UiT The Arctic University of Norway

Faculty of Biosciences, Fisheries, and Economics
Department of Arctic and Marine Biology

# Phylogeny of *Somniosus* sleeper sharks: insights from newly sequenced mitochondrial genomes

Jacob Max Christensen

Master's thesis in Biology BIO-3950, May 2022

# Table of Contents

# Acknowledgements

To my supervisors Dr. Kim Præbel and Dr. Arve Lynghammar. Your guidance has led me through two years of confusion. Thank you for always finding time to fit me into your busy schedules!

To Dr. Shripathi Bhat. You're a wizard. You can do things on a computer that I couldn't dream of in a million years.

To the entire Research Group for Genetics staff. Without your support in the lab, this project would not have been possible.

To Dr. Samuel Iglésias, Ms. Charlotte Chazeau, and the entire Museum National D'historie Naturelle. My sincere thanks for providing *Somniosus antarcticus* tissue samples. This study would have been incomplete without your support.

To Andrew Williston and the Museum of Comparative Zoology (Harvard). You provided tissue samples that can be found nowhere else in the world. My deepest thanks for providing the rare *Somniosus longus*.

# Foreword

The purpose of this study was to investigate the taxonomic and phylogenetic positions of sleeper sharks within *Somniosus*. This project has been supervised by professor Kim Præbel and associate professor Arve Lynghammar. The newly sequenced mitogenomes in this study will be added to NCBI GenBank and this study will be synthesized into a manuscript suitable for peer review.

This article is written in a modified APA format.

# 1 Abstract

Mitochondrial DNA (mtDNA) has been extensively used to explore phylogenetics since the its use became available. This science is considered useful in discerning relationships and evolutionary histories that are not possible with morphological studies alone. The sleeper sharks (genus: *Somniosus*) are a group of poorly understood, long-lived, iconic species whose taxonomic status within the genus have been argued about for many years. The aim of this study was to sequence and characterize mitogenomes of all *Somniosus* species to gain a better phylogenetic understanding of the genus and to discriminate the delimitations between species. Using next-generation sequencing, complete to nearly-complete mitogenomes were assembled of all species within *Somniosus*. Clear divergence was shown between the subgenera *S. Somniosus* and *S. Rhinoscymnus*, estimated at 7.8 million years ago, with no discernable differentiation between *Somniosus pacificus* and *Somniosus antarcticus*, nor between *Somniosus rostratus* and *Somniosus longus.* The results suggest a panmictic population of *Somniosus pacificus* and *Somniosus antarcticus*, as well as between *Somniosus rostratus* and *Somniosus longus*. Combined, the results of this study show that a thorough revision is needed of the *Somniosus* genus to improve taxonomic resolution and allow more informed conservation related decisions to be made.

# 2 Introduction

Understanding the evolutionary history and relationships among groups of organisms is important in the field of conservation to secure biodiversity and to ensure cost-effective management. These understandings are also important to biologists to create a more detailed picture and understanding of the connectivity of life on Earth. The development of using phylogenetics in conservation biology has shifted from a species or ecosystem approach to a biodiversity approach of conserving the "tree of life" (Mace et al., 2003). Phylogenetic trees, therefor, contain important information that can be used to develop conservation strategies. The sleeper sharks of the genus *Somniosus* are a group of poorly understood deep-water sharks, primarily known to the public by the Greenland shark: the longest living vertebrate in the world [392 ± 120 years] (Nielsen et al., 2016).

## 2.1 Mitogenomic phylogenetics

Before the advent of molecular tools, phylogenetics relied on the work of taxonomists using phenotypic variation to classify organisms (Lee & Palci, 2015). With the addition of molecular phylogenetics to the evolutionary toolkit, organisms can now be classified from distinct and discrete molecular sequences (Brown, 2002). Molecular techniques can be used to differentiate between phenotypically similar species and even uncover cryptic species hidden to the eye (e.g., Jadin et al., 2012).

The mitogenome is a popular target for phylogenetic analyses due to its small size, availability, maternal inheritance, high mutation rate, and lack of recombination (Duchêne et al., 2011; Gupta et al., 2015; Tannman, 1999). Historically, phylogenetic studies utilizing the mitochondrion have focused on specific portions of the genome: e.g., cytochrome oxidase I (COX-1), cytochrome *b* (Cyt *b*), and the non-coding control region (CR) (e.g., Choi et al., 2014; Duchêne et al., 2011; Farias et al., 2001; Kitamura et al., 1996). Owing to lower costs in genetic sequencing, the full mitogenome has become more popular for inferring phylogenies and may provide to be more informative than short sequences (Havird & Santos, 2014).

The typical vertebrate mitochondrial genome (mitogenome) is a circular molecule including 13 protein-coding genes (PCs), two ribosomal RNA genes (rRNAs), and 22 transfer RNA genes (tRNAs) (Pereira, 2000). While short-length mtDNA phylogenetic analyses are common, tree topologies may differ depending on the region used due to variations in mutation rates between genetic regions (Haney et al., 2010). Deep coalescence occurs when a gene

diverges before a species (Zhu et al., 2016) and can create what is known in phylogenetics as "phylogenetic noise": a case where deep coalescence is followed by long branches (Townsend et al., 2012). In other words, a rapidly mutating genetic region has more "pull" on phylogenetic analyses than more slowly mutating locations. This noise can be minimized by using larger proportions of DNA, including the full mitogenome (Duchêne et al., 2011; Meiklejohn et al., 2014; Rohland et al., 2007).

Advances in next-generation sequencing (NGS) technologies, such as shotgun sequencing on a high throughput platform, have made it possible to recover sequences from highly fragmented DNA (Marshall et al., 2017). High throughput sequencing technologies have advanced to the point where the error rate is below 1% (Dohm et el., 2008), and can provide an accurate, timely, and cost-effective method for sequencing challenging genomes.

## 2.2 The *Somniosus* genus

*Somniosus* contains five species (Table 1) and is sometimes divided into two subgenera: *Somniosus Somniosus* and *Somniosus Rhinoscymnus* (Yano et al., 2004), however the subgenera are not commonly used (e.g., de Loyola Fernández et al., 2017). The five species are defined primarily by morphological characters; mainly size, dentition, and vertebral morphology (e.g., Hsu et al., 2020; Parin & Kotlyar, 2007). The distinction between the subgenera is based primarily on size. *Somniosus Somniosus* species grow to a much larger size [7.56 m, *S. microcephalus*] (MacNeil, et al., 2012) than *Somniosus Rhinoscymnus* species [1.43 m, *S. rostratus*] (Carpenter, n.d.). *Somniosus Somniosus* contains the Greenland shark, *S.* (*S.*) *microcephalus*, the Pacific sleeper shark, *S.* (*S.*) *pacificus*, and the southern sleeper shark, *S.* (*S.*) *antarcticus*. *Somniosus Rhinoscymnus* contains the little sleeper shark, *S.* (*R.*) *rostratus*, and the frog shark, *S.* (*R.*) *longus* (Yano et al., 2004).

Known from holotypes, two additional members of *Somniosus* may exist. An undescribed species known as the longnose sleeper shark was reported in Portugal (Hamlett, 2011), but the holotype was destroyed in a fire and no additional material has been collected (Yano et al., 2004). An additional species named *Somniosus cheni* was described in 2020 (Hsu et al., 2020) and is beginning to be accepted in the literature (e.g., Pollerspöck & Straube, 2021).

**Table 1**

*Taxonomy of genus Somniosus*

| **Order:** | **Squaliformes** | |
|---|---|---|
| **Family:** | Somniosidae | |
| **Genus:** | *Somniosus* | |
| **Subgenus:** | *S. Somniosus* | *S. Rhinoscymnus* |
| **species:** | ***S. (S.) microcephalus*** (Bloch & Scheider, 1801)<br>Greenland shark | ***S. (R.) rostratus*** (Risso, 1827)<br>Little sleeper shark |
| | ***S. (S.) pacificus*** (Bigelow & Schroeder, 1944)<br>Pacific sleeper shark | ***S. (R.) longus*** (Tanaka, 1912)<br>Frog shark |
| | ***S. (S.) antarcticus*** (Whitley, 1939)<br>Southern sleeper shark | |

The main confusion within *Somniosus* revolves around spatial distribution and identification. Yano et al., 2004 provides a detailed report on the confusion around the species delimitations within *Somniosus*. In short, *S. microcephalus* and *S. pacificus* have generally been considered separate species within the subgenus *S. Somniosus*. *Somniosus antarcticus*, however, has been periodically synonymized with both *S. microcephalus* (e.g., Bass et al., 1976) and *S. pacificus* (see Last & Stevens, 1994). Within the subgenus *S. Rhinoscymnus*, *S. longus* has been periodically synonymized as *S. rostratus* (e.g., Francis et al., 1988). Some early publications identified *S. rostratus* and *S. longus* as belonging to their own genus separate from *Somniosus*: *Rhinoscymnus*, *Brevisomniosus*, *Heteroscymnus*, or *Scymnus* (e.g., Higashi et al., 1955), with the occasional publication reinstating this understanding (Welton & Goedert, 2016). The current understanding is that all five species are distinct within *Somniosus* (Weigmann, 2016).

Studies using mtDNA have shown small divergence between *S. microcephalus* and *S. pacificus* (Cyt *b*, ≈ 1.8 % sequence divergence, Murray et al., 2008) and haplotype sharing between the two (Hussey et al., 2015). Murray et al., 2008 also showed haplotype sharing between *S. pacificus* and *S. antarcticus* utilizing the same gene, Cyt *b*, with no structural genetic differences between populations. To date, there is no published research regarding the molecular phylogenetics of the subgenus *S. Rhinoscymnus*. The current *Somniosus*

phylogenetic tree has many uncertainties, and further genetic investigation will help alleviate these issues.

## 2.3 Range and distribution

*Somniosus* species have a circumglobal distribution and all species are thought to share similar habitats in their respective regions – continental and insular shelves, slopes, and seamounts (Hamlett, 2011). With few exceptions, they have not been targeted in the fishing industry and most specimens are obtained as bycatch from long-line or trawling operations (Chazeau et al., 2019; Ebert 2015; List & Stevens 2019; Yano et al., 2004). The large *Somniosus* species seemingly exhibit a form of Rapoport's Rule (Stevens, 1996), in which their bathymetric range at polar regions is throughout the water column including the photic zone, but as latitude decreases, they are almost exclusively found at mesopelagic and bathypelagic depths (Benfield et al., 2008). All *Somniosus* species have an affinity for deep, cold water, with recurring instances of individuals found long outside their typical range (Ebert et al., 2013).

### 2.3.1 *Somniosus microcephalus*

The IUCN reports *S. microcephalus* as having a range restricted to the northern Atlantic and Arctic oceans (Kulka et al., 2020). It is found in the seas surrounding Greenland, Iceland, and Svalbard, along the coasts of Norway, the Barents Sea, and further south in the North Sea around Ireland (Ebert et al., 2013; MacNeil et al., 2012). However, this species has been documented in the Gulf of Mexico on multiple occasions, even being recorded in the southern Columbian Caribbean (Acero et al., 2018; Benfield et al., 2008; Benz et al., 2007). *Somniosus microcephalus* has a known bathymetric range of 0 m – 2,992 m (Benz et al., 2007; MacNeil et al., 2012; Mecklenburg et al., 2018; Stokesbury et al., 2005).

### 2.3.2 *Somniosus pacificus*

*Somniosus pacificus* is currently reported to have a distribution in the north Pacific (Froese & Pauly, 2017). This range extends east from the coast of Baja California to the Sea of Japan (Compagno, 1984; Kang et al., 2015). There is evidence of this species' range reaching into the Arctic, due to a dead 229 cm long individual found washed up on shore at Point Hope, Alaska in 1998 (Benz et al., 2004). Walter et al., 2017 reports an *S. pacificus* individual (identified by genetic signature) caught on the Mid-Atlantic Ridge near the Azores archipelago, providing evidence of the capability of *S. pacificus* to inhabit the Atlantic. This species has

been observed at the surface to depths of 2,205 m, making its known bathymetric range 0 – 2,205 m (Yeh & Drazen, 2009).

### 2.3.3 *Somniosus antarcticus*

The current understanding of *S. antarcticus* as a distinct species places its range in the South Atlantic, Indo-Pacific, and Southern Oceans (Last & Stevens, 2009; Ebert et al., 2013). *Somniosus antarcticus* has been reported from latitudes as high as 35°42'S in Namibian waters to the Antarctic, the sea mounts of Tasmania, Australia, and possibly Chilean and Patagonian waters (Yano et al., 2004). *Somniosus antarcticus* inhabits continental and insular shelves, with a current records indicating a bathymetric range of 245 – 1,836 m (Chazeau et al., 2019; Ebert, 2015).

### 2.3.4 *Somniosus rostratus*

*Somniosus rostratus* is known primarily from the Mediterranean Sea and adjoining Atlantic Ocean (Ebert et al., 2013). Two individuals caught off the northwest coast of Cuba expand its range across the Atlantic (as cited in Acero et al., 2018). Irmak & Özden, 2021 reports an *S. rostratus* individual caught at a depth of at least 2,500 m, and Capape et al., 2020 reports an individual caught in a trawl at 120 m, making the known bathymetric range of this species 120 m – 2,500 m.

### 2.3.5 *Somniosus longus*

*Somniosus longus* specimens have been collected around Japan, with one specimen most likely misidentified as *S. rostratus* collected off the coast of New Zealand (Francis et al., 1988; Yano et al., 2004). This species has also been reported (but doubtfully) in Chilean waters (Bustamante et al., 2014; Parin et al., 1997). The bathymetric range of *Somniosus longus* is only known from a handful of specimens caught on outer continental shelves and slopes at depths ranging from 250 – 1,160 m (Ebert et al., 2013; Yano et al., 2004).

Targeted assessments of *Somniosus* species are challenging. Acquisition of *Somniosus* specimens is opportunistic and usually comes by the way of fishery bycatch (Ebert et al., 2013). When caught, identification of individuals is difficult, with high levels of intraspecific morphological variability (MacNeil et al., 2012), and interspecific distinguishing characteristics often markedly overlapping (Benz et al., 2007). Ranges are ill-defined (see

Acero et al., 2018), and haplotype sharing has been documented between species (Hussey et al., 2015; Murray et al., 2008).

Range overlap, hybridization, and the propensity of individuals to be found far outside their typical range is problematic in a genus where species identification is difficult. Reliable taxonomic data is extremely important to conservation sciences and biology in general. Taxonomic misrepresentation can lead to gaps and errors in the understanding of ecological relationships (Vecchione et al., 2000). Unchecked, these taxonomic misrepresentations can cascade, causing future problems in biodiversity assessments and ecological management decisions (Bortolus, 2008). Large-scale fishing operations have the power to upset marine ecosystems, so reliable data concerning fishery management practices is essential. Because of its uniqueness and its tendency to be caught as collateral damage from deep-water fishing, the Greenland shark has acquired iconic status in conservation dialogue and is often seen in the media (e.g., Snider, 2022). While lacking the iconic status of the Greenland shark, the other *Somniosus* species operate in similar habitats and may perform similar roles in their respective ecosystems (Ebert et al., 2013). Clearing up the taxonomic uncertainty of *Somniosus* will allow for more informed conservation related decisions to be made with all *Somniosus* species.

## 2.4 Goals and hypotheses

The goals of this study were to (1) sequence and characterize mitogenomes of all *Somniosus* species, (2) infer phylogeny within *Somniosus* and evaluate the status of species within the genus, and (3) compare the use of circularized mitogenomes to short sequences in phylogenetic assessments. It is hypothesized that (1) the mitogenomes of newly sequenced *Somniosus* species will be annotated similarly to the already known *S. microcephalus* and *S. pacificus* mitogenomes, (2) due to known hybridization and overlapping ranges within the genus (e.g., Hussey et al., 2015; Walter et al., 2017), the current status of five distinct species will not be supported, and (3) the use of full mitogenomes will have higher taxonomic support than the use of multiple genes in phylogenetic analyses.

# 3 Materials and Methods

## 3.1 Sample collection

All *S. microcephalus*, *S. pacificus*, and *S. rostratus* samples were in the possession of Research Group for Genetics (RGG) at the University of Tromsø (UiT) prior to the start of this study. No additional samples of these species were used. *Somniosus antarcticus* and *S. longus* samples were acquired on request from outside sources (Table A1).

DNA extractions were carried out on 42 *Somniosus* samples in this study: 21 *S. microcephalus*, 10 *S. pacificus*, five *S. antarcticus*, four *S. rostratus*, and two *S. longus* (Table A2). *Somniosus longus* samples were used to create two extractions per sample because of their limited availability. Due to poor quality extractions, limitations in sequence data, or use in other projects, three *S. microcephalus*, three *S. pacificus*, three *S. antarcticus*, two *S. rostratus*, and one *S. longus* sample were used in the final phylogenetic analysis (Table 2). GS60 was provided by Dr. Kim Præbel in the form of a circularized mitogenome.

## 3.2 DNA extraction and sequencing

Extraction of DNA from all ethanol-fixed soft tissue samples was carried out using a DNeasy Blood & Tissue Kit (Qiagen, Germany) according to the manufacturer's instructions. Extraction of the RNAlater-fixed samples was performed using a DNeasy Blood & Tissue kit (Qiagen, Germany) according to the manufacturer's instructions, with an extra AW2 wash to ensure all salts were removed prior to elution. Formalin-fixed samples were extracted using QIAamp DNA FFPE Tissue Kit (Qiagen, Germany) according to the manufacturer's instructions. To extract DNA from the *S. longus* vertebrae, a method of ancient DNA extraction was performed using the method outlined in Dabney et al., 2013 with modifications. Briefly, sandpaper was used to remove the outer surface of the vertebrae before a pre-digestion step (see Damgaard et al., 2015). Two 50 mg samples of vertebrae powder were extracted using a bastard file and placed into individual tubes. The samples were then incubated overnight in extraction buffer consisting of proteinase K and EDTA (0.5 M, pH 8.0) at 37 °C. Lysates of each sample were bound with 10 mL of PB buffer and centrifuged through a MinElute column (Qiagen, Germany). Lastly, the samples were washed twice with 720 µL PE buffer and eluted with 20 µL of EB buffer.

Following extractions, DNA quantity was assessed on an Invitrogen Qubit 4 fluorometer using a dsDNA HS Assay Kit (Thermo Fisher Scientific, USA) per the manufacturer's instructions, and quality was assessed using agarose gel electrophoresis (1 %, 80 V, and 1 hour). Using this data, the best quality extractions were chosen to have libraries built for sequencing, with the intent of three samples per species.

Libraries were built using an NEBNext Ultra II FS DNA Library Prep Kit for Illumima (New England Biolabs, USA) per the manufacturer's instructions. The libraries were then assessed on an a 2100 Bioanalyzer System using an Agilent high sensitivity DNA kit with 2100 Expert software (Agilent, USA) to measure sequence length. The obtained sequences averaged 460 bp in length (minus the *S. longus* samples, GS516A = 225 bp, GS516B = 220 bp).

Next, quantitative PCR (qPCR) of the libraries was followed using an NEBNext Library Quant Kit for Illumima (New England Biolabs, USA). The libraries were diluted to 2 nM and pooled in equimolar proportions. Due to the short sequence length of *Somniosus longus* libraries, all *S. longus* libraries were combined in a separate pool to be sequenced as a degraded sample.

Sequencing was performed by Novogene (UK) Company Ltd. on an Illumina NovaSeq 6000 (New England Biolabs, USA) on two lanes of an S4 chip with 150 paired-end chemsitry.

## 3.3 Raw data processing

Raw data processing and mitogenome assembly were performed by Dr. Shripathi Bhat. The demultiplexed raw sequence data in fstqc format received from Novogene were quality checked for base quality and the presence of adapters using *FastQC v0.11.9* (Andrews, 2010). The *FastQC* outputs were collated in *MultiQC v1.12* (Ewels et al., 2016) to facilitate results for all samples in one single output. *Cutadapt* (Martin, 2011) was used to trim adapters and poor-quality reads. *Cutadapt* trims adapters and bases when the base quality score is 20 or less from paired reads and discards reads when ≥ 5 % of the bases are not called (i.e., Ns). The trimmed fastq files were then re-analyzed in *FastQC* to confirm adapter removal and quality trimming.

**Table 2**

*Overview of samples sequenced for this project*

| ID | Species | Location | Tissue type | Fixation method | Use in study |
|---|---|---|---|---|---|
| **GS60** | *S. microcephalus* | Clavering Island, Greenland | Muscle/cartilage | Ethanol | Yes |
| **GS80** | *S. microcephalus* | Nuuk, Greenland | Muscle | Ethanol | Yes |
| **GS159** | *S. microcephalus* | Andørya, Norway | Muscle | Ethanol | Yes |
| **GS216** | *S. rostratus* | Mediterranean Sea, Italy | Muscle | Ethanol | Yes |
| **GS218** | *S. rostratus* | Mediterranean Sea, Italy | Muscle | Ethanol | Yes |
| **GS219** | *S. rostratus* | Mediterranean Sea, Italy | Muscle | Ethanol | No |
| **GS500** | *S. pacificus* | Pacific | Skin/muscle/cartilage | RNAlater | Yes |
| **GS502** | *S. pacificus* | Pacific | Skin/muscle/cartilage | RNAlater | Yes |
| **GS505** | *S. pacificus* | Pacific | Skin/muscle/cartilage | RNAlater | Yes |
| **GS510** | *S. antarcticus* | Chatham Islands | Skin/muscle | Ethanol | Yes |
| **GS513** | *S. antarcticus* | Crozet Islands | Skin/muscle | Ethanol | Yes |
| **GS514** | *S. antarcticus* | Crozet Islands | Skin/muscle | Ethanol | Yes |
| **GS516A** | *S. longus* | No data | Vertebrae | Dry | Yes |
| **GS516B** | *S. longus* | No data | Vertebrae | Dry | No |

*Note.* GS516 had two libraries made due to short lengths of DNA extracts: one fragmented (A) and one unfragmented (B). GS219 and GS516B were not used due to poor quality Novogene output. Accurate location data was not available for all samples.

## 3.4 Mitogenome assembly and preparation of assembled mitochondrial sequences

Genomes were assembled using *NOVOPlasty v4.3.1* (Dierckxsens et al., 2017). *NOVOPlasty* is a seed extend based assembler used to assemble short reads into circular organelle genomes. The assembly by *NOVOPlasty* is initiated by a seed which is iteratively extended from both ends. A seed is a sequence of a mitochondrial gene or whole mitogenome from a same species or closely related species. *NOVOPlasty* will look for similarities between the short reads and seed to retrieve one sequence read of the targeted genome. That read is then used to find overlapping reads, further extending the read into a circular genome or the largest

contigs. The seed was prepared from five sequences available on the National Center for Biotechnology Information (NCBI) GenBank (Sayers et al., 2021) (Table A3). *NOVOPlasty* was then run with the prepared seed and produced one of two results: circularized mitogenomes or mitochondrial contigs.

To maintain standard start and stop regions for alignment purposes, a *Somniosus microcephalus* mitogenome, NC_049864.1, was downloaded from NCBI GenBank and used as a reference mitogenome. The complete and nearly complete mitogenomes were adjusted for read orientation and aligned using the *MUSCLE* algorithm (Kumar et al., 2018) implemented in *MEGA11* to the reference mitogenome using default parameters. Long stretches of repeats, which were most likely artifacts of the assembly process, were found in GS216, GS218, and GS516. These repeats could not be verified, so they were removed in order to create comparable alignments. The assembled sequences were annotated using *MitoAnnotator* (Iwasaki et al., 2013).

After alignment, the reference mitogenome was removed, the complete mitogenome of *Squaliolus aliae*, KU873080.1, was added as the outgroup, and sequences were re-aligned. To provide a comparison to using the complete to nearly-complete mitogenomes, COX-1, Cyt *b*, and CR regions were extracted from all individuals' mitogenomes and aligned individually for further analyses. COX-1 and Cyt *b* were chosen due to their previous use in shark phylogenetics (Vélez-Zuazo & Agnarsson, 2011; Straube et al., 2015; Murray et al., 2008). CR was chosen because it has a high mutation rate and is highly variable compared to other regions of the mitogenome (Zhou et al., 2014). A concatenated alignment with of the COX-1, Cyt *b*, and CR sequences was prepared in *MEGA* for use in haplotype analysis and measuring evolutionary divergence. For consistency, the analyses involving alignments of the complete to nearly-complete mitogenomes will be referred to as "mitogenome" analyses, and those involving COX-1, Cyt *b*, and CR sequences will be referred to as "short-sequence" analyses henceforth. Short-sequence and mitogenome alignments of the subgenera were also created in *MEGA11* for site annotation.

## 3.5 Estimates of Evolutionary Divergence between Sequences

The proportion (*p*) of nucleotide sites at which two sequences being compared are different (p-distances) were calculated using *MEGA* (Tamura et al., 2021) between all samples. The rate variation among sites was modeled with a gamma distribution (shape parameter = 4). Codon positions included were 1st+2nd+3rd+Noncoding. All ambiguous positions were

removed for each sequence pair (pairwise deletion option). Values were averaged for interspecific and intraspecific comparisons in *Microsoft Excel* (Guerrero, 2018) using the AVERAGE and STDEVP functions.

The annotated sequence files were uploaded to *MEGA* (Tamura et al., 2021) and polymorphisms were visually verified. For use in comparison between the whole genus and subgenera, conserved, variable, and parsimony-informative sites were annotated in *MEGA* for both short-sequence and complete to nearly-complete mitogenome alignments.

## 3.6 Phylogenetic analyses

### 3.6.1 Bayesian Inference tree modeling

To model phylogenetic relationships, Bayesian Inference (BI) analyses were used in this study. Four data sets were created for tree modeling: one alignment of the complete to nearly-complete mitogenomes, one alignment of COX-1, one alignment of Cyt *b*, and one alignment of CR. The best partition scheme and nucleotide substitution model for each alignment was determined using *jModelTest2.1.10* (Darriba et al., 2012). Akaike Information Criteria (AIC), Bayesian Inference Criteria (BIC), and Decision Tree (DT) scores were calculated, using the model with highest support from these criteria (Table A13).

Two different BI analyses were run on *BEAST v2.6.7* (Bouckaert et al., 2014) for comparison: mitogenome analysis using the complete and nearly-complete mitogenomes and short-sequence analysis using COX-1, Cyt *b*, and CR. (*BEAST* is a program used for Bayesian phylogenetic analyses of molecular sequences. *BEAST* has a selection of multiple site and clock models and uses Markov chain Monte Carlo to average over tree space (Bouckaert et al., 2014).) For *BEAST* workflow, see the appendix (Figure A3). BI analyses were conducted under the following conditions: nucleotide substitution models determined by *jModelTest*, relaxed uncorrelated lognormal clocks (Drummond et al., 2006) with initial rates of $7.0 \times 10^{-6}$ substitutions/site/year (Dudgeon et al., 2012) to allow for independent evolutionary rates among lineages, random starting trees, and birth – death speciation models with lineage splitting and fossil date calibration. Birth – death was chosen as the speciation model because fossil dates were included and incomplete sampling of the genus is possible, as it assumes extinction may occur at any time as well as speciation. The Markov chain Monte Carlo was run for 100 million generations with a 20 % pre-burn-in, producing 100,001 trees. *Tracer v1.7.2* (Rambaut et al., 2018) was used to verify the data, and all parameters had effective sample size

(ESS) values over 200. (ESS values over 200 indicate high-quality posterior distribution.) Maximum clade credibility trees were generated and annotated in *TreeAnnotator v2.6.7* with a 10 % burn-in. *FigTree v1.4.4* (http://tree.bio.ed.ac.uk/software/FigTree/) was used to create visual models of the phylogenetic trees.

### 3.6.2 Age Calibration

The age of the *Dalatiidae – Etmopteridae* & *Somniosidae* split, ≈ 131.91 million years ago (Ma) (Flammensbeck et al., 2018), was used to calibrate the node age between the ingroup and outgroup. Two fossils were also taken into consideration when calibrating age within the ingroup *Somniosus*: **Rhinoscymnus viridiadamas* and *Somniosus gonzalezi*. Both fossils are members of *Somniosus* and date to the Oligocene Epoch (Welton & Goedert, 2016). *Rhinoscymnus viridiadamas* has an estimated age of 28 Ma and clearly shows to be a member of the *S. Rhinoscymnus* subgenus, while *S. gonzalezi* has an age of estimated age of 30 Ma and clearly shows to be a member of the *S. Somniosus* subgenus (Welton & Goedert, 2016). From this information, an exponentially distributed prior of 30 Ma was chosen for the ingroup, with a soft upper bound of 33.9 Ma, the beginning of the Oligocene (Cohen et al., 2013).

The formation of the Isthmus of Panama is commonly used as a calibration date for divergence between Atlantic and Pacific taxa (e.g., Barber & Bellwood, 2005). In this case it could have been used to calibrate the time of divergence between *S. microcephalus* and *S. pacificus*/*S. antarcticus*. This option, however, was not chosen due to hybridization events that currently occur (Murray et al., 2008; Hussey et al., 2015), *Somniosus* individuals found far outside their supposed ranges (Walter et al., 2017; Benz et al., 2007; Benfield et al., 2008; Acero et al., 2018), and that the opening of the Bering Strait could have provided a route for gene flow after the formation of the Isthmus of Panama (Sher, 1999). Age calibration parameters can be viewed in Table A14.

*Note*. The authors describe the subgenera *S. Somniosus* and *S. Rhinoscymnus* as distinct genera: *Somniosus* and *Rhinoscymnus*.

## 3.7 Haplotype analysis

Two haplotype networks (mitogenome and concatenated short-sequence) were constructed with *PopART v1.7* (Leigh et al., 2015), using minimum spanning networks and default settings with the outgroup removed.

# 4 Results

## 4.1 Illumina output

Illumina sequencing resulted in at least 100 million reads for all samples aside from GS216 (4,977,436 reads). Average organelle coverage ranged from 5,980 (GS216) to 25,256 (GS505) (Table 3).

**Table 3**

*High-throughput shotgun sequencing genome coverage from NOVOPlasty4.3.1*

| Sample ID | Species | Total number of reads | Total reads aligned as mitochondrial genome (%) | Average organelle coverage |
|---|---|---|---|---|
| **GS80** | *S. microcephalus* | 126,484,506 | 2.51% | 28,707 |
| **GS159** | *S. microcephalus* | 192,047,318 | 0.34% | 5,960 |
| **GS216** | *S. rostratus* | 4,977,436 | 13.34% | 5,980 |
| **GS218** | *S. rostratus* | 521,541,560 | 0.44% | 19,419 |
| **GS500** | *S. pacificus* | 115,084,304 | 1.96% | 20,402 |
| **GS502** | *S. pacificus* | 118,637,112 | 1.98% | 21,192 |
| **GS505** | *S. pacificus* | 566,683,078 | 0.49% | 25,256 |
| **GS510** | *S. antarcticus* | 154,761,530 | 1.29% | 18,017 |
| **GS513** | *S. antarcticus* | 522,419,522 | 0.38% | 18,031 |
| **GS514** | *S. antarcticus* | 224,208,864 | 0.98% | 19,883 |
| **GS516** | *S. longus* | 207,420,224 | 0.83% | 15,067 |

*Note*. This table contains only samples used in the final analysis. GS60 was provided as a circularized mitogenome, and thus not included in the sequencing of this study.

## 4.2 Characterization of *Somniosus* mitogenomes

Shotgun sequencing yielded circularized mitogenomes eight for of the eleven samples and equally comparable contigs for the remaining two (Table 4). This thesis reports the first ever circular mitogenomes of *S. antarcticus* and draft mitogenomes of *S. rostratus* and *S. longus*. The genomes consisted of 16,730, 16,725 – 16,728, and 16,725 bps for *S. antarcticus*, *S. rostratus*, and *S. longus* respectively. All annotated mitogenomes can be viewed in the appendix (Tables A4 – A12).

All obtained mitogenomes consisted of the characteristic vertebrate 13 protein-coding (PC) genes, two rRNA genes, 22 tRNA genes, and a control region between $tRNA^{Pro}$ and $tRNA^{Phe}$, in the typical configuration. One protein-coding gene (ND6) and eight of the 22

tRNAs were coded on the light (L) strand. The two rRNA genes, the remaining 12 protein coding genes, and remaining 14 tRNAs were coded on the heavy (H) strand. GC content of all mitogenomes was ≈ 39 % (Table 4).

### 4.2.1 Sequence feature of protein coding genes

Two start codons (ATG and GTG) were detected in PC genes and were of identical use throughout the genus. ATG was the most common, accounting for 12 out of the 13 PC genes. GTG only appeared as the start codon in COX-1. Two stop codons (TAG and TAA) and two incomplete stop codons (TA- and T--) were detected in PC genes, with variation in usage between sequences. TAA was the most common stop codon, used in ≥ seven of the twelve PC genes within all sequences. TAG was used as the stop codon of ND1 in *S. microcephalus*, *S. rostratus*, and *S. longus*, while TAA was used in *S. pacificus* and *S. antarcticus.* TAA was used as the stop codon in ND3 in all individuals except GS500 and GS510, where TAG was used. Incomplete stop codon TA- was used in ND2 and ATP6 of all species, and incomplete stop codon T-- was used in COX-2 and ND4 of all species. Two deviations from typical *Somniosus*

**Table 4**

*Summarized genome composition of annotated sequences.*

| Sample ID | Species | Mitogenome length (bp) | Gene composition (PC, rRNA, tRNA) | GC content (%) | Circularized? (Y/N) |
|---|---|---|---|---|---|
| **GS60*** | *S. microcephalus* | 16730 | 13, 2, 22 | 39.25 | Y |
| **GS80** | *S. microcephalus* | 16730 | 13, 2, 22 | 39.34 | Y |
| **GS159** | *S. microcephalus* | 16730 | 13, 2, 22 | 39.32 | Y |
| **GS500** | *S. pacificus* | 16731 | 13, 2, 22 | 39.34 | Y |
| **GS502** | *S. pacificus* | 16728 | 13, 2, 22 | 39.32 | Y |
| **GS505** | *S. pacificus* | 16730 | 13, 2, 22 | 39.32 | Y |
| **GS510** | *S. antarcticus* | 16730 | 13, 2, 22 | 39.32 | Y |
| **GS513** | *S. antarcticus* | 16730 | 13, 2, 22 | 39.24 | Y |
| **GS514** | *S. antarcticus* | 16730 | 13, 2, 22 | 39.26 | Y |
| **GS216** | *S. rostratus* | 16727 | 13, 2, 22 | 39.19 | N |
| **GS218** | *S. rostratus* | 16725 | 13, 2, 22 | 39.21 | N |
| **GS516** | *S. longus* | 16725 | 13, 2, 22 | 39.21 | N |

*Note.* Protein-coding genes are abbreviated as PC. *The circularized mitogenome of GS60 was provided by Dr. Kim Præbel.

sequence length were detected in *S. pacificus* samples; GS500 had a single insertion in ND6 (523 bp vs. 522 bp), and GS502 had two deletions in ND5 (1,831 bp vs 1,833 bp).

### 4.2.2 Sequence feature of tRNAs, rRNAs, and non-coding regions

All tRNAs were equal in length (67 – 75 bp) for all samples in this study. Single nucleotide deletions were detected in both rRNAs of the *S. Rhinoscymnus* subgenus. 12s rRNA was 951 bp in length in the *S. Somniosus* subgenus, but 950 bp in length in the *S. Rhinoscymnus* subgenus, and 16s rRNA was 1,675 bp in length in *S. Somniosus*, but 1,674 bp in length in *S. Rhinoscymnus*. With the exception of GS500 and GS502, variation in mitogenome length not accounted for by rRNA sequence length was due to variation in CR length.

### 4.2.3 Intergenetic space

Spacer DNA and overlapping genes were detected in 16 identical locations in all genomes. Thirteen regions of spacer DNA were identified (e.g., a gap of three nucleotides between ND3 and $tRNA^{Arg}$), and overlapping genes were detected between ATP8 and ATP6 (10 nucleotides), between ND4L and ND4 (seven nucleotides), and between ND5 and ND6 (four nucleotides).

## 4.3 Short-sequence coverage

The mitogenome alignment with all individuals was 16,733 bp long, whereas the short-sequence alignment consisted of 2,703 bp (16.15 %). The short-sequence alignment covered 15.94 % of conserved sites, 20.90 % of variable sites, and 20.33 % of parsimony-informative sites within the mitogenome alignment. Similar coverage was seen between the subgenera *S. Somniosus* and *S. Rhinoscymnus*. Site variation between alignments and analyses can be viewed in the appendix (Table A15).

## 4.4 Estimates of genetic divergence using p-distances

The average p-distance was low throughout the genus in the mitogenome analysis, with the largest values between *S. Somniosus* and *S. Rhinoscymus* species. The smallest p-distance was 0.000958 between *S. rostratus* and *S. longus*, while the largest was between *S. longus* and

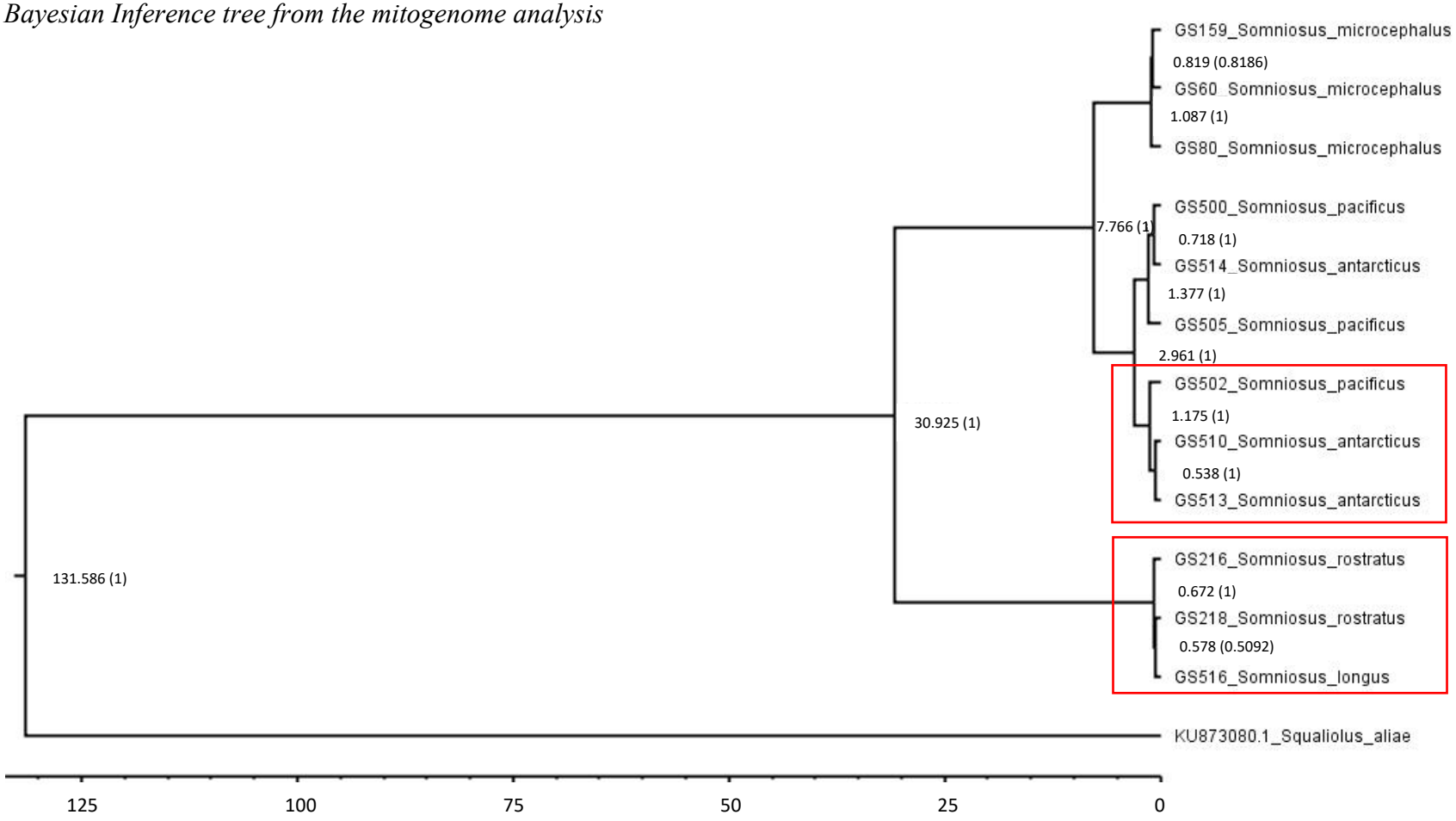
*S. pacificus*, at 0.035563. Within *S. Somniosus*, all interspecific p-distances were under 0.01. The same patterns were observed in the short-sequence analysis with > 60 % higher values. Intraspecific p-distances were under 0.005 for all species in both analyses. Notedly, the intraspecific p-distance in *S. pacificus* (0.003189, mitogenome; 0.004686, short-sequence) was higher than the interspecific p-distance between *S. pacificus* and *S. antarcticus* (0.002544, mitogenome; 0.003330, short-sequence). All p-distances can be viewed in the appendix (Tables A16 – A20).

## 4.5 Phylogenetic relationships

Tree topologies obtained by Bayesian Inference with both the mitogenome and short-sequence analyses revealed monophyletic subgenera *Somniosus* and *S. Rhinoscymnus* with distinct division between the two (Figures 1 & 2). Three distinct clades within *Somniosus* were revealed in both analyses: one comprised solely of *S. microcephalus*, one comprised of both *S. pacificus* and *S. antarcticus*, and one comprised of *S. rostratus* and *S. longus*. Within the *S. pacificus*/*S. antarcticus* clade there was no apparent delineation in grouping between the two species. Similarly, the *S. rostratus*/*S.longus* clade showed very little divergence between the two species. Posterior probability scores on the mitogenome tree were high, with all internal nodes between branches having a value of one, but slightly lower on the short-sequence tree on nodes between terminals (Figures 1 & 2). The mitogenome analysis suggested a *S. microcephalus – S. pacificus* & *S. antarcticus* split date of 7.8 Ma (Figure 1), while the short-sequence analysis suggested the same split of having occurred at 10.153 Ma (Figure 2). Additionally, the groupings of terminals within the *S. pacificus*/*S. antarcticus* and S. *Rhinoscymnus* clades varied slightly between the two analyses (Figures 1 & 2).

**Figure 1**

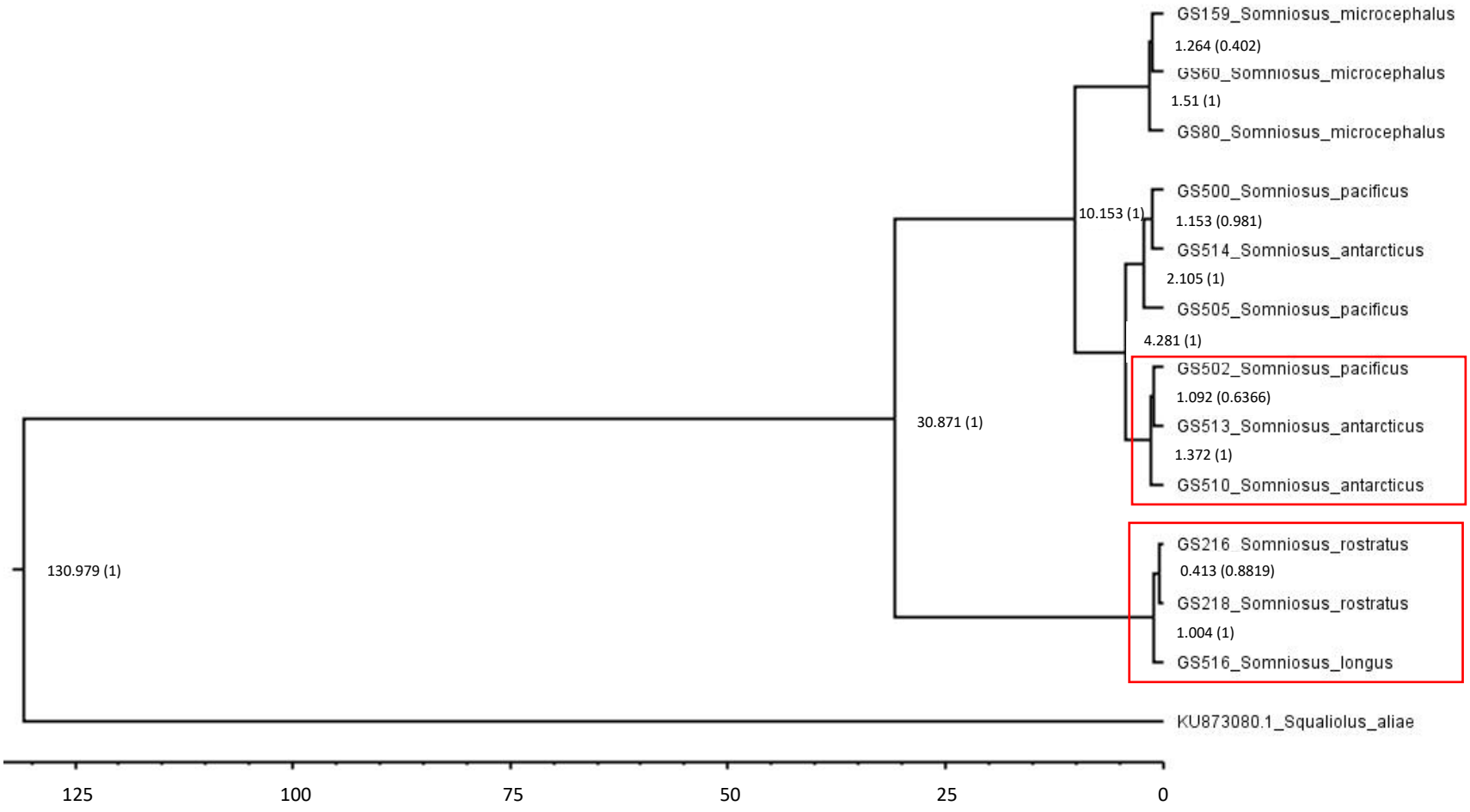
*Bayesian Inference tree from the mitogenome analysis*




*Note*. Node labels represent millions of years before present and posterior probabilities (PP). Red boxes are shown around clades with differing terminal arrangement from the short-sequence tree.

**Figure 2**

*Bayesian Inference tree from the* short-sequence *analysis*



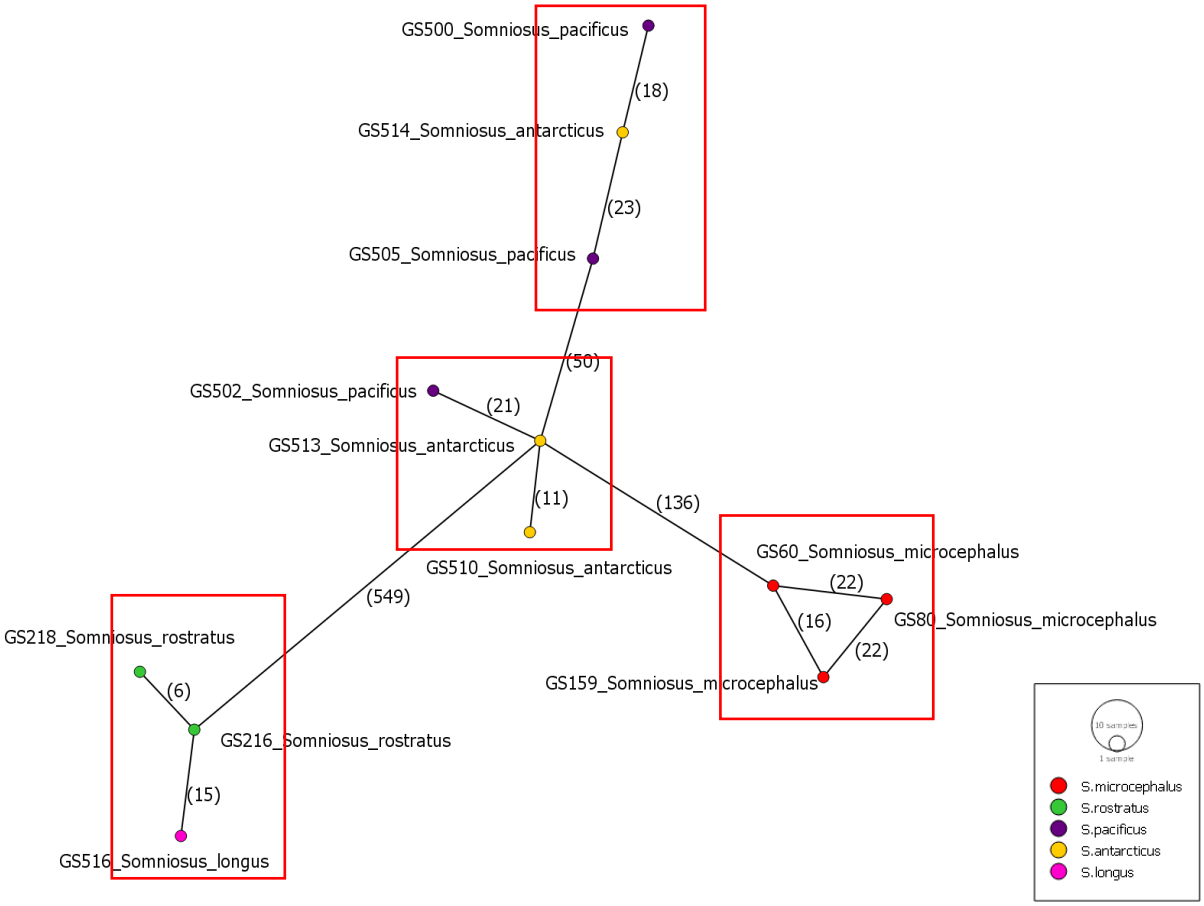

*Note*. Node labels represent millions of years before present and posterior probabilities (PP). Red boxes are shown around clades with differing terminal arrangement from the mitogenome tree.

## 4.6 Haplotype analyses

Haplotype analyses revealed four clusters: one *S. microcephalus* cluster, two *S. pacificus/S. antarcticus* clusters, and one *S. rostratus*/*S. longus* cluster (Figures 3 & 4) in both the mitogenome and short-sequence analyses. The resolution of the short-sequence analysis was less defined than the mitogenome analysis and produced a network with much less distinction between the clusters than the mitogenome analysis.

**Figure 3**

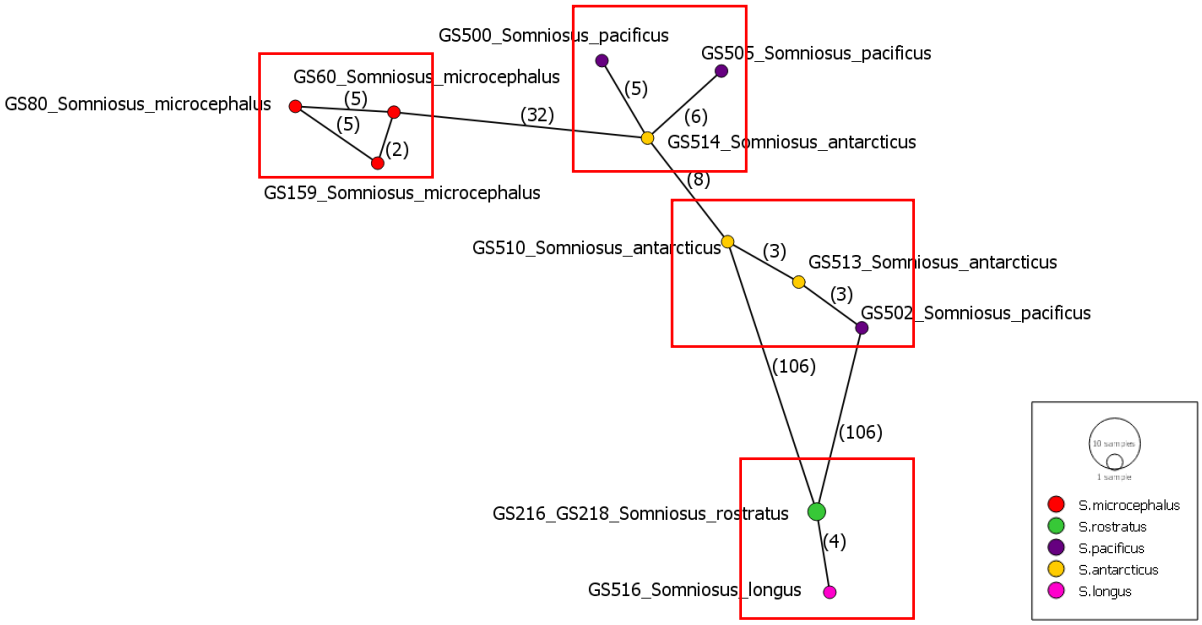
*Haplotype network of the mitogenome analysis*




*Note*. Numbers on branches represent the number of mutational changes between sequences. Clusters are shown in red boxes.

**Figure 4**

*Haplotype network of the* short-sequence *analysis*




*Note*. Numbers on branches represent the number of mutational changes between sequences. Clusters are shown in red boxes. A single point is shown for *S. rostratus* due to near-identical sequences within the three genetic regions.

# 5 Discussion

## 5.1 Genome annotation

The first circularized mitogenomes of *S. antarcticus* and draft *S. rostratus* and *S. longus* mitogenomes were presented in this study. All circularized mitogenomes fell within two bp in length of the reference *S. microcephalus* mitogenome (16,730 bp). The low GC content and incomplete stops codons found in all sequences are not uncommon in sharks (e.g., Doane et al., 2018; Yang et al., 2016; Chen et al., 2014), and were similar to what has been previously reported in *S. microcephalus* (Santaquiteria et al., 2017). The incomplete TA- and T-- stop codons were presumedly modified to TAA by post-transcriptional polyadenylation (Anderson et al., 1981).

## 5.2 Phylogenetic Analyses

The phylogenetic analyses conducted in this study were the first performed on all current *Somniosus* species. The positions of species within subgenera were consistent with Yano et al., 2004 placing *S. microcephalus*, *S. pacificus*, *and S. antarcticus* in *Somniosus Somniosus* and *S. rostratus* and *S. longus* in *Somniosus Rhinoscymnus*. Both mitogenome and short-sequence analyses produced tree topologies including a clade of *S. pacificus* and *S. antarcticus*, without any delimitation between the two species, asserting the findings of Murray et al., 2008 that showed haplotype sharing between the two, and suggesting a panmictic population. *Somniosus rostratus* and *S. longus* also formed a clade with very little divergence, suggesting the two species to be of a singular population.

The formation of the Isthmus of Panama is commonly used to date divergences between Atlantic and Pacific taxa (e.g., Barber & Bellwood, 2005), despite near constant debate over its true date of formation (see Coates et al., 2013). This study reports the split between *S. microcephalus* and *S. pacificus/S. antarcticus* (7.8 Ma, mitogenome analysis; 10.2 Ma, short-sequence analysis) millions of years prior to recent estimations of the formation of the Isthmus of Pamana (2.8 Ma, O'Dea et al., 2016), suggesting either an unrelated divergence or an inaccurate consensus of when the isthmus formed.

## 5.3 Mitogenome vs. short-sequence phylogenies

The short-sequence tree showed both lower posterior probabilities and deeper divergence than the mitogenome tree, correlating with increased p-distances in the short-sequence analysis (> 60 % increase from the mitogenome analysis, Table A20). The increased signal from the short-sequence analyses may be misleading and accounted for as noise. Building on Duchêne et al., 2011, in that adding more sequence data can improve dating and tree topology, the mitogenome tree topology likely most accurately reflects the true phylogeny of *Somniosus*. In addition, short genetic regions may vary in their usefulness for determining phylogeny depending on lineage and may possess phylogenetic signal that is not representative of other regions (Havird & Santos, 2014). Caution should be taken when basing hypotheses regarding the evolution of lineages solely on phylogenies derived from short sequences. Basing a phylogeny on a single gene or short sequence simply ignores the remainder of the genome and portrays relationships between genetic regions, not genomes.

## 5.4 Delimitation of *Somniosus* species and conservation

In 1963 Ernst Mayr described species as "groups of actually or potentially interbreeding natural populations which are reproductively isolated from other such groups", and the Biological Species Concept (BSC) was born (Mayr, 2013). Using the BSC within *Somniosus*, the picture is not entirely clear. Mate recognition is key to this concept, and it is known from previous studies that *S. Somniosus* species recognize each other and reproduce (Hussey et al., 2015; Murray et al., 2008).

Under the Phylogenetic Species Concept (PSC) (Eldredge & Cracraft, 1980; Cracraft, 1983), a species is defined as "the smallest diagnosable cluster of individual organisms within which there is a parental pattern of ancestry and descent'', or as Nelson & Platnick, 1981 phrased it, "simply the smallest detected samples of self-perpetuating organisms that have unique sets of characters." Applying the PSC in molecular phylogenetics, fixed sets of alleles are often used to distinguish between species (Nixon & Wheeler, 1990), and in theory, if discriminatory molecular markers are present, the subdivision of a species is plausible (Qin et al., 2017).

By applying the PSC to this study, the easiest distinction to be made in this study was between the two subgenera. Deletions were seen in the rRNA genes of *S. Rhinoscymnus* (Tables A4 – A12) and > 0.035 p-distances between the subgenera (Tables A18 – A19).

Another distinction to be made was between *S. microcephalus* and *S. pacificus*/*S. antarcticus*; different stop codons were used on ND1, and *S. microcephalus* showed a ≈ 0.009 p-distance from *S. pacificus* and *S. antarcticus*. Although relatively low, this was the highest p-distance between members of the same subgenera in this study.

Hebert et al., 2004 proposed a standard interspecific threshold of 10 times the intraspecific variation to delimit a species for use in DNA barcoding. Threshold values used to delimit the species, however, are arbitrary and vary between taxa (e.g., Qin, 2017). This present study showed how results based off short sequences can be misleading, and while the method in Hebert et al., 2004 only applied to the COX-1 gene, applying it to the mitogenomes in this study would result in only the delimitation between the subgenera. The p-distance measurements here showed no distinction between *S. pacificus* and *S. antarcticus* or *S. rotratus* and *S. longus*. *Somniosus pacificus* sequences even had higher intraspecific p-distances (both mitogenome and short-sequence analyses) than interspecific p-distances compared to *S. antarcticus* sequences (Tables A18 & A19). Furthermore, the p-distance of *S. rostratus* from *S. longus* was extremely low at 0.000958, and only 19 variable sites were found between the three alignments. This study brings to serious question the validity of *S. antarcticus* and *S. longus* as distinct from *S. pacificus* and *S. rostratus*.

Although useful, the PSC can be widely interpreted, and potentially lead to taxonomic inflation (Isaac et al., 2004). An example of this can be seen a subspecies of mountain zebra, *Equus zebra*: the Cape mountain zebra, *Equus zebra zebra*. Habitat fragmentation over the course of 300 years eliminated gene flow between three populations, which resulted in fixed alleles to be present in all three populations (Moodley & Harley, 2005). If applied in a narrow sense, the PSC would provide justification to assign species status to these fragmented populations. Another case in point is of the famous critically endangered northern white rhino. Northern and southern white rhinos are currently classified as subspecies of *Ceratotherium simum*, however, there was argument in the past to classify them separate as individual species (Groves et al., 2010), creating headaches for conservationists advocating for the introduction of southern white rhinos into the north to bolster the northern population. Harley et al., 2016 showed a 0.009 p-distance between the two subspecies using complete mitogenomes and concluded that the previous designations as subspecies should be kept. It is of note that the p-distance between the two subspecies of white rhino is equivalent to the p-distance between *S. microcephalus* and *S. pacificus*, greater than the p-distance between *S. microcephalus* and *S. antarcticus* (Table A18), and much higher than that between *S. rostratus* and *S longus*. Applying the standards of Harley et al., 2016 indicates that the *Somniosus* subgenera may be

comprised of subspecies or populations, as opposed to the current taxonomic rankings within the genus.

Subspecies have been used taxonomically to describe genetically distinct, geographically separated populations within the same species that freely interbreed where ranges overlap (Wilson & Brown, 1953). Being genetically distinct is the defining characteristic that has been used to differentiate subspecies from populations. Being genetically distinct also places subspecies at a higher level of importance in conservation than populations with biodiversity considered. The use of molecular data in the systematics of *Somniosus* brings long-standing uncertainties to light. The mitogenome analyses provided by this study suggests that *S. antarcticus* and *S. longus* may reflect neither subspecies nor species, but populations. Describing and recognizing distinct species from a few morphological characters may be useful in taxonomy, but it hides the evolutionary history of the group and can have downstream effects on conservation and management.

‘

# 6 Conclusions

The identification of species, subspecies, and populations is crucial to fisheries management and conservation. It is estimated that a quarter of the world's sharks are threatened with extinction (Dulvy et al., 2014). Sleeper sharks are poorly understood and certainly affected by anthropogenic pressures. No *Somniosus* species are currently targeted by the fishing industry, but their primary direct threat is incidental bycatch (Ebert 2015; Chazeau et al., 2019; List & Stevens 2019; Yano et al., 2004). Aside from the Northwest Atlantic Fisheries Organization's prohibition on directed fishing for *S. microcephalus* (NAFO, 2018), there are no international bodies that govern sleeper shark harvesting (Finucci et al., 2020; Finucci, 2018; Finucci et al., 2018; Rigby et al., 2021). Additionally, climate change is likely to influence the distribution and carrying capacity of the polar *Somniosus* species due to its effect on sea ice quantity, dynamics, and distribution (Edwards et al., 2019), as well as the potential to open up new fishing grounds due to decreased seasonal ice cover (Serreze & Meier, 2019). Clearing up the taxonomic uncertainty of *Somniosus* and recognizing the genetic diversity within the genus will allow more informed conservation and management related decisions to be made.

This study has attempted to clarify the delimitations of species within *Somniosus* and expound upon uncertainties within the genus. Complete or nearly complete mitogenomes of *S. antarcticus*, *S. rostratus*, and *S. longus* are now available for the first time. All newly sequenced species' mitogenomes were > 96 % similar to the *S. microcephalus* and *S. pacificus* mitogenomes, supporting the first hypothesis of this thesis.

There is clear evidence of a division between the subgenera *S. Somniosus* and *S. Rhinoscymnus*. There is also clear evidence of the splitting of *S. microcephalus* from *S. pacificus* and *S. antarcticus* that can support the current status of *S. microcephalus* as a distinct species, albeit having a low p-distance (0.09) from sister species. This study, however, found no evidence of genetic divergence between *S. antarcticus* and *S. pacificus*, nor was any evidence found to support the genetic discrimination *S. longus* and *S. rostratus*, which supports the second hypothesis of this thesis.

The discrepancies between full mitogenome and short-sequence analyses supported the third hypothesis of this thesis. Highlighted in this study are the possible shortcomings and misrepresentations of using short, highly variable sequences of DNA in phylogenetic analyses. The short-sequence analyses present in this study showed higher divergences between taxa, while accounting for a small fraction of mitogenome sequence lengths. Variable sites and p-distances were over-represented in all short-sequence analyses.

Although mitogenome analyses cannot per se be used to delimit species, the findings presented do not support the current status of five distinct species within *Somniosus*. Concurrence between mitochondrial, nuclear, and morphological assessments should be seen before reaching any conclusions about phylogenetic relationships. Further research involving nuclear DNA will be needed to investigate the complex relationships within the *Somniosus* genus. Complete mitogenomic analyses of sleeper sharks, as opposed to short-sequence analyses, may provide beneficial information for decoding nuanced genetic differences and allow for more robust, data driven efforts to conserve these elusive circumglobally distributed predators.

# 7 Future Perspectives

While NGS sequencing methods are accurate and cost effective, the only way to confirm results is by verifying with Sanger sequencing selected regions within the assembled mitogenomes. To confirm the phylogenetic relationships viewed in this study, sequence data should first be verified. While this study also used only Bayesian inference as a tool for estimating the phylogenetic relationships within *Somniosus*, other methods are available. Another method such as Maximum Likelihood could be used as a comparison to verify the BI generated results. This study was based solely on mtDNA sequences. To reinforce a mitochondria-based phylogeny, highly conserved nuclear DNA region (e.g., recombination-activating gene 1) could be added, with the added benefit of being able to identify hybrids among species. Similarly, added strength to the phylogenetic inference could be obtained by mapping the short-reads of the nuclear DNA regions to a *Somniosus* reference genome, which would allow calling single nucleotide polymorphic (SNP) markers for use in the analysis. The mitochondrial sequences obtained in this study present an important resource for future use in the conservation genetics of sleeper shark species.

# Works Cited


Acero, P. A., Polo-Silva, C. J., León, J., & Puentes, V. (2018). First report of a sleeper shark (*Somniosus sp*.) in the southern Colombian Caribbean. *Journal of Applied Ichthyology*, *34*(4), 981-983.

Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H., Coulson, A. R., Drouin, J., ... & Young, I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature*, *290*(5806), 457-465.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.

Barber, P. H., & Bellwood, D. R. (2005). Biodiversity hotspots: evolutionary origins of biodiversity in wrasses (Halichoeres: Labridae) in the Indo-Pacific and new world tropics. *Molecular Phylogenetics and Evolution*, *35*(1), 235-253.

Bass, A. J., Aubrey, J.D., Kistnasamy, N. (1976). Shark of the east coast of south Africa VI. The families Oxynotidae, Squalidae. Dalatiidae. and Echinorhinidae. *Investigational Report Oceanographic Research Institute*, *45*, 1-103.

Benfield, M. C., Thompson, B. A., & Caruso, J. H. (2008). The second report of a sleeper shark (*Somniosus* (*Somniosus*) *sp*.) from the bathypelagic waters of the northern Gulf of Mexico. *Bulletin of Marine Science*, *82*(2), 195.

Benz, G. W., Hocking, R., Kowunna, A., Bullard, S. A., & George, J. C. (2004). A second species of Arctic shark: Pacific sleeper shark *Somniosus pacificus* from Point Hope, Alaska. *Polar Biology*, *27*(4), 250-252.

Benz, G. W., Hoffmayer, E. R., Driggers, W. B., Allen, D., Bishop, L. E., & Brown, D. A. (2007). First record of a sleeper shark in the western Gulf of Mexico and comments on taxonomic uncertainty within *Somniosus* (*Somniosus*). *Bulletin of Marine Science*, *80*(2), 343-351.

Bortolus, A. (2008). Error cascades in the biological sciences: the unwanted consequences of using bad taxonomy in ecology. *AMBIO: A journal of the human environment*, *37*(2), 114-118.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., ... & Drummond, A. J. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*(4), e1003537.

Brown, T. A. (2002). Molecular phylogenetics. In *Genomes. 2nd edition*. Wiley-Liss.

Bustamante, C., Vargas-Caro, C., & Bennett, M. B. (2014). Not all fish are equal: functional biodiversity of cartilaginous fishes (Elasmobranchii and Holocephali) in Chile. *Journal of Fish Biology*, *85*(5), 1617-1633.

Capape, C., Rafrafi-Nouira, S., Diatta, Y., & Reynaud, C. (2020). FIRST RECORD OF LITTLE SLEEPER SHARK, *SOMNIOSUS ROSTRATUS* (ELASMOBRANCHII: SQUALIFORMES: SOMNIOSIDAE), FROM THE TUNISIAN COAST, CENTRAL MEDITERRANEAN SEA. *Acta Ichthyologica et Piscatoria*, *50*(4).

Carpenter, K. E. (n.d.). *Somniosus Rostratus*. FishBase. Retrieved May 11, 2022, from https://www.fishbase.se/summary/Somniosus-rostratus

Chazeau, C., Iglésias, S. P., Péron, C., Gasco, N., Martin, A., & Duhamel, G. (2019). THE KERGUELEN PLATEAU: MARINE ECOSYSTEM+ FISHERIES.

Chen, H., Ding, W., Chen, X., & Ai, W. (2016). Complete mitochondrial genome and the phylogenetic position of the smalleye pygmy shark *Squaliolus aliae* (Squaliformes: Dalatiidae). *Mitochondrial DNA Part B*, *1*(1), 304-305.

Chen, X., Liu, M., Grewe, P. M., Kyne, P. M., & Feutry, P. (2014). Complete mitochondrial genome of the Critically Endangered speartooth shark *Glyphis glyphis* (Carcharhiniformes: Carcharhinidae).

Choi, N. R., Seo, D. W., Jin, S. D., Sultana, H., Heo, K. N., & Lee, J. H. (2014). Phylogenetic Analysis using mtDNA D-loop Sequences in Korean Native Ducks. *Korean Journal of Poultry Science*, *41*(4), 235-240.

Coates, A. G., & Stallard, R. F. (2013). How old is the Isthmus of Panama?. *Bulletin of Marine Science*, *89*(4), 801-813.

Cohen, K.M., Finney, S.C., Gibbard, P.L. & Fan, J.-X. (2013; updated) The ICS International Chronostratigraphic Chart. Episodes 36: 199-204.

Cracraft, J. (1983). Species concepts and speciation analysis. In *Current Ornithology* (pp. 159-187). Springer, New York, NY.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M. T., Weihmann, A., Nickel, B., ... & Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences*, *110*(39), 15758-15763.

Damgaard, P. B., Margaryan, A., Schroeder, H., Orlando, L., Willerslev, E., & Allentoft, M. E. (2015). Improving access to endogenous DNA in ancient bones and teeth. *Scientific Reports*, *5*(1), 1-12.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, *9*(8), 772-772.

de Loyola Fernández, I., Báez, J. C., García-Barcelona, S., Camiñas, J. A., Ortiz de Urbina, J. M., & Macías, D. (2017). Length-weight relationships of kitefin shark *Dalatias licha*, and little sleeper shark *Somniosus rostratus* from the western Mediterranean Sea, and long snouted lancetfish *Alepisaurus ferox* from the eastern North Atlantic Ocean. *Turkish Journal of Fisheries and Aquatic Sciences*, *17*(5), 1073-1076.

Dierckxsens, N., Mardulyn, P., & Smits, G. (2017). NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research*, *45*(4), e18-e18.

Doane, M. P., Kacev, D., Harrington, S., Levi, K., Pande, D., Vega, A., & Dinsdale, E. A. (2018). Mitochondrial recovery from shotgun metagenome sequencing enabling phylogenetic analysis of the common thresher shark (*Alopias vulpinus*). *Meta Gene*, *15*, 10-15.

Dohm, J. C., Lottaz, C., Borodina, T., & Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Research*, *36*(16), e105.

Drummond, A. J., Ho, S. Y. W., Phillips, M. J., & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology*, *4*(5), e88.

Duchêne S, Archer FI, Vilstrup J, Caballero S, Morin PA (2011) Mitogenome Phylogenetics: The Impact of Using Single Regions and Partitioning Schemes on Topology, Substitution Rate and Divergence Time Estimation. *PLoS ONE*, 6(11): e27138.

Dudgeon, C. L., Blower, D. C., Broderick, D., Giles, J. L., Holmes, B. J., Kashiwagi, T., ... & Ovenden, J. R. (2012). A review of the application of molecular genetics for fisheries management and conservation of sharks and rays. *Journal of Fish Biology*, *80*(5), 1789-1843.

Dulvy, N. K., Fowler, S. L., Musick, J. A., Cavanagh, R. D., Kyne, P. M., Harrison, L. R., ... & White, W. T. (2014). Extinction risk and conservation of the world's sharks and rays. *elife*, *3*, e00590.

Ebert, D.A. 2015. Deep–Sea Cartilaginous Fishes of the Southeastern Atlantic Ocean. FAO Species Catalogue for Fishery Purposes No. 9.

Ebert, D.A., Fowler, S. and Compagno, L. 2013. *Sharks of the World*. Wild Nature Press, Plymouth.

Edwards, J. E., Hiltz, E., Broell, F., Bushnell, P. G., Campana, S. E., Christiansen, J. S., ... & Hussey, N. E. (2019). Advancing research for the management of long-lived species: a case study on the Greenland shark. *Frontiers in Marine Science*, 87.

Eldredge, N., & Cracraft, J. (1980). *Phylogenetic patterns and the evolutionary process: method and theory in comparative biology* (Vol. 349). New York: Columbia University Press.

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047-3048.

Farias, I. P., Ortí, G., Sampaio, I., Schneider, H., & Meyer, A. (2001). The cytochrome b gene as a phylogenetic marker: the limits of resolution for analyzing relationships among cichlid fishes. *Journal of Molecular Evolution*, *53*(2), 89-103.

Finucci, B. (2018). *Somniosus antarcticus*. *The IUCN Red List of Threatened Species* 2018: e.T41857A68643703.

Finucci, B., Cheok, J., Cotton, C.F., Kulka, D.W., Neat, F.C., Rigby, C.L., Tanaka, S. & Walker, T.I. (2020). *Somniosus rostratus*. *The IUCN Red List of Threatened Species* 2020: e.T161432A124484522.

Finucci, B., Francis, M.P. & Tanaka, S. (2018). *Somniosus longus*. *The IUCN Red List of Threatened Species* 2018: e.T161552A116739120.

Flammensbeck, C. K., Pollerspöck, J., Schedel, F. D., Matzke, N. J., & Straube, N. (2018). Of teeth and trees: A fossil tip-dating approach to infer divergence times of extinct and extant squaliform sharks. *Zoologica Scripta*, *47*(5), 539-557.

Ford, J. K., Ellis, G. M., Matkin, C. O., Wetklo, M. H., Barrett-Lennard, L. G., & Withler, R. E. (2011). Shark predation and tooth wear in a population of northeastern Pacific killer whales. *Aquatic Biology*, *11*(3), 213-224.

Francis, M. P., Stevens, J. D., & Last, P. R. (1988). New records of *Somniosus* (Elasmobranchii: Squalidae) from Australasia, with comments on the taxonomy of the genus. *New Zealand Journal of Marine and Freshwater Research*, *22*(3), 401-409.

Froese, R., & Pauly, D. (Eds). (2017). *FishBase: Somniosus pacificus Bigelow & Schroeder, 1944*. Retrieved from www.fishbase.org.

Groves, C. P., Fernando, P., & Robovský, J. (2010). The sixth rhino: a taxonomic re-assessment of the critically endangered northern white rhinoceros. *PLoS ONE*, *5*(4), e9703.

Guerrero, H. (2018). *Excel Data Analysis: Modeling and Simulation*. Springer.

Gupta, A., Bhardwaj, A., Sharma, P., Pal, Y., & Kumar, M. (2015). Mitochondrial DNA-a tool for phylogenetic and biodiversity search in equines. *Journal of Biodiversity & Endangered Species*, *1*, 006.

Hamlett, W. C. (2011). Genus Somniosus. In *Reproductive Biology and Phylogeny of Chondrichthyes* (pp. 518-547). CRC Press.

Haney, R. A., Silliman, B. R., & Rand, D. M. (2010). Effects of selection and mutation on mitochondrial variation and inferences of historical population expansion in a Caribbean reef fish. *Molecular Phylogenetics and Evolution*, *57*(2), 821-828.

Harley, E. H., de Waal, M., Murray, S., & O'Ryan, C. (2016). Comparison of whole mitochondrial genome sequences of northern and southern white rhinoceroses (Ceratotherium simum): the conservation consequences of species definitions. *Conservation Genetics*, *17*(6), 1285-1291.

Havird, J. C., & Santos, S. R. (2014). Performance of single and concatenated sets of mitochondrial genes at inferring metazoan relationships relative to full mitogenome data. *PLoS ONE*, *9*(1), e84080.

Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S., Francis, C. M., & Godfray, C. (2004). Identification of birds through DNA barcodes. *PLoS Biology*, *2*(10), e312.

Higashi, H., Hirao, S., Yamada, J., & Kikuchi, R. (1955). The vitamin A content of *Heteroscymnus longus* Tanaka, a kind of deep sea shark. *The Journal of Vitaminology*, *1*(3), 200-207.

Hsu, H. H., Lin, C. Y., & Joung, S. J. (2020). *Somniosus* (*Rhinoscymnus*) *cheni sp*. nov., A New Species of Sleeper Shark (Squaliformes: Somniosidae) from Eastern Taiwan, with Aspects of Embryo Biology. *Zoological Studies*, *59*.

Hussey, N. E., Cosandey-Godin, A., Walter, R. P., Hedges, K. J., VanGerwen-Toyne, M., Barkley, A. N., ... & Fisk, A. T. (2015). Juvenile Greenland sharks *Somniosus microcephalus* (Bloch & Schneider, 1801) in the canadian arctic. *Polar Biology*, *38*(4), 493-504.

Irmak, E., & Özden, U. (2021). A rare shark for the Mediterranean: *Somniosus rostratus* (Risso, 1827)(Chondrichthyes: Somniosidae) from the coast of Turkey. *Zoology in the Middle East*, *67*(3), 274-276.

Isaac, N. J., Mallet, J., & Mace, G. M. (2004). Taxonomic inflation: its influence on macroecology and conservation. *Trends in Ecology & Evolution*, *19*(9), 464-469.

Iwasaki, W., Fukunaga, T., Isagozawa, R., Yamada, K., Maeda, Y., Satoh, T. P., ... & Nishida, M. (2013). MitoFish and MitoAnnotator: a mitochondrial genome database of fish with an accurate and automatic annotation pipeline. *Molecular Biology and Evolution*, *30*(11), 2531-2540.

Jadin, R. C., Townsend, J. H., Castoe, T. A., & Campbell, J. A. (2012). Cryptic diversity in disjunct populations of Middle American Montane Pitvipers: a systematic reassessment of *Cerrophidion godmani*. *Zoologica Scripta*, *41*(5), 455-470.

Kang, C. B., Lee, W. J., Kim, J. K., & Jung, H. G. (2015). A new record of the Pacific sleeper shark, *Somniosus pacificus* (Squaliformes: Somniosidae) from the western margin of the East Sea, Korea. *Korean Journal of Ichthyology*, *27*(1), 45-49.

Kitamura, T., Takemura, A., Watabe, S., Taniuchi, T., & Shimizu, M. (1996). Molecular phylogeny of the sharks and rays of superorder Squalea based on mitochondrial cytochrome b gene. *Fisheries Science*, *62*(3), 340-343.

Kulka, D.W., Cotton, C.F., Anderson, B., Derrick, D., Herman, K. & Dulvy, N.K. (2020). *Somniosus microcephalus*. *The IUCN Red List of Threatened Species* 2020: e.T60213A124452872.

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547.

Last, P.R. and Stevens, J.D. (1994). *Sharks and Rays of Australia*. CSIRO Publishing, Australia.

Last, P.R. and Stevens, J.D. (2009). *Sharks and Rays of Australia. Second Edition*. CSIRO Publishing, Australia.

Lee, M. S., & Palci, A. (2015). Morphological phylogenetics in the genomic age. *Current Biology*, *25*(19), R922-R929.

Leigh, J. W., & Bryant, D. (2015). POPART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110-1116.

List, I. R., & Stevens, A. Southern Sleeper Shark, *Somniosus antarcticus*. (2019).

Mace, G. M., Gittleman, J. L., & Purvis, A. (2003). Preserving the tree of life. S*cience*, *300*(5626), 1707-1709.

MacNeil, M. A., McMeans, B. C., Hussey, N. E., Vecsei, P., Svavarsson, J., Kovacs, K. M., ... & Fisk, A. T. (2012). Biology of the Greenland shark *Somniosus microcephalus*. *Journal of Fish Biology*, *80*(5), 991-1018.

Margaryan, A. (2020). *Somniosus microcephalus* isolate DM359 mitochondrion, complete genome. DNAmark Project.

Marshall, C., Sturk-Andreaggi, K., Daniels-Higginbotham, J., Oliver, R. S., Barritt-Ross, S., & McMahon, T. P. (2017). Performance evaluation of a mitogenome capture and Illumina sequencing protocol using non-probative, case-type skeletal samples: Implications for the use of a positive control in a next-generation sequencing procedure. *Forensic Science International: Genetics*, *31*, 198-206.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. Journal*, *17*(1), 10-12.

Mayr, E. (1963). Animal species and evolution. In *Animal Species and Evolution*. Belknap Press of Harvard University Press.

Mecklenburg, C. W., Johannesen, E., Byrkjedal, I., Christiansen, J. S., Dolgov, A. V., Karamushko, O. V., ... & Wienerroither, R. M. (2018). Marine Fishes of the Arctic Region Volume 2.

Meiklejohn, K. A., Danielson, M. J., Faircloth, B. C., Glenn, T. C., Braun, E. L., & Kimball, R. T. (2014). Incongruence among different mitochondrial regions: a case study using complete mitogenomes. *Molecular phylogenetics and evolution*, *78*, 314-323.

Moodley, Y., & Harley, E. H. (2005). Population structuring in mountain zebras (*Equus zebra*): the molecular consequences of divergent demographic histories. *Conservation Genetics*, *6*(6), 953-968.

Murray, B. W., Wang, J. Y., Yang, S. C., Stevens, J. D., Fisk, A., & Svavarsson, J. (2008). Mitochondrial cytochrome b variation in sleeper sharks (Squaliformes: Somniosidae). *Marine Biology*, *153*(6), 1015-1022.

NAFO. (2018). Amendments to NAFO CEM – Measure to Conserve Greenland Sharks. 40th ANNUAL MEETING OF NAFO. Northwest Atlantic Fisheries Organization NAFO/COM Doc. 18-17.

Nelson, G., and N. I. Platnick. (1981). Systematics and biogeography: Cladistics and vicariance. New York: *Columbia University Press*.

Nielsen, J., Hedeholm, R. B., Heinemeier, J., Bushnell, P. G., Christiansen, J. S., Olsen, J., ... & Steffensen, J. F. (2016). Eye lens radiocarbon reveals centuries of longevity in the Greenland shark (Somniosus microcephalus). *Science*, *353*(6300), 702-704.

Nixon, K. C., & Wheeler, Q. D. (1990). An amplification of the phylogenetic species concept. *Cladistics*, *6*(3), 211-223.

O'Dea, A., Lessios, H. A., Coates, A. G., Eytan, R. I., Restrepo-Moreno, S. A., Cione, A. L., ... & Jackson, J. B. (2016). Formation of the Isthmus of Panama. *Science Advances*, *2*(8), e1600883.

Ono, K. A., Steinbeiser, C. M., Coco, A. B., Sheehan, M. J., Beck, A. J., Dufault, M. N., ... & Olson, Z. H. (2019). Detecting spiny dogfish in grey seal diets. *Conservation Genetics Resources*, *11*(4), 481-485.

Parin, N. V., & Kotlyar, A. N. (2007). On finding of shark of the genus *Somniosus* (Squalidae) at the submarine ridge of Nazca (southeastern Pacific). *Journal of Ichthyology*, *47*(8), 669-672.

Parin, N. V., Mironov, A. N., & Nesis, K. N. (1997). Biology of the Nazca and Sala y Gomez submarine ridges, an outpost of the Indo-West Pacific fauna in the eastern Pacific Ocean: composition and distribution of the fauna, its communities and history. *Advances in Marine Biology* (Vol. 32, pp. 145-242). Academic Press.

Pereira, S. L. (2000). Mitochondrial genome organization and vertebrate phylogenetics. *Genetics and Molecular biology*, *23*(4), 745-752.

Pollerspöck, J., & Straube, N. (2021). Bibliography Database of Living/Fossil Sharks, Rays and Chimaeras (Chondrichtyes: Elasmobranchii, Holocephali)—*List of Valid Extant Species; List of Described Extant Species; Statistic. World Wide Web electronic publication, Version*, *3*, 2021.

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., ... & Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, *55*(4), 595-609.

Qin, Y., Li, M., Cao, Y., Gao, Y., & Zhang, W. (2017). Molecular thresholds of ITS2 and their implications for molecular evolution and species identification in seed plants. *Scientific Reports*, *7*(1), 1-8.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, *67*(5), 901-904.

Rigby, C.L., Derrick, D., Dyldin, Y.V., Ebert, D.A., Herman, K., Ho, H., Hsu, H., Ishihara, H., Jeong, C.-H., Pacoureau, N., Semba, Y., Tanaka, S., Volvenko, I.V., Walls, R.H.L. & Yamaguchi, A. (2021). *Somniosus pacificus*. *The IUCN Red List of Threatened Species* 2021: e.T161403A887942.

Rohland, N., Malaspinas, A. S., Pollack, J. L., Slatkin, M., Matheus, P., & Hofreiter, M. (2007). Proboscidean mitogenomics: chronology and mode of elephant evolution using mastodon as outgroup. *PLoS Biology*, *5*(8), e207.

Santaquiteria, A., Nielsen, J., Klemetsen, T., Willassen, N. P., & Præbel, K. (2017). The complete mitochondrial genome of the long-lived Greenland shark (*Somniosus microcephalus*): characterization and phylogenetic position. *Conservation Genetics Resources*, *9*(3), 351-355.

Sayers, E. W., Cavanaugh, M., Clark, K., Pruitt, K. D., Schoch, C. L., Sherry, S. T., & Karsch-Mizrachi, I. (2021). GenBank. *Nucleic Acids Research*, *49*(D1), D92-D96.

Serreze, M. C., & Meier, W. N. (2019). The Arctic's sea ice cover: trends, variability, predictability, and comparisons to the Antarctic. *Annals of the New York Academy of Sciences*, *1436*(1), 36-53.

Sher, A. (1999). Traffic lights at the Beringian crossroads. *Nature*, *397*(6715), 103-104.

Snider, M. (2022, April 8). Rare Greenland shark washes aground in UK and found to be first to have meningitis. *USA Today*.

Stevens, G. (1996). Extending Rapoport's rule to Pacific marine fishes. *Journal of Biogeography*, *23*(2), 149-154.

Stokesbury, M. J., Harvey-Clark, C., Gallant, J., Block, B. A., & Myers, R. A. (2005). Movement and environmental preferences of Greenland sharks (*Somniosus microcephalus)* electronically tagged in the St. Lawrence Estuary, Canada. *Marine Biology*, *148*(1), 159-165.

Straube, N., Li, C., Claes, J. M., Corrigan, S., & Naylor, G. J. (2015). Molecular phylogeny of Squaliformes and first occurrence of bioluminescence in sharks. *BMC Evolutionary Biology*, *15*(1), 1-10.

Taanman, J. W. (1999). The mitochondrial genome: structure, transcription, translation and replication. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1410*(2), 103-123.

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, *38*(7), 3022-3027.

Townsend, J. P., Su, Z., & Tekle, Y. I. (2012). Phylogenetic signal and noise: predicting the power of a data set to resolve phylogeny. *Systematic Biology*, *61*(5), 835.

Vecchione, M., Mickevich, M. F., Fauchald, K., Collette, B. B., Williams, A. B., Munroe, T. A., & Young, R. E. (2000). Importance of assessing taxonomic adequacy in determining fishing effects on marine biodiversity. *ICES Journal of Marine Science*, *57*(3), 677-681.

Vélez-Zuazo, X., & Agnarsson, I. (2011). Shark tales: a molecular species-level phylogeny of sharks (Selachimorpha, Chondrichthyes). *Molecular Phylogenetics and Evolution*, *58*(2), 207-217.

Vella, A., Vella, N., & Schembri, S. (2017). A molecular approach towards taxonomic identification of elasmobranch species from Maltese fisheries landings. *Marine Genomics*, *36*, 17-23.

Walter, R. P., Roy, D., Hussey, N. E., Stelbrink, B., Kovacs, K. M., Lydersen, C., ... & Fisk, A. T. (2017). Origins of the Greenland shark (*Somniosus microcephalus*): Impacts of ice-olation and introgression. *Ecology and Evolution*, *7*(19), 8113-8125.

Weigmann, S. (2016). Annotated checklist of the living sharks, batoids and chimaeras (Chondrichthyes) of the world, with a focus on biogeographical diversity. *Journal of Fish Biology*, *88*(3), 837-1037.

Welton, B. J., & Goedert, J. L. (2016). New fossil species of *Somniosus* and *Rhinoscymnus* (Squaliformes: Somniosidae), deep water sharks from Oligocene rocks of Western Washington State, USA. *NM Mus. Nat. Hist. Sci. Bull*, *74*, 309-326.

Wilson, E. O., & Brown, W. L. (1953). The subspecies concept and its taxonomic application. *Systematic Zoology*, *2*(3), 97-111.

Yang, L., Matthes-Rosana, K. A., & Naylor, G. J. (2016). Complete mitochondrial genome of the blacknose shark *Carcharhinus acronotus* (Elasmobranchii: Carcharhinidae). *Mitochondrial DNA Part A*, *27*(1), 169-170.

Yano, K., Stevens, J. D., & Compagno, L. J. (2004). A review of the systematics of the sleeper shark genus *Somniosus* with redescriptions of *Somniosus* (*Somniosus*) *antarcticus* and *Somniosus* (*Rhinoscymnus*) *longus* (Squaliformes: Somniosidae). *Ichthyological Research*, *51*(4), 360-373.

Yeh, J., & Drazen, J. C. (2009). Depth zonation and bathymetric trends of deep-sea megafaunal scavengers of the Hawaiian Islands. *Deep Sea Research Part I: Oceanographic Research Papers*, *56*(2), 251-266.

Zhou, X., Yu, Y., Li, Y., Wu, J., Zhang, X., Guo, X., & Wang, W. (2014). Comparative analysis of mitochondrial genomes in distinct nuclear ploidy loach *Misgurnus anguillicaudatus* and its implications for polyploidy evolution. *PloS ONE*, *9*(3), e92033.

Zhu, J., Yu, Y., & Nakhleh, L. (2016). In the light of deep coalescence: revisiting trees within networks. *BMC Bioinformatics*, *17*(14), 271-282.

# Appendix

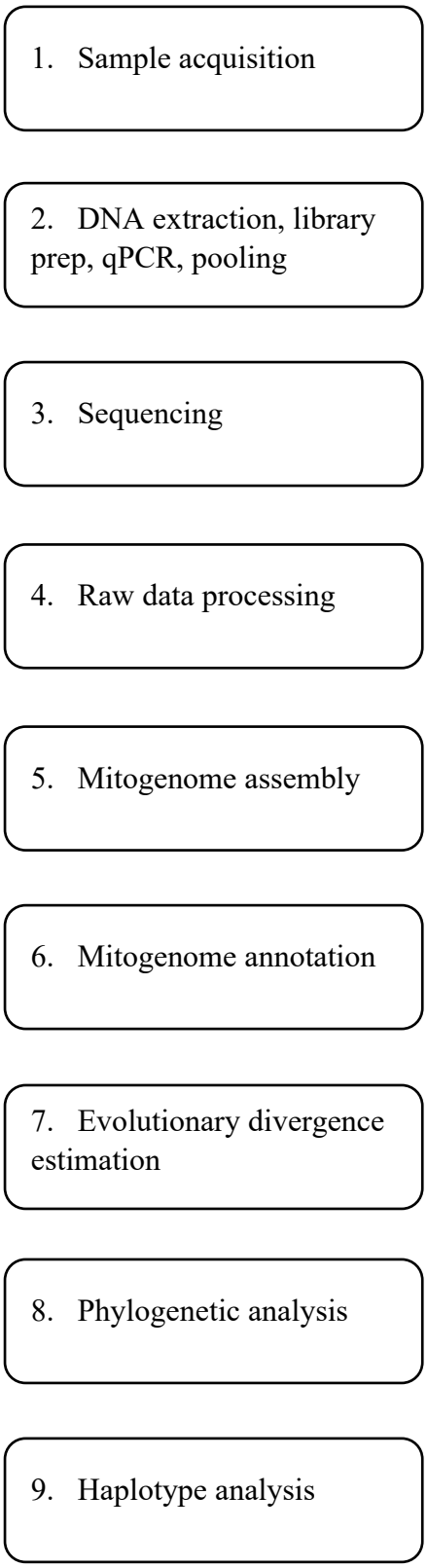



**Figure A1**

*Overview of materials and methods*

*Note.*

1. Sample GS510 was acquired from Dr. Samuel Iglésias (MNHN). Samples GS511, GS512, GS513, and GS514 were acquired from Ms. Charlotte Chazeau, (MNHN). Samples GS515 and GS516 were acquired from Andrew Williston (MCZ Harvard).

3. Sequencing was performed by Novogene (UK)

4. Raw data processing was performed by Dr. Shripathi Bhat (UiT RGG)

2., 5-9. All other procedures were performed with support from RGG staff.

**Table A1**

*Overview of specimens acquired outside of UiT*

| ID | Institutional ID | Species | Acquired from |
|---|---|---|---|
| **GS510** | BPS 0799 | *S. antarcticus* | Dr. Samuel Iglésias, Museum National D'historie Naturelle, France |
| **GS511** | ALC-634-0015 | *S. antarcticus* | Ms. Charlotte Chazeau, Museum National D'historie Naturelle, France |
| **GS512** | ALC-634-0002 | *S. antarcticus* | Ms. Charlotte Chazeau, Museum National D'historie Naturelle, France |
| **GS513** | ALC-675-0011 | *S. antarcticus* | Ms. Charlotte Chazeau, Museum National D'historie Naturelle, France |
| **GS514** | ALC-639-0001 | *S. antarcticus* | Ms. Charlotte Chazeau, Museum National D'historie Naturelle, France |
| **GS515** | MCZ:Ich:39650 | *S. longus* | Andrew Williston, Museum of Comparative Zoology Harvard, USA |
| **GS516** | MCZ:153660 | *S. longus* | Andrew Williston, Museum of Comparative Zoology Harvard, USA |

**Table A2**

*Overview of all specimens used in this study*

| ID | Species | Sequenced (+/-) | Used in study (+/-) |
|---|---|---|---|
| **GS60** | *S. microcephalus* | - | + |
| **GS64** | *S. microcephalus* | + | - |
| **GS67** | *S. microcephalus* | + | - |
| **GS80** | *S. microcephalus* | + | + |
| **GS115** | *S. microcephalus* | + | - |
| **GS121** | *S. microcephalus* | + | - |
| **GS122** | *S. microcephalus* | + | - |
| **GS159** | *S. microcephalus* | + | + |
| **GS201** | *S. microcephalus* | - | - |
| **GS202** | *S. microcephalus* | - | - |
| **GS203** | *S. microcephalus* | - | - |
| **GS204** | *S. microcephalus* | - | - |
| **GS205** | *S. microcephalus* | - | - |
| **GS206** | *S. microcephalus* | - | - |
| **GS207** | *S. microcephalus* | - | - |
| **GS216** | *S. rostratus* | + | + |
| **GS217** | *S. rostratus* | - | - |
| **GS218** | *S. rostratus* | + | + |
| **GS219** | *S. rostratus* | + | - |
| **GS282** | *S. microcephalus* | - | - |
| **GS283** | *S. microcephalus* | - | - |
| **GS284** | *S. microcephalus* | - | - |
| **GS285** | *S. microcephalus* | - | - |
| **GS286** | *S. microcephalus* | - | - |
| **GS288** | *S. microcephalus* | - | - |
| **GS303** | *S. microcephalus* | - | - |
| **GS500** | *S. pacificus* | + | + |
| **GS501** | *S. pacificus* | - | - |
| **GS502** | *S. pacificus* | + | + |
| **GS503** | *S. pacificus* | - | - |
| **GS504** | *S. pacificus* | - | - |
| **GS505** | *S. pacificus* | + | + |
| **GS506** | *S. pacificus* | - | - |
| **GS507** | *S. pacificus* | - | - |
| **GS508** | *S. pacificus* | - | - |
| **GS509** | *S. pacificus* | - | - |
| **GS510** | *S. antarcticus* | + | + |
| **GS511** | *S. antarcticus* | - | - |
| **GS512** | *S. antarcticus* | - | - |
| **GS513** | *S. antarcticus* | + | + |
| **GS514** | *S. antarcticus* | + | + |
| **GS515** | *S. longus* | - | - |
| **GS516** | *S. longus* | + | + |

**Figure A2**

*Laboratory Workflow*

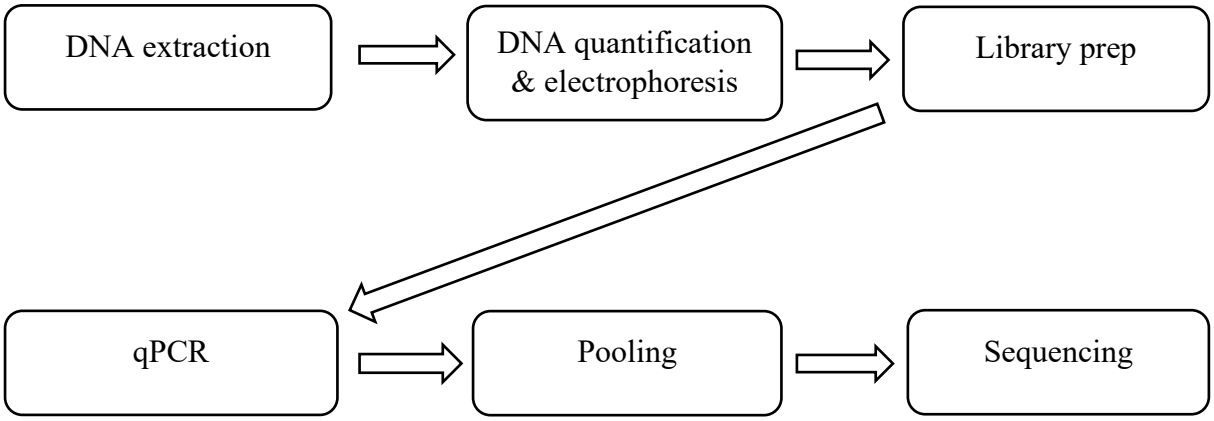



**Table A3**

*NCBI GenBank Sequences used*

| GenBank accession number | Species | Gene | Region | Use | Reference |
|---|---|---|---|---|---|
| **EF090963.1** | *Somniosus antarcticus* | Cyt *b* | 1 – 703 | *NOVOPlasty* Seed | Murray et al., 2008 |
| **KP059873.1** | *Somniosus microcephalus* | Cyt *b* | 1 – 702 | *NOVOPlasty* Seed | Hussey et al., 2015 |
| **HQ260437.1** | *Somniosus pacificus* | Cyt *b* | 1 – 745 | *NOVOPlasty* Seed | Ford et al., 2011 |
| **KY296981.1** | *Squalus acanthias* | Cyt *b* | 1 – 796 | *NOVOPlasty* Seed | Ono et al., 2019 |
| **KY909801.1** | *Somniosus rostratus* | ND2 | 1 – 990 | *NOVOPlasty* Seed | Vella et al., 2017 |
| **NC_049864.1** | *Somniosus microcephalus* | Full mitogenome | 1 – 16,730 | Reference mitogenome | Margaryan, 2020 |
| **KU873080.1** | *Squaliolus aliae* | Full mitogenome | 1 – 16,717 | Outgroup | Chen et al., 2016 |

**Table A4**

*Annotation of GS60, GS80, and GS159, Somniosus microcephalus mitogenomes*

| GS60, GS80, GS159 *Somniosus microcephalus* | | | | | | | |
|---|---|---|---|---|---|---|---|
| Gene | Strands | Gene | | | | | Intergenetic space |
| | | Start (bp) | Stop (bp) | Size (bp) | Start codon | Stop codon | |
| $tRNA^{Phe}$ | H | 1 | 69 | 69 | | | |
| 12S rRNA | H | 70 | 1,020 | 951 | | | |
| $tRNA^{Val}$ | H | 1,021 | 1,092 | 72 | | | |
| 16S | H | 1,093 | 2,767 | 1,675 | | | |
| $tRNA^{Leu\ (UAA)}$ | H | 2,768 | 2,842 | 75 | | | |
| ND1 | H | 2,843 | 3,817 | 975 | ATG | TAG | 2 |
| $tRNA^{Ile}$ | H | 3,820 | 3,889 | 70 | | | 1 |
| $tRNA^{Gln}$ | L | 3,891 | 3,962 | 72 | | | |
| $tRNA^{Met}$ | H | 3,963 | 4,031 | 69 | | | |
| ND2 | H | 4,032 | 5,077 | 1,046 | ATG | TA- | |
| $tRNA^{Trp}$ | H | 5,078 | 5,146 | 69 | | | 1 |
| $tRNA^{Ala}$ | L | 5,148 | 5,216 | 69 | | | |
| $tRNA^{Asn}$ | L | 5,217 | 5,290 | 74 | | | |
| OL | - | 5,291 | 5,327 | 37 | | | |
| $tRNA^{Cys}$ | L | 5,328 | 5,394 | 67 | | | 1 |
| $tRNA^{Tyr}$ | L | 5,396 | 5,465 | 70 | | | 1 |
| COX-1 | H | 5,467 | 7,023 | 1,557 | GTG | TAA | |
| $tRNA^{Ser\ (UGA)}$ | L | 7,024 | 7,094 | 71 | | | 3 |
| $tRNA^{Asp}$ | H | 7,098 | 7,167 | 70 | | | 4 |
| COX-2 | H | 7,172 | 7,862 | 691 | ATG | T-- | |
| $tRNA^{Lys}$ | H | 7,863 | 7,936 | 74 | | | 1 |
| ATP8 | H | 7,938 | 8,105 | 168 | ATG | TAA | -10 |
| ATP6 | H | 8,096 | 8,778 | 683 | ATG | TA- | |
| COX-3 | H | 8,779 | 9,564 | 786 | ATG | TAA | 2 |
| $tRNA^{Gly}$ | H | 9,567 | 9,636 | 70 | | | |
| ND3 | H | 9,637 | 9,987 | 351 | ATG | TAA | 3 |
| $tRNA^{Arg}$ | H | 9,991 | 10,060 | 70 | | | |
| ND4L | H | 10,061 | 10,357 | 297 | ATG | TAA | -7 |
| ND4 | H | 10,351 | 11,731 | 1,381 | ATG | T-- | |
| tRNAHis | H | 11,732 | 11,800 | 69 | | | |
| $tRNA^{Ser\ (GCU)}$ | H | 11,801 | 11,867 | 67 | | | |
| $tRNA^{Leu\ (UAG)}$ | H | 11,868 | 11,939 | 72 | | | |
| ND5 | H | 11,940 | 13,772 | 1,833 | ATG | TAA | -4 |
| ND6 | L | 13,769 | 14,290 | 522 | ATG | TAG | |
| $tRNA^{Glu}$ | L | 14,291 | 14,360 | 70 | | | 4 |
| Cyt *b* | H | 14,365 | 15,510 | 1,146 | ATG | TAA | 1 |
| $tRNA^{Thr}$ | H | 15,512 | 15,584 | 73 | | | 2 |
| $tRNA^{Pro}$ | L | 15,587 | 15,655 | 69 | | | |
| CR | - | 15,656 | 16,730 | 1,075 | | | |

*Note*. H = heavy strand; L = light strand; Intergenetic space = positive values are spacer DNA, negative values are overlapping genes; OL = origin of light strand replication.

**Table A5**

*Annotation of GS216, Somniosus rostratus mitogenome*

| GS216 *Somniosus rostratus* | | | | | | | |
|---|---|---|---|---|---|---|---|
| Gene | Strands | Gene | | | | | Intergenetic space |
| | | Start (bp) | Stop (bp) | Size (bp) | Start codon | Stop codon | |
| $tRNA^{Phe}$ | H | 1 | 69 | 69 | | | |
| 12S rRNA | H | 70 | 1,019 | 950 | | | |
| $tRNA^{Val}$ | H | 1,020 | 1,091 | 72 | | | |
| 16S | H | 1,092 | 2,765 | 1,674 | | | |
| $tRNA^{Leu\ (UAA)}$ | H | 2,766 | 2,840 | 75 | | | |
| ND1 | H | 2,841 | 3,815 | 975 | ATG | TAG | 2 |
| $tRNA^{Ile}$ | H | 3,818 | 3,887 | 70 | | | 1 |
| $tRNA^{Gln}$ | L | 3,889 | 3,960 | 72 | | | |
| $tRNA^{Met}$ | H | 3,961 | 4,029 | 69 | | | |
| ND2 | H | 4,030 | 5,075 | 1,046 | ATG | TA- | |
| $tRNA^{Trp}$ | H | 5,076 | 5,144 | 69 | | | 1 |
| $tRNA^{Ala}$ | L | 5,146 | 5,214 | 69 | | | |
| $tRNA^{Asn}$ | L | 5,215 | 5,288 | 74 | | | |
| OL | - | 5,289 | 5,325 | 37 | | | |
| $tRNA^{Cys}$ | L | 5,326 | 5,392 | 67 | | | 1 |
| $tRNA^{Tyr}$ | L | 5,394 | 5,463 | 70 | | | 1 |
| COX-1 | H | 5,465 | 7,021 | 1,557 | GTG | TAA | |
| $tRNA^{Ser\ (UGA)}$ | L | 7,022 | 7,092 | 71 | | | 3 |
| $tRNA^{Asp}$ | H | 7,096 | 7,165 | 70 | | | 4 |
| COX-2 | H | 7,170 | 7,860 | 691 | ATG | T-- | |
| $tRNA^{Lys}$ | H | 7,861 | 7,934 | 74 | | | 1 |
| ATP8 | H | 7,936 | 8,103 | 168 | ATG | TAA | -10 |
| ATP6 | H | 8,094 | 8,776 | 683 | ATG | TA- | |
| COX-3 | H | 8,777 | 9,562 | 786 | ATG | TAA | 2 |
| $tRNA^{Gly}$ | H | 9,565 | 9,634 | 70 | | | |
| ND3 | H | 9,635 | 9,985 | 351 | ATG | TAA | 3 |
| $tRNA^{Arg}$ | H | 9,989 | 10,058 | 70 | | | |
| ND4L | H | 10,059 | 10,355 | 297 | ATG | TAA | -7 |
| ND4 | H | 10,349 | 11,729 | 1,381 | ATG | T-- | |
| tRNAHis | H | 11,730 | 11,798 | 69 | | | |
| $tRNA^{Ser\ (GCU)}$ | H | 11,799 | 11,865 | 67 | | | |
| $tRNA^{Leu\ (UAG)}$ | H | 11,866 | 11,937 | 72 | | | |
| ND5 | H | 11,938 | 13,770 | 1,833 | ATG | TAA | -4 |
| ND6 | L | 13,767 | 14,288 | 522 | ATG | TAG | |
| $tRNA^{Glu}$ | L | 14,289 | 14,358 | 70 | | | 4 |
| Cyt *b* | H | 14,363 | 15,508 | 1,146 | ATG | TAA | 1 |
| $tRNA^{Thr}$ | H | 15,510 | 15,582 | 73 | | | 2 |
| $tRNA^{Pro}$ | L | 15,585 | 15,653 | 69 | | | |
| CR | - | 15,654 | 16,727 | 1,074 | | | |

*Note*. H = heavy strand; L = light strand; Intergenetic space = positive values are spacer DNA, negative values are overlapping genes; OL = origin of light strand replication.

**Table A6**

*Annotation of GS218, Somniosus rostratus mitogenome*

| GS218 *Somniosus rostratus* | | | | | | | |
|---|---|---|---|---|---|---|---|
| Gene | Strands | Gene | | | | | Intergenetic space |
| | | Start (bp) | Stop (bp) | Size (bp) | Start codon | Stop codon | |
| $tRNA^{Phe}$ | H | 1 | 69 | 69 | | | |
| 12S rRNA | H | 70 | 1,019 | 950 | | | |
| $tRNA^{Val}$ | H | 1,020 | 1,091 | 72 | | | |
| 16S | H | 1,092 | 2,765 | 1,674 | | | |
| $tRNA^{Leu\ (UAA)}$ | H | 2,766 | 2,840 | 75 | | | |
| ND1 | H | 2,841 | 3,815 | 975 | ATG | TAG | 2 |
| $tRNA^{Ile}$ | H | 3,818 | 3,887 | 70 | | | 1 |
| $tRNA^{Gln}$ | L | 3,889 | 3,960 | 72 | | | |
| $tRNA^{Met}$ | H | 3,961 | 4,029 | 69 | | | |
| ND2 | H | 4,030 | 5,075 | 1,046 | ATG | TA- | |
| $tRNA^{Trp}$ | H | 5,076 | 5,144 | 69 | | | 1 |
| $tRNA^{Ala}$ | L | 5,146 | 5,214 | 69 | | | |
| $tRNA^{Asn}$ | L | 5,215 | 5,288 | 74 | | | |
| OL | - | 5,289 | 5,325 | 37 | | | |
| $tRNA^{Cys}$ | L | 5,326 | 5,392 | 67 | | | 1 |
| $tRNA^{Tyr}$ | L | 5,394 | 5,463 | 70 | | | 1 |
| COX-1 | H | 5,465 | 7,021 | 1,557 | GTG | TAA | |
| $tRNA^{Ser\ (UGA)}$) | L | 7,022 | 7,092 | 71 | | | 3 |
| $tRNA^{Asp}$ | H | 7,096 | 7,165 | 70 | | | 4 |
| COX-2 | H | 7,170 | 7,860 | 691 | ATG | T-- | |
| $tRNA^{Lys}$ | H | 7,861 | 7,934 | 74 | | | 1 |
| ATP8 | H | 7,936 | 8,103 | 168 | ATG | TAA | -10 |
| ATP6 | H | 8,094 | 8,776 | 683 | ATG | TA- | |
| COX-3 | H | 8,777 | 9,562 | 786 | ATG | TAA | 2 |
| $tRNA^{Gly}$ | H | 9,565 | 9,634 | 70 | | | |
| ND3 | H | 9,635 | 9,985 | 351 | ATG | TAA | 3 |
| $tRNA^{Arg}$ | H | 9,989 | 10,058 | 70 | | | |
| ND4L | H | 10,059 | 10,355 | 297 | ATG | TAA | -7 |
| ND4 | H | 10,349 | 11,729 | 1,381 | ATG | T-- | |
| $tRNA^{His}$ | H | 11,730 | 11,798 | 69 | | | |
| $tRNA^{Ser\ (GCU)}$ | H | 11,799 | 11,865 | 67 | | | |
| $tRNA^{Leu\ (UAG)}$ | H | 11,866 | 11,937 | 72 | | | |
| ND5 | H | 11,938 | 13,770 | 1,833 | ATG | TAA | -4 |
| ND6 | L | 13,767 | 14,288 | 522 | ATG | TAG | |
| $tRNA^{Glu}$ | L | 14,289 | 14,358 | 70 | | | 4 |
| Cyt *b* | H | 14,363 | 15,508 | 1,146 | ATG | TAA | 1 |
| $tRNA^{Thr}$ | H | 15,510 | 15,582 | 73 | | | 2 |
| $tRNA^{Pro}$ | L | 15,585 | 15,653 | 69 | | | |
| CR | - | 15,654 | 16,725 | 1,072 | | | |

*Note*. H = heavy strand; L = light strand; Intergenetic space = positive values are spacer DNA, negative values are overlapping genes; OL = origin of light strand replication.

**Table A7**

*Annotation of GS516, Somniosus longus mitogenome*

| GS516 *Somniosus longus* | | | | | | | |
|---|---|---|---|---|---|---|---|
| Gene | Strands | Gene | | | | | Intergenetic space |
| | | Start (bp) | Stop (bp) | Size (bp) | Start codon | Stop codon | |
| $tRNA^{Phe}$ | H | 1 | 69 | 69 | | | |
| 12S rRNA | H | 70 | 1,019 | 950 | | | |
| $tRNA^{Val}$ | H | 1,020 | 1,091 | 72 | | | |
| 16S | H | 1,092 | 2,765 | 1,674 | | | |
| $tRNA^{Leu\ (UAA)}$ | H | 2,766 | 2,840 | 75 | | | |
| ND1 | H | 2,841 | 3,815 | 975 | ATG | TAG | 2 |
| $tRNA^{Ile}$ | H | 3,818 | 3,887 | 70 | | | 1 |
| $tRNA^{Gln}$ | L | 3,889 | 3,960 | 72 | | | |
| $tRNA^{Met}$ | H | 3,961 | 4,029 | 69 | | | |
| ND2 | H | 4,030 | 5,075 | 1,046 | ATG | TA- | |
| $tRNA^{Trp}$ | H | 5,076 | 5,144 | 69 | | | 1 |
| $tRNA^{Ala}$ | L | 5,146 | 5,214 | 69 | | | |
| $tRNA^{Asn}$ | L | 5,215 | 5,288 | 74 | | | |
| OL | - | 5,289 | 5,325 | 37 | | | |
| $tRNA^{Cys}$ | L | 5,326 | 5,392 | 67 | | | 1 |
| $tRNA^{Tyr}$ | L | 5,394 | 5,463 | 70 | | | 1 |
| COX-1 | H | 5,465 | 7,021 | 1,557 | GTG | TAA | |
| $tRNA^{Ser\ (UGA)}$ | L | 7,022 | 7,092 | 71 | | | 3 |
| $tRNA^{Asp}$ | H | 7,096 | 7,165 | 70 | | | 4 |
| COX-2 | H | 7,170 | 7,860 | 691 | ATG | T-- | |
| $tRNA^{Lys}$ | H | 7,861 | 7,934 | 74 | | | 1 |
| ATP8 | H | 7,936 | 8,103 | 168 | ATG | TAA | -10 |
| ATP6 | H | 8,094 | 8,776 | 683 | ATG | TA- | |
| COX-3 | H | 8,777 | 9,562 | 786 | ATG | TAA | 2 |
| $tRNA^{Gly}$ | H | 9,565 | 9,634 | 70 | | | |
| ND3 | H | 9,635 | 9,985 | 351 | ATG | TAA | 3 |
| $tRNA^{Arg}$ | H | 9,989 | 10,058 | 70 | | | |
| ND4L | H | 10,059 | 10,355 | 297 | ATG | TAA | -7 |
| ND4 | H | 10,349 | 11,729 | 1,381 | ATG | T-- | |
| tRNAHis | H | 11,730 | 11,798 | 69 | | | |
| $tRNA^{Ser\ (GCU)}$ | H | 11,799 | 11,865 | 67 | | | |
| $tRNA^{Leu\ (UAG)}$ | H | 11,866 | 11,937 | 72 | | | |
| ND5 | H | 11,938 | 13,770 | 1,833 | ATG | TAA | -4 |
| ND6 | L | 13,767 | 14,288 | 522 | ATG | TAG | |
| $tRNA^{Glu}$ | L | 14,289 | 14,358 | 70 | | | 4 |
| Cyt *b* | H | 14,363 | 15,508 | 1,146 | ATG | TAA | 1 |
| $tRNA^{Thr}$ | H | 15,510 | 15,582 | 73 | | | 2 |
| $tRNA^{Pro}$ | L | 15,585 | 15,653 | 69 | | | |
| CR | - | 15,654 | 16,725 | 1,072 | | | |

*Note*. H = heavy strand; L = light strand; Intergenetic space = positive values are spacer DNA, negative values are overlapping genes; OL = origin of light strand replication.

**Table A8**

*Annotation of GS500, Somniosus pacificus mitogenome*

| GS500 *Somniosus pacificus* | | | | | | | |
|---|---|---|---|---|---|---|---|
| Gene | Strands | Gene | | | | | Intergenetic space |
| | | Start (bp) | Stop (bp) | Size (bp) | Start codon | Stop codon | |
| $tRNA^{Phe}$ | H | 1 | 69 | 69 | | | |
| 12S rRNA | H | 70 | 1,020 | 951 | | | |
| $tRNA^{Val}$ | H | 1,021 | 1,092 | 72 | | | |
| 16S | H | 1,093 | 2,767 | 1,675 | | | |
| $tRNA^{Leu\ (UAA)}$ | H | 2,768 | 2,842 | 75 | | | |
| ND1 | H | 2,843 | 3,817 | 975 | ATG | TAA | 2 |
| $tRNA^{Ile}$ | H | 3,820 | 3,889 | 70 | | | 1 |
| $tRNA^{Gln}$ | L | 3,891 | 3,962 | 72 | | | |
| $tRNA^{Met}$ | H | 3,963 | 4,031 | 69 | | | |
| ND2 | H | 4,032 | 5,077 | 1,046 | ATG | TA- | |
| $tRNA^{Trp}$ | H | 5,078 | 5,146 | 69 | | | 1 |
| $tRNA^{Ala}$ | L | 5,148 | 5,216 | 69 | | | |
| $tRNA^{Asn}$ | L | 5,217 | 5,290 | 74 | | | |
| OL | - | 5,291 | 5,327 | 37 | | | |
| $tRNA^{Cys}$ | L | 5,328 | 5,394 | 67 | | | 1 |
| $tRNA^{Tyr}$ | L | 5,396 | 5,465 | 70 | | | 1 |
| COX-1 | H | 5,467 | 7,023 | 1,557 | GTG | TAA | |
| $tRNA^{Ser\ (UGA)}$ | L | 7,024 | 7,094 | 71 | | | 3 |
| $tRNA^{Asp}$ | H | 7,098 | 7,167 | 70 | | | 4 |
| COX-2 | H | 7,172 | 7,862 | 691 | ATG | T-- | |
| $tRNA^{Lys}$ | H | 7,863 | 7,936 | 74 | | | 1 |
| ATP8 | H | 7,938 | 8,105 | 168 | ATG | TAA | -10 |
| ATP6 | H | 8,096 | 8,778 | 683 | ATG | TA- | |
| COX-3 | H | 8,779 | 9,564 | 786 | ATG | TAA | 2 |
| $tRNA^{Gly}$ | H | 9,567 | 9,636 | 70 | | | |
| ND3 | H | 9,637 | 9,987 | 351 | ATG | TAG | 3 |
| $tRNA^{Arg}$ | H | 9,991 | 10,060 | 70 | | | |
| ND4L | H | 10,061 | 10,357 | 297 | ATG | TAA | -7 |
| ND4 | H | 10,351 | 11,731 | 1,381 | ATG | T-- | |
| tRNAHis | H | 11,732 | 11,800 | 69 | | | |
| $tRNA^{Ser\ (GCU)}$ | H | 11,801 | 11,867 | 67 | | | |
| $tRNA^{Leu\ (UAG)}$ | H | 11,868 | 11,939 | 72 | | | |
| ND5 | H | 11,940 | 13,772 | 1,833 | ATG | TAA | -4 |
| ND6 | L | 13,769 | 14,291 | 523 | ATG | TAG | |
| $tRNA^{Glu}$ | L | 14,292 | 14,361 | 70 | | | 4 |
| Cyt *b* | H | 14,366 | 15,511 | 1,146 | ATG | TAA | 1 |
| $tRNA^{Thr}$ | H | 15,513 | 15,585 | 73 | | | 2 |
| $tRNA^{Pro}$ | L | 15,588 | 15,656 | 69 | | | |
| CR | - | 15,657 | 16,731 | 1,075 | | | |

*Note*. H = heavy strand; L = light strand; Intergenetic space = positive values are spacer DNA, negative values are overlapping genes; OL = origin of light strand replication.

**Table A9**

*Annotation of GS502, Somniosus pacificus mitogenome*

| GS502 *Somniosus pacificus* | | | | | | | |
|---|---|---|---|---|---|---|---|
| Gene | Strands | Gene | | | | | Intergenetic space |
| | | Start (bp) | Stop (bp) | Size (bp) | Start codon | Stop codon | |
| $tRNA^{Phe}$ | H | 1 | 69 | 69 | | | |
| 12S rRNA | H | 70 | 1,020 | 951 | | | |
| $tRNA^{Val}$ | H | 1,021 | 1,092 | 72 | | | |
| 16S | H | 1,093 | 2,767 | 1,675 | | | |
| $tRNA^{Leu\ (UAA)}$ | H | 2,768 | 2,842 | 75 | | | |
| ND1 | H | 2,843 | 3,817 | 975 | ATG | TAG | 2 |
| $tRNA^{Ile}$ | H | 3,820 | 3,889 | 70 | | | 1 |
| $tRNA^{Gln}$ | L | 3,891 | 3,962 | 72 | | | |
| $tRNA^{Met}$ | H | 3,963 | 4,031 | 69 | | | |
| ND2 | H | 4,032 | 5,077 | 1,046 | ATG | TA- | |
| $tRNA^{Trp}$ | H | 5,078 | 5,146 | 69 | | | 1 |
| $tRNA^{Ala}$ | L | 5,148 | 5,216 | 69 | | | |
| $tRNA^{Asn}$ | L | 5,217 | 5,290 | 74 | | | |
| OL | - | 5,291 | 5,327 | 37 | | | |
| $tRNA^{Cys}$ | L | 5,328 | 5,394 | 67 | | | 1 |
| $tRNA^{Tyr}$ | L | 5,396 | 5,465 | 70 | | | 1 |
| COX-1 | H | 5,467 | 7,023 | 1,557 | GTG | TAA | |
| $tRNA^{Ser\ (UGA)}$ | L | 7,024 | 7,094 | 71 | | | 3 |
| $tRNA^{Asp}$ | H | 7,098 | 7,167 | 70 | | | 4 |
| COX-2 | H | 7,172 | 7,862 | 691 | ATG | T-- | |
| $tRNA^{Lys}$ | H | 7,863 | 7,936 | 74 | | | 1 |
| ATP8 | H | 7,938 | 8,105 | 168 | ATG | TAA | -10 |
| ATP6 | H | 8,096 | 8,778 | 683 | ATG | TA- | |
| COX-3 | H | 8,779 | 9,564 | 786 | ATG | TAA | 2 |
| $tRNA^{Gly}$ | H | 9,567 | 9,636 | 70 | | | |
| ND3 | H | 9,637 | 9,987 | 351 | ATG | TAA | 3 |
| $tRNA^{Arg}$ | H | 9,991 | 10,060 | 70 | | | |
| ND4L | H | 10,061 | 10,357 | 297 | ATG | TAA | -7 |
| ND4 | H | 10,351 | 11,731 | 1,381 | ATG | T-- | |
| tRNAHis | H | 11,732 | 11,800 | 69 | | | |
| $tRNA^{Ser\ (GCU)}$ | H | 11,801 | 11,867 | 67 | | | |
| $tRNA^{Leu\ (UAG)}$ | H | 11,868 | 11,939 | 72 | | | |
| ND5 | H | 11,940 | 13,770 | 1,831 | ATG | TAA | -4 |
| ND6 | L | 13,767 | 14,288 | 522 | ATG | TAG | |
| $tRNA^{Glu}$ | L | 14,289 | 14,358 | 70 | | | 4 |
| Cyt *b* | H | 14,363 | 15,508 | 1,146 | ATG | TAA | 1 |
| $tRNA^{Thr}$ | H | 15,510 | 15,582 | 73 | | | 2 |
| $tRNA^{Pro}$ | L | 15,585 | 15,653 | 69 | | | |
| CR | - | 15,654 | 16,728 | 1,075 | | | |

*Note*. H = heavy strand; L = light strand; Intergenetic space = positive values are spacer DNA, negative values are overlapping genes; OL = origin of light strand replication.

**Table A10**

*Annotation of GS505, Somniosus pacificus mitogenome*

| GS505 *Somniosus pacificus* | | | | | | | |
|---|---|---|---|---|---|---|---|
| Gene | Strands | Gene | | | | | Intergenetic space |
| | | Start (bp) | Stop (bp) | Size (bp) | Start codon | Stop codon | |
| $tRNA^{Phe}$ | H | 1 | 69 | 69 | | | |
| 12S rRNA | H | 70 | 1,020 | 951 | TGT | | |
| $tRNA^{Val}$ | H | 1,021 | 1,092 | 72 | | | |
| 16S | H | 1,093 | 2,767 | 1,675 | ACA | | |
| $tRNA^{Leu\ (UAA)}$ | H | 2,768 | 2,842 | 75 | | | |
| ND1 | H | 2,843 | 3,817 | 975 | ATG | TAA | 2 |
| $tRNA^{Ile}$ | H | 3,820 | 3,889 | 70 | | | 1 |
| $tRNA^{Gln}$ | L | 3,891 | 3,962 | 72 | | | |
| $tRNA^{Met}$ | H | 3,963 | 4,031 | 69 | | | |
| ND2 | H | 4,032 | 5,077 | 1,046 | ATG | TA- | |
| $tRNA^{Trp}$ | H | 5,078 | 5,146 | 69 | | | 1 |
| $tRNA^{Ala}$ | L | 5,148 | 5,216 | 69 | | | |
| $tRNA^{Asn}$ | L | 5,217 | 5,290 | 74 | | | |
| OL | - | 5,291 | 5,327 | 37 | | | |
| $tRNA^{Cys}$ | L | 5,328 | 5,394 | 67 | | | 1 |
| $tRNA^{Tyr}$ | L | 5,396 | 5,465 | 70 | | | 1 |
| COX-1 | H | 5,467 | 7,023 | 1,557 | GTG | TAA | |
| $tRNA^{Ser\ (UGA)}$ | L | 7,024 | 7,094 | 71 | | | 3 |
| $tRNA^{Asp}$ | H | 7,098 | 7,167 | 70 | | | 4 |
| COX-2 | H | 7,172 | 7,862 | 691 | ATG | T-- | |
| $tRNA^{Lys}$ | H | 7,863 | 7,936 | 74 | | | 1 |
| ATP8 | H | 7,938 | 8,105 | 168 | ATG | TAA | -10 |
| ATP6 | H | 8,096 | 8,778 | 683 | ATG | TA- | |
| COX-3 | H | 8,779 | 9,564 | 786 | ATG | TAA | 2 |
| $tRNA^{Gly}$ | H | 9,567 | 9,636 | 70 | | | |
| ND3 | H | 9,637 | 9,987 | 351 | ATG | TAA | 3 |
| $tRNA^{Arg}$ | H | 9,991 | 10,060 | 70 | | | |
| ND4L | H | 10,061 | 10,357 | 297 | ATG | TAA | -7 |
| ND4 | H | 10,351 | 11,731 | 1,381 | ATG | T-- | |
| tRNAHis | H | 11,732 | 11,800 | 69 | | | |
| $tRNA^{Ser\ (GCU)}$ | H | 11,801 | 11,867 | 67 | | | |
| $tRNA^{Leu\ (UAG)}$ | H | 11,868 | 11,939 | 72 | | | |
| ND5 | H | 11,940 | 13,772 | 1,833 | ATG | TAA | -4 |
| ND6 | L | 13,769 | 14,290 | 522 | ATG | TAG | |
| $tRNA^{Glu}$ | L | 14,291 | 14,360 | 70 | | | 4 |
| Cyt *b* | H | 14,365 | 15,510 | 1,146 | ATG | TAA | 1 |
| $tRNA^{Thr}$ | H | 15,512 | 15,584 | 73 | | | 2 |
| $tRNA^{Pro}$ | L | 15,587 | 15,655 | 69 | | | |
| CR | - | 15,656 | 16,730 | 1,075 | | | |

*Note*. H = heavy strand; L = light strand; Intergenetic space = positive values are spacer DNA, negative values are overlapping genes; OL = origin of light strand replication.

**Table A11**

*Annotation of GS510, Somniosus antarcticus mitogenome*

| GS510 *Somniosus antarcticus* | | | | | | | |
|---|---|---|---|---|---|---|---|
| Gene | Strands | Gene | | | | | Intergenetic space |
| | | Start (bp) | Stop (bp) | Size (bp) | Start codon | Stop codon | |
| $tRNA^{Phe}$ | H | 1 | 69 | 69 | | | |
| 12S rRNA | H | 70 | 1,020 | 951 | TGT | | |
| $tRNA^{Val}$ | H | 1,021 | 1,092 | 72 | | | |
| 16S | H | 1,093 | 2,767 | 1,675 | ACA | | |
| $tRNA^{Leu\ (UAA)}$ | H | 2,768 | 2,842 | 75 | | | |
| ND1 | H | 2,843 | 3,817 | 975 | ATG | TAA | 2 |
| $tRNA^{Ile}$ | H | 3,820 | 3,889 | 70 | | | 1 |
| $tRNA^{Gln}$ | L | 3,891 | 3,962 | 72 | | | |
| $tRNA^{Met}$ | H | 3,963 | 4,031 | 69 | | | |
| ND2 | H | 4,032 | 5,077 | 1,046 | ATG | TA- | |
| $tRNA^{Trp}$ | H | 5,078 | 5,146 | 69 | | | 1 |
| $tRNA^{Ala}$ | L | 5,148 | 5,216 | 69 | | | |
| $tRNA^{Asn}$ | L | 5,217 | 5,290 | 74 | | | |
| OL | - | 5,291 | 5,327 | 37 | | | |
| $tRNA^{Cys}$ | L | 5,328 | 5,394 | 67 | | | 1 |
| $tRNA^{Tyr}$ | L | 5,396 | 5,465 | 70 | | | 1 |
| COX-1 | H | 5,467 | 7,023 | 1,557 | GTG | TAA | |
| $tRNA^{Ser\ (UGA)}$ | L | 7,024 | 7,094 | 71 | | | 3 |
| $tRNA^{Asp}$ | H | 7,098 | 7,167 | 70 | | | 4 |
| COX-2 | H | 7,172 | 7,862 | 691 | ATG | T-- | |
| $tRNA^{Lys}$ | H | 7,863 | 7,936 | 74 | | | 1 |
| ATP8 | H | 7,938 | 8,105 | 168 | ATG | TAA | -10 |
| ATP6 | H | 8,096 | 8,778 | 683 | ATG | TA- | |
| COX-3 | H | 8,779 | 9,564 | 786 | ATG | TAA | 2 |
| $tRNA^{Gly}$ | H | 9,567 | 9,636 | 70 | | | |
| ND3 | H | 9,637 | 9,987 | 351 | ATG | TAG | 3 |
| $tRNA^{Arg}$ | H | 9,991 | 10,060 | 70 | | | |
| ND4L | H | 10,061 | 10,357 | 297 | ATG | TAA | -7 |
| ND4 | H | 10,351 | 11,731 | 1,381 | ATG | T-- | |
| tRNAHis | H | 11,732 | 11,800 | 69 | | | |
| $tRNA^{Ser\ (GCU)}$ | H | 11,801 | 11,867 | 67 | | | |
| $tRNA^{Leu\ (UAG)}$ | H | 11,868 | 11,939 | 72 | | | |
| ND5 | H | 11,940 | 13,772 | 1,833 | ATG | TAA | -4 |
| ND6 | L | 13,769 | 14,290 | 522 | ATG | TAG | |
| $tRNA^{Glu}$ | L | 14,291 | 14,360 | 70 | | | 4 |
| Cyt *b* | H | 14,365 | 15,510 | 1,146 | ATG | TAA | 1 |
| $tRNA^{Thr}$ | H | 15,512 | 15,584 | 73 | | | 2 |
| $tRNA^{Pro}$ | L | 15,587 | 15,655 | 69 | | | |
| CR | - | 15,656 | 16,730 | 1,075 | | | |

*Note*. H = heavy strand; L = light strand; Intergenetic space = positive values are spacer DNA, negative values are overlapping genes; OL = origin of light strand replication.

**Table A12**

*Annotation of GS516 and GS514, Somniosus antarcticus mitogenomes*

| GS513, GS514 *Somniosus antarcticus* | | | | | | | |
|---|---|---|---|---|---|---|---|
| Gene | Strands | Gene | | | | | Intergenetic space |
| | | Start (bp) | Stop (bp) | Size (bp) | Start codon | Stop codon | |
| $tRNA^{Phe}$ | H | 1 | 69 | 69 | | | |
| 12S rRNA | H | 70 | 1,020 | 951 | TGT | | |
| $tRNA^{Val}$ | H | 1,021 | 1,092 | 72 | | | |
| 16S | H | 1,093 | 2,767 | 1,675 | ACA | | |
| $tRNA^{Leu\ (UAA)}$ | H | 2,768 | 2,842 | 75 | | | |
| ND1 | H | 2,843 | 3,817 | 975 | ATG | TAA | 2 |
| $tRNA^{Ile}$ | H | 3,820 | 3,889 | 70 | | | 1 |
| $tRNA^{Gln}$ | L | 3,891 | 3,962 | 72 | | | |
| $tRNA^{Met}$ | H | 3,963 | 4,031 | 69 | | | |
| ND2 | H | 4,032 | 5,077 | 1,046 | ATG | TA- | |
| $tRNA^{Trp}$ | H | 5,078 | 5,146 | 69 | | | 1 |
| $tRNA^{Ala}$ | L | 5,148 | 5,216 | 69 | | | |
| $tRNA^{Asn}$ | L | 5,217 | 5,290 | 74 | | | |
| OL | - | 5,291 | 5,327 | 37 | | | |
| $tRNA^{Cys}$ | L | 5,328 | 5,394 | 67 | | | 1 |
| $tRNA^{Tyr}$ | L | 5,396 | 5,465 | 70 | | | 1 |
| COX-1 | H | 5,467 | 7,023 | 1,557 | GTG | TAA | |
| $tRNA^{Ser\ (UGA)}$ | L | 7,024 | 7,094 | 71 | | | 3 |
| $tRNA^{Asp}$ | H | 7,098 | 7,167 | 70 | | | 4 |
| COX-2 | H | 7,172 | 7,862 | 691 | ATG | T-- | |
| $tRNA^{Lys}$ | H | 7,863 | 7,936 | 74 | | | 1 |
| ATP8 | H | 7,938 | 8,105 | 168 | ATG | TAA | -10 |
| ATP6 | H | 8,096 | 8,778 | 683 | ATG | TA- | |
| COX-3 | H | 8,779 | 9,564 | 786 | ATG | TAA | 2 |
| $tRNA^{Gly}$ | H | 9,567 | 9,636 | 70 | | | |
| ND3 | H | 9,637 | 9,987 | 351 | ATG | TAA | 3 |
| $tRNA^{Arg}$ | H | 9,991 | 10,060 | 70 | | | |
| ND4L | H | 10,061 | 10,357 | 297 | ATG | TAA | -7 |
| ND4 | H | 10,351 | 11,731 | 1,381 | ATG | T-- | |
| tRNAHis | H | 11,732 | 11,800 | 69 | | | |
| $tRNA^{Ser\ (GCU)}$ | H | 11,801 | 11,867 | 67 | | | |
| $tRNA^{Leu\ (UAG)}$ | H | 11,868 | 11,939 | 72 | | | |
| ND5 | H | 11,940 | 13,772 | 1,833 | ATG | TAA | -4 |
| ND6 | L | 13,769 | 14,290 | 522 | ATG | TAG | |
| $tRNA^{Glu}$ | L | 14,291 | 14,360 | 70 | | | 4 |
| Cyt *b* | H | 14,365 | 15,510 | 1,146 | ATG | TAA | 1 |
| $tRNA^{Thr}$ | H | 15,512 | 15,584 | 73 | | | 2 |
| $tRNA^{Pro}$ | L | 15,587 | 15,655 | 69 | | | |
| CR | - | 15,656 | 16,730 | 1,075 | | | |

*Note*. H = heavy strand; L = light strand; Intergenetic space = positive values are spacer DNA, negative values are overlapping genes; OL = origin of light strand replication.

**Table A13**

*jModelTest2.1.10 model selection and partition scheme overview*

| Alignment (G=4) | AIC | BIC | DT | Model selected | Gamma | Kappa | Ra | Rb | Rc | Rd | Re | Rf |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Full mitogenome** | GTR+I+G | GTR+G | GTR+G | GTR+G | 0.21 | N/A | 2.265 | 11.278 | 1.691 | 0.145 | 15.826 | 1.000 |
| **COX-1** | GTR+G | HKY+G | HKY+G | HKY+G | 0.15 | 9.15 | N/A | N/A | N/A | N/A | N/A | N/A |
| **Cyt b** | GTR+G | HKY+G | HKY+G | HKY+G | 0.13 | 39.0 | N/A | N/A | N/A | N/A | N/A | N/A |
| **CR** | GTR+G | HKY+G | HKY+G | HKY+G | 0.47 | 2.90 | N/A | N/A | N/A | N/A | N/A | N/A |



*Note*. The model was selected based on the which model had the most support from the three criteria (AIC, BIC, and DT); AIC = Akaike Information Criteria; BIC = Bayesian Inference Criteria; DT = Decision Tree; Ra = rate AC; Rb = rate AG; Rc = rate AT; Rd = rate CG; Re = rate CT; Rf = rate GT.

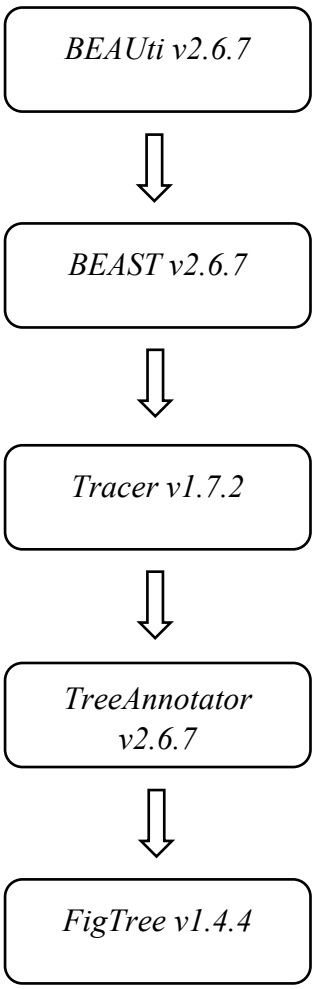



**Figure A3**

*BEAST v2.6.7 Workflow*

*Note.* BEAUti, BEAST, and TreeAnnotator are all a part of the BEAST package.

**Table A14**

*Overview of calibrations used for dating phylogenies*

| Calibration point | Type | Age (Ma) | Distribution | Parameters | Bounds (Ma) | References |
|---|---|---|---|---|---|---|
| ***S. Somniosus* / *S. Rhinoscymnus* split** | Fossil | ≈ 30; ≈ 28 | Exponential | Offset = 30<br>Mean = 0.96 | Max age: 33.9 (beginning of Oligocene)<br>Min age: 30 | Cohen et al., 2013<br><br>Welton & Goedert, 2016 |
| ***Dalatiidae* / *Etmopteridae* & *Somniosidae* split** | Fossil tip- and node-dating | 131.91 | Normal | Mean = 131.91<br>SD = 7 | Max age: 152.6<br>Min age: 110.74 | Flammensbeck et al., 2018 |

*Note.* Ma = millions of years before present; SD = standard deviation.

**Table A15**

*Site variation between alignments and analyses***.**

| **Alignment** | ***Somniosus*** | | | **(*S.*) *Somniosus*** | | | **(*S.*) *Rhinoscymnus*** | | |
|---|---|---|---|---|---|---|---|---|---|
| **Analysis** | Full | Concat. | Concatenated coverage (%) | Full | Concat. | Concatenated coverage (%) | Full | Concat. | Concatenated coverage (%) |
| **Length** | 16,733 | 2,703 | 16.15% | 16,731 | 2,703 | 16.16 | 16,727 | 2,703 | 16.16 |
| **Conserved** | 16,000 | 2,550 | 15.94% | 16,500 | 2,648 | 16.05 | 16,708 | 2,699 | 16.15 |
| **Variable** | 732 | 153 | 20.90% | 230 | 55 | 23.91 | 19 | 4 | 0.21 |
| **Pars-info.** | 664 | 135 | 20.33% | 164 | 37 | 22.56 | 0 | 0 | 0 |

*Note.* Three alignments are shown: genus *Somniosus*, and subgenera (*S.*) *Somniosus* and (*S.*) *Rhinoscymnus*; Conserved, variable, and parsimony-informative site rates are shown.; Full = mitogenomic analysis; Concat = short-sequence analysis; Concatenated coverage = percentage of the mitogenomic alignment that the short-sequence alignment covers; Conserved = number of conserved sites; Variable = number of variable sites; Pars-info. = number of parsimony-informative sites.

## Table A16

*Between-sample p-distances (Mitogenomic analysis)*

| | **GS60** | **GS80** | **GS159** | **GS500** | **GS502** | **GS505** | **GS510** | **GS513** | **GS514** | **GS216** | **GS218** | **GS516** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **GS60 *S. microcephalus*** | | | | | | | | | | | | |
| **GS80 *S. microcephalus*** | 0.001317 | | | | | | | | | | | |
| **GS159 *S. microcephalus*** | 0.000958 | 0.001317 | | | | | | | | | | |
| **GS500 *S. pacificus*** | 0.009305 | 0.009853 | 0.009487 | | | | | | | | | |
| **GS502 *S. pacificus*** | 0.009061 | 0.009610 | 0.009365 | 0.004145 | | | | | | | | |
| **GS505 *S. pacificus*** | 0.009062 | 0.009365 | 0.009121 | 0.001977 | 0.003481 | | | | | | | |
| **GS510 *S. antarcticus*** | 0.008695 | 0.009243 | 0.008999 | 0.003903 | 0.001557 | 0.003120 | | | | | | |
| **GS513 *S. antarcticus*** | 0.008393 | 0.008940 | 0.008696 | 0.003844 | 0.001318 | 0.003060 | 0.000658 | | | | | |
| **GS514 *S. antarcticus*** | 0.008941 | 0.009366 | 0.009000 | 0.001137 | 0.003602 | 0.001437 | 0.003361 | 0.003301 | | | | |
| **GS216 *S. rostratus*** | 0.034511 | 0.034826 | 0.034568 | 0.035160 | 0.034381 | 0.034513 | 0.034449 | 0.034264 | 0.034646 | | | |
| **GS218 *S. rostratus*** | 0.036276 | 0.036591 | 0.036332 | 0.036991 | 0.036207 | 0.036342 | 0.036278 | 0.036158 | 0.036476 | 0.000360 | | |
| **GS516 *S. longus*** | 0.036857 | 0.037172 | 0.036914 | 0.037573 | 0.036788 | 0.036924 | 0.036859 | 0.036740 | 0.037058 | 0.000899 | 0.001018 | |

*Note.* P-distance is the proportion (p) of nucleotide sites at which two sequences being compared are different.

## Table A17

*Between-sample p-distances (short-sequence analysis)*

| | GS60 | GS80 | GS159 | GS500 | GS502 | GS505 | GS510 | GS513 | GS514 | GS216 | GS218 | GS516 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **GS60 *S. microcephalus*** | | | | | | | | | | | | |
| **GS80 *S. microcephalus*** | 0.001854 | | | | | | | | | | | |
| **GS159 *S. microcephalus*** | 0.000741 | 0.001854 | | | | | | | | | | |
| **GS500 *S. pacificus*** | 0.013150 | 0.014670 | 0.013525 | | | | | | | | | |
| **GS502 *S. pacificus*** | 0.013525 | 0.015045 | 0.013900 | 0.005210 | | | | | | | | |
| **GS505 *S. pacificus*** | 0.013913 | 0.015434 | 0.014288 | 0.003342 | 0.005585 | | | | | | | |
| **GS510 *S. antarcticus*** | 0.013525 | 0.015045 | 0.013900 | 0.004462 | 0.001482 | 0.004836 | | | | | | |
| **GS513 *S. antarcticus*** | 0.013150 | 0.014669 | 0.013525 | 0.004838 | 0.001112 | 0.005212 | 0.001112 | | | | | |
| **GS514 *S. antarcticus*** | 0.012389 | 0.013906 | 0.012764 | 0.001854 | 0.004089 | 0.002225 | 0.003343 | 0.003717 | | | | |
| **GS216 *S. rostratus*** | 0.042912 | 0.043708 | 0.043300 | 0.044543 | 0.041264 | 0.044545 | 0.041264 | 0.041688 | 0.043726 | | | |
| **GS218 *S. rostratus*** | 0.045247 | 0.046044 | 0.045635 | 0.047292 | 0.044409 | 0.047293 | 0.044409 | 0.044836 | 0.046472 | 0.000000 | | |
| **GS516 *S. longus*** | 0.046885 | 0.047683 | 0.047273 | 0.048936 | 0.046043 | 0.048938 | 0.046043 | 0.046472 | 0.048114 | 0.001487 | 0.001482 | |

*Note.* P-distance is the proportion (p) of nucleotide sites at which two sequences being compared are different.

**Table A18**

*Interspecific p-distances (mitogenomic analysis)*

| | ***S. microcephalus*** | | ***S. pacificus*** | | ***S. antarcticus*** | | ***S. rostratus*** | | ***S. longus*** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| ***S. microcephalus*** | | | | | | | | | | |
| ***S. pacificus*** | 0.009259 | 0.000245 | | | | | | | | |
| ***S. antarcticus*** | 0.008828 | 0.000273 | 0.002544 | 0.001096 | | | | | | |
| ***S. rostratus*** | 0.034113 | 0.000820 | 0.034186 | 0.000896 | 0.033976 | 0.000860 | | | | |
| ***S. longus*** | 0.035461 | 0.000129 | 0.035563 | 0.000313 | 0.035364 | 0.000121 | 0.000958 | 0.000059 | | |

*Note*. P-distance is the proportion (p) of nucleotide sites at which two sequences being compared are different; SD = standard deviation.

**Table A19**

*Interspecific p-distances (short-sequence analysis)*

| | ***S. microcephalus*** | | ***S. pacificus*** | | ***S. antarcticus*** | | ***S. rostratus*** | | ***S. longus*** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| ***S. microcephalus*** | | | | | | | | | | |
| ***S. pacificus*** | 0.013935 | 0.000698 | | | | | | | | |
| ***S. antarcticus*** | 0.013442 | 0.000780 | 0.003330 | 0.001540 | | | | | | |
| ***S. rostratus*** | 0.042244 | 0.001084 | 0.042614 | 0.001839 | 0.041564 | 0.001619 | | | | |
| ***S. longus*** | 0.044765 | 0.000302 | 0.045382 | 0.001221 | 0.044395 | 0.000799 | 0.001482 | 0.000002 | | |

*Note*. P-distance is the proportion (p) of nucleotide sites at which two sequences being compared are different; SD = standard deviation.

**Table A20**

*Inflation of in interspecific p-distance of short-sequence analysis*

| | ***S. microcephalus*** | ***S. pacificus*** | ***S. antarcticus*** | ***S. rostratus*** | ***S. longus*** |
|---|---|---|---|---|---|
| ***S. microcephalus*** | | | | | |
| ***S. pacificus*** | 66.4454194 | | | | |
| ***S. antarcticus*** | 65.6717769 | 76.4097889 | | | |
| ***S. rostratus*** | 80.7532175 | 80.2219305 | 81.743834 | | |
| ***S. longus*** | 79.2154368 | 78.364982 | 79.6574114 | 64.596399 | |

*Note*. The percent increase in p-distance value of the short-sequence analysis over the mitogenome analysis is shown.

**Table A21**

*Intraspecific p-distances*

| | **Full mitogenome** | | **Short-sequence** | |
|---|---|---|---|---|
| | Mean | SD | Mean | SD |
| ***S. microcephalus*** | 0.001196 | 0.000169 | 0.001480 | 0.000523 |
| ***S. pacificus*** | 0.003189 | 0.000901 | 0.004686 | 0.000971 |
| ***S. antarcticus*** | 0.002431 | 0.001254 | 0.002713 | 0.001144 |
| ***S. rostratus*** | 0.000359 | 0.000000 | 0.000000 | 0.000000 |
| ***S. longus*** | N/A | N/A | N/A | N/A |

*Note*. P-distance is the proportion (p) of nucleotide sites at which two sequences being compared are different; SD = standard deviation.