7-1-2014

A Mitogenomics View of the Population Structure and Evolutionary History of the Basking Shark Cetorhinus maximum

Kimberly A. Finnegan
Nova Southeastern University, atwaterk5@gmail.com

NOVA SOUTHEASTERN UNIVERSITY OCEANOGRAPHIC CENTER

A MITOGENOMICS VIEW OF THE POPULATION STRUCTURE AND EVOLUTIONARY HISTORY OF THE BASKING SHARK, *Cetorhinus maximus*

By

Kimberly A. Finnegan

Submitted to the Faculty of
Nova Southeastern University Oceanographic Center
in partial fulfillment of the requirements for
the degree of Master of Science with a specialty in:

Marine Biology

Nova Southeastern University

July 2014

Thesis of
Kimberly A. Finnegan

Submitted in Partial Fulfillment of the Requirements for the Degree of

Masters of Science:

Marine Biology

Nova Southeastern University
Oceanographic Center

July 2014

Approved:

Thesis Committee

Major Professor :_____
Mahmood S. Shivji Ph.D.

Committee Member :_____
George T. Duncan Ph.D.

Committee Member :_____
Jose V. Lopez Ph.D.

Table of Contents **Page**

A Mitogenomics View of the Population Structure and Evolutionary History of the Basking shark, *Cetorhinus maximus*

Abstract

The basking shark, *Cetorhinus maximus*, has historically been a target of international fisheries, leading to well-documented declines in parts of its global distribution. Currently, the basking shark is listed as globally 'Vulnerable' and regionally 'Endangered' (North Pacific and Northeast Atlantic) on the IUCN Red List of Threatened Species, rendering the species an international conservation priority. Here, we assessed the global matrilineal genetic population structure and evolutionary history of the basking shark by completing the first whole mitochondrial genome sequence level survey of animals sampled from three globally widespread geographic regions: the western North Atlantic (n = 11), the eastern North Atlantic (n = 11), and within New Zealand territorial waters (n = 12). Despite the relatively large amount of sequence data assessed (~16,669 bp per individual), whole mitogenome analyses showed no evidence of population differentiation (Φ_{ST} = -0.047, P > 0.05) and very low nucleotide diversity (π = 0.0014 ± 0.000) across a global seascape. The absence of population structure across large distances and even between ocean basins is indicative of long-dispersal by this species, including an ability to cross known biogeographic barriers known to differentiate populations of other highly vagile pelagic fishes. Notably, evolutionary analyses of the mitogenome sequences revealed two globally sympatric but evolutionary divergent lineages, with a Bayesian framework estimated coalescence time of ~2.46 million years ago. Coalescent-based Bayesian skyline analysis uncovered subtle evidence of

Pleistocene demographic flux for this species, including a potential decline in female effective population size. Thus, historical population changes may be responsible for the occurrence of the two highly divergent, yet sympatric lineages, as population declines may have resulted in the loss of intermediate haplotypes and resulted in an overall loss of genetic diversity. This work supports the recognition of basking sharks as a single matrilineal global population, and as such requires the application of a cooperative multiagency and international approach to fisheries management to conserve this highly vulnerable and ecologically unique species.

Keywords: whole mitochondrial genome; *Cetorhinus maximus*; sympatric lineages; genetic diversity; conservation

Introduction

Scientific advances in DNA sequencing technology have allowed for the efficient sequencing of whole mitochondrial genomes (mitogenomes), thus increasing the utility of these whole molecules as non-recombining, fast-evolving, population genetic and phylogeographic markers (Ma *et al*. 2012; Teacher *et al.* 2012; Winklemann *et al*. 2013). The use of whole mitogenome sequences to infer population structure and evolutionary relationships within species may help to eliminate a variety of the limitations inherent in sequencing shorter, single mitochondrial regions (e.g., low diversity, substitutionally constrained genes, and variable mutation rates among regions), allowing for more accurate and intraspecific relationships to be inferred (Subramanian *et al.* 2009; Morin *et al.* 2010; Knaus *et al.* 2011). In fact, numerous studies suggest that increasing the number of surveyed regions and/or the length of resolved sequences, may be more beneficial than simply increasing the number of individuals sequenced at a single locus when attempting to resolve population structure and increase phylogenetic accuracy (Saitou & Nei 1986; Ruvolo *et al.* 1991; Cummings *et al.* 1997; DeFilippis & Moore 2000; Rokas & Carroll 2005; Morin *et al.* 2010). As such, researchers have begun adopting whole mitogenomic sequences to resolve the population genetic structure and evolutionary divergences within and among species with notable success (McGowen *et al.* 2009; Xiong *et al.* 2009; Vilstrup *et al.* 2011; Jacobsen *et al.* 2014; Karlsen *et al.* 2014). For instance, in some cases where previous single mitochondrial region/gene studies have found little genetic heterogeneity among individuals (Hoelzel *et al.* 2002; LeDuc *et al.* 2008), use of the entire mitogenome has yielded increased statistical and phylogenetic resolution and in

some cases has identified the presence of highly divergent historical lineages, suggesting the potential presence of species complexes or distinct sub-species (Morin *et al.* 2010). Such applications underscore the potential utility of the mitogenome as a powerful genetic marker, particularly for species where past studies have demonstrated poor resolution using fewer or shorter-length genetic markers.

The basking shark, *Cetorhinus maximus*, is globally distributed and is the world's second largest fish species, attaining a maximum length of up to 12 meters (Castro 2011). With little known about this enigmatic species, its massive size and elusive nature has historically inspired numerous accounts of mythical "sea monster" sightings. When the remains of a decomposing basking shark was discovered onshore in 1808, its peculiar appearance, particularly its large body size and gill morphology, spawned the tale of Scotland's Strongsay Beast; more recently, in 1977 a Japanese fishing vessel claimed discovery of a plesiosaur or prehistoric "sea-serpent" when their trawler encountered the decayed remains of what was ultimately determined to be a basking shark (Kuban 1997; Towerie 2014). While such bizarre encounters inspire the imagination of many, human-basking shark interactions have historically been much more mundane, typically consisting of intensive targeted and non-targeted fishing.

While historically abundant, widespread overexploitation of basking sharks has led to severe global declines. Harpoon fisheries within the eastern North Atlantic selectively targeted basking sharks for their meat and liver oil through World War II. Hundreds, if not thousands, of basking sharks were killed from 1955 to 1969 within the Canadian Pacific Ocean as part of a basking shark eradication program aimed at reducing their interference with salmon gillnet and trawling fishing gear (Parker & Stott 1965;

Went & Súilleabháin 1967; COSEWIC 2007). Within the western South Pacific, similar declines have been noted due to extensive bycatch that occurred during the 1980s and 1990s by deep-water trawl and setnet fisheries (Francis & Duffy 2002; Francis & Smith 2010; Francis & Sutton 2012). In response to severe declines in basking shark abundance (Kunlik 1988, CITES 2002), fishery closures were implement to curb additional drastic decreases in numbers (Musik *et al.* 2000). However, despite numerous closures, basking sharks remain at high risk of global overexploitation due to the highly lucrative international fin trade; the sale of a single, large basking shark fin can fetch up to $57, 000 USD within East Asian Markets (Clarke 2004; Magnussen *et al.* 2007). Consequently, the IUCN Red List of Threatened Species has listed the basking shark as globally "Vulnerable" and regionally "Endangered" in both the eastern North Pacific and eastern North Atlantic (Fowler 2005). In addition, tighter restrictions on the international trade of basking sharks has also occurred, with their addition in 2003 to Appendix II of the Convention on International Trade in Endangered Species (CITES) of Wild Fauna and Flora (CITES 2002).

In addition to harvest closures and trade restrictions, proper management and conservation of exploited species requires a thorough understanding of their genetic connectivity and population dynamics (von der Heyden *et al.* 2014). Despite its high profile nature and precarious conservation status, limited such information exists for basking sharks. To date only two studies have examined the genetic basis of connectivity in basking sharks. Hoelzel (2001) in a preliminary genetic survey of only 17 animals found no population structure based on the mitochondrial protein coding gene, Cytochrome *b*. In a subsequent study using the non-coding mitochondrial control region

(CR; 1085 bp, $n = 62$), Hoelzel *et al.* (2006) again found an absence of significant genetic population structure among eight globally distributed basking shark sample sites (non-significant and negative Φ_{ST}), as well as extremely low levels of genetic diversity ($\pi = 0.0013 \pm 0.0009$). In contrast to the genetic homogeneity found across the basking shark's distribution, a CR genetic survey of the population genetic structure of the whale shark (*Rhincodon typus*), another globally distributed and highly migratory filter-feeding shark, revealed significant genetic differentiation between western Atlantic and Indo-Pacific ocean individuals ($\Phi_{ST} = 0.107$, $P < 0.001$) (Castro 2007), as well as much higher levels of genetic diversity ($\pi = 0.0110 \pm 0.006$). The contrast between these two very large, pelagic, highly vagile, filter-feeding sharks highlights the possible uniqueness of the basking shark's apparent high global genetic connectivity and low mitochondrial genetic diversity. In fact, to date, no other globally distributed elasmobranch has demonstrated such high levels of matrilineal genetic connectivity and low diversity across such vast spatial scales.

While inter-ocean basin movements have yet to be documented for the basking shark, spatial movement data is consistent with potentially high connectivity across large spatial scales. For examples, pop-up satellite archival tag (PSAT) data have documented several migrations exceeding 5, 000 kilometers (km) (Gore *et al.* 2008; Johnston 2014), including the trans-Atlantic migration (~ 9, 500 km) of a basking shark originally tagged within the British Isles travelling to the coastal waters of Newfoundland, Canada (Gore *et al.* 2008). Moreover, Californian researchers documented the passage of a basking shark across the East Pacific Barrier, as a single individual travelled a total distance of ~ 4, 000 km from the waters of the U.S. West Coast to the island of Hawai'i (Lee 2012). In

addition to large distances, recent work has also shown that basking sharks are capable of crossing what was presumed to be a likely thermal barrier for this species due to basking shark's preference for cooler temperate waters (Castro 2011). A study examining the movement of basking sharks within the western North Atlantic discovered that during winter months PSAT tagged basking sharks travelled at mesopelagic depths (200-1000 meters) while crossing the tropical waters of the equator (Skomal *et al.* 2009). These extensive movements indicate that presumed biological barriers present for many other smaller shark species (thermal waters, deep ocean depths) appear to have little effect on the migration patterns and lifecycles of basking sharks.

In the light of recent tracking efforts revealing the broad spatial extent of basking shark oceanic movements (Sims & Quayle 1998; Gore *et al.* 2008; Skomal *et al.* 2009), it is not entirely surprising that Hoelzel et al. (2006) found little genetic differentiation among global collections. However, this study only assayed genetic variation present at a single mitochondrial region (i.e. CR), and the question of whether basking sharks truly represent a single, globally panmictic population remains open. Thus to investigate this species' global genetic population structure and phylogeography, further examination using more data is warranted. Herein, whole mitogenome sequences are used to explore numerous key uncertainties concerning the basking shark's global genetic connectivity, including: (1) the identification of its global genetic population structure; (2) the relative signal of genetic diversity among different mitochondrial regions; and (3) the examination of the demographic history of this unique oceanic species.

Materials and Methods

Sampling locations

Basking shark tissue samples (fin or muscle) were collected from by-caught, stranded, or free-ranging individuals within the western North Atlantic (WNA), eastern North Atlantic (ENA), and New Zealand territorial waters (NZ) (Figure 1). All samples were stored in 95% ethanol prior to genomic DNA extraction. Genomic DNA was extracted and purified from ~25mg of tissue using the QIAgen DNeasy® Blood and Tissue Kit (Qiagen Inc., Valencia, CA).

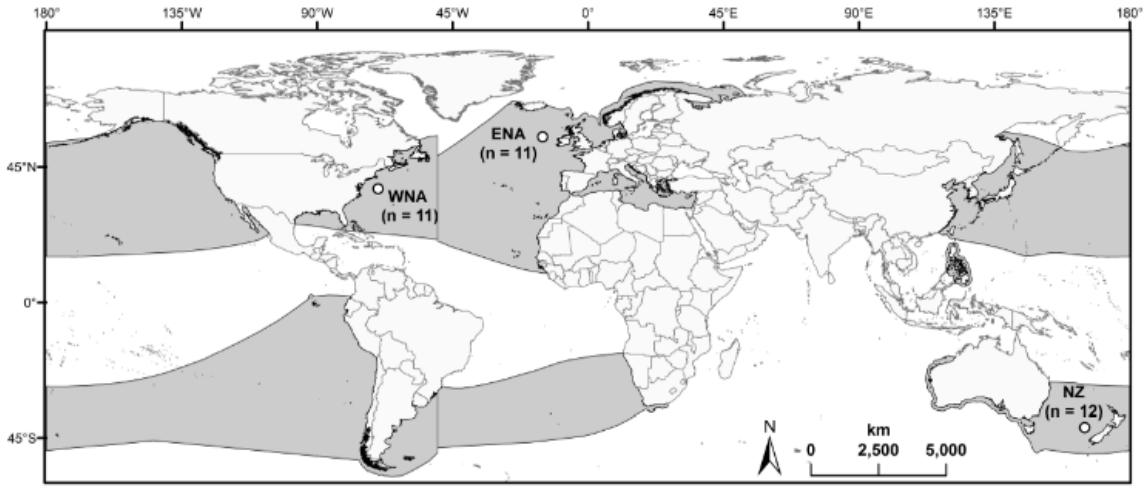


Figure 1. Map depicting the global geographic distribution (grey shaded regions) of the basking shark, *Cetorhinus maximus* (adapted from IUCN 2013) as well as sample collection sites and sizes. Sample sizes refer to the total number of complete mitogenomes.

Mitogenome Sequencing and Assembly

The complete mitochondrial genomes (mitogenomes) of 34 basking sharks (GenBank Accession Numbers: KM096969 – KM096989) (WNA, n = 11; NZ, n =12; ENA, n =11) were amplified via polymerase chain reaction (PCR) in overlapping segments (500 bp – 3.5 kbp), initially utilizing 120 primers designed from a previously published *Mitsukuina owstoni* reference mitogenome (GenBank accession number NC_011825; goblin shark reference) (Appendix A) and two universal 16S rRNA primers (Palumbi *et al.* 1991). An additional 28 basking shark-specific PCR primers were designed using the software Primer3 (Koressaar & Remm 2007; Untergrasser *et al.* 2012) to close sequence gaps between contigs and to achieve double-stranded DNA sequence coverage of the entire mitogenomes (Appendix B).

All PCR amplifications were performed using touchdown cycling conditions (Don *et al.* 1991) in a total PCR reaction volume of 50 ul containing: 1uL of unquantified DNA, 0.2mM of each dNTP (GE Healthcare Inc., United Kingdom), 1X Coral Load PCR buffer (Qiagen Inc.), 1.0U of HotStar *Taq*™ DNA Polymerase (Qiagen Inc.), and 10pM of the Forward and Reverse primers. Electrophoresis of the resulting PCR products was performed using a 1.2% agarose gel to confirm both proper sizing of mitochondrial amplicons and absence of contamination in the negative controls (no genomic DNA). PCR products were purified using the QIAquick ® PCR Purification Kit (Qiagen Inc.) as per the manufacturer's protocols. All cycle sequencing reactions were carried out using the Applied Biosystems BigDye Terminator v3.1 Cycle Sequencing Kit (Life Technologies Inc., Foster City, CA) and purification of the resultant cycle sequencing products was performed using the DyeEx 2.0 Spin Kit (Qiagen Inc.). Electrophoresis of

purified cycle sequencing products was performed on a 3130 Applied Biosystems Genetic Analyzer (Life Technologies Inc.).

All contigs were manually checked and assembled using a previously published basking shark mitogenome generated in our laboratory as a comparative reference sequence (Hester *et al.* 2013) (GenBank Accession Number KF597303). DNA sequence alignments were performed using the software Geneious Pro 5.6.5 (Biomatters Inc., San Francisco, CA) (Drummond *et al.* 2012) and manually checked to ensure proper alignment.

Datasets analyzed

The complete basking shark mitogenome alignment (n = 34) was trimmed to generate four distinct datasets for analysis:

I. An alignment of complete mitogenomes (16,664 - 16,670 bp).

II. An alignment containing the complete concatenated sequences of all 13 protein-coding genes (11, 429 bp).

III. Individual alignments for specific region [i.e., each of the 13 protein-coding genes and each of the three non-protein coding regions (CR, 12S and 16S rRNA genes).

IV. Preliminary statistical analyses (AMOVA, h, hd, and π) of a subset of the surveyed *C. maximus* mitogenomes (n = 30) revealed a low level of global genetic diversity and no population structure (data not shown). Since specific regions of the mitogenome can accumulate substitutions differentially, providing incongruent evolutionary signal (Zhang et al. 2013; Meiklejohn et al. 2014), an

additional 70 *C. maximus* individuals were sequenced at three mitochondrial DNA protein-coding genes in an attempt to improve the statistical resolution of my analyses: ATP synthase protein 8, (ATP8) (168 bp); Cytochrome c oxidase subunit II, (CO2) (691 bp); and NADH dehydrogenase 3 (ND3) (349 bp). Note: these three genes were selected for additional sequencing as they possessed the highest nucleotide diversity (based on preliminary analyses); however, upon analysis of all 34 complete mitogenomes, the protein-coding genes possessing the highest nucleotide diversity had changed (see Results: "Genetic variation and population structure analyses"). Dataset IV, therefore, is comprised of the concatenated sequences of these three genes, hereafter referred to as the 'three most variable protein-coding genes' (1, 208bp) and/or Dataset IV.

Diversity and genetic differentiation analyses

Diversity indices, including: haplotype diversity (*hd*), nucleotide diversity (π), and the number of unique haplotypes (*h*), were estimated using the software DNAsp 5.10.1 (Librado & Rozas 2009) for datasets I - III. For these same datasets, the number of polymorphic sites was estimated using the program GenAlEx 6.5 (Peakall & Smouse 2006, 2012).

To investigate basking shark global population structure, a hierarchical analysis of molecular variance (AMOVA) was performed for all four datasets using Arlequin 3.5.1.3 (Excoffier & Lischer 2010) to estimate F_{ST} (Weir & Cockerham 1984) and Φ_{ST} (Excoffier 1992). Pairwise estimates of differentiation (Φ_{ST}) among *a priori* geographic sampling locations were also generated using Arlequin for datasets I - IV. Significance

values for both the AMOVA and pairwise divergences were estimated using 1,000 simulations.

Phylogenetic analyses

The evolutionary relationships among complete mitogenome haplotypes (Dataset I) were inferred by generating a statistical parsimony network using the software TCS 1.21 (Clement *et al.* 2000). To further resolve the level of genetic differentiation among resolved haplotypes, estimates of genetic distance [pairwise uncorrected *p* and Kimura 2-parameter (K2P)] (Nei & Kumar 2000) were calculated in MEGA 5.0 (Tamura *et al.* 2011) for datasets I and III.

The most suitable model of nucleotide evolution for Dataset II was determined using the software jModelTest 2.1.4 (Guindon & Gascuel 2003; Darriba *et al*. 2012). Contingent on assumptions of the downstream analysis, one of two methods was utilized to select the most appropriate model of evolution: the Akaike Information Criterion (AIC) (Akaike 1974; Burnham & Anderson 2002) for use in a maximum likelihood (ML) analysis framework or the Bayesian Information Criterion (BIC) (Schwartz 1978) for use in a Bayesian analysis framework.

To further assess the evolutionary relationships among surveyed *C. maximus* haplotypes, maximum likelihood (ML) and Bayesian inference phylogenetic trees were constructed using the 13 concatenated protein-coding genes (Dataset II). An Unrooted ML tree was constructed using the software plug-in PhyML 3.0 (Burnham & Anderson 2002) for Geneious Pro. Statistical support for nodes was inferred by conducting 1000 nonparametric bootstrap replicates. A Bayesian inference (BI) haplotype tree was

constructed using MrBayes (Huelzenbeck & Ronquist 2001) executed as a plug-in feature in Geneious Pro, and rooted with the concatenated protein-coding sequence of the white shark [*Carcharodon carcharias* (NC_022415)], the closest related (Heinicke *et al.* 2009) available Lamniformes whole mitogenome sequence. BI analyses were conducted using four heated chains (default heating values) consisting of a Markov Chain Monte Carlo (MCMC) chain length of 5, 000, 000 generations; sampling of chains was performed every 100 generations, and the first 25% of each of the sampled chains were discarded as burnin. Convergence was assumed once posterior probability effective sample size (ESS) values exceeded 200.

Evolutionary and Demographic History

To test for historical demographic changes in population size, the neutrality estimators R_2 (Ramos-Onsins & Rozas 2002) and Fu's F_s (Fu 1996) were generated for Dataset I using the software DNAsp and Arlequin, respectively.

Time to Most Recent Common Ancester (TMRCA) of all *C. maximus* haplotypes was estimated using the Bayesian software BEAST 1.8 (Drummond *et al.* 2012). Analyses were performed using a partitioned dataset comprising unique *C. maximus* concatenated protein-coding gene haplotypes (Dataset II, n = 22) as well as mitogenomic concatenated 13 protein-coding sequences derived from three additional lamniform taxa [*Carcharias taurus* (KF_569943), *Carcharodon carcharias* (NC_022415), and *Isurus oxyrinchus* (NC_022691)]. Three nodes were utilized to perform temporal calibration of molecular evolution rates based on previously published phylogenetic relationships among extant lamniforms (Heinicke *et al.* 2009) and *C.*

maximus fossil evidence (Schultze 2012). The TMRCA of the basking shark was estimated assuming a lognormal prior distribution for the Cetorhinidae node, with a mean of 0.5 million years ago (Mya) and a standard deviation of 3.0 Mya. Such broad priors were employed to accommodate: (i) the potential of a recent coalescent event, and (ii) fossil evidence suggesting a Pliocene species origin (2.5- 5.3 Mya) (Schultze 2012). As per Heinicke *et al.* (2009), priors for the TMRCA of extant members of the family Lamnidae (herein: *C. carcharias, I. oxyrinchus*) were set assuming a normal distribution with a mean of 109 million years ago (Mya) and a standard deviation of 17.0 Mya. Priors for the TMRCA for all analyzed sequences (coalescence of the basal *C. taurus* with all other lamniform sequences derived) were set assuming a normal distribution with a mean of 119 Mya and a standard deviation of 17.0 Mya (Heinicke *et al.* 2009). Analyses were performed implementing a random starting tree, and assuming the following settings: the SRD06 nucleotide substitution model (Shapiro *et al.* 2006), a Bayesian relaxed clock, uncorrelated lognormal rate heterogeneity across branches, unlinked substitution models between data partitions, and a birth-death process speciation tree prior. Posterior distributions were estimated via 100, 000, 000 MCMC iterations, sampled every 10, 000 iterations. The first 25% of all samples were discarded as burn-in. Convergence was assessed using the software Tracer 1.6 (Rambaut & Drummond 2007) and was assumed to have occurred once ESS values exceeded 200.

Bayesian skyline plots (BSP) were generated in BEAST for Dataset II (concatenated protein-coding genes, $n = 34$) to detect any historical demographic changes in mean female population size (N_{Ef} x generation time). Analyses were performed assuming the SRD06 nucleotide substitution model and a strict molecular clock. The

mutation rate was fixed using the estimated mutation rate derived from the above Bayesian TMRCA analyses. Skyline plots were generated with a MCMC chain consisting of 100, 000, 000 iterations with sampling occurring every 10, 000 iterations. The first 25% of the total iterations were discarded as burn-in. Convergence was assessed using the software Tracer and assumed to have occurred once ESS values exceeded 200.

Results

Genetic variation and population structure analyses

Sequencing of 34 *C. maximus* individuals provided genome sizes of 16, 664 - 16, 670 bp. The data revealed 22 distinct haplotypes and a GC-content (guanine-cytosine content) of 40.65%. A total of 130 single nucleotide variations were detected, including seven insertion or deletions (indels) and 123 single nucleotide polymorphisms (SNPs) (Appendix C). Overall mitogenome haplotype and nucleotide diversities were 0.970 ± 0.014 and 0.0014 ± 0.000, respectively (Table 1).

Concatenation of all 13 protein-coding genes generated an alignment of 11, 429 bp and yielded 22 unique haplotypes. An identical haplotype diversity of 0.970 ± 0.14 was estimated along with a comparably low nucleotide diversity of 0.0018 ± 0.000 (Table 1). Single nucleotide variations (n = 114) for all protein-coding loci included: six indels and 108 SNPs (Appendix D).

Haplotype diversities for all individual genes [protein-coding (n = 13) and non-protein coding regions (n = 3)] ranged from a low of 0.266 ± 0.092 [ATP synthase F0 subunit 6 (ATP6); 684 bp] to high of 0.845 ± 0.047 [NADH dehydrogenase 5 (ND5); 1,

830 bp]. Estimates of nucleotide diversity spanned 0.0004 ± 0.000 (16S rRNA; 1, 666 bp) to 0.0081 ± 0.002 (ATP8; 168 bp) (Table 1).

Results from the *C. maximus* whole mitogenome AMOVA produced an overall non-significant Φ_{ST} of -0.047 (P = 0.786) and non-significant F_{ST} value of 0.013 (P =0.226), suggesting an absence of geographic partitioning of molecular variance (Table 1). AMOVAs generated for datasets II - IV demonstrated a similar lack of genetic population structure (Table 1, Appendix E). Estimated pairwise values of Φ_{ST} among the three global collections also revealed no significant differences ($P > 0.05$) upon analysis of datasets I - III (Table 2, Appendix F).

Table 1. Global summary statistics for *C. maximus* mitogenome survey including: number of individuals (*n*), sequence length in base pairs (bp), number of haplotypes (*h*), haplotype diversity (*hd*), nucleotide diversity (π), and overall ϕ_{ST} (Exocoffier 1992) and F_{ST} (Weir & Cockerham 1984) values.

Data set (*n* = 34)	bp	*h*	*hd* ± SD	π ± SD	ϕ_{ST}	F_{ST}
Whole Mitogenome	16 669	22	0.970 ± 0.014	0.0014 ± 0.000	-0.047	0.013
Concatenated Protein-Coding Genes	11 429	22	0.970 ± 0.014	0.0018 ± 0.000	-0.050	0.013
CR	1 048	5	0.710 ± 0.052	0.0013 ± 0.000	-0.027	-0.015
Cyt b	1 144	5	0.321 ± 0.100	0.0016 ± 0.001	-0.030	-0.017
ND2	1 044	3	0.399 ± 0.151	0.0012 ± 0.000	-0.022	-0.049
ND4	1 381	7	0.734 ± 0.063	0.0016 ± 0.000	0.008	0.019
ATP8	168	4	0.540 ± 0.820	0.0081 ± 0.002	-0.072	-0.040
CO2	691	9	0.676 ± 0.083	0.0026 ± 0.000	-0.056	-0.014
ND5	1 830	11	0.845 ± 0.047	0.0024 ± 0.000	-0.030	0.001
ATP6	684	3	0.266 ± 0.092	0.0016 ± 0.001	-0.072	-0.067
CO1	1 554	3	0.399 ± 0.001	0.0012 ± 0.000	-0.070	-0.020
CO3	786	7	0.458 ± 0.104	0.0018 ± 0.000	-0.053	-0.017
ND1	975	4	0.549 ± 0.085	0.0016 ± 0.000	-0.069	-0.022
ND3	349	3	0.551 ± 0.085	0.0006 ± 0.000	-0.075	0.009
ND4L	297	3	0.308 ± 0.092	0.0012 ± 0.000	-0.038	0.007
ND6	526	7	0.635 ± 0.085	0.0016 ± 0.000	-0.059	-0.025
12S	954	5	0.551 ± 0.086	0.0011 ± 0.000	-0.030	-0.010
16S	1 666	4	0.362 ± 0.099	0.00036 ± 0.000	-0.033	-0.021

Table 2. *C. maximus* population-level pairwise values of Φ_{ST} for the whole mitogenome data set (upper triangular matrix) and for the concatenated protein-coding gene data set only (lower triangular matrix).

Location	WNA	NZ	ENA
WNA	-	-0.06689	-0.04705
NZ	-0.04893	-	-0.03166
ENA	-0.06790	-0.03567	-

WNA = western North Atlantic, NZ= New Zealand, ENA= eastern North Atlantic

Phylogenetic analyses

The TCS generated statistical parsimony network for the whole mitogenome dataset failed to connect all *C. maximus* mitogenomic haplotypes at the 95% probability level; however, all were successfully joined when the statistical probability level was lowered to 90% (Figure 2). At the 90% probability level, analyses resolved two evolutionarily distinct mitogenome clades separated by 75 unsampled haplotypes. Clade I (Figure 2) was composed of three unique haplotypes generated from sequencing four individuals obtained from across the species sampling distribution (WNA, $n = 1$; NZ, $n = 2$; ENA, $n = 1$); clade II consisted of all other resolved haplotypes ($n = 19$). Despite the presence of the two evolutionarily distinct lineages, no geographic association of specific haplotypes was detected, with both lineages being sympatric in all three geographic locations sampled.

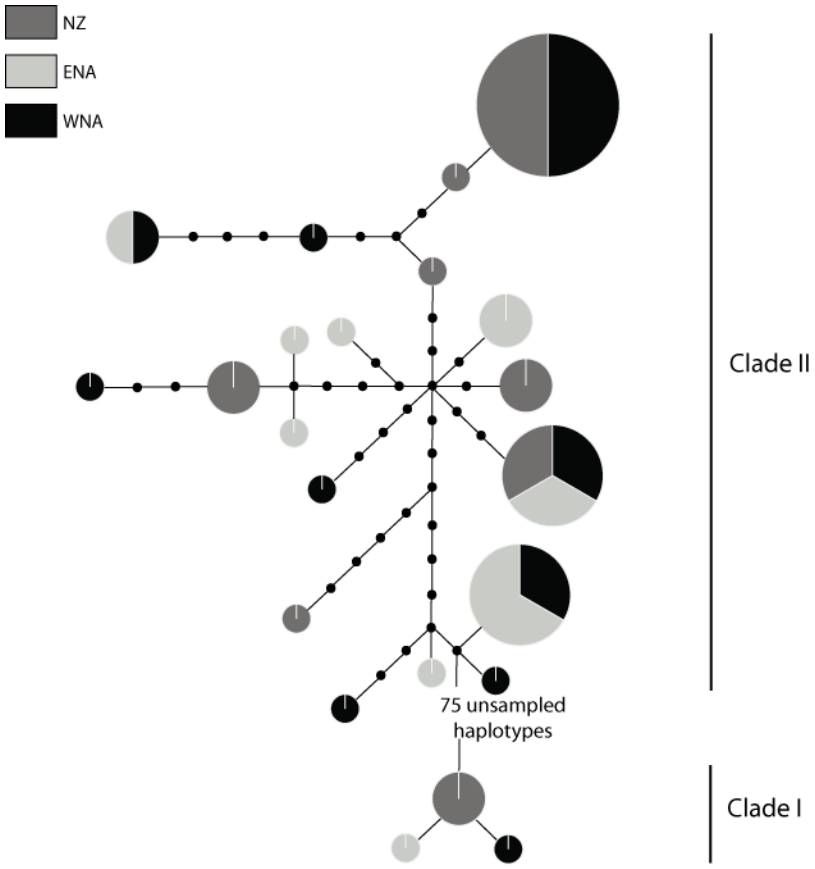


Figure 2. TCS 1.21 (Clement *et al.* 2000) statistical parsimony network depecting the relationships among whole mitogenome haplotypes joined at the 90% probability level. Circle size is proportional to haplotype frequency, black connecting lines represent single mutational steps, and small black circles (●) represent hypothetical unsampled haplotypes.

The pairwise genetic distance values (*p*, K2P) estimated among whole mitogenome haplotypes ranged from 0 – 0.5%. Individual gene/region genetic distance estimates range from 0 – 3.6 %, with ATP8 possessing the highest levels of divergence (3.6 %) and ND3 (NADH dehydrogenase 3) demonstrating the lowest level of

differentiation (0.3 %) (Data not shown). Maximum estimates of divergence (*p*) among resolved haplotypes at the four most commonly surveyed mitochondrial protein-coding loci in elasmobranchs (see Hoelzel 2001; Quattro *et al.* 2006; Naylor *et al.* 2012) were as follows: (i) CR (0.4 %), COI (0.6 %), ND2 (0.5 %), and Cyt b (0.7 %).

The most suitable model of nucleotide evolution chosen for the concatenated protein-coding gene dataset (II) was identified as the general time-reversible model with invariable sites (GTR + I), according to AIC method as implemented in jModeltest. The resultant ML tree (incorporating the GTR + I model) resolved two distinct and strongly supported (1000 bootstraps) monophyletic lineages (Figure 3), consisting of the same two clades identified with whole mitogenome TCS analysis.

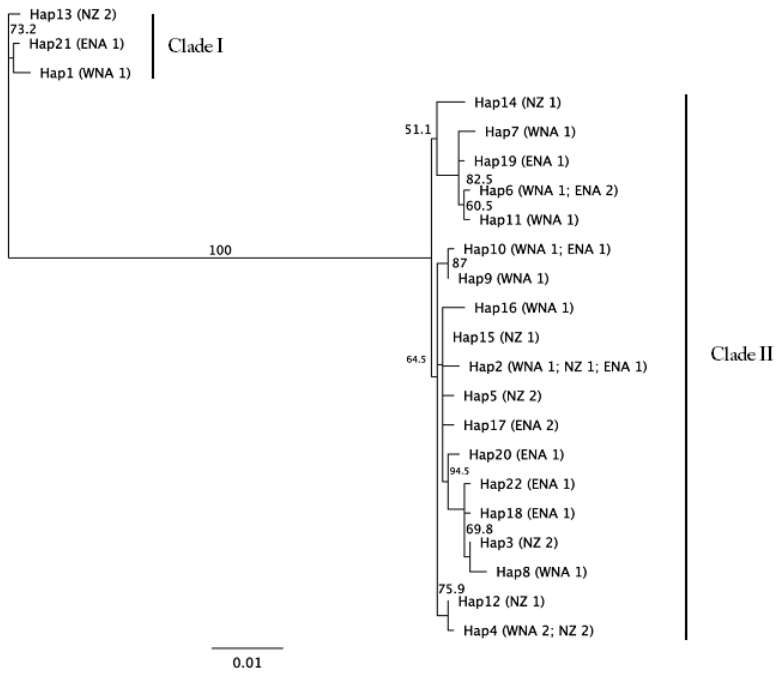


Figure 3. Unrooted Maximum Likelihood phylogenetic analysis of *C. maximus* concatenated protein-coding gene haplotypes. Bootstrap values over 50% are indicated for each node.

According to jModeltest's BIC model selection criteria, the Tamura-Nei (Trn) model of evolution was identified as the most appropriate for Dataset II. Similar to the ML analysis, the resultant BI tree showed no evidence of geographic genetic population structure (Appendix G).

Nuclear DNA Sequencing

As two discrete and highly divergent sympatric mitochondrial lineages were identified via TCS, ML and BI analyses (see Figures 2 and 3; Appendix G), post-hoc analyses were performed to investigate the potential for concordant nuclear differentiation between lineages. A 494 bp region of the nuclear ribosomal internal transcribed spacer II region (ITS2) was sequenced from a subset of four individuals from each of the two lineages. Amplification and sequencing was performed as described in Pank *et al.* (2001). Nuclear DNA sequencing resolved only a single haplotype, as all individuals possessed 100% sequence identity.

Evolutionary and Demographic History

Mean coalescence time to TMRCA for all basking shark Dataset II haplotypes was estimated by BEAST to be 2.46 Mya (CI: 0.8636 – 3.6038 Mya). Additionally, this TMRCA corresponded to the coalescence of the two distinct mitochondrial clades. Mean estimates of the TMRCA for all temporally calibrated nodes was consistent with priors (Data not shown) and a mutation rate of 1.74×10^{-9} substitutions per site per year was estimated by BEAST for all included individuals.

BSPs revealed a largely stable historical effective population size (Figure 4); however, an increase in the mean female population size (N_{Ef} x generation time) was detected starting around 116, 000 years before present (ybp) with a max peak at 33, 100 ybp, and followed by a population decline occurring just prior to the start of the Holocene at ~ 16, 000 ybp. The final decline in female population size continues on to ~ time 0 or present day. It is important to note, however, that confidence intervals (95% Highest Posterior Density) surrounding the mean female population size were quite large (Figure 4).

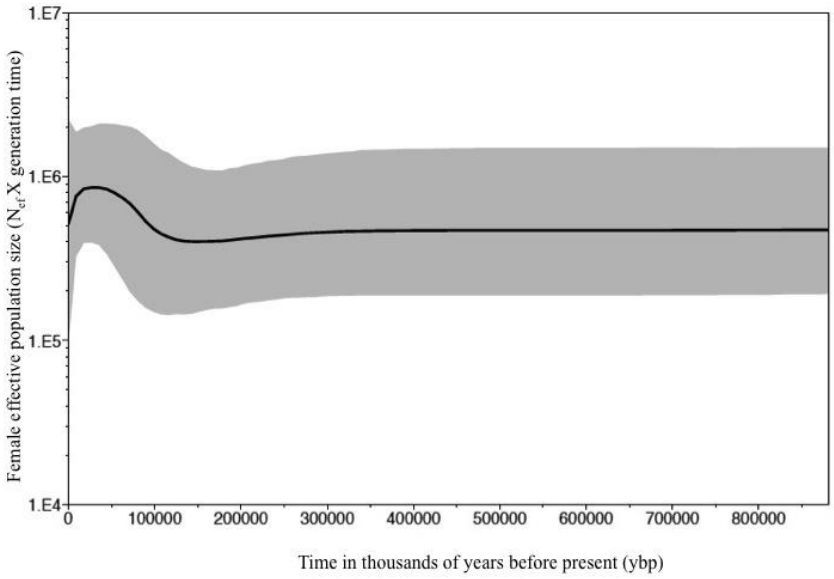


Figure 4. Bayesian skyline plot (BSP) of the *C. maximus* concatenated protein-coding data set. The y-axis indicates the female effective population size x generation time and the x-axis depicts time in years before present (ybp). The solid black line indicates the mean Nef estimate and shading depicts the 95% highest posterior density limits.

In contrast to the BSPs, the neutrality tests, Fu's F_S (Fu 1996) and R_2 (Ramos-Onsins & Rozas 2002) failed to identify a population expansion ($P > 0.05$) for any of the surveyed locations (Appendix H).

Discussion

Global Population Structure

The vast majority of globally distributed pelagic fishes are genetically delineated into multiple distinct populations, rather than a single panmictic unit, despite the innate ability to migrate substantial oceanic distances (Viñas *et al.* 2004a, 2007; Boustany *et al.* 2008; Bradman *et al.* 2011). Pelagic sharks appear no exception to this paradigm; population genetic surveys of pelagic sharks have demonstrated considerable levels of genetic heterogeneity among distinct portions of their global distribution (Heist *et al.* 1996; Castro *et al*. 2010; Benavides *et al.* 2011) suggesting that factors other than dispersal ability may shape their genetic connectivity. Most notably, a mitochondrial CR survey of the whale shark (*Rhincodon typus*), the largest extant fish species, found significant differentiation between western Atlantic and Indo-Pacific ocean basins, despite data documenting the migration of whale sharks across wide tracts of open ocean (Sequeira *et al.* 2013). Nevertheless, in contrast to the majority of other surveyed globally distributed and highly migratory pelagic fishes, all three mitochondrial genetic surveys of the basking shark to date have shown no evidence of genetic population structure across its global range (Hoelzel 2001; Hoelzel *et al*. 2006; this study)

The two previous population genetic studies of the basking shark analyzed two separate mitochondrial loci, the protein coding gene Cyt b (550 bp; n = 17) (Hoelzel 2001) and the non-protein coding CR (1,085 bp; n = 62) (Hoelzel *et al.* 2006), and both studies found no evidence of significant genetic differences among globally distributed sampling locations. Both these studies, however, sequenced only a small number of individuals and each study independently analyzed only a single mitochondrial region, suggesting that these previous studies may have suffered from limited power to detect a low level of population structure. In contrast, the present study employed a whole mitogenome approach to explore the global genetic population structure of the basking shark, and additionally employed an enhanced sample set (n = 100), utilizing a concatenated subset of three of the most variable mitochondrial protein-coding genes, to resolve genetic structure. Interestingly, whole mitogenome results confirmed previous findings suggesting extremely low levels of genetic diversity ($\pi = 0.0014 \pm 0.00036$) as well as non-significant genetic population structure ($\Phi_{ST} = -0.047$, $P > 0.05$; $F_{ST} = 0.01317$, $P > 0.05$). Furthermore, when tested independently, individual protein-coding genes (n = 13), rRNA genes (n = 2), the non-coding CR, and both concatenated datasets (II and IV), similarly demonstrated an absence of population genetic structure ($\Phi_{ST} = -0.075$ to $\Phi_{ST} = 0.008$, $P > 0.05$; $F_{ST} = -0.067$ to $F_{ST} = 0.019$; $P > 0.05$). Two genetic differentiation estimators (Φ_{ST} and F_{ST}) were utilized in this survey because Φ_{ST} values calculate molecular distance between haplotypes (Excoffier *et al.* 1992) while F_{ST} analyses provide a measure of the haplotype frequency distribution among surveyed genetic populations (Weir & Cockerham 1984). Non-significant values (across multiple

estimators) exhibited by all datasets, suggest high levels of global matrilineal genetic connectivity.

While global panmixia is rare, a limited number of marine pelagic species have also demonstrated a similar absence of genetic population structure. The teleost, wahoo (*Acanthocybium solandri*), an oceanic species that occupies waters ranging from warm temperate to tropical **(**Collette & Nauen 1983), has shown high genetic connectivity across its global range. Wahoo genetic surveys found no significant genetic population structure (nuclear $F_{ST} = 0.0125$; mtDNA $\Phi_{ST} \leq 0.0025$) across its distribution using three separate loci, including two well-studied maternally inherited mtDNA makers (Cyt b, and CR) and a single bi-parentally inherited nuclear DNA marker (Lactate dehydrogenase Subunit A intron 6) (Garber *et al.* 2005; Theisen *et al.* 2008). This genetic homogeneity among global collections was attributed to the extensive dispersal capacity of this species, across both juvenile and adult life stages. Studies have shown that wahoo often spawn near open-ocean currents, which likely facilitates transport and dispersal of their pelagic larvae (Iversen & Yoshida 1957; Matsumot 1968; Wollam 1969). Adults are also known to undertake long seasonal migrations (>1,000 km) (Franks 1998; Oxenford *et al.* 2003) and conventional tag-recapture data have recorded movement of individuals up to ~ 2, 750 km (NMFS 1999). More recently, Winkelmann *et al.* (2013) completed a global phylogenetic study of the giant squid (*Architeuthis* spp.) employing the entire mitogenome (20, 331 bp); however, despite a whole mitogenomic approach, results revealed high connectivity across this species' entire distribution, as well as low levels of genetic diversity ($\pi = 0.0066 \pm 0.0005$).The authors hypothesized that the giant squid's global genetic connectivity was likely a result of dispersal occurring across many life

stages, including, passive drift of the giant squid's small pelagic paralarvae via oceanic currents and widespread migration by adults (Winkelmann *et al.* 2013).

Identifying the potential mechanisms responsible for the basking shark's global matrilineally-based panmixia and associated low levels of genetic diversity remains challenging. Unlike marine teleosts or invertebrates, elasmobranchs do not possess a larval life stage (Hamlett 1999) eliminating the potential for passive gene flow or dispersal of early life stages via currents. Similar to other lamnoids, basking sharks are ovoviviparous (Kunzlik 1988) undergoing internal fertilization with young born fully formed, and approximately 1.8 m in length (Sund 1943). Thus, active dispersal by juveniles and/or adults must be largely responsible for the high genetic connectivity of this species; however, it is important to note that to date, little if anything is known about the behavior and movements of juvenile basking sharks. Tracking datasets comprised of mostly adult basking sharks have demonstrated that their movements appear to be largely unhindered by recognized biogeographic barriers, some of which have been previously suggested to inhibit gene flow of other smaller shark species (Keeney & Heist 2006; Schultz *et al.* 2008). For instance, cold ocean temperatures are believed to be largely limiting to some shark species, including the nurse (*Ginglymostoma cirratum*) and great hammerhead sharks (*Sphyrna mokarran*) (Rosa *et al.* 2006; Denham *et al.* 2007). Movement data for basking sharks, however, indicate that they are capable of withstanding a wide range of water temperatures (Sims *et al.* 1998, 2006; Gore *et al.* 2008), including extremely cold waters (5° C) found at mesopelagic depths (200 - 1000 m) (Skomal *et al.* 2009). In fact, deep- and cold-water seasonal migrations by basking sharks are believed to occur regularly and may be linked to their feeding behavior, as

areas of high prey abundance (zoo- and phytoplankton blooms) occur seasonally in temperate waters (Sims *et al.* 2006). The ability to withstand temperate ocean conditions, combined with a capacity for both deep- and long-distance migrations, may be a factor facilitating the inter-ocean movements of basking sharks.

Markers for Population Genetics Studies

The adoption of a whole mitogenomic approach to infer genetic connectivity, allowed for the unique opportunity to survey the genetic variability present within individual mitochondrial regions (rRNA genes, protein-coding genes and noncoding CR) and thus to identify the most variable parts within the basking shark mitogenome. Currently, the overwhelming majority of research evaluating the phylogeography, genetic diversity, and population genetic structure of elasmobranchs have relied heavily upon utilizing the protein-coding genes Cyt *b* (1, 144bp) and ND2 (1, 044 bp), as well as the non-coding CR (1, 048 bp) (Duncan *et al.* 2006; Keeney & Heist 2006; Stow *et al.* 2006; Castro *et al.* 2007; Chabot & Allen 2009; Lim *et al.* 2010; Pereyra *et al.* 2010; Karl *et al.* 2011; Veríssimo *et al.* 2012). These three mitochondrial regions are typically selected for a number of key mutational and evolutionarily significant reasons: (i) Cyt *b* is a stable, constantly evolving marker and is also one of the most well studied vertebrate protein-coding mitochondrial genes (Johns & Avise 1998); (ii) ND2 is one of the fastest evolving protein-coding genes (Broughton & Reneau 2006) allowing for small variations between closely related species, sub-species, and cryptic species to be detected (Naylor *et al.* 2012); and (iii) CR is the only non-coding region present in the vertebrate mitogenome, and is therefore theoretically under less mutational constraints than coding genes,

allowing for much higher rates of substitution (Saccone *et al.* 1987; Wan *et al.* 2004; Diniz *et al.* 2005). Hoezel *et al.* (2006) reported a CR basking shark haplotype diversity of 0.720 ± 0.028 as well as the lowest nucleotide diversity (0.0013 ± 0.0009) in any shark species to date. Interestingly, the CR diversity indices in the present study are nearly identical (h = 0.710 ± 0.052 and π = 0.0013 ± 0.000) to those of Hoezel *et al.* (2006). Such low levels of CR diversity are surprising, and beg the question of whether another mitochondrial region may be a more suitable (i.e. more variable) marker to resolve the genetic population structure of basking sharks and/or other elasmobranchs.

The mitochondrial regions exhibiting the highest levels of nucleotide diversity after sequencing 34 individuals were unexpectedly three protein-coding genes: ATP8 (168 bp), CO2 (691 bp), and ND5 (1, 830 bp). Interestingly, these three loci comprised 48 (36.9%) of all 130 identified mitogenomic polymorphisms. Furthermore, analysis revealed that the protein-coding gene, ND2, possessed the second lowest level of nucleotide diversity (0.00117 ± 0.00038) among all coding and non-coding mitochondrial region. In fact, the ND2 nucleotide diversity was lower than that of Cytochrome oxidase I (COI) (π = 0.00124 ± 0.00043), the universal animal DNA barcoding gene, elected as a universal species identifier due to its slow rate of evolution and low intra-species variability (Hebert *et al.* 2003). Based on such unexpected findings, additional fine-scale research surveying the variability present across the elasmobranch mitogenome is suggested to determine which mitochondrial genetic markers possess the highest power to determine population genetic connectivity.

Sympatric Lineages

Unexpectedly, two distinct and globally sympatric matrilineages were identified via phylogenetic analysis (TCS, ML and BI) of the basking shark mitogenomic data set. Bayesian analysis suggested that the two lineages were highly evolutionarily divergent, possessing a coalescence time of ~2.46 Mya. While the presence of two sympatric (and highly divergent) lineages could indicate the existence of a second, cryptic basking shark species, this hypothesis was dismissed for two reasons. First, 100% sequence identity was found between the two sympatric lineages upon sequencing a subset of individuals at the nuclear locus ITS2, which suggested a lack of reproductive isolation between the two lineages (Haine *et al.* 2006). Previous research on elasmobranchs has successfully uncovered reproductively isolated lineages when utilizing mitochondrial DNA analyses in conjunction with ITS2 sequences (Richards *et al.* 2009; Pinhal *et al.* 2012). Second, the maximum pairwise genetic distance (*p*) between the whole mitogenome haplotypes of the two divergent matrilineal lineages (0.5%) is much smaller than the (whole mitogenome) distance between other established lamnid sister-species, such as the pelagic (*Alopias pelagicus*) and big eye thresher sharks (*Alopias superciliosus*) (9.6%) and the shortfin mako (*Isurus oxyrinchus*) and longfin mako sharks (*Isurus paucus*) (9.8%). In addition, COI sequences exhibited a maximum pairwise genetic distance of 0.6% among all basking sharks haplotypes, which is again, much smaller than the sequence divergence between other established elasmobranch sister taxa (Serra-Pereira *et al*. 2011; Karl *et al.* 2012a) including between the newly discovered Atlantic cryptic hammerhead shark (*Sphyrna gilberti* sp. nov) and the globally distributed scalloped

hammerhead shark (*Sphyrna lewini*) (3.0-3.6 % sequence divergence) (Quattro *et al.* 2006).

Aside from cryptic speciation, several key evolutionary or demographic mechanisms may be employed to explain the occurrence of the two globally distributed and sympatric basking shark maternal lineages. First, some pelagic teleosts exhibit highly divergent, yet sympatric matrilineal lineages across parts of their range, including: wahoo *(Acanthocybium solandri*), swordfish (*Xiphias gladius),* Atlantic bluefin tuna (*Thunnus thynnuss*), and Atlantic bonito (*Sarda sarda)*; however, within these fishes their sympatric lineage presence is restricted to within a single ocean basin (Viñas *et al.* 2004b, 2007; Alvarado Bremer *et al.* 2005; Garber *et al.* 2005; Boustany *et al.* 2008). In these cases, the presence of the two sympatric clades was hypothesized to be the outcome of historical separation (allopatric divergence) followed by secondary contact, through unidirectional gene flow of individuals from the Indo-Pacific into the Atlantic. A similar hypothesis may be invoked to explain the presence of the two matrilineal basking shark sympatric lineages; however, for the sympatric lineages to occur globally, gene flow must have been historically or contemporarily bi-directional.

A second, and perhaps more likely explanation for the presence of sympatric matrilines, may be the historical occurrence of demographic population changes. In some instances, such events can lead to the presence of unusual patterns of genetic variation, including the coexistence of evolutionary distinct lineages, as seen here. Such a pattern can arise through the random loss of diversity (i.e. intermediate or connecting haplotypes), caused by large reduction in population size, thereby creating the opportunity for highly divergent haplotypes (lineages) to coexist within a single

population and/or species (Vincek *et al.* 1997; Li & Roossinck 2004; Broughton & Reneau 2006; Johnson *et al.* 2007). These types of large population declines can also lead to low levels of genetic variability across the entire genome, which is consistent with the extremely low diversity indices found in the present study. In fact, to explain the low levels of CR genetic diversity found in their basking shark study, Hoelzel *et al.* (2006) suggested that a historical population bottleneck may be responsible.

The population history revealed by the basking shark BSP (of Dataset II; 11, 429 bp) was also largely consistent with the hypothesis of historical demographic changes. However, such changes were subtle and credibility intervals were quite wide. The BSP suggests that basking sharks underwent a population expansion ~ 116, 000 ybp, followed by a very recent decline (~16, 500 ybp). Interestingly, the bulk of the BSP (> 165, 000 ybp) showed that this species endured a long period of population stability, which preceded the more recent demographic flux. However, Karl *et al*. (2012) suggested that such a pattern could simply be an artifact of the coalescent analysis itself, and that the more recent population size changes may have been sufficient to eliminate evidence of historical changes, thereby creating the false illusion of historical stability. Nevertheless, the BSP generated herein supports the hypothesis that demographic changes (including declines) have occurred, which may have resulted in the random loss genetic diversity, and ultimately could be responsible for the presence of the two sympatric and evolutionarily divergent matrilines. Admittedly, however, the BSP possessed wide credibility intervals and noted the occurrence of only subtle changes (Figure 4). Furthermore, BSPs possess numerous potential sources of error, including sample size, mutation rates, and molecular clocks; thus, any demographic findings revealed by the

BSP can only be considered inferences of historical events affecting basking shark population structure.

With respect to the above inferences, it is important to consider several key caveats. Despite the presence of numerous, presumably functionally independent, regions within the mitogenome, this genome behaves as only a single maternally-inherited locus, and represents the coalescent genealogy or history of only one genetic marker for this species (Ho & Shapiro 2011). Additionally, when performing demographic history analyses, inferences on the timing of historical events are wholly dependent upon the input of the generational mutation rate (μ) of any marker employed (Palsbøll *et al.* 2012). In the present study, a lineage-specific mutation rate was estimated using divergence times from previous molecular analysis and dated fossils to estimate minimum divergence times between lineages. Recurrent mutations in mitochondrial genes and the utilization of multiple species during calculation may lower mutation rates in phylogenetic studies (Santos *et al.* 2005; Phillips *et al.* 2009); thus, the estimated mutation rate of 1.74×10^{-3} per site/million years may be lower than the true value.

Conservation implications

Several key findings from this work highlight the need for strong conservation and management actions for basking sharks, especially in light of the continued exploitation of this enigmatic species (Clarke 2004; Compagno *et al.* 2005). Low mitogenome-wide genetic diversity, globally sympatric, yet evolutionary divergent matrilineages, and the likely occurrence of historical demographic changes, all suggest that basking sharks may have irreversibly lost substantial mitochondrial genetic diversity,

which ultimately may render this species more susceptible to disease and less likely to adapt to future environmental or anthropogenic changes (Reusch & Wood 2007). Furthermore, as exploitation continues, this species will become increasingly vulnerable through the additional loss of genetic diversity.

Future work is needed to investigate the global genetic connectivity and demographic history of this species using additional nuclear (bi-parentally inherited) genetic markers, and much larger sample sizes to investigate the potential for fine-scale genetic geographic heterogeneity and population structure in this seemingly panmictic marine species. Given the basking shark's ability to travel large ocean distances across numerous biogeographic and political boundaries, and its seemingly panmictic genetic population structure, a cooperative multiagency and global approach to its conservation and management is required in order to ensure the persistence of this highly vulnerable and ecologically important species.

References:

Akaike H. 1974 A new look at the statistical model identification. *IEEE T. Automat. Contr.* **19**, 716-723. (doi:10.1109/TAC.1974.1100705, MR 0423716)

Alvarado Bremer JR, Viñas J, Mejuto J, Ely B, Pla C. 2005 Comparative phylogeography of Atlantic bluefin tuna and swordfish: the combined effects of vicariance, secondary contact, introgression, and population expansion on the regional phylogenies of two highly migratory pelagic fishes. *Mol. Phylogenet. Evol.* **36**, 169-187. (doi: 10.1016/j.ympev.2004.12.011)

Benavides MT, Horn RL, Feldheim KA, Shivji MS, Clarke SC, Wintner S, Natanson L, Braccini M, Boomer JJ, Gulak SJB, Chapman DD. 2011 Global phylogeography of the dusky shark *Carcharhinus obscurus*: implications for fisheries management and monitoring the shark fin trade. *Endanger. Species Res.* **14**, 13-22. (doi:10.3354/esr00337)

Boustany AM, Reeb CA, Block BA. 2008 Mitochondrial DNA and electronic tracking reveal population structure of Atlantic bluefin tuna (*Thunnus thynnus*). *Mar. Biol.* **156**, 13-24. (doi: 10.1007/s00227-008-1058-0)

Bradman H, Grewe P, Appleton B. 2011 Direct comparison of mitochondrial markers for the analysis of swordfish population structure. *Fish. Res.* **109**, 95-99. (doi: 10.1016/j.fishres.2011.01.022)

Broughton RE, Reneau PC. 2006 Spatial covariation of mutation and nonsynonymous substitution rates in vertebrate mitochondrial genomes. *Mol. Bio. Evol.* **23**, 1516-1524. (doi: 10.1093/molbev/msl013)

Buonaccorsi V, McDowell JR, Graves J. 2001 Reconciling patterns of inter-ocean molecular variance from four classes of molecular markers in blue marlin (*Makaira nigricans*). *Mol. Ecol.* **10**, 1179–1196. (doi: 10.1046/j.1365-294X.2001.01270.x)

Burnham KP, Anderson DR. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach (2nd ed.). New York, NY: Springer-Verlag New York, Inc.

Castro ALF, Stewart BS, Wilson SG, Hueter RE, Meekan MG, Motta PJ, Bowen BW, Karl SA. 2007 Population genetic structure of Earth's largest fish, the whale shark

(*Rhincodon typus*). *Mol. Ecol.* **16**, 5183-5192. (doi: 10.1111/j.1365-294X.2007.03597.x)

Castro JI. 2011 *The Sharks of North America*. New York, NY: Oxford University Chabot CL, Allen LG. 2009 Global population structure of the tope (*Galeorhinus galeus*) inferred by mitochondrial control region sequence data. *Mol. Ecol.* **18**, 545-552. (doi: 10.1111/j.1365-294X.2008.04047.x)

Press.

CITES. 2002 *Consideration of proposals for amendment of Appendices I and II. Proposal: Inclusion of Basking Shark (Cetorhinus maximus) on Appendix II of CITES.* Prop. 12.36.

Clarke S. 2004 *Shark Product Trade in Hong Kong and Mainland China and Implementation of the CITES Shark Listings*. Hong Kong, China: TRAFFIC East Asia.

Clement M, Posada D, Crandall K. 2000 TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* **9**, 1657-1660.

Collette B, Nauen C. 1983 FAO Species Volume 2. Scombrids of the world. An annotated and illustrated catalogue of tunas, mackerels, bonitos and related species known to date. FAO Fisheries Synopsis, 125. FAO/UNDP, Rome, Italy.

Compagno LJV, Dando M, Fowler S. 2005 Sharks of the World. London, UK: HarperCollins Publishers Ltd

COSEWIC 2007. COSEWIC assessment and status report on the basking shark *Cetorhinus maximus* (Pacific population) in Canada. Committee on the Status of Endangered Wildlife in Canada. Ottawa. **7**, 34.

Cummings MP, Otto SP, Wakeley J. 1995 Sampling properties of DNA sequence data in phylogenetic analysis. *Mol. Bio. Evol.* **12**, 814-822.

Darriba D, Taboada GL, Doallo R, Posada D. 2012 jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods*. **9**, 772-772. (doi:10.1038/nmeth.2109)

DeFilippis VR, Moore WS. 2000 Resolution of phylogenetic relationships among recently evolved species as a function of amount of DNA sequence: An empirical study based on woodpeckers (Aves: Picidae). *Mol Phylogenet. Evol*. **16**, 143-160. (doi:10.1016/j.bbr.2011.03.031)

Denham J, Stevens J, Simpfendorfer CA, Heupel MR, Cliff G, Morgan A, Graham R, Ducrocq M, Dulvy ND, Seisay M, Asber M, Valenti SV, Litvinov F, Martins P, Lemine Ould Sidi M, Tous P, Bucal D. 2007 *Sphyrna mokarran*. In: IUCN 2013. IUCN Red List of Threatened Species. Version 2013.2. See www.iucnredlist.org.

Diniz FM, Maclean N, Ogawa M, Cintra IH, Bentzen P. 2005 The hypervariable domain of the mitochondrial control region in Atlantic spiny lobsters and its potential as a marker for investigating phylogeographic structuring. *Mar. Biotechnol.* **7**, 462-473. (doi: 10.1007/s10126-004-4062-5)

Don R, Cox P, Wainwright B, Baker K, Mattick J. 1991 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res.* **19**, 4008.

Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Field M, Heled J, Kearse M, Markowitz S, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A. 2012 Geneious v5.6, Available from http://www.geneious.com

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012 Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969-1973. (doi: 10.1093/molbev/mss075)

Duncan KM, Martin AP, Bowen BW, De Couet HG. 2006 Global phylogeography of the scalloped hammerhead shark (*Sphyrna lewini*). *Mol. Ecol.* **15**, 2239-2251. (doi: 10.1111/j.1365-294X.2006.02933.x)

Excoffier L, Smouse PE, Quattro JM. 1992 Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479-491.

Excoffier L, Lischer HEL. 2010 Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Res*. **10**, 564-567. (doi: 10.1111/j.1755-0998.2010.02847.x)

Fowler SL. 2005 *Cetorhinus maximus*. In: IUCN 2013. IUCN Red List of Threatened Species. Version 2013.1. See www.iucnredlist.org

Francis MP, Duffy C. 2002 Distribution, seasonal abundance and bycatch of basking sharks (*Cetorhinus maximus*) in New Zealand, with observations on their winter habitat. *Mar. Biol.* **140**, 831-842. (doi: 10.1007/s00227-001-0744-y)

Francis MP, Smith MH. 2010 Basking shark (*Cetorhinus maximus*) bycatch in New Zealand fisheries, 1994–95 to 2007–08. *N. Z. Aquat. Environ. Biodiv. Rep.* **49**, 57.

Francis MP, Sutton P. 2012 Possible factors affecting bycatch of basking sharks (*Cetorhinus maximus*) in New Zealand trawl fisheries. *NIWA Client Report Prepared for Department of Conservation No. WLG2012-48*.

Franks JS. 1998 Anatomy of a wahoo. *Big Game Fish. J. Fall*, pp 33-43

Fu YX. 1996 Estimating the age of the common ancestor of a DNA sample using the number of segregating sites. *Genetics* **144**, 829-838.

Knaus BJ, Cronn R, Liston A, Pilgrim K, Schwartz MK. 2011 Mitochondrial genome sequences illuminate maternal lineages of conservation concern in a rare carnivore. *BMC Ecol.* **11**, 10. (doi:10.1186/1472-6785-11-10)

Kuban GJ. 1997 Sea-monster or shark? An analysis of a supposed plesiosaur carcass netted in 1977. *NCSE Reports* **17**, 16-28.

Garber AF, Tringali MD, Franks JS. 2005 Population genetic and phylogeographic structure of wahoo, *Acanthocybium solandri*, from the western central Atlantic and central Pacific Oceans. *Mar. Biol.* **147**, 205-214. (doi: 10.1007/s00227-004-1533-1)

Gore MA, Rowat D, Hall J, Gell FR, Ormond RF. 2008 Transatlantic migration and deep mid-ocean diving by basking shark. *Bio. Lett.* **4**, 395-398. (doi:10.1098/rsbl.2008.0147)

Guindon S, Gascuel O. 2003 A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol. 52*(**5**), 696-704. (doi: 10.1080/10635150390235520)

Haine ER, Martin J, Cook JM. 2006 Deep mtDNA divergences indicate cryptic species in a fig-pollinating wasp. *BMC Evol. Biol.* **6**, 83. (doi: doi:10.1186/1471-2148-6-83)

Hamlett WC. 1999 *Sharks, skates, and rays: the biology of elasmobranch fishes*. Baltimore, Maryland: JHU Press

Hebert PD, Cywinska A, Ball SL. 2003 Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B.* **270**, 313-321. (doi: 10.1098/rspb.2002.2218)

Heinicke MP, Naylor GJP, Hedges SB. 2009 *Cartilaginous fishes (Chondrichthyes).* pp. 320-7. New York, NY: Oxford University Press.

Heist EJ, Musick JA, Graves JE. 1996 Genetic population structure of the shortfin mako (*Isurus oxyrinchus*) inferred from restriction fragment length polymorphism analysis of mitochondrial DNA. *Can. J. Fish. Aquat. Sci*. **53**, 583-588 (doi: 10.1139/f95-245)

Hester J, Atwater K, Bernard A, Francis M, Shivji MS. 2013 The complete mitochondrial genome of the basking shark *Cetorhinus maximus* (Chondrichthyes, Cetorhinidae). *Mitochondr. DNA* **0**, 1-2. (doi:10.3109/19401736.2013.845762)

Ho SY, Shapiro B. 2011 Skyline- plot methods for estimating demographic history from nucleotide sequences. *Mol. Ecol. Resour*. **11**, 423-434. (doi: 10.1111/j.1755-0998.2011.02988.x)

Hoelzel AR. 2001 Shark fishing in fin soup. *Conserv. Genet.* **2**, 69-72. (doi: 10.1023/A:1011590517389)

Hoelzel AR, Natoli A, Dahlheim ME, Olavarria C, Baird RW, Black NA. 2002 Low worldwide genetic diversity in the killer whale (*Orcinus orca*): implications for demographic history. *Proc. R. Soc. Lond. B.* **269**, 1467-1473. (doi: 10.1098/rspb.2002.2033)

Hoelzel AR, Shivji MS, Magnussen J, Francis MP. 2006 Low worldwide genetic diversity in the basking shark (*Cetorhinus maximus*). *Bio. Lett.* **2**, 639-642. (doi:10.1098/rsbl.2006.0513)

Huelsenbeck JP, Ronquist F. 2001 MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754-755. (doi:10.1093/bioinformatics/17.8.754)

Iversen ES, Yoshida HO. 1957 Notes on the biology of the wahoo in the Line Islands. *Pac. Sci.* **11**, 370-379

Jacobsen MW, Pujolar JM, Gilbert MTP, Moreno-Mayar JV, Bernatchez L, Als TD, Lobon-Cervia J, Hansen MM. 2014 Speciation and demographic history of Atlantic eels (*Anguilla anguilla* and *A. rostrata*) revealed by mitogenome sequencing. *Heredity*. (doi:10.1038/hdy.2014.44)

Johns GC, Avise JC. 1998 A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome b gene. *Mol. Bio. Evol.* **15**, 1481-1490.

Johnson JA, Dunn PO, Bouzat JL. 2007 Effects of recent population bottlenecks on reconstructing the demographic history of prairie- chickens. *Mol. Ecol.* **16**, 2203-2222. (doi: 10.1111/j.1365-294X.2007.03285.x)

Johnston M. 2014 Shark Tracker – Banba "Bob". See http://www.baskingshark.ie.

Koressaar T, Remm M. 2007 Enhancements and modifications of primer design program Primer3. *Bioinformatics* **23**, 1289-91. (doi:10.1093/bioinformatics/btm091)

Karl SA, Castro ALF, Lopez JA, Charvet P, Burgess GH. 2011 Phylogeography and conservation of the bull shark (*Carcharhinus leucas*) inferred from mitochondrial and microsatellite DNA. *Conserv. Genet.* **12**, 371-382. (doi: 10.1007/s10592-010-0145-1)

Karl SA, Castro AL, Garla RC. 2012a Population genetics of the nurse shark (*Ginglymostoma cirratum*) in the western Atlantic. *Mar. Biol.* **159**, 489-498. (doi: 10.1007/s00227-011-1828-y)

Karl SA, Toonen RJ, Grant WS, Bowen BW. 2012b Common misconceptions in molecular ecology: echoes of the modern synthesis. *Mol. Ecol.* **21**, 4171-4189. (doi: 10.1111/j.1365-294X.2012.05576.x)

Karlsen BO, Emblem Å, Jørgensen TE, Klingan KA, Nordeide JT, Moum T, Johansen SD. 2014 Mitogenome sequence variation in migratory and stationary ecotypes of North-east Atlantic cod. *Marine genomics*. **15**, 103-108. (doi:10.1016/j.margen.2014.01.001)

Keeney DB, Heist EJ. 2006 Worldwide phylogeography of the blacktip shark (*Carcharhinus limbatus*) inferred from mitochondrial DNA reveals isolation of western Atlantic populations coupled with recent Pacific dispersal. *Mol. Ecol.* **15**, 3669-3679. (doi: 10.1111/j.1365-294X.2006.03036.x)

Kunzlik PA. 1988 *The basking shark. Scottish Fisheries Information Pamphlet*. Department of Agriculture and Fisheries for Scotland, Aberdeen.

LeDuc RG, Robertson KM, Pitman RL. 2008 Mitochondrial sequence divergence among Antarctic killer whale ecotypes is consistent with multiple species. *Biol. Lett.* **4**, 426-429. (doi:10.1098/rsbl.2008.0168)

Lee M. 2012 Shark's journey a first for science. In *The San Diego Union-Tribune*. See http://www.utsandiego.com/news/2012/Feb/13/sharks-journey-first-science/2/?#article-copy.

Li H, Roossinck MJ. 2004 Genetic bottlenecks reduce population variation in an experimental RNA virus population. *J. Virol.* **78**, 10582-10587. (doi: 10.1128/JVI.78.19.10582-10587.2004)

Librado P, Rozas J. 2009 DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451-1452. (doi:10.1093/bioinformatics/btp187)

Lim DD, Motta P, Mara K, Martin AP. 2010 Phylogeny of hammerhead sharks (Family Sphyrnidae) inferred from mitochondrial and nuclear genes. *Mol. Phylogenet. Evol.* **55**, 572-579. (doi:10.1016/j.ympev.2010.01.037)

Ma C, Yang P, Jiang F, Chapuis MP, Shali Y, Sword GA, Kang LE. 2012 Mitochondrial genomes reveal the global phylogeography and dispersal routes of the migratory locust. *Mol. Ecol.* **21**, 4344-4358. (doi: 10.1111/j.1365-294X.2012.05684.x)

Magnussen JE, Pikitch EK, Clarke SC, Nicholson C, Hoelzel AR, Shivji MS. 2007 Genetic tracking of basking shark products in international trade. *Anim. Conserv.* **10**, 199-207. (doi: 10.1111/j.1469-1795.2006.00088.x)

Matsumot WM. 1968 Morphology and distribution of larval wahoo *Acanthocybium solandri* (Cuvier) in central Pacific Ocean. *USFWS Fish. Bull.* **66**, 299

McGowen MR, Spaulding J, Gatesy J. 2009 Divergence date estimation and a comprehensive molecular tree of extant cetaceans. *Mol. Phylogenet. Evol.* **53**, 891–906. (doi: 10.1016/j.ympev.2009.08.018)

Meiklejohn KA, Danielson MJ, Faircloth BC, Glenn TC, Braun EL, Kimball RT. 2014 Incongruence among different mitochondrial regions: A case study using complete mitogenomes. *Mol. Phylogenet. Evol*. **78**, 314-323. (doi: 10.1016/j.ympev.2014.06.003)

Morin PA, Archer FI, Foote AD, Vilstrup J, Allen EE, Wade P, Durban J, Parsons K, Pitman R, Lewyn L, Bouffard P, Nielsen SCA, Rasmussen M, Willerslev E, Gilbert MTP, Harkins T. 2010 Complete mitochondrial genome phylogeographic analysis of killer whales (*Orcinus orca*) indicates multiple species. *Genome Res.* **20**, 908-916. (doi:10.1101/gr.102954.109)

Musick JA, Burgess G, Cailliet G, Camhi M, Fordham S. 2000 Management of sharks and their relatives (Elasmobranchii). *Fisheries* **25**, 9-13. (doi: 10.1577/1548-8446(2000)025<0009:MOSATR>2.0.CO;2)

Naylor GJ, Caira JN, Jensen K, Rosana KAM, White WT, Last PR. 2012 A DNA sequence-based approach to the identification of shark and ray species and its implications for global elasmobranch diversity and parasitology. *Bull. Amer. Mus. Nat. Hist.* 1-262. (doi: http://dx.doi.org/10.1206/754.1)

Nei M, Kumar S. 2000 *Molecular Evolution and Phylogenetics*. New York, NY: Oxford University Press.

NMFS. 1999 Billfish newsletter and cooperative tagging program report. Prepared by David Holtz, Fishery Biologist Southwest Fisheries Science Center, PO Box 271, La Jolla, CA 92038-0271

Oxenford HA, Murray PA, Luckhurst BE. 2003 The biology of wahoo (*Acanthocybium solandri*) in the western central Atlantic. *Gulf Caribb Res.* **15**, 33-34

Palumbi S, Martin A, Ramano S, McMillan WO, Stice L, Grabowski G. 1991 The Simple Fool's Guide to PCR, version 2. 46 pp. Honolulu, Hawaii: University of Hawaii Zoology Department.

Palsbøll PJ, Zachariah Peery M, Olsen MT, Beissinger SR, Berube M. 2013 Inferring recent historic abundance from current genetic diversity. *Mol. Ecol.* **22**, 22-40. (doi: 10.1111/mec.12094)

Pank M, Stanhope M, Natanson L, Kohler N, Shivji M. 2001 Rapid and simultaneous identification of body parts from the morphologically similar sharks *Carcharhinus*

obscurus and *Carcharhinus plumbeus* (Carcharhinidae) using multiplex PCR. *Mar. Biotechnol*. **3**, 231-240. (doi: 10.1007/s101260000071)

Parker HW, Stott FC. 1965 Age, size and vertebral calcification in the basking shark, *Cetorhinus maximus* (Gunnerus). *Zool. Med. Leiden*. **40**, 305-319.

Peakall R, Smouse PE. 2006 GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* **6**, 288-295. (doi:10.1111/j.1471-8286.2005.01155.x)

Peakall R, Smouse PE. 2012 GenAlEx 6.5: Population genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**, 2537-2539. (doi:10.1093/bioinformatics/bts460)

Pereyra S, García G, Miller P, Oviedo S, Domingo A. 2010 Low genetic diversity and population structure of the narrownose shark (*Mustelus schmitti*). *Fish. Research*. **106**, 468-473. (doi:10.1016/j.fishres.2010.09.022)

Phillips CD, Trujillo RG, Gelatt TS, Smolen MJ, Matson CW, Honeycutt RL, Patton JC, Bickham JW. 2009 Assessing substitution patterns, rates and homoplasy at HVRI of Steller sea lions, *Eumetopias jubatus*. *Mol. Ecol.* **18**, 3379-3393. (DOI: 10.1111/j.1365-294X.2009.04283.x)

Pinhal D, Shivji MS, Vallinoto M, Chapman DD, Gadig OBF, Martins C. 2012 Cryptic hammerhead shark lineage occurrence in the western South Atlantic revealed by DNA analysis. *Mar. Biol.* **159**, 829-836. (doi: 10.1007/s00227-011-1858-5)

Quattro JM, Stoner DS, Driggers WB, Anderson CA, Priede KA, Hoppmann, EC, Campbell NH, Duncan KM, Grady JM. 2006 Genetic evidence of cryptic speciation

within hammerhead sharks (Genus Sphyrna). *Mar. Biol.* **148**, 1143-1155. (doi: 10.1007/s00227-005-0151-x)

Rambaut A, Drummond AJ. 2007 Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer

Ramos-Onsins SE, Rozas J. 2002 Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol*. **19**, 2092-2100.

Reusch TB, Wood TE. 2007 Molecular ecology of global change. *Mol. Ecol.* **16**, 3973-3992. (doi: DOI: 10.1111/j.1365-294X.2007.03454.x)

Richards VP, Henning M, Witzell W, Shivji MS. 2009 Species delineation and evolutionary history of the globally distributed spotted eagle ray (*Aetobatus narinari*). *J. Hered.* **100**, 273-283. (doi: 10.1093/jhered/esp005)

Rokas A, Carroll SB. 2005 More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. *Mol. Bio. Evol.* **22**, 1337-1344. (doi: 10.1093/molbev/msi121)

Rosa RS, Castro ALF, Furtado M, Monzini J, Grubbs RD. 2006 *Ginglymostoma cirratum*. In: IUCN 2013. IUCN Red List of Threatened Species. Version 2013.2. See www.iucnredlist.org.

Ruvolo M, Disotell TR, Allard MW, Brown WM, Honeycutt RL. 1991 Resolution of the African hominoid trichotomy by use of a mitochondrial gene sequence. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1570-1574. (doi: 10.1073/pnas.88.4.1570)

Saccone C, Attimonelli M, Sbisa E. 1987 Structural elements highly preserved during the evolution of the D-loop-containing region in vertebrate mitochondrial DNA. *J. Mol. Evol.* **26**, 205-211. (doi: 10.1007/BF02099853)

Saitou N, Nei M. 1986 The number of nucleotides required to determine the branching order of three species, with special reference to the human chimpanzee gorilla divergence. *J. Mol. Evol.* **24**, 189-204. (doi: 10.1007/BF02099966)

Santos C, Montiel R, Sierra B, Bettencourt C, Fernandez E, Alvarez L, Lima M, Abade A, Aluja MP. 2005 Understanding differences between phylogenetic and pedigree-derived mtDNA mutation rate: a model using families from the Azores Islands (Portugal). *Mol. Biol. Evol.* **22**, 1490-1505. (doi: 10.1093/molbev/msi141)

Schultz JK, Feldheim KA, Gruber SH, Ashley MV, McGovern TM, Bowen BW. 2008 Global phylogeography and seascape genetics of the lemon sharks (genus Negaprion). *Mol. Ecol.* **17**, 5336-5348. (doi: 10.1111/j.1365-294X.2008.04000.x)

Schultze HP. (Ed.) 2012 *Handbook of Paleoichthyology: Chondrichthyes.-Mesozoic and Cenozoic Elasmobranchii: Teeth/by H. Cappetta. Vol. 3E*. München, Germany: Verlag Dr. Friedrich Pfeil.

Schwarz GE. 1978 Estimating the dimension of a model. *Ann. Stat.* **6,** 461-464. (doi: 10.1214/aos/1176344136.MR 468014)

Sequeira AMM, Mellin C, Meekan MG, Sims DW, Bradshaw CJA. 2013 Inferred global connectivity of whale shark *Rhincodon typus* populations. *J. Fish. Biol.* **82**, 367-389. (doi: 10.1111/jfb.12017)

Serra- Pereira B, Moura T, Griffiths AM, Serrano Gordo L, Figueiredo I. 2011 Molecular barcoding of skates (Chondrichthyes: Rajidae) from the southern Northeast Atlantic. *Zool. Scr.* **40**, 76-84. (doi: 10.1111/j.1463-6409.2010.00461.x)

Shapiro B, Rambaut A, Drummond AJ. 2006 Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol. Biol. Evol.* **23**, 7-9. (doi: 10.1093/molbev/msj021)

Sims DW, Quayle VA. 1998 Selective foraging behaviour of basking sharks on zooplankton in a small-scale front. *Nature* **393**, 460-464. (doi:10.1038/30959)

Sims DW. 2006 Threshold foraging behavior of basking sharks on zooplankton: life on an energetic knife-edge? *Proc. R. Soc. B.* **266**, 1437-1443. (doi: 10.1098/rspb.1999.0798)

Skomal GB, Zeeman SI, Chisholm JH, Summers EL, Walsh HJ, McMahon KW, Thorrold SR. 2009 Transequatorial migrations by basking sharks in the western Atlantic Ocean. *Curr. Bio.* **19**, 1019-1022. (doi:10.1016/j.bbr.2011.03.031)

Stow A, Zenger K, Briscoe D, Gilings M, Peddemors V, Otway N, Harcourt R. 2006 Isolation and genetic diversity of endangered grey nurse shark (*Carcharias taurus*) populations. *Biol. Lett.* **2**, 308-311. (doi:10.1098/rsbl.2006.0441)

Subramanian S, Denver DR, Millar CD, Heupink T, Aschrafi A, Emslie SD, Baroni C, Lambert DM. 2009 High mitogenomic evolutionary rates and time dependency. *Trends Genet.* **25**, 482-286. (doi: 10.1016/j.bbr.2011.03.031)

Sund O. 1943. Et brugdelbarsel. *Naturen* **67**, 285-286.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011 MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731-2739. (doi: 10.1093/molbev/msr121)

Teacher AG, André C, Merilä J, Wheat CW. 2012 Whole mitochondrial genome scan for population structure and selection in the Atlantic herring. *BMC Evol. Biol.* **12**, 248. (doi:10.1186/1471-2148-12-248)

Theisen TC, Bowen BW, Lanier W, Baldwin JD. 2008 High connectivity on a global scale in the pelagic wahoo, *Acanthocybium solandri* (tuna family Scombridae). *Mol. Ecol.* **17**, 4233-4247. (doi: 10.1111/j.1365-294X.2008.03913.x)

Towrie S. 2014 Monsters of the Deep – The Strongsay Beast. See http://www.orkneyjar.com/folklore/seabeasts.htm

Untergrasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012 Primer3 - new capabilities and interfaces. *Nucleic Acids Res.* **40,** e115. (doi: 10.1093/nar/gks596)

Veríssimo A, McDowell JR, Graves JE. 2012 Genetic population structure and connectivity in a commercially exploited and wide-ranging deepwater shark, the leafscale gulper (*Centrophorus squamosus*). *Mar. Freshwater. Res.* **63**, 505-512. (doi:10.1071/MF11237)

Vilstrup JT, Ho SYW, Foote AD, Morin PA, Kreb D, Krützen M, Parra GJ, Robertson KM, de Stephanis R, Verborgh P, Willerslev E, Orlando L, Gilber MTP. 2011 Mitogenomic phylogenetic analyses of the Delphinidae with an emphasis on the Globicephalinae. *BMC Evol. Bio.* **11**, 65. (doi: 10.1186/1471-2148-11-65)

Viñas J, Bremer JA, Pla C. 2004a Inter-oceanic genetic differentiation among albacore (*Thunnus alalunga*) populations. *Mar. Biol.* **145**, 225-232. (doi: 10.1007/s00227-004-1319-5)

Viñas J, Alvarado Bremer JR, Pla C. 2004b Phylogeography of the Atlantic bonito (*Sarda sarda*) in the northern Mediterranean: the combined effects of historical vicariance, population expansion, secondary invasion, and isolation by distance. *Mol. Phylogenet. Evolu.* **33**, 32-42. (doi: 10.1016/j.ympev.2004.04.009)

Viñas J, Bremer JA, Mejuto J, de la Serna JM, García-Cortés B, Pla C. 2007 Swordfish genetic population structure in the North Atlantic and Mediterranean. *Collect. Vol. Sci. Pap. ICCAT*, **61**, 99-106.

Vincek V, O'Huigin C, Satta Y, Takahata N, Boag PT, Grant PR, Grant BR, Klein J. 1997 How large was the founding population of Darwin's finches? *Proc. R. Soc. B.* **264**, 111-118. (doi: 10.1098/rspb.1997.0017)

von der Heyden S, Beger M, Toonen RJ, van Herwerden L, Juinio-Meñez MA, Ravago-Gotanco R, Fauvelot C, Bernardi G. 2014 The application of genetics to marine management and conservation: examples from the Indo-Pacific. *B. Mar. Sci.* **90**, 123-158. (doi: 10.5343/bms.2012.1079)

Wan QH, Wu H, Fujihara T, Fang SG. 2004 Which genetic marker for which conservation genetics issue? *Electrophoresis* **25**, 2165-2176. (doi: 10.1002/elps.200305922)

Weir BS, Cockerham, CC. 1984 Estimating F-Statistics for the analysis of population structure. *Evolution.* **38**, 1358-1370.

Went AEJ, Súilleabháin, ÓS. 1967 Fishing for sun-fish or basking sharks in Irish waters. *P. Roy. Irish. Acad. C.* **65**, 91-115.

Winkelmann I, Campos PF, Strugnell J, Cherel Y, Smith PJ, Kubodera T, Allcock L, Kampmann ML, Schroeder H, Guerra A, Norman M, Finn J, Ingrao D, Clarke M,

Gilbert MTP. 2013 Mitochondrial genome diversity and population structure of the giant squid *Architeuthis*: genetics sheds new light on one of the most enigmatic marine species. *Proc. R. Soc. B.* **280**. (doi:10.1098/rspb.2013.0273)

Wollam MB. 1969 Larval wahoo, *Acanthocybium solandri*, from the straits of Yucatan and Florida. *Fla. Dep. Nat. Resourc. Leafl. Ser*. **4**, 1-7

Xiong Y, Brandley MC, Xu S, Zhou K, Yang G. 2009 Seven new dolphin mitochondrial genomes and a time-calibrated phylogeny of whales. *BMC Evol. Bio*. **9**, 20. (doi: 10.1186/1471-2148-9-20)

Zhang P, Liang D, Mao RL, Hillis DM, Wake DB, Cannatella DC. 2013 Efficient sequencing of anuran mtDNAs and a mitogenomic exploration of the phylogeny and evolution of frogs. *Mol. Biol. Evol.* **30**, 1899-1915. (doi: 10.1093/molbev/mst091)

Appendix A. List of PCR and sequencing primers used herein. Headings include gene/region name (Gene), primer name, primer sequence (5'-3'), and notes on primer origin, including citation where available (Description) origin (J. Hester unpublished).

Gene	Primer Name	Primer Sequence (5'-3')	Description
12S	Cmax 12SF1	CGTCTATATACCGCCGTCGT	*C. maximus* species-specific
	Cmax12SF2	TGTCCCAAACCCACCTAGAG	*C. maximus* species-specific
	Cmax12SR1	GAGCCAGTTTCAGAAAAGCTC	*C. maximus* species-specific
	Cmax12SR2	CTCAAGCTTACGCTTATGTGTC	*C. maximus* species-specific
16S	Elasmo16S-F3	TCGGCAAACACAAACTCCGCC	Elasmobranch 16S forward
	Elasmo16S-R2	TGTTTTTGGTAAACAGGCGGAG	Elasmobranch 16S reverse
	Elasmo16S-R3	GAGTTTGTGTTTGCCGAGTTCC	Elasmobranch 16S reverse
	16Sarl(F)	CGCCTGTTTATCAAAAACAT	Palumbi 1996
	16Srlh(R)	CCGGTCTGAACTCAGATCACGT	Palumbi 1996
	Lam16SF3	TCTTTTAAATGAAGACCCGTATGAAAGGCATCACG	Lamniform 16S forward
	Cmax16SF4	ACCGGAGAAATCCAGGTCAGTTTCTATCTATG	*C. maximus* species-specific
	Cmax16SF5	GGAGCCAATACCCAAGGCACGCTCCATTTTC	*C. maximus* species-specific
	Cmax16SF6	GGAGGCAAGTCGTAACATGGT	*C. maximus* species-specific
	Cmax16SF7	GTGGCAAAAGAGTGGGAAGACT	*C. maximus* species-specific
	Cmax16SF8	AACTCGGCAAACACAAACTCCGCC	*C. maximus* species-specific
	Cmax16SR2	CACGAGTAGATCAATTTCATTGATTAGAAAATAG	*C. maximus* species-specific
	Cmax16SR3	GACCTTAATAAGGTGGACTTATTGGTGATG	*C. maximus* species-specific
	Cmax16SR4	CAATGTTGGGGCCTTTGAGGAGT	*C. maximus* species-specific

Appendix A continued.

ND1	CmaxND1F1	CTTACGAAGTGAGCCTTG	*C. maximus* species-specific
	CmaxND1R1	GGATTATCCAAGGCTCACTTCG	*C. maximus* species-specific
	CmaxND1R2	CATATTAAGGCTAGGGGTC	*C. maximus* species-specific
	CmaxND1R3	GGTACAGGTTCAAGTCCTG	*C. maximus* species-specific
ND2	CmaxND2F1	CTATCTACCATTGCTCTTATAACCCTCCTTTCC	*C. maximus* species-specific
	CmaxND2F2	TAATCGTTGACTATTTTCTACAAACCATAA	*C. maximus* species-specific
	CmaxND2F3	CTGTCCGCGGANCTACAACCCGCTACTTAATTC	*C. maximus* species-specific
	CmaxND2F4	GAAGATTATTACAAATGCATGAGCTGTCAC	*C. maximus* species-specific
	CmaxND2F5	CACCACAGGACTCATCCTATCT	*C. maximus* species-specific
	CmaxND2F6	GGACCACTTTGATAGAGTGGG	*C. maximus* species-specific
	CmaxND2F7	TTTCCCTCGGAGGATTACCT	*C. maximus* species-specific
	CmaxND2F8	CTTTTACCTGCGCCTATGCT	*C. maximus* species-specific
	CmaxND2R2	GTAAAAGAATAGACTGAGGAGGGTTATTATAG	*C. maximus* species-specific
	CmaxND2R3	GGTTAATGTTGTGGCATAGCATAGGCGCAGG	*C. maximus* species-specific
	CmaxND2R4	CCAAGATGTGCAATAGAGGAAT	*C. maximus* species-specific
	CmaxND2R5	GTAGGATCGAGGCCTATTTGTC	*C. maximus* species-specific
	CmaxND2R6	CGGCTCGAATTAGGAGGCTTA	*C. maximus* species-specific
	CmaxND2R7	GGAGTCAGTGTGAACTGATG	*C. maximus* species-specific
ASN	CmaxASNF1	CAATCCAGCGAACTTTTACC	*C. maximus* species-specific
CO1	CmaxCOIF1	TAATATAAGCTTTTGACTCCTCCCTCCTTC	*C. maximus* species-specific
	CmaxCOIF2	GCTGGAGCCGGAACTGGCTGAACAGTATACCCT	*C. maximus* species-specific
	CmaxCOIF3	GGTCACCCAGAAGTTTA	*C. maximus* species-specific
	CmaxCOIF4	CCAGCCATCTCCCAGTATCA	*C. maximus* species-specific
	CmaxCOIF5	CGTGTGATCAATTCTAGTCAC	*C. maximus* species-specific
	CmaxCOIF6	ATGAAGCCACCAGCCTCT	*C. maximus* species-specific
	CmaxCOIF7	CTGATTCCCCTTAATATCTGGC	*C. maximus* species-specific

Appendix A continued.

	CmaxCOIF8	CATCCTCTGAGAAGCATTTGCC	*C. maximus* species-specific
	CmaxCOIF9	ACTCCATGGCTGCCCCCCACCC	*C. maximus* species-specific
	CmaxCOIR1	GAAGGAGGGAGGAGTCAAAAGCTTTATTA	*C. maximus* species-specific
	CmaxCOIR2	GGTTCCGATCGGTAAGCAATATTGTGATGCCGGCT	*C. maximus* species-specific
	CmaxCOIR4	CTCCCGCCTTTGTTCTTACGGCGGGAGAAAG	*C. maximus* species-specific
	CmaxCOIR5	GTTCGCTGGATTGAACAC	*C. maximus* species-specific
	CmaxCOIR6	CTGGGTGACCGAAGAATCAG	*C. maximus* species-specific
	CmaxCOIR7	CCAAGAAGTGATCCGGGTTGGCC	*C. maximus* species-specific
	CmaxCOIR10	CCTCCATGAAGGGTTGCTAA	*C. maximus* species-specific
	CmaxCOIR11	TCCCAGCAAGACCTAGGAAA	*C. maximus* species-specific
	CmaxCOIR12	GTCGTGGTATCCCAGCAAGA	*C. maximus* species-specific
CO2	MowsCO2F1	ATGGCACACCCCTCACAATTAGG	*M. owstoni* forward primer
	CmaxCO2R4	ATTGGCCTGGGGTTAAGTCT	*C. maximus* species-specific
	CmaxCO2R6	AGAATAATGGCGGGGAGAAT	*C. maximus* species-specific
	CmaxCO2R8	GGGCATAAAACTGTGGTTGG	*C. maximus* species-specific
LYS	CmaxLysF1	GAAGCTAAATTGGGCCTAGCG	*C. maximus* species-specific
ATP8	MowsATP8F1	CCTGAAACTGACCATGATCA	*M. owstoni* forward primer
ATP6	CmaxATP6F2	GCCGTAGCAATAATTCAAGC	*C. maximus* species-specific
	CmaxATP6R2	TGATAAATCAACCTTGGAGGGC	*C. maximus* species-specific
	CmaxATP6R3	GGGATTCCAAGGAAAGAGGGGCTT	*C. maximus* species-specific
	CmaxATP6R6	GCTACGGCAACTTCTAGGGTTG	*C. maximus* species-specific
	CmaxATP6R7	TGCAATTGTTGGTCGGTTTA	*C. maximus* species-specific
	CmaxATP6R8	GTTGGGTTGTGGGCGTAAAGGTA	*C. maximus* species-specific

CO3	ElasmoCO3F1	CCTTTATGTATCCATCTATTGATGAGGCTC	Elasmobranch CO3 forward
	ElasmoCO3F2	CCATCTATTGATGAGGCTCATA	Elasmobranch CO3 forward
	CmaxCO3F5	AGCCCTCCAAGCCATAGAAT	*C. maximus* species-specific
	CmaxCO3R3	CAGAAGACGTAATGAGGCTC	*C. maximus* species-specific
	CmaxCO3R7	GGCAAGGCTTGAGTGGTAAA	*C. maximus* species-specific
ND3	CmaxND3R1	AGGATTAGAGAAATCAGGGCCGTA	*C. maximus* species-specific
ND4L	CmaxND4LF2	AGCCACTTCACGTTCTCA	*C. maximus* species-specific
	CmaxND4LR1	CACAGGCTGAGAATGTGAGG	*C. maximus* species-specific
ND4	CmaxND4F4	GGCCTAGTCGCAGGAGCGATCCT	*C. maximus* species-specific
	CmaxND4F5	CTATGATGACTAGCCTGCCT	*C. maximus* species-specific
	CmaxND4F8	TCACAGCCTCCTACTCCCTTT	*C. maximus* species-specific
	CmaxND4F9	GCCTCCACCTTATTCCAGTC	*C. maximus* species-specific
	CmaxND4R4	CTGTGGATTCGCTCGTAGTTAGTG	*C. maximus* species-specific
	CmaxND4R5	GGCTGTAGGTAGTAGTTATGGG	*C. maximus* species-specific
	CmaxND4R6	CATTTCGGTTGCGGAGAATGCC	*C. maximus* species-specific
	CmaxND4R7	AGGAGGCAGGCTAGTCATCA	*C. maximus* species-specific
	CmaxND4R8	GTTGCGGAGAATGCCATAAT	*C. maximus* species-specific
ND5	ElasmoND5F1	TCAAAAGACGCCATCATTGAA	Elasmobranch ND5 forward
	CmaxND5F2	CAATCATCCACCATCAGC	*C. maximus* species-specific
	CmaxND5F3	ACCCTAATTGATCCTCATCCCATG	*C. maximus* species-specific
	CmaxND5F4	TGGCTGATGACACAGCCGAA	*C. maximus* species-specific
	CmaxND5R6	TGTCAAATGGTCCAATGTTCA	*C. maximus* species-specific
	ElasmoND5R1	TCGGCTGTATCATCAACCAA	Elasmobranch ND5 reverse

Appendix A continued.

	CmaxND5R2	CTGGAATAAGGTGGAGGCT	*C. maximus* species-specific
	CmaxND5R3	TCGTGTCCTTGACCTCTCTCGG	*C. maximus* species-specific
	CmaxND5R4	GAGGAAAATAAGGCGGAGGC	*C. maximus* species-specific
	CmaxND5R5	GCTTCTGCTTGGAGTTGCATCA	*C. maximus* species-specific
	CmaxND5R6	TTGATGTTGAGAAGGCGATG	*C. maximus* species-specific
	CmaxND5intF2	AGGGGGACTCCACAAACTCT	*C. maximus* species-specific
	CmaxND5extR1	GCTGATGGTGGATGATTGTTGG	*C. maximus* species-specific
ND6	CmaxND6F1	CCCCACCAATAACCCACAC	*C. maximus* species-specific
	CmaxND6R1	TGCTGATGGTGGATGATTGT	*C. maximus* species-specific
	CmaxND6R2	GCTGCATTAGGTTTGGTCGT	*C. maximus* species-specific
Cyt b	CmaxCytBF3	CTGAGTCCTAATAGCCGATATACTAATCCTAACCT	*C. maximus* species-specific
	CmaxCytBF5	CAAGCACATTACTCATCTCGACTACATCAC	*C. maximus* species-specific
	CmaxCytBF6	CATGAGTAACAATTGTCAAGTAGACCAACAC	*C. maximus* species-specific
	CmaxCytBF7	CACATTCTTCACTTTATCGGGCAT	*C. maximus* species-specific
	CmaxCytBF9	AGTTCCCCTCCTCCACACTT	*C. maximus* species-specific
	CmaxCytBF10	CAATGAATCTGAGGCGGTTT	*C. maximus* species-specific
	CmaxCytBR2	GAGTAGGGTTAGGATGCCTAGAAGGATGATTA	*C. maximus* species-specific
	CmaxCytBR3	GAACTAGTATGAGGATGAGAATGGAGAATAGCAGA	*C. maximus* species-specific
	CmaxCytBR4	GTGAGTGGGATCACAATGAGGAATAGAGAG	*C. maximus* species-specific
	CmaxCytBR5	GGCTTACAAGGCCGACGCTTTAAGTTAAGC	*C. maximus* species-specific
	CmaxCytBR6	GAGGCTCCATTGGCGTGG	*C. maximus* species-specific
	CmaxCytBR7	GAGCCTGTTTCATGTAGGAAGAG	*C. maximus* species-specific
	CmaxCytBR9	AGGAGAAGGCTAGGGAGACG	*C. maximus* species-specific

Appendix A continued.

PRO	CmaxProF1(2)	CCAAGATTCTGCCCAAACTG	*C. maximus* species-specific
CR	CmaxCRF1	GACCTCGACATCTGTCTA	*C. maximus* species-specific
	CmaxCRF2	TCGTCCTTGACCGTCTCAAGAT	*C. maximus* species-specific
	CmaxCRF3	CCCCCTCCCCCTAATATACA	*C. maximus* species-specific
	CmaxCRF6	GGTCGAGGTGTAGCAAATGAA	*C. maximus* species-specific
	CmaxCRR1	GTCGAGATGAGTAATGTGCTTG	*C. maximus* species-specific
	CmaxCRR2	GTCAATTGGTGGGGATCAAC	*C. maximus* species-specific
	CmaxCRR3	ATATATGTCCGGCCCTCGTT	*C. maximus* species-specific
	CmaxCRR4	AATGTATGTGGGCCATGTCA	*C. maximus* species-specific

*All primers were designed by J. Hester for this study unless stated otherwise

Appendix B. Additional *C. maximus*-specific PCR primers (*n* = 28) created in Primer3 (Unterrasser *et al.* 2012; Koressaar *et al*. 2007) including gene/region, primer name, primer sequence (5'-3'), sequencing direction and description of origin.

Gene	Primer Name	Primer Sequence (5'-3')	Description
ND1	CmaxND1F2	ACTCCGCAAAGGCCCCAACA	*C. maximus* species-specific
	CmaxND1R4	GGCGAATGAGCCTCCTGCGTA	*C. maximus* species-specific
	CmaxND1R5	GGCCTGCTGTAGCTGTGGGG	*C. maximus* species-specific
ND2	CmaxND2F9	CCAGGCTCTGCCACACTAGCC	*C. maximus* species-specific
	CmaxND2F10	AGGCTCTGCCACACTAGCCACA	*C. maximus* species-specific
	CmaxND2R8	TGGCTGGGGTGACTGGGAGG	*C. maximus* species-specific
	CmaxND2R9	GGCTGGGGTGACTGGGAGGA	*C. maximus* species-specific
CO1	CmaxCO1F10	ATTGCCATCCCCACGGGTGT	*C. maximus* species-specific
	CmaxCO1F11	TGCCATCCCCACGGGTGTAA	*C. maximus* species-specific
	CmaxCO1R13	GGGGCAGCCGTGGAGTCATT	*C. maximus* species-specific
	CmaxCO1R14	GGGCAGCCGTGGAGTCATTCG	*C. maximus* species-specific
CO2	CmaxCO2F2	TCCCATGGAGTCGCCTGTTCG	*C. maximus* species-specific
	CmaxCO2F3	CCCATGGAGTCGCCTGTTCG	*C. maximus* species-specific
	CmaxCO2F4	AGCTGTACCGGCCCTAGGAGT	*C. maximus* species-specific
	CmaxCO2R9	AGGGTGCGGATTAAGTTGAGGCA	*C. maximus* species-specific
	CmaxCO2R10	GGGTGCGGATTAAGTTGAGGCA	*C. maximus* species-specific
	CmaxCO2R11	TGGTCAGTTTCAGGGCTCAGGT	*C. maximus* species-specific

Appendix B continued.

ND5	CmaxND5F5	GTTGCCGGCGTCTTCCTGCT	*C. maximus* species-specific
	CmaxND5F6	AGTCGTTGCCGGCGTCTTCC	*C. maximus* species-specific
	CmaxND5F7	GTCGTTGCCGGCGTCTTCCT	*C. maximus* species-specific
	CmaxND5R7	GGTGAGGATTAGGGCTCAGGCGT	*C. maximus* species-specific
ND6	CmaxND6F2	ACCACACGCAAAGTCCCCCA	*C. maximus* species-specific
	CmaxND6F3	CCGCAGCATGCGCACTTACC	*C. maximus* species-specific
	CmaxND6F4	CGCAGCATGCGCACTTACCC	*C. maximus* species-specific
	CmaxND6R4	GGCGCTTGCTGCTGAGCCTT	*C. maximus* species-specific
	CmaxND6R5	GCTTATACGGCGGCGCTTGC	*C. maximus* species-specific
	CmaxND6R6	GCTGGGGTCGGGTGTGGGTT	*C. maximus* species-specific
	CmaxND6R7	AGGCCAGCAGGAGACCAGCA	*C. maximus* species-specific

Appendix C. Sequence variation within *C. maximus* mitogenome haplotypes (*n* = 22). Relative sequence position (bp) and gene/region names are listed. Indels are indicated by a dash (-).

Locus	12S	12S	12S	12S	12S	12S	12S	16S	16S	16S	16S	ND1	ND1	ND1	ND1	ND1
bp	137	192	327	410	433	498	652	1115	2394	2550	2725	2930	2979	2980	3044	3118
Hap1	C	A	T	A	G	-	T	T	G	A	A	G	G	C	A	C
Hap2	T	G	T	G	G	C	T	C	A	A	A	G	A	T	A	T
Hap3	T	G	T	G	G	C	T	C	A	A	A	G	A	T	A	T
Hap4	T	G	T	G	G	C	T	C	A	A	A	G	A	T	G	T
Hap5	T	G	T	G	G	C	T	C	A	A	A	G	A	T	A	T
Hap6	T	G	T	G	T	C	T	C	A	A	A	A	A	T	A	T
Hap7	T	G	T	G	T	C	T	C	A	A	A	A	A	T	A	T
Hap8	T	G	T	G	G	C	T	C	A	A	A	G	A	T	A	T
Hap9	T	G	T	G	G	C	T	C	A	A	A	G	A	T	A	T
Hap10	T	G	T	G	G	C	T	C	A	A	A	G	A	T	A	T
Hap11	T	G	T	G	T	C	T	C	A	A	A	A	A	T	A	T
Hap12	C	A	T	A	G	-	T	T	G	A	A	G	G	C	A	C
Hap13	T	G	A	G	G	C	T	C	A	A	G	G	A	T	A	T
Hap14	T	G	T	G	G	C	T	C	A	A	A	G	A	T	A	T
Hap15	T	G	T	G	G	C	C	C	A	A	A	G	A	T	A	T
Hap16	T	G	T	G	G	C	T	C	A	A	A	G	A	T	A	T
Hap17	T	G	T	G	G	C	T	C	A	A	A	G	A	T	A	T
Hap18	T	G	T	G	G	C	T	C	A	G	A	G	A	T	A	T
Hap19	T	G	T	G	T	C	T	C	A	A	A	A	A	T	A	T
Hap20	T	G	T	G	G	C	T	C	A	A	A	G	A	T	A	T
Hap21	C	A	T	A	G	-	T	T	G	A	A	G	G	C	A	C
Hap22	T	G	T	G	G	C	T	C	A	A	A	G	A	T	A	T

Appendix C continued.

Locus	ND1	ND1	ND1	*tRNA*^{Gln}	ND2	ND2	ND2	ND2	ND2	CO1	CO1	CO1	CO1	CO1	CO1	CO1
bp	3301	3424	3784	3901	4524	4638	4854	4899	5056	5713	5779	5980	6079	6286	6439	6763
Hap1	C	A	A	C	A	T	C	A	T	G	G	A	T	T	C	C
Hap2	T	A	G	T	G	C	T	G	C	G	A	G	C	C	A	T
Hap3	T	A	G	T	G	C	T	G	C	A	A	G	C	C	A	T
Hap4	T	G	G	T	G	C	T	G	C	G	A	G	C	C	A	T
Hap5	T	A	G	T	G	C	T	G	C	G	A	G	C	C	A	T
Hap6	T	A	G	T	G	C	T	A	C	G	A	G	C	C	A	T
Hap7	T	A	G	T	G	C	T	G	C	G	A	G	C	C	A	T
Hap8	T	A	G	T	G	C	T	G	C	G	A	G	C	C	A	T
Hap9	T	A	G	T	G	C	T	G	C	G	A	G	C	C	A	T
Hap10	T	A	G	T	G	C	T	G	C	G	A	G	C	C	A	T
Hap11	T	A	G	T	G	C	T	A	C	G	A	G	C	C	A	T
Hap12	C	A	A	T	A	T	C	A	T	G	G	A	T	T	C	C
Hap13	T	A	G	T	G	C	T	G	C	G	A	G	C	C	A	T
Hap14	T	A	G	T	G	C	T	G	C	G	A	G	C	C	A	T
Hap15	T	A	G	T	G	C	T	G	C	G	A	G	C	C	A	T
Hap16	T	A	G	T	G	C	T	G	C	G	A	G	C	C	A	T
Hap17	T	A	G	T	G	C	T	G	C	G	A	G	C	C	A	T
Hap18	T	A	G	T	G	C	T	G	C	G	A	G	C	C	A	T
Hap19	T	A	G	T	G	C	T	G	C	G	A	G	C	C	A	T
Hap20	T	A	G	T	G	C	T	G	C	G	A	G	C	C	A	T
Hap21	C	A	A	C	A	T	C	A	T	G	G	A	T	T	C	C
Hap22	T	A	G	T	G	C	T	G	C	G	A	G	C	C	A	T

Appendix C continued.

Locus	CO1	CO1	CO2	CO2	CO2	CO2	CO2	CO2	CO2	CO2	CO2	CO2	CO2	CO2	CO2	CO2
bp	6907	6934	7295	7391	7403	7439	7520	7526	7591	7725	7766	7775	7834	7835	7836	7938
Hap1	A	C	A	A	T	A	A	A	T	G	C	A	T	A	G	G
Hap2	G	T	G	A	T	A	G	G	T	G	C	G	T	A	G	A
Hap3	G	T	G	A	T	A	G	G	T	G	C	G	T	A	G	A
Hap4	G	T	G	A	T	A	G	G	T	G	C	G	T	A	G	A
Hap5	G	T	G	A	T	A	G	G	T	G	C	G	T	A	G	A
Hap6	G	T	G	A	C	A	G	G	T	G	T	G	T	A	G	A
Hap7	G	T	G	A	C	A	G	A	T	G	C	G	T	A	G	A
Hap8	G	T	G	A	T	A	G	G	C	G	C	G	T	A	G	A
Hap9	G	T	G	G	T	A	G	G	T	G	C	G	T	A	G	A
Hap10	G	T	G	G	T	A	G	G	T	G	C	G	-	-	-	A
Hap11	G	T	G	A	C	A	G	G	T	G	C	G	T	A	G	A
Hap12	A	C	A	A	T	A	A	A	T	G	C	A	T	A	G	G
Hap13	G	T	G	A	C	G	G	G	T	A	C	G	T	A	G	A
Hap14	G	T	G	A	T	A	G	G	T	G	C	G	T	A	G	A
Hap15	G	T	G	A	T	A	G	G	T	G	C	G	T	A	G	A
Hap16	G	T	G	A	T	A	G	G	T	G	C	G	T	A	G	A
Hap17	G	T	G	A	T	A	G	G	T	G	C	G	T	A	G	A
Hap18	G	T	G	A	T	A	G	G	T	G	C	G	T	A	G	A
Hap19	G	T	G	A	C	A	G	G	T	G	C	G	T	A	G	A
Hap20	G	T	G	A	T	A	G	G	T	G	C	G	T	A	G	A
Hap21	A	C	A	A	T	A	A	A	T	G	C	A	T	A	G	G
Hap22	G	T	G	A	T	A	G	G	T	G	C	G	T	A	G	A

Appendix C continued.

Locus	ATP8	ATP8	ATP8	ATP8	ATP8	ATP6	ATP6	ATP6	ATP6	ATP6	ATP6	CO3	CO3	CO3	CO3	CO3
bp	7971	7973	8013	8040	8059	8195	8426	8458	8602	8620	8662	8800	8872	9004	9040	9136
Hap1	G	T	T	A	C	C	T	G	C	C	C	C	C	T	C	C
Hap2	A	T	C	G	T	T	C	G	T	T	T	C	T	C	T	C
Hap3	A	T	C	G	T	T	C	G	T	T	T	C	T	C	T	C
Hap4	A	T	C	G	T	T	C	G	T	T	T	C	T	C	T	C
Hap5	A	T	C	G	T	T	C	G	T	T	T	T	T	C	T	C
Hap6	A	T	C	G	C	T	C	G	T	T	T	C	T	C	T	C
Hap7	A	T	C	G	C	T	C	C	T	T	T	C	T	C	T	C
Hap8	A	T	C	G	T	T	C	G	T	T	T	T	T	C	T	C
Hap9	A	T	C	G	T	T	C	G	T	T	T	C	T	C	T	C
Hap10	A	T	C	G	T	T	C	G	T	T	T	C	T	C	T	C
Hap11	A	T	C	G	C	T	C	G	T	T	T	C	T	C	T	C
Hap12	G	T	T	A	C	C	T	G	C	C	C	C	C	T	C	T
Hap13	A	T	C	G	C	T	C	G	T	T	T	C	T	C	T	C
Hap14	A	T	C	G	T	T	C	G	T	T	T	C	T	C	T	C
Hap15	A	C	C	G	T	T	C	G	T	T	T	C	T	C	T	C
Hap16	A	T	C	G	T	T	C	G	T	T	T	C	T	C	T	C
Hap17	A	T	C	G	T	T	C	G	T	T	T	T	T	C	T	C
Hap18	A	T	C	G	T	T	C	G	T	T	T	C	T	C	T	C
Hap19	A	T	C	G	C	T	C	G	T	T	T	C	T	C	T	C
Hap20	A	T	C	G	T	T	C	G	T	T	T	C	T	C	T	C
Hap21	G	T	T	A	C	C	T	G	C	C	C	C	C	T	C	C
Hap22	A	T	C	G	T	T	C	G	T	T	T	T	T	C	T	C

Appendix C continued.

Locus	CO3	CO3	CO3	CO3	ND3	ND3	ND3	ND3	ND4L	ND4	ND4	ND4	ND4	ND4
bp	9205	9236	9415	9427	9811	9866	9867	9868	10133	10235	10339	10375	10433	10591
Hap1	T	G	T	G	G	C	T	A	T	T	C	A	C	G
Hap2	T	G	C	A	A	C	T	A	T	T	T	G	T	A
Hap3	T	G	C	A	A	-	-	-	T	T	T	G	T	A
Hap4	T	G	C	A	A	C	T	A	T	T	T	G	T	A
Hap5	T	G	C	A	A	C	T	A	T	C	T	G	T	A
Hap6	T	G	C	A	A	C	T	A	T	T	T	G	T	A
Hap7	T	G	C	A	A	C	T	A	T	T	T	G	T	A
Hap8	C	G	C	A	A	C	T	A	C	C	T	G	T	A
Hap9	T	G	C	A	A	-	-	-	T	T	T	G	T	A
Hap10	T	G	C	A	A	-	-	-	T	T	T	G	T	A
Hap11	T	G	C	A	A	C	T	A	T	T	T	G	T	A
Hap12	T	G	C	G	G	C	T	A	T	T	C	A	C	G
Hap13	T	G	C	A	A	C	T	A	T	T	T	G	T	A
Hap14	T	G	C	A	A	-	-	-	T	T	T	G	T	A
Hap15	T	G	C	A	A	C	T	A	T	T	T	G	T	A
Hap16	T	G	C	A	A	-	-	-	T	T	T	G	T	A
Hap17	T	G	C	A	A	C	T	A	T	C	T	G	T	A
Hap18	T	G	C	A	A	C	T	A	T	T	T	G	T	A
Hap19	T	G	C	A	A	C	T	A	T	T	T	G	T	A
Hap20	T	G	C	A	A	C	T	A	T	C	T	G	T	A
Hap21	T	G	C	G	G	C	T	A	T	T	C	A	C	G
Hap22	T	A	C	A	A	C	T	A	T	C	T	G	T	A

Appendix C continued.

Locus	ND4	ND4	ND4	ND4	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5
bp	10738	10753	10786	11170	11227	11614	11690	11965	11968	12102	12172	12279	12354	12361
Hap1	A	A	C	A	G	T	A	A	T	G	C	G	G	G
Hap2	G	A	C	G	A	T	A	A	C	A	T	A	A	G
Hap3	G	G	C	G	A	T	A	A	T	A	T	A	A	G
Hap4	G	A	C	G	A	T	A	A	T	A	T	A	A	G
Hap5	G	A	C	G	A	T	A	A	T	A	T	A	A	A
Hap6	G	A	C	G	G	T	A	G	T	A	T	A	A	G
Hap7	G	A	C	G	G	T	A	G	T	A	T	A	A	G
Hap8	G	A	C	G	A	T	A	A	T	A	T	A	A	A
Hap9	G	A	C	G	A	T	A	A	T	A	T	A	A	G
Hap10	G	A	C	G	A	T	A	A	T	A	T	A	A	G
Hap11	G	A	C	G	G	T	A	G	T	A	T	A	A	G
Hap12	A	A	C	A	G	T	A	A	T	G	C	G	G	G
Hap13	G	A	C	G	A	T	A	A	T	A	T	A	A	G
Hap14	G	A	C	G	A	T	A	A	T	A	T	A	A	G
Hap15	G	A	C	G	A	T	A	A	T	A	T	A	A	G
Hap16	G	G	C	G	A	T	A	A	T	A	T	A	A	G
Hap17	G	A	C	G	A	T	A	A	T	A	T	A	A	G
Hap18	G	A	C	G	A	T	G	A	T	A	T	A	A	G
Hap19	G	A	C	G	G	T	A	G	T	A	T	A	A	G
Hap20	G	A	C	G	A	C	A	A	T	A	T	A	A	G
Hap21	A	A	T	A	G	T	A	A	T	G	C	G	G	G
Hap22	G	A	C	G	A	T	A	A	T	A	T	A	A	G

Appendix C continued.

Locus	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5
bp	12490	12507	12543	12545	12648	12739	12762	12804	13047	13077	13233	13299	13534	13566
Hap1	G	G	C	C	G	C	T	T	T	G	A	T	C	G
Hap2	G	C	T	T	A	T	T	C	C	A	G	T	C	A
Hap3	G	C	T	C	A	T	T	C	C	A	A	T	C	A
Hap4	G	C	T	C	A	T	T	C	C	A	A	T	C	A
Hap5	G	C	T	C	A	T	T	C	C	A	A	T	C	A
Hap6	G	C	T	C	A	T	T	C	C	A	A	T	C	A
Hap7	G	C	T	C	A	T	T	C	C	A	A	T	C	A
Hap8	G	C	T	C	A	T	T	C	C	A	A	T	C	A
Hap9	G	C	T	C	A	T	T	C	C	A	A	T	C	A
Hap10	G	C	T	C	A	T	T	C	C	A	A	T	T	A
Hap11	G	C	T	C	A	T	T	C	C	A	A	T	C	A
Hap12	G	G	C	C	G	C	T	T	T	A	A	T	C	G
Hap13	G	C	T	C	A	T	T	C	C	A	A	A	C	A
Hap14	G	C	T	C	A	T	T	C	C	A	A	T	C	A
Hap15	G	C	T	C	A	T	C	C	C	A	A	T	C	A
Hap16	G	C	T	C	A	T	T	C	C	A	A	T	C	A
Hap17	G	C	T	C	A	T	T	C	C	A	A	T	C	A
Hap18	G	C	T	C	A	T	T	C	C	A	A	T	C	A
Hap19	C	C	T	C	A	T	T	C	C	A	A	T	C	A
Hap20	G	C	T	C	A	T	T	C	C	A	A	T	C	A
Hap21	G	G	C	C	G	C	T	T	T	A	A	T	C	G
Hap22	G	C	T	C	A	T	T	C	C	A	A	T	C	A

Appendix C continued.

Locus	ND5	ND5	ND5	ND6	ND6	ND6	ND6	ND6	ND6	Cyt b	Cyt b	Cyt b	Cyt b	Cyt b
bp	13584	13654	13728	13851	13895	13982	14102	14174	14227	14447	14557	14596	14638	14758
Hap1	G	T	G	G	C	G	G	C	C	T	T	G	T	T
Hap2	A	C	G	G	C	G	G	T	C	T	C	A	C	C
Hap3	A	C	G	G	C	G	G	T	C	T	C	A	C	C
Hap4	A	C	G	G	C	G	G	T	C	T	C	A	C	C
Hap5	A	C	C	G	C	G	G	T	T	T	C	A	C	C
Hap6	A	C	G	G	C	G	G	T	C	T	C	A	C	C
Hap7	A	C	G	G	C	G	G	T	C	T	C	A	C	C
Hap8	A	C	C	G	C	G	G	T	T	T	C	A	C	C
Hap9	A	C	G	G	C	G	A	T	C	T	C	A	C	C
Hap10	A	C	G	G	C	G	A	T	C	T	C	A	C	C
Hap11	A	C	G	G	C	A	G	T	C	T	C	A	C	C
Hap12	G	T	G	G	T	G	G	C	C	T	T	G	T	C
Hap13	A	C	G	G	C	G	G	T	C	T	C	A	C	C
Hap14	A	C	G	G	C	G	G	T	C	T	C	A	C	C
Hap15	A	C	G	A	C	G	G	T	C	T	C	A	C	C
Hap16	A	C	G	G	C	G	G	T	C	T	C	A	C	C
Hap17	A	C	C	G	C	G	G	T	T	C	C	A	C	C
Hap18	A	C	G	G	C	G	G	T	C	T	C	A	C	C
Hap19	A	C	G	G	C	G	G	T	C	T	C	A	C	C
Hap20	A	C	G	G	C	G	G	T	C	T	C	A	C	C
Hap21	G	T	G	G	C	G	G	C	C	T	T	G	T	C
Hap22	A	C	C	G	C	G	G	T	T	T	C	A	C	C

Appendix C continued.

Locus	Cyt b	Cyt b	Cyt b	Cyt b	Cyt b	Cyt b	CR	CR	CR	CR
bp	14839	14879	15071	15163	15245	15316	15753	16021	16210	16365
Hap1	C	C	C	G	C	A	C	G	A	G
Hap2	T	T	T	G	T	G	T	A	G	A
Hap3	T	T	T	G	T	G	C	A	G	G
Hap4	T	T	T	G	T	G	T	A	G	A
Hap5	T	T	T	G	T	G	T	A	G	A
Hap6	T	T	T	G	T	G	T	A	G	G
Hap7	T	T	T	G	T	G	T	A	G	A
Hap8	T	T	T	G	T	G	T	A	G	A
Hap9	T	T	T	G	T	G	T	A	G	G
Hap10	T	T	T	G	T	G	T	A	G	G
Hap11	T	T	T	G	T	G	T	A	G	G
Hap12	C	C	C	G	C	A	C	G	A	G
Hap13	T	T	T	G	T	G	T	A	G	G
Hap14	T	T	T	G	T	G	T	A	G	A
Hap15	T	T	T	G	T	G	T	A	G	A
Hap16	T	T	T	G	T	G	C	A	G	G
Hap17	T	T	T	G	T	G	T	A	G	A
Hap18	T	T	T	G	T	G	T	A	G	A
Hap19	T	T	T	G	T	G	T	A	G	G
Hap20	T	T	T	A	T	G	T	A	G	A
Hap21	C	C	C	G	C	A	C	G	A	G
Hap22	T	T	T	G	T	G	T	A	G	A
Hap1	C	C	C	G	C	A	C	G	A	G

Appendix D. Sequence variation within *C. maximus* concatenated protein-coding gene haplotypes (*n* = 22). Relative sequence position (bp) and gene name are listed. Indels are indicated by a dash (-).

Gene	ND1	ND1	ND1	ND1	ND1	ND1	ND1	ND1	ND2	ND2	ND2	ND2	ND2	CO1	CO1
bp	2930	2979	2980	3044	3118	3301	3424	3784	4524	4638	4854	4899	5056	5713	5779
Hap1	G	G	C	A	C	C	A	A	A	T	C	A	T	G	G
Hap2	G	A	T	A	T	T	A	G	G	C	T	G	C	G	A
Hap3	G	A	T	A	T	T	A	G	G	C	T	G	C	A	A
Hap4	G	A	T	G	T	T	G	G	G	C	T	G	C	G	A
Hap5	G	A	T	A	T	T	A	G	G	C	T	G	C	G	A
Hap6	A	A	T	A	T	T	A	G	G	C	T	A	C	G	A
Hap7	A	A	T	A	T	T	A	G	G	C	T	G	C	G	A
Hap8	G	A	T	A	T	T	A	G	G	C	T	G	C	G	A
Hap9	G	A	T	A	T	T	A	G	G	C	T	G	C	G	A
Hap10	G	A	T	A	T	T	A	G	G	C	T	G	C	G	A
Hap11	A	A	T	A	T	T	A	G	G	C	T	A	C	G	A
Hap12	G	G	C	A	C	C	A	A	A	T	C	A	T	G	G
Hap13	G	A	T	A	T	T	A	G	G	C	T	G	C	G	A
Hap14	G	A	T	A	T	T	A	G	G	C	T	G	C	G	A
Hap15	G	A	T	A	T	T	A	G	G	C	T	G	C	G	A
Hap16	G	A	T	A	T	T	A	G	G	C	T	G	C	G	A
Hap17	G	A	T	A	T	T	A	G	G	C	T	G	C	G	A
Hap18	G	A	T	A	T	T	A	G	G	C	T	G	C	G	A
Hap19	A	A	T	A	T	T	A	G	G	C	T	G	C	G	A
Hap20	G	A	T	A	T	T	A	G	G	C	T	G	C	G	A
Hap21	G	G	C	A	C	C	A	A	A	T	C	A	T	G	G
Hap22	G	A	T	A	T	T	A	G	G	C	T	G	C	G	A

Appendix D continued.

Gene	CO1	CO1	CO1	CO1	CO1	CO1	CO1	CO2	CO2	CO2	CO2	CO2	CO2	CO2	CO2	CO2
bp	5980	6079	6286	6439	6763	6907	6934	7295	7391	7403	7439	7520	7526	7591	7725	7766
Hap1	A	T	T	C	C	A	C	A	A	T	A	A	A	T	G	C
Hap2	G	C	C	A	T	G	T	G	A	T	A	G	G	T	G	C
Hap3	G	C	C	A	T	G	T	G	A	T	A	G	G	T	G	C
Hap4	G	C	C	A	T	G	T	G	A	T	A	G	G	T	G	C
Hap5	G	C	C	A	T	G	T	G	A	T	A	G	G	T	G	C
Hap6	G	C	C	A	T	G	T	G	A	C	A	G	G	T	G	T
Hap7	G	C	C	A	T	G	T	G	A	C	A	G	A	T	G	C
Hap8	G	C	C	A	T	G	T	G	A	T	A	G	G	C	G	C
Hap9	G	C	C	A	T	G	T	G	G	T	A	G	G	T	G	C
Hap10	G	C	C	A	T	G	T	G	G	T	A	G	G	T	G	C
Hap11	G	C	C	A	T	G	T	G	A	C	A	G	G	T	G	C
Hap12	A	T	T	C	C	A	C	A	A	T	A	A	A	T	G	C
Hap13	G	C	C	A	T	G	T	G	A	C	G	G	G	T	A	C
Hap14	G	C	C	A	T	G	T	G	A	T	A	G	G	T	G	C
Hap15	G	C	C	A	T	G	T	G	A	T	A	G	G	T	G	C
Hap16	G	C	C	A	T	G	T	G	A	T	A	G	G	T	G	C
Hap17	G	C	C	A	T	G	T	G	A	T	A	G	G	T	G	C
Hap18	G	C	C	A	T	G	T	G	A	T	A	G	G	T	G	C
Hap19	G	C	C	A	T	G	T	G	A	C	A	G	G	T	G	C
Hap20	G	C	C	A	T	G	T	G	A	T	A	G	G	T	G	C
Hap21	A	T	T	C	C	A	C	A	A	T	A	A	A	T	G	C
Hap22	G	C	C	A	T	G	T	G	A	T	A	G	G	T	G	C

Appendix D continued.

Gene	CO2	CO2	CO2	CO2	CO2	ATP8	ATP8	ATP8	ATP8	ATP8	ATP6	ATP6	ATP6	ATP6	ATP6	ATP6
bp	7775	7834	7835	7836	7938	7971	7973	8013	8040	8059	8195	8426	8458	8602	8620	8662
Hap1	A	T	A	G	G	G	T	T	A	C	C	T	G	C	C	C
Hap2	G	T	A	G	A	A	T	C	G	T	T	C	G	T	T	T
Hap3	G	T	A	G	A	A	T	C	G	T	T	C	G	T	T	T
Hap4	G	T	A	G	A	A	T	C	G	T	T	C	G	T	T	T
Hap5	G	T	A	G	A	A	T	C	G	T	T	C	G	T	T	T
Hap6	G	T	A	G	A	A	T	C	G	C	T	C	G	T	T	T
Hap7	G	T	A	G	A	A	T	C	G	C	T	C	C	T	T	T
Hap8	G	T	A	G	A	A	T	C	G	T	T	C	G	T	T	T
Hap9	G	T	A	G	A	A	T	C	G	T	T	C	G	T	T	T
Hap10	G	-	-	-	A	A	T	C	G	T	T	C	G	T	T	T
Hap11	G	T	A	G	A	A	T	C	G	C	T	C	G	T	T	T
Hap12	A	T	A	G	G	G	T	T	A	C	C	T	G	C	C	C
Hap13	G	T	A	G	A	A	T	C	G	C	T	C	G	T	T	T
Hap14	G	T	A	G	A	A	T	C	G	T	T	C	G	T	T	T
Hap15	G	T	A	G	A	A	C	C	G	T	T	C	G	T	T	T
Hap16	G	T	A	G	A	A	T	C	G	T	T	C	G	T	T	T
Hap17	G	T	A	G	A	A	T	C	G	T	T	C	G	T	T	T
Hap18	G	T	A	G	A	A	T	C	G	T	T	C	G	T	T	T
Hap19	G	T	A	G	A	A	T	C	G	C	T	C	G	T	T	T
Hap20	G	T	A	G	A	A	T	C	G	T	T	C	G	T	T	T
Hap21	A	T	A	G	G	G	T	T	A	C	C	T	G	C	C	C
Hap22	G	T	A	G	A	A	T	C	G	T	T	C	G	T	T	T

Appendix D continued.

Gene	CO3	CO3	CO3	CO3	CO3	CO3	CO3	CO3	CO3	ND3	ND3	ND3	ND3	ND4L	ND4	ND4
bp	8800	8872	9004	9040	9136	9205	9236	9415	9427	9811	9866	9867	9868	10133	10235	10339
Hap1	C	C	T	C	C	T	G	T	G	G	C	T	A	T	T	C
Hap2	C	T	C	T	C	T	G	C	A	A	C	T	A	T	T	T
Hap3	C	T	C	T	C	T	G	C	A	A	-	-	-	T	T	T
Hap4	C	T	C	T	C	T	G	C	A	A	C	T	A	T	T	T
Hap5	T	T	C	T	C	T	G	C	A	A	C	T	A	T	C	T
Hap6	C	T	C	T	C	T	G	C	A	A	C	T	A	T	T	T
Hap7	C	T	C	T	C	T	G	C	A	A	C	T	A	T	T	T
Hap8	T	T	C	T	C	C	G	C	A	A	C	T	A	C	C	T
Hap9	C	T	C	T	C	T	G	C	A	A	-	-	-	T	T	T
Hap10	C	T	C	T	C	T	G	C	A	A	-	-	-	T	T	T
Hap11	C	T	C	T	C	T	G	C	A	A	C	T	A	T	T	T
Hap12	C	C	T	C	T	T	G	C	G	G	C	T	A	T	T	C
Hap13	C	T	C	T	C	T	G	C	A	A	C	T	A	T	T	T
Hap14	C	T	C	T	C	T	G	C	A	A	-	-	-	T	T	T
Hap15	C	T	C	T	C	T	G	C	A	A	C	T	A	T	T	T
Hap16	C	T	C	T	C	T	G	C	A	A	-	-	-	T	T	T
Hap17	T	T	C	T	C	T	G	C	A	A	C	T	A	T	C	T
Hap18	C	T	C	T	C	T	G	C	A	A	C	T	A	T	T	T
Hap19	C	T	C	T	C	T	G	C	A	A	C	T	A	T	T	T
Hap20	C	T	C	T	C	T	G	C	A	A	C	T	A	T	C	T
Hap21	C	C	T	C	C	T	G	C	G	G	C	T	A	T	T	C
Hap22	T	T	C	T	C	T	A	C	A	A	C	T	A	T	C	T

Appendix D continued.

Gene	ND4	ND4	ND4	ND4	ND4	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5
bp	10375	10433	10591	10738	11170	11227	11614	11690	11965	11968	12102	12172	12279	12354
Hap1	A	C	G	A	A	C	A	G	T	A	A	T	G	G
Hap2	G	T	A	G	A	C	G	A	T	A	A	C	A	A
Hap3	G	T	A	G	G	C	G	A	T	A	A	T	A	A
Hap4	G	T	A	G	A	C	G	A	T	A	A	T	A	A
Hap5	G	T	A	G	A	C	G	A	T	A	A	T	A	A
Hap6	G	T	A	G	A	C	G	G	T	A	G	T	A	A
Hap7	G	T	A	G	A	C	G	G	T	A	G	T	A	A
Hap8	G	T	A	G	A	C	G	A	T	A	A	T	A	A
Hap9	G	T	A	G	A	C	G	A	T	A	A	T	A	A
Hap10	G	T	A	G	A	C	G	A	T	A	A	T	A	A
Hap11	G	T	A	G	A	C	G	G	T	A	G	T	A	A
Hap12	A	C	G	A	A	C	A	G	T	A	A	T	G	G
Hap13	G	T	A	G	A	C	G	A	T	A	A	T	A	A
Hap14	G	T	A	G	A	C	G	A	T	A	A	T	A	A
Hap15	G	T	A	G	A	C	G	A	T	A	A	T	A	A
Hap16	G	T	A	G	G	C	G	A	T	A	A	T	A	A
Hap17	G	T	A	G	A	C	G	A	T	A	A	T	A	A
Hap18	G	T	A	G	A	C	G	A	T	G	A	T	A	A
Hap19	G	T	A	G	A	C	G	G	T	A	G	T	A	A
Hap20	G	T	A	G	A	C	G	A	C	A	A	T	A	A
Hap21	A	C	G	A	A	T	A	G	T	A	A	T	G	G
Hap22	G	T	A	G	A	C	G	A	T	A	A	T	A	A

Appendix D continued.

Gene	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5	ND5
bp	12361	12490	12507	12543	12545	12648	12739	12762	12804	13047	13077	13233	13299	13534
Hap1	G	G	G	C	C	G	C	T	T	T	G	A	T	C
Hap2	G	G	C	T	T	A	T	T	C	C	A	G	T	C
Hap3	G	G	C	T	C	A	T	T	C	C	A	A	T	C
Hap4	G	G	C	T	C	A	T	T	C	C	A	A	T	C
Hap5	A	G	C	T	C	A	T	T	C	C	A	A	T	C
Hap6	G	G	C	T	C	A	T	T	C	C	A	A	T	C
Hap7	G	G	C	T	C	A	T	T	C	C	A	A	T	C
Hap8	A	G	C	T	C	A	T	T	C	C	A	A	T	C
Hap9	G	G	C	T	C	A	T	T	C	C	A	A	T	C
Hap10	G	G	C	T	C	A	T	T	C	C	A	A	T	T
Hap11	G	G	C	T	C	A	T	T	C	C	A	A	T	C
Hap12	G	G	G	C	C	G	C	T	T	T	A	A	T	C
Hap13	G	G	C	T	C	A	T	T	C	C	A	A	A	C
Hap14	G	G	C	T	C	A	T	T	C	C	A	A	T	C
Hap15	G	G	C	T	C	A	T	C	C	C	A	A	T	C
Hap16	G	G	C	T	C	A	T	T	C	C	A	A	T	C
Hap17	G	G	C	T	C	A	T	T	C	C	A	A	T	C
Hap18	G	G	C	T	C	A	T	T	C	C	A	A	T	C
Hap19	G	C	C	T	C	A	T	T	C	C	A	A	T	C
Hap20	G	G	C	T	C	A	T	T	C	C	A	A	T	C
Hap21	G	G	G	C	C	G	C	T	T	T	A	A	T	C
Hap22	G	G	C	T	C	A	T	T	C	C	A	A	T	C

Appendix D continued.

Locus	ND5	ND5	ND5	ND5	ND6	ND6	ND6	ND6	ND6	ND6	Cyt b	Cyt b	Cyt b	Cyt b
bp	13566	13584	13654	13728	13851	13895	13982	14102	14174	14227	14447	14557	14596	14638
Hap1	G	G	T	G	G	C	G	G	C	C	T	T	G	T
Hap2	A	A	C	G	G	C	G	G	T	C	T	C	A	C
Hap3	A	A	C	G	G	C	G	G	T	C	T	C	A	C
Hap4	A	A	C	G	G	C	G	G	T	C	T	C	A	C
Hap5	A	A	C	C	G	C	G	G	T	T	T	C	A	C
Hap6	A	A	C	G	G	C	G	G	T	C	T	C	A	C
Hap7	A	A	C	G	G	C	G	G	T	C	T	C	A	C
Hap8	A	A	C	C	G	C	G	G	T	T	T	C	A	C
Hap9	A	A	C	G	G	C	G	A	T	C	T	C	A	C
Hap10	A	A	C	G	G	C	G	A	T	C	T	C	A	C
Hap11	A	A	C	G	G	C	A	G	T	C	T	C	A	C
Hap12	G	G	T	G	G	T	G	G	C	C	T	T	G	T
Hap13	A	A	C	G	G	C	G	G	T	C	T	C	A	C
Hap14	A	A	C	G	G	C	G	G	T	C	T	C	A	C
Hap15	A	A	C	G	A	C	G	G	T	C	T	C	A	C
Hap16	A	A	C	G	G	C	G	G	T	C	T	C	A	C
Hap17	A	A	C	C	G	C	G	G	T	T	C	C	A	C
Hap18	A	A	C	G	G	C	G	G	T	C	T	C	A	C
Hap19	A	A	C	G	G	C	G	G	T	C	T	C	A	C
Hap20	A	A	C	G	G	C	G	G	T	C	T	C	A	C
Hap21	G	G	T	G	G	C	G	G	C	C	T	T	G	T
Hap22	A	A	C	C	G	C	G	G	T	T	T	C	A	C

Appendix D continued.

Locus	Cyt b	Cyt b	Cyt b	Cyt b	Cyt b	Cytb	CR	CR	CR	CR
bp	14758	14839	14879	15071	15163	15245	15316	15753	16021	16210
Hap1	T	C	C	C	G	C	A	C	G	A
Hap2	C	T	T	T	G	T	G	T	A	G
Hap3	C	T	T	T	G	T	G	C	A	G
Hap4	C	T	T	T	G	T	G	T	A	G
Hap5	C	T	T	T	G	T	G	T	A	G
Hap6	C	T	T	T	G	T	G	T	A	G
Hap7	C	T	T	T	G	T	G	T	A	G
Hap8	C	T	T	T	G	T	G	T	A	G
Hap9	C	T	T	T	G	T	G	T	A	G
Hap10	C	T	T	T	G	T	G	T	A	G
Hap11	C	T	T	T	G	T	G	T	A	G
Hap12	C	C	C	C	G	C	A	C	G	A
Hap13	C	T	T	T	G	T	G	T	A	G
Hap14	C	T	T	T	G	T	G	T	A	G
Hap15	C	T	T	T	G	T	G	T	A	G
Hap16	C	T	T	T	G	T	G	C	A	G
Hap17	C	T	T	T	G	T	G	T	A	G
Hap18	C	T	T	T	G	T	G	T	A	G
Hap19	C	T	T	T	G	T	G	T	A	G
Hap20	C	T	T	T	A	T	G	T	A	G
Hap21	C	C	C	C	G	C	A	C	G	A
Hap22	C	T	T	T	G	T	G	T	A	G

Appendix E. Global summary statistics for *C. maximus* Dataset IV ((concatenated protein-coding genes ATP8, CO2, ND3) including: number of individuals (*n*), sequence length in base pairs (bp), number of haplotypes (*h*), haplotype diversity (*hd*), nucleotide diversity (π), and overall ϕ_{ST} (Exocoffier 1992) and F_{ST} (Weir & Cockerham 1984) values.

n	bp	*h*	*hd* ± SD	π ± SD	ϕ_{ST}	F_{ST}
100	1 208	15	0.852 ± 0.041	0.0023 ± 0.000	0.025	0.020

Appendix F. *C. maximus* population-level pairwise values of Φ_{ST} (upper triangular matrix) and statistical probability values, p-values (lower triangular matrix) for Dataset III (individual protein coding genes, individual rRNA genes, and non-coding control region).

12S rRNA

Location	WNA	NZ	ENA
WNA	-	-0.0046	-0.0921
NZ	0.3253	-	-0.0055
ENA	0.9909	0.3063	-

16S rRNA

Location	WNA	NZ	ENA
WNA	-	-0.0585	-0.0290
NZ	0.9909	-	-0.0173
ENA	0.6847	0.4505	-

ATP synthase F0 subunit 6 (ATP6)

Location	WNA	NZ	ENA
WNA	-	-0.0901	-0.0627
NZ	0.9909	-	-0.0697
ENA	0.8108	0.9909	-

ATP synthase F0 subunit 8 (ATP8)

Location	WNA	NZ	ENA
WNA	-	-0.0618	-0.0925
NZ	0.8018	-	-0.0663
ENA	0.9909	0.8378	-

Appendix F continued.

Cytochrome c oxidase I (CO1)

Location	WNA	NZ	ENA
WNA	-	-0.0717	-0.0758
NZ	0.9909	-	-0.0552
ENA	0.7117	0.7928	-

Cytochrome c oxidase II (CO2)

Location	WNA	NZ	ENA
WNA	-	-0.0475	-0.0835
NZ	0.8108	-	-0.0403
ENA	0.9909	0.7117	-

Cytochrome c oxidase III (CO3)

Location	WNA	NZ	ENA
WNA	-	-0.0420	-0.0690
NZ	0.5315	-	-0.0512
ENA	0.9909	0.6487	-

Cytochrome b (Cyt b)

Location	WNA	NZ	ENA
WNA	-	-0.0646	-0.0829
NZ	0.7117	-	-0.0614
ENA	0.9909	0.7297	-

NADH dehydrogenase 1 (ND1)

Location	WNA	NZ	ENA
WNA	-	0.0034	-0.1000
NZ	0.3423	-	0.0034
ENA	0.9909	0.3874	-

Appendix F continued.

NADH dehydrogenase 2 (ND2)

Location			
	WNA	NZ	ENA
WNA	-	-0.0658	-0.1000
NZ	0.8198	-	-0.0658
ENA	0.9909	0.7389	-

NADH dehydrogenase 3 (ND3)

Location			
	WNA	NZ	ENA
WNA	-	-0.0896	0.0845
NZ	0.9909	-	0.0535
ENA	0.3514	0.3333	-

NADH dehydrogenase 4 (ND4)

Location			
	WNA	NZ	ENA
WNA	-	-0.0553	-0.0545
NZ	0.6937	-	-0.0105
ENA	0.7387	0.3153	-

NADH dehydrogenase 4L (ND4L)

Location			
	WNA	NZ	ENA
WNA	-	-0.0470	0.0122
NZ	0.7568	-	-0.0598
ENA	0.5225	0.5856	-

NADH dehydrogenase 5 (ND5)

Location			
	WNA	NZ	ENA
WNA	-	-0.0490	-0.0881
NZ	0.7117	-	-0.0460
ENA	0.9909	0.6937	-

Appendix F continued.

NADH dehydrogenase 6 (ND6)

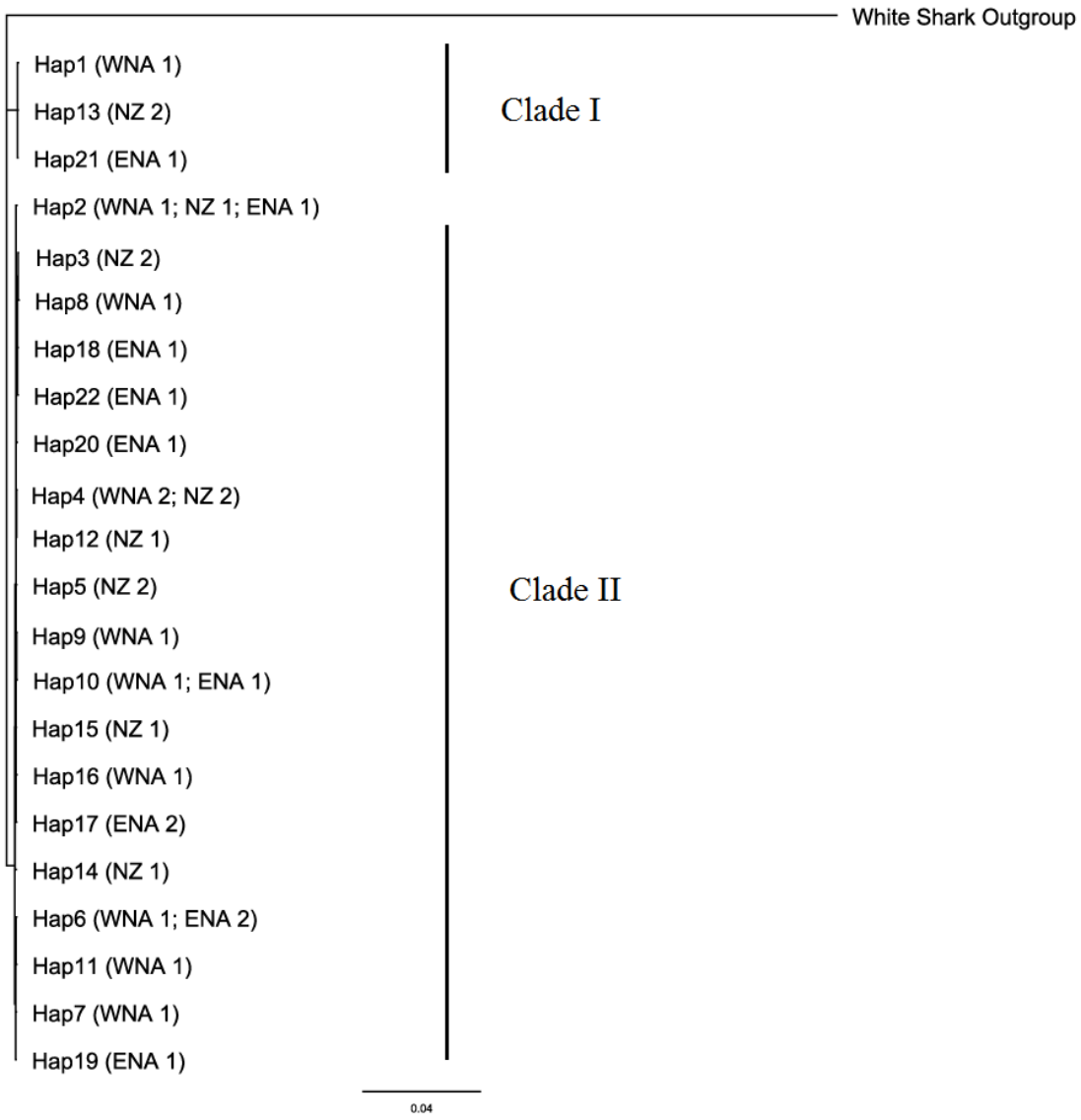
Location			
	WNA	NZ	ENA
WNA	-	0.0037	-0.0560
NZ	0.3694	-	-0.0362
ENA	0.9909	0.6847	-

Control Region

Location			
	WNA	NZ	ENA
WNA	-	-0.0564	-0.0344
NZ	0.6487	-	-0.0011
ENA	0.7027	0.3784	-

WNA = western North Atlantic, NZ= New Zealand, ENA= eastern North Atlantic

Appendix G. Bayesian inference of *C. maximus* concatenated protein-coding gene haplotypes.



Appendix H. *C. maximus* neutrality statistics for the whole mitochondrial genome (Dataset I): Fu's F_S (Fu 1996) and R_2 (Ramierez-Soriano & Rozas 2002).

Sample Collection Location	Fu's (*Fs*)	R_2
WNA (*n* = 11)	-0.18627	0.1959
NZ (*n* = 12)	3.99158	0.1421
ENA (*n* = 11)	0.94529	0.1214

WNA = western North Atlantic, NZ= New Zealand, ENA= eastern North Atlantic

Acknowledgements

I would like to express my gratitude to my advisor, Dr. Mahmood Shivji, for his leadership, support, and expertise in the field of population genetics. The opportunity to study such a charismatic shark species in a highly regarded laboratory was once in a lifetime. I would also like to thank my committee members, Dr. Jose Lopez for his proficiency in genetics and Dr. George Duncan for his enthusiasm for everything DNA. A big thank you to funding sources for supporting my research: Save Our Seas Foundation, Guy Harvey Ocean Foundation, FISH Florida Foundation, the IWFA Scholarship Fund, Guy Harvey Research Institute, Nova Southeastern University, AES Student Travel Grant and SGA Student Travel Grant Award.

A large part of this project including: the writing, presentations, thought processes, and beautiful maps, would not have been possible without the help of Dr. Andrea Bernard. I owe you so many trips to Tim Hortons.

Also, a big thank you to past and present labmates: Al Pickard, Shara Teter, Teagen Gray, Christine Testerman, Cassie Ruck, Jeremy Vaudo, and Derek Burkholder. Thank you for your help in the lab and all of the great laughs.

Last (and most importantly) all my family and friends, who encouraged me these past three years. Specifically I would like to thank: My family, The Finnegan-Carrine-Zandi clan, Michelle, Matt, Dave, Packer, and Jack.