTAA CTGAACCTGA ATAAATACTA ACTCCTTCGA CATCAACATT
 12241 AGCTAAATG TTGACTTCTA CTCAATTATT TTTACCCCCA TTGCCCCCTA CGTAACTTGG
 12301 TCCATCCTAG AATTGCAATG ATGATACATG CATGCAGACCA AAACATAAA CGCGCTTCTC
 12361 AAGTATTCAC TGACATTTCT GATTGCCATC ATCGTTCTCG TAACCGCAA CAATATGTT
 12421 CAACTATTCA TCGGTTGAGA GGGGTGAGGG ATCATATCGT TCCTCCTTAT CGGGTGGTGG
 12481 TTCGGACGGG CGGATGCTAA CACTGCAGCA CTTCAAGCAG TACTTTATAA CCGAGTTGGG
 12541 GACATCGGTC TTATCTTCG CATGGCTGTA ATAGCAGCAA ACCTAAACTC CTGGGAAATG
 12601 CAACAAATCT TTGTAACCTC CAAAGACATA GACTTAACCTT ATCCTCTCCT CGGGCTCATC
 12661 ATCCGAGCAA CCGGAAGTC GGCTCAATTG GGTCTTCAACC CTCGACTACC CTCCGCTATG
 12721 GAGGGTCTCA CACCGCTCTC TGCCCTACTT CATTCTAGCA CTATAGTCGT TGCGGGCATT
 12781 TTCTGTGGT TCGGAATGAG GGGCCCTCTG GAAAACAACC AAACCTGCCCT CACAACCTGC
 12841 CTGTGCTCG GGGCCCTAC AACACTTTG ACTGCAACCT GTGCCCTCAC AAAAAATGAT
 12901 ATTAAAAAAA TCGTCGCATT CTCCACATCA AGCCAGCTTG GACTTATGAT GGTAACCATC
 12961 GGACTAAACC AACCACAACT AGCGTTCTT CACATTGCA CACACGCTTT CTTCAAGGCC
 13021 ATACTATTTT TATGCTCCGG ATCTATTATT CATACTCTTA ATGACGAACA AGACATCCGA
 13081 AAAATAGGAG GCATACACCA CCTTGCCTT TTTACATCTT CTTGCCTAAC TATTGGCAGT
 13141 CTTCCCTTA CAGGGGGCCC CTTCTTAGCT GGCTTCTTCT CTAAAGATGC TATCATCGA
 13201 GCACTAAACA CATCCCACCT AAACGCCATG GCCCTAACTT TAACCCCTCCT AGCCACCTCT
 13261 TTACAGCAA TTACAGCTT CCGAGTAGT TTTTTGTC CCATGGGTCA CCCCCGGTTT
 13321 AATCCCTCT CCCCCATCAA CGAAAATAAC CCAGCAGTGA TTAACCCACT TAAGCGATTA
 13381 GCATGAGGAA GCATGTTGC AGGCTTAACCTA ATTACTCTA ACATTACACC CCTAAA
 13441 CCTGTTATAT CTATAACCCC TCTCCTTAA CTAGCTGCTC TTGAGCTAAC AATCCTTGGC
 13501 CTACTTGTG CTATGAACT CGCCATGCTA ACTAATAAGC ATATAAAATC AATACCGAC
 13561 ACACACGTAC ACCACTCTC CACCTGTGA GGCTTCTTCC CGGGAGTAAT GCACCGCCTA
 13621 GCCCCTAAAT TACTCTTCTT CCTAGGCCAA AACGTTGCCA ACCAACAGT AGATCAA
 13681 TGGCTAGAAA AAATCGGCC CAAAGCATT GAGTCTCTA ACCGACCCCT GATCTCTTCT
 13741 ACAAGTAACA TTCACTGGG AGTAATCAAG ACATATCTA CCCTTTCTCCT CCGACACTA
 13801 GCCCTATAA TCCCACATT CATCGCTAA ACTGCCCGCA ACGCCCTCG ACTTATA
 13861 CGAGTCACACT CGAGAACCC AAACAAAGTT AAAAGCAATA CCCACGCACT AATAATTAA
 13921 AACCACCCCC CAGAAGATA TATTAAGCT ACACCTCCGA TATCCCCCCC AACACGGAG
 13981 AACTCCGCA ACTCATCTGC TAGGCCAA GACGACTCGT ATCACCCCCC TCAGAACAGC
 14041 CCAGAAACTA ACATTACCCC CCCTGTATAA ACTACCCCCC AAACCAATTAC CGACCAACTA
 14101 CCCCAGCTCT CCGGATACGG CTCAGCAGCT AACCGCTGCG AGTATGCAA TACA
 14161 ATCCCCACCA AGTAATTTA GAACAAAAGT AGTGACAAAAA AAGGCCCCC ATGCCCAACT
 14221 AATaTACAC ACCCCATCCC AGCTACTACA ACCAACACCA AAGCAGCGAA ATAAGGAGAT
 14281 GGATTAGAAG CAACTGCAAC TAACCTAAC ACAGAGAGAA ATAAA
 14341 AAAGTCATAA TTCTGCCAG GACTTTAACC AGGACTAATG GCTTGAAAAA CCACCGTTGT
 14401 TATTCAACTA CAAGAACCTT AATGGCCAGT CTACGTAAT CCCCACCTCT TCTAAA
 14461 GCAACAGTATG CTTCAGTGA CCTCCAGCG CCCTCTAATC TCTCCGTTTG GTGAAAC
 14521 GGGCTCTTC TAGGACTCTG TTTAGTAAAC CAAATGCTA CGGGCTTGT TCTAGCCATA
 14581 CACTATACAT CAGATATTG TACTGCCCTC ACCTCCGTGG CTCACATCTG CGGGGACGTC
 14641 AACTACGGCT GACTTACCG AAGCATTCA GCAACAGCGC CATCATTCTT TTTCATTG
 14701 ATCTACCTAC ACATCGCCG AGGACTCTAC TATGGCTCAT ACCTCTATAA AGAAACATGA
 14761 ACTATTGGAG TCGTCTCTT CCTCCTCGTA ATAATAACGG CCTTCGTCGG ATACGTCTC
 14821 CCTTGAGGAC AAATGTCGTT TTGAGGTCA ACCGTTATTA CCAACCTCCT ATCCGCC
 14881 CCCTATGTTG GCAACACCC AGTCCAAATGG ATTGAGGTG GTTTTCTGT AGACAATGCT
 14941 ACCCTTACCC GGTTCTTGC CTTCCACTC CTTTCCCCCT TCATCATCGC CGCCGCAACA
 15001 GTAATTCACTC TACTATTCTT TCACGAAACT GGCTCAAATA ATCCCACCGG CTTAA
 15061 GACTCAGACAA AGGTTCCCTT CCACCCCTAC TTCACTTACA AAGACCTT CTTGTTG
 15121 GTGCTCTCA CTGCATTAGC ATCCCTCGCT TTGTTTCAAC CTAATCTTT AGGAGACCA
 15181 GATAACTTTA CCCCCGCAAA TCCACTGTT ACACCCCCAC ACATTA
 15241 TTCTCTTTG CTTACGCTAT TCTCCGGTCC ATCCCAACAA AACTTGGAGG CGTTCTGCC
 15301 CTTTATTTT CGATCTTAGT TCTCATGCT GTCCCCCTTC TTCACACCTC TAA
 15361 AGCCTAATGT TCCGCCCTAC AACGCAATTG TTGTTCTGGT CTCTGGTAGC AGAC
 15421 ATTCTAACTT GAATTGGAGG AATGCCGTAA GAAACACCCCT CGTTTATCAT TGG
 15481 GCATCACTCC TTAACTCTGC CCTCTTCTGT GTCTGATCC CAA
 15541 AATAAAAGTCC TTGGATGAAA ATGCACTAGT AGTCAGCGC ACAGAGCCCC
 15601 ACTGGCCGTC GGAGGTTAAA ATCCTCCCTA CTGCTCAAAG AAAGGAGATT
 15661 CCCCTAACTC CCAAAAGCTAG GATTCTAGCA TAA
 15721 AAAACAAAGG ATTTTATAT ACATGATGT ATTAACACCA TAT
 15781 TGTGTAATGT ACTAGGACAT ACATGATAA TCACCTAA
 15841 ACCACCAATT TTAAC
 15901 GAAATTCCAG GAC
 15961 CAAGACTCAA AATTCACTC ACTTAAAAA
 16021 TCCCTTAATG ATA
 16081 GAATTATTCC TGGCATTGG T

16141 TTTCATCGAC ACTTACATAA GTTAATGTTG ATAATACATA CGACTCGTTA CCCACCAAGC
 16201 CGGGCGTTCA CTCCAGCGG TAAGGGGTTC TCTTTTTTT TTTCTTTCA CTTGACATT
 16261 CAAAGGCAT CCAGCCAACA AAAACGTTCA AAGGGGGAGC ACTTTTCTTG CACGCGGCGT
 16321 ACATAGTATC CATGTATTA GTCTTTATTAA AAAAATAAC ATAAAGGGAT ATCATGGCA
 16381 TAAAATGTG CTCCCTTCTT CTATCATCCC CAGGATACCC CTTTTTTCG CGTAAAACCC
 16441 CCCTACCCCC CTAAACTCCT GAAGTTGCTA AAACCCCTGA AAACCCCCCG GAAACAGGAC
 16501 AAACCTCTGG TAGCTTAAA AAGGCAATT CTTTCACC TACTTTCAA CCCTAAAATA
 16561 CCCCTGACCA AACATAAGCC CACCAGTTAT TTTAATAATG GTACTTTCA ACCCTAAAAT
 16621 ACCCCTGGCC AAAGATAAGC CCACCAGTTA TTTAATAAT GTACTTTTC AACCTAAAAA
 16681 TACCCCTGAC CAAAGATAAG CCCACCAGTT ATTATAATAA TGTTACTTT CAACCTAAA
 16741 ATACCCCTGA CAAAGATAA GCCCACCAAGT TTTTAATAA ATGGTACTTT TCAACCTAA
 16801 AATACCCCTG GCCAAAGATA AGCCCACAG TTATTTAAT AATGGTACTT TTCAACCTA
 16861 AAATACCCCT GGCCAAAGAT AAGCCACCA GTTATTTAA TAATGGTACT TTCAACCT
 16921 AAAATACCCC TGGCAAAGA TAAGCCCACC AGTTATTTTA ATAATGGTAC TTTCAACCC
 16981 AAAATACCC CTGGCAAAG ATAAGCCCAC CAGTTATTTT AATAATGGTA CTTTTCAACC
 17041 CTAAAATACC CCTGGCCAA GATAAGGCCA CCAGTTATTT TAATAATGGT ACTTTCAAC
 17101 CCTAAAATAC CCCTGGCCA AGATAAGCCC ACCAGTTATT TTAATAATGG TACTTTCAA
 17161 CCCTGGCCA AAGATAAGCC CACCAGTTA TTTAATAATG GTACTTTCA
 17221 ACCCTAAAAT ACCCCTGGCC AAAGATAAGC CCACCAGTTA TTTAATAAT GTACTTTTC
 17281 AACCTAAAAT TACCCCTGGC CAAAGATAAG CCCACCAGTT ATTATAATAA TGTTACTTT
 17341 CAACCTAAA ATACCCCTGG CCAAAGATAA GCCCACCAAGT TTTTAATAA ATGGTACTTT
 17401 TCAACCTAA AATACCCCTG GCCAAAGATA AGCCCACAG TTATTTAAT AATGGTACTT
 17461 TCTAACCTAA AAATACCCCT GGCCAAAGAT AAGCCACCA GTTATTTAA TAATGGTACT
 17521 TTTCAACCTT AAAATACCCC TGGCAAAGA TAAGCCCACC AGTTATTTTA ATAATGGTAC
 17581 TTTCAACCC TAAAATACCC CTGGCAAAG ATAAGCCCAC CAGTTATTTT AATAATGGTA
 17641 CTTTCAACC CTAAAATACC CCTGGCCAA GATAAGGCCA CCAGTTATTT TAATAATGGT
 17701 ACTTTCAAC CCTAAAATAC CCCTGGCCA AGATAAGCCC ACCAGTTATT TTAATAATGG
 17761 TACTTTCAA CCCTAAAATA CCCCTGGCCA AAGATAAGCC CACCAGTTAT TTAATAATG
 17821 GTACTTTCA ACCCTAAAAT ACCCCTGGCC AAAGATAAGC CCACCAGTTA TTTAATAAT
 17881 GTACTTTTC AACCCTAAA TGACTGGGAC CTGTAACGCC GTGTACATCTT CTTCATGATC
 17941 AAAAAACTGC TGTTTAGATT ATTCAAGTA TTCTAGCAT CATCCCAACT ACGAAGGGCA
 18001 CAAACAAAAA ACCGAGAAGA CTTGCTCTC CATTCAAAAA CTATTTTTT TTAACCGACA
 18061 AGCCCTAAA CCAACCTCT ATTATCCTAA AACTTGACAA GACTAGGGAA ACACCTTTG
 18121 GTGTAACAGA TTTTAACT

Reinhardtius hippoglossoides 3 17895 bp

1 GCTAACGTAG CTTAATTTAA GCATAACACT GAAGATGTTA AGATGGGCC TAGAAAGCCC
 61 CGGGAGCACA AAGGCTTGGT CCTGACTTTA CTGTCGACTT TAACCTAAACT TACACATGCA
 121 AGTATCGCC CCCCTGTGAG AATGCCACA ACTCCCTGCT TGGGAACCTAG GAGCTGGTAT
 181 CAGGCACAAG CCCCGTAGC CCACGACGCC TTGCTTAGCC ACACCCCTAA GGGAACTCAG
 241 CAGTGATAAA TATTAAGCCA TAAGTGAAAAA CTTGACTTTAG TTAAGGTTAA GAGGGCGGT
 301 AAAACTCGTG CCAGGCCACCG CGGTTATACG AGAGGCCAA GTTGACAAAC AACGGCGTAA
 361 AGAGTGGTTA GGGGATTAC TAAACTAGAG CGAACGCTT TCAGAGCTGT TATACGCACC
 421 CGAAAGTATG AAACCCAAATT ACGAAAGTAG CTCTACTCAT CTCGAACCCA CGAAAGCTAA
 481 GAAACAAACT GGGATTAGAT ACCCCACTAT GCTTAGCCCT AAACATCGAT TGCACTATAC
 541 ACTCCATATC CGCCCGGGAA TTATGAACGT CAGTTAAAAA CCCCAGGAC TTGGCGGTG
 601 TTAACATCCA CCTAGAGGAG CCTGTTCTAG AACCGATAAC CCCCCTTAA CCTCACCTC
 661 TCTTGTNTA TCCGCCTATA TACCACCGTC GTCACTTAC CCTGTGAAGG CTTTACAGTA
 721 AGCAAATTC GCACAGCCA AAACGTCAAGG TCGAGGTGA GTGAATGAGG GGGGAAGAAA
 781 TGGGCTACAT TTGCTAAACA TAGAAACAC GAATGTTGCA TGAAACATG CAACTGAAGG
 841 AGGATTTAGC AGTAACCGAG AAATAGAGCG TCCCGCTGA ACTGGCCCTA AAGCGCGCAC
 901 ACACCGCCCG TCACCCCTCC CAAGCCCCCT GAACAAACT ATTAAAAACCC CAACACCCG
 961 CGAAGGGGAG GAAAGTCGA ACATGGTAAG TGTACCGGG GGTGCGCTTG GAAAAATCAG
 1021 AGTGTAGCTA AGTTAGATTA CAGCATCTA CTTAACCGGA GAAGACGCC GTGCAAGTCG
 1081 GACCACCTG ACGCTTATA GCTAGCCAA CCCCTTAACA CAACAAACCC CCATTTATAA
 1141 CCCCTAAAGC ACGAACACC CACGTAGCTA AACCAATTCTC CCCCCTAAGT CCAGGCGATA
 1201 AAAAAGGAAA TTTGGAGCA TAGAAAAAGT ACCGCAAGGG AAAGCTGAAA GAGAGATGAA
 1261 AAAGCCAGT AAAGCTTAAG AAAGCAGAGC TTAAAGCTCG TACCTTTTG ATCATGATTT
 1321 AGCTAGCACT TTCAAGCAA GAGAACCTAA AGTTGTAAAC CCGGAAACTG AGTGAGCTAC
 1381 TCCAAGACAG CCTATTAAAGG GGGCAACCC GTCTCTGTGG CAAAAGAGTG GGAAGAGCTT
 1441 TGAGTAGAGG TGACAAACCT ACCGAACATCA GTTATAGCTG GTTGCCTGTG ATTGAATAG
 1501 AAGTTCAGCC CCCTGGCTC TCCACTCATG CACTTTATTA ACCCTTCAGA TGCAATGAGG
 1561 ACCCAGGGGT GTTAGTCAAA GGGGGTACAG CCCCTTTGAT ACAAGACACA ACTTTCTCAG
 1621 TAGGATAAAAG ATCATAATAA AATAAGGACA GATGTTTTAG TGGGCCTAAA AGCAGCCACC
 1681 TTAACAGAAA GCGTTAACAG TCAAACATGA AGCCCTCCCG TTATACCGAT AACCTATCT
 1741 TAATCCCCA CATTAAACAG GCCCTCTTAT GCATCACATA GGAACGACTA TGCTAATATG
 1801 AGATAATAAGA GGGTTAAC ACCCTCTCCT TGACATGTG TAAATCGGAA CGGACCCCC
 1861 ACCGAATCTT AACCGCCCCA ATCAAAGAGG GTATTGGAAA CTACCACAAA TTCAGGCCAG
 1921 AAAAACATCC AATGATAAAC CGTTAACCCCC ACACCTGGTGT GCCCAAAGG AAAGACCGAA
 1981 GGGGGGAGAA GGAACCTGGC AAACATACCC CAAGCCTCGC CTGTTACCA AAAACATCGC
 2041 CTCTTGCATA ACCACAGTAT AAGAGTCCC GCCTGCCAG TGACCATATA GTCAACGGC
 2101 CGCGGTATTT TGACCGTGA AAGGTAAGCT AATCACTTGT CTTTTAAATG AAGACCTGTA
 2161 TGAATGGCAT AACGAGGGCT TAACTGCTCT CTTCCCCCGG TCAATGAAAT TGATCTCCC
 2221 GTGAGAAGC GGGGATTAAA CCATAAGACG AGAAGACCT GTGGAGCTTT AGACACACAG
 2281 GTGGAGGATG TCAAATACCC CAACTAAGG GCCTGAACTA AGTGGAAACCT GCCTTGATGT
 2341 CTTCGTTGG GGCGACCATG GGGAAATCAA AACCCCCACG TGGAAAGGGG GCACACCCCT
 2401 AAGTTACTTC TTCTCCCGCA AGCCAGAGCA ACAGCTCTAA CCAGCAGAAA TTCTGACCAA
 2461 ACTGATCCGG TAAAACCGAT CAACGAACCA AGTTACCTA GGGATAACAG CGCAATCCCC

2521 TTTTAGAGCC CATATCGACA AGGGGGTTA CGACCTCGAT GTTGGATCAT GACATCCTAA
 2581 TGGTGCAGCC GCTATTAAGG GTTCGTTGT TCAACGATTA AAGTCCTACC TGATCTGAGT
 2641 TCAGACCGGA GTAATCCAGG TCAGTTCTA TCTATGACAT GATCTTTCT AGTACGAAAG
 2701 GACCGAAAAG AAGGGGCCA TGCTAAAAGT ACGCCTCACCC CACCTAAAT GAAAAAAATCT
 2761 AAACTAGGCA AAAGGGCATA ACCCTTTGC TGGAGATAAC AGCAAGTTGG GGTGGCAGAG
 2821 CCCGGCTAAT GCAAAAGACC TAAGCCCTT CCACAGAGGT TCAAATCCTC TCCTTAACCA
 2881 TGATTTAAC TCTCATTACC CATATTATTA ACCCCCCTAAC CTTTATTGTA CCTGTATTGC
 2941 TAGCGTAGC ATTCCCTCACC CTCCCTGAAC GGAAAGTCTCCTC AGGCTACATA CAACTCCGAA
 3001 AAGGGCCTAA CATCGTAGG CCTTACGGCC TCCTTCACCC CATTGCTGAC GGCCTAAAC
 3061 TCTTTATTAA AGAACCCGGT CGGGCTTCCA CCGCATCCCC CGGTTCTGTT CTCCTGGCC
 3121 CTATACTCGC ACTCACACTG GCCCTAACCC TTGAGCCCC GATGCCCTTC CGTACCCCTG
 3181 TTGTGGACCT TAACCTCGGT ATCCTTTTA TTCTAGCACT ATCTAGCCTT GCAGTATACT
 3241 CTATTCTGG GTCTGGTTGA GCATCCAATT CAAAATACGC TTTAGTAGGA GCACTGCGGG
 3301 CTGTTGCACA ACAAACTTCT TACGAAGTTA GCCTAGGACT CATCCTCCCT AACATCATCA
 3361 TCTTACCGG AGGATTACA CTTCAACCTT TAAACACGCC CCAAGAGGCC ATCTGACTAG
 3421 TTGTACCCCG CTGACCCCTA GCTGCCATAT GATACATCTC CACTTGTGCC GAAACAAACC
 3481 GTGCACCCCT CGACCTAACCA GAGGGGAAT CTGAACCTGT TTCAAGGCTTC AACGTTAGAAT
 3541 ACAGCAGGAGG GCCCCCTCGCT TTGTTCTTT TAGCCGAATA CTCAAATATC CTCCTAAATA
 3601 ATACCCCTGTC CGCAACATTA TTTTGGCG CTCACCCACAT TCCCTCTATC CCAGAATTAA
 3661 CTAGCATTAA TATTATAACT AAAGCAGCCC TTCTCTCAAT TGTTTTCTG TGAGTCCGAG
 3721 CTTCGTACCC GCGATTCCGT TACGACCAAC TTATGCACCT CATTGAAAG AACTTTCTC
 3781 CACTGACTTT AGCACTAGTT ATTTGACACT TAGCGCTCCC TATTGCACTT GCTGGTCTCC
 3841 CACCACAACT ATAAACCGAG GAGCTGTGCG TGAATTAAA GGGCCACTTT GATAGAGTGA
 3901 ATTATGGGG TAAATACCC CCAACTCT TAGAAAAGAG GGGATCGAAC CCAACCTGAA
 3961 GAGATCAAA CTCTCGTGC TTCTCTACCA CCACCTCCCTA GTAAAGTCAG CTAAACAAGC
 4021 TCTTGGGCC ATACCCCAAC CATGTAGTT CAAATCCTC CTTTACTAAT GAACCCCTAC
 4081 ATCTTGACCA CTCTTATATT TGGTTGGC CTAGGAACAA CACTTACGTT TGCAAGCTCA
 4141 CACTGACTCC TCGCTGAAT AGGCCTCGAA ATTAATACAC TAGCCATTAT TCCACTAATA
 4201 GCCCAACATC ATCACCCCG AGCAGTCGA GCTACTACTA AATACTTTCT CGCACAAGCC
 4261 ACAGCAGGCC CTACCCCTT GTTTCAGGT ACCACCAACG CATGACTAAC AGGCCAATGA
 4321 GACATCCAAC AAATAACACA CCCCCCCCCC ACAACAATAA TGTTTATTGC CCTAGCCTA
 4381 AAAATTGGAC TAGCCAAAT ACACCTGTG CTTCCAGAAAG TGCTTCAGG ACTAGACCTC
 4441 ACCACCGGAC TTATCCTCTC GACTTGACAA AAACCTGAC CTTTGCCTT ATTACTACAA
 4501 ATTCAAGCAA CAAACCTAC GCCTCTAAC ATTATTGGCT TGCTATCCAC CTCGTTGGA
 4561 GGATGAGGTG GTCTCAATCA AACTCAATT CGTAAAGTCC TCGCCTACTC CTCGATTGCC
 4621 CATCTTGGCT GAATAATACT AATCATTCAAG TTTTCAGCC TTCTCAATCT TGAAAGTAAA
 4681 CTCACGTATT TTAATGAC TTTTCAGCA TTCTCAATCT TGAAAGTAAA CAAGGCTACC
 4741 ACTGTCAATG CCCTCGCAAT CTCACTAACCA AAAACCCCTG CCCTCACAGC CTTAGCACCC
 4801 CTAGTCTTAT TGTCCTTGG TGGCCCTCTT CCCCCCTACCG GCTTCATACC CAAATGATT
 4861 ATTCTTCAAG AGCTGACTAA GCAAGACCTC CCAGCCCTTG CTACCTCGC CGCATTAACT
 4921 GCCCTCTAA GCCTTACTT CTACTTGCCTC CTCTCGTAC CGATAACCCT GACAATGTT
 4981 CCTAACAAAC TCGTCGGGT TACCCCCCTGA CGGGTCTATT CTCCCCAATT TACGCTTCCC
 5041 CTCGCCATGT CCACTCGAGC AACCAACCTT CTCTCTCCAC TAGCCCCCTGC TGCTGTAGCA
 5101 CTACTCATAG CTAAAGGGAC TTAGGCTAGC ACTTAGACCA ATGGCCTTCA AAGCCATCAG
 5161 CGTGAGTGAA AATCTCTAG TCCCTGTTAA GACTTGCAGG CTATTATCCC ACATCTTCTG
 5221 CATGCAAAAG AGACACTTTA ATAAAGCTAA AACCTTTCA TAGAGGAAGG CCTTGATCCT
 5281 ACAAAATCTT AGTTAACAGC TAAGCGCCCA ATCCAGCGAG CATCTATCTA CTTTCTCCG
 5341 CTTGATTTAC CGAGTAAAAA GCGGGAGAAA GCCCCGGAG GCGATTAGCC TGCTTCTTA
 5401 GATTGCAAT CTAACATGTA ACACCCAGA GCTGATAAG AAGAGGGCTT GCACCTCTGT
 5461 CTATGGTCT ACAATCCACC GCTTAACCTCA GCCATCCTAC CTGTGGCAAT CACACGTG
 5521 TTTTCTCGA CCAATCACCA AGACATCGGC ACCCTCTATC TGTTTCTG TGCTGAGCC
 5581 GTTATAGTGG GAACAGGCT AAGTCTGTT ATTCCGGCAG AACTAAGCCA ACCTGGGGCT
 5641 CTCCCTGGAG ACGACAAAT TTATAACGTA ATCGTCACCC CACACGCCCT TGTAATAATC
 5701 TTTTTTATAG TAATACCCAT TATGATCGGG GGTTCAGGA ACTGGCTTAT TCCACTAATA
 5761 ATTGGAGGCC CAGATATGGC TTTCCCTCGA ATAAATAACA TGAGTTCTG ACTTCTTCCC
 5821 CCATCCTTC TTCTCTCTT AGCCTCTCA GGTGTTGAAG CTGGGGCAGG TACGGGGTGA
 5881 ACCGTTTATC CACCACTAGC TGGTAATCTG GCCCACGCC GAGCATCCGT TGACCTAAC
 5941 ATCTTCTCAC TTCACCTTGC AGGGATTTCG TCAATTCTGG GGGCAATTAA CTTTATTACT
 6001 ACCATCATCA ACATGAAACC CACAACAGT ACTATGACCA AAATCCATT ATTGTTGAA
 6061 GCCGTTCTAA TTACCGCCGT ACTTCTCTT CTGCTCCCTC CGTCTTAGC CGCAGGGATT
 6121 ACAATGCTAC TAAACAGCC CAAACCTAACCA ACAACCTTTT TTGACCCCCGC CGGAGGTGGT
 6181 GACCCCATCC TCTATCAAC CCTATTCTGA TTCTTGGCC ACCCAGAGGT ATATATTCTC
 6241 ATTCTTCCAG GCTTCGGAAT AATTCCAC ATTGTTGCAT ACTATGCAGG TAAAAAAAGAA
 6301 CTTTTGGTT ACATAGGAAT AGTCTGAGCT ATAATAGCCA TTGGACTCCTT GGGATTCTT
 6361 GTATGGGCC ATCACATGTT TACAGTCGGG ATAGACGTCG ACACACAGGC CTACTTTACA
 6421 TCTGCCACAA TAATCATTGC AATCCGACCC GGCCTAAAG TCTTCTGACTCGAACC
 6481 CTCCACGGG GAAGCATCAA ATGAGAACCG CCACCTCTT GGGCCCTCGG CTTTATTCTC
 6541 CTCTTACAG TAGGGCTT TACTGGCATT GTCTTGGCTA ACTCCTCTCT TGACATTGTT
 6601 CTGCTGACA CATACTATGT AGTAGCCAC TTCCACTATG TACTATCTAT GGGTGTGTA
 6661 TTGCGCAATCG TTGCGCCCT CGTCCACTGA TTCCACTAT TTACAGGTTA TACCCCTCAC
 6721 TCCACATGAA CAAAAATCCA CTTCGGCTG ATATTATTG GGGTCAATCT AACATTCTC
 6781 CCCAACACT TCCTGGCCT AGCCGAATG CCCGACGAT ACTCAGACTA CCCAGACGCC
 6841 TATACCCCTT GAAACACTGT CTCATCAATC GGATCCCAA TGTCGCTCGT CGCTGTAATC
 6901 TTATTTTGT TCATTATTG AGAACGATTT ACAGCCAAAC GAGAAGTTGG GGCAGTAGAA
 6961 CTTACTGCAA CTAACATCGA ATGACTTTAC GGCTGCCCTC CACCCCTACCA CACATTGAA
 7021 GAGCCGCAT TCGTCAAGT TCGAATAAT TCGAACAAAC TAACGAGAAA GGGAGGAGTT
 7081 GAACCCCAT CAATTGGTT CAAGCCAACC ACATAACCGC TCTGTCACCT TCTTCACAAAC
 7141 ACACCAATAA GATACTAGTA AAACGCTATA ACACGCTT GTCAAGGCAG AATTGTGGGT
 7201 TAAATCCCCG CGTATCTAA ACGAAATGGC ACATCCCTCA CAACTCGGAT TTCAAGACGC

7261 AGCTTCACCC TTAATAGAAG AACTACTCA CTTCCACGAT CACGCCTAA TAATTGTAAT
 7321 TCTCATCAGC ACGATAGTAC TTTATATTAT TGCGGCTATA GTTACAGCCA AACTAACAGA
 7381 TAAACTGTT CTAGATTCTC AAGAAATTGA AATTATCTGA ACAGTTCTCC CAGCTGTAT
 7441 TCTTATTCTC ATCGCCCTAC CATCTCTCG AATTGGTAC TTAATAGACC AAATTAATGA
 7501 TCCCCACTTA ACAATCAAAG CCATGGCCA CCAGTGGTAC TGAAGCTAGC AATACACAGA
 7561 CTATGAAGAC CTCGGTTTG ACTCTTACAT AATCCCTACA CAAGACCTGA CCCCCGGACA
 7621 GTTCCGACTA CTTGAAGCAG ATCACCGAT AGTAATCCCG GTAGAATCTC CTATTGCGT
 7681 TCTAATCTC GCTGAAGATG TATTACATC CTGAGCAGTC CCCTCCCTGG CGTAAAAGT
 7741 TGACGCCGTA CCTGGACGAT TAAATCAAGC AACCTTATT GTAGCCGGC CAGGCGTATT
 7801 TtATGGTCAA TGCTGTGAAA TTTGGGAGC GAACCATAGT TTATACCCA TCGTTGTAGA
 7861 AGCAGTCCCC CTTGAACACT TTGAAAATCT GTCTTCACTA ATAATTGAAG ACGCCTCGCT
 7921 AAGAAGCTAA TAAGTACAAG CGTTAGCCTT TTAAGCTAA GACTGGTGCCTA ACACAACAC
 7981 CCTTAGCGAC ATGCCTAAC TGAACCCCGC ACCCTGATT GCAATTCTTGT TTTCTCTTG
 8041 AATAATCTT TTAACTGTTA TTCCCCCAA AGTTTAGCA CATACTTACCC CCAATGAACC
 8101 TACCTCCAA AGCACACAA AACCTAAAC AGAACCTGA AACTGACCAT GATACTAAC
 8161 TTTTTGACG AATTATATAC CCCAGTATAC CTCGGTATTG CGCTAATTG ACTAGCAATT
 8221 ACCCTCCCT GAATTCTGTT CCCAACCCCT CAAGCCCGCT GATTAACAA CCGTGTGCTA
 8281 ACACCTCAAG GATGGTTAT TAGCCGCTTC ACCTCACAC TTCTTCTTCC CCTAAACCCC
 8341 GGGGGCCACA AATGAGCAGT TTTATTAC TCATTAATGG TTTTCTTATT TTCTATTAAAC
 8401 ATACTAGGAC TCCTTCATA TACCTTCACA CCAACAAACCC AACTCTCCCT TAATATGGC
 8461 CTCGCAGTTC CCCTCTGATT AGCAACTGTC ATTATTGGAA TGCAGAAATCA ACCTACACAT
 8521 GCCCTAGGTC ACCTCTTCC AGAACCGACC CCTACTGCTC TCATCCCTGT ACTCATCATC
 8581 ATTGAAACAA TTAGCCTATT TATTGGACCC CTCGCTCTCG GTGTTGACT AACAGCGAAT
 8641 CTCACAGCAG GCCACCTCTC TATCAACTC ATTGACACAG CTGCTTGTG CTTCTTCCG
 8701 CTCATACCTG TTGTTGCTAT CTTAACACG GTGGTCTCT tcCTtCTTAC TCTGCTAGAA
 8761 GTCGCGGTCG CCATGATCCA AGCCTATGTC TTTGTTCTAC TCTTAAGCCT ATATCTACAA
 8821 GAGAACGTTT AATGGCCAT CAAGCACACC CATACACAT AGTCGACCCCC AGTCCATGAC
 8881 CCCTCACGGG GGCTATTGCT GCCCTATTGA TAACATCTGG CTTGCTATC TGATTTCACT
 8941 TCCATCCAC AACCCATTATA ACTATTGGGA CAGTCCTCTC CACCTTAACA ATCTTCAAT
 9001 GATGACGGGA CGTCTGACG GAAGGCACAT TTCAAGGACA CCACACTCCC CCTGTTCAAA
 9061 TAGGCCTCCG ATACGGAATA ATCTTATTAA TTACCTCAGA AGTCTGTT TTCTTAGGCT
 9121 TCTTCTGAGC CTTTTACCC TCAAGCCTAG CACCTACTCC AGAAACTAGGA GGCTTCTGAC
 9181 CCCCAGCAGG CATTACCCCT TTAGACCCAT TTGAAGTCCC GCTTCTTAAC ACAGCAGTCC
 9241 TCCTTGCCTC TGGTGTAACT GTTACCTGAG CACATCACAG CATTATAGAG GGTAAACGGA
 9301 AACAAACTAT TCAATCTT GCCCTACAA TCTTAACCGG GGGTTACTTC ACTTTCTCC
 9361 AAGGCCCTGA GTATGTCGA GCCCCCTTCA CCATTGAGA CGGAGTTTAC GGTGCTACAT
 9421 TCTTCTGTC CACCGGCTTC CACGGCTCTC ACGTCTTAAT TGGCACCTCA TTCTTAGCCG
 9481 TTTGTCCTGTT ACGCAAATC CTTCACCAT TTACATCAAAC CACCACTTT GGGTTGAAG
 9541 CAGCCCATG ATACTGACAC TTGAGACG TCGTCTGACT TTCTCTCTAT ATCTCTATT
 9601 ACTGATGAGG ATCTTAAATCT TCCTAGTATC AAATCTAGTA TAAGTACTT CCAATCACCC
 9661 GGTCTGGTT AAAGTCAAG GGAAGATAAT GAGCCTTCTT CTAACCATCA TTCAATCGC
 9721 CACCCCTCTC TCAACAGTAC TTGCCATTGT ATCCTTTGA CTCCCCAAA TTACACCAGA
 9781 CCATGAAAG CTATCACCAT ACGAATGCGG CTTGACCCG ATAGGCTCTG CCCGACTACC
 9841 TTTTCGCTG CGATTTTTC TCATGCCAT CCTCTTCTC CTGTTGACT TAGAGATTC
 9901 ACTCCCTCTC CCCCTCCCT GAGGAGACCA ATTAGCATCG CCACTACTCA CATTCACTTG
 9961 AGCTACAGCT GTCTAGGGC CTTCAACCT TGGCTCAT TACGAATGAA TGCAAGGAGG
 10021 CCTAGAGTGA GCTGAATAGG TGGTTAGTCT AAGAAAAACA TTGTTGTTCC GCTAAAAAC
 10081 TTGTGGTTA AATCCGCAAC CGCTTAATGA CCCCCACACA CTTGCTTC TCCTCAGCCT
 10141 TTCTCTGGG TTTAACAGGC CTGGCATTTC ACCGGTTCCA CCTCCTTCC GCTTATTAT
 10201 GTCTGAAGG AATAATACTA TCCCTATTAA TTGCCCTCTC CCTATGAACA CTCCAACCTG
 10261 ACTCAACAAA CTTTGACCA TCTCCTATGC TTCTACTTGC ATTTTCAGCC TGCAGGCAA
 10321 CGCCGCCCT CGCCCTGCTA GTAGCCACTG CTGGAACCC CGGAACCGAC CGACTACAAA
 10381 GCCTAAATCT TCTCAATGC AAAAATCTT AATCCAACCA CCTATGCTAA TCCCAACAGC
 10441 CTGGTTACTC AAACCCAGT GACTCTGGCC CATAACCTTG ATGTATAGCT TTTGCATCTC
 10501 CTTAATTAGC CTTTCATGAT TGAAGAACCT CTCAGAACCC GGCTGATCAT CACTTAACCT
 10561 CTTTATAGCC ACCGACTCCC TATCAACCC CCTCCTCGTT CTCACATGCT GACTGCTACC
 10621 ACTAATGATT TTGGCAAGCC AAAAACACAC AGCCTCAGAG CCTCTTGGTC GCCAGCGCAT
 10681 GTACATTACA CTTCTCGCT CACTCAGTT CTTCTGATT CTAGCATTAA GCGCCACTGA
 10741 GTGGTAAATG TTTTACGAA TATTGAGC CACCCCTCATC CCACTCTGA TTATTATTAC
 10801 CCGCTGAGGG AACCAAACAG AACGCCCTAA TGCAAGGACC TACTTTCTT TTACACATT
 10861 GGCAGGCTCA CTACCCCTC TCGTTGCTT ACTCTTCTC CAGAACACAT CGGGCACCC
 10921 GTCGTTGCTA ACCCTACACT ATACAGACCC ACTCACTCTC TCATCTTATG CAGACAAATT
 10981 ATGATGAGCA GGCTGTCTC TAGCATTCT CGTAAAATA CCTCTCTAT GAGTCCACCT
 11041 GTGACTCCCT AAAGCCACG TTGAAGCCCC CATTCAGGC TCAATGATCC TTGCAAGGGT
 11101 TCTCTTAAAAA CTGGGAGGAT ATGGCATGAT CCGCATAATA ACAATATTAG AGCCTCTAAC
 11161 TAAGGAGCTA AGCTACCCCT TCATTGCTT TGCACTTGG GGGTAATCA TAACTGGCTC
 11221 AATTGGTCA CGACAGACG ACCTTAAGTC CCTAATTGCTG TACTCATCTG TAAGTCACAT
 11281 GGGCTAGTG GCCGGGGAA TTCTTATCA GTCACCTCTGG GGCTTAACAG GGGCCCTTAC
 11341 CCTAATAATT GCACACGGTC TTACCTCATC AGCCCTCTTC TGCCTAGCAA ACACAAACTA
 11401 TGAACGAACC CACAGTCGA CGATAGTACT AGCACGAGGG CTTCAAGTAG CTCTACCTT
 11461 GATGGGCACT TGATGATTAA TTTCTAGCTT AGCCAACCTA GCCCTTCCAC CACTACCTAA
 11521 CCTCATAGGT GAACTAATAA TTATTACCTC TCTATTCAAC TGATCCTGGT GAACTCTAGC
 11581 ATTAACCGGG GCAGGCACTC TAATTACAGC TGTTACTCC CTTTATATGT TCTTAATAAC
 11641 TCAACGGGAA CCTCTCCAG CACATCATC TGCGCTTGAG CCATCTCAC CCGAGAACAA
 11701 CCTCTTATT GCACTCTCAT TTATCCCTT GATCTCTCTC GTGCTTAAGC CCGAGCTGAT
 11761 CTGAGGTTGG ACTGCTGTA GGTGTTAGTTT AAACAAAAAC ATTAGATTGT GATTCTAAAA
 11821 ATAAGGGTTA AAATCCCTC TCCCACCGAG AGAGGCTCGC AGCAATGGGA ACTGCTAATT
 11881 CCCATGACCT TGGTGGACC CCCaGGCTCA CTCGAGAACG TCCTAAAGGA TAACAGCTCA
 11941 TCCGTTGGTC TTAGGAACCA AAAACTCTG GTGCAAATCC AAGTAGCAGC TATGCACCC

12001 ACCTCCCTAA TGATTCGTC AAGCCTGATC ACAATTTTG CATTGCTAGC CTATCCCCTA
 12061 GTCACCACAA TACGACCCAC GCCCCAAGAA CCTCAGTGGG CTTCTACCCA TGTTAAAACA
 12121 GCTGTAAAAA TAGCTTCTT TGTTAGTCTG TTACCCCTGA CCCTATTTCT TAATGAGGGT
 12181 GCCGAGACAA TCGTACTAA CTGAACCTGA ATAAATACTA ACTCCTTCGA CATCAACATT
 12241 AGCCTAAAAT TTGACTTCA CTCAATTATT TTTACCCCA TTGCCCTCTA CGTAACCTGG
 12301 TCCATCCTAG AATTGCGATT ATGATACATG CATGCAGACC CAAACATAAA CCGCTTCTC
 12361 AAGTATCTAC TGACATTCT GATTGCCATA ATCGTTCTCG TAACCGCAA CAATATGTTT
 12421 CAACTATTCA TCGGTTGAGA GGGTGTGGG ATCATATCGT TCCTCCTTAT CGGGTGGTGG
 12481 TTCGGACGGG CGGATGCTAA CACTGCAGCA CTTCAAGCAG TACTTTATAA CCGAGTTGGG
 12541 GACATCGTC TTATCTTCG CATGGCTGA ATAGCAGCAA ACCTAAACTC CTGGGAAATG
 12601 CAACAAATCT TTGTAAACCTC CAAAGACATA GACTTAACCTT ATCCTCTCTC CGGGCTCATC
 12661 ATCGCAGCAA CCGGCAAGTC GGCTCAATT GGCTTCACC CCTGACTACC CTCCGCTATG
 12721 GAGGGTCTTA CACCGGTCTC TGCCCTACTT CATCTAGCA CTATAGTCGT TGCCGGCATT
 12781 TTCTGCTGG TCCGATGAG CCCCCCTCTG GAAAACAACC AAAC TGCCCTCAC ACACAAATGAT
 12841 CTGTGCCCTG GGGCCCTAC AACACTGTT ACTGCAACCT GTGCCCTCAC ACACAAATGAT
 12901 ATTAAAAAAA TCGTGCATT CTCCACATCA AGCCAGCTTG GACTTATGAT GGTAACCATC
 12961 GGACTAAACCA AACCACAACT AGCGTTCTT CACATTGCA CACACGCTT CTTCAAGGCC
 13021 ATGCTATTTT TATGCTCGG ATCTATTATT CATACTCTTA ATGACGAACA AGACATCCGA
 13081 AAAATAGGAG GCATACACCA CTTGCGCCT TTTACATCTT CTTGCCTAAC TATTGGCACT
 13141 CTTGCCCTTA CAGGGACCCC CTTCTTAGCT GGCTTCTCT CAAAGATGC TATCATCGAA
 13201 GCACTAAACCA CATCCCACCT AAACGCCATA GCCCTAACTT TAACCCCTCCT AGCCACCTCT
 13261 TTACAGCAA TTTACAGCTT TCGAGTAGTA TTTTTGTGC CATGGGCCA CCCCCGGTTT
 13321 ATCCCCCTCT CCCCCATCAA CGAAAACAAAC CCAGCAGTGA TTAACCCACT TAAGCGATTA
 13381 GCATGAGGAA GCATCGTTG AGGCTACTA ATTACTCTA ACATTACACC CCTAAAAAAC
 13441 CCTGTTATAT CTATACCCCC TCTCTTTAA CTAGCTGCTC TTGAGTAAAC AATCCTGGC
 13501 CTACTTGTG CTATGAACT CGCCATGCTA ACTAATAAAGC ATATATAAATC AATACCGAC
 13561 ACACACGTAC ACCACTTCTC CACCTGTTA GGCTTCTTCC CCGGAGTAAT GCACCGCTA
 13621 GCCCTAAAT TACTCTCTT CCTAGGCCA AACGTTGCCA ACCAAACAGT AGATCAAAC
 13681 TGGCTAGAAA AAATCGGCC CAAAGGCTT GAGTCCCTTA ACCGACCCCT GATCTCTCT
 13741 ACAAGTAACA TTCAGGGAG AGTAATCAAG ACATATCTCA CCCTTTCTCCT CCTGACACTA
 13801 GCCCTTATAA TCCCACTATT CATCGCTTAA ACTGCCCCTG ACGCCCCCTCG ACTTATACCC
 13861 CGAGTCAACT CGAGAACCCAA AAACAAAGTT AAAAGCAATA CCCACGCACT AATAATTAAT
 13921 AACCACCCCC CAGAAGATA TATTAAGCT ACACCTCTAA TATCCCCCG AACACGGAG
 13981 AACTCCGCCA ACTCATCTGC TAGGACCCAA GAGCAGCTGT ATCACCCCCC TCAGAACAGC
 14041 CCAGAAACTA ACATTACCCC TCCTGTATAA ACTACCCCCAT AAACCATTAC CGACCAACTA
 14101 CCCAGCTCT CCGGATACGG CTCAGCAGCT AACGCTGCGG AGTATGCAA TACAACCTAGC
 14161 ATCCCAACCA AGTAAATTAA GAACAAACAA AGTGCACAAA AAGGCCCCCCC ATGCCCAACT
 14221 ATATACCCAC ACCCCATCCC AGCTACTACA ACCAAACCCA AAGCAGCGAA ATAAGGAGAT
 14281 GGATTAGAAG CAACTGCAC TAACCTAAC AGCAGAGAAA ATAAACTAA ATATATAGCA
 14341 AAAGTCATAA TTCTGCCAG GACTTTAACCG AGGACTAATG GCTTGAAAAA CCACCGTTGT
 14401 TATTCAACTA CAAGAACCT AATGGCCAGT CTACGTAAT CCCACCCCTCT TCTAAAATC
 14461 GCAAACGATG CTTTAGTTGA CCTCCAGGCC CCCTCTAATA TCTCCGTTTG GTGAAACTTT
 14521 GGGCTCTTC TAGGACTCTG TTTAGTAACC CAAATTGCTA CCGGCTTGT TCTAGCCATA
 14581 CACTATACAT CAGATATTGC TACTGCCCTC ACCTCCGTGG CTACATCTG CGGGGACGTC
 14641 AACTACGGCT GACTTATCCG AAGCATTCTAT GCCAACGGCG CATCATTCTT TTTCATTGTC
 14701 ATCTACCTAC ACATCGGCCG AGGACTCTAC TATGGCTCAT ACCTCTATAA AGAAACATGA
 14761 ACTATTGGAG TCGTCTCTT CCTCTCTGT ATAATAACGG CCTTCGTCGG ATACGTCTC
 14821 CCTTGAGGAC AAATGTCGT TTGAGGTCA ACCGTTATTA CCAACCTCTC ATCCGCCGTC
 14881 CCCTATGTT GCAACACCT AGTCAAATGG ATTGAGGTG GTTTTCTGT AGACAATGCT
 14941 ACCCTTACCC GGTTCTTGC CTTCCACTTC CTTTCCCTCT TCATCATCGC CGCCGCAACA
 15001 GTAATTCTAC TACTATTCTC TCACGAAACT GGCTCAATAA ATCCACCGGG CTTAAACTCA
 15061 GACTCAGACA AGGTTCCCT CCACCTTAC TTCACTTACA AAGACCTCTT AGGTTTGCA
 15121 GTGCTCTCA CTGCAATTAGC ATCCCTCGCT TTGTTTCACTAATCTTTT AGGAGACCCA
 15181 GATAACTTTA CCCCCCGAAA TCCACTTGT ACACCCCCAC ACATTAACCC AGAGTGTAC
 15241 TTCTCTTTG CTTACGCTAT TCTCCGGTCC ATCCCAAACA AACTTGGAGG CGTTCTGCC
 15301 CTTTATTTT CGATCTTAGT TCTCATGCTT GTCCCCCTCC TTCACACCTC TAAACAAACGC
 15361 AGCCTAATGT TCCGCCCTC AACGCAATT TTGTTCTGGT CTCTGGTAGC AGACGTTATA
 15421 ATTCTAACTT GAATCGGAG AATGCCGTA GAACACCCCT CGTTATCAT TGGACAAGTA
 15481 GCATCACTCC TTTACTCTG CCTCTTCTA GTCCTGATCC CAACAGCAGG TTACTAGAG
 15541 AATAAAGTAA TTGGATGAAA ATGCACTAGT AGCTCAGCGC ACAGAGCCCC AGTCTTGAA
 15601 ACTGGCCGTC GGAGGTTAA ATCCTCTTA CTGCTCAAAG AAAGGAGATT TCAACTCTA
 15661 CCCCTAACTC CCAAAGCTAG GATTCTAGCA TTAAACTATT CTTGCGGCCA TAATGTATGT
 15721 AAAACAAAGG ATTTTATAT ACATGTATGT ATTAACACCA TATATTATA GTAACCATT
 15781 TGTGTATGT ACTAGGACAT ACATGTATAA TCACCTAACAA TAGTAATATA GCACTCATTC
 15841 ACCACCATTT TTAACTAAGA AATTACTAGA ACCCTGACTT CACATTCAAG TTACATATGT
 15901 GAAATTCCAG GACCAGCGA CATTAAAGAC CGAACACAAAC ACTCACCCTG CAAGTTATAC
 15961 CAAGACTCAA AATTCTACCT ACTAAACCA CCAACCCAAAT AGAGCCTAC CAACCGGTGA
 16021 TCCCTTAATG ATAACCTCTA TTGAGGGTGA GGGACAAAAA TTGTTGGGGT TTCACTCGGT
 16081 GAATTATTC CTTGAGTTGG TTCTACTTC AGGGCCATTG ATTGATATTA TCCCTCACAC
 16141 TTTCATCGAC ACTTACATAA GTTAATGTT ATAATACATA CGACTCGTTA CCCACCAAGC
 16201 CGGGCGTTCA CTCCAGCGGG TAAGGGGTT CTTTTTTTT TTTCCTTCA CTTGACTTTT
 16261 CAAAGgGcaT CCAGCCAAACAA AAGGGGGAGC actTTTCTG caCGCGGcGt
 16321 ACATAGtATC CaTGTATTA gTCTTCTTA AAAGGATAAC ATAAAGGGAT ATcaTGtGcA
 16381 TAAAATATGtG CTCCCTTCTT CTATCatCCC CaGGATACCC CTTTTTTG cGTAAAACCC
 16441 CCCtACCCCCC CTAAACTCTC GAAGTGTCA AaACCCCTGA AAACCCCCCG gAACAcGGAC
 16501 AAACCTCTGG TAgCTTAgAA AAGGCAATT CTTTCAACC TACTTTCAA CCCTAAAATA
 16561 CCCCTGacaA AAGAtAAGcC CaCCAGTTAT TTTAATAATG GTACTTTCA ACCCTAAAAT
 16621 ACCCCTGACC AAAGATAAGC CCACCAGTTA TTTAATAAT GGTACTTTTC AACCCCTAAAA
 16681 TACCCCTGAC CAAAGATAAG CCCACCAGTT ATTTAATAA TGGTACTTTT CAACCCCTAAA

16741 ATACCCCTGA CCAAAGATAA GCCCACCA GTTATTTAATA ATGGTACTTT TCAACCCCTAA
 16801 AATACCCCTG ACCAAAGATA AGCCCCACAG TTATTTTAAT ATGGTACTT TTCAACCCCTA
 16861 AAATACCCCT GACCAAGAT AAGCCACCA GTTATTTAA TAATGGTACT TTTCAACCCCT
 16921 AAAATACCCC TGACCAAGA TAAGCCCAC AGTTATTTA ATAATGGTAC TTTCAACCC
 16981 TAAAATACCC CTGACCAAG ATAAGCCCAC CAGTTATTT ATAATGGTA CTTTCACCC
 17041 CTAAAATACC CCTGACCAA GATAAGCCC ACAGTTATTT TAATAATGGT ACTTTCAAC
 17101 CCTAAAATAC CCCTGACCA AGATAAGCCC ACCAGTTATT TTAATAATGG TACTTTCAA
 17161 CCCTAAAATA CCCCTGACCA AAGATAAGCC CACCGTTAT TTAATAATGG TACTTTCA
 17221 ACCCTAAAATA ACCCTGACCA AAAGATAAGC CCACCGTTA TTTAATAATGG TACTTTCA
 17281 AACCTAAAATA TACCCCTGAC CAAAGATAAG CCCACCGTT ATTAAATAA TGTTACTTT
 17341 CAACCTAAA ATACCCCTGA CCAAAGATAA GCCCACCA GTTATTTAATA ATGGTACTTT
 17401 TCAACCCCTAA AATACCCCTG ACCAAAGATA AGCCCACAG TTATTTTAAT ATGGTACTT
 17461 TTCAACCCCTA AAATACCCCT GACCAAGAT AAGCCCACCA GTTATTTAA TAATGGTACT
 17521 TTTCAACCCCT AAAATACCCC TGACCAAGA TAAGCCCAC AGTTATTTA ATAATGGTAC
 17581 TTTTCACCC TAAAATACCC CTGACCAAG ATAAGCCCAC CAGTTATTT ATAATGGTA
 17641 CTTTCACCC CTAAAATGAC TGGGACCTGT AACGCCGTGT CATCTCTTC ATGATCAAAA
 17701 AACTGCTGTT TAGATTATT CAAGTATTCT TAGCATCAT CCAACTACGA AGGGCACAAA
 17761 ACAAAACCG AGAAGACTT GCTCTCATT CAAAAGACTAT CTTTTTTAA CCGACAAGCC
 17821 CTTAAACCAA CCTTCTATTA TCCTAAACT TGACAAGACT AGGGAAACAC CTTTGGTGT
 17881 AACAGATATT TAACT

Reinhardtius hippoglossoides 4 18078 bp

1 GCTAACGTAG CTTAATTAAA GCATAACACT GAAGATGTTA AGATGGGCC TAGAAAGCCC
 61 CGGGAGCACA AAGGCTTGG CCGTACCTTA CTGTCGACTT TAACCTAAACT TACACATGCA
 121 AGTATCGCC CCCCTGTGAG AATGCCACA ACTCCCTGCT TGGGAACCTAG GAGCTGGTAT
 181 CAGGCACAAG CCCCGTAGC CCACGACGCC TTGCTTAGCC ACACCCCTAA GGGAACTCAG
 241 CAGTGATAAA TATTAAGCCA TAAGTGAAAA CTTGACTTAG TTAAGGTTAA GAGGGCCGGT
 301 AAAACTCGTG CCAGCCACCG CGGTTATACG AGAGGCCAA GTTGACAAAC AACGGCGTAA
 361 AGAGTGGTTA GGGGATTTC TAAACTAGAG CCGAACGCTT TCAGAGCTGT TATACGCACC
 421 CGAAAGTATG AAACCCAAATT ACGAAAGTAG CTCTACTCAT CCTGAAACCCA CGAAAGCTAA
 481 GAAACAAACT GGGATTAGAC ACCCCACTAT GCTTAGCCCT AAACATCGAT TGCACATATAC
 541 ACTCCATATC CGCCCCGGAA TTATGAAGCT CAGTTAAAC CCAAAGGAC TTGGCGGTGC
 601 TTAACATCCA CCTAGAGGAG CCTGTTCTAG AACCGATAAC CCCCGTTAAA CCTCACCCCTC
 661 TCTTGTTTA TCCGCCTATA TACCACCGTC GTCAGCTTAC CCTGTGAAGG CTTTACAGTA
 721 AGCAAAATTG GCACAGCCC AAACGTCAAG TCGAGGTGTA GTGAATGAGG GGGGAAGAAA
 781 TGGGCTACAT TTGCTAAACA TAGAAACAC GAATGTTGCA TTGAAACATG CAACTGAAGG
 841 AGGATTAGC AGTAAACCGG AAATAGACGG TCCCGCTGA ACTGGCCCTA AAGCGCGCAC
 901 ACACGGCCCG TCACCCCTCC CAAGCCCCCT GAACCTAAACT ATTAAAACCC CAACAACCCG
 961 CGAAGGGAG GAAAGTCGTA ACATGGTAAG TGTACCGGAA GGTGCGCTTG GAAAATCAG
 1021 AGTGTAGCTA AGTTAGATTA CAGCATCTA CTTACACCGA GAAGACGCC GTGCAAGTCG
 1081 GACCACCTG ACGCCTATTA GCTAGCCAA CCCCTTAACA CAACAAACCC CCATTTATAA
 1141 CCCCTAAAGC ACGAACACC CACGTAGCTA AACCAATTCTC CCCCCTAAGT CCAGGCGATA
 1201 AAAAAGGAAA TTTGGAGCA TAGAAAAAGT ACCGCAAGGG AAAGCTGAA GAGAGATGAA
 1261 AGAGCCCTAGT AAAGCTTAAG AAAGCAGAGC TTAAAGCTCG TACCTTTTGTC ATCATGATTT
 1321 AGCTGACT TTCAAGCAA GAGAACCTAA AGTTGTAAAC CCGGAAACTCG AGTGAGCTAC
 1381 TCCAAGACAG CCTATTATA GGGCAACCC GTCTCTGTGG CAAAGAGGT GGAAGAGCTT
 1441 TGAGTAGAGG TGACAAACCT ACCGAACATCA GTTATAGCTG GTTGCCTGTG AATTGAATAG
 1501 AAGTTCAGCC CCCTGGGTT CTCACCATCG CACTTTATTT ACCCTTCAGA TGCAATGAGG
 1561 ACCCAGGGGT GTTAGTCAAA GGGGGTACAG CCCCTTTGAT ACAAGACACA ACTTTCTCAG
 1621 TAGGATAAAAG ATCATAATAA AATAAGGACA GATGTTTTAG TGGGCCTAAA AGCAGCCACC
 1681 TTAACAGAAA GCGTTAAAGC TCAAACATGA AGCCCTCCCG TTATACCGAT AACCTATCT
 1741 TAATCCCCCA CATTAAACAG GCCCTCTTAT GCATCACATA GGAACGACTA TGCTAATATG
 1801 AGTAATAAGA GGGTATAACC ACCCTCTCCT TGACATGTG TAAATCGGAA CGGACCCCCC
 1861 ACCGAATCTT AACCGCCCCA ATCAAAGAGG GTATTGGAAA CTACCACAAA TTCAGGCCAG
 1921 AAAAACATCC AATGATAACC CGTTAACCCCC ACACTGGTGT GCCCAAAGG AAAGACCGAA
 1981 GGGGGGAGAA GGAACCTGGC AAACATACCC CAAGCCTCGC CTGTTTACCA AAAACATCGC
 2041 CTCTTGCTA ACCACAGTAT AAGAGGTCCC GCCTGCCAG TGACCATATA GTCAACGGC
 2101 CGCGGTATT TGACCGTGC AAGGTAGCTG AATCACTTGT CTTTTAAATG AAGACCTGTA
 2161 TGAATGGCATT AACGAGGGCT TAACTGCTC CTTCCCCCGG TCAATGAAAT TGATCTCCCC
 2221 GTGCAGAACG GGGGATTAA CCATAAGACG AGAAGACCT ATGGAGCTTT AGACACACAG
 2281 GTGGACCATG TCAAATACCC CAAACTAAGG GCCTGACTA AGTGGAAACCT GCCTTGATGT
 2341 CTTCCGTTGG GGCGACCATG GGGAAATCAA AACCCCCACG TGGAAAGGGA GCACACCCCT
 2401 AAGTTACTTC TTCTCCCGCA AGCCAGAGCA ACAGCTCTAA CCAGCAGAAA TTCTGACCAA
 2461 ACTGATCCGG TAAAACCGAT CAACGAACCA AGTACCTCTA GGGATAACAG CGCAATCCCC
 2521 TTTTAGAGCC CATATCGACA AGGGGTTTA CGACCTCGAT GTTGGATCAG GACATCTAA
 2581 TGGTGCAGCC GCTATTAAAGG GTTGTGTTGT TCAACGATT AAGTCCTACG TGATCTGAGT
 2641 TCAAGCAGCGA GTAATCCAGG TCAGTTCTA TCTATGACAT GATCTTTCT AGTACGAAAG
 2701 GACCGAAAAAG AAGGGCCCA TGCTAAAGT ACGCCTCACC CCCACCTAAT GAAAAAAATCT
 2761 AAACTAGGCA AAAGGGCATC ACCCTTTGTC TGGAGATAAC AGCAAGTTGG GGTGGCAGAG
 2821 CCCGGCTAAT GCAAAGACCA TAAGCCCTT CCACAGAGGT TCAAATCCTC TCCTTAACCA
 2881 TGATTCACAC TCTCATTACCATATTATA ACCCCCTAAAC CTTTATTGTA CCTGTATTGC
 2941 TAGCGTAGC ATTCCCTCACC CTCCCTGAC GAAAAGTCCT AGGCTACATA CAACTCCGAA
 3001 AAGGGCTAA CATCGTAGGG CCTTACGCC CCCTTCACCC CATTGCTGAC GGCCTAAAC
 3061 TCTTATTAA AGAACCCGT CCGCCTCCCA CCGCATCCCC CGTCTGTTC CTCCCTGGCCC
 3121 CTATACTCGC ACTCAACTG CCGCTGACCC TTTGAGCCCG ATGCCCCCTC CCGTACCC
 3181 TTGTGGACCT TAACCTGGGT ATCCTTTTA TTCTAGCCT ATCTAGCCTT GCAGTATACT
 3241 CTATTCTGG GTCTGGTGA GCATCCAATT CAAAATACGC TTTAGTAGGA GCACTCGGG
 3301 CTGTGACACA AACAAATTCT TACGAAGTTA GCCTAGGACT CATCCTCCTT AACATCATCA

3361 TCTTTACCGG AGGATTACA CTTCAGACCT TTAACACGGC CCAAGAGGCC ATCTGACTAG
 3421 TTGTACCCCG CTGACCCCTC GCTGCCATAT GATAACATCTC CACTCTGCC GAAACAAACC
 3481 GTGCACCCCT CGACCTAACAA GAGGGGAAT CTGAACCTGT TTCAGGCTTC AACGTAGAAT
 3541 ACGCAGGAGG GCCCTTCGCT TTGTTCTTT TAGCCGAATA CTCAAATATC CTCCAATAA
 3601 ATACTCTGTC CGCAACATTA TTTTGGCG CCTCCCACAT TCCCTCTATC CCAGAATTAA
 3661 CTAGCATTA TaTTaTAATC AAAGCAGCCC TTCTCTCAAT TGTTTTCTG TGAGTCCGAG
 3721 CTTCTGTACCC GCGATTCCGT TACGACCAAC TTATGCACCT CATTGAAAG AACTTTCTC
 3781 CACTGACTTT AGCACTAGTT ATTTGACACT TAGCGCTCCC TATTGCACTTC GCTGGCTC
 3841 CACCACAACT ATAAACGCAG GAGCTGTGCC TGAATTAAA GGGCACTTT GATAGAGTGA
 3901 ATTATGGGG TTAAATCCC CCCAACCTCT TAGAAAAGAG GGGATCGAAC CCAACCTGAA
 3961 GAGATCAAAA CTCTCTGTG TTCCCTACCA CCACCTCCCTA GTAAGGTAG CTAACAAAGC
 4021 TCTTGGGCC ATACCCCAAC CATGTAGTT CAAATCCTTC CTTTACTAAT GAACCCCTAC
 4081 ATCTTGACCA CTCTTATATT TGGTTGGC CTAGGAACAA CACTTACGTT TGCAAGCTCA
 4141 CACTGACTCC TCGCTTGAAT AGGCCTCGAA ATTAATACAC TAGCCATTAT TCCACTAATA
 4201 GCCCCAACATC ATCACCCCCG AGCAGTCGAA GCTACTACTA AATACCTTCT CGCACAAGCC
 4261 ACAGCAGCCG CTACCCCTCA GTTGTCAAGT ACCACCAAGC CATGACTAAC AGGCCAATGA
 4321 GACATCCAAC AAATAACACCCCCCTTACAAATAAATC TTGTTATTGC CCTAGCACTA
 4381 AAAATTGGAC TAGCACCAT ATCACTCTGA CTTCAGAAAG TGCTTCAAGG ACTAGACCTC
 4441 ACCACCGGAC TTATCCTCTC GACTTGACAA AAACCTGAC CTTTGCCTT ATTACTACAA
 4501 ATTCAAGCAA CAAACCTAC GCCTCTAAC ATTATTGGCT TGCTATCCAC CTCGTTGGA
 4561 GGATGAGGTG GTCTCAATCA AACTCAATT CGTAAAGTCC TCGCCTACTC CTCGATTGCC
 4621 CATCTTGGCT GAATAATACT AATATTCAAG TTTTCAGCC TCCTCGCCCT CTTGCCCTA
 4681 CTCACGTATT TTACTATGAC TTTTCAGCA TTCTCTAAC TTAAAGTAAA CAAGGCTACC
 4741 ACTGTCAATG CCCTCGCAAT CTATGAAACA AAAACCCCTG CCCTCACAGC CCTAGCACCC
 4801 CTAGCTTAT TGTCCTTGG TGGCTCTCTT CCCCTTACCG GCTTCATACC CAAATGATT
 4861 ATTCTTCAAG AACTGACTAA GCAAGACCTC CCAGCCCTG CTACCCCTCG CGCATTAACT
 4921 GCCCTCTTAA GCCTTACTT CTACTTGCCTC CTCTCGTAGC CGATAACCCCT GACAATGTT
 4981 CCTAACAAACC TCGTCGGGT TACCCCTGA CGGTTCTATT CTCCCCAATT TACGCTTCCC
 5041 CTCGCCATGT CCACTCGAGC AACCCCTT CTTCTCCAC TAGCCCTG TGCTGTAGCA
 5101 CTACTCATAG CTTAAGGGAC TTAGGCTAGC ACTTAGACCA ATGGCTTCA AAGCCATCAG
 5161 CGTGAGTGA AATCTCTCA GCTCTGTAA GACTTGCAGG CTATTATCCC ACATCTTCTG
 5221 CATGCAAAGC AGACACTTAA ATTAAGCTAA AACCTTTCTA GATAGGAAGG CCTTGATCCT
 5281 ACAAAATCTT AGTTAACAGC TAAGGCCCA ATCCAGCAG CATCTATCTA CTTTCTCCG
 5341 CTTGATTTAC CGAGTAAAAA GCGGGAGAAA GCCCCGGAG GCGATTAGCC TGCTTCTTA
 5401 GATTGCAAT CTAACATGTA ACACCCAGA GCTTGATAAG AAGAGGGCTT GCACCTCTGT
 5461 CTATGGTCT ACAATCCACC GCTTAACCTCA GCCATCCTAC CTGTGGCAAT CACACGTGTA
 5521 TTTTCTCGA CCAATCACA AGACATCGGG ACCCTCTATC TGTTATTG TGCTGAGGCC
 5581 GTTATAGTGG GAACAGGCT AAGTCTGTT ATTCGGGAG AACTAAGCCA ACCTGGGCT
 5641 CTCCCTGGAG ACGACAAAT TTATAACGTA ATCGTCACCC CACACGCCCT TGTAATAATC
 5701 TTTTTATAG TAATACCCAT TATgATCGGG GGTTTGGAA ACTGGCTTAT TCCACTAATA
 5761 ATTGGAGGCC CAGATATGGC TTCCCTCGA ATAAATAACA TGAGTTCTG ACTTCTTCCC
 5821 CCATCCTTC TTCTCTCTT AGCCTCTCA GGTGTTGAAG CTGGGGCAGG TACGGGGTGA
 5881 ACCGTTTATC CACCACTAGC TGGTAATCTG GCCCACGCC GAGCATCCGT TGACCTAAC
 5941 ATCTTCTCAC TTCACTTGC AGGGATTTCG TCAATTCTGG GGGCAATTAA CTTTATTACT
 6001 ACCATCATCA ACATGAAACC CACAACAGTT ACTATGTACC AAATCCCTT ATTGTTTGA
 6061 GCCGTTCTAA TTACCGCCGT ACTTCTCTT CTGCTCCCTC CGCTCTTAGC CGCAGGGATT
 6121 ACAATGCTAC TAACAGACCG CAACCTAAC ACAACCTTT TTGACCCCGC CGGAGGTGGT
 6181 GACCCCATCC TCTATCAAC CCTATTCTGA TTCTTGGCC ACCCAGAGGT ATATATTCTC
 6241 ATTCTTCCAG GCTTCGGAAT AATTCCCAC ATTGTTGCAT ACTATGCAGG TAAAAAGAA
 6301 CCTTTGGTT ACATAGGAAT AGTCTGAGCT ATAATAGCCA TTGGACTCCT GGGATTCTT
 6361 GTATGGCCAA ATCACATGTT TACAGTCGGG ATAGACGTCG ACACACGAGC CTACTTACA
 6421 TCTGCCACAA TAATCATGTC AATCCGACC GGGTAAAG TCTTCTGACTCGAAC
 6481 CTCCACGGGG GAAGCATCAA ATGAGAACAGC CCACCTCTT GGGCCCTCGG CTTTATTTC
 6541 CTCTTACAG TAGGGGCT TACTGGCATT GTCTTGGCTA ACTCCTCTCT TGACATTGTT
 6601 CTGCATGACA CATACTATGT AGTAGGCCAC TTCCACTATG TACTATCTAT GGGTGTGTA
 6661 TTCGCAATCG TTGCGCCCT CGTCCACTGA TTCCCCTAT TTACAGGTTA TACCCCTCAC
 6721 TCCACATGAA CAAAAATCCA CTTCGGCTG ATATTATTG GGGTCAATCT AACATTCTC
 6781 CCCAACACT TCCTGGCCT AGCCGAGATG CCCGAGCAT ACTCAGACTA CCCAGACGCA
 6841 TATACCTTGA GAAACACTGT CTCATCAATC GGATCCCTAA TGTCGCTCGT CGCTGTAATC
 6901 TTATTTGTG TCATATTG AGAACGATT ACAGCCAAAC GAGAAGTTGG GGCAGTAGAA
 6961 CTTACTGCAA CTAACATCGA TGAACCTTAC GGCTGCCCTC CACCCCTACCA CACATTGAA
 7021 GAGCCGCAT TCGTCAAAGT TCGAATAAAAT TCGAACAAAC TAACGAGAAA GGGAGGAGTT
 7081 GAACCCCAT CAATTGGTT CAAGCCAACC ACATAACCGC TCTGTCACTT TCTTCACAA
 7141 ACACCAATAA GATACTAGTA AAACGCTATA ACACCTGCCT GTCAAGGAG AATTGTGGT
 7201 TAAATCCCCG CGTATCTAA ACGAAATGCC ACATCCCTCA CAACTCGGAT TTCAAGACGC
 7261 AGCTTCACCC TTAATAGAAG AACTACTTCA CTTCCACGAT CACGCCCTAA TAATTGTAAT
 7321 TCTCATCAGC ACGATAGTAC TTATATTG TGCGCTTAA TTACAGGCC AACTAACAGA
 7381 TAAACTGTT CTAGATTCTC AAGAAATGAA ATTATCTGA ACAGTTCTCC CAGCTGTAT
 7441 TCTTATTCTC ATCGCCCTAC CATCTCTCGG ATTATTGTC TTAATAGACG AAATTAATGA
 7501 TCCCCACTTA ACAATCAAAG CCATGGCCA CCAGTGGTAC TGAAGCTAGG AATACACAGA
 7561 CTATGAAGAC CTCGGTTTG ACTCTTACAT AATCCCTACCA CAAGACCTGA CCCCTGGACA
 7621 GTTCCGACTA CTTGAAGGAG ATCACCGGAT AGTAATCCCG GTAGAATCTC CTATTGCGGT
 7681 TCTAATCTCA GCTGAAGATG TATTACATT CTGAGCAGTC CCCTCCCTGG GCGTAAAAGT
 7741 TGACGCCGTA CCTGGACGCTA TAAATCAAGC AACCTTATT GTTAGCCGGC CAGGCGTATT
 7801 TTATGGTCAA TGCTCTGAAA TTGCGGAGC GAACCATAGT TTATACCCA TCGTTGTAGA
 7861 AGCAGTCCCC CTTGAACACT TTGAAAATG GTCTTCACTA ATAATTGAAAG ACGCCTCGCT
 7921 AAGAAGCTAA TAAGTACAAG CGTTAGCCTT TTAAGCTAAAG GACTGGTGCCTAACAC
 7981 CCTTAGCGAC ATGCCTCAAC TGAACCCCGC ACCCTGATT GCAATTCTTGT TTTCTCTTG
 8041 AATAATCTT TTAACTGTTA TTCCCCCAA AGTTTAGCA CATACTTACCA CCAATGAACC

8101 TACCTCCAA AGCACACAAA AACCTAAAC AGAACCCCTGA AACTGACCCT GATACTAAGC
 8161 TTTTTGACC AATTATATC CCCAGTATAC CTCGGTATTC CGCTAATTGC ACTAGCAATT
 8221 ACCCTCCCCT GAATTCTGTT CCCAACCCCT CAAGCCCGCT GATTAAACAA CGGTGTGCTA
 8281 ACACCTCAAG GATGGTTAT TAGCCGCTTC ACCTCACAC TTCTTCTTCC CCTAAACCCC
 8341 GGGGCCACA AATGAGCAGT TTTATTAC TCATTAATGG TTTTCTTATT TTCTATTAAAC
 8401 ATACTAGGAC TCCTTCATA TACCTCACA CCAACAAACCC AACTCTCCCT TAATATGGC
 8461 CTCGAGTTC CCCTCTGATT AGCAACTGTC ATTATTGGAA TGCAGAAATCA ACCTACACAT
 8521 GGCCTAGGTC ACCTCTTCC AGAAGGCCAC CCTACTGCTC TCATCCCTGT ACTCATCATC
 8581 ATTGAAACAA TTAGGCTATT TATTGACCCCT CTCGCTCTCG TGTTTGACT AACAGCGAAT
 8641 CTCACAGCAG GCCACCTCT TATTCAACTC ATTGACACAG CTGCTTGTG CCTTCTTCCG
 8701 CTCATACCTG TTGTTGCTAT CTTAACACAG GTGGTCTCT TCTTCTTAC TCTGCTAGAA
 8761 GTCGCGGTCG CCATGATCCA AGCCTATGTC TTTGTTCTAC TCTTAAGCCT ATATCTACAA
 8821 GAGAACGTTT AATGGCCAT CAAGCACACC CATACACAT AGTCGACCCCC AGTCCATGAC
 8881 CCCTCACGGG GGCTATTGCT GCCCTATTGA TAACATCTGG CCTTGCTATC TGATTTCACT
 8941 TCCATTCCAC AACCCCTATA ACTATTGGGA CAGTCCTCT CACCTTAACA ATCTTCCAAT
 9001 GATGACGGGA CGTCGTACCA GAAGGCCACAT TTCAAGGACA CCACACTCCC CCTGTTCAA
 9061 TAGGCCTCCG ATACGGAATA ATCTTATTAA TTACCTCAGA AGTTCTGTT TTCTTAGGCT
 9121 TCTTCTGAGC CTTTACCCAC TCAAGGCTAG CACCTACTCC AGAAACTAGGA GGCTTCTGAC
 9181 CCCCAGCAGG CATTACCCCT TTAGACCCAT TTGAAGTCCC GCTTCTTAAC ACAGCAGTCC
 9241 TCCTTGCCTC TGGTGTAACT GTTACCTGAG CACATCACAG CATTATAGAG GGTAAACGGA
 9301 AACAAACTAT TCAATCTT GCCCTACAA TCTTAACCGG GGGTTACTTC ACTTTCTCC
 9361 AAGGCCCTGA GTATGTCGA GCCCCCTTC CAATTGCGAGA CGGAGTTTAC GGTGCTACAT
 9421 TCTTCTGAGC CACGGCTTC CACGGCTTC ACGCTTAAT TGTCACCTCA TTCTTAGCCG
 9481 TTTGTCGTT ACGCCAAATC CTTACCAATT TTACATCAA CCACCACTTT GGGTTGAAG
 9541 CAGCCGATG ATACTGACAC TTTCAGACG TCGTCTGACT TTCTCTCTAT ATCTCTATT
 9601 ACTGATGAGG ATCTTAATCT TCCTAGTATC AAATCTAGTA TAAGTGAATT CCAatCACCC
 9661 GGTCTGGTT AAAGCCAAG GGAAGATAAT GAGCCTTCTT CTAACCCTCA TTTCAATCGC
 9721 CACCCCTCTC TCAACAGTAC TTGCCATTGT ATCCTTTGTA CTCCCCAAA TTACACCAGA
 9781 CCATGAAAG CTATCACCAT ACGAATGGG CTTGACCCG ATAGGCTCTG CCCGACTTAC
 9841 TTTTCGCTG CGATTTTTC TCATGCCAT CCTCTTCTC CTGTTGACT TAGAGATTGC
 9901 ACTCCTCTC CCCCTCCCT GAGGAGACCA ATTAGCATGCC CCACTACTCA CATTCACTTG
 9961 AGCTACAGCT GTCTACGGC TTCTAACCTC TGGCCTCATT TACGAATGAA TGCAAGGAGG
 10021 CCTAGAGTGA GCTGAATAGG TGGTTAGTCT AAGAAAAACA TTGATTTCG GCTAAAAAAC
 10081 TTGTGGTTA AATCCGCAAC CGCTTAATGA CCCCCACACA CTTGCTCTC TCCTCAGCCT
 10141 TTCTCTGGG TTTAACAGGC CTGGCATTTC ACCGGTTCCA CCTCCTTTCC GCTTTATTAT
 10201 GTCTTGAAGG AATAATACTA TCCCTATTAA TTGCCCTCTC CCTATGAACA CTCCAACTTG
 10261 ACTCAACAAA CTTTCAGCA TCTCCTATGC TTCTACTTGC ATTTTCAGCC TGCGAAGCAA
 10321 GCGCCGGCTC CGCCCTGCTA GTAGCCACTG CTCGAACCCA CGGAACCGAC CGACTACAAA
 10381 GCCTAAATCT TCTCAATGC TAAAAAATCTT AATCCAACA CCTATGCTAA TCCCCTCAGC
 10441 CTGGTTACTC AAACCCAGCT GACTCTGGCC CATAACCTTC ATGTATAGCT TTTGCATCTC
 10501 CTTAATTAGC CTTTCATGAT TGAAGAACCT CTCAGAACCC GGCTGATCAT CACTTAACCT
 10561 CTTTATAGCC ACCGACTCCC TATCAACCCCC CCTCCTCGTT CTCACATGCT GGCTGCTACC
 10621 ACTAATGATT TTGGCAAGCC AAAAACACAC AGCCTCAGAG CCTCTTGGTC GCCAGCGCAT
 10681 GTACATTACA CTTCTCGCT CACTCAGTT CTTCTGATT CTAGCATTTA GCGCCACTGA
 10741 GCTGGTAAATG TTTTACGAA TATTGAGG CACCCCTCATC CCAACTCTGA TTATTATTAC
 10801 CCGCTGAGGG AACCAACACAG AACGCTTAA TGCAAGGGACC TACTTTCTTT TTACACATT
 10861 GGCAGGCTCA CTACCCCTC TCGTTGCTTT ACTCTTCTC CAGAACACAT CGGGCACCC
 10921 GTCGTTGCTA ACCCTACACT ATACAGACCC ACTCACTCTC TCATCTTATG CAGACAAATT
 10981 ATGATGAGCA GGCTGTCTCT TAGCATTCTC CGTAAAATA CCTCTCTATG GAGTCCACCT
 11041 GTGACTCCCT AAAGCCACG TTGAAGCCCC CATTGCAAGGC TCAATGATCC TTGCAGCGGT
 11101 TCTCTTAAAAA CTGGGAGGGT ATGGCATGAT CCGCATAATA ACAATATTAG AGCCTCTAAC
 11161 TAAGGAGCTA AGCTACCCCT CTTATGCTT TGCACTTTGG GGGTAATCA TAACTGGCTC
 11221 AATTGTCATA CGACAGACAG ACCTTAAGTC CCTAATTGCT TGACTCATCTG TAAGTCACAT
 11281 GGGCTAGTG GCGGGGGAA TTCTTATTCA GTTACCTCTG GGCTTAACAG GGGCCCTTAC
 11341 CCTAATAATT GCACACGGTC TTACCTCATC AGCCCTCTTC TGCTTAGCAA ACACAAACTA
 11401 TGAACGAACC CACAGTCGA CGATAGTACT AGCACGAGGG CTTCAAGTAG CTCTACCTATT
 11461 GATGGCCACT TGATGATTAA TTCTCTGTT AGCCAACCTA GCCCTTCCAC CACTACCTAA
 11521 CCTCATAGGT GAACTAATAA TTATTACCTC TCTATTAAC TGATCTGTT GAACTCTAGC
 11581 ATTAACCGGG GCGACACTC TAATTACAGC TGGTTACTCC CTTTATATGT TCTTAATAAC
 11641 TCAACGGGG CCTCTCCAG CACATCAT TGCCTTGTAG CCATCTCACA CCGGAGAACAA
 11701 CCTCTTATTGC CACTTTCATC TTATCCCCCTT GATTCTCTC GTGCTTAAGC CCGAGCTGAT
 11761 CTGAGGTTGG ACTGCTGTA GGTGTTAGTT TAACAAAAAC ATTAGATTGT GATTCTAAAA
 11821 ATAAGGGTTA AAATCCCTC TCCCACCGAG AGAGGCGCG AGCAATGGGA ACTGCTAATT
 11881 CCCATGACCT TGGTTGGAC CCCAGGCTCA CTCGAGAACG CCTCTAAAGGA TAACAGCTCA
 11941 TCCGGTGGTC TTAGGAACCA AAAACTCTTG GTGCAAATCC AAGTAGCAGC TATGCACCCC
 12001 ACCTCCCTAA TGATTCGTC AAGCCTGATC ACAATCTTG CATTGCTAGC CTATCCCCTA
 12061 GTCACCCAAA TACGACCCAG GCCCCAACAGA CCTCAGTGGG CCTCTACCCA TGTTAAACAA
 12121 GCTGTTAAAAA TAGCTTCTC TGTAGTCTG TTACCCCTGA CCTCTATTCT TAATGAGGGT
 12181 GCGGAGACAA TCGTTACTAA CTGAACCTGA ATAAATACTA ACTCCTCTGA CATCAACATT
 12241 AGCCTAAAT TTGACTCTA CTCAAATTATT TTTACCCCA TTGCTCTCTA CGTAACCTGG
 12301 TCCATCCTAG AATTGCAATT ATGATACATG CATGCAGACC CAAACATAAA CGCTTCTTC
 12361 AAGTATCTAC TGACATTCTC GATTGCCATA ATCGTTCTCG TAACCGCAAA CAATATGTTT
 12421 CAACTATTCA TCGGGTGGAGA GGGTGTGGGG ATCATATCTG TCCTCTTAT CGGGTGGTGG
 12481 TTCGGACGGG CGGATGCTAA CACTGAGCA CTTCAAGCAG TACTTTATAA CCGAGTTGGG
 12541 GACATCGGTC TTATCTTCGC CATGGCTTGA ATAGCAGCAA ACATAAAACTC CTGGGAAATG
 12601 CAACAAATCT TTGTAACCTC CAAAGACATA GACTTAACCTT ATCCTCTCTC CGGGCTCATC
 12661 ATCGCAGCAA CGGGCAAGTC GGCTCAATTG GGTCTTCACC CCGTACTACC CTCCGCTATG
 12721 GAGGGTCTA CACCGGTCTC TGCCCTACTT CATCTCTGCA CTATAGTCGT TGCCGGCATT
 12781 TTCTGCTGG TCCGAATGAG CCCCTCCTG GAAAACAACC AAACGCCCT CACAACCTGC

12841 CTGTGCCTCG GGGCCCTTAC AACACTGTTT ACTGCAACCT GTGCCCTCAC ACAAAATGAT
 12901 ATTAAAAAAA TCGTGCATT CTCCACATCA AGCCAGCTTG GACTTATGAT GGTAACCAC
 12961 GGACTAAACC AACCACAAC TCGTGCATT CACATTTGCA CACACGCTT CTTCAAGGCC
 13021 ATGCTATTT TATGCTCGG ATCTATTATT CATACTCTTA ATGACGAACA aGACATCCGA
 13081 AAAATAGGAG GCATACACCA CCTTGCCTT CTTACATCCT CTTGCCTAAC TATTGGCA
 13141 CTTGCCCTTA CAGGGACCC CTTCTTAGCT GGCTTCTTCT CTAAAGATGC CATCATCGA
 13201 GCACTAAACA CATCCCACCT AAACGCCCTGA GCCCTAACTT TAACCCCTCT AGCCACCTCT
 13261 TTCACAGCAA TTTACAGCTT TCGAGTAGT TTTTTGTC CCATGGGTCA CCCCCGGTTT
 13321 ATACCCCTCT CCCCCATCAA CGAAAATAAc CCAGCAGTGA TAAACCCACT TAAGCGATTA
 13381 GCATGAGGA GCATCGTGC AGGCTACTA ATTACTTCA ACATTACACC CCTAAAACC
 13441 CCTGTTATAT CTATACCCCC TCTCTTTAA CTAGCTGCTC TTGCAGTAAC AATCCTTGGC
 13501 CTACTTGTG CTATGGAAC CGCCATGCTA ACTAATAAGC AACATAAAATC AATACCGAC
 13561 ACACACGTAC ACCACTTCTC CACCTGTTA GGCTTCTTCC CCGGAGTAAT GCACCGCCTA
 13621 GCCCTTAAT TACTCTCTT CCTAGGCCA AACGTTGCCA ACCAAACAGT AGATCAAACC
 13681 TGGCTAGAAA AAATCGCCC CAAAGGCTT GAGTCTCTA ACCGACCCCT GATCTCTTCT
 13741 ACAAGTAACA TTCAGCGAGG AGTAATCAAG ACATATCTCA CCCTTTCTC CTCGACACTA
 13801 GCCCTTATAA TCCCCACTT CATCGCTTA ACTGCCCGA ACGCCCTCG ACTTATACCC
 13861 CGAGTCACT CGAGAACCCAC AAACAAAGTT AAAAGCAATA CCCACGCACT AATAATTAAT
 13921 AACCACCCCC CAGAAGAATA TATTAAGGCT ACACCTCCGA TATCCCCCG AACACCGAG
 13981 AACTCCGCCA ACTCATCTGC TAGGACCCAA GAGCAGCTGT ATCACCCCCC TCAGAACAGC
 14041 CCAGAAACTA ACATTACCCC TCCTGTATAA ACTACCCCCAT AAACCATTAC CGACCAACTA
 14101 CCCAGCTCT CCGGATACGG CTCAGCAGCT AACGCTGCCG AGTATGCAAAT TACAACCTAGC
 14161 ATCCCAACCA AGTAAATTAA GAACAAACT AGTACAGGAA AAGGCCCCCC ATGCCCAACT
 14221 AATATACAC ACCCCATCCAC AGCTACTACA ACCAAACCCA AAGCAGCGAA ATAAGGAGAT
 14281 GGATAGAAG CAACTGCAAC TAACCTAAC ACAGAGAGAAA ATAAAATCAA ATATATAGCA
 14341 AAAGTCATAA TTCTGCCAG GACTTTAACG AGGACTATG GCTTGAAAAA CCACCGTGT
 14401 TATTCAACTA CAAGAACCT AATGGCCAGT CTACGTAAT CCCACCCCTCT TCTAAAATC
 14461 GCAAACGATG CTTTAGTTGA CCTCCAGCC CCCCTCTAATA TCTCCGTTTG GTGAAACTTT
 14521 GGGCTCTTC TAGGACTCTG TTTAGTAACC CAAATTGCTA CCGGCTTGT TCTAGCCATA
 14581 CACTATACAT CAGATATTGC TACTCCCTC ACCTCCGTTG CTACATCTC CCGGGACGTC
 14641 AACTACGGCT GACTTATCCG AAGCATTCTAC GCCAACGGCG CATCATTCTT TTTCATTG
 14701 ATCTACCTAC ACATCGGCCG AGGACTCTAC TATGGCTCAT ACCTCTATAA AGAAACATGA
 14761 ACTATTGGAG TCGTCCTCTT CCTCTCGTA ATAATAACGG CTTCTGCTGG ATACGTCTC
 14821 CCTTGAGGAC AAATGTCGT TTGAGGTGA ACCGTTATTA CAAACCTCTC ATCCGCCGTC
 14881 CCCTATGTT GCAACACCT AGTCAATGG ATTGAGGTG GTTTTCTGT AGACAATGCT
 14941 ACCCTTACCC GGTTCTTGC CTTCCACTTC CTTTCCCTCT TCATCATCGC CGCCGCAACA
 15001 GTAACTCATC TACTATTCTC TCACGAAACT GGCTCAATA ATCCACCCGG CTTAAACTCA
 15061 GACTCAGACA AGGTTCCCTT CCACCCCTAC TTCACTTACA AAGACCTCTT AGGTTtGCA
 15121 GTGCTCTCA CTGCATTAGC ATCCCTCGT TTGTTtCAC CTAATCTTT AGGAGACCA
 15181 GATAACTTTA CCCCCGCAAA TCCACTGTT ACACCCCCAC ACATTAACCC AGAGTGTATAC
 15241 TTCTCTTTG CTTACGCTAT TCTCCGGTCC ATCCCAAACA AACTTGGAGG CGTTCTG
 15301 CTTTTATTTT CGATCTTAGT TCTCATGTT GTCCCCCTCC TTCACACCTC TAAACAACGC
 15361 AGCCTAATGT TCCGCCCTC AACGCAATT TTGTTCTGG CTCTGGTAGC AGACGTTATA
 15421 ATTCTAACTT GAATTGGAGG AATGCCGTA GAACACCCCT CGTTATCAT TGGACAAGTA
 15481 GCATCACTCC TTTACTTCG CCTCTTCTG GTCTGTATCC CAACAGCAGG TTACTAGAG
 15541 AATAAGTCC TTGGATGAAA ATGCACTAGT AGCTCAGCGC AGAGGCCCC AGTCTGTAA
 15601 ACTGGCCGTC GGAGGTTAAAT CTCCTCCCTA CTGCTCAAAG AAAGGAGATT TCAACTCTA
 15661 CCCCTAACTC CCAAAGCTAG GATTCTAGCA TTAAACTATT CTtTGCGGCA TAATGTATGT
 15721 AAAACAAAGG ATTTTATAT ACATGTATGT ATTAACACCA TATATTATA GTAACCAATT
 15781 TGTGTAATGT ACTAGGACAT ACATGTATAA TCACCTAACAA TAGTAATATA GCACTCATTC
 15841 ACCACCATTT TTAACTAAA AATTACTAGA ACCCTGACTT CACATTCAAG TTACATATGT
 15901 GAAATTCAG GACCAGCGA CATTAAAGAC CGAACACAAAC ACTCACCCTG CAAGTTATAC
 15961 CAAGACTCAA AATTCACTC ATTAAACCCAT CCAACCCAACT AGAGCCTAC CAACGGTGA
 16021 TCCCTTAATG ATAACCTCTA TTGAGGGTGA GGGACAAAAA TTGTGGGGGT TTCACTCGGT
 16081 GAATTATTCC TGGCATTGG TTCTACTTC AGGGCCATTG ATTGATATTAA TCCCTCACAC
 16141 TTTCATCGAC ACTTACATAA GTTAATGTT ATAATACATA CGACTCGTTA CCCACCAAGC
 16201 CGGGCGTTCA CTCCAGCGGG TAAGGGGTT TCTTTTTTT TTTCTTTCA CTTGACATT
 16261 CAAAGTGCAT CCAGGCCAAC AAAACGGTCA AAGGGGGAGC ACTTTTCTTG CACGCCGGT
 16321 ACATAGTATC CATGTTATAA GTCTTTATAA AAAGGATAAAC ATAAAGGGAT ATcaTGTG
 16381 TAAAAATGTG CTCCCTTCTT CTATCATCCC CAGGAtaCCC CTTTTTTCG CGTAAAACCC
 16441 CCCtACCCCCC CTAAACTCt gAAAGTGTGA AaACCCCTGA AAACCCCCCG gAAACAGGAC
 16501 AAACCTCTGG tagcTtAAAA AAGGcAATTAA CTTTtcAAC TACTTTtCAA CCCTAAAATA
 16561 CCCCTGACCA AAGATAAGCC CACCAAGTTAT TTTAATAATG GTACTTTTCA ACCCTAAAT
 16621 ACCCTGACC AAAGATAAGC CCACCAAGTTA TTTAATAAT GGTACTTTCA ACCCTAAAT
 16681 TACCCCTGAC CAAAGATAAG CCCACCAAGT ATTAAATAA TGGTACTTTT CAACCCCTAA
 16741 ATACCCCTGA CCAAAGATAA GCCCCACAGT TATTTAATAA ATGGTACTTT TCAACCCCTAA
 16801 ATACCCCTG ACCAAAGATA AGGCCACAG TTATTTAAT ATGGTACTTT TTCAACCCCTAA
 16861 AAATACCCCT GACCAAGATA AAGCCACCA GTTATTTTAAT TAATGGTACT TTTCAACCC
 16921 AAAATACCCC TGACCAAGA TAAGCCCAC AGTTATTTTA ATAATGGTAC TTTCAACCC
 16981 TAAAATACCC CTGACCAAG ATAAGCCCAC CAGTTATTTT ATAATGGTAA CTTTCAACCC
 17041 CTAAAATACC CCTGACCAAA GATAAGCCC CCAGTTATTT TAATAATGGT ACTTTTCAAC
 17101 CCTAAAATAC CCCTGACCA AGATAAGCCC ACCAGTTATTT TTAATAATGG TACTTTTCA
 17161 CCCTAAAATA CCCCTGACCA AGATAAGCC CACCAAGTTAT TTAATAATGG GTACTTTTCA
 17221 ACCCTAAAAT ACCCCCTGACC AAAGATAAGC CCACCAAGTTA TTTAATAAT GGTACTTTCA
 17281 AACCTAAAAA TACCCCTGAC CAAAGATAAG CCCACCAAGT ATTAAATAA TGGTACTTTT
 17341 CAAACCTAAA ATACCCCTGA CCAAAGATAA GCCCACCAAGT TATTTAATAA ATGGTACTTT
 17401 TCAACCCCTAA AATACCCCTG ACCAAAGATA AGCCCACAG TTATTTTAAT AATGGTACTT
 17461 TTCAACCCCTA AAATACCCCT GACCAAGATA AAGCCCACCA GTTATTTAA TAATGGTACT
 17521 TTCAACCCCT AAAATACCCC TGACCAAGA TAAGCCCACCA AGTTATTTTA ATAATGGTAC

17581 TTTTCAACCC TAAAATACCC CTGACCAAAG ATAAGCCCAC CAGTTATTT AATAATGGTA
 17641 CTTTTCAACC CTAAAATACC CCTGACCAAAG ATAAGCCCAC CCAGTTATTT TAATAATGGT
 17701 ACTTTTCAC CCTAAAATAC CCCTGACCAAAG ATAAGCCCAC ACCAGTTATTT TTAATAATGG
 17761 TACTTTCAA CCCTAAAATA CCCCTGACCAAAG ATAAGCCCAC CACCAAGTTATTT TTAATAATGG
 17821 GTACTTTTA ACCCTAAAAT GACTGGGACCG TGTAACGCCG TGTCATCTTC TTCACTGATCA
 17881 AAAAACTGCT GTTTAGATTA TTCAGTAT TCCTAGCATC ATCCCAACTA CGAAGGGCAC
 17941 AAAACAAAAA CCGAGAGAC TTTGCTCTCC ATTCAAAAC TATCTTTTT TAACCGACAA
 18001 GCCCTTAAAC CAACCTCTA TTATCCTAA ACTTGACAAG ACTAGGGAAA CACCTTTG
 18061 TGTAACAGAT ATTAACT

>*Scophthalmus maximus* 7678 bp

AACCTGAAAGATGTTAAGATGGGCCCTAGAAAGCCCCATGAGCACAAAGGCTTGGCCTG
 ACTTTACTATCAACTTATCTAAACTTATACATGCAAGTTCCgCCCCCCCCGTGAGGATG
 CCCACAGTCCCTGCTGGGAACAAGGAGCTGGTATCAGGCACACCCCCCCCCGCCAC
 GACACCTGCTTAGCCACACCCCCAAGGGTACTCAGCAGTATAAATTAGCCATGAG
 TGAAAACCTGACTTAAATTATAGATAAGAGGGCGGTAAAACCTGGGCCAGGCCCGGT
 TAGACAAGAGGCCAAGTGGATAGACAACCGCGTAAAGGGTGGTTAGGGGTACAACACAA
 CTAAGGCTGAATGCTCAAGACTGTTATGCGTTCAGAGCTTGAAGAACACACTACGA
 AAGTGGCTTACTAGACCTGAACCCACGAAGCTAAGGAACAAACTGGGATTAGATACCC
 CACTATGCTTAGCCCTAACATGATTAAATTGCCCTATTAAATCCGCCGGGATTACG
 AACATAAGTTAAAACCCAAGGACTTGGCGGTCTTATATCCACCTAGAGGAGCCTGT
 TCTAGAACCGATAAACCCCGTTAACCTCACCCCTCTCTGTATATCCGCCTATATACCG
 CCGTCGTACGCTAACCTGTGAAGGAACACAGTAAGGAGATTGGTACAACCCAAACG
 TCAGGTCGAGGTGAGTGTAGTGTAGTGGGAAGAAATGGCTACATTGCTAGTTAGCA
 AATACGAACGATGATTGAAATATGCAACTGAAGGAGGATTAGTAGTAAGCTAAAATA
 GAGTGGTTAGCTGAAATGGCCCTAACGGCGCACACACCGCCGTACCCCTCCCGAGC
 TACAAAGTAAGCAGAGATTAAACCTGTACCTTTGCACTCATGATTAGCTAGCATCTC
 AGGCAAGGAGCGCTTAGCTGAGACCCGAAACTAAGTGAAGCTACTCCAAGACAGCCTA
 TAAATTAGGGCAAACCCGCTCTGTGGCAAAGAGTGGGAAGAGCTCGAGTAGAGGTGA
 TAAACCTACCGAACTTAGTTAGCTGTTGCTGGGAATGGATAGGAGTTGAGCCTCC
 CGGATTCTCATTTCACCTTGTGTTACATACCTTCGGAAGCCTAGAGAAACCAAGAGAGT
 TATTCAAAAGGGGTACAGCCCTTGTGAAACAGCACACCTTGAAGCTGGTAAAGAT
 CATATTATTAAGAGGAAACTGTGTTAGTTGGCCTAAAGCAGCCACCTCAAAAGAAAG
 CGTTAAAGCTTACGTGATTGCTGATTGCTACGATTGATCATTTAATCTACCCCTT
 TGATTTACCAAGCCACCCATGCTAACATAGAGGAGACCATGCTAACATATGATAATAAGA
 GAACCCCTGTTCTCTCTAGCACACGCATAGCTGGAACGGACCCCCCAGCAATATT
 AACGGCCCAAATAAGAGGGTACTGAAGAAGGACCTGAAGACCAGAAAACCTTCAATA
 AACATACCGTTAACCTACACTGGTGTCCCCCTGAGAAAGACTAATAGAAAGAGAAGGAA
 CTCGGCAAAATACTTACAGCCTGCCCTGTTACCAAAAACATGCCCTTGGCAAACAC
 TGAAGTATAAGAGGCTGGCCAGTGCACCCAGTTCAACGGCGCGTATTTGA
 CCGTGCAGGGTAGCTAATCACTTGTCTTAAATGGAGACCTGTATGAATGGCATAAC
 GAGGGCTTAAGTGTCTCTCAAGTCAATGAAGTGTATCTCCCGTGCAGAACGGGG
 GATATTCACATAAGCAGAGAACCCCTATGGAGCTGAGACGATAGGGCAGCCATGTTA
 AGTACCCCCCCCCTACCGGACCAAACATAATGACCCCTGCTTAATGTCTCGGGTGGG
 GCGACCCCTGGGAATACAAAACCCCATGTGGAATAGGAGCACCCCCCTTAACCTCCTC
 CAAATGAGAGCCACACCTCCTAAACAGAACTTTGACCAAAATGATCCGGCAAAGCC
 GATCAACGGACCAAGTACCTCTAGGATAAACAGCGCAATCCCTTTAGGCTTATCG
 ACAGGGGGTTACGACCTCGATGTTGAGTCAGGACATCTAATGGTGCAGCGCTATT
 AGGGTCTGTTGTTCAACGATTAAGCTTACGTGATCTGAGTTGAGCAGCCGAATCC
 AGGTCAAGTTCTATCTATGAAATGACCTTCTAGTAGCAGAAAGGACCGAAAAGGCGAAGC
 CCCTGCTACAGGCACGCCCTACCTGTATGAAACATCTAAACAGGCAAGAGAAC
 ATACCCCTCAACCATAGAGAATGGTAAGCTAAGATGGCAGAACCCGGTTATGGCAAAGA
 CCTAAGCCCTTCTATGGGGTCAAGTCCCCCTTAACTATGACTTCAACCCCTTCA
 CACATGCTCTTAACTCTAACTTATTGTCAGTATTACTAGCTGTAGCATCCCTCA
 CCCTTATTGAGGCAAAGTGTAGGCTATATACAACACTCGGAAGGGGCCAATATTGTTG
 GCCCATACGGGCTCTCAACCCATGCCGATGGAATCAAGCTTTCTTAAAGGAAACCTG
 TTGACCTCGACCTCTTCCCTGTCTTTCTTATGGCCCATTGCTAGCGCTACTC
 TTGCTTAAACCTATGAACACCCATCCCTACCCCTACCCGTCGCTGACCTCAGCTAA
 GCATCCCTTTATTGCGCATTTCAGCCTCGCGTACTCTATTTAGGGTCAGGAT
 GAGCCTCCAACCTCGAAGTACGCCTATTGGGCTCTCGAGCGCTCGCACAAACTATT
 CCTACAGGTTAGCCTAGGACTCATCTCTAAATGTTATCATTTCTCAGGAGGTTCA
 CGCTCAAACATTAACTACAGCCAGAGGGCAGTTGACTGCTCCCTCCCTCG
 TTGCGCGATATGGTATGTCCTACCTGGCTGAAACAAATCGGGCCCCCTCGACCTCA
 CGGAGGGGGAGTCTGAGCTTGTCTGGCTTAACGAGATAATGCAAGGGGGACCTTTG
 CCCTTATTGAGGCAAACACTCCAACATTATTGATAAAACACATTGCTGCCGAC
 TGTGTTAGGAGCCTCTCACATTCCAACATTCCCTGAAGAATAGCAACCATTACCAT
 CTAAGCAGCCCTTATCTATTGCTTCTTGAGTAGCAGCTTGTATCCCCGGTTCC
 GATATGACCAAACATTGACCTTAACTTGAAGAGGTTTACCCCTAACACTAGCGTTGG
 TTATGGTACACCTTGCTACCTCCATCGCGTTGCTGGCTACCCCTCAGCTTGTATC
 GGAGCCGTGCCGATGAAATTAGGACCACTTGTAGAGGTACTATGGGGTTAGAGTCC
 CCCCGCCTCTTAGAAAGAAGGGGTTGAACCCCTCTGAAGAGAGTCAAACACTCTGGT
 CTCCACTACACCACTCTAGTAACGTACGTAATTAGCTTGGGCCATACCCAA
 CCATGCAGGTTAAACCTACCGTTGCTGCATGCCTCACCCCTACGCTTTGCTGTTCTG
 GTATCCACATTAATCTCGGCCACCAACTAGCTTACCGAGCTCGCATTGACTCCCTGCC
 TGGATTGGACTAGAGATAACACCATCGCTATTCTACCCCTCTTAGCAAGAGTATCAC
 CCCCGAGGCTATGAAAGGCCCTCAAAATTCTACATACAGGAGTACAGCTTCTAGCACT
 TTACTCTTGCTGATGCACCCCTTAACGCGATGATTAACAGGAGCAGTGGGCCATACCCAGGAC
 TTCCCTCTGCTGGCACCATGCTATTATGCCTTAACATTAAGGTTGGTCTCGCC
 CCCCTACATGCCGACTACCCGAAGCTTACAGGGATTAAGGTTAGTAAACAGATTAC

ATTGCTTGTCAAGGAAATTGCGAGTTAACCCCCCGTACCTTAATAAGAATGGCCCT
 CCCCTCGCAGCTAGCCTCCAAGACCCAGGCTCCCCCTAATAGAGGAGCTGCTCCATT
 CCACGACCACGCCCTAATAGTTGTCTTTAATTAGCACAATAGTTCTTACATTATTGT
 TGCATAGTTACAGCTAAATTAAACGGACAACCTGTTGACTCACAAGAAATTGAAAT
 TATCTGAACAGTACTTCCTGCAATTATTCTAATTCTTATTGCCCTGCCCTCGGAT
 TCTTATTAAAGACGAAATTATGACCCCCACCTACAATTAAAGCCTTAGGCCACCA
 GTGATATTGAAGTTATGAGTATACCGATTATGAAGATCTGGCTTGACTCCTATATAAT
 TCCTACTCATGACCTCACCCCTGGACAGTCCGACTCTAGAAGCCGACCACCGTATGGT
 TATCCCCTAGAATCTGGCTGTACTTATTCTGCTGAAGACGTCCTCCTACTCTG
 AGCCGTCCTGCCCTGGCATTAAGCCGACGCGACTCCCGCCTAAACCAAACAC
 CTTTATTGTTAACCGCCCCGGTGTCTTATGGACAATGTTCCGAAATCTGGCGCAA
 TCACAGCTTCATACCTATTGTTGGAGGCCATCCCTAGAACATTTGAAAATGAAAC
 CTCACTAATAAGTGAAGAAACTTCGCTAAGTAGCTAACAGGCCCTAGCGTTAGCCTTT
 AACGCTAAAGAATGGTGAATGCCAACACCCCTAGCGATAACCTCAACTCTGCCCGCC
 CGTATTGCTATCTTACCTTACCTGACTTGTCTTAAACACTAACATTGTTCCAAAAA
 TCCTAGGCCCCACGTAACGCTACTTCAAGTGTGAGAAAATCTAAAAAAC
 CCCCTGATCTTGACCATGACATTAAGCGCTTCGACCAATTCTGTCCCCACTTTGCT
 CGGTATTCCCTTAATTGCTCTTGCCCTCATTTACCCGACTCCCTAACCGACACCTCA
 GCCCGCACCTCTCACCGCTGACTAGCCAACCGACTCTTACATTCAAGAACTGATTCA
 TAATCTGAATAATTAGTCAAGTTATTCTCCCTAAACCCAGGGGTCAAAATGAGCAGC
 CCTCTTATATCACTTATGATTTTATTACCCCTAAATATACTTGGCTTACTTCCCTA
 CACCTTACCCCTACAACACAACCTCTCCCTAGCCTAGGCCCTCGCATTCCCTTATGATT
 AGCCACCGTGATTCTGGTATACGAAATCAACCAACAGGAGCTTGGCCACCTCTTCC
 AAAAGGTACCCCTACACTCTTACCTTACCGTACTTACGTTGATGAAACAATTAGTGTCTT
 TATTGACACTTCTTCTTAGGAGTGTGCACTTACCGGAATCTCACAGCCGGCACCTCT
 TATTCACTTAAATTCTCAGCGTATTAGTAATGCTACTTCAAGCATCTCTATTGCTAT
 CCCTACAATGGTTATCCTCTTCTTACGCTGCTAGAAACAGCCGGTGTGTTATTCA
 GGCATGTTGCTTCTGCTTAAAGCTATATTACAGAGAAACATTAAATGGCACT
 CAAGCACACCCCTATCATTAGTAGAGCCCCAGGCCATGGCCCTCACAGGCCATTGCA
 GCCCTTATAATTACATCGGAACGCCATCTGATTTCACCTCCACCTCACCTCA
 ACTCTGGCATAGCTGCTAATTAACTTAAACAGTTACGTTACATGAGGCCACACAGCA
 TTATTGAAGGAGAACGGAAACAGGCCATCCAAGCCCTCACCTAACAAATTAACTAGGCG
 GGTATTTTACTTCTTCAAGCCATAGAGTACTGCGAAGCCCTTACCATCGCCGATG
 GCGTGTATGGGCTACCTCTTGTGCAACAGGCTTCCATGGACTCATGCTTAAATTG
 GCTCCACCTTCTAGGCATCTGCCACCTGCGCAAGTCAAACACCATTAGTCAACCC
 ATCACTTTGGGTCAGCAGGCCCTGACTGACACTTTGTTGATGTTGCTGATTAC
 TCCCTTACATCTCATCTATGTTGAGGCTCTAAACACAGTCAAGTACGGTAATTATTAA
 TAAACTTATACCTAACCTTAAACCCCTAATAAGGCCACCCCTAATATTAAACATTAA
 GCAATCTGAGCAGCACCTAGACCCCTATGCTTACATAGCACGCCCTACCTCTAGTAT
 TAAATAAGTATAAGTCACTTCAATCACCTGGCTTGGTTAAATCAAGGAAGGTAAT
 GAACCTTAATCATAGCCATGCCGCCATTGCCCTACACTTCAACAACTCCTGCTATCGT
 ATCTTTGACTTCTCAAAATAAGCCCGACTATGAGAAAATCTGCCCTATGAATGTGG
 TTTGACCCGTCGGCTCAGCTGGTTACCTTTCAATACGATTCTTTAATCGCTAT
 TCTATTCTACTCTCGATCTGGAGATCGCATTACTCTCACCTCTCCCTGAGGAGTCCA
 ACTTTATCGCCCTTAACACCCCTATCTGAGCCTCCGGTTTTGTTATTGCT
 CTCCTCTGTGAACACTCAACTAGCTAACCAACTTTCAGCATCCCCCATACTCTG
 CTTGCATTTCAGCTGTGAAGCAAGCCGGACTGCCCTTCTAGTAGTACTGCTCGA
 ACTCACGGGACCGACGATTACAAAGCTAAACTTACTCCAATGCTAAAATTCTAATCC
 CAACACTTATGTAATCCAAACAGCCTGACTAATCAAACCTAACTGACTCTGCCCTAA
 CTCGTTATAGCTCTGCATCTCTTAATTAGCCTCTGTTGATTTAAACCTGTCAG
 AAACCGGCTGATCTCACTTAACCTGTTATAGCTACCGACTCTCTATCAACACCCCTCC
 TCGCCTTACATGCTGACTTACCTACTGATGCTCTGGCAAGGCCAAACACAGCTT
 TAGAACCCCTGGGGCTAACCGCATGTCACACTCTCACTCTGCATCACTCCAGTTTTC
 TGATTTAGCATTAGCGCCACTGAACTATCATGTTTATGTAATTATTGAAGCAGTC
 TCATCCCAACACTGATTATTACCCGCTGAGGTAACCAAACAGAACGCTAAATGCGG
 GAACCTACTTCTATTATACATTAGCGGCTCCCTCCCTCTGTTGCTCTCT
 TACTTCAAACACATCCGGCACTCTTCAACTCTGCAATTACCGACCCATTTC
 CTCTATCATCTATGAGACAAATTATGATGAGCGGGTGTCTTAGCATTCTCGTTA
 AAATGCCACTGATGGGTAACCCCTGGCTCCCAAAGCCCATGTTGAAGCCCCAATCG
 CAGGCTCAATAATTCTGAGCAGCTACTAAACTAGGGGATACGGCATAATCCG

>Pleuronectes_plateissa 15012 bp

GCTAACGCTTAATTAAAGCATAACACTGAAAGATGTTAAGATGGCCCTAGAACGCC
 CGGGGGCACAAGGCTGGCTCTGACTTACTGTCACCTTAACAACTTACATGCA
 ACTATCCGCCCCCTGTGAGAATGCCAACACTCCCTGCTGGGAACTAGGAGCCGGTAT
 CAGGCACAAGCCAGTAGCCCACGACGCCCTGCTAGCCACACCCCTCAAGGGACCTCAG
 CAGTGATAAATATTAGCCATAAGTGAAGCTTAACATCCACCTAGAGGAGCTGTTCTA
 GAACCGATAACCCCGTTAACCTCACCTCTCTGTTTATCCGCTATATTACCCACCGT
 CGTCAGCTTACCCGTGAAGGATTTACAGTAAGCAAATTGGCACAGGCCAAACGTCAG
 GTCGAGGTGAGTGAATGGGGGAGAAATGGGCTACATTGCTAAGTATAGCAAACA
 CGATGTTGATTGAAACATGCACTGAGCTAAGGAGGATTAGCAGTAAGCAGGAAGTAGAGC
 GTCCCGCTGAAACTGCCCTAAAGCGCACACACCGCCGTCACCCCTCCCAAGCCCC
 TGAACATACTAATTAAAGCCAACAACCCCGCAAGGGGAGGAAAGTCGTAACATGGTAA
 GTGTACNGGAAGGTGACTTGGAAAATCAGAGCGTACCTAAGATAGATAACAGCATCTCA
 CTTACACCAGAAGACGTCGGATGCCCTGACGCCATTAGCTAGCCCC
 CCCCTTAACACAAACCCCTAACCTAACCCCTAAACCCAGGAAACACCCACGTA
 AACCATCTCCCCCTAACCTGAGCGATAAAAAGGAAATTGGAGGCAATAGAAAAAGT
 ACCGCAAGGGAAAGCTGAAAGAGGATGAAAAGCCAGTAAAGCTTAGAGAAGCAGAGC
 TTAAGCTGTAACCTTTGCTCATGATTAGCTAGCATTcAAGCGAAGCAAGAGaAC

CTAAAGTTGTAACCCGAAACTGAGTGAAGCTACTCCAAGACAGCCTATTTATAGGGCGA
ACCCGTCTCTGTGCCAAAAGACTGGGAAGAGCTTGAGTAGAGGTGACAAACCTACCGAA
CTTAGTTATAGCTGGCTGTGAATTGAATAGAAGTTAGCAGCCCCCTGGATTCTCCACT
CATGCACTTTATTTAACCTTCAGACCAATGAGAAACCCAGGGGTGTTAGTCAAAGGGGG
TACAGCCCCTTGAAACAAGACACAACCTTCCCAGCAGGATAAGATCATATTCAAATAA
GGACAGATGTTTAGTGGGCTAAAGCAGCCACCTAACAGAAAGCGTTAAAGCTCAA
CATGAAGCCCTCCGTTATACCGATAACCTTATCTTAATCCCCACATTTAACAGGCCCT
CCTATGCATCACATAGGAACGACTATGCTAATATGAGTAAAGAAAGTATAACCAACTT
CTCTTGCACATGTGTTAAACCGAACGGACCCCCCACCAGAATCTAACGGAAATACCCCC
AGCTAAGGGCTGAACTAATGGAAACTGCGTTGATGTTGGGCGACCATGGG
GAATAACAAAACCCACAGTGGAAAGGGAGCACACCCCTAAGTTACTCTTCTCCGCAAG
CCAGAGCAACAGCTAACAAGCAGAAATTCTGACCAAACGTATCCGTTAAACCGATCA
ACGAACCAAGTTACCCCTAGGGATAACAGCGCAATCCCCCTTAGAGGCCATATCGACAAG
GGGTTTACGACCTCGATGTGGATCAGGACATCTAATGTTGAGCCGCTATTAGGGT
TCGTTGTTCAACGATTAAAGTCTACGTGATCTGAGTTGAGCCGAGTAATCCAGGTC
AGTTCTATCTACATGACATGTTCTAGTACGAAAGGAGCGAAAAGAAGGGGCCATG
CTAAAGTGACGCCCTACCCCCCACCTAATGAAAAAAACTAAACTAGACAAAGGGCATAAC
CATTTGCTGGAGATAACAGCAAGTGGGGTGGCAGAGCCGGCTAATGCAAAAGACCTA
AGCCCTTCCACAGAGGTTCAAGTCTCTCTTAACATGATTCAACCCCTATTACGCA
TATTATTAACCCACTAGCCTCATTGTCGGTTTGTAGCCGTAGCATTCTAACCC
CCTGAACGGAAAGTCTAGGCTATATACAACACTACGAAAAGGCCTAACATCGTGGGCC
TTAGGCCTCTTCAACCCATCGCTGACGGCTAAAGCTATTATCAAAGAGCCGTTCG
CCCTTCGACTCTCGCCGCTCTGTTCTCGTAGGCCCTGACTCACACTAGC
ACTTACACTCTGAGCCCCATGCTTCCCTACCCCTGTTGACCTTAACCTAGGCAT
TTTGTATCTAGCATTATCCAGCCTTCAGTTACTCATCTGGGTAGGGTGGC
ATCTAATTCAAAGTATGCACTGGTAGGAGCGTACGGCTGTTGCCAGACAACTTCTA
CGAAGTAAGCCTCGGACTCATCTTACTAACATCATTATCTTACCGGAGGGTACGCT
CCAGACCTTAAACGGCCAAGAGGCTTTGACTGTTGACTACCCGCTTGACCACGGC
TGCTATATGATACATCTCACGCTGCAAGTACGTTACTCATCTGGGTAGGGTGGC
GGGGAGTCGGAGCTTGCTCAGGCTCAACGTTAGAGTACGCCGGAGGGCTTCTGCTT
ATTCTTTAGCGAGTACTAACATCTCTAACATTTAGCTATCCGAAACACTATT
CTTAGGGCCTCCACATTCCATCCCAGCTTGTGGCTGAGCTCATACCCGATTCTGTT
TGACCACTTATGCACCTCATTTGAAAGAAACTTCTCCCTGACGCTAGCGTAGTTAT
TTGACACTTAGCGCTCCATTGCAATTGCTGGCCTCCACACAACTGTAAGCGGGGA
GCTGTGCTGAATTAAAGGGCCACTTTGATAGA_gTGAATCATGGGGTTAAAGTCCCC
CAACTCCTAGAAAAGAAGGGATGCAACCCAACTGAGAGATCAAACACTTCGCT
CCTACACCACTCTCTAGTAAAGTCAGCTAAAGGCTCTGGGCCCATAACCCAAACCA
TGTAGTTAAAACCTGCTTGTCTAGTAAACCCCTACATTGACCAACCTCTATTG
GTTGGGCTAGGTACAACACTCACATTGCAAGTTCGCACTGACTCTCGTTGAATGG
GCCTCGAAATTAAACACACTGGCATTATTCCATTAAATGGCTCAACATCACCACCGAG
CAGTAGAAGCTACTAAATTTTTAGCACAAGCCACAGCAGCCGACCCCTCT
TTGCTAGCACCACCAACGCCGTACTTACCGGCCAGTGAGACATCCAACAGATGACCCACC
CACTCCCACCAATAATTGTATTGCTGGCCTAGCATTTAAAGGACTGGCACCATGC
ACTCTGACTCCAGAGGTTCTCAAGGCTAGACCTTACCCGGCTCATCCTATCAA
CCTGCAAAGGCTGACCTTGTCTTGTCTAACATTAAACACCAACCCACAC
CTCTGATCATTATTGCTGCTTCTGCCCTGTGGGGGTGAGGGGGCTAAACCAAA
CTCAGCTACGTTAAACCTTGCCTACTCATGATTGCAACCTGGGTGAATGACTTA
TTCTCAATTTCACCCTCAGTGTAAAGCACTCGCAATATCGTAGGAAAC
TCCCGCTTACAGCCCTAGCGCTCTGTTTACTTCATCGGTGGCCTCCCCACT
CGCTGGCTTACGCAAATGATTCTCAAGGACTACCCGGCTCATCCT
ACTTGCTACCATCGTGCATGACTGCCCTTAAAGCTTACTTACCTACGCTCTC
ATATGCAATTGACCCCTAACATGTTCTAACACCTCGTTGGTGTACCCCTGACGATT
TTACTCCCCCTCAATTCAACCTCCCCCTGCCGTCTCACTGCAACTACCCCTTCT
ACCAACTAGCCCTGCTGCCGTGGCACTCCATCACTTAGGGACTTAGGCTAGCAATTAG
ACCAACGGCCTCAAGCCGTGAGCTGAAATCTCAGTCCCTGTTAGACTTG
CGGGACATTATCCCACATCTCTGATGCAACAGCAGACACTTTAAGCTAAAGCCTT
TCTAGATAGGAAGGCCATTGATCTAACAAATCTTAGTAAACAGCTAAGCGCCCCATCCGG
CGACGCTATCTACTTTCCGCCAGTTAGTAACAAAAGGCGGGAAAAGCCCGGC
AGACGATTAGTCTGCTTCTAGTTGCAATCTAACATGTAACACCCAGGGCTGATA
AGAAGAGGGCTGACCTCTGTCTATGGCTTACAATCCACCGCTTAACCTACG
ACCTGTGGCAATCACAGTGTATTCTCGACCAATCAAAGACATCGGCACCCCTCTA
TCTCGTATTGCTGGCTGAGCCGAATAGTGGGAGCGACCAATTATAATGTAATCGTCAC
AGAGCTAAGCCAACCTGGGCTCTCTGGAGACGACCAATTATAATGTAATCGTCAC
CCACACGCCCTGTAATAATCTTCTTATAGTAAACCAATTATGATTGGGGTTGG
AAACTGACTTATCCATTAAATATTGGGCCCGGATATGGCCTCTCGAATAATAA
TATGAGCTTCTGACTTACCCCCCTCTCTGCTTCTCTGGCTTCAAGGTGTTGA
AGCCGGGGCGGGAAACAGGGTGAACCGTATACCCCTTACTGAGTGGAAACCTAGCACACGC
TGGGGCATCCGTAGACCTCACAAATTCTCTCCACCTTGGCGAATTCTCATCAATTCT
AGGGGCAATCAACTCATTACCATCATCAACATGAAGCCTACAGCAGTCACTATGTA
CCAAATCCCACATTGTTGAGCCGTACTAATTACGCCGTCTTCTCTCCCT
TCCCGTTTAGGCCGTGGCATTACAATGCTACTAACAGACCGCAACCTAAACACACCTT
CTTGACCCCTGCTGGAGGGGGTGAACCCCATCCTCAACACCTATTCTGATTCTTGG
CCACCCAGGGTATACTTTAATTCTCCAGGCTTGGGATGATTCTCAGATCGTTGC
ATACTATGCAAGGTTAGAAAAGAACCCCTTGGCTATATAGGAATAGTCTGAGCTATGAG
TATTGGACTCTGGCTTCTGTTGAGGTTAATGCTACATTGACATTCTCCCCAACACTT
CCTGGCTGGCAGGTACCTCGACGGTATTGAGTACCCCTGATGCAATACCCCTTG
AAATACTGTCTCATCAATCGGGTCAACTAATGTCCTGTTGCTGTGATTAAATT
TATTATTTGAGAAGCATTACTGCCAACGAGAAGTGGGGCAGTAGAACTAACCTCAAC

CAATATTGAATGACTTACGGCTGCCCTCACCCTATCACACATTGAGAGGCCGCATT
CGTCAAGTCGTATAAATTGAAACGCTAACGAGAAAGGGAGGAGTTGAACCCCCATCA
ATTGGTTCAAGCCAACCACATAACCGCTGTCACTTCTCAACAACACACCAATAAGA
TACTAGTAAACGTTAACACTGCCTTGTCAAGGCAGAATTGTGGGTTCAACCCCGCG
TATCTAAACGCAATGGCACATCCATCACAACTGGGATTCAGGACGCAGTCCCCCT
AATAGAAGAACTACTTCACTCCATGATCACGCCCTAACATTGTTATTCTCATTAGCAC
GATAGTACTATATTATTGTGGCAATGGTCACGGCCAACTAACCGATAAGCTTGTGTT
AGACTCTCAAGAAATTGAGGTTATCTGAACAGTCTCCCGAGCTATTATTCTCATCTCAT
CGCCCTCCCATCACTCGAATTATTAATAGACCAAATTAGACCCCCACTGAC
AATAAAGCCTGGCCACCAATGGTACTGAAGCTACGAATCACAGACTACCAAGACCT
CGGTTTGACTCATATAATCCCAACTCAAGACCTAACCCCTGGACAATTCCGACTACT
AGAGGCAGACCATCGAATGGTAATTCTGTAGAATCCCCAATTGAGTTCTTATCTCAGC
TGAGGATGTCTACACTCTGAGCAGTCCCCCTCCCTGGCGTAAAGTAGACGCCGTACC
TGGACGGTAAACCAAGCAACCTTTATTGTTAGCCGACCCGGCGTGTCTACGCCAATG
CTCTGAAATTGCGGAGCAACACCAGCTTATGCCCCTCGTTGTAGAGGCAGTCCACT
TGACCACTTGTGAGCTTCTCGTAATAATTGAAAGCAGCCTCCGTAAGAAGCTAATA
AGTACAAGCGTTAGCCTTAAAGCTAAAGACTGGTGCTTAACACCACCCCTAGGCCACAT
GCCTCAACTGAACCCCGCACCCGTGATTGCAATTGTTGGTTCTCTGAATAATCCTTT
AACTGTCATTCCCCAAAAGTTTGTCTACACCTACCCATAATGAGCCAACCTCCAAAG
CACACAAAACCTAAACAGAACCCCTGAAACTGACCATGATACTAACGCTTGTGATCAA
TTTATATCCCCGTATTCTGGTATTCACTAACCGTCAATTGCACTCGCAATTACCCGCCCTGA
ACACTCTCCGACCCCGTGTCCCGTCACTAACACAACCCGCTAACCTCAACTTCAAGGA
TGATTATTAGCGGCTTAACTCACAACCTCTCCGCTAACCCGGGGGACACAA
TGAGCACTCTATTGCTCACTGATGCTTATCTCATTTCTATTACATGCTTGGACTC
CTTCGTAACCCCTCACACCCACAACCGCAACTTCACTAACCTCGCCCTCGAGTTCCG
CTCTGATTAGCAACCGTCAATTGGGATGCGAAACCAACACAGTTGCTTAGGCCAC
CTTCTCCAGAAGGAACCTCCACCCCTCTAACCCCTGACTAACATCATCGAAACAAATT
AGCCTATTATTGACCTCTCGCTTGGTGTGACTAACAGCAACCTCACAGGGG
CACCTGCTTATTCAACTCATGCAACAGCTGCCCTCGCTTGGCTTATGCCGTT
GTTGCTATTAAACACAAAGCTTCTCTGCTTACTCTACTAGAAGTAGCAGTCGCT
ATAATTCAAGCCTATGTCTTGTATTACTCTAACGCTTACCTAACAGAAAAGCTAA
TGGCCATCAAGCACCCCATACACATAGTCGACCCCTAGCCATGACCCCTCACGGGG
CTATTGCTGCCATTGATAACATCTGCTTGTCTATGATTCCACTCCACCAA
CCCTTATAACTATGGAACAGCTCTTATTAAACGTTCTCAATGATGACGTGATG
TCGTCGAGAAGGCACATTCAAGGACACCAACTCCCTCCGTTAAAAAGGCTCCGAT
ACGGAATAATTCTTATTACCTCAGAAGTTCTATTCTCTAGGTTCTGGCAT
TTTACCACTCAAGCCTGGCACCCACCCCTGAAGTCCACTCTAAACACAGCAGTCTACTGCT
TCACCCCTTGGACCCATTGAGTCCACTCTAAACACAGCAGTCTACTGCT
GCCACTGTATTGACCATGAGGCTTACATGAGGTTAACGAAACAGCAATT
AATCCCTGCCCTAACATCTACTCGGGGACTTCACTTCACTTCTCCAAGGCTTGAGT
ACCACGAAGCCCTTCACTATTGCAAGCAGGGTTACGGTGCCACATTCTTGCTGCCA
CCGGCTTCCACGGACTTCACGTCTTAGTGGCTCTCTTAGCCGTTGTCTACTAC
GCCAATTCTACACCATTTACATGACCCACACTTGGATTGAAGCAGCCGATGAT
ACTGACACTTGTAGACGCTGTACTCTCTCTATCTTACTGATGAGGAT
CTTAATCTCTCTAGTCAAAACTCTAGTAAAGTGACTTCAACCTAACCTGGTTAA
AATCCAAGGGAAAGTAAAGCCTTCTAACCATCTAACATTACCGCCACATGAGAAACT
CTCACCATACGAATGTGGCTTGACCCCATAGGTTCCCGGACTACCTTCTCCCTACG
ATTTTTCTCATCGCAATCTCTTCTTCGACTGGAGATTGCTCTTACTCC
CCTCCCTGAGGGGACCAACTAGCATCCCGTACTTACATTACTTGAGCTACGCCGT
CCTATCCCTTGCCCTGCCTCATCTACGAATGAATACAAGGAGGCCCTAGAATGAGC
TGAATAGGGTTAGTCTAAAGAAAACATTGAGTTGGCTCTAAACATTGGTTAA
TCCCAACCGCTTAATGACCCACACATTGCTCTCTCTCAGCCTTTCTAGGTT
TAACAGGCCCTGGCATTCACCCCGTTCCACCTCTCCGCTACTGTGCCCTGAAGGGA
TAATACTATCCATTGTTGGCTCTCTCTATGAACCCCTCAACTGGACTCAACTAACT
TTTCAGCATCTCCATTGCTCTCTTGCTATTTCAGCCTGCGAAGCAAGGCCGGCTCG
CCCTCTAGTGCCACTGCCGAACCCACGGGACCGACTGCAAAGCTAACCTCC
TCCATGCTAAATTCTAACCCACATTGCTAACCTAACAGCCTGACTACTCAA
ACCTAGCTGGCTGTGACCCCTGACTTTATTATAGCTCTGCATTCCCTAAATTAGCCT
CTCATGACTAAAGAACCTCTCAGAGACGGCTGATCCTCACTCAATGTTCATGCCAC
CGACTCCCTGTCACACCACTCCCTGCTCACATGCTGCCCTCTCCATTGATAATT
GGCAAGCCAAAGCATACAGCTCTGAGCCCTCGGCCACAACGTTATACATCACACT
CCTCGCTCACTCCAATTCTCTAATCTAGCATTAGGCCACTGAACTAGTAATT
CTACGTAATTGAGCTACTCTCATCCCCACCCCTTATCATTATTACCCGCTGAGGTTAA
CCAACACGAAACGCCCTAACATGAGGGACTACTCCCTCTACACCCCTGGCGGGCTC
TCCCTCTTGTGCGCTTGCTTACTCCAGAACACATGCCGACTCTCTATTGAC
CCTCCACTACGGGACCCCTAGCCCTATCATCTCTGAGAACAAATTGTTGAGGG
CTGCTTTGGCTTCTAGGTTAGATACCACTTATGGCTTCACCTCTGGCTCCCTAA
ACGACACGTTGAAGCCCAATTGCAAGCTCCATAATCTTGCACTGGCTTACTTAA
AGGGGGTTACGGCATAATTGCGATAATAACAAACTAGAGCCTCTAACTAAAGAATTAG
CTACCCATTCTTGCTTGTACTCTGGGTGATCATGACGGCTCAATTGCTACG
CCAACACGACCTTAAGTCCCTAACCTCATCTGAGGCCACATGGGCTCGTAGC
TGGAGGGATTCTCATCTAACATCACCCTGAGGCCCTCACAGGAGCAGTACACTTAAATTG
ACATGGCCTCACCTCCTCAGGCCCTATTCTGCCCTCGCAAAACACAAACTATGAGCGGACTCA
CAGCGTACAATGGTCTGGCACGAGGACTCAAGTGGCTTACCCCTAACGCACTTG
ATGATTATCTCTAGTCTGGCAACCTCGCCCTGGCACCCTGCCTAACCTCATGGGG
GCTCATATTACACTCTCTCTTGTACTGATCTGGTGAACCCCTGGTATTAAACAGGTT
TGGCACTCTTACAGCCGTTACAGCCTCTACATATTCTTAATAACTCAACGAGGCC
TCTCCCAACACATGTAATTGCACTTGAACCATCACACACCCGAGAGCATCTTATTGC

ACTTCATCTTACCCCTCTTATTCTCCTCGCTTAAGCCCGAGCTAATCTGAGGGTGAAC
 TGCTGTAGGTATAGTTAACAAAACATTAGATTGTGATTCTAGAAATAAGGGTTAAA
 ATCCCTTTCCCACCGAGAGAGGCTCGCAGCAATGGGAACTGCTAATTCCCAGACCTTG
 GTTGGACCCCCAGGCTCACTCGAAGGGCTCCTAAAGGATAACAGCTCATCCGTGGTCTT
 AGGAACAAAAACTCTGGTGCACATCCAAGTAGCAGCTATGCACCCACCTCCCTAATG
 ATCTCATCAAGCTGGTACAAATCTTGCTTACAGGCTACCCCTGGTACCCACAATT
 CGTCAACACCCGGGGCCCTCACTGGCAACATCCCATGTCAAAACAGCTGTAAAATA
 GCTTCTTGTAGCCTTACCCCTCGTCTGTTAATGAAGGTGCAGAGACGATT
 GTTACTAATGAACTGAAATAACAAACAGCTCAACATCAACATTAGCCTAAATTT
 GACTTATTCAATTTACACCAATGCCCTTACGTACCTGGTCTATTCTAGAG
 TTTGCATCATGATATATGCACGCCGACCGTAAACCGCTTTCAAGTACCTCTG
 ACATTTCTAATTGCCATGATCATTAGTAACCGCAAACAACATATTCACTCTTATC
 GGTGGGAGGGTAGGAATCATGTCCTCCCTCATCGGTGATGGTACGGACGGCG
 GACGCTAACACTGCACTCAAGCAGTACTGTACAACCGAGTCGGGATATTGGCTT
 ATCTCGCATGGCTGAATAGCAACAAACCTAAATTCTGAGAAATACAACAAATT
 gCGACCTTACAGGCTGAATAGCAACAAACCTAAATTCTGAGAAATACAACAAATT
 GGGAGTCAGCCCAGTTGGACTCCACCCCTGGTACCCCTGGCATGGAGGGTCTACA
 CCGGTATCTGCCCTACTTCATTCTAGCACTATGGTTGGCGCATTTCTGCTGGT
 CGAATAAGCCCCCTCTGGAGATAACCGACCGCCCTACAACCTGCTTGTCTCGGT
 GCCCTACGACACTATTACTGCAACTGTGCCCTACCCAGAAATGATATTAAAAAA
 CGTGCATTCTCATCCAGCCAACCTGGCTCATATAAGTAACCATCGGCCGAACCA
 GCCACAACTAGCATTCTCACATCTGCCACACAGCCTCTTAAAGCCATACATTCT
 TTGCTCGGGTCTATTATTCACAGCCTAACAGCAGGAGACATCCGAAAATAGGCG
 CATGCACCACCTTGCACCTTACATCATCTGCCTAAAGATGCTATCATCGAAC
 AGGAACCCCCCTCTAGCGGTTCTCTAAAGATGCTATCATCGAAC
 ATCCCACCTAACGCCGAGCCCTAACCCCTACCCGTGGCCACATCTTCACAGCAAT
 CTATAGCTCCGAGTAATTCTTGTACCCATAGGCCACCCCGATTAGCCACTCTC
 CCCCATCAACGAGAACACCAGCAGTAATTAAACCCACTAAACGACTAGCATGGCAG
 TATCATTGCAGGACTACTAATTACCTCGAACATTACCCCTAAACACCTGTTATGTC
 TATGCTCCCCCTCTTAAAGCTAGGCCCTCGCAGTTACAATCTGGCCTACTTATTG
 TATAGAACCTGGCATGTCACCAACAAAGTCAAGTCAACCTTAAAGGTACA
 TAACCTCTCAACATGCTAGGCTTCTCCCGCGTTGATCGCTAAACCTTAAGTT
 AGGCCTAGTTCTGCCAAGACATTGCTAACAGATAAGTAGATCAGACCTGACTAGAGAA
 AGCAGGCCCAAGGCCATTGTATCCTCTAACCTCCCCCTGGTTCTCTACAGTAATAT
 TCAACGAGGAATAATCAAAACGTACCTAACACTTTCTTTAACACTAGCCCTTATAAT
 CCCAACACTCATCCCTAACGCCGCTCGCAACGCCCTCGGCTAACCCACGAGTTACTC
 AAGGACTAACAAACAGGTGAGAACACCCACGCACAAATAATTAAATCACCCCC
 AGAAGAGTATATAAGCCACCCCTCCAATATCACCCGAAACACGGAGAATCGGCAA
 CTCATCTACTAGAACCCACGACTCATATCACCCACCTCAGAACACCCAGAGACTAA
 GACTACCCCTCCACGTTAACCAACCCGTATACCATACCGACCAGCTACCCAGCTCTC
 AGGGTATGGCTCAGCGGCTAACGCCGCTGAATATGCAAATAACACTAACATCCGCCAA
 GTAAATTAGGAAAAGAAACTAAGGACAAAAGGGCCCCATGCCGACTAAACACCACA
 CCCATGCCAGTACCCACAAACCAACGCCAGAACAGTAGGGGACGGTTAGAAGC
 AACGTCAACTAACCCAAATCAAGAGAAAATAAAACTAATATAACAAAAGTCATAAT
 TCTCGCAGGACTTAAACCGAGACTATGGCTTGAAAACCCACCGTGTATTCAACTAC
 AAGAACCTTAATGGCAACCTCGTAAATCCCACCCCTTCTAAATCGCAAACGATG
 TTTAGTCGATCTCCAGCCCCCTCTAACATCTGTGATGAAACTTGGCTCCCTCTT
 AGGACTCTGTTAGTGAACATTGCTACCGGTTATTCTAGCCATGCACTATACATC
 TGATATTGCTACTGCTTCACCTCCGTGACACATCTGCCGGACGTCAACTACGGCTG
 ACTTATCCGAAGCATTCTGCAACGGCGCATCATTCTTTTCAATTGCTACTCTCA
 TATCGGCCGAGGTCTACTATGCTCTTACCTTATAAGGAAACATGAACTATCGGAGT
 TGTCTACTGCTCTCGTAAATGACACGCTTCTGGCTACGTCCTCCCTGGAGGACA
 AATGTCATTGAGGTCAACCGTCACTAACCTTATCTGCCGTCCCTATGTCGG
 AGGCACACTTGTCCAGTGGATTGAGGTGGCTTTCTGTAGACAATGCAACCTCACCG
 GTTCTTGCACTCCACTTCTCTCCATTATCATTCGCCCGCAACGGTATCCACCT
 ACTCTTCTTCAGAAACTGGTCAAAATAACCCACCCGACTCAACTCAGACTGATAAA
 AGTCCCTTCAACCCCTACTCACATACAAAGACCTTACGGCTCGAGTCCTCCTCAC
 TGCACGGCTTACTCGCCCTATTTCCCAATCTTCTGGAGACCCAGAACACTTCAC
 GCCTGCCAACCCGCTCGTACGCCACACATCAAGGAGACTGATACTCTCTTGT
 CtACGCCATTCTCGCTCCATCCCAACAAAGCTTGGGGCGTACTTGGCCCTTTATTCTC
 TATCTCGTTCTCATGCTGCCCTTCTCCACACCTCTAAACACGAAGCTGATATT
 TCGCCCTGTAACACAATTCTGTTCTGGTCTAGTAGCTGACGTAATAATTCTGACCTG
 AATTGGAGGAATGCCGTAGAACACCCCTTCGTTATCATGGACAAGTAGCATCTCTCAT
 CTACTTCTCCCTTTCTAGTCCTGATCCAACCTGCAAGGCTGAATGGAAAATAAGTCCT
 TGGATGAAAATGCACTAGTAGCTCAGCGTCCAGAGGCGAGTCTTGAAACTGACCGTCG
 GAGGTTAAAATCCCCCTACTGCTCAAAGAAAGGAGATTCAACTCCCTACCCCTAACTC
 CCAAGCTAGGATTCTAGCTAAACATTCTTGGCGTCAAACATGTCATGAAAGCT
 TTCATGTACATGTTGATAAAACACCATATTTATAGTAACCACTTATATAATGAACT
 AGGACATTCTGTTGAAACACCAACTGCTAACACTCACGACAATATTGAACT
 TAAGAAATACTAGAACCTAACGAAGCACTAACCTCACATTAGTATACTTCCAGGACAGT
 CGAAATTAAAGACCGAACACACACTCATCGAGCTGAGTTACCAAGACTCAAACCC
 TCAAGTCAAAACCGTATGTTAGTAAGACCTACCAACCGGTGATTACTTAATGATAACGGT
 TATTGAAGGTGAGGGACAAAAATTGTGGGGTTTCACTCCGTGATTATTCTGGCATT
 GGTCTACTTCAGGGCATAAAATGATTATTCCCCACACTTCATCGACGCTTACATA
 AGTTAATGTTGATAATACATACGACTCGTACCCAGCAAGCCGGCGTTACTCCAGCG
 GTAAGGGTTCT

>Solea_vulgaris 13420 bp

GCTAGCGTAGCTTAACAAAGCATAACACTGAAGATGTTAAGATGGCCCTAGAAAGGCC

CGCAAGCACAAAGGCTGGTCTGACTTATTGTCAGCTAGCTATACTACATGC
AGTCTCCGACCCCGTGAGAATGCCGCTCCCTTATGAGAACTCGGACCAGGTATCA
GGCACACCCCTAGTAGCCCAGCACACCTGTTAGCCACACCCCAAGGGATTCA
GATAAATATTAAGCCATAAGTAAAACCTGACTTAATTAGGCCAAGAGAGCGGTAAAA
CTCGTGCAGCACCAGCGGTGAGACGAGAGGCTAAGTGCACAAACGGCACAAAGCGT
GGTAAGGAATTACAGACTAAAGTAAACGCTCCAAAGCTGTTACGCACTCGAAA
GTATGAAATTCAACCCACGAAAGTAACCTAACACCCCTGAACCCACGAAAGCTAAGGAAC
AAACTGGGATTAGATAACCCCACTATGCTTAGCCATAAACATCGATCTTACACAAAATC
CGCCCGGGAAACTACAACACGAGTTAAAACCCAAAGGACTGGCGGTGCTAACATCCA
CCTAGAGGAGGCTGTTAGAACCGCATAACCCCCGTTAACCTCACCCCCCTAGCTAA
AACCGCCTATATACCAACCGTCGCCAGCTACCCCTGCACAAAGCCCCAAAGTAAGCAA
ACTACGTCAAGACGTCAGGTCAGGTTAGCGCATGGGGGGAGAAAGATGGCTACA
TTCGCTACCCCTAGCGTATAACGAATGACTCGCGTGAACACCGCTTGAGGAGGATT
AGAAGTAAGAAAAAAATAGAGTGTGTTCTGAAACATGGACTAAAGCGCGTACATACCG
CCCGTCACTCTCTCTGAATTACCTTATTTGTTAGTATAACACCATCCAAATAGAG
GAGAGGAAAGCTGTAACATGGTAAGTGTACTGGAAAGTCAGTGGATAATCAGAGTATA
GCTAAATAGTATAGCATCTCCTTACACTGAGAAATCATCGTGCACACCGAATTACCC
GACACCGACAGCTAGCCCCAATATTTAAACAAACAAATCATTATAAAACCCCTAAA
ACACCCAAAGCAATTAAATCAAACCATTTTACTCTAGTACGGGAGACGGAAAGGACAA
AAAGGAGCTATAGATAGAGTACCGCAAGGAACGCTGAAAAAGAGATGAAAGAGCCAGT
TAAGACCATCAAAGCAGAGATTAATCCTGTACCTTGATCATGATTTAGCTAGAAAA
TTCAAGCAAAGAGCCTTTAGTTGAAACCCCGAAACTGAGTGAAGCTACTCCAAGACAGC
CTATAATTAGGCACACCCGCTCTGTGCAAAAGAGTGGGAAGAGGCTTGAGTAGAGG
TGACAGACTACCGAATCTAGTATAGCTGGTGCCTGTGAATTGATAATAGTCAGCC
TCCAGCTTCTAATTCCACAAACCTTGTGACCAAGGATTCCAAGAAACCGAGAGAGT
TAGTCAAAGGGGTACAGCCCTTTGAAACAGACACAACTTTATAGGAGGGTAAAGATC
ATACTATTAAAGGAATTGGCCTGGTGGGCTGAAAGCAGCCATCTAACAGAAAGCGT
TATAGCTCAAGCATCCTTACCCCTAACAGATTAGATAACTCTATCTAACCCCTATCT
TACCCAGGCCCTCTATTACAGAGACTATGCTAATATGAGTAATAAGAAAGACAC
CATCTTCTCACACCATACGTAAACGGGACACCCGGGAAACCTAACAGGACC
CAACCAAAGAGGAAATTGAAAAAAACACAGAAAAACCTCTAACATATTACCGT
TAATCCCACACAGGATGCAATTAAATGAAAGACTAAAGATAAGAGAAGGAACCTGGCA
AAACCTAACAAAGCCTCGCTGTTACAAAACATGCCCTTGTGACCAACAAAAATAAG
AGGTCTGCCTGCCAGTGCACACAAGTTAACGGCCGCGGTATTTGACCGTGCTAAGGT
AGCGCAATCACTGTCTTTAAATGAAGACCCGTATGAATGGCAAACAGAGGGCTTA
GTCCTCTATCCAGTCATGAAATTGATCTCCCGTGAGAAGCGGGGATACCA
AGACGAGAAGACCTATGGGCTTAGACAACAATTCAAGCCGTCCTCCAAATTCTAA
AACGAACACACTAACGGCCCTGTATTAAATGTCCTGGGTCGCGGATCAGCG
AACCAAGTTACCTAGGGATAACAGCGCAATCCTCTTTAGAGCCATATGCACAAAGGG
GTTACGACCTCGATGTTGATCAGGACATCTAACGCGCAGCAGCTATTAGGGT
TTTGTCAACGATTAAGTCTACGTGATCTGAGTTCAACCGGAATAATCAGGT
TTCTATCTATGCTACAATCTTCTAGTACGAAAGGCCAAAAGAGAGGCTATGCTA
ATGGAAGCCTACCCCAACTAGTGAAGAAATCTAAACTAGATAAAGGGGTAACCC
ACGTTAGAGAAAATAACATGTTAGGTGCAAGGCCTGGTACTGCAAGAGACCTAA
TCTCTACGGAGGTTCAAATCTCCCTCAACTGTGATCTAACCTAACCTCATCC
TACACCTTGTGACCAATTGCCCCAGTACTCTAGCCGTCGCCCTCACCCTCTGG
AACGAAAGACAATCAGTTATACAACTGCGAAAGGCCAAATACGTTGGACCC
GACTGCTCAACCAATGCCGACGGCATTAAACTATTAAAGAACCGTTCGCC
CCACAGCATCTCCCTCATTTATCTCGCACCAGTTCTAGCTGACTCTGCTCTG
CCCTCTGACTTCTTACCCCTCCCTATTGATCTTAACCTAGGAATT
TTGTCCTCGACTCTAGCTGACTCAATTAGGGGGATGAGCTCTA
ACTCGAAATATGCCCTATTGGAGCCCTAGGCGTAGCACAAACTATCTCTACGAAG
TGAACCTGGACTAATCTAATCTCGTCACCATTAGTGTGAGAATTGCA
CTTTGCCACATCCAAGAGAGTACATGGATAATTTCCTCTGTGACCCCTAGCG
TATGGTTATTCAACATTAGCAGAACCAACCGTGACCCCTTGATCTACTGAGGGAG
AACATGAGACTATCGGGATTTAACCTGAGTGAATATGCA
TTCTGCTGAGTACTCCAACTCCTCTAATAAATACCCCTCTCGTATTGTT
GTGCCTCTCATACCCGCCCTCCCTGAACTAACCACTTAAATATATAA
CCCTTTATGTGAGTCTCTGACTGAGCATCTCCCGGTTCTGTTATGACC
AACTAATGCATCTAACCTGAAAAACTTCTCCCTAACATTGGCAATGATTG
ACTTAGCCATCCCCATTGCACTCGCCGATTGCCCGCAGCTCTAACAGAGCT
CTGAACATAAGGGCCTTGAAGAGTGAACCATGGGTTAAACCCCTAAC
TAGAAAGAAGGGAAATTGAAACCTTACTGAAGAGATCAAAC
CCACTCTCTAGTGAGGTGAGCTAAAGCTGTAGGGCCATACCC
AGACCCCTACCTTCAACTAAAGTGCACACTAATTTCTACTCT
TCTGGCTCCACAGTAACTTCAAGCTCACACTGACTCTCGCT
AAATTAAATGCAATTGCTCAACTATGGCTGCCACATT
AGCTGCCACAAACTTCCCTCGCTCAAGCGACCGCAGGCC
CATAACCAATGCCGATAACCGGCAATGAACAATT
CCTGACATTAGTGAATTGGCCCTAGC
ACTACCAAGAGATAAGACTTGCA
CTTAATTAAAGCTCTAGACAGGTAGGGCT
ACAGCTAACGGCTAAAC
AGCCGGAGGAAATCCGGCAGGA
AACACCCCGGACTTG
CACTAAA
CAAAGACATCGGCAC
143

CCTAAGCCTACTAATTCGAGCTGAACCTAACGCCAACCGGGCTCCCTCTAGGAGACGATCA
AATCTATAATGTTGTCGTTACTGCACATGCCCTCGTAATGATTTCTTATAGTAA
AGTAATGATTGGAGGCTTCGAAACTGACTCATCCCCATAATGATCGGGGCCCCGATAT
AGCATTCCCGAATAAACATAAGTTTGACTCCTCCCCCTGCTTTCTGCTACT
TCTCACCTCATCGTTGAGGCCGGGGGGACAGGATGAACTGTCTACCCCCCTCT
ATCAAGTAACCTGCCATGCAGGTGCATCGTCGACCTAACATCTCTCCCTCACCT
CGCCGGAGTGTCTCAATCCTGGGAATTAACCTTACACCATCATTAATATGAA
ACCTGCCACTAACAAATACCAAATACCCCTATTGTTGATCTGTTAATTACCGC
TGTCTCTACTCTACCCATCGACTTAGCTGCGGGTATTACAAATACTTAAACCGA
CCGAAACCTAACACACTCTTGCACCTGGGGGGAGAGACCCGTCTTACCA
GCACATTCTGATTCTCGCCACCCAGAAGTCTATCTTATTCTCCAGGCTTCGG
AATGATTCTCATCATCGCATTACTCGGGAAAAAGAACCATCGGTTACATAGG
CATGGTCTGAGAATAATGCCAATCGGCCACTAGGATTATTGTCGAGCACATCACAT
GTTACAGTGAATGGACGTTGACACCCGAGCATACTTACATCGCCACAATAATTAT
TGCATCCCCACAGGTAAAAGTGTAGCTGACTGACCAACTGCATGGGGAAAAT
TACCTGAGACACCCACTCCGTGACACTAGGTTATTCTCTTACTGTTGGAGG
ACTAACCGGCAATTCTATCATGGTAAACTTCTCTAGATATTCTCACGACACATATT
TGTAGTAGCACATTCCACTACGTTCTCCATGGAGCAGTCTTCAATTAGCAGG
CTTCGTACACTGATTCCCCACTTTCAAGGCCGTAACCACCTGCCACTTCTTAA
GACACTAGTAAAAGACATTACATCTCTGTCAAGGAAAATTGGGGTTGAATCCCCG
CGTGCCTAACCCACATGGCACATCCCACACAACCAGGATTCAAGATGCACTTCCCC
CTAATGGAGGAACCTACACTTACGACCCACACCTAAAGTAAATTCTTATGATCAGC
ACATTGGCTTACATTAATCACCATAGCCATAGCTAAACAGACATACTCGT
TTAGATTCTCAAGAAATTGAGATTCTGAACATTCTGCTGCAATCGTACTTACTT
ATTGCACTCCCCACTCCGAATCCTACTTAATAGACAGAAGTCAGTGTACCCACTTA
ACAATCAAAGCAACTGGGACCAATGATACTGAAGCTACGAATACTCAGACTATGAGAAC
CTAGAATTGACTCTACATGATCCCAACACAAGACCTACTCCAGGCCATTCCGGCTT
CTGGAAGCCGACCACCGAATGGTAATTCCGTGGAATCTTATTGAGCCCTAGTCTCC
GCTGAAGCAGTCTCACTCTGAGCTGTCCCGCTTAAAGGATTAAAGTGTGCAATT
CCCGGACGGTGAACCAAACACTTACGGCAGGCCACCCAGGATTCTCGGTCAA
TGCTGAAATCTGAGGCAATCACAGCTTACCTGCTATTGAGCTATTGAGCTGTACCT
CTCCACACTTTGAAAAGTGAACCTCTTAATAATGAAGACGCTTCGTTGAGAAGCTAA
ATAGGTTCAAGCTTAGCCTTTAAGCTAAAGAATGGTACTGCCGCCACCTCAGCGA
TATGCCTCAACTAAACCCAGCCCTGGTTCTTAATCTTAGTCTTCACTGACTAGTGT
CTTAACAATTCTCCACCTAAAATCCTCAACCACACTTCCCAATGAGCCAACCCAGC
CAGCTAAAAAGCCCTAACTAACTCTGAAACTGACCATGACTTAAAGCTTTAAT
CAATTGCACTCCCCGTCTACTAGGACTCCCTTAATTGCCCTAGCAATCTGCTACCC
TGAGCCCTTCTCCCTCCCTGCCCTACCGTTGACTTAATAACCGCATACTATCTCCAA
GGTGATTATCGAGGATTACTAAACATTAGTACTCTCCCTCAGCCCTGGGGCCAT
AAATGAGCCCTTATTCACTCACTAACTAGTCTACCTAAATTACAATTACCTGCTAGGC
CTACTCCCTACACCTCACCCCTACCACACAACCTCTGCTTAATATAGGCTAGCAATC
CCCTATGACTCGAACAGTATTATTGAAATGCGAAATGCCAACCCACTTTAGCC
CACTCCCTCCGAAGGCACCTCACCTTCTTATCCCTGTTAGTAATTATCGAAACA
ATTAGCCTATTCTCGCCCTGGCTTAGGAGTTCGACTAACCGCTAATCTAACAGCT
GGCACCTACTTACTGCAATAGCAACCAATGACTCTCCCTTATGCT
TCAGTAGCCATTCTCACTGGTACCCCTACTTTTATTAAACTCTCTAGAAATTGCCGTA
GCCATAATCCAGGCTACGTTCTCTTCTTAAAGCTCTACCTACAAGAAAACGTC
TAATGGCCCATCAAGCACACCCCTTCACATAGTCGATCCAAGCCCTGACCCCTACCG
GAGCCACAGGAGCTATGCTAACAGCTCAGGACTTGCAATATGATTCACTCCACTACA
CCTCCCTCTCTTCTGGACTCACCTCTAGTTTACAGTCATCCAATGATGACGAG
ATGTTATTGAGAAGGAACCTTCAAGGCCACACACTCCCCCGTACAAAATGCCCTCC
GATTGGGATAATCCTTTATTGCTCAGAAGTCCTTCTCGTTGAGTTCTGAG
CATCTACACCGCTACGCTGCCCAACTATAGACCTAGGAAATACTTGACCCCCAACAG
GAATCACCCCTGGACCCATTGAACTCCACTACTCAACACAGCGTTCTCTAGCAT
CCGGAGTAACAGTCACCTGAGCTCACCACAGCATTAGAAGGGAACCGGCCAACGTA
TTCAATCCCTAGCTCAACCACCTCTAGGCTTATTTCACTTCCCTCCAGGCCATAG
AATACTACGAAGCCCCCTCACTCTGGTGTGGAGTCTAGGAACACTTTCTTGTAG
CCACAGGATTCCACCGATTACACGTAATCATGGAAACACTTCCCTATTGATGCCCTCC
TTCGACAAACAGAACACCTTCACATGACAGCACCACCTTGGCTTGAAGCAGCTGCTT
GATACTGACACTTGTAGATGTTGCTGACTATTCTATATCTCTATTACTGATGAG
GATCTTAATCTTCTAGTACAAGTAGTACATGTGACTTCAATCCTGATCTGGTTAA
AATCCAAGGAAGAATAATAAACCTAATTACAACAGCCCTTTATTACCGCAGTACTATC
ACTTATCCTCATACTAGTCGCTTTGACTCCCCAACGACTACGAGAAACT
ATGCCATACGAATGGCTTGCACCTCTAGGATCTGCTGCCCTGCCCTTCCCTCG
ATTCTTCTAGGCCATCTTCTTCTTCTATTGACCTGAAATGCCCTCTCTCC
CCTCCCTGAGGAATCAATAACGACCCCTAGTAACCTTGCCTGAGCTATGGCTCT
TCTACTCTCTAACCTAGGCTCTGCTTACGAATGGCCAAGGGCTTGAATGAGC
CGAATAAGCTATTGTTAGGAAACATTGATTGCTTCAAAACTTGTGTTAA
GTCCCAATGGCTATGACCCCATCCATTGCTTCTCATCAGCTTATTCTTGG
TTAACAGGGCTAGCATCCATCGATCCCACCTGCTCTGCCCTATCTGCTAGAACGA
ATGATAGTCTCACTATTATGCCCTCCCTGTGGACAATACAACACTAAACCTCC
TTCTAAAACCTAGATTGATTCTAAAAACAGGTTAAACCCCTTGTGACCG
GAGAGGCTGACAGCAATGACAACGCTAATTTCATCTCCTGGTAGACCCAGGGCT
CACTCAAAGCTTAAAGGATAATGTTCATCCATTGGCTTACGGAAACAAAACCTTG
GTGCAAGTCCAAGTAGCAGCTATGCACTGCCCTCAATTATAACGACCTCTGCTT
ATTATCTGCCCTCTCACATTCCCCCTATAACTACCCCTGTACCCAAAACCCACGAC
CCAGAATGAGCCCTCTCCACGTTAAACAGCAGTAAAATGGCTTTTGTAAAGCCTT
CTACCACTCTTGTGCTCACCCTGGCTAGAAGTTTACTACATCTGATCCTGA
ATAAAATATTAACCTGTTGATATTAACATTAGCCTAAATTGACTTCACTCAATCATT

TTCACCCCAATGCCCTACGTTACGTACTATTCTAGAATTGCATCCTGATACATA
AGCTCTGACCCCCACATGAATCGGTTCTCAAGTACCTACTAATTTCATAATA
ATTACCCCTGTAACAGCAAATAATCTATTCAACTATTATGGCTGAGAAGGAGTTGGA
ATCATATCATTTCTAATTGGGTGTTGATACGCCGGAGCGGAATACCGCAGCC
CTTCAAGCAGTACTTATAACCGAGTCGGAGACATCGGACTAATCTTGCTATAGCATGA
ATAGCCACAAACCTTAACATGAGAACTCAACAAATCTCACAACCTCAAAAGATTT
GACTAACCTCCCGCCCTAGGACTAATCCTGCTGCAGCAGGAAAATCAGCACAAATC
GGACTTCACCCCTGACTGCCATCTGCTATGGAAGGCCAACACCGGTATCTGCCCTACTC
CATTCAAGCACAATGCTGAGCCGAATTCTTCTGCGCATGAGCCCTATCTTA
GAAAACAGTAAACAGCTTAACTACTGCGCTGCGTAGGAGCTATCACTACCCCTTC
ACGCCACCTGCGACTTACCCAAATGATACAAAAAAATTGAGCTTCTACTTCT
AGTCAACTAGGACTAATGATAGTCACTATTGGCCTTAACCAACCCAACTTGCCTCCTT
CACATTGACCCCACGCACTTCAAGGCCATACTTTCTTATGCTCTGGCTCAACTATC
CACAGCCTAAACAAACGAGCAAGACATTGAAAAATAGGCGGAATACACCACCTCACACCT
TTTACCTCTCATCTTACATCGCAGCCTAGCCCTAACAGGGACCCCTTCCTAGCC
GGCTTCTCTCAAAAGAGCCTAAATTGAGCAGCTAACACACCTAACCTAACGGCTGG
GCCCTTGCTTAACCTTGCTGCTACCGCCTACTGCAATTACAGCCTCCGAGTGATT
TTCTTGCTCTCAATAGGCTTCTCTGATTTAACACATTACCCAAATTATGAGCCCTCTA
CTCTGCTGTAATCAACCCAATAAACGACTTGCGATGAGGAAGCATTATGAGCCCTCTA
ATTACTCTAACCTCACTCTATAAAACACCTATCATAACAATACCACCCCTCTCAA
CTAGCTGCCCTAGCAGTCACATTCTAGGACTACTTACAGCCCTAGAACTTGCCCAACTC
ACCAACACCCAAAAAAAGCTATTCTGCAACTTCCCTCACCCTACTCTCAACATGCTT
GGCTTTTCCCCTCATTAAGTCCACCCTTATTCTCTAAGCTAAACTTAATTTAGGGCAA
ACAATTGCCCACAAATAGTGCACCAACCTGACTAGAGAAAAGCGGGACCAAAGCCGTT
TCTCTCTCAACATTCCCTAGCGCTCCGCTAGCAACATCCAACAAGGTTATAATTAAA
ACCTATTTCACCTATTCTACTAACCTTAACCCCTAGCAGTCCTGCCGTCTAATTAAA
TAGGAGCAAGAGGCCACGACTCTGCCCTCGAGTGAGCTCAACACAAATAAGTTA
GGAGAAGAGCTCCCCACTTAAATCAACATCCCCCACCCAAGAGTACATCAAAGCCA
CCCCCTCCCATATCCCCCGCAACACAGAGAAATCAACTAATCATCCATCACACCCACG
ACGACTCATACCAACCCCTCAGAAAAAGAACACCAACAACACCCAAACAAATAAA
CTACACCAGAAATTAGTACCAACGACTGCTTAAGCTTCGGGAAAGGGGTCCGAGCAA
GAGCGCTGATAAAGCAAACACTACTAGCATCCCCCAAGTAAATTAAAGAATAAGACTA
AAGATAGGAAGGAGCCACCGTGTCAAGATAAAATACCGCACCTATGCTGTAAGACTA
CTAACCCCAAAGCCCCAAAGAAGGGTGAGGGGTTGGAAGACAATAACGACAAACCTAAAA
TTATAGAAAACAAAATAATACATAATAAGTCCATAATTCTACCCAGGACTCTAACCA
GGACTAATGGCTTAAAAACCCACCGTTGTATTCAACTACAAGAACCTAATGGCCAACC
TCCGTAATACCTTCTCTTAAATCGCAACCGATGCTTACTGCTGATCTCCAGCCC
CCTCTAACATCTCTGCTGATGAAACTTTGGTCCCTTCTGAGACTCTGTTAGTGA
AAATTGCTACCGCTTATTCTTAGCCATGCACTATACATCTGATATTGCTACTGCTTCA
CCTCGTTGCACACATCTGCCGGACGTCACACTACGGTCACTTATCGGAACCATCATG
CCAACGGCGCATCTTCTTCTACCGCATCTACCTCATATCGGGCAGGTCTCTACT
ATGGCTCTTACCTTATAAGGAAACATGAACATCGGAGTTGTTACTGCTTCTCGTAA
TAATGACAGCCTCTGGCTACGTCTCCCTTGAGGACAATGTCAATTGAGGTGCAA
CCGTCACTAACCTTATCTGCCGTCCCTATGCGGAAGCACACTTGTCCAGTGG
TTTGAGGGCTTTCTGAGACAATGCAACCCCTCACCGGTTCTTGCTGATCTCCACTTCC
TCTCCCATTATCATCGGGCGCACAGGTGATCCACCTACTCTTCTTCAAGAAACTG
GTTCAAATAACCCACCGGGCTCAACTCAGACTCTGATAAGTCCCTTCAACCCACT
TCACATACAAAGACCTTCTAGGCTTCAGTCAGTCTCTACTGCACTGGCTTCACTGCC
TATTTCCCCAAATCTTCTAGGAGACCCAGACAACCTCACGCTGCAAACCCGCTGTCA
CGCCCCCACACATCAAGCCAGAGTGAATTCTCTTGTCTACGCCATTCTCGCTCCA
TCCCAAACAGCTGGCGCTACTTGCCTTCTTATTCTCTATCTCGTTCTCATGCTCG
TCCCTTCTCCACACCTCTAAACACAGAGCAGCTGATATTCTGCCCTGTAACACATTCC
TGTCTGGTCTCTAGTGTGAGCTAATATTCTGACTGAATTGGAGGAATGCCGTAG
AACACCCCTTCGTTATCATCGGACAAGTAGCATCTCATCTACTTCTCCCTTCTCTAG
TCCTGATCCCAACTGCAGGCTGAATGGAAAATAAGTCTTGGATGAAATGCACTAGTA
GCTCAGCGTCCAGGCCCCAGTCTGAAACTGGCGTGGAGGTTAAACCTCCAC
TGCTTCAAAGAAAGGAGATTCAACTCTACCCCTAACCTCAAAGCTAGGATTCTAGCA
TTAAACTATTCTTGTGCGTCAAACATGTACATGAAAGCTTCATGTACATGTATGTAAT
AACACCATATTATAGTAAACACTTATATAATGAAATAGGACATTCTCATGTATAATAA
CCTAATCTAGTAAATATAGCACTCaCGgCAATATTAACTAAGAAACactAGAACCTAA
CGAACGACTAACCTCACATTAGTATACTTCCAGGcCCAGTCGAAATTAAAGACCGAACAC
AACACTCATCAGTCGAGTTATCCAAGACTCAAATCCCTCAAGTCAAAAGCCTGAtGTA
GTAAGAgCcTACCAACCGGTGATTactTAATGATAACGGTATTGAAGGTGAGGGACAAA
AATTGTGGGGGTTTCatCCGTGATTATCctGcCaTTGGTCTCTGTTTATATGT
ATGAAAATAATACATATATGTATTACCATTAATCA

Genetic Markers in Atlantic Halibut: Tools for Breeding Success

Kenneth Mjelle, Tor E Jørgensen, Bård O. Karlsen, Steinar D. Johansen, Truls Moum

Bodø Regional University, Bodø, Norway



The Background

Animal mitochondrial DNA (mtDNA) is maternally inherited and supposedly lacks recombination. It is a useful genetic marker for intra-specific studies, due to its small size and conserved gene content, yet high nucleotide substitution rate.

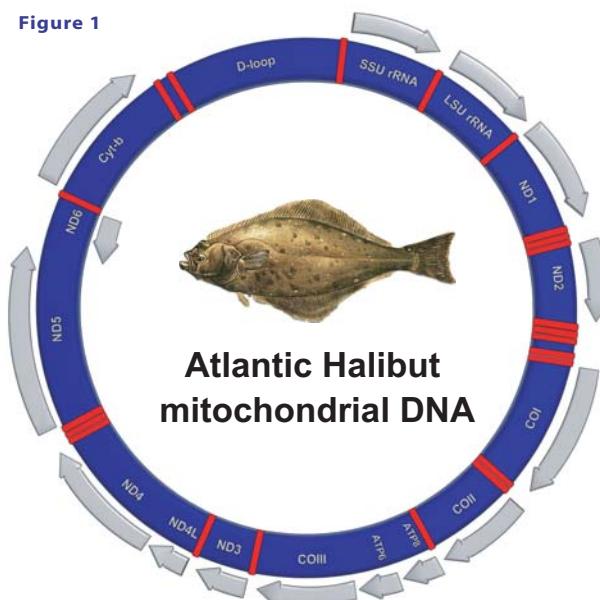
The aim of the project is to identify single nucleotide polymorphisms (SNPs) among Atlantic halibut (*Hippoglossus hippoglossus*) mitochondrial genomes. We screen for SNPs by determining the complete primary DNA sequence of the mitochondrial genome in four Norwegian halibut individuals.

The Results

Figure 1

The Genome

The mitochondrial genome of Atlantic halibut has a typical vertebrate gene content and organization. It contains the usual 13 protein coding genes, 22 tRNA genes, two rRNA genes, a non-coding control region, and short intergenic spacers (Fig. 1). The genome size is approximately 17400 bp and varies slightly due to a variable number of tandem repeats in the control region.



The SNPs

The sequence analyses revealed SNPs in all of the 13 protein coding genes except ATPase 6. Fig. 2 shows the distribution of 42 SNPs found among the four halibut individuals. The ND1 and ND2 genes seem to contain clusters of SNPs that may provide a target for further analyses. A larger sample of halibuts will be screened for SNPs along the complete mitochondrial genome.

Figure 2

	ND1	ND2	CO I	CO II	ATP6	CO III	ND3	ND4L	ND4	ND5	ND6	Cyt b
Position	3316	3668	3719	3848	4383	4462	4500	4506	4521	5629	5967	15047
Halibut 1	A	A	C	A	A	G	A	A	G	C	A	G
Halibut 2	G	A	C	A	A	A	G	A	C	G	A	G
Halibut 3	A	A	C	A	A	G	A	A	G	C	G	A
Halibut 4	G	G	T	G	G	G	A	G	G	G	G	G

The Conclusion

The mitogenomics approach is used to investigate the molecular evolution of the mitochondrial genomes and develop genetic markers to be applied in marine ecology, i.e. to investigate population relatedness, introgression events and dispersal histories. Also, assessments of inter- and intraspecific genetic variations of Atlantic halibut and related species are important for halibut breeding. Mitochondrial SNPs should prove useful as a fast screening and diagnostic tool in breeding programs.

Acknowledgement

We would like to thank Arild Eeg at Risør Akvarium for providing a halibut tissue sample.